# On Choquet's Theorem

Thesis Report MAT499

Venkata Subramaniam J
Under the supervision of Prof. Taduri S.S.R.K. Rao
Shiv Nadar Institute of Eminence

Spring 2025

# Contents

## Acknowledgements

I would like to thank the continued guidance of Professor Taduri S.S.R.K. Rao in the course of my thesis, and for helping me explore the interesting field of functional analysis.

**Abstract**

Choquet's theorem is a result for some convex sets in metrisable spaces that provides a method to represent points as 'barycentres' of probability measures supported by its extreme points. It generalises the Krein-Milman theorem in a setting of integrals over some measures, and provides a translation of some geometric properties into measure theoretical analytical ones. In this sense it is a measure theoretic version of the Krein-Milman theorem, or an *integral representation* of the Krein-Milman theorem.

There are many applications for the theorem, including in mathematical physics and economics. One such application in the theory of economics is shown in 5.

The theorem we ultimately aim to show is the following.

**Theorem** (Choquet)**.** *Suppose $Q$ is a metrisable compact convex subset of a locally convex topological vector space $X$. Let $x_0 \in Q$. Then there is a probability measure $\mu$ on $Q$ whose barycentre is $x_0$ and $\mu$ is supported by the extreme points of $Q$.*

# 1 Introduction

**Definition 1.1** (Extreme Set)**.** Let $X$ be a vector space and $K \subseteq X$. A set $\varnothing \neq S \subseteq K$ is said to be an *extreme set* of $K$ if for every $x \in K$, $y \in K$ and $t \in (0,1)$ where

$$(1-t)x + ty \in S$$

we have $x \in S$ and $y \in S$.

And an *extreme point* of $K$ are the extreme sets in it that are singletons. The set of all such points will be denoted $E(K)$. A classical result of the kind which we wish to generalise is the following.

**Theorem 1.2** (Carathéodory)**.** *If $X \subseteq \mathbb{R}^n$ is compact convex, then every $x \in X$ is a convex combination of at most $n+1$ extreme points of $X$.*

While the Krein-Milman (2.3) theorem generalises this into infinite dimensions, to get a constructive way to represent these points, it is not enough. That is what motivates us to search for integral representations, using Borel measures.

Stated first are some preliminary results from functional analysis and measure theory, namely the Hahn-Banach extension theorem and the Riesz representation theorem, both of which will be used extensively in the subsequent sections.

**Theorem 1.3** (Hahn-Banach Extension)**.** *Let $X$ be a normed space, and $M$ be a subspace. Let $f$ be a continuous linear functional on $M$. Then there is an $f$ , a continuous linear functional on $X$ such that $\tilde{f} = f$ on $M$ and $\|\tilde{f}\| = \|f\|$.*

**Theorem 1.4.** *If $X$ is a locally compact Hausdorff space, then every bounded linear functional $\Phi$ on $C_0(X)$ is represented by a unique regular Borel measure $\mu$ such that*

$$\Phi f = \int_X f \, \mathrm{d}\mu$$

*for every $f \in C_0(X)$. Moreover,*

$$\|\Phi\| = |\mu|(X)$$

*And $\mu$ is positive iff $\Phi$ is positive.*

Let us motivate the main theorem by an example that makes it clear why such a representation exists in some cases.

**Example 1.5.** Let $\Omega$ be a compact Hausdorff space and consider $C(\Omega)$, the continuous real functionals on $\Omega$.

Denote by $C(\Omega)^*_1$ the closed unit ball in $C(\Omega)^*$. Now, we consider the subspace $A(\Omega)$, that is, the affine continuous real functionals on $\Omega$.

Note that $1 \in A(\Omega)$. Now consider a positive functional $\Lambda \in A(\Omega)^*$ such that $\|\Lambda\| = 1$ and $\Lambda(1) = 1$. By the Hahn-Banach extension theorem, we are able to get a $\tilde{\Lambda} \in C(\Omega)^*$ such that $\|\tilde{\Lambda}\| = \|\Lambda\|$ and $\tilde{\Lambda}(1) = 1$, and $\tilde{\Lambda} = \Lambda$ on $A(\Omega)^*$.

However, by the Riesz representation theorem, we know that there is a unique regular positive (as the functional is positive) Borel measure such that

$$\tilde{\Lambda}(a) = \int_\Omega a \, \mathrm{d}\mu$$

for every $a \in C(K)$ and $\mu(X) = \|\tilde{\Lambda}\| = 1$. Then $\mu$ is a probability measure. Restricting $a$ back to $A(K)$, we get a unique representation $\mu$ for every $a \in A(K)$ such that

$$\Lambda(a) = \int_\Omega a \, \mathrm{d}\mu$$

So we have really identified these elements of norm 1 with the probability measures on $K$, $\mathscr{P}(K)$.

# 2 Krein-Milman Theorem

**Theorem 2.1** (Hahn-Banach Separation)**.** *Suppose $A$ and $B$ are disjoint nonempty convex subsets of a locally convex topological vector space $X$. Then if $A$ is compact, $B$ is closed, and $X$ is locally convex, then there exist $\lambda \in X^*$, $\gamma_1 \in \mathbb{R}$, $\gamma_2 \in \mathbb{R}$ such that*

$$\text{Re } \Lambda x < \gamma_1 < \gamma_2 < \text{Re } \Lambda y$$

*for every $x \in A$ and for every $y \in B$*

This well known result (see 3.4 from [4] for a proof) will be used to prove the following theorem.

**Theorem 2.2.** *Let $X$ be a topological vector space on which $X^*$ separates points. Suppose $A \subseteq X$ and $B \subseteq X$ are disjoint nonempty compact convex sets in $X$. Then there exists $\Lambda \in X^*$ such that*

$$\sup_{x \in A} \text{Re}(\Lambda x) < \inf_{x \in B} \text{Re}(\Lambda y)$$

*for every $x \in A$, $y \in B$*

*Proof.* Note first that if $X$ is locally convex, then $X^*$ separates points, so we are using a weaker assumption.

Denote by $X_w$ the space $X$ imbued with the weak topology. Then $A$ and $B$ are compact in $X_w$, as they are compact in the strong topology, and they are convex sets, implying they are weakly compact.

Since the weak topology is Hausdorff, they are also weakly closed, and as $X_w$ is locally convex, we may apply 2.1 to obtain a $\Lambda \in (X_w)^*$ such that

$$\sup_{x \in A} \operatorname{Re}(\Lambda x) < \inf_{x \in B} \operatorname{Re}(\Lambda y)$$

But we know that $(X_w)^* = X^*$, so $\Lambda \in X^*$ and we have obtain our required map. ∎

Now we are ready to prove the Krein-Milman theorem. Note that by $\operatorname{co}(\cdot)$ we denote the *convex hull* of a set, or the smallest convex set containing it.

**Theorem 2.3** (Krein-Milman)**.** *Let $X$ be a topological vector space on which $X^*$ separates points. If $\varnothing \neq K \subseteq X$ is a compact convex set, then $K = \overline{\operatorname{co}}(E(K))$.*

*Proof.* Let $\mathcal{P}$ be the set of compact extreme sets of $K$. Since $K$ itself belongs to the set, it is nonempty. We will use two properties of $\mathcal{P}$.

a) If $L$ is a nonempty subcollection of $\mathcal{P}$, then its intersection $S = \cap_{l \in L} l \in \mathcal{P}$, if $S \neq \varnothing$

b) If $S \in \mathcal{P}$, $\Lambda \in X^*$, $\mu$ is the maximum of $\operatorname{Re} \Lambda$ on $S$ and

$$S_\Lambda = \{x \in S : \operatorname{Re} \Lambda x = \mu\}$$

then $S_\Lambda \in \mathcal{P}$.

To show a), note that $S$ is compact as we have a Hausdorff space, and $S$ is an extreme set, so $S \in \mathcal{P}$.

Now b) will be shown. Let $tx + (1-t)y = z \in S_\Lambda$, $x \in K$, $y \in K$, $0 < t < 1$. Since $z \in S$ and $S \in \mathcal{P}$, we have that $x \in S$ and $y \in S$. Thus $\operatorname{Re} \Lambda x \leq \mu$, and $\operatorname{Re} \Lambda y \leq \mu$. Since $\operatorname{Re} \Lambda z = \mu$ and $\Lambda$ is linear, we have that $\mu = \operatorname{Re} \Lambda x = \operatorname{Re} \Lambda y$. Thus we have that $x \in S_\Lambda$ and $y \in S_\Lambda$, so b) has been proven.

Now, let $S \in \mathcal{P}$ and let

$$\mathcal{P}' = \{P \in \mathcal{P} : P \subseteq S\}$$

since $S \in \mathcal{P}'$, $\mathcal{P}' \neq \varnothing$. Now let $(\mathcal{P}', \subseteq)$ be a partially ordered set, and let $\Omega$ be a maximal totally ordered subset of $\mathcal{P}'$, and let $M$ be the set

$$M = \bigcap_{O \in \Omega} O$$

since $\Omega$ has compact sets with the finite intersection property, $M \neq \varnothing$. By a), we know that $M \in \mathcal{P}'$. Since $\Omega$ is maximal, there is no proper subset of $M$ in $\mathcal{P}$. Now, apply b) so that every $\Lambda \in X^*$ is constant on $M$. Since $X^*$ separates points, $M$ only has one point and thus $M \in E(K)$. Thus we have shown that $E(K) \cap S \neq \varnothing$ for every $S \in \mathcal{P}$, or, every compact extreme set of $K$ contains an extreme point of $K$.

As $K$ is compact convex, we know $\overline{\text{co}}(E(K)) \subseteq K$, and so $\overline{\text{co}}(E(K))$ is compact.

Let, if possible, $x_0 \in K$ such that $x_0 \notin \overline{\text{co}}(E(K))$. Then by 2.2, we have a $\Lambda \in X^*$ such that Re $\Lambda x <$ Re $\Lambda x_0$ for every $x \in \overline{\text{co}}(E(K))$. If $K_\Lambda$ is defined as in b), then $K_\Lambda \in \mathcal{P}$, but then $K_\Lambda$ is disjoint from $\overline{\text{co}}(E(K))$, which is a contradiction to $E(K) \cap S \neq \varnothing$ with $S = K$, so $K = \overline{\text{co}}(E(K))$. ■

However, the Krein-Milman theorem does not suffice for every case. This can be illustrated by the following example.

**Example 2.4.** Consider the space Hilbert space $\mathscr{H} = \ell^2$, and the closed unit ball

$$B = \ell^2_1 = \left\{ \{x_n\}_{n\geq 1} \subseteq \mathbb{C} : \sum |x_n|^2 \leq 1 \right\}$$

We wish to find the extreme points of $B$. Clearly an extreme point must have norm 1, as if $\|x\| < 1$ for some $x \in B$ then it is the convex combination of some two other points in the ball. So the extreme points must have $\|x\| = 1$.

Now, we show that points of norm 1 are in fact extreme points. Let $x \in B$ such that $\|x\| = 1$. Suppose for the sake of contradiction that there are points $y = \{y_n\} \neq z = \{z_n\} \in B$ and some $0 < t < 1$ such that $x = ty + (1-t)z$. Then, we have that $\|x\|^2 = 1$ and from the parallelogram law,

$$\begin{aligned} \|x\|^2 &= \|ty + (1-t)z\|^2 \\ &= \|ty\|^2 + \|(1-t)z\|^2 + 2t(1-t)\langle y, z\rangle \\ &= t^2\|y\|^2 + (1-t)^2\|z\|^2 + 2t(1-t)\langle y, z\rangle \end{aligned}$$

and as $\|y\|^2 \leq 1$ and $\|z\|^2 \leq 1$,

$$\begin{aligned} 1 = \|x\|^2 &\leq t^2 + (1-t)^2 + 2t(1-t)\langle y, z\rangle \\ &= 1 - 2t + 2t^2 + 2t(1-t)\langle y, z\rangle \\ &= 1 - 2t(1-t)(\langle y, z\rangle - 1) \end{aligned}$$

thus $\langle y, z\rangle \geq 1$, but from Cauchy-Schwartz, $|\langle y, z\rangle| \leq 1$, so $\langle y, z\rangle = 1$. And so finally

$$\begin{aligned}
&\|y - z\|^2 \\
=&\|y\|^2 + \|z\|^2 - 2\langle y, z\rangle \\
\leq&1 + 1 - 2(1) \\
=&0
\end{aligned}$$

so $\|y - z\| = 0$, and $y = z$ which is a contradiction.

We have shown that $E(B) = \{\{x_n\} \in B : \|\{x_n\}\| = 1\}$. But from Banach-Alaoglu we know $B$ is weakly compact convex and so we get that the closed convex hull $\overline{\text{co}}(E(B)) = B$ as $E(B)$ is weakly dense in $B$. So in this case we are not getting a useful identification from the Krein-Milman theorem.

It is clear that a sharpening of the theorem is necessary, as by taking the closure we forego a way to get a concrete constructive representation.

Milman's theorem is also relevant, for which an alternate proof will also be provided using the developed theory.

**Theorem 2.5** (Milman)**.** *If $K$ is a compact set in a locally convex topological vector space $X$, and if $\overline{\text{co}}(K)$ is compact, then every extreme point of $\overline{\text{co}}(K)$ lies in $K$.*

*Proof.* Suppose an extreme point $p$ of $\overline{\text{co}}(K)$ is such that $p \notin K$. Then there is a convex balanced neighbourhood $V$ of 0 in $X$ such that

$$(p + \bar{V}) \cap K = \varnothing$$

and let $x_1, \ldots x_n \in K$ such that

$$K \subseteq \bigcup_{i=1}^{n}(x_i + V)$$

then each $A_i = \overline{\text{co}}(K \cap (x_i + V))$ is compact convex as $A_i \subseteq \overline{\text{co}}(K)$. And $K \subseteq A_1 \cup \cdots \cup A_n$. But then $\text{co}(A_1 \cup \cdots \cup A_n)$ is also compact, and

$$\overline{\text{co}}(K) \subseteq \overline{\text{co}}(A_1 \cup \cdots \cup A_n) = \text{co}(A_1 \cup \cdots \cup A_n)$$

But the reverse inclusion is also true since $A_i \subseteq \overline{\text{co}}(K)\ \forall\, i$. So,

$$\overline{\text{co}}(K) = \text{co}(A_1 \cup \cdots \cup A_n)$$

Let $p = t_1 y_1 + \cdots + t_N y_N$ where $y_j \in A_i$ and $t_j > 0$, $\sum t_j = 1$. So

$$p = t_1 y_1 + (1 - t)\frac{t_2 y_2 + \cdots + t_N y_N}{t_2 + \cdots + t_N}$$

shows $p$ as a convex combination of points in $\overline{\text{co}}(K)$, and since $p$ is an extreme point of $\overline{\text{co}}(K)$, we must have $y_1 = p$. Thus for some $i$

$$p \in A_i \subseteq x_i + \bar{V} \subseteq K + \bar{V}$$

which is a contradiction. ■

# 3 Integral Representations

Some integral representation theorems will be shown, leading up to a way to generalise the Krein-Milman theorem.

First, we will define 'vector-valued' integration. Let $Q$ be a measure space with real or complex measure $\mu$. Let $f : Q \to X$ be a function where $X$ is a topological vector space. We wish to associate a value to

$$\int_Q f \, \mathrm{d}\mu$$

**Definition 3.1.** Let $Q$ be a measure space and $\mu$ be a measure (real or complex) on it. Let $X$ be a topological vector space on which $X^*$ separates points. Let $f : Q \to X$ be a function such that $\Lambda f$ is integrable for every $\Lambda \in X^*$. Note that if $q \in Q$ then $(\Lambda f)(q) = \Lambda(f(q))$. If there exists $y \in X$ such that

$$\Lambda y = \int_Q (\Lambda f) \, \mathrm{d}\mu$$

for each $\Lambda \in X^*$ then we define

$$\int_Q f \, \mathrm{d}\mu = y$$

The requirement for this definition is that

$$\Lambda \left( \int_Q f \, \mathrm{d}\mu \right) = \int_Q (\Lambda f) \, \mathrm{d}\mu$$

for a $\Lambda \in X^*$ as $\Lambda$ is linear it preserves sums, and integration is the limit of sums. This is the motivation for the definition.

This value is well defined. Let us show this.

$X^*$ separates points, so there is at most one $y$. Thus the value is unique and well defined. Next the existence of it must be ensured. For our purposes it is enough to show it for cases where $Q$ is compact with continuous $f$. Then $f(Q)$ is compact, and $\overline{\text{co}}(f(Q))$ must be compact.

A probability measure is defined to be a positive measure of total mass 1, or $\mu(X) = 1$.

Recall also the Jordan decomposition theorem, that any complex measure may be decomposed into four positive probability measures.

We furnish the existence with the following theorem

**Theorem 3.2.** *Let $X$ be a topological vector space on which $X^*$ separates points and let $\mu$ be a Borel probability measure on a compact Hausdorff space $Q$. Then if $f : Q \to X$ is continuous and $\overline{\text{co}}(f(Q))$ is compact, then*

$$y = \int_Q f \, \mathrm{d}\mu$$

*exists.*

*Proof.* We only need to prove this for real spaces, since if $\nu$ is any positive Borel measure on $Q$, then $\lambda\nu$ is a probability measure for some $\lambda \in \nu$. Thus, by the Jordan decomposition theorem, we may extend it to any complex measure.

Thus suppose $X$ to be a real space. Let $H = \text{co}(f(Q))$. Let

$$L = \{\Lambda_1, \ldots, \Lambda_n\}$$

be a finite collection in $X^*$ and let

$$E_L = \{y \in \overline{H} : \Lambda y = \int_Q (\Lambda f) \, \mathrm{d}\mu \;\; \forall \Lambda \in L\}$$

Now let $\Lambda_i \in L$ and

$$E_{\Lambda_i} = \{y \in \overline{H} : \Lambda_i y = \int_Q (\Lambda_i f) \, \mathrm{d}\mu\}$$

and consider a net $\{y_n\} \subseteq E_{\Lambda_i}$ such that $y_n \to y$ for a $y \in \overline{H}$. We wish to show $y \in E_{\Lambda_i}$.

From the continuity of $\Lambda$ we have

$$\Lambda y = \lambda \lim y_n = \lim \Lambda y_n = \int_Q (\Lambda f) \, \mathrm{d}\mu$$

and thus $y \in E_{\Lambda_i}$ and the set is closed. We have $E_L = \bigcup_{i=1}^n E_{\Lambda_i}$ so $E_L$ is closed. Since $\overline{H}$ is compact by hypothesis, $E_L$ is compact. If we show $E_L \neq \varnothing$ for any $L$, then the set of all such $E_L$ has the finite intersection property, and their intersection is nonempty. Thus there exists $y$ in this intersection, which would imply $\Lambda y = \int_Q \Lambda f \, \mathrm{d}\mu$ for every $\Lambda \in X^*$. Thus we will now prove $E_L \neq \varnothing$.

Regard $L$ as a mapping $L : X \to \mathbb{R}^n$ such that

$$L(x) = (\Lambda_1(x), \ldots, \Lambda_n(x))$$

and put $K = L(f(Q))$.

Then put

$$m_i = \int_Q (\Lambda_i f) \, \mathrm{d}\mu \quad (1 \leq i \leq n)$$

And define $m = (m_1, \ldots, m_n)$. It will now be shown that $m \in co(K)$.

Since $L$ and $f$ are continuous maps and so $L \circ f$ must be a continuous map. $Q$ is compact, so $K$ is compact. Since $K \subseteq \mathbb{R}^n$, we know that $co(K)$ is also compact. Now, $\mathbb{R} \setminus co(K)$ is also convex, and since $co(K)$ is open (by form of convex hull in $\mathbb{R}^n$) we have that $\mathbb{R} \setminus co(K)$ is a closed convex set. Now, we may apply 2.1 to obtain $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$\sum_{i=1}^{n} c_i u_i < \sum_{i=1}^{n} c_i t_i$$

for $u = (u_i)_{1 \leq i \leq n} \in K$. Thus for $q \in Q$

$$\sum_{i=1}^{n} c_i \Lambda_i f(q) < \sum_{i=1}^{n} c_i t_i$$

Now, we integrate the left side of the inequality over the probability measure $\mu$ to obtain $\sum c_i m_i < \sum c_i t_i$ and so $t \neq m$. Thus, we must have $m \in co(K)$. Since $K = L(f(Q))$ and $L$ is linear, $m = Ly$ for some $y \in H = co(f(Q))$. For this $y$ we get

$$\Lambda_i y = m_i = \int_Q (\Lambda_i f) \, \mathrm{d}\mu$$

and thus $y \in E_L$. So $E_L \neq \varnothing$ and by earlier results, the proof is complete. ∎

Now we are ready to finally get a characterisation of points in a compact convex sets in terms of their representations.

**Theorem 3.3.** *Suppose $X$ is a topological vector space on which $X^*$ separates points, $Q \subseteq X$ is compact and $\overline{H} = \overline{\text{co}}(Q)$ is compact. Then $y \in \overline{H}$ iff there is a regular Borel probability measure $\mu$ on $Q$ such that*

$$y = \int_Q x \, \mathrm{d}\mu \, (x)$$

*Proof.* Take $X$ to be a real space. Consider $C(Q)$, the space of real continuous functions on $Q$ as a normed space with the sup-norm (making it Banach). Now, by 1.4, the space $C(Q)^*$ of real functionals may be unique identified with the space of real Borel measures on $Q$ that are differences of regular positive measures. Thus we will define the mapping

$$\phi : C(Q)^* \to X$$

by

$$\phi(\mu) = \int_Q x \, \mathrm{d}\mu \, (x)$$

Let $P$ be the set of all regular Borel probability measures on $Q$. We will thus prove that $\phi(P) = \overline{H}$.

Let $x \in Q$ and consider the unit mass

$$\delta_x = \begin{cases} 0 & x \notin Q \\ 1 & x \in Q \end{cases}$$

Then clearly $\delta_x \in P$. Since $\phi(\delta_x) = \int_Q x \, \mathrm{d}\delta_x \, (x) = x$, we have $Q \subseteq \phi(P)$. $\phi$ is linear and $P$ is convex so $H \subseteq \phi(P)$. By the previous theorem we get $\phi(P) \subseteq \overline{H}$. Since we have $\phi(P) \subseteq \overline{H}$ and $H \subseteq \phi(P)$, if $\phi(P)$ is a closed set then $\phi(P) = H$. We will prove this.

First, it will be shown that $P$ is w* compact in $C(Q)^*$. First, note that

$$P \subseteq \left\{ \mu : \left| \int_Q h \, \mathrm{d}\mu \right| < 1, \ \|h\| < 1 \right\}$$

and this set is w* compact by the Banach-Alaoglu theorem. If $h \in C(Q)$ and $h \geq 0$ define

$$E_h = \left\{ \mu : \int_Q h \, \mathrm{d}\mu \geq 0 \right\}$$

The map $\mu \to \int h \, \mathrm{d}\mu$ is continuous so every $E_h$ must be w* closed. The same is applied to say that the set

$$E = \left\{ \mu : \int_Q 1 \, \mathrm{d}\mu = 1 \right\}$$

is w* closed. Since $P$ is the intersection of $E$ and every $E_h$, it is also w* closed and since it is a subset of a w* compact set, it is w* compact.

Now we prove $\phi$ is continuous when $C(Q)^*$ is given the w* topology and $X$ the weak topology.

Note that it is a linear map. Every weak neighbourhood of 0 has a set of the form

$$W = \{y \in X : |\Lambda_i y| < r_i \ 1 \le i \le n\}$$

where $\Lambda_i \in X^*$ and $r_i > 0$. If we restrict $\Lambda_i$ to $Q$ then

$$V = \left\{\mu \in C(Q)^* : \left|\int_Q \Lambda_i \,\mathrm{d}\mu\right| < r_i \ 1 \le i \le n\right\}$$

is a w* neighbourhood of 0 in $C(Q)^*$, but then $\int_Q \Lambda_i \,\mathrm{d}\mu = \Lambda_i(\int_Q x \,\mathrm{d}\mu\,(x)) = \Lambda_i \phi(\mu)$ by definition and so $\phi(V) \subseteq W$ and so $\phi$ is continuous.

Thus we have that $\phi(P)$ is weakly compact and thus weakly closed. So it is also originally (strongly) closed and so $\phi(P) = H$. ■

Through this, we may reformulate the Krein-Milman theorem in terms of integral representations, but first we define some terms.

**Definition 3.4.** Let $\mu$ be a positive regular Borel measure on the compact Hausdorff space $Q$. Let $S \subseteq Q$ be Borel. Then $\mu$ is supported by $S$ if $\mu(X \setminus S) = 0$.

**Definition 3.5.** Let $Q$ be a compact Hausdorff subset of a locally convex topological vector space $X$. $x$ is said to be the barycentre of a positive regular Borel measure $\mu$ on $Q$ if $\Lambda x = \int_Q \Lambda \,\mathrm{d}\mu$ for every $\Lambda \in X^*$.

**Theorem 3.6.** *Let $X$ be a locally convex topological vector space, and let $Q \subseteq X$ be compact convex. Then every $x \in Q$ is the barycentre of a probability measure $\mu$ on $Q$ supported on $\overline{E(Q)}$.*

*Proof.* Let $H = \overline{E(Q)}$, then $x \in \overline{\mathrm{co}}(H)$. By 3.3, there is a probability measure with $x$ barycentre. By the Hahn-Banach extension theorem, extend $\mu$ to $Q$, and we are done. ■

Now, to motivate the necessity to generalise this theorem, we need a representation theorem supported on extreme points, rather than the closure of the extreme points.

**Lemma 3.7.** *Suppose $Q$ is a compact convex subset of a topological vertor space $X$, and suppose that $Q$ is metrisable. Then $E(Q)$ is a $G$-$\delta$ set.*

*Proof.* Let $d$ be a metric on $Q$. Define

$$F_n = \left\{x : x = \frac{1}{2}(y+z), \ y, z \in Q, \ d(y,z) \ge \frac{1}{n}\right\}$$

Then clearly each $F_n$ is a closed set and $x \in Q$ is not an extreme point iff $x \in F_n$ for some $n \in \mathbb{N}$. So, $E(Q)^c$ is an $F$-$\sigma$ set and thus $E(Q)$ is a $G$-$\delta$ set. ■

$\delta_x$, the dirac measure, always represents any $x$. So, if $x$ is not an extreme point, there exist other measures whose barycentre is $x$. Thus a point is extreme iff $\delta_x$ is the only measure representing it.

This statement is known as the theorem of Bauer.

**Theorem 3.8** (Bauer)**.** *Let $X$ be a locally convex topological vector space and $Q$ be a compact convex subset. Then $x \in E(Q)$ iff $\delta_x$ is the only probability measure on $Q$ whose barycentre is $x$.*

*Proof.* Suppose $x \in E(Q)$ and is the barycentre of $\mu$. As $\mu$ is regular(it is a regular positive Borel measure), we wish to show that $\mu(D) = 0$ for every compact $D \subseteq Q \setminus \{x\}$, to show that $\mu$ is supported by $\{x\}$.

Suppose, for contradiction, that $\mu(D) > 0$ for one such compact set. Then there is a point $y \in D$ such that $\mu(U \cap Q) > 0$ for every neighbourhood $U$ of $y$, as $D$ is compact. Let $U$ be a closed convex neighbourhood such that $K := U \cap Q \subseteq Q \setminus \{x\}$. Then $K$ is compact convex and $0 < r = \mu(K) < 1$, as if $\mu(K) = 1$, then its barycentre will be in $K$.

Define the Borel measures $\mu_1, \mu_2$ such that $\mu_1(B) = \frac{1}{r}\mu(B \cap K)$ and $\mu_2(B) = \frac{1}{1-r}\mu(B \cap (Q \setminus K))$ for every Borel $B \subseteq Q$. Let $x_i$ be the barycentre of $\mu_i$, then since $\mu_1(K) = 1$, we have $x_1 \in K$ and so $x \neq x_1$, and we also have $\mu = r\mu_1 + (1-r)\mu_2$, so $x = rx_1 + (1-r)x_2$, which is a contradiction. ■

Now we may also offer an alternative proof of 2.5. Suppose $x \in E(\overline{\text{co}}(K))$. By 3.3, we have a measure $\mu$ on $K$ such that $x$ is the barycentre of $\mu$. By 3.8, $\mu = \delta_x$ and so $x \in K$ clearly.

# 4 Choquet's Theorem

We have finally developed the machinery to state the main theorem of Choquet, and the condition we put is metrisability.

**Theorem 4.1** (Choquet)**.** *Suppose $X$ is a metrisable compact convex subset of a locally convex topological vector space $E$. Let $x_0 \in X$. Then there is a probability measure $\mu$ on $X$ whose barycentre is $x_0$ and $\mu$ is supported by the extreme points of $X$.*

Before we begin the proof, it is necessary to establish some definitions and lemmata.

Let $C$ be convex.

**Definition 4.2.** A function $f : C \to \mathbb{R}$ is said to be *affine* if

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y)$$

for each $x, y \in C$ and $0 \leq t \leq 1$.

**Definition 4.3.** A function $f : C \to \mathbb{R}$ is said to be *convex* if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

for each $x, y \in C$ and $0 \leq t \leq 1$. $f$ is said to be concave if $-f$ is convex.

**Definition 4.4.** A function $f : C \to \mathbb{R}$ is said to be *strictly convex* if $f$ is convex and

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

for $x \neq y \in C$ and $0 < t < 1$.

**Definition 4.5.** A function $f : X \to \mathbb{R}$ is said to be upper semi-continuous if $f^{-1}((-\infty, r))$ is an open set for all real $r$. $f$ is lower semi-continuous if $-f$ is upper semi-continuous.

Denote by $A$ the set of affine continuous functions on the space $X$ and note it is a subspace of $C(X)$. Also note that it contains the constant functions, and all functions of the form $x \mapsto f(x) + r$ for every $f \in C(X)$.

It is useful first to establish the following lemma about upper semi continuous functions.

**Lemma 4.6.** *If $f$ is upper semi-continuous, then $f$ is Borel measurable and the set*

$$F = \{(x, r) \in X \times \mathbb{R} : f(x) \geq r\}$$

*is closed*

*Proof.* The first assertion will be proved as follows. Let $A = \{x \in X : f(x) \geq a\}$. Then $A = \bigcap_{n=1}^{\infty} \{x \in X : f(x) > a - \frac{1}{n}\}$. These are all open sets as $f$ is upper semi-continuous, and so $A$ is open, and thus $f$ is (Borel) measurable.

Now the second assertion is shown. Suppose there is a sequence $\{(x_n, r_n)\}_n$ in $F$ and $(x_n, r_n) \to (x, r)$, and suppose for contradiction that $(x, r) \notin F$. Then there is some $\delta > 0$ such that $f(x) < x - \delta$. But we know $f^{-1}((-\infty, r - \delta))$ open. So $x \in f^{-1}((-\infty, r - \delta))$. But then $F$ is endowed with the product topology and so

$$f^{-1}((-\infty, r - \delta))$$

contains the tail of $\{x_n\}$ and $(-\infty, r - \delta)$ contains the tail of $\{r_n\}$. So there exists some $N$ such that $f(x_N) < r - \delta$ and $r_N > r - \delta$. But then we get a contradiction from $(x_N, r_N) \in F$. ■

We now define an important special function.

**Definition 4.7.** If $f$ is a bounded function on $X$ and $x \in X$, then $\bar{f}$ is called the upper envelope of $f$ and is defined as

$$\bar{f} = \inf\{h(x) : h \in A, h \geq f\}$$

The following lemma about properties of $\bar{f}$ is now established.

**Lemma 4.8.** *Let $f$ be bounded on a locally convex, compact $X$. Then*

1. *$\bar{f}$ is concave, bounded and upper semi-continuous*
2. *$f \leq \bar{f}$ and if $f$ is concave and upper semi continuous then $f = \bar{f}$*
3. *if $g$ is also bounded then $\overline{f+g} \leq \bar{f} + \bar{g}$ and $|\bar{f} - \bar{g}| \leq \|f - g\|$ while $\overline{f+g} = \bar{f} + \bar{g}$ if $g \in A$. If $r > 0$ then $\overline{rf} = r\bar{f}$.*

*Proof.* Several statements in the lemma are obvious from the definition, so only the necessary ones are proven. First, 1. is proven. Note that for $0 < t < 1$ the set

$$\{th(x)+(1-t)h(y) : h \in A,\ h \geq f\} \subseteq \{th_1(x)+(1-t)h_2(y) : h_1, h_2 \in A,\ h_1, h_2 \geq f\}$$

And so we have that $\bar{f}(tx+(1-t)y) \geq \inf\{th_1(x)+(1-t)h_2(y) : h_1, h_2 \in A,\ h_1, h_2 \geq f\}$. But then this is the same as $t\bar{f}(x) + (1-t)\bar{f}(y)$, so $-\bar{f}$ is convex, and $\bar{f}$ is concave. Now we may show it is bounded. As $f$ is bounded we have $\bar{f} \leq M$ for some $M > 0$ as constant functions are affine, and similarly from below, so $\bar{f}$ is bounded.

Let $r \in \mathbb{R}$ and $y \in \{x \in X : \bar{f}(x) > r\}$. Then there is some $h \in A$ such that $h \geq f$ and $\bar{f}(y) < h(y) < r$. Now, $y \in \{x \in X : h(x) > r\}$ as it is open ($h$ is affine continuous) so for any such $x$, $\bar{f}(x) \leq h(x) < r$. So $\{x \in X : h(x) > r\} \subseteq \{x \in X : \bar{f}(x) > r\}$, making $\bar{f}$ upper semi-continuous.

Now, suppose $f$ is concave and upper semi-continuous. Then by 4.6 $F$ is closed and convex. Suppose $f(x_1) < \bar{f}(x_1)$ at some point, then by Hahn-Banach separation, there is a continuous linear functional $\Lambda$ such that there exists $\lambda$ s.t.

$$\sup \Lambda(X) < \lambda < \Lambda(x_1, \bar{f}(x_1))$$

so, as $\Lambda(x_1, f(x_1)) < \Lambda(x_1, \bar{f}(x_1))$, we have $\Lambda(0,1) > 0$. So the function $h$ such that $h(x) = r$ if $\Lambda(x,r) = \lambda$ exists and $h \in A$. And also, $f < h$, but $h(x_1) < \bar{f}(x_1)$ which is a contradiction. Thus, there is no such point $x_1$ and $f = \bar{f}$.

Now let us show 3.

$$\begin{aligned}(\bar{f}+\bar{g})(x) &= \inf\{h_1(x) \mid h_1 \geq f,\ h_1 \in A\} \\ &\quad + \inf\{h_2(x) : h_2 \geq g,\ h_2 \in A\} \\ &= \inf\{(h_1+h_2)(x) : h_1 \geq f,\ h_2 \geq g,\ h_1, h_2 \in A\}\end{aligned}$$

So as $h_1 + h_2 > f + g$, $\overline{f+g}$ is the inf of a bigger set and so $\overline{f+g} \leq \bar{f} + \bar{g}$. So also

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

As constant functions are affine, $f \leq \|f\|$ so $\bar{f} \leq \|f\|$. So $\bar{f} - \bar{g} \leq \overline{f-g} \leq \|f-g\|$. We may do the same by interchanging $f$ and $g$, to obtain $|\bar{f} - \bar{g}| \leq \|f-g\|$.

Now consider the case $g \in A$. Then $g = \bar{g}$ and by the previous part, $\overline{f+g} \leq \bar{f} + g$, If $h \in A$ and $h \geq f+g$, then $h - g \geq f$ so $h - g \geq \bar{f}$. Thus $\bar{f}(x) \leq \inf\{h(x) - g(x) : h \in A,\ h \geq f+g\} = \inf\{h(x) : h \in A,\ h \geq f+g\} - g(x) = \overline{f+g}(x) - g(x)$, so $\bar{f} + g = \bar{f} + \bar{g} \leq \overline{f+g}$, and we already had the other inequality, so there is equality.

The final assertion is thus shown. Let $r > 0$. For any $h \in A$, $h \geq rf \iff \frac{h}{r} \geq f$, so $\bar{f} = \frac{\overline{rf}}{r}$.

■

Now we are finally ready to prove the main theorem.

*Proof to Choquet's theorem 4.1.* $X$ is metrisable so $C(X)$ is separable, and so $A$ is separable. So choose a sequence $\{h_n\} \subseteq A$ such that $\|h_n\| = 1$ and $\{h_n\}_{n=1}^{\infty}$ is dense in $A_1$. And so $\{h_n\}_{n=1}^{\infty}$ separates points as well, as if $h_n(x_1) = h_n(x_2)$, then for any $f \in A_1$ there is a subsequence of $h_n$ converging to it. Which implies $f(x_1) = f(x_2)$.

Let $f = \sum_{n=1}^{\infty} 2^{-n} h_n^2$. This clearly exists uniformly since $C(X)$ is a complete space and the Cauchy criterion. So $f \in C(X)$ and we wish to show it is strictly convex. If $x \neq y$ then $h_n(x) \neq h_n(y)$ for some $n$, so $h_n$ is nonconstant on $[x, y]$. Now, $\alpha \mapsto \alpha^2$ is a strictly convex function, and for $0 < t < 1$,

$$th_n^2(x) + (1-t)h_n^2(y) > (th_n(x) + (1-t)y)^2 = h_n^2(tx + (1-t)y)$$

thus, there is some $\delta > 0$ such that

$$2^{-n} h_n^2(tx + (1-t)y) < 2^{-n}\left(th_n^2(x) + (1-t)h_n^2(y)\right) - \delta$$

For $m \neq n$, $h_n^2$ is convex. So let $m > n$. Then

$$\begin{aligned} &\sum_{i=1}^{m} 2^{-i} h_i^2(tx + (1-t)y) \\ <&\sum_{i=1}^{m} 2^{-i}\left(th_i^2(x) + (1-t)h_i^2(y)\right) - \delta \\ =&t\sum_{i=1}^{m} 2^{-i} h_i^2(x) + (1-t)\sum_{i=1}^{m} h_i^2(y) - \delta. \end{aligned}$$

Now let $m \to \infty$, then from this we get that $f$ is strictly convex.

Let $B = A + \mathbb{R}f$, that is a subspace of $C(X)$ generated by $A$ and $f$. Let $p$ be a functional defined on $C(X)$as $p(g) = \bar{g}(x_0)$. This is subadditive as

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

And also, from 3. of 4.8, $p$ satisfies $p(rg) = rp(g)$ if $r \geq 0$. For $h \in A$ and $r \in \mathbb{R}$, define the functional on $B$, $\phi(h + rf) = h(x_0) + r\bar{f}(x_0)$. We wish to show that $\phi(h + rf) \leq \overline{h + rf}(x_0)$ for every $h$. $r$.

Let $r \geq 0$. Then

$$\begin{aligned}\phi(h + rf) &= (h + r\bar{f})(x_0) \\ &= \overline{h + rf}(x_0) = p(h + rf)\end{aligned}$$

by 3. of 4.8. $f$ is convex so $rf$ is is concave when $r < 0$ and $\overline{rf} = rf$. Now, using 2. and 3. both, and $rf(x_0) \geq r\bar{f}(x_0)$ when $r < 0$,

$$\begin{aligned}\phi(h + rf) &= h(x_0) + r\bar{f}(x_0) \leq h(x_0) + rf(x_0) \\ &= h(x_0) + \overline{rf}(x_0) = \overline{h + rf}(x_0) = p(h + rf)\end{aligned}$$

Thus, we have shown $\phi$ is dominated by $p$.

By the Hahn-Banach extension theorem, we may extend this to a linear functional $m \in C(X)^*$ such that $m(g) \leq \bar{g}(x_0)$ for $g \in C(X)$ and $m(h+rf) = \phi(h + rf)$ when restricted to $B$.

If $g \in C(X)$ and $g \leq 0$, then $0 \geq \bar{g}(x_0) = p(g) \geq m(g)$, as $0 \in A$. Thus, this means that $m$ is nonpositive on nonpositive functions, and so it is a positive continuous linear functional.

Applying the Riesz representation theorem, there is a positive regular Borel measure $\mu$ on $X$ such that $m(g) = \int_X g \, d\mu$ for $g \in C(X)$. Since $1 \in A$,

$$1 = \bar{1}(x_0) = m(1) = \int_X 1 \, d\mu = \mu(X)$$

and so $\mu$ is a probability measure. And,

$$\int_X f \, d\mu = m(f) = \bar{f}(x_0).$$

$f \leq \bar{f}$, so then clearly $\int_X f \, d\mu \leq \int_X \bar{f} \, d\mu$. So $\mu$ represents $x_0$. We must show it is supported on the extreme points.

But if $h \in A$ and $h \geq f$, then $h \geq \bar{f}$ and so

$$h(x_0) = m(h) = \int_X h \, d\mu \geq \int_X \bar{f} \, d\mu$$

By definition of $\bar{f}$, $\bar{f}(x_0) \geq \int_X \bar{f}\,\mathrm{d}\mu$, and so

$$\int_X f\,\mathrm{d}\mu = \int_X \bar{f}\,\mathrm{d}\mu \tag{1}$$

Let $E = \{x : f(x) = \bar{f}(x)\}$, i.e. those points where the function coincides with the upper envelope. Then $E$ is Borel. Then, we have $\mu(E^c) = 0$, i.e., the function $\bar{f} - f \geq 0$ is 0 a.e.

Let $x = \frac{y+z}{2}$, where $y \neq z$. Then as $f$ is strictly convex, we have

$$f(x) < \frac{1}{2}y + \frac{1}{2}z \leq \frac{\bar{f}(y) + \bar{f}(z)}{2} \leq \bar{f}(x)$$

and so this is contained in the set of extreme points, so $\mu$ is supported by the extreme points. In fact $E$ is precisely the set of extreme points. This completes the proof.

■

# 5 Applications

In this section an application from [2] is shown, which uses this representation in the field of financial mathematics.

**Definition 5.1.** A topological vector space $X$ is said to be a Polish space if $X$ is separable, metrisable, and complete as a metric space.

Denote by $C_b(X)$ the space of real continuous bounded functions on $X$. Then we define the following.

**Definition 5.2.** Let $X$ be a Polish space, and $u : C_b(X) \to \mathbb{R}$ be a function. We call $u$ a concave utility monetary function if

1. $u$ is a concave function
2. for every $f \in C_b(X)$ such that $f$ is nonnegative, we have $u(f)$ to be nonnegative
3. $u$ is monetary, i.e. for every $a \in \mathbb{R}$, there is $u(f + a) = u(f) + a$ for every $f \in C_b(X)$
4. $u$ is positively homogeneous, i.e., for every $f \in C_b(X)$ and $\lambda \geq 0$, we have $u(\lambda f) = \lambda u(f)$
5. $u$ satisfies the **Fatou property**, i.e., for every sequence of functions $\{f_n\}_{n\geq 1} \subseteq C_b(X)$ such that $f_n(x) \to f(x)$ for every $x \in X$, and $f_n$ is decreasing, we have $u(f_n) \to u(f)$ and $u(f_n)$ is decreasing, for $f \in C_b(X)$.

To generalise it to $C(X)$, the space of real continous functions on $X$, which may not necessarily be Banach, the following is true.

The topology endowed on the space is of uniform convergence on compact sets. Then, we have

**Theorem 5.3.** *Let $u : C(X) \to \mathbb{R}$ be a function. Suppose that*

1. *$u$ is a concave function*
2. *$u$ is positively homogeneous*
3. *$u$ is a monetary function*
4. *$u$ is monotone, i.e., $f \leq g \implies u(f) \leq u(g)$*

*then $u$ is Fatou.*

Using this, we get to the main result.

**Theorem 5.4.** *Suppose $X$ is a Polish space and $u : C(X) \to \mathbb{R}$ is a function such that*

1. *$u$ is a concave function*
2. *$u$ is positively homogeneous*
3. *$u$ is a monetary function*
4. *$u$ is monotone*

*Then there exists a compact set $L \subseteq X$, and a convex set $\mathscr{S} \subseteq X$ of probability measures (positive regular Borel measures of total mass $1$) with support contained in $L$, i.e. $\mu \in \mathscr{S} \implies \mu(L) = 1$ such that for every $f \in C(X)$,*

$$u(f) = \inf\{\mu(f) : \mu \in \mathscr{S}\}$$

This is an application of the representation we did earlier. For a complete proof of these statements refer to [2].

# Bibliography


[1] Gustave Choquet. *Lectures on Analysis: Representation Theory*. Vol. 2. W.A Benjamin, 1969.

[2] Freddy Delbaen. “Monetary Utility Functions On $C_b(X)$ Spaces”. In: *International Journal of Theoretical and Applied Finance* (2024), p. 2350033.

[3] Robert R Phelps. *Lectures on Choquet’s theorem*. Springer, 2002.

[4] W. Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1991. ISBN: 9780070542365.

[5] W. Rudin. *Real and Complex Analysis*. Higher Mathematics Series. McGraw-Hill, 1974. ISBN: 9780070542334.