Krein-Milman Theorem and its applications

Bachelor's degree final project

Author:
Antonio Zarauz Moreno

Advisor:
Dr. Juan Carlos Navarro Pascual

Degree in Mathematics

July, 2016
University of Almeria

Contents

Abstract in English

The present degree's project intends to be a brief guide for those who want to begin a path in convex set theory and functional analysis. This ambitious project has also been motivated by the achievement of Collaboration's scholarship for the student and it is a first approach to the contents that have been developed during the whole academic course.

The title of this work, *Krein-Milman theorem and its applications*, contains the essence of our purpose, which is delving into the previous theorem and its requirements: topological vector spaces, convexity, measure theory, and so on and so forth. The reference of the original paper can be consulted in [1].

The first chapter introduces clearly the essential concepts of convex analysis, and also the main ideas of this topic in finite-dimensional spaces, due to Carathéodory and Minkowski. In that environment, the writer has included some graphic examples (and counterexamples) which are supposed to be profitable for the reader. The structure of the chapter and the obtaining of the main results has been elaborated via [2], [3] and [4]. Specifically:

- In the first section we have used [2] to introduce the definitions, whereas the rest of result has been developed by the student.

- In the second section lies one of the most important theorems of the chapter (Carathéodory's theorem), and it also incorporates a long list of remarks to show the importance and the improvable facts of the theorem. In the following section, we discuss the notion of dimension of a convex set and its relative interior. To write these sections, the writer has selected [4].

- To conclude the chapter, sections four and five comprise concepts of paramount importance for the next chapter, such as extreme point, face or exposed point. One can highlight the canonical way to build faces via continuous linear functionals, the existence of extreme points for compact convex sets, the Carathéodory-Minkowski's theorem and one of the most beautiful applications in finite-dimensional theory, which is the existence of extreme values of linear functionals over compact convex sets in extreme points of the domain. To elaborate this sections we have consulted [3].

The second chapter is devoted to the exposition of Krein-Milman theorem, giving a wide introduction to the infinite-dimensional spaces via several examples of canonical spaces. Specifically,

- The enriching list of examples of extreme points has been developed in order to get the reader used to the main strategies which lies into this theory; the sixth example is inspired by [9] and [13].

- Before getting into Krein-Milman theorem, there are some considerations about the origin of the main algebraic concepts involved in this theory, which are balancedness, absorbency and the own definition of convexity, through the algebraic

characterisation of unit balls in seminormed spaces (it has been required [14]). Furthermore, since compactness is one of the most cultivated concepts in General Topology, it is reasonable to study the convex hull of a compact set, giving a solution to that problem in an infinite-dimensional context.

- The proof of the mentioned theorem is based on [6]. Some interesting remarks are made after the theorem, which delve into the conditions applied on the theorem and the notation of some other authors. Finally, it will be discussed one more detail about compactness, which reflects the fact that those sets contain every extreme point of the closure of their convex hull, i.e., Milman theorem (also based on [6]).

- The applications of Krein-Milman theorem has been adapted from [6] (necessary condition for duality, representation theorem), [12] (Stone-Cech compactification), [13] (Banach-Stone theorem) and [3] (Stone-Weierstrass theorem).

Resumen en español

El presente trabajo de fin de grado pretende ser una breve guía para aquellos que quieren dar unos primeros pasos en la teoría de conjuntos convexos y análisis funcional. Este ambicioso proyecto ha sido motivado por la obtención de la beca de Colaboración por parte del alumno, y es una primera aproximación a los contenidos que han sido desarrollados durante el curso académico.

El título de este trabajo, *El teorema de Krein-Milman y sus aplicaciones*, contiene la esencia de nuestro objetivo, que consiste en analizar en profundidad el anterior teorema y todos sus prerrequisitos: espacios vectoriales topológicos, convexidad, teoría de la medida, etcétera. La referencia original puede ser consultada en [1].

El primer capítulo introduce de forma clara los conceptos esenciales del análisis convexo, y también las principales ideas de esta materia en espacios finito-dimensionales, debidos a Carathéodory y Minkowski. En este contexto, el se han incluido numerosos ejemplos gráficos (y contraejemplos) que serán productivos para el lector. La estructura del capítulo y la obtención de los principales resultados ha sido elaborada a partir de [2], [3] and [4]. Concretamente:

- En la primera sección se ha usado [2] para introducir definiciones, mientras que el resto de resultados han sido desarrollados por el estudiante.
- La segunda sección alberga uno de los teoremas más importantes del capítulo (teorema de Carathéodory) y también incorpora una larga lista de observaciones para ensalzar su importancia y los detalles mejorables del mismo. En la siguiente sección, se discute la noción de dimensión de un conjunto convexo y su interior relativo. Para escribir estas secciones, se ha escogido [4].
- Para concluir el capítulo, las secciones cuatro y cinco comprenden conceptos de primordial importancia para el siguiente capítulo, tales como punto extremo, cara de un conjunto convexo o punto expuesto. Resaltamos además la forma canónica de construir caras mediante funcionales lineales y continuos, la existencia de puntos extremos para conjuntos compactos y convexos, el teorema de Carathéodory-Minkowski y una de las más relucientes aplicaciones en un contexto finito-dimensional, que es la existencia de valores extremos en funcionales lineales y continuos sobre compactos convexos. Para elaborar estas secciones, hemos consultado [3].

El segundo capítulo está dedicado a la exposición del teorema de Krein-Milman, dando una amplia introducción a los espacios infinito-dimensionales mediante ejemplos sobre espacios canónicos. Concretamente,

- La enriquecedora lista de ejemplos sobre puntos extremos ha sido desarrollada para introducir al lector en las principales estrategias que subyacen en esta teoría; el sexto ejemplo está inspirado en [9] y [13].
- Antes de adentrarnos en el teorema de Krein-Milman, hacemos algunas consideraciones sobre el origen de las principales definiciones algebraicas tratadas en

el tema, tales como equilibrio, absorbencia y el propio concepto de convexidad, a través de la caracterización algebraica de bolas unidad en espacios seminormados (ha sido requerido [14]). Además, dado que la compacidad es una de las más cultivadas de la Topología General, es razonable estudiar la envolvente convexa de conjuntos compactos, dando una solución contundente a tal problema en espacios de dimensión infinita.

- La demostración del mencionado teorema está basada en [6]. Algunas observaciones oportunas son propuestas tras el teorema, que profundizan en las condiciones aplicadas en la hipótesis del teorema y la notación de algunos autores. Finalmente, será considerado un detalle adicional sobre compacidad que refleja el hecho de que tales conjuntos contienen cada puntos extremo del cierre de su envolvente convexa, i.e., el teorema de Milman (también basado en [6]).

- Las aplicaciones del teorema de Krein-Milman han sido adaptadas de [6] (condición necesaria de dualidad, teorema de representación), [12] (compactificación de Stone-Cech), [13] (teorema de Banach-Stone) y [3] (teorema de Stone-Weierstrass).

1 Basic concepts and finite-dimensional theory

The first chapter will be devoted to the exposition of several elementary notions related to convex and functional analysis. We start our path to Krein-Milman theorem proving its famous finite-dimensional preceding; i.e., Carathéodory-Minkowski theorem.

Recall that a topological vector space is a pair (X,τ) where X is a vector space over the field $\mathbb{K}=\mathbb{R}\vee\mathbb{C}$, and τ is a compatible topology with the vector structure in X; that is, the maps $(x,y)\mapsto x+y$ and $(\alpha,x)\mapsto\alpha x$ are continuous from $X\times X$ onto X and from $\mathbb{K}\times X$ onto X respectively, considering the product topology in each space.

Secondly, a normed space is a pair $(X,\|\cdot\|)$ where X is a vector space and $\|\cdot\|$ a norm in X. Since the topology induced by the norm is compatible with the vector structure, normed space form a strongly relevant example of topological vector spaces. There also are other structures which are compatible with the norm, such as the weak topology of a normed space X, denoted by ω, and the weak-star topology, written as ω^*. As usual, we write X instead of (X,τ) or $(X,\|\cdot\|)$ when we are making reference to a topological vector space or a normed space, respectively.

Let n be a natural number and X a Hausdorff topological vector space with $\dim(X)=n$. Then, every linear bijection from $\mathbb{K}^n$ onto X is bicontinuous, hence X is isomorphic as a vector space to $\mathbb{K}^n$ and homeomorphic as a topological vector space to the Euclidean space. However, the notation X for finite-dimensional spaces will be used during the whole chapter, since it will make easier the step of abstraction given in the following chapters.

1.1 Convex sets and maps

Definition 1.1. *In a vector space X over $\mathbb{K}$, a subset $A\subset X$ is convex if, given $x,y\in A$ and $t\in[0,1]$,*

$$\{tx+(1-t)y\colon t\in]0,1[\}\subset A. \tag{1.1}$$

Definition 1.2. *If $A\subset X$ is a convex set, a function $f:A\to\mathbb{R}$ is said to be convex (resp. concave) if the following inequality holds for each $x,y\in A$ and $t\in[0,1]$:*

$$f(tx+(1-t)y)\le tf(x)+(1-t)f(y)\quad (resp.\, f(tx+(1-t)y)\ge tf(x)+(1-t)f(y)). \tag{1.2}$$

If the inequalities 1.2 are strict, f is strictly convex (resp. strictly concave), and if it is an equality in both cases, f is affine.

The following examples can be easily checked by the reader.

Examples 1.1.

- *Any segment (either open or closed) is a convex set:*

$$S_{x,y}=\{tx+(1-t)y\colon t\in[0,1]\},\ \overset{\circ}{S}_{x,y}=\{tx+(1-t)y\colon t\in]0,1[\}$$

where x,y are arbitrary points of a vector space.

- *Any hyperplane is a convex set*

$$H = \{x \in X \colon f(x) = \lambda\},$$

where $f : X \to \mathbb{R}$ *is a linear functional on a vector space and* $\lambda \in \mathbb{R}$*, and every half-space*

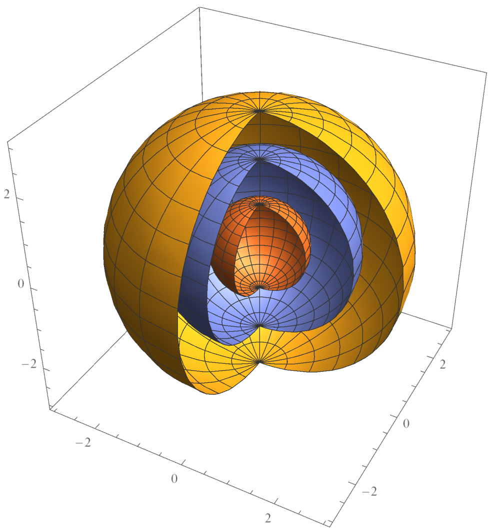
$$H_1 = \{x \in X \colon f(x) \leq \lambda\}, H_2 = \{x \in X \colon f(x) \geq \lambda\}$$

is also a convex set.

- *Any ball (either open or closed) in a normed space* X *is a convex set*

$$B(x,r) = \{y \in X \colon \|y - x\| < r\}, \overline{B}(x,r) = \{y \in X \colon \|y - x\| < r\},$$

Actually, every set with the form $C = F \cup B(x,r)$, $F \subset \partial\overline{B}(x,r) := S(x,r)$ *is convex.*



- *Let* X *be a vector space and consider* $A \subset X$*. The convex hull of* A *is the intersection of all the convex subsets of* X *containing* A*. It is clear that* $A \subset co(A)$ *and* $co(A)$ *is convex (see proposition 1.1); in addition, we state that* $co(A)$ *admits the expression*

$$co(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i \colon n \in \mathbb{N}, x_i \in A, \lambda_i \in \mathbb{R}_0^+, \forall i = \{1, \ldots, n\}, \sum_{i=1}^{n} \lambda_i = 1 \right\}, \tag{1.3}$$

In fact, let V *be the set in the right side of the equation (1.3).*

To verify that $co(A) \subset V$ *it is enough to show that* V *is convex (note that* $A \subset V$*). Given* $a = \sum_{i=1}^{n} \alpha_i y_i$ *and* $b = \sum_{i=1}^{m} \beta_i z_i$ *elements in* V *(suppose without loss of generality that* $n \leq m$*), for each* $t \in [0,1]$ *we make the change of variables*

$$\lambda_i = \begin{cases} (1-t)\alpha_i & \text{if} \quad i = 1, \ldots, n \\ t\beta_i & \text{if} \quad i = n+1, \ldots, n+m \end{cases}, x_i = \begin{cases} (1-t)y_i & \text{if} \quad i = 1, \ldots, n \\ tz_i & \text{if} \quad i = n+1, \ldots, n+m \end{cases}$$

so we obtain $(1-t)a+tb = \sum_{i=1}^{n+m} \lambda_i x_i$ *where* $\sum_{i=1}^{n+m} \lambda_i = 1$ *and* $x_i \in A$ *for every* $i = 1,\ldots,n+m$.

The other inclusion will be proved by induction. For $n = 1$ *it is clear, so long as* $A \subset co(A)$ *and* $co(A)$ *is a convex set. Suppose that the statement holds for* $n \in \mathbb{N}$ *and let* $x = \sum_{i=1}^{n+1} \lambda_i x_i$. *If* $\lambda_{n+1} = 0 \vee \lambda_{n+1} = 1$ *is straightforward. Otherwise,*

$$x = (1-\lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i + \lambda_{n+1} x_{n+1},$$

which is a convex combination of $\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i$ *and* x_{n+1}, *both elements of* A *by induction hypothesis.*

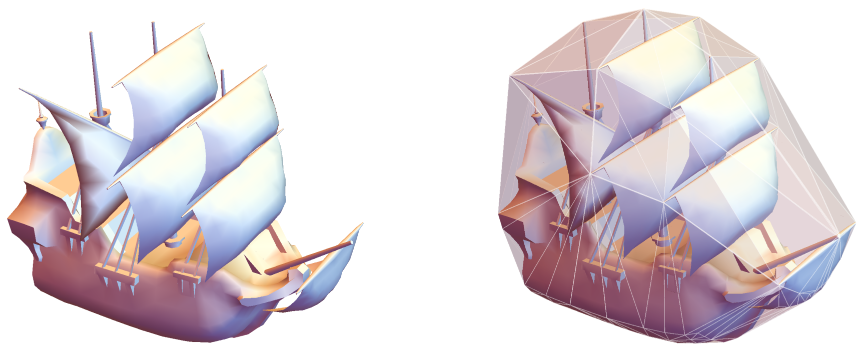


Figure 1.1: Convex hull of a galleon.

- *In particular, if* A *is a finite union of convex sets; i.e.,* $A = \cup_{i=1}^{n} A_i$ *with* A_i *convex for every* $i = 1,\ldots,n$, *then one can choose every point of the previous convex linear combination in each* A_i:

$$co(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i,\ x_i \in A_i,\ \lambda_i \in \mathbb{R}_0^+,\ \forall i \in \{1,\ldots,n\},\ \sum_{i=1}^{n} \lambda_i = 1 \right\}. \tag{1.4}$$

First of all, the reader should appreciate that the number n *is fixed under these circumstances. It is easy to check that the set*

$$E = \left\{ \sum_{i=1}^{n} \lambda_i x_i,\ x_i \in A_i,\ \lambda_i \in \mathbb{R}_0^+,\ \forall i \in \{1,\ldots,n\},\ \sum_{i=1}^{n} \lambda_i = 1 \right\}$$

is convex. On the one hand, the inclusion $A \subset E$ *and the convexity of* E *implies that* $co(A) \subset E$. *On the other hand, the previous example shows that* $E \subset co(A)$.

- *In a similar way, we define the real affine hull of* $A \subset X$ *as*

$$aff(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i :\ n \in \mathbb{N},\ x_i \in A,\ \lambda_i \in \mathbb{R},\ \forall i \in \{1,\ldots,n\},\ \sum_{i=1}^{n} \lambda_i = 1 \right\}, \tag{1.5}$$

and it also verifies that it is the least affine space which contains A, and $aff(A) = aff(co(A))$.

- *Every affine map is linear-convex by definition. Reciprocally, we can suppose that $t < 0$ without losing generality (if $t > 1$ we can interchange the role of x and y). Then,*

$$f(y) = f\left(\underbrace{\frac{1}{1-t}}_{\in[0,1]}(tx+(1-t)y) + \left(1 - \frac{1}{1-t}\right)x\right).$$

Using the affinity of f, this reduces to

$$f(y) = \frac{1}{1-t}f(tx+(1-t)y) + \left(1-\frac{1}{1-t}\right)f(x) \Rightarrow f(tx+(1-t)y) = tf(x)+(1-t)f(y).$$

Proposition 1.1. *Let $\mathcal{C}$ the family of all the convex sets of a vector space X. Then,*

1. *Whatever $\{C_i\}_{i\in I} \subset \mathcal{C}$ is $\cap_{i\in I} C_i \in \mathcal{C}$.*
2. *$\mathcal{C}$ satisfies that $A+B \in \mathcal{C}$, $\lambda A \in \mathcal{C}$ for all $A, B \in \mathcal{C}$ and $\lambda \in \mathbb{R}$. In addition, $(\lambda+\mu)A = \lambda A + \mu A$ for every $\lambda, \mu \in \mathbb{R}$ such that $\lambda\mu \geq 0$.*[1]
3. *$co(\cdot) : X \to \mathcal{C}$ is a monotone and additive operator.*
4. *A is convex iff $A = co(A)$.*

Proof. 1. Given $x, y \in \cap_{i\in I} C_i$, since $x, y \in C_i$ for all $i \in I$, $tx+(1-t)y \in C_i$ for all $i \in I$ and $t \in [0,1]$. Hence $tx + (1-t)y \in \cap_{i\in I} C_i$ for every $t \in [0,1]$.

2. Let $x, y \in A+B$ and $t \in [0,1]$. We can express $x = a_x + b_x$, $y = a_y + b_y$ with $a_x, a_y \in A$, $b_x, b_y \in B$. Then,

$$tx+(1-t)y = t(a_x+b_x)+(1-t)(a_y+b_y) = [ta_x+(1-t)a_y]+[tb_x+(1-t)b_y] \in A+B, \forall t \in [0,1].$$

Furthermore, for every $\lambda \in \mathbb{R}$ and $x, y \in A$,

$$t(\lambda x) + (1-t)(\lambda y) = \lambda[tx+(1-t)y] \in \lambda A, \forall t \in [0,1].$$

The last statement is checked as follows: the implication $(\lambda+\mu)A \subset \lambda A + \mu A$ is clear thanks to the distributive law in $\mathbb{K}$; reciprocally, it is straightforward when $\lambda = 0 \vee \mu = 0$. Otherwise, let $\lambda a \in \lambda A$, $\mu a' \in \mu A$, then

$$\lambda a + \mu a' = \frac{\lambda}{\lambda+\mu}(\lambda+\mu)a + \frac{\mu}{\lambda+\mu}(\lambda+\mu)a'.$$

Since the previous equality is a convex combination of elements in $(\lambda+\mu)A$, the result lies in that set, hence $\lambda A + \mu A \subset (\lambda+\mu)A$.

[1] Indeed, one can characterise $\mathcal{C}$ through this property, in an even easier way than the implication we have already proved. In fact, if A is not convex, we can find $x \in X$, $\lambda = \mu = \frac{1}{2}$ satisfying that $x \in A$ but $x \notin \frac{1}{2}A + \frac{1}{2}A$.

3. The monotony derives from the definition. To see that $\mathrm{co}(A+B) \subset \mathrm{co}(A)+\mathrm{co}(B)$, given $x \in \mathrm{co}(A+B)$, there exist $\{\lambda_i\}_{i=1}^k \subset \mathbb{R}_0^+$, $\{a_i\}_{i=1}^k \subset A$ and $\{b_i\}_{i=1}^k \subset B$ such that $\sum_{i=1}^k \lambda_i = 1$ and

$$x = \sum_{i=1}^{k} \lambda_i (a_i + b_i) = \sum_{i=1}^{k} \lambda_i a_i + \sum_{i=1}^{k} \lambda_i b_i \in \mathrm{co}(A) + \mathrm{co}(B).$$

On the other hand, given $a \in \mathrm{co}(A)$ and $b \in \mathrm{co}(B)$, there exists $\{\lambda_i\}_{i=1}^k, \{\mu_j\}_{j=1}^m \subset \mathbb{R}_0^+$, $\{a_i\}_{i=1}^k \subset A$ and $\{b_j\}_{j=1}^m \subset B$ such that $\sum_{i=1}^k \lambda_i = \sum_{j=1}^m \mu_j = 1$ and

$$a + b = \underbrace{\sum_{j=1}^{m} \mu_j}_{1} \left(\sum_{i=1}^{k} \lambda_i a_i \right) + \underbrace{\sum_{i=1}^{k} \lambda_i}_{1} \left(\sum_{j=1}^{m} \mu_j b_j \right) = \sum_{i,j} \lambda_i \mu_j (a_i + b_j) \in \mathrm{co}(A+B)$$

since $\sum_{i,j} \lambda_i \mu_j = \left(\sum_{i=1}^k \lambda_i\right)\left(\sum_{j=1}^m \mu_j\right) = 1$ and $\lambda_i \mu_j \geq 0$ for each i, j.

4. $\Rightarrow$) Since A is convex and $A \subset A$, we have that $\mathrm{co}(A) \subset A$.

$\Leftarrow$) It is clear that $\mathrm{co}(A)$ is convex, hence A is convex too. □

Remarks 1.1.

- *The union of convex sets is not necessarily a convex set, as the next figure based on examples 1.1 shows:*

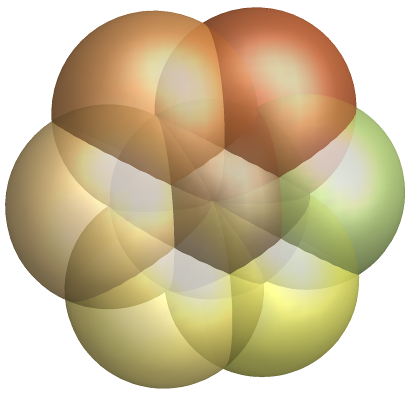


Figure 1.2: Counterexample for the union of convex sets.

- *The intersection of hyperplanes gives us the solution of a linear system of equations, and the intersection of half-spaces gives us a polyhedron; a bounded polyhedron is called a polytope. In particular, the k-simplex ($k \leq N+1$) determined by a set of affine-independent points $\{x_i\}_{i=1}^k \subset X$ is $\Delta^k\left(\{x_i\}_{i=1}^k\right) = co\left(\{x_i\}_{i=1}^k\right)$ (figure 1.3).*

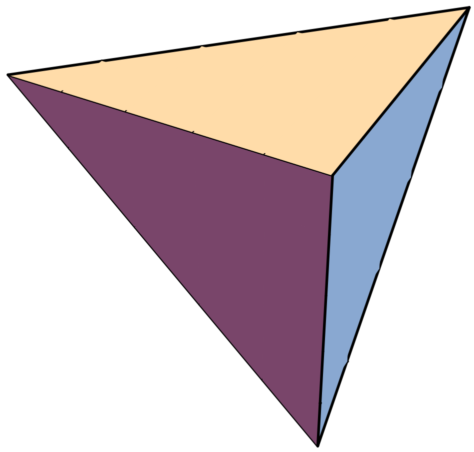


Figure 1.3: Example of 4-simplex in $X = \mathbb{R}^3$ (tetrahedron).

1.2 Carathéodory's theorem

Let X be a vector space. The definition of convex hull of a subset $A \subset X$ brings us a characterisation of convex sets through proposition 1.1. However, there is no limit in the number of elements involved in the representation of each $x \in \mathrm{co}(A)$. In this sense, Carathéodory's theorem states that every point $x \in \mathrm{co}(A)$ can be expressed with $n+1$ points of A as much.

The following lemma shows, in particular, the highest number of linear-independent elements.

Lemma 1.1. *A set of points $\{x_i\}_{i=1}^k \subset X$ is affine-dependent iff there exists $\{\lambda_i\}_{i=1}^k \in \mathbb{R}$ such that $\sum_{i=1}^k \lambda_i = 0 < \sum_{i=1}^k |\lambda_i|$ and $\sum_{i=1}^k \lambda_i x_i = 0$.*

Proof. Since $\{x_i\}_{i=1}^k \subset X$ is affine-dependent, we have that $\{x_j - x_1\}_{j=2}^k$ is linear-dependent as a set of vectors, so there exists $\{\alpha_j\}_{j=2}^k$ with $\sum_{j=2}^k |\alpha_j| > 0$ and

$$0 = \sum_{j=2}^k \alpha_j (x_j - x_1) = \left(-\sum_{j=2}^k \alpha_j \right) x_1 + \sum_{j=2}^k \alpha_j x_j \left(\left\{ \begin{array}{lcl} \lambda_1 & = & -\sum_{j=2}^k \alpha_j \\ \lambda_j & = & \alpha_j, \, j = 2, \ldots, k \end{array} \right. \right) = \sum_{i=1}^k \lambda_i x_i.$$

The collection of numbers $\{\lambda_i\}_{i=1}^k$ achieves what we desired. $\square$

Theorem 1.1 (Carathéodory)**.** *If $\dim(X) = n$, $A \subset X$ and $x \in co(A)$, then x is a convex combination of affine-independent points from A (in particular, $n+1$ as much).*

Proof. Let $x \in \mathrm{co}(A)$ such that

$$x = \sum_{i=1}^{k} \lambda_i x_i$$

with $\{x_i\}_{i=1}^k \subset A$, $\{\lambda_i\}_{i=1}^k \subset \mathbb{R}_0^+$ and $\sum_{i=1}^k \lambda_i = 1$ to be the shortest expression of x in terms of elements of A. By *reductio ad absurdum*, suppose that $\{x_i\}_{i=1}^k$ are affine-dependent. The previous lemma 1.1 shows that there exists $\{\alpha_i\}_{i=1}^k \subset \mathbb{R}$ satisfying

$$\sum_{i=1}^{k} \alpha_i = 0 < \sum_{i=1}^{k} |\alpha_i|, \sum_{i=1}^{k} \alpha_i x_i = 0.$$

It can also be considered, without losing generality, that

$$\frac{\lambda_k}{\alpha_k} = \min_{i=1,\dots,k} \left\{ \frac{\lambda_i}{\alpha_i} : \alpha_i > 0 \right\},$$

the objective now is looking for a linear combination of x in terms of $\{x_i\}_{i=1}^{k-1}$ to find a contradiction:

$$x = \sum_{i=1}^{k} \lambda_i x_i = \sum_{k=1}^{k-1} \left(\lambda_i - \frac{\lambda_k}{\alpha_k}\alpha_i \right) x_i + \underbrace{\sum_{i=1}^{k-1} \frac{\lambda_k}{\alpha_k}\alpha_i x_i + \lambda_k x_k}_{0} = \sum_{k=1}^{k-1} \left(\lambda_i - \frac{\lambda_k}{\alpha_k}\alpha_i \right) x_i.$$

Calling $\xi_i = \lambda_i - \frac{\lambda_k}{\alpha_k}\alpha_i$ for each $i = 1, \dots, k-1$, it is clear that $\xi_i \geq 0$ because of the assumption over $\frac{\lambda_k}{\alpha_k}$. Finally,

$$\sum_{i=1}^{k-1} \xi_i = \left(\lambda_k + \sum_{i=1}^{k-1} \lambda_i \right) - \left(\lambda_k + \sum_{i=1}^{k-1} \frac{\lambda_k}{\alpha_k}\alpha_i \right) = 1 - \frac{\lambda_k}{\alpha_k} \underbrace{\sum_{i=1}^{k} \alpha_i}_{0} = 1.$$

□

In spite of its usefulness, this result does not give any information about the points we select to express some $x \in \mathrm{co}(A)$. In fact, as we have already seen, this result is valid for every vector space (with no topological structure). However, it would be desirable to obtain a more powerful result with the aid of a suitable structure.

The next step in the process will be the choice of a *reduced* group of points P of a set A satisfying $\mathrm{co}(P) = A$. That is a first approach to what Krein-Milman theorem will state in next chapter for locally convex topological vector spaces:

$$\overline{\mathrm{co}(P)} = A.$$

The mentioned set A would have to be compact and convex. There are some reasons why it is fixed the attention in that kind of sets:

- Convexity's hypothesis is clear, since our purpose is the *reconstruction* of the set A through its convex hull (or its closed convex hull if necessary) of a distinguished subset P of A.
- Compactness is also required, since one can find examples of closed and bounded convex sets which has no extreme points (see subsection 2.1). As well as in many other branches of Mathematics, this is a convenient hypothesis to ensure the existence of the previous set $P \subset A$ satisfying the desired condition.

In the rest of the section, we will assume that X is a finite-dimensional topological (only required for proposition 1.2) vector space; this is enough to prove Carathéodory-Minkowski theorem, even if the previous results can be discussed in more general structures.

The next results show the topological properties of $\mathrm{co}(\cdot)$ as an operator over $\mathcal{C}$.

Lemma 1.2. *Given $A \subset X$ convex, $x \in int(A)$ and $y \in \overline{A}$, is $\overset{\circ}{S}_{x,y} \subset int(A)$.*

Proof. Let $t \in]0,1[$ be fixed; we have to show that $tx + (1-t)y \in \mathrm{int}(A)$. By translation if necessary we can assume that $tx + (1-t)y = 0$, in particular $y = \alpha x$ where $\alpha < 0$. Since the mapping $\omega \mapsto \alpha\omega$ is a homeomorphism of X and $x \in \mathrm{int}(A)$, $y \in \overline{A}$, there exists $z \in \mathrm{int}(A)$ such that $\alpha z \in A$.

Let $\mu = \frac{\alpha}{\alpha-1}$; then $\mu \in]0,1[$ and

$$\mu z + (1-\mu)\alpha z = 0.$$

Then, the set

$$U = \{\mu\omega + (1-\mu)\alpha z \colon \omega \in \mathrm{int}(A)\}$$

is a 0-neighbourhood, as long as $\omega \mapsto \mu\omega + (1-\mu)\alpha z$ is a homeomorphism of X mapping $z \in \mathrm{int}(A)$ onto 0. But $\omega \in \mathrm{int}(A)$ and $\alpha z \in A$ imply that $U \subset A$ for being A convex, and $0 \in \mathrm{int}(A)$. □

Proposition 1.2. *Given a set $A \subset X$, we have:*

1. *$int(A)$ and $\overline{A}$ are convex sets if A is convex.*
2. *$co(\cdot)$ maps open sets into open sets.*
3. *$co(\cdot)$ maps bounded sets into bounded sets.*
4. *$co(\cdot)$ maps compact sets into compact sets.*[2]
5. *$co(\cdot)$ maps precompact sets into precompact sets.*
6. *If A is convex and $int(A) \neq \emptyset$, then $int(A) = int(\overline{A})$ and $\overline{int(A)} = \overline{A}$.*

Proof. 1. For any $x, y \in \mathrm{int}(A)$, we have that $\overset{\circ}{S}_{x,y} \subset \mathrm{int}(A)$ by lemma 1.2, so $S_{x,y} \subset \mathrm{int}(A)$ and $\mathrm{int}(A)$ is convex. Furthermore, given $x, y \in \overline{A}$ and $t \in [0,1]$, there exist sequences

[2]Here is required that X is a finite-dimensional space.

$\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}$ in A which converge to x and y respectively. By convexity of A, the family of sequences

$$z_{n,t} = tx_n + (1-t)y_n, \quad n \in \mathbb{N},$$

belong to A and converge to $tx + (1-t)y$ for each $t \in [0,1]$. Hence $S_{x,y} \subset \overline{A}$ and $\overline{A}$ is convex.

2. Consider $z \in \mathrm{co}(A)$. Then exist $\{\lambda_i\}_{i=1}^k \subset \mathbb{R}_0^+$ and $\{x_i\}_{i=1}^k \subset A$ satisfying

$$\sum_{i=1}^k \lambda_i = 1, z = \sum_{i=1}^k \lambda_i x_i.$$

Since A is open, there are $\{\delta_i\}_{i=1}^k \subset \mathbb{R}^+$ such that $B_i := B(x_i, \delta_i) \subset A$ for every $i = 1, \ldots, k$. Calling $\delta := \min_{i=1,\ldots,k}\{\delta_i\}$, it is clear that

$$B(z,\delta) \subset \sum_{i=1}^k \lambda_i B_i \subset \mathrm{co}(A).$$

3.Let $M \in \mathbb{R}$ such that $\|x\| \leq M$. Then choosing $y \in \mathrm{co}(A)$, there exist $\{x_i\}_{i=1}^k$ in A and $\{\lambda_i\}_{i=1}^k$ satisfying $y = \sum_i \lambda_i x_i$. Using the triangle inequality we conclude that $\|y\| \leq M$.
4. Let $n = \dim(X)$ and consider the map

$$F : [0,1]^{n+1} \times A^{n+1} \longrightarrow X$$

given by

$$F(\lambda_1, \ldots, \lambda_{n+1}, x_1, \ldots, x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i x_i$$

It is clear that

$$\Gamma = \{(\lambda_1, \ldots, \lambda_{n+1}) \in [0,1]^{n+1} : \sum_{i=1}^{n+1} \lambda_i = 1\}$$

is compact, so $\Gamma \times A^{n+1}$ is compact in $[0,1]^{n+1} \times A^{n+1}$. Applying theorem 1.1, $F(\Gamma \times A^{n+1}) = \mathrm{co}(A)$. Since F is continuous, $\mathrm{co}(A)$ is compact too.

5. Given $A \subset X$ precompact and $\varepsilon \in \mathbb{R}^+$, there exists a finite set $\mathcal{S} \subset A$ satisfying

$$A \subset \bigcup_{x\in\mathcal{S}} B(x,\varepsilon).$$

Using the previous result[3], $\mathrm{co}(\mathcal{S})$ is compact and $\mathrm{co}(A) \subset \mathrm{co}(\mathcal{S}) + B(x,\varepsilon)$ since $\mathrm{co}(\mathcal{S}) + B(x,\varepsilon)$ is convex and contains A. Hence it can be found a finite set $\mathcal{S}_1 \subset \mathrm{co}(\mathcal{S})$ such that

$$\mathrm{co}(\mathcal{S}) = \bigcup_{x\in\mathcal{S}_1} B\left(x, \frac{\varepsilon}{2}\right).$$

Now it follows that $\mathrm{co}(A) \subset \cup_{x\in\mathcal{S}_1} B(x,\varepsilon)$, showing that $\mathrm{co}(A)$ is precompact.

[3]It can be used, so long as $\mathcal{S}$ is compact and $\dim(\mathrm{co}(\mathcal{S})) < +\infty$.

6. The inclusion $\operatorname{int}(A) \subset \operatorname{int}(\overline{A})$ is trivial. On the other hand, given $z \in \operatorname{int}(\overline{A})$ and $x \in \operatorname{int}(A)$ with $z \neq x$ (if $z = x$ is obvious), consider $r > 0$ such that $B(z,r) \subset \overline{A}$ and the point

$$\omega = z + \frac{r}{2}\frac{z-x}{\|z-x\|} \in B(z,r) \subset \overline{A}. \tag{1.6}$$

Using lemma 1.2 is $\overset{\circ}{S}_{x,\omega} \subset \operatorname{int}(A)$. Solving the equation 1.6 for z we have $z = tx + (1-t)\omega$ where $t = \frac{r}{r+2\|z-x\|} \in]0,1[$ and $z \in \overset{\circ}{S}_{x,\omega} \subset \operatorname{int}(A)$.

To prove the other equality, it is clear that $\overline{\operatorname{int}(A)} \subset \overline{A}$. Reciprocally, given $x \in \operatorname{int}(A)$ and $z \in \overline{A}$, is $\overset{\circ}{S}_{x,z} \subset \operatorname{int}(A)$ by lemma 1.2. Hence, taking any sequence $\{x_i\}_{i\in\mathbb{N}} \subset \overset{\circ}{S}_{x,z}$ which converges to z, we conclude that $z \in \overline{\operatorname{int}(A)}$. □

Remark 1.1.

Even in finite-dimensional spaces, the application co(·) does not always map closed sets to closed sets. To give an example,

$$A = \left\{ \left(\pm n, \frac{1}{n} \right) : n \in \mathbb{N} \right\} \subset \mathbb{R}^2$$

is closed but $\mathcal{X} = \{(x,0) : x \in \mathbb{R}\} \in \overline{co(A)} - co(A)$.

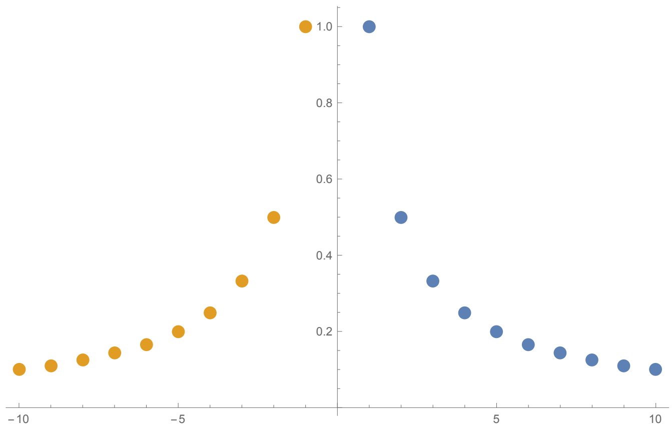


Figure 1.4: Illustration of A.

1.3 Dimension of a convex set

Suggestively, the dimension of a vector space is the number of elements which has every base that can be defined on it. Nonetheless, this concept is reserved to subspaces (in particular the own space). As it could be appreciated in examples 1.1, the definition of affine hull gives us the possibility to assign a *dimension* to convex sets. During this section, we will assume that X is a finite-dimensional topological vector space with $\dim(X) = n$, hence we can suppose that the norm employed is the Euclidean one, which comes from its respective scalar product.

Definition 1.3. *Let X be a convex set. For any convex set $A \subset X$, the dimension of A is the dimension of its affine hull:*

$$dim(A) = dim(aff(A)).$$

There arises now the problem of studying the convex set in its affine hull, in order to get more information. For this circumstance appears the next definition.

Definition 1.4. *In a topological vector space X, the relative interior of a convex set $A \subset X$ is the interior of A in the induced topology by its affine hull. The collection of relative interior points of A is denoted by $ri(A)$.*

It should be appreciated that $\operatorname{int}(A)$ and $\operatorname{ri}(A)$ are not the same concepts: in fact, given $X = \mathbb{R}^3$ and A any unit disk, i.e.,

$$A = \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 < 1, z = 0\},$$

we have that A is convex, but $\operatorname{int}(A) = \emptyset$ and $\operatorname{ri}(A) = A$ since $\operatorname{aff}(A) = \{(x, y, z) \in \mathbb{R}^3 \colon z = 0\}$.

Proposition 1.3. *Let X be a topological vector space and A be a non-empty convex subset of X. Then,*

1. *$ri(A) \neq \emptyset$.*
2. *$aff(A) = aff(ri(A))$.*
3. *$\overline{A} = \overline{ri(A)}$.*

Proof. *1.* First of all, lemma 1.2 shows that $\operatorname{ri}(A)$ is convex. We can suppose without losing generality that $0 \in A$ and $\dim(A) = m$, $0 \leq m \leq n = \dim(X)$.

If $m = 0$ it is trivial, so long as $A = \operatorname{aff}(A) = \{0\}$ and $\operatorname{ri}(A) = \{0\}$. Otherwise, we can find $\{x_i\}_{i=1}^m$ linear-independent vectors that span $\operatorname{aff}(A)$ (i.e., forming a basis for $\operatorname{aff}(A)$). Consider

$$Y = \left\{ x \in A \colon x = \sum_{i=1}^m \lambda_i x_i, \sum_{i=1}^m \lambda_i < 1, \lambda_i > 0, \forall i = 1, \ldots, m. \right\}.$$

We want to state that Y is open relative to $\operatorname{aff}(A)$. To do that, fix $y \in Y$ and let $x \in \operatorname{aff}(A)$. Let M be the $n \times m$-matrix which columns are $\{x_i\}_{i=1}^m$ and λ, $\overline{\lambda}$ the unique m-dimensional vectors such that

$$y = M\overline{\lambda}, x = M\lambda.$$

Due to the fact that $M^t M$ is a symmetric and positive definite matrix, we can find $\gamma \in \mathbb{R}^+$ satisfying

$$||x - y||^2 = ||M(\lambda - \overline{\lambda})||^2 = (M(\lambda - \overline{\lambda}))^t (M(\lambda - \overline{\lambda})) = (\lambda - \overline{\lambda})^t M^t M(\lambda - \overline{\lambda}) \geq \gamma ||\lambda - \overline{\lambda}||^2.$$

Since $y \in Y$, the vector $\overline{\lambda}$ lies in the open set

$$E = \left\{ (\lambda_1, \ldots, \lambda_m) \colon \sum_{i=1}^m \lambda_i < 1, \lambda_i > 0, \forall i = 1, \ldots, m \right\}.$$

From the above calculations we see that if x lies in a suitable small ball centered at y, the corresponding vector λ lies in E, implying that $x \in Y$.

This means that Y contains the intersection of $\operatorname{aff}(A)$ and an open ball centred at y, so Y is open relative to $\operatorname{aff}(A)$. Note that every point $y \in Y$ is a relative interior point of A, and hence $\operatorname{ri}(A) \neq \emptyset$.

2. Our previous construction of Y gives us that $\operatorname{aff}(Y) = \operatorname{aff}(A)$, and since $Y \subset \operatorname{ri}(A)$, we see that $\operatorname{aff}(A) = \operatorname{aff}(\operatorname{ri}(A))$.

3. It is clear that $\operatorname{ri}(A) \subset A \Rightarrow \overline{\operatorname{ri}(A)} \subset \overline{A}$. On the other hand, let $y \in \overline{A}$ and $x \in \operatorname{ri}(A)$. If $x = y$, it is done. Otherwise, we know that $\overset{\circ}{S}_{x,y} \subset \operatorname{ri}(A)$.

Consider the sequence

$$\left\{\frac{1}{n}x + \left(1 - \frac{1}{n}\right)y\right\}_{n\in\mathbb{N}} \subset \operatorname{ri}(A).$$

This sequence converges to y, hence $y \in \overline{\operatorname{ri}(A)}$ and $\overline{A} \subset \overline{\operatorname{ri}(A)}$. □

A detailed reading of the last proposition gives us an explicit expression of the relative interior of a convex set given by the convex hull of affine-independent points:

$$\operatorname{ri}[\operatorname{co}(\{x_0,\ldots,x_k\})] = \left\{\sum_{i=0}^{k} \lambda_i x_i : \sum_{i=0}^{k} \lambda_i = 1,\ \lambda_i > 0,\ \forall i = 0,\ldots,k\right\}.$$

1.4 Extreme points

We devote the most important section of this chapter to the introduction of the concept of extreme point.

Definition 1.5. *An extreme point of a convex set A in a vector space X is a point $x \in A$ satisfying, for every $y,z \in A$:*

$$x \in S_{y,z} \Rightarrow x = y \vee x = z.$$

We will denote as $ext(A)$ the set of extreme points of A.

In other words, an extreme point is a point which is not contained in any non-trivial segment of points of A.

Examples 1.2.

- *The extreme points of a polyhedron are their own vertexes.*
- *In a closed Euclidean ball $A = \overline{B}(x,r)$, $ext(A) = S(x,r)$. This example shows that $ext(A)$ may not be necessarily finite.*
- *Consider the following subset of $X = \mathbb{R}^3$:*

  $$A = co(\{(\pm1,\pm1,\pm1)\} \cup \{(\cos\alpha, \pm(1+\sin\alpha), 0) : a \in [0,\pi]\}).$$

  In this case, the points $\{(\pm1,\pm1,0)\} \notin ext(A)$ since they are contained in any segment with the form $S_{(\pm1,\pm1,r),(\pm1,\pm1,-r)} \subset A$, $0 < r \leq 1$. Now it can be appreciated that $ext(A)$ may not be closed (see figure 1.5).

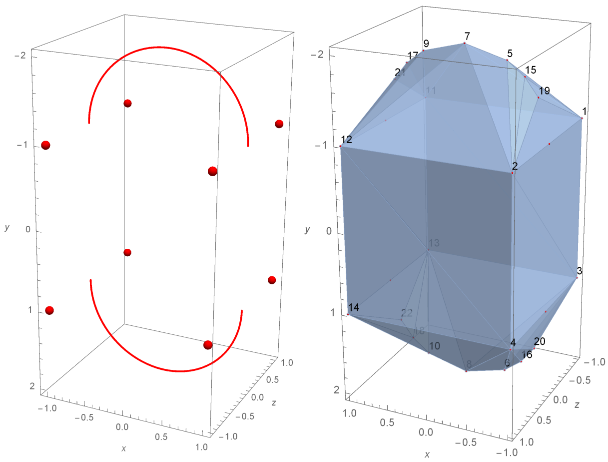


Figure 1.5: Plot of A and $\mathrm{co}(A)$.

A more general notion is derived from the previous concept.

Definition 1.6. *Let $A \subset X$ a convex set in a vector space. A subset $F \subset A$ is said to be a face of A if it is a convex set and, for every $x, y \in A$,*

$$\mathring{S}_{x,y} \cap F \neq \emptyset \Rightarrow S_{x,y} \subset F.$$

A proper face $F \subset A$ satisfies $F \neq A$.

Extreme points are one-point faces of A. A canonical way proper faces are constructed is via linear functionals.

Proposition 1.4. *Let $A \subset X$ a convex set in a vector space and $f : A \longrightarrow \mathbb{R}$ an affine functional with $\sup_{x \in A} f(x) = \alpha < +\infty$. Then, if*

$$F = \{y \in A \colon f(y) = \alpha\} \tag{1.7}$$

is a non-empty set, is a face of A. In particular, when X is a topological vector space, any linear and continuous functional defines a face over a compact convex subset $A \subset X$ in a topological vector space.

Proof. It is clear that F is convex by linearity of f. Given $y, z \in A$ with $\mathring{S}_{y,z} \subset F$, we have

$$\left\{ \begin{array}{ccc} tf(y) + (1-t)f(z) & = & \alpha \\ f(y) & \leq & \alpha \\ f(z) & \leq & \alpha \end{array} \right. \Rightarrow f(y) = f(z) = \alpha$$

and $S_{y,z} \subset F$. □

If X is a topological vector space and the functional f which appeared in proposition 1.4 is nonzero, linear, continuous and defined in X, the set given by the equation (1.7) is called an *exposed set*; in particular, if it is a singleton, we call the point an *exposed point*.

Remarks 1.2. *Let X be a topological vector space.*

- *In addition to the previous proposition, if the functional f is nonconstant in A, then the mentioned face is proper.*
- *Every exposed set F is closed (in the relative topology of A) by the own definition. In particular, if X is Hausdorff and $A \subset X$ is compact, so is F.*
- *Every exposed point is an extreme point, but the reciprocal is not true in general. As an example, consider $X = \mathbb{R}^3$ and*

  $$A = \left\{(x,y,z) \in \mathbb{R}^3 \colon x^2 + y^2 \leq 1, -2 \leq z \leq 0,\right\} \bigcup \left\{(x,y,z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 \leq 1\right\}.$$

  Every point in the set $\left\{(x,y,z) \in \mathbb{R}^3 \colon x^2 + y^2 = 1, z = 0\right\}$ is an extreme one, but it is not an exposed one since the unique supporting hyperplane is not a singleton (figure 1.6).

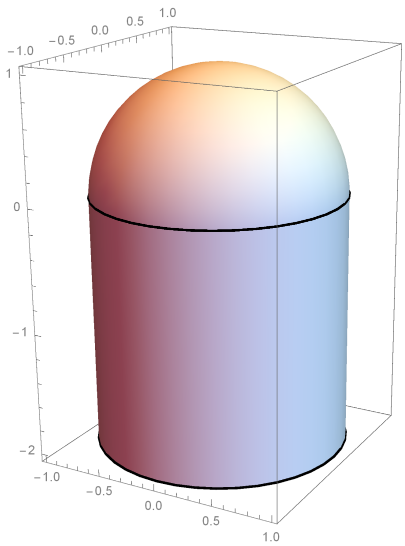


Figure 1.6: Example of non-exposed and extreme points.

Recall now the geometric version of Hahn-Banach theorem:

Theorem 1.2 (Existence of supporting functionals)**.** *Let X be a topological vector space and A a closed convex subset of X such that $int(A) \neq \emptyset$. Then, for any x_0 in the boundary of A, there exists a nonzero linear and continuous functional f such that*

$$Ref(x_0) = \max_{x \in A} Ref(x).$$

If $\alpha = \mathrm{Re} f(x_0)$, the affine hyperplane in $X_{\mathbb{R}}$ given by $H = \{x \in X \colon \mathrm{Re} f(x) = \alpha\}$ contains x_0 and isolates A; we say that the functional f or the hyperplane H supports the set A in the point x_0. One can easily notice that the supporting hyperplane may not be unique (any vertex of a regular polyhedron admits infinite many of them).

This is an appropriate moment to define the convex bodies: a subset A of a topological vector space X is said to be a convex body if it is closed, convex and $\mathrm{int}(A) \neq \emptyset$.

Proposition 1.5. *Let X be a topological vector space and A a convex subset of X. Any proper face $F \subset A$ lies in the boundary of A. Conversely, if A is a convex body, then every point of its boundary is contained in a proper face.*

Proof. Let $x \in F$ and $y \in A - F$. The set $B = \{t \in \mathbb{R} \colon tx + (1-t)y \subset A\}$ is contained in $[0,1]$ but it can not include any $t > 1$ for if it did, x would be an interior point of a segment in A with at least one point in $A - F$. Hence

$$\left\{(1+n^{-1})x + n^{-1}y\right\}_{n\in\mathbb{N}}$$

is a sequence in $X - A$ which converges to x, i.e. $x \in \overline{A} \cap \overline{X - A} = \partial A$.

Reciprocally, let us assume that A is a convex body and x_0 a point in its boundary. In light of theorem 1.2, there exists a continuous functional $f \neq 0$ such that $\alpha = \sup_{y\in A} \mathrm{Re} f(y) = \mathrm{Re} f(x_0)$. In addition, according to proposition 1.5, the set $\{x \in A \colon \mathrm{Re} f(x) = \alpha\}$ defines a proper face of A which contains x_0, so long as if $\mathrm{Re} f$ was constant in A, it would be constant in X. □

Corollary 1.1. *If X is a finite-dimensional topological vector space the dimension of any proper face F of a convex set $A \subset X$ is strictly less than* $\dim(A)$.

Proof. If $\dim(F) = \dim(A)$, then $V = \mathrm{aff}(A) = \mathrm{aff}(F)$, hence $\mathrm{ri}(F) \neq \emptyset$. But F lies in the boundary of A relative to V by proposition 1.5, so we have a contradiction. □

Proposition 1.5 highlights the importance of compact sets, so long as it is needed the existence of boundary points (closed sets) and their abundance (bounded sets). Henceforth, we will also restrict the term "face" to indicate a closed set, even if there exist nonclosed faces in infinite dimensional spaces. We conclude this section with an observation about the transitivity of the faces in a convex set.

Proposition 1.6. *Let $A \subset X$ be a convex set in a vector space and F a face of A. Let $B \subset F$. Then B is a face of F iff it is a face of A. In particular, $x \in F$ is in $ext(F)$ iff it is also in $ext(A)$, i.e.,*

$$ext(F) = F \cap ext(A).$$

Proof. $\Rightarrow$) Suppose that $B \subset F$ is a face, $x \in B$ and $x \in \mathring{S}_{y,z} \subset A$. Since $x \in F$ and F is a face, we have that $y, z \in F$. Hence $y, z \in B$ and so B is a face of A.

$\Leftarrow$) If $B \subset A$ is a face, $x \in B$ and $x \in \mathring{S}_{y,z} \subset F \subset A$, then $y, z \in F \subset A$ and consequently $y, z \in B$ for being B a face of A. Thus, B is a face of F. □

1.5 Carathéodory-Minkowski theorem

The final section of this chapter will introduce us to Carathéodory-Minkowski theorem in finite-dimensional spaces. The existence of extreme points will be given by the compactness of the convex set.

Taking into account that, for a finite-dimensional vector space X, there exists a unique Hausdorff topology compatible with the vector structure, and that topology is induced by any norm, we can use the Euclidean one for our purposes.

Lemma 1.3. *Let $A \subset X$ a compact convex set.*

1. *Every compact convex set $A \subset X$ has at least one extreme point.*

2. *If $f : A \to \mathbb{R}$ is an affine functional which attains a unique maximum in $a \in A$, a is an extreme point.*

Proof. *1.* Since A is compact and $\|\cdot\| : X \to \mathbb{R}$ is continuous in A, it attains its maximum in $a \in A$. Suppose, without losing generality, that $a = \frac{1}{2}(x+y)$ for some $S_{x,y} \subset A$. Then,

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

Then $\|a\| = \frac{1}{2}(\|x\|+\|y\|)$.

- If $\|x\| < \|y\| \vee \|y\| \leq \|x\|$, then $\|a\| < \|y\| \vee \|a\| < \|x\|$, which is a contradiction with the choice of a.

- If $\|x\| = \|y\|$, the strict convexity of the Euclidean norm concludes that $a = x = y$.

2. It is an easy consequence of proposition 1.4. □

The final theorem of this chapter is ready to be introduced now.

Theorem 1.3 (Carathéodory-Minkowski)**.** *Let $A \subset X$ a compact convex subset of a finite-dimensional space X (with $dim(X) = n$). Then,*

$$A = co(ext(A)),$$

namely, every $a \in A$ is a convex combination of $n+1$ extreme points in A as much.

Proof. It will be done by induction on the dimension n. For $n = 0$, A is a point and the result is obvious. Let assume the theorem for $n < d$. It can be also supposed, without loss of generality that $\operatorname{int}(A) \neq \emptyset$. Otherwise we can find an affine variety of lower dimension ($< d$) that contains the set A such that $\operatorname{ri}(A)$ is non-empty (proposition 1.3). Since the dimension will be less than d, the result follows from induction hypothesis.

Let a an element in the boundary of A. By proposition 1.5, there exists a face F such that $a \in F$ isolating A. Since $\dim(F) < d$ (corollary 1.1) and any face of a compact convex set is a compact convex set, by induction hypothesis is $a \in \operatorname{co}(\operatorname{ext}(F)) \subset \operatorname{co}(\operatorname{ext}(A))$ (propositions 1.1 and 1.6).

Suppose $a \in \operatorname{int}(A)$. Since A is bounded, there exist x, y in the boundary of A satisfying $a \in \overset{\circ}{S}_{x,y} \subset A$. As it has been proved, $x, y \in \operatorname{co}(\operatorname{ext}(A))$, and since $\operatorname{co}(\operatorname{ext}(A))$ is convex, $a \in \operatorname{co}(\operatorname{ext}(A))$. □

After having established Carathéodory-Minkowski's theorem, it can be fathomed that for every non-empty compact convex subset $A \subset X$, the set of its extreme points, $\operatorname{ext}(A)$, is always non-empty and it can also be justified the assertion that the convex hull of $\operatorname{ext}(A)$, $\operatorname{co}(\operatorname{ext}(A))$, is always closed (in fact, compact).

The following corollary highlights the importance of Carathéodory-Minkowski's theorem in linear optimization.

Corollary 1.2. *Let $A \subset X$ be a compact convex set and $f : A \to \mathbb{R}$ a linear (continuous) functional. Then f attains its maximum (or its minimum) at an extreme point of A.*

Proof. By lemma 1.3 A has an extreme point. Plus, since f is continuous and A compact, f attains its maximum (the other case can be reduced to consider the functional $-f$) in a point $a \in A$. Then a is a convex combination of some extreme points of A, i.e.,

$$a = \sum_{i=1}^{n+1} \lambda_i x_i, \sum_{i=1}^{n+1} |\lambda_i| > 0, \sum_{i=1}^{n+1} \lambda_i = 1, \{x_i\}_{i=1}^{n+1} \subset \mathrm{ext}(A).$$

Hence,

$$f(a) = \sum_{i=1}^{n+1} \lambda_i f(x_i) \leq \sum_{i=1}^{n+1} \lambda_i f(a) = f(a).$$

It is clear that, for every $\lambda_i \neq 0$, $f(x_i) = f(a)$, and hence f attains its maximum at an extreme point. $\square$

2 Krein-Milman theorem

Now we have a slight background about finite-dimensional theory, we want to extend those results to an arbitrary dimension (with appropriate considerations). Indeed, we will extend the domain of the space we are going to develop the theory, and we will consider locally convex Hausdorff topological vector spaces; specifically, Krein-Milman theorem states that:

> *"Every non-empty compact convex subset of a locally convex Hausdorff topological vector space is the topological closure of the convex hull of its extreme points".*

This generalisation adds up within our purpose of studying convex sets, since it is only needed the structure of vector space[1]. In that sense, a first approach to the problem will be given by the algebraic characterisation of unit balls in normed spaces. This step will bring us naturally the main algebraic concepts involved in this theory, which are *balancedness*, *absorbency* and the own definition of convexity. The immediate generalisation of those results will make us able to extend the domain to TVS.

The essential tool in the proof will be, as well as in many other results of functional analysis, Zorn's lemma. The next reminder of the mentioned statement will be useful for the reader:

Theorem 2.1 (Zorn). *Let X be a preordered set. If each chain in X has an upper bound, then X has at least one maximal element.*

A further relation between the axiom of choice, Zorn's lemma and Zermelo's theorem can be found in [10].

Another useful result will be Hahn-Banach theorem, specifically one form of it which is known as the Geometric Hahn-Banach theorem. It states that:

Theorem 2.2 (Hahn-Banach). *Let X be a locally convex topological vector space over $\mathbb{K} = \mathbb{R} \vee \mathbb{C}$. If A, B are convex, non-empty disjoint subsets of X, A compact and B closed, then there exists a continuous linear map $f : X \to \mathbb{K}$ and $s, t \in \mathbb{R}$ satisfying*

$$Re(f(a)) < t < s < Re(f(b)), \quad \forall a \in A, \forall b \in B.$$

In particular, X^* separates points of X.

To begin with, we will develop some examples of extreme points in infinite-dimensional spaces to show not only the differences between both cases, but also the main strategies that can be developed during the analysis of this theory.

2.1 Examples of extreme points

The most usual examples of topological vector spaces are normed spaces $(X, \|\cdot\|)$. Plus, the study of extreme points in some closed ball $\overline{B}(x, r)$ for $x \in X$ and $r > 0$ can be reduced to the study of $\overline{B}(0, 1)$. For those reasons, our first examples will be devoted in that environment.

[1] To revise some concepts related to (locally convex) topological vector spaces, (cf. [6]).

Even if proposition 1.5 shows that extreme points can be only found in $S(0,1)$, it is interesting to consider the wide variety of enriching strategies which take place during the particular verifications that $\overset{\circ}{B}_X$ has no extreme points. In order to establish some helpful notation for an arbitrary topological vector space X over a field $\mathbb{K}$, the following concepts will be abbreviated as follows:

$$B_X := \overline{B}(0,1), \quad S_X := S(0,1), \quad \mathbb{T} := \{\lambda \in \mathbb{K} \colon |\lambda| = 1\}, \quad \mathrm{E}_X := \mathrm{ext}(B_X).$$

$L_p([0,1])$ spaces

The first example of extreme points (which also contains infinite many of them) will show the difference between extreme points in $L_p([0,1])$ spaces, taking into consideration whether $p=1$ or $1<p<+\infty$. If $p=1$, there are no extreme points in B_{L_1}, whereas if $1<p<+\infty$ we have that $\mathrm{E}_{L_p} = S_{L_p}$.

- First let $f \in B_{L_1}$ with $\|f\|_1 = 0$. Then,

$$f = 0 = \frac{1}{2}(1) + \frac{1}{2}(-1)$$

almost everywhere. Since 1 and -1 are different functions in B_1, $f \in \overset{\circ}{S}_{-1,1} \subset B_{L_1}$, hence $f \notin \mathrm{E}_{L_1}$. Now consider $f \in B_{L_1}$ with $0 < \|f\|_1 \le 1$ and the function $F : [0,1] \to \mathbb{R}$ given by

$$F(x) = \int_0^x |f(t)|\,dt, \forall x \in [0,1].$$

Then, F is continuous and satisfies $F(0)=0$, $F(1) = \|f\|_1 \le 1$. Hence, by the Intermediate Value Theorem there exists $\xi \in]0,1[$ satisfying $F(\xi) = \frac{1}{2}F(1)$.

Now for each $t \in [0,1]$, define

$$f_1(t) = \chi_{[0,\xi]} 2f(t), \quad f_2(t) = \chi_{]\xi,1]} 2f(t).$$

It is clear that $f_1, f_2 \in B_{L_1}$ and, for every $t \in [0,1]$:

$$f(t) = \chi_{[0,\xi]} f(t) + \chi_{]\xi,1]} f(t) = \frac{1}{2}\left(\chi_{[0,\xi]} 2f(t)\right) + \frac{1}{2}\left(\chi_{]\xi,1]} 2f(t)\right) = \frac{1}{2}f_1(t) + \frac{1}{2}f_2(t).$$

Finally, it has to be checked that $f \in \overset{\circ}{S}_{f_1,f_2}$:

$$\begin{aligned}
\|f - f_1\|_1 &= \int_0^1 |(f-f_1)(t)|\,dt \ge \int_0^\xi |(f-f_1)(t)|\,dt = \int_0^\xi |(f-2f)(t)|\,dt \\
&= \int_0^\xi |f(t)|\,dt = \frac{1}{2}\int_0^1 |f(t)|\,dt = \frac{1}{2}\|f\|_1 > 0. \\
\|f - f_2\|_1 &= \int_0^1 |(f-f_2)(t)|\,dt \ge \int_\xi^1 |(f-f_2)(t)|\,dt = \int_\xi^1 |(f-2f)(t)|\,dt \\
&= \int_\xi^1 |f(t)|\,dt = \frac{1}{2}\int_0^1 |f(t)|\,dt = \frac{1}{2}\|f\|_1 > 0.
\end{aligned}$$

- The first part of the strategy is going to be followed now is slightly different, so long as B_{L_p} has infinite many extreme points. To show that $\mathrm{E}_{L_p} = S_{L_p}$, it will be required the fact that the map
$$x \mapsto \|x\|_p^p$$
is strictly convex for each $1 < p < +\infty$, which means in particular that, for every different $f, g \in L_p([0,1])$ and $t \in]0,1[$,
$$\|tf + (1-t)g\|_p^p < t\|f\|_p^p + (1-t)\|g\|_p^p.$$
In fact, suppose that there exists $f \in S_{L_p}$ with $f = \frac{1}{2}(f_1 + f_2)$ for some different $f_1, f_2 \in B_{L_p}$. Then,
$$\|f\|_p^p = \left\|\frac{1}{2}(f_1 + f_2)\right\|_p^p < \frac{1}{2}\|f_1\|_p^p + \frac{1}{2}\|f_2\|_p^p \leq \frac{1}{2} + \frac{1}{2} = 1,$$
hence there is a contradiction since $\|f\|_p = 1 = \|f\|_p^p$.
The implication $S_{L_p} \subset \mathrm{ext}(B_{L_p})$ has been proved. To check the equality, it only has to be considered the proposition 1.5.

c_0 space

The next example will be devoted to the space c_0 of convergent sequences to 0. First of all, given any $x = \{x_n\}_{n\in\mathbb{N}} \in B_{c_0}$, we have that $|x_n| \leq 1$ for all $n \in \mathbb{N}$ and there also exists $m \in \mathbb{N}$ satisfying
$$|x_n| < \frac{1}{2}, \quad \forall n \geq m.$$
Let y, z be the elements given by
$$y_n = \begin{cases} x_n & \text{if} \quad n \neq m \\ x_m + \frac{1}{2} & \text{if} \quad n = m \end{cases}, z_n = \begin{cases} x_n & \text{if} \quad n \neq m \\ x_m - \frac{1}{2} & \text{if} \quad n = m \end{cases}.$$
It is clear that $y, z \in B_{c_0}$, $y \neq z$ and $x = \frac{1}{2}(y+z)$, hence x is not an extreme point.

ℓ_1 space

Now we are going to consider the closed unit ball B_{ℓ_1} in the space of sequences ℓ_1 whose series is absolutely convergent. To check that
$$\mathrm{E}_{\ell_1} = \{e_{\alpha,k} := \{e_{\alpha,k}(n)\}_{n\in\mathbb{N}}\}_{k\in\mathbb{N}}$$
with $\alpha \in \mathbb{T}$ and $e_{\alpha,i}(j) = \alpha\delta_{ij}$, suppose that there exist $a = \{a_i\}_{i\in\mathbb{N}}, b = \{b_j\}_{j\in\mathbb{N}} \in B_{\ell_1} - \{0\}$ with $e_{\alpha,k} = \frac{1}{2}(a+b)$. Then,
$$e_{\alpha,k}(k) = \alpha = \frac{1}{2}(a_k + b_k) \quad \wedge \quad e_{\alpha,k}(m) = 0 = \frac{1}{2}(a_m + b_m),\, m \neq k.$$
In particular, $|a_m| = |b_m|$ for each $m \neq k$ and $|a_k| = |2 - b_k| \geq |2 - |b_k|| = 2 - |b_k|$, hence
$$\|a\| = 2 - |b_k| + \sum_{\substack{k=1 \\ k\neq m}}^{+\infty} |b_k| \geq 2 - |b_k| > 1,$$

and we have a contradiction.

It only lefts to prove that $0 \notin \mathrm{E}_{\ell_1}$, but that is trivial, so long as it can be considered any $e_{\alpha,k} := a$ and its opposite $-e_{\alpha,k} := b$ for any $k \in \mathbb{N}$ (which are different) satisfying $0 = \frac{1}{2}(a+b)$.

Now, let $x = \{x_n\}_{n\in\mathbb{N}} \in B_{\ell_1} - \{0\}$ with $\|x\| < 1$ and consider $0 < \varepsilon < 1 - \|x\|$. Defining

$$y_n = \begin{cases} x_1 + \varepsilon & \text{if} \quad n = 1 \\ x_n & \text{if} \quad n > 1 \end{cases}, \; z_n = \begin{cases} x_1 - \varepsilon & \text{if} \quad n = 1 \\ x_n & \text{if} \quad n > 1 \end{cases},$$

we have that $y \neq z$ and

$$\|y\| \leq \varepsilon + \sum_{n=1}^{+\infty} |x_n| < 1, \; \|z\| \leq \varepsilon + \sum_{n=1}^{+\infty} |x_n| < 1,$$

hence $y, z \in B_{\ell_1}$ and $x = \frac{1}{2}(y+z)$.

Finally, if $\|x\| = 1$ and there exist at least two $i, j \in \mathbb{N}$ such that $0 \notin \{x_i, x_j\}$, define $0 < \varepsilon < \min\{|x_i|, |x_j|\}$ and

$$y_n = \begin{cases} x_n & \text{if} \quad n \notin \{i,j\} \\ x_i\left(1 + \frac{\varepsilon}{|x_i|}\right) & \text{if} \quad n = i \\ x_j\left(1 - \frac{\varepsilon}{|x_j|}\right) & \text{if} \quad n = j \end{cases}, \; z_n = \begin{cases} x_n & \text{if} \quad n \notin \{i,j\} \\ x_i\left(1 - \frac{\varepsilon}{|x_i|}\right) & \text{if} \quad n = i \\ x_j\left(1 + \frac{\varepsilon}{|x_j|}\right) & \text{if} \quad n = j \end{cases}.$$

We have that $y \neq z$ and

$$\|y\| = |x_i|\left(1 + \frac{\varepsilon}{|x_i|}\right) + |x_j|\left(1 - \frac{\varepsilon}{|x_j|}\right) + \sum_{\substack{n=1 \\ n\notin\{i,j\}}}^{+\infty} |x_n| = |x_i| + \varepsilon + |x_j| - \varepsilon + \sum_{\substack{n=1 \\ n\notin\{i,j\}}}^{+\infty} |x_n| = \sum_{n=1}^{+\infty} |x_n| = 1$$
$$\|z\| = |x_i|\left(1 - \frac{\varepsilon}{|x_i|}\right) + |x_j|\left(1 + \frac{\varepsilon}{|x_j|}\right) + \sum_{\substack{n=1 \\ n\notin\{i,j\}}}^{+\infty} |x_n| = |x_i| - \varepsilon + |x_j| + \varepsilon + \sum_{\substack{n=1 \\ n\notin\{i,j\}}}^{+\infty} |x_n| = \sum_{n=1}^{+\infty} |x_n| = 1$$

hence $y, z \in B_{\ell_1}$ and $x = \frac{1}{2}(y+z)$.

ℓ_∞ space

Let B_{ℓ_∞} be the closed unit ball in ℓ_∞, and define

$$\mathcal{C} = \{x := \{x_n\}_{n\in\mathbb{N}} \in B_{\ell^\infty} : x_n \in \mathbb{T}, \forall n \in \mathbb{N}\}.$$

To verify that $\mathrm{E}_{\ell_\infty} = \mathcal{C}$, consider $x \in \mathcal{C}$. Since every $x_n \in \mathbb{T}$ is an extreme point of $B_{\mathbb{K}}$, it follows that $x \in \mathrm{E}_{\ell_\infty}$. Now select $x \in B_{\ell_\infty} - \mathcal{C}$; there exists $m \in \mathbb{N}$ such that $|x_m| < 1$. Defining $0 < \varepsilon < 1 - |x_m|$ and

$$y_n = \begin{cases} x_m + \varepsilon & \text{if} \quad n = m \\ x_n & \text{if} \quad n \neq m \end{cases}, \; z_n = \begin{cases} x_m - \varepsilon & \text{if} \quad n = m \\ x_n & \text{if} \quad n \neq m \end{cases},$$

we have that $y \neq z$ and

$$\|y\| = \sup_{\substack{n\in\mathbb{N} \\ n\neq m}} \{\{|x_n|\} \cup \{|x_m + \varepsilon|\}\} \leq 1, \; \|z\| = \sup_{\substack{n\in\mathbb{N} \\ n\neq m}} \{\{|x_n|\} \cup \{|x_m - \varepsilon|\}\} \leq 1,$$

hence $y, z \in B_{\ell_\infty}$ and $x = \frac{1}{2}(y+z)$.

$C(X)$ space

Let X be a compact Hausdorff space. Then, the space $C(X)$ of continuous functionals $f : X \to \mathbb{K}$ is a normed space with $\|f\|_\infty = \sup_{x\in X}|f(x)|$. Consider $B_{C(X)}$ the unit closed ball in $C(X)$, and

$$\mathcal{C} = \left\{f \in B_{C(X)} \colon Im(f) \subset \mathbb{T}\right\}.$$

To show that $\mathcal{C} \subset \mathrm{E}_{C(X)}$, it only has to be noticed that every $f(x) \in \mathbb{T}$ is an extreme point of $B_{\mathbb{K}}$.

Conversely, given $f \in B_{C(X)}$ with $|f(x)| < 1$ for some $x \in X$, by continuity of f it can be found

$$0 < \varepsilon < 1 - |f(x)|$$

and $\delta > 0$ satisfying the previous inequality for every $x \in B(x,\delta)$. Given $b \in B(x,\delta)$, Urysohn's lemma states that there exists $\phi \in B_{C(X)}$ such that $\phi(b) = 1$ and $\phi|_{B(x,\delta)^C} = 0$. Let $g = \varepsilon\phi \in B_{C(X)}$; then

$$|(f \pm g)(y)| = \begin{cases} |(f \pm \varepsilon\phi)(y)| < 1 - \varepsilon + \varepsilon = 1 & \text{if} \quad y \in B(x,\delta) \\ |f(y)| \leq 1 & \text{if} \quad y \in B(x,\delta)^C \end{cases} \Rightarrow \|f \pm g\|_\infty \leq 1,$$

hence $f \pm g \in B_{C(X)}$ and $f = \frac{1}{2}[(f+g) + (f-g)]$.

Probability measures on a compact Hausdorff space

The next example concerns the extreme points of the set of probability measures on a (locally) compact Hausdorff topological space X. This is an example of paramount importance related to the integral representation theorems of convex sets. The objective of this section is proving that $\mathrm{E}_{P(X)} = \{\varepsilon_x \colon x \in X\}$, where $P(X)$ is the set of probability measures on X and ε_x is the Borel measure which equals 1 on any Borel subset of X which contains x, and equals 0 otherwise.

First it is going to be remarked the definition of a Radon measure.

Definition 2.1. *Let X be a locally compact Hausdorff topological space. A Radon measure μ on X is a Borel measure satisfying the following conditions:*

- $\mu(K) < +\infty$ *for every compact set $K \subset X$.*
- *Outer regularity: For each Borel set $A \subset X$,*

$$\mu(A) = \inf\{\mu(U) \colon U \text{ open}, A \subset U\}.$$

- *Inner regularity: For each open set $U \subset X$,*

$$\mu(U) = \sup\{\mu(K) \colon K \text{ compact}, K \subset U\}.$$

A probability measure is a Radon measure on X having total mass equal to 1.

Roughly speaking, the support of a Radon measure μ is the biggest closed subset of X for which every open neighbourhood of every point of the set has positive measure. In order to define it, consider $\{U_i\}_{i\in I}$ the family of open set such that $\mu(U_i) = 0$. Then, the set $U = \cup_{i\in I} U_i$ is clearly open and $\mu(U) = 0$. Using the inner regularity of μ, it is enough proving the result for every $K \subset U$ compact. Since K is compact, there exists $i_1, \ldots, i_n \in I$ satisfying

$$K \subset \bigcup_{j=1}^{n} U_{i_j}.$$

Hence,

$$\mu(K) \leq \mu\left(\bigcup_{j=1}^{n} U_{i_j}\right) \leq \sum_{j=1}^{n} \mu(U_{i_j}) = 0.$$

Definition 2.2. *The support of a Radon measure μ is the complement of the set U previously defined.*

Given $x \in X$, recall that the support of the Dirac mass δ_x at x is $\mathrm{supp}(\delta_x) = \{x\}$. In addition, the converse also holds.

Lemma 2.1. *Let μ be a Radon measure on a locally compact Hausdorff topological space X and $x \in X$. If $supp(\mu) = \{x\}$, then $\mu = c\,\varepsilon_x$ where $c = \mu(X)$.*

Proof. By the definition of support of a measure,

$$\mu(\{x\}) = \mu(X) - \mu(X - \{x\}) = \mu(X) = c.$$

To show that $\mu = c\varepsilon_x$, let $A \in \mathcal{B}(X)$. If $x \notin A$, $A \subset X - \{x\}$ and $\mu(A) = 0 = c\varepsilon_x(A)$. On the other hand, if $x \in A$, then

$$\mu(A) = \mu(A - \{x\}) + \mu(\{x\}) = \mu(\{x\}) = \mu(X) = c\varepsilon_x(A).$$

Thus, $\mu = c\varepsilon_x$ for all $x \in X$. □

The next lemma highlights the importance of the support of a measure in terms of integration theory.

Lemma 2.2. *Let μ be a Radon measure on a locally compact Hausdorff topological space X. If $f \in B_{C(X)}$ and $f(x) > 0$ for some $x \in supp(\mu)$, then*

$$\int_X f\,d\mu > 0.$$

Proof. Consider the set

$$U = \left\{y \in X \colon f(y) > \frac{f(x)}{2}\right\}.$$

It is clear that U is open, $x \in U$ and $f > \frac{f(x)}{2}\chi_U$. Hence,

$$\int_X f\,d\mu > \int_U \frac{f(x)}{2}\,d\mu = \frac{f(x)}{2}\mu(U) > 0.$$

The latter holds because if $\mu(U) = 0$, then it would be contained in the largest open set having μ-measure zero, in particular x. On the other hand, by definition of $\mathrm{supp}(\mu)$, it follows that $x \notin \mathrm{supp}(\mu)$, which is a contradiction. □

Now it is time to get on with the desired result.

Proposition 2.1. *Let μ be a Radon measure on a locally compact Hausdorff topological space X. Then $P(X)$ is a convex set and*

$$E_{P(X)} = \{\varepsilon_x \colon x \in X\}.$$

Proof. First of all, it is clear that $P(X)$ is convex. The previous equality is going to be checked by double implication.

$\Rightarrow$) Let $\mu \in \mathrm{E}_{P(X)}$. By lemma 2.1 it suffices to show that $\mathrm{supp}(\mu) = \{x\}$ for some $x \in X$. Suppose that there exist $x, y \in X$ with $x \neq y$ satisfying $x, y \in \mathrm{supp}(\mu)$. Since every compact Hausdorff topological space is normal, there exist $U, V \in \mathcal{B}(X)$ such that $x \in U$, $y \in V$ and $\overline{U} \cap \overline{V} = \emptyset$. By Urysohn's lemma, there exists $f \in B_{C(X)}$ such that $f|_U = 1$ and $f|_V = 0$. Given that $f(x) = 1 > 0$, lemma 2.2 shows that

$$0 < \int_X f\, d\mu := \lambda.$$

Furthermore, it will be shown that $\lambda < 1$. For this, note that $f \leq \chi_{X-V}$, which implies

$$\lambda \leq \mu(X - V) = \mu(X) - \mu(V) = 1 - \mu(V) < 1,$$

where $\mu(V) > 0$ since $y \in \mathrm{supp}(\mu)$.

Now set

$$\mu_1 = \frac{1}{\lambda} f \mu, \quad \mu_2 = \frac{1}{1-\lambda}(1-f)\mu.$$

Then $\mu_1, \mu_2 \in P(X)$, $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$ and

$$\begin{aligned} \mu_1(U) &= \frac{1}{\lambda}\int_U f\, d\mu = \frac{\mu(U)}{\lambda} > 0, \\ \mu_2(U) &= \frac{1}{1-\lambda}\int_U (1-f)\, d\mu = 0, \end{aligned}$$

hence $\mu_1 \neq \mu_2$, which contradicts the fact that $\mu \in \mathrm{E}_{P(X)}$.

$\Leftarrow$) Assume to reach a contradiction that, given $x \in X$, $\varepsilon_x \notin \mathrm{E}_{P(X)}$. Then there exist $\mu_1, \mu_2 \in B_{P(X)} - \{\varepsilon_x\}$ satisfying

$$\varepsilon_x = t\mu_1 + (1-t)\mu_2,\ 0 < t < 1.$$

We have

$$1 = \varepsilon_x(\{x\}) = t\mu_1(\{x\}) + (1-t)\mu_2(\{x\}).$$

Since $\mu_1(\{x\}), \mu_2(\{x\}) \in [0,1]$ and $0 < t < 1$, we must have $1 = \mu_1(\{x\}) = \mu_2(\{x\})$. But both are probability measures, which implies that

$$\mu_1(X - \{x\}) = \mu_2(X - \{x\}) = 0$$

and $\varepsilon_x = \mu_1 = \mu_2$. $\square$

The extension of proposition 2.1 to signed measures is very natural to obtain, as long as this result can be understood as the study of a face in the general problem. Taking into consideration that, if $x \in \mathrm{E}_X$, then $\mathbb{T}x \subset \mathrm{E}_X$ for every normed space X, a similar reasoning to the previous proposition brings the following theorem.

Theorem 2.3 (Arens-Kelley). *Let $C(X)$ be the space of continuous functions on the compact Hausdorff space X. For each $x \in X$ let $\delta_x \in C(X)^*$ defined by*

$$\delta_x(f) = f(x), \forall f \in C(X).$$

Then

$$E_{C(X)^*} = \mathbb{T}\{\delta_x : x \in X\}.$$

Characterisation of strictly convex spaces by extreme points

Geometrically, a normed space $(X, \|\cdot\|)$ is strictly convex when its unit sphere S_X does not contain segments.

Definition 2.3. *Given a normed space $(X, \|\cdot\|)$, it is said to be strictly convex if, for every $x, y \in S_X$ with $\|x + y\| = 2$, it implies that $x = y$.*

Strictly convex spaces have an important characterisation in terms of extreme points, which is going to be developed in the following lines.

Proposition 2.2. *Let $(X, \|\cdot\|)$ be a normed space. Then X is strictly convex if, and only if, $E_X = S_X$.*

Proof. $\Rightarrow$) Let $x \in S_X$ and suppose that there exist $y, z \in B_X$ with $y \neq z$ satisfying $x = \frac{1}{2}(y + z)$. It is obvious that $2 = \|y + z\|$; furthermore,

$$\begin{aligned} 1 = \|x\| &= \frac{1}{2}\|y + z\| \leq \frac{1}{2}(\|y\| + \|z\|) \leq \frac{1}{2} 2 = 1 \\ &\Rightarrow 1 = \frac{1}{2}(\|y\| + \|z\|) \Rightarrow 2 = \|y\| + \|z\| \\ &\Rightarrow \|y\| = 1 = \|z\| \end{aligned}$$

Hence, by strictly convexity of X, it implies that $y = z$, which is a contradiction.

$\Leftarrow$) Given $x, y \in S_X$, suppose that $\|x + y\| = 2$. Then, the element $u = \frac{1}{2}(x + y)$ satisfies

$$\|u\| = \frac{1}{2}\|x + y\| = 1,$$

hence $u \in \mathrm{E}_X$ and $x = y$. □

The previous examples conclude that $L^p([0,1])$ $(1 < p < +\infty)$ are strictly convex spaces, whereas the rest of them are not strictly convex. It is also important the consideration of the following sufficient condition about this property.

Proposition 2.3. *Every normed space $(X, \|\cdot\|)$ satisfying the parallelogram identity is strictly convex.*

Proof. Let $x, y \in S_X$ with $\|x + y\| = 2$. Then,

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

Since $x, y \in S_X$, $\|x\| = \|y\| = 1$, hence $\|x\| + \|y\| - 2 = 0$, which means that $\|x - y\| = 0$ and $x = y$. □

2.2 Characterisation of unit balls in normed spaces

As it has been detailed in the introduction of this chapter, the main goal is giving an algebraic characterisation to a certain subset $A \subset X$ of a TVS for being the unit ball of a seminorm p over X. Since unit balls are one of the most important subsets of a seminormed space, it would be desirable to have a powerful knowledge of their properties. In addition, the requirements of the seminorm given by A will be satisfied by those properties, hence there are more than enough reasons to start with this section.

In order not to lose sight of the main purpose, after the establishment of the results it is going to be analysed every used property.

Definition 2.4. *Let $A \subset X$ be a subset of a TVS. A is said to be absorbing if, for every $x \in X$, there exists $r(x) := r \in \mathbb{R}^+$ satisfying $x \in rA$. Plus, A is said to be balanced if, for every $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$, is $\alpha A \subset A$.*

The main properties of those kind of sets can be easily checked for the reader, and it can be also consulted in [6]. One of the most important ones is that every 0-neighbourhood is absorbing (it will be used in theorem 2.6).

Examples 2.1.

- *The unit ball B_X of any normed space X is balanced, absorbing and convex. In fact, $0 \in B_X$ and for any $x \in X - \{0\}$ it can be chosen $r(x) = \|x\|$ satisfying $x \in \|x\| B_X$, or equivalently $\frac{x}{\|x\|} \in B_X$. Furthermore, given $\alpha \in \mathbb{K}$ such that $|\alpha| \leq 1$, it is clear by definition that $\alpha B_X \subset B_X$.*

- *The set $A = \{0\} \cup S_X$ of any normed space X is absorbing but not balanced. Indeed, a detailed lecture of the previous example shows that $\frac{x}{\|x\|} \in S_X$ for any $x \in X - \{0\}$. However, for every $\alpha \in \mathbb{K}$ such that $0 < |\alpha| < 1$, $\alpha A \not\subset A$.*

- *Given $x \in X$ with $\dim(X) > 1$, the set $S_{0,x}$ is balanced but not absorbing, since for each $\alpha \in \mathbb{K}$ such that $|\alpha| \leq 1$, $\alpha S_{0,x} = S_{0,\alpha x} \subset S_{0,x}$ so long as $\alpha x = (1-\alpha)0 + \alpha x \in S_{0,x}$ and every segment is convex. Nonetheless, for every $y \notin \operatorname{span}(\{x\})$, it does not exist $r \in \mathbb{R}$ satisfying that $y \in r S_{0,x}$ by the own definition of linear span.*

To begin with, it can be appreciated that given $A \subset X$ a convex absorbing set, the real number

$$p_A(x) = \inf\{r \in \mathbb{R}^+ : x \in rA\} \tag{2.1}$$

exists for each $x \in X$. In fact, $\{r \in \mathbb{R}_0^+ : x \in rA\}$ is non-empty and 0 is a lower bound for each $x \in X$. Hence one can define the map $p_A : X \to \mathbb{R}_0^+$ given by the equation (2.1). The main problem is what kind of properties must A satisfy to make q_A a seminorm.

Proposition 2.4. *Given the previous real number $q_A(x)$ for each $x \in X$ and $A \subset X$ a convex absorbing set:*

1. *The map q_A is a quasi-seminorm; i.e.,*

$$\begin{array}{rcl} q_A(tx) & = & t q_A(x) \\ q_A(x+y) & \leq & q_A(x) + q_A(y) \end{array}, \forall x, y \in X, \forall t \in \mathbb{R}_0^+.$$

2. *One has the inclusions*

$$\{x \in X : q_A(x) < 1\} \subset A \subset \{x \in X : q_A(x) \leq 1\}.$$

3. *In addition, if A is balanced, then q_A is a seminorm.*

The map q_A is called the Minkowski fuctional associated to A.

Proof. 1. since A is absorbing, it contains the zero vector, so in fact $0 \in \lambda A$ for all $\lambda \in \mathbb{R}^+$ and $p_A(0) = 0$. If $t > 0$,

$$p_A(tx) = \inf\{r \in \mathbb{R}^+ : tx \in rA\} = \inf\{r \in \mathbb{R}^+ : x \in \frac{r}{t}A\} = t\inf\{r \in \mathbb{R}^+ : x \in rA\} = tp_A(x).$$

On the other hand, given $x, y \in X$, we only have to check that $x+y \in (p_A(x)+p_A(y))A$, but that is clear in light of proposition 1.1, so long as $x+y \in p_A(x)A + p_A(y)A = (p_A(x)+p_A(y))A$ when A is convex.

2. To prove the first inclusion, suppose $x \in X$ satisfying $q_A(x) < 1$. There exists some $t \in]0,1[$ such that $x \in tA$. In particular, the vector $a = \frac{1}{t}x$ belongs to A, hence

$$ta + (1-t)0 = x \in A.$$

The second inclusion is obvious, so long as the set $\{t > 0 : x \in tA\} \ni 1$ for every $x \in A$.

3. Assume A is balanced. Since $\mathbb{T}A = A$, for any $x \in X$, $\lambda > 0$ and $\gamma \in \mathbb{T}$ one has the equivalences

$$x \in \lambda A \Leftrightarrow \gamma x \in \gamma\lambda A \Leftrightarrow \gamma x \in \lambda A.$$

It is clear by definition that

$$p_A(\gamma x) = p_A(x), \forall x \in X, \forall \gamma \in \mathbb{T}.$$

In addition, since $\{0\} \cup \mathbb{T}$ is absorbing in $\mathbb{K}$, for any $\alpha \in \mathbb{K}$ there exist $\gamma \in \mathbb{T}$ and $|\alpha| \in \mathbb{R}_0^+$ such that $\alpha = \gamma|\alpha|$. Finally,

$$p_A(\alpha x) = p_A(\gamma|\alpha|x) = p_A(|\alpha|x) = |\alpha|p_A(x).$$

□

The second part of proposition 2.4 shows that the set A would be contained between the open and the closed ball of the seminormed space. The most desirable property would be the identification of the topological interior (resp. closure) of the set A with the open ball (resp. closed ball) of that seminormed space. The next proposition will be provided under the most general circumstances; in particular, the previous remark holds.

Proposition 2.5. *Let $A \subset X$ be a non-empty convex absorbing subset of a TVS X. The following conditions are equivalent:*

1. *p_A is continuous.*

2. *$0 \in int(A)$.*

Under those circumstances, $int(A) = \{x \in X : p_A(x) < 1\}$ and $\overline{A} = \{x \in X : p_A(x) \leq 1\}$.

Proof. $\Rightarrow$) Since A is absorbing, in particular $0 \in A$. In addition, by continuity of p_A, $0 \subset \{x \in X : p_A(x) < 1\} \subset \operatorname{int}(A)$.

$\Leftarrow$) If A is a 0-neighbourhood, then $\varepsilon A \subset p_A^{-1}([0, \varepsilon])$ for every $\varepsilon \in \mathbb{R}_0^+$, which means that p_A is continuous at 0; i.e., continuous.

Now suppose that p_A is continuous. On the one hand, it is clear that $\{x \in X : p_A(x) < 1\} \subset \operatorname{int}(A)$. On the other hand, let $x \in \operatorname{int}(A)$ and $\{\lambda_n\}_{n\in\mathbb{N}}$ be a sequence of numbers in $]1, +\infty[$ which converges to 1. Then $\{\lambda_n x\}_{n\in\mathbb{N}} \to x$, hence there exists $N \in \mathbb{N}$ such as $\lambda_N x \in \operatorname{int}(A) \subset A$, and $p_A(x) \leq \frac{1}{x_N} < 1$.

To verify that $\overline{A} = \{x \in X : p_A(x) \leq 1\}$, the inclusion $\supset$ is straightforward since $A \subset \overline{A}$. Given $x \in \overline{A}$, there exists $\{x_n\}_{n\in\mathbb{N}} \subset A$ which converges to x. Hence

$$p_A(x_n) \leq 1, \forall n \in \mathbb{N},$$

and using the continuity of p_A it is concluded that $p_A(x) \leq 1$. □

In a nutshell, is has been proved the equivalence of the following statements for every TVS X and every seminorm $p : X \to \mathbb{R}_0^+$:

1. p is continuous.
2. The set $\{x \in X : p(x) < 1\}$ is open.
3. $0 \in \operatorname{int}(\{x \in X : p(x) < 1\})$.
4. p is continuous at 0.
5. There exists a continuous seminorm $q : X \to \mathbb{R}_0^+$ such that $p \leq q$.

The correspondence between convex balanced 0-neighbourhoods and continuous seminorms shows the equivalence between locally convex spaces and TVS which topology is associated to a family of seminorms.

2.3 Krein-Milman theorem

During the development of the previous chapter, most of the definitions and the results has been stated without considering the dimension of the vector space X. Nevertheless, some of them required that condition, such as the fact that $\operatorname{co}(\cdot)$ maps compact sets into compact sets. This property was of paramount importance in the statement of Carathéodory-Minkowski's theorem, since it concluded that $\operatorname{co}(\operatorname{ext}(A))$ is compact when $A \subset X$ is compact.

The previous operator does not have that property in infinite-dimensional spaces, as the next example shows:

Example 2.1. *Consider the TVS ℓ_2 of square summable sequences. The set*

$$A = \{\frac{1}{k} e_{1,k}\}_{k\in\mathbb{N}} \cup \{0\}$$

of sequences defined in subsection 2.1 and its null limit is compact[2]. *The sequence*

$$x = \left(0, \frac{1}{2}, \frac{1}{1}, \frac{1}{2^2}, \frac{1}{2}, \ldots, \frac{1}{2^n}, \frac{1}{n}, \ldots\right) \in \ell_2$$

and it is easy to check that

$$x = \lim_{n\to+\infty} \left(\sum_{k=1}^{n} \frac{1}{2^k}\right)^{-1} \left(\sum_{k=1}^{n} \frac{1}{2^k} \frac{e_{1,k}}{k}\right).$$

Hence $x \in \overline{co(A)}$, but $x \notin co(A)$ (since $co(A) \subset c_{00}$ and x has infinite many nonzero values), so $co(A)$ is not closed (in particular, compact).

However, under some additional restrictions over the compact set it can be obtained the compactness of $\mathrm{co}(A)$.

Proposition 2.6. *Let $A \subset X$ be a compact subset with a finite number of convex components; i.e., $A = \cup_{i=1}^{N} A_i$ with A_i compact and convex for all $i = 1, \ldots, N$. Then $co(A)$ is compact.*

Proof. According to examples 1.1,

$$\mathrm{co}(A) = \left\{\sum_{i=1}^{n} \lambda_i x_i \colon n \in \mathbb{N}, x_i \in A_i, \lambda_i \in \mathbb{R}_0^+, \sum_{i} \lambda_i = 1\right\}.$$

Now one can reduce the result to the application of proposition 1.2 to the map

$$\begin{array}{cccc} F: & \mathbb{K}^n \times X^n & \longrightarrow & X \\ & ((\lambda_1, \ldots, \lambda_n), (x_1, \ldots, x_n)) & \longmapsto & \displaystyle\sum_{i=1}^{n} \lambda_i x_i \end{array}$$

□

For that reason, in the following theorem it will be considered the topological closure of the convex hull of the extreme points (it was explained in detail after theorem 1.1).

Theorem 2.4 (Krein-Milman)**.** *Let X be a locally convex Hausdorff TVS, and $\emptyset \neq A \subset X$ a compact convex set. Then, $ext(A)$ is nonempty and*

$$\overline{co(ext(A))} = A.$$

Proof. Let

$$\mathcal{F} = \{B \subset A \colon B \text{ is a closed face}\}.$$

Notice that $A \subset \mathcal{F}$, hence $\mathcal{F}$ is nonempty. Defining an order in $\mathcal{F}$ as follows

$$B_1 \leq B_2 \Longleftrightarrow B_2 \subset B_1,$$

[2]In general, it is a basic result from General Topology that, for every countable set $\{x_n\}_{n\in\mathbb{N}}$ which converges to a point x (not necessarily unique), the set $A = \{x_n\}_{n\in\mathbb{N}} \cup \{x\}$ is compact.

given a chain $\mathcal{C} = \{B_i\}_{i\in I}$ it is clear that $\cap_{i\in I} B_i \in \mathcal{C}$ and $B_k \leq \cap_{i\in I} B_i$ for every $k \in I$. By lemma 2.1, there exists a maximal element $K \in \mathcal{F}$. The next step is proving that K consists of a single point, which implies that $\text{ext}(A) \neq \emptyset$.

Suppose that K consists of multiple points, and let $x, y \in K$ with $x \neq y$. By theorem 2.2, there exists a lineal functional $f \in X^*$ for which $f(x) < f(y)$. Then let

$$K_1 = \left\{ x \in K : f(x) = \max_{k\in K} f(k) \right\},$$

which is also a face of K (proposition 1.4) and by proposition 1.6 a face of A ($K_1 \in \mathcal{F}$). But $x \notin K_1$ so long as $f(x) < f(y)$, hence $K_1 \subset K$ and $K < K_1$, contradicting the maximality of K.

Finally, it is going to be checked that $\overline{\text{co}(\text{ext}(A))} = A$. Since A is compact and convex[3], $\overline{\text{co}(\text{ext}(A))} \subset A$. Suppose that there exists $x \in A - \overline{\text{co}(\text{ext}(A))}$. Following a similar strategy, by theorem 2.2 it can be found a linear functional $\varphi \in X^*$ satisfying

$$\varphi(y) < \varphi(x), \quad \forall y \in \overline{\text{co}(\text{ext}(A))}. \tag{2.2}$$

Then define

$$A_1 = \left\{ a \in A : \varphi(a) = \max_{\alpha\in A} \varphi(\alpha) \right\}.$$

This is a face of A due to proposition 1.4, which is nonempty by continuity of φ. Since A_1 is nonempty and convex, it can be selected $a_1 \in \text{ext}(A_1)$. Then a_1 is also an extreme point of A by proposition 1.6, in particular $a_1 \in \text{ext}(A) \subset \text{co}(\text{ext}(A)) \subset \overline{\text{co}(\text{ext}(A))}$. Using the equation (2.2),

$$\varphi(a_1) < \varphi(x) \leq \max_{\alpha\in A} \varphi(\alpha) = \varphi(a_1),$$

which is a contradiction. □

Remarks 2.1.
- *Some authors use the the notation $\overline{\text{co}}(A)$ instead of $\overline{\text{co}(A)}$, even if both are the same subset of X. Indeed, defining $\overline{\text{co}}(A)$ as the least closed convex set which contains A, the inclusion $\overline{\text{co}(A)} \subset \overline{\text{co}}(A)$ is clear by lemma 1.2, and the other inclusion is a consequence of proposition 1.2.*

- *In the previous theorem, it suffices to consider a TVS X on which X^* separates points, since the following result holds:*

  Lemma 2.3. *Suppose X is a TVS on which X^* separates points. Suppose A, B are disjoint, nonempty, compact, convex sets in X. Then there exists $\Lambda \in X^*$ such that*

  $$\sup_{a\in A} Re(\Lambda a) < \inf_{b\in B} Re(\Lambda b).$$

  It is also known that every locally convex TVS X satisfies that X^* separates points, so this is a weaker condition.

- *The convexity of A was only used to show that $\overline{\text{co}(\text{ext}(A))}$ is compact (see proposition 2.6). It may happen in this situation that $\overline{\text{co}(A)}$ contains extreme points which are not*

[3]Notice that it is the first time it is been used that A is convex.

in A. To give an example (see [6]), considering the Banach space $C([0,1])$, the element $\Lambda \in C([0,1])^*$ given by

$$\Lambda f = \int_0^1 f(s)\,ds$$

belongs to $ext(X \cap \overline{co(A)})$, where

$$A = \{\delta_t : t \in [0,1]\} \subset C([0,1])^*, \quad X = span(\{\Lambda\} \cup A) \subset C([0,1])^*, \quad \delta_t(f) = f(t),$$

but $\Lambda \notin A$.

The next theorem shows that this pathology cannot occur if $\overline{\text{co}(A)}$ is compact.

Theorem 2.5 (Milman)**.** *If A is a compact set in a locally convex t.v.s. X and $\overline{co(A)}$ is also compact, then every extreme point of $\overline{co(A)}$ lies in A.*

Proof. Assume that some extreme point p of $\overline{\text{co}(A)}$ is not in A. Then there exists a convex balanced neighbourhood V of 0 in X such that

$$(p + \overline{V}) \cap A = \emptyset. \tag{2.3}$$

Since A is compact, it can be found $\{x_i\}_{i=1}^N$ in X satisfying $A \subset \bigcup_{i=1}^N (x_i + V)$. Each set

$$A_i = \overline{\text{co}(A \cap (x_i + V))},\ i = 1, \dots, N, \tag{2.4}$$

is compact and convex so long as $A_i \subset \overline{\text{co}(A)}$ for each $i = 1, \dots, N$. In addition, $K \subset \cup_{i=1}^N A_i$. Proposition 2.6 shows therefore that

$$\overline{\text{co}(A)} \subset \overline{\text{co}\left(\bigcup_{i=1}^N A_i\right)} = \text{co}\left(\bigcup_{i=1}^N A_i\right),$$

but the opposite inclusion holds also, since $A_i \subset \overline{\text{co}(A)}$ for each $i = 1, \dots, N$. Thus,

$$\overline{\text{co}(A)} = \text{co}\left(\bigcup_{i=1}^n A_i\right). \tag{2.5}$$

In particular,

$$p = \sum_{i=1}^N \lambda_i a_i = \lambda_1 a_1 + (1 - \lambda_1) \frac{\sum_{i=2}^N \lambda_i a_i}{\sum_{i=2}^N a_i}, \quad a_i \in A_i, \forall i = 1, \dots, N. \tag{2.6}$$

The previous equality exhibits p as a convex linear combination of two points of $\overline{\text{co}(A)}$ by equation (2.5). Since p is supposed to be an extreme point of $\overline{\text{co}(A)}$, is is concluded from equation (2.6) that $p = a_1$. Thus, for some $i = 1, \dots, N$:

$$p \in A_i \overset{eq.(2.4)}{\subset} x_i + \overline{V} \subset A + \overline{V},$$

which contradicts equation (2.3). □

2.4 Applications of Krein-Milman theorem

Now that it is known Krein-Milman theorem, it would be desirable to know some applications of it. Even if its beauty can be appreciated on previous sections, the most powerful tools will be given by applications.

Necessary condition for duality

This result is the one which brings, in particular, a necessary condition to check whether a TVS is a dual space or not, in terms of the extreme points of the closed unit ball.

Theorem 2.6. *[Banach-Alaoglu] Let X be a TVS and V a neighbourhood of 0. Then the set*

$$K = \{\Lambda \in X^* \colon |\Lambda x| \le 1, \forall x \in V\}$$

is ω^*-compact.

Proof. Since V is absorbing, for each $x \in X$ there exists $r(x) \in \mathbb{R}$ such that $x \in r(x)V$. Hence

$$|\Lambda x| \le r(x), x \in X, \Lambda \in X^*.$$

Let $D_x = \{\alpha \in \mathbb{K} \colon |\alpha| \le r(x)\}$ and $P = \prod_{x \in X} D_x$. Since each D_x is compact, P is compact with the product topology τ_P by Tychonoff's theorem.

The elements of P are functions f on X (linear or not) that satisfy

$$|f(x)| \le r(x), \forall x \in X,$$

thus $K \subset X^* \cap P$. It follows that K inherits two topologies, ω^* and τ_P. To prove the theorem it has to be checked that:

- K is a closed subset of P with τ_P-topology: Suppose that f_0 is in the τ_P-closure of K. To see that f_0 is linear, let $x, y \in X$ and $\alpha, \beta \in \mathbb{K}$. The set

  $$P_{f_0,\varepsilon} = \{f \in P \colon |(f - f_0)(z)| < \varepsilon, z \in \{x, y, \alpha x + \beta y\}\}, \varepsilon \in \mathbb{R}^+,$$

  is a τ_P-neighbourhood of f_0, hence there exists $f \in K \cap P_{f_0,\varepsilon}$. Since f is linear,

  $$f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y) = (f_0 - f)(\alpha x + \beta y) + \alpha(f - f_0)(x) + \beta(f - f_0)(y)$$

  and then

  $$\left|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)\right| < (1 + |\alpha| + |\beta|)\varepsilon.$$

  The arbitrariness of ε concludes that f_0 is linear. In addition, if $x \in V$ and $\varepsilon > 0$, the same argument shows that there is an $f \in K$ such that $|(f - f_0)(x)| < \varepsilon$, and since $|f(x)| \le 1$ by definition of K, it follows that $|f_0(x)| \le 1$. Hence $f_0 \in K$.

- ω^*-topology and τ_P coincide on K: It has been already proved that $K \subset P$ is compact with τ_P topology. Now fix $\Lambda_0 \in K$ and choose

  $$\{x_i\}_{i=1}^n \subset X, \delta \in \mathbb{R}^+, n \in \mathbb{N}.$$

Considering the families of sets given by

$$\begin{aligned} W_{X^*,n} &= \{\Lambda \in X^* \colon |\Lambda x_i - \Lambda_0 x_i| < \delta, i = 1, \ldots, n\} \\ W_{P,n} &= \{f \in P \colon |(f - \Lambda_0)(x_i)| < \delta, i = 1, \ldots, n\} \end{aligned}$$

they form a local base for the ω^*-topology of X^* at Λ_0 and a local base for the τ_P-topology of P at Λ_0 respectively. Since $K \subset X^* \cap P$, it concludes that

$$W_{X^*,n} \cap K = W_{P,n} \cap K, \quad \forall n \in \mathbb{N}.$$

Finally, P is compact and $K \subset P$ closed with ω^*-topology (equal to τ_P-topology on K), so K is compact with ω^*-topology. □

In particular, when X is a normed space and $V = B_X$, which is also a neighbourhood of 0, it follows that B_{X^*} is ω^*-compact, hence every normed space in which B_X has no extreme points can not be a dual space. This statement is easily obtained applying theorem 2.4 to the set K defined in the previous result.

Representation theorem

The following result is an introduction to the integral representation of compact convex sets (analysed in the next chapter), which contains transcendent results such as Loewner theorem. Furthermore, it allows to prove some important inequalities in convex operator theory.

Proposition 2.7. *Let $\mathcal{F}$ be a compact convex subset of $C(\mathbb{R},\mathbb{R})$ (with the topology of punctual convergence), and suppose that ext$(\mathcal{F}) = \{h_\lambda \colon \lambda \in [0,1]\}$. Then, for every $f \in \mathcal{F}$ there exists a probability measure μ defined in $[0,1]$ such that*

$$f(t) = \int_0^1 h_\lambda(t) d\mu(\lambda).$$

Proof. Let $f \in \mathcal{F}$. Since $\mathcal{F}$ is compact, there exists a net $\{f_i\}_{i \in I}$ which converges to f. Applying theorem 2.4, each f_i is a finite convex linear combination of elements of $\mathcal{F}$,

$$f_i = \sum_{k=1}^{n_i} \alpha_k h_{\lambda_k}, \quad 0 < \sum_{k=1}^{n_i} |\alpha_k|, \sum_{k=1}^{n_i} \alpha_k = 1.$$

However, since

$$f_i = \sum_{k=1}^{n} \alpha_k h_{\lambda_k} = \int_0^1 h_\lambda(t)\, d\mu_i(\lambda), \text{ where } \mu_i(\lambda) = \begin{cases} \alpha_k & \text{if} \quad \lambda = \lambda_k, k = 1, \ldots, n_i \\ 0 & \text{if} \quad \lambda \neq \lambda_k, \forall k = 1, \ldots, n_i \end{cases},$$

it can be considered the net $\{\mu_I\}_{i \in I} \subset B_{P([0,1])}$. By Riesz-Markov-Kakutani representation theorem and theorem 2.6, $B_{P([0,1])}$ is ω^*-compact, hence there exists a subnet $\{\mu_{\sigma(i)}\}$ which converges with ω^*-topology to a probability measure μ. In other words, it has just been verified that

$$f(t) = \int_0^1 h_\lambda(t)\, d\mu(\lambda).$$

□

The Stone-Cech compactification

First of all, it will be recalled the definition of compactification of a locally compact Hausdorff topological space.

Definition 2.5. *Let X be a topological space. A compactification of X is a pair (Y,i), where Y is a compact Hausdorff topological space and $i : X \to Y$ an homeomorphism over $i(X)$ such that $i(X)$ is dense in Y.*

It should be introduced some previous notation to approach the problem. As we did in section 2.1, if $x \in X$, let δ_x be defined by

$$\begin{array}{rccc} \delta_x : & C(X) & \longrightarrow & \mathbb{K} \\ & \delta_x(f) & \longmapsto & f(x) \end{array}$$

It is easy to see that $\delta_x \in C(X)^*$ and $\|\delta_x\| = 1$. Now let Δ be defined by

$$\begin{array}{rccc} \Delta : & X & \longrightarrow & C(X)^* \\ & x & \longmapsto & \delta_x \end{array}$$

If $\{x_i\}_{i\in I}$ is a net in X and $\{x_i\} \to x$, then $\{f(x_i)\} \to f(x)$ for every $f \in C(X)$. This says that $\{\delta_{x_i}\} \xrightarrow{\omega^*} \delta_x$ in $C(X)^*$, hence $\Delta : X \to (C(X)^*, \omega^*)$ is continuous. Is Δ a homeomorphism of X onto $\Delta(X)$?

Proposition 2.8. *The map $\Delta : X \to (\Delta(X), \omega^*)$ is a homeomorphism if, and only if, X is completely regular.*

Proof. $\Rightarrow$) Suppose that Δ is a homeomorphism. Since $B_{C(X)^*}$ is ω^*-compact by theorem 2.6, it is completely regular. Plus, since $\Delta(X) \subset B_{C(X)^*}$, $\Delta(X)$ is also completely regular, hence X is completely regular.

$\Leftarrow$) Assume X is completely regular. If $x_1 \neq x_2$, then by theorem 2.2 there is an $f \in C(X)$ such that $f(x_1) = 1$ and $f(x_2) = 0$; thus $\delta_{x_1}(f) \neq \delta_{x_2}(f)$ and Δ is injective. Now it only lefts to prove that Δ is a homeomorphism over its image, and it is going to be showed appreciating that Δ is an open map.

Let U be an open subset of X and $x_0 \in U$. Since X is completely regular, by theorem 2.2 there is an $g \in C(X)$ satisfying $g(x_0) = 1$ and $g \equiv 0$ on $X - U$. Let

$$V_1 = \{\mu \in C(X)^* : g(u) > 0\}.$$

Then V_1 is ω^*-open in $C(X)^*$ and $V_1 \cap \Delta(X) = \{\delta_x : g(x) > 0\}$. So if $V = V_1 \cap \Delta(X)$, V is ω^*-open in $\Delta(X)$ and $\delta_{x_0} \in V \subset \Delta(U)$. Since x_0 was arbitrary, $\Delta(U)$ is open in $\Delta(X)$. $\square$

Let X any topological space and consider the Banach space $C(X)$. Unless some assumption is made regarding X, it may be that $C(X)$ is *small*. If, for example, it is assumed that X is completely regular, then $C(X)$ has *many* elements. The next result says that this assumption is also necessary in order for $C(X)$ to be *large*.

Theorem 2.7 (Stone-Cech compactification). *If X is completely regular, then there is a compact space βX such that:*

1. *There is a continuous map $\Delta : X \to \beta X$ with the property that $\Delta : X \to \Delta(X)$ is a homeomorphism.*

2. *$\Delta(X)$ is dense in βX.*

3. *If $f \in C(X)$, then there is a continuous map $f^\beta : \beta X \to \mathbb{K}$ such that $f^\beta \circ \Delta = f$.*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\Delta"] & \mathbb{K} \\ \beta X \arrow[ur, "f^\beta"'] & \end{tikzcd}$$

Moreover, if Ω is a compact space having there properties, then Ω is homeomorphic to βX.

Proof. Let $\Delta : X \to C(X)^*$ be the map defined by $\Delta(x) = \delta_x$, and let βX be the ω^*-closure of $\Delta(X)$ in $C(X)^*$. By theorem 2.6 and the fact that $\|\delta_x\| = 1$ for all $x \in X$, βX is compact. Using proposition 2.8, (1) holds, and part (2) is true by definition, hence it only remains to show (3).

Fix $f \in C(X)$ and define $f^\beta : \beta X \to \mathbb{K}$ given by $f^\beta(\tau) = \tau(f)$ for every $\tau \in \beta X$. Clearly f^β is continuous and

$$f^\beta \circ \Delta(x) = f^\beta(\delta_x) = \delta_x(f) = f(x),$$

hence (3) holds.

To show that βX is unique, assume that Ω is a compact space such that:

1. There is a continuous map $\pi : X \to \Omega$ with the property that $\pi : X \to \pi(X)$ is a homeomorphism.

2. $\pi(X)$ is dense in Ω.

3. If $f \in C(X)$, then there is a continuous map $\tilde{f} : \Omega \to \mathbb{K}$ such that $\tilde{f} \circ \pi = f$.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "\pi"] & \mathbb{K} \\ & \Omega \arrow[u, "\tilde{f}"'] \end{tikzcd}$$

Define $g : \Delta(X) \to \Omega$ given by $g(\Delta(x)) = \pi(x)$, i.e., $g = \pi \circ \Delta^{-1}$. The idea is to extend g to a homeomorphism of βX onto Ω.

If $\tau_0 \in \beta X$, then (2) implies that there is a net $\{x_i\}_{i \in I}$ in X such that $\Delta(x_i) \to \tau_0$ in βX. Now $\{\pi(x_i)\}_{i \in I}$ is a net in Ω and since Ω is compact, there is an $\omega_0 \in \Omega$ satisfying that $\{\pi(x_i)\} \overset{cl}{\to} \omega_0$. If $F \in C(\Omega)$, let $f = F \circ \pi$; so $f \in C(X)$ (and $F = \tilde{f}$). Also,

$$\{f(x_i)\}_{i \in I} = \{\delta_{x_i}(f)\}_{i \in I} \to \tau_0(f) = f^\beta(\tau_0),$$

but it is also true that $f(x_i) = (F \circ \pi)(x_i) \overset{cl}{\to} F(\omega_0)$. Hence $F(\omega_0) = f^\beta(\tau_0)$ for any $F \in C(\Omega)$. This implies that ω_0 is the unique cluster point of $\{\pi(x_i)\}_{i \in I}$; thus $\{\pi(x_i)\}_{i \in I} \to \omega_0$.

Let $g(\tau_0) = \omega_0$. It must be shown that the definition of $g(\tau_0)$ does not depend on the net $\{x_i\}_{i \in I}$ in X such that $\{\Delta(x_i)\}_{i \in I} \to \tau_0$, but that is clear since $\{\pi(x_i)\}_{i \in I} \to \omega_0$ for every net $\{x_i\}_{i \in I}$ in X satisfying $\{\Delta(x_i)\}_{i \in I} \to \tau_0$ in βX.

To summarise, it has been shown that there is a function $g : \beta X \to \Omega$ such that if $f \in C(X)$, then $f^\beta = \tilde{f} \circ g$.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\Delta"'] \arrow[dr, "\pi"] & \mathbb{K} \\
\beta X \arrow[r, "g"'] \arrow[ur, "f^\beta"'] & \Omega \arrow[u, "\tilde{f}"']
\end{tikzcd}$$

To show that g is continuous, let $\{\tau_i\}_{i\in I}$ be a net in βX such that $\{\tau_i\}_{i\in I} \to \tau$. If $F \in C(\Omega)$, let $f = F \circ \pi$; so $f \in C(X)$ and $\tilde{f} = F$. Also, $\{f^\beta(\tau_i)\}_{i\in I} \to f^\beta(\tau)$. But $\{(F \circ g)(\tau_i)\}_{i\in I} = \{f^\beta(\tau_i)\}_{i\in I} \to f^\beta(\tau) = (F \circ g)(\tau)$. Proposition 2.8 shows that $\{g(\tau_i)\}_{i\in I} \to g(\tau)$ in Ω, thus g is continuous.

To verify that g is injective, it only has to be noticed that Δ^{-1} and π are homeomorphisms (onto their images). Since $g(\beta X) \supset g(\Delta(X)) = \pi(X)$, $g(\beta X)$ is dense in Ω. But $g(\beta X)$ is compact, so g is bijective; i.e., a homeomorphism. $\square$

The compact set βX obtained in the preceding theorem is called the *Stone-Cech compactification* of X. By properties (1) and (2), X can be considered as a dense subset of βX and the map Δ can be taken to be the inclusion map. With this convention, (3) can be interpreted as saying that every bounded continuous function on X has a continuous extension to βX.

Banach-Stone theorem

In this section it is going to be proved the Banach-Stone theorem, which gives a characterisation of isometries between $C(X)$ and $C(Y)$, where X and Y are compact Hausdorff spaces. It is required for the reader to have a slightly knowledge of adjoint operators; in order to obtain some useful information, (cf. [6], [12]).

To begin with, note that if X and Y are compact spaces, $\tau : Y \to X$ is a continuous map and

$$\begin{array}{rccc} \Lambda : & C(X) & \longrightarrow & C(Y) \\ & f & \longmapsto & f \circ \tau \end{array}$$

then Λ is a bounded linear map and $\|\Lambda\| = 1$. In fact, for any $f \in C(X)$,

$$\begin{aligned} \|\Lambda f\|_\infty &= \sup_{y\in Y}\{|\Lambda f(y)|\} = \sup_{y\in Y}\{|(f\circ\tau)(y)|\} \le \sup_{x\in X}\{|f(x)|\} = \|f\|_\infty \\ \|\Lambda\| &= \sup_{f\in B_{C(X)}}\{\|\Lambda f\|_\infty\} \le \sup_{f\in B_{C(X)}}\{\|f\|_\infty\} \end{aligned}$$

and if $f \equiv \lambda \in B_{\mathbb{K}}$ the equality holds. Moreover, Λ is an isometry if and only if τ is surjective. To check the previous affirmation, it is obvious that the implication ($\Leftarrow$) holds. The converse follows from the fact that $X^* \subset C(X)$ separates points.

Finally, if Λ is a surjective isometry, then τ is injective; i.e., a homeomorphism. Indeed, if $y_0, y_1 \in Y$ and $y_0 \neq y_1$, then there exists $g \in C(Y)$ such that $g(y_0) = 0$ and $g(y_1) = 1$. Let $f \in C(X)$ satisfying $\Lambda f = f \circ \tau = g$. Hence $f(\tau(y_0)) = 0$ and $f(\tau(y_1)) = 1$, so $\tau(y_0) \neq \tau(y_1)$.

To summarize, if $\tau : Y \to X$ is a homeomorphism and $\alpha : Y \to \mathbb{K}$ is a continuous function such that $\alpha(Y) \subset \mathbb{T}$, then

$$\begin{array}{rccc} T : & C(X) & \longrightarrow & C(Y) \\ & f & \longmapsto & \alpha(y) f(\tau(y)) \end{array}$$

is a surjective isometry. The next result gives a converse to this.

Theorem 2.8 (Banach-Stone)**.** *If X and Y are compact Hausdorff spaces and $T : C(X) \to C(Y)$ is a surjective isometry, then there is a homeomorphism $\tau : Y \to X$ and a function $\alpha \in C(Y)$ such that $\alpha(Y) \subset \mathbb{T}$ and*

$$Tf(y) = \alpha(y) f(\tau(y))$$

for all $f \in C(X)$ and $y \in Y$.

Proof. The linear map $T^* : C(Y)^* \to C(X)^*$ is readily seen to be a surjective isometry. Thus, the ω^* continuity of T^*[4] and theorem 2.6 conclude that T^* is a ω^* homeomorphism from $B_{C(Y)^*}$ onto $B_{C(X)^*}$ that distributes over convex combinations. Hence (since T^* is linear and isometric)

$$T^*(\mathrm{E}_{C(Y)^*}) = \mathrm{E}_{C(X)^*}.$$

By theorem 2.3 this implies that for every $y \in Y$ there is a unique $\tau(y)$ in X and a unique scalar $\alpha(y)$ such that $|\alpha(y)| = 1$ and

$$T^*(\delta_y) = \alpha(y)\delta_{\tau(y)}.$$

By the uniqueness, $\alpha : Y \to \mathbb{K}$ and $\tau : Y \to X$ are well-defined functions. It only lefts to prove that α is continuous and τ is a homeomorphism.

To show that α is continuous, let $\{y_i\}_{i\in I}$ a net in Y with $\{y_i\}_{i\in I} \to y$. Then, $\{\delta_{y_i}\}_{i\in I} \xrightarrow{\omega^*} \delta_y$ in $C(Y)^*$, hence

$$\{\alpha(y_i)\delta_{\tau(y_i)}\}_{i\in I} = \{T^*(\delta_{y_i})\}_{i\in I} \xrightarrow{\omega^*} T^*(\delta_y) = \alpha(y)\delta_{\tau(y)}. \tag{2.7}$$

In particular, $\{\alpha(y_i)\}_{i\in I} \to \alpha(y)$.

By equation (2.7), it is clear that given $\{y_i\}_{i\in I}$ a net in Y with $\{y_i\}_{i\in I} \to y$,

$$\{\delta_{\tau(y_i)}\}_{i\in I} = \left\{\alpha(y_i)^{-1}[\alpha(y_i)\delta_{\tau(y_i)}]\right\}_{i\in I} \to \delta_{\tau(y)},$$

and using proposition 2.8 this implies that $\{\tau(y_i)\}_{i\in I} \to \tau(y)$, so that $\tau : Y \to X$ is continuous.

If $y_1, y_2 \in Y$ and $y_1 \neq y_2$, then $\overline{\alpha(y_1)}\delta_{y_1} \neq \overline{\alpha(y_2)}\delta_{y_2}$. Since T^* is injective, it implies that $\tau(y_1) \neq \tau(y_2)$ and so τ is one-to-one. If $x \in X$, then the fact that T^* is surjective implies that there is a $\mu \in M(Y)$ such that $T^*\mu = \delta_x$. It must be that $\mu \in \mathrm{E}_{M(Y)}$ so that $\mu = \beta\delta_y$ for some $y \in Y$ and $\beta \in \mathbb{T}$. Thus

$$\delta_x = T^*(\beta\delta_y) = \beta\alpha(y)\delta_{\tau(y)} \Rightarrow \left\{ \begin{array}{ccc} \beta & = & \overline{\alpha(y)} \\ \tau(y) & = & x \end{array} \right.$$

Therefore τ is a continuous bijection, Y compact, and X Hausdorff. Hence τ must be a homeomorphism. $\square$

[4]Note that $T^*(C(Y)^*) \subset C(X)^*$

Stone-Weierstrass theorem

Last but not least important is this application of Krein-Milman theorem, which concerns the subalgebras A of $C_{\mathbb{R}}(X)$, the real-valued functions on X, which are dense in $\|\cdot\|_\infty$. Just the existence of extreme points in compact convex sets is powerful. We now provide a proof of an analytic result that would seem to have no direct connection to the Krein–Milman theorem. Before the proof of the theorem, recall that $C_{\mathbb{R}}(X)^*$ is the space of real signed measures on X with the total variation norm; that is, for any $\mu \in C_{\mathbb{R}}(X)^*$, there is a set unique up to μ-measure zero sets $B \subset X$, so $\mu \uparrow B \geq 0$, $\mu \uparrow X - B \leq 0$ and $\|\mu\| = \mu(B) + |\mu(X-B)|$.

Theorem 2.9 (Stone-Weierstrass). *Let X be a compact Hausdorff space. Let A be a subalgebra of $C_{\mathbb{R}}(X)$ so that for any $x, y \in X$ and $\alpha, \beta \in \mathbb{R}$, there exists $f \in A$ so $f(x) = \alpha$ and $f(y) = \beta$. Then A is dense in $C_{\mathbb{R}}(X)$ in $\|\cdot\|_\infty$.*

Proof. Define
$$L = \{\mu \in C_{\mathbb{R}}(X) \colon \|\mu\| \leq 1, \, \mu(f) = 0, \forall f \in A\}$$
By theorem 2.6 L is a compact (and convex) set.If $L \neq \{0\}$, then L has an extreme point μ with $\|\mu\| = 1$.

The next step is proving that, if $g \in A \cap B_{C_{\mathbb{R}}(X)}$, then $g = 0$ a.e. $d\mu$ or $(1-g) = 0$ a.e. $d\mu$. First notice that $g\, d\mu \in L$ for any $\mu \in L$ since

$$\int f(g\, d\mu) = \int (fg)\, d\mu = 0, \quad \wedge \quad \|g\, d\mu\| \leq \|g\|_\infty \|\mu\|.$$

If $g \in A \cap B_{C_{\mathbb{R}}(X)}$, then

$$\|g\, d\mu\| + \|(1-g)\, d\mu\| = \int_B g\, d\mu + \int_B (1-g)\, d\mu + \int_{X-B} g\, d\mu + \int_{X-B} (1-g)\, d\mu = \|\mu\|,$$

so μ is a convex combination of elements in L and we have the desired assertion.

The last step consist in showing that $\mathrm{supp}(d\mu)$ is a singleton. If $\mathrm{supp}(d\mu)$ has two points x, y, we can pick $f \in A$ wth $f(x) = 1$ and $f(y) = 2$. Thus $g = \frac{f^2}{\|f\|_\infty^2}$ satisfies $g \in A \cap B_{C_{\mathbb{R}}(X)}$ and $g(x) \in]0, \frac{1}{4}[$, and so $g \in \frac{1}{4} B_{C_{\mathbb{R}}(X)}$ in a neighbourhood U of x. But $\mu(U) \neq 0$, which is a contradiction according to the last paragraph. Finally, $\mu \in L$ and $f(x) = 0$ for all $f \in A$, violating the assumption about $f(x) = \alpha$ can have any real value α.

This contradiction implies that $L = \{0\}$ which, by the Hahn-Banach theorem, implies that A is dense. □

The Stone–Weierstrass theorem does not hold if $C_{\mathbb{R}}(X)$ is replaced by $C(X)$, the complex-valued function. The canonical example of a nondense subalgebra of $C(X)$ with the α, β property is the analytic functions on $\mathbb{D}$.

Bibliography

[1] Krein, M., and Milman, D. *On extreme points of regular convex sets.* Studia Mathematica 9.1 (1940): 133-138. <http://eudml.org/doc/219061>.

[2] Dimitri P. Bertsekas. *Convex analysis and optimization*, 1^{st} edition. Athena Scientific, Belmont, MA, 2003.

[3] Barry Simons. *Convexity: An analytic viewpoint*, 1^{st} edition. Cambridge Tracts in Mathematics and Mathematical Physics, No. 47, Cambridge University Press, New York, 2011.

[4] Javier Pérez online notes http://www.ugr.es/~jperez/docencia/GeomConvexos/cap1.pdf.

[5] Walter Rudin. *Principles of Mathematical Analysis*, 3^{rd} edition. McGraw-Hill Education, 1976.

[6] Walter Rudin. *Functional analysis*, 2^{nd} edition. McGraw-Hill Education, 1991.

[7] Robert R. Phelps. *Lectures on Choquet's theorem*, 2^{nd} edition. Lecture Notes in Mathematics, No. 1757. Springer-Verlag, Berlin, 2001.

[8] Gustave Choquet. *Representation theory*, 1^{st} edition. Lectures on Analysis, No. 2. W.A. Benjamin, New York-Amsterdam, 1969.

[9] Nicolas Bourbaki. *Integration I*, 1^{st} edition. Elements of Mathematics, Ch. 1-6. Springer-Verlag, Berlin, 2001.

[10] James Dugundji. *Topology*, 1^{st} edition. Allyn and Bacon series in Advanced Mathematics. Allyn and Bacon, Boston, 1966.

[11] Jaroslav Lukes. *Integral Representation Theory: Applications to Convexity, Banach Spaces and Potential Theory*, 1^{st} edition. De Gruyter Studies in Mathematics, No. 35. De Gruyter, Berlin, 2010.

[12] John B. Conway. *A course in Functional Analysis*, 2^{nd} edition. Graduate texts in mathematics, No. 96. Springer-Verlag, New York, 1990.

[13] N. Dunford, J. T. Schwartz. *Linear operators, Part I: General Theory*, 1^{st} edition. Interscience, New York, 1952.

[14] B. Cascales, J. M. Mira, J. Orihuela, M. Raja. *Análisis funcional*, 1^{st} edition. Textos universitarios, No. 1. Electolibris, Murcia, 2010.