FUNDAMENTUM

LILIANA PAVEL

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

EDITURA UNIVERSITĂŢII DIN BUCUREŞTI

LILIANA PAVEL

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

LILIANA PAVEL

AN INTRODUCTION TO FUNCTIONAL ANALYSIS

Editura Universității din București
2000

Referenţi ştiinţifici: **Acad. prof. dr. ROMULUS CRISTESCU**
Conf. dr. GHEORGHE GRIGORE

Descrierea CIP a Bibliotecii Naţionale
PAVEL, LILIANA
An introduction to functional analysis / Liliana Pavel
- Bucureşti, Editura Universităţii din Bucureşti, 2000
p. 216; cm 20
Bibliogr.
ISBN 973-575-413-4

517.98(075.8)

Contents

Preface

This elementary text is an introduction to functional analysis. The book covers only a limited number of topics, but they are sufficient to lay a foundation in functional linear analysis, which, partly because of its many applications has become a very popular mathematical dicipline interesting for applied mathematicians, probabilists, classical and numerical analysts. It grew out of my attempts to present the material in a way that was interesting and understandable to second-third year graduate students who are taking a course in this subject.

The only background material needed is what is usually covered in a one-year graduate level course analysis and an acquittance with linear algebra. However, to reach as large an audience as possible, the material is generally self-contained: any lack of knowledge can be compensated for by referring to Preliminaries, to Appendices and the references therein.

This book consists, basically of three parts. All chapters deal exclusively with linear problems. We begin with introductory results on vector spaces and linear operators (Chapter 1), and with basic facts from the theory of normed spaces and bounded linear operators on Banach spaces (Chapter 2 and Chapter 3). We continue with a chapter on the geometry of Hilbert spaces (Chapter 4), then proceed to the study of bounded linear operators acting on these spaces (Chapter 5), and to the elementary spectral theory of compact self-adjoint operators (Chapter 6). The last part of the text concentrates on locally convex spaces (Chapter 7). We offer a large selection of examples, applications and exercises (complete solutions of the exercises could be found in [8]).

We hope that exposure to this material will stimulate the students to expand their knowledge of functional analysis.

I would like to take this opportunity to thank my teachers in functional analysis of the University of Bucharest, Prof. R. Cristescu, I. Colojoara and Gh. Grigore.

Preliminaries

The first purpose of this introductory chapter is to establish the notation and the terminology that will occur throughout the book. We shall also present here some very well-known basic topology results.

0.1 Sets and functions

We assume that the reader is familiar with the basic concepts of set theory. Besides the usual signs as apartenence, inclusion, union and intersection, we will denote the complement of a set A (in X) with $X \setminus A$ or $C_X A$. Usually, the symbols $\mathbb{R}$, $\mathbb{C}$ are used for the set of real numbers, respectively complex numbers. $\mathbb{N}$ is the set of positive integers (not including zero).

A sets collection $\{A_i\}_{i\in I}$ is said to be a *partition* of the set X if $\bigcup_{i\in I} A_i = X$, $A_i \neq \emptyset$, $\forall i \in I$ and $A_i \cap A_j = \emptyset$, $\forall i \neq j$. A sets family $\{Z_i\}_{i\in I}$ is a *cover* for X if $X \subset \bigcup_{i\in I} Z_i$.

We will use words "*function*", "*mapping*" or "*application*" interchangeably. A function from a set X to another set Y, is denoted by $f : X \longrightarrow Y$, or $x \longmapsto f(x)$. If $A \subset X$, then $f(A) = \{f(x) \mid x \in A\}$ is a subset of Y and $f^{-1}(B) = \{x \mid f(x) \in B\}$ is a subset of X if $B \subset Y$. $f(X)$ will usually be called the *range* of f. X is called the *domain* of f. If $g : X \longrightarrow Y$ and $f : Y \longrightarrow Z$, the *composition* of f with g, $f \circ g$ is defined from X to Z, by $(f \circ g)(x) = f(g(x))$, for all x in X. The *identity* function from X to X, $x \longmapsto x$ is denoted by I_X, or, when is no danger of confusion, only by I. A function $f : X \longrightarrow Y$ will be called *injective* (or one-one) if for each y in Y there is at most x in X such that $f(x) = y$; f is called *surjective* (or onto) if

$f(X) = Y$. If f is both injective and surjective, we will say that it is *bijective*. In this case, there exists a function, from Y to X, called the *inverse* of f, written f^{-1} such that $f \circ f^{-1} = I_Y$ and $f^{-1} \circ f = I_X$. The restriction of $f : X \longrightarrow Y$ to the subset $A \subset X$ is denoted by $f|_A$. The *characteristic function* of a subset A of X, χ_A is defined as follows: $\chi_A(x) = 1$, if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$.

If $\{X_i\}_{i \in I}$ is a family of sets, their *Cartesian product* $\prod_{i \in I} X_i$ is the set of all mappings $x : I \longrightarrow \bigcup_{i \in I} X_i$ such that $x(i) \in X_i$, $\forall i \in I$. If $i \in I$ we define the *i-th projection* or coordinate mapping $pr_i : \prod_{i \in I} X_i \longrightarrow X_i$ by $pr_i(x) = x(i)$ (denoted by x_i). In the particular case of two sets, X_1, X_2, the Cartesian product is denoted by $X_1 \times X_2$, thus it may be identified with the set of ordered pairs (x_1, x_2), with $x_1 \in X_1$, $x_2 \in X_2$. When $X_i = X$, $\forall i \in I$, we write X^I instead of $\prod_{i \in I} X_i$.

A (binary) *relation* in a set X is just a subset $\mathcal{R}$ of $X \times X$; it is customary, though, to use a relation sign, such as $\leq$ (or as $\sim$), to indicate the relation. Thus, $(x, y) \in \mathcal{R}$ is written $x \leq y$ (or $x \sim y$). A relation is said to be *transitive* if $\forall x$, y, $z \in X$, $(x, y) \in \mathcal{R}$ and $(y, z) \in \mathcal{R}$ implies $(y, z) \in \mathcal{R}$; *reflexive* if $\forall x \in X$, $(x, x) \in \mathcal{R}$; *symmetric* if $\forall x$, $y \in X$, $(x, y) \in \mathcal{R}$ implies $(y, x) \in \mathcal{R}$ and *antisymmetric* if $\forall x$, $y \in X$, $(x, y) \in \mathcal{R}$ and $(y, x) \in \mathcal{R}$ implies $x = y$.

An *equivalence relation*, written $\sim$, is a relation which is reflexive, symmetric and transitive. If $\mathcal{R}$ is an equivalence relation on X and $x \in X$, the set of elements of X related to a given $x \in X$ is called the *equivalence class* of x modulo $\mathcal{R}$, denoted usually as $\hat{x}$. Any equivalence relation $\mathcal{R}$ on X determines a partition of the set X. The set $\{\hat{x} \mid x \in X\}$ will be denoted by $X/\mathcal{R}$. Conversely, any partition of the set X determines an equivalence relation on X. The mapping from X onto $X/\mathcal{R}$, $x \mapsto \hat{x}$ is called the *canonical surjection* (denoted usually by π).

A *partial ordering* (or order) on X is a relation $\mathcal{R}$, written $\leq$, $\prec$ which is transitive, reflexive, and antisymmetric. Whenever $\mathcal{R} = X \times X$, the set X is called *totally ordered*.

An (partial) ordered set is a pair $(X, \leq)$, where X is a nonempty set and "$\leq$ " is an order on X. An ordered set is called *directed* (to the right) if $\forall x$, $y \in X$, $\exists z \in X$ such that $x \leq z$ and $y \leq z$.

For example, $\mathbb{R}$ with its usual order is totally ordered. If $\mathcal{A}$ is a family of subsets of the set X the usual order on $\mathcal{A}$ is $C \leq D \Leftrightarrow C \subset D$; if $\mathcal{F}$ is a family of real mappings on X one considers on $\mathcal{F}$ the partial ordering

$f \leq g \Leftrightarrow f(x) \leq g(x), \forall x \in X$.

Let $(X, \leq)$ be an ordered set and $A \subset X$. An element $p \in X$ is called *upper bound* (or majorant) for A if $y \leq p$ for all $y \in Y$. *Lower bound* (or minorant) is defined analogously. If $m \in X$ and $m \leq x$ implies $x = m$,(so, m has no proper majorants) we say that m is a *maximal element* of X, and analogously, if $s \in X$ and $x \leq s$ implies $x = s$, we say that s is a *minimal element* of X.

Let us say that the ordered set $(X, \leq)$ is *inductively ordered* if each totally ordered subset if X (in the order induced from X), has a majorant in X. *Zorn's lemma* states that every inductively ordered set has a maximal element.

A subset A of the ordered set $(X, \leq)$ is said to be *bounded from above* (or majorized), if it has a majorant. If there exists a majorant a which belongs to A, then a is called the *greatest element* of A. The notions of set bounded from bellow (or minorized) and smallest element are appearent. The subset A is said to be *bounded* if it is bounded from above and from below. One says that A has a least upper bound, (respectively a greatest lower bound) if it is bounded from above and the set of its majorants has a smallest element, α (respectively the set of its minorants has a greatest element, β). The element α is called the *least upper bound* of A or *supremum* of A, denoted usually by sup A; analogously the element β is called the *greatest lower bound* or *infimum* of A and is denoted by inf A. The least upper bound (respectively the greatest lower bound), if it exists, is unique.

Finally, a *net* in a set X is a function $x : \Delta \longrightarrow X$ where Δ is an ordered set directed to the right. As usually, we denote $x(\alpha)$ by x_α and the net by $(x_\alpha)_{\alpha \in \Delta}$. If Δ is the set of natural numbers, with its usual order, the net is called *sequence* and is denoted by $(x_n)_n$.

0.2 Topological spaces

Generalities. Usually, we will use the symbol τ for a topology and if X is a topological space with the topology τ, we will denote this space by (X, τ). If Y is a subset of X, we will denote by $\tau|_Y$ the *relativization* of τ to Y. For a subset A of the topological space (X, τ), the closure, respectively the interior will be denoted by $\overline{A}$, respectively $\overset{\circ}{A}$. A subset $A \subset X$ is said to be *dense* in X if $\overline{A} = X$. A topological vector space X is said to be *separable* if there is a countable dense subset of X. A subset $A \subset X$ is called *nowhere dense* in

X if $\overline{A}$ has an empty interior.

If (X, τ) is a topological space, a family $\mathcal{B} \subset \tau$ is called a *base* for τ if $\forall D \in \tau$, D is the union of sets of $\mathcal{B}$. A family $\mathcal{B}$ of sets is a base for some topology for the set $X = \bigcup_{B \in \mathcal{B}} B$ if and only if $\forall A,\ B \in \mathcal{B}$ and $\forall x \in A \cap B$, $\exists C \in \mathcal{B}$ such that $x \in C$ and $C \subset A \cap B$.

A family $\mathcal{S} \subset \tau$ is a *subbase* for the topology τ if the family of finite intersections of members of $\mathcal{S}$ is a base for τ. Every nonempty family $\mathcal{S}$ is a subbase for some topology for the set $X = \bigcup_{S \in \mathcal{S}} S$. It is the smallest topology containing $\mathcal{S}$ and it is uniquely determined by $\mathcal{S}$ (it is called the topology generated by $\mathcal{S}$).

If $\{(X_i, \tau_i)\}_{i \in I}$ is a family of topological spaces, $\tau_i = \{D^{(i)}\}$, then, the collection of sets of $X = \prod\limits_{i \in I} X_i$, $\{pr_i^{-1}(D^{(i)})\}_{i \in I}$ is a subbase for some topology for X, called the *product topology*, $\prod\limits_{i \in I} \tau_i$.

Let (X, τ) be a topological space, $\mathcal{R}$ an equivalence relation on X and $\pi : X \longrightarrow X/\mathcal{R}$, the canonical surjection. The family

$$\hat{\tau} = \{\hat{D} \subset X/\mathcal{R} \mid \pi^{-1}(\hat{D}) \in \tau\}$$

is a topology for $X/\mathcal{R}$ called the *quotient topology* (modulo $\mathcal{R}$).

If x is a point of the topological space (X, τ), and $\mathcal{V}_x$ is the neighbourhood system of x, then, a family of neighbourhoods $\mathcal{B}_x$ of x is called a *base for the neighbourhood system* of x (or a *fundamental system of neighbourhoods* of x) if $\forall V \in \mathcal{V}_x$, $\exists B \in \mathcal{B}_x$, $B \subset V$.

Theorem 0.2.1 *Let (X, τ) be a topological space and for each $x \in X$ let $\mathcal{V}_x$ be the family of all neighbourhoods of x. Then :*
V1) If $V \in \mathcal{V}_x$, then $x \in V$
V2) If $V \in \mathcal{V}_x$ and $V \subset W$, then $W \in \mathcal{V}_x$;
V3) If $V, W \in \mathcal{V}_x$, then $V \cap W \in \mathcal{V}_x$;
V4) If $V \in \mathcal{V}_x$, then there is a member W of $\mathcal{V}_x$ such that $W \in \mathcal{V}_y$, for each y in V.
Conversely, if to each $x \in X$, there is a nonempty family $\mathcal{V}_x$ satisfying V1)-V4), then the family τ of all sets G, such that $G \in \mathcal{V}_x$ whenever $x \in G$ is a topology on X; τ is the unique topology for X such that $\mathcal{V}_x$ is precisely the neighbourhood system of x relative to the topology τ.

The topological space (X, τ) is a *Hausdorff* space, if $\forall x, y \in X$, $x \neq y$, there are disjoint sets $U \in \mathcal{V}_x$ and $V \in \mathcal{V}_y$. Notice that $(\prod\limits_{i \in I} X_i, \prod\limits_{i \in I} \tau_i)$ is Hausdorff if and only if for all i in I, (X_i, τ_i) is Hausdorff.

Let (X, τ) and (Y, σ) be topological spaces. A function $f : X \longrightarrow Y$ is said to be *continuous* if $f^{-1}(A) \in \tau$ for every A in σ. It is said to be *continuous at a point* x in X if $f^{-1}(A) \in \mathcal{V}_x$ for every $A \in \mathcal{V}_{f(x)}$. A function is continuous if and only if it is continuous at every point. A function $f : X \longrightarrow Y$ is *open* if $f(A) \in \sigma$ for every A in τ. A *homeomorphism* is a bijective function $f : X \longrightarrow Y$ that is both open and continuous, equivalently, both f and f^{-1} are continuous functions. It is clear that the composition of two continuous or open functions produces a function of the same type.

A net $\{x_\alpha\}_{\alpha \in \Delta} \subset (X, \tau)$ is said to be *convergent* to $x \in X$ ($x_\alpha \longrightarrow x$) if for each $V \in \mathcal{V}_x$, $\exists \alpha_V \in \Delta$ such that $\forall \alpha \geq \alpha_V$ implies $x_\alpha \in V$. The point x is called the *limit* of the net $(x_\alpha)_\alpha$, written,

$$x = \lim_\alpha x_\alpha$$

A function $f : (X, \sigma) \longrightarrow (Y, \sigma)$ is continuous at $x \in X$ if and only if

$$\forall \{x_\alpha\}_{\alpha \in \Delta}, \quad x_\alpha \longrightarrow x \quad \Longrightarrow \quad f(x_\alpha) \longrightarrow f(x).$$

For a complex (real)-valued function on X, the *support* of f is the subset of X, supp $f = \overline{\{x \in X \mid f(x) \neq 0\}}$.

Metric spaces. A *metric space* is a set M and a real-valued function $d(\cdot, \cdot)$ on $M \times M$ (called metric on M) which satisfies:

i) $d(x, y) \geq 0$;

ii) $d(x, y) = 0$ if and only if $x = y$;

iii) $d(x, y) = d(y, x)$;

iv) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

When it is not clear from the context which metric we are talking about, we will denote the metric space by (M, d). The set $\{x \in X \mid d(x, y) < r\}$ is called the *open ball*, $B(y, r)$ of radius r about the point y. The metric topology on a metric space (M, d), τ_d is defined as follows: a set $G \subset X$ is open if and only if $\forall y \in G$, $\exists r > 0$ such that $B(y, r) \subset G$. The topology τ_d is Hausdorff.

A sequence $(x_n)_n$ of a metric space (M, d) converges to an element x if and only if $d(x, x_n) \longrightarrow 0$ as $n \to \infty$, that means, for given $\varepsilon > 0$ there exists n_ε so that $n \geq n_\varepsilon$ implies $d(x, x_n) < \varepsilon$; x is called the limit of the sequence $(x_n)_n$, written,

$$x = \lim_n x_n$$

A sequence $(x_n)_n$ of a metric space (M, d) is called a *Cauchy sequence* if $\forall \varepsilon > 0$, there is n_ε so that $n, m \geq n_\varepsilon$ implies $d(x_n, x_m) < \varepsilon$. Any convergent sequence is Cauchy. A metric space in which all Cauchy sequences converge is called *complete*.

Let $A \subset X$ be. The *distance* from $x \in X$ to A is defined by

$$d(x, A) = \inf_{y \in A} d(x, y).$$

Clearly, $d(x, A) = 0$ if and only if x in the closure of A (in the metric topology τ_d).

An element x is in the closure of A if there exists a sequence $(x_n)_n \subset A$ such that $(x_n)_n$ converges to x.

If X, Y are metric spaces and f is a function from X to Y, then f is continuous at $x \in X$ if and only if for each sequence $(x_n)_n \subset X$, $x_n \longrightarrow x$ it results that $f(x_n) \longrightarrow f(x)$.

A topological space (X, τ) is said to be *metrizable* if there is a metric d on X such that $\tau = \tau_d$.

Compact sets. A subset K of a topological space (X, τ) is *compact* if and only if each open cover has a finite subcover. Each closed subset of a compact set is itself compact. If (X, τ) is Hausdorff, each compact set is closed. The topological space (X, τ) is compact if X is a compact set.

A family $\{F_\alpha\}_{\alpha \in I} \subset X$ of sets has the *finite intersection property* if the intersection of the members of each finite subfamily of $\{F_\alpha\}_{\alpha \in I}$ is nonempty. If (X, τ) is Hausdorff, the set $K \subset X$ is compact if and only if each family of closed subsets of X, $\{F_\alpha\}_{\alpha \in I} \subset X$ which has the finite intersection property on K,

$$K \cap \big(\bigcap_{\alpha \in J} F_\alpha \big) \neq \emptyset, \quad \forall J \subset I, \ J \text{ finite}$$

has a nonempty intersection on K, that means

$$K \cap \big(\bigcap_{\alpha \in I} F_\alpha \big) \neq \emptyset$$

It is well known the next theorem:

Theorem 0.2.2 *(Tychonoff's theorem) The product space $(\prod_{i \in I} X_i, \prod_{i \in I} \tau_i)$ is compact if and only if $\forall i \in I$, the space (X_i, τ_i) is compact.*

Further, the next characterization of compacity in metric spaces will be useful.

Theorem 0.2.3 *Let* (M, d) *a metric space and* τ_d *the metric topology on* M. *A subset* A *of* M *is compact if and only if each sequence of elements in* A, *has a subsequence which converges to an element of* A.

Chapter 1

Vector spaces and linear operators

1.1 Vector spaces

In this section we will introduce some elementary notions of linear algebra.

Next, we will denote by $\mathbb{K}$ one of the sets of the real numbers, $\mathbb{R}$, or of the complex numbers, $\mathbb{C}$. If $\alpha = a + bi \in \mathbb{C}$, $\operatorname{Re} \alpha = a$ and $\operatorname{Im} \alpha = b$. The conjugate of α, is $\overline{\alpha} = a - bi$.

Definition. A *vector space* (linear space) over $\mathbb{K}$ is a set X with a binary operation (addition), $(x, y) \longrightarrow x + y$ and with a mapping (scalar multiplication) defined on $\mathbb{K} \times X$, $(\alpha, x) \longrightarrow \alpha x$ satisfying the conditions:
1) $x + (y + z) = (x + y) + z,\ \forall x,\ y,\ z \in X$;
2) $x + y = y + x,\ \forall x,\ y \in X$;
3) $\exists 0 \in X$ such that $0 + x = x + 0,\ \forall x \in X$;
4) $\forall x \in X$, $\exists(-x) \in X$ such that $x + (-x) = (-x) + x = 0$;
5) $(\alpha + \beta)x = \alpha x + \beta x,\ \forall \alpha,\ \beta \in \mathbb{K},\ \ x \in X$;
6) $\alpha(x + y) = \alpha x + \alpha y, \forall \alpha \in \mathbb{K},\ \ x, y \in X$;
7) $(\alpha\beta)x = \alpha(\beta x), \forall \alpha,\ \beta \in \mathbb{K},\ \ x \in X$;
8) $1x = x,\ \forall x \in X$.
When $\mathbb{K} = \mathbb{R}$, X is called a *real vector space*, and, when $\mathbb{K} = \mathbb{C}$, X is said to be a *complex vector space*.

Remark. We notice that $\mathbb{K}$ is a vector space over $\mathbb{K}$, the scalar multiplication being the multiplication of $\mathbb{K}$.

Notations. If A, $B \subset X$ and $\Lambda \subset \mathbb{K}$ we will denote by

$$A + B = \{a + b \mid a \in A, b \in B\} \text{ and } \Lambda A = \{\alpha a \mid \alpha \in \Lambda, a \in A\}.$$

Definition. A subset Y of the vector space X is called a *linear subspace* of X if $Y + Y \subset Y$ and $\mathbb{K}\, Y \subset Y$.

Clearly, the intersection of a family of linear subspaces of X is also a linear subspace, thus we have the following definition:

Definition. Let X be a vector space and A a subset of X. The intersection of all linear subspaces of X containing A is called the *linear subspace spanned by* A, $\operatorname{Sp} A$.

Proposition 1.1.1 *Let X be a vector space and A a subset of X. Then,*

$$\operatorname{Sp} A = \{z \in X \mid z = \sum_{j=1}^{n} \alpha_j x_j,\ \alpha_j \in \mathbb{K}\ ,\ x_j \in A,\ n \in \mathbb{N}\ \}$$

Proof. Clearly $\operatorname{Sp} A \subset \{z \in X \mid z = \sum_{j=1}^{n} \alpha_j x_j,\ \alpha_j \in \mathbb{K},\ x_j \in A,\ n \in \mathbb{N}\ \}$ since this set is a linear subspace which contains A. Conversely, if Y is a linear subspace such that $A \subset Y$, it follows that every $z = \sum_{j=1}^{n} \alpha_j x_j$, $\alpha_j \in \mathbb{K}$, $x_j \in A$, $n \in \mathbb{N}$, is in Y, therefore the converse inclusion holds.

Definition. A subset C of a vector space X is said to be a *convex* set if for any two points x, y of C and any real number t, $0 \leq t \leq 1$, the point $tx + (1 - t)y$ is in C.

Definition. For $A \subset X$, the *convex hull* of A, denoted by co A, is defined by

$$\text{co } A = \bigcap_{C \supset A,\ C \text{ convex}} C$$

Remark. The convex hull of A is the smallest convex set containing A.

Definition. Let X be a vector space. The finite set of X, $\{x_k\}_{1 \leq k \leq n}$ is called *linearly independent* if $\sum_{j=1}^{n} \alpha_j x_j = 0$, $\alpha_j \in \mathbb{K}$ implies $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. A subset B of X is said to be linearly independent if each finite subset of B is linearly independent.

Definition. A subset B of X is called an *algebraic basis* of X if B is linearly independent and if for every $x \in X$, there exist $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{K}$ and $x_1, x_2, ...x_n \in B$ such that $x = \sum_{j=1}^{n} \alpha_j x_j$.

By a Zorn's lemma argument one can prove immediately:

Proposition 1.1.2 *Each vector space has an algebraic basis. Moreover, each linearly independent set in a vector space is contained in a basis.*

We have to notice that in a vector space X all bases have the same cardinal, called the *dimension* of the vector space X. The space is called finite dimensional if it has a finite dimensional basis, and otherwise it is called infinite dimensional.

We will end this section with two useful remarks that follow immediately by the definition of the vector spaces.

Remarks. 1. *Let $\{X_j\}_{j \in J}$ be a family of vector spaces over the same field $\mathbb{K}$. Then, $X = \prod_{j \in J} X_j$ becomes a vector space over $\mathbb{K}$ with the following operations:*

$$(x_j)_{j \in J} + (y_j)_{j \in J} = (x_j + y_j)_{j \in J}$$

$$\alpha(x_j)_{j \in J} = (\alpha x_j)_{j \in J}$$

2. *Let X be a vector space and Y a linear subspace of X. We define on X an equivalence relation, $x \sim y \iff x - y \in Y$. The echivalence class of x, $x + Y$ will be denoted by $\hat{x}$, and the set $\{\hat{x} \mid x \in X\}$ will be denoted by X/Y. X/Y, endowed with the operations*

$$\widehat{x} + \widehat{y} = \widehat{x + y},$$

$$\alpha \widehat{x} = \widehat{\alpha x}$$

is a vector space.

1.2 Linear functionals and operators

Let X, Y be vector spaces over the same field $\mathbb{K}$.

Definition. A mapping U from X to Y is said to be a *linear operator* if it is additive, $(U(x + y) = U(x) + U(y),\ \forall x,\ y \in X)$ and homogeneous $(U(\alpha x) = \alpha U(x),\ \forall \alpha \in \mathbb{K},\ x \in X)$.

Notation. If U is a linear operator from X to Y and $x \in X$, we shall often write Ux instead of $U(x)$.

Remark. If $U : X \longrightarrow Y$ is linear, then,

$$U(\sum_{j=1}^{n} \alpha_j x_j) = \sum_{j=1}^{n} \alpha_j U(x_j),\ \forall \alpha_j \in \mathbb{K},\ x_j \in X,\ n \in \mathbb{N}$$

It is easy to see that if U, V are linear operators from X to Y and $\alpha \in \mathbb{K}$, then $U + V : X \longrightarrow Y$,

$$(U+V)(x) = U(x) + V(x), \quad \forall x \in X$$

and $\alpha U : X \longrightarrow Y$,

$$(\alpha U)(x) = \alpha U(x), \quad \forall x \in X$$

are linear operators, too. Thus, denoting by $\mathcal{L}(X, Y)$ the set of all linear operators from X to Y, the next proposition is appearent.

Proposition 1.2.1 *The set $\mathcal{L}(X, Y)$ with the usual addition and scalar multiplication is a vector space over the field $\mathbb{K}$. $\mathcal{L}(X, Y)$ is called the space of linear operators from X to Y.*

Remark. *If $X = Y$ we write $\mathcal{L}(X)$ for $\mathcal{L}(X, Y)$.*

Notations. 1. If U, $V \in \mathcal{L}(X)$, we will often set UV instead of $U \circ V$ (which clearly is in $\mathcal{L}(X)$) and I instead of I_X.

2. Let U be in $\mathcal{L}(X, Y)$. We denote by

$$\operatorname{Ker} U = \{x \in X \mid U(x) = 0\},$$

called the *kernel* of U.

Proposition 1.2.2 *Let U be in $\mathcal{L}(X, Y)$. Then Ker U is a linear subspace of X. In addition, U is injective if and only if* $\operatorname{Ker} U = \{0\}$.

The proof of this result is immediate.

Remark. We notice that if U is an invertible linear operator, $U : X \longrightarrow Y$, then the inverse, $U^{-1} : Y \longrightarrow X$ is also a linear operator.

Definition. Two vector spaces X, Y are called *isomorphic* if there exists an invertible linear operators from X to Y.

Remarks. 1. If $B = \{x_j\}_{j\in J}$ is an algebraic basis of X and $F = \{y_j\}_{j\in J}$ a family of elements in Y, then there exists a unique linear operator $U : X \longrightarrow$ $\longrightarrow Y$ such that $U(x_j) = y_j$, $\forall j \in J$. Indeed, as every $x \in X$ can be uniquely represented as $x = \sum_{j\in F} \alpha_j x_j$, $F \subset J$, F finite, we define $U(x) = \sum_{j\in F} \alpha_j f_j$.

In particular if X and Y are two vector spaces with the same finite algebraic dimension, there exists an invertible linear operator from X to Y.

2. If X is a finite dimensional vector space and $B = \{x_1, x_2, ..., x_n\}$ is a basis of X, then $U \in \mathcal{L}(X)$ is uniquely determined by the elements $y_j = U(x_j)$, $j = 1, 2, .., n$. Every y_j, $j = 1, 2, .., n$ can be represented as $y_j = \sum_{k=1}^{n} \alpha_{kj} x_k$. The matrix $(\alpha_{jk})_{1\leq j,k\leq n}$ is called the matrix of the operator U in the basis B, and often we identify the operator with its matrix. Clearly, if

$$x = \sum_{j=1}^{n} \alpha_j x_j$$

and

$$U(x) = y = \sum_{j=1}^{n} \beta_j x_j,$$

we have

$$\beta_j = \sum_{k=1}^{n} \alpha_{jk} x_k, \quad j = 1, 2, .., n$$

Definition. Let X be a vector space over $\mathbb{K}$. A *functional* on X is a mapping from X to $\mathbb{K}$. A ***linear functional*** on X is a functional which is a linear operator from X to the vector space $\mathbb{K}$.

Notation. The vector space $\mathcal{L}(X, \mathbb{K})$ is denoted X'.

Remark. Whenever we will need to emphasize that a functional is defined on a real (complex) vector space we will call it a real (complex) functional.

Proposition 1.2.3 *Let Y be a linear subspace of the vector space X, $Y \subsetneq X$. The following are equivalent:*
(i) Y is a maximal subspace of X (with respect to the inclusion order);
(ii) For every $x_o \in X \setminus Y$, the linear subspace Sp $(Y \cup \{x_o\})$ *coincides to X;*

(iii) $\dim X/Y = 1$;
(iv) There exists $f \in X'$ *such that* $Y = \text{Ker } f$.

Proof. (i)⇒(ii) As $Y \subset \text{Sp }(Y \cup \{x_o\})$ and Y is maximal, it follows that $\text{Sp }(Y \cup \{x_o\})$ coincides to X.

(ii)⇒(i) If Z is a linear subspace that contains Y, $Y \subsetneq Z$, there exists $z \in Z \setminus Y$. Then, from (ii) $\text{Sp }(Y \cup \{z\}) = X$. From the inclusions

$$\text{Sp }(Y \cup \{z\}) \subset Z \subset X,$$

it follows that $Z = X$, thus, indeed Y is maximal.

(ii)⇒(iii) If $x_o \in X \setminus Y$, we will show that $X/Y = \{\alpha\widehat{x_o} \mid \alpha \in \mathbb{K}\}$. Let $\widehat{x_o}$ be in X/Y, $\widehat{x_o} \neq \widehat{0}$; therefore there exists $x \in X \setminus Y$, $x \in \widehat{x_o}$. As $X = \text{Sp }(Y \cup \{x_o\})$, $x = y + \alpha x_o$ for some $\alpha \in \mathbb{K}$, $\alpha \neq 0$ and $y \in Y$. Then, $\widehat{x_o} = \widehat{y + \alpha x_o} = \alpha\widehat{x_o}$.

(iii)⇒(iv) Since $\dim X/Y = 1$, there exists an invertible linear operator $U : X/Y \longrightarrow \mathbb{K}$. If $\pi : X \longrightarrow X/Y$ is the canonical surjection, then defining $f = U \circ \pi$, clearly f is a linear functional whose kernel is Y.

(iv)⇒(ii) Let $x_o \in X \setminus Y$ and x be arbitrary in X. We have to prove that there exist $y \in Y$ and $\alpha \in \mathbb{K}$ such that $x = y + \alpha x_o$. As $Y = \text{Ker } f$, where $f \in X'$, setting $\alpha = f(x)\ f(x_o)^{-1}$ and $y = x - f(x)\ f(x_o)^{-1} x_o$ everything is clear.

1.3 The Hahn-Banach extension theorem

In dealing with vector spaces endowed with a topology, one often needs to construct linear functionals with certain properties. In order to do that, first one defines the linear functional on a subspace of the vector space where it is easy to verify the desired properties; second, one uses an extension theorem which ensures that any such functional can be extended to the whole space while retaining the desired properties. One of the fundamental results about the extension of functionals, is the Hahn-Banach extension theorem. We begin with some general notions.

Definition. Let X be a vector space over the field $\mathbb{K}$. A ***sublinear functional*** on X is a real-valued function p on X which is subadditive ($p(x+y) \leq \leq p(x) + p(y), \forall x, y \in X$) and positive homogeneous ($p(tx) = tp(x)$, $\forall t \geq 0$,

$\forall x \in X$).

Definition. Let X be a vector space over the field $\mathbb{K}$. A real-valued subadditive function p on X is said to be a *seminorm* if $p(\alpha x) = |\alpha| p(x)$, $\forall \alpha \in \mathbb{K}$, $\forall x \in X$.

A sublinear functional resembles a seminorm, except that the second condition is only supposed to hold for positive scalars, thus a seminorm is in particular a sublinear functional. We remark, that, if p is a seminorm, $p(x) \geq 0$, $\forall x \in X$ (since for any sublinear functional $p(0) = 0$, $-p(-x) \leq p(x)$, and when p is seminorm $p(-x) = p(x)$).

Theorem 1.3.1 *(The* **Hahn-Banach extension theorem***) Let X be a real vector space, p a sublinear functional on X. Suppose that f is a linear functional defined on a subspace Y of X which satisfies $f(x) \leq p(x), \forall x \in Y$. Then, there is a linear functional $\tilde{f} : X \longrightarrow \mathbb{R}$, satisfying $\tilde{f}(x) \leq p(x), \forall x \in$ $\in X$, such that $\tilde{f}(x) = f(x), \forall x \in Y$.*

Proof. The idea of the proof is the following. First we will show that for $x_o \in X \setminus Y$, we can extend f to the space spanned by x_o and Y, $\mathrm{Sp}(Y \cup$ $\cup \{x_o\})$. Then, by a Zorn's lemma argument we prove that this process can be continued to extending f to the whole space X.

For arbitrary y', y'' in Y we have

$$f(y') - f(y'') = f(y' - y'') \leq p((y' + x_o) - (y'' + x_o)) \leq$$

$$\leq p(y' + x_o) + p(-(y'' + x_o)),$$

so,

$$-f(y'') - p(-(y'' + x_o)) \leq -f(y') + p(y' + x_o)$$

It follows that the set $A = \{-p(-(y + x_o)) - f(y) \mid y \in Y\} \subset \mathbb{R}$ has an upper bound and the set $B = \{p(y + x_o) - f(y) \mid y \in Y\} \subset \mathbb{R}$ has a lower bound. Let us denote $\sup A$ by c' and $\inf B$ by c''. As $c' \leq c''$, we may consider a real number $c \in [c', c'']$. Thus,

$$-p(-y - x_o) - f(y) \leq c \leq p(y + x_o) - f(y), \quad \forall y \in Y$$

We now define a real functional g on $\mathrm{Sp}(Y \cup \{x\}_o)$, by

$$g(y + \lambda x_o) = f(y) + \lambda c, \quad y \in Y, \lambda \in \mathbb{R}$$

It can be easily verified that this functional is linear. We have to see that

$$(*) \quad g(y + \lambda x_o) \leq p(y + \lambda x_o), \quad y \in Y,\ \lambda \in \mathbb{R}$$

If $\lambda = 0$, is clear. Suppose that $\lambda > 0$; after a division by λ, the above inequality becomes

$$g(\frac{1}{\lambda}y + x_o) \leq p(\frac{1}{\lambda}y + x_o) \Longleftrightarrow f(\frac{1}{\lambda}y) + c \leq p(\frac{1}{\lambda}y + x_o),$$

or equivalently,

$$c \leq p(\frac{1}{\lambda}y + x_o) - f(\frac{1}{\lambda}y)$$

If $\lambda < 0$, dividing by $(-\lambda) > 0$, the inequality $(*)$, is equivalent to

$$g(-\frac{1}{\lambda}y - x_o) \leq p(-\frac{1}{\lambda}y - x_o) \Longleftrightarrow f(-\frac{1}{\lambda}y) - c \leq p(-\frac{1}{\lambda}y - x_o),$$

and, finally, to

$$p(-\frac{1}{\lambda}y - x_o) - f(-\frac{1}{\lambda}y) \leq c$$

Therefore, the functional g defined on $\mathrm{Sp}(Y \cup \{x_o\})$ is an extension of f to $\mathrm{Sp}(Y \cup \{x_o\})$ such that $g(z) \leq p(z)$, $\forall z \in \mathrm{Sp}(Y \cup \{x_o\})$.

We now proceed with the Zorn's lemma argument. Let $\mathcal{F}$ be the collection of extensions g of f, $g : Z \longrightarrow \mathbb{R}$ which satisfy $g(z) \leq p(z)$ on the subspace Z, where they are defined. We partially order $\mathcal{F}$ by setting $g_1 \prec g_2$ if g_2 is defined on a larger subspace than g_1, and $g_2(z) = g_1(z)$ where they are both defined. We claim that $(\mathcal{F}, \prec)$ is inductively ordered, i.e. each totally ordered subset of $\mathcal{F}$ has an upper bound. Indeed, let $(g_\alpha)_\alpha$ a totally ordered family in $\mathcal{F}$, $g_\alpha : Z_\alpha \longrightarrow \mathbb{R}$. Define $\tilde{g} : \bigcup_\alpha Z_\alpha \longrightarrow \mathbb{R}$ by $\tilde{g}(z) = g_\alpha(z)$ if $z \in Z_\alpha$. The mapping $\tilde{g}$ is well defined, since for every α_1, α_2, $(g_\alpha)_\alpha$ being a totally ordered set, we have that $Z_{\alpha_1} \subset Z_{\alpha_2}$ (or conversely) and $g_{\alpha_2}(z) = g_{\alpha_1}(z)$ on Z_{α_1}. Clearly, $g_\alpha \prec \tilde{g}$ so each totally ordered subset has an upper bound. By Zorn's lemma, we conclude that $\mathcal{F}$ has a maximal element $\tilde{f}$, defined on some subspace X' of X, satisfying $\tilde{f}(x) \leq p(x)$ on the subspace X'. But X' must be all X, since if we had $X' \subsetneq X$, the first part of the proof applied to the functional $\tilde{f} : X' \longrightarrow \mathbb{R}$ would give a domination extension of $\tilde{f}$ to the space $\mathrm{Sp}(X' \cup \{x_o'\})$, where $x_o' \in X \setminus X'$, contradicting the maximality of $\tilde{f}$. Thus, the extension $\tilde{f}$ is defined on the whole X, which ends the proof.

Corollary 1.3.1 *If p is a sublinear functional on the real vector space X, then, for every $x_o \in X$, there is a linear functional on X such that $f(x_o) = p(x_o)$ and $f(x) \leq p(x)$, $\forall x \in X$.*

Proof. Let us denote by Y the subspace of X spanned by $\{x_o\}$ and define on it the linear functional $f(\lambda x_o) = \lambda p(x_o)$, $\lambda \in \mathbb{R}$. This functional satisfies that $f(x_o) = p(x_o)$ and $f(x) \leq p(x)$, $\forall x \in Y$. This follows by

$$f(\lambda x_o) = \lambda p(x_o) = p(\lambda x_o),$$

if $\lambda > 0$, and by

$$f(\lambda x_o) = \lambda p(x_o) \leq -\lambda p(-x_o) = p(\lambda x_o),$$

if $\lambda < 0$.

By the previous theorem the functional f can be extended to the whole space.

Theorem 1.3.2 (*Complex Hahn-Banach extension theorem*) *Let X be a complex vector space, p a seminorm on X. Suppose that f is a complex linear functional defined on a subspace Y of X satisfying $|f(x)| \leq p(x), \forall x \in Y$. Then, there is a complex linear functional $\tilde{f} : X \to \mathbb{C}$, which satisfies $|\tilde{f}(x)| \leq p(x), \forall x \in X$, such that $\tilde{f}(x) = f(x), \forall x \in Y$.*

Proof. The functional f can be represented as

$$f(x) = f_1(x) + i f_2(x), \quad \forall x \in Y,$$

where f_1, f_2 are real linear functionals on Y ($f_1(x) = \operatorname{Re} f(x)$ and $f_2(x) = \operatorname{Im} f(x)$). Since

$$i(f_1(x) + i f_2(x)) = i f(x) = f(ix) = f_1(ix) + i f_2(ix)$$

it follows that $f_2(x) = -f_1(ix)$, thus

$$f(x) = f_1(x) - i f_1(ix), \quad \forall x \in Y.$$

Now, we consider the real linear functional f_1 on Y and since

$$f_1(x) \leq |f_1(x)| \leq |f(x)| \leq p(x)$$

f_1 has a real linear extension $\tilde{f}_1$ to whole X obeying $\tilde{f}_1(x) \leq p(x)$ (by the Hahn-Banach extension theorem). Setting

$$\tilde{f}(x) = \tilde{f}_1(x) - i\tilde{f}_1(ix), \quad \forall x \in X,$$

clearly the complex functional $\tilde{f}$ is an additive extension of f. For an arbitrary complex number $a + bi$ and $x \in X$, we have

$$\tilde{f}((a+bi)x) = \tilde{f}_1(ax + ibx) - i\tilde{f}_1(aix - bx) =$$

$$= a\tilde{f}_1(x) + b\tilde{f}_1(ix) - ai\tilde{f}_1(ix) + bi\tilde{f}_1(x) =$$

$$= (a+bi)\tilde{f}_1(x) - (a+bi)i\tilde{f}_1(ix),$$

which proves the linearity of $\tilde{f}$. To complete the proof, we need only to see that $|\tilde{f}(x)| \leq p(x), \forall x \in X$. If we let $\theta = \arg\ \tilde{f}(x)$ and use the fact that $\mathrm{Re}\ \tilde{f} = \tilde{f}_1$, we see that

$$|\tilde{f}(x)| = \tilde{f}(x)\ e^{-i\theta} = \tilde{f}(e^{-i\theta}x) = \tilde{f}_1(e^{-i\theta}x) \leq$$

$$\leq p(e^{-i\theta}x) = |e^{-i\theta}x|\ p(x) = p(x)$$

Similarly to the real case we have the next corollary:

Corollary 1.3.2 *If p is a seminorm on the complex vector space X, then, for every $x_o \in X$, there is a linear functional on X such that $f(x_o) = p(x_o)$ and $|f(x)| \leq p(x)$, $\forall x \in X$.*

1.4 Exercises

1. Let X be a real (respectively complex) vector space, p a sublinear functional (respectively a seminorm) on X, and $\mathcal{F}$ the set of all real linear (respectively complex) functionals on X dominated (respectively in absolute value) by p. Show that $p(x) = \sup\{f(x) \mid f \in \mathcal{F}\}$ (respectively $p(x) = \sup\{|f(x)| \mid f \in \mathcal{F}\}$), $\forall x \in X$.

2. Let X be a real vector space, $x_o \in X$, p a sublinear functional on X. Show that for every $\lambda \in [-p(-x_o), p(x_o)]$ there exists a real linear functional on X such that $f(x_o) = \lambda$ and $|f(x)| \leq p(x)$, $\forall x \in X$.

3. Let X be a real vector space, p_1, p_2 sublinear functionals on X and f a real linear functional on X such that $f(x) \leq p_1(x) + p_2(x)$, $\forall x \in X$. Show that there exist f_1, f_2 linear functionals on X such that $f(x) = f_1(x) + f_2(x)$, $f_1(x) \leq p_1(x)$, $f_2(x) \leq p_2(x)$, $\forall x \in X$.

4 (F.F. Bonsall) Let X be a real vector space, $C \subset X$ such that $C + C \subset C$ and $\mathbb{R}_+ C \subset C$, $p : C \to R$ a sublinear functional, $q : C \to R$ a functional such that $q(x + y) \geq q(x) + q(y)$, $\forall x, y \in C$. We have $p(x) \leq$ $\leq q(x), \forall x \in C$. Show that there exists a real linear functional f on X such that $q(x) \leq f(x)$, $\forall x \in C$ and $f(x) \leq p(x)$, $\forall x \in X$.

5 (H. Nakano) Let X be a real vector space, Y a subspace of X, f a linear functional on X and p a subadditive, positive ($p(x) \geq 0$, $\forall x \in X$) functional on X with $\lim_{\lambda \searrow 0} p(\lambda x) = 0$, $\forall x \in X$ such that $f(x) \leq p(x)$, $\forall x \in X$. Show that there exists $\tilde{f}$ a real linear functional on X which is an extension of f with the property $\tilde{f}(x) \leq p(x)$, $\forall x \in X$.

6. Let p be a seminorm on the linear space X.

a) Show that Ker p is a linear subspace of X.

b) Show that the real mapping $\hat{p}$ on $X/\text{Ker } p$ defined by $\hat{p}(\hat{x}) = p(x)$, $x \in \hat{x}$ is a norm on $X/\text{Ker } p$.

c) Show that each non zero seminorm on $\mathbb{R}$ is a norm.

7. Let p be a seminorm on the linear space X and Y a linear subspace of X. Show that $p(\hat{x}) = \inf\{p(u) \mid u \in \hat{x}\}$ is a seminorm on X/Y.

Chapter 2

Banach spaces

2.1 Normed spaces, Banach spaces

Next X is a vector space over the field $\mathbb{K}$.

Definition. A *norm* on X is a function $\|\cdot\| : X \longrightarrow \mathbb{R}$ satisfying the conditions:

i) $\|x\| = 0 \Rightarrow x = 0$,

ii) $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in X$;

iii) $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\forall \alpha \in \mathbb{K}, \forall x \in X$.

Definition. A *normed space* is a pair $(X, \|\cdot\|)$ where X is a vector space and $\|\cdot\|$ is a norm on X.

Every normed space $(X, \|\cdot\|)$ is a metric space under the metric $d(x, y) = \|x - y\|$. The norm topology on X, $\tau_{\|\,\|}$ is the metric topology defined by this metric. Consequently, a sequence $(x_n)_n \subset X$ is convergent to $x \in X$ if given $\varepsilon > 0$, there is n_ε so that $\forall n \geq n_\varepsilon$, $\|x_n - x\| < \varepsilon$; $(x_n)_n \subset X$ is a Cauchy sequence if for every $\varepsilon > 0$, there exists n_ε such that $\forall n, m \geq n_\varepsilon$, $\|x_n - x_m\| < \varepsilon$.

Definition. If $(X, \tau_{\|\,\|})$ is complete, the normed space $(X, \|\cdot\|)$ is called a *Banach space*.

Remark. By the properties ii) and iii) of the norm, it follows

$$|\|x\| - \|y\|| \leq \|x - y\|, \quad \forall x, y \in X,$$

which shows that the real mapping on $(X, \tau_{\|\,\|})$, $x \mapsto \|x\|$ is continuous. It is also easy to see that the mapping $(x, y) \longmapsto x + y$ from $X \times X$ (endowed with the product topology) to X is continuous and for every $\alpha \neq 0$, the mapping $x \longmapsto \alpha x$ from X to X is a homeomorphism.

Notations. For $x_o \in X$ and $r > 0$ we denote by

$$B(x_o, r) = \{x \in X \mid \|x - x_o\| < r\},$$

the ball in X with center x_o and radius r and by $\overline{B}(x_o, r)$ its closure, which, one can easily show, coincides to $\{x \in X \mid \|x - x_o\| \leq r\}$. For $x_o = 0$, sometimes we will write $B(r)$ instead of $B(0, r)$; we have $B(x_o, r) = B(r) + \{x_o\}$.

Remark. Note that for any point x_o, the family $\{B(x_o, r)\}_{r>0}$ is a fundamental system of neighbourhoods of x_o.

Definition. A subset A of X is said to be *bounded* if there exists $\alpha > 0$ such that $\|x\| \leq \alpha$, $\forall x \in A$.

Remark. The closure of any bounded set is also bounded

Definition. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the vector space X are called *equivalent* (written $\|\cdot\|_1 \sim \|\cdot\|_2$) if there exist the constants $\alpha > 0$ and $\beta > 0$ such that

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1, \qquad \forall x \in X.$$

Proposition 2.1.1 *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ two norms on the vector space X. The following are equivalent:*
1) The norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent;
2) The topologies $\tau_{\|\cdot\|_1}$ and $\tau_{\|\cdot\|_2}$ coincide.

The proof of this result is appearent.

Definition. Two normed spaces X, Y over the same field $\mathbb{K}$ are said to be *isometrically isomorphic* if there exists a linear bijection $U : X \longrightarrow Y$ which preserves the norm, i.e. $\|U(x)\| = \|x\|$, $\forall x \in X$

Remark. A linear mapping from X to Y such that $\|U(x)\| = \|x\|$, $\forall x \in X$ will be called further a (linear) *isometry* on X

Definition. Let $(x_n)_n$ be a sequence in the normed space $(X, \|\ \|)$. The pair $((x_n)_n, (s_n)_n)$, where $s_n = \sum_{k=1}^{n} x_k$ is called the *series* corresponding to the sequence $(x_n)_n$, denoted by $\sum_{n\geq 1} x_n$.
The series $\sum_{n\geq 1} x_n$ is said to be *convergent*, and the sequence $(x_n)_n$ *summable*, if the sequence $(s_n)_n$ is convergent; the limit of this sequence is called the

sum of the series and is denoted by $\sum_{n=1}^{\infty} x_n$.
The series $\sum_{n\geq 1} x_n$ is said to be *absolutely convergent*, and the sequence $(x_n)_n$ *absolutely summable*, if the series $\sum_{n\geq 1} \|x_n\|$ is convergent.
The series $\sum_{n\geq 1} x_n$ is said to be *unconditional convergent*, if for each permutation σ of $\mathbb{N}$, the series $\sum_{n\geq 1} x_{\sigma(n)}$ is convergent.

It is important to have criteria to determine whether normed linear spaces are complete. Such a criterion is given by the next proposition.

Proposition 2.1.2 *Let $(X, \|\cdot\|)$ be a normed space. Then, $(X, \|\cdot\|)$ is Banach if and only if every absolutely convergent series in X is convergent.*

Proof. Suppose first that X is complete and let $\sum_{n\geq 1} x_n$ be an absolutely convergent series in X. Then, for $\forall \varepsilon > 0$, $\exists n_\varepsilon$ such that $\sum_{n=n_\varepsilon}^{\infty} \|x_n\| < \varepsilon$. Then, for every $n > m \geq n_\varepsilon$ we have

$$\|\sum_{k=1}^{n} x_k - \sum_{k=1}^{m} x_k\| = \|\sum_{k=m+1}^{n} x_k\| \leq \sum_{k=m+1}^{n} \|x_k\| \leq \sum_{n=n_\varepsilon}^{\infty} \|x_n\| < \varepsilon$$

We conclude that $(\sum_{k=1}^{n} x_k)_n$ is Cauchy, therefore, since X is Banach, convergent.

Conversely, let $(x_n)_n$ be a Cauchy sequence in X. Then, we can obtain a subsequence $(x_{k_n})_n$ such that

$$\|x_{k_{n+1}} - x_{k_n}\| < \frac{1}{2^{n+1}}, \quad \forall n$$

It follows that the series $\sum_{n\geq 1} \|x_{k_{n+1}} - x_{k_n}\|$ is convergent. Then, we know that the series $\sum_{n\geq 1}(x_{k_{n+1}} - x_{k_n})$ is convergent, too. Let us denote by x the sum

$$x = x_{k_1} + \sum_{n=1}^{\infty}(x_{k_{n+1}} - x_{k_n})$$

As for every positive integer $m \geq 2$,

$$x_{k_1} + \sum_{n=1}^{m-1}(x_{k_{n+1}} - x_{k_n}) = x_{k_m}$$

it follows that the subsequence $(x_{k_n})_n$ of the sequence $(x_n)_n$ is convergent. Taking into account that a Cauchy sequence which contains a convergent subsequence is itself convergent, the proof is finished.

Further, we shall show that the product of a family of normed spaces and the quotient of a normed space by a closed subspace can be naturally endowed with a structure of normed space.

Theorem 2.1.1 *Let $(X_j, \|\cdot\|_j)$, $j = 1, 2, \ldots, n$ be normed spaces. Then,*

$$\|(x_1, x_2, \ldots, x_n)\| = \max_{j=1,2,\ldots,n} \|x_j\|_j$$

is a norm on $X = \prod_{j=1}^n X_j$, and the topology $\tau_{\|\cdot\|}$ coincides to the product topology on X.
If, in addition each $(X_j, \|\cdot\|_j)$, $j = 1, 2, \ldots, n$ is a Banach space, the normed space $(X, \|\cdot\|)$ is also Banach.

Proof. It is not difficult to show that the real-valued mapping

$$(x_1, x_2, \ldots, x_n) \longmapsto \max_{j=1,2,\ldots,n} \|x_j\|_j$$

defined on $X = \prod_{j=1}^n X_j$, is a norm. If $B(\varepsilon)$, $B_j(\varepsilon)$ are the ε-balls in X, respectively X_j, $j = 1, 2, \ldots, n$ we also have that, $\forall \varepsilon > 0$,

$$B(\varepsilon) = \prod_{j=1}^n B_j(\varepsilon)$$

which proves that the $\tau_{\|\cdot\|}$ coincides to the product topology. Further let $((x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}))_m$ be a Cauchy sequence in X. By

$$\|x_j^{(k)} - x_j^{(l)}\|_j \leq \max_{j=1,2,\ldots,n} \|x_j^{(k)} - x_j^{(l)}\|_j, \ \forall k, l \in \mathbb{N}$$

it follows that for every $j = 1, 2, \ldots, n$, the sequence $(x_j^{(m)})_m \subset X_j$ is Cauchy. As $(X_j, \|\cdot\|_j)$, $j = 1, 2, \ldots, n$ is Banach, there exists $x_j \in X_j$, the limit of the sequence $(x_j^{(m)})_m$. Clearly, with an usual convergence argument,

$$\|(x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}) - (x_1, x_2, \ldots, x_n)\| \longrightarrow 0$$

as $m \to \infty$.

Theorem 2.1.2 *Let $(X, \|\cdot\|)$ be a normed space and Y a closed subspace of X. Then,*

$$\|\hat{x}\| = \inf_{x \in \hat{x}} \|x\|$$

is a norm on X/Y. If $(X, \|\cdot\|)$ is Banach, the space X/Y, with this norm, is also Banach.

Proof. By the properties of the norm on X, it follows immediately that the defined mapping on X/Y is subadditive and $\|\widehat{\alpha x}\| = |\alpha| \cdot \|\widehat{x}\|$. It remains to show that $\|\widehat{x}\| = 0$ involves $\widehat{x} = 0$. If $\|\widehat{x}\| = 0$, there exists a sequence $(x_n)_n \subset x + Y$ such that $x_n \longrightarrow 0$, as $n \to \infty$. As $x + Y$ is a closed set, we have that $0 \in x + Y$, so $\widehat{x} = 0$.

Now, let $(\widehat{x}_n)_n \subset X/Y$ be a Cauchy sequence so that the series $\sum_{n \geq 1} \|\widehat{x}_n\|$ converges. Accordingly to Proposition 2.1.2, in order to prove that X/Y is Banach, we have to show that the series $\sum_{n \geq 1} \widehat{x}_n$ converges too. By the definition of the norm on X/Y, for every $n \in \mathbb{N}$, there exists $y_n \in \widehat{x}_n$ such that

$$\|y_n\| < \|\widehat{x}_n\| + \frac{1}{2^n}$$

It follows that the series $\sum_{n \geq 1} \|y_n\|$ is convergent. As X is Banach, the series $\sum_{n \geq 1} y_n$ is also convergent. If

$$x = \sum_{n=1}^{\infty} y_n$$

by the inequality

$$\|\widehat{\sum_{k=1}^{n} y_k} - \widehat{x}\| \leq \| \sum_{k=1}^{n} y_k - x\|$$

it results that the series $\sum_{n \geq 1} \widehat{x}_n$ converges in X/Y to $\widehat{x}$.

2.2 Examples of Banach spaces

In this section we shall give some examples of Banach spaces, which will be useful further.

1. Finite dimensional normed spaces. The space $\mathbb{K}^n$ can be normed in many ways. Of major interest are the p-norms, for $p \in [1, \infty)$. Let us define on $\mathbb{K}^n$ the real mapping $(\xi_1, \xi_2, \ldots, \xi_n) \longmapsto \|(\xi_1, \xi_2, \ldots, \xi_n)\|_p$,

$$\|(\xi_1, \xi_2, \ldots, \xi_n)\|_p = (\sum_{k=1}^{n} |\xi_k|^p)^{\frac{1}{p}}$$

This mapping is a norm. In order to prove that we need some preliminary results.

(1) Let p, $q > 1$ such that $1/p + 1/q = 1$. Then,

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad \forall a, b \in \mathbb{K}$$

Indeed, we have to prove that the inequality is true only when $a > 0$, $b > 0$ (since if one of them is zero, it is obviously verified). Considering the real mapping on $(0, \infty)$, $f(x) = x^p/p - x$, this has its minimum $-1/q$ (in $x = 1$), thus

$$\frac{x^p}{p} - x \geq \frac{-1}{q}, \quad \forall x > 0$$

or, equivalently,

$$\frac{x^p}{p} + \frac{1}{q} \geq x, \quad \forall x > 0$$

If in the above inequality one inserts $x = ab^{1-q}$ and one multiplies by b^q, the inequality is proven.

(2) (`Hölder's inequality`) Let p, $q > 1$ such that $1/p + 1/q = 1$ and ξ_i, η_i $(i = 1, 2, ..., n)$ in $\mathbb{K}$. Then,

$$\sum_{i=1}^{n} |\xi_i \eta_i| \leq \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}}$$

Suppose first that

$$\left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \neq 0 \text{ and } \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}} \neq 0$$

Then, in the previous result (1), we may set

$$a = \frac{|\xi_i|}{\left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}}}, \quad b = \frac{|\eta_i|}{\left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}}}$$

and, we obtain

$$\frac{|\xi_i| \cdot |\eta_i|}{\left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}}} \leq \frac{|\xi_i|^p}{p\left(\sum_{i=1}^{n} |\xi_i|^p\right)} + \frac{|\eta_i|^q}{q\left(\sum_{i=1}^{n} |\eta_i|^q\right)}, \; i = 1, 2, ..., n$$

Summarizing, it follows that

$$\sum_{i=1}^{n} |\xi_i| \cdot |\eta_i| \leq \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}} \left(\frac{1}{p} + \frac{1}{q}\right) = \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |\eta_i|^q\right)^{\frac{1}{q}}$$

If

$$\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}=0 \text{ or } \left(\sum_{i=1}^{n}|\eta_i|^q\right)^{\frac{1}{q}}=0,$$

everything is clear, since the inequality becomes $0 \le 0$.

(3) (**Minkowski's inequality**) Let $p \in [1,\infty)$ and ξ_i, η_i $(i = 1, 2, ..., n)$ in $\mathbb{K}$. Then,

$$\left(\sum_{i=1}^{n}|\xi_i+\eta_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}|\eta_i|^p\right)^{\frac{1}{p}}$$

The inequality is immediate if $p = 1$ or $\sum_{i=1}^{n}|\xi_i|^p = 0$ or $\sum_{i=1}^{n}|\eta_i|^p = 0$. It is enough to prove the inequality for ξ_i, $\eta_i \ge 0$ and at least one of the ξ_i, respectively η_i is not zero. U ing the Hölder's inequality, we have

$$\sum_{i=1}^{n}(\xi_i+\eta_i)^p = \sum_{i=1}^{n}(\xi_i+\eta_i)^{p-1}\xi_i + \sum_{i=1}^{n}(\xi_i+\eta_i)^{p-1}\eta_i \le$$

$$\le \left(\sum_{i=1}^{n}(\xi_i+\eta_i)^p\right)^{\frac{p-1}{p}} \left(\sum_{i=1}^{n}\xi_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}(\xi_i+\eta_i)^p\right)^{\frac{p-1}{p}} \left(\sum_{i=1}^{n}\eta_i^p\right)^{\frac{1}{p}} =$$

$$= \left(\sum_{i=1}^{n}(\xi_i+\eta_i)^p\right)^{\frac{p-1}{p}} \left[\left(\sum_{i=1}^{n}\xi_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}\eta_i^p\right)^{\frac{1}{p}}\right]$$

thus, dividing by $\left(\sum_{i=1}^{n}(\xi_i+\eta_i)^p\right)^{\frac{p-1}{p}}$, the proof is finished.

Now, let us return to the space $\mathbb{K}^n$ and to the mapping

$$(\xi_1, \xi_2, ..., \xi_n) \longmapsto \|(\xi_1, \xi_2, ..., \xi_n)\|_p.$$

The first two properties of the norm are immediate, and the fact that

$$\|(\xi_1, \xi_2, ..., \xi_n) + (\eta_1, \eta_2, ..., \eta_n)\|_p \le \|(\xi_1, \xi_2, ..., \xi_n)\|_p + \|(\eta_1, \eta_2, ..., \eta_n)\|_p$$

follows from Minkowski's inequality.

Moreover, $(\mathbb{K}^n, \|\cdot\|_p)$ is a Banach space. Take an arbitrary Cauchy sequence $(x_l)_l$ in $\mathbb{K}^n$, $x_l = (\xi_1^{(l)}, \xi_2^{(l)}, ..., \xi_n^{(l)})$. For any fixed natural k,

$$|\xi_k^{(l)} - \xi_k^{(l)}| \le \|x_l - x_m\|_p \longrightarrow 0$$

as $l, m \to \infty$, so, $(\xi_k^{(l)})_l$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$.is complete, for each k there is a number ξ_k, with $\xi_k = \lim_l \xi_k^{(l)}$. We set $x = (\xi_1, \xi_2, ..., \xi_n) \in \mathbb{K}$ and since

$$\|x_l - x\|_p = \left(\sum_{k=1}^{n} |\xi_k^{(l)} - \xi_k|^p \right)^{\frac{1}{p}}$$

it follows that $\|x_l - x\|_p \longrightarrow 0$ as $l \to \infty$.

Taking into account that any finite dimensional vector space X is isomorphic to $\mathbb{K}^n$ (the isomorphism is $x \longmapsto (\xi_1, \xi_2, ..., \xi_n)$, where $x = \sum_{k=1}^{n} \xi_k e_k$, and $\{e_k\}_{1 \le k \le n}$ is a fixed basis of the n-dimensional vector space X) we may conclude that $(X, \|\cdot\|_p)$ is Banach, where

$$\| \sum_{k=1}^{n} \xi_k e_k \|_p = (\sum_{k=1}^{n} |\xi_k|^p)^{\frac{1}{p}}$$

2. Bounded numerical sequences spaces. Let $l_{\mathbb{K}}^{\infty}$ the set of bounded sequences $x = (\xi_n)_{n \in \mathbb{N}}$ in $\mathbb{K}$. With the usual sum and product for sequences, $l_{\mathbb{K}}^{\infty}$ is a vector space over $\mathbb{K}$. We define on $l_{\mathbb{K}}^{\infty}$ the mapping $x \longmapsto \|x\|$,

$$\|x\| = \sup_{n \in \mathbb{N}} |\xi_n|$$

which, clearly is a norm on $l_{\mathbb{K}}^{\infty}$. Let us denote by $c_{\mathbb{K}}$ the linear subspace of $l_{\mathbb{K}}^{\infty}$ containing all convergent sequences, $c_{\mathbb{K}}^{o}$ the linear subspace of $l_{\mathbb{K}}^{\infty}$ of sequences coverging to zero (which is still a linear subspace of $c_{\mathbb{K}}$), s_{oo} the linear subspace of $l_{\mathbb{K}}^{\infty}$ of all sequences which are zero, except a finite number of their terms (which, clearly is a subspace of $c_{\mathbb{K}}^{o}$ and $c_{\mathbb{K}}$). Then $(l_{\mathbb{K}}^{\infty}, \|\cdot\|)$, $(c_{\mathbb{K}}, \|\cdot\|)$, $(c_{\mathbb{K}}^{o}, \|\cdot\|)$ are Banach spaces (consequently, $c_{\mathbb{K}}$ is a closed subspace of $l_{\mathbb{K}}^{\infty}$ and $c_{\mathbb{K}}^{o}$ is a closed subspace of $c_{\mathbb{K}}$). The normed space $(s_{oo}, \|\cdot\|)$ is not Banach, and it is dense in $c_{\mathbb{K}}^{o}$, $\overline{s_{oo}} = c_{\mathbb{K}}^{o}$.

Let us prove first that $(l_{\mathbb{K}}^{\infty}, \|\cdot\|)$ is Banach. Let $(x_n)_n$ be a Cauchy sequence in $(l_{\mathbb{K}}^{\infty}, \|\cdot\|)$, $x_n = (\xi_k^{(n)})_{k \ge 1}$. Fix $k \in \mathbb{N}$; then

$$|\xi_k^{(n)} - \xi_k^{(m)}| \le \|x_n - x_m\| \longrightarrow 0,$$

as n, $m \to \infty$, so $(\xi_k^{(n)})_{n \ge 1}$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, for each $k \in \mathbb{N}$, there is a (unique) number ξ_k, with

$$\xi_k = \lim_{n \to \infty} \xi_k^{(n)}$$

Denoting by x the sequence $(\xi_k)_k$ we will show that $x \in l^\infty_{\mathbf{K}}$. Clearly, as each Cauchy sequence is bounded, there is a positive $\alpha > 0$ such that $\|x_n\|_\infty < \alpha$, $\forall n \in \mathbb{N}$. Then,

$$|\xi_k^{(n)}| \leq \alpha, \ \forall n \in \mathbb{N}, \quad \forall k \in \mathbb{N}$$

Thus,

$$|\xi_k| = \lim_n |\xi_k^{(n)}| \leq \alpha, \ \forall k \in \mathbb{N},$$

that is, $x \in l^\infty_{\mathbf{K}}$. It remains to show that $\|x_n - x\| \longrightarrow 0$, as $n \to \infty$. Fix $\varepsilon > 0$, and pick n_ε so $n, m \geq n_\varepsilon$ implies $\|x_n - x_m\| < \varepsilon/2$. Then,

$$|\xi_k^{(n)} - \xi_k^{(m)}| < \frac{\varepsilon}{2}, \quad \forall k \in \mathbb{N}$$

Fix k arbitrary and fix also $n \geq n(\varepsilon)$, in the above inequality. As $m \to \infty$, we have

$$|\xi_k^{(n)} - \xi_k| \leq \frac{\varepsilon}{2}, \forall k \in \mathbb{N} \ , \quad \forall n \geq n(\varepsilon)$$

Consequently,

$$\|x_n - x\| = \sup_k |\xi_k^{(n)} - \xi_k| \leq \frac{\varepsilon}{2} < \varepsilon, \quad \forall n \geq n(\varepsilon)$$

In order to prove that $(c_{\mathbf{K}}, \|\cdot\|)$ (respectively $(c^o_{\mathbf{K}}, \|\cdot\|)$) is Banach we shall show that $c_{\mathbf{K}}$, (respectively $c^o_{\mathbf{K}}$) is a closed subspace of $(m_{\mathbf{K}}, \|\cdot\|)$. Consequently, let us start with $x = (\xi_k)_k \in \overline{c_{\mathbf{K}}}$ and prove that x is Cauchy in $\mathbb{K}$, thus convergent. Pick the sequence $(x_n)_n \subset c_{\mathbf{K}}$, where $x_n \longrightarrow x$. It follows, given $\varepsilon > 0$ there is $n(\varepsilon) \in \mathbb{N}$ so that $\forall\ n \geq n(\varepsilon)$ and $\forall k \in \mathbb{N}$

$$|\xi_k^{(n)} - \xi_k| < \frac{\varepsilon}{3}$$

As $x_{n(\varepsilon)}$ is Cauchy, there is $k(\varepsilon) \in \mathbb{N}$ as $\forall k,\ l \geq k(\varepsilon)$, $|\xi_k^{(n(\varepsilon))} - \xi_l^{(n(\varepsilon))}| < \varepsilon/3$. Then, for k, $l \geq k(\varepsilon)$ we have

$$|\xi_k - \xi_l| \leq |\xi_k - \xi_k^{(n(\varepsilon))}| + |\xi_k^{(n(\varepsilon))} - \xi_l^{(n(\varepsilon))}| + |\xi_l^{(n(\varepsilon))} - \xi_l| < \varepsilon$$

Similarly, we prove that $\overline{c^o_{\mathbf{K}}} = c^o_{\mathbf{K}}$. Given $x = (\xi_k)_k \in \overline{c^o_{\mathbf{K}}}$, we pick a sequence $(x_n)_n \subset c^o_{\mathbf{K}}$, where $x_n \longrightarrow x$. It follows, given $\varepsilon > 0$ there is $n(\varepsilon) \in \mathbb{N}$ so that $\forall\ n \geq n(\varepsilon)$ and $\forall k \in \mathbb{N}$,

$$|\xi_k^{(n)} - \xi_k| < \frac{\varepsilon}{2}$$

As $x_{n(\varepsilon)}$ is convergent to zero, there is $k(\varepsilon) \in \mathbb{N}$ such that for each $k \geq k(\varepsilon)$,

$$|\xi_k^{(n(\varepsilon))}| < \varepsilon/2$$

Then, for $k \geq k(\varepsilon)$ we have

$$|\xi_k| \leq |\xi_k - \xi_k^{(n(\varepsilon))}| + |\xi_k^{(n(\varepsilon))}| < \varepsilon.$$

thus, the sequence $x = (\xi_k)_k$ converges to zero.

Further we find a Cauchy sequence in s_{00} that does not converge in s_{00}. Let us take the sequence $x_n = (\xi_k^{(n)})_k$ where

$$\xi_k^{(n)} = \begin{cases} \frac{1}{2^k} & k \leq n \\ 0 & k > n \end{cases}$$

and notice that for arbitrary $m > n$,

$$x_m - x_n = (0, 0, \dots, 0, 1/2^{n+1}, \dots, 1/2^m, 0 \dots), \text{ so } \|x_m - x_n\| = 1/2^{n+1}$$

thus $(x_n)_n$ is Cauchy in s_{00}. Suppose that $\|x_n - x\| \longrightarrow 0$, for some $x = (\xi_1, \dots, \xi_k, 0, \dots)$. Then, for each $n > k$

$$x_n - x = (\frac{1}{2} - \xi_1, \frac{1}{2^2} - \xi_2, \dots, \frac{1}{2^k} - \xi_k, \frac{1}{2^{k+1}}, \dots, \frac{1}{2^n}, 0, \dots)$$

so $\|x_n - x\| \geq 1/2^{k+1}$, and as $n \to \infty$, $\lim_n \|x_n - x\| \geq 1/2^{k+1}$ (one contradicts $\|x_n - x\| \longrightarrow 0$, as $n \to \infty$).

Finally, we have to prove that $\bar{s}_{00} = c_{\mathbb{K}}^0$. Let x be in $c_{\mathbb{K}}^0$, $x = (\xi_k)_k$. Given $\varepsilon > 0$, it is enough to pick $\tilde{x} \in s_{00}$, so that $\|x - \tilde{x}\| < \varepsilon$. As $\lim_k \xi_k = 0$ we find $k(\varepsilon)$ so that $k \geq k(\varepsilon)$, $|\xi_k| < \varepsilon/2$. Set $\tilde{x} = (\xi_1, \dots, \xi_{k(\varepsilon)}, 0, \dots)$ (that belongs to s_{00}). Then

$$\|x - \tilde{x}\| = \|(0, \dots, 0, \xi_{k(\varepsilon)+1}, \dots, \xi_k \dots)\| = \sup_{k \geq k(\varepsilon)+1} |\xi_k| \leq \frac{\varepsilon}{2} < \varepsilon.$$

3. $l_{\mathbb{K}}^p$ spaces. Let $p \in [1, \infty)$ be. Further we will consider the set of all p absolutely summable numerical sequences, i.e. $l_{\mathbb{K}}^p = \{(\xi_n)_{n \geq 1} \subset \mathbb{K} | \sum_{n \geq 1} |\xi_n|^p$ is convergent$\}$. Under the usual addition and scalar multiplication $l_{\mathbb{K}}^p$ is a vector space (the fact that the sum of two sequences in $l^p_{\mathbb{K}}$ is still in $l^p_{\mathbb{K}}$ follows from the Minkowski's inequality). The mapping

$$(\xi_n)_{n\geq 1} \longmapsto \left(\sum_{n=1}^{\infty} |\xi_n|^p\right)^{\frac{1}{p}}$$

is a norm on $l^p_{\mathbb{K}}$, denoted by $\|\cdot\|_p$. The first two properties of the norm are immediate, and $\|(\xi_n)_{n\geq 1} + (\eta_n)_{n\geq 1}\|_p \leq \|(\xi_n)_{n\geq 1}\|_p + \|(\eta_n)_{n\geq 1}\|_p$ results from the Minkowski's inequality.

Moreover, $(l^p_{\mathbb{K}}, \|\cdot\|_p)$ is Banach. In order to prove that, we take an arbitrary Cauchy sequence $(x_n)_n$ in $l^p_{\mathbb{K}}$, $x_n = (\xi_k^{(n)})_k$. For any fixed natural k,

$$|\xi_k^{(n)} - \xi_k^{(m)}| \leq \|x_n - x_m\|_p \longrightarrow 0 \text{ as } n, m \to \infty$$

so, $(\xi_k^{(n)})_n$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, for each k there is a number ξ_k, with

$$\xi_k = \lim_n \xi_k^{(n)}$$

We notice first that the numerical sequence $x = (\xi_k)_k$ belongs to $l^p_{\mathbb{K}}$. Indeed, since $(x_n)_n$ is Cauchy, it is also bounded, thus there is a positive $\alpha > 0$, such that

$$\sum_{k=1}^{m} |\xi_k^{(n)}|^p \leq \alpha^p, \quad \forall n, m \in \mathbb{N}$$

For fixed $m \in \mathbb{N}$, as $n \to \infty$, we have

$$\sum_{k=1}^{m} |\xi_k|^p \leq \alpha^p,$$

which proves that $x \in l^p_{\mathbb{K}}$.

Thus, if we can show that $\|x_n - x\|_p \longrightarrow 0$ as $n \to \infty$, we can conclude that $(l^p_{\mathbb{K}}, \|\cdot\|_p)$ is Banach. Given $\varepsilon > 0$, we find n_ε so $n, m \geq n_\varepsilon$ implies $\|x_n - x_m\|_p < \varepsilon/2$. Then, for each arbitrary natural l, we have

$$\sum_{k=1}^{l} |\xi_k^{(n)} - \xi_k^{(m)}|^p < \left(\frac{\varepsilon}{2}\right)^p$$

and, as $m \to \infty$,

$$\sum_{k=1}^{l} |\xi_k^{(n)} - \xi_k|^p < \left(\frac{\varepsilon}{2}\right)^p$$

From this inequality, as $l \to \infty$, it results

$$\sum_{k=1}^{\infty} |\xi_k^{(n)} - \xi_k|^p \le \left(\frac{\varepsilon}{2}\right)^p$$

which involves $\|x_n - x\|_p < \varepsilon$ for each $n \ge n_\varepsilon$.

4. Bounded functions spaces. Given T an arbitrary nonempty set, let $\mathcal{B}_{\mathbb{K}}(T)$ be the bounded functions $x : T \longmapsto \mathbb{K}$. Let us define on this vector space the real-valued mapping $x \longmapsto \|x\|_\infty$, where

$$\|x\|_\infty = \sup_{t \in [a,b]} |x(t)|$$

It is immediate that this mapping is a norm on $\mathcal{B}_{\mathbb{K}}(T)$. We claim that $(\mathcal{B}_{\mathbb{K}}(T), \|\cdot\|_\infty)$ is Banach. Indeed, let $(x_n)_n$ be a Cauchy sequence. Then, for any fixed $t \in [a,b]$,

$$|x_n(t) - x_m(t)| \le \|x_n - x_m\|_\infty \longrightarrow 0$$

as $n, m \to \infty$, so $(x_n(t))_n$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, for each t, there is a (unique) number, $x(t)$ with $x_n(t) \longrightarrow x(t)$,as $n \to \infty$. Clearly, as each Cauchy sequence is bounded, there is a positive $\alpha > 0$ such that $\|x_n\|_\infty < \alpha$, $\forall n \in \mathbb{N}$. Then,

$$|x(t)| = \lim_n |x_n(t)| \le \alpha, \quad \forall t \subset T,$$

thus the function x on $[a,b]$ defined by

$$x(t) = \lim_n x_n(t)$$

is bounded.

Let us prove that, given $\varepsilon > 0$, we find n_ε so $n \ge n_\varepsilon$ implies $\|x_n - x\|_\infty < \varepsilon$, that is $(x_n)_n$ converges to x in $(\mathcal{B}_{\mathbb{K}}(T), \|\cdot\|_\infty)$. Fix $\varepsilon > 0$, and pick n_ε so $n, m \ge n_\varepsilon$ implies $\|x_n - x_m\|_\infty < \varepsilon/3$. Then,

$$\|x - x_{n_\varepsilon}\|_\infty = \sup_{t \in [a,b]} \lim_{n \to \infty} |x_n(t) - x_{n_\varepsilon}(t)| \le$$

$$\le \sup_{t \in [a,b]} \sup_{n \ge n_\varepsilon} |x_n(t) - x_{n_\varepsilon}(t)| = \sup_{n \ge n_\varepsilon} \|x_n - x_{n_\varepsilon}\|_\infty \le \frac{\varepsilon}{3},$$

and, consequently, if $n \geq n_\varepsilon$,

$$\|x - x_n\|_\infty \leq \|x - x_{n_\varepsilon}\|_\infty + \|x_n - x_{n_\varepsilon}\|_\infty \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon$$

We notice that $(\mathcal{B}_{\mathbb{K}}(\mathbb{N}), \|\cdot\|_\infty) = (l^\infty_{\mathbb{K}}, \|\cdot\|)$ (Example 2).

5. **Continuous functions spaces.** Let $\mathcal{C}_{\mathbb{K}}[a,b]$ be the continuous functions on $[a,b]$ (to $\mathbb{K}$) with the norm

$$\|x\|_\infty = \sup_{t \in [a,b]} |x(t)|$$

Then, $(\mathcal{C}_{\mathbb{K}}[a,b], \|\cdot\|_\infty)$ is Banach. In order to prove that, it is enough to show that the subspace $\mathcal{C}_{\mathbb{K}}[a,b]$ of $(\mathcal{B}_{\mathbb{K}}[a,b], \|\cdot\|_\infty)$ is closed. Then let $(x_n)_n$ be a sequence in $\mathcal{C}_{\mathbb{K}}[a,b]$ which converges to x. If we can show that the function x is continuous on $[a,b]$, we can conclude that $(x_n)_n$ converges to x in $(\mathcal{C}_{\mathbb{K}}[a,b], \|\cdot\|_\infty)$. We are thus left proving that x is continuous at each fixed $t \in [a,b]$. Given $\varepsilon > 0$ we want to find $\delta > 0$ so $|s-t| < \delta$ implies $|x(s) - x(t)| < \varepsilon$. Pick n_ε so that $\|x_{n_\varepsilon} - x\|_\infty < \varepsilon/3$. Since x_{n_ε} is continuous at t, there is $\delta > 0$ so that $|s-t| < \delta$ implies $|x_{n_\varepsilon}(s) - x_{n_\varepsilon}(t)| < \varepsilon/3$. Then $|s-t| < \delta$ implies

$$|x(s)-x(t)| \leq |x(s)-x_{n_\varepsilon}(s)|+|x_{n_\varepsilon}(s)-x_{n_\varepsilon}(t)|+|x_{n_\varepsilon}(t)-x(t)| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon$$

thus x is continuous. Usually the norm on $\mathcal{C}_{\mathbb{K}}[a,b]$, will be denoted by $\|\cdot\|$ instead of $\|\cdot\|_\infty$.

6. $L^p_{\mathbb{K}}$ **spaces.** Let $(X, \mathcal{X}, m)$ be a measure space and $p \in [1, \infty)$. (For more about measure spaces one can see Appendix C.) The space $L^p_{\mathbb{K}}(X)$ consists of all m-equivalence classes of $\mathcal{X}$-measurable $\mathbb{K}$ -functions x for which $|x|^p$ is integrable (has finite integral with respect to m over X). Two functions are m-equivalent if they are equal m-almost everywhere (we will denote the equivalence class of x by the same symbol, x). We set

$$\|x\|_p = \left(\int |x|^p \,\mathrm{d}\, m\right)^{\frac{1}{p}}$$

We shall show that $(L^p_{\mathbb{K}}(X), \|\cdot\|_\infty)$ is a Banach space. It is understood that the vector operations between the elements of $L^p_{\mathbb{K}}(X)$ are defined pointwise: the sum of the equivalence classes containing x and y is the equivalence class containing $x + y$ and similarly for the product αx. The fact that if $x, y \in L^p_{\mathbf{K}}(X)$ the sum is still in $L^p_{\mathbf{K}}(X)$ follows from

$$|x + y|^p \leq [2 \ \sup\{|x|, |y|\}]^p \leq 2^p \left(|x|^p + |y|^p\right)$$

We have to notice that in the particular case where m is the counting measure on all subsets of $\mathbb{N}$, the $L^p_{\mathbf{K}}$-spaces can be identified with the sequence spaces $l^p_{\mathbf{K}}$ (Example 3). In this case each equivalence class contains one element.

In order to establish that $\|\cdot\|_\infty$ yields a norm on $L^p_{\mathbf{K}}$ we shall need the Hölder and Minkowski inequalities for functions. Recall us, that for p, $q \geq 1$ such that $1/p + 1/q = 1$ we have

$$|ab| \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}, \quad \forall a, b \in \mathbb{K}$$

(Example 1(1)). Suppose that $x \in L^p_{\mathbf{K}}$ and $y \in L^q_{\mathbf{K}}$, and that $\|x\|_p \neq 0$ and $\|y\|_q \neq 0$. The product xy is measurable, and the above inequality with $a = |x(t)|/\|x\|_p$ and $b = |y(t)|/\|y\|_q$ implies that

$$\frac{|x(t)y(t)|}{\|x\|_p\|y\|_q} \leq \frac{|x(t)|^p}{p\|x\|_p} + \frac{|y(t)|^q}{q\|y\|_q}$$

Moreover, on integrating we obtain the Hölder's inequality:

$$\int |xy| \,\mathrm{d}\, m \leq (\int |x|^p \,\mathrm{d}\, m)^{\frac{1}{p}} (\int |y|^q \,\mathrm{d}\, m)^{\frac{1}{q}}$$

which, in the particular case of $p = q = 1/2$, involves the Cauchy-Schwarz inequality:

$$|\int xy \,\mathrm{d}\, m| \leq (\int |x|^2 \,\mathrm{d}\, m)^{\frac{1}{2}} (\int |y|^2 \,\mathrm{d}\, m)^{\frac{1}{2}}$$

Now, we will obtain the Minkowski's inequality,

$$(\int |x + y|^p \,\mathrm{d}\, m)^{\frac{1}{p}} \leq (\int |x|^p \,\mathrm{d}\, m)^{\frac{1}{p}} + (\int |y|^p \,\mathrm{d}\, m)^{\frac{1}{p}}$$

(that is $\|x + y\|_p \leq \|x\|_p + \|y\|_p$).

We have already seen that $x + y \in L^p_{\mathbf{K}}$. Moreover,

$$|x + y|^p = |x + y| \cdot |x + y|^{p-1} \leq |x| \cdot |x + y|^{p-1} + |y| \cdot |x + y|^{p-1}$$

Since $x + y \in L^p_{\mathbf{K}}$, then $x + y \in L^1_{\mathbf{K}}$; since $p = (p-1)q$ it follows that $|x+y|^{p-1} \in L^q_{\mathbf{K}}$. Hence we can apply Hölder's inequality to infer that

$$\int |x| \cdot |x+y|^{p-1} \,\mathrm{d}\,m \leq \|x\|_p \left[\int |x+y|^{(p-1)q} \,\mathrm{d}\,m\right]^{\frac{1}{q}} = \|x\|_p \|x+y\|_p^{\frac{p}{q}}$$

If we treat the second term on the right similarly, we have

$$\int |y| \cdot |x+y|^{p-1} \,\mathrm{d}\,m \leq \|y\|_p \|x+y\|_p^{\frac{p}{q}}$$

thus, finally

$$\int |x+y|^p \,\mathrm{d}\,m \leq \int |x| \cdot |x+y|^{p-1} \,\mathrm{d}\,m + \int |y| \cdot |x+y|^{p-1} \,\mathrm{d}\,m \leq$$

$$\leq \|x\|_p \|x+y\|_p^{\frac{p}{q}} + \|y\|_p \|x+y\|_p^{\frac{p}{q}} = (\|x\|_p + \|y\|_p) \|x+y\|_p^{\frac{p}{q}}$$

If $\|x+y\|_p = 0$, Minkowski's inequality is trivial. If $\|x+y\|_p \neq 0$, we can divide the above inequality by $\|x+y\|_p^{p/q}$; since $p - p/q = 1$ we obtain Minkowski's inequality.

Next, we shall prove that the normed space $(L^p_{\mathbf{K}}(X), \|\cdot\|_\infty)$ is Banach. This result is known as Riesz- Fisher theorem. We will show that each absolutely convergent series of elements in $L^p_{\mathbf{K}}(X)$ is convergent to an element in $L^p_{\mathbf{K}}(X)$ (Proposition 2.1.2). Let $(x_n)_n$ be a sequence in $L^p_{\mathbf{K}}(X)$ such that $\sum_{n=1}^\infty \|x_n\|_p = \alpha < \infty$. We define the sequence of functions $(y_n)_n$ by

$$y_n(t) = \sum_{k=1}^n |x_k(t)|$$

By Minkowski's inequality, we have

$$\|y_n\|_p \leq \sum_{k=1}^n \|x_n\|_p$$

so, it follows that

$$\int (y_n)^p \,\mathrm{d}\,m \leq \alpha^p$$

Given $t \in X$, the increasing sequence $(y_n(t))_n \subset \mathbb{R}_+ \cup \{\infty\}$ has a limit in $\mathbb{R}_+ \cup \{\infty\}$, denoted further by $y(t)$. Since

$$\sum_{n=1}^\infty |x_n(t)| = y(t)$$

the function on X to $\mathbb{R}_+ \cup \{\infty\}$, $t \longmapsto y(t)$ is measurable. By Fatou's lemma, we obtain

$$\int y^p \, \mathrm{d}\, m \leq \alpha^p$$

so y^p is m-integrable, and consequently y is m-finite almost everywhere. Since the functions series $\sum_{n=1}^{\infty} x_n(t)$ is m-almost everywhere absolutely convergent, it results that the series $\sum_{n=1}^{\infty} x_n(t)$ is m-almost everywhere convergent; denote its sum (when the series is convergent) by $s(t)$. Let us denote by A the set of all $t \in X$ with $y(t) = \infty$ and define a function on X as follows: $s(t) = 0$ if $t \in A$, and, otherwise $s(t) = \sum_{n=1}^{\infty} x_n(t)$. Hence we obtain a function s from X to $\mathbb{R}$ with the property that m-almost everywhere, $s(t) = \lim_n \sum_{k=1}^{n} x_k(t)$. Clearly, since s is μ-almost everywhere the limit of a sequence of measurable functions is itself measurable. By the inequality

$$|\sum_{k=1}^{n} x_k(t)| \leq y(t)$$

as $n \to \infty$, we obtain that $|s(t)| \leq y(t)$, so we can conclude $s \in L^p_{\mathbb{K}}(X)$.

By taking into account that

$$|\sum_{k=1}^{n} x_k(t) - s(t)| \leq 2^p [g(t)]^p,$$

the function $2^p[g(t)]^p$ is integrable and that $|\sum_{k=1}^{n} x_k(t) - s(t)| \longrightarrow 0$ m-almost everywhere, one can use the Dominated convergence theorem to infer

$$\int |\sum_{k=1}^{n} x_k(t) - s(t)|^p \, \mathrm{d}\, m \longrightarrow 0,$$

so, equivalently, $\| \sum_{k=1}^{n} x_k - s\|_p \longrightarrow 0$, as $n \to \infty$. This proves that the series $\sum_{n\geq 1} x_n$ converges to s in the norm of $L^p_{\mathbb{K}}(X)$.

2.3 Finite dimensional normed spaces

2.3.1 The equivalence of the norms

The next result shows that all norms on a finite dimensional vector space are equivalent, consequently, the norm topologies coincide.

Theorem 2.3.1 *Let X be a vector space over the field $\mathbb{K}$ of finite dimension, $\dim_{\mathbb{K}} X = n < \infty$ and $\{e_k\}_{1\le k\le n}$ an algebraic basis of X. Then,*
1) The real-valued mapping on X, $x \longmapsto \|x\|_2$, where for $x = \sum_{j=1}^{n} \xi_j e_j$,

$$\|x\|_2 = \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{\frac{1}{2}}$$

is a norm on X.
2) Any other norm on X is equivalent to $\|\cdot\|_2$.

Proof. 1) It follows immediately from Minkowski's inequality, (2.2, Example 1, (3)).

2) Let $\|\cdot\|$ be an other norm on X. Then, for an arbitrary $x = \sum_{j=1}^{n} \xi_j e_j$, we have with the properties of the norm and Hölder' s inequality,

$$\|x\| = \|\sum_{j=1}^{n} \xi_j e_j\| \le \sum_{j=1}^{n} |\xi_j| \cdot \|e_j\| \le$$

$$\le \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{j=1}^{n} \|e_j\|^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{n} \|e_j\|^2\right)^{\frac{1}{2}} \|x\|_2$$

Thus, we have only to see that there exists $\beta > 0$ such that $\|x\|_2 \le \beta \|x\|$, $\forall x \in X$. Suppose otherwise, so, for each $\beta > 0$, there exists $x_\beta \in X$, with $\|x_\beta\|_2 > \beta\|x_\beta\|$; in particular, for every positive integer m, $\exists x_m \in X$, with $\|x_m\|_2 > m\|x_m\|$. Let us consider the sequence $(y_m)_{m\ge 1}$, where

$$y_m = \frac{1}{\|x_m\|_2}\, x_m$$

We remark that $\|y_m\|_2 = 1$ and $\|y_m\| < 1/m$, $\forall m \ge 1$. Every y_m can be written as

$$y_m = \sum_{j=1}^{n} \eta_j^{(m)} e_j, \quad \forall m \in \mathbb{N}$$

For an arbitrary fixed $j \in \{1, 2, ..., n\}$, from

$$|\eta_j^{(m)}| \le \left(\sum_{j=1}^{n} |\eta_j^{(m)}|^2\right)^{\frac{1}{2}} = \|y_m\|_2 = 1, \forall m \in \mathbb{N}$$

it follows that the numerical sequence $(\eta_j^{(m)})_{m\geq 1}$ is bounded. As every bounded numerical sequence has a convergent subsequence, we may conclude, by a diagonal procedure, that for $\forall j \in \{1, 2, ..., n\}$, there exists a convergent subsequence of $(\eta_j^{(m)})_{m\geq 1}$, denoted for simplicity, $(\gamma_j^{(m)})_m$, $\gamma_j^{(m)} \longrightarrow \gamma_j$. Further, we define in $\mathbb{K}$ the sequence $(z_m)_{m\geq 1}$,

$$z_m = \sum_{j=1}^{n} \gamma_j^{(m)} e_j,$$

and the element

$$z = \sum_{j=1}^{n} \gamma_j e_j$$

Clearly, $\|z\|_2 = 1$, since $(z_m)_{m\geq 1}$ is a subsequence of $(y_m)_{m\geq 1}$, thus

$$1 = \lim_m \|z_m\|_2 = \lim_m \left(\sum_{j=1}^{n} |\gamma_j^{(m)}|^2\right)^{\frac{1}{2}} =$$

$$= \left(\sum_{j=1}^{n} \lim_m |\gamma_j^{(m)}|^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{n} |\gamma_j|^2\right)^{\frac{1}{2}} = \|z\|_2$$

On the other hand, for an arbitrary $m \in \mathbb{N}$,

$$\|z\| \leq \|z - z_m\| + \|z_m\| = \|\sum_{j=1}^{n} (\gamma_j - \gamma_j^{(m)}) e_j\| + \|z_m\| \leq$$

$$\leq \sum_{j=1}^{n} |\gamma_j - \gamma_j^{(m)}| \cdot \|e_j\| + \|z_m\| \leq \sum_{j=1}^{n} |\gamma_j - \gamma_j^{(m)}| \|e_j\| + \frac{1}{m}$$

As $m \to \infty$, one obtains that $\|z\| = 0$, so $z = 0$, which contradicts $\|z\|_2 = 1$.

It follows that there exists $\beta > 0$ such that $\|x\|_2 \leq \beta \|x\|$, $\forall x \in X$, therefore every norm on X is equivalent to $\|\cdot\|_2$.

Corollary 2.3.1 *Every finite dimensional normed space $(X, \|\cdot\|)$ is a Banach space.*

Proof. Let $\{e_k\}_{1\leq k\leq n}$ an algebraic basis of X. Since the norm $\|\cdot\|$ is equivalent to $\|\cdot\|_2$, it is enough to prove that the space is complete under the metric defined by the norm $\|\cdot\|_2$. Take an arbitrary Cauchy sequence, $(x_m)_m$; each x_m is written uniquely as $\sum_{j=1}^{n} \xi_j^{(m)} e_j$. For every $j \in \{1, 2, ..., n\}$, the sequence $(\xi_j^{(m)})_m$ is Cauchy, as it results from

$$|\xi_j^{(k)} - \xi_j^{(m)}| \leq \|x_k - x_m\|_2$$

Then, as $(\mathbb{K}, |\cdot|)$ is complete, there exists $\xi_j \in \mathbb{K}$, $\xi_j = \lim_m \xi_j^{(m)}$. Let us define $x = \sum_{j=1}^{n} \xi_j e_j$ and prove that $x_m \longrightarrow x$. Since $(x_m)_m$ is Cauchy, it results for any $\varepsilon > 0$, $\exists m_\varepsilon \in \mathbb{N}$, such that $\forall k, m \geq m_\varepsilon$, $\|x_k - x_m\|_2 =$ $= (\sum_{j=1}^{n} |\xi_j^{(k)} - \xi_j^{(m)}|^2)^{1/2} < \varepsilon/2$. If we fix an arbitrary $m \geq m_\varepsilon$ and we pass to the limit with respect to k in the previous inequality, we get that

$$(\sum_{j=1}^{n} |\xi_j^{(m)} - \xi_j|^2)^{1/2} \leq \varepsilon/2 < \varepsilon, \quad \forall m \geq m_\varepsilon,$$

so, $\|x_m - x\|_2 < \varepsilon$, $\forall m \geq m_\varepsilon$. We have shown that $(x_m)_m$ is convergent.

Corollary 2.3.2 *Each finite dimensional subspace Y of a normed space $(X, \|\cdot\|)$ is closed.*

Proof. By Corollary 2.3.1, the normed space $(Y, \|\cdot\|)$ is Banach. Let y be in the closure of Y, so there exists a sequence $(y_n)_n \subset Y$, converging to y. The sequence $(y_n)_n$ is Cauchy in the complete space Y, thus there exists $z \in Y$, the limit of the sequence $(y_n)_n$. By the uniqueness of the limit (in Hausdorff spaces) it results that $z = y \in Y$.

2.3.2 Compact sets in finite dimensional normed spaces

Here we will see that in finite dimensional vector spaces the compact sets are exactly the bounded closed sets.

Lemma 2.3.1 *Let $(X, \|\cdot\|)$ be a normed space and Y a closed proper subspace of X $(Y \subsetneq X)$. Then, for every $\varepsilon > 0$ there is an element x_ε in $X \setminus Y$ such that $\|x_\varepsilon\| = 1$ and $\|x_\varepsilon - y\| > 1 - \varepsilon$, $\forall y \in Y$.*

Proof. Take an $\varepsilon > 0$, $(\varepsilon < 1)$. Since $Y \subsetneq X$, there exists $x \in X \setminus Y$. Denote by d the distance from x to Y, which, since Y is closed, is strictly positive, thus, $d = \inf_{y \in Y} \|x - y\| > 0$. As $d < d(1+\varepsilon)$, there is an element $y_\varepsilon \in Y$ such that $d \le \|x - y_\varepsilon\| < d(1+\varepsilon)$. Defining x_ε as

$$x_\varepsilon = \frac{1}{\|x - y_\varepsilon\|} \cdot (x - y_\varepsilon)$$

we claim that $x_\varepsilon \notin X \setminus Y$. Otherwise, if $x_\varepsilon \in X \setminus Y$, it follows that

$$x = x_\varepsilon \|x - y_\varepsilon\| + y_\varepsilon \in Y,$$

which contradicts the choice of $x \in X \setminus Y$.

In addition, $\|x_\varepsilon\| = 1$, and, for an arbitrary $y \in Y$ we have

$$\|x_\varepsilon - y\| = \|\frac{1}{\|x - y_\varepsilon\|} \cdot (x - y_\varepsilon) - y\| =$$

$$= \frac{1}{\|x - y_\varepsilon\|} \cdot \|x - (y_\varepsilon + y\|x - y_\varepsilon\|)\| > \frac{1}{\|x - y_\varepsilon\|} \cdot d > \frac{1}{d(1+\varepsilon)} \cdot d = \frac{1}{1+\varepsilon} > 1 - \varepsilon$$

Theorem 2.3.2 *(**The Riesz theorem**) Let $(X, \|\cdot\|)$ be a normed space over the field $\mathbb{K}$. The following are equivalent:*
(1) X is finite dimensional;
(2) Every bounded closed subset of X is compact.

Proof. (1) $\Rightarrow$ (2) Suppose that the dimension of X is n and consider $\{e_k\}_{1 \le k \le n}$ an arbitrary basis of X. Let A be a bounded closed subset of X and $(x_m)_m$ a sequence of elements in A. Since $\|\cdot\| \sim \|\cdot\|_2$, it follows that there exists $\alpha > 0$ such that $\|x_m\|_2 \le \alpha$, $\forall m \in \mathbb{N}$. Each $x_m = \sum_{i=1}^n \xi_i^{(m)} e_i$, and, from

$$|\xi_i^{(m)}| \le \left(\sum_{i=1}^n |\xi_i^{(m)}|^2 \right)^{\frac{1}{2}} = \|x_m\|_2 \le \alpha, \quad \forall m \in \mathbb{N}$$

it results that the sequence $(\xi_i^{(m)})_m$ is bounded for every $i = 1, 2, ..., n$, therefore it has a convergent subsequence $(i = 1, 2, ..., n)$. Then, by a diagonal procedure we obtain a subsequence $(z_m)_m$ of $(x_m)_m$, where $z_m = \sum_{i=1}^n \gamma_i^{(m)} e_i$

and $\gamma_i^{(m)} \longrightarrow \gamma_i$, as $m \to \infty$, $i = 1, 2, ..., n$. Setting $z = \sum_{i=1}^{n} \gamma_i e_i$, we have that $z_m \longrightarrow z$. Indeed,

$$\|z_m - z\|_2 = \left(\sum_{i=1}^{n} |\gamma_i^{(m)} - \gamma_i|^2 \right)^{\frac{1}{2}} \longrightarrow 0,$$

as $m \to \infty$. Moreover, because A is closed, $z \in A$. Therefore we have proven that A is compact, since we have shown that every sequence of elements of A contains a convergent subsequence.

(2) $\Rightarrow$ (1) Suppose that X is not finite dimensional. Take an arbitrary $x_1 \in X$, $x_1 \neq 0$ and let Y_1 denote the linear subspace spanned by x_1. The subspace Y_1 is finite dimensional; by the Corollary 2.3.2 it results that Y_1 is closed. As X is not finite dimensional, $Y_1 \subsetneq X$. Then, by the previous lemma $\exists x_2 \notin Y_1$, $\|x_2\| = 1$, $\|x_2 - y\| > 1/2$, $\forall y \in Y_1$; in particular, $\|x_2 - x_1\| >$ $> 1/2$. Further, let Y_2 the linear subspace spanned by $\{x_1, x_2\}$. With similar arguments as above, $\exists x_3 \notin Y_2$, $\|x_3\| = 1$, $\|x_3 - y\| > 1/2$, $\forall y \in Y_2$; in particular, $\|x_3 - x_i\| > 1/2$, $i = 1, 2$. Continuing this process, by induction, we obtain a sequence $(x_m)_m$ in X such that $\|x_m\| = 1$ and $\|x_m - x_k\| > 1/2$, $\forall m$, $k \in \mathbb{N}$. It follows that the closed unit ball $\overline{B}(1)$ contains a sequence which has no convergent subsequences, so there exists a bounded closed set $(\overline{B}(1))$ which is not compact (one contradicts 2)).

Corollary 2.3.3 *Let $(X, \|\cdot\|)$ be a normed space over the field $\mathbb{K}$ such that the closed unit ball of X is compact. Then, X is finite dimensional.*

Proof. Let A be a bounded closed subset of X. Then, $\exists \alpha > 0$ such that $\|x\| \leq \alpha$, $\forall x \in A$, or, equivalently, $1/\alpha \ A \subset \overline{B}(1)$. It results that $1/\alpha \ A$ is compact, since it is a closed subset of the compact set $\overline{B}(1)$. As the mapping $x \longmapsto 1/\alpha \cdot x$ is an homeomorphism on X, we have that A is compact too. By the Riesz theorem, it results that X is finite dimensional.

We will end this section with an application, an approximation result in normed space.

Proposition 2.3.1 *Let $(X, \|\cdot\|)$ be a normed space over the field $\mathbb{K}$ and Y a finite dimensional subspace of X. Then, for every $x \in X$, there exists an element $y_x \in Y$ such that*

$$d(x, Y) = \|x - y_x\|,$$

(y_x is called the closest element to x in Y).

Proof. Let us denote by $d = d(x, Y)$ (which is $\inf_{y \in Y} \|x - y\|$). Then, for every $n \in \mathbb{N}$, there exists $y_n \in Y$ such that $d \leq \|x - y_n\| < d + 1/n$. Then, as Y is closed in X (Corollary 2.3.2), it follows that A, the closure of the set $\{y_n \mid n \in \mathbb{N}\}$ is still in Y. In addition, A is bounded. Since Y is finite dimensional, by the Riesz theorem, it follows that A is compact, thus each sequence in A has a convergent subsequence, in particular $(y_n)_n$. Then, there exists a convergent subsequence $(y_{n_k})_k$ of $(y_n)_n$. Let us denote its limit by y_x; clearly, since A is closed $y_x \in A \subset Y$. By

$$d \leq \|x - y_{n_k}\| < d + \frac{1}{n_k}$$

it follows, as $k \to \infty$, that $d = \|x - y_x\|$.

Example. Consider the normed space $(\mathcal{C}_{\mathbb{K}}[a, b], \|\cdot\|_{\infty})$ and the subspace of $\mathcal{C}_{\mathbb{K}}[a, b]$ of all polynomials of degree less than n, $\mathcal{P}_n[a, b]$. Clearly, $\mathcal{P}_n[a, b]$ is finite dimensional (a basis for it is $\{1, t, t^2, ..., t^n\}$). By the above proposition, for every continuous function $x \in \mathcal{C}_{\mathbb{K}}[a, b]$, there exists a polynomial

$$p(t) = \sum_{k=1}^{n} \lambda_k^{(x)} t^k,$$

where $\lambda_k^{(x)} \in \mathbb{K}$, $k = 1, 2, ..., n$, such that

$$\inf_{\lambda_1, \lambda_2, ..., \lambda_n} \|x - \sum_{k=1}^{n} \lambda_k t^k\|_{\infty} = \|x - \sum_{k=1}^{n} \lambda_k^{(x)} t^k\|_{\infty}$$

2.4 Exercises

1. Prove that the closure of every bounded set in a normed space is still bounded.

2. Let X be a normed space and $(x_n)_n$ a sequence in X converging to zero. Then, there is a sequence $(\mu_n)_n$ in $\mathbb{K}$ such that $|\mu_n| \longrightarrow \infty$ and $\mu_n x_n \longrightarrow 0$.

3. Let $(X_j, \|\cdot\|_j)$, $j = 1, 2, ..., m$ be normed spaces and $X = \prod_{j=1}^{m} X_j$ their normed product space ($\|(x_1, x_2, ..., x_m)\| = \max_{j=1,2,...,m} \|x_j\|_j$). Show that the norm of X, is equivalent to

$$\|(x_1, x_2, ..., x_m)\|_p = \left(\sum_{j=1}^{m} |x_j|^p\right)^{\frac{1}{p}}$$

where $p \in [1, \infty)$.

4. A normed space X admits a topological (Schauder) basis if there is a sequence $(x_n)_n \subset X$ (called topological basis of X or Schauder basis) such that each $x \in X$ can be written uniquely as $x = \sum_{n=1}^{\infty} \alpha_n x_n$, where $\alpha_n \in \mathbb{K}$. Show that $(e_n)_n$ (where $e_n = (\delta_k^{(n)})_{k \geq 1}$) is a topological basis in $l_{\mathbf{K}}^p$, $p \in [1, \infty)$.

5. Show that $c_{\mathbf{K}}^o$ and $c_{\mathbf{K}}$ admit topological bases.

6. Show that $\mathcal{C}_{\mathbf{K}}([0,1])$ admits a topological basis.

7. Each normed space admitting a topological basis is separable.

8. Show that $\mathcal{B}_{\mathbf{K}}(T)$ is separable if and only if T is finite. In particular, $l_{\mathbf{K}}^{\infty}$ is not separable.

9. Prove that each linear subspace of a separable normed space is separable.

10. a) Show that the real valued mapping

$$(\xi_1, \xi_2, \ldots, \xi_n) \longmapsto \|(\xi_1, \xi_2, \ldots, \xi_n)\|_{\infty},$$

where

$$\|(\xi_1, \xi_2, \ldots, \xi_n)\|_{\infty} = \max_{1 \leq i \leq n} |\xi_i|$$

is a norm on $\mathbb{K}^n$.

b) The real-valued mapping

$$p \longmapsto \|(\xi_1, \xi_2, \ldots, \xi_n)\|_p$$

is decreasing on $[1, \infty)$ and

$$\inf_{p \in [1,\infty)} \|(\xi_1, \xi_2, \ldots, \xi_n)\|_p = \lim_{p \to \infty} \|(\xi_1, \xi_2, \ldots, \xi_n)\|_p = \|(\xi_1, \xi_2, \ldots, \xi_n)\|_{\infty}$$

11. Show that $l_{\mathbf{K}}^p$ is a linear subspace of $l_{\mathbf{K}}^{\infty}$ which is not closed; its closure in $l_{\mathbf{K}}^{\infty}$ is $c_{\mathbf{K}}^o$.

12. Let $1 \leq p < q \leq \infty$ be. Show that $l_{\mathbf{K}}^p \subset l_{\mathbf{K}}^q$, $\|\cdot\|_p \geq \|\cdot\|_q$ and the subspace $l_{\mathbf{K}}^p$ is not closed in $l_{\mathbf{K}}^q$.

Chapter 3

Bounded operators on Banach spaces

3.1 The normed space $\mathcal{B}(X, Y)$

Next, X and Y are two normed spaces, over the same field $\mathbb{K}$. We use the same symbol, $\|\cdot\|$, for the norm of X and of Y.

Definition. Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be two normed spaces. A mapping $T : X \longrightarrow Y$ is said to be *bounded* if and only if there is $M > 0$ so that

$$\|T(x)\| \leq M\|x\|, \quad \forall x \in X$$

Proposition 3.1.1 *Let T be a linear operator between two normed spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$. The following are equivalent:*
(1) The mapping T is continuous on X;
(2) The mapping T is continuous at zero;
(3) The mapping T is bounded.

Proof. As $(1) \Rightarrow (2)$ is clear, let us begin with the converse implication Take an arbitrary $x_o \in X$. Given $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ such that $\|u\| < \delta_\varepsilon$ implies $\|T(u)\| < \varepsilon$. Setting here $u = x - x_o$ and using the linearity of T one obtains that for given $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ such that $\|x - x_o\| < \delta_\varepsilon$ implies $\|T(x) - T(x_o)\| < \varepsilon$, so T is continuous at x_o.

Suppose further that T is continuous at zero ((2)) and that it is not bounded. Then, for each positive integer m there is $x_m \in X$ so that

$\|T(x_m)\| > m\|x_m\|$. The sequence $(y_m)_m$ defined by

$$y_m = \frac{1}{m\|x_m\|} \cdot x_m$$

converges clearly to zero and $\|T(y_m)\| > 1$, which contradicts the continuity of T at zero. Thus we have seen that (2) $\Rightarrow$ (3). The implication (3) $\Rightarrow$ (1) is obvious.

Notation. The set of all bounded linear operator between two normed spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ will be denoted by $\mathcal{B}(X, Y)$, and if $X = Y$, $\mathcal{B}(X)$. In the particular case of $(Y, \|\cdot\|) = (\mathbb{K}, |\cdot|)$, so the bounded linear operators are bounded linear functionals, the space $\mathcal{B}(X, Y)$ will be denoted by X^* (X^* is called the *dual* of the normed space $(X, \|\cdot\|)$).

Remark. The set $\mathcal{B}(X, Y)$ is clearly a subspace of $\mathcal{L}(X, Y)$.

Proposition 3.1.2 *The real mapping on $\mathcal{B}(X, Y)$, $T \longmapsto \|T\|$, defined by*

$$\|T\| = \sup_{\|x\| \leq 1} \|T(x)\|$$

is a norm on $\mathcal{B}(X, Y)$.

Proof. Since T is bounded, we have that the set (of positive reals) $\{\|T(x)\| \mid \|x\| \leq 1\}$ is upper bounded, so the mapping is well defined. If $\|T\| = 0$, then, for every $x \neq 0$, $T((1/\|x\|)\, x) = 0$, thus $T(x) = 0$. It follows $T = 0$. If T_1, $T_2 \in \mathcal{B}(X, Y)$ and $\|x\| \leq 1$, we have

$$\|(T_1 + T_2)(x)\| = \|T_1(x) + T_2(x)\| \leq \|T_1(x)\| + \|T_2(x)\| \leq \|T_1\| + \|T_2\|$$

thus,

$$\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$$

For arbitrary $\alpha \in \mathbb{K}$ and $T \in \mathcal{B}(X, Y)$, by

$$\|\alpha T\| = \sup_{\|x\| \leq 1} \|(\alpha T)(x)\| = \sup_{\|x\| \leq 1} |\alpha| \cdot \|T(x)\| = |\alpha| \sup_{\|x\| \leq 1} \|T(x)\|$$

it follows that

$$\|\alpha T\| = |\alpha| \cdot \|T\|$$

Remarks. 1. *Next, always the vector space $\mathcal{B}(X,Y)$ will be endowed with the above norm, called the operator norm.*

2. By the definition of the norm on $\mathcal{B}(X,Y)$, it results that for $T \in \mathcal{B}(X,Y)$,

$$\|T(x)\| \leq \|T\| \cdot \|x\|, \quad \forall x \in X$$

Proposition 3.1.3 *Let T be in $\mathcal{B}(X,Y)$. Then*

$$\|T\| = \inf \{M > 0 \mid \|T(x)\| \leq M\|x\|, \ \forall x \in X\} =$$

$$= \sup_{\|x\|=1} \|T(x)\| = \sup_{\|x\|<1} \|T(x)\| = \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|}$$

Proof. Let us denote by

$$\|T\|_1 = \sup_{\|x\|=1} \|T(x)\|, \ \|T\|_2 = \sup_{\|x\|<1} \|T(x)\|,$$

$$\|T\|_3 = \inf \{M > 0 \mid \|T(x)\| \leq M\|x\|, \ \forall x \in X\} \text{ and } \|T\|_4 = \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|}$$

Clearly, $\|T\|_1 \leq \|T\|$ and $\|T\|_2 \leq \|T\|$.

Let x be arbitrary in X with $\|x\| \leq 1$. Since $\|n/(n+1)x\| < 1$ we have that

$$\|T\left(\frac{n}{n+1}x\right)\| \leq \|T\|_2$$

It follows that

$$\|Tx\| \leq \frac{n+1}{n}\|T\|_2, \quad \forall n \in \mathbb{N}$$

which implies $\|T\| = \|T\|_2$

Now we check that $\|T\|_3 \leq \|T\|_1$. For every $x \neq 0$, since the norm of $x/\|x\|$ is 1, we have

$$\|T\left(\frac{x}{\|x\|}\right)\| \leq \|T\|_1$$

thus $\|Tx\| \leq \|T\|_1\|x\|$, so $\|T\|_3 \leq \|T\|_1$. Further, if x is in X, with $\|x\| = 1$, $\|Tx\| \leq M$, thus $\|T\|_1 \leq \|T\|_3$. We have proven that $\|T\|_3 = \|T\|_1$.

If $\|x\| \leq 1$, we can conclude similarly that $\|T\| \leq \|T\|_3 = \|T\|_1$.

In addition, by $\|Tx\| \leq \|T\|_4 \|x\|$, it follows that $\|T\|_3 \leq \|T\|_4$. As obviously $\|T\|_4 \leq \|T\|_1$, we have checked all the equalities.

Examples. 1. Let M_α be the operator on $\mathcal{C}_K([0,1])$ defined by

$$M_\alpha(x) = \alpha(t) \cdot x(t),$$

where $\alpha \in \mathcal{C}_K([0,1])$. Clearly, M_α is linear and, since for every $x \in \mathcal{C}_K([0,1])$,

$$\|M_\alpha(x)\| = \sup_{t \in [0,1]} |\alpha(t) \cdot x(t)| \leq \|\alpha\| \cdot \|x\|,$$

bounded. Thus, $M_\alpha \in \mathcal{B}(\mathcal{C}_K([0,1]))$. Moreover, $\|M_\alpha\| \leq \|\alpha\|$. On the other hand, as the norm of the function x_o on $[0,1]$ defined by $x_o(t) = 1, \forall t \in [0,1]$ is equal to one, we have

$$\|M_\alpha\| = \sup_{\|x\| \leq 1} \|M_\alpha(x)\| \geq \|M_\alpha(x_o)\| = \sup_{t \in [0,1]} |\alpha(t)| = \|\alpha\|$$

It follows that $\|M_\alpha\| = \|\alpha\|$.

2. Suppose that $(\lambda_n)_n$ is in $l^\infty_{\mathbf{K}}$ and define the mapping T on $l^p_{\mathbf{K}}$ by

$$T((\xi_n)_n) = (\lambda_n \xi_n)_n$$

We notice that for each $x = (\xi_n)_n \in l^p_{\mathbf{K}}$, the sequence $(\lambda_n \xi_n)_n$ is also p-summable, since

$$\left(\sum_{n=1}^{\infty} |\lambda_n \xi_n|^p \right)^{\frac{1}{p}} \leq (\sup_n |\lambda_n|) \; \|(\xi_n)_n\|$$

It follows that the operator $T \in \mathcal{B}(l^p_{\mathbf{K}})$ and $\|T\| \leq \sup_n |\lambda_n|$. Moreover, for every n, pick in $l^p_{\mathbf{K}}$ the sequence $e_n = (\delta_k^{(n)})_k$, which, obviously has $\|e_n\| = 1$. Then,

$$\|T\| = \sup_{\|x\| \leq 1} \|T(x)\| \geq \|T(e_n)\| = |\lambda_n|,$$

therefore $\|T\| \leq \sup_n |\lambda_n|$. We have shown that

$$\|T\| = \sup_n |\lambda_n|$$

3. On $\mathcal{C}_K([0,1])$ define the mapping $x \longmapsto Tx$, with

$$Tx(s) = \int_0^1 k(s,t) x(t) \, \mathrm{d}t, \quad \forall s \in [0,1],$$

where $k \in \mathcal{C}_K([0,1] \times [0,1])$. It is easy to see that

$$Tx \in \mathcal{C}_K([0,1]), \quad \forall x \in \mathcal{C}_K([0,1])$$

We shall show that $T \in \mathcal{B}(\mathcal{C}_{\mathbf{K}}([0,1]))$ and

$$\|T\| = \sup_{s \in [0,1]} \int_0^1 |k(s,t)| \, \mathrm{dt}$$

(The operator T is called the integral operator of kernel k).

As the linearity of T is immediate, let us check its continuity. For each $x \in \mathcal{C}_{\mathbf{K}}([0,1])$,

$$\|Tx\| = \sup_{s \in [0,1]} |Tx(s)| \leq \sup_{s \in [0,1]} \int_0^1 |k(s,t)| \cdot |x(t)| \, \mathrm{dt} \leq \|x\| \cdot \sup_{s \in [0,1]} \int_0^1 |k(s,t)| \, \mathrm{dt},$$

thus $T \in \mathcal{B}(\mathcal{C}_{\mathbf{K}}([0,1]))$ and

$$\|T\| \leq \sup_{s \in [0,1]} \int_0^1 |k(s,t)| \, \mathrm{dt}$$

In order to compute the norm of the operator T, we notice that the function $t \longmapsto \int_0^1 |k(s,t)| \, \mathrm{dt}$ is continue on $[0,1]$, so there exists $s_o \in [0,1]$ such that

$$\sup_{s \in [0,1]} \int_0^1 |k(s,t)| \, \mathrm{dt} = \int_0^1 |k(s_o,t)| \, \mathrm{dt}$$

Let f_o be the function on $[0,1]$ defined by $f_o(t) = \text{sign } k(s_o,t)$ (where, as usually, for a real number a, sign $a = 1, 0$, respectively -1 if $a > 0$, $= 0$, respectively < 0). This function is measurable, thus, by Luzin's Theorem (Appendix C), for every $\varepsilon > 0$, there is a continuous real function, g_ε on $[0,1]$ such that the measure of the set $\{t \in [0,1] \mid f_o(t) \neq g_\varepsilon(t)\}$ is less than ε ; moreover g_ε can be chosen such that

$$\|g_\varepsilon\| \leq \sup_{t \in [0,1]} |f_o(t)| = 1$$

It follows that we can obtain a sequence $(g_n)_n \subset \mathcal{C}_{\mathbf{K}}([0,1])$ such that for each n, the measure of the set $A_n = \{t \in [0,1] \mid f_o(t) \neq g_n(t)\}$ is less than $1/n$ and $\|g_n\| \leq 1$. We claim

$$\int_0^1 |k(s_o,t)| \, \mathrm{dt} = \lim_n \int_0^1 k(s_o,t) g_n(t) \, \mathrm{dt}$$

as it results from

$$\left|\int_0^1 k(s_o,t)f_o(t)\,\mathrm{dt} - \int_0^1 k(s_o,t)g_n(t)\,\mathrm{dt}\right| \le \int_0^1 |k(s_o,t)|\cdot|g_n(t)-f_o(t)|\,\mathrm{dt} \le$$

$$\le \int_{A_n} |k(s_o,t)|\cdot(|g_n(t)|+|f_o(t)|)\,\mathrm{dt} \le 2\sup_{t\in[0,1]}|k(s_o,t)|\cdot\frac{1}{n}\longrightarrow 0$$

Then, since

$$Tg_n(s_o) \le |Tg_n(s_o)| \le \|Tg_n\| \le \|T\|\cdot\|g_n\| \le \|T\|,$$

we have

$$\int_0^1 |k(s_o,t)|\,\mathrm{dt} = \lim_n \int_0^1 k(s_o,t)g_n(t)\,\mathrm{dt} = \lim_n Tg_n(s_o) \le \|T\|$$

Theorem 3.1.1 *If Y is a Banach space, $\mathcal{B}(X,Y)$ is a Banach space.*

Proof. Let $(T_n)_n \subset \mathcal{B}(X,Y)$ be a Cauchy sequence. We must prove that there is a bounded linear operator T so that $\|T_n - T\| \longrightarrow 0$. Since for each arbitrary $x \in X$,

$$\|T_nx - T_mx\| \le \|T_n - T_m\|\cdot\|x\|, \quad \forall n,m\in\mathbb{N}$$

it results that $(T_nx)_n$ is a Cauchy sequence in Y, $\forall x\in X$. As Y is complete, $(T_nx)_n$ converges to a (unique) element $y\in Y$. Define $Tx = y$. It is easy to check that T is a linear operator:

$$T(\alpha x+\beta y) = \lim_{n\to\infty} T_n(\alpha x+\beta y) = \lim_{n\to\infty}(\alpha T_nx+\beta T_ny) =$$

$$= \alpha\lim_{n\to\infty}T_nx + \beta\lim_{n\to\infty}T_ny = \alpha Tx+\beta Ty, \quad \forall\alpha,\beta\in\mathbb{K}\ ,\ x,y\in X$$

Since $(T_n)_n$ is a Cauchy sequence in $\mathcal{B}(X,Y)$, it is bounded, thus, there is $M>0$ so that $\|T_n\|\le M$, $\forall n$. As, for each $x\in X$, $\|T_nx\|\le\|T_n\|\cdot\|x\|$, we have

$$\|Tx\| = \lim_{n\to\infty}\|T_nx\| \le M\|x\|,$$

so $T\in\mathcal{B}(X,Y)$. We must still show that $T_n\longrightarrow T$ in the operator norm. Given $\varepsilon>0$, there is n_ε so that $n,m\ge n_\varepsilon$ implies $\|T_n-T_m\|<\varepsilon$. As for arbitrary x in X we have

$$\|T_nx - T_mx\| \le \|T_n - T_m\|\cdot\|x\|$$

it follows that $n, m \geq n_\varepsilon$ implies

$$\|T_n x - T_m x\| < \varepsilon \cdot \|x\|, \quad \forall x \in X$$

For fixed $n \geq n_\varepsilon$, as $m \to \infty$ the previous inequality becomes

$$\|T_n x - Tx\| \leq \varepsilon \cdot \|x\|, \quad \forall x \in X$$

which implies

$$\|T_n - T\| \leq \varepsilon, \quad \forall n \geq n_\varepsilon$$

therefore $T_n \longrightarrow T$ in $\mathcal{B}(X, Y)$.

In addition, as the norm mapping is continuous on $\mathcal{B}(X, Y)$, we have that $\|T_n\| \longrightarrow \|T\|$.

Remark. One can show whenever $\mathcal{B}(X, Y)$ is a Banach space, Y is a Banach space (Exercise 3).

Corollary 3.1.1 *The dual space of a normed space is a Banach space.*

Proof. It follows from the previous theorem since $(\mathbb{K}, |\cdot|)$ is Banach and $X^* = \mathcal{B}(X, \mathbb{K})$

Further, we shall state and prove the theorem concerning the extension by continuity.

Theorem 3.1.2 *(Extension by continuity) Let $(X, \|\cdot\|)$ be a normed space and X_o a dense subspace of X. Suppose that T is a bounded linear operator from the normed space X_o to a Banach space $(Y, \|\cdot\|)$. Then, T can be uniquely extended to a bounded linear operator $\widetilde{T} : X \longrightarrow Y$. Moreover $\|T\| = \|\widetilde{T}\|$.*

Proof. For each $x \in X$, there is a sequence of elements $(x_n)_n$ in X_o with $x_n \longrightarrow x$ as $n \to \infty$. Since $(x_n)_n$ converges, it is Cauchy, so given $\varepsilon > 0$, we can find n_ε so that $n, m \geq n_\varepsilon$ implies $\|x_n - x_m\| \leq \varepsilon / \|T\|$. Then,

$$\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| \leq \|T\| \cdot \|x_n - x_m\| \leq \varepsilon$$

which proves that $(Tx_n)_n$ is a Cauchy sequence in Y. Since Y is Banach, $Tx_n \longrightarrow y$ for some y. Set $\widetilde{T}x = y$. We must first show that this definition is independent of the chosen sequence $x_n \longrightarrow x$. Let $(y_n)_n$ be an other sequence which converges to x. Then

$$0 = T(\lim_n (x_n - y_n)) = \lim_n T(x_n - y_n) = \lim_n Tx_n - \lim_n Ty_n$$

therefore

$$\lim_n Tx_n = \lim_n Ty_n$$

We show first that $\widetilde{T}$ so defined is linear. Indeed for $x, y \in X$, there are some sequences $(x_n)_n$, $(y_n)_n$ in X_o with $x_n \longrightarrow x$, $y_n \longrightarrow x$ as $n \to \infty$. Since $x_n + y_n \longrightarrow x + y$, we have

$$\widetilde{T}(x+y) = \lim_n T(x_n + y_n) = \lim_n Tx_n + \lim_n Ty_n = \widetilde{T}(x) + \widetilde{T}(y)$$

It is also easy to check that

$$\widetilde{T}(\alpha x) = \alpha \widetilde{T}(x), \quad \forall \alpha \in \mathbb{K},\ x \in X$$

Next, we shall prove that $\widetilde{T}$ is bounded. For arbitrary $x \in X$, $x = \lim_n x_n$, $(x_n)_n \subset X_o$, we have

$$\|\widetilde{T}x\| = \| \lim_n T(x_n)\| = \lim_n \|T(x_n)\| \leq \lim_n \|T\| \cdot \|x_n\| = \|T\| \cdot \|x\|$$

Thus $\widetilde{T}$ is bounded and $\|\widetilde{T}\| \leq \|T\|$.

Clearly, $\widetilde{T}$ is an extension of T, because, for $x \in X_o$, $x = \lim_n x_n$, where $x_n = x$, $\forall n$. Then,

$$\widetilde{T}x = \lim_n Tx_n = Tx$$

Now, we have to prove that $\widetilde{T}$ defined above is the unique bounded linear extension of T to X. Suppose that there is an other linear operator $\widehat{T}$ on X enjoying the same properties like $\widetilde{T}$. For each $x \in X$, let $(x_n)_n$ be in X_o, $x_n \longrightarrow x$. Then, as $\widehat{T} \in \mathcal{B}(X, Y)$, we have

$$\widehat{T}(x) = \lim_n \widehat{T}(x_n) = \lim_n T(x_n) = \widetilde{T}x$$

so $\widehat{T}$ coincides to $\widetilde{T}$.

In order to prove that $\|\widetilde{T}\| = \|T\|$, since we have already seen that $\|\widetilde{T}\| \leq$ $\leq \|T\|$, it is enough to check that $\|\widetilde{T}\| \geq \|T\|$. As

$$\|\widetilde{T}\| = \sup_{x \in X,\ \|x\| \leq 1} \|\widetilde{T}x\| \geq \sup_{x \in X_o,\ \|x\| \leq 1} \|\widetilde{T}x\| = \sup_{x \in X_o,\ \|x\| \leq 1} \|Tx\| = \|T\|$$

everything is proven.

3.2 Bounded linear functionals

3.2.1 Hahn-Banach extension theorem in normed spaces and its consequences

Theorem 3.2.1 *(The Hahn–Banach extension theorem in normed spaces)* *Let $(X, \|\cdot\|)$ be a normed space and X_o a subspace of X. Suppose that f is a continuous linear functional on X_δ. Then, f can be extended to a continuous linear functional $\widetilde{f}$ on X. Moreover $\|f\| = \|\widetilde{f}\|$.*

Proof. As $|f(x)| \leq \|f\| \cdot \|x\|$, $\forall x \in X_o$, and $p(x) = \|f\| \cdot \|x\|$ is a seminorm on X, by Hahn-Banach extension theorem, we conclude that there is an extension $\widetilde{f}$ of f to the whole space so that $|\widetilde{f}(x)| \leq \|f\| \cdot \|x\|$, $\forall x \in X$. It follows that $\|\widetilde{f}\| \leq \|f\|$. On the other hand

$$\|\widetilde{f}\| = \sup_{x\in X,\ \|x\|\leq 1} \|\widetilde{f}(x)\| \geq \sup_{x\in X_o,\ \|x\|\leq 1} \|\widetilde{f}(x)\| = \sup_{x\in X_o,\ \|x\|\leq 1} \|f(x)\| = \|f\|$$

thus $\|\widetilde{f}\| = \|f\|$.

Corollary 3.2.1 *For every $x_o \in X$, there is a continuous linear functional on X such that $f(x_o) = \|x_o\|$ and $\|f\| = 1$.*

Proof. The existence of $f \in X^*$ such that $f(x_o) = \|x_o\|$ and $\|f\| \leq 1$ is proven, by Corollary 1.3.1 and Corollary 1.3.2, where $p(x) = \|x\|$. If we recall that f is the extension of the bounded linear functional f_o defined on the subspace of X spanned by $\{x_o\}$ by $f_o(\lambda x_o) = \lambda\|x_o\|$, which whenever $x_o \neq 0$ has its norm one, it follows that $\|f\| = 1$. If $x_o = 0$, everything is clear.

The next results are immediate consequences of the above corollary.

Corollary 3.2.2 *Let x_o be in the normed space X ($X \neq \{0\}$) such that $f(x_o) = 0$ for every $f \in X^*$. Then $x_o = 0$.*

Corollary 3.2.3 *Let X be a normed space, $X \neq \{0\}$. Then, $X^* \neq \{0\}$.*

Corollary 3.2.4 *Let X be a normed space, let Y be a closed, linear subspace of X, and let x_o be a point in X that is not in Y. Let $\delta = \operatorname{dist}(x_o, Y)$. Then, there is an element $f \in X^*$ such that*

$$\|f\| = \frac{1}{\delta}\ ,\ \ f(x_o) = 1,\ \text{and}\ f(y) = 0,\ \forall y \in Y.$$

Proof. Set $Z = \text{Sp}\, Y \cup \{x_o\}$, and define f_o on Z by $f_o(y + \lambda x_o) = \lambda$. Once we have shown that f_o has norm $1/\delta$ on Z then we can take f to be any element of X^* whose restriction to Z is f_o and whose norm is $1/\delta$ (such extension exists by Hahn-Banach theorem). Now,

$$\|y + \lambda x_o\| = |\lambda| \cdot \|x_o - \frac{-1}{\lambda} y\| \geq |\lambda| \cdot \delta = \delta \cdot |f_o(y + \lambda x_o)|$$

thus, $\|f_o\| \leq 1/\delta$. On the other hand, there is a sequence $(y_n)_n \subset Y$ so that $\|x_o - y_n\| \longrightarrow \delta$. Then,

$$\frac{1}{\|x_o - y_n\|}\,(x_o - y_n) \in Z, \text{ and } \|\frac{1}{\|x_o - y_n\|}\,(x_o - y_n)\| = 1 \quad \forall n,$$

so

$$\|f_o\| = \sup_{\|z\| \leq 1} |f(z)| \geq |f(\frac{1}{\|x_o - y_n\|}\,(x_o - y_n)| = \frac{1}{\|x_o - y_n\|} \longrightarrow \frac{1}{\delta}$$

It follows that $\|f_o\| = 1/\delta$.

Next we describe the duals of the Banach spaces $l^p_{\mathbf{K}}$, $p \geq 1$.

Example. If $p > 1$ the space $(l^p_{\mathbf{K}})^*$ is isometrically isomorphic to $l^q_{\mathbf{K}}$ (where q is the conjugate of p, $1/p + 1/q = 1$). The dual space of $l^1_{\mathbf{K}}$ is isometrically isomorphic to $l^\infty_{\mathbf{K}}$.

In order to prove the above assertions, we first remark that for each $x = (\xi_k)_k \in l^p_{\mathbf{K}}$ $(p \geq 1)$, the series $\sum_{n \geq 1} \xi_n e_n$ is convergent in $l^p_{\mathbf{K}}$ and its sum is x (where, for each n, e_n is the element in $l^p_{\mathbf{K}}$ defined by $e_n = (\delta^{(n)}_k)_k$, $\delta^{(n)}_k = 0$, if $k \neq n$ and $\delta^{(n)}_k = 1$ if $k = n$). Now, if $f \in (l^p_{\mathbf{K}})^*$, we have

$$f((\xi_k)_k) = \sum_{n=1}^{\infty} \xi_n f(e_n)$$

Let us denote by y_f the numerical sequence $(\eta_n)_n$ where $\eta_n = f(e_n)$, $\forall n$ and to show that if $p > 1$, $y_f \in l^q_{\mathbf{K}}$ and if $p = 1$, $y_f \in l^\infty_{\mathbf{K}}$.

Take first the case $p > 1$. For each n, let x_n be the numerical sequence

$$x_n = (|\eta_1|^{q-2}\eta_1, |\eta_2|^{q-2}\eta_2, ..., |\eta_n|^{q-2}\eta_n, 0, 0, ..., 0, ...)$$

Clearly, $x_n \in l^p_{\mathbf{K}}$ and

$$\|x_n\| = \left(\sum_{k=1}^{n} |\eta_k|^{(q-1)p}\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{n} |\eta_k|^{q}\right)^{\frac{1}{p}}$$

so, it follows that

$$\sum_{k=1}^{n} |\eta_k|^q = f(x_n) \leq \|f\| \cdot \|x_n\| = \|f\| \left(\sum_{k=1}^{n} |\eta_k|^q\right)^{\frac{1}{p}}$$

and, finally, after a division,

$$\left(\sum_{k=1}^{n} |\eta_k|^q\right)^{\frac{1}{q}} \leq \|f\|$$

Hence, we can conclude that $y_f \in l^q_{\mathbf{K}}$ and $\|y_f\| \leq \|f\|$. We also notice that, by the Hölder's inequality it follows

$$|f(x)| = |\sum_{n=1}^{\infty} \xi_n \eta_n| \leq \left(\sum_{n=1}^{\infty} |\eta_n|^q\right)^{\frac{1}{q}} \left(\sum_{n=1}^{\infty} |\xi_n|^q\right)^{\frac{1}{q}} = \|y_f\| \cdot \|x\|$$

thus $\|y_f\| = \|f\|$.

If $p = 1$, let x_n be the numerical sequence

$$x_n = (\delta_k^{(n)} \operatorname{sign} \eta_k)_k,$$

(where sign $\eta_k = 1, 0$, respectively -1 if $\eta_k > 0$, $= 0$, respectively < 0). Clearly, $x_n \in l^{\infty}_{\mathbf{K}}$ and $\|x_n\| \leq 1$. Taking into account that $f \in (l^1_{\mathbf{K}})^*$, we obtain

$$|\eta_n| = f(x_n) \leq \|f\| \cdot \|x_n\| \leq \|f\|$$

so $y_f \in l^{\infty}_{\mathbf{K}}$ and $\|y_f\| \leq \|f\|$. In addition,

$$|f(x)| \leq \sum_{n=1}^{\infty} |\xi_n \eta_n| \leq |f(x)| = \left(\sum_{n=1}^{\infty} |\xi_n|\right) \|y_f\| = \|y_f\| \cdot \|x\|$$

thus $\|f\| = \|y_f\|$.

We have proven that the mapping Ψ from $(l^p_{\mathbf{K}})^*$ onto $l^q_{\mathbf{K}}$, (respectively from $(l^1_{\mathbf{K}})^*$ onto $l^{\infty}_{\mathbf{K}}$) defined by $\Psi(f) = y_f$ is an isometric isomorphism.

3.2.2 The canonical embedding of a normed space into its bidual. Reflexive Banach spaces

Given a normed space $(X, \|\cdot\|)$ we form the Banach dual space X^* and, by iteration, we obtain the bidual space X^{**}. It consists, of course, of all bounded linear functionals on X^*. For each fixed $x \in X$ define $\widehat{x}(f)$ to be $f(x)$ for all $f \in X^*$. It is clear that $\widehat{x}$ is a linear functional on X^*, and since

$$|\widehat{x}(f)| = |f(x)| \leq \|f\| \cdot \|x\|$$

we see that $\widehat{x} \in X^{**}$. Hence we can define a map $\phi : X \longrightarrow X^{**}$ by letting $\phi(x) = \widehat{x}$.

Proposition 3.2.1 *The mapping $\phi : X \longrightarrow X^{**}$ defined by $\phi(x) = \widehat{x}$ is an isometric isomorphism from X onto a linear subspace of X^{**}.*

Proof. It is enough to check that $\|\widehat{x}\| = \|x\|$. We have already seen that $\|\widehat{x}\| \leq \|x\|$. On the other hand, by Corollary 3.2.1, there is $f_o \in X^*$ so that $\|f_o\| = 1$ and $f_o(x) = \|x\|$. It follows that

$$\|\widehat{x}\| = \sup_{\|f\| \leq 1} |\widehat{x}(f)| \geq |\widehat{x}(f_o)| = |f_o(x)| = \|x\|$$

thus we conclude that $\|\widehat{x}\| = \|x\|$.

Definition. The isometric isomorphism ϕ from X onto a linear subspace X^{**} defined by $\phi(x) = \widehat{x}$ is called the *canonical embedding* of X into X^{**}.

Remark. *We often identify X with its image in X^{**}, $\phi(X) = \{\widehat{x} \mid x \in X\}$.*

Definition. A normed space is said to be a *reflexive* normed space if the canonical embedding maps the space onto its bidual.

Example. The spaces $l^p_{\mathbb{K}}$, $p \geq 1$ are reflexive. Indeed, let Ψ be the canonical isometric isomorphism from $(l^p_{\mathbb{K}})^*$ onto $l^q_{\mathbb{K}}$ and Φ be the canonical isometric isomorphism from $(l^q_{\mathbb{K}})^*$ onto $l^p_{\mathbb{K}}$ (see Example in 3.2.1). It follows that, for $x = (\xi_k)_k \in l^p_{\mathbb{K}}$ and $y = (\eta_k)_k \in l^q_{\mathbb{K}}$,

$$\Psi^{-1}(y)(x) = \Phi^{-1}(y)(x) = \sum_{n=1}^{\infty} \xi_n \eta_n$$

Let $x^{**} \in (l^p_{\mathbb{K}})^{**}$; we have to prove that there is an element $x \in l^p_{\mathbb{K}}$ so that $\widehat{x}(f) = x^{**}$, $\forall f \in (l^p_{\mathbb{K}})^*$. We set

$$x = \Phi(x^{**} \circ \Psi^{-1})$$

It follows that

$$\Phi^{-1}(x)(y) = x^{**}(\Psi^{-1}(y)), \ \forall y \in l^q_{\mathbb{K}},$$

thus, for each $f \in (l^p_{\mathbb{K}})^*$,

$$x^{**}(f) = x^{**}(\Psi^{-1}(y)) = \Phi^{-1}(y)(x) = \sum_{n=1}^{\infty} \xi_n \eta_n = f(x) = \hat{x}(f).$$

The next proposition, will enable us to prove that $l^1_{\mathbb{K}}$ is not a reflexive Banach space.

Proposition 3.2.2 *Let $(X, \|\cdot\|)$ be a normed space.*
1) If X^* is separable, then X is separable;
2) If X is separable and reflexive, then X^* is separable.

Proof. 1) Let $(f_n)_n$ be a countable dense subset of X. Since

$$\|f_n\| = \sup_{\|x\|=1} |f_n(x)|,$$

we can pick $x_n \in X$ with $\|x_n\| = 1$ and

$$|f_n(x_n)| \geq \frac{1}{2}\|f_n\|$$

Set X_1 the closure of the linear space spanned by the set $\{x_n | \; n \in \mathbb{N}\}$. By

$$X_1 = \overline{\{\sum_{j=1}^{n} \rho_j x_j \mid \rho_j \in \mathbb{Q}, \; n \in \mathbb{N}\}}$$

(where $\mathbb{Q}$ is the set of the rationals), it follows that X_1 is separable. Let us prove that $X_1 = X$. Suppose that there is $x_o \in X \setminus X_1$. By Corollary 3.2.4, there exists $f_o \in X^*$ so that its restriction at X_1 is zero and $f_o(x_o) = 1$. Then, for each n we have

$$\frac{1}{2}\|f_n\| \leq |f_n(x_n)| = |(f_n - f_o)(x_n)| \leq \|f_n - f_o\|$$

thus,

$$\|f_o\| \leq \|f_n - f_o\| + \|f_n\| \leq 3\|f_n - f_o\|$$

Taking into account that $f_o \in \overline{\{f_n | \; n \in \mathbb{N}\}}$, it follows that $f_o = 0$, which contradicts $f_o(x_o) = 1$.

2) Since X is reflexive and separable, we infer that X^{**} is separable. As $X^{**} = (X^*)^*$, applying 1) we obtain that X^* is separable.

The next example shows that there exists Banach spaces which are not reflexive.

Example. The spaces $l^1_{\mathbb{K}}$ and $l^\infty_{\mathbb{K}}$ are not reflexive.

Suppose that $l^1_{\mathbb{K}}$ is reflexive. Since it is also separable (a countable dense subset of $l^1_{\mathbb{K}}$ is the set $\{\sum_{j=1}^n \rho_j e_j \mid \rho_j \in \mathbb{Q},\ n \in \mathbb{N}\}$), it follows by the previous proposition that $(l^1_{\mathbb{K}})^*$ is a separable space. As $(l^1_{\mathbb{K}})^*$ is isometrically isomorphic to $l^\infty_{\mathbb{K}}$, it results that $l^\infty_{\mathbb{K}}$ is separable too. But this contradicts the non-separability of $l^\infty_{\mathbb{K}}$ (see Exercise 8, Ch. 3). The fact that $l^\infty_{\mathbb{K}}$ is not separable can be shown directly as follows. Fist, we remark that the subset E of $l^\infty_{\mathbb{K}}$, consisting of sequences $(\xi_n)_n$, with $\xi_n \in \{0,1\}$ is not countable. Suppose that A is a dense subset of $l^\infty_{\mathbb{K}}$. Then, for each $x \in E$, we can pick y_x in $A \cap B(x, 1/2)$. The mapping $x \longmapsto y_x$ from E to A is injective, since, if $x_1 \neq x_2$, we have that $\|x_1 - x_2\| = 1$, thus $y_{x_1} \neq y_{x_2}$. Consequently, the set A can not be countable.

As an immediate application to the embedding of a normed space into its bidual one can establish the existence of a completion for each normed space.

Definition. Let $(X, \|\cdot\|)$ be a normed space. A Banach space that has a dense, linear subspace isometrically isomorphic to $(X, \|\cdot\|)$ is called a *completion* of $(X, \|\cdot\|)$.

The existence of a completion of a normed space $(X, \|\cdot\|)$ can be established in several ways. The most pedestrian is to imitate the construction of the real numbers from the rationals, and define the completion to be the space of Cauchy sequences in X modulo the space of null sequences. We choose instead to let the completion be the closure of X embedded in its bidual (Banach) space X^{**}.

Theorem 3.2.2 *To each normed space*$(X, \|\cdot\|)$ *there is a completion of* $(X, \|\cdot\|)$, $(\widetilde{X}, \|\cdot\|)$ *uniquely determined up to isometric isomorphism.*

Proof. Let ϕ be the canonical embedding of $(X, \|\cdot\|)$ into X^{**}. Clearly, $(X, \|\cdot\|)$ and $\phi(X)$ are isometrically isomorphic and $\phi(X)$ is a dense linear subspace of its closure in $(X^{**}, \|\cdot\|)$. But $(X^{**}, \|\cdot\|)$ is a Banach space. Hence, the closure of $\phi(X)$ is also a Banach space. It follows that $\overline{\phi(X)}$ is a completion of $(X, \|\cdot\|)$.

Next we prove the uniqueness of the completion up to an isometrics isomorphism. If $\widetilde{X}_1$ and $\widetilde{X}_2$ are two Banach spaces for which we have isometric embeddings $T_j : X \longrightarrow \widetilde{X}_j$, $j = 1, 2$, of X as a dense subspace, then $T = T_2 T_1^{-1}$ is an isometry of $T_1(X)$ onto $T_2(X)$. By Theorem 3.1.2, T can be extended by continuity to $\widetilde{X}_1$; consequently we obtain an isometry $\widetilde{T}$ of $\widetilde{X}_1$ onto $\widetilde{X}_2$ (because $\widetilde{T}(\widetilde{X}_1)$ is both closed and dense in $\widetilde{X}_2$).

Example. Let us consider on s_{oo} the norm $\|(\xi_k)_k\| = \sup_k |\xi_k|$. We have already proven that $(s_{oo}, \|\cdot\|)$ is not a Banach space and that s_{oo} is a dense linear subspace of c_o (2.2, Example 2). It follows that the completion of $(s_{oo}, \|\cdot\|)$ is c_o.

3.3 The Baire category theorem and its consequences

3.3.1 The Baire category theorem

In Banach space theory it is of great interest to know when sets have nonempty interiors, as we shall see by proving some of the most important theorems about bounded operators on normed spaces.

Theorem 3.3.1 *(The Baire category theorem) A complete metric space (X, d) (in particular a Banach space) is never the union of a countable number of nowhere dense sets.*

Proof. The idea of the proof is as follows: suppose that the complete metric space $X = \bigcup_{n=1}^{\infty} A_n$, with each A_n nowhere dense. We will construct a Cauchy sequence $(x_m)_m$ which stays away from each A_n, so, its limit (which exists in X by completeness) is in no A_n, thereby contradicting the statement $X = \bigcup_{n=1}^{\infty} A_n$.

Since A_1 is nowhere dense, the closure of the set $C_X \overline{A_1}$ is X, so we can find $x_1 \notin \overline{A_1}$. Then, there exists an open ball B_1 about x_1 with radius smaller than $1/2$ so that $B_1 \cap A_1 = \emptyset$.

Since A_2 is nowhere dense, and $x_1 \in X = \overline{C_X \overline{A_2}}$, it follows that $B_1 \cap$ $\cap C_X \overline{A_2} \neq \emptyset$, so we can find $x_2 \in B_1 \setminus \overline{A_2}$. Pick an open ball B_2 about x_2 with radius smaller than $1/2^2$ so that $B_2 \cap A_2 = \emptyset$, $\overline{B_2} \subset B_1$. Proceeding inductively, we find $x_n \in B_{n-1} \setminus \overline{A_n}$ and choose an open ball B_n about x_n, with radius smaller than $1/2^n$ satisfying $B_n \cap A_n = \emptyset$, $\overline{B_n} \subset B_{n-1}$.

The sequence $(x_n)_n$ is a Cauchy sequence. Indeed, given $\varepsilon > 0$, we pick n_ε so that $2^{n_\varepsilon - 1} < \varepsilon$. Then, $m, n \geq n_\varepsilon$ implies that x_m, $x_n \in B_{n_\varepsilon}$ and we have

$$d(x_m, x_n) \leq d(x_m - x_{n_\varepsilon}) + d(x_n - x_{n_\varepsilon}) \leq 2 \cdot 2^{-n_\varepsilon} = 2^{n_\varepsilon - 1} < \varepsilon$$

Now, taking into account that X is complete, the sequence $(x_n)_n$ converges in X; let

$$x = \lim_n x_n$$

Since $X = \bigcup_{n=1}^{\infty} A_n$, there exists n_o so that $x \in A_{n_o}$. By $B_{n_o} \cap A_{n_o} = \emptyset$, it follows that $x \notin B_{n_o}$. On the other hand, for $n \geq n_o + 1$, $x_n \in B_{n_o+1}$, so,

$$x = \lim_n x_n \in \overline{B_{n_o+1}} \subset B_{n_o},$$

which contradicts $x \notin B_{n_o}$.

The above theorem has an immediate application.

Application. Let X be a vector space that has an infinite countable algebraic basis. Then, equipped with any norm, X can not be a Banach space. Suppose that there is a norm on X, $\|\cdot\|$ so that $(X, \|\cdot\|)$ is Banach. Take $B = \{e_n \mid n \in \mathbb{N}\}$ a countable algebraic basis of X and set

$$X_n = \mathrm{Sp}\ \{e_1, e_2, \ldots e_n\}, \quad \forall n$$

Clearly, each X_n is a closed, proper, linear subspace of X (since it is finite dimensional), and $X = \bigcup_{n=1}^{\infty} X_n$. By the Baire category theorem, it follows that we can pick n_o so that the interior of X_{n_o} is nonempty. So, in X we have a proper linear subspace X_{n_o} such that $\overset{\circ}{X}_{n_o} \neq \emptyset$. But, this is not possible because, if $\overset{\circ}{X}_{n_o} \neq \emptyset$, there is $x_o \in X_{n_o}$ and $r > 0$ so that $B(x_o, r) \subset X_{n_o}$; it follows that each x in X is in X_{n_o} (since $x_o + (r/2\|x\|)x \in X_{n_o}$), hence $X_{n_o} = X$. So, once we obtained that the interior of a proper linear subspace of a normed space is nonempty, we reach a contradiction. The statement is proven.

Hence we can infer that there is no norm on s_{oo} or on the vector space of all polynomials on $[0, 1]$, $\mathcal{P}([0, 1])$, so that s_{oo} or $\mathcal{P}([0, 1])$ be Banach spaces.

3.3.2 Principle of uniform boundedness

Theorem 3.3.2 *(Principle of uniform boundedness) Let $(X, \|\cdot\|)$ be a Banach space. Let $\{T_\iota\}_{\iota\in I}$ be a family of bounded linear operators from X to some normed linear space $(Y, \|\cdot\|)$ such that for each $x \in X$, the subset of Y, $\{T_\iota(x) \mid \iota \in I\}$ is bounded. Then, the family $\{T_\iota\}_{\iota\in I}$ is bounded in $\mathcal{B}(X,Y)$.*

Proof. For each $n \in \mathbb{N}$ let $A_n = \{x \in X \mid \|T_\iota(x)\| \leq n, \ \forall \iota \in I\}$. By the hypothesis each x is in some A_n, that is $X = \bigcup_{n=1}^{\infty} A_n$. Moreover each A_n is closed since

$$A_n = \bigcap_{\iota\in I} g_\iota^{-1}([0,n]),$$

where g_ι is the continuous real-valued mapping on X, $g_\iota = \|\cdot\| \circ T_\iota$. By the Baire category theorem, some A_n has a nonempty interior, i.e. $\exists n_o$ so that $\overset{\circ}{A}_n \neq \emptyset$. It follows that $\overset{\circ}{A}_n$ contains the ball $B(x_o, r)$. Given an arbitrary $x \neq 0$, the element

$$y_x = x_o + \frac{r}{2\|x\|}\, x$$

is in the ball $B(x_o, r)$. Thus, for all $x \in X$ and $\iota \in I$,

$$\|T_\iota(x)\| = \|T_\iota\left(\frac{2\|x\|}{r}(y_x - x_o)\right)\| = \frac{2\|x\|}{r}\|T_\iota(y_x) - T_\iota(x_o)\| \leq$$

$$\leq \frac{2\|x\|}{r}\left(\|T_\iota(y_x)\| + \|T_\iota(x_o)\|\right) \leq \frac{4n_o}{r}\|x\|$$

and we can conclude that

$$\|T_\iota\| \leq \frac{4n_o}{r}\ , \quad \forall \iota \in I$$

Remark. The proof of the above theorem strongly uses that X is Banach. The following example shows that the uniform boundedness principle does not work when X is not complete.

Example. Let $(f_n)_n$ be the sequence of linear functionals on $s_{oo} \subset (l^\infty_{\mathbb{K}}, \|\cdot\|)$ defined by

$$f_n((\xi_k)_k) = n\xi_n$$

Clearly, each f_n is in s_{oo}^* because of

$$\|f_n((\xi_k)_k)\| \leq n\|(\xi_k)_k\|$$

and for each $x = (\xi_k)_k$, the set $\{f_n((\xi_k)_k) \mid n \in \mathbb{N}\}$ is bounded in $\mathbb{K}$ (since if $\xi_k = 0$, $\forall k \geq k_o$ it follows that $f_n((\xi_k)_k) = 0$, $\forall n \geq k_o$). On the other hand, as $f_n((\delta_k^{(n)})_k) = n$, it follows that $\|f_n\| = n$, so the set $\{f_n \mid n \in \mathbb{N}\}$ is not bounded in s_{oo}^*. This occurs since s_{oo} is not Banach.

Definition. A sequence $(T_n)_n$ in $\mathcal{B}(X,Y)$ is said to be *pointwise convergent* if for each $x \in X$ the sequence $(T_n x)_n$ is convergent in Y.

Remarks. 1. *If $(T_n)_n$ is a pointwise convergent sequence in $\mathcal{B}(X,Y)$, then, the mapping T from X to Y defined by*

$$T(x) = \lim_n T_n(x),$$

is clearly well defined (by the uniqueness of the limit) and also linear (by the linearity of each T_n and by the continuity of the sum and the multiplication on normed spaces). The linear operator T from X to Y is called the pointwise limit of the sequence $(T_n)_n$.

2. Let $(T_n)_n$ be a convergent sequence in the normed space $\mathcal{B}(X,Y)$ and let T be its norm limit. Then, since for each x in X,

$$\|T_n(x) - T(x)\| \leq \|T_n - T\| \cdot \|x\|$$

it follows that the sequence $(T_n)_n$ converges pointwise to T. The converse is not generally true, as it results from the following example.

Example. Let $(T_n)_n$ be the sequence in $\mathcal{B}(l^p_{\mathbf{K}})$ defined by

$$T_n((\xi_k)_k) = (0, 0, ..., 0, \xi_n, \xi_{n+1}, ...)$$

Each T_n is defined by the bounded numerical sequence $(\lambda_k)_k$ where $\lambda_k = 0$ if $k < n$ and $\lambda_k = 1$ if $k \geq n$ (Example 2 in 3.1), thus $\|T_n\| = 1$, $\forall n$. For any $x = (\xi_k)_k \in l^p_{\mathbf{K}}$ we have that $\lim_n T_n(x) = 0$, since

$$\|T_n(x)\|^p = \sum_{m=n}^{\infty} |\xi_m|^p \longrightarrow 0,$$

as $n \to \infty$. It follows that $(T_n)_n$ converges pointwise to the null operator. Suppose that there is T in $\mathcal{B}(l^p_{\mathbf{K}})$ so that $(T_n)_n$ converges in norm to T. By the previous remark, necessarily, $T = 0$, thus $\|T_n\| \longrightarrow 0$, as $n \to \infty$, which contradicts $\|T_n\| = 1$, $\forall n$.

The next theorem makes clear the fact that the pointwise limit of sequences in $\mathcal{B}(X,Y)$ (X Banach) is also in $\mathcal{B}(X,Y)$. As we have pointed out (see the above remark and example), that does not mean that the pointwise convergence implies the convergence in $\mathcal{B}(X,Y)$.

Theorem 3.3.3 *(The Banach–Steinhauss theorem) Let $(X, \|\cdot\|)$ be a Banach space, Y be some normed linear space and $(T_n)_n \subset \mathcal{B}(X,Y)$ be a pointwise convergent sequence of bounded linear operators from X to Y. Then, T, the pointwise limit of the sequence $(T_n)_n$ is a bounded linear operator from X to Y. In addition the numerical sequence $(\|T_n\|)_n$ is bounded and $\|T\| \leq \sup_n \|T_n\|$.*

Proof. Let $(T_n)_n$ be a pointwise convergent sequence in the space $\mathcal{B}(X,Y)$ and let T be its pointwise limit. We have already noticed that T is linear. As for each x the sequence $(T_n(x))_n$ is convergent, it follows that the set $\{T_n(x) \mid n \in \mathbb{N}\}$ is bounded in Y, so by the principle of uniform boundedness, there is $M > 0$, so that $\|T_n\| \leq M$, $\forall n$. Then, for each x and n,

$$\|T_n(x)\| \leq \|T_n\| \cdot \|x\| \leq M\|x\|$$

and we have

$$\|T(x)\| = \lim_n \|T_n(x)\| \leq M\|x\|, \ \forall x \in X$$

It follows in addition that

$$\|T\| \leq \sup_n \|T_n\|$$

We give next a typical application of the Banach-Steinhauss theorem.

First, let us notice that, generally a mapping $f(x,y)$ on the product $X \times Y$ of two topological spaces, separately continuous (that is, for each x, $f(x,\cdot)$ is continuous on Y, and for each y, $f(\cdot,y)$ is continuous on X) is not necessarily jointly continuous (continuous on $X \times Y$ with the product topology). Even on $\mathbb{R}^2$ we have the standard example,

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

In the particular case of the linear mappings on the product of two Banach spaces (called bilinear mappings), the separate continuity implies the joint continuity. More precisely we have the next proposition:

Proposition 3.3.1 *Let X and Y be Banach spaces and let B a separately bilinear mapping from $X \times Y$ to $\mathbb{K}$. Then, B is jointly continuous.*

Proof. Let $(x_n, y_n)_n$ be a sequence in $X \times Y$ which converges to zero. Define on Y the (countable) family of bounded linear functionals,

$$f_n(y) = B(x_n, y)$$

Since $B(\cdot, y)$ is continuous on X, it follows that $(f_n(y))_n$ converges pointwise to zero for each y, so, by Banach-Steinhauss theorem, there exists $M > 0$ so that $|f_n(y)| \leq M\|y\|$, $\forall n$. Then,

$$|B(x_n, y_n)| = |f_n(y_n)| \leq M\|y_n\| \longrightarrow 0$$

3.3.3 The open mapping theorem

An other consequence of the Baire category theorem is the next fundamental theorem.

Theorem 3.3.4 *(Open mapping theorem) Let X, Y be Banach spaces. Suppose that T is a bounded linear operator from X onto Y. Then, if A is an open set in X, $T(A)$ is open in Y (T is an open mapping from X onto Y).*

Proof. We shall proceed in 3 steps.

1) *Let T be a bounded linear operator from the normed space X onto the Banach space Y. Then, for each positive $r > 0$, there is $\alpha_r > 0$ so that $B(\alpha_r) \subset \overline{T(B(r))}$.*

We notice first that for each arbitrary $r > 0$, X can be written as $\bigcup_{n=1}^{\infty} nB(r/2)$. Then, $T(X) = \bigcup_{n=1}^{\infty} nT(B(r/2))$, and, since T is onto, we have $Y = \bigcup_{n=1}^{\infty} nT(B(r/2))$. As Y is complete, by the Baire category theorem, it follows that there is a natural n_o so that the closure of $n_oT(B(r/2))$ has a nonempty interior, so, taking into account that the mapping on Y, $y \longmapsto n_oy$ is a homeomorphism, it follows that $\overset{\circ}{\overline{T(B(r/2))}} \neq \emptyset$. Let us denote this open nonempty set by U. Then, $U - U$ is an open neighbourhood of zero (since $U - U = \bigcup_{u \in U}(u - U)$ and $y \longmapsto u - y$ is a homeomorphism of Y). It results that there is $\alpha_r > 0$ so that $B(\alpha_r) \subset U - U$. Using the continuity of the mapping from $Y \times Y$ to Y, $(y, z) \longmapsto y - z$, and the linearity of T, we have the next inclusions

$$B(\alpha_r) \subset U - U = \overset{\circ}{\overline{T(B(r/2))}} - \overset{\circ}{\overline{T(B(r/2))}} \subset \overline{T(B(r/2))} - \overline{T(B(r/2))} \subset$$

$$\subset \overline{T(B(r/2)) - T(B(r/2))} = \overline{T(B(r/2) - B(r/2))} = \overline{T(B(r))}$$

which end the proof of the first step.

Using 1), we shall establish, the most difficult part of the proof:

2) *Let X, Y be Banach spaces and T a bounded linear operator from X onto Y. Then, for each positive $r > 0$, there is $\delta_r > 0$ so that*

$$B(\delta_r) \subset T(B(r))$$

Let $r > 0$ be. For each positive integer k, denote by r_k the positive number $r/2^{k+2}$. By 1), $\forall k$, there is $\alpha_{r_k} > 0$ so that $B(\alpha_{r_k}) \subset \overline{T(B(r_k))}$. Since $T(B(r_k)) \subset B(\|T\| r_k)$,we may suppose that the sequence $(\alpha_{r_k})_k$ converges to zero.

We will show that $B(\alpha_{r_o}) \subset T(B(r))$, consequently, the desired δ_r is α_{r_o}. Let y be arbitrary in $B(\alpha_{r_o}) \subset \overline{T(B(r_o))}$. Then, since $B(y, \alpha_{r_1}) \cap T(B(r_o))$ is nonempty, we can pick $x_o \in B(r_o)$ so that $T(x_o) \in B(y, \alpha_{r_1})$. It follows that the element x_o has the properties

$$\|x_o\| < r_o \text{ and } y - T(x_o) \in B(\alpha_{r_1}) \subset \overline{T(B(r_1))}$$

Then, since $B(y - T(x_o), \alpha_{r_2}) \cap T(B(r_1))$ is nonempty, we can pick $x_1 \in B(r_1)$ so that $T(x_1) \in B(y - T(x_o), \alpha_{r_2})$. Then, x_1 has the properties

$$\|x_1\| < r_1 \text{ and } y - T(x_o) - T(x_1) \in B(\alpha_{r_2}) \subset \overline{T(B(r_2))}$$

By induction, we choose a sequence $(x_n)_n$ so that

$$\|x_n\| < r_n \text{ and } \|y - \sum_{k=0}^{n} T(x_k)\| < \alpha_{r_{n+1}}$$

Further, define for each n natural

$$z_n = \sum_{k=0}^{n} x_k$$

and let us check that the sequence $(z_n)_n$ is Cauchy. This follows from

$$\|z_{n+p} - z_n\| = \| \sum_{k=n+1}^{n+p} x_k\| \le \sum_{k=n+1}^{n+p} \|x_k\| \le \sum_{k=n+1}^{n+p} \frac{r}{2^{k+2}} =$$

$$= \frac{r}{4} \sum_{k=n+1}^{n+p} \frac{1}{2^k} = \frac{r}{2^{n+2}}(1 - \frac{1}{2^p}) < \frac{r}{2^{n+2}}, \ \forall n, p \in \mathbb{N}$$

Since X is Banach, there exists the limit of the sequence $(z_n)_n$; denote it by x. By

$$\|x\| = \| \lim_n \sum_{k=0}^{n} x_k\| \le \lim_n \sum_{k=0}^{n} \|x_k\| \le \lim_n \sum_{k=0}^{n} \frac{r}{2^{k+2}} = \lim_n \frac{r}{2}(1 - \frac{1}{2^{n+1}}) = \frac{r}{2}$$

it results that $x \in B(r)$. Further, since

$$\|y - \sum_{k=0}^{n} T(x_k)\| < \alpha_{r_{n+1}}, \quad \forall n,$$

we have that $\|y - T(z_n)\| < \alpha_{r_{n+1}}$, $\forall n$, and using the continuity of T, that implies

$$\|y - T(\lim_n z_n)\| \le \lim_n \alpha_{r_{n+1}}$$

which shows that $y = T(x)$, thus, $y \in T(B(r))$.

3) *The proof* follows immediately from the above statement, 2). Let D be an open set in X. We have to show that $T(D)$ is open in Y. Let y be arbitrary in $T(D)$, so $y = T(x)$ with x in D. Since D is open, it contains a ball with center x and radius $r > 0$, $B(x,r) = x + B(r)$. By the second step 2), there exists $\delta = \delta_r > 0$ so that $B(\delta) \subset T(B(r)$. It follows that

$$T(x) + B(\delta) \subset T(x) + T(B(r)) = T(x + B(r)) = T(B(x,r)) = T(D).$$

In practice, one rarely uses the open mapping theorem directly but rather its consequences. One of them is the next theorem.

Theorem 3.3.5 *(Inverse mapping theorem) A continuous linear bijection of one Banach space onto another has a continuous inverse.*

Proof. If T is defined from X to Y, we have to show that $(T^{-1})^{-1}(D)$ is open in Y for each arbitrary open set $D \subset X$. This follows by the Open mapping theorem, since for each $D \subset X$, $(T^{-1})^{-1}(D) = T(D)$.

3.3.4 Closed graph theorem

Definition. Let T be a mapping of a subset D of a normed linear space X into a normed linear space Y. The *graph* of T, denoted by G_T is defined as

$$G_T = \{(x,y)|(x,y) \in D \times Y,\ y = Tx\}$$

The mapping T is said to be closed if its graph is closed in the normed space $X \times Y$.

Remark. *The graph of the mapping* $T : D_T \subset X \longrightarrow Y$ *is closed if and only if for each arbitrary sequence* $(x_n)_n \subset D_T$ *so that* $x_n \longrightarrow x$ *and* $T(x_n) \longrightarrow y$, as $n \to \infty$, *it follows that* $x \in D_T$ *and* $y = T(x)$.

Theorem 3.3.6 *(Closed graph theorem) Let* X *and* Y *be Banach spaces and* T *a linear operator from* X *into* Y. *Then* T *is bounded if and only if the graph of* T *is closed.*

Proof. Suppose that G_T is closed. Then, since T is linear, G_T is a subspace of the Banach space $X \times Y$. By assumption G_T is closed and thus is a Banach space in the norm induced by the norm of $X \times Y$. Consider the continuous linear maps $P_1 : X \times Y \longrightarrow X$ and $P_2 : X \times Y \longrightarrow Y$ defined by $P_1(x,y) = x$, respectively $P_2(x,y) = y$. By

$$\|P_1(x,y)\| = \|x\| \le \max(\|x\|,\|y\|) \text{ and } \|P_2(x,y)\| = \|y\| \le \max(\|x\|,\|y\|)$$

it follows that the mappings P_1, P_2 are bounded. The restriction of P_1 to G_T is a bounded linear operator from the Banach space G_T onto the Banach space Y, thus, by the Theorem 3.3.5, $P_1^{-1} : Y \longrightarrow G_T$ is a bounded linear operator. But, $T = P_2 \circ P_1^{-1}$, so T is continuous. The converse is trivial.

Remark. To avoid the confusions, we emphasize that the mapping T in the previous theorem is implicitly assumed to be defined on the Banach space X. There are examples of linear operators with closed graph which are not continuous (thus, their domain is not Banach).

Example. Let X be the linear subspace of $l^2_{\mathbb{K}}$ of all sequences $(\xi_n)_n \in l^2_{\mathbb{K}}$ satisfying $\sum_{n\ge1} n^2|\xi_n|^2 < \infty$ and $T : X \longrightarrow l^2_{\mathbb{K}}$ the linear operator defined by $T((\xi_n)_n) = (n\xi_n)_n$. The graph of T is closed. Indeed let $(x_n)_n \subset X$, $x_n = (\xi_k^{(n)})_k$ so that $x_n \longrightarrow x$, $x = (\xi_k)_k$ and $T(x_n) \longrightarrow y$, as Since the sequence $(T(x_n))_n$ is convergent in $l^2_{\mathbb{K}}$, it is also bounded, thus there is a positive α such that

$$\sum_{k=1}^{l} k^2|\xi_k^{(n)}|^2 \leq \alpha^2, \quad \forall n \in \mathbb{N},\ \forall l \in \mathbb{N}.$$

Let $l \in \mathbb{N}$ be arbitrary fixed. Since for each $k \in \mathbb{N}$,

$$\lim_n \xi_k^{(n)} = \xi_k$$

it follows that

$$\lim_n \sum_{k=1}^{l} k^2|\xi_k^{(n)}|^2 \leq \alpha^2, \text{ thus } \sum_{k=1}^{l} k^2|\xi_k|^2 \leq \alpha^2$$

We have proved that $x \in X$.

Further, since $(T(x_n))_n$ is Cauchy, for each $\varepsilon > 0$, we can pick n_ε such that m, $n \geq n_\varepsilon$ implies $\|T(x_n) - T(x_m)\| < \varepsilon/2$, so

$$\sum_{k=1}^{\infty} k^2|\xi_k^{(n)} - \xi_k^{(m)}|^2 \leq \frac{\varepsilon^2}{4}$$

For fixed arbitrary l and $n \geq n_\varepsilon$ we have

$$\sum_{k=1}^{l} k^2|\xi_k^{(n)} - \xi_k^{(m)}|^2 \leq \frac{\varepsilon^2}{4}$$

and passing here to the limit with respect to $m \to \infty$,

$$\sum_{k=1}^{\infty} k^2|\xi_k^{(n)} - \xi_k|^2 \leq \frac{\varepsilon^2}{4}$$

Thus, $\|T(x_n) - T(x)\| < \varepsilon$, $\forall n \geq n_\varepsilon$. By the uniqueness of the limit, it results that $T(x) = y$. This ends the proof of the fact that T is a closed operator.

The linear operator T is not bounded since the sequence $(e_n)_n \subset X$, $e_n = (\delta_k^{(n)})_k$, has $\|e_n\| = 1$, $\forall n$ and, on the other hand $\|Te_n\| = n$, $\forall n$.

This is happening since the space X is not Banach.

3.4 Fixed point theorems

Definition. Let V be a map on a set X. A point $x \in X$ for which $Vx = x$ is called a *fixed point* of V.

Definition. Let (X, d) be a metric space. A map $V : X \longrightarrow X$ for which $d(Vx, Vy) \leq d(x, y)$ is called a *contraction*. If there is $\alpha < 1$ for which $d(Vx, Vy) \leq \alpha \cdot d(x, y)$, V is called a *strict contraction*.

The next result is valid in complete metric spaces, with the same proof. In order to simplify the notations and since the examples will be connected to normed spaces, we will state and prove it in Banach spaces.

Theorem 3.4.1 *(Contraction mapping principle) A strict contraction V on a Banach space $(X, \|\cdot\|)$ has a unique fixed point.*

Proof. We first notice that a contraction is automatically continuous, since for each $\varepsilon > 0$, there is $\delta_\varepsilon = \alpha^{-1}\varepsilon$ so that $\|x - y\| < \delta_\varepsilon$ implies $\|Vx - Vy\| < \varepsilon$. Now, let x_o be arbitrary and let define the sequence $(x_n)_n$ by $x_n = V^n(x_o)$ (where $V^1 = V$ and $V^{n+1} = V^n \circ V$, $\forall n$). We will prove that $(x_n)_n$ is Cauchy. For each n, we have

$$\|x_n - x_{n-1}\| = \|Vx_{n-1} - Vx_{n-2}\| \leq \alpha\|x_{n-1} - x_{n-2}\| \leq$$

$$\leq \alpha^2\|x_{n-2} - x_{n-3}\| \leq \ldots \leq \alpha^{n-1}\|x_1 - x_o\|$$

Thus if $n > m$,

$$\|x_n - x_m\| \leq \sum_{k=m+1}^{n} \|x_k - x_{k-1}\| \leq \frac{\alpha^m}{1-\alpha}\|x_1 - x_o\|$$

Given $\varepsilon > 0$, since $\alpha^n \longrightarrow 0$, there is n_ε so that $\alpha^{n_\varepsilon} < (1 - \alpha)\varepsilon$. Then, $n > m \geq n_\varepsilon$ implies that $\|x_n - x_m\| < \varepsilon$, so $(x_n)_n$ is Cauchy. Thus, $x_n \longrightarrow x$, for some x. Since V is continuous,

$$Vx = \lim_n Vx_n = \lim_n x_{n+1} = x$$

so x is a fixed point of V.

Let us prove the uniqueness of the fixed point of V. If $Vx = x$ and $Vy = y$, then

$$\|x - y\| = \|Vx - Vy\| \leq \alpha\|x - y\|$$

Since $\alpha < 1$ and $\|x - y\| \geq 0$, we conclude that $\|x - y\| = 0$, consequently, $x = y$.

Theorem 3.4.2 *Let $(X, \|\cdot\|)$ be a Banach space, T a bounded linear operator on X and y_o arbitrary in X. Suppose that $\|T\| < 1$. Then, the equation*

$$x = Tx + y_o$$

has a unique solution in X.

Proof. The operator $V : X \longrightarrow X$, defined by $Vx = Tx + y_o$ is a strict contraction, since

$$\|Vx - Vy\| = \|Tx + y_o - Ty - y_o\| =$$

$$= \|Tx - Ty\| \leq \|T\| \cdot \|x - y\|, \quad \forall x, y \in X,$$

and $\|T\| < 1$. By the Contraction mapping principle, it follows that V has a unique fixed point, thus there is a unique $x^* \in X$ so that $Tx^* + y_o = x^*$.

Application. Consider an infinite system of linear equations

$$\sum_{k=1}^{\infty} \alpha_{jk} x_k = \eta_j, \ j \in \mathbb{N}$$

We are interested in solving it, so to find a sequence $(\xi_k)_k$ such that for each $j \in \mathbb{N}$ the series $\sum_{k\geq 1} \alpha_{jk}\xi_k$ be convergent with the sum η_j. We will discuss how the above theorem can be applied to prove the existence of the solutions of an infinite system of linear equations, under certain conditions. First, notice that the system can be written

$$x_j = \sum_{k=1}^{\infty} \gamma_{jk} x_k + \eta_j, \ j \in \mathbb{N}$$

where

$$\gamma_{jk} = \begin{cases} -\alpha_{jk} & j \neq k \\ 1 - \alpha_{jj} & j = k \end{cases}$$

We suppose that there exists $\rho \in (0, 1)$ such that

$$\sum_{k=1}^{\infty} |\gamma_{jk}| \leq \rho, \quad \forall j \in \mathbb{N}$$

and that the sequence $(\eta_j)_j$ is bounded, $(\eta_j)_j \in l^\infty_{\mathbb{K}}$. Consider in the Banach space $l^\infty_{\mathbb{K}}$ the equation $x = Tx + y_o$, where T is the linear operator on $l^\infty_{\mathbb{K}}$ defined by

$$T(\xi_j)_j = (\sum_{k=1}^{\infty} \gamma_{jk}\xi_k)_j$$

and $y_o = (\eta_j)_j$. Since for each $j \in \mathbb{N}$

$$|\sum_{k=1}^{\infty} \gamma_{jk}\xi_k| \leq \sum_{k=1}^{\infty} |\gamma_{jk}| \cdot |\xi_k| \leq \left(\sum_{k=1}^{\infty} |\gamma_{jk}|\right) \|(\xi_k)_k\| \leq \rho\|(\xi_k)_k\|$$

it follows that T is well defined, is bounded and its norm is less than ρ, thus $\|T\| < 1$. By the Theorem 3.4.2, we can infer that the equation $x = Tx + y_o$ (so, the infinite system of linear equations $x_j = \sum_{k=1}^{\infty} \gamma_{jk}x_k + \eta_j$, $j \in \mathbb{N}$) has a unique solution in $l^\infty_{\mathbb{K}}$.

Next we prove a version of the Leray-Schauder-Tychonoff theorem (which states that each continuous map on a nonempty compact convex subset of a locally convex space (Chapter 7) has a fixed point). First, we make a definition.

Definition. Let X and Y be vector spaces, A a convex subset of X. A map $T : A \longrightarrow Y$ is called an *affine* linear map on A if

$$T(tx + (1-t)y) = tT(x) + (1-t)T(y), \ \forall x, y \in A, \ \forall t, \ 0 \leq t \leq 1$$

Theorem 3.4.3 *Let A be a nonempty compact convex subset of a normed space X. Let T be a continuous affine map of A into itself. Then, T has a fixed point.*

Proof. We pick an arbitrary x_o in A and define the sequence $(x_n)_n$ by

$$x_n = \frac{1}{n}\sum_{j=0}^{n-1} T^j(x_o)$$

Since A is convex, each $x_n \in A$. As A is compact, some subsequence $(x_{n'})_{n'}$ of $(x_n)_n$ converges to a limit x. We wish to show that x is a fixed point of T. Suppose that $Tx - x \neq 0$. By Corollary 3.2.1, there exists $f \in X^*$ such that $f(Tx - x) \neq 0$. Since A is compact and f is continuous on X,

$$\sup_{x\in A} |f(x)| = M < \infty$$

Thus,

$$|f(Tx_n - x_n)| = |f(\frac{1}{n}T^n x_o - \frac{1}{n}x_o)| \leq \frac{2}{n}M$$

and we conclude that

$$\lim_n |f(Tx_n - x_n)| = 0$$

As a result $f(Tx-x) = \lim_n |f(Tx_n - x_n)| = 0$, which contradicts $f(Tx-x) \neq 0$.

The last fixed-point theorem that we consider deals with a whole family of maps.

Theorem 3.4.4 *(The Markov-Kakutani theorem) Let $(X, \|\cdot\|)$ be a normed space and A a nonempty compact convex subset of X. Let $\mathcal{U}$ be a family of commuting affine maps of A into itself; that is $UVx = VUx$ for all $U, V \in \mathcal{U}$ and $x \in A$. Then, $\mathcal{U}$ has a common fixed point (there exists an $x \in A$ so that $Ux = x$, $\forall U \in \mathcal{U}$).*

Proof. For each finite subset $\mathcal{F} \subset \mathcal{U}$, let

$$A_{\mathcal{F}} = \{x \in A \mid Ux = x \text{ for all } U \in \mathcal{F}\} = \bigcap_{U \in \mathcal{F}} (U - I)^{-1}\{0\}$$

(where I is the identity operator). Since the U are all continuous, each $A_{\mathcal{F}}$ is closed, and clearly $A_{\mathcal{F}_1} \cap A_{\mathcal{F}_2} = A_{\mathcal{F}_1 \cup \mathcal{F}_2}$. Thus, if we can show each $A_{\mathcal{F}}$ is nonempty, $\bigcap_{\mathcal{F}} A_{\mathcal{F}} \neq \emptyset$, by the finite intersection property, so there is an x with $Ux = x$ for all $U \in \mathcal{U}$.

We have only to prove that $A_{\mathcal{F}} \neq \emptyset$, $\forall \mathcal{F} \subset \mathcal{U}$, $\mathcal{F}$ finite. We proceed by induction with respect to the number of sets in $\mathcal{F}$. If $\mathcal{F} = \{U\}$,

$$A_{\{U\}} = \{x \in A \mid Ux = x\},$$

which is nonempty, by the previous theorem. Suppose that $A_{\mathcal{F}} \neq \emptyset$ and let $V \in \mathcal{U}$. Since the $U \in \mathcal{F}$ are affine linear $A_{\mathcal{F}}$ is convex. In addition $V(A_{\mathcal{F}}) \subset A_{\mathcal{F}}$, because $U \in \mathcal{F}$ implies $U(Vx) = V(Ux) = Vx$ when $Ux = x$. Since $A_{\mathcal{F}}$ is nonempty, convex, compact, and $V : A_{\mathcal{F}} \longrightarrow A_{\mathcal{F}}$, there is an $x \in A_{\mathcal{F}}$ with $Vx = x$, that is $A_{\mathcal{F} \cup \{V\}} \neq \emptyset$. The theorem is proved.

Remark. The both above theorems are valid in a larger context, such as locally convex spaces (Chapter 7), with the same proof.

3.5 Compact operators

In this section we exhibit an important class of linear operators which especially arise in the study of integral equations. Let X, Y be normed spaces.

Definition. A linear operator T from X to Y is said to be *compact* if T takes bounded sets in X into relative compact sets in Y.

Remark. *By the characterization of the compact sets in metric spaces, one may say that a linear operator T from X to Y is compact if and only if for each bounded sequence $(x_n)_n$ in X, the sequence $(Tx_n)_n$ has a subsequence which converges in Y. So, in order to prove that T is compact we have to see that for each sequence $(x_n)_n$ in X, with $\|x_n\| = 1$ the sequence$(Tx_n)_n$ has a subsequence which converges in Y.*

Examples. 1. Let T be in $\mathcal{B}(X, Y)$, of *finite rank* (that means, the subspace $T(X)$ of Y is finite dimensional), then T is compact. (In particular, each bounded linear operator from X to a finite dimensional space is compact.) Indeed, let $(x_n)_n$ be in X, with $\|x_n\| = 1$. Then, the sequence $(Tx_n)_n$ is a bounded sequence in a finite dimensional space, hence, by Theorem 2.3.2, it has a convergent subsequence.

2. The integral operator, $T : \mathcal{C}_{\mathbb{K}}([0,1]) \longrightarrow \mathcal{C}_{\mathbb{K}}([0,1])$,

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\,\mathrm{dt}, \quad x \in \mathcal{C}_{\mathbb{K}}([0,1]),$$

where $k \in \mathcal{C}_{\mathbb{K}}([0,1] \times [0,1])$, is compact.

Let A be a bounded set in $\mathcal{C}_{\mathbb{K}}([0,1])$, so there is $\alpha > 0$ such that $\|x\| \leq \alpha$, $\forall x \in A$. We have to prove that $T(A)$ is relative compact in $\mathcal{C}_{\mathbb{K}}([0,1])$. By Ascoli's theorem (Appendix A), it is enough to check that $T(A)$ is uniformly bounded and equicontinuous. We have,

$$|Tx(s)| \leq \int_0^1 |k(s,t)| \cdot |x(t)|\,\mathrm{dt} \leq \|x\| \cdot \int_0^1 |k(s,t)|\,\mathrm{dt}$$

thus, since A is bounded and

$$\sup_{s \in [0,1]} \int_0^1 |k(s,t)|\,\mathrm{dt} < \infty,$$

$T(A)$ is bounded.

We now prove that $T(A)$ is equicontinuous at an arbitrary $s_o \in [0,1]$. As $k : [0,1]\times[0,1]$ is uniformly continuous, it follows that for given $\varepsilon > 0$, there is $\delta_\varepsilon > 0$ such that $|s-s'| < \delta_\varepsilon$ and $|t-t'| < \delta_\varepsilon$ implies $|k(s,t)-k(s',t')| < \varepsilon/2\alpha$. Then, if $|s - s_o| < \delta_\varepsilon$,

$$|Tx(s) - Tx(s_o)| \leq \int_0^1 |k(s,t) - k(s_o,t)| \cdot |x(t)| \, \mathrm{dt} \leq \alpha \cdot \frac{\varepsilon}{2\alpha} < \varepsilon.$$

3. The identity operator I on an infinite normed space is not compact. Indeed, if one suppose that I is compact, it results that the unit ball in X is relative compact, which contradicts the Riesz theorem (2.3.2).

Notation. Further, the set of all compact operators from X to Y is denoted by $\mathcal{K}(X,Y)$. For $\mathcal{K}(X,X)$ we use the shorthand $\mathcal{K}(X)$.

Proposition 3.5.1 *1) $\mathcal{K}(X,Y)$ is a linear subspace of the space $\mathcal{B}(X,Y)$. 2) If $S \in \mathcal{K}(X)$ and $T \in \mathcal{B}(X)$, the operators ST and TS are compact.*

Proof. 1) By the definition, obviously, each compact operator is bounded. Let S, $T \in \mathcal{K}(X,Y)$, and $(x_n)_n$ a sequence in X, $\|x_n\| = 1$. Then, $(Sx_n)_n$ has a subsequence convergent $(Sx_{n'})_{n'}$. Since, T is compact $(Tx_{n'})_{n'}$ has a convergent subsequence $(Tx_{n''})_{n''}$. Thus $((S+T)x_{n''})_{n''}$ converges and $S + T \in \mathcal{K}(X,Y)$.

2) Given $(x_n)_n$ a sequence in X, $\|x_n\| = 1$, the sequence $(Tx_n)_n$ is bounded. It follows, from the compacity of S that $(STx_n)_n$ has a subsequence convergent, i.e. ST is compact.

Given $(x_n)_n$ a sequence in X, $\|x_n\| = 1$, the sequence $(Sx_n)_n$ has a subsequence convergent $(Sx_{n'})_{n'}$. Therefore, by the continuity of T, $(TSx_{n'})_{n'}$ converges.

Theorem 3.5.1 *Suppose Y is a Banach space. Then, $\mathcal{K}(X,Y)$ is a closed subspace of $(\mathcal{B}(X,Y), \|\cdot\|)$.*

Proof. Let $(T_n)_n$ be a sequence of compact operators which converges in $\mathcal{B}(X,Y)$ to T and $(x_n)_n$ a sequence in X, $\|x_n\| = 1$. In order to show that $(Tx_n)_n$ has a subsequence convergent, we employ a diagonalization procedure, as follows. Since, T_1 is compact, there exists a subsequence $(x_{1n})_n$ of $(x_n)_n$ such that $(T_1x_{1n})_n$ converges. Since, T_2 is compact, there exists a subsequence $(x_{2n})_n$ of $(x_{1n})_n$ such that $(T_2x_{2n})_n$ converges. Continuing in this manner, we obtain, for each integer $k \geq 2$, a subsequence $(x_{kn})_n$ of

$(x_{(k-1)n})_n$ such that $(T_k x_{kn})_n$ converges. We claim that , the "diagonal" sequence $(Tx_{nn})_n$ converges, so T is compact. Let us simplify the notations, by setting $v_n = x_{nn}$.

Now, for arbitrary, j, l, $n \in \mathbb{N}$ we have:

$$\|Tv_j - Tv_l\| \leq \|Tv_j - T_n v_j\| + \|T_n v_j - T_n v_l\| + \|T_n v_l - Tv_l\|$$

Given $\varepsilon > 0$, since $(T_n)_n \longrightarrow T$, $\|T_{n_\varepsilon} - T\| < \varepsilon/3$, for some n_ε. Now, $(T_{n_\varepsilon} v_j)_j$ is convergent, thus there exists $j_\varepsilon \in \mathbb{N}$ such that

$$\|T_{n_\varepsilon} v_j - T_{n_\varepsilon} v_l\| < \varepsilon/3, \quad \forall j, l \geq j_\varepsilon$$

Since, for $\forall j$, $l \geq j_\varepsilon$,

$$\|Tv_j - Tv_l\| \leq 2\|T - T_{n_\varepsilon}\| + \|T_{n_\varepsilon} v_j - T_{n_\varepsilon} v_l\| < \varepsilon,$$

it follows, from the completeness of Y that $(Tv_n)_n$ converges, i.e. T is compact.

3.6 Exercises

1. Let X, Y be normed space and $T : X \longrightarrow Y$ a linear operator. The following are equivalent:

(i) $T \in \mathcal{B}(X, Y)$;

(ii) $\exists G \subset X$, $G \neq \emptyset$, G open such that the set $T(G)$ is bounded;

(iii) $\forall (x_n)_n \longrightarrow 0$, the sequence $(T(x_n))_n$ is bounded.

2 Let T be a linear operator from the Banach space X to some normed space Y. Show that $T \in \mathcal{B}(X, Y)$ if and only if $\forall x \in X$ if $(x_n)_n \longrightarrow x$,then

$$\|T(x)\| \leq \liminf_n \|T(x_n)\|$$

3. Let X, Y be normed space such that $\mathcal{B}(X, Y)$ is a Banach space. Then, Y is Banach.

4. Show that the operator T defined on $\mathcal{C}_{\mathbf{K}}([0,1])$ by

$$(Tx)(s) = \int_0^1 e^{s-t} x(t)\, \mathrm{dt}$$

is in $\mathcal{B}(\mathcal{C}_{\mathbf{K}}([0,1]))$ and find its norm.

5. The same problem as **4.** if the operator T is defined on $\mathcal{C}_{\mathbb{K}}([0,1])$ by

$$(Tx)(s) = \int_0^1 s^n t^m x(t)\,\mathrm{dt}$$

or by

$$(Tx)(s) = \int_0^1 x(t)\ \sin\pi(s-t)\,\mathrm{dt}$$

6. Let T be the operator defined on $L^2_{\mathbb{K}}([0,1])$ by $Tx(t) = tx(t)$.
a) Show that $T \in \mathcal{B}(L^2_{\mathbb{K}}([0,1]))$and find $\|T\|$;
b) Show that the operator T is not compact.
7. Let T be the operator defined on $L^2_{\mathbb{K}}([0,1])$ by

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\,\mathrm{dt}$$

where $k \in L^2_{\mathbb{K}}([0,1]\times[0,1])$. Show that:
a) $T \in \mathcal{B}(L^2_{\mathbb{K}}([0,1]))$ and

$$\|T\| \le \left(\int_0^1 \int_0^1 |k(s,t)|^2\,\mathrm{ds}\,\mathrm{dt}\right)^{\frac{1}{2}}$$

b) T is compact.
8. Let T be the operator defined on $l^p_{\mathbb{K}}$, $p \in [1,\infty)$ by

$$T((\xi_n)_n) = (\lambda_n \xi_n)_n,$$

where $(\lambda_n)_n \subset \mathbb{K}$ is an arbitrary bounded numerical sequence. Show that:
a) $T \in \mathcal{B}(l^p_{\mathbb{K}})$ and $\|T\| = \sup\limits_n |\lambda_n|$;
b) T is compact if and only if $\lambda_n \longrightarrow 0$, as $n \to \infty$;
9. Let T be the operator from $l^\infty_{\mathbb{K}}$ to $l^2_{\mathbb{K}}$ defined by

$$T((\xi_n)_n) = \left((\frac{1}{n})\xi_n\right)_n$$

Show that T is a bounded linear well defined operator and find its norm. Is T injective? Is T surjective?

10. a) Show that the operator $T : (\mathcal{C}^2_{\mathbb{K}}([0,1]), \|\cdot\|) \longrightarrow (\mathcal{C}_{\mathbb{K}}([0,1]), \|\cdot\|)$ (where the norm is the same on the both spaces, namely $\|x\| = \sup\limits_{t\in[0,1]} |x(t)|$), defined by

$$(Tx)(t) = x''(t) + (\sin\ \pi t)x'(t) + tx(t)$$

is a linear operator that is not bounded

b) Show that the operator T from $(\mathcal{C}^2_{\mathbb{K}}([0,1]), \|\cdot\|^{\sim})$ to $(\mathcal{C}_{\mathbb{K}}([0,1]), \|\cdot\|)$ (where $\|x\|^{\sim} = \|x\| + \|x'\| + \|x''\|$, $\forall x \in \mathcal{C}^2_{\mathbb{K}}([0,1])$) defined by the same formula is bounded.

11. Let T be a linear operator from the Banach space $(X, \|\cdot\|)$ to the normed space $(Y, \|\cdot\|)$. Set $p(x) = \|x\| + \|Tx\|$ (clearly p defines a norm on X). The following are equivalent:

(i) $T \in \mathcal{B}(X,Y)$;

(ii) The initial norm on X, $\|\cdot\|$, and the norm p are equivalent;

(iii) (X,p) is Banach.

12. If $X = \mathcal{C}^1_{\mathbb{K}}([0,1])$ (the space of all differentiable real-valued functions with the first derivative continuous with the norm inherited from $\mathcal{C}_{\mathbb{K}}([0,1])$) and $Y = \mathcal{C}_{\mathbb{K}}([0,1])$, let D be the linear mapping from X to Y defined by $Dx = x'$ (where x' is the derivative of x). Show that D is closed and it is not bounded. Does this example contradict the closed graph theorem?

13. Let T be a linear operator from the Banach space X to the Banach space Y such that for every $f \in Y^*$, the functional $f \circ T$ is in X^*. Show that $T \in \mathcal{B}(X,Y)$.

14. Prove that a Banach space is reflexive if and only if its dual is reflexive. Show that $l^{\infty}_{\mathbb{K}}$ is not reflexive.

15. Prove that $\mathcal{C}_{\mathbb{K}}([0,1]), \|\cdot\|)$ is not reflexive.

16. Let $p \in [1,\infty)$ be. Show that $(\mathcal{C}_{\mathbb{K}}([0,1]), \|\cdot\|_p)$ is not complete and its completion is $L^p_{\mathbb{K}}([0,1])$.

17. Let $(X, \|\cdot\|)$ be a normed space such that there exists T a compact operator on X invertible in $\mathcal{B}(X)$. Show that X is finite dimensional.

18. Let $(X, \|\cdot\|)$ be a normed space. Show that X is finite dimensional if and only if X^* is finite dimensional.

19. Let $(X, \|\cdot\|)$ be a normed space and x in X arbitrary. Show that

$$\|x\| = \sup_{f \in X^*, \|f\| \le 1} |f(x)|$$

20. Let f be the functional on $c_{\mathbb{K}}$ defined by $f((\xi_n)_n) = \lim_n \xi_n$. Show that $f \in c^*_{\mathbb{K}}$ and $\|f\| = 1$. Infer from here that $c^o_{\mathbb{K}}$ is Banach.

21. Let $(X, \|\cdot\|)$ be a normed space and f in X^*, $f \neq 0$. Show that, for each $x \in X$,

$$d(x, \text{Ker } f) = \frac{|f(x)|}{\|f\|}$$

22. Let f be the functional on $c^o_{\mathbf{K}}$ defined by

$$f((\xi_n)_n) = \sum_{m=1}^{\infty} \frac{1}{2^{m-1}} \xi_m$$

Show that

a) $f \in (c^o_{\mathbf{K}})^*$ and $\|f\| = 2$;

b) $\forall x \in \text{Ker } f$, $d(x, \text{Ker } f) \neq \|x - y\|$, $\forall y \in \text{Ker } f$.

Chapter 4

Hilbert Spaces

4.1 Definition of Hilbert space and elementary properties

Definition. A vector space X over the field $\mathbb{K}$ is called an *inner product space* if there is a $\mathbb{K}$-valued function $< \cdot, \cdot >$ on $X \times X$ such that:
i) $< \alpha x + \beta y, z >= \alpha < x, z > +\beta < y, z >$, $\forall\, x,\, y,\, z \in X$, $\alpha,\, \beta \in \mathbb{K}$;
ii) $< x, y >= \overline{< y, x >}$, $\forall\, x,\, y \in X$;
iii) $< x, x >> 0$, $\forall x \in X,\ x \neq 0$.
The function $< \cdot, \cdot >$ is called an inner product.

Remark. We note that the previous properties of $< \cdot, \cdot >$ imply
1) $< x, \alpha y + \beta z >= \overline{\alpha} < x, y > +\overline{\beta} < x, z >$, $\forall\, x,\, y,\, z \in X$, $\alpha,\, \beta \in \mathbb{K}$;
2) $< x, x > \geq 0$, $\forall x \in X$;
3) $< x, 0 >= 0$, $\forall x \in X$.
We introduce the shorthand $\|x\| = \sqrt{< x, x >}$. We will shortly see that $\|\cdot\|$ is in fact a norm.

Examples. 1. For $x = (\xi_1, \xi_2, ..., \xi_n)$ and $y = (\eta_1, \eta_2, ..., \eta_n)$ in $\mathbb{K}^n$ define

$$< x, y >= \sum_{i=1}^{n} \xi_i \overline{\eta_i}$$

Then, $< \cdot, \cdot >$ is an inner product on $\mathbb{K}^n$.

2. The $\mathbb{K}$-valued function $< \cdot, \cdot >$ defined on $l^2_{\mathbb{K}} \times l^2_{\mathbb{K}}$ by

$$< x, y >= \sum_{i=1}^{\infty} \xi_i \overline{\eta_i}$$

where $x = (\xi_n)_n$, $y = (\eta_n)_n$ is an inner product.

3. Let us define for arbitrary elements x, $y \in L^2_{\mathbb{K}}([a, b])$

$$< x, y >= \int_a^b x(t)\overline{y(t)}\, \mathrm{dt}$$

Then, $(L^2_{\mathbb{K}}([a, b]), < \cdot, \cdot >)$ is an inner product space.

Proposition 4.1.1 *Let X be an inner product space and x, y in X. Then:*
1) $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ *(Parallelogram law).*
2) When $\mathbb{K} = \mathbb{R}$

$$< x, y >= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

and when $\mathbb{K} = \mathbb{C}$

$$< x, y >= \frac{1}{4} \sum_{k=0}^{3} i^k \|x + i^k y\|^2$$

(Polarization identities).

Proof. The equalities are immediate from the properties of inner product.

We now develop those geometrical notions that extend from finite dimensional spaces to arbitrary inner product spaces.

Definition. Two vectors, x and y, in an inner product space X are called *orthogonal*, written $x \perp y$, if $< x, y >= 0$. The vector x is said to be *orthogonal to a set* $A \subset X$ if $x \perp y$ for all y in A; we denote by $A^{\perp}$ the set of vectors in X which are orthogonal to A. A family $\{x_{\iota} \mid \iota \in I\}$ of vectors in X is called orthogonal if $x_{\iota} \perp x_{\tau}$, $\iota \neq \tau$. If in addition, $\|x_{\iota}\| = 1$, $\forall \iota$, then the family is called *orthonormal.*
Remark. An orthonormal family $\{x_{\iota} \mid \iota \in I\}$ of vectors in X is linearly independent since $0 = \sum_{k=1}^{n} \alpha_k x_k$ implies $0 =< \sum_{k=1}^{n} \alpha_k x_k, x_k >= \alpha_k$.

Proposition 4.1.2 *(Pythagoras identity) Let X be an inner product space and x, y in X. If $x \perp y$, then*

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

Proof. For x, y in X, $x \perp y$, we have

$$\|x+y\|^2 = < x+y, x+y > =$$

$$= \|x\|^2 + < x, y > + < y, x > + \|y\|^2 = \|x\|^2 + \|y\|^2$$

Theorem 4.1.1 *Let $\{x_k\}_{1 \leq k \leq n}$ be an orthonormal family in an inner product space X. Then for all $x \in X$,*

$$\|x\|^2 = \sum_{k=1}^{n} | < x, x_k > |^2 + \|x - \sum_{k=1}^{n} < x, x_k > \|^2$$

Proof. We write x as

$$x = \sum_{k=1}^{n} < x, x_k > x_k + (x - \sum_{k=1}^{n} < x, x_k > x_k)$$

A short computation based on the properties of the inner product shows that $x - \sum_{k=1}^{n} < x, x_k > x_k$ and $\sum_{k=1}^{n} < x, x_k > x_k$ are orthogonal. Thus, by the Pythagoras identity,

$$\|x\|^2 = \| \sum_{k=1}^{n} < x, x_k > x_k \|^2 + \|x - \sum_{k=1}^{n} < x, x_k > x_k \|^2$$

Taking into account that the family $\{x_k\}_{k=1}^{n}$ is orthonormal, we have

$$\| \sum_{k=1}^{n} < x, x_k > x_k \|^2 = \sum_{k=1}^{n} | < x, x_k > |^2$$

which proves the equality.

Next result follows evidently by the above theorem:

Corollary 4.1.1 *(Bessel's inequality) Let $\{x_k\}_{1 \leq k \leq n}$ be an orthonormal family in an inner product space X. Then for all $x \in X$,*

$$\|x\|^2 \geq \sum_{k=1}^{n} | < x, x_k > |^2$$

Corollary 4.1.2 *(The Cauchy Schwarz inequality) If x and y are vectors in an inner product space X, then*

$$| < x, y > | \leq \|x\| \cdot \|y\|$$

Proof. *The case $y = 0$ is trivial, so suppose $y \neq 0$. The vector $y/\|y\|$ by itself forms an orthogonal set, so applying Bessel's inequality to any $x \in X$ we get*

$$\|x\|^2 \geq | < x, \frac{y}{\|y\|} > |^2 = \frac{| < x, y > |^2}{\|y\|^2}$$

Theorem 4.1.2 *Every inner product space X is a normed linear space with the norm $\|x\| =< x, x >^{\frac{1}{2}}$.*

Proof. All the properties of the norm, except the triangle inequality, for $\|\cdot\|$ follow immediately from the properties of inner products. Suppose x, $y \in X$. Then,

$$\|x + y\|^2 =< x, x > + < x, y > + < y, x > + < y, y >=$$

$$=< x, x > +2Re \ < x, y > + < y, y >\leq$$

$$\leq< x, x > +2| < x, y > |+ < y, y >\leq$$

$$\leq< x, x > +2 < x, x >^{\frac{1}{2}}< y, y >^{\frac{1}{2}} + < y, y >$$

by the Cauchy-Schwarz inequality. Thus,

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2$$

This theorem shows that we have a natural metric,

$$d(x, y) =< x - y, x - y >^{\frac{1}{2}}$$

in X. We thus have the notions of convergence, completeness and density defined for metric spaces.

Definition. A complete inner product space is called a *Hilbert* space.

Remark. *We note that an inner product space is a Hilbert space if $(X, \|\cdot\|)$ is a Banach space.*

Remark. In a Hilbert space the application $(x, y) \longmapsto< x, y >$ from $X \times X$ to $\mathbb{K}$ is continuous. Indeed, let $(x_n)_n$ and $(y_n)_n$ be such that $x =\lim\limits_n x_n$ and $y =\lim\limits_n y_n$. By the Cauchy-Schwarz inequality

$$| < x_n, y > - < x, y > | = | < x_n - x, y > | \leq \|x_n - x\| \cdot \|y\|$$

thus, $(< x_n, y >)_n$ converges to $< x, y >$. By

$$| < x_n - x, y_n - y > | \leq \|x_n - x\| \cdot \|y_n - y\|$$

it results that $< x_n - x, y_n - y > \longrightarrow 0$, as $n \to \infty$. Everything is clear now by the next equality

$$< x_n, y_n > - < x, y >=$$

$$=< x_n - x, y_n - y > + < x_n, y > + < x, y_n > -2 < x, y >$$

Examples. **1.** $\mathbb{K}^n$ is a n-dimensional Hilbert space.

2. $l^2_{\mathbf{K}}$ is an infinite dimensional Hilbert space.

3. $L^2_{\mathbf{K}}([a, b])$ is an infinite dimensional Hilbert space, since the functions 1, t, $t^2, \ldots..$ are linearly independent.

4. The inner product space $\mathcal{P}$ of all polynomials with

$$< x, y >= \int_a^b x(t)\overline{y(t)}\, \mathrm{dt}$$

is not complete. To see this, let

$$P_n(t) = \sum_{j=0}^{n} \frac{1}{2^j} t^j$$

Then, $(P_n)_n$ converges in $L^2_{\mathbf{K}}([a, b])$ to $y(t) = 1/(1 - \frac{1}{2}t)$. So, $(P_n)_n$ is a Cauchy sequence in $\mathcal{P}$ which does not converge to a vector in $\mathcal{P}$ since $y \notin \mathcal{P} \subset \subset L^2_{\mathbf{K}}([0, 1])$.

The next two theorems will be useful in the Gram-Schmidt ortogonalization procedure (see Theorem 4.3.3).

Theorem 4.1.3 *Let $\{x_k\}_{1 \leq k \leq n}$ be an orthonormal family in an inner product space X and Y the linear space spanned by the set $\{x_k\}_{1 \leq k \leq n}$. Then, for each $x \in X$, the vector $y_x = \sum_{k=1}^n < x, x_k > x_k$ is the unique vector in Y with the property $d(x, Y) = \|x - y_x\|$ (the closest element to x). In addition $d(x, Y) = (\|x\|^2 - \|x - \sum_{k=1}^n < x, x_k > x_k\|^2)^{\frac{1}{2}}$ and $x - y_x \perp Y$*

Proof. Take an arbitrary element in Y, $z = \sum_{k=1}^n \alpha_k x_k$, $\{\alpha_k\}_{1 \leq k \leq n} \subset \mathbb{K}$. Then

$$\|x - z\|^2 = \|x - \sum_{k=1}^{n} \alpha_k x_k\|^2 =< x - \sum_{k=1}^{n} \alpha_k x_k, x - \sum_{k=1}^{n} \alpha_k x_k >=$$

$$= \|x\|^2 - < x, \sum_{k=1}^{n} \alpha_k x_k > - < \sum_{k=1}^{n} \alpha_k x_k, x > + < \sum_{k=1}^{n} \alpha_k x_k, \sum_{k=1}^{n} \alpha_k x_k > =$$

$$= \|x\|^2 - \sum_{k=1}^{n} \alpha_k \overline{< x, x_k >} + \sum_{k=1}^{n} \overline{\alpha_k} < x, x_k > + \sum_{k=1}^{n} \alpha_k \overline{\alpha_k} =$$

$$= \|x\|^2 - \sum_{k=1}^{n} < x, x_k > \overline{< x, x_k >} + \sum_{k=1}^{n} (\alpha_k - < x, x_k >)(\overline{\alpha_k} - \overline{< x, x_k >}) =$$

$$= \|x\|^2 - \sum_{k=1}^{n} | < x, x_k > |^2 + \sum_{k=1}^{n} |\alpha_k - < x, x_k > |^2 =$$

$$= \|x\|^2 - \sum_{k=1}^{n} | < x, x_k > |^2 + \|y_x - z\|^2$$

Thus,

$$d(x, Y) = \inf_{z \in Y} \|x - z\| \geq (\|x\|^2 - \sum_{k=1}^{n} | < x, x_k > |^2)^{\frac{1}{2}}$$

and, the equality holds if and only if $z = y_x$.

A short computation based on the properties of inner products shows that $x - \sum_{k=1}^{n} < x, x_k > x_k$ is orthogonal to Y.

Theorem 4.1.4 *Let Y be a subspace of X. Suppose $x \in X$ and $y \in Y$. Then $x - y \perp Y$ if and only if $\|x - y\| = d(x, Y)$.*

Proof. Let z be arbitrary in Y, $z \neq y$. If $x - y \perp Y$, by the Pythagoras identity,

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

so, $\|x - y\| = d(x, Y)$.

Conversely, suppose $\|x - y\| \leq \|x - z\|$ for all $z \in Y$. Since Y is a subspace, $y + \lambda z$ is in Y for all $z \in Y$ and $\lambda \in \mathbb{K}$. Therefore,

$$\|x - y\|^2 \leq \|x - (y + \lambda z)\|^2 = < x - y - \lambda z, x - y - \lambda z > =$$

$$= \|x - y\|^2 - 2Re\ \lambda < z, x - y > + |\lambda|^2 \|z\|^2$$

Hence,

$$2Re\ \lambda < z, x - y > \leq |\lambda|^2 \|z\|^2$$

Set $\lambda = r\overline{< z, x - y >}$, where r is a real number. We get

$$2r| < z, x - y > |^2 \leq r^2 | < z, x - y > |^2 \|z\|^2,$$

and, since r is arbitrary, it follows that $< z, x - y > = 0$.

4.2 Projections onto subspaces

Proposition 4.2.1 *For every subset A in X, $A^{\perp}$ is a closed subspace of X and $A^{\perp} = \overline{SpA}^{\perp}$.*

Proof. The fact that $A^{\perp}$ is a subspace of X follows from the linearity of the inner product. Let $x \in \overline{A^{\perp}}$, so $x_n \longrightarrow x$, for some sequence $(x_n)_n$ in $A^{\perp}$. For arbitrary $y \in A$, $< x_n, y > \longrightarrow < x, y >$. As $< x_n, y >= 0$ it results that $< x, y >= 0$, therefore $x \in A^{\perp}$. The equality $A^{\perp} = \overline{SpA}^{\perp}$ can be immediately proven by the same kind of arguments.

Next, let Y be a closed subspace of X. The closed subspace $Y^{\perp}$ is called the *orthogonal complement* of Y. The following theorem shows that there are vectors orthogonal to any closed proper subspace, indeed there are enough such that $X = \{y+z \mid y \in Y,\ z \in Y^{\perp}\}$. This important geometric property is one of the main reasons that Hilbert spaces are easier to handle than Banach spaces.

Lemma 4.2.1 *Let X be a Hilbert space and Y a closed subspace of X. Then for each $x \in X$ there is in Y a unique element closest to x.*

Proof. Let us denote by $d = d(x, Y)$. Choose a sequence $(y_n)_n \subset Y$, so that

$$d \leq \|x - y_n\| < d + \frac{1}{n}$$

Then for $n, p \in \mathbb{N}$,

$$\|y_{n+p} - y_n\|^2 = \|(x - y_{n+p}) - (x - y_n)\|^2 =$$

$$= 2\|x - y_{n+p}\|^2 + 2\|x - y_n\|^2 - \|x - y_{n+p} + x - y_n\|^2 =$$

$$= 2\|x - y_{n+p}\|^2 + 2\|x - y_n\|^2 - 4\|x - \frac{y_{n+p} + y_n}{2}\|^2 \leq$$

$$\leq 4(d + \frac{1}{n})^2 - 4d^2 \longrightarrow 0, \text{ as } n \to \infty.$$

The second equality follows from the parallelogram law; the inequality follows from the fact that $(y_{n+p} + y_n)/2 \in Y$. Thus $(y_n)_n$ is Cauchy and since Y is closed, $(y_n)_n$ converges to an element y_x of Y. It results immediately that

$\|x - y_x\| = d$. Suppose z is an other element in Y such that $\|x - z\| = d$. Then

$$\|z - y_x\|^2 = \|(x - z) - (x - y_x)\|^2 =$$

$$= 2\|x - z\|^2 + 2\|x - y_x\|^2 - 4\|x - \frac{z + y_x}{2}\|^2 \leq 4d^2 - 4d^2 = 0,$$

which proves the uniqueness.

Remark. It is clear from the proof that the previous result holds in a more general situation, namely when Y is a closed convex subset of X.

Theorem 4.2.1 *(**Projection theorem**) Let X be a Hilbert space and Y a closed subspace of X. Then every $x \in X$ can be uniquely written $x = y + z$, where $y \in Y$ and $z \in Y^{\perp}$.*

Proof. Let x be in X. Then, by the above lemma there is a unique element $y_x \in Y$ closest to x. Define $z = x - y_x$, then we clearly have $x = y + z$. We shall check that $z \perp Y$, equivalently $< z, v >= 0$, $\forall v \in Y$. For arbitrary $t \in Y$ and $\lambda \in \mathbb{K}$.

$$\|z\|^2 \underset{t \in Y}{\leq} \|x - t\|^2 \underset{t = y_x + \lambda v}{=} \|z - \lambda v\|^2 =$$

$$= \|z\|^2 - \lambda \overline{< z, v >} - \overline{\lambda} < z, v > + \lambda\overline{\lambda}\|v\|^2$$

Setting $\lambda =< z, v > / \|v\|^2$, it follows

$$\|z\|^2 \leq \|z\|^2 - \frac{|< z, v >|^2}{\|v\|^2} - \frac{|< z, v >|^2}{\|v\|^2} + \frac{|< z, v >|^2}{\|v\|^2},$$

thus $|< z, v >|^2 \leq 0$ which implies that $< z, v >= 0$, $\forall v \in Y$.

Uniqueness is left as exercise.

As an immediate consequence of this theorem we have:

Corollary 4.2.1 *Let X be a Hilbert space and Y a closed proper subspace of X. Then $Y^{\perp} \neq \{0\}$.*

Theorem 4.2.2 *(**The Riesz lemma**) Let X be a Hilbert space and f a continuous linear functional on X ($f \in X^*$). Then*

1) There exists an element a_f in X such that $f(x) =< x, a_f >$, $\forall x \in X$;

2) The element a_f is unique with the property 1);

3) $\|f\| = \|a_f\|$.

Proof. For $f \in X^*$, Ker f is a closed subspace of X. The case Ker $f = X$ is trivial since $f \equiv 0$ and we set $a_f = 0$. So suppose there exists $z \neq 0$, $z \in$ Ker $f^{\perp}$. For arbitrary $x \in X$, the element

$$x - \frac{f(x)}{f(z)} z \in \text{Ker } f$$

therefore

$$< x - \frac{f(x)}{f(z)} z, z >= 0$$

It results that

$$< x, z > - \frac{f(x)}{f(z)} \|z\|^2 = 0, \text{ or equivalently } f(x) =< x, \frac{\overline{f(z)}}{\|z\|^2} z > .$$

Setting

$$a_f = \frac{\overline{f(z)}}{\|z\|^2} z$$

1) is proved. The uniqueness is obvious (2)). 3) follows by

$$|f(x)| = | < x, a_f > | \leq \|x\| \cdot \|a_f\| \Rightarrow \|f\| \leq \|a_f\|$$

and, on the other hand, by

$$\|f\| = \sup_{\|x\|=1} |f(x)| \geq f(\frac{1}{\|a_f\|} a_f) = \|a_f\|$$

Remark. We note that the Cauchy-Schwarz inequality shows that the converse of the Riesz lemma is true. Namely, each $y \in X$ defines a continuous linear functional f_y on X by $f_y(x) =< x, y >$.

4.3 Orthonormal bases

We shall begin with some preliminaries facts.

Lemma 4.3.1 *Let X be a Hilbert space and $(x_n)_n$ an orthogonal sequence in X such that the series $\sum_{n\geq 1} \|x_n\|^2$ converges. Then, the series $\sum_{n\geq 1} x_n$ is unconditional convergent. Moreover, the sum of the series $\sum_{n\geq 1} x_n$ does not depend of the order of the terms.*

Proof. If $s_n = \sum_{k=1}^n x_k$ and $t_n = \sum_{k=1}^n \|x_k\|^2$, by the Pythagoras identity,

$$\|s_{n+p} - s_n\|^2 = t_{n+p} - t_n, \quad \forall n,\ p \in \mathbb{N}$$

which shows that $(s_n)_n$ is a Cauchy sequence, and since X is a Hilbert space, the series $\sum_{n\geq 1} x_n$ converges. Take, σ a permutation on $\mathbb{N}$. Because the numerical series $\sum_{n\geq 1} \|x_n\|^2$ is unconditional convergent, we have, as before that $\sum_{n\geq 1} x_{\sigma(n)}$ converges. We only must verify that $x = y$, where

$$x = \sum_{n=1}^{\infty} x_n \text{ and } y = \sum_{n=1}^{\infty} x_{\sigma(n)}$$

Clearly, $x - y \in \overline{\mathrm{Sp}\{x_n \mid n \in \mathbb{N}\}}$. On the other hand, for each $m \in \mathbb{N}$,

$$< x - y, x_k >= \lim_{n\to\infty} < \sum_{k=1}^{n} x_k - \sum_{k=1}^{n} x_{\sigma(k)}, x_m >= 0,$$

therefore, $x - y \in \overline{\mathrm{Sp}\{x_n \mid n \in \mathbb{N}\}}^{\perp}$. It follows that

$$x - y \in \overline{\mathrm{Sp}\{x_n \mid n \in \mathbb{N}\}}^{\perp} \cap \overline{\mathrm{Sp}\{x_n \mid n \in \mathbb{N}\}},$$

so $x = y$.

Lemma 4.3.2 *Let X be a Hilbert space and $\{x_\iota \mid \iota \in I\}$ an orthonormal family in X. Then, for each $x \in X$, the subset of I,*

$$I_x = \{\iota \in I \mid\ < x, x_i > \neq 0\}$$

is at most countable.

Proof. Take $\varepsilon > 0$, and denote by $I_x^{(\varepsilon)} = \{\iota \in I \mid | < x, x_\iota > | \geq \varepsilon\}$. Suppose that $I_x^{(\varepsilon)}$ is an infinite set. Then, for $\forall n \in \mathbb{N}$, $\exists \iota_1, \iota_2, ..., \iota_n \in I_x^{(\varepsilon)}$, that means $| < x, x_{\iota_k} > | \geq \varepsilon$, $\forall k = 1, ..., n$. By Bessel's inequality it results that $\sum_{k=1}^n | < x, x_{\iota_k} > |^2 \leq \|x\|^2$ which implies that $n\ \varepsilon^2 \leq \|x\|^2$, $\forall n \in \mathbb{N}$ (contradiction). It follows for each $\varepsilon > 0$, $I_x^{(\varepsilon)}$ is empty or finite, and, as $I_x = \bigcup_{n\geq 1} I_x^{(1/n)}$, the lemma is proven.

Remark. *If $\{x_\iota \mid \iota \in I\}$ is an orthonormal family in X, for each $x \in X$ we can consider the numerical family $\{< x, x_\iota > \mid \iota \in I\}$, called the family of Fourier coefficients of x with respect to the orthonormal family $\{x_\iota \mid \iota \in I\}$.*

Notations. Let X be a Hilbert space, $\{x_\iota \mid \iota \in I\}$ an orthonormal family in X, $x \in X$ and $I_x = \{\iota \in I \mid \ <x, x_i> \neq 0\}$. If I_x is finite we denote by

$$\sum_{\iota \in I} <x, x_\iota> x_\iota \ , \text{(respectively} \sum_{\iota \in I} |<x, x_\iota>|^2 \ ),$$

the sum

$$\sum_{\iota \in I_x} <x, x_\iota> x_\iota \ , \text{(respectively} \sum_{\iota \in I_x} |<x, x_\iota>|^2)$$

If I_x is not finite, then, by Lemma 4.3.2 there is a bijection $\sigma : \mathbb{N} \longrightarrow I_x$. The sequence $(<x, x_{\sigma(n)}> x_{\sigma(n)})_n$ satisfies the hypothesis of Lemma 4.3.1, since

$$\sum_{n \geq 1} \| <x, x_{\sigma(n)}> x_{\sigma(n)} \|^2 = \sum_{n \geq 1} |<x, x_{\sigma(n)}>|^2$$

and this series is (unconditional) convergent as it follows by Bessel'inequality,

$$t_n = \sum_{k=1}^{n} |<x, x_{\sigma(k)}>|^2 \leq \|x\|^2$$

It results that the series $\sum_{n \geq 1} <x, x_{\sigma(n)}> x_{\sigma(n)}$ converges in X and its sum does not depend of the choice of σ. Recall that the sum of the numerical series $\sum_{n \geq 1} |<x, x_{\sigma(n)}>|^2$ is also independent of the choice of the bijection σ, because this series converges absolutely, so unconditionally. Therefore we may denote by

$$\sum_{\iota \in I} <x, x_\iota> x_\iota, \text{(respectively} \sum_{\iota \in I} |<x, x_\iota>|^2)$$

the sum of the series

$$\sum_{n \geq 1} <x, x_{\sigma(n)}> x_{\sigma(n)}, \text{(respectively} \sum_{n \geq 1} |<x, x_{\sigma(n)}>|^2)$$

since it does not depend of the choice of σ. Finally we have to notice that

$$\sum_{\iota \in I} |<x, x_\iota>|^2 \leq \|x\|^2,$$

(*the Bessel's inequality for an arbitrary orthonormal family*).

Definition. An orthonormal subset $B = \{x_\iota \mid \iota \in I\}$ of X is said to be an *orthonormal basis* (or a *complete orthonormal system*) if there is no other orthonormal subset of X which contains B as a proper subset (that means B is a maximal orthonormal set of X with respect to the inclusion order).

Theorem 4.3.1 *Every Hilbert space $X \neq \{0\}$ has an orthonormal basis.*

Proof. Let $\mathcal{C}$ be the collection of all orthonormal sets of X. $\mathcal{C}$ is nonempty (since $X \neq \{0\}$ implies $\{(1/\|x\|)x\} \in \mathcal{C}$). We order $\mathcal{C}$ by inclusion: $B_1 \prec B_2$ if and only if $B_1 \subset B_2$. That is a partial order relation on $\mathcal{C}$. We observe that $(\mathcal{C},\prec)$ is inductively ordered (since if $(B_\alpha)_\alpha$ is a totally ordered family, $B_0 = \bigcup_\alpha B_\alpha$ is an upper bound of it). By Zorn's lemma we conclude that there exists B a maximal element of $(\mathcal{C},\prec)$, so an orthonormal basis for X.

Definition. A subset A of X is called a *total set* of X if $x \perp A \Rightarrow x = 0$.

Theorem 4.3.2 *Suppose $B = \{x_\iota \mid \iota \in I\}$ is an orthonormal set in the Hilbert space X. The following statements are equivalent:*
(1) B is an orthonormal basis;
(2) B is a total set;
(3) $\overline{\mathrm{Sp}\ B} = X$;
(4) For each $x \in X$, $x = \sum_{\iota\in I} < x, x_\iota > x_\iota$ (Fourier's development of x with respect to the basis B);
(5) For each $x \in X$, $\| x\|^2 = \sum_{\iota\in I} | < x, x_\iota > |^2$ (Parseval's equality).

Proof. By the Theorem 4.1.1, for arbitrary $n \in \mathbb{N}$, we have

$$\|x\|^2 - \sum_{k=1}^{n} | < x, x_{\sigma(k)} > |^2 = \|x - \sum_{k=1}^{n} < x, x_{\sigma(k)} > x_{\sigma(k)}\|^2$$

(where σ is a bijection of $\mathbb{N}$ onto I_x). It follows that $x = \sum_{\iota\in I} < x, x_\iota > x_\iota$ if and only if $\| x\|^2 = \sum_{\iota\in I} | < x, x_\iota > |^2$, thus (4)$\Leftrightarrow$(5).

The fact that (4)$\Rightarrow$ (3) is obvious. Suppose now $x \perp B$, therefore $x \in$ $\in B^\perp = \overline{\mathrm{Sp}\ B}^\perp$. It results that (3)$\Rightarrow$(2).

We show that (2)$\Rightarrow$ (4). Let $x \in X$, and $z = \sum_{\iota\in I} < x, x_\iota > x_\iota$. In order to see that $x = z$ we check that $x - z \perp B = \{x_\iota \mid \iota \in I\}$. If $\iota_o \in I_x$,

$$< x - z, x_{\iota_o} >=< x, x_{\iota_o} > - < \sum_{\iota\in I} < x, x_\iota > x_\iota, x_{\iota_o} >=$$

$$=< x, x_{\iota_o} > - < x, x_{\iota_o} >= 0$$

If $\iota_o \notin I_x$,

$$< x - z, x_{\iota_o} >= 0- < \sum_{\iota\in I} < x, x_\iota > x_\iota, x_{\iota_o} >= 0 - 0 = 0$$

Further, if (1) is valid, and $x \in X$, $x \neq 0$, then

$$B' = \{\frac{1}{\|x\|} x\} \cup \{x_\iota \mid \iota \in I\}$$

is an orthonormal set which contains B (one contradicts the maximality of B). It results $x = 0$. Thus, (1)$\Rightarrow$ (2). The converse implication is immediate since if B' is an orthonormal set such that $B \subsetneq B'$, there exists $x \in B' \setminus B$. It follows $x \neq 0$ and $x \perp B$ (contradiction with (1)).

Examples. **1.** In $\mathbb{K}^n$, the canonical algebraic basis $B = \{e_k\}_{1 \leq k \leq n}$, $e_k = (\delta_1^{(k)}, \delta_2^{(k)}, ..., \delta_n^{(k)})$ (where, for $j, k \in \mathbb{N}$, $\delta_j^{(k)} = 1$ if $j \neq k$ and $\delta_j^{(k)} = 0$ otherwise) is obviously an orthonormal set and also a total set of the Hilbert space $\mathbb{K}^n$. Thus, it is an orthonormal basis, too.

2. In $l^2_{\mathbb{K}}$ the set $B = \{e_n \mid n \in \mathbb{N}\}$ where $e_n = (\delta_k^{(n)})_k$, is an orthonormal set. B is also a total set, since if $x = (\xi_k)_k \in l^2_{\mathbb{K}}$ is orthogonal to B, it results that $< x, e_n >= 0$, $\forall n \in \mathbb{N}$. As $< x, e_n >= \xi_n$, it follows $\xi_n = 0$, $\forall n \in \mathbb{N}$, so $x = 0$. By the previous theorem, we conclude that $B = \{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis in $l^2_{\mathbb{K}}$.

3. In $L^2_{\mathbb{K}}([-\pi, \pi])$ the countable set $B = \{e_0, e_{11}, e_{12}, e_{21}, e_{22}, ...\}$ where

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_{n1}(t) = \frac{1}{\sqrt{\pi}} \cos\ nt, \quad e_{n2}(t) = \frac{1}{\sqrt{\pi}} \sin\ nt, \quad n \in \mathbb{N},$$

is an orthonormal basis. The fact that B is an orthogonal set can be easily checked. In order to prove that B is an orthonormal basis we have to see that $\overline{\text{Sp}\ B} = L^2_{\mathbb{K}}([-\pi, \pi])$ (Theorem 4.3.2), or equivalently, to show that for each $x \in L^2_{\mathbb{K}}([-\pi, \pi])$ and $\varepsilon > 0$, there exists a trigonometric polynomial, $p : [-\pi, \pi] \longrightarrow \mathbb{K}$ such that $\|p - x\|_2 \leq \varepsilon$.

First, take an arbitrary x in $L^2_{\mathbb{R}}([-\pi, \pi])$ and $\varepsilon > 0$. As $\mathcal{C}_{\mathbb{R}}([-\pi, \pi])$ is dense in $L^2_{\mathbb{R}}([-\pi, \pi]$, $\exists y \in \mathcal{C}_{\mathbb{R}}([-\pi, \pi])$ such that $\|x - y\|_2 \leq \varepsilon/3$. We show now that there exists $z \in \mathcal{C}_{\mathbb{R}}([-\pi, \pi])$ with $z(-\pi) = z(\pi)$ and $\|y - z\|_2 \leq \varepsilon/3$. If $y = 0$ if we take $z = 0$ everything is clear. Suppose further that $y \neq 0$; thus $\sup \{|y(t)| \mid t \in [-\pi, \pi]\} \neq 0$. Let z be the real mapping on $[-\pi, \pi]$ defined as follows

$$z(t) = \begin{cases} \alpha t + \beta & \text{if } t \in [-\pi, -\pi + \delta) \\ y(t) & \text{if } t \in [-\pi + \delta, \pi] \end{cases}$$

where $\delta \in (0, 2\pi)$, α, $\beta \in \mathbb{R}$ will be determined such that the next three conditions hold: $\|y - z\|_2 \leq \varepsilon/3$, the mapping z continuous, $z(-\pi) = z(\pi)$.

It follows that for $t \in [-\pi, \pi]$ and $\alpha,\ \beta,\ \delta$ like we have just said, we have $|z(t)| \leq \sup\ \{|y(t)| \mid t \in [-\pi, \pi]\} \neq 0$ (since the linear map $u : [-\pi, -\pi + \delta]$, $u(t) = \alpha t + \beta$ is monotone and $u(-\pi) = y(-\pi),\ u(-\pi + \delta) = y(-\pi + \delta)$).

Thus, let us find $\alpha,\ \beta,\ \delta$ satisfying the above mentioned conditions. We notice that

$$\|y - z\|_2^2 = \int_{-\pi}^{\pi} |y(t) - z(t)|^2 \,\mathrm{dt} = \int_{-\pi}^{-\pi+\delta} |y(t) - z(t)|^2 \,\mathrm{dt} \leq$$

$$\leq \int_{-\pi}^{-\pi+\delta} (|y(t)| + |z(t)|)^2 \,\mathrm{dt} \leq 4\,\delta \sup_{t \in [-\pi,\pi]} |y(t)|^2,$$

so, $\|y - z\|_2 \leq 2\sqrt{\delta} \sup\limits_{t \in [-\pi,\pi]} |y(t)|$. Taking

$$\delta < \min\left(\frac{\varepsilon^2}{\sup\limits_{t \in [-\pi,\pi]} |y(t)|^2}, 2\pi\right)$$

and $\alpha,\ \beta$ the solutions of the next system

$$\begin{cases} \alpha(-\pi) + \beta = y(-\pi) \\ \alpha(-\pi + \delta) + \beta = y(-\pi + \delta) \end{cases}$$

we have $\|y - z\|_2 \leq \varepsilon/3$.

From the Weierstrass second approximation theorem (Appendix B), there exists p a trigonometric polynomial such that

$$\sup_{t \in [-\pi,\pi]} |z(t) - p(t)| < \frac{\varepsilon}{3\sqrt{2\pi}}$$

Then,

$$\|y - z\|_2 \leq \left(\int_{-\pi}^{\pi} |z(t) - p(t)|^2 \,\mathrm{dt}\right)^{\frac{1}{2}} \leq \frac{\varepsilon}{3}$$

Finally, it results

$$\|x - p\|_2 = \|x - y + y - z + z - p\|_2 \leq$$

$$\leq \|x - y\|_2 + \|y - z\|_2 + \|z - p\|_2 < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

If x is complex-valued, it follows applying the above approach to Re f and Im f that $\overline{\mathrm{Sp}\ B} = L^2_{\mathbb{C}}([-\pi, \pi])$.

It results that the series

$$\alpha_o + \sum_{n\geq 1}(\alpha_n \cos\ nt + \beta_n \sin\ nt)$$

where

$$\alpha_o = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\,\mathrm{dt},\quad \alpha_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos\ nt\ \mathrm{dt},\quad \beta_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin\ nt\ \mathrm{dt},$$

converges in $L^2_{\mathbf{K}}([-\pi,\pi])$ to x,

$$\lim_n \int_{-\pi}^{\pi} |x(t) - \alpha_o - \sum_{k=1}^{n}(\alpha_k \cos\ kt + \beta_k \sin\ kt)|^2\,\mathrm{dt} = 0$$

In addition,

$$\int_{-\pi}^{\pi} |x(t)|^2\,\mathrm{dt} = 2\pi|\alpha_o|^2 + \pi\sum_{n=1}^{\infty}|\alpha_n|^2 + |\beta_n|^2$$

(Parceval's equality).

4. Since

$$\cos\ nt = \frac{e^{int} + e^{-int}}{2} \text{ and } \sin\ nt = \frac{e^{int} - e^{-int}}{2}$$

it follows that $\overline{\mathrm{Sp}\,\{e^{int} \mid n \in \mathbb{Z}\}} = L^2_{\mathbb{C}}([-\pi,\pi])$ (where $\mathbb{Z}$ is the set of all integers). Hence $\{1/\sqrt{2\pi}e^{int} \mid n \in \mathbb{Z}\}$ is an orthonormal basis in $L^2_{\mathbb{C}}([-\pi,\pi])$, the Fourier series is $\sum_{n\in\mathbb{Z}} c_n e^{int}$, where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}\ \mathrm{dt},\quad n \in \mathbb{Z}$$

converges in $L^2_{\mathbb{C}}([-\pi,\pi])$ to x and by Parceval's equality

$$\int_{-\pi}^{\pi} |x(t)|^2\,\mathrm{dt} = 2\pi\sum_{n=-\infty}^{\infty}|c_n|^2$$

Theorem 4.3.3 *Let $(u_n)_n$ be an arbitrary sequence of linearly independent vectors in the inner product space X. Then, there exists an orthonormal sequence $(v_n)_n$ such that for each $m \in \mathbb{N}$, the linear space spanned by $\{u_j\}_{1\leq j\leq m}$ coincides to the one spanned by $\{v_j\}_{1\leq j\leq m}$.*

Proof. Let us construct the sequence $(v_n)_n$ as follows: take $v_1 = w_1$, where $w_1 = (1/\|u_1\|)\, u_1$, and for arbitrary $n \in \mathbb{N}$, $n \geq 2$,

$$v_n = \frac{1}{\|w_n\|} w_n, \text{ where } w_n = u_n - \sum_{k=1}^{n-1} < u_n, v_k > v_k$$

Applying Theorem 4.1.3 and Theorem 4.1.4, we notice that for each $n \in \mathbb{N}$, the vector

$$\sum_{k=1}^{n-1} < u_n, v_k > v_k,$$

is the closest element to u_n in the linear space spanned by $\{v_k\}_{1 \leq k \leq n-1}$ and therefore w_n is orthogonal to this subspace of X. It is now clear that the sequence $(v_n)_n$ constructed above satisfies the requirements of the theorem.

Remark. The process used in the proof of the previous theorem for constructing an orthonormal sequence $(v_n)_n$ from an arbitrary sequence of independent vectors $(u_n)_n$ is known as the *Gram-Schmidt orthogonalization procedure.*

Definition. Two Hilbert spaces X_1 and X_2 are called *isomorphic* if there exists a linear mapping U from X_1 onto X_2 such that

$$< U(x), U(y) >=< x, y >, \quad \forall x, y \in X_1.$$

The following theorem allows us to characterizes separable Hilbert spaces (which frequently arise in practice) up to isomorphism.

Theorem 4.3.4 *A Hilbert space X is separable if and only if X has a countable orthonormal basis B. If there are n elements in B, then X is isomorphic to $\mathbb{K}^n$ and if there are countably many elements in B, then X is isomorphic to $l^2_{\mathbb{K}}$.*

Proof. Suppose first that X is separable and let $(y_n)_n$ a countable dense subset of X. By throwing out some of the y_n 's we can get a subset of independent vectors $(x_n)_n$ which spans the same subspace as $(y_n)_n$,

$$\text{Sp}\{x_n \mid n \in \mathbb{N}\} = \text{Sp}\{y_n \mid n \in \mathbb{N}\},$$

so $\text{Sp}\{x_n \mid n \in \mathbb{N}\}$ is dense in X. Further, by applying the Gram-Schmidt orthogonalization procedure to the sequence $(x_n)_n$ one obtains a countable orthonormal set B which spans the same subspace as $(x_n)_n$. Thus $\overline{\mathrm{Sp}\ B} = X$, and by Theorem 4.3.2, it results that B is an orthonormal basis of X.

Conversely, if $B = \{x_n \mid n \in \mathbb{N}\}$ is a countable orthonormal basis of X, then the set

$$A = \{\sum_{k=1}^{n} \rho_k x_k \mid n \in \mathbb{N},\ \rho_k = r_k + i s_k,\ r_k, s_k \in \mathbb{Q}\}$$

(of all finite linear combinations of the x_n with rational coefficients) is dense in X. Since this set is countable, X is separable.

If $B = \{x_k\}_{1 \le k \le n}$, each $x \in X$ is $\sum_{k=1}^{n} < x, x_k > x_k$, so taking $U : X \longrightarrow$ $\longrightarrow \mathbb{K}^n$,

$$U(x) = (< x, x_1 >, < x, x_2 >, \ldots, < x, x_n >)$$

everything is clear. Similarly, if $B = \{x_n \mid n \in \mathbb{N}\}$, using the Parseval identity it results that $(< x, x_n >)_n \in l^2_{\mathbb{K}}$, $\forall x \in X$, therefore $U : X \longrightarrow l^2_{\mathbb{K}}$,

$$U(x) = (< x, x_n >)_n$$

is well defined. We have to check that U preserves the inner product. Let $x, y \in X$, so

$$< U(x), U(y) > = \sum_{n=1}^{\infty} < x, x_k > \overline{< y, x_k >} =$$

$$= \lim_{n \to \infty} \sum_{k=1}^{n} < x, x_k > \overline{< y, x_k >} = \lim_{n \to \infty} < x, \sum_{k=1}^{n} < y, x_k > x_k > = < x, y >,$$

and the proof is finished.

Example. Using this theorem it is clear that all the spaces in the previous examples are separable Hilbert spaces. Here we have got an example of non separable inner-product space, the space of almost periodic functions. Recall that a complex valued function which is continuous on $\mathbb{R}$ is *almost periodic* if it is the uniform limit on $\mathbb{R}$ of a sequence of trigonometric polynomials of the form $\sum_{k=1}^{n} a_k e^{i\lambda_k t}$, λ_k real. Denote by $\mathcal{A}$ the set of all almost periodic functions. $\mathcal{A}$ endowed with the usual operations of addition and scalar multiplication becomes a vector space over $\mathbb{C}$. It can be shown that the mapping on $\mathcal{A} \times \mathcal{A}$,

$$< x, y > = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t)\overline{y(t)}\,\mathrm{dt}$$

is well defined and it is an inner product on $\mathcal{A}$. Since $\{e^{i\lambda t} \mid \lambda \in R\}$ is an uncountable orthonormal set, thus $\mathcal{A}$ is not separable.

We conclude this section by exhibiting how Hilbert spaces arose naturally from problems in classical analysis. If $x : [-\pi, \pi] \longrightarrow \mathbb{R}$ is an integrable function on $[-\pi, \pi]$ we can define the numbers

$$\alpha_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t)\,\mathrm{dt}, \quad \alpha_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \cos\ nt\ \mathrm{dt}, \quad n \in \mathbb{N},$$

$$\beta_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \sin\ nt\ \mathrm{dt}, \quad n \in \mathbb{N}.$$

The formal series $\alpha_0 + \sum_{n\geq 1}(\alpha_n \cos\ nt + \beta_n \sin\ nt)$ is called the Fourier series of the function x.

The classical problem is: for which x and in what sense does the Fourier series of x converge to x? This problem, originated with Fourier (1811) has had a rich history, being the point of depart for an entire branch of analysis, the abstract harmonic analysis. The most known answer to this question is the next theorem.

Theorem 4.3.5 *Suppose that $x : \mathbb{R} \longrightarrow \mathbb{R}$ is periodic of period 2π and is continuously differentiable. Then, the Fourier series of x converges uniformly to x.*

This theorem gives sufficient conditions for the Fourier series of a function to converge uniformly. But, finding the exact class of functions whose Fourier series converge uniformly or converge pointwise has proven to be a hard problem. We can, however, get a nice answer to this question if we change our notion of "convergence" and this is just where Hilbert spaces come in.

Theorem 4.3.6 *For any $x \in L^2_{\mathbb{R}}([-\pi, \pi]$ the series*

$$\alpha_0 + \sum_{n\geq 1}(\alpha_n \sin\ nt + \beta_n \cos\ nt)$$

converges to x in $(L^2_{\mathbb{R}}([-\pi, \pi], \|\cdot\|_2)$. In addition

$$\int_{-\pi}^{\pi} |x(t)|^2\,\mathrm{dt} = 2\pi\alpha_0^2 + \pi \sum_{n=1}^{\infty}(\alpha_n^2 + \beta_n^2)$$

Proof. We have proven (Example 3, 4.3) that $B = \{e_0, e_{11}, e_{12}, e_{21}, e_{22}, \dots\}$ where

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_{n1}(t) = \frac{1}{\sqrt{\pi}} \cos\ nt, \quad e_{n2}(t) = \frac{1}{\sqrt{\pi}} \sin\ nt, \quad n \in \mathbb{N},$$

is an orthonormal basis in $L^2_{\mathbb{R}}([-\pi, \pi]$. By Theorem 4.3.2 it results for each $x \in L^2_{\mathbb{R}}([-\pi, \pi]$ that

$$x = < x, e_o > + \sum_{n=1}^{\infty} (< x, e_{n1} > e_{n1} + < x, e_{n2} > e_{n2})$$

(the convergence of the series is in $L^2_{\mathbb{R}}([-\pi, \pi])$. By the definition of the inner product in $L^2_{\mathbb{R}}([-\pi, \pi]$,

$$< x, e_0 > = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x(t)\, \mathrm{dt}, \quad < x, e_{n1} > = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x(t) \cos\ nt\ \mathrm{dt}, \quad n \in \mathbb{N},$$

$$< x, e_{n2} > = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x(t) \sin\ nt\ \mathrm{dt}, \quad n \in \mathbb{N}.$$

If we denote

$$\alpha_0 = \frac{< x, e_0 >}{\sqrt{2\pi}}, \quad \alpha_n = \frac{< x, e_{n1} >}{\sqrt{\pi}}, \quad \beta_n = \frac{< x, e_{n2} >}{\sqrt{\pi}}$$

the theorem is proven. In addition, the Parseval's equality gives

$$\|x\|_2^2 = \frac{1}{2\pi} (\int_{-\pi}^{\pi} x(t)\, \mathrm{dt})^2 + \sum_{n=1}^{\infty} \frac{1}{\pi} [(\int_{-\pi}^{\pi} x(t) \cos\ nt\ \mathrm{dt})^2 + (\int_{-\pi}^{\pi} x(t) \sin\ nt\ \mathrm{dt})^2],$$

thus,

$$\int_{-\pi}^{\pi} |x(t)|^2\, \mathrm{dt} = 2\pi\alpha_0^2 + \pi \sum_{n=1}^{\infty} (\alpha_n^2 + \beta_n^2).$$

4.4 Exercises

1. Let X be an inner product space. Show that $| < x, y > | = \|x\| \cdot \|y\|$ if and only if x and y are linearly dependent.

2. (**Theorem Banach-Saks**) Let X be a Hilbert space and $(x_n)_n$ a bounded sequence in X. Prove that there exist $x \in X$ and a subsequence $(x_{n_k})_k$ such that

$$\lim_{k \to \infty} \| \frac{x_{n_1} + x_{n_2} + \dots x_{n_k}}{k} - x \| = 0$$

3. Let X be a Hilbert space and $A \subset X$. Prove that:

a) $A \subset A^{\perp\perp}$;

b) $A^{\perp\perp} = \overline{\text{Sp } A}$;

c) $A^{\perp\perp\perp} = A^{\perp}$;

d) If A is a closed linear subspace of X, then $A^{\perp\perp} = \overline{A}$.

4. Let $(X, \|\cdot\|)$ be a normed space such that

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \quad \forall x, y \in X$$

Show that there exists an inner product $< \cdot, \cdot >$ on X such that

$$\|x\| = < x, x >^{1/2}, \quad \forall x \in X$$

(the norm is generated by an inner product).

5. Prove that $l^p_{\mathbf{K}}$ $(p \in [1, \infty))$ is a Hilbert space if and only if $p = 2$.

6. Prove that $L^p_{\mathbf{K}}([0,1])$ $(p \in [1, \infty))$ is a Hilbert space if and only if $p = 2$.

7. Show that each Hilbert space is reflexive.

8. (The Gram determinant) Let $\{y_1, y_2, \dots, y_n\}$ be a basis for the subspace $Y \subset X$. Prove that for $x \in E$,

$$d(x, Y) = \left(\frac{g(y_1, y_2, \dots, y, y)}{g(y_1, y_2, \dots, y_n)} \right)^{\frac{1}{2}}$$

and

$$0 < g(y_1, y_2, \dots, y_n) \le \|y_1\|^2 \cdot \|y_2\|^2 \dots \|y_n\|^2,$$

where

$$g(y_1, y_2, \dots, y_n, y) = det \begin{pmatrix} < y_1, y_1 > & . & . & . & < y_n, y_1 > & < y, y_1 > \\ . & & & & . & . \\ . & & & & . & . \\ . & & & & . & . \\ < y_1, y_n > & . & . & . & < y_n, y_n > & < y, y_n > \\ < y_1, y > & . & . & . & < y_n, y > & < y, y > \end{pmatrix}$$

and $g(y_1, y_2, \dots, y_n) = det(< y_i, y_j >)$ (the Gram determinant corresponding to $\{y_1, y_2, \dots, y_n\}$).

In addition, the closest element to x in Y, y_x is given by

$$y_x = -\frac{1}{g(y_1, y_2, \dots, y_n,)} \times$$

$$\times \det \begin{pmatrix} <y_1, y_1> & <y_2, y_1> & \dots & <y_n, y_1> & <y, y_1> \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ <y_1, y_n> & <y_2, y_n> & \dots & <y_n, y_n> & <y, y_n> \\ y_1 & y_2 & \dots & y_n & 0 \end{pmatrix}$$

9. (`Hadamard's inequality for a determinant`) Let $A = (a_{ij})_{ij}$ be an $n \times n$ matrix of complex (real) numbers. Show that

$$|det\ A|^2 \leq \prod_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2.$$

10. Write the Fourier series of the function $x(t) = t^2$ and using the Parceval's equality in $L^2_{\mathbb{R}}([-\pi, \pi]$ find the sum of the numerical series $\sum_{n \geq 1} \frac{1}{n^4}$.

11. Show that the Legendre polynomial, $(\varphi_n)_n$,

$$\varphi_n(t) = \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{dt}^n} (t^2 - 1)^n$$

are obtained by applying the Gram-Schmidt ortogonalization procedure to the sequence $(u_n)_n$, $u_n(t) = t^n$, $n \in \mathbb{N} \cup \{0\}$ in the Hilbert space $L^2_{\mathbb{R}}([-1, 1])$.

12. Let $\varphi_1, \varphi_2, \ldots$ an orthonormal basis for $L^2_{\mathbb{K}}([a, b])$. Then $\varphi_{ij}(t, s) = \varphi_i(t)\overline{\varphi_j}(s)$, $i, j = 1, 2, \ldots$ is an orthonormal basis for $L^2_{\mathbb{K}}([a, b] \times [a, b])$.

Chapter 5

Linear operators on Hilbert spaces

5.1 The correspondence between sesquilinear forms and operators. The adjoint and its properties

Definition. A *sesquilinear form* on a vector space X over the field $\mathbb{K}$ is a map $B : X \times X \longrightarrow \mathbb{K}$ that is linear in the first variable and conjugate linear in the second, i.e.

i) $B(\alpha x + \beta y, z) = \alpha B(x,z) + \beta B(y,z)$, $\forall x,\ y,\ z \in X,\ \alpha,\ \beta \in \mathbb{K}$;

ii) $B(x, \alpha y + \beta z) = \overline{\alpha} B(x,y) + \overline{\beta} B(x,z)$, $\forall\ x,\ y,\ z \in X,\ \alpha,\ \beta \in \mathbb{K}$.

A sesquilinear form B is said to be *self-adjoint* (or *hermitian*) if $B(x,y) = \overline{B(y,x)}$, $\forall x, y \in X$, and *positive* if $B(x,x) \geq 0$, $\forall x \in X$.

Remark. An inner product on X is a positive, self-adjoint sesquilinear form such that $B(x,x) = 0$ implies $x = 0$ $(x \in X)$.

Remark. If B is a positive, self-adjoint sesquilinear form on X, then

$$0 \leq B(\lambda x + y, \lambda x + y) = |\lambda|^2 B(x,x) + 2Re\lambda\, B(x,y) + B(y,y), \forall x, y \in X, \lambda \in \mathbb{K}$$

thus, if $B(x,x) \neq 0$ (or $B(y,y) \neq 0$) setting here

$$\lambda = \frac{\overline{B(x,y)}}{B(x,x)},$$

one obtains the *generalized Cauchy-Schwarz inequality*,

$$|B(x,y)| \leq B(x,x)^{\frac{1}{2}} B(y,y)^{\frac{1}{2}}, \quad \forall x, y \in X$$

When both $B(x,x)$ and $B(y,y)$ are zero, the above inequality results setting $\lambda = -\overline{B(x,y)}$.

Further, suppose that $(X, < \cdot, \cdot >)$ is a Hilbert space.

Definition. A sesquilinear form on the Hilbert space X is said to be bounded if there exists a positive constant $M > 0$ such that

$$|B(x,y)| \leq M \; \|x\| \; \|y\|, \quad \forall x, y \in X$$

Notation. Note that the set of all bounded sesquilinear forms on X equipped with the usual functions addition and scalar multiplication is a vector space, denoted here by $\mathcal{SB}(X)$. The real map on $\mathcal{SB}(X)$,

$$B \longmapsto \|B\| = \inf \{M > 0 \mid |B(x,y)| \leq M \; \|x\| \; \|y\|, \forall x, y \in X\},$$

is a norm on $\mathcal{SB}(X)$. An easy computation leads to

$$\|B\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |B(x,y)| = \sup_{\|x\|=1, \|y\|=1} |B(x,y)| = \sup_{\|x\| \neq 0, \|y\| \neq 0} \frac{|B(x,y)|}{\|x\| \cdot \|y\|}$$

Next theorem is of historical interest (apart from being quite a useful result). It should be recalled that the spectral theory was developed by Hilbert (the beginning of our century) as a theory for quadratic and bilinear forms, but even the simplest computation with the expression of the forms as infinite matrices had the tendency to be very complicated. One of the reasons (but not only) to von Neumann's success was his consistent use of the operator concept to tackle problems in Hilbert spaces.

Proposition 5.1.1 *There is a bijective, isometric correspondence between operators in* $\mathcal{B}(X)$ *and forms in* $\mathcal{SB}(X)$*, given by* $T \longmapsto B_T$*, where*

$$B_T(x,y) = < x, Ty >, \quad \forall x, y \in X$$

Proof. If $T \in \mathcal{B}(X)$, then clearly B_T is a sesquilinear form on X. The boundedness of B results from

$$|B_T(x,y)| = |< x, Ty >| \leq \|T\| \cdot \|x\| \cdot \|y\| , \quad \forall x, y \in X,$$

thus $\|B_T\| \leq \|T\|$. On the other hand, inserting $x \rightsquigarrow Ty$ in the definition of B_T we obtain

$$\|Ty\|^2 = |B_T(Ty, y)| \leq \|B_T\| \cdot \|Ty\| \cdot \|y\|$$

which implies that $\|B_T\| \leq \|T\|$. It follows $\|B_T\| = \|T\|$.

We have only to see that the mapping $T \mapsto B_T$ is onto. For B in $\mathcal{SB}(X)$ and $y \in X$ let us consider the functional on X, defined by

$$f_y^B(x) = B(x,y), \quad x \in X,$$

which, evidently is in X^*. Applying to it the Riesz lemma, we obtain a unique element in X, Ty, such that

$$f_y^B(x) =< x, Ty >, \ \forall x \in X \text{ and } \|f_y^B\| = \|Ty\|$$

The map $y \mapsto Ty$ is linear, as it results by

$$< z, T(\alpha x + \beta y) >= f^B_{\alpha x + \beta y}(z) = B(x, \alpha x + \beta y)$$

$$= \overline{\alpha} B(z,x) + \overline{\beta} B(z,y) = \overline{\alpha} < z, Tx > + \overline{\beta} < z, Ty >$$

$$=< z, \alpha Tx + \beta Ty > \ \forall x, \ y, \ z \in X, \ \alpha, \ \beta \in \mathbb{K}$$

Moreover,

$$\|Ty\| = \|f_y^B\| = \sup_{\|x\|=1} |B(x,y)| \leq \|B\| \cdot \|y\|,$$

therefore T is bounded. As $B(x,y) =< x, Ty >$, $\forall x, y \in X$, the proposition is proven.

Corollary 5.1.1 *Let T be a linear operator on X. Then, T is bounded if and only if there exists a constant $M > 0$ such that*

$$|< x, Ty >| \leq M \ \|x\| \ \|y\|, \quad \forall x, y \in X$$

In addition,

$$\|T\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |< x, Ty >| =$$

$$= \sup_{\|x\|=1, \|y\|=1} |< x, Ty >| = \sup_{\|x\| \neq 0, \|y\| \neq 0} \frac{|< x, Ty >|}{\|x\| \cdot \|y\|} =$$

$$= \inf \{ M > 0 \mid |< x, Ty >| \leq M \ \|x\| \ \|y\|, \quad \forall x, y \in X \}$$

Proof. It follows by the equivalence, $T \in \mathcal{B}(X) \Leftrightarrow B_T \in \mathcal{SB}(X)$.

Remark. Since $| < y, Tx > | = |\overline{< Tx, y >}| = | < Tx, y > |$ the statement of the previous corollary can be rewritten: $T \in \mathcal{B}(X) \Leftrightarrow \exists M > 0$ such that $| < Tx, y > | \leq M \; \|x\| \; \|y\|, \forall x, y \in X$.

Theorem 5.1.1 *(The Hellinger–Toeplitz theorem) Each linear operator T on a Hilbert space $(X, < \cdot, \cdot >)$ satisfying $< x, Ty >=< Tx, y >$ for all x, y in X, is bounded.*

Proof. We will prove that the graph of T is closed. Suppose that $(x_n)_n \longrightarrow x$ and $(Tx_n)_n \longrightarrow y$. For any $z \in X$,

$$< z, y >=\lim_n < z, Tx_n >=\lim_n < Tz, x_n >=< Tz, x >=< z, Tx >$$

Thus, $y = Tx$ and the graph of T is closed. By the Closed graph theorem it follows that $T \in \mathcal{B}(X)$.

Theorem 5.1.2 *To each $T \in \mathcal{B}(X)$ there is a unique $T^* \in \mathcal{B}(X)$ such that*

$$< Tx, y >=< x, T^*y >, \quad \forall x, y \in X$$

In addition $\|T\| = \|T^*\|$.

Proof. For arbitrary T in $\mathcal{B}(X)$ we consider the sesquilinear form

$$(x, y) \stackrel{B}{\longmapsto} < Tx, y >$$

which clearly is in $\mathcal{SB}(X)$. By Proposition 5.1.1, it follows that there exists an operator, $T^* \in \mathcal{B}(X)$ such that $B_{T^*} = B$, i.e.

$$< Tx, y >=< x, T^*y >, \quad \forall x, y \in X$$

Moreover, $\|B_{T^*}\| = \|T^*\|$, and on the other hand

$$\|B_{T^*}\| = \inf \{M > 0 \mid |B_{T^*}(x, y)| \leq M \; \|x\| \; \|y\|, \forall x, y \in X\} =$$

$$= \inf \{M > 0 \mid |(Tx, y)| \leq M \; \|x\| \; \|y\|, \forall x, y \in X\} =$$
$$= \inf \{M > 0 \mid |(x, Ty)| \leq M \; \|x\| \; \|y\|, \forall x, y \in X\} = \|T\|$$

Definition. The operator $T^* \in \mathcal{B}(X)$ enjoying the property

$$< Tx, y > = < x, T^*y > \quad \forall x, y \in X,$$

is called the *adjoint* of the operator T. We say that the operator T is *self-adjoint* if $T = T^*$.

The set of all self-adjoint operators on the Hilbert space X will be denoted further by $\mathcal{A}(X)$.

Remark. *If A is a self-adjoint operator, it is easy to see that for each $x \in X$, $< Ax, x > \in \mathbb{R}$. Moreover $< Ax, x > \in [-\|A\|, \|A\|]$, $\forall x \in X$ with $\|x\| \leq 1$.*

Notations. For A a self-adjoint operator, denote

$$m_A = \inf_{\|x\| \leq 1} < Ax, x > \text{ and } M_A = \sup_{\|x\| \leq 1} < Ax, x >$$

Examples. 1. If $T \in \mathcal{B}(\mathbb{K}^n)$, $T \sim (a_{ij})_{1 \leq i,j \leq n}$, then, $T^* \sim (a^*_{ij})_{1 \leq i,j \leq n}$, where $a^*_{ij} = \overline{a_{ji}}$, thus the matrix of T^* is the adjoint of the matrix of T.

We remark that the terminology introduced in the above definition is the same terminology as the one developed in linear algebra. Thus, T is self-adjoint if its matrix is an hermitian matrix.

2. Let T be the operator on $l^2_{\mathbb{K}}$ defined by $T((\xi_n)_n) = (\lambda_n \xi_n)_n$, $(\xi_n)_n \in l^2_{\mathbb{K}}$, (where $(\lambda_n)_n \in l^\infty_{\mathbb{K}}$).

By Example 2 in the section 3.1, T is well defined and $T \in \mathcal{B}(l^2_{\mathbb{K}})$. The adjoint of T is defined by $T^*((\xi_n)_n) = (\overline{\lambda_n} \xi_n)_n$.

Indeed, we have for $e_n = (\delta_k^{(n)})_k$,

$$< Te_n, e_n > = < \lambda_n e_n, e_n > = < e_n, \overline{\lambda_n} e_n >,$$

thus $T^*(e_n) = \overline{\lambda_n} e_n$. Then,

$$T^*((\xi_n)_n) = T^*(\sum_{n=1}^{\infty} \xi_n e_n) = \sum_{n=1}^{\infty} \xi_n T^* e_n = \sum_{n=1}^{\infty} \overline{\lambda_n} \xi_n e_n = (\overline{\lambda_n} \xi_n)_n$$

The operator T is self-adjoint if and only if $(\lambda_n)_n \subset \mathbb{R}$.

3. Given the infinite matrix $(a_{ij})_{i,j=1}^{\infty}$, where $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|^2 < \infty$, define the operator A on $l^2_{\mathbb{K}}$, by

$$A(\xi_i)_i = (\sum_{j=1}^{\infty} a_{ij} \xi_j)_i, \quad (\xi_i)_i \in l^2_{\mathbb{K}}$$

Since

$$|\sum_{j=1}^{\infty} a_{ij}\xi_j| \leq \sum_{j=1}^{\infty} |a_{ij}| \cdot |\xi_j| \leq (\sum_{j=1}^{\infty} |a_{ij}|^2)^{\frac{1}{2}} (\sum_{j=1}^{\infty} |\xi_j|^2)^{\frac{1}{2}},$$

it results that

$$\|A(\xi_i)_i\|^2 = \sum_{i=1}^{\infty} |\sum_{j=1}^{\infty} a_{ij}\xi_j|^2 \leq \left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}|^2 \right) \|(\xi_i)_i\|^2,$$

so $A \in \mathcal{B}(l^2_{\mathbb{K}})$ and

$$\|A\|^2 \leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}|^2$$

Note that the condition $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}|^2 < \infty$ is not a necessary condition for A to be bounded (since, for example the identity matrix, does not satisfy this condition and gives rise to $A = I$).

Taking in $l^2_{\mathbb{K}}$ the standard basis, $\{e_n \mid n \in \mathbb{N}\}$,

$$a_{ij} = < Ae_j, e_i >$$

since

$$< A^* e_j, e_i > = < e_j, Ae_i > = \overline{a_{ji}},$$

and

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\overline{a_{ji}}|^2 = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}|^2,$$

it results that the adjoint of A, is the operator defined by the infinite matrix $(\overline{a_{ji}})_{i,j=1}^{\infty}$.

We have to notice that this operator is compact. To show that , let us consider for each $n \in \mathbb{N}$, the operator A_n on $l^2_{\mathbb{K}}$, defined by

$$A_n(\xi_i)_i = (\sum_{j=1}^{n} a_{1j}\xi_j, \sum_{j=1}^{n} a_{2j}\xi_j, \dots, \sum_{j=1}^{n} a_{nj}\xi_j, 0, 0 \dots, 0, ..),$$

where $\alpha = (\xi_i)_i \in l^2_{\mathbb{K}}$. Since A_n is of finite rank, it is compact. In addition,

$$\|A - A_n\|^2 \leq \sum_{i=n+1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}|^2 + \sum_{j=n+1}^{\infty}\sum_{i=1}^{\infty} |a_{ij}|^2 \longrightarrow 0$$

Thus, by Theorem 3.5.1, $A \in \mathcal{K}(l^2_{\mathbb{K}})$.

4. Let $k \in L^2_{\mathbb{K}}([a,b] \times [a,b])$ and define the operator T on $L^2_{\mathbb{K}}([a,b])$ by

$$(Tx)(t) = \int_a^b k(t,s)x(s)\,\mathrm{ds}, \quad x \in L^2_{\mathbb{K}}([a,b])$$

The integral exists since for each t, $k(t,s)x(s)$ is Lebesgue measurable on $[a,b]$ and by, Cauchy-Schwarz inequality,

$$\int_a^b |k(t,s)x(s)|\,\mathrm{ds} \leq \left(\int_a^b |k(t,s)|^2\,\mathrm{ds}\right)^{\frac{1}{2}} \left(\int_a^b |x(s)|^2\,\mathrm{ds}\right)^{\frac{1}{2}}$$

Thus,

$$\|Tx\|^2 \leq \int_a^b \left(\int_a^b |k(t,s)x(s)|\,\mathrm{ds}\right)^2 \mathrm{dt} \leq \left(\int_a^b \int_a^b |k(t,s)|^2\,\mathrm{ds}\;\;\mathrm{dt}\right) \|x\|^2$$

Hence $T \in \mathcal{B}(L^2_{\mathbb{K}}([a,b]))$ and

$$\|T\|^2 \leq \int_a^b \int_a^b |k(t,s)|^2\,\mathrm{ds}\;\;\mathrm{dt} = \|k\|^2$$

By Fubini's theorem,

$$< Tx, y >= \int_a^b \left(\int_a^b k(t,s)x(s)\,\mathrm{ds}\right) \overline{y(t)}\;\;\mathrm{dt} =$$

$$= \int_a^b x(s) \left(\overline{\int_a^b \overline{k(t,s)}y(t)\,\mathrm{dt}}\right) \mathrm{ds} =< x, y^* >,$$

where

$$y^*(s) = \int_a^b \overline{k(t,s)}y(t)\,\mathrm{dt}$$

It follows, that

$$(T^*x)(t) = \int_a^b \overline{k(t,s)}x(s)\,\mathrm{ds}, \quad x \in L^2_{\mathbb{K}}([a,b])$$

Notice that this operator is compact too.

Theorem 5.1.3 *(Properties of the adjoint) For $T, V \in \mathcal{B}(X)$ and $\alpha, \beta \in$ $\in \mathbb{K}$, we have*
1) $(\alpha T + \beta V)^* = \overline{\alpha} T^* + \overline{\beta} V^*$;

2) $(T\ V)^* = V^*T^*$;
3) $T^{**} = T$;
4) $\|T^*T\| = \|T\|^2$;
5) T *is invertible in* $\mathcal{B}(X)$ *if and only if* T^* *is invertible in* $\mathcal{B}(X)$; *in addition* $(T^*)^{-1} = (T^{-1})^*$;
6) $\text{Ker } T^* = (T(X))^{\perp}$ *and* $(\text{Ker } T)^{\perp} = \overline{T^*(X)}$;
7) $T \in \mathcal{K}(X)$ *if and only if* $T^* \in \mathcal{K}(X)$.

Proof. In order to prove the first three properties, we shall use the correspondence between sesquilinear forms and bounded linear operators on X.

$$B_{(\alpha T+\beta V)^*}(x,y) =< x, (\alpha T + \beta V)^* y >=$$

$$=< (\alpha T + \beta V)x, y >= \alpha < Tx, y > +\beta < Vx, y >=$$

$$=< x, \overline{\alpha} T^* y > + < x, \overline{\beta} V^* y >= B_{(\overline{\alpha} T^* + \overline{\beta} V^*)}(x,y), \quad \forall x, y \in X,$$

thus 1) holds. For 2), we have

$$B_{(TV)^*}(x,y) =< x, (TV)^* y >=< (TV)y, x >=$$

$$=< V(x), T^* y >=< x, V^* T^* y >= B_{V^* T^*}(y), \quad \forall x, y \in X$$

The third property results from

$$B_{T^{**}}(x,y) =< x, ((T)^*)^* y >=< T^* x, y >= \overline{< y, T^* x >} =$$

$$= \overline{< Ty, x >} =< x, Ty >= B_T(x,y), \quad \forall x, y \in X$$

Further, we prove 4). Clearly,

$$\|T^*T\| \leq \|T^*\| \cdot \|T\| = \|T\|^2$$

On the other hand, for each $y \in Y$,

$$\|Ty\|^2 =< Ty, Ty >=< y, T^*Ty >\leq \|y\|^2 \|T^*T\|,$$

therefore $\|T\| \leq \|T^*T\|$.

T is invertible in $\mathcal{B}(X)$ if there exists in $\mathcal{B}(X)$ an operator, (denoted by T^{-1}) such that $T\ T^{-1} = I$, $T^{-1}T = I$, (where I is the identity operator); hence, using 2), it follows $(T^{-1})^*T^* = I$, $T^*(T^{-1})^* = I$. These equalities are equivalent to the existence of the inverse of the mapping T^*, the inverse of it being $(T^{-1})^*$. As $T^{-1} \in \mathcal{B}(X)$, we have also that $(T^{-1})^* \in \mathcal{B}(X)$.

For 5), from the defining identity $< Tx, y > = < x, T^*y >$ we see that if $y \in \text{Ker } T^*$, then $y \in (T(X))^\perp$. Conversely, if $y \in (T(X))^\perp$, then $T^*y \in \in X^\perp = \{0\}$. The other relation can be proven similarly.

7) Suppose T is compact and let $(x_n)_n$ a sequence in X, $\|x_n\| = 1$. Since, TT^* is compact (Proposition 3.5.1), there exists a subsequence $(x_{n'})_{n'}$ of $(x_n)_n$ such that $(TT^*x_{n'})_{n'}$ converges. Then, for arbitrary $n', m' \in \mathbb{N}$,

$$\|T^*x_{n'} - T^*x_{m'}\| = < TT^*(x_{n'} - x_{m'}), x_{n'} - x_{m'} > \leq$$

$$\leq \|TT^*(x_{n'} - x_{m'})\| \cdot \|x_{n'} - x_{m'}\| \leq 2\|TT^*x_{n'} - TT^*x_{m'}\|,$$

which means that $(T^*x_{n'})_{n'}$ is Cauchy, thus, by the completeness of X, convergent.

5.2 The numerical radius

Definition. For each $T \in \mathcal{B}(X)$ we define the *numerical radius* of T by

$$|||T||| = \sup_{\|x\|=1} | < Tx, x > |$$

Remark. It is obvious to show that the numerical radius can be calculated with

$$|||T||| = \sup_{\|x\|\leq 1} | < Tx, x > | = \sup_{x \neq 0} \frac{| < Tx, x > |}{\|x\|^2}$$

Theorem 5.2.1 *Let X be a complex Hilbert space. Then,*
1) The real mapping $T \longmapsto |||T|||$ on $\mathcal{B}(X)$ is a norm;
2) For each $T \in \mathcal{B}(X)$,

$$\frac{1}{2}\|T\| \leq |||T||| \leq \|T\|$$

3) For each $T \in \mathcal{B}(X)$, $\|T^2\| \leq \|T\|^2$.

Proof. First we will prove the second assertion. By the Cauchy-Schwarz inequality we have

$$| < Tx, x > | \leq \|Tx\| \cdot \|x\|,$$

thus $| < Tx, x > | \leq \|T\| \cdot \|x\|^2$, which implies that $|||T||| \leq \|T\|$.

In order to obtain that $\|T\|$ is dominated by $2 \cdot |||T|||$, we consider at the beginning the identity

$$< T(x+y), x+y > - < T(x-y), x-y >=$$

$$= 2 < Tx, y > +2 < Ty, x >, \quad x, y \in X$$

If x and y are unit vectors, using the above identity and the Parallelogram law, one obtain

$$2| < Tx, y > + < Ty, x > | \leq |||T||| \; (\|x+y\|^2 + \|x-y\|^2) =$$

$$= 2 \; |||T||| \; (\|x\|^2 + \|y\|^2) = 4 \; |||T|||$$

Therefore, $\mathrm{Re}(< Tx, y > + < Ty, x >) \leq 2|||T|||$. If one inserts here $y \rightsquigarrow$ $\rightsquigarrow (1/\|Tx\|)Tx$, we have

$$(*) \quad \|Tx\| + \mathrm{Re} \frac{1}{\|Tx\|} < T^2x, x > \leq 2 \; |||T|||, \quad \forall x \in X, \; \|x\| = 1$$

Further, if $| < T^2x, x > | = e^{i\theta} < T^2x, x >$, $\exists \gamma \in \mathbb{C}$ such that $\gamma^2 = \theta$. As the previous inequality holds for each $T \in \mathcal{B}(X)$, it is valid also for $T \rightsquigarrow \gamma T$:

$$\|Tx\| + \mathrm{Re} \frac{1}{\|Tx\|} < \gamma^2 T^2 x, x > \leq 2 \; |||T|||, \quad \forall x \in X, \; \|x\| = 1$$

Thus we have obtained that

$$\|Tx\| + \frac{1}{\|Tx\|} | < T^2x, x > | \leq 2 \; |||T|||, \quad \forall x \in X, \; \|x\| = 1,$$

which implies $\|Tx\| \leq 2 \; |||T|||$, $\forall x \in X$, $\|x\| = 1$, so $\|T\| \leq 2 \; |||T|||$.

Now, 1) is clear because evidently, $T \longmapsto |||T|||$ is a seminorm and if $|||T||| = 0$, by 2), $\|T\| = 0$, therefore $T = 0$.

In order to prove 3) we observe that

$$\|Tx\| + \frac{1}{\|Tx\|} | < T^2x, x > | \leq 2 \; |||T||| \iff$$

$$\iff 0 \leq 2\,|||T||| \cdot \|Tx\| - \|Tx\|^2 - |<T^2x, x>|$$

So, if we continue to calculate in the right-hand side, we have

$$0 \leq -(|||T||| - \|Tx\|)^2 + |||T|||^2 - |<T^2x, x>| \leq$$
$$\leq |||T|||^2 - |<T^2x, x>|,\ \forall x \in X,\ \|x\| = 1$$

which implies

$$|<T^2x, x>| \leq |||T|||^2,\ \forall x \in X,\ \|x\| = 1$$

and ends the proof.

Corollary 5.2.1 *Let X be a Hilbert space over the field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) and A a self-adjoint operator on X. Then, $\|A\| = |||A|||$.*
Moreover $\|A\| = \max\{|m_A|, |M_A|\}$.

Proof. The equality $(*)$ in the proof of the previous theorem holds for each $T \in \mathcal{B}(X)$. If we set here $T \rightsquigarrow A$, since $A = A^*$ we have,

$$\|Ax\| + \frac{1}{\|Ax\|}|<A^2x, x>| \leq 2\,|||A|||, \quad \forall x \in X,\ \|x\| = 1,$$

whence $\|Ax\| + \|Ax\| \leq 2\,|||A|||$. Consequently, $\|A\| \leq |||A|||$, which combined with 2) of the Theorem 5.2.1 shows that $\|A\| = |||A|||$.

Corollary 5.2.2 *Let X be a Hilbert space over the field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) and A a self-adjoint operator on X such that $<Ax, x>= 0$, $\forall x \in X$. Then $A = 0$.*

Proof. If $<Ax, x>= 0$, $\forall x \in X$ it follows that $|||A||| = 0$, so $\|A\| = 0$.

5.3 Some special classes of operators on Hilbert spaces

5.3.1 Normal operators, unitary operators

Definition. $T \in \mathcal{B}(X)$ is said to be *normal* if $TT^* = T^*T$.

Remark. Obviously, each self-adjoint operator is normal.

Proposition 5.3.1 *Let X be a complex Hilbert space and $T \in \mathcal{B}(X)$ a normal operator. Then, $|||T||| = \|T\|$.*

Proof. Since T is normal, for each $n \in \mathbb{N}$, $(TT^*)^n = T^n(T^*)^n$. Then,

$$\|T\|^{2^n} = (\|TT^*\|^{\frac{1}{2}})^{2^n} = (\|TT^*\|^{2^n})^{\frac{1}{2}} = \|(TT^*)^{2^n}\|^{\frac{1}{2}} =$$
$$= \|T^{2^n}(T^*)^{2^n}\|^{\frac{1}{2}} = \|T^{2^n}(T^{2^n})^*\|^{\frac{1}{2}} = \|T^{2^n}\|$$

By the Theorem 5.2.1, it follows that $\|T^{2^n}\| \leq 2|||T^{2^n}|||$, and applying the same result, 3) we have finally that

$$\|T\|^{2^n} \leq 2|||T|||^{2^n} \iff \|T\| \leq 2^{\frac{1}{2^n}}|||T|||, \quad \forall n \in \mathbb{N},$$

whence $\|T\| \leq |||T|||$. As we have also $|||T||| \leq \|T\|$, the proposition is proven.

Proposition 5.3.2 *Let T be in $\mathcal{B}(X)$. Then, the following statements are equivalent:*
1) T is normal;
2) $\|Tx\| = \|T^*x\|$, $\forall x \in X$.

Proof. 1)$\Rightarrow$ 2) results by

$$\|Tx\|^2 = < Tx, Tx > = < x, T^*Tx > =$$
$$= < x, TT^*x > = < T^*x, T^*x > = \|T^*x\|^2$$

In order to prove the converse, we recall that it is enough to show that $< (TT^* - T^*T)x, x > = 0$, $\forall x \in X$, because $TT^* - T^*T$ is self-adjoint (Corollary 5.2.2). By 2), we have

$$< (TT^* - T^*T)x, x > = < TT^*x, x > - < T^*Tx, x > = \|Tx\|^2 - \|T^*x\|^2 = 0,$$

Definition. A *unitary* operator is a map of X onto itself such that

$$< Ux, Uy > = < x, y >, \quad \forall x, y \in X$$

Remark. *Each unitary operator U on X is linear and continuous, so U is* in $\mathcal{B}(X)$. *First we check the linearity of U. Let t be arbitrary in X; since U is onto, there exists $z \in X$ such that $U(z) = t$. Then, everything is clear from the next computation.*

$$< U(\alpha x + \beta y), t > = < U(\alpha x + \beta y), Uz > = < \alpha x + \beta y, z > =$$
$$= \alpha < x, z > + \beta < y, z > = \alpha < Ux, Uz > + \beta < Uy, Uz > =$$
$$= < \alpha Ux, t > + < \beta Uy, t > = < \alpha Ux + \beta Uy, t >, \quad \forall x, y \in X, \alpha, \beta \in \mathbb{K}$$

Continuity of U obviously results from the definition if in the equality $< Ux, Uy > = < x, y >$, $\forall x, y \in X$, we set $x = y$. Moreover, $\|U\| = 1$.

Proposition 5.3.3 *Let U be in $\mathcal{B}(X)$. The following are equivalent:*
(1) U is a unitary operator;
(2) U is onto and $\|Ux\| = \|x\|$, $\forall x \in X$;
(3) $UU^* = U^*U = I$;
(4) U^* is a unitary operator.

Proof. Clearly (1) $\Rightarrow$ (2). It follows from the Polarization identities that an isometric map preserves the inner product, because

$$4 < Ux, Uy >= \sum_{k=0}^{3} i^k \|U(x + i^k y)\|^2 = \sum_{k=0}^{3} i^k \|x + i^k y\|^2 = 4 < x, y >$$

(and similarly in the real case), so (2) $\Rightarrow$ (1).

From (1) U is onto and injective, so $\exists U^{-1}$ such that $UU^{-1} = U^{-1}U = I$. On the other hand

$$< Ux, Uy >=< x, y > \iff < x, U^*Uy >=< x, y >, \quad \forall x, y \in X,$$

therefore, $U^*U = I$. As U is onto, it follows that $U^{-1} = U^*$, so (1) $\Rightarrow$ (3). The converse is clear, since (3) means in particular that U is onto and, in addition

$$< Ux, Uy >=< x, U^*Uy >=< x, y >$$

Now, if U^* is unitary, then, taking into account that (1) $\Leftrightarrow$ (3),

$$U^*(U^*)^* = (U^*)^*U^* = I,$$

i.e. $U^*U = UU^* = I \Leftrightarrow U$ unitary.

Remark. *Note that the product (composition) of unitary operators is again unitary, so that the set $\mathcal{U}(X)$ of unitary operators on X is a group (a subgroup of the general linear group $\mathcal{GL}(X)$ of invertible operators in $\mathcal{B}(X)$).*

5.3.2 Positive operators, the square root of a positive operator

Definition. An operator $A \in \mathcal{A}(X)$ is said to be *positive* (written $A \geq 0$) if $< Ax, x >\geq 0$, $\forall x \in X$.

The set of all positive operators is denoted by $\mathcal{A}_+(X)$.

Remarks. 1. If $T \in B(X)$, then the operators TT^* and T^*T are clearly positive.

2. For A, B in $\mathcal{A}_+(X)$, their sum, $A+B$ is also in $\mathcal{A}_+(X)$.

3. If $A \in \mathcal{A}_+(X)$ and $t \geq 0$, the operator $tA \in \mathcal{A}_+(X)$.

4. If both A and $-A$ are in $\mathcal{A}_+(X)$, then $A = 0$. Indeed, If A and $-A$ are in $\mathcal{A}_+(X)$, then $< Ax, x >= 0, \forall x \in X$, so $x = 0$.

5. We will order $\mathcal{A}(X)$ by

$$A_1 \leq A_2 \Longleftrightarrow A_2 - A_1 \in \mathcal{A}_+(X).$$

$(\mathcal{A}(X), \leq)$ is a partial ordered set. We note that, if $A_1, A_2 \in \mathcal{A}_+(X)$, and $A_1 \leq A_2$ then $\|A_1\| \leq \|A_2\|$. This follows from

$$A_1 \leq A_2 \Longleftrightarrow < A_1x, x > \leq < A_2x, x >, \forall x \in X \Longleftrightarrow |||A_1||| \leq |||A_2|||.$$

6. For each $A \in \mathcal{A}_+(X)$, the sesquilinear form corresponding to it (Proposition 5.1.1) is positive self-adjoint, so we have the generalized Cauchy-Schwarz inequality,

$$| < Ax, y > | \leq < Ax, x >^{\frac{1}{2}} < Ay, y >^{\frac{1}{2}}, \quad \forall x, y \in X.$$

(Let us make the convention to call this inequality the Cauchy-Schwarz inequality corresponding to the positive operator A).

We add that, if we set in the above inequality $y = Ax$, it results

$$\|Ax\| \leq \|A\|^{\frac{1}{2}} < Ax, x >^{\frac{1}{2}}, \quad \forall x \in X.$$

Theorem 5.3.1 *Let $(A_n)_n$ be a sequence in $\mathcal{A}(X)$ with the following properties:*
1) $A_n \leq A_{n+1}$, $\forall n$;
2) $\exists B \in \mathcal{A}(X)$ satisfying $A_n \leq B$, $\forall n$.
Then, there exists $A \in \mathcal{A}(X)$ such that for each $x \in X$, $A_n(x) \longrightarrow A(x)$, as $n \to \infty$, (the sequence $(A_n)_n$ converges pointwise to A).
Moreover, $A = \sup\limits_n A_n$.

Proof. We may evidently assume that $(A_n)_n \subset \mathcal{A}_+(X)$ (otherwise, we replace in the next arguments A_n by $A_n - A_1$). For arbitrary fixed $x \in X$, consider the real numerical sequence $(< A_n x, x >)_n$ and notice that it is upper bounded (by $< Bx, x >$) and increasing. It results that this sequence is convergent, fact which enables us to show that $(A_n x)_n$ is a Cauchy sequence in X. Consequently, for $m, n \in \mathbf{N}$, $m \geq n$, using the Cauchy-Schwarz inequality corresponding to the positive operator $A_m - A_n$, we derive

$$\|A_m x - A_n x\|^4 =< (A_m - A_n)x, (A_m - A_n)x >^2 \leq$$

$$\leq < (A_m - A_n)x, x > \cdot < (A_m - A_n)^2 x, (A_m - A_n)x > \leq$$

$$\leq < (A_m - A_n)x, x > \cdot \|A_m - A_n\|^4 \cdot \|(A_m - A_n)^2 x\|^2$$

It follows,

$$\|A_m x - A_n x\|^2 \leq 2\|B\| \ < (A_m - A_n)x, x >,$$

therefore $(A_n x)_n$ is a Cauchy sequence in the Hilbert space X. Then, there exists $A(x) = \lim_{n\to\infty} A_n(X)$; clearly, as the pointwise limit of a sequence of self-adjoint operators is itself in $\mathcal{A}(X)$ and as

$$< A_n x, x > \leq < A_{n+m} x, x >, \quad \forall n, m$$

it results that A is an upper bound of $(A_n)_n$ in $\mathcal{A}(X)$. If C is an other upper bound of this sequence, we have $< A_n x, x > \leq < Cx, x >$, whence, as $n \to \infty$, $A \leq C$; this means that $A = \sup_n A_n$.

Theorem 5.3.2 *Let A, B be in $\mathcal{A}_+(X)$ such that $AB = BA$. Then $AB \in$ $\in \mathcal{A}_+(X)$.*

Proof. Take $A \in \mathcal{A}_+(X)$, $A \neq 0$ (if $A = 0$ everything is clear). Let us define inductively the sequence of operators $(A_n)_n$ by

$$A_1 = \frac{1}{\|A\|} A, \quad A_{n+1} = A_n - A_n^2, \quad \forall n \geq 2$$

The sequence of continuous operators $(A_n)_n$ has the following properties:

1) $0 \leq A_n \leq I$, $\forall n$;
2) $A_n B = B A_n, \forall n$;
3) $A_1 x = \sum_{n=1}^{\infty} A_n^2 x$, $\forall x \in X$.

Let us suppose that we have already proven that the sequence $(A_n)_n$ enjoys the above properties. Then, $\forall x \in X$

$$< ABx, x >= \|A\| < \sum_{n=1}^{\infty} A_n^2(Bx), x >= \|A\| < \sum_{n=1}^{\infty} A_n B A_n x, x >=$$

$$= \|A\| \sum_{n=1}^{\infty} < BA_n x, A_n x > \geq 0$$

and the proposition is proven.

Now, we show that $(A_n)_n$ has the properties 1)-3). We prove 1) by induction (with respect to n). If $n = 1$, by the previous remarks 3), $0 \leq A_1$. Since $\|A\| = |||A|||$, $< Ax, x > \leq \|A\| < x, x >$, $\forall x \in X$, thus $A_1 \leq I$. Further, we suppose that $0 \leq A_n \leq I$, and we have to obtain $0 \leq A_{n+1} \leq I$. This follows from

$$I - A_{n+1} = I - A_n + A_n^2 = (I - A_n) + A_n^2$$

and

$$A_{n+1} = A_n - A_n^2 = A_n(I - A_n)^2 + (I - A_n)A_n^2$$

taking into account that the sum of two positive operators is also a positive operator and that the operators $I - A_n$, A_n^2, $A_n(I - A_n)^2$, $(I - A_n)A_n^2$ are positive. An easy computation shows that indeed, $A_n(I - A_n)^2$, $(I - A_n)A_n^2$ are in $\mathcal{A}_+(X)$:

$$< A_n(I - A_n)^2 x, x >=< (I - A_n)A_n(I - A_n)x, x >=$$

$$=< A_n(I - A_n)x, (I - A_n)x > \geq 0$$

and

$$< (I - A_n)A_n^2 x, x >=< A_n(I - A_n)A_n x, x >=< (I - A_n)A_n x, A_n x > \geq 0 \;)$$

The property 2) can be easy checked also by induction, using the assumption $AB = BA$. For 3), we have, inductively,

$$A_1 = A_2 + A_1^2 = A_3 + A_2^2 + A_1^2 = \sum_{k=1}^{n} A_k^2 + A_{n+1}, \quad \forall n \in \mathbb{N},$$

Then,

$$\sum_{k=1}^{n} A_k^2 = A_1 - A_{n+1} \leq A_1, \quad \forall n \in \mathbb{N}.$$

Thus the sequence $(\sum_{k=1}^{n} A_k^2)_n$ is increasing and bounded, so by the Theorem 5.3.1, for each $x \in X$, the series $\sum_{k\geq 1} A_k^2 x$ converges in X.

On the other hand, from the above inequality, we also have for each $x \in X$

$$\sum_{k=1}^{n} \|A_k x\|^2 = \sum_{k=1}^{n} < A_k x, A_k x >= \sum_{k=1}^{n} < A_k^2 x, x >=$$

$$=< \sum_{k=1}^{n} A_k^2 x, x > \leq < A_1 x, x >,$$

so, it results that the numerical series $\sum_{k=1}^{\infty} \|A_k x\|^2$ is convergent. Therefore, for each $x \in X$,

$$\lim_{n\to\infty} A_n x = 0$$

But,

$$A_1 x = \sum_{k=1}^{n} A_k^2 x + A_{n+1} x, \quad \forall n \in \mathbb{N}.$$

As $n \to \infty$, it follows that $A_1 x = \sum_{k=1}^{\infty} A_k^2 x$, and 3) is proven.

Theorem 5.3.3 *(Square root theorem) To each positive operator $A \in$ $\in \mathcal{A}_+(X)$ there is a unique positive operator $B \in \mathcal{A}_+(X)$, (called the square root of A) satisfying $B^2 = A$. Moreover, B commutes with every bounded operator commuting with A.*

Proof. If $A = 0$, we can evidently take $B = 0$. We assume further that $A \neq 0$. Then, there is no loss of generality assuming that $\|A\| = 1$ (since, otherwise, we may consider the operator $(1/\|A\|)A$ and if its square root is B', the one of A is obviously $\sqrt{\|A\|} \cdot B'$).

We define inductively a sequence of operators by

$$B_1 = 0, \quad B_{n+1} = B_n + \frac{1}{2}(A - B_n^2), \quad n \geq 1$$

and we proceed by induction with respect to n in order to establish that $(B_n)_n$ is a bounded increasing sequence of self-adjoint operators. Indeed, clearly $B_1 = 0 \in \mathcal{A}(X)$, and, assuming that $B_n \in \mathcal{A}(X)$, B_{n+1} is also in $\mathcal{A}(X)$ because it is sum of two self-adjoint operators. Now, we see that $B_n \leq I$, $\forall n$. This results from

$$I - B_{n+1} = I - B_n - \frac{1}{2}A + \frac{1}{2}B_n^2 = \frac{1}{2}(I - B_n)^2 + \frac{1}{2}(I - A)$$

since $I - A \geq 0$. Suppose that $B_{n-1} \leq B_n$ $(n \geq 2)$ and compare B_n with B_{n+1}.

$$B_{n+1} - B_n = B_n + \frac{1}{2}(A - B_n^2) - B_{n-1} - \frac{1}{2}(A - B_{n-1}^2) =$$

$$= B_n - B_{n-1} - \frac{1}{2}(B_n^2 - B_{n-1}^2) =$$

$$= (B_n - B_{n-1})\frac{1}{2}[(I - B_n) + (I - B_{n-1})] \geq 0,$$

(we have used Theorem 5.3.2 since the operators

$$(B_n - B_{n-1}) \text{ and } [(I - B_n) + (I - B_{n-1})]$$

evidently commute, so satisfy its hypothesis).

It follows, by Theorem 5.3.1, that there exists $B \in \mathcal{A}_+(X)$ the pointwise limit of the sequence $(B_n)_n$. For arbitrary fixed $x \in X$,we have

$$\|B_n^2 x - B^2 x\| \leq \|B_n(B_n x) - B(Bx)\| =$$

$$= \|B_n(B_n x) - B_n(Bx) + B_n(Bx) - B(Bx)\| \leq$$

$$\leq \|B_n\| \cdot \|B_n x - Bx\| + \|B_n(Bx) - B(Bx)\| \longrightarrow 0$$

which shows that $B_n^2 x \longrightarrow B^2 x$, as $n \to \infty$. Then, passing to the limit with respect to n in

$$B_{n+1} = B_n + \frac{1}{2}(A - B_n^2),$$

it follows that $B^2 = A$.

We prove further that the operator B constructed above is the unique with the property $B^2 = A$. If C is an other operator in $\mathcal{A}_+(X)$ satisfying $C^2 = A$, we claim that it commutes with B. This can be proven by induction, as it results from the next computation (because $AC = CA$) :

$$CB_{n+1} = CB_n + \frac{1}{2}(CA - CB_n^2) = B_n C + \frac{1}{2}(AC - B_n^2 C) = B_{n+1}C$$

Now, in order to show that $Cx = Bx$, $\forall x \in X$, because of

$$\|Cx - Bx\|^2 = < (C - B)x, (C - B)x > = < x, (C - B)^2 x >$$

it is enough to establish that $(C-B)(C-B)x=0$. Taking into account that C and B are in $\mathcal{A}_+(X)$ and commute each other, and denoting $(C-B)x$ by y, we have

$$0 \leq < Cy, y > + < By, y > = < (B+C)y, y > =$$

$$= < (B+C)(B-C)x, (B-C)x > = < (B^2-C^2)x, x > = 0$$

Therefore $< Cy, y > = 0$ and $< By, y > = 0$. But, by the Cauchy-Schwarz inequality corresponding to the positive operator C, and similarly to B, it results

$$\|Cy\|^2 = < Cy, Cy > \leq < Cy, y >^{\frac{1}{2}} < C^2y, Cy >^{\frac{1}{2}} = 0,$$

that implies $Cy = By = 0$, and, finally $(C-B)y = 0$ (the desired equality).

In order to end the proof, we have to notice that if T is a bounded operator which commutes with A, then, clearly, by the defining relation of the sequence $(B_n)_n$, it follows that $TB_n = B_nT$, so for $n \to \infty$, $TB = BT$.

5.3.3 Projections, partial isometries. The polar decomposition of a bounded operator

Definition. A bounded operator P on X is said to be a *projection* if $P^2 = P$. An *orthogonal projection* is a projection satisfying $P^* = P$.

Convention. Since orthogonal projections arise more frequently than non-orthogonal ones, *we normally use the word projection to mean orthogonal projection.* So, further, $P \in \mathcal{B}(X)$ is a projection if $P^2 = P$ and $P^* = P$.

Remark. Each projection is a positive operator. This results from

$$< Px, x > = < P^2x, x > = < Px, Px > = \|Px\|^2, \ \forall x \in X$$

Notation. Let Y be a closed subspace of X. By the projection theorem, each $x \in X$ can be uniquely written $x = x_1 + x_2$, with $x_1 \in Y$ and $x_2 \in Y^{\perp}$. Let P_Y be the operator on X defined by $P_Y(x) = x_1$.

The next theorem sets up a one to one correspondence between projections and closed subspaces.

Theorem 5.3.4 *For any closed subspace Y of X, the operator P_Y is a projection. Conversely, for each projection P there is a closed subspace Y of X such that $P_Y = P$.*

Proof. We see first that $P_Y \in \mathcal{B}(X)$. Let x, y be in X, so $x = x_1 + x_2$, $y = y_1 + y_2$ with x_1, $y_1 \in Y$, x_2, $y_2 \in Y^\perp$ and α, $\beta \in \mathbb{K}$. Since Y and $Y^\perp$ are subspaces, $\alpha x_1 + \beta y_1 \in Y$ and $\alpha x_2 + \beta y_2 \in Y^\perp$. Then,

$$P_Y(\alpha x + \beta y) = \alpha x_1 + \beta y_1 = \alpha P_Y(x) + \beta P_Y(y)$$

which means that P_Y is linear. The continuity results from

$$\|P_Y(x)\|^2 = \|x_1\|^2 \leq \|x_1\|^2 + \|x_2\|^2 = \|x\|^2$$

In addition we have

$$P_Y^2(x) = P_Y(P_Y(x)) = P_Y(x_1) = x_1 = P_Y(x)$$

and

$$< P_Y(x), y >=< x_1, y_1 + y_2 >=$$

$$=< x_1 + x_2, y_1 > + < x_1, y_2 >=< x, P_Y(y) >$$

Therefore, P_Y is a projection.

Conversely, for $P \in \mathcal{A}(X)$, satisfying $P^2 = P$, let us denote by Y the subspace $P(X)$. Notice that Y is closed. Indeed, let $y \in \overline{Y}$ be, thus there exists $(x_n)_n \subset X$ such that $Px_n \longrightarrow y$. Then,

$$y = \lim_n Px_n = \lim_n P(Px_n) = P(\lim_n Px_n) = P(y),$$

so, clearly, $y \in Y$. We conclude that $Y = \overline{Y}$.

Further, we show that $P = P_Y$. For an arbitrary x in X, $x = x_1 + x_2$, $x_1 \in P(X)$, $(x_1 = Pz,\ z \in X)$, $x_2 \in P(X)^\perp = \text{Ker } P$ we have

$$P(x) = P(Pz + x_2) = Pz + Px_2 = Pz = x_1 = P_Y x$$

Remark. *For each closed subspace Y, the operator P_Y defined above is called the projection onto the subspace Y, so any projection is the projection onto its own image.*

Definition. An operator $V \in \mathcal{B}(X)$ is said to be a *partial isometry* if

$$\|Vx\| = \|x\|, \quad \forall x \in (\text{Ker } V)^\perp$$

$(\text{Ker } V)^\perp$ is called the *initial subspace* of V and $V(X)$, the *final subspace* of V.

A partial isometry whose initial subspace coincides to X is an *isometry* of X.

Remark. If V is a partial isometry, the operator $V : (\text{Ker } V)^{\perp} \longrightarrow V(X)$ is an unitary operator.

Examples. 1. Clearly, each projection is a partial isometry. More generally, if $V \in \mathcal{B}(X)$ such that $V^*V = P$ for some projection P, then we see from the relation $\|Vx\|^2 =< P(x), x >= \|x\|^2$, that V is isometric onto $P(X)$ and 0 on $P(X)^{\perp}$, so V is a partial isometry with the initial subspace $P(X)^{\perp}$ and final subspace $P(X)$.

2. If X is a separable Hilbert space with orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$, the unilateral shift operator,

$$S(\sum_{n=1}^{\infty} \xi_n e_n) = \sum_{n=1}^{\infty} \xi_n e_{n+1}$$

is an isometry of X onto the subspace $(\text{Sp } e_1)^{\perp}$, and its kernel is Sp e_1. Consequently, S is a partial isometry. We see from this example that, contrary to the case of finite-dimensional Hilbert spaces, in infinite-dimensional Hilbert spaces one may have isometries that are not unitaries (because they are not surjective).

Next theorem will make clear the relationship between projections and partial isometries. Moreover, it shows that the set of partial isometries on a Hilbert space has a particular algebraic structure under the multiplication (composition) of operators.

Theorem 5.3.5 *For $V \in \mathcal{B}(X)$, the following are equivalent:*
(1) V is a partial isometry;
(2) V^*V is a projection;
(3) VV^* is a projection;
(4) $VV^*V = V$;
(5) $V^*VV^* = V^*$;
(6) V^* is a partial isometry.

Proof. If V is a partial isometry, then we show that V^*V is the projection onto the subspace $(\text{Ker } V)^{\perp}$. First, if x is arbitrary in X, $x = x_1 + x_2$, where $x_1 \in \text{Ker } V$ and $x_2 \in (\text{Ker } V)^{\perp}$, so $(V^*V)x = (V^*V)(x_1 + x_2) =$ $= (V^*V)x_2$. We have only to see that the restriction of V^*V to $(\text{Ker } V)^{\perp}$

is the identity operator on $(\text{Ker } V)^{\perp}$, or equivalently, since $V^*V - I$ is self-adjoint, $< (V^*V - I)x, x >= 0$, $\forall x \in (\text{Ker } V)^{\perp}$ which is clear since V is a partial isometry. Conversely, if V^*V is a projection, then V^*V is the projection onto the subspace $V^*V(X)$. By

$$V^*V(X) = \overline{V^*V(X)} = (\text{Ker } V^*V)^{\perp} = (\text{Ker } V)^{\perp}$$

it results that for each $x \in (\text{Ker } V)^{\perp}$,

$$\|Vx\|^2 =< Vx, Vx >=< x, V^*Vx >=< x, x >,$$

so, indeed V is a partial isometry. We have seen that (1) $\Leftrightarrow$ (2).

Suppose now that V^*V is a projection, and prove that VV^* is also a projection. This means, since VV^* is clearly self-adjoint to verify that $(VV^*)^2 = VV^*$, i.e.

$$V(V^*V - I)V^*x = 0, \quad \forall x \in X,$$

or, equivalently, $(V^*V - I)V^*x \in \text{Ker } V, \forall x \in X$. As $\text{Ker } V = \text{Ker } VV^*$, to check the above inequality means to check

$$V^*V(V^*V - I)V^*x = 0, \quad \forall x \in X,$$

or, equivalently

$$V^*VV^*VV^*x = VV^*V^*x, \quad \forall x \in X$$

This is true, since V^*V is a projection; thus, (2) $\Rightarrow$ (3). Similarly, (3) $\Rightarrow$ (2).

Now, we prove that (2) $\Rightarrow$ (4). For arbitrary x in X, $x = x_1 + x_2$, where $x_1 \in \text{Ker } V = \text{Ker } V^*V$ and $x_2 \in (\text{Ker } V)^{\perp} = V^*V(X)$,

$$VV^*Vx = VV^*Vx_2 = Vx_2 = Vx$$

Conversely, we have

$$(V^*V)^2 = (V^*V)(V^*V) = V^*(VV^*V) = VV^*,$$

therefore VV^* is a projection.

Similarly, (3) $\Leftrightarrow$ (5). Intertwining the role of V with V^*, it is clear that the statement (6) is equivalent to the others.

Next, we will prove the polar decomposition theorem which is an analogues for operators on Hilbert spaces to the decomposition of each complex number z as $z = |z|\, e^{i \arg z}$.

Theorem 5.3.6 *(Polar decomposition theorem) To each operator T in $\mathcal{B}(X)$, there exist the operators A, V in $\mathcal{B}(X)$ with the following properties:*
1) $T = VA$;
2) $A \in \mathcal{A}_+(X)$
3) V *is a partial isometry;*
4) Ker $A =$ Ker V*;*
5) A, V *are unique satisfying* 1) – 4).
Moreover, $V(X) = \overline{T(X)}$.

Proof. As $T^*T \in \mathcal{A}_+(X)$, one may consider its square root, $A \in \mathcal{A}_+(X)$. Clearly, $\|Ax\| = \|Tx\|$, $\forall x \in X$ and Ker $A =$ Ker T. On the subspace of X, $A(X)$, we define the map $V_o : A(X) \longrightarrow T(X)$, by

$$V_o(Ax) = Tx, \quad x \in X,$$

(which is well defined because Ker $A =$ Ker T). Since V_o is linear and bounded we may extend it by continuity to $\overline{A(X)}$, so let us denote by V_1 this extension, $V_1 : \overline{A(X)} \longrightarrow \overline{T(X)}$. It results that for $y \in \overline{A(X)}$, $y = \lim\limits_n Ax_n$,

$$V_1(y) = \lim_n V_o(Ax_n) = \lim_n Tx_n$$

Next, we define on X a linear operator, V such that

$$Vy = \begin{cases} V_1(y) & \text{if} \quad y \in \overline{A(X)} \\ 0 & \text{if} \quad y \in \overline{A(X)}^{\perp} = \text{Ker } A \end{cases}$$

which is bounded since, for each $x \in X$, $x = y + z$, $y \in \overline{A(X)}$, $z \in \overline{A(X)} =$ $=$ Ker A,

$$\|Vx\| = \|V(y+z)\| = \|V_1(y)\| \leq \|V_1\| \cdot \|y\|$$

In addition, $V(X) = \overline{T(X)}$.

We will show that Ker $A =$ Ker V. By the definition of V, Ker $A \subset$ Ker V. For the converse inclusion, let $x \in$ Ker V, $x = y + z$, $y \in \overline{A(X)}$, $z \in \overline{A(X)} =$ $=$ Ker A, so $Vy = 0$. As $Ax = Ay$, we have only to see that, to each $y \in \overline{A(X)}$ with $Vy = 0$ it results $Ay = 0$. Thus, consider $y = \lim\limits_n Ax_n$. We have

$$0 = Vy = V_1(y) = \lim_n V_0(Ax_n) = \lim_n Tx_n,$$

therefore $(Tx_n)_n$ converges to 0. On the other hand, $\|Ax_n\| = \|Tx_n\|$, which combined with $Tx_n \longrightarrow 0$ and $Ax_n \longrightarrow y$, proves that $y = 0$, so $Ay = 0$.

Finally, for $y \in \overline{A(X)} = (\text{Ker } A)^{\perp}$, $y = \lim_n Ax_n$,

$$\|Vy\| = \|V_1 y\| = \lim_n \|V_o(Ax_n)\| = \lim_n \|Tx_n\| = \lim_n \|Ax_n\| = \|y\|$$

and

$$VAy = V_1 Ay = V_o Ay = Ty$$

We conclude that the operators A, V have the properties $1) - 4)$. If B, W are other two operators with the same properties, W^*W is the projection onto $(\text{Ker } W)^{\perp} = \overline{B(X)}$ and also $T = WB$. Then,

$$A^2 = T^*T = (WB)^*(WB) = B^*(W^*W)B = B^2$$

By the uniqueness of the square root, it follows that $A = B$. Using further that $WA = VA$, we have that the partial isometries W, V coincide on their initial subspace, $\overline{A(X)}$, so, everywhere.

Remark. The preceding result (due to von Newmann), enables us to regard the partial isometry V as a generalized "sign" of T and A as the "absolute value", of T.

The next corollary is immediate:

Corollary 5.3.1 *If T is invertible in $\mathcal{B}(X)$, the partial isometry in its polar decomposition is unitary.*

5.4 Matrix representations of bounded operators

Next, it is shown how to associate a matrix with a given bounded operator $T \in \mathcal{B}(X)$ on a separable Hilbert space X. Let $\{e_n \mid n \in \mathbb{N}\}$be an orthonormal basis; then, for $x \in X$, $x = \sum_{n=1}^{\infty} < x, e_n > e_n$. From the linearity and continuity of T, we have

$$Tx = \sum_{n=1}^{\infty} < x, e_n > Te_n$$

On the other hand, the development in Fourier series of each Te_n $(n \in \mathbb{N})$ is

$$Te_n = \sum_{k=1}^{\infty} < Te_n, e_k > e_k$$

By the above two relations it follows that

$$Tx = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} < x, e_n >< Te_n, e_k > e_k = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} < x, e_n >< Te_n, e_k > e_k$$

Thus, we have

$$\begin{pmatrix} < Te_1, e_1 > & < Te_2, e_1 > & \cdots \\ < Te_1, e_2 > & < Te_2, e_2 > & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} < x, e_1 > \\ < x, e_2 > \\ \vdots \end{pmatrix} = \begin{pmatrix} < Tx, e_1 > \\ < Tx, e_2 > \\ \vdots \end{pmatrix}$$

This matrix equation leads to the following definition:

Definition. Let T be a bounded linear operator on a separable Hilbert space X and $\{e_n \mid n \in \mathbb{N}\}$ be an orthonormal basis of X. The *matrix corresponding to T and the orthonormal basis* $\{e_n \mid n \in \mathbb{N}\}$ is defined by

$$a_{ij} =< Te_j, e_i >, \quad i, j \in \mathbb{N}.$$

Example. Let $X = L^2_{\mathbb{K}}([-\pi, \pi])$ and $T \in \mathcal{B}(L^2_{\mathbb{K}}([-\pi, \pi]))$, defined by

$$(Tx)(t) = a(t)x(t),$$

with $a : [-\pi, \pi] \longrightarrow \mathbb{K}$, a bounded complex-valued Lebesgue measurable function. The "doubly infinite" matrix $(a_{jk})_{j,k=-\infty}^{\infty}$ corresponding to T and the orthonormal basis $e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}$, $n \in \mathbb{Z}$, is obtained as follows. Let

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} a(t)e^{-int}\,\mathrm{dt}\,.$$

Then,

$$a_{jk} =< Ae_k, e_j >= \frac{1}{2\pi}\int_{-\pi}^{\pi} a(t)e^{i(k-j)t}\,\mathrm{dt} = a_{j-k}$$

Thus the matrix is

$$\begin{pmatrix} \ddots & \ddots & \ddots & & & & & \\ \cdots & a_1 & a_0 & a_{-1} & \cdots & & & \\ & \cdots & a_1 & a_0 & a_{-1} & \cdots & & \\ & & \cdots & a_1 & a_0 & a_{-1} & \cdots & \\ & & & & \ddots & \ddots & \ddots & \end{pmatrix}$$

which is called the Toeplitz matrix.

Definition. A bounded linear operator T on a separable Hilbert space X is said to be *diagonalizable* if there is an orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$ for X such that the matrix corresponding to T and the orthonormal basis $\{e_n \mid n \in \mathbb{N}\}$ is diagonal.

Remark. The previous definition may be generalized for operators defined on arbitrary Hilbert spaces, as follows: a bounded linear operator T on X is said to be diagonalizable if there is an orthonormal basis $\{x_j \mid j \in J\}$ for X and a set $\{\lambda_j \mid j \in J\}$ in $\mathbb{K}$ (necessarily bounded) such that

$$Tx = \sum_j \lambda_j < x, x_j > x_j, \quad \forall x \in X$$

5.5 Exercises

1. For arbitrary $n \in \mathbb{N}$, let T_n defined on $l^2_{\mathbb{K}}$ by $T_n((\xi_k)_k) = (\xi_{n+1}, \xi_{n+2,\dots})$.

a) Show that $T_n \in \mathcal{B}(l^2_{\mathbb{K}})$, $\forall\, n \in \mathbb{N}$, and find T_n^*;

b) Show that $\lim\limits_{n\to\infty} T_n(x) = 0$, $\forall x \in l^2_{\mathbb{K}}$, but there exists $x_o \in l^2_{\mathbb{K}}$ such that the sequence $(T_n^*(x_o))_n$ does not converge in $l^2_{\mathbb{K}}$.

(From b), it results that the mapping $T \longmapsto T^*$ on $\mathcal{B}(X)$ endowed with the pointwise convergence topology is not continue.)

2. In $l^2_{\mathbb{K}}$ the standard orthonormal basis is denoted by $\{e_n \mid n \in \mathbb{N}\}$. Let $(\lambda_n)_n \subset \mathbb{K}$ be an arbitrary numerical sequence.

a) Show that there exists uniquely $T \in \mathcal{B}(l^2_{\mathbb{K}})$ such that

$$T(e_n) = \lambda_n e_n, \quad \forall n \in \mathbb{N}$$

if and only if $(\lambda_n)_n \in l^\infty_{\mathbb{K}}$.

In addition $T((\xi_k)_k) = (\lambda_k \xi_k)_k$, $\forall (\xi_k)_k \in l^2_{\mathbb{K}}$ and $\|T\| = \sup\limits_n |\lambda_n|$.

b) If $(\lambda_n)_n \in l^\infty_{\mathbb{K}}$, find T^*.

c) Give necessary and sufficient conditions for the sequence $(\lambda_n)_n \in l^\infty_{\mathbb{K}}$ such that the operator T be self-adjoint, respectively normal, respectively unitary, respectively projection.

d) Show that the operator T is compact if and only if the sequence $(\lambda_n)_n$ converges to 0.

e) If $T = T^*$, find m_T and M_T.

f) Show that T is compact if and only if $\lambda_n \longrightarrow 0$, as $n \to \infty$.

3. a) Show that the unilateral shift operator on $l^2_{\mathbb{K}}$, $S((\xi_k)_k) = (0, \xi_1, \xi_2, ...)$ is in $\mathcal{B}(l^2_{\mathbb{K}})$, find $\|S\|$, and then S^*.

b) Find the matrix corresponding to S and to the standard orthonormal basis of $l^2_{\mathbb{K}}$;

c) Show that S is a partial isometry and find its initial, and respectively final subspace.

4. Let $A : L^2_{\mathbb{K}}([0,1]) \to L^2_{\mathbb{K}}([0,1])$ be the operator defined by

$$Ax(t) = t \cdot x(t), \quad t \in [0,1]$$

a) Show that A is a positive operator;

b) Find m_A , M_A and the norm of A.

c) Find the square root of A.

d) Is A compact?

5. Let X be a complex Hilbert space and $T \in \mathcal{B}(X)$. Show that T is self-adjoint if and only if $< Tx, x > \in \mathbb{R}$, $\forall x \in X$.

6. Let P be a projection on the Hilbert space X. Then

a) $0 \leq < Px, x > \leq \|x\|^2$, $\forall x \in X$;

b) If $P \neq 0$, then $\|P\| = 1$;

c) Ker $P = \{x \in X \mid < Px, x >= 0\}$;

d) $P(X) = \{x \in X \mid Px = x\}$.

7. a) Find for an arbitrary projection P, m_P and M_P.

b) Find for an arbitrary projection the square root and its polar decomposition.

8. Write the polar decomposition for an arbitrary partial isometry.

9. Let Y, Z be two closed subspaces of the Hilbert space X and P, Q the projections onto Y, respectively Z. Then PQ is a projection if and only if $PQ = QP$. Under these conditions, PQ is the projection onto $Y \cap Z$.

10. Let Y, Z be two closed subspaces of the Hilbert space X and P, Q the projections onto Y, respectively Z. Then, the following are equivalent:

(1) $Y \subset Z$;

(2) $QP = P = PQ$;

(3) $\|Px\| \leq \|P\|$;

(4) $P \leq Q$.

11. Let Y be a closed subspace of the Hilbert space X, P the projections onto Y, and $T \in \mathcal{B}(X)$. Then,

a) $T(Y) \subset Y \iff PTP = TP$;

b) $T(Y) \subset Y$ and $T^*(Y) \subset Y \iff PT = TP$.

12. Let A be in $\mathcal{A}(X)$ and define $R = A + iI$. Show that:

a) R is normal;

b) $\|Rx\|^2 = \|Ax\|^2 + \|x\|^2$ and R is injective;

c) R is invertible in $\mathcal{B}(X)$;

d) The Cayley transform of A, $U = R^*R^{-1}$ is unitary.

13. Let X be a separable Hilbert space. Show that $T \in \mathcal{B}(X)$ is diagonalizable if and only if there is a unitary operator U from $l^2_{\mathbf{K}}$ onto X such that the corresponding matrix to UTU^{-1} and to standard orthonormal basis of $l^2_{\mathbf{K}}$ is diagonal.

13. For a bounded linear operator T on a Hilbert space X, the following are equivalent:

1) $\exists \alpha > 0$ such that $\|Tx\| \geq \alpha\|x\|$, $\forall x \in X$;

2) $\exists S \in \mathcal{B}(X)$ such that $ST = I$.

14. a) Let A be a self-adjoint operator on a Hilbert space X such that $\exists \alpha > 0$ with $\|Ax\| \geq \alpha\|x\|$, $\forall x \in X$. Show that A is invertible in $\mathcal{B}(X)$.

b) Denote by X the linear space of all complex polynomials equipped with the inner product $< p, q >= \int_0^1 p(t)\overline{q(t)}\,dt$ and consider the map $T : X \longrightarrow X$, $Tp(t) = (1 + t)p(t)$. Show that $< Tp, q >=< p, Tq >$, $\forall p, q \in X$, $\|Tp\| \geq \|p\|$, $\forall p \in X$, $T(X)$ is dense in X, but $T(X) \neq X$. Does this result contradict a)?

15. Let X be a Hilbert space and $T \in \mathcal{B}(X)$. The following are equivalent:

(1) T is compact;

(2) There exists a sequence of operators of finite rank $(T_n)_n \in \mathcal{B}(X)$ such that $(T_n)_n$ converges in the operator norm to T.

16. Let X be a separable Hilbert space. Show that $T \in \mathcal{B}(X)$ is diagonalizable if and only if there is a unitary operator U from $l^2_{\mathbf{K}}$ onto X such that the corresponding matrix to UTU^{-1} and the standard orthonormal basis of $l^2_{\mathbf{K}}$ is diagonal.

Chapter 6

Elementary spectral theory

6.1 Invertible elements in Banach algebras

For each Banach (Hilbert) space X, the Banach space of all bounded linear operators on X, $(\mathcal{B}(X), \|\cdot\|)$ is naturally endowed with a multiplication, the composition of operators. Therefore, $(\mathcal{B}(X), \cdot)$ is an algebra with a submultiplicative norm ($\|ST\| \leq \|S\| \cdot \|T\|$, $\forall S, T \in \mathcal{B}(X)$), such that $(\mathcal{B}(X), \|\cdot\|)$ is a Banach space. It will be our pattern for an abstract topological-algebraic structure, called Banach algebra. We will focus here on the properties connected to the invertibility of elements, useful in particular in $\mathcal{B}(X)$.

Definition. A *normed algebra* over the field $\mathbb{K}$ is a normed space $(\mathcal{A}, \|\cdot\|)$ over the field $\mathbb{K}$, equipped with a multiplication, $(S, T) \longmapsto ST$, such that
i) $(\mathcal{A}, \cdot)$ has an algebraic structure of algebra, i.e.
  i1) $\lambda(ST) = (\lambda S)T) = S(\lambda T)$, $\forall S, T \in \mathcal{A}, \forall \lambda \in \mathbb{K}$;
  i2) $(S + T)V = SV + TV$, $\forall S, T, V \in \mathcal{A}$;
  i3) $S(T + V) = ST + SV$, $\forall S, T, V \in \mathcal{A}$
ii) The norm of $\mathcal{A}$ is submultiplicative, i.e. $\|ST\| \leq \|S\| \cdot \|T\|$, $\forall S, T \in \mathcal{A}$.

Normed algebra $\mathcal{A}$ is said to be *unital* if there is $I \in \mathcal{A}$, (necessarily unique) called the *unity* of $\mathcal{A}$, such that $IT = TI = T$, $\forall T \in \mathcal{A}$ and *commutative* if $ST = TS$, $\forall S, T \in \mathcal{A}$.

A normed algebra is called a *Banach algebra* if the normed space $(\mathcal{A}, \|\cdot\|)$ is Banach.

Remark. If $\mathcal{A}$ is a unital normed algebra, $\mathcal{A} \neq \{0\}$, then we may suppose, without lose the generality that $\|I\| = 1$ (see Exercise 1).

Definition. In a unital normed algebra $\mathcal{A}$, an element U is called *invertible* if there are the elements $S, T \in \mathcal{A}$ such that $US = TU = I$.

Notation. In the above definition, the elements S,T are necessarily equal, as it follows from

$$S = IS = (TU)S) = T(US) = TI = T.$$

Then, we denote by $U^{-1} = S = T$ and we call it the *inverse* of the element U. Note that the inverse is necessarily unique.

The set of all invertible elements in $\mathcal{A}$, denoted here by $\mathcal{G}(\mathcal{A})$, endowed with the inherited multiplication of $\mathcal{A}$ is evidently, a group (called when $\mathcal{A} = \mathcal{B}(X)$ the general linear group).

Further, in this section $\mathcal{A}$ will always be a unital Banach algebra over the field $\mathbb{K}$, with unity I, $\|I\| = 1$.

Theorem 6.1.1 *Suppose $T \in \mathcal{A}$ and $\|T\| < 1$. Then, $I - T \in \mathcal{G}(\mathcal{A})$, and*

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n \quad (T^0 = I)$$

Moreover,

$$\|(I - T)^{-1}\| \leq \frac{1}{1 - \|T\|}$$

Proof. Let $S_n = \sum_{k=0}^{n} T^k$. Since the norm is submultiplicative we have

$$\|S_n\| \leq \sum_{k=0}^{n} \|T^k\| \leq \sum_{k=0}^{n} \|T\|^k,$$

and, as $\|T\| < 1$, it follows that the series $\sum_{n\geq 0} T^n$ converges. In addition,

$$(I - T)S_n = S_n(I - T) = I - T^{n+1} \longrightarrow I$$

as $n \to \infty$, since $\|T^{n+1}\| \leq \|T\|^{n+1}$, and $\|T\| < 1$. It results

$$(I - T)(\sum_{n=0}^{\infty} T^n) = (\sum_{n=0}^{\infty} T^n)(I - T) = I$$

which proves that $I - T$ is invertible and $(I-T)^{-1} = \sum_{n=0}^{\infty} T^n$. Finally,

$$\|(I-T)^{-1}\| \leq \sum_{n=0}^{\infty} \|T\|^n = \frac{1}{1-\|T\|}$$

Corollary 6.1.1 *1) Suppose $T \in \mathcal{A}$ and $\|I-T\| < 1$. Then, $T \in \mathcal{G}(\mathcal{A})$, and*

$$T^{-1} = \sum_{n=0}^{\infty} (I-T)^n$$

2) Let $T_o \in \mathcal{A}$ be invertible. Suppose $T \in \mathcal{A}$ and $\|T - T_o\| < 1/\|T_o^{-1}\|$. Then, $T \in \mathcal{G}(\mathcal{A})$,

$$T^{-1} = \sum_{n=0}^{\infty} [T_o^{-1}(T_o - T)]^n T_o^{-1}$$

and

$$\|T^{-1} - T_o^{-1}\| \leq \frac{\|T_o^{-1}\|^2 \|T - T_o\|}{1 - \|T_o^{-1}\| \cdot \|T - T_o\|}$$

Proof. 1) Everything is clear by the previous theorem, replacing $T \rightsquigarrow \rightsquigarrow I - T$.

2) Since

$$T = T_o - (T_o - T) = T_o[I - T_o^{-1}(T_o - T)]$$

and

$$\|T_o^{-1}(T_o - T)\| \leq \|T_o^{-1}\| \cdot \|(T_o - T)\| < 1$$

with the Theorem 6.1.1 it follows that T is invertible and

$$T^{-1} = [I - T_o^{-1}(T_o - T)]^{-1} T_o^{-1} = \sum_{n=0}^{\infty} [T_o^{-1}(T_o - T)]^n T_o^{-1},$$

or equivalently,

$$T^{-1} - T_o^{-1} = \sum_{n=1}^{\infty} [T_o^{-1}(T_o - T)]^n T_o^{-1}$$

Using the submultiplicity of the norm, we have

$$\|T^{-1} - T_o^{-1}\| \leq \|T_o^{-1}\| \sum_{n=1}^{\infty} \|T_o^{-1}(T_o - T)\|^n \leq$$

$$\leq \|T_o^{-1}\| \sum_{n=1}^{\infty} (\|T_o^{-1}\| \cdot \|T_o - T\|)^n = \frac{\|T_o^{-1}\|^2 \|T - T_o\|}{1 - \|T_o^{-1}\| \cdot \|T - T_o\|}$$

Corollary 6.1.2 *The multiplicative group $\mathcal{G}(\mathcal{A})$ is an open subgroup of $\mathcal{A}$ and the map $T \longmapsto T^{-1}$ on $\mathcal{G}(\mathcal{A})$ is a homeomorphism of $\mathcal{G}(\mathcal{A})$.*

Proof. By the Corollary 6.1.1, 2), if $T_o \in \mathcal{G}(\mathcal{A})$, the open ball

$$B(T_o, \frac{1}{\|T_o^{-1}\|}) \subset \mathcal{G}(\mathcal{A}),$$

thus $\mathcal{G}(\mathcal{A})$ is open. In addition, as

$$\|T^{-1} - T_o^{-1}\| \leq \frac{\|T_o^{-1}\|^2 \|T - T_o\|}{1 - \|T_o^{-1}\| \cdot \|T - T_o\|}, \quad \forall T \in B(T_o, \frac{1}{\|T_o^{-1}\|})$$

it follows that the map $T \longmapsto T^{-1}$ on $\mathcal{G}(\mathcal{A})$ and its inverse are continuous.

Remark. The above results hold in the particular case of $\mathcal{A} = \mathcal{B}(X)$. Taking into account that the invertibility of an element in the algebra $\mathcal{B}(X)$ is equivalent to its bijectivity and that the convergence in $\mathcal{B}(X)$ implies the pointwise convergence, we have to point out:

1) For each $T \in \mathcal{B}(X)$ with $\|T\| < 1$, the operator $I - T$ is bijective and

$$(I - T)^{-1} y = \sum_{n=0}^{\infty} T^n y, \quad \forall y \in X$$

(Theorem 6.1.1).

2) Let $T_o \in \mathcal{B}(X)$ be invertible. Suppose $T \in \mathcal{B}(X)$ and $\|T - T_o\| < < 1/\|T_o^{-1}\|$. Then T is invertible, and

$$T^{-1} y = \sum_{n=0}^{\infty} [T_o^{-1}(T_o - T)]^n T_o^{-1} y, \quad \forall y \in X$$

(Corollary 6.1.1).

Examples. 1. Infinite systems of linear equations.

We will relate the previous results to certain infinite systems of linear equations

$$\sum_{j=1}^{\infty} a_{ij} x_j = \gamma_i, \ i = 1, 2, \ldots.$$

where $\gamma = (\gamma_k)_k \in l^2_{\mathbf{K}}$ is given and $x = (x_k)_k$ is the desired solution in $l^2_{\mathbf{K}}$. Clearly, any equation $Tx = \gamma$, where $T \in \mathcal{B}(X)$ and X is a separable Hilbert space, can be written as an infinite system of linear equations. Indeed, let $\{e_n \mid n \in \mathbb{N}\}$ be an orthonormal basis for X. Then, with the matrix representation of the operator T, (section 5.4), the equation $Tx = \gamma$ becomes the system

$$\sum_{j=1}^{\infty} < Te_j, e_i >< x, e_j >=< \gamma, e_i >, \ i = 1, 2, ...$$

A natural approach to the problem of finding a solution to an infinite system of linear equations is to approximate the solution by solutions to finite sections of the system. The next proposition, which is rather easy to prove (being an immediate consequence of Theorem 6.1.1), is nevertheless very useful.

Proposition 6.1.1 *Suppose $(a_{ij})_{i,j=1}^{\infty}$ is an infinite matrix satisfying*

$$\sum_{i,j=1}^{\infty} |a_{ij}|^2 < 1.$$

Then, the system of equations

$$x_i - \sum_{j=1}^{\infty} a_{ij} x_j = \gamma_i, \ i = 1, 2, ...$$

has a unique solution $\eta = (\eta_k)_k \in l^2_{\mathbf{K}}$ for every $(\gamma_k)_k \in l^2_{\mathbf{K}}$.
The truncated system of equations

$$x_i - \sum_{j=1}^{n} a_{ij} x_j = \gamma_i, \ i = 1, 2, ..., n$$

has a unique solution $(\eta_1^{(n)}, ..., \eta_n^{(n)})$ and $(z_n)_n = ((\eta_1^{(n)}, ..., \eta_n^{(n)}, 0, 0...))_n$ converges in $l^2_{\mathbf{K}}$ to η as $n \to \infty$.

Proof. Define A on $l^2_{\mathbf{K}}$ by $A(\alpha_i)_i = (\sum_{j=1}^{\infty} a_{ij}\alpha_j)_i$. We know that $A \in \mathcal{B}(l^2_{\mathbf{K}})$ and $\|A\|^2 \le \sum_{i,j=1}^{\infty} |a_{ij}|^2 < 1$. Hence we have that $I - A$ is invertible (Theorem 6.1.1), which ensures for each $\gamma \in l^2_{\mathbf{K}}$ the existence and the uniqueness of the solution in $l^2_{\mathbf{K}}$ of the discussed system.

To approximate this solution, we consider, for each $n \in \mathbf{N}$, the operator A_n on $l^2_{\mathbf{K}}$, defined by

$$A_n(\xi_i)_i = (\sum_{j=1}^{n} a_{1j}\xi_j, \sum_{j=1}^{n} a_{2j}\xi_j, ..., \sum_{j=1}^{n} a_{nj}\xi_j, 0, 0...., 0, ..),$$

where $(\xi_i)_i \in l^2_{\mathbb{K}}$. Then, $\|A\|^2 \leq \sum_{i,j=1}^{n} |a_{ij}|^2 < 1$ and

$$\|(I-A)-(I-A_n)\|^2 = \|A-A_n\|^2 \leq \sum_{i=n+1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|^2 + \sum_{j=n+1}^{\infty}\sum_{i=1}^{\infty}|a_{ij}|^2 \to 0$$

as $n \to \infty$. Hence, $I - A_n$ is invertible, and, by Corollary 6.1.1 we have also

$$\|(I-A)^{-1} - (I-A_n)^{-1}\| \to 0$$

Taking $(\eta_1^{(n)}, \eta_2^{(n)}, ...) = (I - A_n)^{-1}\gamma$, it follows from the definition of A_n that $(\eta_1^{(n)}, ...\eta_n^{(n)})$ is the unique solution of the finite linear system $x_i - \sum_{j=1}^{n} a_{ij}x_j =$ $= \gamma_i$, $i = 1, 2, ..., n$ and $\eta_j^{(n)} = \gamma_j$, for $j > n$. Hence,

$$(I-A)^{-1}\gamma = \lim_{n\to\infty}(I-A_n)^{-1}\gamma = \lim_{n\to\infty}(\eta_1^{(n)}, ...\eta_n^{(n)}, \gamma_{n+1}, \gamma_{n+2}, ...)$$

Taking into account that

$$\lim_{n\to\infty}\|(\eta_1^{(n)}, ...\eta_n^{(n)}, \gamma_{n+1}, \gamma_{n+2}, ...) - (\eta_1^{(n)}, ..\eta_n^{(n)}, 0, 0....)\| = \lim_{n\to\infty}\sum_{i=n+1}^{\infty}\gamma_i = 0$$

(since $\gamma \in l^2_{\mathbb{K}}$), it follows that the solution, $(I - A)^{-1}\gamma$ is the limit in $l^2_{\mathbb{K}}$ of the sequence $(z_n)_n = ((\eta_1^{(n)}, ..., \eta_n^{(n)}, 0, 0...))_n$.

2) Integral equation of the second kind.

We consider the integral equation (of second kind),

$$x(t) - \int_a^b k(t,s)x(s)\,\mathrm{ds} = g(t)$$

(here " = " means equal almost everywhere), where $g \in L^2_{\mathbb{K}}([a, b])$ is given, $k \in L^2_{\mathbb{K}}([a, b] \times [a, b])$ and x is the desired solution in $L^2_{\mathbb{K}}([a, b])$. Clearly one may write the above equation as

$$(I - K)x = g,$$

where K is the integral operator, defined on $L^2_{\mathbb{K}}([a, b])$ by

$$(Kg)(t) = \int_a^b k(t,s)g(s)\,\mathrm{ds}, \quad g \in L^2_{\mathbb{K}}([a,b])$$

It is known that $K \in \mathcal{B}(L^2_{\mathbb{K}}([a, b]))$ and its norm is less than $\|k\|$ in $L^2_{\mathbb{K}}([a, b]\times$ $\times[a, b])$. Suppose that $\|k\| < 1$.

Proposition 6.1.2 *Given $k \in L^2_{\mathbb{K}}([a,b] \times [a,b])$, with $\|k\| < 1$, the equation*

$$x(t) - \int_a^b k(t,s)x(s)\,\mathrm{ds} = g(t), \quad (a.e.)$$

has a unique solution $f \in L^2_{\mathbb{K}}([a,b])$ for every $g \in L^2_{\mathbb{K}}([a,b])$.
Moreover, for each $n \in \mathbb{N}$, the operator K^n is still an integral operator,

$$(K^n g)(t) = \int_a^b k_n(t,s)g(s)\,\mathrm{ds}, \quad g \in L^2_{\mathbb{K}}([a,b]),$$

with $k_n \in L^2_{\mathbb{K}}([a,b] \times [a,b])$ and the solution of the equation is given by

$$f(t) = g(t) + \int_a^b \widetilde{k}(t,s)g(s)\,\mathrm{ds},$$

where, $\widetilde{k}$ is the sum in $L^2_{\mathbb{K}}([a,b])$ of the series $\sum_{n=1}^{\infty} k_n$.

Proof. By Theorem 6.1.1, since $\|K\| < 1$, $I - K$ is invertible, from which it follows that the integral equation of the second kind has for each $g \in L^2_{\mathbb{K}}([a,b])$ a unique solution in $L^2_{\mathbb{K}}([a,b])$, $(I-K)^{-1}g$. In addition,

$$(I-K)^{-1}g = \sum_{n=0}^{\infty} K^n g$$

By Fubini's theorem,

$$(K^2 g)(t) = \int_a^b k(t,r)Kg(r)\,\mathrm{dr} = \int_a^b k(t,r)\left(\int_a^b k(r,s)g(s)\,\mathrm{ds}\right)\mathrm{dr} =$$

$$= \int_a^b g(s)\left(\int_a^b k(t,r)k(r,s)\,\mathrm{dr}\right)\mathrm{ds} = \int_a^b k_2(t,s)g(s)\,\mathrm{ds},$$

where

$$k_2(t,s) = \int_a^b k(t,r)k(r,s)\,\mathrm{dr}$$

Using the Cauchy-Schwarz inequality, it results that

$$|k_2(t,s)|^2 \le \left(\int_a^b |k(t,r)|^2\,\mathrm{dr}\right)\left(\int_a^b |k(r,s)|^2\,\mathrm{dr}\right),$$

therefore,

$$\int_a^b |k_2(t,s)|^2 \,\mathrm{ds} \le \left(\int_a^b |k(t,r)|^2 \,\mathrm{dr}\right)\left(\int_a^b\int_a^b |k(r,s)|^2 \,\mathrm{dr}\,\mathrm{ds}\right)$$

and

$$\int_a^b\int_a^b |k_2(t,s)|^2 \,\mathrm{ds}\,\mathrm{dt} \le \left(\int_a^b |k(t,r)|^2 \,\mathrm{dr}\,\mathrm{dt}\right)\left(\int_a^b\int_a^b |k(r,s)|^2 \,\mathrm{dr}\,\mathrm{ds}\right)$$

thus $k_2 \in L^2_{\mathbb{K}}([a,b]\times[a,b])$ and $\|k_2\| \le \|k\|^2$.

Proceeding in this manner, an induction argument shows that

$$(K^n g)(t) = \int_a^b k_n(t,s)g(s)\,\mathrm{ds}, \quad \forall n \in \mathbb{N}$$

and $\|k_n\| \le \|k\|^n$, where

$$k_1(t,s) = k(t,s), \text{ and } k_n(t,s) = \int_a^b k(t,r)k_{n-1}(r,s)\,\mathrm{dr}, \quad \forall n \in \mathbb{N}$$

is in $L^2_{\mathbb{K}}([a,b]\times[a,b])$.

Since $\|k\| < 1$, the series $\sum_{n\ge 1}\|k\|^n$ converges, which, combined with $\|k_n\| \le \|k\|^n$ ensures the convergence in $L^2_{\mathbb{K}}([a,b]\times[a,b])$ of the series $\sum_{n\ge 1} k_n$. Let $\widetilde{k} \in L^2_{\mathbb{K}}([a,b]\times[a,b])$ be its sum, and $\widetilde{K}$ the integral operator with kernel $\widetilde{k}$. Then,

$$\|\sum_{j=1}^n K^j g - \widetilde{K}g\| = \|(\sum_{j=1}^n K^j - \widetilde{K})g\| \le \|\sum_{j=1}^n k^j - \widetilde{k}\| \cdot \|g\| \longrightarrow 0,$$

as $n \to \infty$. It results that

$$(I-K)^{-1}g = \sum_{n=0}^{\infty} K^n g = (I+\widetilde{K})g,$$

which shows that the solution of the equation in discussion is given by

$$f(t) = g(t) + \int_a^b \widetilde{k}(t,s)g(s)\,\mathrm{ds}$$

.

We shall end with a concrete example. To solve the integral equation

$$x(t) - \lambda \int_0^1 e^{t-s} x(s)\, \mathrm{ds} = g(t), \quad g \in L^2_{\mathbb{K}}([0,1]),$$

where, $|\lambda| < 1$, let $k_1(t,s) = \lambda\, e^{t-s}$. A short computation shows that $\|k_1\| <$ < 1. We have,

$$k_2(t,s) = \lambda^2 \int_0^1 e^{t-r} e^{r-s}\, \mathrm{dr} = \lambda^2 e^{t-s},$$

and, inductively, $k_n(t,s) = \lambda^n e^{t-s}$. Then, the solution of the equation is, by the above proposition,

$$f(t) = g(t) + \sum_{n=1}^{\infty} \lambda^n \int_0^1 e^{t-s} g(s)\, \mathrm{ds} = g(t) + \frac{\lambda}{1-\lambda} \int_0^1 e^{t-s} g(s)\, \mathrm{ds}$$

Even the series $\sum_{n\geq 1} \lambda^n$ converges only for $|\lambda| < 1$, a straightforward computation shows that for all $\lambda \neq 1$,

$$g(t) + \frac{\lambda}{1-\lambda} \int_0^1 e^{t-s} g(s)\, \mathrm{ds}$$

is still a solution of the discussed equation.

6.2 Spectrum, definition, elementary properties, spectral radius

6.2.1 Spectrum

At the beginning, let $\mathcal{A}$ be a unital Banach algebra over the field $\mathbb{K}$, with unity I ($\|I\| = 1$).

Definition. For every $T \in \mathcal{A}$, we define the *spectrum* of T, denoted by $\sigma(T)$, as the set

$$\sigma(T) = \{\lambda \in \mathbb{K} \mid \lambda I - T \notin \mathcal{G}(\mathcal{A})\}$$

The complement of $\sigma(T)$ is the *resolvent set* of T, denoted by $\rho(T)$. The mapping from $\rho(T)$ to $\mathcal{A}$, defined by

$$R(T;\lambda) = (\lambda I - T)^{-1}$$

is called the *resolvent function* of T.

Remark. When the algebra $\mathcal{A}$ is real ($\mathbb{K} = \mathbb{R}$), it is possible to find $T \in \mathcal{A}$, with $\sigma(T) = \emptyset$. For example, taking $\mathcal{A} = \mathcal{L}(\mathbb{R}^2)$, the element

$$T \rightsquigarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

has its spectrum void. Later we will see that this does not happen if the algebra is complex.

Proposition 6.2.1 *For every $T \in \mathcal{A}$, the resolvent set of T, $\rho(T)$ is an open subset of $\mathbb{C}$.*

Proof. Let $\lambda_o \in \rho(T)$ and λ be arbitrary in $\mathbb{K}$. Then,

$$|\lambda - \lambda_o| = \|(\lambda I - T) - (\lambda_o I - T)\|,$$

therefore, by Corollary 6.1.1, if

$$\|(\lambda I - T) - (\lambda_o I - T)\| \leq \frac{1}{\|(\lambda_o I - T)^{-1}\|},$$

$\lambda I - T$ is invertible, so $\lambda \in \rho(T)$. It follows that

$$\mathcal{B}(\lambda_o, \|(\lambda_o I - T)^{-1}\|^{-1}) \subset \rho(T).$$

Further, in this section, $\mathcal{A}$ is supposed to be a *complex* unital Banach algebra ($\mathbb{K} = \mathbb{C}$), $\mathcal{A} \neq \{0\}$. In order to obtain the main result of this section, which shows that $\sigma(T) \neq \emptyset$, $\forall T \in \mathcal{A}$, we need some preliminary results.

Lemma 6.2.1 *For every $T \in \mathcal{A}$, the resolvent function of T has the following properties:*
1) $R(T; \cdot)$ is continuous on $\rho(T)$;
2) For λ, $\mu \in \rho(T)$,

$$R(T;\lambda) - R(T;\mu) = (\mu - \lambda)\ R(T;\lambda)\ R(T;\mu)$$

(the Hilbert equation);
3) $R(T;\cdot)$ is an holomorphic function on $\rho(T)$
4) $\lim\limits_{|\lambda| \to \infty} R(T;\lambda) = 0$.

Proof. 1) follows from the continuity of the map $T \longmapsto T^{-1}$ (Corollary 6.1.2) which shows that $R(T;\cdot)$ is continuous.

2) results from the next short computation, by multiplying at the right by $R(T;\mu)$:

$$I = (\mu I - T)\ R(T;\mu) = [(\mu - \lambda)I + \lambda I - T]\ R(T;\mu) =$$

$$= (\mu - \lambda)\ R(T;\mu) + (\lambda I - T)\ R(T;\mu)$$

3) Given $\lambda_o \in \rho(T)$, by 1), 2), we have

$$\lim_{\lambda \to \lambda_o} \frac{R(T;\lambda) - R(T;\lambda_o)}{\lambda - \lambda_o} = \lim_{\lambda \to \lambda_o} -R(T;\lambda)\ R(T;\lambda_o) = -R(T;\lambda_o)^2$$

which shows that $R(T;\cdot)$ is holomorphic (it posses a derivative wherever it is defined).

(4) Given $\lambda \in \rho(T)$ such that $|\lambda| > \|T\|$, by Theorem 6.1.1, since $\|T\ /\ \lambda\| < 1$, there exists $(I - T/\lambda)^{-1}$ and

$$\left(I - \frac{T}{\lambda}\right)^{-1} = \sum_{n=0}^{\infty} \left(\frac{T}{\lambda}\right)^n,$$

or, equivalently, $(\lambda I - T) \in \mathcal{G}(\mathcal{A})$ and

$$(\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \frac{1}{\lambda} \left(\frac{T}{\lambda}\right)^n$$

Estimating the norm of $(\lambda I - T)^{-1}$ it results that

$$\|R(T;\lambda)\| = \|\ (\lambda I - T)^{-1}\ \| \le \sum_{n=0}^{\infty} \frac{1}{|\lambda|} \left\|\frac{T}{\lambda}\right\|^n =$$

$$= \frac{1}{|\lambda|} \cdot \frac{1}{1 - \frac{\|T\|}{|\lambda|}} = \frac{1}{|\lambda| - \|T\|} \longrightarrow 0,$$

as $|\lambda| \to \infty$, so 4) holds.

By the proof of the above proposition, it follows that

Corollary 6.2.1 *1)* $\sigma(T) \subset \overline{B}(0, \|T\|)$;
2) For every $\lambda \in \|T\|$,

$$(\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$$

Lemma 6.2.2 *(The Liouville's theorem on Banach spaces) Let $\mathcal{O}(D, \mathcal{A})$ the space of holomorphic functions from $D = \overset{\circ}{D} \subset \mathbb{C}$ to $\mathcal{A}$. Then,*
1) For every $x \in \mathcal{O}(D, \mathcal{A})$ and $f \in \mathcal{A}^*$, the function $f \circ x$ is still in $\mathcal{O}(D, \mathcal{A})$.
2) Each bounded $x \in \mathcal{O}(\mathbb{C}, \mathcal{A})$ is constant.

Proof. 1) Given $f \in \mathcal{A}^*$, $x \in \mathcal{O}(D, \mathcal{A})$ and $\lambda_o \in D$, we have

$$\left\| \frac{(f \circ x)(\lambda) - (f \circ x)(\lambda_o)}{\lambda - \lambda_o} - f(x'(\lambda_o)) \right\| = \left\| f \left(\frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} - x'(\lambda_o) \right) \right\| \leq$$

$$\leq \|f\| \left\| \frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} - x'(\lambda_o) \right\|,$$

(where $x'(\lambda_o)$ is the derivative of x at λ_o). Thus $f \circ x \in \mathcal{O}(D, \mathcal{A})$ and

$$(f \circ x)'(\lambda_o) = f(x'(\lambda_o))$$

2) Let $x \in \mathcal{O}(\mathbb{C}, \mathcal{A})$, $\|x(\lambda)\| \leq M$, $\forall \lambda \in \mathbb{C}$ and f arbitrary in $\mathcal{A}^*$. Clearly, by

$$\|(f \circ x)(\lambda)\| \leq \|f\| \; \|x(\lambda)\| \leq M \; \|f\|,$$

we have $f \circ x \in \mathcal{O}(\mathbb{C}, \mathbb{C})$, therefore by the Liouville's theorem (see Appendix D), $f \circ x$ must be constant,

$$(f \circ x)(\lambda) = (f \circ x)(0), \quad \forall \lambda \in \mathbb{C}$$

This enables us to conclude that for a given arbitrary λ in $\mathbb{C}$,

$$f(x(\lambda) - x(0)) = 0, \quad \forall f \in \mathcal{A}^*$$

Then, by Corollary 3.2.2, $x(\lambda) - x(0) = 0$, $\forall \lambda \in \mathbb{C}$, so x is constant on $\mathbb{C}$.

Theorem 6.2.1 *For every element T in a complex unital Banach algebra $\mathcal{A}$, the spectrum of T is a compact, nonempty subset of $\mathbb{C}$.*

Proof. By Proposition 6.2.1, $\rho(T)$ is open and, by Corollary 6.2.1, $\sigma(T)$ is contained in $\overline{B}(0, \|T\|)$. It follows that, the complement of $\rho(T)$, $\sigma(T)$ is closed, and, as it is also bounded, it is compact.

Suppose that $\sigma(T) = \emptyset$. Then, the resolvent function of T is an analytic function defined on the whole $\mathbb{C}$ and it is bounded since $\lim_{|\lambda| \to \infty} R(T; \lambda) = 0$.

By the Lemma 6.2.2, we have that $R(T;\cdot)$ is constant, this means, because its limit at ∞ is zero that $R(T;\lambda) = 0$, $\forall\lambda \in \mathbb{C}$. Hence,

$$I = R(T;\lambda)^{-1}R(T;\lambda) = 0,$$

contradiction ($\mathcal{A} \neq \{0\}$).

As a nice consequence of the above theorem we have the next corollary, known as Gelfand-Mazur theorem.

Corollary 6.2.2 *If $\mathcal{A}$ is a division ring (i.e. $\mathcal{G}(\mathcal{A}) = \mathcal{A}\setminus\{0\}$), then, $\mathcal{A}$ is isometrically isomorphic to $\mathbb{C}$.*

Proof. For each T in $\mathcal{A}$, the spectrum of T is nonempty, thus $\exists\lambda \in \sigma(T)$, such that $\lambda I - T$ is not invertible. It follows that $\lambda I - T = 0$, therefore $T = \lambda I$. In addition, $\|T\| = \|\lambda I\| = |\lambda|$.

Proposition 6.2.2 *Suppose $0 \notin \sigma(T)$ (T is invertible). Then*

$$\sigma(T^{-1}) = \{\frac{1}{\lambda} \mid \lambda \in \sigma(T)\}$$

Proof. Let $\lambda \in \sigma(T^{-1})$ (necessarily $\lambda \neq 0$). Then,

$$\lambda I - T^{-1} \notin \mathcal{G}(\mathcal{A}) \iff I - \frac{1}{\lambda}T^{-1} \notin \mathcal{G}(\mathcal{A}) \iff$$

$$\iff T^{-1}(T - \frac{1}{\lambda}I) \notin \mathcal{G}(\mathcal{A}) \iff T - \frac{1}{\lambda}I \notin \mathcal{G}(\mathcal{A}),$$

thus $1/\lambda \in \sigma(T)$.

6.2.2 The spectral radius

The next lemma will enable us to define the spectral radius.

Lemma 6.2.3 *Suppose that $\mathcal{A}$ is a unital algebra equipped with a submultiplicative norm. Then, for every $T \in \mathcal{A}$, there exists*

$$\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\geq 1} \|T^n\|^{\frac{1}{n}}$$

Proof. Let $m \geq 1$ be a fixed positive integer and $n \geq 1$, arbitrary in $\mathbb{N}$. Then, there exist the positive integers $q(n)$, $r(n)$ such that

$$n = mq(n) + r(n) \text{ and } 0 \leq r(n) \leq m - 1$$

It follows that

$$\lim_{n\to\infty} \frac{r(n)}{n} = 0 \text{ and } \lim_{n\to\infty} \frac{q(n)}{n} = \frac{1}{m}$$

Taking into account that the norm is submultiplicative, we have

$$\|T^n\|^{\frac{1}{n}} \leq \|T^m\|^{\frac{q(n)}{n}} \|T\|^{\frac{r(n)}{n}},$$

hence

$$\limsup_n \|T^n\|^{\frac{1}{n}} \leq \limsup_n \|T^m\|^{\frac{q(n)}{n}} \|T\|^{\frac{r(n)}{n}} =$$

$$= \|T^m\|^{\lim\limits_{n\to\infty} \frac{q(n)}{n}} \|T\|^{\lim\limits_{n\to\infty} \frac{r(n)}{n}} = \|T^m\|^{\frac{1}{m}}$$

It results that

$$\limsup_n \|T^n\|^{\frac{1}{n}} \leq \inf_{m\geq 1} \|T^m\|^{\frac{1}{m}}$$

which combined with

$$\inf_{n\geq 1} \|T^m\|^{\frac{1}{m}} \leq \liminf_n \|T^n\|^{\frac{1}{n}}$$

shows that there exists $\lim\limits_{n\to\infty} \|T^n\|^{\frac{1}{n}}$ and this limit is equal to $\inf\limits_{n\geq 1} \|T^n\|^{\frac{1}{n}}$.

Definition. Suppose that $\mathcal{A}$ is a complex unital Banach algebra. For every $T \in \mathcal{A}$, we define the *spectral radius* of T (denoted $\|T\|_\sigma$), by

$$\|T\|_\sigma = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$$

Lemma 6.2.4 *For each λ such that $|\lambda| > \|T\|_\sigma$ the series*

$$\sum_{n\geq 0} \frac{1}{\lambda^n} T^n$$

is convergent.

Proof. Using the Cauchy-Hadamard theorem (see Appendix D), we obtain that the series $\sum_{n\geq 0}\|T^n\|\ t^n$converges for each t with

$$|t| < \frac{1}{\limsup\limits_{n} \|T^n\|^{\frac{1}{n}}} = \frac{1}{\|T\|_\sigma}$$

Since for each λ, with $|\lambda| > \|T\|_\sigma$, we have

$$0 \leq \frac{1}{|\lambda|} \leq \frac{1}{\|T\|_\sigma},$$

it follows that the series

$$\sum_{n\geq 0} \|\frac{1}{\lambda^n}T^n\|$$

converges. Thus, the series

$$\sum_{n\geq 0} \frac{1}{\lambda^n}T^n$$

is absolutely convergent, hence convergent.

Remark. Since for each λ, with $|\lambda| > \|T\|_\sigma$ the series $\sum_{n\geq 0}(1/\lambda^n)T^n$ is convergent, it results that

$$\lim_{n\to\infty} \frac{1}{\lambda^n}T^n = 0$$

Next result shows that the spectral radius of T is the smallest positive real r such that $\sigma(T) \subset \overline{B}(0,r)$.

Theorem 6.2.2 *Let T be in the complex unital Banach algebra $\mathcal{A}$. Then,*
1) Every λ with $|\lambda| > \|T\|_\sigma$ is in $\rho(T)$ and

$$R(T;\lambda) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$$

2) The spectral radius of T is

$$\|T\|_\sigma = \sup_{\lambda\in\sigma(T)} |\lambda|$$

Proof. 1) By the previous lemma, if $|\lambda| > \|T\|_\sigma$, the series

$$\sum_{n\geq 0} \frac{1}{\lambda^n} T^n$$

is convergent. Considering, for any positive integer $n \geq 1$,

$$S_n = \sum_{k=0}^{n} \frac{1}{\lambda^k} T^k,$$

we have

$$\left(I - \frac{T}{\lambda}\right) S_n = S_n \left(I - \frac{T}{\lambda}\right) = I - \left(\frac{T}{\lambda}\right)^{n+1}$$

Hence,

$$\left(I - \frac{T}{\lambda}\right) \left(\sum_{n=0}^{\infty} \frac{T^n}{\lambda^n}\right) = I$$

It results that the element $\lambda I - T$ is invertible, therefore $\lambda \in \rho(T)$ and

$$(\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$$

2) From 1), clearly, $\sigma(T) \subset \overline{B}(0, \|T\|_\sigma)$. We show that, for each $r > 0$, $\|T\|_\sigma \leq \sup\limits_{\lambda \in \sigma(T)} |\lambda| + r$. In order to do that, let us denote by $\xi_r = \sup\limits_{\lambda \in \sigma(T)} |\lambda| + r$ and, observe that $\xi_r \notin \sigma(T)$. Then, $\xi_r \in \rho(T)$ and there exists

$$R(T; \xi_r) = (\xi_r I - T)^{-1}$$

On the other hand, for any $|\lambda| > \|T\|_\sigma$,

$$R(T; \lambda) = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} T^n$$

By the uniqueness of the development in Laurent series (Appendix D), it follows that

$$R(T; \xi_r) = \sum_{n=0}^{\infty} \frac{1}{\xi_r^{n+1}} T^n,$$

thus

$$\lim_{n \to \infty} \left\| \frac{1}{\xi_r^n} T^n \right\| = 0.$$

Hence, one can conclude that $\|T^n\|^{\frac{1}{n}} \leq \xi_r$ (for n sufficiently large), so $\|T\|_\sigma \leq$ $\leq \xi_r$, which ends the proof.

6.2.3 More about the spectrum in involutive Banach algebras

Suppose, as before, that $\mathcal{A}$ is a complex unital Banach algebra.

Definition. A map $T \longmapsto T^*$ from $\mathcal{A}$ to $\mathcal{A}$ is called an *involution* if
i) $(\alpha T + \beta S)^* = \overline{\alpha} T^* + \overline{\beta} S^*, \quad \forall T, S \in \mathcal{A}, \alpha, \beta \in \mathbb{C}$;
ii) $(TS)^* = S^* T^*, \quad \forall T, S \in \mathcal{A}$;
iii) $T^{**} = T, \quad \forall T \in \mathcal{A}$.

The involutions that naturally occur in Banach algebras are isometric,

$$\|T^*\| = \|T\|, \quad \forall T \in \mathcal{A}$$

Definition. A Banach algebra $\mathcal{A}$ endowed with an involution is called an *involutive Banach algebra.*

If the involution is isometric and

$$\|TT^*\| = \|T\|^2, \quad \forall T \in \mathcal{A}$$

the algebra is called a $\mathbb{C}^*-$*algebra.*

Remarks. 1. *In every unital involutive algebra, $I^* = I$.*

2. *If $T \in \mathcal{G}(\mathcal{A})$, then, a short computation shows that $T^* \in \mathcal{G}(\mathcal{A})$ and $(T^*)^{-1} = (T^{-1})^*$.*

We note that the algebra of all bounded operators on a Hilbert space X, $\mathcal{B}(X)$ is a $\mathbb{C}^*-$algebra. Some of the particular terminology of this algebra is inherited in general involutive Banach algebras.

Definition. Let $\mathcal{A}$ be a unital involutive Banach algebra. An element $T \in \mathcal{A}$ is said to be *normal* if $TT^* = T^*T$. An element $U \in \mathcal{A}$ is said to be *unitary* if $UU^* = U^*U = I$. An element $A \in \mathcal{A}$ is called *self-adjoint* if $A^* = A$. A *projection* is a self-adjoint element $P \in \mathcal{A}$ such that $P^2 = P$.

Theorem 6.2.3 *Let $\mathcal{A}$ be a unital involutive algebra. Then, for every $T \in$ $\in \mathcal{A}$,*

$$\sigma(T^*) = \{\overline{\lambda} \mid \lambda \in \sigma(T)\}.$$

Proof. Taking into account the above second remark, we have

$$\lambda \in \sigma(T^*) \Longleftrightarrow \lambda I - T^* \in \mathcal{G}(\mathcal{A}) \Longleftrightarrow$$

$$\Longleftrightarrow (\lambda I - T^*)^* = \overline{\lambda} I - T \in \mathcal{G}(\mathcal{A}) \Longleftrightarrow \overline{\lambda} \in \sigma(T)$$

Theorem 6.2.4 *Suppose that $\mathcal{A}$ is a $\mathbb{C}^*-$algebra. Then,*
1) For every $U \in \mathcal{A}$, U unitary, $\sigma(U) \subset \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$;
2) For every $A \in \mathcal{A}$, A self-adjoint, $\sigma(A) \subset \mathbb{R}$;
3) If $T \in \mathcal{A}$, T normal, $\|T\|_\sigma = \|T\|$.

Proof. 1) If U is unitary, then $\|U\| = \|U^*\| = 1$. It follows that $\sigma(U) \subset$ $\subset \overline{B}(0,1)$, which implies that any $\lambda \in \sigma(U)$ satisfies $|\lambda| \le 1$. On the other hand, since $U^* = U^{-1}$,

$$\sigma(U^*) = \{\frac{1}{\lambda} \mid \lambda \in \sigma(U)\} \subset \overline{B}(0,1),$$

thus, any $\lambda \in \sigma(U)$ satisfies

$$\frac{1}{|\lambda|} \le 1 \iff |\lambda| \ge 1$$

We conclude that $\sigma(U) \subset \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$.

2) First, we note that, for each $\alpha \in \mathbb{C}$,

$$\sigma(T + \alpha I) = \sigma(T) + \alpha$$

Then, if $a + bi \in \sigma(T)$ $(a, b \in \mathbb{R})$, for each $r \in \mathbb{R}$, the complex number $a + bi + ri$ is in $\sigma(T) + ri = \sigma(T + riI) \subset \overline{B}(0, \|T + riI\|)$. We have,

$$|a + bi + ri|^2 \le \|T + riI\|^2 \quad \|(T + riI)(T + riI)^*\| =$$
$$= \|T^2 - r^2 I\| \le \|T\|^2 + r^2$$

Hence, $2br \le \|T\|^2 - a^2$, $\forall r \in \mathbb{R}$,and we infer $b = 0$.

3) If T is normal, then $\|T\|^2 = \|T^2\|$. Indeed,

$$\|T^2\|^2 = \|(T^2)(T^2)^*\| = \|(TT^*)(TT^*)^*\| = \|TT^*\|^2 = \|T\|^4$$

By an induction argument we obtain that

$$\|T^{2^n}\|^{\frac{1}{2^n}} = \|T\|, \quad \forall n \in \mathbb{N},$$

and because of,

$$\|T^{2^n}\|^{\frac{1}{2^n}} \longrightarrow \|T\|_\sigma,$$

as $n \to \infty$, everything is clear.

Taking into account that $\mathcal{B}(X)$ (where X is a Hilbert space) is a $\mathbb{C}^*-$algebra, for bounded operators on Hilbert space all the above results are valid, hence we have:

Theorem 6.2.5 *Let X be a Hilbert space. Then,*
1) $\forall T \in \mathcal{B}(X)$, $\sigma(T) \neq \emptyset$; $\sigma(T)$ is compact and contained in $\overline{B}(0, \|T\|_\sigma)$;
2) $\forall T \in \mathcal{B}(X)$, $\rho(T)$ is open;
3) If $0 \notin \sigma(T)$, $\sigma(T^{-1}) = \{1/\lambda \mid \lambda \in \sigma(T)\}$;
4) $\sigma(T^*) = \{\overline{\lambda} \mid \lambda \in \sigma(T)\}$;
5) For every U unitary, $\sigma(U) \subset \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$;
6) For every A self-adjoint, $\sigma(T) \subset \mathbb{R}$;
7) For every T normal, $\|T\|_\sigma = \|T\|$.

6.3 The spectrum of compact self-adjoint operators

In this section X is a Hilbert space over the field $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Definition. For each $T \in \mathcal{B}(X)$, the *point spectrum* of T, denoted by $\sigma_p(T)$, is the set

$$\sigma_p(T) = \{\lambda \in \mathbb{K} \mid \lambda I - T \text{ is not injective}\}$$

Remarks. 1. Clearly $\sigma_p(T)$ is contained in $\sigma(T)$.
2. A number $\lambda \in \sigma_p(T)$, if and only if there exists $x \neq 0$ such that $(\lambda I - T)x = 0$, or equivalently, the subspace Ker $(\lambda I - T) \neq \{0\}$.

Definition. A number λ in $\sigma_p(T)$ is called an *eigenvalue* of T and the elements which are not zero of Ker $(\lambda I - T)$ are called the *eigenvectors* corresponding to λ.

Remark. If $X = \mathbb{K}^n$, then, since on finite dimensional spaces a linear operator is injective if and only if it is bijective, it follows that, for each linear operator T, $\sigma_p(T) = \sigma(T)$. In addition, $\lambda \in \sigma(T)$, if and only if

$$det\ (\lambda I - T) = 0,$$

so, λ is a root of this algebraic equation.

Lemma 6.3.1 *Let $(\lambda_n)_n$ be a sequence of eigenvalues such that $\lambda_n \neq \lambda_m$, $\forall m \neq n$ and $x_n \in \text{Ker}(\lambda_n I - T)$, $\forall n$. Then, the set $(x_n)_n$ is linearly independent.*

Proof. We will proceed by induction. For $n = 1$, since each nonzero element form itself a linearly independent set, the statement is true. Suppose, that all hold for a positive integer n, and let λ_{n+1} be an eigenvalue, $\lambda_{n+1} \neq \lambda_m$, $\forall m = 1, 2, ...n$, $x_{n+1} \in \mathrm{Ker}(\lambda_{n+1}I - T)$. If the set $\{x_k\}_{1 \leq k \leq n+1}$ is linearly dependent, then there exist $\alpha_1, \alpha_2, ...\alpha_n$ (not all zero) such that

$$x_{n+1} = \sum_{k=1}^{n} \alpha_k x_k,$$

thus, $T(x_{n+1}) = \sum_{k=1}^{n} \alpha_k T(x_k)$, i.e.

$$\lambda_{n+1} x_{n+1} = \sum_{k=1}^{n} \alpha_k \lambda_k x_k.$$

Combining the above relations, it follows that

$$\sum_{k=1}^{n} \alpha_k (\lambda_{n+1} - \lambda_k) x_k = 0$$

in which not all the coefficients are zero that means the set $\{x_k\}_{1 \leq k \leq n}$ is linearly dependent (contradiction).

Remark. Given a sequence of eigenvalues $(\lambda_n)_n$ such that $\lambda_n \neq \lambda_m$, $\forall m \neq n$ and $x_n \in \mathrm{Ker}(\lambda_n I - T)$, $\forall n$, let X_n be the space spanned by$\{x_k\}_{1 \leq k \leq n}$, $\forall n$. Then, with the previous lemma, obviously, X_n is strictly contained in X_{n+1}, $\forall n$.

Further T always will be a self-adjoint operator, $T = T^*$.

Proposition 6.3.1 *The point spectrum $\sigma_p(T)$ of a self-adjoint operator T is contained in the interval $[m_T, M_T]$, and, for every λ, $\mu \in \sigma_p(T)$, $\lambda \neq \mu$, the subspaces* Ker $(\lambda I - T)$ *and* Ker $(\mu I - T)$ *are orthogonal each other.*

Proof. For an arbitrary $\lambda \in \sigma_p(T)$,

$$\lambda < x, x > = < Tx, x >$$

for some $x \neq 0$. Then,

$$m_T \leq \inf_{x \neq 0} \frac{< Tx, x >}{< x, x >} \leq \lambda \leq \sup_{x \neq 0} \frac{< Tx, x >}{< x, x >} = M_T$$

Now, for $x \in \text{Ker}\ (\lambda I - T)$ and $y \in \text{Ker}\ (\mu I - T)$ we have

$$\lambda < x, y > = < \lambda x, y > = < Tx, y > =$$

$$= < x, Ty > = < x, \mu y > = \mu < x, y >,$$

which, since $\lambda \neq \mu$, gives $< x, y > = 0$, therefore $\text{Ker}\ (\lambda I - T) \perp \text{Ker}\ (\mu I - T)$.

Proposition 6.3.2 *Let T be a self-adjoint operator on the Hilbert space X. Then, $\lambda \in \sigma_p(T)$ if and only if $\overline{(\lambda I - T)(X)} \neq X$.*

Proof. Given λ in $\sigma_p(T)$, $\text{Ker}\ (\lambda I - T) \neq \{0\}$,which is equivalent to

$$\text{Ker}(\lambda I - T)^{\perp} \neq X$$

It follows, since

$$\text{Ker}(\lambda I - T)^{\perp} = \overline{(\lambda I - T)^{*}(X)}$$

and $\lambda I - T$ is self-adjoint that $\overline{(\lambda I - T)(X)} \neq X$.

Conversely, if $\overline{(\lambda I - T)(X)} \neq X$, then, there exists $x_o \neq 0$, x_o belonging to

$$\overline{(\lambda I - T)(X)}^{\perp} = \text{Ker}(\lambda I - T)^{*} = \text{Ker}(\overline{\lambda} I - T)$$

It follows that $\overline{\lambda} x_o = T x_o$, thus $\overline{\lambda} \in \sigma_p(T) \subset \mathbb{R}$. Then $\lambda = \overline{\lambda} \in \sigma_p(T)$.

Remark. For every $T \in \mathcal{B}(X)$ one defines the *residual spectrum* of T by

$$\sigma_r(T) = \{\lambda \in \mathbb{K} \mid \lambda I - T \text{ is injective and } \overline{(\lambda I - T)(X)} \neq X\}$$

The previous proposition shows that if T is a self-adjoint operator, its residual spectrum is empty.

Proposition 6.3.3 *Let T be a self-adjoint operator on the Hilbert space X. Then, $\lambda \in \rho(T)$ if and only if there exists $\mu > 0$ such that*

$$\|(\lambda I - T)x\| \geq \mu \|x\|, \quad \forall x \in X.$$

Proof. Suppose that $\lambda \in \rho(T)$, i.e. there exists $(\lambda I - T)^{-1} \in \mathcal{B}(X)$. Since $\lambda I - T \in \mathcal{B}(X)$, $\exists M > 0$ such that

$$\|(\lambda I - T)x\| \leq M\|x\|, \quad \forall x \in X,$$

in which, setting $x = (\lambda I - T)^{-1}y$, one obtains

$$\mu\|x\| \leq \|(\lambda I - T)x\|, \quad \forall x \in X,$$

(where $\mu = 1/M$).

Conversely, if there exists $\mu > 0$ such that

$$\|(\lambda I - T)x\| \geq \mu\|x\|, \quad \forall x \in X,$$

$\lambda I - T$ is injective. Then $\lambda \notin \sigma_p(T)$, and, by Proposition 6.3.2, $\overline{(\lambda I - T)(X)} =$ $= X$. It follows that for every $y \in X$, there exists $(x_n)_n \subset X$ such that

$$y_n = \lambda x_n - Tx_n \longrightarrow y,$$

as $n \to \infty$. By

$$\|x_n - x_m\| \leq \frac{1}{\mu}\|y_n - y_m\|,$$

it results that $(x_n)_n$ is a Cauchy sequence, which means, since X is complete, that $\exists x = \lim\limits_n x_n$. Consequently, $y = \lambda x - Tx$. We have proven that $\lambda I - T$ is surjective, and, as it is also injective, all is done.

As an immediate consequence of this theorem, we have:

Corollary 6.3.1 *Let T be a self-adjoint operator on the Hilbert space X. Then, $\lambda \in \sigma(T)$ if and only if there exists a sequence $(x_n)_n$, with $\|x_n\| = 1$ such that $\lambda x_n - Tx_n \longrightarrow 0$, as $n \to \infty$.*

Proposition 6.3.4 *For every self-adjoint operator T, the spectrum of T is contained in $[m_T, M_T]$ and m_T, $M_T \in \sigma(T)$.*

Proof. Suppose that $\lambda < m_T$, so $m_T - \lambda > 0$. Then, since for arbitrary $x \in X$,

$$< (T - \lambda I)x, x >=< Tx, x > -\lambda < x, x >\geq (m_T - \lambda)\|x\|^2$$

it results that $T - \lambda I \in \mathcal{A}_+(X)$. Thus, by the Cauchy-Schwarz inequality, we have

$$\|(T - \lambda I)x\| \cdot \|x\| \geq < (T - \lambda I)x, x > \geq (m_T - \lambda)\|x\|^2, \ \forall x \in X,$$

and using Proposition 6.3.3, it follows that $\lambda \in \rho(T)$.

Similarly, if $\lambda > M_T$, the operator $\lambda I - T$ is positive, because of

$$< (\lambda I - T)x, x > = \lambda < x, x > - < Tx, x > \geq (\lambda - M_T)\|x\|^2$$

and by the same argument as above, one obtains

$$\|(\lambda I - T)x\| \cdot \|x\| \geq < (\lambda I - T)x, x > \geq (\lambda - M_T)\|x\|^2, \quad \forall x \in X,$$

therefore $\lambda \in \rho(T)$.

We have seen that $\sigma(T) \subset [m_T, M_T]$. Further, we prove that $m_T \in \sigma(T)$. Since $m_T = \inf\limits_{\|x\|=1} < Tx, x >$, there exists a sequence $(x_n)_n$, with $\|x_n\| = 1$ such that

$$< Tx_n, x_n > \longrightarrow m_T,$$

as $n \to \infty$. If we check that $(m_T I - T)x_n \longrightarrow 0$, by the Corollary 6.3.1, it follows $m_T \in \sigma(T)$. In the next computation one uses the Cauchy-Schwarz inequality corresponding to the positive operator $T - m_T I$. Thus, we have,

$$\|(T - m_T I)x_n\|^4 = < (T - m_T I)x_n, (T - m_T I)x_n >^2 \leq$$

$$\leq < (T - m_T I)x_n, x_n > \cdot < (T - m_T I)^2 x_n, (T - m_T I)x_n > \leq$$

$$\leq < (T - m_T I)x_n, x_n > \cdot \|T - m_T I\| \cdot \|(T - m_T I)x_n\|^2$$

which, combined with $< (T - m_T I)x_n, x_n > \longrightarrow 0$, shows that

$$(m_T I - T)x_n \longrightarrow 0$$

Similarly, $M_T \in \sigma(T)$.

6.4 Spectral properties of compact self-adjoint operators

In order to obtain the description of the spectrum of a compact self-adjoint operator on a Hilbert space, we give first a result concerning the spectrum of compact operators. We have to notice that the next proposition remains valid if X is only a Banach space.

Proposition 6.4.1 *Let T be a compact operator on the Hilbert space X. Then, the point spectrum of T, $\sigma_p(T)$ is an (at most) countable set with zero as the only possible accumulation point.*

Proof. Suppose that there exists $\lambda_o \neq 0$ an accumulation point of the point spectrum of T. Then, there exists $(\lambda_n)_n \subset \sigma_p(T)$, $\lambda_n \neq \lambda_m$, $\lambda_n \longrightarrow \lambda_o$. Let $(x_n)_n$ be a sequence of corresponding eigenvectors,

$$x_n \in \text{Ker}(\lambda_n I - T), \quad \forall n$$

and X_n the linear subspace of X spanned by $\{x_1, x_2, ..., x_n\}$. Clearly, X_n is closed (it is finite dimensional) and X_{n-1} is strictly contained in X_n. For every $n \geq 2$, since $X_{n-1} \subsetneq X_n$, by the Projection theorem, there exists $x_n^o \in X_n \cap X_{n-1}^{\perp}$, $\|x_n^o\| = 1$. Then, for $x \in X_{n-1}$,

$$\|x_n^o - x\|^2 =< x_n^o - x, x_n^o - x >=$$

$$= \|x_n^o\|^2 - < x_n^o, x > - < x, x_n^o > + \|x\|^2 =$$

$$= \|x_n^o\|^2 + \|x\|^2 \geq \|x_n^o\|^2 = 1,$$

therefore $\|x_n^o - x\| \geq 1$, $\forall x \in X_{n-1}$.

Now, we consider the sequence $((1/\lambda_n)x_n^o)_n$, which, as $\lambda_n \longrightarrow \lambda_o \neq 0$, is bounded and we show that the sequence

$$\left(T\left(\frac{1}{\lambda_n}x_n^o\right)\right)_n$$

has no convergent subsequences. Indeed, for every n, $x_n^o \in X_n$, thus,

$$x_n^o = \sum_{k=1}^{n} \alpha_k x_k, \ (\alpha_k \in \mathbb{K}),$$

which implies that

$$\alpha_n x_n = x_n^o - \sum_{k=1}^{n-1} \alpha_k x_k,$$

and applying T,

$$Tx_n^o = \sum_{k=1}^{n} \alpha_k \lambda_k x_k.$$

It follows that

$$\frac{1}{\lambda_n}Tx_n^o = \sum_{k=1}^{n-1}\alpha_k\frac{\lambda_k}{\lambda_n}x_k + \alpha_n x_n.$$

Then, for an arbitrary $x \in X_{n-1}$, we obtain

$$(*)\quad \|\frac{1}{\lambda_n}Tx_n^o - x\| = \|\sum_{k=1}^{n-1}\alpha_k\frac{\lambda_k}{\lambda_n}x_k + \alpha_n x_n - x\| =$$

$$= \|\sum_{k=1}^{n-1}\alpha_k\frac{\lambda_k}{\lambda_n}x_k + x_n^o - \sum_{k=1}^{n-1}\alpha_k x_k - x\| =$$

$$= \|x_n^o - (\sum_{k=1}^{n-1}\alpha_k(1-\frac{\lambda_k}{\lambda_n})x_k + x)\| \geq 1,$$

since

$$\sum_{k=1}^{n-1}\alpha_k(1-\frac{\lambda_k}{\lambda_n})x_k + x \in X_{n-1}$$

Taking into account that

$$\frac{1}{\lambda_{n-1}}Tx_{n-1}^o \in X_{n-1},$$

and inserting it in the above formula $(*)$, it results that

$$\|\frac{1}{\lambda_n}Tx_n^o - \frac{1}{\lambda_{n-1}}Tx_{n-1}^o\| \geq 1,$$

and as, for $m < n-1$, $X_m \subsetneq X_{n-1}$, it follows

$$\|\frac{1}{\lambda_n}Tx_n^o - \frac{1}{\lambda_m}Tx_m^o\| \geq 1,$$

which shows that $(T(1/\lambda_n)x_n^o)_n$ has no a convergent subsequence (contradiction, because of T is compact).

Thus, we have proven that, if there exists an accumulation point of the point spectrum, it has to be only zero. Further, we notice that

$$\sigma_p(T) \subset \sigma(T) \subset \overline{B}(0, \|T\|) =$$

$$= \bigcup_{n=1}^{\infty}\{\lambda \in \mathbb{K} \mid \frac{1}{n+1}\|T\| \leq |\lambda| \leq \frac{1}{n}\|T\|\}\bigcup\{0\},$$

thus,

$$\sigma_p(T) \subset \bigcup_{n=1}^{\infty} \{\lambda \in \sigma_p(T) \mid \frac{1}{n+1}\|T\| \le |\lambda| \le \frac{1}{n}\|T\|\} \bigcup \{0\}$$

Each set

$$\{\lambda \in \sigma_p(T) \mid \frac{1}{n+1}\|T\| \le |\lambda| \le \frac{1}{n}\|T\|\},$$

$(n \in \mathbb{N})$, can be only finite (otherwise, since it is infinite and bounded, it follows that it has an accumulation point which, necessarily is not zero; one contradicts what we have proven). We conclude, that $\sigma_p(T)$ is (at most) countable, as a subset of an (at most) countable set.

Proposition 6.4.2 *Let T be a compact self-adjoint operator. Then, each $\lambda \in \sigma(T)$, $\lambda \neq 0$ is in $\sigma_p(T)$, thus*

$$\sigma_p(T) \subset \sigma(T) \subset \sigma_p(T) \cup \{0\}.$$

Proof. Given $\lambda \in \sigma(T)$, $\lambda \neq 0$ there exists a sequence $(x_n)_n$, $\|x_n\| = 1$ such that $\lambda x_n - Tx_n \longrightarrow 0$. Since T is compact, there exists a subsequence $(x_{n'})_{n'}$ of $(x_n)_n$ such that $(Tx_{n'})_{n'}$ is convergent. Then,

$$x_{n'} = \frac{1}{\lambda}(\lambda x_{n'} - Tx_{n'}) + \frac{1}{\lambda}Tx_{n'}$$

converges. Let x be its limit. As $\|x_{n'}\| = 1$, clearly, $x \neq 0$, and since

$$\lambda x_{n'} - Tx_{n'} \longrightarrow 0,$$

it follows that $\lambda x = Tx$, therefore $\lambda \in \sigma_p(T)$.

Remark. If X has infinite dimension and T is compact, then $0 \in \sigma(T)$. By the previous proposition it follows that $\sigma(T) = \sigma_p(T) \cup \{0\}$.

Corollary 6.4.1 *For every compact self-adjoint operator $T \neq 0$, the point spectrum $\sigma_p(T)$ is nonempty. Moreover, there exists $\lambda \in \sigma_p(T)$, $\lambda \neq 0$.*

Proof. Since $T \neq 0$, $\|T\| = \max(|m_T|, |M_T|) \neq 0$, therefore at least one of m_T, M_T is not zero. By Proposition 6.3.4, m_T, $M_T \in \sigma(T)$, thus by the Proposition 6.4.2, there exists $\lambda \in \sigma_p(T)$, $\lambda \neq 0$.

Proposition 6.4.3 *If T is compact, for every $\lambda \neq 0$, $\lambda \in \sigma_p(T)$, the subspace Ker $(\lambda I - T)$ has finite dimension.*

Proof. Suppose $\lambda \neq 0$, thus

$$\operatorname{Ker}(\lambda I - T) = \operatorname{Ker}\,(I - \frac{1}{\lambda}T)$$

Let us denote by $T_1 = (1/\lambda)T$ and by N_1 the closed subspace Ker $(I - T_1)$. Clearly, T_1 is compact and $T_1(N_1) \subset N_1$. If $\overline{B}_1$ is the unit ball of N_1, $T_1(\overline{B}_1) = \overline{B}_1$. Since $\overline{B}_1$ is bounded in X, and T_1 is compact, $T_1(\overline{B}_1)$ is relative compact in X, so $\overline{B}_1$ is compact in N_1. By Corollary 2.3.3, as the closed unit ball of N_1 is compact, it results that N_1 is finite dimensional.

Definition. For each $\lambda \in \sigma_p(T)$, the dimension of the corresponding subspace of eigenvectors, Ker $(\lambda I - T)$ is called the *multiplicity* of the eigenvalue λ.

We will state now one of the main theorems of this section, which gives a complete description of the spectrum of a compact self-adjoint operator on a Hilbert space.

Theorem 6.4.1 *(The Riesz–Schauder theorem) Let T be a compact self-adjoint operator on the Hilbert space X. Then, the spectrum of T, $\sigma(T)$ is a countable set having no accumulation points except perhaps $\lambda = 0$. Further, any nonzero $\lambda \in \sigma(T)$ is an eigenvalue of finite multiplicity.*

Proof. All follows combining the results of Propositions 6.4.1, 6.4.2 and 6.4.3.

The next theorem is known as *the Fredholm alternative.*

Theorem 6.4.2 *Let T be a compact self-adjoint operator on the Hilbert space X and $\lambda \in \mathbb{K} \setminus \{0\}$. Then*
1) If $\lambda \notin \sigma_p(T)$, the equation $(\lambda I - T)x = z$ has a unique solution, for any $z \in X$.
2) If $\lambda \in \sigma_p(T)$, the equation $(\lambda I - T)x = z$ has solutions if and only if $z \in (\operatorname{Ker}(\lambda I - T))^{\perp}$.

Proof. 1) Suppose $\lambda \neq 0$, $\lambda \notin \sigma_p(T)$. Then, by Proposition 6.4.2, $\lambda \notin$ $\notin \sigma(T)$, thus, $\lambda I - T$ is surjective and injective, therefore for every $z \in X$, there exists a unique $x \in X$ such that $(\lambda I - T)x = z$.

2) Let $\lambda \in \sigma_p(T)$. Clearly, $T(\text{Ker }(\lambda I - T)) \subset \text{Ker }(\lambda I - T)$, and since T is self-adjoint, $T(\text{Ker }(\lambda I - T)^\perp) \subset \text{Ker }(\lambda I - T)^\perp$. Let us consider the restriction of T to the subspace $\text{Ker }(\lambda I - T)^\perp$,

$$T_1 : \text{Ker }(\lambda I - T)^\perp \to \text{Ker }(\lambda I - T)^\perp$$

If $\lambda \in \sigma(T_1)$, $\exists x_o \neq 0$, $x_o \in \text{Ker }(\lambda I - T)^\perp$ such that $Tx_o = \lambda x_o$, thus $x_o \in \text{Ker }(\lambda I - T)$ (contradiction). Hence $\lambda \notin \sigma(T_1)$, or equivalently, $\lambda I - T_1$ is invertible on $\text{Ker }(\lambda I - T)^\perp$, which proves that for each $z \in (\text{Ker }(\lambda I - T))^\perp$ there exists $x \in (\text{Ker }(\lambda I - T))^\perp$ such that $(\lambda I - T)x = z$.

Conversely, if the equation $(\lambda I - T)x = z$ has a solution $x \in X$, then we can write $x = x_1 + x_2$, with $x_1 \in \text{Ker }(\lambda I - T)$, $x_2 \in \text{Ker }(\lambda I - T)^\perp$. So,

$$(\lambda I - T)(x_1 + x_2) = (\lambda I - T)(x_2) \in \text{Ker }(\lambda I - T)^\perp.$$

Notation. Next, T is a compact self-adjoint operator on the Hilbert space X. Then, by the Riesz-Schauder theorem, the set $\sigma_p(T)\setminus\{0\}$ is a nonempty countable set. Therefore, we may consider this set as a sequence $(\lambda_n)_n$. We suppose that in this sequence every eigenvalue repeats itself as many times as its multiplicity (the dimension of the corresponding subspace of eigenvectors) is. In every subspace of eigenvectors we choose an orthonormal basis. Since for $\lambda_n \neq \lambda_m$, $\text{Ker }(\lambda_n I - T) \perp \text{Ker }(\lambda_m I - T)$, it follows that, considering the set of all orthonormal bases of the spaces of corresponding eigenvectors to the sequence of eigenvalues $(\lambda_n)_n$, we obtain an orthonormal sequence of eigenvectors, $(x_n)_n$, $x_n \in \text{Ker }(\lambda_n I - T)$, $\forall n$.

With these notations, we have the next theorem concerning the spectral representation of compact self-adjoint operators.

Theorem 6.4.3 *(The Hilbert-Schmidt theorem) For each $x \in X$,*

$$Tx = \sum_{n=1}^{\infty} \lambda_n < x, x_n > x_n$$

Proof. Let us denote by Y the linear space spanned by $(x_n)_n$. We prove that

$$\text{Ker } T = Y^\perp$$

The inclusion $\text{Ker } T \subset Y^\perp$ is clear whenever T is injective; if it is not, $0 \in \sigma_p(T)$, and, since $\lambda_n \neq 0$, $\forall n$, it results that

$$\text{Ker } T \perp \text{Ker }(\lambda_n I - T), \qquad \forall n,$$

thus, $\forall x \in \text{Ker } T$ is orthogonal to Y, so $x \in Y^\perp$.

For the converse, we note first that, because $(x_n)_n$ are all eigenvectors, $T(Y) \subset Y$, thus, T being self-adjoint, $T(Y^\perp) \subset Y^\perp$. Consider the restriction of T to $Y^\perp$, $T_1 : Y^\perp \to Y^\perp$ and suppose that $T_1 \neq 0$. T_1 is a compact self-adjoint operator on the Hilbert space $Y^\perp$, therefore, by Corollary 6.4.1, it has an eigenvalue $\lambda \neq 0$. It follows that there exists $x_\lambda \in Y^\perp$, $x_\lambda \neq 0$ such that $Tx_\lambda = \lambda x_\lambda$. Then, $x_\lambda \in Y \cap Y^\perp$, therefore $Y \cap Y^\perp \neq \{0\}$ (contradiction). Consequently, $T_1 = 0$, that means, the restriction of T to $Y^\perp$ is zero, or equivalently, $Y^\perp \subset \text{Ker } T$.

Hence, every $x \in X$ can be written (uniquely) as $x = x_o + y$, where $x_o \in \text{Ker } T = Y^\perp$ and $y \in \overline{Y}$. In addition, since $(x_n)_n$ is an orthonormal set in $\overline{Y}$ such that the closure of the linear space spanned by $(x_n)_n$ coincides to $\overline{Y}$, it follows that $(x_n)_n$ is an orthonormal basis for the Hilbert space $\overline{Y}$, therefore

$$y = \sum_{n=1}^{\infty} < y, x_n > x_n = \sum_{n=1}^{\infty} < x, x_n > x_n$$

Then,

$$Tx = T(x_o + \sum_{n=1}^{\infty} < x, x_n > x_n) =$$

$$= Tx_o + \sum_{n=1}^{\infty} < x, x_n > Tx_n = \sum_{n=1}^{\infty} \lambda_n < x, x_n > x_n$$

Remark. The numerical sequence $(\lambda_n)_n$ converges to zero. Indeed, otherwise, there exists a subsequence $\lambda_{n'} \longrightarrow \lambda_o \neq 0$. Since $(x_{n'})_{n'}$ is bounded, there exists a subsequence $(x_{n''})_{n''}$ such that $T(x_{n''}) = \lambda_{n''} x_{n''}$ converges. It results finally that $(x_{n''})_{n''}$ is convergent (contradiction, since $\|x_n - x_m\| = = \sqrt{2}$).

Remark. The previous theorem states that each compact self-adjoint operator on a Hilbert space is diagonalizable (see section 5.4). In the particular case of finite dimensional spaces, it results that every hermitian matrix is diagonalizable.

6.5 Exercises

1. Let $\mathcal{A}$ be a unital normed algebra and I its unity. Show that:

a) $\|I\| \geq 1$;

b) There exists a norm $\|\cdot\|'$ on $\mathcal{A}$ such that $\|\cdot\|' \sim \|\cdot\|$, $\|I\|' = 1$ and $(\mathcal{A}, \|\cdot\|')$ is a unital normed algebra.

2. Let $\mathcal{A}$ be a normed algebra (over the field $\mathbb{K}$). Show that $\mathcal{A} \times \mathbb{K}$ can be organized as a unital normed algebra such that the map $T \longmapsto (T, 0)$ be an algebra homeomorphism from $\mathcal{A}$ to $\mathcal{A} \times \mathbb{K}$.

3. Let $\mathcal{A}$ be a normed algebra and $(A_n)_n$, $(B_n)_n$ Cauchy sequences. Show that $(A_n B_n)_n$ is a Cauchy sequence.

4. Let T be a locally compact space and $C_\infty(T) = \{x : T \to \mathbb{K} \mid x$ continuous, and $\forall \varepsilon > 0$, $\exists Q(\varepsilon)$ compact such that $|x(t)| < \varepsilon$, $\forall t \notin Q(\varepsilon)\}$ with the usual norm, $\|x\| = \sup\{|x(t)| \mid t \in T\}$. Define the multiplication by $(xy)(t) = x(t)y(t)$ and an involution by $x \longmapsto x^*$, $x^*(t) = \overline{x(t)}$. Show that $C_\infty(T)$ is a commutative non-unital Banach $\mathbb{C}^*$-algebra.

5. Define on $L^1_{\mathbb{C}}(\mathbb{R})$ a multiplication by

$$(x * y)(t) = \int_{\mathbb{R}} x(t-s)y(s)\,\mathrm{ds}$$

Prove that $(L^1_{\mathbb{C}}(\mathbb{R}), *)$ is a commutative non-unital Banach algebra.

6. Let $\mathcal{A}$ be a unital normed algebra.

a) Show that for every $S, T \in \mathcal{A}$,

$$\sigma(ST) \cup \{0\} = \sigma(TS) \cup \{0\}$$

b) Consider in the algebra $\mathcal{B}(l^2_{\mathbb{K}})$, $T(\xi_n)_n = (\xi_n)_{n \geq 2}$. Show that $\sigma(TT^*) \neq \sigma(T^*T)$ (therefore, generally $\sigma(ST) \neq \sigma(TS)$)

7. Let $\mathcal{A}$ be a unital normed algebra and $P \in \mathcal{A}$, $P \notin \{0, I\}$, $P^2 = P$. Show that $\sigma(P) = \{0, 1\}$.

8. Let $\mathcal{A}$ be a unital Banach algebra. Show that, for every $T \in \mathcal{A}$.

$$\|T\|_\sigma = \inf\left\{t > 0 \mid \left(\left(\frac{x}{t}\right)^n\right)_n \text{ is a bounded sequence}\right\}$$

9. Let $T \in \mathcal{B}(l^2_{\mathbb{K}})$ be the operator defined by $T(\xi_n)_n = (\xi_n)_{n \geq 2}$.

a) Find $\sigma(T)$, $\sigma_p(T)$, $\sigma(T^*)$, $\sigma_p(T^*)$.

b) Is the operator T^* compact?

10. Define on the space $L^2_{\mathbb{K}}[-1, 1]$, the operator $x \longmapsto Tx$, where

$$Tx(t) = tx(t), \quad \forall t \in [-1, 1]$$

a) Show that $T \in \mathcal{A}(L^2_{\mathbb{K}}[-1, 1])$ and find m_T, M_T, $\|T\|$;

b) Find $\sigma(A)$, $\sigma_p(A)$;

c) Show that $T(L^2_{\mathbb{K}}[-1,1])$ is not a closed subspace of $L^2_{\mathbb{K}}[-1,1]$;

d) Prove that T is not compact.

11. Define on the space $l^2_{\mathbb{K}}$, the operator T, by

$$T(\xi_n)_{n\geq 1} = (\frac{2}{1}\xi_2, \frac{3}{2}\xi_3, ..., \frac{n+1}{n}\xi_{n+1}, ...)$$

a) Show that $T \in \mathcal{B}(l^2_{\mathbb{K}})$ and find $\|T\|$;

b) Show that $\sigma_p(T) = \{\lambda \in \mathbb{K} | \ |\lambda| \leq 1\}$;

c) Find $\|T\|_\sigma$;

d) Find $\sigma(T)$;

e) Is the operator T normal?

f) Is the operator T compact?

12. Let X be a separable infinite Hilbert space and $\{e_n \mid n \in \mathbb{N}\}$ an orthonormal basis. For a bounded numerical sequence $(\lambda_n)_n$ let T be the bounded linear operator defined by $Te_n = \lambda_n e_n, \forall n$. Show that $\sigma_p(T) = \{\lambda_n \mid n \in \mathbb{N}\}$ and $\sigma(T) = \overline{\{\lambda_n \mid n \in \mathbb{N}\}}$.

13. Let U be the integral operator on $L^2_{\mathbb{K}}[0,1]$ defined by

$$Ux(s) = \int_0^1 stx(t)\,\mathrm{ds}$$

Find $\sigma(U)$, $\sigma_p(U)$.

14. Let T be the operator on $\mathcal{C}_{\mathbb{K}}([0,1])$, defined by

$$Tx(s) = \int_0^s x(t)\,\mathrm{dt}$$

a) Show that $T \in \mathcal{B}(\mathcal{C}_{\mathbb{K}}[0,1])$ and find $\|T\|$;

b) Find $\|T\|_\sigma$;

c) Find $\sigma_p(T)$, $\sigma(T)$.

15. Let X be a separable infinite Hilbert space and $D \subset X$ a bounded open set. Show that there exists $T \in \mathcal{B}(X)$ such that $\sigma_p(T) = D$ and $\sigma(T) = \overline{D}$.

Chapter 7

Locally convex spaces

7.1 Topological vector spaces

Definition. A vector space X over the field $\mathbb{K}$, ($\mathbb{K} = \mathbb{R}$, $\mathbb{C}$) equipped with a topology τ such that:

i) the mapping $(x, y) \longmapsto x + y$ from $X \times X$ to X is continuous (with respect to the product topology on $X \times X$);

ii) the mapping $(\alpha, x) \longmapsto \alpha x$ from $\mathbb{K} \times X$ to X is continuous (with respect to the product topology on $\mathbb{K} \times X$);

is called a *topological vector space.*

The topology τ is said to be *compatible* with the structure of vector space X.

Notation. Both, the origin of X and the number zero in $\mathbb{K}$ will be denoted by the same symbol, 0, since, by the context, no confusion could be made.

Remarks. **1.** *If X is a topological vector space, then, clearly, for every $x_o \in X$ and $\alpha \in \mathbb{K}$, $\alpha \neq 0$ the mappings $x \longmapsto x + x_o$ and $x \longmapsto \alpha x$ are homeomorphisms (of X onto X). Consequently, if V is a neighbourhood of the origin, then the sets $V + x_o = \{x + x_o \mid x \in V\}$ and $\alpha V = \{\alpha x \mid x \in V\}$ are also neighbourhoods of the origin.*

2. *By the Remark 1, it follows whenever $\mathcal{B}$ is a fundamental system of neighbourhoods of the origin in X, setting $\mathcal{B}_{x_o} = \{B + x_o \mid B \in \mathcal{B}\}$ we get a fundamental system of neighbourhoods for x_o.*

3. *Take $x_o \in X$, $\alpha_o \in \mathbb{K}$, arbitrary fixed. By*

$$\alpha x - \alpha_o x_o = (\alpha - \alpha_o)(x - x_o) + (\alpha - \alpha_o)x_o + \alpha_o(x - x_o)$$

it follows that ii) in the above definition holds if and only if i) holds and the mappings $(\alpha, x) \longmapsto \alpha x$ (from $\mathbb{K} \times X$ to X), $\alpha \longmapsto \alpha x_o$ (from $\mathbb{K}$ to X), $x \longmapsto \alpha_o x$ (from X to X) are continuous at $(0,0)$, respectively at 0.

Examples. 1. Each vector space X endowed with the indiscrete topology, $\tau = \{\emptyset, X\}$ is a topological vector space.

2. Every normed space is a topological vector space.

3. If $U : X \longrightarrow Y$ is a linear operator from the vector space X to the topological vector space (Y, σ), then $\tau = \{U^{-1}(D) \mid D \in \sigma\}$ is a topology compatible with the structure of the vector space X, thus X is a topological vector space.

4. If $\{X_j, \tau_j\}_{j \in J}$ is a family of topological vector spaces, then

$$X = \prod_{j \in J} X_j$$

with the product topology, $\tau = \prod_{j \in J} \tau_j$ is a topological vector space.

In order to prove that, consider first the spaces

$$X \times X \text{ and } Z = \prod_{j \in J} (X_j \times X_j)$$

(with the product topology) and notice that the mapping

$$g : X \times X \longrightarrow Z, \quad g(\{x_j\}_{j \in J}, \{y_j\}_{j \in J}) = \{(x_j, y_j)\}_{j \in J}$$

is a homeomorphism. On the other hand, for each $j \in J$, the mapping

$$h_j : X_j \times X_j \longrightarrow X_j, \quad h_j(x_j, y_j) = x_j + y_j$$

is continuous, so also

$$h : Z \longrightarrow X, \quad h(\{(x_j, y_j)\}_{j \in J}) = \{x_j + y_j\}_{j \in J}$$

is continuous. It follows that the map $h \circ g$ from $X \times X$ to X, i.e. $(x, y) \longmapsto$ $\longmapsto x + y$ is continuous. Similarly, one can show that the application $(\alpha, x) \longmapsto \alpha x$ from $\mathbb{K} \times X$ to X is continuous.

5. If X is a vector space, $X \neq \{0\}$, then X with the discrete topology, is not a topological vector space. Indeed, let us suppose that the application $\lambda \longmapsto \lambda x_o$ (where $x_o \in X$, $x_o \neq 0$) is continuous at zero. It follows that for each neighbourhood of zero in X, in particular for $V = \{0\}$, there exists

$\delta_V > 0$, such that $\forall\lambda$, $|\lambda| < \delta_V$, $\lambda x_o \in V$ (contradiction, since for $\lambda \neq 0$, $\lambda x_o \neq 0$, thus $\lambda x_o \notin V = \{0\}$).

Remark. Let Y be a subspace of a topological vector space X. Then, its closure is also a linear subspace. This follows from the continuity of the mappings $f : X \times X \longrightarrow X$; $f(x,y) = x + y$ and $g : \mathbb{K} \times X \longrightarrow X$, $g(\alpha, x) = \alpha x$, thus $f(\overline{Y}, \overline{Y}) \subset \overline{f(Y,Y)}$ and $g(\mathbb{K}, \overline{Y}) \subset \overline{g(\mathbb{K}, Y)}$.

Definition. If X is a (topological) vector space, a subset A of X is said to be *balanced* if and only if $\alpha A \subset A$, $\forall \alpha \in \mathbb{K}$, $|\alpha| \leq 1$. A subset A of X is said to be an *absorbing set* if, for any $x \in X$, there exists $\alpha_x > 0$, such that $\forall \alpha \in \mathbb{K}$, $0 < |\alpha| \leq \alpha_x$, A contains αx.

Remarks. 1. If A is balanced, then A is an absorbing set $\iff \forall x \in X$, there exists $\alpha_x > 0$, such that $\alpha_x x \in A$. This is clear taking into account that, if $\alpha \in \mathbb{K}$, $0 < |\alpha| \leq \alpha_x$, then $|\alpha/\alpha_x| \leq 1$, therefore

$$\alpha x = \frac{\alpha}{\alpha_x} \cdot \alpha_x \cdot x \in \frac{\alpha}{\alpha_x} A \subset A$$

2. If A is a balanced set, then, for every λ, $\mu \in \mathbb{K}$, $|\lambda| < |\mu|$, $\lambda A \subset \mu A$. In particular, if $|\alpha| = 1$, $\alpha A = A$. It is also immediate, by the continuity of the mapping $g : \mathbb{K} \times X \longrightarrow X$, $g(\alpha, x) = \alpha x$, that $\overline{A}$ is also balanced.

Examples. 1. In a normed space the closed unit ball $\overline{B}(0,1)$ is a balanced absorbing set.

2. In a topological vector space, every neighbourhood of the origin is an absorbing set.

3. In the normed space $l^2_{\mathbb{K}}$, the set

$$A = \{(\xi_n)_n \in l^2_{\mathbb{K}} \mid \sum_{n=1}^{\infty} n|\xi_n|^2 \leq 1\}$$

is balanced, but it is not an absorbing set.

Proposition 7.1.1 *Let A be a set in the topological space X. Then*

$$\overline{A} = \bigcap_{W \in \mathcal{B}} (A + W),$$

(for every $\mathcal{B}$ a fundamental system of neighbourhoods of the origin in X).

Proof. Let $\mathcal{B}$ be a fundamental system of neighbourhoods of the origin in X. Then, $\{x - W\}_{W\in\mathcal{B}}$ is a fundamental system of neighbourhoods of x (x arbitrary in X). Then,

$$x \in \overline{A} \Longleftrightarrow A \cap (x - W) \neq \emptyset, \quad \forall W \in \mathcal{B} \Longleftrightarrow$$

$$\Longleftrightarrow \exists y_W \in W, x \in A + y_W \Longleftrightarrow x \in \bigcap_{W\in\mathcal{B}} (A + W)$$

Proposition 7.1.2 *Let X be a topological vector space. Then, X is Hausdorff if and only if $\forall x \neq 0$, $\exists V \in \mathcal{V}_0$ such that $x \notin V$.*

Proof. X is Hausdorff if and only if

$$\Delta_X = \{(x, y) \in X \times X \mid x = y\}$$

is closed in $X \times X$ (with respect to the product topology). As the mapping

$$h : X \times X \longrightarrow X, \quad h(x, y) = x - y$$

is continuous, in order to prove that Δ_X is closed we have to see that the set $\{0\}$ is closed in X. Let x be in the closure of $\{0\}$ and suppose that $x \neq 0$, so consequently, $-x \neq 0$, too. Then, by assumption, there exists $V \in \mathcal{V}_0$ such that $-x \notin V$. It follows that $\{0\} \cap (x + V) = \emptyset$, so $x \notin \overline{\{0\}}$ (contradiction). We conclude that the set $\{0\}$ coincides to its closure, thus $\{0\}$ is closed.

The converse is obvious.

Proposition 7.1.3 *Let X be a topological vector space. Then, there exists $\mathcal{B}$ a fundamental system of neighbourhoods of the origin with the following properties:*
1) Every $B \in \mathcal{B}$ is balanced;
2) Every $B \in \mathcal{B}$ is closed;
3) For every $B \in \mathcal{B}$, there exists $B_1 \in \mathcal{B}$ such that $B_1 + B_1 \subset B$;
4) For every $B \in \mathcal{B}$ and $\alpha \in \mathbb{K}$, $\alpha \neq 0$, $\alpha B \in \mathcal{B}$.

Proof. First we notice that for each $V \in \mathcal{V}_0$, there exists $W \in \mathcal{V}_0$, W balanced, $W \subset V$. Indeed, take an arbitrary $V \in \mathcal{V}_0$. By the continuity of the mapping $(\alpha, x) \longmapsto \alpha x$ at $(0, 0)$, it follows that $\exists \theta > 0$ and $\exists W_1 \in \mathcal{V}_0$ such that for every $\alpha \in \mathbb{K}$, $|\alpha| \leq \theta$ and $x \in W_1$, $\alpha x \in V$. Then, setting $W = \bigcup_{|\alpha|\leq\theta} \alpha W_1$, clearly $W \in \mathcal{V}_0$, W balanced and $W \subset V$.

Next we will prove that each $V \in \mathcal{V}_0$ contains a balanced closed neighbourhood of zero. By the continuity of the addition, there exists $W \in \mathcal{V}_0$ such that $W + W \subset V$. We saw that W can be supposed balanced. We focus on showing that $\overline{W} \subset V$. So, let $z \in \overline{W}$ be. As the mapping $(x, y) \longmapsto x - y$ is continuous at (z, z) and $W \in \mathcal{V}_0$, there exists $U \in \mathcal{V}_z$ such that $U - U \subset W$. On the other hand, since $z \in \overline{W}$, $U \cap W \neq \emptyset$. Set $z_o \in U \cap W$. Then,

$$z = (z - z_o) + z_o \in (U - U) + W \subset W + W \subset V$$

Setting $\mathcal{B} = \{W \in \mathcal{V}_0 | \ W \text{ balanced}, W \text{ closed}\}$, everything is clear.

Theorem 7.1.1 *Let X be a (nonempty) vector space and $\mathcal{B} \subset \mathcal{P}(X)$ with the following properties:*
1) Every $B \in \mathcal{B}$ is an absorbing balanced set;
2) For any B_1, $B_2 \in \mathcal{B}$, there exists $B \in \mathcal{B}$ such that $B \subset B_1 \cap B_2$;
3) For any $B \in \mathcal{B}$, there exists $B_1 \in \mathcal{B}$, such that $B_1 + B_1 \subset B$;
Then, there exists τ, a topology on X such that, endowed with this topology X is a topological vector space and $\mathcal{B}$ is a fundamental system of neighbourhoods of the origin for τ.

Proof. Let us set

$$\mathcal{V} = \{V \subset X \mid \exists B \in \mathcal{B}, B \subset V\}$$

and for any $x \in X$,

$$\mathcal{V}_x = \{x + V \mid V \in \mathcal{V}\}$$

We need to prove first that $\forall x \in X$, $\mathcal{V}_x$ has the properties of the system of neighbourhoods of a point in an arbitrary topology:

V1) $x \in U, \forall U \in \mathcal{V}_x$;

V2) If $U \in \mathcal{V}_x$ and $U \subset V$, then $V \in \mathcal{V}_x$;

V3) For every U, $V \in \mathcal{V}_x$, $U \cap V \in \mathcal{V}_x$;

V4) If $U \in \mathcal{V}_x$, $\exists V \in \mathcal{V}_x$ such that $U \in \mathcal{V}_y$, $\forall y \in V$.

To show V1), let $U \in \mathcal{V}_x$. Then, $\exists V \in \mathcal{V}$ such that $U = V + x$, so $\exists B \in \mathcal{B}$, B balanced such that $B + x \subset U$. As each balanced set contains the origin, it follows that $x \in U$.

The second property V2) easily results by the definition of $\mathcal{V}$: If $U \in \mathcal{V}_x$ and $U \subset V$, there is $B \in \mathcal{B}$ such that $B + x \subset U \subset V$, hence $V \in \mathcal{V}_x$. Taking into account the property 2) of $\mathcal{B}$ and the definitions of $\mathcal{V}_x$, V3) follows immediately. For V4), let $U \in \mathcal{V}_x$, and $B \in \mathcal{B}$ such that $x + B \subset U$. Then, (3)), $\exists B_1 \in \mathcal{B}$, $B_1 + B_1 \in \mathcal{B}$. Setting $V = B_1 + x$, we have for each $y \in V$ that $U \in \mathcal{V}_y$, because

$$B_1 + y \subset B_1 + B_1 + x \subset B + x \subset U$$

Then, by a well known theorem (see Preliminaries), there exists a topology τ on X such that, $\forall x \in X$, $\mathcal{V}_x = \mathcal{V}_x^\tau$. Further, we will prove that this topology is compatible with the structure of vector space X. It is almost evident, by 3) that the map $(x, y) \longmapsto x + y$ is continuous at each (x_o, y_o). We will use the Remark 3 to check the continuity of the mapping $(\alpha, x) \longmapsto \alpha x$ (from $\mathbb{K} \times X$ to X). Because each B is an absorbing balanced set, obviously, the mappings $(\alpha, x) \longmapsto \alpha x$ (from $\mathbb{K} \times X$ to X) and $\alpha \longmapsto \alpha x_o$ (from $\mathbb{K}$ to X) are continuous at $(0,0)$ and, respectively at 0 (where x_o is arbitrary in X). The only thing to check is that the mapping $x \longmapsto \alpha_o x$ (from X to X) is continuous at 0 (where α_o is arbitrary in $\mathbb{K}$). Take $V \in \mathcal{V}$. By the property 3) of $\mathcal{B}$, we can find inductively, for each natural number n, $V_n \in \mathcal{V}$, such that $2^n V_n \subset V$. Take n_o natural such that $|\alpha_o| < 2^{n_o}$. Then, $\alpha_o V_{n_o} \subset V$, which ends the proof.

Definition. Let X be a topological vector space. A subset A of X is said to be *bounded* if for every $V \in \mathcal{V}_0$, there is $\lambda > 0$ such that $\lambda A \subset V$.

The definition involves the next properties regarding bounded sets:

Proposition 7.1.4 *Let X be a topological vector space. Then:*
1) If $A \subset X$, A bounded and $B \subset A$, then B is bounded;
2) If $A \subset X$, A bounded and $\alpha \in \mathbb{K}$, then αA is bounded;
3) If A, B are bounded subsets of X, then $A + B$ is bounded;
4) Every finite subset of X is bounded;
5) Each convergent sequence of X is a bounded set.

Proposition 7.1.5 *Let A be a subset of a topological vector space. Then, A is bounded if and only if for any sequence $(x_n)_n \subset A$ and any decreasing numerical sequence $\lambda_n \searrow 0$, the sequence $(\lambda_n x_n)_n$ converges to zero.*

Proof. Suppose that A is bounded and take arbitrary sequences $(x_n)_n \subset A$ and $\lambda_n \searrow 0$. Let V be a neighbourhood of the origin, which, without restricting the generality may be supposed balanced (Proposition 7.1.3). Then,

$\exists\lambda > 0$, such that $\lambda x_n \in V$, $\forall n$. Since $\lambda_n \searrow 0$, there exists a positive integer n_o such that $\lambda_n < \lambda$, $\forall n \geq n_o$. As V is balanced, it follows that

$$\lambda_n x_n = \frac{\lambda_n}{\lambda} \cdot \lambda x_n \in V, \quad \forall n \geq n_o$$

Conversely, supposing A is not bounded, it results that $A \not\subseteq nV$, for some $V \in \mathcal{V}_0$ and arbitrary $n \in \mathbb{N}$. Thus, we can pick $(x_n)_n \subset A$ and

$$(\frac{1}{n}\, x_n)_n \not\subseteq V$$

(contradiction).

Remark. By the previous proposition, it follows that an infinite subset A of X is bounded if and only if each countable subset of A is bounded.

7.2 Locally convex spaces

Let X be a vector space and $\{p_j\}_{j\in\mathcal{A}}$ be a family of seminorms on X. There is a standard method of defining a topology on X compatible with the structure of vector space X by means of this family, as follows. Let $\mathcal{F}(\mathcal{A})$ be the family of all finite subsets J of $\mathcal{A}$. For any $J \in \mathcal{F}(\mathcal{A})$ and $\varepsilon \in (0,\infty)$, let $W_{J,\varepsilon} \subset X$,

$$W_{J,\varepsilon} = \{x \in X \mid p_j(x) < \varepsilon, \quad \forall j \in J\}$$

and consider $\mathcal{W} = \{W_{J,\varepsilon} \mid J \in \mathcal{F}(\mathcal{A}), \varepsilon \in (0,\infty)\}$.

Proposition 7.2.1 *The family $\mathcal{W} \subset \mathcal{P}(X)$ has the following properties:*
1) Every $W \in \mathcal{W}$ is an absorbing balanced set;
2) For any W_1, $W_2 \in \mathcal{W}$, there exists $W \in \mathcal{W}$ such that $W \subset W_1 \cap W_2$;
3) For any $W \in \mathcal{W}$, there exists $W_1 \in \mathcal{W}$, such that $W_1 + W_1 \subset \mathcal{W}$;
4) Each $W \in \mathcal{W}$ is convex.

Proof. Let $W_{J,\varepsilon} \in \mathcal{W}$. If $\lambda \in \mathbb{K}$ and $|\lambda| < 1$, then for $x \in W_{J,\varepsilon}$,

$$p_j(\lambda x) = |\lambda| \cdot p_j(x) \leq p_j(x) < \varepsilon, \quad \forall j \in J,$$

thus $W_{J,\varepsilon}$ is balanced. Take an arbitrary $x \in X$. If $p_j(x) = 0$, $\forall j \in J$, then $\alpha x \in W_{J,\varepsilon}$ $(\forall\alpha \in \mathbb{K})$. If $\max_{j\in J} p_j(x) \neq 0$, then

$$\frac{\varepsilon}{\max_{j\in J} p_j(x)} x \in W_{J,\varepsilon},$$

so, $W_{J,\varepsilon}$ is an absorbing set. We have checked that 1) holds.

2) is proven by

$$W_{J_1 \cup J_2, \min(\varepsilon_1, \varepsilon_2)} \subset W_{J_1,\varepsilon_1} \cap W_{J_2,\varepsilon_2}$$

The third property is also clear since

$$W_{J,\varepsilon/2} + W_{J,\varepsilon/2} \subset W_{J,\varepsilon}$$

As any seminorm is positive homogeneous and subadditive, obviously each $W \in \mathcal{W}$ is convex.

Definition. A topological vector space X is said to be *locally convex* if there exists a fundamental system of convex neighbourhoods of the origin of X.

Theorem 7.2.1 *Let X be a vector space and $\{p_j\}_{j\in\mathcal{A}}$ be a family of seminorms on X. Then,*

1) There exists τ a topology on X compatible with the structure of vector space X such that $\mathcal{W}$ is a τ-fundamental system of neighbourhoods of the origin of X;

2) X endowed with the topology τ is a locally convex space;

3) The topology τ is the coarsest topology on X compatible with the structure of vector space X such that each p_j, $j \in \mathcal{A}$, is continuous on X for this topology;

4) The topology τ is Hausdorff if and only if $\{p_j\}_{j\in\mathcal{A}}$ satisfies the following separation condition: for each $x \in X$, $x \neq 0$, there is some $j \in \mathcal{A}$, such that $p_j(x) \neq 0$;

5) If $\{p_j\}_{j\in\mathcal{A}}$ is a directed family (i.e. $\forall p_{j_1}, p_{j_2}$ there exists $j \in \mathcal{A}$ such that $p_{j_1}, p_{j_2} \leq p_j$), then a fundamental system of neighbourhoods for the topology τ is $\mathcal{W} = \{W_{j,\varepsilon} \mid j \in \mathcal{A}, \varepsilon \in (0,\infty)\}$ where $W_{j,\varepsilon} = \{x \in X \mid p_j(x) < \varepsilon\}$.

Proof. 1) and 2) result directly from the previous proposition combined with Theorem 7.1.1. Let us consider an other topology on X compatible with its vector structure, $\tau' \subset \tau$ such that each p_j, $j \in \mathcal{A}$, is continuous on (X, τ'). As $p_j^{-1}(-\infty, \varepsilon) = W_{\{j\},\varepsilon}$, it results that $\tau \subset \tau'$, therefore τ and τ' coincide, so 3) holds. The fourth statement follows immediately by Proposition 7.1.2, taking into account that $p_j(x) = 0$, $\forall\, j \in \mathcal{A}$ involves $x \in \bigcap_{W\in\mathcal{W}} W$.

If $\{p_j\}_{j\in\mathcal{A}}$ is a directed family of seminorms, then for each $J \in \mathcal{F}(\mathcal{A})$, there exists $l \in \mathcal{A}$ such that $p_l \geq p_j, \forall j \in J$ (such l exists because the family of seminorms is directed). We infer that

$$W_{\{l\},\varepsilon} \subset W_{J,\varepsilon} \ ,$$

which ends the proof.

Remark. Let $\{p_j\}_{j\in\mathcal{A}}$ be a family of seminorms on X. Then the family of seminorms $\{q_J\}_{J\in\mathcal{F}(\mathcal{A})}$, where

$$q_J(x) = \max_{j\in J} p_j(x)$$

is a directed family of seminorms which evidently defines the same topology as $\{p_j\}_{j\in\mathcal{A}}$. Hence, there is no loss in generality assuming that the family of seminorms which defines the locally convex topology in the above theorem is directed.

Further, we will prove that the topology of any locally convex space is defined by a family of seminorms.

Remarks. 1. Let X be a locally convex space. Then, each neighbourhood of the origin, contains a balanced convex neighbourhood. Indeed, let V be a neighbourhood of zero and W' a convex neighbourhood of zero, $W' \subset V$. Setting

$$W = \bigcap_{|\lambda|\geq 1} \lambda W'$$

we have that W is balanced. Indeed, take x in W, so $x \in \lambda W'$, $\forall\lambda$, $|\lambda| \geq 1$ and μ such that $|\mu| \leq 1$. Then, as $|\lambda/\mu| \geq 1$, it follows that $x \in \lambda/\mu W'$, $\forall\lambda$, $|\lambda| \geq 1$; that means $\mu x \in \lambda W'$, $\forall\lambda$, $|\lambda| \geq 1$. We conclude that $\mu W \subset W$, $\forall\mu$, $|\mu| \leq 1$.

2. *Let X be a locally convex space and $\mathcal{W}$ a fundamental system of balanced convex neighbourhoods of the origin. For every $W \in \mathcal{W}$, the gauge function of W is the mapping defined on X by*

$$p_W(x) = \inf\{\mu > 0 \mid x \in \mu W\}, \quad x \in X.$$

Notice that since W is an absorbing set, for every $x \in X$, there exists $\alpha_x > 0$ such that

$$(\frac{1}{\alpha_x}, +\infty) \subset \{\mu > 0 \mid x \in \mu W\},$$

thus p_W is a well defined ($\{\mu > 0 \mid x \in \mu W\} \neq \emptyset$) mapping from X to $\mathbb{R}$.

The next lemma shows that whenever W is convex, p_W is subadditive, and if in addition W is balanced, p_W is a seminorm.

Lemma 7.2.1 *Let X be a locally convex space and W a balanced convex neighbourhood of the origin. Then, the gauge function of W, p_W is a seminorm.*

Proof. Let $x, y \in X$ and $\varepsilon > 0$, ε arbitrary. Then, $x \in (p_W(x) + \varepsilon)W$, $y \in (p_W(y) + \varepsilon)W$, and accordingly,

$$\frac{1}{p_W(x) + \varepsilon} x \in W \text{ and } \frac{1}{p_W(y) + \varepsilon} y \in W$$

Since W is convex and

$$\lambda = \frac{p_W(x) + \varepsilon}{p_W(x) + p_W(y) + 2\varepsilon} < 1$$

we have that

$$\lambda \frac{1}{p_W(x) + \varepsilon} x + (1 - \lambda) \frac{1}{p_W(y) + \varepsilon} y \in W,$$

thus $x + y \in (p_W(x) + p_W(y) + 2\varepsilon)W$. It follows that

$$p_W(x + y) \le p_W(x) + p_W(y) + 2\varepsilon$$

Because ε is arbitrary, it follows that p_W is subadditive.

Now, let $t > 0$. We have for any $x \in X$,

$$p_W(tx) = \inf\{\mu > 0 \mid tx \in \mu W\} = \inf\{\mu > 0 \mid x \in \frac{\mu}{t} W\} =$$

$$\underset{\nu = \mu \backslash t}{=} \inf\{t\nu > 0 \mid x \in \nu W\} = t\, p_W(x).$$

If β is complex and $|\beta| = 1$, since W is balanced, $\beta W = W$, therefore

$$p_W(\beta x) = \inf\{\mu > 0 \mid \beta x \in \mu W\} = \inf\{\mu > 0 \mid x \in \mu W\} = p_W(x)$$

Take now an arbitrary complex $\alpha \neq 0$. As $\alpha = |\alpha|\beta$, where $|\beta| = 1$, it follows that

$$p_W(\alpha x) = p_W(|\alpha|\beta x) = |\alpha|\, p_W(\beta x) = |\alpha|\, p_W(x)$$

Obviously, if $\alpha = 0$, $p_W(0x) = 0$, which ends the proof.

Lemma 7.2.2 *Let W be a balanced convex neighbourhood of the origin in the locally convex space X and p_W the gauge function of W. Then,*

$$\{x \in X \mid p_W(x) < 1\} \subset W \subset \{x \in X \mid p_W(x) \le 1\}$$

Proof. Let x be in X such that $p_W(x) < 1$. Then, there exists $\mu < 1$ such that $x \in \mu W$, and, since W is balanced, it follows that $x \in W$, thus the first inclusion is proven. The second is clear.

Theorem 7.2.2 *Let (X, τ) be a locally convex space and $\mathcal{W}$ a fundamental system of balanced convex neighbourhoods of the origin. Then, the topology of X, τ is defined by the family of seminorms $\{p_W\}_{W \in \mathcal{W}}$ (where p_W is the gauge function of W).*

Proof. For every $W \in \mathcal{W}$, we have by Lemma 7.2.2 that

$$\{x \in X \mid p_W(x) < 1\} \subset W,$$

therefore the topology defined by the family of seminorms $\{p_W\}_{W \in \mathcal{W}}$, is stronger than τ. Conversely, to show that τ is stronger than the topology defined by the family of seminorms $\{p_W\}_{W \in \mathcal{W}}$, consider

$$V = \{x \in X \mid p_{W_i}(x) < \varepsilon, \ i = 1, 2, ..., n\}.$$

Clearly, $(\varepsilon/2)W$, where

$$W = W_1 \cap W_2 .... \cap W_n$$

is a τ-neighborhood of the origin and, in addition it is contained in V. Indeed, if $x \in (\varepsilon/2)W$, then $(2/\varepsilon)x \in W$, which implies that

$$p_{W_i}\left(\frac{2}{\varepsilon}x\right) \leq 1, \quad \forall i = 1, 2, ..., n,$$

or, equivalently,

$$p_{W_i}(x) \leq \frac{2}{\varepsilon} < \varepsilon, \quad \forall i = 1, 2, ..., n.$$

Remark. The above theorem shows that the topology of any locally convex space is defined by a family of seminorms.

Theorem 7.2.3 *(Kolmogorov' theorem) Let (X, τ) be a Hausdorff topological vector space. Then, there exists a norm, $\|\cdot\|$ on X such that $\tau_{\|\cdot\|} = \tau$ if and only if there exists a bounded convex τ-neighbourhood of the origin of X.*

Proof. Clearly, if (X, τ) is a normed space, $\tau = \tau_{\|\cdot\|}$, then the unit ball $B(0,1) = \{x \in X \mid \|x\| < 1\}$ is a bounded convex τ-neighborhood of the origin of X.

Conversely, suppose that W is a bounded convex τ-neighborhood of the zero in X. We may consider that W is also balanced (otherwise, we can replace it by $\bigcap_{|\lambda| \geq 1} \lambda W$). Thus the gauge function of W is a seminorm. We will show that p_W is a norm. Let x be such that $p_W(x) = 0$. Then, $x \in \mu W$, $\forall \mu > 0$. Taking into account that W is bounded, for each τ- neighbourhood V of zero, $\mu W \subset V$ for some $\mu > 0$. It follows that x belongs to each τ-neighbourhood of the origin, and, as (X, τ) is Hausdorff, this involves that $x = 0$.

Now, we have only to show now that the topology defined by the norm p_W coincides to τ. Because of the inclusion

$$\varepsilon W \subset \{x \in X \mid p_W(x) \leq \varepsilon\}, \quad \forall \varepsilon > 0,$$

we infer that τ is weaker than the topology defined by the norm p_W . We claim that it is also stronger than this topology. Indeed, if V is a τ- neighbourhood of the origin, $\lambda W \subset V$ for some $\lambda > 0$. It follows that

$$\{x \in X \mid p_W(x) < \lambda\} \subset \{x \in X \mid p_W\left(\frac{x}{\lambda}\right) < 1\} \subset$$

$$\subset \lambda\{x \in X \mid p_W(x) < 1\} \subset \lambda W \subset V,$$

therefore

$$\{x \in X \mid p_W(x) < \lambda\} \subset V,$$

which ends the proof.

7.3 Weak topologies

If $(X, \|\cdot\|)$ is a normed space and X^* is its dual space, then the family of seminorms $\{p_f\}_{f \in X^*}$ on X, where $p_f(x) = |f(x)|$ for all $x \in X$, define a locally convex topology on X denoted by $\sigma(X, X^*)$ (or ω) and called the *weak topology* on X. Notice that, by Corollary 1.3.1 to the Hahn-Banach theorem, the topology $\sigma(X, X^*)$ is Hausdorff (the family of seminorms $\{p_f\}_{f \in X^*}$ satisfies the separation condition).

Theorem 7.3.1 *Let $(X, \|\cdot\|)$ be a normed space. Then, the weak topology on X, $\sigma(X, X^*)$ is weaker than the norm topology, $\tau_{\|\cdot\|}$. The two topologies coincide if and only if the space X is finite dimensional.*

Proof. Let $W(f_1, f_2, ..., f_n; \varepsilon)$ a $\sigma(X, X^*)$-neighbourhood topology of zero. Since $f_j \in X^*$, $j = 1, 2, ..., n$, we can pick $\delta_j > 0$ such that for every $x \in B(\delta_j)$, $|f_j(x)| < \varepsilon$, $j = 1, 2, ..., n$. Then, setting

$$\delta = \min_{1 \leq j \leq n} \delta_j,$$

clearly, $B(\delta)$ is contained in $W(f_1, f_2, ..., f_n; \varepsilon)$. Thus the weak topology on X, $\sigma(X, X^*)$ is weaker than the norm topology, $\tau_{\|\cdot\|}$.

If $\dim_{\mathbf{K}} X = n$, let $\{e_1, e_2, ..., e_n\}$ be an algebraic basis of X. We may suppose that $\|e_i\| = 1$, $i = 1, 2, ..., n$. Consider for every $i = 1, 2, ..., n$ the linear functional on X defined by

$$f_i(\sum_{j=1}^{n} \xi_j e_j) = \xi_i.$$

From

$$|f_i(\sum_{j=1}^{n} \xi_j e_j)| = |\xi_i| \leq \|(\xi_1, \xi_2, ..., \xi_n)\|_1,$$

since the norms are equivalent on X, it follows that $f_i \in X^*$, $i = 1, 2, ..., n$. Then, the $\sigma(X, X^*)$-neighborhood topology of zero $W(f_1, f_2, ..., f_n; \varepsilon/n)$ is contained in $B(\varepsilon)$ as it results by

$$\|x\| = \|\sum_{j=1}^{n} \xi_j e_j\| \leq \sum_{j=1}^{n} |\xi_j| < \varepsilon$$

We have proved that in the case of finite dimensional spaces, $\sigma(X, X^*)$ coincides to the norm topology, $\tau_{\|\cdot\|}$.

Next, suppose that the weak topology on X, $\sigma(X, X^*)$ coincides to the norm topology, $\tau_{\|\cdot\|}$. It follows that there exist $f_1, f_2, ..., f_n \in X^*$ and $\delta > 0$ such that

$$W(f_1, f_2, ..., f_n; \delta) = \{x \mid |f_j(x)| < \delta\} \subset \overline{B}(1)$$

Then, we can infer that

$$\bigcap_{j=1}^{n} \text{Ker } f_j = \{0\}$$

Indeed, let x be with $f_j(x) = 0$, $j = 1, 2, ..., n$; it results that $f_j(mx) = 0$, $\forall m \in \mathbb{N}$, $j = 1, 2, ..., n$, so $mx \in \overline{B}(1)$, $\forall m \in \mathbb{N}$, or equivalently $\|x\| \leq 1/m$, $\forall m \in \mathbb{N}$, i.e. $x = 0$.

Further, since $\bigcap_{j=1}^{n} \text{Ker } f_j = \{0\}$, we have that

$$\bigcap_{j=1}^{n} \text{Ker } f_j \subset \text{Ker } f, \quad \forall f \in X^*$$

Then, for every $f \in X^*$, we can find $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{K}$ such that

$$f(x) = \sum_{j=1}^{n} \alpha_j f_j(x)$$

as it results from the sequel. Define on X the $\mathbb{K}^n$-valued function, by

$$g(x) = (f_1(x), f_2(x), ..., f_n(x))$$

and set $X_o = g(X)$. Clearly, X_o is a linear subspace of $\mathbb{K}^n$. Next, we consider the $\mathbb{K}$-valued mapping h on X_o defined by

$$h(y_1, y_2, ..., y_n) = f(x),$$

where x is chosen in X such that

$$f_j(x) = y_j, \quad j = 1, 2, ..., n$$

Notice that the mapping h is well defined since if $f_j(z) = y_j$, $j = 1, 2, ..., n$, then $z - y \in \text{Ker } f_j$, $j = 1, 2, ..., n$, so

$$z - y \in \bigcap_{j=1}^{n} \text{Ker } f_j \subset \text{Ker } f$$

The mapping h is linear; if H is a linear extension of h to the whole space $\mathbb{K}^n$, there exist $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{K}$ such that

$$H(y_1, y_2, ..., y_n) = \sum_{j=1}^{n} \alpha_j y_j$$

In particular, for every $x \in X$, we have

$$f(x) = h(g_1(x), g_2(x), ..., g_n(x)) =$$

$$= H(g_1(x), g_2(x), ..., g_n(x)) = \sum_{j=1}^{n} \alpha_j g_j(x)$$

Thus, as each $f \in X^*$ is in Sp $\{f_1, f_2, ..., f_n\}$, we have proven that X^* is finite dimensional. This fact implies that X is also finite dimensional.

Now consider the dual space X^* of the normed space X. This has a norm topology and a weak topology also, $\sigma(X^*, X^{**})$. Another useful topology on X^*, is the locally convex topology, (obviously Hausdorff), defined by the family of seminorms $\{p_x\}_{x \in X}$, where $p_x(f) = |f(x)|$ for all $f \in X^*$. This topology, denoted by $\sigma(X^*, X)$ (or ω^*) is called the *weak* topology* on X^*. We notice that since $X \subset X^{**}$ (Proposition 3.2.1), $\sigma(X^*, X)$ is weaker than $\sigma(X^*, X^{**})$.

Remark. *We may regard X^* as a subset of the product $\mathbb{K}^X$. Taking into account that a typical $\sigma(X^*, X)$-neighborhood of zero contains a set of the form*

$$\{f \in X^* | \ |f(x_j)| < \varepsilon, \ \forall j, \ 1 \leq j \leq n\},$$

clearly the weak topology on X^* is the relativization to X^* of the product topology of $\mathbb{K}^X$.*

Theorem 7.3.2 *(Alaoglu's theorem) Let $(X, \|\cdot\|)$ be a normed space. Then, the unit ball of X^*, $B = \{f \in X^* | \ \|f\| \leq 1\}$ is compact for the weak* topology.*

Proof. For each $x \in X$ we set $B(x) = \{f(x) \mid f \in B\}$. Notice that

$$B \subset \prod_{x \in X} B(x).$$

The idea of the proof is the following. First, we will show that $\prod_{x \in X} B(x)$ is compact for the weak* topology. Our arguments here will be based on Tychonoff's theorem (Preliminaries), which states that in the product topology a product of sets is compact if and only if each of them is compact and on the fact that the weak* topology on $X^*$is the relativization to $X^*$of the product topology of $\mathbb{K}^X$. Second we will prove that B is ω^*-closed, which, since a closed subset of a compact set is also compact, will end the proof.

So, start showing that for each $x \in X$, $B(x)$ is compact. Since

$$|f(x)| \leq \|x\|, \quad \forall f \in B$$

it follows that $B(x) \subset \{\lambda \in \mathbb{K} | \; |\lambda| \le \|x\|\}$, therefore $B(x)$ is bounded in $\mathbb{K}$. Further, let $(f_n(x))_n$ be a sequence in $B(x)$ that converges to $\lambda \in \mathbb{K}$. Since

$$|f_n(x)| \longrightarrow |\lambda|$$

we have that $|\lambda| \le \|x\|$, so $|\lambda| \cdot \|x\|^{-1} \le 1$. By the Corollary 1.3.1 to the Hahn-Banach theorem, there exists $f_o \in X^*$ with $\|f_o\| = 1$ and $f_o(x) = \|x\|$. Let consider the linear functional on X defined by

$$f(y) = |\lambda| \cdot \|x\|^{-1} f_o(y), \quad \forall y \in X$$

We remark that $f(x) = \lambda$ and $f \in B$ since by

$$|f(y)| = |\lambda| \cdot \|x\|^{-1} |f_o(y)| \le |\lambda| \cdot \|x\|^{-1} \|y\|$$

we have that $f \in X^*$ and

$$\|f\| \le |\lambda| \cdot \|x\|^{-1} \le 1$$

We have just proved that $\lambda \in B(x)$, this shows that $B(x)$ is closed. Hence, we may infer by Tichonoff's theorem that $\prod_{x \in X} B(x)$ is compact for the relativization of the product topology on X^*, so for the ω^* topology.

Finally, let us verify that B is closed for the ω^* topology. Let f_o be in the ω^*-closure of B. We have to prove that f_o is in B, that means f_o is linear, bounded and its norm is less than one. Take arbitrary $x_1, x_2 \in X$, α, β in $\mathbb{K}$ and $\varepsilon > 0$ and consider the $\sigma(X^*, X)$-neighbourhood of f_o,

$$W = f_o + W_{\{(\alpha x_1 + \beta x_2),\, x_1,\, x_2\},\, \varepsilon}$$

It follows that for each $f \in W$, $|f(\alpha x_1 + \beta x_2) - f_o(\alpha x_1 + \beta x_2)| < \varepsilon$, $|f(x_1) - f_o(x_1)| < \varepsilon$ and $|f(x_2) - f_o(x_2)| < \varepsilon$.

As f_o is in the ω^*-closure of B, there exists $f_1 \in \in W \cap B$. By

$$|f_o(\alpha x_1 + \beta x_2) - \alpha f_o(x_1) - \beta f_o(x_2)| =$$

$$= |f_o(\alpha x_1 + \beta x_2) - f_1(\alpha x_1 + \beta x_2) + \alpha f_1(x_1) + \beta f_1(x_2) - \alpha f_o(x_1) - \beta f_o(x_2)| \le$$

$$\le |f_o(\alpha x_1 + \beta x_2) - f_1(\alpha x_1 + \beta x_2)| + |\alpha| \cdot |f_o(x_1) - f_1(x_1)| + |\beta| \cdot |f_o(x_2) - f_1(x_2)| \le$$

$$\le \varepsilon + |\alpha|\varepsilon + |\beta|\varepsilon = \varepsilon(1 + |\alpha| + |\beta|),$$

since ε is arbitrary, it follows that f_o is linear. Finally, consider the $\sigma(X^*, X)$-neighbourhood of f_o,

$$W = \{f \in X^* | \ |f(x) - f_o(x)| < \varepsilon\}$$

(with x in X and $\varepsilon > 0$ arbitrary). As f_o is in the ω^*-closure of B, there exists $f_1 \in W \cap B$. By

$$|f_o(x)| \leq |f_o(x) - f_1(x)| + |f_1(x)| < \varepsilon + \|f_1\| \cdot \|x\| < \varepsilon + \|x\|,$$

it follows that $|f_o(x)| \leq \|x\|$, $\forall x \in X$, i.e. f_o is bounded and $\|f_o\| \leq 1$. The theorem is proven. .

7.4 Linear functionals on locally convex spaces

At the beginning, let us make an useful remark.

Remark. A linear functional on a topological vector space X is continuous on X if it is continuous at zero. Indeed, if $f : X \longrightarrow \mathbb{K}$, is continuous at zero, $\forall \varepsilon > 0$, $\exists W \in \mathcal{V}_0$ such that for all x in W, $|f(x)| < \varepsilon$. Then $W + x_o$ is a neighbourhood of x_o (x_o arbitrary in X), and for all x in $W + x_o$ we have that $|f(x) - f(x_o)| < \varepsilon$.

Proposition 7.4.1 *Let X be a locally convex space such that its topology is defined by means of the directed family of seminorms $(p_j)_{j \in \mathcal{A}}$ and $f : X \longrightarrow \mathbb{K}$ linear. Then, the functional f is continuous on X if and only if $\exists j_o \in \mathcal{A}$ and $\eta > 0$ such that $|f(x)| \leq \eta \ p_{j_o}(x)$, $\forall x \in X$.*

Proof. Let f be continuous on X. Then, there exists a neighbourhood $W_{\{j_o\},\varepsilon_o}$ such that for all $x \in W_{\{j_o\},\varepsilon_o}$, $|f(x)| < 1$. Take now an arbitrary $x \in X$. Evidently, for each $\theta > 0$,.

$$x' = \frac{\varepsilon_o}{p_{j_o}(x) + \theta} \ x \in W_{\{j_o\},\varepsilon_o}$$

which, implies that

$$|f(x)| < \frac{1}{\varepsilon_o}(p_{j_o}(x) + \theta)$$

Since $\theta > 0$ is arbitrary, the previous inequality implies that

$$|f(x)| \leq \eta\, p_{j_o}(x), \forall x \in X,$$

(where $\eta = 1/\varepsilon_o$).

The converse is clear.

The next result is an immediate consequence of the Hahn-Banach extension theorem.

Theorem 7.4.1 *Let X be a locally convex space, let Y be a closed subspace of X, and let x_o be any point of X that $x_o \notin Y$. Then, there is a continuous linear functional f on X such that its restriction to Y is zero and $f(x_o) = 1$.*

Proof. Let x_o be in X, $x_o \notin Y$, thus $-x_o \notin Y$. Since Y is closed there exists $W \in \mathcal{V}_0$ such that $(W - x_o) \cap Y = \emptyset$. Then, $\{x \mid p(x) < \varepsilon\} \subset W$ for some seminorm p on X and $\varepsilon > 0$. If y is arbitrary in Y, as y is not in $W - x_o$, evidently $y + x_o \notin W$. It follows that $p(y + x_o) \geq \varepsilon$. Thus, we have

$$\frac{1}{\varepsilon} p(y + x_o) \geq 1, \quad \forall y \in Y.$$

If Z is the linear space spanned by $Y \cup \{x_o\}$, we define here the linear functional $f_o : Z \longrightarrow \mathbb{K}$, by

$$f_o(y + \lambda x_o) = \lambda, \quad y \in Y,\ \lambda \in \mathbb{K}.$$

Then,

$$|f_o(y + \lambda x_o)| = |\lambda| \leq |\lambda| \cdot p(\frac{y}{\lambda} + x_o) \cdot \frac{1}{\varepsilon} = \frac{1}{\varepsilon} \cdot p(\frac{y}{\lambda} + x_o),$$

equivalently,

$$|f_o(z)| \leq p_1(z), \quad \forall z \in Z$$

where p_1 is the seminorm $1/\varepsilon \cdot p$. By the Hahn-Banach extension theorem, $\exists f : X \longrightarrow \mathbb{K}$, a linear extension of f_o such that $|f(x)| < p_1(x)$, $\forall x \in X$. Applying Proposition 7.4.1, it follows that f is also continuous. Since f is an extension of f_o defined as above, clearly $f(y) = 0$, for each y in Y and $f(x_o) = 1$.

Corollary 7.4.1 *Let X be a Hausdorff locally convex space and x_o in X, $x_o \neq 0$. Then, there exists f a continuous linear functional on X such that $f(x_o) \neq 0$.*

Proof. One can apply the previous theorem for the closed subspace $Y = \{0\}$ and the point $x_o \notin Y$.

An immediate consequence of this corollary is:

Corollary 7.4.2 *Let X be a Hausdorff locally convex space and x, y in X, $x \neq y$. Then, there exists f a continuous linear functional on X such that $f(x) \neq f(y)$.*

The next theorem, apart from being an interesting result, will be an useful tool further. Recall that if X is a vector space over $\mathbb{C}$, a real linear functional on X is a mapping $f : X \longrightarrow \mathbb{R}$ which is additive and real homogeneous, i.e. $f(ax) = af(x)$, $\forall a \in \mathbb{R}$, $x \in X$.

Theorem 7.4.2 *Let X be a locally convex space, let C be a closed convex subset of X, and let x_o be any point of X that $x_o \notin C$. Then, there is a continuous, real linear functional f on X such that*

$$f(x_o) < \sup_{x\in C} f(x).$$

Proof. Since C is closed and $x_o \notin C$, there exists a balanced convex open neighbourhood of zero V such that $(x_o - V) \cap C = \emptyset$. Let us define

$$\widetilde{C} = \bigcup_{x\in C} (x + V)$$

and notice that clearly, $\widetilde{C}$ is an open convex set including C. In addition, $x_o \notin \widetilde{C}$ (otherwise, $\exists x \in C$ such that $x_o \in x+V$; it follows that $(x_o - V) \cap C \neq \emptyset$, contradiction). We shall show that there is no loss in generality in assuming that $0 \in \widetilde{C}$. Indeed, let z be any point in the open set $\widetilde{C}$. Clearly, $\widetilde{C} - z$ has zero in its interior and $x_o - z \notin \widetilde{C} - z$. If we can find a continuous, real linear functional $f : X \longrightarrow \mathbb{R}$ such that $f(x_o - z) > \sup\ \{f(y) \mid y \in \widetilde{C} - z\}$, then

$$f(x_o) > \sup\ \{f(y) \mid y \in \widetilde{C}\} \geq \sup\ \{f(y) \mid y \in C\}$$

and we would have proven the theorem.

Thus we may consider that $\widetilde{C}$ is a convex neighbourhood of zero, thus $\widetilde{C}$ is an absorbing convex set. Defining $p_{\widetilde{C}}$ the gauge function of $\widetilde{C}$, clearly (see Lemma 7.2.1) this is well defined ($\widetilde{C}$ is an absorbing set), subadditive ($\widetilde{C}$ is convex), and $p_{\widetilde{C}}(\alpha x) = \alpha p_{\widetilde{C}}(x)$, $\forall \alpha \in \mathbb{R}$, $\alpha \geq 0$, $x \in X$ (only for positive scalars since $\tilde{C}$ is not necessarily balanced). Regard X as a vector space over $\mathbb{R}$, let Y be the linear space spanned by $\{x_o\}$ and define $f_o : Y \longrightarrow \mathbb{R}$, by

$$f_o(\lambda x_o) = \lambda \cdot p(x_o), \quad \lambda \in \mathbb{R}$$

Let $y \in Y$, $y = \lambda x_o$. If $\lambda \geq 0$,

$$f_o(y) = f_o(\lambda x_o) = \lambda \cdot p(x_o) = p(\lambda x_o) = p(y)$$

For $\lambda < 0$,

$$f_o(y) = f_o(\lambda x_o) = \lambda \cdot p(x_o) < 0 \leq p(\lambda x_o) = p(y),$$

so, we infer that $f_o(y) \leq p(y)$, $\forall y \in Y$.

By the Hahn Banach extension theorem, there exists an extension of f_o, $f : X \longrightarrow R$, real linear such that $f(x) \leq p(x)$, $\forall x \in X$. Then,

$$f(x_o) = p(x_o) > 1 \geq \sup\{p(x) \mid x \in \tilde{C}\} \geq$$

$$\geq \sup\{f(x) \mid x \in \tilde{C}\} \geq \sup\{f(x) \mid x \in C\}.$$

Moreover f is continuous. Clearly, we can pick $U \subset \tilde{C}$, U balanced, thus

$$p_{\tilde{C}}(x) \leq p_U(x), \quad \forall x \in X$$

Then $f(x) \leq p_U(x)$, and $f(-x) \leq p_U(-x)$, $\forall x \in X$, accordingly,

$$|f(x)| \leq p_U(x), \quad \forall x \in X$$

Using Proposition 7.4.1 it results that f is continuous.

7.5 Extreme points, the Krein-Milman theorem

The next theorem (due to Krein and Milman) proves that the compact subsets of locally convex spaces have a useful geometric property. At the beginning let us introduce some more notions.

Definition. Let X be a vector space over the field $\mathbb{K}$ and let A be a nonempty subset of X. A nonempty subset B of A is said to be an *extreme subset* of A if a proper convex combination $\lambda x + (1 - \lambda)y$, $0 < \lambda < 1$ of two points x

and y of A lies in B only if both x and y are in B. An extreme subset of A consisting of just one point is called an *extreme point* of A.

Remark. Notice that $\omega \in A$ is an extreme point of A if the conditions x, $y \in A$, λ a real number such that $0 < \lambda < 1$ and $\lambda x + (1-\lambda)y = \omega$ imply that $x = y = \omega$.

Examples. 1. Let A be the solid unit square in $\mathbb{R}^2$,

$$A = \{(x, y) \mid 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$$

Then, the boundary of A,

$$B = \{(x, y) | x = 0,\ y = 0,\ x = 1,\ y = 1\}$$

is an extreme subset of A and the vertices of the square are extreme points of it.

2. Let A be the unit closed ball in $\mathbb{R}^2$, $A = \{(x, y) | x^2 + y^2 \leq 1\}$. Then, each point of the unit circle is an extreme point of A.

Notation. We will write $\mathcal{E}xtr\ A$ for the set of all extreme points of A.

Theorem 7.5.1 *(**The Krein-Milman theorem**) Let X be a Hausdorff, locally convex space and let K be any compact subset of X. Then,*
(1) $\mathcal{E}xtr\ K \neq \emptyset$;
(2) $\overline{\mathrm{co}\ (\mathcal{E}xtr\ K)} = \overline{\mathrm{co}\ K}$.

Proof. Let $\mathcal{P}$ be the family of all closed, extreme subsets of K and remark that $\mathcal{P}$ has the following properties:

(α) $\mathcal{P}$ is not the empty family (since it contains K).

(β) If $(S_\iota)_{\iota \in \Upsilon} \subset \mathcal{P}$ such that $\bigcap_{\iota \in \Upsilon} S_\iota \neq \emptyset$, then $\bigcap_{\iota \in \Upsilon} S_\iota \in \mathcal{P}$.

Indeed, since $\bigcap_{\iota \in \Upsilon} S_\iota$ is a closed subset of the compact set K, it follows that $\bigcap_{\iota \in \Upsilon} S_\iota$ is compact too. Moreover, for any x, $y \in K$ and λ, $0 < \lambda < 1$ with $\lambda x + (1-\lambda)y \in \bigcap_{\iota \in \Upsilon} S_\iota$ we have that $\lambda x + (1-\lambda)y \in S_\iota$, hence, since S_ι is an extreme subset of K, x, $y \in S_\iota$, $\forall \iota \in \Upsilon$, or, equivalently x, $y \in \bigcap_{\iota \in \Upsilon} S_\iota$.

(γ) Let S be in $\mathcal{P}$, f a continuous, real linear functional on X, and

$$\mu = \sup\{f(x) \mid x \in S\}$$

Then, the set of the points of S where f attains its maximum,

$$S_f = \{x \in S \mid f(x) = \mu\}$$

is in $\mathcal{P}$. Clearly, S_f is compact as a closed subset, ($S_f = f^{-1}(\{\mu\})$) of the compact K. Further, for any $x, y \in K$ and λ, $0 < \lambda < 1$ with $\lambda x + (1-\lambda)y \in$ $\in S_f$ it results first that $\lambda x + (1-\lambda)y \in S$. It follows, $f(x) \leq \mu$ and $f(y) \leq \mu$. Then,

$$\mu = f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y) \leq \lambda\mu + (1-\lambda)\mu = \mu,$$

which allows us to conclude that $f(x) = \mu$, $f(y) = \mu$, so $x, y \in S_f$.

Now we will prove the statements of the theorem in two steps.

I) First we show that for any S in $\mathcal{P}$, $S \cap \mathcal{E}xtr\ K \neq \emptyset$ (which, in particular will get $\mathcal{E}xtr\ K \neq \emptyset$).

Thus, for S arbitrary fixed in $\mathcal{P}$, let us denote by $\mathcal{S}$ the family of all subsets of S which are in $\mathcal{P}$ which clearly is not the empty family since $S \in \mathcal{P}$. We introduce a partial order, " $\prec$ " on $\mathcal{S}$,

$$T_1 \prec T_2 \iff T_2 \subset T_1$$

and notice that each chain $\mathcal{C} = (T_i)_{i\in I}$ in the partial ordered set $\mathcal{S}$, has an upper bound, as it results from the next. For any finite subset J of I, since $\mathcal{C}$ is totally ordered

$$\bigcap_{i\in J} T_i \neq \emptyset$$

Then, the family of compact sets $(T_i)_{i\in I}$ has the finite intersection property the

$$\bigcap_{i\in I} T_i \neq \emptyset$$

By the property (β) of $\mathcal{P}$, it follows that $\bigcap_{i\in I} T_i$ is in $\mathcal{P}$, thus, $\bigcap_{i\in I} T_i \in \mathcal{S}$. In addition

$$T_i \prec \bigcap_{i\in I} T_i, \quad \forall i \in I$$

Hence, by Zorn's lemma, $\mathcal{S}$ has a maximal element, T_o (so, $T_o \subset S$ and T_o is an extreme subset of K).

We will prove that T_o consists of a single point, $T_o = \{\omega\}$. Suppose there exist x, y in T_o, $x \neq y$; then by Corollary 7.4.1, there exists a continuous, real linear functional f on X such that $f(x) \neq f(y)$. If

$$\mu = \sup\{f(x) \mid x \in T_o\},$$

then,

$$T_{of} = \{z \in T_o \mid f(z) = \mu\} \subsetneqq T_o$$

By the property (γ) of $\mathcal{P}$, it follows that $T_{of} \in \mathcal{S}$. As $T_o \prec T_{of}$, our assumption allows us to a contradiction (since T_o is maximal).

Now, we can conclude that $S \cap \mathcal{E}xtr\ K \neq \emptyset$, because ω is an extreme point of K, and it is also contained in S.

II) Second, it remained to see that $\overline{\mathrm{co}(\mathcal{E}xtr\ K)} = \overline{\mathrm{co}\, K}$. Clearly,

$$\overline{\mathrm{co}(\mathcal{E}xtr\ K)} \subset \overline{\mathrm{co}\, K}$$

For the converse inclusion there is enough to check that

$$K \subset \overline{\mathrm{co}(\mathcal{E}xtr\ K)}$$

Otherwise, there is x_o in K,

$$x_o \notin \overline{\mathrm{co}(\mathcal{E}xtr\ K)}$$

By Theorem 7.4.2, there exists a continuous, real linear functional f on X such that

$$f(x_o) > \sup \{f(x) \mid x \in \overline{\mathrm{co}(\mathcal{E}xtr\ K)}\}$$

Let $\mu = \sup \{f(x) \mid x \in K\}$ and $K_f = \{x \in K \mid f(x) = \mu\}$. Then, $K_f \in \mathcal{P}$, and by the first step,

$$K_f \cap \mathcal{E}xtr\ K \neq \emptyset$$

This is a contradiction, since for any $x \in \mathcal{E}xtr\ K$, $f(x) < f(x_o) \leq \mu$, thus $x \notin K_f$.

Application to the Krein Milman theorem

We will prove that each probability measure on a Hausdorff compact space T can be approximated pointwise on $\mathcal{C}_{\mathbb{C}}(T)$ by measures with finite support on T. Let us be more specific.

If $\mathcal{C}_{\mathbb{C}}(T)$ is as usually the Banach space of all continuous complex functions on a Hausdorff compact space T, (endowed with the norm, $\|x\| = \sup_{t \in T} |x(t)|$), the dual of $\mathcal{C}_{\mathbb{C}}(T)$, denoted by $\mathcal{M}(T)$ is called the space of *Radon complex measures* on T. The set of all *probability measures* on T, $\mathcal{M}_p(T)$, is $\mathcal{M}_p(T) =$ $= \{\mu \in \mathcal{M}(T) \mid \|\mu\| \leq 1,\ \mu(\mathbf{1}) = 1\}$ (where $\mathbf{1} \in \mathcal{C}_{\mathbb{C}}(T)$, $\mathbf{1}(t) = 1$, $\forall t \in T$). We will consider on $\mathcal{M}(T)$ the ω^*-topology.

Theorem 7.5.2 *Let $\mathcal{M}(T)$ be the space of Radon complex measures on a Hausdorff compact space T, and $\mathcal{M}_p(T)$ be the set of all probability measures on T. Then,*

1) $\mathcal{M}_p(T)$ is a ω^*-compact convex subset of $\mathcal{M}(T)$;

2) $\mathcal{E}xtr\ (\mathcal{M}_p(T)) = \{\delta_t \mid t \in T\}$, where $\delta_t(x) = x(t)$, $\forall x \in \mathcal{C}_{\mathbb{C}}(T)$ (δ_t is called the unit point mass at t, or the Dirac measure at t).

Proof. 1) Clearly, $\mathcal{M}_p(T)$ is a ω^*-closed convex subset of the unit ball of $\mathcal{M}(T)$ that by Alaoglu's theorem is ω^*-compact. It follows that $\mathcal{M}_p(T)$ is also ω^*-compact.

2) First, we prove the following properties of the elements of $\mathcal{M}_p(T)$:

i) If $x \in \mathcal{C}_{\mathbb{C}}(T)$, and $x = \overline{x}, (x(t) = \overline{x(t)}, \forall t)$, then $\mu(x) \in \mathbb{R}, \forall \mu \in \mathcal{M}_p(T)$;

ii) If $x \in \mathcal{C}_{\mathbb{C}}(T)$, and $x \geq 0$, then $\mu(x) \geq 0$, $\forall \mu \in \mathcal{M}_p(T)$;

iii) If $x \in \mathcal{C}_{\mathbb{C}}(T)$, and $\mu \in \mathcal{M}_p(X)$, then $|\mu(x)| \leq \mu(|x|)$.

Let $x \in \mathcal{C}_{\mathbb{C}}(T)$, with $x = \overline{x}$ and suppose that $\mu(x) = a + ib$, $b \neq 0$. Then, for $\forall n \in \mathbb{N}$, we have

$$|\mu(x + ibn)| \leq \|x + ibn\|$$

Taking into account that $|\mu(x + ibn)| = a^2 + b^2(1+n)^2$ and $|x(t) + itn| = |x(t)|^2 + b^2n^2$, $\forall t \in T$, the previous inequality gives

$$a^2 + b^2(1+n)^2 \leq \|x\|^2 + b^2n^2$$

which enables us to conclude that $b = 0$, so i) holds.

Let $x \geq 0$ be and $\mu \in \mathcal{M}_p(T)$. Then,

$$\|x\| - \mu(x) = \mu(\|x\| - x) \leq \|x\| - x \leq \|x\|$$

which shows that $\mu(x) \geq 0$.

By i), $\mu(\operatorname{Re} x) = \operatorname{Re} \mu(x)$, $x \in \mathcal{C}_{\mathbb{C}}(T)$. Then, iii) follows from

$$|\mu(x)| = \mu(x)e^{-i \arg \mu(x)} = \mu(x \cdot e^{-i \arg \mu(x)}) = \operatorname{Re} \mu(x \cdot e^{-i \arg \mu(x)}) =$$

$$= \mu(\operatorname{Re} e^{-i \arg \mu(x)} x) \leq \mu(|x \cdot e^{-i \arg \mu(x)}|) = \mu(x)$$

We have also to remark that an immediate consequence of ii) is

ii$'$) If $x, y \in \mathcal{C}_{\mathbb{C}}(T)$, and $x \geq y$, then $\mu(x) \geq \mu(y)$, $\forall \mu \in \mathcal{M}_p(T)$.

Further, we will prove that each $\mu \in \mathcal{E}xtr\ (\mathcal{M}_p(T))$ is a multiplicative functional,

$$\mu(xy) = \mu(x)\mu(y), \quad \forall x,\ y \in \mathcal{C}_{\mathbb{C}}(T)$$

First, we notice, since the linear space spanned by $\{x \in \mathcal{C}_{\mathbb{C}}(T) \mid 0 \leq x \leq 1\}$ coincides to $\mathcal{C}_{\mathbb{C}}(T)$, there is sufficiently to prove that $\mu(xy) = \mu(x)\mu(y)$, only for x, $y \in \mathcal{C}_{\mathbb{C}}(T)$, $\|x\| \leq 1$, $\|y\| \leq 1$.

Let $x \in \mathcal{C}_{\mathbb{C}}(T)$ be such that $0 \leq x \leq 1$. Set $\alpha = \mu(x)$. Using ii), clearly $\alpha \in [0,1]$. If $\alpha = 0$, we have $\mu(x) \cdot \mu(y) = \alpha \cdot \mu(y) = 0$, and on the other hand, by

$$0 \leq |\mu(xy)| \leq \mu(\|y\| \cdot x) = \|y\| \cdot \mu(x) = 0,$$

$\mu(xy) = 0$, thus the equality holds. If $\alpha = 1$, by

$$0 \leq \mu((1-x)y)| \leq \|y\| \cdot \mu(1-x) = 0,$$

it follows that $\mu(y) = \mu(xy)$, equivalently, $\mu(xy) = \mu(y) \cdot 1 = \mu(y)\mu(x)$.

Now, for an arbitrary $\alpha \in (0,1)$ we define on $\mathcal{C}_{\mathbb{C}}(T)$ the functionals φ and ψ by

$$\varphi(y) = \alpha^{-1}\mu(xy), \quad \forall y \in \mathcal{C}_{\mathbb{C}}(T)$$

and respectively,

$$\psi(y) = \alpha^{-1}\mu[(1-x)y], \quad \forall y \in \mathcal{C}_{\mathbb{C}}(T)$$

Clearly, φ and ψ are linear and $\varphi(\mathbf{1}) = \psi(\mathbf{1}) = 1$. In addition,

$$|\mu(xy)| \leq \mu(|xy|) = \mu(x|y|) \leq \|y\| \cdot \alpha,$$

which implies that $|\varphi(y)| \leq \|y\|$, $\forall y \in \mathcal{C}_{\mathbb{C}}(T)$, so $\|\varphi\| \leq 1$.

Similarly,

$$|\mu((1-x)y)| \leq \mu(|xy|) = \mu((1-x)|y|) \leq \|y\| \cdot (1-\alpha),$$

thus, also $|\psi(y)| \leq \|y\|$, $\forall y \in \mathcal{C}_{\mathbb{C}}(T)$, hence $\|\psi\| \leq 1$.

We have obtained that φ, $\psi \in \mathcal{M}_p(T)$. But $\mu = \alpha\varphi + (1-\alpha)\psi$ and $\mu \in \mathcal{E}xtr\ (\mathcal{M}_p(T))$. It results that $\mu = \varphi = \psi$, therefore $\mu(y) = \mu(x)^{-1}\mu(xy)$.

The next step in our proof is to show that for each $\mu \in \mathcal{E}xtr\ (\mathcal{M}_p(T))$, there exists $t \in T$ such that Ker $\mu \subset$ Ker δ_t. Otherwise, suppose that for each t in W, there is $x^{<t>}$ in Ker μ with $x^{<t>}(t) \neq 0$. Then, there exists a neighbourhood of t, $V^{<t>}$ such that $x^{<t>}$ is strictly positive on $V^{<t>}$. As T is compact and contained in $\bigcup_{t\in T} V^{<t>}$, there exist $t_1, t_2, \ldots, t_n$ in T such that

$$X \subset \bigcup_{i=1}^{n} V^{<t_i>}$$

For each $i = 1, 2, .., n$, denote by x_i the function $x^{<t_i>}$, and by V_{t_i} the neighbourhood of t_i, $V^{<t_i>}$. We have that $x_i(s) > 0$, $\forall s \in V_{t_i}$. Then, the function

$$x = \sum_{i=1}^{n} x_i\overline{x_i}$$

has clearly the property that $x(t) \neq 0$, $\forall t \in T$. Hence,

$$1 = \mu(x \cdot \frac{1}{x}) = \mu(x) \cdot \mu(\frac{1}{x}) = 0,$$

(contradiction). It follows that Ker $\mu \subset$ Ker δ_t.

Then, for each $x \in \mathcal{C}_{\mathbb{C}}(T)$,

$$x - \mu(x) \cdot \mathbf{1} \in \operatorname{Ker} \mu \subset \operatorname{Ker} \delta_t,$$

thus, $x(t) = \mu(x)$. Consequently, we have obtained, that each μ belonging to $\mathcal{E}xtr\ (\mathcal{M}_p(T))$, is also in $\{\delta_t \mid t \in T\}$.

In order to end the proof we need to see that $\forall t \in T$,

$$\delta_t \in \mathcal{E}xtr\ (\mathcal{M}_p(T))$$

Let φ, $\psi \in \mathcal{M}_p(T)$ and $\alpha \in (0,1)$ such that $\delta_t = \alpha\varphi + (1-\alpha)\psi$. By,

$$\alpha|\varphi(y)| \leq \alpha\varphi(|y|) \leq \delta_t(|y|) = |y(t)|$$

it results that Ker $\varphi \supset$ Ker δ_t, thus as above $\varphi(y) = \delta_t(y)$, $\forall\ y \in \mathcal{C}_{\mathbb{C}}(T)$. Similarly one prove that $\psi = \delta_t$.

Corollary 7.5.1 *Each probability measure on a Haudorff compact space T can be approximated pointwise on $\mathcal{C}_{\mathbb{C}}(T)$ by measures with finite support on T.*

Proof. Combining the previous theorem with the Krein Milman theorem,

$$\mathcal{M}_p(T) = \overline{\operatorname{co}\{\delta_t \mid t \in T\}}^{\omega^*}$$

Then, for each $\mu \in \mathcal{M}_p(T)$, there exists a net $(\theta_\beta)_\beta \subset \operatorname{co}\{\delta_t \mid t \in T\}$ such that $(\theta_\beta)_\beta$ converges to μ in the ω^*-topology. Taking into account that the elements of $\operatorname{co}\{\delta_t \mid t \in T\}$ are finite linear combinations of Dirac measures, the corollary is proven.

7.6 Fréchet spaces

In this section we will describe a special class of locally convex spaces, of those locally convex spaces whose topologies can be defined by means of a countable family of seminorms. Let us begin by a preliminary result.

Lemma 7.6.1 *Let (X,τ) a Hausdorff locally convex space, the topology τ being defined by means of the countable family of seminorms $(p_n)_{n\in\mathbb{N}}$. For each $x \in X$ define*

$$q(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{p_n(x)}{1+p_n(x)}$$

1) Then:
i) $q(x) = 0$ if and only if $x = 0$;
ii) $q(-x) = q(x)$, $\forall x \in X$;
iii) $q(x+y) \leq q(x) + q(y)$, $\forall x, y \in X$.
2) For each $x, y \in X$ define $d(x,y) = q(x-y)$. Then, d is a translation invariant metric on X and the topology defined by d coincides to τ.

Proof. 1) Clearly, for each $x \in X$ the series

$$\sum_{n\geq 1} \frac{1}{2^n} \cdot \frac{p_n(x)}{1+p_n(x)}$$

is convergent, since

$$\frac{1}{2^n} \cdot \frac{p_n(x)}{1+p_n(x)} \leq \frac{1}{2^n}, \quad \forall n \in \mathbb{N}$$

Suppose $q(x) = 0$. Then, because of

$$\frac{1}{2^n} \cdot \frac{p_n(x)}{1+p_n(x)} \leq q(x), \quad \forall n \in \mathbb{N}$$

it follows that for each $n \in \mathbb{N}$, $p_n(x) = 0$. As the family of seminorms $(p_n)_{n\in\mathbb{N}}$ satisfies the separation condition ((X,τ) is a Hausdorff space) we have $x = 0$.
ii) is evident by

$$p_n(-x) = p_n(x)$$

For each $x, y \in X$, taking into account that

$$p_n(x+y) \leq p_n(x) + p_n(y)$$

and using the fact that the function $t \longmapsto t(1+t)^{-1}$ is increasing on $[0,\infty)$ we have

$$\frac{1}{2^n} \cdot \frac{p_n(x+y)}{1+p_n(x+y)} \leq \frac{1}{2^n} \cdot \frac{p_n(x)+p_n(y)}{1+p_n(x)+p_n(y)} =$$

$$= \frac{1}{2^n}\left(\frac{p_n(x)}{1+p_n(x)+p_n(y)} + \frac{p_n(y)}{1+p_n(x)+p_n(y)}\right) \leq$$

$$\leq \frac{1}{2^n}\left(\frac{p_n(x)}{1+p_n(x)} + \frac{p_n(y)}{1+p_n(y)}\right), \ \forall n \in \mathbb{N}$$

Summarizing, iii) results.

An immediate computation shows that d is a translation invariant metric. Let us denote by τ_d the topology defined on X by d. We shall investigate the relationship between the τ_d- and τ-neighbourhoods of zero in X. We prove first that for each $\varepsilon > 0$, there exists a finite set $F = \{i_1, i_2, ..., i_n\} \subset \mathbb{N}$ and $\delta > 0$ such that

$$W_{F,\delta} \subset B_q(\varepsilon)$$

(where, as in section 7.2, $W_{F,\delta} = \{x \in X \mid p_j(x) < \delta, \ j \in F\}$ and, as usually $B_q(\varepsilon) = \{x \in X \mid q(x) < \varepsilon\}$).

We shall consider the increasing family of seminorms $(p_n')_{n\in\mathbb{N}}$,

$$p_n'(x) = \max\{p_1(x), \ p_2(x), ..., p_n(x)\}$$

and we notice that the locally convex topology defined by means of this family coincides to τ. For each $x \in X$ define

$$q'(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{p_n'(x)}{1+p_n'(x)}$$

Since

$$p_n'(x) \geq p_n(x), \quad \forall x \in X$$

and the function $t \longmapsto t(1+t)^{-1}$is increasing on $[0,\infty)$ we have that $q'(x) \geq$ $\geq q(x)$, $\forall x \in X$, hence, for each $\varepsilon > 0$,

$$B_{q'}(\varepsilon) \subset B_q(\varepsilon)$$

It follows it is enough to show that for each $\varepsilon > 0$, $\exists \eta > 0$ and $m \in \mathbb{N}$, such that

$$W_{\{p_m'\},\eta} \subset B_{q'}(\varepsilon)$$

Let $\varepsilon > 0$ be; then $2^{-k} < \varepsilon$ for some $k > 0$. We show that

$$W_{\{p_{k+1}'\},2^{-(k+1)}} \subset B_{q'}(\varepsilon)$$

Indeed, let $x \in X$ such that

$$p'_{k+1}(x) < 2^{-(k+1)}$$

As $(p'_n)_{n\in\mathbb{N}}$ is an increasing family it follows that

$$p'_j(x) < 2^{-(k+1)}, \quad \forall j = 1, 2, ..., k+1$$

Then,

$$q'(x) = \sum_{j=1}^{\infty} \frac{1}{2^j} \cdot \frac{p'_j(x)}{1+p'_j(x)} \leq \sum_{j=1}^{k+1} \frac{1}{2^j} p'_j(x) + \sum_{j=k+2}^{\infty} \frac{1}{2^j} <$$

$$< \frac{1}{2^{k+1}} \cdot \sum_{j=1}^{k+1} \frac{1}{2^j} + \frac{1}{2^{k+1}} < \frac{1}{2^{k+1}} \cdot \sum_{j=1}^{\infty} \frac{1}{2^j} + \frac{1}{2^{k+1}} = \frac{1}{2^k} < \varepsilon$$

Setting $\eta = 2^{-(k+1)}$ and $m = k+1$, it follows that

$$W_{\{p'_m\},\eta} \subset B_{q'}(\varepsilon)$$

Thus τ is stronger than τ_d.

Now, let $W_{\{i_1,i_2,...i_n\},\varepsilon}$. Let k be natural such that $2^{-k} < \varepsilon$. We have

$$B_q\left(\frac{1}{2^{i_1+i_2+...+i_n+k+1}}\right) \subset W_{\{i_1,i_2,...i_n\},2^{-k}} \subset W_{\{i_1,i_2,...i_n\},\varepsilon}$$

Indeed, let x be with

$$q(x) < \frac{1}{2^{i_1+i_2+...+i_n+k+1}} \iff \sum_{j=1}^{\infty} \frac{1}{2^j} \cdot \frac{p_j(x)}{1+p_j(x)} < \frac{1}{2^{i_1+i_2+...+i_n+k+1}}$$

As

$$\frac{1}{2^{i_l}} \cdot \frac{p_{i_l}(x)}{1+p_{i_l}(x)} < q(x), \qquad \forall l = 1, 2, ..., n$$

we have

$$\frac{1}{2^{i_l}} \cdot \frac{p_{i_l}(x)}{1+p_{i_l}(x)} < \frac{1}{2^{i_l} \cdot 2^{\sum_{j=1,\neq i_l}^{n} i_j + k + 1}}$$

from where it follows that

$$1 + p_{i_l}(x) > p_{i_l}(x) \cdot 2^{\sum_{j=1,\neq i_l}^{n} i_j + k + 1},$$

thus,

$$p_{i_l}(x) < \frac{1}{2^{\sum_{j=1,\neq i_l}^{n} i_j + k + 1}} < \frac{1}{2^k}$$

We conclude that the topology τ_d is stronger than τ, so finally the two topologies coincide.

Theorem 7.6.1 *Let (X, τ) a Hausdorff locally convex space. Then, the following are equivalent:*
(1) The topology τ is metrizable;
(2) There exists $\mathcal{B}$ a countable, fundamental system of τ-neighbourhoods of zero;
(3) There exists $\mathcal{P}$ a countable family of seminorms that satisfies the separation condition such that the topology defined by $\mathcal{P}$ coincides to τ.

Proof. It is clear that (1) implies (2). Assuming (2), it is no loss in generality supposing that all neighbourhoods in $\mathcal{B} = \{U_n\}_n$ are convex and balanced. Let p_n be the gauge function of U_n for each n. By Theorem 7.2.1, clearly the topology τ is defined by means of the countable family of seminorms $\mathcal{P} = \{p_n\}_n$, then (2) implies (3). (3) involves (1) follows from the previous lemma.

Definition. A locally convex space whose topology is metrizable is called a *metrizable locally convex* space. A complete, metrizable, locally convex space is called a *Fréchet* space.

Example. In this example, we shall denote by $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $n \in \mathbb{N}$, $\mathbb{R}^n$ is, as usually endowed with the euclidean norm, $\|\cdot\|_2$. For the closed ball of radius $m > 0$, we shall write as usually $\overline{B}(m)$. Let be $C^\infty(\mathbb{R}^n)$ the space of all numerical functions on $\mathbb{R}^n$ which have partial derivatives of any order. If $l = (l_1, l_2, ..., l_n) \in \mathbb{N}_0^n$, $l_i \in \mathbb{N}_0$ is a $n-$multi-index, we write $|l|$ for $l_1 + l_2 + ... + l_n$. For each arbitrary function $x \in C^\infty(\mathbb{R}^n)$ and each multi-index l, let

$$D^l x = \frac{\partial^{|l|} x}{\partial t_1^{l_1} \partial t_2^{l_2} ... \partial t_n^{l_n}}$$

Further, if we write $\operatorname{supp} x$ for the support of $x \in C^\infty(\mathbb{R}^n)$, let us denote by $\mathcal{D}_m = \{x \in C^\infty(\mathbb{R}^n) \mid \operatorname{supp} x \subset \overline{B}(m)\}$. We can define a countable family of seminorms on $\mathcal{D}_m$, $(p_j^{(m)})_{j \in \mathbb{N}_0}$ by

$$p_j^{(m)}(x) = \sup_{t \in \overline{B}(m)} \{|D^l x(t)| \mid l \in \mathbb{N}_0^n, \ |l| < j\}$$

The family $(p_j^{(m)})_{j \in \mathbb{N}_0}$ is a countable family of seminorms that satisfies the separation condition; it defines a metrizable, locally convex topology on $\mathcal{D}_m$, τ_m. Moreover $(\mathcal{D}_m, \tau_m)$ is a Fréchet space. Indeed, let $(x_k)_k$ be a Cauchy sequence, $(x_k)_k \subset \mathcal{D}_m$. Thus, for each $j \in \mathbb{N}_0$ and $\varepsilon > 0$, there exists $k(j, \varepsilon)$ such that $\forall i,\ k \geq k(j, \varepsilon)$,

$$|D^l x_i(t) - D^l x_k(t)| < \varepsilon,\ \forall\ l \in \mathbb{N}_0^n,\ |l| < j,\ \forall t \in \overline{B}_m,$$

Hence, it follows that $(D^l x_k)_k$ is uniformly Cauchy on $\overline{B}(m)$, so uniformly convergent to the continuous function y^l ($\forall\ l \in \mathbb{N}_0^n$, $|l| < j$). It results that $y^l = D^l y_0$, $\forall\ l \in \mathbb{N}_0^n$, $|l| < j$. As, this judgement is valid for each $j \in \mathbb{N}_0$, we can conclude that $y_0 \in \mathcal{D}_m$, and the fact that $(\mathcal{D}_m, \tau_m)$ is a Fréchet space is proved.

7.7 Inductive limits of locally convex spaces

Let X be a vector space over the field $\mathbb{K}$ and $(X_j)_{j \in J}$ a family of linear subspaces of X with the following properties:

i) $X = \bigcup_{j \in J} X_j$;

ii) The family $(X_j)_{j \in J}$ is directed by inclusion (i.e. $\forall j_1,\ j_2 \in J$, $\exists j_3 \in J$ such that $X_{j_1} \subset X_{j_3}$ and $X_{j_2} \subset X_{j_3}$);

iii) For each $j \in J$, the space X_j is endowed with a locally convex topology, τ_j such that if $X_j \subset X_l$, the trace of τ_l on X_j is weaker than τ_j.

Let us consider further $\mathcal{W}$, the family of all balanced, convex subsets W of X enjoying the next property: $\forall j \in J$, the set $W \cap X_j$ is a τ_j-neighbourhood of zero in X_j. Clearly, the family $\mathcal{W}$ satisfies the hypotheses of the Theorem 7.1.1, accordingly, there exists τ_{ind} a topology compatible with the vector structure of X, which in addition is locally convex, such that $\mathcal{W}$ is a fundamental system of neighbourhoods of the origin for this topology.

The space X endowed with the topology τ_{ind} defined as above is called the *inductive limit* of the spaces $(X_j, \tau_j)_{j \in J}$ and is denoted by

$$(X, \tau_{ind}) = \underrightarrow{\lim}\ (X_j, \tau_j)$$

Proposition 7.7.1 *Let* $(X, \tau_{ind}) = \underrightarrow{\lim}\ (X_j, \tau_j)$ *be and for* $\forall j \in J$, *let* l_j *be the canonical mapping which maps* X_j *in* X, $l_j(x) = x$. *Then,* τ *is the strongest locally convex topology such that* l_j *is continuous,* $\forall j \in J$.

Proof. Let τ' be a locally convex topology such that

$$l_j : (X_j, \tau_j) \longrightarrow (X, \tau')$$

is continuous, $\forall j \in J$. Then, for each arbitrary τ'-neighbourhood of zero, U,

$$l_j^{-1}(U) = U \cap X_j$$

is a τ_j-neighbourhood of zero, thus $U \in \mathcal{W}$. It follows that τ_{ind} is stronger than τ'.

Proposition 7.7.2 *Let* $(X, \tau_{ind}) = \varinjlim (X_j, \tau_j)$ *be and* f *be a linear functional on* X, $f : X \longrightarrow \mathbb{K}$. *Then, the functional* f *is continuous if and only if for each* $j \in J$,

$$f \circ l_j : (X_j, \tau_j) \longrightarrow \mathbb{K}$$

is continuous.

Proof. Clearly, by the previous proposition, if f is continuous, then, $\forall j \in J$, $f \circ l_j$ is continuous. Conversely, suppose that $\forall j \in J$, $f \circ l_j$ is continuous, so, for each $\varepsilon > 0$, $(f \circ l_j)^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\})$ is a τ_j-neighbourhood of zero, $\forall j \in J$. Since

$$(f \circ l_j)^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\}) = l_j^{-1} \circ (f^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\}))$$

we have that, for every $j \in J$, $f^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\}) \cap X_j$ is a τ_j-neighbourhood of zero ($\forall j \in J$). Evidently $f^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\}) \cap X_j$ is balanced and convex. It follows that $f^{-1}(\{\lambda \in \mathbb{K} \mid |\lambda| < \varepsilon\})$ is a τ_{ind}-neighbourhood of zero, which proves that f is continuous.

Example. (*The distributions space, distributions*) Let us denote by

$$\mathcal{D} = \{x \in C^\infty(\mathbb{R}^n) \mid \operatorname{supp} x \text{ compact}\}$$

and consider the family of vector subspaces of $\mathcal{D}$, $(\mathcal{D}_m, \tau_m)_{m \in \mathbb{N}}$ (where for each $m \in \mathbb{N}$, $(\mathcal{D}_m, \tau_m)$ is the Fréchet space of the example in the section 7.6). It is easy to see that $\mathcal{D} = \bigcup_m \mathcal{D}_m$,and that $\mathcal{D}_m \subset \mathcal{D}_{m+1}$, $\forall m$, thus the family $(\mathcal{D}_m, \tau_m)_{m \in \mathbb{N}}$ is directed by inclusion. The equality

$$\{x \in \mathcal{D}_m \mid p_j^{(m)}(x) < \varepsilon\} = \{x \in \mathcal{D}_m \mid p_j^{(m+1)}(x) < \varepsilon\} \cap \mathcal{D}_m$$

shows that for each $m \in \mathbb{N}$ the relativization of the topology τ_{m+1} to $\mathcal{D}_m$ coincides to τ_m. Then, we can define $(\mathcal{D}, \tau_{ind})$, the inductive limit of $(\mathcal{D}_m, \tau_m)_{m \in \mathbb{N}}$.

The space $(\mathcal{D}, \tau_{ind})$ *is called the distributions space* (the Schwartz space).

A continuous linear functional on $(\mathcal{D}, \tau_{ind})$ *is called a distribution on* $\mathbb{R}^n$. By the Proposition 7.7.2 it follows that $f : (\mathcal{D}, \tau_{ind}) \longrightarrow \mathbb{K}$ is a distribution if and only if $\forall m \in \mathbb{N}$, $f \circ l_m : (\mathcal{D}_m, \tau_m) \longrightarrow \mathbb{K}$ is continuous, therefore $\forall m \in \mathbb{N}$, $\exists j \in \mathbb{N}_0$, and $\exists \mu > 0$ such that $|f(x)| \le \mu\, p_j^{(m)}(x)$, $\forall x \in \mathcal{D}_m$.

7.8 Exercises

1. Let T be a Hausdorff topological space and $\mathcal{C}_{\mathbb{K}}(T)$ the linear space of all continuous functions, $x : T \longrightarrow \mathbb{K}$. Denote by $\mathcal{K}(T)$ the family of compact subsets of T. For $Q \in \mathcal{K}(T)$ and $\varepsilon > 0$ define

$$W_{Q,\varepsilon} = \{x \in \mathcal{C}_{\mathbb{K}}(T) \mid \sup_{t \in Q} |x(t)| < \varepsilon\}$$

Show that there exists τ a topology compatible with the vector structure of $\mathcal{C}_{\mathbb{K}}(T)$ such that the family $\mathcal{W} = \{ W_{Q,\varepsilon} | \ Q \in \mathcal{K}(T), \ \varepsilon > 0\}$ is a fundamental system of τ-neighbourhoods (this topology is known as the topology of the compact convergence).

2. Let $X = \{x \in \mathcal{C}_{\mathbb{K}}([0,1]) \mid \exists \alpha \in (0,1]$ such that $x(t) = 0, \ \forall t \in (0.\alpha]\}$ endowed with the usual norm topology of $\mathcal{C}_{\mathbb{K}}([0,1])$. Show that the set

$$D = \{x \in X \mid n \cdot |x(1/n)| \leq 1, \ \forall n \in \mathbb{N}\}$$

is an absorbing, balanced, convex set, but D is not a neighbourhood of zero in X.

3. Let X be a topological vector space and $f : X \longrightarrow \mathbb{K}$, f linear. The following are equivalent:

(1) The functional f is continuous;

(2) Ker f is a closed subspace of X.

4. Let (X, τ) be a locally convex space, τ being defined by means of a family of seminorms $(p_j)_{j \in J}$. Show that the net $(x_\alpha)_{\alpha \in A} \subset X$ converges to $x \in X$ if and only if $\forall j \in J$ the net $(p_j(x_\alpha - x))_{\alpha \in A}$ converges to zero.

5. Let (X, τ) be a locally convex space, τ being defined by means of a family of seminorms $(p_j)_{j \in J}$. Show that the set $B \subset X$ is τ-bounded if and only if $\forall j \in J$, $\exists \lambda_j \geq 0$ such that $p_j(x) \leq \lambda_j$, $\forall x \in B$.

6. Let (X, τ) be a locally convex space, τ being defined by means of the directed family of seminorms $\mathcal{P} = (p_j)_{j \in J}$ and q a seminorm on X.

a) The following are equivalent:

(1) The seminorm q is continuous;

(2) The seminorm q is continuous at zero;

(3) The set $\{x \in X \mid q(x) < 1\}$ is a τ-neighborhood of zero.

(4) $\exists c > 0$ and $p \in \mathcal{P}$ such that $q(x) \leq cp(x)$, $\forall x \in X$.

b) If q is a continuous seminorm on X, then the family $\mathcal{P} \cup \{q\}$ defines the same topology on X as $\mathcal{P}$.

7. Let $s_{\mathbb{K}}$ be the linear space of all numerical sequences. Let $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ three family of seminorms defined as follows:

$$\mathcal{P} = \{p_n|\ n \in \mathbb{N}\}, \quad p_n((\xi_k)_k) = |\xi_n|$$

$$\mathcal{Q} = \{q_n|\ n \in \mathbb{N}\}, \quad q_n((\xi_k)_k) = \max_{1 \le k \le n} |\xi_k|$$

$$\mathcal{R} = \{r_n|\ n \in \mathbb{N}\}, \quad r_n((\xi_k)_k) = \sum_{k=1}^{n} |\xi_k|$$

Show that:

a) $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ satisfy the separation condition;

b) $\mathcal{Q}$, $\mathcal{R}$ are directed and $\mathcal{P}$ is not directed;

c) The locally convex topologies defined by means of the families $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ coincide.

d) The locally convex topology defined at c) coincides to the product topology on $s_{\mathbb{K}} = \mathbb{K}^{\mathbb{N}}$.

e) The space $s_{\mathbb{K}}$ equipped with the previous topology is a Fréchet space.

8. Let T be a Hausdorff topological space and $\mathcal{C}_{\mathbb{K}}(T)$ the linear space of all continuous functions, $x : T \longrightarrow \mathbb{K}$. . Show that the topology defined at the Exercise1 coincides to the locally convex topology defined by the family of seminorms $(p_Q)_{Q \in \mathcal{K}(T)}$, where

$$p_Q(x) = \sup_{t \in Q} |x(t)|$$

9. Let (X, τ) be a topological vector space locally bounded (i.e. there exists a τ-bounded neighborhood of zero). Show that (X, τ) is metrizable. Considering the case of the space $s_{\mathbb{K}}$ (Exercise7) show that, in general, a metrizable topological vector space is not locally bounded.

10. Show that $s_{\mathbb{K}}$ is not normable.

11. Show that $\mathcal{C}_{\mathbb{K}}(T)$, T noncompact (Exercise 8) is not normable.

12. Show that the product of the family of normed space $(X_j)_{j \in J}$, $X_j \neq \neq \{0\}$ is normable if and only if J is finite.

Appendix A

Equicontinuity

We notice that, as usually, a sequence $(x_n)_n \subset \mathcal{C}_{\mathbb{K}}([a,b])$ is said to be pointwise convergent if for each $t \in [a,b]$, the numerical sequence $(x_n(t))_n$ is convergent in $\mathbb{K}$; the sequence $(x_n)_n \subset \mathcal{C}_{\mathbb{K}}([a,b])$ is said to be uniform convergent to the $\mathbb{K}$-valued function x defined on $[a,b]$ if for each $\varepsilon > 0$, there is n_ε such that $n \geq n_\varepsilon$ implies $|x_n(t) - x(t)| < \varepsilon$, $\forall t \in [a,b]$. As the function x is necessarily continuous, this is equivalent to the fact that the sequence $(x_n)_n$ is convergent in the normed space $(\mathcal{C}_{\mathbb{K}}([a,b]), \|\cdot\|)$ $(\|x\| = \sup\limits_{t \in [a,b]} |x(t)|)$ to $x \in \mathcal{C}_{\mathbb{K}}([a,b])$.

Definition. Let $\mathcal{F}$ be a family of functions from a metric space (M,d) to another metric space (N,ρ). We say $\mathcal{F}$ is an *equicontinuous* family at t if and only if for each $\varepsilon > 0$, there is $\delta_{\varepsilon,t} > 0$ such that $d(t,t') < \delta_{\varepsilon,t}$ implies $\rho(x(t), x(t')) < \varepsilon$.
The family $\mathcal{F}$ is said to be an equicontinuous family on M if it is an equicontinuous family at each $t \in M$.

Theorem A.0.1 *Let be $(x_n)_n$ a sequence of functions from one metric space to another with the property that the family $\mathcal{F} = \{x_n \mid n \in \mathbb{N}\}$ is equicontinuous. Suppose that $(x_n)_n$ converges pointwise to x. Then, x is continuous.*

Theorem A.0.2 *Let $(x_n)_n$ be a sequence of functions from one metric space (M,d) to another metric space (N,ρ) with N complete such that the family $\mathcal{F} = \{x_n \mid n \in \mathbb{N}\}$ is equicontinuous. Suppose that $(x_n)_n$ converges pointwise on a dense subset of M. Then $(x_n)_n$ converges pointwise on M.*

Definition. Let $\mathcal{F}$ be a family of functions from a metric space (M,d) to another metric space (N,ρ). We say $\mathcal{F}$ is a *uniformly equicontinuous* family if and only if for each $\varepsilon > 0$, there is $\delta_\varepsilon > 0$ such that $d(t,t') < \delta_\varepsilon$ implies $\rho(x(t),x(t')) < \varepsilon$.

Theorem A.0.3 *Let $(x_n)_n$ be a sequence of functions from $[a,b]$ to some metric space (N,ρ) such that the family $\mathcal{F} = \{x_n \mid n \in \mathbb{N}\}$ is uniformly equicontinuous. Then $(x_n)_n$ converges uniformly on $[a,b]$.*

We notice that the above theorems can be proved immediately by an ε argument.

Theorem A.0.4 *(Ascoli's theorem) Let $(x_n)_n$ be a sequence of functions from $[a,b]$ to some metric space (N,ρ) such that the family $\mathcal{F} = \{x_n \mid n \in \mathbb{N}\}$ is uniformly bounded and equicontinuous. Then there is a subsequence $(x_{n'})_{n'}$ of $(f_n)_n$ such that $(x_{n'})_{n'}$ converges uniformly on $[a,b]$*

Proof. Let $(q_m)_m$ be a numbering of the rationals. Since $\{x_n \mid n \in \mathbb{N}\}$ is uniformly bounded, $|x_n(q_m)| \leq M, \forall n,m$. Thus, by the diagonalization trick, we can find a subsequence with $(x_{n'}(q_m))_{n'}$ converges as $n' \to \infty$ for each m. By Theorem A.0.2, the sequence $(x_{n'})_{n'}$ converges pointwise everywhere and then, by Theorem A.0.3, $(x_{n'})_{n'}$ converges uniformly.

Remark. Ascoli's theorem shows that a subset A of the Banach space $(C_{\mathbf{K}}[a,b], \|\cdot\|)$ is relative compact if it is bounded and equicontinuous.

Appendix B

Weierstrass approximation theorems

We state here Weierstrass approximation theorems, used in our approach especially in the section concerning orthonormal bases .

Theorem B.0.5 *(Weierstrass approximation theorem) If x is a complex (real) valued function which is continuous on $[a, b]$, then for every $\varepsilon > 0$ there exists a polynomial p such that*

$$|x(t) - p(t)| < \varepsilon \text{ for all } t \in [a, b]$$

Remark. The above theorem shows that the linear subspace of polynomials of the normed linear space $(\mathcal{C}_{\mathbb{K}}[0,1], \|\cdot\|)$ is dense in $(\mathcal{C}_{\mathbb{K}}[0,1], \|\cdot\|)$.

Theorem B.0.6 *(Weierstrass second approximation theorem). If x is a complex (real) valued function which is continuous on $[-\pi, \pi]$, and $x(-\pi) =$ $= x(\pi)$, then for every $\varepsilon > 0$ there exists a trigonometric polynomial*

$$T_n(t) = \sum_{j=0}^{n} (a_j \cos\ jt + b_j \sin\ jt)$$

such that

$$|x(t) - T_n(t)| < \varepsilon \textit{ for all } t \in [a, b]$$

Appendix C

Measure spaces

Definition. A triple $(X, \mathcal{X}, m)$ is called a *measure space* if
i) $(X, \mathcal{X})$ is a σ-ring, i.e. $\mathcal{X}$ is a family of subsets of the set X such that:
i1) $X \in \mathcal{X}$; i2) $C_X B \in \mathcal{X}$, $\forall B \in \mathcal{X}$; i3) $\bigcup_{n=1}^{\infty} B_n \in \mathcal{X}$, for any countable family of sets $(B_n)_n \in \mathcal{X}$ $(n = 1, 2, ...)$.
ii) m is a non-negative, σ-additive measure defined on $\mathcal{X}$, i.e.
ii1)) $m(B) \geq 0$, $\forall B \in \mathcal{X}$;
ii2) $m(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} m(B_n)$ for any disjoint sequence $(B_n)_n \in \mathcal{X}$.

Definition. A real- (or complex-) valued function $x(s)$ defined on X is said to be $\mathcal{X}$-*measurable* or, short, measurable if the following condition is satisfied: for any open set $G \subset \mathbb{R}$ (or $\mathbb{C}$), the set $x^{-1}(G)$ belongs to $\mathcal{X}$ (it is permitted that $x(s)$ takes the value ∞).

Definition. Let $(X, \mathcal{X}, m)$ and $(X', \mathcal{X}', m')$ be two measure spaces. We denote by $\mathcal{X} \times \mathcal{X}'$ the smallest σ-ring of subsets of $X \times X'$ which contains all the sets of the form $B \times B'$, where $B \in \mathcal{X}$, $B' \in \mathcal{X}'$. It is proved that there exists a uniquely determined non-negative measure m defined on $\mathcal{X} \times \mathcal{X}'$ such that

$$(m \times m')(B \times B') = m(B)m(B').$$

$m \times m'$ is called the *product measure* of m and m'.

A property pertaining to points s of X is said to hold m-almost everywhere (m-a.e.), if it holds for those s which form a set $A \in \mathcal{X}$ with $m(A) = 0$.

Definition. Let X be a closed subset of $\mathbb{R}^n$. The *Borel subsets* of X are the members of the smallest σ-ring $\mathcal{B}$ of subsets of X which contains every compact set of X. The non-negative *Borel measure* on X is a non-negative, σ-additive measure defined on $\mathcal{B}$, such that the measure of every compact set is finite.

Example. (The Lebesgue measure) Suppose X is the real line $\mathbb{R}$ or a closed interval of $\mathbb{R}$. Define a function m on the family of all countable unions of disjoint open intervals (which is just the family of open sets), by

$$m\left(\bigcup_{n=1}^{\infty}(a_n, b_n)\right) = \sum_{n=1}^{\infty} m(b_n - a_n)$$

This m has a uniquely determined extension to a non-negative Borel measure on X, called the Lebesgue measure,

$$m(B) = \inf\{m(I) \mid B \subset I,\ I \text{ open}\} = \sup\{m(K) \mid K \subset B,\ K \text{ compact}\}$$

The *Lebesgue measure* in $\mathbb{R}^n$ is obtained from the n-tuple of the one-dimensional Lebesgue measures through the process of forming the product measure.

Theorem C.0.7 *(Luzin's theorem) Let x be a real-valued function defined on the Lebesgue measurable set A. Then x is (Lebesgue) measurable if and only if for every $\varepsilon > 0$, there is a closed subset $F_\varepsilon \subset A$ such that $m(A \setminus F_\varepsilon) < \varepsilon$ and the function x is continuous on F_ε.*

Further we shall define and give some results concerning the integral with respect to a certain non-negative measure.

Definition. A real- (or complex-) valued function $x(s)$ defined on X is said to be *finitely-valued* if it is a finite non-zero constant on each of a finite number, say n, of disjoint $\mathcal{X}$-measurable sets B_j and zero on $X \setminus \bigcup_{j=1}^{n} B_j$. Let the value of $x(s)$ on B_j be denoted by x_j. Then x is m-integrable over X if $\sum_{j=1}^{n} |x_j|\ m(B_j) < \infty$, and the value $\sum_{j=1}^{n} x_j\ m(B_j)$ is defined as the integral of x over X with respect to the measure m, denoted by

$$\int_X x(s)\, \mathrm{d}m(s), \text{or, in short}, \int_X x\, \mathrm{d}m$$

A real- (or complex-) valued function $x(s)$ defined m-a.e. on X is said to be m-integrable over X if there exists a sequence $(x_n)_n$ of finitely-valued integrable functions converging to x m-a.e. and satisfying $\forall \varepsilon > 0$, there is n_ε such that $m, n \geq n_\varepsilon$ implies

$$\int_X |x_n(s) - x_m(s)| \, \mathrm{dm}(s) < \varepsilon$$

It is proved that, if the function x is m-integrable in the sense of the above definition, a finite

$$\lim_n \int_X x_n(s) \, \mathrm{dm}(s)$$

exists and the value of this limit is independent of the choice of the approximating sequence $(x_n)_n$.

Definition. The *integral* of the function x over X with respect to the measure m is defined by

$$\int_X x(s) \, \mathrm{dm}(s) = \lim_n \int_X x_n(s) \, \mathrm{dm}(s)$$

Notation. If $(X, \mathcal{X}, m)$ is a measure space, the set of all m-integrable functions over X is a linear space (with the usual sum and scalar multiplication), denoted by $\mathcal{L}^1(X, m)$; the linear space of equivalence classes of functions in $\mathcal{L}^1(X, m)$ equal m-a.e. is denoted by $L^1(X, m)$.

The following crucial theorems hold:

Theorem C.0.8 *(Monotone convergence theorem) If $x_n \in \mathcal{L}^1(X, m)$,*

$$0 \leq x_1 \leq x_2 \leq \ldots \text{ and } x(s) = \lim_n x_n(s),$$

then $x \in \mathcal{L}^1(X, m)$ if and only if

$$\lim_n \int_X |x_n(s)| \, \mathrm{dm}(s) < \infty$$

and in that case

$$\lim_n \int_X |x_n(s) - x(s)| \, \mathrm{dm}(s) = 0$$

and

$$\lim_n \int_X |x_n(s)| \, \mathrm{dm}(s) = \int_X |x(s)| \, \mathrm{dm}(s)$$

Theorem C.0.9 *(Dominated convergence theorem) If $x_n \in L^1(X,m)$,*

$$x(s) = \lim_n x_n(s),$$

m-a.e. and if there is $g \in L^1(X,m)$ with $|x_n(s)| < g(s)$ m-a.e., for all n, then $x \in L^1(X,m)$ and

$$\lim_n \int_X |x_n(s) - x(s)| \, \mathrm{dm}(s) = 0$$

Theorem C.0.10 *(Fatou's lemma) If $x_n \in \mathcal{L}^1(X,m)$, each $x_n(s) \geq 0$, and if*

$$\liminf_n \int_X |x_n(s)| \, \mathrm{dm}(s) < \infty,$$

then

$$x(s) = \liminf_n x_n(s) \in \mathcal{L}^1(X,m)$$

and

$$\int_X |x(s)| \, \mathrm{dm}(s) \leq \liminf_n \int_X |x_n(s)| \, \mathrm{dm}(s)$$

Theorem C.0.11 *(Fubini's theorem) Let $(X, \mathcal{X}, m)$ and $(X', \mathcal{X}', m')$ be two measure spaces, and $m \times m'$ the product measure of m and m'. Let x be a measurable function on $X \times X'$. Then*

$$\int_X \left(\int_{X'} |x(s,t)| \, \mathrm{d}\, m'(t) \right) \mathrm{dm}(s) < \infty$$

if and only if

$$\int_{X'} \left(\int_X |x(s,t)| \, \mathrm{dm}(s) \right) \mathrm{d}\, m'(t) < \infty$$

and if one (and thus both) of these integrals is finite, then

$$\int_X \left(\int_{X'} x(s,t) \, \mathrm{d}\, m'(t) \right) \mathrm{dm}(s) = \int_{X'} \left(\int_X |x(s,t)| \, \mathrm{dm}(s) \right) \mathrm{d}\, m'(t)$$

Theorem C.0.12 *Let $[a,b]$ a closed interval in $\mathbb{R}$. For any Lebesgue integrable function $x(s)$, and $\varepsilon > 0$, there is a continuous function g_ε on $[a,b]$ such that*

$$\int_{[a,b]} |x(s) - g_\varepsilon(s)| \, \mathrm{dm}(s) < \varepsilon$$

Appendix D

Holomorphic functions

In the sequel, x is a complex valued function defined on a open set D of $\mathbb{C}$.

Definition. The function $x : D \longrightarrow \mathbb{C}$ has *derivative* at $a \in D$ if there exists

$$\lim_{t \to a} \frac{x(t) - x(a)}{t - a}$$

The value of this limit is called the derivative of x at a, written $x'(a)$.

Definition. The function $x : D \longrightarrow \mathbb{C}$ is called *holomorphic* (*analytic*) if it posses a derivative wherever the function is defined.

Definition. A *power series* is of the form $\sum_{n \geq 0} a_n t^n$, where the coefficients a_n and the variable t are complex. A *Laurent series* is of the form $\sum_{n \geq -\infty} a_n t^n$, where the coefficients a_n and the variable t are complex.

Theorem D.0.13 *(Abel's theorem) For every power series $\sum_{n=0}^{\infty} a_n t^n$, there exists a number $\rho \in [0, \infty]$, called the radius of convergence such that the series converges absolutely for every $|t| < \rho$. In $|t| < \rho$ the sum of the series is an analytic function. The derivative can be obtained by termwise differentiation, and the derived series has the same radius of convergence.*

Theorem D.0.14 *(Hadamard's formula) The radius of convergence of a power series $\sum_{n=0}^{\infty} a_n t^n$ is given by*

$$\rho = \frac{1}{\limsup_{n} |a_n|^{\frac{1}{n}}}$$

Theorem D.0.15 *The function $x : D \longrightarrow \mathbb{C}$ is holomorphic if and only if for each $t_o \in D$, there exists a ball with center t_o and radius δ, $B(t_o, \delta)$ such that on this ball x coincides to the sum of a power series,*

$$x(t) = \sum_{n=0}^{\infty} a_n (t - t_o)^n, \quad \forall t \in B(t_o, \delta)$$

For the holomorphic functions defined on

$$\Omega(t_o; r, R) = \{t \in B(t_o, R) \mid r < t < R\},$$

where $0 \leq r < R$, we have the next result of global representation in Laurent series.

Theorem D.0.16 ***(Laurent's development theorem)*** *Let x be an holomorphic function on $\Omega(t_o; r, R)$. Then, there exists a (unique) Laurent series converging on a set that contains $\Omega(t_o; r, R)$ and that coincides to x on $\Omega(t_o; r, R)$,*

$$x(t) = \sum_{n=-\infty}^{\infty} a_n (t - t_o)^n, \quad \forall t \in \Omega(t_o; r, R)$$

Theorem D.0.17 ***(Liouville's theorem)*** *Let x be a bounded holomorphic function defined on the whole complex $\mathbb{C}$. Then, x is necessarily constant.*

Bibliography

[1] Akilov, G.P., and Kantorovici, L.V., *Functional Analysis* (translated in romanian from russian), Editura Stiintifica si Enciclopedica, 1977.

[2] Bourbaki, N., *General Topology* (2 vols), Herman, Paris, and Addison-Wesley, Reading Mass., 1966.

[3] Colojoara, I., and Foias, C., *Theory of generalized spectral operators*, Gordon and Breach, New-York, 1968.

[4] Cristescu, R., *Topological Vector Spaces*, Noordhoff Int. Publ. Leyden, The Netherlands, 1977.

[5] Cristescu, R., *Notions of Linear Functional Analysis* (in romanian), Editura Academiei Române, Bucuresti,1998.

[6] Dunford, N., and Schwartz, J.T., *Linear Operators*, Part I: *General Theory*, Interscience, New York, 1958.

[7] Dunford, N., and Schwartz, J.T., *Linear Operators*, Part II: *Spectral Theory*, Interscience, New York, 1963.

[8] Farkas, W., and Pavel, L., *Functional Analysis-Exercises and Problems* (in romanian), Editura Universitatii Bucuresti, 1994.

[9] Folland, G., *Real Analysis*, Wiley-Interscience, 1984.

[10] Gheorghiu, N., *Introduction to functional analysis* (in romanian), Editura Academiei, Bucuresti, 1974.

[11] Gohberg, I., and Goldberg, S., *Basic Operator Theory*, Birkhäuser Boston, 1981.

[12] Halmos, P.R., *Measure Theory*, Van Nostrand, Princeton, N.J., 1950.

[13] Niculescu, C., and Popa, N., *Elements of Banach spaces theory* (in romanian), Editura Academiei, Bucuresti, 1981.

[14] Pedersen, G., *Analysis Now*, Springer-Verlag, New-York Berlin Heidelberg, 1989.

[15] Reed, M., and Simon, B., *Methods of modern mathematical physics*, Vol. I: *Functional Analysis*, Academic Press, New York, 1980.

[16] Rudin, W., *Functional Analysis*, McGraw-Hill, 1973.

[17] Schaefer, H.H., *Topological vector spaces*, Springer-Verlag, New-York Berlin Heidelberg, 1971.

[18] Taylor, A.E., and Lay, D.C., *Introduction to Functional Analysis*, 2nd Ed.,Wiley, New York, 1980.

[19] DeVito, C., *Functional Analysis,* Academic Press, New York, 1978.

[20] Yosida, K., *Functional Analysis,* 2nd Ed., Springer-Verlag, Berlin and New-York, 1968.

Index

Tiparul s-a executat sub cda 662/1999
la Tipografia Editurii Universității din București

DATA RESTITUIRII

16. OCT. 2002	15 IAN 2007	
30. OCT. 2002		
20 IAN. 2[illegible]	05 DEC. 2018	
	06 DEC. 2019	
23 IAN. 2003	06 DEC. 2019	
29 IAN. 2003		
30 IAN. 2003		
28. MAI. 200[illegible]		
28. MAI. 200[illegible]		
23. IAN. 2004		
15. IUN. 2005		
20. NOV. 20[illegible]		

FUNDAMENTUM

This elementary text is an introduction to functional analysis. The book covers only a limited number of topics, but they are sufficient to lay a foundation in functional linear analysis, which, partly because of its many applications has become a very popular mathematical discipline interesting for applied mathematicians, probabilists, classical and numerical analysts. It grew out of my attempts to present the material in a way that was interesting and understandable to second-third year graduate students who are taking a course in this subject.

The only background material needed is what is usually covered in a one-year graduate level course analysis and an acquittance with linear algebra. However, to reach as large an audience as possible, the material is generally self-contained: any lack of knowledge can be compensated for by referring to Preliminaries, to Appendices and the references therein.

ISBN 973-575-413-4 Lei 33400