Choquet theorem and its applications

Master's final project

Author:

Antonio Zarauz Moreno

Advisor:

Dr. Juan Carlos Navarro Pascual

MSc in Mathematics

July, 2017

University of Almeria

Contents

Abstract

This work contains a clear introduction to integral representation theory, showing some classical results combined with a list of elaborated examples. The choice of this topic has not been made by chance; one of the main reasons is the fact that it comprises a lot of areas of pure mathematics, such as Functional Analysis, Topology or Measure Theory. In addition, it is the natural continuation of the degree's final project carried out with the same advisor.

In order to be more precise with the contents of these pages, we proceed to make a detailed exposition of the outline of every chapter:

- The first chapter summarises briefly all the necessary prerequisites, which can be extended in the aforementioned dissertation. The first part contains the definition of Radon measure and topological vector space motivated by the definition of normed space, as well as other considerations concerning weak topologies, convex sets and the convex hull of a set. Secondly, a few classical theorems of Functional Analysis are introduced, so long as they will be taken into account throughout this theory.

- The second chapter is intended to introduce the reader to integral representation theory, revisiting Minkowski-Caratheodory and Krein-Milman theorems in this setting, and explaining in depth the example of the space $C(K)$, which has enough *tools* by itself to be distinguished as an almost trivial situation. We find out that the reformulation of Krein-Milman theorem does not meet our needs, as long as it relies on the closure of the extreme points of the given set, and in some cases compact convex sets in Banach spaces coincides with the closure of its extreme points.

- The third chapter makes a concise review of Choquet theorem, both in metrizable and nonmetrizable cases, using the results shown in the previous chapters. It is particularly interesting to point out the examples carried out in the nonmetrizable setting to motivate the new construction of the support of a measure, as well as the provided applications to Rainwater and Haydon theorems. Among the main reasons to distinguish the metrizable setting, we may point out that, in those spaces, we can always find a Radon measure with total support, the equivalence of the metrizability with the existence of strictly convex functions on A, or the measurability of the set of extreme points of a compact convex subset. We also show a beautiful characterization of the extreme points of a compact convex set in terms of the upper envelope of a bounded function.

- The last chapter is devoted to the study of the uniqueness of the integral representation, which leads us to motivate the definition of simplex in infinite dimensional spaces. Also, a new viewpoint of Riesz theorem is obtained in order to prove the cornerstone of this chapter Choquet-Meyer theorem.

Preliminaries

1

The reader will be assumed to have a strong background in topology, measure theory and functional analysis. However, given the relevance of some concepts and results, it would be useful to introduce them in order to know more precisely the applications we have given to them.

1.1 Topological vector spaces

Topological vector spaces are a generalisation of normed spaces in which we can obtain (under some conditions such as local convexity or metrizability) similar properties and results. 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.

Now we define the convex hull of a set A in a vector space X.

Definition 1.1. *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:*

$$co(A)=\left\{\sum_{i=1}^{n}\lambda_i x_i:\ n\in\mathbb{N},\ x_i\in A,\ \lambda_i\in\mathbb{R}_0^+,\ \forall i=\{1,\dots,n\},\ \sum_{i=1}^{n}\lambda_i=1\right\},\tag{1.1}$$

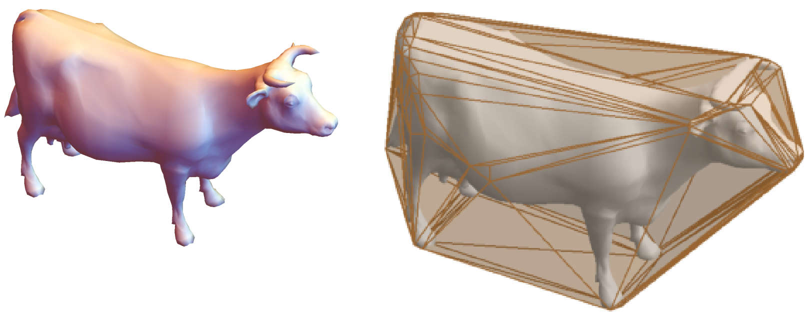


Figure 1.1: Convex hull of a cow.

The properties of this operator on vector spaces are given below.

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 $A = co(A)$.*

When we deal with topological vector spaces, we can go one step further.

Proposition 1.2. *Given a set $A \subset X$, where X is a topological vector space, 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 when X is locally convex.*
4. *$co(\cdot)$ maps compact sets into compact sets.*[2]
5. *$co(\cdot)$ maps precompact sets into precompact sets if X is locally convex.*
6. *If A is convex and $int(A) \neq \emptyset$, then $int(A) = int(\overline{A})$ and $\overline{int(A)} = \overline{A}$.*

For more details about the proof of these results and related topics, one can consult [3] and [5].

1.2 Measure theory

Definition 1.2. *Let X be a set and Σ a σ-algebra over X. A function $\mu : \Sigma \to [-\infty, +\infty]$ is a measure if it satisfies the following conditions:*

1. *For every $E \in \Sigma$, $\mu(E) \geq 0$.*
2. *$\mu(\emptyset) = 0$.*
3. *For every collection $\{E_n\}_{n=1}^{+\infty} \subset \Sigma$ pairwise disjoint, $\mu(\cup_{n=1}^{+\infty} E_n) = \sum_{n=1}^{+\infty} \mu(E_n)$.*

If only the second and third conditions of the definition of measure above are met, and μ takes on at most one of the values $\pm\infty$, then μ is called a signed measure. The members of Σ are called measurable sets.

The stronger condition over X is required, the more enriching properties we can obtain from the measure X. In particular, if we assume that X is a Hausdorff topological space, which is the setting we are concerned on, one can define the Borel σ-algebra over X to be the one generated by the open sets of the topology, $\mathcal{B}(X)$. In order to combine measures and topology, the next definition is introduced.

Definition 1.3. *A Radon measure is a measure μ defined over $\mathcal{B}(X)$ of a Hausdorff topological space X satisfying the following properties:*

[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 $x \notin \frac{1}{2}A + \frac{1}{2}A$.

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

1. *For every point $x \in X$ there exists a neighbourhood which has finite measure.*
2. *For every Borel set U, $\mu(U) = \sup\{\mu(K)\colon K \subset U, K$ compact$\}$.*

The second property is called *inner regularity*. When X is locally compact, every finite Radon measure is also *outer regular*; that is, for every Borel set U,

$$\mu(U) = \inf\{\mu(K)\colon U \supset K, K \text{ compact}\}$$

Definition 1.4. *A topological space X is called a Radon space if every finite measure defined on $\mathcal{B}(X)$ is a Radon measure.*

For instance, Euclidean spaces are Radon spaces, and separable complete metric spaces are Radon spaces as well.

The emerged condition of regularity in Radon measures has something to do with continuous linear functionals with compact support. Let X be a locally compact Hausdorff space and $\mathcal{C}_c(X,\mathbb{K})$ the space of continuous, compactly supported functions from X into $\mathbb{K}$. For any compact subset $K \subset X$, denote by $\mathcal{C}_c(X,K;\mathbb{K})$ the space of those functions which vanishes outside K. This space, endowed with the sup-norm, becomes a Banach space.

Definition 1.5. *A $\mathbb{K}$-linear form μ on $\mathcal{C}_c(X,\mathbb{K})$ is called a Radon measure whenever, for every compact subset $K \subset X$, its restriction to $\mathcal{C}_c(X,K;\mathbb{K})$ is continuous.*

We can prove, with the help of Riesz representation theorem, that any non-negative and bounded Radon measure in this sense is the restriction to $\mathcal{C}_c(X)$ of the integral with respect to a unique (non-finite) Radon measure (in the sense of definition 1.3). To find more information about this, one may check [2].

1.3 Functional analysis

In this section we will essentially introduce the required theorems to develop the integral representation theory. First of all, a couple of results concerning compact convex set on locally convex topological vector spaces are given.

Theorem 1.1 (Minkowski-Carathéodory)**.** *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.

Theorem 1.2 (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*

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

We will devote the following chapter to revisit both results and formulate them into the integral-representation setting. A few applications of Krein-Milman theorem will be required. The first of them, Banach-Alaoglu theorem brings lots of examples of compact sets in infinite-dimensional spaces with the appropiate 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.

Theorem 1.3. *[Banach-Alaoglu] Let X be a TVS and V a neighbourhood of* 0*. Then the set*

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

is ω^*-compact.

An example of description of extreme points is given by Arens-Kelley theorem, who studied the unit ball of $C(K)^*$.

Theorem 1.4 (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, $E_{C(X)^*} = \mathbb{T}\{\delta_x : x \in X\}$.

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, as it would be appreciated in a carefully reading of the proof of the next theorem.

Theorem 1.5 (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 $\|\cdot\|_\infty$.*

2 Introduction to integral representation

After having established Krein-Milman theorem, one can get more powerful results introducing some concepts about measure theory. To be more precise, the main purpose of this chapter is to reformulate the the most profitable results of the previous ones, such as Carathéodory and Krein-Milman theorem, in terms of measures *supported* by the set of extreme points of a given compact convex set, instead of employing convex linear combinations.

2.1 Krein-Milman theorem revisited

To begin with, let us rewrite Minkowski-Carathéodory theorem (1.1) to show with a first example the meaning of a *representing* measure. Given a compact convex subset A of a finite-dimensional topological vector space X ($\dim(X) = n$) and $a \in A$, then there exists $\{x_i\}_{i=1}^{n+1}$ extreme points and $\{\lambda_i\}_{i=1}^{n+1} \subset \mathbb{R}_0^+$ with $\sum_{i=1}^{n+1} \lambda_i = 1$ such that $a = \sum_{i=1}^{n+1} \lambda_i x_i$. If δ_{x_i} is the *point mass* at x_i, we will denote by ε_{x_i} the Borel measure which equals to 1 on any Borel set which contains x_i, and equals 0 otherwise. Let $\mu = \sum_{i=1}^{n+1} \lambda_i \varepsilon_{x_i}$; then μ is a Borel measure on A, $\mu \geq 0$ and $\mu(A) = 1$. Furthermore, for any continuous functional $f \in X^*$,

$$f(a) = \sum_{i=1}^{n+1} \lambda_i f(x_i) = \int_A f \, d\mu.$$

This last assertion is what we mean when we say that μ *represents* $a \in A$.

Definition 2.1. *Suppose that A is a nonempty compact subset of a locally convex space X, and that μ is a probability measure on A. A point $x \in X$ is said to be represented by μ if*

$$f(x) = \int_A f \, d\mu, \quad \forall f \in X^*.$$

It will be denoted sometimes by $\mu(f)$ instead of $\int_A f \, d\mu$ (other terminology: "x is the barycenter of μ", "x is the resultant of of μ").

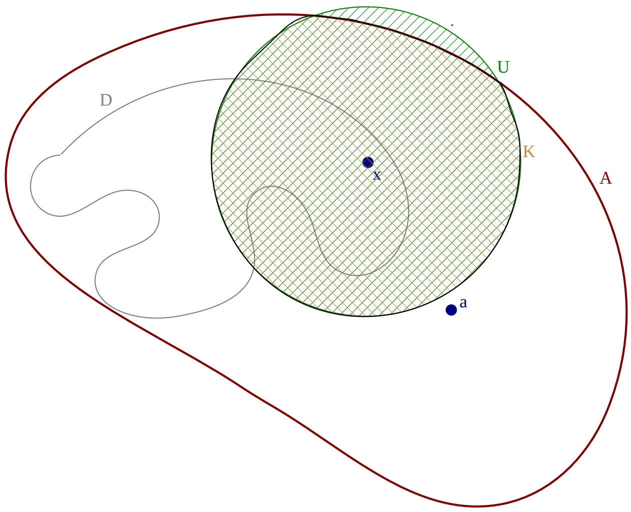
The restriction that X be locally convex is to ensure the existence of sufficiently many functionals in X^* to separate points; this guarantees that there is at most one point represented by μ. Note that each $x \in X$ is trivially represented by ε_x; the most interesting factor brought out by the above example is that, for every compact convex subset A of a finite-dimensional space X, every point of A may be represented by a probability measure which is *supported* by the extreme points of A.

Definition 2.2. *If μ is a nonnegative Borel measure on the compact Hausdorff space X and S is a Borel subset of X, μ is said to be supported by S if $\mu(X \setminus S) = 0$.*

This is one of the fundamental parts we are going to deal with in the development of this chapter. Indeed, the reformulation and extension of Krein-Milman theorem will be given by the existence of a family of Borel measures $\{\mu_a\}_{a \in A}$ supported by the extreme points of a compact convex set A which represent every $a \in A$. It is also important to recall that the extreme points are characterised by the fact that they only have one representing measure.

Proposition 2.1. *Suppose that A is a compact convex subset of a locally convex space X, and let $a \in A$. Then a is an extreme point of A if, and only if, the point mass ε_a is the only probability measure on A which represents a.*

Proof. $\Rightarrow$) Suppose that $a \in \operatorname{ext}(A)$ and that the measure μ represents a. By regularity of μ, to see that μ is supported by $\{a\}$ it suffices to show that $\mu(D) = 0$ for each compact set $D \subset A \backslash \{a\}$. Suppose $\mu(D) > 0$ for some such D; from compactness of D it follows that there is some point $x \in D$ such that $\mu(U \cap A) > 0$ for every neighbourhood U of x. Choose U to be a closed convex neighbourhood of x such that $K := U \cap A \subset A \backslash \{a\}$.



The set K is compact and convex, and $0 < r := \mu(K) < 1$, so long as if $\mu(K) = 1$ the resultant a of μ would be in K. Now define μ_1 and μ_2 Borel measures con A given by

$$\mu_1(B) = \frac{1}{r}\mu(B \cap A), \quad \mu_2(B) = \frac{1}{1-r}\mu(B \cap (A \backslash K)).$$

Let a_i be the resultant of μ_i; since $\mu_1(K) = 1$, $a_1 \in K$ and $a_1 \neq a$. Furthermore,

$$\mu = r\mu_1 + (1-r)\mu_2 \Rightarrow a = ra_1 + (1-r)a_2,$$

contradiction.

$\Leftarrow$) Suppose that $a \notin \operatorname{ext}(A)$; then there exists $n \in \mathbb{N}$, $\{a_i\}_{i=1}^n \subset \operatorname{ext}(A)$ and $\{\lambda_i\}_{i=1}^n \subset \mathbb{R}_0^+$ such that $\sum_{i=1}^n \lambda_i = 1$ and

$$a = \sum_{i=1}^n \lambda_i a_i.$$

Then $\mu = \sum_{i=1}^n \lambda_i \varepsilon_{a_i} \neq \varepsilon_a$ also represents a. □

Example 2.1. *In order to become acquainted with the results we are considering, we dedicate a first (infinite-dimensional) example to the space of continuous functions on a compact Hausdorff space. Let Y be a compact Hausdorff space, $C(Y)$ the Banach space of all continuous real-valued functions on Y and $A \subset X := C(Y)^*$ the set of all continuous linear functions on $C(Y)$ such that $L(1) = 1 = \|L\|$. Then A is a compact convex subset of the locally convex space X in its w^*-topology. In fact, the first equality give us the convexity of A, and the second one the compactness:*

- *Convexity: Let $L, L' \in A$ and $t \in]0,1[$. Then, $L_t = tL + (1-t)L'$ is also a linear functional on X satisfying $L_t(1) = tL(1) + (1-t)L'(1) = 1$ and $\|L_t\| \leq t\|L\| + (1-t)\|L'\| = 1$, but since $\|L_t(1)\| = 1$, the equality holds.*

- *Compactness: A is a closed subset of S_X, hence a closed subset of B_X; then by Banach-Alaoglu theorem (1.3) we conclude that A is a w^*-compact subset.*

The Riesz-Markov-Kakutani theorem assets that to each L in A there corresponds a unique probability measure μ on Y such that

$$L(f) = \int_Y f\,d\mu, \quad \forall f \in C(Y).$$

Also, Y is homeomorphic with the set of extreme points of A by Arens-Kelley theorem (1.4), hence one may consider μ as a probability measure on $\mathcal{B}(A)$ which vanish on the open set $A \setminus Y$[1], so μ is supported by the set of extreme points of A.

One only need to recall that X^*, the space of w^*-continuous linear functionals on $C(Y)^*$, consists precisely of those functionals $L \mapsto L(f)$ ($f \in C(Y)$) in order to see that this is a representation theorem of the type we are considering.

There are some considerations that it should be made under the previous example, as long as they are not necessarily affirmative in more general situations.

- The compactness of the set of extreme points of A let us to use, in particular, that it is a Borel set.

- The representation was unique; this is a more settle detail that will be studied in depth later.

It is clear that any probability measure μ on Y defines, by $f \mapsto \int_X f\,d\mu$, a linear functional on $C(Y)$ which is in X. This fact is true under fairly general circumstances; first one should recall that it suffices for the intersection of an arbitrary family of compact sets to be nonempty that any finite intersection is nonempty as well, which let us to define a finite rank function in the following result.

Proposition 2.2. *Suppose that Y is a compact subset of a locally convex space X, and that $A := \overline{co(Y)}$ is compact. If μ is a probability measure on Y, then there exists a unique point $a \in A$ which is represented by μ, and the function $\mu \mapsto \int_Y f\,d\mu$ is an affine w^*-continuous map from $C(Y)^*$ into A.*

Proof. The goal is showing that the compact convex set A contains a point a such that $f(a) = \int_Y f\,d\mu$ for each $f \in X^*$. Let $f \in X^*$ and consider the closed hyperplane

$$H_f = \{a \in A \colon f(a) = \mu(f)\}.$$

To check that $\cap_{f \in X^*} H_f$ is nonempty, it will be used that A is compact to define, given a finite family $\{f_i\}_{i=1}^n \subset X^*$, the following function (see the previous comments to the statement of the proposition)

$$\begin{array}{rccc} T\colon & A & \longrightarrow & \mathbb{R}^n \\ & a & \longmapsto & (f_1(a), \ldots, f_n(a)). \end{array}$$

T is a continuous linear map, hence TA is compact and convex.

[1] The set Y is now considered via the inclusion $y \mapsto \delta_y$.

It suffices to verify that $p = (\mu(f_1),\dots,\mu(f_n)) \in TA$. If $p \notin TA$, there exists a linear functional on $\mathbb{R}^n$ which separates p and TA; representing this functional by $b = (b_1,\dots,b_n)$ this means that

$$(b,p) > \sup\{(b,Ta)\colon a \in A\}.$$

If we define $g \in X^*$ given by $g(x) = \sum_{i=1}^n b_i f_i$, the previous inequality becomes

$$\int_Y g\,d\mu > \sup g(A).$$

Since $Y \subset A$ and $\mu(Y) = 1$, this is impossible.

To prove the second assertion, let $\{\mu_i\}_{i\in I}$ be a net of probability measures on Y which converges w^* in $C(A)^*$ to the probability measure μ, and let $\{x_i\}_{i\in I}$ and x be their respective resultants. Thanks to the compactness of X, to show that $\{x_i\}_{i\in I} \to x$ it suffices to check that every subnet $\{x_j\}_{j\in J}$ of $\{x_i\}_{i\in I}$ converges to x. But $\{x_j\}_{j\in J} \to y$, which means that

$$f(x_j) = \mu_j(f) \to \mu(f) = f(x), \quad \forall f \in X^*,$$

and since the latter separates points of A, $y = x$. □

Note that the compactness of A may be avoided when the space meets some conditions, such as complete spaces. In order to rewrite Krein-Milman theorem, we describe the closed convex hull of a compact set in terms of barycenters of measures; it is just an extension of Proposition 2.1

Proposition 2.3. *Let Y be a compact subset of a locally convex space X. A point $x \in X$ lies in $A := \overline{co}(Y)$ if, and only if, there exists a probability measure μ on Y which represents x.*

Proof. $\Leftarrow$) Let μ be a probability measure on Y which represents $x \in X$. For every $f \in C(Y)$,

$$f(x) = \mu(f) \le \sup f(Y) \le \sup f(A).$$

Since A is closed and convex, $x \in A$.
$\Rightarrow$) For every $x \in Y$ there exists a net $\{x_i\}_{i\in I}$ in the convex hull of Y such that $\{x_i\}_{i\in I} \to x$. Now we write

$$x_i = \sum_{k=1}^{n_i} \lambda_k x_k^i, \quad \sum_{k=1}^{n_i} \lambda_k = \sum_{k=1}^{n_i} |\lambda_k| = 1, \{x_k^i\}_{k=1}^{n_i} \subset Y$$

hence every x_i can be represented by the measure $\mu_i = \sum_{k=1}^{n_i} \lambda_k \varepsilon_{x_k^i}$. Since the set of probability measures on Y can be identified with a w^*-compact subset of $C(Y)^*$ (Riesz theorem), there exists a subnet $\{\mu_\alpha\}$ of $\{\mu_i\}$ such that

$$\mu_\alpha \xrightarrow{w^*} \mu \in C(Y)^*$$

In particular, for every $f \in X^*$ we have that $f|_Y \in C(Y)$, thus

$$f(x) = \lim f(x_i) = \lim f(x_\alpha) = \lim \int_Y f\,d\mu_\alpha = \int_Y f\,d\mu.$$

□

This proposition makes it easy to reformulate the Krein-Milman theorem. To be more precise, we will show the equivalence between Krein-Milman theorem and the statement:

Every point of a compact convex subset A of a locally convex space X is the barycentre of a probability measure on A which is supported by the closure of the extreme points of A.

Suppose first that Krein-Milman holds and let $x \in A$. If $Y := \overline{\text{ext}}(A)$, then $x \in \overline{\text{co}}Y$. Thanks to the above proposition, x is the barycentre of a probability measure μ on Y. Extending the measure (in the obvious way) to A, the result holds. Conversely, if $x \in A$ and we define Y as in the previous implication, by proposition (2.3) x lies in the closed convex hull of Y, hence in the closed convex hull of the extreme points of A.

It would be convenient to recall the Milman's classical *converse* to the Krein-Milman theorem, which states that the closure of $\text{ext}(A)$ is the smallest closed subset of A generating A.

Theorem 2.1 (Milman)**.** *Suppose that A is a compact convex subset of a locally convex space X, that $Z \subset X$ and that $A = coext(Z)$. Then $ext(A) \subset \overline{Z}$.*

Proof. Let $Y = \overline{Z}$ and suppose that $x \in \text{ext}(A)$. By proposition 2.3 there exists a measure μ on Y which represents x; by proposition 2.1 $\mu = \varepsilon_x$, hence $x \in Y$. □

It is now clear that Krein-Milman reformulation does not meet our needs, so long as it would be desirable for the measure to be supported by the extreme points of the compact convex set, instead of the closure of the mentioned set. In fact, V. L. Klee proved in 1957 that (with the appropiate considerations) that almost every compact convex set in a Banach space *is* the closure of its extreme points.

The problem of finding such measures arises mainly from the measurability of the set of extreme points. When the set is metrizable, the next proposition brings an affirmative answer.

Proposition 2.4. *If A is metrizable, compact convex subset of a topological vector space, then $ext(A)$ is a G_δ set.*

Proof. Consider the sets

$$F_n = \left\{ x \in A \colon x = \frac{1}{2}(y+z),\, y, z \in A,\, d(y,z) \geq \frac{1}{n} \right\}$$

It is obvious that every F_n is a closed set, and that $\text{ext}(A) = (\cup_{n\in\mathbb{N}} F_n)^c = \cap_{n\in\mathbb{N}} F_n^c$. □

3 Choquet theorem

The chapter devoted to this theorem will be carried out in two different sections, according to the metrizability of the compact convex set A. It ensures, among other things, the measurability of the set of extreme points of A, hence it makes easier the study of integral representation theorems. In the latter part, we study the general setting, for which we will have to redefine the concept of support of a measure.

3.1 Metrizable case

First of all, we will denote by $\mathcal{A}$ the set of affine functions on a compact (metrizable) space A, which is a subspace of $C(A)$ that contains the constant functions and separates points of A. The following concept will be a cornerstone in the proof of Choquet's theorem, since it will allow us to define a subadditive and positive homogeneous functional on the subspace $\mathcal{A}$ to extend it applying Hahn-Banach theorem.

Definition 3.1. *If f is a bounded function on A and $a \in A$, the upper envelope of f is the function*

$$\overline{f} = \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq f\}$$

A few visual examples are shown now to introduce the idea intuitively.

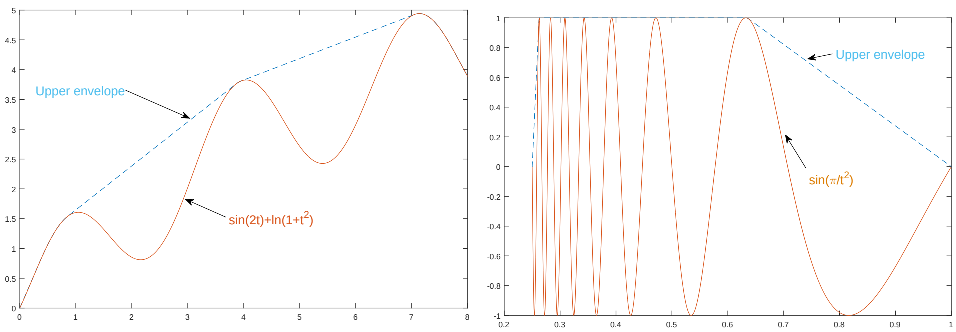


Figure 3.1: Upper envelope of some continuous functions.

The following properties hold for the upper envelope of a bounded function.

Proposition 3.1. *Let f, g be bounded functions on A and $r \geq 0$. Then,*

1. *$\overline{f}$ is concave, bounded and upper semicontinuous.*
2. *$f \leq \overline{f}$ and the equality holds if f is concave and upper semicontinuous.*
3. $\overline{f+g} \leq \overline{f} + \overline{g}$.
4. *If $g \in \mathcal{A}$,* $\overline{f+g} = \overline{f} + \overline{g} = \overline{f} + g$.
5. $|\overline{f} - \overline{g}| \leq \|f - g\|$.
6. $\overline{rf} = r\overline{f}$.

Proof.

1. To check that $\overline{f}$ is concave, for every $a, a' \in A$ and $t \in [0,1]$,

$$\begin{array}{rcl} \overline{f}((1-t)a+ta') & = & \inf\{h((1-t)a+ta') \colon h \in \mathcal{A} \wedge h \geq f\} \\ & \geq & \inf\{(1-t)h(a)+th(a') \colon h \in \mathcal{A} \wedge h \geq f\} \\ & = & (1-t)\inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq f\} + t\inf\{h(a') \colon h \in \mathcal{A} \wedge h \geq f\} \\ & = & (1-t)\overline{f}(a) + t\overline{f}(a') \end{array}$$

   To see that f is bounded, recall that the function $h(x) = \sup f(A) < +\infty$ is a constant (hence affine) function on A and $f(a) \leq h(a)$ for every $a \in A$, thus $\overline{f}(a) \leq h(a) < +\infty$.

   The function $\overline{f}$ is also upper semicontinuous. In fact, for every $\lambda \in \mathbb{R}$ and $x \in \overline{f}^{-1}(]-\infty,\lambda[)$, let g be the element of X^* such that $g(x) = \frac{f(x)}{2}$. Then the set $(g+\frac{\lambda}{2})^{-1}(]-\infty,\lambda[)$ is an open neighbourhood of x in $\overline{f}^{-1}(]-\infty,\lambda[)$.

2. The inequality $f \leq \overline{f}$ is straightforward by definition. To prove the other assertion, we have to notice that $K = \{(a,r) \colon \underline{f(a)} \geq r\}$ is closed and convex in the locally convex space $X \times \mathbb{R}$, so if $f(a_0) < \overline{f}(a_0)$ at some point $a_0 \in A$, the separation theorem would provide the existence of a linear functional $L \in (X \times \mathbb{R})^*$ such that

$$\sup L(K) < \lambda < L(a_0, \overline{f}(a_0)).$$

   In particular, $L(0,1) > 0$ and the affine function $h(a) = r$ if $L(a,r) = \lambda$ on A exists and $f < h$ with $h(a_0) < \overline{f}(a_0)$, contradiction.

3. For every $a \in A$,

$$\begin{array}{rcl} \overline{f+g}(a) & = & \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq (f+g)\} \\ & \leq & \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq f\} + \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq g\} \\ & = & \overline{f}(a) + \overline{g}(a) \end{array}$$

4. For every $a \in A$, since the sum of two affine functions is affine,

$$\begin{array}{rcl} \overline{f+g}(a) & = & \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq (f+g)\} \\ & = & \inf\{(h+g)(a) \colon h \in \mathcal{A} \wedge h \geq f\} \\ & = & g(a) + \inf\{h(a) \colon h \in \mathcal{A} \wedge h \geq f\} \\ & = & g(a) + \overline{f}(a) \end{array}$$

5. Since $f \leq \|f\|$,

$$\overline{f} = \overline{(f-g)+g} \leq \overline{(f-g)} + g,$$

   so $\overline{f} - \overline{g} \leq \overline{f-g} \leq \|f-g\|$. Interchanging f and g yields the result.

6. It is direct from the fact that multiplying a function by a non negative number preserves its condition of concave.

$\square$

Remarks 3.1.

- *The first statement shows in particular that every function $\overline{f}$ is Borel measurable.*
- *To illustrate the first assertion of* (3)*, one only has to consider $g = -f$ and $f(a) \neq 0$ as an example of the given inequality.*

- *An immediate conclusion of* (3) *and* (6) *is that, given a point $x_0 \in A$, the functional p on $C(A)$ given by*
$$p(g) = \overline{g}(x_0)$$
is subadditive and positively homogeneous.

Now we are ready to introduce the metrizable version of Choquet theorem.

Theorem 3.1. *Suppose that A is a metrizable compact convex subset of a locally convex space X, and that $x_0 \in A$. Then there is a probability measure on A which represents x_0 and is supported by the extreme points of A.*

Proof. Since A is metrizable, $C(A)$ is separable, and we may choose a dense unitary sequence in $\mathcal{A}$, namely $\{h_n\}_{n\in\mathbb{N}}$. In particular, it separates points of A. Let $f = \sum_{n\in\mathbb{N}} \frac{1}{2^n} h_n^2$; by CS-compactness of the unit ball of a Banach space, this element belongs to $\mathrm{B}_{C(A)}$. In addition, f is a strictly convex function: suppose that $x, y \in A$, $x \neq y$, then $h_n(x) \neq h_n(y)$ for some $n \in \mathbb{N}$, hence h_n^2 is strictly convex and so is f.

Let B be the subspace $\mathcal{A} + \operatorname{span}\{f\}$. The functional on B given by $h + rf \mapsto h(x_0) + r\overline{f}(x_0)$ is dominated by p (see previous remark) on B, hence by Hahn-Banach theorem there exists a linear functional m on $C(A)$ such that

$$m(g) \leq \overline{g}(x_0) \text{ for } g \text{ in } C(A) \text{ and } m(h + rf) = h(x_0) + r\overline{f}(x_0) \text{ on } B.$$

In order to use the Riesz representation theorem, we need to prove that m is continuous, but that is clear from the fact that if $g \in C(A)$, $g \leq 0$, then $m(g) \leq \overline{g}(x_0) \leq 0$. Thus, there is a nonnegative regular Borel measure μ on X satisfying $m(g) = \mu(g)$ for $g \in C(A)$. It only lefts to prove that μ is a probability measure and is supported by the extreme points of A:

- Probability measure: Since $g \equiv 1$ on A belongs to $\mathcal{A} \subset B$, $1 = m(1) = \mu(1)$.

- Supported by $\operatorname{ext}(A)$: We will show that μ vanishes on the complement of the set $\mathcal{E} = \{a \in A\colon f(a) = \overline{f}(a)\}$, and then the fact that $\mathcal{E} \subset \operatorname{ext}(A)$. On the one hand, $f \leq \overline{f}$, so $\mu(f) \leq \mu(\overline{f})$; on the other hand, if $h \in \mathcal{A}$ and $h \geq f$, then $h \geq \overline{f}$ and consequently
$$h(x_0) = m(h) = \mu(h) \geq \mu(\overline{f}),$$
which implies by definition that $\mu(\overline{f}) \leq \overline{f}(x_0) = m(f) = \mu(f)$.

  The set $\mathcal{E}$ is contained in $\operatorname{ext}(A)$, so long as if given $x = \frac{1}{2}(y+z)$ with $y, z \in A$, then the strict convexity of f implies that
$$f(x) < \frac{1}{2}(f(y) + f(z)) \leq \frac{1}{2}(\overline{f}(y) + \overline{f}(z)) \leq \overline{f}(x).$$

□

It is useful (and beautiful from a geometric viewpoint) to remark that the previous set $\mathcal{E}$ actually coincides with $\operatorname{ext}(A)$. In addition, it will provide us a good motivation to introduce the nonmetrizable setting. To that end, it would be appropriate to identify the measures that *attain the same value in $\mathcal{A}$* in the following sense.

Definition 3.2. *If μ and λ are probability measures such that $\mu(f) = \lambda(f)$ for each $f \in \mathcal{A}$, we will write $\mu \sim \lambda$.*

We already know thanks to proposition (2.1) that ε_x is the only measure that represent any extreme point x, hence the following result will bring us the desired equality $\mathcal{E} = \mathrm{ext}(A)$.

Proposition 3.2. *If f is a continuous function on the compact convex set A, then for each $x \in A$,*

$$\overline{f}(x) = \sup\{\mu(f)\colon\ \mu \sim \varepsilon_x\}.$$

Proof. Let $f'(x) = \sup\{\mu(f)\colon\ \mu \sim \varepsilon_x\}$; we must show that $\overline{f} = f'$. First of all, it is clear from the definition that f' is concave. To prove that f' is upper semicontinuous, fix $r > 0$ and let $\{x_i\}_{i\in I} \subset A$ be a net converging to x with each $f'(x_i) \geq r$. To check that $f'(x) \geq r$, suppose that $\varepsilon > 0$ and for each $i \in I$ choose $\mu_i \sim \varepsilon_{x_i}$ with $\mu_i(f) > r - \varepsilon$. By w^*-compactness, there exists a probability measure μ and a subnet $\{\mu_j\}_{j\in J}$ which converges w^* to μ. If $g \in \mathcal{A}$, then

$$g(x_j) = \mu_j(g) \to \mu(g),$$

and since $g(x_j) \to g(x)$ we see that $\mu(g) = g(x)$ for every $x \in A$ and therefore $\mu \sim \varepsilon_x$. Finally,

$$r - \varepsilon \leq \lim \mu_j(f) = \mu(f) \leq f'(x).$$

Since the choice of ε was arbitrary, we conclude that $f'(x) \geq r$.

Since f' is upper semicontinuous, $\{(x,r)\colon\ f'(x) \geq r\}$ is a closed convex subset of $X \times \mathbb{R}$; using a similar argument as in proposition (3.1), we get that $\overline{f} \leq f'$. On the other hand, if $h \in \mathcal{A}$, $x \in A$ and $h \geq f$, then for any $\mu \sim \varepsilon_x$,

$$h(x) = \mu(h) \geq \mu(f).$$

It follows that $f'(x) \leq h(x)$, so $f' \leq \overline{f}$. □

3.2 Non-metrizable case

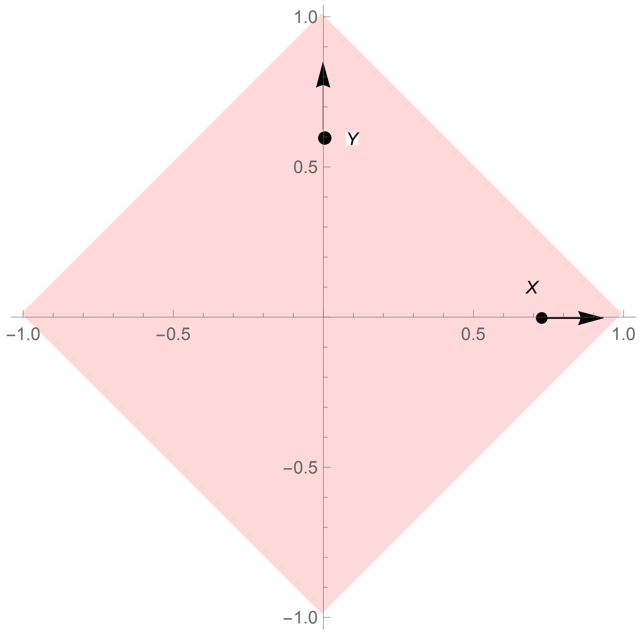
Suppose that A is a nonmetrizable compact convex subset of a locally convex space X. Examples shown by Beshop-de Leeuw [7] reveal that the set of extreme points can be, far from a G_δ set, *as bad as one may desire*, and hence it does not make any sense considering measures supported in a nonmeasurable set. It turns out that the best adaptation for a reformulation of a similar result relays on an adaptation of the definition of *supported by* for Borel measures. An alternative definition may require that μ vanish on every Borel set which is disjoint from the set of extreme points, but (again) Bishop and de Leeuw have shown that it is not always possible to obtain representing measures with this property. If, however, one demands only that μ vanish on the Baire subsets of A which contain no extreme points, then representation results can be obtained.

This step leaves unanswered the question of *approaching* the set of extreme points using Baire sets; in other words, how to how to find a net of measures such that their supports are every time closer to the extreme points. To that purpose, it will be defined an order in the class of regular Radon measures in the following way: given $A \subset X$ a compact convex set and μ, λ Radon measures on A, $\mu \succ \lambda$ if and only if $\mu(f) \geq \lambda(f)$ for every convex function f on A. The agreement of this definition with our original goal may be heuristically verified with the next examples:

Examples 3.1.

- *Consider f a convex function and $A \subset \mathbb{R}^2$ the square of vertexes $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$. For a fixed $(x,y) \in A \backslash \{(0,0)\}$ define*

$$\mu_{(x,y)} = \frac{1}{4}[\delta_{(x,0)} + \delta_{(-x,0)} + \delta_{(0,y)} + \delta_{(0,-y)}].$$



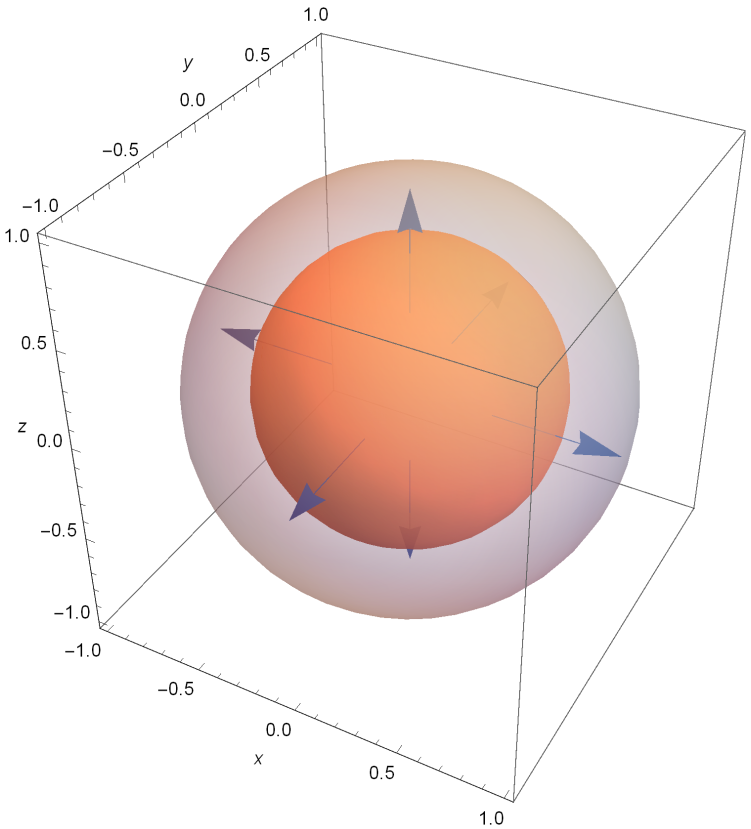
Now we check that if $t_x \leq x$ and $t_y \leq y$, then $\mu_{(x,y)} \succ \mu_{(t_x,t_y)}$, hence for every increasing sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$ which converge to 1, the sequence $\{\mu_{(x_n,y_n)}\}_{n\in\mathbb{N}}$ is increasing as well (with the order $\succ$ previously defined).

$$\begin{aligned}
\int_A f\, d\mu_{(x,y)} &= \frac{1}{4}[f(x,0)+f(-x,0)+f(0,y)+f(0,-y)] \\
&= \frac{1}{4}\left[\frac{x+t_x}{2x}f(x,0)+\frac{x-t_x}{2x}f(-x,0)\right]+\frac{1}{4}\left[\frac{x-t_x}{2x}f(x,0)+\frac{x+t_x}{2x}f(-x,0)\right] \\
&+ \frac{1}{4}\left[\frac{y+t_y}{2y}f(0,y)+\frac{y-t_y}{2y}f(0,-y)\right]+\frac{1}{4}\left[\frac{y-t_y}{2y}f(0,y)+\frac{y+t_y}{2y}f(0,-y)\right] \\
&\geq \frac{1}{4}f\left(\frac{x+t_x}{2x}x+\frac{x-t_x}{2x}(-x),0\right)+\frac{1}{4}f\left(\frac{x-t_x}{2x}x+\frac{x+t_x}{2x}(-x),0\right) \\
&+ \frac{1}{4}f\left(0,\frac{y+t_y}{2y}y+\frac{y-t_y}{2y}(-y)\right)+\frac{1}{4}f\left(0,\frac{y-t_y}{2y}y+\frac{y+t_y}{2y}(-y)\right) \\
&= \frac{1}{4}\left[f(t_x,0)+f(-t_x,0)+f(0,t_y)+f(0,-t_y)\right] \\
&= \int_A f\, d\mu_{(t_x,t_y)}
\end{aligned}$$

- *Let n be a natural number, $A = B_{\mathbb{R}^n}$ and*

$$S^n = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

For any $0 < r \leq 1$, define the measures $\mu_r = \frac{1}{|rS^n|}\chi_{rS^n}\, dx_1 \dots dx_n$. Once again, it can be shown that given $0 < r_1 \leq r_2 \leq 1$, then $\mu_{r_1S^n}(f) \leq \mu_{r_2S^n}(f)$ for any convex function f.



For that purpose note that, for $0 < r \leq 1$, using the generalised spherical coordinates and Fubini's theorem:

$$\begin{aligned}\mu_{rS^n}(f) &= \int_A f\,d\mu = \frac{1}{|rS^n|}\int_{rS^n} f\,dx_1\dots dx_n = C\int_{rS^n} f\,d\theta_1\dots d\theta_{n-1}\\ &= C\int_0^{\pi}\left(\dots\left(\int_0^{\pi}\left(\int_0^{2\pi} f\,d\theta_{n-1}\right)d\theta_{n-2}\right)\dots\right)d\theta_1\end{aligned}$$

The next lemma will provide us a useful inequality:

Lemma 3.1. *Given $0 < r_1 \leq r_2 \leq 1$ and a convex function f,*

$$\underbrace{f(r_1,\theta_1,\dots,\theta_{n-1})}_{f_{r_1^+}}+\underbrace{f(-r_1,\theta_1,\dots,\theta_{n-1})}_{f_{r_1^-}} \leq f(r_2,\theta_1,\dots,\theta_{n-1})+f(-r_2,\theta_1,\dots,\theta_{n-1}).$$

Proof. Using the decomposition $r_1 = \frac{r_2+r_1}{2r_2}r_2 + \frac{r_2-r_1}{2r_2}(-r_2)$ and the convexity of f,

$$f_{r_2^+}+f_{r_2^-} = \left(\frac{r_2+r_1}{2r_2}f_{r_2^+}+\frac{r_2-r_1}{2r_2}f_{r_2^-}\right)+\left(\frac{r_2-r_1}{2r_2}f_{r_2^+}+\frac{r_2+r_1}{2r_2}f_{r_2^-}\right) \geq f_{r_1^+}+f_{r_1^-}.$$

□

Integrating the previous quantity for $\theta_{n-1} \in [0,\pi]$ and any given $0 < r \leq 1$,

$$I = \int_0^{\pi} f_{r^+}\,d\theta_{n-1} + \int_0^{\pi} f_{r^-}\,d\theta_{n-1}$$

If we use the change of variables $(r,\theta_1,\dots,\theta_{n-1}) \mapsto (-r,\theta_1,\dots,\theta_{n-1})$ *in the second integral,*

$$I = \int_0^{\pi} f_{r^+}\, d\theta_{n-1} - \int_0^{-\pi} f_{r^-}\, d\theta_{n-1} = \int_0^{\pi} f_{r^+}\, d\theta_{n-1} + \int_{-\pi}^{0} f_{r^-}\, d\theta_{n-1} = \int_0^{2\pi} f\, d\theta_{n-1}.$$

Hence integrating the equation in (3.1) in $\theta_{n-1} \in [0,\pi]$ *we conclude that*

$$\int_0^{2\pi} f\, d\theta_{n-1} \le \int_0^{2\pi} f\, d\theta_{n-1}$$

The Choquet-Bishop-de Leeuw theorem is shown now to devote the rest of this section to its proof.

Theorem 3.2. *Suppose that A is a compact convex set of a locally convex space X and that $x_0 \in A$. Then there exists a probability measure μ on A which represents x_0 and which vanishes on every Baire subset of A which is disjoint from the set of extreme points of A.*

The necessary tools for proving the theorem version will be exposed below. Let C denote the set of all convex functions on a compact convex set in a locally convex space $A \subset X$. The subspace $C - C$ is a lattice under the usual partial ordering in $C(A)$. Since it contains $\mathcal{A}$, the set of all affine functions on A, $C-C$ separates points of A and is dense in $C(A)$ (in the norm topology) by Stone-Weierstrass theorem. Now we formally introduce the order relation introduced at the beginning of the section.

Definition 3.3. *If λ and μ are nonnegative Borel measures on A, it is said that $\lambda \succ \mu$ if $\lambda(f) \ge \mu(f)$ for every $f \in C$.*

This relation is readily seen to be reflexive and transitive; the antisymmetry comes from the fact that $C-C$ is dense in $C(A)$. Note that if $f \in \mathcal{A}$, then both f and $-f$ lie in C, so that $\lambda \succ \mu$ implies $\lambda(f) = \mu(f)$; i.e., λ and μ represent the same linear functional on $\mathcal{A}$. It is also well worth noting that if $\mu \sim \varepsilon_x$, then $\mu \succ \varepsilon_x$: indeed, if $f \in -C$, then $\overline{f} = f$ and hence

$$f(x) = \inf\{h(x)\colon\, h \in \mathcal{A}, h \ge f\} = \inf\{\mu(h)\colon\, h \in \mathcal{A}, h \ge f\} \ge \mu(f).$$

According to the previous examples, one may expect that the *maximal measures* (maximal with respect to the order "$\succ$") have their support every time *closer* to the set of extreme points of A.

An application of Zorn's lemma leads us to prove that for any given nonnegative measure λ, there exists a maximal measure μ so that $\mu \succ \lambda$.

Lemma 3.2. *If λ is a nonnegative measure, there exists a maximal measure μ with $\mu \succ \lambda$.*

Proof. Suppose that $\lambda \ge 0$ and let $Z = \{\mu\colon\, \mu \ge 0 \wedge \mu \succ \lambda\}$. To find a maximal element of Z, let W be a chain in Z. One may regard W as a net (the directed "index set" being the elements of W themselves) which is contained in the w^*-compact set $\{\mu \ge 0\colon\, \mu(1) = \lambda(1)\}$. Thus, there exists $\mu_0 \ge 0$ and a subnet $\{\mu_i\}_{i\in I} \subset W$ satisfying

$$\{\mu_i\}_{i\in I} \xrightarrow{w^*} \mu_0.$$

If $\mu_1 \in W$, it follows from the definition of a subset that eventually $\mu_i \succ \mu_1$, and hence $\mu_0 \succ \mu_1$. We have proved that μ_0 is an upper bound of W that belongs to Z; by Zorn's lemma, Z contains a maximal element.

It is only left to see that this measure is maximal, but that comes from the fact that if $\eta \succ \mu$, then $\eta \succ \lambda$ and $\eta \in Z$, so $\mu = \eta$. □

The idea of the proof is simple: If $x_0 \in A$, choose a maximal measure μ so that $\mu \succ \varepsilon_{x_0}$. As noted above, μ represents x_0; it remains to show that the maximality of the measure implies that μ vanishes on the Baire sets which contain no extreme points. The first step in that direction is contained in the following result.

Proposition 3.3. *If μ is a maximal measure on A, then $\mu(f) = \mu(\overline{f})$, for each continuous function f on A.*

Proof. Let $f \in C(A)$ be a continuous function and define the functional L on $\operatorname{span}(f)$ by

$$L(rf) = r\mu(\overline{f}).$$

Define the sublinear functional p on $C(A)$ by $p(g) = \mu(\overline{g})$. If $r \geq 0$, then $L(rf) = p(rf)$, while if $rz0$, then

$$0 = \overline{rf - rf} \leq \overline{rf} + \overline{(-rf)} = \overline{rf} - r\overline{f},$$

hence $L(rf) = \mu(r\overline{f}) \leq \mu(\overline{rf}) = p(rf)$. Thus, $L \leq p$ on $\operatorname{span}(f)$ and there exists an extension L' of L to $C(A)$ such that $L' \leq p$.

If $g \leq 0$, then $\overline{g} \leq 0$, so

$$L'(g) \leq p(g) = \mu(\overline{g}) \leq 0.$$

It follows that $L' \geq 0$ and hence there exists a nonnegative measure ν on A such that $L'(g) = \nu(g)$ for each $g \in C(A)$. If g is convex, then $-g$ is concave and $-g = \overline{-g}$, si

$$\nu(-f) \leq p(-g) = \mu(\overline{-g}) = \mu(-g),$$

i.e., $\nu \succ \mu$. Since μ is maximal, we must have $\mu = \nu$, and therefore $\mu(f) = \nu(f) = L(f) = \mu(\overline{f})$. □

It will be seen in the next chapter that the converse to this result is also true (lemma 4.3); furthermore, this last proposition implies that the maximal measure μ is supported by $\{x\colon f(x) = \overline{f}(x)\}$. As shown by proposition (3.2), each of these sets contains the extreme points of A. If C contained a stricly convex function f_0, we would have that

$$\operatorname{ext}(A) = \{x\colon f(x) = \overline{f}(x)\},$$

and the proof will be complete. However, the existence of a strictly convex function on A implies its metrizability, see [10]. Instead, we prove that $\operatorname{ext}(A)$ is the intersection of all the sets of the form $\{x\colon f(x) = \overline{f}(x)\}$ for $f \in C$. Indeed, if $f(x) = \overline{f}(x)$ for each $f \in C$ and $x = \frac{1}{2}(y+z)$ with $y, z \in A$, then

$$f(y) + f(z) \geq 2f(x) = 2\overline{f}(x) \geq \overline{f}(y) + \overline{f}(z) \geq f(y) + f(z),$$

i.e., $f(x) = \frac{1}{2}[f(y) + f(z)]$ for each $f \in C$. The same equality holds for $f \in -C$, hence for each element of $C - C$. Since the latter subspace is dense in $C(A)$, we must have $x = y = z$.

To show that any maximal measure μ vanishes on the Baire sets which are disjoint from $\operatorname{ext}(A)$, it suffices to show that $\mu(D) = 0$ for every compact G_δ set D

disjoint from $\mathrm{ext}(A)$. It will be helpful later if we merely assume that D is a compact subset of a G_δ set which is disjoint from $\mathrm{ext}(A)$. Since D is compact, one can use Uryshon's lemma to choose a nondecreasing sequence $\{f_n\}_{n\in\mathbb{N}}$ of continuous functions on A with

$$-1 \le f_n \le 0, \quad f_n(D) = -1, \quad \lim_{n\to+\infty} f_n(x) = 0, \forall x \in \mathrm{ext}(A).$$

Now we show that if μ is maximal, then $\lim_{n\to+\infty}\mu(f_n) = 0$; as a corollary we have that $\mu(D) = 0$. The next two technical lemmas are required.

Lemma 3.3. *Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a bounded sequence of concave upper semicontinuous functions on A with $\liminf_{n\to+\infty} f_n(x) \ge 0$ for each $x \in \mathrm{ext}(A)$. Then $\liminf_{n\to+\infty} f_n(x) \ge 0$ for every $x \in A$.*

Proof. Assume first that A is metrizable. If $x \in A$, choose a probability measure $\mu \sim \varepsilon_x$ which is supported by $\mathrm{ext}(A)$. By Fatou's lemma, using the hypothesis, $\liminf_{n\to+\infty}\mu(f_n) \ge 0$. Since each f_n is concave and upper semicontinuous, $f_n = \overline{f}_n$, so that

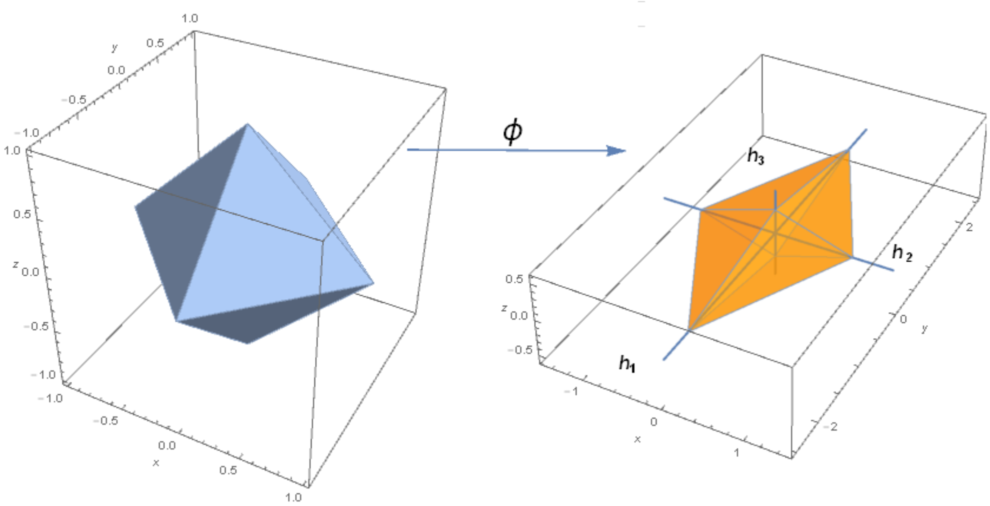
$$f_n(x) = \inf\{h(x)\colon h \in \mathcal{A}, h \ge f_n\} = \inf\{\mu(h)\colon h \in \mathcal{A}, h \ge f_n\} \ge \mu(f_n).$$

Thus, $\liminf_{n\to+\infty} f_n(x) \ge \liminf_{n\to+\infty}\mu(f_n) \ge 0$.

Suppose now, turning to the general case, that $x \in A$ and for each $n \in \mathbb{N}$ choose $h_n \in \mathcal{A}$ satisfying

$$h_n \ge f_n, \quad h_n(x) < f_n(x) + \frac{1}{n}.$$

Define $\phi : A \to \mathbb{R}^{\mathbb{N}}$ by $\phi(y) = \{h_n(y)\}_{n\in\mathbb{N}}$. Since ϕ is affine and continuous, $A' = \phi(A)$ is a compact convex set of the metrizable space $\mathbb{R}^{\mathbb{N}}$.



Given $x' \in A'$, the set $\phi^{-1}(x')$ is compact and convex in A; by the Krein-Milman theorem it has an extreme point y. It is clear that $x' \in \mathrm{ext}(A')$ implies that $y \in \mathrm{ext}(A)$. Since $\pi_n(x') = h_n(y) \ge f_n(y)$, where $\pi_n : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ is the *n-th coordinate* projection, we have

$$\liminf_{n\to+\infty}\pi_n(x') \ge \liminf_{n\to+\infty} f_n(y) \ge 0, \quad \forall x' \in \mathrm{ext}(A').$$

The functions π_n are affine and continuous on the metrizable set A', so from the first part of the proof we conclude that $\liminf_{n\to+\infty}\pi_n(x') \ge 0$ for every $x' \in \mathrm{ext}(A')$.

Taking $x' = \phi(x)$,

$$0 \le \liminf_{n\to+\infty} \pi_n(\phi(x)) = \liminf_{n\to+\infty} h_n(x) = \liminf_{n\to+\infty} f_n(x),$$

which completes the proof. $\square$

Lemma 3.4. *If μ is a maximal measure on A and $\{f_n\}_{n\in\mathbb{N}}$ is a nondecreasing sequence in $C(A)$ such that $-1 \le f_n \le 0$ and $\lim_{n\to+\infty} f_n(x) = 0$ for each $x \in ext(A)$, then $\lim_{n\to+\infty} \mu(f_n) = 0$.*

Proof. Consider the sequence $\{\overline{f}_n\}_{n\in\mathbb{N}}$ of concave upper semicontinuous functions. Since $-1 \le f_n \le \overline{f}_n \le 0$, we have $\lim_{n\to+\infty} \overline{f}_n(x) = 0$ if $x \in \mathrm{ext}(A)$. Plus, The sequence $\{\overline{f}_n\}_{n\in\mathbb{N}}$ is nondecreasing, so $\lim_{n\to+\infty} \overline{f}_n(x)$ exists for each $x \in A$. It follows from the previous lemma that $\lim_{n\to+\infty} \overline{f}_n(x) = 0$ for every $x \in A$. From the Lebesgue bounded convergence theorem it follows that $\lim_{n\to+\infty} \mu(\overline{f}_n) = 0$; from proposition (3.3) we have $\mu(\overline{f}_n) = \mu(f_n)$, which completes the proof. $\square$

Thus, it has been shown that any maximal measure μ on A vanishes on the Baire subsets of $A\backslash\mathrm{ext}(A)$. Plus, something slightly different has been proved:

A maximal measure μ vanishes on any G_δ subset of A contained in $A\backslash ext(A)$.

This is quite relevant, since it shows in particular that a maximal measure is supported by any closed set which contains $\mathrm{ext}(A)$, and hence the Choquet-Bishop-de Leeuw theorem generalises the Krein-Milman theorem.

The next reformulation of the Choquet-Bishop-de Leeuw theorem will be more useful in terms of its applications.

Corollary 3.1. *Suppose that A is a compact convex subset of a locally convex space X, and denote by $\mathcal{S}$ the σ-ring of subsets of A generated by $ext(A)$ and the Baire sets. Then for each point $x_0 \in A$ there exists a probability measure μ on $\mathcal{S}$ that represents x_0 and $\mu(ext(A)) = 1$.*

Proof. By the Choquet-Bishop-de Leeuw theorem, there exists a Borel measure λ which represents x_0 and vanishes on the Baire subsets of $A\backslash\mathrm{ext}(A)$. We only have to extend λ to a nonnegative measure μ on $\mathcal{S}$ and show that $\mu(\mathrm{ext}(A)) = 1$. For that purpose, observe that any set $S \in \mathcal{S}$ is of the form

$$S = (B_1 \cap \mathrm{ext}(A)) \cup (B_2 \cap (X\backslash\mathrm{ext}(A)),$$

where B_1 and B_2 are Baire sets. Defining $\mu(S) = \lambda(B_1)$, then μ is well defined and $\mu(\mathrm{ext}(A)) = \lambda(A) = 1$. $\square$

3.3 Applications

This last section is devoted to a couple of theorems that take the most of this strengthened version of Krein-Milman theorem. The first one characterises the weak convergence of a sequence $\{x_n\}_{n\in\mathbb{N}}$ in a normed space X in terms of its boundedness and the extreme points of the unit ball of its dual space X^*.

One of the first results concerning this idea emerged in the space $C(K)$ of continuous real-valued functions on a compact set K. A classical theorem states that

A sequence $\{f_n\}_{n\in\mathbb{N}} \subset C(K)$ converges weakly to f if and only if the sequence is uniformly bounded and $\lim_{n\to+\infty} f_n(x) = f(x), \quad \forall x \in K$.

Thanks to Arens-Kelley theorem, it's already known that the extreme points of the corresponding unit ball are the functionals of the form $f \mapsto \pm f(x)$, hence this result is seen to be a special case of Rainwater theorem.

Theorem 3.3 (Rainwater)**.** *Let X be a normed space and suppose that $\{x_n\}_{n\in\mathbb{N}} \subset X$. Then $\{x_n\}_{n\in\mathbb{N}} \overset{\omega}{\to} x \in X$ if and only if the sequence is bounded and $\lim_{n\to+\infty} f(x_n) = f(x)$ for every extreme point f of the unit ball of X^*.*

Proof. Let $J : X \to X^{**}$ denote the natural isometry of X into X^{**}. If $\{x_n\}_{n\in\mathbb{N}} \overset{\omega}{\to} x \in X$, the sequence $\{(Jx_n)(f)\}_{n\in\mathbb{N}}$ is bounded and hence by the uniform boundedness theorem shows that $\{Jx_n\}_{n\in\mathbb{N}}$ is bounded in norm; i.e., $\{x_n\}_{n\in\mathbb{N}}$ is bounded in norm.

Conversely, suppose that $\{Jx_n\}_{n\in\mathbb{N}}$ is bounded and that $f(x_n) = (Jx_n)(f) \to (Jx)(f) = f(x)$ for each $f \in \mathrm{E}_{X^*}$; it suffices to show that, for any $g \in \mathrm{B}_{X^*}$, $(Jx_n)(g) \to (Jx)(g)$. Since (B_{X^*}, w^*) is compact (and convex), by the Bishop-de Leeuw theorem there exists a σ-ring $\mathcal{S}$ of subsets of B_{X^*} with $\mathrm{E}_{X^*} \subset \mathcal{S}$ and a probability measure μ on $\mathcal{S}$ such that

$$\mu(\mathrm{B}_{X^*} \backslash \mathrm{E}_{X^*}) = 0 \text{ and } L(g) = \int L\, d\mu, \text{ for each } w^*\text{-continuous affine function } L \text{ on } \mathrm{B}_{X^*}.$$

In particular, $(Jx_n)(g) = \int Jx_n\, d\mu$ and $(Jx)(g) = \int Jx\, d\mu$. Furthermore, $\{Jx_n\}_{n\in\mathbb{N}}$ converges to Jx on B_{X^*} μ-a.e.,so by the Lebesgue bounded convergence theorem $\int Jx_n\, d\mu \to \int Jx\, d\mu$. □

The second application deals with arbitrary Banach spaces. Note that, even if the separability of the space X is not assumed, is a corollary of the theorem.

Theorem 3.4 (Haydon)**.** *Let X be a real Banach space and let K be a w^* compact convex subset of X^* such that ext(K) is norm separable. Then K is the norm closed convex hull of its extreme points.*

Proof. Let $M = \sup\{\|f\| \colon f \in K\}$, suppose that $\varepsilon > 0$ and let $\{f_n\}_{n\in\mathbb{N}}$ be a norm dense subset of ext(K). For each $n \in \mathbb{N}$, Let B_n denote the intersection with K of the closed ball of radius $\frac{\varepsilon}{3}$ centred at f_i. Thus, each B_n is w^*-compact and convex, and

$$\mathrm{ext}(K) \subset \cup_{n\in\mathbb{N}} B_n.$$

Let $f \in K$ and μ a maximal probability measure on K with resultant $r(\mu) = f$. Since $\cup_{n\in\mathbb{N}} B_n$ is a w^* F_σ set, we have $\mu(\cup_{n\in\mathbb{N}} B_n) = 1$. Let $n_0 \in \mathbb{N}$ such that, if $D = \cup_{n=1}^{n_0} B_n$, then $\mu(D) > 1 - \frac{\varepsilon}{3M}$. Then μ can be decomposed as

$$\mu = \lambda\mu_1 + (1-\lambda)\mu_2,$$

where $\lambda = \mu(D)$ and μ_1, μ_2 are probability measures defined on K by

$$\lambda\mu_1 = \mu|_D, \quad (1-\lambda)\mu_2 = \mu|_{(K\backslash D)}.$$

Then $f = r(\mu) = \lambda r(\mu_1) + (1-\lambda) r(\mu_2)$. Since $r(\mu_2) \in K$, we have

$$\|f - \lambda r(\mu_1)\| = (1-\lambda)\|r(\mu_2)\| \leq \frac{\varepsilon}{3M} M = \frac{\varepsilon}{3}.$$

In light of the fact that μ_1 is a probability measure supported by D, the point $r(\mu_1)$ lies in the convex hull of D, which is w^*-compact. Hence $r(\mu_1) = \sum_{n=1}^{n_0} \lambda_n g_n$, where $g_n \in B_n$, $\sum_{n=1}^{n_0} \lambda_n = \sum_{n=1}^{n_0} |\lambda_n| = 1$. Let $h = \sum_{n=1}^{n_0} \lambda_n f_n \in \mathrm{co}(\mathrm{ext}(K))$; this point satisfies that $\|r(\mu_1) - h\| \leq \frac{\varepsilon}{3}$, and consequently

$$\|f - h\| \leq \|f - \lambda r(\mu_1)\| + \|f - (1-\lambda) r(\mu_1)\| + \|r(\mu_1) - h\| \leq \varepsilon.$$

Thus, $\mathrm{co}(\mathrm{ext}(K))$ is norm dense in K. $\square$

4 Uniqueness

As we described in the previous chapter, there is a slight difference between the concept of support of a measure in the metrizable and nonmetrizable setting. It would be desirable to obtain a theorem which characterises those compact convex sets A with the property that to each point there exists a unique measure that represents the point and is supported by the set of extreme points. Choquet has proved such a theorem for metrizable compact convex sets, but there is no affirmative answer in the general case. On the other hand, Choquet and Meyer have characterised those sets with the property that to each point there corresponds a unique *maximal* measure that represents the point. Since maximal measures are *supported* by the set of extreme points, it would seem that this answers the question; Mokobodzi showed that uniqueness of maximal representing measures does no imply uniqueness of representing measures which vanishes on Baire subsets of $A \backslash \mathrm{ext}(A)$.

4.1 Introduction: Cones

For our present purposes, it will be convenient to assume that our compact convex set A is contained in a closed hyperplane that misses the origin. There is not lost of generality in this assumption, so long as one can embed the locally convex space X in $X \times \mathbb{R}$ with the product topology via $X \times \{1\}$.

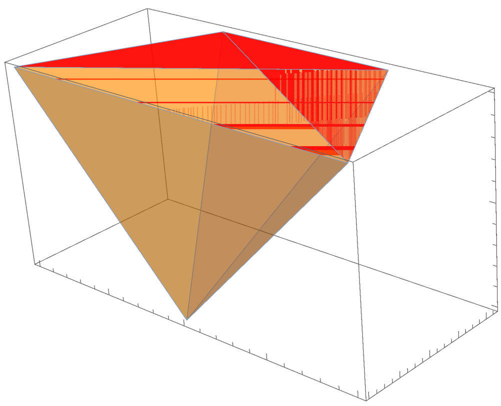


Figure 4.1: Example of embedding of a compact convex set $A \subset \mathbb{R}^2$.

The reason for doing this is that the question of uniqueness is more natural when A is the *base* of a convex cone P; i.e., when there is a convex cone P such that $y \in P$ if and only if there exists a unique $\alpha \geq 0$ and $x \in A$ such that $y = \alpha x$. If A is contained in a hyperplane which misses the origin, take $P = \tilde{A} = \{\alpha x \colon \alpha \geq 0, x \in A\}$ is the cone generated by A. Reciprocally, if A is a base for a cone P, then $0 \notin A$, so by the separation theorem there exists a continuous linear functional f on X and $\beta > 0$ such that $f(x) \geq \beta$ for all $x \in A$. Then the set

$$A' = \{x' \in P \colon f(x') = \beta\}$$

is also a base for P which is affinely homeomorphic to A under the map $x \mapsto \frac{\beta x}{f(x)}$, for every $x \in A$.

It seems natural to induce a (traslation invariant) partial ordering from the structure of a cone P into the space X: given $x, y \in X$, $x \geq y$ if and only if $x - y \in P$. If P has a base B, then $P \cap (-P) = \{0_X\}$, hence

$$x \geq y \wedge y \geq x \Rightarrow x = y.$$

In addition, if x and y are in the subspace $P - P$ generated by P, then there exists $z \in P$ so that $z \geq x$ and $z \geq y$; the element z is called the *upper bound* of x and y in P. We say that $z (= x \vee y)$ is the *least* upper bound for x and y if $z \leq w$ whenever $w \geq x$ and $w \geq y$; this notion clearly extends the idea of supremum.

If a convex set A, not necessarily compact, is the base of a cone $\tilde{A}$, we call A a *simplex* if the space $\tilde{A} - \tilde{A}$ is a *vector lattice* in the ordering induced by A; that is, if each pair $x, y \in \tilde{A} - \tilde{A}$ has a least upper bound $x \vee y$ in $\tilde{A} - \tilde{A}$. Equivalently, $\tilde{A} - \tilde{A}$ is a vector lattice if and only if each pair x, y has a *greatest lower bound* (dual concept of least upper bound), denoted by $x \wedge y = -(-x \vee -y)$. Now we establish a relation between the ordering in $\tilde{A}$ and $\tilde{A} - \tilde{A}$.

Proposition 4.1. *Given a convex set A, $\tilde{A} - \tilde{A}$ is a vector lattice if, and only if, $\tilde{A}$ is a lattice.*

Proof. The implication ($\Leftarrow$) has already been introduced. On the other hand, suppose that each pair $x, y \in \tilde{A}$ has a least upper bound $x \vee y$ in $\tilde{A}$. If

$$x = x_1 - x_2, y = y_1 - y_2 \in \tilde{A} - \tilde{A},$$

let $z = (x_1 + y_2) \vee (y_1 + x_2) - (x_2 + y_2)$; it only lefts to prove that z is the least upper bound of x, y in $\tilde{A} - \tilde{A}$. Since

$$z - x = (x_1 + y_2) \vee (y_1 + x_2) - (x_1 + y_2),$$

we have that $z \geq x$; similarly $z \geq y$. If $w = w_1 - w_2 \in \tilde{A} - \tilde{A}$ with $w \geq x$, $w \geq y$, we must show that $w \geq z$. The first two inequalities imply that

$$w_1 + w_2 + y_2 \geq w_2 + x_1 + y_2, \quad w_1 + x_2 + y_2 \geq w_2 + x_2 + y_1.$$

Using the translation invariance,

$$w - z = (w_1 + x_2 + y_2) - [(w_2 + x_2 + y_1) \vee (w_2 + x_1 + y_2)] \geq 0.$$

$\square$

It is readily seen that being a simplex is an *intrisic* property of A; that is, if A is contained in a hyperplane which misses the origin in X, if A_1 is similarly situated in X_1, and if there exists a one-to-one affine map of A onto A_1, then this map may be extended in the obvious way to a one-to-one additive, order preserving map which carries $\tilde{A}$ onto $\tilde{A}_1$, so that on of these cones is lattice if and only if the other is a lattice.

4.2 Riesz representation theorem revisited

One of the milestones in the first example of this work, related to the space $C(K)$ of continuous functions on a compact set, was the Riesz representation theorem. It gave us, in particular, the uniqueness of the measure which represents any given point. In order to extend it to a more general setting, it will be reformulated in terms of simplices. Suppose that Y is a compact Hausdorff space and let A be the compact convex set of all probability measures on Y. As it was noted in the introduction, the Riesz theorem can be formulated as follows: To each point of A there exists a unique representing measure which is supported by $\mathrm{ext}(A) = \phi(Y)$. The uniqueness assertion can be considered to be a consequence of the fact that A is a simplex; i.e., that *the cone of all nonnegative measures on Y has the set A of probability measures as a base and is a lattice in the usual ordering*.

The next lemma is required to prove the main result of this section due to Choquet and Meyer. In particular, the *decomposition lemma*, which is the last assertion of the following result.

Lemma 4.1. *Suppose that V is a vector lattice.*

1. *For each $x, y, z \in V$, $(x+z) \wedge (y+z) = (x \wedge y) + z$.*
2. *If $x \geq 0$, $y \geq 0$ and $z \geq 0$, then $(x+y) \wedge z \leq (x \wedge z) + (y \wedge z)$.*
3. *If $\{x_i\}_{i \in I}$ and $\{y_j\}_{j \in J}$ are finite sequences of nonnegative elements of V, and if*
$$\sum_{i \in I} x_i = \sum_{j \in J} y_j,$$
then there exists $\{z_{ij}\}_{\substack{i \in I \\ j \in J}}$ such that for every $i \in I$ and $j \in J$,
$$x_i = \sum_{j \in J} z_{ij}, \quad y_j = \sum_{i \in I} z_{ij}.$$

Plus, as a consequence of the monotone convergence theorem, one can obtain this result.

Lemma 4.2. *If $f \in C(A)$, let*
$$G = \{g \colon g \in -C \wedge g \geq f\}.$$
Then $\overline{f} = \inf\{f : g \in G\}$, G is directed (downward) by $\geq$, and $\mu(\overline{f}) = \inf\{\mu(g) \colon g \in G\}$ for any nonnegative measure μ on A.

As an application of this lemma, the converse to proposition (3.3) holds.

Lemma 4.3. *A positive measure μ on A is maximal if, and only if, $\mu(f) = \mu(\overline{f})$, for each continuous convex function f on A.*

Proof. In light of proposition (3.3), it lefts to prove that if $\mu(f) = \mu(\overline{f})$ for each $f \in C$, then μ is maximal. Choose a maimal measure λ with $\lambda \succ \mu$. Then for $f \in C$,
$$\lambda(\overline{f}) = \lambda(f) \geq \mu(f) = \mu(\overline{f}).$$
If $g \in -C$, then $\lambda(g) \leq \mu(g)$, so by the previous lemma
$$\lambda(\overline{f}) = \inf\{\lambda(g) \colon g \in -C, g \geq f\} \leq \inf\{\mu(g) \colon g \in -C, g \geq f\} = \mu(\overline{f}).$$
Since $C - C$ is dense in $C(A)$, it follows that $\lambda(f) = \mu(f)$ for every $f \in C$, and hence μ is maximal. $\square$

Another technical result concerning vector lattices is needed. Suppose that P_1 and P_2 are cones in a vector space X with $P_1 \subset P_2$, with partial orderings $\leq_1$ and $\leq_2$. It is said that P_1 is an *hereditary subcone* of P_2 if $y \in P_1$, $x \in P_2$ and $x \leq_2 y$ imply $x \in P_1$.

Lemma 4.4. *If P_2 is a lattice in the ordering $\leq_2$ and P_1 is an hereditary subcone of P_2, then P_1 is a lattice in the ordering $\leq_1$.*

Proof. Suppose that $x, y \in P_1$ and let $z = x \vee y$. Then $z \leq_2 x$, so $z \in P_1$, and we will show that if $w \leq_1 x$ and $w \leq_1 y$ then $w \leq_1 z$. Since $P_1 \subset P_2$, it is clear that $w \leq_2 x$ and $w \leq_2 y$, hence $0 \leq_2 w \leq_2 z$. IT follows that $z - w \in P_2$ and that $z - w \leq_2 z$; by the hereditary property, $z - w \in P_1$, so that $w \leq_1 z$. □

As a consequence, the set Q of all nonnegative maximal measures on A is a subcone of the cone P of al nonnegative measures on A. Plus, the convex set

$$Q_1 = \{\mu \colon \mu \in Q, \mu(A) = 1\}$$

is a base for Q, and Q_1 is a simplex. One can think of this result as an injection of our set A and the respective cone P in the space $C(A)^*$; in fact, this is one of the implications of the Choquet-Meyer uniqueness theorem.

To prove this corollary, we check first that Q is closed for sums and multiplication by nonnegative scalars. Suppose that λ and μ are maximal measures; by proposition (4.3), $(\lambda + \mu)(f) = (\lambda + \mu)(\overline{f})$ for each continuous convex function f on A, so $\lambda + \mu$ is maximal. Similarly, $r\mu$ is maximal for every $\mu \in Q$ and $r \geq 0$. Since Q_1 is the intersection of Q with the probability measures, it is clearly a base for Q using the previous considerations. To see that Q_1 is a simplex, we must show that Q is a lattice in its natural ordering. By the previous lemma it suffices to show that Q is hereditary in P. Suppose then that $0 \leq \lambda \leq \mu$ and $\mu \in Q$. LEt λ_1 be a maximal measure with $\lambda_1 \succ \lambda$. Then

$$\lambda_1 + (\mu - \lambda) \succ \lambda + (\mu - \lambda) = \mu;$$

since μ is maximal, $\mu = \lambda_1 = \mu - \lambda$, so that $\lambda = \lambda_1 \in Q$.

4.3 Choquet-Meyer theorem

Theorem 4.1 (Choquet-Meyer)**.** *Suppose that A is a nonempty compact convex subset of a locally convex space X. The following assertions are equivalent:*

1. *A is a simplex.*
2. *If $f \in C$, then $\overline{f} \in \mathcal{A}$.*
3. *If μ is a maximal measure on A with resultant x, and if $f \in C$, then $\overline{f}(x) = \mu(f)$.*
4. *For every $f \in C$ and $g \in C(A)$, $\overline{f + g} = \overline{f} + \overline{g}$.*
5. *For each $x \in A$ there is a unique maximal measure μ_x such that $\mu_x \sim \varepsilon_x$.*

Proof. (1) $\Rightarrow$ (2): It will be required a sharpened version of proposition (3.2).

Lemma 4.5. *If $f \in C(A)$, then for every $x \in A$,*

$$\overline{f}(x) = \sup\{\mu(f) \colon \mu \textit{ is discrete and } \mu \sim \varepsilon_x\}.$$

Proof of the lemma:. By a *discrete measure* we mean a measure which is a finite convex combination of measures of the form ε_y. It follows for the aforementioned proposition that it suffices to prove that, for every $f \in C(A)$, $x \in A$, $\mu \sim \varepsilon_x$ and $\varepsilon > 0$, there exists a discrete measure λ such that $\lambda \sim \varepsilon_x$ and $\mu(f) - \lambda(f) < \varepsilon$.

For that purpose, cover A by a finite number of closed convex neighbourhoods $\{U_i\}_{i=1}^n$ such that $|f(y) - f(z)| < \frac{\varepsilon}{2}$ for every $y, z \in U_i \cap A$. Let

$$V_1 = U_1 \cap A,\ V_i = (U_i \cap A) \backslash (V_1 \cup \dots V_{i-1}),\ i = 2, \dots, n.$$

Then the family $\{V_i\}_{i=1}^n \subset \mathcal{B}$ is pairwise disjoint; for those that $\mu(V_i) \neq 0$, we can obtain a probability measure λ_i on A supported by V_i, by defining

$$\lambda_i(B) = \mu(V_i)^{-1} \mu(B \cap V_i),\ \forall B \in \mathcal{B}.$$

Let x_i be the resultant of λ_i. Since V_i $V_i \subset U_i \cap A$, the latter must contain x_i. Define $\lambda = \sum \mu(V_i) \varepsilon_{x_i}$; if h is a continuous affine function on A, then

$$\lambda(h) = \sum \mu(V_i) \lambda_i(h) = \sum \int_{V_i} h \, d\mu = \mu(h) = h(x),$$

hence $\lambda \sim \varepsilon_x$. Furthermore,

$$\mu(f) - \lambda(f) = \sum \left(\int_{V_i} f \, d\mu - \mu(V_i) f(x_i) \right) = \sum \left(\int_{V_i} (f - f(x_i)) \, d\mu \right) < \varepsilon \sum \mu(V_i) = \varepsilon.$$

△

Suppose that $x_1, x_2 \in A$, $\alpha_1, \alpha_2 > 0$ with $\alpha - 1 + \alpha_2 = 1$ and $f \in C$. If $z = \alpha_1 x_1 + \alpha_2 x_2$, it will be shown that $\overline{f}(z) = \alpha_1 \overline{f}(x_1) + \alpha_2 \overline{f}(x_2)$. Since $\overline{f}$ is concave, it suffices to show that $\overline{f}$ is also convex. According to the previous lemma, choose a discrete measure $\mu \sim \varepsilon_z$; then there exists a finite convex combination in A so that

$$\alpha_1 x_1 + \alpha_2 x_2 = z = \sum_{j=1}^n \beta_j y_j.$$

Applying lemma (4.1) to the elements $\alpha_i x_i$ and $\beta_j y_j$ of $\tilde{A}$ one can choose $z'_{ij} \in \tilde{A}$ with

$$\alpha_i x - i = \sum_{j=1}^n z'_{ij}, \quad \beta_j y_j = z'_{1j} + z'_{2j}.$$

Each $z'_{ij} = \gamma_{ij} z_{ij}$ with $\gamma_{ij} \geq 0, z_{ij} \in A$, and hence $x_i = \sum_{j=1}^n \alpha_1^{-1} \gamma_{ij} z_{ij}$ is a convex combination of elements of A. It follows that the right side represents a discrete measure $\mu_i \sim \varepsilon_{x_i}$, and therefore $\overline{f}(x_i) \geq \mu_i(f) = \sum_{j=1}^n \alpha_i^{-1} \gamma_{ij} f(z_{ij})$. On the other hand, $\mu(f) = \sum_{j=1}^n \beta_j f(y_j)$ and for each $j = 1, \dots, n$,

$$f(y_j) = f(\beta_1^{-1} \gamma_{1j} z_{1j} + \beta_j^{-1} \gamma_{2j} z_{2j}) \leq \beta_j^{-1} \gamma_{1j} f(z_{1j}) + \beta_j^{-1} \gamma_{2j} f(z_{2j}),$$

so $\mu(f) \leq \alpha_1 \mu_1(f) + \alpha_2 \mu_2(f) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2)$. Taking the supremum over all the possible discrete measures gives the conclusion.
(2) $\Rightarrow$ (3): If μ is maximal and $f \in C$, then $\mu(f) = \overline{f}(x)$. Since f is affine and upper semicontinuous, the next lemma implies that $\mu \sim \varepsilon_x$, and $\mu(\overline{f}) = \overline{f}(x)$.

Lemma 4.6. *Suppose that f is an affine upper semicontinuous function on A and that $\mu \sim \varepsilon_x$. Then $\mu(f) = f(x)$.*

Proof of the lemma:. It suffices to prove that the family H of all $h \in \mathcal{A}$ such that $H > f$ is directed downward and that $f = \inf\{h\colon h \in H\}$. Indeed, if this be true, then we have $\mu(f) = \inf\{\mu(H)\colon h \in H\}$ for any μ; in particular, if $\mu \sim \varepsilon_x$, then $\mu(f) = \inf\{h(X)\colon h \in H\} = f(x)$.

To see that H is directed downward, suppose that $h_1 > f$ and $h_2 \in f$ with $h_!, h_2 \in \mathcal{A}$; it will be found $h \in \mathcal{A}$ with $h \leq h_1$ and $h \leq h_2$. To this end, define

$$J = \{(x,r)\colon x \in A, r \leq f(x)\}, J_i = \{(x,r)\colon x \in A, r = h_i(x)\}, i = 1, 2.$$

Since f is affine and upper semicontinuous, J is closed and convex, while the continuity of h_i implies that J_i is compact. Furthermore, $J \cap \operatorname{co}(J_1 \cup J_2) = \emptyset$, with J_3 compact. By the separation theorem applied to 0 and $J_3 - J$, there exists a linear functional L on $X \times \mathbb{R}$ with $L(J) < \inf L(J_3) = \alpha$. The function $h : A \to \mathbb{R}$ defined by $L(x, h(X)) = \alpha$ meets our needs. A similar argument shows that $f = \inf\{h\colon h \in H\}$. △

(3) $\Rightarrow$ (4): Suppose that $f, g \in C$ and that $x \in A$. Choose a maximal measure $\mu \sim \varepsilon_x$; by hypothesis we get

$$(\overline{f+g})(x) = \mu(f+g) = \mu(f) + \mu(g) = \overline{f}(X) + \overline{g}(x).$$

(4) $\Rightarrow$ (5): Suppose that $x \in A$ and consider the functional defined for $f \in C$ by $f \mapsto \overline{f}(x)$. This is positive-homogeneous, and by hypothesis it is also additive. From this it follows that $m(f-g) = \overline{f}(x) - \overline{g}(x)$ defines a linear functional m on the subspace $C - C$, and using the properties of the upper envelope we have $|m(f-g)| \leq \|f - g\|$. Thus, m is uniformly continuous on the dense subspace $C - C$ of $C(A)$ and hence there exists a unique extension to a continuous linear functional of norm at most 1 on $C(A)$. Since $m(1) = 1$, this functional is given by a probability measure, which we denote by μ_x. Since, for $f \in C$, we have $\mu_x(f) = m(f) = \overline{f}(x)$, using proposition (3.2), $\mu_x(f) = \sup\{\mu(f)\colon \mu \sim \varepsilon_x\}$; i.e., $\mu_x > \mu$ whenever $\mu \sim \varepsilon_x$. It follows that μ_x is the unique maximal measure which represents x.

(5) $\Rightarrow$ (1): It is an immediate conclusion of the observation made before the theorem. □

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] Michael Simonnet. *Measures and probabilities*, 1^{st} edition. Lecture Notes in Mathematics, No. 1757. Springer-Verlag, New York, 1996.

[3] H. H. Schaefer. *Topological vector spaces*, 2^{nd} edition. Graduate texts in mathematics, No. 3. Springer-Verlag, New York, 1999.

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

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

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

[7] E. Beshop, K De Leeuw. *The representations of linear functionals by measures on sets of extreme points*, Annales de l'institut Fourier, tome 9 (1959), p. 305-331.

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

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

[10] Michel Hervé. *Sur les representations intégrales a l'aide des points extrémaux dans un ensemble compact convexe metrizable*, C.R. Acad. Sci (Paris) 253 (1976), 377-385.

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

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

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