———— * ————

Convex sets and their integral representations

———— * ————

Bachelor Thesis in Mathematics.
Department of Mathematical Science,
University of Copenhagen

Bachelorprojekt i Matematik.
Institut for Matematiske Fag,
Københavns Universitet

08.06.12

Author:
Anna Munk Ebbesen

Advisor:
Magdalena Musat

Abstract

This Bachelor Thesis in Mathematics is concerned with establishing integral representation theorems for convex sets in terms of their extreme points. The idea of considering a point in a metrizable compact convex set X as being represented by a probability measure concentrated on the set of extreme points $\mathrm{Ext}(X)$ of X is due to Choquet, and the whole theory is called Choquet theory.

In the special case when $\mathrm{Ext}(X)$ is a closed set, the representation theorem is equivalent to the Krein-Milman theorem, which is a fundamental result in functional analysis with a broad range of applications. In this thesis, we prove the Krein-Milman theorem and explain how to reformulate it as an integral representation theorem. We also prove Choquet's theorem in the metrizable case.

Resumé

Dette bachelorprojekt i Matematik handler om at etablere sætninger for repræsentationsintegraler for konvekse mængder i forhold til deres ekstreme punkter, $\mathrm{Ext}(X)$. Choquet var den første, der betragtede et punkt i en kompakt, metriserbar og konveks mængde X som værende repræsenteret af et sandsynlighedsmål, der er koncentreret på $\mathrm{Ext}(X)$. Teorien omkring dette emne kaldes meget naturligt for Choquet teori.

Hvis $\mathrm{Ext}(X)$ er en lukket mængde, så er integralrepræsentationen ækvivalent med Krein-Milmans Theorem, der er et afgørende resultat i funktionalanalyse med en bred vifte af anvendelser. I dette projekt beviser vi Krein-Milmans Theorem og omformulerer det til vores nye sprog med repræsentationsintegraler. Derudover viser vi Choquets Theorem i det metriserbare tilfælde.

Introduction

Before reading this thesis the reader should gather some basic knowledge of functional analysis. In particular the different versions of the Hahn-Banach Theorems, the Riesz Representation Theorem and properties of nets, since we will juggle around with these things quite alot but will not spent time discussing them.

With this clarified we can continue to state the main question of this thesis.

> "*If X is a compact convex subset of a locally convex Hausdorff topological vector space E and $x \in X$, does there exist a probability measure μ on X which is supported by the extreme points of X and which represents x?*"

The first stop on our quest for the answer is at one of the fundamental results about compact convex sets in $\mathbb{R}^n$ (due to Minkowski). It is that their points can be represented as convex combinations of extreme ones. In Section 1 we formulate and prove a more refined version of this theorem, known as the Carathéodory theorem. To do that, we need some basic results from convex analysis, for this [2] has been very helpful.

Next we move on to the Krein-Milman Theorem i Section 2. The proof of this is mostly inspired by [3].

Third stop on the road is the introduction to integral representation theorems. This section is a big jump in the right direction to the answer of our question. We reformulate the Krein-Milman Theorem from Section 2 in our new language. This section is based on [5].

Final step in our quest is Section 4. Here we prove Choquet's Theorem in the metrizable case and give an answer to our main question. This section is also based on [5].

Acknowledgements

I would like to thank my advisor Magdalena Musat for a lot of work and help. She managed to be (more or less) patient with me, even though I made it quite difficult at times. I will give an ekstra applaus for her big work effort in the last week, where I took most of her time. Furthermore, I would like to thank Kristian Knudsen Olesen for constant help with both big and small questions.

Rasmus Bryder and Marcus Dorph De Chiffre have read and commented on parts of the thesis, which have only made it better so a great thank to them as well.

A special thank to Katrine Lykke Jensen, Kevin Aguyar Brix, Mikkel Strunge, Nikolaj Borum and Lise Frellesen who all took their time to look out for me and make me dinner in the last week, when deadline was nearby.

Last but not least a great thank to Andreas Midjord for company in the late nights and for always answering even the most stupid questions.

Contents

1 Carathéodory's Theorem

A more refined version of Minkowski's Theorem, known as the Carathéodory theorem, is the object of study of this section. In particular, we will prove Carathéodory's Theorem. To do this, we need to know a few things about convex analysis of which, in particular the notions of *relative interior* and *supporting hyperplane* for a convex set in E. Throughout this section, when we write E, it will be an n-dimensional real vector space - this means E is homeomorphic to $\mathbb{R}^n$.

1.1 Background on convexity

Before we can do anything interesting we need to develop some background knowledge on convex sets.

Recall that a non-empty subset X of E is ***convex*** if $\alpha x + (1-\alpha)y \in X$ for all $x, y \in X$ and $0 < \alpha < 1$.

In the following proposition we state some important and very useful facts about convexity.

Proposition 1.1. *Let X_i be a convex set*

(a) The intersection $\bigcap_{i \in I} X_i$ of any collection $\{X_i \mid i \in I\}$ is convex.

(b) The set λX_i is convex for any $\lambda \in \mathbb{R}$.

(c) The closure, $\overline{X_i}$, and the interior, $\operatorname{int}(X_i)$, is convex.

Proof. The proof is straightforward and will be omitted. □

In the definition of convexity we had $\alpha \in (0, 1)$; if, instead, we let $\alpha \in \mathbb{R}$ we get the defintion of an affine set. More precisely,

Definition 1.2. *A non-empty subset X of E is **affine** if $\alpha x + (1-\alpha)y \in X$ for all $x, y \in X$ and $\alpha \in \mathbb{R}$.*

With this at hand we can continue to define the convex (respectively affine) hull of X. It is defined to be the smallest convex (respectively affine) set containing X. We denote the convex hull of X by $\operatorname{conv}(X)$ and the affine hull of X by $\operatorname{aff}(X)$. It turns out that we can describe them in another way, as well. Let X be a subset of E. Then

- $\operatorname{conv}(X) = \{\sum_{i=1}^n \alpha_i x_i \mid x_i \in X, \alpha_i > 0, \sum_{i=1}^n \alpha_i = 1, n \in \mathbb{N}\}$,
- $\operatorname{aff}(X) = \{\sum_{i=1}^n \alpha_i x_i \mid x_i \in X, \alpha_i \in \mathbb{R}, \sum_{i=1}^n \alpha_i = 1, n \in \mathbb{N}\}$.

We note that $\operatorname{aff}(X) = \operatorname{aff}(\operatorname{conv}(X))$.

As a consequence of the properties set forth in Proposition 1.1, we obtain the following result, which will be useful later.

Lemma 1.3. *For any non-empty subset X of E,*

$$\overline{\operatorname{conv}(X)} = \overline{\operatorname{conv}(\overline{X})}.$$

Proof. It is clear that $\overline{\operatorname{conv}(X)} \subseteq \overline{\operatorname{conv}(\overline{X})}$. For the other inclusion, notice that $\operatorname{conv}(X) \supseteq X$, hence $\overline{\operatorname{conv}(X)} \supseteq \overline{X}$. As $\overline{\operatorname{conv}(X)}$ is convex by Proposition 1.1 (c), it will also contain $\operatorname{conv}(\overline{X})$. Furthermore, as $\overline{\operatorname{conv}(X)}$ is closed, it will also contain $\overline{\operatorname{conv}(\overline{X})}$. □

Another important concept associated to a convex set $X \subseteq E$ is the one of *relative interior.*

Definition 1.4. *Let X be a subset of E. A point $x \in X$ is said to be a* ***relative interior*** *point of X if there exists an open ball in E centered at x with radius r, $B(x, r) \subseteq E$, such that $B(x, r) \cap \operatorname{aff}(\mathrm{X}) \subseteq X$, i.e., x is an interior point of X relative to $\operatorname{aff}(X)$. The collection of all relative interior points of X is denoted by $\operatorname{ri}(X)$.*

This concept can seem rather strange and meaningless at first. But as the following example shows, $\operatorname{int}(X)$ and $\operatorname{ri}(X)$ need not be equal but sometimes they will be.

Example 1.5. Let $E = \mathbb{R}^3$ and consider the closed unit disc D. For every $x \in D$ there does not exist an $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq D$, hence $\operatorname{int}(D) = \emptyset$.

The relative interior of D is the interior of D according to $\operatorname{aff}(D)$ which in this case will be a plane. Hence, the relative interior of D is the open unit disc. Thus, $\operatorname{int}(D) \neq \operatorname{ri}(D)$.

If we on the other hand consider the closed unit sphere S, then the interior of S will be the open unit sphere. The affine hull of S is all of E and hence the relative interior is the same as the interior.

Proposition 1.6. *Let X be a non-empty convex set in E.*

(a) If $x \in \operatorname{ri}(X)$ and $y \in \overline{X}$, then all points on the line segment connecting x and y, except possibly y, belong to $\operatorname{ri}(X)$.

(b) $\operatorname{ri}(X)$ is a non-empty convex set and

$$\operatorname{aff}(\operatorname{ri}(X)) = \operatorname{aff}(X).$$

(c) $\overline{X} = \overline{\operatorname{ri}(X)}$.

Proof. For part (a), since $x \in \operatorname{ri}(X)$ there exists $\varepsilon > 0$ such that the open sphere $S = \{z \in X \mid \|z - x\|_\infty < \varepsilon\}$ satisfies $S \cap \operatorname{aff}(X) \subseteq X$. Given $y \in \overline{X}$, let for $\alpha \in (0, 1]$, $x_\alpha = \alpha x + (1 - \alpha)y$ and $S_\alpha = \{z \in X \mid \|z - x_\alpha\|_\infty < \alpha\varepsilon\}$.

If $y \in X$, then every point of $S_\alpha \cap \operatorname{aff}(X)$ is a convex combination of some point in $S \cap \operatorname{aff}(X)$ and y. Hence, by the convexity of X, $S_\alpha \cap \operatorname{aff}(X) \subseteq X$, implying that $x_\alpha \in \operatorname{ri}(X)$.

Next, suppose $y \in \overline{X} \backslash X$. Consider a sequence $(x_k)_{k \geq 1} \subseteq X$ that converges to y as $k \to \infty$, and set $x_{k,\alpha} = \alpha x + (1-\alpha)x_k$, for all $\alpha \in (0, 1]$. With the same argument as above, $\{z \in X \mid \|z - x_{k,\alpha}\|_\infty < \alpha\varepsilon\} \cap \operatorname{aff}(X) \subseteq X$ for every $k \geq 1$. Since $x_{k,\alpha} \to x_\alpha$ as $k \to \infty$, we have for large enough k that

$$\left\{z \in X \mid \|z - x_\alpha\|_\infty < \frac{\alpha\varepsilon}{2}\right\} \subseteq \{z \in X \mid \|z - x_{k,\alpha}\|_\infty < \alpha\varepsilon\}.$$

It follows that $\{z \in X \mid \|z - x_\alpha\|_\infty < \alpha\varepsilon/2\} \cap \operatorname{aff}(X) \subseteq X$. Thus $x_\alpha \in \operatorname{ri}(X)$.

Part (b) uses part (a) to show the convexity of $\operatorname{ri}(X)$. For the rest of part (b), assume without loss of generality that $0 \in X$. The affine hull of X is a subspace of E whose dimension will be denoted m.

If $m = 0$ then both X and $\operatorname{aff}(X)$ consist of a single point, which is unique. Hence $\operatorname{ri}(X) \neq \emptyset$ and $\operatorname{aff}(X) = \operatorname{aff}(\operatorname{ri}(X))$.

If $m > 0$, we can find m linearly independent vectors $z_1, \dots, z_m \in X$ that span $\operatorname{aff}(X)$, thus forming a basis for $\operatorname{aff}(X)$. Consider

$$Y = \left\{ x \in X \;\middle|\; x = \sum_{i=1}^{m} \alpha_i z_i, \sum_{i=1}^{m} \alpha_i < 1,\ \alpha_i > 0,\ i = 1, \dots, m \right\}.$$

We want to show that Y is open relative to $\operatorname{aff}(X)$, i.e., for every vector $y \in Y$, there exists an open ball B centered at y such that $B \cap \operatorname{aff}(X) \subseteq Y$. To do that, fix $y \in Y$ and let $x \in \operatorname{aff}(X)$. Let Z be the $n \times m$-matrix whose columns are the vectors $z_1, \dots, z_m$ and let $\bar{\alpha}, \alpha$ be the suitable unique m-dimensional vectors such that $y = Z\bar{\alpha}$ and $x = Z\alpha$. This is possible due to [2, Proposition 1.1.10 (iv)]. The matrix $Z^t Z$ is symmetric and positive definite, and hence by [2, Proposition 1.1.11 (d)], we have for some scalar γ (independent of x and y),

$$\begin{aligned} \|x - y\|^2 = \|Z\alpha - Z\bar{\alpha}\|^2 &= \|Z(\alpha - \bar{\alpha})\|^2 \\ &= (Z(\alpha - \bar{\alpha}))^t (Z(\alpha - \bar{\alpha})) \\ &= (\alpha - \bar{\alpha})^t Z^t Z(\alpha - \bar{\alpha}) \\ &\geq \gamma \|\alpha - \bar{\alpha}\|^2. \end{aligned}$$

Since $y \in Y$, the corresponding vector $\bar{\alpha}$ lies in the open set

$$A = \left\{ (\alpha_1, \dots, \alpha_m) \;\middle|\; \sum_{i=1}^{m} \alpha_i < 1,\ \alpha_i > 0,\ i = 1, \dots, m \right\}.$$

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

This means, that Y contains the intersection of $\operatorname{aff}(X)$ and an open ball centered at y. Hence, Y is open relative to $\operatorname{aff}(X)$. Note that all points of Y are relative interior points of X, and hence $\operatorname{ri}(X) \neq \emptyset$. Our construction of Y gives us that $\operatorname{aff}(Y) = \operatorname{aff}(X)$, and since $Y \subseteq \operatorname{ri}(X)$, we see that $\operatorname{aff}(\operatorname{ri}(X)) = \operatorname{aff}(X)$.

For part (c), we have $\operatorname{ri}(X) \subseteq X$ and hence $\overline{\operatorname{ri}(X)} \subseteq \overline{X}$. Conversely, let $y \in \overline{X}$ and $x \in \operatorname{ri}(\mathrm{X})$. Due to (b), such an x exists. If $x = y$, we are done. Assume instead that $x \neq y$. From (a), we know that $\alpha x + (1-\alpha)y \in \operatorname{ri}(X)$ for all $\alpha \in (0,1)$. Consider the sequence

$$\left(\frac{1}{k}x + \left(1 - \frac{1}{k}\right)y\right)_{k \geq 1} \subseteq \operatorname{ri}(X).$$

This sequence converges to y and hence $y \in \overline{\operatorname{ri}(X)}$. Meaning that $\overline{X} \subseteq \overline{\operatorname{ri}(X)}$, and thus we are done. □

We continue with another concept closely associated with convex sets.

Definition 1.7. *A non-empty convex subset X of E is called a **convex body** if $\operatorname{int}(X) \neq \emptyset$.*

We notice that if X is a convex body, then

$$\overline{\operatorname{int}(\mathrm{X})} = \overline{X}. \tag{1}$$

Indeed, as $\operatorname{int}(X)$ is a non-empty open set, we deduce that $\operatorname{aff}(\operatorname{int}(X)) = E$. Since $\operatorname{aff}(X) \supseteq \operatorname{aff}(\operatorname{int}(X))$, we have $\operatorname{aff}(X) = E$. By the definition of relative interior, it follows that $\operatorname{int}(X) = \operatorname{ri}(X)$. Upon taking the closure on both sides, an application of Proposition 1.6 (c) proves the claim.

Now, given an arbitrary convex subset X of E, let $E_0 = \operatorname{aff}(X)$. Then E_0 is an m-dimensional space for some $m \leq n$ (n is the dimension of E) and X is a convex body relative to E_0 by definition. Hence X is a convex body in E if and only if $E_0 = E$.

1.2 Extreme points

The notion of *extreme points* of a convex set is the central theme of this thesis. We recall the definition.

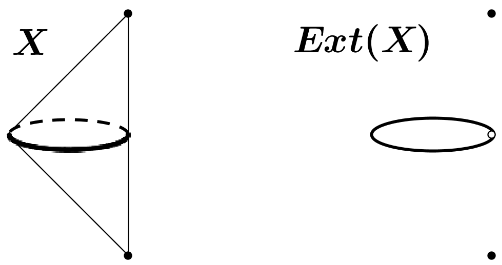
Definition 1.8. *Let X be a non-empty convex set. A point $x \in X$ is an **extreme point** of X if the condition*

$$x = \alpha x_1 + (1-\alpha)x_2$$

for $x_1, x_2 \in X$ and $0 < \alpha < 1$ implies that $x = x_1 = x_2$. The set of extreme points of X is denoted by $\operatorname{Ext}(X)$.

The following example shows that the set of extreme points of a convex set need not be convex, nor closed.

Example 1.9. Let X consist of the union of two closed solid cones sharing a common base (as drawn on the left hand-side of the picture). The drawing on the right represents the extreme points of X. Clearly, $\mathrm{Ext}(X)$ is neither closed, nor convex.



One can see from the picture that $X = \mathrm{conv}(\mathrm{Ext}(X))$, and in particular, it follows that $\mathrm{conv}(\mathrm{Ext}(X))$ is closed. This are general features of compact convex sets in $\mathbb{R}^n$. In fact, a more refined version of this result (namely, that every point of a convex compact subset X of $\mathbb{R}^n$ is a convex combination of at most $n+1$ extreme points in X) is the statement of Carathéodory's theorem which we prove in Section 1.4. The proof relies on the notion of *supporting hyperplane*, which we discuss next.

1.3 Supporting hyperplane

Definition 1.10. *A **supporting hyperplane** H for a non-empty convex set X in E is an affine hyperplane in E such that X is contained in one of the two closed half spaces of E determined by H.*

Recall that any affine hyperplane $H \subseteq E$ can be described as

$$H = \{x \in E \mid f(x) = \alpha\}\,,$$

for some linear functional $f\colon E \to \mathbb{R}$ and some $\alpha \in \mathbb{R}$. In other words, H is a supporting hyperplane for $X \subseteq E$ if there exists $f\colon E \to \mathbb{R}$ linear, $f \neq 0$ and $\alpha \in \mathbb{R}$ such that $H = \{x \in E \mid f(x) = \alpha\}$ and $f(y) \leq \alpha$ for all $y \in X$.

Lemma 1.11. *If H is a supporting hyperplane for $X \subseteq E$ and if $H \cap X \neq \emptyset$, then* $\mathrm{Ext}(H \cap X) \subseteq \mathrm{Ext}(X)$.

Proof. Let $H = \{x \in E \mid f(x) = \alpha\}$ as above, and $x \in \mathrm{Ext}(H \cap X)$. To show that $x \in \mathrm{Ext}(X)$, let $y_1, y_2 \in X$ and $0 < \lambda < 1$ be such that $x = \lambda y_1 + (1-\lambda)y_2$. Due to linearity of f, $\alpha = f(x) = \lambda f(y_1) + (1-\lambda)f(y_2)$. As $f(y_j) \leq \alpha$ for $j = 1, 2$, we must have $f(y_1) = f(y_2) = \alpha$. Hence, $y_1, y_2 \in H \cap X$. Now, we use the fact that $x \in \mathrm{Ext}(H \cap X)$, to deduce from $x = \lambda y_1 + (1-\lambda)y_2$ that $y_1 = y_2 = x$. This shows that $x \in \mathrm{Ext}(X)$, as wanted. □

Proposition 1.12. *If $X \subseteq E$ is a convex body and if $x \in \partial X$, then there exists a supporting hyperplane H for X such that $x \in H$.*

Proof. Let $x \in \partial X$. Since $\mathrm{int}(X)$ is a non-empty convex set, there exists a hyperplane H separating x and $\mathrm{int}(X)$, due to the Hahn-Banach Separation Theorem. Then, the corresponding closed half space containing x, also contains $\overline{X} = \overline{\mathrm{int}(X)}$ (due to (1) on p. 8), showing that H is a supporting hyperplane for X. □

1.4 Carathéodory's Theorem

With all this at hand we are finally ready to prove the announced main result of this section.

Theorem 1.13 (Carathéodory's Theorem)**.** *Let E be an n-dimensional space and let X be a non-empty convex compact subset of E. Then* $\mathrm{Ext}(X) \neq \emptyset$ *and each point $x \in X$ is a convex combination of at most $n+1$ extreme points of X.*

Proof. We will prove this by induction on n. If $n = 1$, then $E = \mathbb{R}$ and $X = [\alpha, \beta]$ for some $\alpha, \beta \in \mathbb{R}$. In this case $\emptyset \neq \mathrm{Ext}(X) = \{\alpha, \beta\}$ and every $x \in [\alpha, \beta] = X$ is of the form $x = \lambda\alpha + (1-\lambda)\beta$ for some $0 \leq \lambda \leq 1$. Hence a convex combination of at most $2 = (n+1)$ extreme points.

Now, suppose $n \geq 2$, and that the claim has been proved whenever E has dimension at most $n-1$. If X is not a convex body, then (see comment following Definition 1.7) $X \subseteq E_0$ for some proper affine subspace E_0 of E and $\dim(E_0) < \dim(E) = n$ and the claim will follow from the induction hypothesis. We may therefore assume that X is a convex body.

First, let $x \in \partial X$, which is non-empty since X is compact. Let $H \subseteq E$ be a supporting hyperplane for X such that $x \in H$. Such a hyperplane exists due to Propostion 1.12. Then $H \cap X$ is convex and compact in H, and $\dim(H) = n-1$. By the induction hypothesis, $\mathrm{Ext}(H \cap X) \neq \emptyset$ and x is a convex combination of at most n extreme points of $H \cap X$. By Lemma 1.11 every extreme point of $H \cap X$ is an extreme point of X, hence $\mathrm{Ext}(X) \neq \emptyset$. Moreover, x is a convex combination of at most n extreme points of X, so in particular of at most $n+1$ extreme points, as wanted.

Now let $x \in \mathrm{int}(X)$ and take $y \in \mathrm{Ext}(X)$ (note that this implies $y \neq x$). Set $L = \{x + t(y-x) \mid t \in \mathbb{R}\} \subseteq E$. Then $L \cap X$ is a compact convex subset of L and $x \in L \cap X$. Choose t_0 to be the smallest $t \in \mathbb{R}$ such that $x + t(y - x) \in L \cap X$ and set $z = x + t_0(y - x)$. Note that $t_0 < 0$ and $z \in \partial X$. Then x belongs to the line segment $[z, y]$, so $x = \lambda z + (1 - \lambda)y$ for some $0 < \lambda < 1$. Since $z \in \partial X$, we know that z is a convex combination of at most n extreme points of X. This implies that $x = \lambda z + (1 - \lambda)y$ is a convex combination of at most $n + 1$ extreme points of X, and we are done. $\square$

Remark 1.14. After having established Carathéodory's theorem, we now know that for any non-empty compact convex subset X of $\mathbb{R}^n$, the set of its extreme points, $\mathrm{Ext}(X)$, is always non-empty and we can also justify the assertion that the convex hull of $\mathrm{Ext}(X)$, $\mathrm{conv}(\mathrm{Ext}(X))$, is always closed (in fact, compact).

2 The Krein-Milman Theorem

By Carathéodory's theorem, we know that if $E = \mathbb{R}^n$, then we can write any point of a compact convex non-empty subset X of E as a convex combination of (finitely many, in this case) extreme points of X. The goal of this section is to prove that an (appropriately formulated) analogous statement holds when $\mathbb{R}^n$ is replaced by a more general locally convex Hausdorff topological vector space E. This statement is known as the *Krein-Milman theorem*.

The notion of convexity (and related concepts) can be introduced naturally in this setting (since E has the structure of a vector space), and, in particular, the definition of extreme points is the same as the one previously given.

We will start by discussing some concrete examples of convex sets and their extreme points.

2.1 Examples of extreme points

The first example shows that in this more general setting (unlike the case of $\mathbb{R}^n$), the set of extreme points may, sometimes, be empty.

Example 2.1. Let K_1 be the closed unit ball of $L_1([0,1], m)$. We will show that $\operatorname{Ext}(K_1) = \emptyset$.

First let $f \in K_1$ with $\|f\|_1 = 0$, then $f = 0$ almost everywhere. Note that $f = 0 = 1/2 \cdot 1 + 1/2 \cdot (-1)$ almost everywhere. Since 1 and -1 are different functions in K_1, f is a proper convex combination of elements of K_1, hence $f = 0 \notin \operatorname{Ext}(K_1)$.

Now, let $f \in K_1$ with $0 < \|f\|_1 \leq 1$. Consider the function $F\colon [0,1] \to \mathbb{R}$ defined by $F(x) = \int_0^x |f(t)| \,\mathrm{d}t$, $x \in [0,1]$. Then F is continuous and satisfies $F(0) = 0$ and $F(1) = 1$. Hence, by the Intermediate Value Theorem there exists $b \in [0,1]$ such that

$$\int_0^b |f(t)| \,\mathrm{d}t = \frac{1}{2} \int_0^1 |f(t)| \,\mathrm{d}t.$$

Now, for $t \in [0,1]$ define $f_1(t) = 1_{[0,b]}(t) 2f(t)$ and $f_2(t) = 1_{(b,1]} 2f(t)$. Then

$$\|f_1\|_1 = \int_0^1 |f_1(t)| \,\mathrm{d}t = \int_0^b 2|f(t)| \,\mathrm{d}t = \int_0^1 |f(t)| \,\mathrm{d}t \leq 1$$

and likewise

$$\begin{aligned} \|f_2\|_1 = \int_0^1 |f_2(t)| \,\mathrm{d}t &= \int_b^1 2|f(t)| \,\mathrm{d}t = \int_0^1 2|f(t)| \,\mathrm{d}t - \int_0^b 2|f(t)| \,\mathrm{d}t \\ &= 2\int_0^1 |f(t)| \,\mathrm{d}t - \int_0^1 |f(t)| \,\mathrm{d}t = \int_0^1 |f(t)| \,\mathrm{d}t \leq 1, \end{aligned}$$

Hence $f_1, f_2 \in K_1$. Now, note that $f(t) = 1_{[0,b]}(t)f(t) + 1_{(b,1]}(t)f(t) = 1/2\left(1_{[0,b]}(t)2f(t)\right) + 1/2\left(1_{(b,1]}(t)2f(t)\right) = 1/2f_1(t) + 1/2f_2(t)$, for all $t \in [0,1]$. We only need to check whether $f_1 = f$ or $f_2 = f$. We have

$$\begin{aligned}
\|f_1 - f\|_1 &= \int_0^1 |f_1(t) - f(t)|\,\mathrm{d}t \geq \int_0^b |f_1(t) - f(t)|\,\mathrm{d}t \\
&= \int_0^b |2f(t) - f(t)|\,\mathrm{d}t = \int_0^b |f(t)|\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^1 |f(t)|\,\mathrm{d}t > 0
\end{aligned}$$

and likewise

$$\begin{aligned}
\|f_2 - f\|_1 &= \int_0^1 |f_2(t) - f(t)|\,\mathrm{d}t \geq \int_b^1 |f_2(t) - f(t)|\,\mathrm{d}t \\
&= \int_b^1 |2f(t) - f(t)|\,\mathrm{d}t = \int_b^1 |f(t)|\,\mathrm{d}t \\
&= \int_0^1 |f(t)|\,\mathrm{d}t - \int_0^b |f(t)|\,\mathrm{d}t \\
&= \int_0^1 |f(t)|\,\mathrm{d}t - \frac{1}{2}\int_0^1 |f(t)|\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^1 |f(t)|\,\mathrm{d}t > 0.
\end{aligned}$$

We conclude that, $f \neq f_1$ and $f \neq f_2$. But then f is a proper convex combination of two elements of K_1, hence f is not an extreme point of K_1. Thus, $\mathrm{Ext}(K_1) = \emptyset$.

The next example concerns the extreme points of the set of *probability measures* on a (locally) compact Hausdorff topological space. These are *Radon measures* with total mass 1. We first recall the definition of a Radon measure, and discuss some results concerning the notion of *support*.

Definition 2.2. *Let X be a locally compact Hausdorff topological space. A **Radon measure** μ on X is a Borel measure, i.e., a non-negative measure on the Borel σ-algebra $B(X)$ on X, satisfying the following conditions:*

- $\mu(K) < \infty$ *for all compact sets* $K \subseteq X$.
- (Outer regularity)
  For each Borel set A in X, $\mu(A) = \inf\{\mu(U) \mid U \text{ open}, \ A \subseteq U\}$.
- (Inner regularity)
  For each open set U in X, $\mu(U) = \sup\{\mu(K) \mid K \text{ compact}, \ K \subseteq U\}$.

As mentioned above a ***probability measure*** is a Radon measure on X having total mass equal to 1.

Now, we continue to define the support of μ. Let $(U_i)_{i\in I}$ be the family of all open sets $U_i \subseteq X$ such that $\mu(U_i) = 0$ for all $i \in I$. Set $N = \bigcup_{i\in I} U_i$. Clearly N is also open. We will show that $\mu(N) = 0$. Since μ is a Radon measure, it is inner regular, so it suffices to show that $\mu(K) = 0$ for all compact sets $K \subseteq N$. Let $K \subseteq \bigcup_{i\in I} U_i$ be compact. There exists $i_1, \dots, i_n \in I$ such that $K \subseteq \bigcup_{k=1}^n U_{i_k}$. Hence, $\mu(K) \leq \mu\left(\bigcup_{k=1}^n U_{i_k}\right) \leq \sum_{k=1}^n \mu\left(U_{i_k}\right) = 0$ as wanted.

The complement of N is called the ***support of*** μ and is denoted by $\operatorname{supp}(\mu)$.

Remark 2.3. For $x \in X$, consider the Dirac mass δ_x at x. It is immediate to see that $\operatorname{supp}(\delta_x) = \{x\}$.

The following shows that the converse also holds.

Lemma 2.4. *Let μ be a Radon measure on a locally compact Hausdorff topological space X. If the support of μ is a singleton, i.e.,* $\operatorname{supp}(\mu) = \{x\}$ *for some $x \in X$, then $\mu = c\delta_x$, where $c = \mu(X)$.*

Proof. Let $x \in X$ and $\operatorname{supp}(\mu) = \{x\}$. By the definition of support of a measure, $\mu(X\backslash\{x\}) = 0$, hence $\mu(\{x\}) = \mu(X) - \mu(X\backslash\{x\}) = \mu(X) = c$. To show that $\mu = c\delta_x$, let $A \in B(X)$. First, if $x \notin A$ then $A \subseteq X\backslash\{x\}$ which means $\mu(A) = 0 = c\delta_x(A)$. Now, if $x \in A$ then $A = \{x\} \cup (A\backslash\{x\})$, and hence

$$\mu(A) = \mu(\{x\}) + \mu(A\backslash\{x\}) = \mu(\{x\}) = c = c\delta_x(A).$$

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

Note that if μ is a probability measure such that $\operatorname{supp}(\mu) = \{x\}$, then $\mu = \delta_x$, as announced in the above Remark.

Lemma 2.5. *Let μ be a Radon measure on a locally compact Hausdorff topological space X. If $f \in C(X)$, $0 \leq f \leq 1$ and $f(x) > 0$ for some $x \in \operatorname{supp}(\mu)$, then $\int_X f \,\mathrm{d}\mu > 0$.*

Proof. Let $\lambda := f(x)$. Set

$$U = \left\{ y \in X \;\middle|\; f(y) > \frac{\lambda}{2} \right\}.$$

Then $x \in U$, the set U is open and $f > \lambda/2 \cdot 1_U$. Hence

$$\int_X f \,\mathrm{d}\mu > \int_X \frac{\lambda}{2} 1_U \,\mathrm{d}\mu = \frac{\lambda}{2}\mu(U) > 0.$$

The latter holds because if $\mu(U) = 0$, then $x \in U \subseteq N$ where N is the largest open set having μ-measure zero. By the definition of $\operatorname{supp}(\mu)$, it follows that $x \notin \operatorname{supp}(\mu)$, which is a contradiction. □

We are now ready to study the extreme points of $P(X)$, the set of probability measures on X, when X is a compact Hausdorff topological space.

Proposition 2.6. *Let X be a compact Hausdorff topological space. Then $P(X)$ is a convex set and*

$$\operatorname{Ext}(P(X)) = \{\delta_x \mid x \in X\}.$$

Proof. It is easy to see that a convex combination of probability measures is a probability measure, hence $P(X)$ is convex. Given $x \in X$, assume to reach a contradiction that $\delta_x \notin \operatorname{Ext}(P(X))$. Then there exists $\mu_1 \neq \delta_x$ and $\mu_2 \neq \delta_x$ such that $\delta_x = \alpha\mu_1 + (1-\alpha)\mu_2$ for some $0 < \alpha < 1$ and $\mu_1, \mu_2 \in P(X)$. We have $\delta_x(\{x\}) = \alpha\mu_1(\{x\}) + (1-\alpha)\mu_2(\{x\})$, which implies $1 = \alpha\mu_1(\{x\}) + (1-\alpha)\mu_2(\{x\})$. For $0 \leq \mu_1(\{x\}), \mu_2(\{x\}) \leq 1$ and $0 < \alpha < 1$, we must have $1 = \mu_1(\{x\}) = \mu_2(\{x\})$. As μ_1 is a probability measure, $\mu_1(X\backslash\{x\}) = 0$, hence $\operatorname{supp}(\mu_1) = \{x\}$ and due to Lemma 2.4, $\mu_1 = \delta_x$, a contradiction. This proves that $\{\delta_x \mid x \in X\} \subseteq \operatorname{Ext}(P(X))$.

To show the converse inclusion, let $\mu \in \operatorname{Ext}(P(X))$. We want to show that there exists $x \in X$ such that $\mu = \delta_x$. By Lemma 2.4, it suffices to show that $\operatorname{supp}(\mu) = \{x\}$ for some $x \in X$. Suppose there exist $x, y \in X$ with $x \neq y$ such that $x, y \in \operatorname{supp}(\mu)$. Since X is a compact Hausdorff topological space, X is normal [4, 4.25]. Hence there exist U, V open subsets of X such that $x \in U$, $y \in V$ and $\overline{U} \cap \overline{V} = \emptyset$. By Urysohn's Lemma [4, 4.15], there exists $f \in C(X)$, $0 \leq f \leq 1$ such that $f|_U = 1$ and $f|_V = 0$. Since $f(x) = 1 > 0$, by Lemma 2.5 we deduce that $0 < \int_X f \, \mathrm{d}\mu =: \lambda$. Furthermore, we want to show that $\lambda < 1$. For this note that $f \leq 1_{X\backslash V}$ which implies $\lambda = \int_X f \, \mathrm{d}\mu \leq \mu(X\backslash V) = \mu(X) - \mu(V) = 1 - \mu(V) < 1$. The latter follows from the fact that if $\mu(V) = 0$ then $V \subseteq N$ (where N is the largest open set of μ-measure zero). But $y \in V$ and $y \in \operatorname{supp}(\mu) = N^c$, a contradiction. Now, set

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

Then $\mu_1, \mu_2 \in P(X)$ and we have $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$. Note that

$$\mu_1(U) = \frac{1}{\lambda}\int_U f \, \mathrm{d}\mu = \frac{\mu(U)}{\lambda} > 0 \quad \text{and} \quad \mu_2(U) = \frac{1}{1-\lambda}\int_U (1-f) \, \mathrm{d}\mu = 0$$

hence $\mu_1 \neq \mu_2$, which contradicts the fact that $\mu \in \operatorname{Ext}(P(X))$. □

2.2 The Krein-Milman Theorem

Now we are ready to prove the main result of this section. To do that we first establish some very helpful results.

Proposition 2.7. *If X is a non-empty compact subset of a locally convex Hausdorff topological vector space E and $f\colon X \to \mathbb{R}$ is upper semicontinuous (i.e, the set $\{x \in X \mid f(x) < a\}$ is open for all $a \in \mathbb{R}$), then f assumes its maximum on X.*

Proof. First we show that the maximum does exist, i.e., $\sup_{x\in X} f(x) < \infty$. Note that

$$X = \bigcup_{n=1}^{\infty} \{x \in X \mid f(x) < n\}.$$

Since X is compact there exists $n_0 \in \mathbb{N}$ such that $X = \{x \in X \mid f(x) < n_0\}$. This shows that $\sup_{x\in X} f(x) \leq n_0 < \infty$, as claimed.

Next, we show that there exists $x_0 \in X$ such that $\sup_{x\in X} f(x) = f(x_0)$. Suppose to reach a contradiction that this is not the case. Then for all $x \in X$, $f(x) < \sup_{x\in X} f(x) =: M$. Then $X = \bigcup_{n=1}^{\infty} \{x \in X \mid f(x) < M - 1/n\}$. Since X is compact, there exists $n_1 \in \mathbb{N}$ so $X = \{x \in X \mid f(x) < M - 1/n_1\}$. This implies $\sup_{x\in X} f(x) \leq M - 1/n_1 < M$, a contradiction. $\square$

Actually, f assumes its maximum at an extreme point. This will be shown soon. But before we continue, we need the concept of a *convex function*.

Definition 2.8. *Let f be a real valued function on a convex set X. Then f is a **convex function** if for each $x, y \in X$*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

*If the inequality is strict, the function is called **strictly convex**.*

If there is equality in the statement above f is said to be an ***affine function***.

Example 2.9. Notice that the function $f\colon x \mapsto x^2$ is strictly convex. This means that for each $x \neq y$ and $s \in (0,1)$, we have

$$f(sx + (1-s)y)^2 < sf(x) + (1-s)f(y)$$

or equivalently

$$(sx + (1-s)y)^2 < sx^2 + (1-s)y^2.$$

This example will be useful later.

We can now continue to show that f (defined as in Proposition 2.7) assumes its maximum at an extreme point.

Theorem 2.10. *If X is a non-empty compact subset of a locally convex Hausdorff topological vector space E and $f\colon X \to \mathbb{R}$ is convex and upper semicontinuous. Then there exists an extreme point of X at which f assumes its maximum value.*

Proof. Define a collection of subsets of X by

$$\mathcal{A} = \{F \subseteq X \mid F \neq \emptyset, F \text{ is closed and if } (a,b) \subseteq X \\ \text{and } (a,b) \cap F \neq \emptyset, \text{then } (a,b) \subseteq F\}.$$

Here (a,b) for $a,b \in F$ means the open segment $\{\lambda a + (1-\lambda)b \mid 0 < \lambda < 1\}$. Note that the set $(a,a) = \{a\}$.

We will now state and prove five properties for $\mathcal{A}$ and use them to finish the proof.

(i) $X \in \mathcal{A}$.

This follows from the fact that X is non-empty and compact.

(ii) If $\{X_i\}_{i\in I} \subseteq \mathcal{A}$ and $\bigcap_{i\in I} X_i \neq \emptyset$, then $\bigcap_{i\in I} X_i \in \mathcal{A}$.

Clearly $\bigcap_{i\in I} X_i$ is non-empty and closed, so it remains to show that if $(a,b) \subseteq X$ and $(a,b) \cap (\bigcap_{i\in I} X_i) \neq \emptyset$ then $(a,b) \subseteq \bigcap_{i\in I} X_i$. For this let $(a,b) \subseteq X$ and $(a,b) \cap (\bigcap_{i\in I} X_i) \neq \emptyset$. Then $\bigcap_{i\in I} ((a,b) \cap X_i) \neq \emptyset$ which means $(a,b) \cap X_i \neq \emptyset$ for all $i \in I$. Since $X_i \in \mathcal{A}$ for all $i \in I$ and $(a,b) \subseteq X$, we have $(a,b) \subseteq X_i$ for all $i \in I$. Thus $(a,b) \subseteq \bigcap_{i\in I} X_i$ and hence $\bigcap_{i\in I} X_i \in \mathcal{A}$, as wanted.

(iii) For each $x \in X$, we have $\{x\} \in \mathcal{A}$ if and only if x is an extreme point of X.

To show this, let $\{x\} \in \mathcal{A}$. Suppose $x = \alpha a + (1-\alpha)b$ for $0 < \alpha < 1$, then $\{x\} \in (a,b)$. By our hypothesis, $(a,b) \subseteq \{x\}$, which means $(a,b) = \{x\}$ which again gives us that $x = a = b$, meaning $x \in \operatorname{Ext}(X)$. Conversely, assume $x \in \operatorname{Ext}(X)$. Then $(a,b) = \{x\}$ for all a,b with $\{x\} \subseteq (a,b)$, hence $(a,b) \cap \{x\} \neq \emptyset$ and $(a,b) \subseteq \{x\}$. Clearly, $\{x\}$ is closed and non-empty, hence $\{x\} \in \mathcal{A}$, as wanted.

(iv) $F' \in \mathcal{A}$ for each $F \in \mathcal{A}$ and every convex upper semicontinuous function g, where

$$F' = \left\{x \in F \;\middle|\; g(x) = \sup\{g(y) \mid y \in F\}\right\}.$$

Indeed, since we know from Proposition 2.7 that there exists some x_0 at which g assumes its maximum value, we deduce that $F' \neq \emptyset$. To show closedness of F', note that

$$F' = \bigcap_{n=1}^{\infty} \left\{x \in F \;\middle|\; g(x) \geq \sup_{y\in F} g(y) - \frac{1}{n}\right\}.$$

Since g is upper semicontinuous, all sets on the right hand-side are closed, hence F' is closed.

To check the last condition, assume that $(a, b) \subseteq X$ and $(a, b) \cap F' \neq \emptyset$. This means there exists $x_0 \in (a, b) \cap F' \subseteq (a, b) \cap F$. Since $F \in \mathcal{A}$, $(a, b) \subseteq F$. Let $x, y \in (a, b)$ be such that $x_0 = \lambda x + (1 - \lambda)y$ for some $0 < \lambda < 1$. Then

$$\begin{aligned} g(x_0) \leq \lambda g(x) + (1-\lambda)g(y) &\leq \lambda \sup_{z \in F} g(z) + (1-\lambda) \sup_{z \in F} g(z) \\ &= \lambda g(x_0) + (1-\lambda) g(x_0) = g(x_0), \end{aligned}$$

using that $x_0 \in (a, b) \cap F'$. Since $g(x_0) = \sup_{z \in F} g(z)$, we must have $g(x) = g(x_0) = g(y)$, so $x, y \in F'$ and hence $(a, b) \in F'$, as wanted.

We define a partial ordering on $\mathcal{A}$ by setting $F_1 \leq F_2$ if and only if $F_1 \supseteq F_2$ and we then continue with the last statement about $\mathcal{A}$.

(v) If $F \in \mathcal{A}$ is a maximal element, then $F = \{a\}$.

Suppose $F \in \mathcal{A}$ is a maxial element. Let $a \in F$ and assume that also $b \in F$ with $a \neq b$. The Hahn-Banach Separation Theorem gives us that there exists a functional $l \colon E \to \mathbb{R}$ such that $l(a) \neq l(b)$. Use (iv) with $g = l$ to conclude that

$$F' = \{x \in F \mid l(x) = \sup_{z \in F} l(z)\}.$$

Since $l(a) \neq l(b)$ then $a \notin F'$ or $b \notin F'$. This means $F' \subset F$ hence $F' > F$, so F is not maximal, a contradiction.

Now we are ready to finish the proof. Let $(F_\alpha)_{\alpha \in A}$ be a linearly ordered family in $\mathcal{A}$. Since the family has the finite intersection property, it has a non-empty intersection due to compactness of X and hence $\bigcap_{\alpha \in A} F_\alpha \neq \emptyset$. By (ii), $\bigcap_{\alpha \in A} F_\alpha \in \mathcal{A}$ and thus it is clear that $\bigcap_{\alpha \in A} F_\alpha$ is an upper bound for the family. This shows that Zorn's Lemma [4, 0.2] applies, thus every $F \in \mathcal{A}$ is majorised by a maximal element. Take, say, $F = \{y \in X \mid f(y) = \sup_{x \in X} f(x)\}$. From (i) $X \in \mathcal{A}$, hence (iv) gives us that $F \in \mathcal{A}$. If $\{a\} \subseteq F$ is maximal, then $a \in \operatorname{Ext}(X)$ according to (iii), as wanted. □

Using Theorem 2.10, we can easily prove the Krein-Milman Theorem.

Theorem 2.11 (Krein-Milman)**.** *Let X be a non-empty compact convex subset of a locally convex Hausdorff topological vector space E. Then*

$$X = \overline{\operatorname{conv}(\operatorname{Ext}(X))}.$$

Proof. Since $\operatorname{Ext}(X) \subseteq X$, we have $\overline{\operatorname{conv}(\operatorname{Ext}(X))} \subseteq \overline{\operatorname{conv}(X)} = X$, as X is convex and closed.

For the converse inclusion suppose there exists $x \in X \backslash \overline{\operatorname{conv}(\operatorname{Ext}(X))}$. By the Hahn-Banach Separation Theorem [1, Theorem 2.3, Chapter 1], there exists a functional f on X and $\alpha \in \mathbb{R}$ such that $f\left(\overline{\mathrm{conv}(\mathrm{Ext}(X))}\right) \leq \alpha$ and $f(x) > \alpha$. Due to Theorem 2.10, there exists an extreme point of X where f assumes its maximum value, hence there is an extreme point of X which is not in $\overline{\mathrm{conv}(\mathrm{Ext}(X))}$, a contradiction. $\square$

Remark 2.12. With the help of Krein-Milman's theorem, we can now say a bit more about the example discussed in Proposition 2.6. Namely, in view of the Riesz representation theorem [4, Theorem 7.2] , the set $P(X)$ of probability measures on a compact Hausdorff topological space X can be identified with a weak*-compact (-convex) subset of $C(X)^*$. Then, by Krein-Milman's theorem, combined with Proposition 2.6, we deduce that

$$P(X) = \overline{\mathrm{conv}\{\delta_x \mid x \in X\}}.$$

3 The Krein-Milman Theorem as an integral representation theorem

We would like to interpret the Krein-Milman theorem as an *integral representation theorem.* This terminology has to be developed first and in the end of this section, we will reformulate the Krein-Milman theorem in our new language.

Definition 3.1. *Suppose that X is a non-empty compact subset of a locally convex Hausdorff topological vector space E, and that μ is a probability measure on X. A point x in E is said to be* ***represented by*** *μ if $f(x) = \int_X f\,\mathrm{d}\mu$ for every continuous linear functional f on E. Often we write $\mu(f) = \int_X f\,\mathrm{d}\mu$. Other terminology is that x is the* ***resultant of*** *μ.*

Remark 3.2. Given $x \in X$ there is always a probability measure representing x, namely δ_x. Uniqueness of representing measures will be discussed later in this section.

Definition 3.3. *If μ is a Radon measure on the compact Hausdorff space X and S is a Borel subset of X, we say that μ is* ***supported by*** *S if $\mu(X\backslash S) = 0$.*

Note that the support of μ, $\operatorname{supp}(\mu)$, defined earlier is a Borel subset of X such that $\mu(X\backslash\operatorname{supp}(\mu)) = 0$, hence μ is supported by $\operatorname{supp}(\mu)$.

We now develop further theory of representing measures.

Theorem 3.4. *Suppose that Y is a compact subset of a locally convex space E, and that the closed convex hull X of Y is compact. If μ is a probability measure on Y, then there exists a unique point $x \in X$ which is represented by μ, and the function $\mu \mapsto$ (resultant of μ) is an affine weak*-continuous map from $P(Y)$ (the probability measures on Y) into X.*

Proof. For the first statement of the theorem, we want to begin with showing that the compact convex set X contains a point x such that $f(x) = \int_Y f\,\mathrm{d}\mu$ for all $f \in E^*$. To do this, given $f \in E^*$, let

$$H_f = \left\{ y \in E \;\middle|\; f(y) = \int_Y f\,\mathrm{d}\mu =: \mu(f) \right\}.$$

We observe that H_f is closed. Now we want to show that $\bigcap_{f\in E^*} H_f \cap X \neq \emptyset$. Since H_f and X are both closed, $H_f \cap X$ is closed for all $f \in E^*$. The set X is compact and therefore it has the finite intersection property. Hence it suffices to show that $\bigcap_{i=1}^n H_{f_i} \cap X \neq \emptyset$ for finitely many $f_1, \dots, f_n \in E^*$, arbitrarily chosen.

To do this, we define $T\colon E \to \mathbb{R}^n$ by $Ty = (f_1(y), f_2(y), \dots, f_n(y))$. The map T is clearly linear and continuous. Since X is compact, TX is compact, as well. To check the convexity of TX, let $t \in (0,1)$ and $a, b \in TX$. Then we can find $x, y \in X$ such that $a = Tx$ and $b = Ty$. Consider

$$ta + (1-t)b = tTx + (1-t)Ty = T\left(tx + (1-t)y\right).$$

The last assertion comes from the fact that T is linear. We know that X is convex, hence $tx + (1-t)y \in X$. Therefore TX is convex, as well.

Now, let $p = (\mu(f_1), \mu(f_2), \ldots, \mu(f_n))$. If $p \in TX$ then there exists $y \in X$ such that

$$(f_1(y), f_2(y), \ldots, f_n(y)) = Ty = p = \left(\int_Y f_1 \,\mathrm{d}\mu, \int_Y f_2 \,\mathrm{d}\mu, \ldots, \int_Y f_n \,\mathrm{d}\mu\right)$$

which is what we wanted.

Assume towards contradiction that $p \notin TX$. Recall that TX is compact and convex. Due to the Hahn-Banach Separation Theorem there exists a linear functional $\Phi\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\Phi(p) > \sup\left\{\Phi(Ty) \mid y \in X\right\} = \sup\left\{\Phi(z) \mid z \in TX\right\}.$$

The space $\mathbb{R}^n$ is a Hilbert space and due to the Riesz Representation Theorem [4, Theorem 5.25], we can find a vector $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ such that

$$\Phi(z) = \langle z, a\rangle,\ z \in \mathbb{R}^n.$$

Since $a, z \in \mathbb{R}^n$, we have $\langle z, a\rangle = \langle a, z\rangle$. Therefore

$$\langle p, a\rangle = \langle a, p\rangle > \sup\left\{\langle a, Tx\rangle \mid x \in X\right\}.$$

We define a linear functional g on E by $g(x) = \langle a, Tx\rangle = \sum_{i=1}^n a_i f_i(x)$ for $x \in E$. Since T is continuous, so is g. We have that

$$\Phi(p) = \langle a, p\rangle = \sum_{i=1}^n a_i \int_Y f_i \,\mathrm{d}\mu = \int_Y \sum_{i=1}^n a_i f_i \,\mathrm{d}\mu = \int_Y g \,\mathrm{d}\mu.$$

This means

$$\int_Y g \,\mathrm{d}\mu = \langle a, p\rangle > \sup\left\{\langle a, Tx\rangle \mid x \in X\right\} = \sup\left\{g(x) \mid x \in X\right\}.$$

But we also have

$$\begin{aligned}\int_Y g \,\mathrm{d}\mu \le \int_Y \sup\left\{g(x) \mid x \in Y\right\} \,\mathrm{d}\mu &= \sup\left\{g(x) \mid x \in Y\right\} \mu(Y)\\ &\le \sup\left\{g(x) \mid x \in X\right\}\end{aligned}$$

This gives us a contradiction and hence, we have $p \in TX$, as wanted.

To show the uniqueness of the represented point, let $x, y \in E$ be represented by μ. This means that for all $f \in E^*$, we have $f(x) = \int_Y f \,\mathrm{d}\mu$ and

$f(y) = \int_Y f \,\mathrm{d}\mu$. In particular $f(x) = f(y)$ for all $f \in E^*$. Due to linearity of f, we have $f(x-y) = 0$. Since f is arbitrary, the Hahn-Banach Separation Theorem then gives us that $x - y = 0$. Therefore $x = y$ and hence the represented point is unique.

To show the second part of the theorem, we want to prove that the map $\Psi \colon P(Y) \to X$, which sends a probability measure μ into its resultant in X, is weak*-continuous and affine. To show it is affine, let μ_1, μ_2 be probability measures and assume x_1 is represented by μ_1 and x_2 is represented by μ_2. This means $f(x_1) = \int_Y f \,\mathrm{d}\mu_1$ and $f(x_2) = \int_Y f \,\mathrm{d}\mu_2$ for all $f \in E^*$. Hence

$$\begin{aligned}\int_Y f \,\mathrm{d}(\lambda\mu_1 + (1-\lambda)\mu_2) &= \lambda \int_Y f \,\mathrm{d}\mu_1 + (1-\lambda) \int_Y f \,\mathrm{d}\mu_2 \\ &= \lambda f(x_1) + (1-\lambda) f(x_2) \\ &= f\left(\lambda x_1 + (1-\lambda) x_2\right)\end{aligned}$$

which means that $\lambda\mu_1 + (1-\lambda)\mu_2$ represents $\lambda x_1 + (1-\lambda)x_2$. Hence we have $\Psi\left(\lambda\mu_1 + (1-\lambda)\mu_2\right) = \lambda x_1 + (1-\lambda)x_2 = \lambda\Psi(\mu_1) + (1-\lambda)\Psi(\mu_2)$ which shows that Ψ is affine.

To show the weak*-continuity of Ψ, assume there is a net $(\mu_\alpha)_\alpha$ of probability measures converging in weak*-topology to μ. By definition $\left(\Psi(\mu_\alpha)\right)_\alpha$ is a net in X. Since X is compact there exists a subnet $\left(\Psi(\mu_\beta)\right)_\beta$ such that $\Psi(\mu_\beta) \to y$ for some $y \in X$ [4, Theorem 4.29]. The net $(\mu_\beta)_\beta$ is a subnet of $(\mu_\alpha)_\alpha$ which clearly converges to the same limit μ. Hence for all $f \in E^*$

$$f(x_\beta) = \int_Y f \,\mathrm{d}\mu_\beta \to \int_Y f \,\mathrm{d}\mu = f(x)$$

where x_β and x are (according to the first part of the theorem) the unique points in X that are represented by μ_β and μ, respectively. This means $f(x-y) = f(x) - f(y) = f(x) - f\left(\lim_\beta x_\beta\right) = f(x) - \lim_\beta f(x_\beta) = f(x) - f(x) = 0$. But then $f(x-y) = 0$ for all $f \in E^*$, hence $x - y = 0$ which implies $x = y$. Since $\left(\Psi(\mu_\alpha)\right)_\alpha$ is a net in the compact set X, we know $\Psi(\mu_\alpha) \to y = x = \Psi(\mu)$, hence Ψ is weak*-continuous. $\square$

The next result gives a characterization of points in the closed convex hull of compact sets in terms of representing measures.

Theorem 3.5. *Suppose Y is a compact subset of a locally convex space E. A point $x \in E$ is in the closed convex hull X of Y if and only if there exists a probability measure μ on Y which represents x.*

Proof. Suppose there exists a probability measure μ on Y which represents x, then for each $f \in E^*$, we have

$$\begin{aligned}f(x) = \mu(f) = \int_Y f \,\mathrm{d}\mu &\leq \int_Y \sup\{f(y) \mid y \in Y\} \,\mathrm{d}\mu \\ &= \sup\{f(y) \mid y \in Y\}\, \mu(Y) = \sup f(Y) \\ &\leq \sup\{f(y) \mid y \in X\} = \sup f(X)\end{aligned}$$

Assume $x \notin X$. Since X is closed and convex, the Hahn-Banach Separation Theorem gives us that there exists $g \in E^*$, such that

$$g(x) > \sup\{g(y) \mid y \in X\}.$$

This is a contradiction to the calculations above. Hence $x \in X$.

To show the converse implication, assume $x \in \overline{\mathrm{conv}(Y)} = X$. Then there exists a net in the convex hull of Y which converges to x [4, Proposition 4.18]. Equivalently, there exists a net $(y_\alpha)_\alpha$ of the form $y_\alpha = \sum_{i=1}^{n_\alpha} \lambda_i^\alpha x_i^\alpha$ where $\lambda_i^\alpha > 0$, $\sum_{i=1}^{n_\alpha} \lambda_i^\alpha = 1$ and $x_i^\alpha \in Y$, which converges to x. Each y_α is represented by the probability measure $\mu_\alpha = \sum_{i=1}^{n_\alpha} \lambda_i^\alpha \delta_{x_i}^\alpha$ according to Remark 3.2. By the Riesz Representation theorem, we know that the set of all probability measures on Y can be identified with a weak*-compact convex subset of $C(Y)^*$. Hence there exists a subnet $(\mu_\beta)_\beta$ of $(\mu_\alpha)_\alpha$, which converges to some probability measure μ on Y. Furthermore each $f \in E^*$ is in $C(Y)$, when restricted to Y, hence

$$\lim_\beta f(y_\beta) = \lim_\beta \int_Y f \,\mathrm{d}\mu_\beta = \int_Y f \,\mathrm{d}\mu.$$

Since $(y_\alpha)_\alpha$ converges to x, so does the subnet $(y_\beta)_\beta$ and therefore $f(x) = \int_Y f \,\mathrm{d}\mu$ for all $f \in E^*$. $\square$

The following result shows that the extreme points of X can be characterized by the fact that they have unique representing measures, namely, the corresponding point-masses.

Proposition 3.6. *Suppose X is a compact convex subset of a locally convex Hausdorff topological vector space E and that $x \in X$. Then x is an extreme point of X if and only if the point mass δ_x is the only probability measure on X which represents x.*

Proof. Suppose x is an extreme point of X and that the measure μ represents x. Then we want to show that μ is supported by the set $\{x\}$, which in turns of Lemma 2.4 will show that $\mu = \delta_x$. To do this it suffices to show that $\mu(D) = 0$ for each compact set D with $D \subseteq X \backslash \{x\}$, due to inner regularity of μ. Suppose instead that $\mu(D) > 0$ for some D. We want to show that there exists a point $y \in D$ such that $\mu(U \cap X) > 0$ for every open neighborhood U of y. Assume towards contradiction that for every y there exists a U_y such that $\mu(U_y \cap X) = 0$. The set $\{U_y\}_{y \in D}$ is an open cover of D. Since D is compact $U_{y_1}, \dots, U_{y_n}$ also covers D for $y_1, \dots, y_n \in D$. This means $D \subseteq U_{y_1} \cup \dots \cup U_{y_n}$. We also have

$$D = D \cap X \subseteq \left(\bigcup_{k=1}^{n} U_{y_k} \right) \cap X = \bigcup_{k=1}^{n} (U_{y_k} \cap X).$$

Due to our first assumption

$$0 < \mu(D) \leq \mu\left(\bigcup_{k=1}^{n} (U_{y_k} \cap X)\right) \leq \sum_{k=1}^{n} \mu\left(U_{y_k} \cap X\right) = n \cdot 0 = 0,$$

a contradiction. Hence, there must exist y such that $\mu(U \cap X) > 0$ for all U. Choose U to be the closed convex neighbourhood of y such that $K = U \cap X \subseteq X \backslash \{x\}$. Since K is a closed subset of the compact set X, K is itself compact, as well as convex, and $0 < r := \mu(K) < 1$. The latter holds because if $\mu(K) = 1$, then the resultant x of μ would be in K, a contradiction. We define Borel measures μ_1 and μ_2 on X by

$$\mu_1(B) = r^{-1}\mu(B \cap K) \quad \text{and} \quad \mu_2(B) = (1-r)^{-1}\mu(B \cap (X \backslash K))$$

for each Borel set B in X. Let x_i be the resultant of μ_i for $i = 1, 2$. Since $\mu_1(K) = 1$, we see that $x_1 \in K$ and hence $x \neq x_1$. We observe that

$$\begin{aligned}
(r\mu_1 + (1-r)\mu_2)(B) &= r\mu_1(B) + (1-r)\mu_2(B) \\
&= r \cdot r^{-1}\mu(B \cap K) + (1-r)(1-r)^{-1}\mu(B \cap (X \backslash K)) \\
&= \mu(B \cap K) + \mu\left(B \cap (X \backslash K)\right) \\
&= \mu\left((B \cap K) \cup (B \cap (X \backslash K))\right) = \mu(B \cap X) = \mu(B).
\end{aligned}$$

Let $f \in E^*$, then using the above calculations, we get

$$\begin{aligned}
f(x) &= \int_X f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\left(r\mu_1 + (1-r)\mu_2\right) \\
&= r \int_X f \, \mathrm{d}\mu_1 + (1-r) \int_X f \, \mathrm{d}\mu_2 = rf(x_1) + (1-r)f(x_2) \\
&= f\left(rx_1 + (1-r)x_2\right).
\end{aligned}$$

Since this holds for all $f \in E^*$, we have $x = rx_1 + (1-r)x_2$. But we also have $x \neq x_1$ and $r \in (0,1)$, hence $x \notin \operatorname{Ext}(X)$, a contradiction. Thus, $\mu(D) \not> 0$ and therefore $\mu(D) = 0$, as wanted. Since μ is regular, $\mu\left(X \backslash \{x\}\right) = \sup\left\{\mu(D) \mid D \subseteq X \backslash \{x\},\ D \text{ compact}\right\} = 0$. Hence, $\mu(\{x\}) = 1$ which means $\mu = \delta_x$.

For the converse implication, if $x \notin \operatorname{Ext}(X)$, then there exist $x_0, x_1 \in X \backslash \{x\}$ and $0 < \alpha < 1$ such that $x = \alpha x_0 + (1-\alpha)x_1$. It is then easily seen that $\mu = \alpha\delta_{x_0} + (1-\alpha)\delta_{x_1}$ represents x, and clearly $\mu \neq \delta_x$ (for instance, $\mu\{x_0\} = \alpha \neq 0 = \delta_x(\{x_0\})$). $\square$

Proposition 3.7. *Suppose that X is a compact convex subset of a locally convex Hausdorff topological vector space, that $Z \subseteq X$ and that X is the closed convex hull of Z. Then the extreme points of X are contained in the closure of Z.*

Proof. Let $Y = \overline{Z}$ and suppose $x \in \mathrm{Ext}(X)$. Since Y is closed and a subset of the compact set X, Y is itself compact. By Proposition 3.5 there exists a probability measure μ on Y which represents x, i.e.,

$$f(x) = \int_Y f \, \mathrm{d}\mu.$$

By Proposition 3.6, $\mu = \delta_x$. Hence $x \in Y$. □

Theorem 3.8. *Every point of a compact convex subset X of a locally convex Hausdorff topological vector space E is represented by a probability measure μ on X which is supported by the closure of the extreme points of X.*

Proof. We will use the Krein-Milman Theorem (Theorem 2.11) to show this theorem. Let $x \in X$ and define $Y = \overline{\mathrm{Ext}(X)}$. By the Krein-Milman Theorem, we have that $x \in X = \overline{\mathrm{conv}(\mathrm{Ext}(X))} \subseteq \overline{\mathrm{conv}(Y)}$. Theorem 3.5 gives us, that there exists a probability measure μ on Y which represents x. We extend μ to X and observe that $\mu(X \backslash Y) = 0$, which means μ is supported by the closure of the extreme points, as wanted. □

We are now ready to reformulate the Krein-Milman theorem as an integral representation result.

Corollary 3.9. *The Krein-Milman Theorem (Theorem 2.11) and Theorem 3.8 are equivalent statements.*

Proof. We have already used the Krein-Milman theorem to prove Theorem 3.8. Conversely, assume the conclusion of Theorem 3.8 holds and let $x \in X$ and define $Y = \overline{\mathrm{Ext}(X)}$. Since X is compact and $\overline{\mathrm{Ext}(X)} \subseteq X$, $\overline{\mathrm{Ext}(X)}$ is compact as well. By assumption every point $x \in X$, in particular $x \in Y$, is represented by a probability measure μ. Then Theorem 3.5 gives us that $x \in \overline{\mathrm{conv}(Y)} = \overline{\mathrm{conv}(\overline{\mathrm{Ext}(X)})} = \overline{\mathrm{conv}(\mathrm{Ext}(X))}$. As this holds for all $x \in X$, we have $X \subseteq \overline{\mathrm{conv}(\mathrm{Ext}(X))}$. But since X is convex and compact (hence closed), we have $X = \overline{\mathrm{conv}(X)}$. We also have $\mathrm{Ext}(X) \subseteq X$, hence $\overline{\mathrm{conv}(\mathrm{Ext}(X))} \subseteq \overline{\mathrm{conv}(X)} = X$, thus $X = \overline{\mathrm{conv}(\mathrm{Ext}(X))}$, as wanted. □

4 Choquet's Theorem

Motivated by the reformulation of Krein-Milman's theorem as an integral representation result, as explained in the previous section, we now consider the following more refined problem:

If X is a compact convex subset of a locally convex Hausdorff topological vector space E and $x \in X$, does there exist a probability measure μ on X which is supported by the extreme points of X and which represents x?

Note that a priori, there is no guarentee that $\mathrm{Ext}(X)$ is a Borel subset of X (as it is required by the definition of "supported by"). In fact, such counterexamples do exist (see [5] for a reference).

In the case where X is metrizable, $\mathrm{Ext}(X)$ will always be a Borel set, as shown by the following:

Theorem 4.1. *If X is a metrizable, compact convex subset of a locally convex Hausdorff topological vector space E, then the extreme points of X form a G_δ set.*

Proof. Suppose the topology on X is given by the metric d and for each $n \in \mathbb{N}$ let $F_n = \left\{x \mid x = 2^{-1}(y+z),\ y,z \in X,\ d(y,z) \geq n^{-1}\right\}$. F_n is closed. We want to show

$$\bigcup_{n=1}^{\infty} F_n = \{x \in X \mid x \notin \mathrm{Ext}(X)\}.$$

Let $x \in F_n$, then $x = 1/2y + 1/2z$ with $d(y,z) \geq 1/n$. This means $y \neq z$ and in particular x is not an extreme point of X.

To show the converse inclusion, let $x \notin \mathrm{Ext}(X)$, then there exists $y, z \in X$ with $y \neq x$, $z \neq x$ and $\lambda \in (0,1)$ such that $x = \lambda y + (1-\lambda)z$. Now $\lambda \leq 1/2$ or $(1-\lambda) \leq 1/2$. Without loss of generality, we can assume $\lambda \leq 1/2$ (because otherwise we just redefine $\lambda' := (1-\lambda)$). If $\lambda = 1/2$ we can choose n such that $1/n < d(y,z)$ and then $x \in F_n$. If $\lambda < 1/2$ then $2\lambda \in (0,1)$ and hence $2\lambda y + (1-2\lambda)z \in X$. Define $y' = 2\lambda y + (1-2\lambda)z$. We observe $1/2z + 1/2y' = 1/2z + \lambda y + 1/2(1-2\lambda)z = 1/2z + \lambda y + 1/2z - \lambda z = \lambda y + (1-\lambda)z = x$, hence all we need to show is that $z \neq y'$. Assume to reach a contradiction that $z = y'$, then $z = y' = 2\lambda y + (1-2\lambda)z = 2\lambda y + z - 2\lambda z$, which means $2\lambda y = 2\lambda z$, hence $y = z$, a contradiction. Therefore we can choose n such that $1/n < d(z, y')$, which means $x \in F_n$.

Now we have shown, that the complement of the extreme points is a countable union of closed sets, i.e., it is an F_σ-set. Hence $\mathrm{Ext}(X)$ is a G_δ-set (countable intersection of open sets). ☐

Furthermore under the metrizability assumption, Choquet has proved that the answer to the above question is affirmative. We will discuss Choquet's theorem in the following. The proof makes use of the notion of upper (affine) envelope of a bounded function, a concept which we now define.

Denote by A the set of all continuous affine functions on X. Clearly $A \subseteq C(X)$.

Definition 4.2. *If f is a bounded function on X and $x \in X$, let $\overline{f}(x) = \inf\{h(x) \mid h \in A,\ h \geq f\}$.*

The function $\overline{f}$ is called the *upper envelope* of f, and it has the following properties:

Proposition 4.3. *Let $\overline{f}$ be the upper envelope of f, then*

(a) $\overline{f}$ is concave, bounded, and upper semicontinuous - hence Borel measurable.

(b) $f \leq \overline{f}$ and if f is concave and upper semicontinuous, then $f = \overline{f}$.

(c) If f, g are bounded, then $\overline{f+g} \leq \overline{f} + \overline{g}$ and $|\overline{f} - \overline{g}| \leq \|f - g\|$, while $\overline{f+g} = \overline{f} + g$ if $g \in A$. If $r > 0$, then $\overline{rf} = r\overline{f}$.

Proof. The proof will be omitted. □

Theorem 4.4 (Choquet)**.** *Suppose that X is a metrizable compact convex subset of a locally convex Hausdorff topological vector space E, and that x_0 is an element of X. Then there is a probability measure μ on X which represents x_0 and is supported by the extreme points of X.*

Proof. The first thing we want to show is the existence of a strictly convex function $f \in C(X)$. For this, observe that since X is metrizable, $C(X)$ is separable and hence so is it's subset A defined above. (We include the proofs of these results in the Appendix). We choose a countable dense subset $S' = \{h'_n \mid n \in \mathbb{N}\}$ consisting of functions of A. For every $n \in \mathbb{N}$, we define $h_n = h'_n/\|h'_n\|_\infty$. We define $S = \{h_n \mid n \in \mathbb{N}\}$. We want to show that S is dense in the unit sphere. Let $h_0 \in K := \{h \in A \mid \|h\|_\infty = 1\}$, then there exists a sequence $(g'_n)_{n\in\mathbb{N}}$ in S' converging to h_0. Thus we have $g'_n/\|g'_n\| \to h_0/\|h_0\| = h_0$ as $n \to \infty$. Since $g'_n/\|g'_n\|_\infty \in S$, S is also dense in K.

Now we want to show that S separates points in X. To do that, we first show it for S'. Let $x, y \in X$ with $x \neq y$ and hence $x - y \neq 0$. Due to the Hahn-Banach Separation Theorem there exists a linear functional ϕ such that $\phi(x-y) \neq 0$ and hence $\phi(x) \neq \phi(y)$. Let $|\phi(x) - \phi(y)| = 2r$. Since ϕ is linear it is in particular affine and therefore in A. Consider the open ball $B(\phi, r) \subseteq A$. Since S' is dense in A and the ball is open and non-empty, it contains a point from S'. Let $h' \in S'$, $x, y \in X$ with $x \neq y$ and $h' \in B(\phi, r)$.

Then

$$\begin{aligned}
\left|h'(x)-h'(y)\right| &= \left|\left(h'(x)-\phi(x)\right)+\left(\phi(x)-\phi(y)\right)+\left(\phi(y)-h'(y)\right)\right| \\
&\geq |\phi(x)-\phi(y)| - \left|\left(h'(x)-\phi(x)\right)+\left(\phi(y)-h'(y)\right)\right| \\
&\geq 2r - \left(\left|h'(x)-\phi(x)\right| + \left|\phi(y)-h'(y)\right|\right) \\
&\geq 2r - \left(\|h'-\phi\|_\infty + \|\phi-h'\|_\infty\right) \\
&= 2r - 2\|h'-\phi\| > 2r-2r = 0.
\end{aligned}$$

Hence $h'(x) \neq h'(y)$, and therefore S' separates points of X. To show that S does this as well, note that $h'(x)/\|h'\| \neq h'(y)/\|h'\|$. Since $h'/\|h'\| \in S$, the set S does separate points of X.

Let $f(x) = \sum_{n=1}^{\infty} 2^{-n} h_n^2(x)$. Since $h_n^2(x) \leq \|h_n^2\| = \|h_n\|^2 = 1$ the series converges pointwise. Consider

$$\begin{aligned}
\left|f(x) - \sum_{n=1}^{N} 2^{-n} h_n^2(x)\right| &= \left|\sum_{n=N+1}^{\infty} 2^{-n} h_n^2(x)\right| \leq \sum_{n=N+1}^{\infty} 2^{-n} \left|h_n^2(x)\right| \\
&\leq \sum_{n=N+1}^{\infty} 2^{-n} \left\|h_n^2\right\| \leq \sum_{n=N+1}^{\infty} 2^{-n}.
\end{aligned}$$

Hence we have

$$\sup_{x\in X} \left|f(x) - \sum_{n=1}^{\infty} 2^{-n} h_n^2(x)\right| \leq \sum_{n=N+1}^{\infty} 2^{-n} \to 0 \quad \text{as} \quad N \to \infty.$$

Hence the series converges uniformly to f, thus $f \in C(X)$.

To see that f is a strictly convex function, we need the result of Example 2.9.

In our case, let $x, y \in X$ with $x \neq y$ and $s \in (0,1)$. Since h_n is affine, we have

$$\begin{aligned}
h_n^2\left(sx+(1-s)y\right) &= \left(h_n\left(sx+(1-s)y\right)\right)^2 \\
&= \left(sh_n(x)+(1-s)h_n(y)\right)^2.
\end{aligned}$$

Thus, we have

$$\left(sh_n(x)+(1-s)h_n(y)\right)^2 < sh_n^2(x)+(1-s)h_n^2(y)$$

when $h_n(x) \neq h_n(y)$. Since S separates points of X we can actually find some $k \in \mathbb{N}$ for which $h_k(x) \neq h_k(y)$ and hence

$$\begin{aligned} f\left(sx + (1-s)y\right) &= \sum_{n=1}^{\infty} 2^{-n} h_n^2\left(sx + (1-s)y\right) \\ &< \sum_{n=1}^{\infty} 2^{-n}\left(sh_n^2(x) + (1-s)h_n^2(y)\right) \\ &= s\left(\sum_{n=1}^{\infty} 2^{-n} h_n^2(x)\right) + (1-s)\sum_{n=1}^{\infty} 2^{-n} h_n^2(y) \\ &= sf(x) + (1-s)f(y), \end{aligned}$$

giving us that f is strictly convex.

Now, let B denote the subspace $A + \mathbb{R}f$ of $C(X)$. Furthermore, let $p\colon C(X) \to \mathbb{R}$ be defined by $p(g) = \overline{g}(x_0)$ for $g \in C(X)$ and $x_0 \in X$. We want to show that p is a sublinear functional. To do this, let $g_1, g_2 \in C(X)$ and note that g_1, g_2 are bounded, since X is compact. Then by Proposition 4.3 (c), we have

$$p(g_1 + g_2) = \overline{g_1 + g_2}(x_0) \leq \overline{g_1}(x_0) + \overline{g_2}(x_0) = p(g_1) + p(g_2).$$

Furthermore, let $r \geq 0$ and note that

$$p(rg_1) = \overline{rg_1}(x_0) = r\overline{g_1}(x_0) = rp(g_1).$$

Hence, p is a sublinear functional, as wanted.

Define a new linear functional ψ on B by $h + rf \to h(x_0) + r\overline{f}(x_0)$ where $h \in A$ and $r \in \mathbb{R}$. We will show that ψ is dominated on B by p, i.e., $\psi(g) \leq p(g)$ for all $g \in B$. This is the same as $\psi(h + rf) = h(x_0) + r\overline{f}(x_0) \leq \overline{h + rf}(x_0)$ for all $h \in A$ and $r \in \mathbb{R}$. Let $r \geq 0$, then by Proposition 4.3 (c), we have that

$$\psi(h + rf) = h(x_0) + r\overline{f}(x_0) = h(x_0) + \overline{rf}(x_0) = \overline{h + rf}(x_0).$$

If $r < 0$, then $h + rf$ is concave. Indeed, since $-h - rf$ is convex by these calculations

$$\begin{aligned} (-h - rf)\left(\lambda x + (1-\lambda)y\right) &= -h\left(\lambda x + (1-\lambda)y\right) - rf\left(\lambda x + (1-\lambda)y\right) \\ &\leq -\lambda h(x) - (1-\lambda)h(y) - \lambda rf(x) - (1-\lambda)rf(y) \\ &= \lambda(-h - rf)(x) + (1-\lambda)(-h - rf)(y). \end{aligned}$$

Hence $h + rf$ is concave and easily seen to be continuous - in particular upper semicontinuous. Hence, due to Proposition 4.3 (b), $\overline{h + rf} = h + rf \geq h + r\overline{f}$, where the latter comes from the fact that $r < 0$.

By the Hahn-Banach Extension Theorem [4, 5.6], then, there exists a linear functional m on $C(X)$ such that $m(g) \leq p(g) = \overline{g}(x_0)$ for $g \in C(X)$

and $m(g) = \psi(g)$ for $g \in B$, i.e., $m(h + rf) = h(x_0) + r\overline{f}(x_0)$ for $h \in A$ and $r \in \mathbb{R}$. We want to show that m is a positive linear functional with $m(1) = 1$. Note that $1 \in A$ since all constant functions a clearly affine, thus $m(1) = m(1 + 0f) = 1$. Furthermore, let $g \in C(X)$ with $g \leq 0$, then by definition $\overline{g}(x_0) = \inf\{h(x) \mid h \in A \text{ and } h \geq g\} \leq 0$. Thus, $m(g) \leq p(g) = \overline{g}(x_0) \leq 0$. Now let $g \geq 0$, then $-g \leq 0$ and $m(-g) \leq 0$ and since m is linear, $-m(g) \leq 0$ and thus $m(g) \geq 0$. All in all, m is a positive linear functional. Note that $C_c(X) = C(X)$ since X is compact. Thus, due to the Riesz Representation Theorem [4, 7.2] there exists a Radon measure μ on X such that

$$m(g) = \int_X g \, \mathrm{d}\mu = \mu(g)$$

for $g \in C(X)$. Notice that

$$\mu(X) = \int_X 1 \, \mathrm{d}\mu = m(1) = 1,$$

hence, μ is a probability measure. Now we need to show that μ represents x_0. Consider $\phi \in E^*$, since ϕ is linear it is in A as well, hence in B. Therefore we have

$$\mu(\phi) = \int_X \phi \, \mathrm{d}\mu = m(\phi) = \phi(x_0),$$

thus μ represents x_0.

The only thing left is to show that μ is supported by the extreme points of X. For this, we notice $f \leq \overline{f}$ by Proposition 4.3 (b) end hence $\mu(f) \leq \mu(\overline{f})$. If $h \in A$ and $h \geq f$, then $h(x) \geq \inf\{h'(x) \mid h' \in A,\ h' \geq f\} = \overline{f}(x)$, thus $h(x_0) = \mu(h) \geq \mu(\overline{f})$. On the other hand, we also have $\mu(f) = m(f) = \overline{f}(x_0) = \inf\{h(x_0) \mid h \in A,\ h \geq f\} \geq \mu(\overline{f})$. Therefore

$$\mu(f) = \mu(\overline{f}).$$

Define $\mathcal{E} = \{x \in X \mid f(x) = \overline{f}(x)\}$. We next show that μ vanishes on the complement of $\mathcal{E}$. This is a standard result in measure theory, but we include the proof. Since $\mu(f) = \mu(\overline{f})$, we have

$$0 = \int_X \overline{f} \, \mathrm{d}\mu - \int_X f \, \mathrm{d}\mu = \int_X (\overline{f} - f) \, \mathrm{d}\mu.$$

As $f \leq \overline{f}$, $\overline{f} - f \geq 0$. Now, for all $n \in \mathbb{N}$, let $E_n = \{x \in X \mid \overline{f} - f \geq 1/n\}$. As $0 \leq 1_{E_n}(\bar{f} - f) \leq \bar{f} - f$, we have

$$0 = \int_X \bar{f} - f \, \mathrm{d}\mu \geq \int_{E_n} \bar{f} - f \, \mathrm{d}\mu \geq \int_{E_n} \frac{1}{n} \, \mathrm{d}\mu = \frac{1}{n}\mu(E_n),$$

so $\mu(E_n) = 0$ for all $n \in \mathbb{N}$. Hence, $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = 0$, but $\bigcup_{n=1}^{\infty} E_n = \{x \in X \mid \bar{f} - f > 0\}$, which shows us that μ vanishes on $\{x \in X \mid \bar{f} - f > 0\} = \{x \in X \mid \bar{f} > f\} = \mathcal{E}^c$.

Therefore, we only need to show that $\mathcal{E}$ is contained in the set of extreme points of X.

For this, let $x \in X$ and suppose $x = \lambda x_1 + (1-\lambda)x_2$ for $\lambda \in (0,1)$ and $x_1, x_2 \in X$ with $x \neq x_1$, $x \neq x_2$. That $\mathcal{E} \subseteq \operatorname{Ext}(X)$ is the same as $\operatorname{Ext}(X)^c \subseteq \mathcal{E}^c$. Let $x \in \operatorname{Ext}(X)^c$. Since f is strictly convex and $f \leq \overline{f}$, we have

$$\begin{aligned} f(x) = f\left(\lambda x_1 + (1-\lambda)x_2\right) &< \lambda f(x_1) + (1-\lambda) f(x_2) \\ &\leq \lambda \overline{f}(x_1) + (1-\lambda)\overline{f}(x_2), \end{aligned}$$

but we also have

$$\begin{aligned} \overline{f}(x) &= \overline{f}(\lambda x_1 + (1-\lambda)x_2) \\ &= \inf\left\{h\left(\lambda x_1 + (1-\lambda)x_2\right) \mid h \in A,\ h \geq f\right\} \\ &= \inf\left\{\lambda h(x_1) + (1-\lambda)h(x_2) \mid h \in A,\ h \geq f\right\} \\ &\geq \inf\left\{\lambda h(x_1) \mid h \in A,\ h \geq f\right\} + \inf\left\{(1-\lambda)h(x_2) \mid h \in A,\ h \geq f\right\} \\ &= \lambda \inf\left\{h(x_1) \mid h \in A,\ h \geq f\right\} + (1-\lambda)\inf\left\{h(x_2) \mid h \in A,\ h \geq f\right\} \\ &= \lambda \overline{f}(x_1) + (1-\lambda)\overline{f}(x_2). \end{aligned}$$

So all in all, we have that $f(x) < \overline{f}(x)$, thus $x \in \mathcal{E}^c$, as wanted. □

Choquet's theory of integral representations of convex sets in terms of their extreme points provides a unified approach to a number of important results in harmonic analysis and functional analysis. In this thesis we have only "scratched the surface", and we hope to pursue these topics further in the future.

A Separability

Recall that a metric space (M, d) is called *separable* if it contains a countable subset $(x_n)_{n\geq 1}$ which is dense in M (i.e., for all $x \in M$ and every $\varepsilon > 0$, there exists $n \geq 1$ such that $d(x, x_n) < \varepsilon$).

The following result provides us with examples of separable metric spaces.

Lemma A.1. *If M is a compact metric space, then M is separable.*

Proof. For every $\varepsilon > 0$, the family of open balls $\{B(x, \varepsilon) \mid x \in M\}$ is an open cover of M. Hence, by compactness of M, there is a finite subfamily covering M. This argument shows, in particular, that for $K \in \mathbb{N}$, there exists a finite set $F_k \subseteq M$ such that $\bigcup_{x\in F_k} B(x, 1/k)$ covers M. Now let $F = \bigcup_{k\in\mathbb{N}} F_k$. Then F is a countable subset of M, so there is an enumeration $(x_n)_{n\geq 1}$ of F. We want to show that $(x_n)_{n\geq 1}$ is dense in M. For this, let $\varepsilon_0 > 0$ and $x \in M$ be given. Then there exists $k_0 \in \mathbb{N}$ such that $1/k_0 < \varepsilon_0$. By construction of the family, $F_{k_0} \subseteq \{x_n \mid n \geq 1\}$ so there exists $n_0 \in \mathbb{N}$ such that

$$x \in B\left(x_{n_0}, \frac{1}{k_0}\right) \subseteq B(x_{n_0}, \varepsilon_0),$$

and we are done. □

The following lemma will be quite useful to us.

Lemma A.2. *If (M, d) is a separable metric space and M_0 is a non-empty subset of M, then M_0 is separable as well.*

Proof. Let $(x_n)_{n\geq 1}$ be dense in M. Set

$$I = \left\{(n, m) \in \mathbb{N} \times \mathbb{N} \mid B\left(x_n, \frac{1}{m}\right) \cap M_0 \neq \emptyset\right\}.$$

Let $y \in M_0$, then for all $m \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that $x_n \in B(y, 1/m)$ which implies that $y \in B(x_n, 1/m)$, i.e., $(n, m) \in I$, thus $I \neq \emptyset$.

Now, for every $(n, m) \in I$, choose $y_{(n,m)} \in B(x_n, 1/m) \cap M_0$. Then the set $\mathcal{C} := \{y_{(n,m)} \mid (n, m) \in I\}$ is a countable subset of M_0. To show it is also dense, let $z \in M_0$ and $0 < \varepsilon < 1$. Choose $x_0 \in B(z, \varepsilon/4) \cap \{x_n \mid n \geq 1\}$, then there exists $N \in \mathbb{N}$ such that $\varepsilon/4 \leq 1/N \leq \varepsilon/2$. Since $z \in B(x_0, 1/N) \cap M_0 \neq \emptyset$, there exists $y_{(n,m)} \in \mathcal{C}$ such that $d(y_{(n,m)}, x_0) < 1/N$. Now, we have that $d(y_{(n,m)}, z) \leq d(y_{(n,m)}, x_0) + d(x_0, z) < 1/N + 1/N \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$, hence $y_{(n,m)} \in B(z, \varepsilon)$ and therefore M_0 has a countable dense subset $\mathcal{C}$, thus is separable. □

We now prove the main result concerning separability:

Theorem A.3. *Suppose that (M, d) is a compact metric space. Then $C(M)$ (the space of real-valued continuous functions on M, equipped with the metric given by the uniform norm) is separable.*

Proof. Since M is a compact metric space, Theorem A.1 gives us that M itself is separable. Hence, there exists a countable dense subset of M, i.e., there exists $(x_n)_{n\geq 1} \subseteq M$ such that for all $x \in M$ and every $\varepsilon > 0$, we can choose x_n with $d_M(x, x_n) < \varepsilon$. We define a sequence of functions $g_n \in C(M)$ by $g_n(x) = d_M(x, x_n)$. Given an arbitrary finite string of indices $(n_1, \dots, n_k)$, we define

$$g_{n_1,\dots,n_k}(x) = g_{n_1}(x) \cdots g_{n_k}(x).$$

Let $\mathcal{A}$ be the set of all linear combinations of $g_{n_1,\dots,n_k}$, i.e,

$$\mathcal{A} = \left\{ a + \sum_{n_1,\dots,n_k} a_{n_1,\dots,n_k} g_{n_1,\dots,n_k} \;\middle|\; a, a_{n_1,\dots,n_k} \in \mathbb{R} \right\},$$

where the above sums are finite. Note that $\mathcal{A}$ is not necessarily countable. Notice that $\mathcal{A}$ contains the constant functions and observe furthermore that $\mathcal{A}$ is clearly an algebra. Our aim is for $\mathcal{A}$ to satisfy the conditions for the Stone-Weierstrass Theorem [4, Theorem 4.45], hence it remains to show that $\mathcal{A}$ separates points of M. Thus, suppose $x, y \in M$ with $x \neq y$ and let $d_M(x, y) = d_M(y, x) = 2\varepsilon > 0$. Due to separability of M, we can choose $x_n \in M$ such that $d_M(x, x_n) < \varepsilon$. Now, consider $g_n \in \mathcal{A}$. By definition, $g_n(x) < \varepsilon$. But, we also have

$$g_n(y) = d_M(y, x_n) \geq d_M(y, x) - d_M(x, x_n) \geq 2\varepsilon - \varepsilon = \varepsilon,$$

thus $g_n(x) \neq g_n(y)$. Now, applying the Stone-Weierstrass Theorem, we get that $\overline{\mathcal{A}} = C(M)$, hence $\mathcal{A}$ is a dense subset of $C(M)$. We still need the subset to be countable. For this consider the set

$$\mathcal{A}' = \left\{ a + \sum_{n_1,\dots,n_k} a_{n_1,\dots,n_k} g_{n_1,\dots,n_k} \mid a, a_{n_1,\dots,n_k} \in \mathbb{Q} \right\}.$$

Obviously, $\mathcal{A}'$ is countable and dense in $\mathcal{A}$, therefore $\mathcal{A}'$ is a countable dense subset of $C(M)$. $\square$

References

[1] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. *Harmonic analysis on semigroups*, volume 100 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Theory of positive definite and related functions.

[2] Dimitri P. Bertsekas. *Convex analysis and optimization*. Athena Scientific, Belmont, MA, 2003. With Angelia Nedić and Asuman E. Ozdaglar.

[3] Gustave Choquet. *Lectures on analysis. Vol. II: Representation theory*. Edited by J. Marsden, T. Lance and S. Gelbart. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[4] Gerald B. Folland. *Real Analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[5] Robert R. Phelps. *Lectures on Choquet's theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001.