A study on Choquet's Theorems and their applications

Teena Thomas

A Thesis submitted to
Indian Institute of Technology Hyderabad
In Partial Fulfillment of the Requirement for
The degree of Master of Science

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Under the guidance of
Dr. Tanmoy Paul
Department of Mathematics

May 2019

DECLARATION

This thesis entitled **"A study on Choquet's Theorems and their applications"** submitted by me to the Indian Institute of Technology, Hyderabad for the award of the degree of Master of Science in Mathematics contains a literature survey of the work done by some authors in this area. The work presented in this thesis has been carried out under the supervision of **Dr. Tanmoy Paul**, Department of Mathematics, Indian Institute of Technology, Hyderabad, Telangana.

I hereby declare that, to the best of my knowledge, the work included in this thesis has been taken from the books mentioned in the Bibliography. No new results have been created in this thesis. I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

Teena 22/05/19

Teena Thomas

MA17MSCST11014

Roll No.

Approval Sheet

This Thesis entitled **"A study on Choquet's Theorems and their applications"** by **Teena Thomas** is approved for the degree of Master of Science (Mathematics) from IIT Hyderabad.

22/5/2019

Dr. Tanmoy Paul
Adviser
Dept. of Mathematics
IITH

Acknowledgement

Firstly, I would like to express my sincere gratitude to my advisor *Dr. Tanmoy Paul, Assistant Professor, Department of Mathematics, IIT Hyderabad*, for his continuous support, patience and encouragement for my project study. His motivation, dedication and expert guidance helped me to understand the concepts involved in this theory really well. Besides my advisor, I would like to thank the teachers of *Department of Mathematics, IIT Hyderabad* for imparting their knowledge and understanding of Mathematics to me. Due to their efforts and encouragement, I had a wonderful learning experience throughout my M.Sc. programme. I also take this opportunity to thank my family and friends for supporting and encouraging me throughout this programme. Above all, I thank the *Almighty* for all the blessings He has showered upon me.

Teena Thomas

Abstract

This project is a literature survey of various theorems and their applications in *Choquet theory.* For a compact convex subset D of a locally convex topological vector space E, each point $x \in D$ is a barycentre of a maximal probability measure on D. This is, in fact, a generalized version of *Minkowski's Theorem* for finite dimensional spaces. This measure exists uniquely if the compact convex set is a *simplex.* If the compact convex set is metrizable then the above measure is supported by the $ext(D)$ but very few information is available if the set is non-metrizable. Measures supported by the extreme points are the maximal measures. For a non-metrizable compact convex set the set of all extreme points may not be of Borel category, hence for such cases, the support of maximal measure can have a non-empty intersection with a Borel set disjoint from the extreme boundary. The *Choquet-Bishop-De Leeuw Theorem*, hence, states that - For an arbitrary locally convex topological vector space, each point of a compact convex subset is represented by a maximal probability measure which gives zero value to all Baire sets disjoint from the extreme points.

Further, we study the analysis of function spaces, namely, $C(K)$ in the context of *Choquet boundary.* If M is a uniform algebra of continuous functions over a compact Hausdorff space K then the state space of M is defined; it is a w^*-compact convex subset of M^*. The extreme points of the state space are precisely the point evaluation functionals. This motivates to define the *Choquet boundary* of M, as a subset of K. *Choquet boundary* is a *boundary* and its closure is the smallest closed boundary for M, called the *Šilov boundary.* Here we also study the notion of peak point and the result that when K is metrizable then the set of all peak points is dense in the *Choquet Boundary.* As an application of this notion, we discuss the well- known result by *Šaškin*, which states that for a *Korovkin* subspace of $C(K)$ the *Choquet boundary* is the whole K and also vice versa.

Contents

Dedicated to my parents

Chapter 0

List of Notations

$\mathbb{N}$	Represents set of natural numbers.
$\mathbb{Z}$	Represents set of integers.
$\mathbb{R}$	Represents set of real numbers.
∂D	Boundary of D, if D is a compact convex subset of a lctvs E.
$\mathbb{F}$	The underlying field ($\mathbb{C}$ or $\mathbb{R}$).
$a \wedge b$	$\min\{a, b\}$.
$a \vee b$	$\max\{a, b\}$.
$\text{conv}(K)$	The convex hull of set K.
$S(\mu)$	The support of the measure μ.
δ_t	the Dirac delta measure on K.
$ext(D)$	The set of all extreme points of a compact convex D.
$C_{\mathbb{R}}(K)$	The set of real valued continuous functions on a compact Hausdorff space K.
$C_{\mathbb{C}}(K)$	The set of complex valued continuous functions on a compact Hausdorff space K.
$A_{\mathbb{R}}(D)$	For a compact convex subset D of a lctvs E, represents the set of all real valued affine functions on D.
$S(D)$	For a compact convex subset D of a lctvs E, $S(D)$ represents the set of all continuous convex functions on D.
$M(K)$	Set of all regular Borel measures on K for a compact Hausdorff set K.

$M^+(K)$	Set of all non-negative regular Borel measures on K for a compact Hausdorff set K.
$K(M)$	The State space of M; the continuous linear functionals on M with unit norm and at 1 its value 1.
$\mathscr{P}(K)$	For a compact Hausdorff set K, represents the set of all probability measures on K. By a *probability measure* we mean a positive measure with total variation norm 1.
$\widehat{f}(x)$	$\inf\{g(x) : g \in -S(D), g \geqslant f\}$, the upper envelope of a bounded f on D. $\hat{f}(x) \geqslant f(x)$ for $x \in D$ and is concave on D.
$\check{f}(x)$	$\sup\{g(x) : g \in S(D), g \leqslant f\}$, the lower envelope of a bounded f on D. $\check{f}(x) \leqslant f(x)$ for $x \in D$ and is convex on D.

Chapter 1

Introduction

The importance of convexity in functional analysis has long been realized, but a comprehensive theory of infinite-dimensional convex sets has hardly existed for more than a decade. In fact, the integral representation theorems of Choquet-Bishop-de Leeuw together with the uniqueness theorem of Choquet inaugurated a new epoch in infinite-dimensional convexity. Initially considered curious and technically difficult, these theorems attracted many mathematicians, and the proofs were gradually simplified and fitted into a general theory. Today Choquet Theory provides a unified approach to study integral representations and its applications in the fields like potential theory, probability, function algebras, operator theory, group representations, and ergodic theory. At the same time, the new concepts and results have made it possible and relevant to ask new questions within the abstract theory itself. Such questions pertain to the interplay between compact convex sets D and their associated spaces, $A_{\mathbb{R}}(D)$, of continuous affine functions; to the duality between faces of D and appropriate ideals of $A_{\mathbb{R}}(D)$; to dominated extension problems for continuous affine functions on faces; and to direct convex sum decomposition into faces, as well as to integral formulas generalizing such decompositions. These problems are of geometric interest in their own right, but they are primarily suggested by applications, in particular to operator theory and function algebras.

In its geometrical form, the Choquet representation theorem can be viewed as an infinite-dimensional generalization of a classical theorem of Minkowski concerning finite dimensional compact convex sets. Indeed, suppose that D is a compact convex subset of a locally convex Hausdorff real topological vector space E. If E is assumed to be finite dimensional, then the Minkowski's theorem asserts that each point x in D is a convex combination (or barycenter) of some finite set of extreme points; that is, there exist positive real numbers $a_1, a_2, \ldots, a_n$ and points $x_1, x_2, \ldots, x_n$ in $ext(D)$ such that $\sum_{i=1}^{n} a_i = 1$ and $x = \sum_{i=1}^{n} a_i x_i$. Furthermore, each point of D admits only one such representation if and only if D is a simplex. If E is assumed to be infinite dimensional then Minkowski's theorem fails. However, the *Krein-Milman Theorem* rescues us and shows that such convex combinations of extreme points are dense in D. If, in addition, D is metrizable then the Choquet's theorem applies and asserts that each point of D is a *barycentre* of a Borel probability measure supported on $ext(D)$. A curious reader must raise voice "Is this measure unique?"

We define Choquet simplex in its classical form by means of the following. Here by $\widetilde{D}$ we denote the convex cone generated by D.

Definition 1.0.1. Let D be a compact convex subset of a lctvs E, lying on a hyperplane H not containing the origin. D is said to be a simplex if the subspace generated by the cone $\widetilde{D}$ viz. $\widetilde{D} - \widetilde{D}$ is a lattice.

Theorem 1.0.2 (Choquet-Meyer)**.** Let D be a compact convex subset of a locally convex space E. D is a simplex if and only if each point of it is a barycentre of a unique maximal Borel probability measure.

Maximality of a measure depends on how closely the measure is supported on the set of extreme points of D, although support of a maximal measure on D is not necessarily contained in the set of extreme points of D. In finite dimensional case Choquet simplexes are precisely the n-simplexes (convex combination of n extreme points) in a space of dimension at most $n - 1$. Most common example of a simplex in an infinite dimensional space is $\mathscr{P}(K)$, where K is a compact Hausdorff space. The list includes the State space (say $K(M)$) of an uniformly closed subspace M of $C_{\mathbb{R}}(K)$ which separates points and containing 1 provided $K(M)$ satisfies a *Uniqueness condition*: For $L \in K(M)$ there exists unique $\mu \in \mathscr{P}(K)$, such that $\mu \circ \phi^{-1}$ is a boundary measure on $K(M)$ and $L(g) = \int_K g(x) d\mu(x)$, $\forall g \in M$. Here $\phi : K \to M^*$ be defined by $\phi(t)(f) = f(t)$.

It is relevant to mention here that for the case of complex scalars the *Uniqueness condition* may not be ensured by this property(Simplex) of the state space although a sophisticated geometry takes place in this case; M satisfies Uniqueness condition if and only if the dual closed unit ball of M^* is *Simplexoid*, i.e. every proper face of it is a Simplex.

The goal of this project is to study the Choquet's Theorems and its applications in Functional Analysis in particular to the theory of convex sets in Banach spaces.

We now give a chapter-wise summary of this thesis. This is a literature survey and most of the results are quoted from [1, 2, 9, 10]. [7, 11, 14] are some good references for the theory of Convex sets in finite dimensional spaces, the first part of Chapter 2 is motivated from these monographs. All our notations are standard and common in the literature. A list of common symbols is given in the Chapter 0.

In Chapter 2 we study *Choquet Integral Representation Theorem* (Theorem 2.2.2). It states that corresponding to each probability measure supported on the extreme points of a compact convex subset there exists a unique point of the compact convex set, called the *barycentre or resultant of the measure*. This association is w^*-continuous. It is rather a deep fact and the central theme of this project is for any point of a compact convex set there exists a measure supported on the extreme points of the set. In the same Chapter, we study the *Completely monotone function*. A representation theorem for the class of bounded completely monotone function is also discussed.

In Chapter 3 the main result is Choquet-Bishop-De Leeuw Theorem(Theorem 3.2.1) for a compact convex metrizable set. It is, in fact, converse of Theorem 2.2.2 when the compact convex set is metrizable. In this Chapter we introduce the space $A_{\mathbb{R}}(D)$ where D is a compact convex subset of a lctvs E. For a bounded function f, we introduce the upper envelope and lower envelope of f. They are concave and convex functions respectively. Several properties of these functions are also discussed.

Chapter 4 is devoted to discussing the Choquet-Bishop-De Leeuw Theorem for the non-metrizable case. The tools developed in the literature have impacts on many other aspects also. This chapter starts with a generalization of Stone-Weierstrass Theorem which asserts a sufficient condition which makes a subspace of $C_{\mathbb{R}}(K)$ uniformly dense in it (Lemma 4.0.1). $M^+(D)$ turns out to be a partially ordered set with respect to the partial ordering defined in Pg.35. A judicious Zornification ensures that each element of $M^+(D)$ is dominated by a maximal measure in it(Lemma 4.1.4). The maximality of a measure discussed at the beginning is analogous to the maximality defined here. In Theorem 4.1.13, it is observed that a maximal measure μ in $M^+(D)$ does not distinguish f and $\hat{f}$ for $f \in C_{\mathbb{R}}(D)$ (Mokobodzki). On the other hand, a point $x \in D$ is an extreme point if and only if $\hat{f}(x) = f(x)$ for $f \in C_{\mathbb{R}}(D)$, see Theorem 4.1.12. This clarifies $S(\mu)$ must be in a small neighborhood of $ext(D)$. The notion of *boundary measure* is introduced for a signed measure. Finally, the most generalized version of Choquet-Bishop-De Leeuw is obtained in Theorem 4.1.16.

Theorem 1.0.3. Let D be a compact convex subset of a lctvs E. Then for each $x \in D$ there exists a measure $\mu \in \mathscr{P}(D)$ such that $x = r(\mu)$ and $\mu(F) = 0$ for any Baire set $F \subseteq D \backslash ext(D)$.

Chapter 4 ends with some applications of these results: The *Rainwater Theorem* (Theorem 4.2.1) is proved, it is shown that in a compact convex set D a probability measure μ is a w^*-limit of a net of discrete probability measure with the same resultant (see proposition 4.2.5) as that for μ. A new proof of Stone-Weierstrass Theorem is also given.

In Chapter 5 we discuss about a special type of compact convex subset of a lctvs E, called *Simplex*(Definition 1.0.1). Various characterizations are discussed when a compact convex set is a Simplex. The most geometric version of Definition 1.0.1 is stated in Theorem 1.0.2. Finally, it is observed that if for a compact convex X, $\dim\left(\tilde{X}-\tilde{X}\right)=n$ then X is a Simplex if and only if X has only n extreme points. Here X is assumed to be a subset of a hyperplane in E not containing the origin.

Let M be a subalgebra of $C_{\mathbb{C}}(K)$ containing constants and separates points. The weak* compact convex $K(M)=\{L\in B_{M^*}:L(1)=1\}$ is called the *State space of* M, an obvious generalization of $\mathscr{P}(K)$. From Milman's converse of Krein-Milman Theorem it follows $ext(K(M))\subseteq\phi(K)$ and this leads to define $B(M)=\{t\in K:\phi(t)\in ext(S(M))\}$. It is a *boundary* for M called the *Choquet boundary*, the smallest boundary contained in any closed boundary. The aim of Chapter 6 is to discuss the analysis of the function spaces viz. $C(K)$ and its uniform algebras in the context of Choquet boundary. Choquet boundary plays a crucial role in the analysis of continuous function spaces; it carries information of the space, a glimpse of it may be found in Theorem 1.0.3 if D is replaced by the State space of a subspace. $\overline{B(M)}$ is the so-called *Silov boundary* for M. The interplay between *peak set* and Choquet boundary are discussed and are explained with various examples. The celebrated Korovkin Theorem is discussed with its full generality in the context of the Choquet boundary.

Four appendixes at the end contain some basic prerequisites which make this Thesis self contained and are also important in their own right.

This project initiates to create curiosity in the fields of *infinite dimensional convexity*, *function spaces of continuous functions* over various domains and other allied areas which have more or fewer connections with the Choquet Theory. We encounter numerous results related to these fields (in Chapter 4, 6) which can be derived as consequences of Choquet's Theorems. In the last few decades, the territory of Choquet's Integral Representation has reached to the subjects like Mathematical Economics, Risk Management, Potential Theory, Game Theory, Operations Research, etc. Interested readers can come across the articles [4–6] to get an overview of the vastness of its applicabilities.

Chapter 2

Choquet Integral representation theorem

We begin this Chapter with two fundamental theorems in Functional analysis, viz. Hahn-Banach theorem and Hahn-Banach separation theorem.

Theorem 2.0.1. (Hahn-Banach theorem) Let M be a proper subspace of a lctvs E and p be a sublinear functional on E. Let f be a linear functional on M such that $Re\ f(x) \leqslant p(x)$ for all $x \in M$. Then f can be extended to whole E with $Re\ f(x) \leqslant p(x)$ for all $x \in E$.

Theorem 2.0.2. (Hahn-Banach separation theorem) Let M be a convex subset of a lctvs E and $x_0 \in E \backslash M$. Then there exists a linear functional $f \in E^*$ and $\lambda \in \mathbb{R}$ such that $Re\ f(x_0) \geqslant \lambda \geqslant \sup_{x \in M} Re\ f(x)$.

A weakening or strengthening of the above separation depends on the topological properties of M. For example if M, N be two disjoint compact convex sets then a strict separation is possible, on the other hand, if M is compact convex and x_0 is an extreme point of M then only above separation is possible.

2.1 Carathéodory's theorem

Proposition 2.1.1. Let K be a non-empty compact convex subset of a lctvs X. Suppose Φ is a real-valued convex and usc funtion on K. Then, Φ attains its supremum on K.

Proof. It is clear that $\sup_{x \in K} \Phi(x) < \infty$. In fact, if $V_\alpha = \{x \in K : \Phi(x) < \alpha\}$ then $(V_\alpha)_{\alpha \in \mathbb{R}}$ is an open cover of K, which in turn, admits a finite subcover for K and hence the result.

Choose $(x_\alpha) \subseteq K$ such that $\Phi(x_\alpha) \to \sup_{x \in K} \Phi(x)$. Since K is compact, there exists a subnet (x_{α_i}) of (x_α) such that for some $x_0 \in K$,

$$x_{\alpha_i} \to x_0.$$

Now,

$$\Phi(x_{\alpha_i}) \to \sup_{x \in K} \Phi(x).$$

Thus,

$$\sup_{x \in K} \Phi(x) \geqslant \Phi(x_0) \geqslant \limsup_i \Phi(x_{\alpha_i}) = \sup_{x \in K} \Phi(x).$$

This implies that $\Phi(x_0) = \sup_{x \in K} \Phi(x)$. Thus, Φ attains its supremum on K. □

Let us recall upper semi continuous function in Chapter 0. Here are few remarks which can be derived directly from the definition.

Proposition 2.1.2. (*a*) Let f be a real valued function of a topological space (X, τ). f is usc function if and only if for any net $x_\alpha \to x_0 \Rightarrow f(x_0) \geqslant \limsup_\alpha f(x_\alpha)$.

(*b*) Let f be a real valued function of a topological space (X, τ). f is lsc function if and only if for any net $x_\alpha \to x_0 \Rightarrow f(x_0) \leqslant \liminf_\alpha f(x_\alpha)$.

(*c*) Let $(X, \|.\|)$ be any normed linear space. We can define the weak* topology on X^* (refer Appendix B). If $\|.\|$ is the induced norm on X^*, then $\|.\| : (X^*, w^*) \longrightarrow \mathbb{R}_{\geqslant 0}$ is a lsc function:

Proof. (*a*) In fact if $x_\alpha \to x_0$ and $\epsilon > 0$ then by the definition of usc function, $x_0 \in \{x \in X : f(x) < f(x_0) + \epsilon\}$ is open in (X, τ). Therefore, by definition of convergence of net, for infinitely many α,

$$f(x_\alpha) < f(x_0) + \epsilon.$$

This implies, for all $\epsilon > 0$

$$\limsup_\alpha f(x_\alpha) \leqslant f(x_0) + \epsilon.$$

This implies,

$$\limsup_\alpha f(x_\alpha) \leqslant f(x_0).$$

Conversely, let $\lambda \in \mathbb{R}$ and $A = \{x \in X : f(x) \geqslant \lambda\}$. We, now, show that A is closed. Let $(x_\alpha) \subseteq A$ such that $x_\alpha \to x_0$. This implies, for all α,

$$f(x_\alpha) \geqslant \lambda.$$

From hypothesis,

$$f(x_0) \geqslant \limsup_{\alpha} f(x_\alpha) \geqslant \lambda.$$

This implies, $x_0 \in A$. Hence, A is closed. This implies, $X \backslash A$ is open. Therefore, f is usc function.

(b) A similar proof like (a) .

(c) By the Banach Alaoglu Theorem $B.1.1$, $\{x^* \in X^* : ||x^*|| \leqslant 1\}$ is w^*-compact in X^*. Hence, $\{x^* \in X^* : \|x^*\| \leqslant 1\}$ is w^*-closed in X^*. This implies, for any $\alpha \in \mathbb{R}$,

$$\{x^* \in X^* : ||x^*|| \leqslant \alpha\} = \{y^* \in X^* : ||y^*|| \leqslant 1\} \text{ is } w^*\text{-closed in } X^*.$$

Hence, $\|.\|$ is a lsc function on X^*.

□

Theorem 2.1.3 (Bauer's Maximum Principle)**.** Let K be a non-empty compact convex subset of a lctvs X. Suppose Φ is a real-valued convex and usc funtion on K. Then, Φ attains its supremum at some extreme point of K.

Proof. Define $\mathscr{S} = \{F \subseteq X : F \neq \phi,\ F$ is closed extreme set of $K\}$. $\mathscr{S}$ is non-empty as $K \in \mathscr{S}$. Now, $\mathscr{S}$ has the following properties:

(i) If $(X_i)_{i\in I} \subseteq \mathscr{S}$ and $\bigcap_{i\in I} X_i \neq \phi$, then $\bigcap_{i\in I} X_i \in \mathscr{S}$.

(ii) For each $F \in \mathscr{S}$ and convex usc function g, let $F' = \{x \in F : g(x) = \sup_{y\in F} g(y)\}$. Then, $F' \in \mathscr{S}$.

Proof of (i): Let $(X_i)_{i\in I} \subseteq \mathscr{S}$ and $\bigcap_{i\in I} X_i \neq \phi$. Since each X_i is closed, so is $\bigcap_{i\in I} X_i$. Let $x, y \in \bigcap_{i\in I} X_i$ such that for some $\lambda \in (0,1)$, $\lambda x + (1-\lambda)y \in \bigcap_{i\in I} X_i$. This implies, for all $i \in I$, $\lambda x + (1-\lambda)y \in X_i$. Hence, for each $i \in I$, $x, y \in X_i$. Therefore, $x, y \in \bigcap_{i\in I} X_i$.

Proof of (ii): First of all, F' is non-empty, by Proposition (2.1.1). Let $x, y \in F$ such that $\lambda x + (1-\lambda)y \in F'$.

$$\begin{aligned}
\sup_{z\in F} g(z) &= g(\lambda x + (1-\lambda)y) \\
&\leqslant \lambda g(x) + (1-\lambda)g(y) \\
&\leqslant \lambda(\sup_{z\in F} g(z)) + (1-\lambda)(\sup_{z\in F} g(z)) \\
&= \sup_{z\in F} g(z).
\end{aligned} \tag{2.1}$$

This implies, $\lambda g(x) + (1-\lambda)g(y) = \sup_{z\in F} g(z)$. Hence, $g(x) = g(y) = \sup_{z\in F} g(z)$. Thus, $x, y \in F'$. Therefore, F' is an extreme set of F and we have assumed F is extreme set of K. This implies that F' is extreme set of K, since extreme sets of a given set are transitive. It can be easily seen that F' is closed subset of K. Therefore, F' is non-empty closed extreme set of K. Hence, $F' \in \mathscr{S}$.

Now, define a partial ordering $\prec$ on $\mathscr{S}$ as:

$$\text{For } F_1, F_2 \in \mathscr{S}, F_1 \prec F_2 \text{ if and only if } F_2 \subseteq F_1.$$

Thus, $(\mathscr{S}, \prec)$ is a partially ordered set. Let $(X_i)_{i\in I} \subseteq \mathscr{S}$ be any chain. Then, by Cantor's Intersection theorem and from property (i) of $\mathscr{S}$, $\phi \neq \bigcap_{i\in I} X_i \in \mathscr{S}$ and clearly, for all $i \in I, X_i \prec \bigcap_{i\in I} X_i$. Thus, $(X_i)_{i\in I}$ has an upper bound. Hence, every chain in $\mathscr{S}$ has an upper bound. Now, by Zorn's Lemma, $(\mathscr{S}, \prec)$ has a maximal element say $F \in \mathscr{S}$.

Claim: If $F \in \mathscr{S}$ is a maximal element of $(\mathscr{S}, \prec)$, then F must be singleton.

Proof of the Claim: Suppose there exists $x, y \in F$ such that $x \neq y$. Then by Hahn Banach theorem, there exists $f \in X^*\backslash\{0\}$ such that $f(x) < f(y)$. Consider, $F' = \{z \in F : f(z) = \sup_{w\in F} f(w)\}$. Clearly, $F' \subsetneq F$, since $x \notin F'$. By property (ii) of $\mathscr{S}$, $F' \in \mathscr{S}$. This implies that $F \prec F'$. But this contradicts the maximality of F. Thus, F must be singleton.

Consider $F' = \{x \in K : \Phi(x) = \sup_{y\in K} \Phi(y)\}$. Then, by property (ii), $F' \in \mathscr{F}$. We can get a chain $(X_i)_{i\in I}$ in $(\mathscr{S}, \prec)$ such that $X_j = F'$, for some $j \in I$. Then, let A be a maximal element of $(\mathscr{S}, \prec)$ such that $X_i \prec A$, for all $i \in I$. In particular, $F' \prec A$, which implies $A \subseteq F'$, Now, by above claim, A is singleton say $A = \{a\}$. It follows easily that a is an extreme point of set K and hence, Φ attains its supremum at a. □

Theorem 2.1.4 (Krein Milman's Theorem). Let K be a non-empty compact convex subset of a lctvs X. Then, $ext(K) \neq \phi$ and $\overline{conv}(ext(K)) = K$.

Proof. From Bauer's Maximum Principle (2.1.3), it is proved that $ext(K) \neq \phi$. Clearly, $\overline{conv}(ext(K)) \subseteq K$. Suppose there exists $x_0 \in K\backslash\overline{conv}(\text{ext}(K)$. Then, by Hahn Banach separation theorem (2.0.2), there exists $f \in X^*\backslash\{0\}$ such that

$$f(x_0) > \sup f(\overline{conv}(\text{ext}(K)).$$

Consider, $M = \{x \in K : f(x) = \sup_{y\in K} f(y)\}$. By Bauer's Maximum Principle (2.1.3), M is non-empty, closed subset of K and has an extreme point. This implies that $\overline{conv}(\text{ext}(K)) \cap M \neq \phi$. Now,

$$\sup_{y\in K} f(y) \geqslant f(x_0) > \sup f(\overline{conv}(\text{ext}(K)).$$

This implies that $\overline{conv}(\text{ext}(K) \cap M = \phi$. This is a contradiction. Hence, $\overline{conv}(ext(K))$ $= K$. □

Theorem 2.1.5 (Carathéodory's theorem). Let X be a lctvs. Let $K \subseteq X$ be compact and convex. Let dim $X = n$. Then, each point in K can be expressed as a convex combination of at most $(n+1)$ extreme points of K.

Proof. We will prove by induction on n. The result is true for $n = 1$ since for $n = 1$, X is a line and K is just a line segment. Suppose the result is true for any lctvs of dimension n. We now prove the result for a lctvs X of dimension $n+1$. Let $K \subseteq X$ be compact and convex. Let $x \in K$.

CASE (1): Let $x \in \partial K$. Then by Hahn Banach separation theorem (2.0.2), there exists a supporting hyperplane H of K in X such that $x \in H \cap K$. Since H is a hyperplane in X,

$$H = \{x \in X : f(x) = \alpha\}, \text{ for some } \alpha \in K \text{ and } f \in X^* \backslash \{0\}.$$

Now, $\dim(H \cap K) \leqslant n$, since $H \cap K \subseteq H$ and dim $X = n$. Now, $H \cap K$ is compact (since it is a closed subset of the compact set K).

CLAIM: $H \cap K$ is convex set.

PROOF OF THE CLAIM: Let $y, z \in H \cap K$ and $\lambda \in (0,1)$. Then, $\lambda y + (1-\lambda)z \in K$, since K is convex set. Since $y, z \in H \Rightarrow f(y) = \alpha$ and $f(z) = \alpha$. Hence,

$$f(\lambda y + (1-\lambda)z) = \lambda f(y) + (1-\lambda)f(z) = \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

This implies $\lambda y + (1-\lambda)z \in H$ and therefore $\lambda y + (1-\lambda)z \in H \cap K$. It follows from here that $H \cap K$ is convex. This completes the proof of the claim.

Therefore, now, by induction hypothesis, there exists $(n+1)$ extreme points $z_1, z_2, ..., z_{n+1}$ such that x can be written as convex combination of $z_1, z_2, ..., z_{n+1}$.

We now claim that $ext(H \cap K) \subseteq ext(K)$. It is sufficient to prove that $H \cap K$ is an extreme set of K and then since $ext(H \cap K) \subset ext(K)$, we arrive at our claim. Let $y, z \in K$ such that there exists $\lambda \in (0,1)$ and $\lambda y + (1-\lambda)z \in H \cap K$. This implies $\lambda y + (1-\lambda)z \in H$ and $\lambda y + (1-\lambda)z \in K$. Thus, $f(\lambda y + (1-\lambda)z) = \alpha \Rightarrow \lambda f(y) + (1-\lambda)f(z) = \alpha$. Suppose without loss of generality, $f(y) \neq \alpha$. Assume without loss of generality that K lies in the half space $\{x \in X : f(x) \leqslant \alpha\}$. Then, $f(y) < \alpha$. Now,

$$\begin{aligned} \alpha &= \lambda f(y) + (1-\lambda) f(z) \\ &\leqslant \lambda f(y) + (1-\lambda)\alpha \\ &< \lambda\alpha + (1-\lambda)\alpha \\ &= \alpha. \end{aligned} \tag{2.2}$$

which gives us a clear contradiction. Similarly, we get a contradiction if $f(z) < \alpha$. This implies $f(y) = f(z) = \alpha$, which in turn, implies that $y, z \in H$. Also, since $y, z \in K$, we get $y, z \in H \cap K$. Therefore, $H \cap K$ is an extreme set of K. Therefore, using a similar result like Remark A.1.4(b), we get that $ext(H \cap K) \subseteq ext(K)$. This implies that $z_1, z_2, ..., z_{n+1}$ are extreme points of K. Therefore, x can be written as convex combination of the extreme points of K namely $z_1, z_2, ..., z_{n+1}$. Thus, the result is true in this case.

Case 2: Let $x \in \text{int}(K)$.

By Krein Milman's theorem (Theorem 2.1.4), $ext(K) \neq \phi$ and thus, there exists $y \in ext(K)$. Consider the line joining y and x and intersecting the boundary of K at point say z. Then, there exists $(n+1)$ extreme points $z_1, z_2, ..., z_{n+1}$ such that z is convex combination of $z_1, z_2, ..., z_{n+1}$ by following the same proof as in Case(1). Finally, we get x to be the convex combination of the points $y, z_1, z_2, ..., z_{n+1}$. Hence, the result is true in the case too. Hence, the result is true for any lctvs of dimension $n+1$.

Thus, by 2^{nd} principle of mathematical induction, the result is true for any lctvs of dimension n, for any $\in \mathbb{N}$. □

2.2 Resultant of a Measure

Let X be a compact convex subset of a lctvs E. The Proposition C.2.2 in Appendix C tells us that the set of all extreme points of set of all probability measures $\mathscr{P}(X)$ in $M(X)$ is the set of all Dirac measures $\{\delta_t : t \in X\}$. Thus, by Krein Milman Theorem 2.1.4,

$$\mathscr{P}(X) = \overline{conv}^{w^*}\{\delta_t : t \in X\}.$$

Definition 2.2.1 (Resultant of a Measure)**.** Let X be a compact convex subset of a lctvs E. Let $\mu \in M(X)$. A point x is said to be a resultant (or barycentre) of μ if for any $f \in E^*$,

$$\int_X f d\mu = f(x).$$

Notation: $x = r(\mu)$.

Theorem 2.2.2. Let Y be a compact subset of a lctvs E. Suppose $\overline{conv}(Y) = X$ is also compact. Let $\mu \in \mathscr{P}(Y)$. Then, there exists a unique $x \in X$ such that $x = r(\mu)$ and the map $r : \mathscr{P}(Y) \longrightarrow X$ defined as $\mu \longmapsto r(\mu)$ is w^*-continuous.

Proof. Let $\mu \in \mathscr{P}(Y) = \overline{conv}^{w^*}\{\delta_t : t \in Y\}$. This implies there exists a net $(\mu_\alpha) \subset conv\{\delta_t : t \in Y\}$ such that $\mu_\alpha \longrightarrow \mu$. Hence, for every α,

$$\mu_\alpha = \sum_{j=1}^{k_\alpha} \beta_j \delta_{t_j^\alpha}$$

for some $t_j^\alpha \in Y$, $\beta_j \geqslant 0$ and $\sum_{j=1}^{k_\alpha} \beta_j = 1$. Now,

$$\sum_{j=1}^{k_\alpha} \beta_j t_j^\alpha \in \overline{conv}(Y) = X.$$

Let

$$x_\alpha = \sum_{j=1}^{k_\alpha} \beta_j t_j^\alpha.$$

For every $f \in E^*$, $f(x_\alpha) = \sum_{j=1}^{k_\alpha} \beta_j f(t_j^\alpha) = \int_Y f d\mu_\alpha$.

Hence, $x_\alpha = r(\mu_\alpha)$. Now, since X is compact, there exists a convergent subnet say (x_{α_i}) such that $x_{\alpha_i} \longrightarrow x$, for some $x \in X$. Let $f \in E^*$. Then,

$$\begin{aligned} f(x) &= \lim_i f(x_{\alpha_i}) \\ &= \lim_i \sum_{j=1}^{k_{\alpha_i}} \beta_j f(t_j^{\alpha_i}) \\ &= \lim_i \sum_{j=1}^{k_{\alpha_i}} \beta_j \delta_{t_j^{\alpha_i}}(f) \\ &= \lim_i \mu_{\alpha_i}(f) \\ &= \lim_i \int_Y f d\mu_{\alpha_i} \\ &= \int_Y f d\mu. \end{aligned} \tag{2.3}$$

Hence, $x = r(\mu)$. Now, suppose there exists another $y \in X$ such that $y = r(\mu)$. This implies, for any $f \in E^*$,

$$f(y) = \int_Y f d\mu = f(x).$$

This is a contradiction, since by Hahn Banach theorem 2.0.1, there exists $g \in E^*$ such that $g(x) \neq g(y)$. This implies there exists unique $x \in X$ such that $x = r(\mu)$.

We now show the map $\gamma : \mathscr{P}(Y) \longrightarrow X$ defined as $\mu \longmapsto r(\mu)$ is w^*-continuous. Let (μ_α) in $\mathscr{P}(Y)$ such that $\mu_\alpha \xrightarrow{w^*} \mu_0$. It is enough to show that any convergent subnet of $(r(\mu_\alpha))$ converges to $r(\mu_0)$. Let $r(\mu_{\alpha_i}) \longrightarrow y$ in X. Then, for any $f \in E^*$,

$$f(r(\mu_{\alpha_i})) \longrightarrow f(y)$$

which implies

$$\int_Y f d\mu_{\alpha_i} \longrightarrow f(y).$$

Now,

$$\int_Y f d\mu_{\alpha_i} \longrightarrow \int_Y f d\mu_0 = f(r(\mu_0)).$$

This implies for all $f \in E^*$, $f(y) = f(r(\mu_0))$. By Hahn Banach theorem 2.0.1, E^* separates points and hence, $y = r(\mu_0)$. Therefore, $r(\mu_{\alpha_i})$ converges to $r(\mu_0)$. Suppose $(r(\mu_\alpha))$ does not converge to $r(\mu_0)$. This implies, there exists a subnet $(r(\mu_{\alpha_i}))$ and a neighborhood U of $r(\mu_0)$ such that for all i, $r(\mu_{\alpha_i}) \notin U$. Since X is compact, there exists a convergent subnet of this subnet say $(r(\mu_{\alpha_{i_j}}))$. Hence, $r(\mu_{\alpha_{i_j}}) \longrightarrow r(\mu_0)$. But, for all j, $r(\mu_{\alpha_{i_j}}) \notin U$, which is a contradiction. This implies, $r(\mu_\alpha) \longrightarrow r(\mu_0)$. □

Theorem 2.2.3. Let Y be a compact subset of a lctvs E. Suppose $\overline{conv}(Y) = X$ is also compact. Then, $x \in X$ if and only if there exists $\mu \in \mathscr{P}(Y)$ such that $x = r(\mu)$.

Proof. The converse is true by the theorem 2.2.2. Now, assume $x \in X = \overline{conv}(Y)$. Then, there exists $(x_\alpha) \subset conv(Y)$ such that $x_\alpha \longrightarrow x$. Hence,

$$x_\alpha = \sum_{i=1}^{k} t_i^\alpha y_i^\alpha$$

where for some $k \in \mathbb{N}$ and all i, $y_i^\alpha \in Y$, $t_i^\alpha \geqslant 0$ and $\sum_{i=1}^k t_i^\alpha = 1$ Define, $\mu_\alpha = \sum_{i=1}^k t_i^\alpha \delta_{y_i^\alpha} \in \mathscr{P}(Y)$. Then, $||\mu_\alpha|| = 1$ and μ_α is a positive measure. Since $\mathscr{P}(Y)$ is w^*-compact, there exists a subnet (μ_{α_i}) in $\mathscr{P}(Y)$ and $\mu_0 \in \mathscr{P}(Y)$ such that $\mu_{\alpha_i} \xrightarrow{w^*} \mu_0$. Let $f \in E^*$. Then,

$$\begin{aligned}\int_Y f d\mu_0 &= \lim_\alpha \int_Y f d\mu_{\alpha_i} \\ &= \lim_\alpha \sum_{i=1}^k t_i^\alpha \int_Y f d\delta_{y_i^\alpha} \\ &= \lim_\alpha f(\sum_{i=1}^k t_i^\alpha y_i^\alpha) \\ &= \lim_\alpha f(x_\alpha) \\ &= f(x).\end{aligned} \tag{2.4}$$

This implies, $x = r(\mu_0)$ where $\mu_0 \in \mathscr{P}(Y)$. □

Remark 2.2.4. Let K be a compact convex subset of a lctvs X. Then, by Krein Milman theorem (2.1.4), $ext(K) \neq \phi$ and $\overline{conv}(ext(K)) = K$. Now, let $Y = \overline{ext(K)} \subset X$ is compact and also, $\overline{conv}(Y) = K$. Hence, by earlier theorem, for any $x \in X$, there exists $\mu \in \mathscr{P}(Y)$ such that $x = r(\mu)$. Thus, $S(\mu) \subset \overline{ext(K)}$.

As an immediate consequence of above Theorem we have the following.

Corollary 2.2.5. Let E be a lctvs. The following statements are equivalent.

(a) (Krein-Milman Theorem) If $X \subseteq E$ is a compact convex set and $Y = ext(X)$, then $X = \overline{conv(Y)}$.

(b) Each $x \in X$ is represented by a $\mu \in \mathscr{P}(X)$ with $S(\mu) \subseteq \overline{Y}$.

Theorem 2.2.6. (Bauer) Let K be a compact convex subset of a lctvs X. Let $x \in \text{ext}(K)$. Then, δ_x is the unique probability measure such that $x = r(\delta_x)$.

Proof. It is clear that $r(\delta_x) = x$. Now, let $\mu \in \mathscr{P}(Y)$ where $Y = \overline{ext(K)}$ such that $x = r(\mu)$. It suffices to show that $S(\mu) = \{x\}$.

Claim: For any $D \subset Y$ such that $x \notin D$ and D is compact, $\mu(D) = 0$.

Proof of the claim: Suppose for some compact $D \subset Y \backslash \{x\}$, $\mu(D) > 0$. Since $D \cap S(\mu) \neq \phi$, let $x' \in D \cap S(\mu)$ (If $S(\mu) \cap D = \phi$, then $D \subset S(\mu)^c$. Since D is compact, there exists open sets $G_1, G_2, ..., G_m$ such that for all $i = 1, 2, ..., m$, $\mu(G_i) = 0$ and $D \subset \bigcup_{i=1}^{m} G_i$. This implies $\mu(D) = 0$, which is a contradiction.) Let $F \subset Y$ be closed set such that $x' \in F$ and $x \notin F$(since X is normal space and Y is closed in X, Y is normal space). Define the following measures: For any measurable set $E \subset Y$, $\mu_1(E) = \mu(E \cap F)$ and $\mu_2(E) = \mu(E \cap Y \backslash F)$. Hence, $S(\mu_1) = F$. Also, $\mu_1 + \mu_2 = \mu$ and $1 = \|\mu\| = \|\mu_1\| + \|\mu_2\|$. By Urysohn's lemma, there exists a continuous function $f : Y \to [0, 1]$ such that $f(y) = 0$, for each $y \in F$ and $f(x) = 1$. So,

$$1 = f(x) = \int_Y f d\mu = \int_F f d\mu_1 + \int_{Y \backslash F} f d\mu_2 = \int_{Y \backslash F} f d\mu_2.$$

This implies that $\|\mu_2\| \neq 0$. Similarly, we can show that $\|\mu_1\| \neq 0$, since $S(\mu_2) = Y \backslash F$. Hence, we have $0 < \|\mu_1\| < 1$ and $0 < \|\mu_2\| < 1$. Now,

$$\mu = \|\mu_1\| \frac{\mu_1}{\|\mu_1\|} + \|\mu_2\| \frac{\mu_2}{\|\mu_2\|}.$$

Let $z_1 = r(\frac{\mu_1}{\|\mu_1\|})$ and $z_2 = r(\frac{\mu_2}{\|\mu_2\|})$. Therefore,

$$x = r(\mu) = \|\mu_1\| z_1 + \|\mu_2\| z_2$$

where $\|\mu_1\| + \|\mu_2\| = 1$ and $0 < \|\mu_1\| < 1$ and $0 < \|\mu_2\| < 1$. This implies $x \notin ext(K)$, which is a contradiction. Hence the claim.

Using above claim, $\mu(Y\backslash\{x\}) = \sup\{\mu(D) : D \subset Y\backslash\{x\} \text{ is compact}\} = 0$.

Hence, $S(\mu) = \{x\}$. This completes the proof. □

A partial converse of Krein-Milman theorem is the following.

Theorem 2.2.7. Let E be a lctvs and $X \subseteq E$ be a compact convex set. Then for any $Z \subseteq X$ with $\overline{conv}(Z) = X$ we have $ext(X) \subseteq \overline{Z}$.

Proof. If $Y = \overline{Z}$ then $X = \overline{conv}(Y)$. Let $x \in ext(X)$ then there exists $\mu \in \mathscr{P}(Y)$ such that $r(\mu) = x$. Now from Theorem 2.2.6 $\mu = \delta_x$. Hence $x \in Y$. □

2.3 An application of Krein-Milman theorem

We conclude this Chapter by an interesting application of Krein-Milman Theorem.

Definition 2.3.1 (Completely monotonic functions)**.** A real valued function f on $(0, \infty)$ is said to be completely monotonic function if it has derivatives $f^{(0)} = f$, $f^{(1)}$, $f^{(2)}$,... of all orders and if $(-1)^n f^{(n)} \geqslant 0$, for all $n \in \mathbb{N} \cup \{0\}$.

Remark 2.3.2. A completely monotonic function is non-negative and non-increasing, as is each of the functions $(-1)^n f^{(n)}$. Examples of such functions are $x^{-\alpha}, e^{-\alpha x}$, for $\alpha \geqslant 0$.

Definition 2.3.3. (Convex cone) Let V be a vector space. Let $C \subset V$. C is said to be a cone if for any positive scalar α and for any $x \in C$, $\alpha x \in C$. A cone C is called convex cone if for any positive scalars α, β and for any $x, y \in C$, $\alpha x + \beta y \in C$.

Example 2.3.4. Examples of convex cone are :

(i) for any vector space V, the empty set, the space V and any linear subspace of V.

(ii) the set of all non-negative continuous functions on $\mathbb{R}$

Sergei Natanovich Bernstein proved a fundamental representation theorem for completely monotonic functions. We will prove the theorem for bounded functions. The proof of this theorem is an application of Krein Milman theorem. Let us denote the one point compactification of $[0, \infty)$ as $[0, \infty]$.

Theorem 2.3.5 (Bernstein)**.** If f is bounded and completely monotonic function on $(0, \infty)$, then there exists a unique non-negative Borel measure μ on $[0, \infty]$ such that $\mu([0, \infty]) = f(0^+)$ and for each $x > 0$,

$$f(x) = \int_0^\infty e^{-\alpha x} d\mu(\alpha).$$

Proof. (The converse of the above theorem is true. Assume we have a function defined as above. Then, by Leibniz Integral rule, differentiation under the integral sign is possible. Hence, for any $n = 0, 1, 2, ...$, for each $x > 0$,

$$f^{(n)}(x) = (-1)^n \int_0^\infty e^{-\alpha x} d(\alpha^n \mu).$$

Therefore, since μ is non-negative Borel measure and $e^{-\alpha x} > 0$, for each $x > 0$,

$$(-1)^n f^{(n)}(x) = \int_0^\infty e^{-\alpha x} d(\alpha^n \mu) \geqslant 0.$$

This implies f is completely monotonic function on $(0, \infty)$. Now, define for all $n \in \mathbb{N}$ and $\alpha \in [0, \infty]$,

$$g_n(\alpha) = e^{-\alpha/n}.$$

Then,

$$|g_n(\alpha)| \leqslant 1 \text{ and } \lim_{n\to\infty} g_n(\alpha) = e^{-\alpha/n} = 1.$$

By dominated convergence theorem,

$$\begin{aligned} &\lim_{n\to\infty} \int_0^\infty e^{-\alpha/n} d\mu(\alpha) = \int_0^\infty 1 d\mu(\alpha) \\ \Longrightarrow &\lim_{n\to\infty} f(1/n) = \mu([0,\infty]) \\ \Longrightarrow &\lim_{x\to 0^+} f(x) = \mu([0,\infty]) \\ \Longrightarrow &f(0^+) = \mu([0,\infty]). \end{aligned} \tag{2.5}$$

Hence, $f(0^+) < \infty$. This implies f is bounded function.)

Let us begin with the sketch of the proof. Let us denote CM to be the convex cone of all completely monotonic functions f such that $f(0^+) < \infty$. ($f(0^+)$ exists always since f is completely monotonic function, although it may be infinite). Let $K = \{f \in CM : f(0^+) \leqslant 1\}$. We will prove that K is convex. Now, if $f \in CM$, $f \neq 0$, then $\frac{f}{f(0^+)} \in K$. Thus, it suffices to prove the theorem for elements of K.

Let E be the space of all real valued infinitely differentiable functions on $(0, \infty)$. Then, $K \subset E$. Now, we digress a little from the proof sketching and define a topology on E. For all $n \in \mathbb{N}$, let $K_n = [\frac{1}{n}, n]$. Then, $(0, \infty) = \bigcup_{n\in\mathbb{N}} [\frac{1}{n}, n]$. For all $f \in E$, define the seminorms

$$p_{m,n}(f) = \sup\{|f^{(k)}(x)| : x \in K_n, 0 \leqslant k \leqslant m\}.$$

Then, the countable collection of seminorms $\{p_{m,n}(f) : m, n \in \mathbb{N}\}$ gives rise to a topology on E say τ. (E, τ) has the following properties:

(i) (E, τ) is lctvs.

(ii) (E, τ) is metrizable and has Heine Borel property.

(iii) $A \subset E$ is bounded if and only if every $p_{m,n}(f)$ is bounded on A.

(iv) $f_n \xrightarrow{\tau} f$ if and only if for each compact set $K \subset (0, \infty)$, $\|(f_n - f)\|_\infty \to 0$ in K.
If $f_n \xrightarrow{\tau} f$, it follows easily that $f_n^{(k)} \xrightarrow{\tau} f^{(k)}$, for all $k = 1, 2, 3, \ldots$.

We, now, continue with the sketch of the proof. We will show that K is compact in this topology defined on E so that we can apply Krein Milman theorem on K. Futhermore, we will show that the extreme points of K are precisely the functions $x \longmapsto e^{-\alpha x}$, $\alpha \in [0, \infty]$. (Let us define $e^{-\infty x} = 0$, for each $x \in (0, \infty)$). It is easy to see that the $ext(K)$ is compact since it is homeomorphic to $[0, \infty]$. Applying Krein Milman theorem to K, each $f \in K$ can be represented by a Borel probability measure m supported on $ext(K)$. This measure m can then be carried to a measure μ on $[0, \infty]$ and the evaluational functionals $f \longmapsto f(x)$, for each $x > 0$ are continuous on E. We combine all these facts and obtain the desired representation of f. We can prove the uniqueness of such a Borel measure on $[0, \infty]$ by applying Stone Weierstrass theorem on the subalgebra of $C[0, \infty]$ generated by the exponentials.

We, now, begin the proof. First we show that K is convex. Let $f, g \in K$ and $\lambda \in (0, 1)$, then $\lambda f + (1 - \lambda)g \in CM$ since CM is a convex cone. Also,

$$(\lambda f + (1 - \lambda)g)(0^+) = \lambda f(0^+) + (1 - \lambda)g(0^+) \leqslant \lambda + (1 - \lambda) = 1.$$

Therefore, $\lambda f + (1 - \lambda)g \in K$. This implies K is convex set. Secondly, we show that K is compact in (E, τ). From (ii), it is enough to show that K is closed and bounded. Let $(f_n) \subset K$ such that $f_n \xrightarrow{\tau} f$. Then, from (iv), we get $f_n^{(k)} \xrightarrow{\tau} f^{(k)}$, for all $k = 1, 2, 3, \ldots.$ Hence, it follows easily from here that,

$$(-1)^n f^{(n)} \geqslant 0, \text{for all } n = 0, 1, 2, \ldots \text{ and}$$

$$f(0^+) = \lim_{x \to 0} f(x) = \lim_{x \to 0} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{x \to 0} f_n(x) = \lim_{x \to 0} f_n(0^+) \leqslant 1.$$

Hence, $f \in K$. This implies K is closed. In order to show that K is bounded, from (iii), it suffices to show that for each $m \in \mathbb{N}$ and $n \in \mathbb{N} \cup \{0\}$, $\sup\{p_{m,n}(f) : f \in K\}$ is finite i.e. to show $\sup \bigcap_{k=0}^{m} \{|f^k(x)| : x \in K_n, f \in K\}$ is finite. For this, it suffices to show $\sup\{|f_m(x)| : x \in K_n, f \in K\}$, for all $m \geqslant 0$ and $n \geqslant 1$. Clearly, the following lemma will prove this fact.

Lemma 2.3.6. Let $I_n = \{(-1)^n f^{(n)} : f \in K\}$, for $n = 0, 1, 2, \ldots.$ Then, for each $a > 0$ and each $n \geqslant 0$, the functions in I_n are bounded above on $[a, \infty)$ by $\delta_{an} \stackrel{def}{=} a^{-n} 2^{(n+1)(n/2)}$

Proof. We prove by mathematical induction. It is clear that the lemma is true for $n = 0$ as the functions in $I_0 = K$ are bounded above by 1. Assume that the lemma is true for I_n i.e. for each $a > 0$ and each $n \geqslant 0$, the functions in I_n are bounded above on $[a, \infty)$ by $\delta_{an} \stackrel{def}{=} a^{-n} 2^{(n+1)(n/2)}$. Now, consider $I_{n+1} = \{(-1)^{(n+1)} f^{(n+1)} : f \in K\}$. Let $a > 0$. Since the functions in I_{n+1} are non-increasing, it is enough to show that $(-1)^{(n+1)} f^{(n+1)}(a) \leqslant \delta_{a(n+1)}$. By Mean Value theorem applied to f^n on $[\frac{a}{2}, a]$, there exists c with $\frac{a}{2} < c < a$ such that

$$\left(\frac{a}{2}\right) f^{(n+1)}(c) = f^{(n)}(a) - f^{(n)}\left(\frac{a}{2}\right).$$

Hence,

$$(-1)\left(\frac{a}{2}\right) f^{(n+1)}(c) = f^{(n)}\left(\frac{a}{2}\right) - f^{(n)}(a).$$

Therefore,

$$\begin{aligned} (-1)^{n+1}\left(\frac{a}{2}\right) f^{(n+1)}(c) &= (-1)^n f^{(n)}\left(\frac{a}{2}\right) - (-1)^n f^{(n)}(a) \\ &\leqslant (-1)^n f^{(n)}\left(\frac{a}{2}\right) \\ &\leqslant \delta_{(a/2)(n)} \text{ ......by induction hypothesis.} \end{aligned} \tag{2.6}$$

Hence,

$$\begin{aligned} (-1)^{n+1} f^{(n+1)}(c) &\leqslant \left(\frac{a}{2}\right)^{-(n+1)} 2^{(n+1)(n/2)} \\ &= (a)^{-(n+1)} 2^{(n+1)(n/2)+(n+1)} \\ &= (a)^{-(n+1)} 2^{(n+2)(n+1/2)} \\ &= \delta_{a(n+1)}. \end{aligned} \tag{2.7}$$

Since, $(-1)^{n+1} f^{(n+1)}$ is non-increasing and $c < a$,

$$(-1)^{n+1} f^{(n+1)}(a) \leqslant \delta_{a(n+1)}.$$

Hence, the result follows. □

Next step will be to identify the extreme points of K.

Lemma 2.3.7. The extreme points of K are those functions f of the form for each $\alpha \in [0, \infty]$, for all $x > 0$, $f(x) = e^{-\alpha x}$.

Proof. Let $f \in ext(K)$. Let $x_0 > 0$. Define, for all $x > 0$, $u(x) = f(x + x_0) - f(x_0)f(x)$. We will prove that $f \pm u \in K$. Then, since $f \in ext(K)$, $u = 0$. Hence, we get, $f(x + x_0) = f(x_0)f(x)$, for each $x, x_0 > 0$. Since f is continuous, this implies either $f = 0$ (the case

$\alpha = \infty$) or $f(x) = e^{-\alpha x}$, for some α. Since $-f'(x) = \alpha e^{-\alpha x} \geqslant 0$, we must have $\alpha \geqslant 0$. Thus, now, we are left to show $f \pm u \in K$. We have,

$$(f+u)(0^+) = f(0^+)(1 - f(x_0)) + f(x_0) \leqslant 1$$

and

$$(f-u)(0^+) = f(0^+)(1 - f(x_0)) - f(x_0) \leqslant 1$$

since $f(0^+), f(x_0) \leqslant 1$. Also, we have,

$$(-1)^n (f+u)^{(n)}(x) = (-1)^n f^{(n)}(x)(1 - f(x_0)) + (-1)^n f^{(n)}(x + x_0) \geqslant 0$$

and

$$(-1)^n (f-u)^{(n)}(x) = [(-1)^n f^{(n)}(x) - (-1)^n f^{(n)}(x + x_0)] + (-1)^n f(x_0) f^{(n)}(x) \geqslant 0$$

since $(-1)^n f^{(n)}$ is non-increasing.

To prove the converse, for $r > 0$, consider the transformation $T_r : K \longrightarrow K$ defined as $T_r(f)(x) = f(rx)$. Clearly, T_r is well-defined. It is also easily seen that T_r is one-one, onto and preserves convex combinations. We claim that T_r takes $ext(K)$ to $ext(K)$. Let $f \in ext(K)$. Define, for all $x > 0$, $h(x) = f(r/x)$. Suppose there exists $h_1, h_2 \in K$ such that

$$h = \frac{h_1 + h_2}{2}.$$

Then, for each $x > 0$,

$$h(rx) = \frac{h_1(rx) + h_2(rx)}{2}.$$

This implies

$$f(x) = \frac{h_1(rx) + h_2(rx)}{2}.$$

Since $f \in ext(K)$, for each $x > 0$, $f(x) = h_1(rx)$ and $f(x) = h_2(rx)$. Hence, for each $x > 0$, $f(x/r) = h_1(x)$ and $f(x/r) = h_2(x)$. Hence, for each $x > 0$, $h(x) = h_1(x)$ and $h(x) = h_2(x)$. This implies $h \in ext(K)$ and $T_r(h) = f$. To prove reverse inclusion, consider $f \in ext(K)$. Suppose there exists $h_1, h_2 \in ext(K)$ such that

$$T_r(f) = \frac{h_1 + h_2}{2}.$$

For each $x > 0$,

$$T_r(f)(x/r) = \frac{h_1(x/r) + h_2(x/r)}{2}$$

This implies

$$f(x) = \frac{h_1(x/r) + h_2(x/r)}{2}.$$

Since $f \in ext(K)$, for each $x > 0$, $f(x) = h_1(x/r)$ and $f(x) = h_2(x/r)$. Thus for each $x > 0$, $f(rx) = h_2(x)$ and $f(rx) = h_2(x)$. Hence,

$$T_r(f) = h_1 \text{ and } T_r(f) = h_2.$$

This implies $T_r(f) \in ext(K)$. Therefore, $T_r(ext(K)) = ext(K)$. We just proved that any extreme point of K is of the form $e^{-\alpha x}$, for some α and hence the image $e^{-\alpha r x}$ of this function under T_r is an extreme point of K. Since this is true for all $r > 0$, all the exponentials are extreme points of K (the constant functions 0 and 1 are clearly extreme points) and hence the proof is complete. □

It can be easily seen that the map $T : [0, \infty] \to K$ defined as $\alpha \longmapsto e^{-\alpha(.)}$ is well-defined and continuous. Since $[0, \infty]$ is compact, its image under T i.e. $ext(K)$ is compact. By Krein Milman representation theorem, for each $f \in ext(K)$, there exists regular Borel measure m supported on $ext(K)$ such that f is resultant of m i.e. for each linear functional in E^*, $L(f) = \int_{ext(K)} L dm$. Now, for each $x > 0$, consider the evaluational functional $L_x(f) = f(x)$, which are continuous on E. Then, we have, for each $x > 0$, $f(x) = \int_{ext(K)} L_x dm$. Define a measure μ for each Borel subset B of $[0, \infty]$ as $\mu(B) = m(T(B))$ i.e. $\mu = m \circ T$. Since $L_x(T(\alpha)) = e^{-\alpha x}$, we have for each $x > 0$,

$$\begin{aligned} f(x) &= \int_{ext(K)} L_x dm \\ &= \int_{T^{-1}(ext(K))} L_x \circ T d(m \circ T) \\ &= \int_0^\infty e^{-\alpha x} d\mu. \end{aligned} \tag{2.8}$$

It is now just left to prove that μ is unique. Suppose there exists another measure ν on $[0, \infty]$ such that for each $x > 0$, $f(x) = \int_0^\infty e^{-\alpha x} d\nu$ and $\nu([0, \infty]) = f(0^+)$. For each $x \geqslant 0$, the function $\alpha \to e^{-\alpha x}$ is continuous on $[0, \infty]$. Let A be the subalgebra of $C([0, \infty])$ generated by these functions, then A consists of finite linear combinations of these same functions. Now, for any $\alpha \in [0, \infty]$,

$$\mu\left(e^{-\alpha(.)}\right) = \int_0^\infty e^{-\alpha x} d\mu = f(x) = \int_0^\infty e^{-\alpha x} d\nu = \nu\left(e^{-\alpha(.)}\right).$$

Therefore, as linear functionals on $C([0, \infty])$, μ and ν are equal on A. Since A separates points of $[0, \infty]$ and for $x = 0$, since the constant function 1 belongs to A, A vanishes nowhere, by Stone Weierstrass theorem, A is dense in $C([0, \infty])$ and so $\mu = \nu$. □

Chapter 3

Choquet's theorem for metrizable compact convex sets

Our next objective is to prove Choquet's theorem for metrizable compact convex sets. Before that, we need to discuss some preliminaries about continuous affine functions defined on a compact convex subset of a lctvs.

3.1 Preliminaries of Affine functions

Definition 3.1.1 (Affine function)**.** Let E_1, E_2 be two linear spaces. A map $\phi : E_1 \to E_2$ is said to be affine if for any two vectors $x, y \in E_1$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta = 1$, $\phi(\alpha x + \beta y) = \alpha\phi(x) + \beta\phi(y)$. Let D be a convex subset of a real linear space E. Then, a real valued function f on D is said to be affine if for $x, y \in D$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta = 1$, $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$, whenever $\alpha x + \beta y \in D$.

Now, let D be a compact convex subset of a lctvs X. Let us recall that the set of all real valued continuous affine functions on D is denoted to be $A_{\mathbb{R}}(D)$ or A when it is unlikely to cause confusion. It is easily seen that the collection $\{\Phi|_D + r : \Phi \in X^*, r \in \mathbb{R}\} \subset A_{\mathbb{R}}(D)$. This collection does not exhaust A and the following examples establish this fact.

Example 3.1.2. Consider $(X = \ell_1, w^*)$ and $D = \{(x_i) \in \ell_1 : \text{for each i} \in \mathbb{N}, |x_i| \leqslant 4^{-i}\}$. Let $\lambda \in (0, 1)$ and $(x_n), (y_n) \in \ell_1$. Since for each $n \in \mathbb{N}$, $|\lambda x_n + (1-\lambda)y_n| \leqslant \lambda|x_n| + (1-\lambda)|y_n| \leqslant \lambda(4^{-n}) + (1-\lambda)(4^{-n}) = 4^{-n}$, $\lambda(x_n) + (1-\lambda)(y_n) = (\lambda x_n + (1-\lambda)y_n) \in D$. This implies D is convex set. Also, since for each $(x_n) \in D$ and $n \in \mathbb{N}$, $x_n \in [-4^{-n}, 4^{-n}]$, $D \cong \prod_{n=1}^{\infty}[-4^{-n}, 4^{-n}]$. Since $\prod_{n=1}^{\infty}[-4^{-n}, 4^{-n}]$ is compact w.r.t. product topology and w^* topology is same as product topology, we can conclude that D is compact in (ℓ_1, w^*).

Define $f : D \to \mathbb{R}$ as

$$f((x_n)) = \sum_{n=1}^{\infty} 2^n x_n.$$

Clearly, f is well-defined as

$$|f(x_n)| \leqslant \sum_{n=1}^{\infty} 2^n |x_n| \leqslant \sum_{n=1}^{\infty} 2^n 4^{-n} = \sum_{n=1}^{\infty} 2^{-n} < \infty.$$

Let $\lambda \in (0,1)$ and $(x_n), (y_n) \in \ell_1$. Then,

$$\begin{aligned} f(\lambda(x_n) + (1-\lambda)(y_n)) &= f((\lambda x_n + (1-\lambda)y_n)) \\ &= \sum_{n=1}^{\infty} 2^n(\lambda x_n + (1-\lambda)y_n) \\ &= \lambda\left(\sum_{n=1}^{\infty} 2^n x_n\right) + (1-\lambda)\left(\sum_{n=1}^{\infty} 2^n y_n\right) \\ &= \lambda f((x_n)) + (1-\lambda) f((y_n)). \end{aligned} \tag{3.1}$$

This implies f is affine function on D. Clearly, f is continuous on D. Thus, $f \in A_{\mathbb{R}}(D)$. We know $(c_0)^* \cong \ell_1$ and given any $F \in (c_0)^*$, there exists $(x_n) \in \ell_1$ such that for each $(\alpha_n) \in c_0$, $F((\alpha_n)) = (x_n)((\alpha_n)) = \sum_{n=1}^{\infty} x_n \alpha_n$. Also, $D \subset c_0$ and c_0 is embedded in ℓ_1. Therefore, clearly, $f \notin c_0 + \mathbb{R}$ and hence $f \notin D^* + \mathbb{R}$.

Example 3.1.3. Consider $(X = \ell_2, w^*)$ and $D = \{(x_i) \in \ell_2 : \text{for each i} \in \mathbb{N}, |x_i| \leqslant 2^{-i}\}$. Then, on similar lines with example 3.1.2, we can show that D is w^*-compact convex subset of ℓ_2. Define $f : D \to \mathbb{R}$ as

$$f((x_n)) = \sum_{n=1}^{\infty} x_n.$$

Clearly, f is well-defined as

$$|f(x_n)| \leqslant \sum_{n=1}^{\infty} |x_n| \leqslant \sum_{n=1}^{\infty} 2^{-n} < \infty.$$

Again, on similar lines with example 3.1.2, we can show $f \in A_{\mathbb{R}}(D)$ and also that $f \notin \{\Phi|_D + r : \Phi \in \ell_2^*, r \in \mathbb{R}\}$.

Remark 3.1.4. $A_{\mathbb{R}}(D)$ is uniformly closed in $C_{\mathbb{R}}(D)$. Let $(f_n) \subset A_{\mathbb{R}}(D)$ such that $f_n \xrightarrow{\|\cdot\|_\infty} f_0$. Let $\lambda \in (0,1)$ and $x, y \in D$. Then,

$$f_0(\lambda x + (1-\lambda)y) = \lim_{n\to\infty} f_n(\lambda x + (1-\lambda)y) = \lambda \lim_{n\to\infty} f_n(x) + (1-\lambda) \lim_{n\to\infty} f_n(y).$$

This implies, $f_0(\lambda x + (1-\lambda)y) = \lambda f_0(x) + (1-\lambda) f_0(y)$. Hence, $f_0 \in A_{\mathbb{R}}(D)$.

Proposition 3.1.5. The subspace $M = X^*|_D + \mathbb{R} \subset C_{\mathbb{R}}(D)$ is uniformly dense in $A_{\mathbb{R}}(D)$.

Proof. Let $f \in A_{\mathbb{R}}(D)$ and $n \in \mathbb{N}$. Consider $W = \{(d, f(d)) : d \in D\}$. Then, $W \subset D \times f(D)$. Let $((d_n, f(d_n))) \subset W$ such that $(d_n, f(d_n)) \to (d, h)$. Since D is compact subset of lctvs X, D is closed in X. Hence, $d \in D$. Now $f \in A_{\mathbb{R}}(D)$ implies $f(d_n) \to f(d)$. Hence, $f(d) = h \in f(D)$. This implies $(d, h) \in W$. Thus, W is closed in $D \times \mathbb{R}$. Let $\lambda \in (0,1)$ and $(d_1, f(d_1)), (d_2, f(d_2)) \in W$. Then since D is convex set and $f \in A_{\mathbb{R}}(D)$,

$$\begin{aligned}\lambda(d_1, f(d_1)) + (1-\lambda)(d_2, f(d_2)) &= (\lambda d_1 + (1-\lambda)d_2, \lambda f(d_1) + (1-\lambda) f(d_2)) \\ &= (\lambda d_1 + (1-\lambda)d_2, f(\lambda d_1 + (1-\lambda)d_2)) \qquad (3.2) \\ &\in W.\end{aligned}$$

Hence, W is convex. Also, W is a closed subset of compact set $D \times f(D)$. Hence, W is compact.

Now, consider $W_1 = \{(d, f(d) + \frac{1}{n}) : d \in D\} \subset D \times (f + \frac{1}{n})(D)$. Then, clearly, W_1 is also compact convex subset of $D \times \mathbb{R}$. Also, $W \cap W_1 = \phi$. Therefore, by Hahn Banach separation theorem, there exists $\Lambda_n \in (X \times \mathbb{R})^*$ and $\lambda_n \in \mathbb{R}$ such that for each $(d, f(d)) \in W$ and $(s, f(s) + \frac{1}{n}) \in W_1$, $\Lambda_n((d, f(d))) < \lambda_n < \Lambda_n((s, f(s) + \frac{1}{n}))$. Let $\Lambda_n = (h_n, p_n) \in X^* \times \mathbb{R}^*$. This implies, for each $d \in D$,

$$h_n(d) + p_n f(d) < \lambda_n < h_n(d) + p_n \left(f(d) + \frac{1}{n} \right).$$

Thus, from the first and third inequalities, $p_n \frac{1}{n} > 0$. Hence, $p_n > 0$. Again, from the above inequality, for each $d \in D$,

$$p_n f(d) < \lambda_n - h_n(d) < p_n \left(f(d) + \frac{1}{n} \right).$$

This implies, for each $d \in D$,

$$0 < \frac{\lambda_n - h_n(d)}{p_n} - f(d) < \frac{1}{n}.$$

Let $\Phi_n = \frac{-h_n}{p_n} \in X^*$ and $r_n = \frac{\lambda_n}{p_n} \in \mathbb{R}^*$. Therefore,

$$\sup_{d \in D} |f(d) - (\Phi_n + r_n)(d)| \leqslant \frac{1}{n}.$$

This implies $(\Phi_n, r_n) \to f$. Therefore, $M = X^*|_D + \mathbb{R}$ is uniformly dense in $A_{\mathbb{R}}(D)$. □

Now we can have a new characterization of the space of continuous functions on K, viz. we claim that $C_{\mathbb{R}}(K) \cong A_{\mathbb{R}}(\mathscr{P}(K))$.

Remark 3.1.6. Proposition 3.1.5 enable us to identify $C_{\mathbb{R}}(K)$ as a space of affine functions on a compact convex subset of a lctvs. Let us recall $\mathscr{P}(K)$, a (w^*) compact convex subset of a lctvs $(M(K), w^*)$. From the above Proposition it is clear that $C_{\mathbb{R}}(K) = (M(K), w^*)^* \hookrightarrow A_{\mathbb{R}}(\mathscr{P}(K))$, which is clearly an embedding. In fact for $f \in C_{\mathbb{R}}(K), \|f\|_\infty = \sup_{t\in K} |f(t)| = \sup_{t\in K} |\tilde{f}(\delta_t)| = \sup_{\mu\in\mathscr{P}(K)} |\tilde{f}(\mu)|$. On the other hand, for any $\varphi \in A_{\mathbb{R}}(\mathscr{P}(K)), \|\varphi\|_\infty = \sup_{\delta_t\in\mathscr{P}(K)} |\varphi(\delta_t)|$. Thus if we define $f(t) = \varphi(\delta_t)$ then $f \in C_{\mathbb{R}}(K)$, which ensures that the other side is also an embedding.

Note 3.1.7. We denote the set of all continuous convex functions on D by S. Hence, $-S$ denotes the collection of continuous concave functions and $A = S \cup -S$. We will write S as $S(D)$ if there is some chance of confusion.

Definition 3.1.8. Let $g : D \to \mathbb{R}$ be a bounded function on D. For each $x \in D$, define,

$$\hat{g}(x) = \inf\{f(x) : f \in -S, f \geqslant g\} \text{ and } \check{g}(x) = \sup\{f(x) : f \in S, f \leqslant g\}.$$

We list out some basic properties of the maps $g \mapsto \hat{g}$ and $g \mapsto \check{g}$ which are required in the subsequent discussion.

Proposition 3.1.9. Let g be a bounded function defined on a closed convex set D. Then,

(a) $\hat{g}$ is concave usc function and $\check{g}$ is convex lsc function.

(b) For each $x \in D$, $\check{g}(x) \leqslant g(x) \leqslant \hat{g}(x)$ and $\hat{g}$, $\check{g}$ are bounded on D.

(c) The map $g \to \hat{g}$ is increasing and sublinear while $g \to \check{g}$ is increasing and superlinear.

(d) $g = \hat{g}$ on D if and only if g is concave and usc function, while $g = \check{g}$ on D if and only if g is convex and lsc function.

(e) If $g \in A_{\mathbb{R}}(D)$ and f is a real valued bounded function on D, then

(i) $\widehat{f+g} = \hat{f} + g$.

(ii) $\widecheck{f+g} = \check{f} + g$.

(f) $\hat{g}(x) = \inf\{a(x) : a \in A_{\mathbb{R}}(D), a \geqslant g\}$ and $\check{g}(x) = \sup\{a(x) : a \in A_{\mathbb{R}}(D), a \leqslant g\}$.

(g) If $g \in C_{\mathbb{R}}(D)$, then $\hat{g}(x) = \sup\{\mu(f) : \mu \in \mathscr{P}(D), r(\mu) = x\}$, for all $x \in D$.

(h) If $x \in D$ then for any two $f, g \in C_{\mathbb{R}}(D)$ $|\hat{f}(x) - \hat{g}(x)| \leqslant \|f - g\|_\infty$.

Proof. (a). Let $x, y \in D$ and $\lambda \in (0, 1)$. Let $f \in -S$ such that $f \geqslant g$ and let $h \in S$ such that $h \leqslant g$. Then, by definition of $-S$, S, $\hat{g}$ and $\check{g}$,

$$f(\lambda x + (1 - \lambda)y) \geqslant \lambda f(x) + (1 - \lambda)f(y) \geqslant \lambda\hat{g}(x) + (1 - \lambda)\hat{g}(y)$$

and

$$h(\lambda x + (1-\lambda)y) \leqslant \lambda h(x) + (1-\lambda)h(y) \leqslant \lambda \check{g}(x) + (1-\lambda)\check{g}(y).$$

Therefore, by definition of $\hat{g}$ and $\check{g}$,

$$\hat{g}(\lambda x + (1-\lambda)y) \geqslant \lambda \hat{g}(x) + (1-\lambda)\hat{g}(y)$$

and

$$\check{g}(\lambda x + (1-\lambda)y) \leqslant \lambda \check{g}(x) + (1-\lambda)\check{g}(y).$$

Thus, $\hat{g}$ is concave and $\check{g}$ is convex. Now, we prove that $\hat{g}$ is usc function. Let $\alpha \in \mathbb{R}$. Let $x_0 \in \{x \in X : \hat{g}(x) < \alpha\}$. Then, $\hat{g}(x_0) < \alpha$. This implies there exists $f \in -S$ such that $f \geqslant g$ and $f(x_0) < \alpha$. Since f is continuous, there exists an open set U containing x_0 such that $f(x) < \alpha$, for each $x \in U$. Thus, $x_0 \in U \subset \{x \in X : \hat{g}(x) < \alpha\}$. This implies $\{x \in X : \hat{g}(x) < \alpha\}$ is an open set in X, for each $\alpha \in \mathbb{R}$. Therefore, $\hat{g}$ is usc function.

Next, we prove $\check{g}$ is lsc function. Let $\alpha \in \mathbb{R}$. Let $x_0 \in \{x \in X : \check{g}(x) > \alpha\}$. Then, $\check{g}(x_0) > \alpha$. This implies there exists $f \in S$ such that $f \leqslant g$ and $f(x_0) > \alpha$. Since f is continuous, there exists an open set U containing x_0 such that $f(x) > \alpha$, for each $x \in U$. Thus, $x_0 \in U \subset \{x \in X : \check{g}(x) > \alpha\}$. This implies $\{x \in X : \check{g}(x) > \alpha\}$ is an open set in X. Thus, for each $\alpha \in \mathbb{R}$, $\{x \in X : \check{g}(x) \leqslant \alpha\}$ is a closed set in X.

(b). Let $x \in D$. Let $f \in -S$ such that $f \geqslant g$ and $h \in S$ such that $h \leqslant g$. Then, by definition of $\hat{g}$ and $\check{g}$,

$$\hat{g}(x) \geqslant g(x) \text{ and } \check{g}(x) \leqslant g(x).$$

Hence, for each $x \in D$, $\check{g}(x) \leqslant g(x) \leqslant \hat{g}(x)$.

Since g is bounded on D and $\check{g} \leqslant g$ on D, $\check{g}$ is bounded on D. Now, since g is bounded, $\|g\|_\infty$ is finite. Consider the concave function, $f(x) = \|g\|_\infty$, for each $x \in D$. Then, $g \leqslant f$ on D. Thus, by definition of $\hat{g}$, for each $x \in D$, $\hat{g}(x) \leqslant f(x)$ i.e. $\hat{g}(x) \leqslant \|g\|_\infty$. Hence, $\hat{g}$ is bounded on D.

(c). Let g_1 and g_2 be any two bounded function on D such that $g_1 \leqslant g_2$. Then,

$$\{f(x) : f \in -S, f \geqslant g_2\} \subset \{f(x) : f \in -S, f \geqslant g_1\}$$

and

$$\{h(x) : h \in S, h \leqslant g_1\} \subset \{h(x) : h \in S, h \leqslant g_2\}.$$

This implies,

$$\inf\{f(x) : f \in -S, f \geqslant g_1\} \leqslant \inf\{f(x) : f \in -S, f \geqslant g_2\}$$

and

$$\sup\{h(x) : h \in S, h \leqslant g_1\} \subset \sup\{h(x) : h \in S, h \leqslant g_2\}.$$

Hence, $\hat{g_1} \leqslant \hat{g_2}$ and $\check{g_1} \leqslant \check{g_2}$ on D. Hence, the maps $g \to \hat{g}$ and $g \to \check{g}$ are increasing.

We now prove that the map $g \to \hat{g}$ is sublinear. Let $\alpha > 0$ and $x \in D$. Then,

$$\begin{aligned} \alpha\hat{g}(x) &= \alpha \inf\{f(x) : f \in -S, f \geqslant g\} \\ &= \inf\{(\alpha f)(x) : \alpha f \in -S, \alpha f \geqslant \alpha g\} \text{ ....since } \alpha > 0 \\ &\geqslant \widehat{\alpha g}(x). \end{aligned} \tag{3.3}$$

and

$$\begin{aligned} \widehat{\alpha g}(x) &= \inf\{f(x) : f \in -S, f \geqslant \alpha g\} \\ &= \inf\{f(x) : f \in -S, \frac{1}{\alpha} f \geqslant g\} \text{ ....since } \alpha > 0 \\ &= \alpha \inf\left\{\left(\frac{1}{\alpha} f\right)(x) : \frac{1}{\alpha} f \in -S, \frac{1}{\alpha} f \geqslant g\right\} \\ &\geqslant \alpha\hat{g}(x). \end{aligned} \tag{3.4}$$

Therefore, for any $\alpha > 0$, $\widehat{\alpha g} = \alpha\hat{g}$. Consider two bounded functions on D say g_1 and g_2. Let $x \in D$. Let $\epsilon > 0$. Then, there exists $h_1, h_2 \in -S$ with $h_1 \geqslant g_1$ and $h_2 \geqslant g_2$ such that

$$h_1(x) < \hat{g_1}(x) + \frac{\epsilon}{2} \text{ and } h_2(x) < \hat{g_2}(x) + \frac{\epsilon}{2}.$$

Then,

$$(h_1 + h_2)(x) < \hat{g_1}(x) + \hat{g_2}(x) + \epsilon$$

where $h_1 + h_2 \in -S$ and $h_1 + h_2 \geqslant g_1 + g_2$. As a result,

$$\widehat{g_1 + g_2}(x) < \hat{g_1}(x) + \hat{g_2}(x) + \epsilon, \text{ for any } \epsilon > 0.$$

Thus, $\widehat{g_1 + g_2}(x) \leqslant \hat{g_1}(x) + \hat{g_2}(x)$, for each $x \in D$. Hence, the map $g \to \hat{g}$ is sublinear.

Next we prove the map $g \to \check{g}$ is superlinear. Let $\alpha > 0$ and $x \in D$. Then,

$$\begin{aligned} \alpha\check{g}(x) &= \alpha \sup\{f(x) : f \in S, f \leqslant g\} \\ &= \sup\{(\alpha f)(x) : \alpha f \in S, \alpha f \leqslant \alpha g\} \text{ ....since } \alpha > 0 \\ &\leqslant \widecheck{\alpha g}(x). \end{aligned} \tag{3.5}$$

and

$$\begin{aligned}\widecheck{\alpha g}(x) &= \sup\{f(x) : f \in S, f \leqslant \alpha g\} \\ &= \sup\{f(x) : f \in S, \frac{1}{\alpha} f \leqslant g\} \text{ ....since } \alpha > 0 \\ &= \alpha \sup\left\{\left(\frac{1}{\alpha} f\right)(x) : \frac{1}{\alpha} f \in S, \frac{1}{\alpha} f \leqslant g\right\} \\ &\leqslant \alpha \check{g}(x).\end{aligned} \tag{3.6}$$

Therefore, for any $\alpha > 0$, $\widecheck{\alpha g} = \alpha \check{g}$. Consider two bounded functions on D say g_1 and g_2. Let $x \in D$. Let $\epsilon > 0$. Then, there exists $h_1, h_2 \in S$ with $h_1 \leqslant g_1$ and $h_2 \leqslant g_2$ such that

$$h_1(x) > \check{g}_1(x) - \frac{\epsilon}{2} \text{ and } h_2(x) > \check{g}_2(x) - \frac{\epsilon}{2}.$$

Then,

$$(h_1 + h_2)(x) > \check{g}_1(x) + \check{g}_2(x) - \epsilon$$

where $h_1 + h_2 \in S$ and $h_1 + h_2 \leqslant g_1 + g_2$. As a result,

$$\widecheck{g_1 + g_2}(x) > \check{g}_1(x) + \check{g}_2(x) - \epsilon, \text{ for any } \epsilon > 0.$$

Thus, $\widecheck{g_1 + g_2}(x) \geqslant \check{g}_1(x) + \check{g}_2(x)$, for each $x \in D$. Hence, the map $g \to \check{g}$ is superlinear.

(d). If $g = \hat{g}$ on D, then clearly from (a), g is concave and usc function. Assume g is concave and usc function. Then, $g \in -S$. Thus, $\hat{g} \leqslant g$ on D and clearly, by definition of $\hat{g}$, $\hat{g} \geqslant g$ on D. Therefore, $g = \hat{g}$ on D.

If $g = \check{g}$ on D, then clearly from (a), g is convex and lsc function. Assume g is convex and lsc function. Then, $g \in S$. Thus, $\check{g} \geqslant g$ on D and clearly, by definition of $\check{g}$, $\check{g} \leqslant g$ on D. Therefore, $g = \check{g}$ on D.

(e).

(i) From (c) and (d), $\widehat{f + g} \leqslant \hat{f} + \hat{g} = \hat{f} + g$.

Now to prove $\hat{f} \leqslant \widehat{f + g} - g$. Let $x \in D$. From (b), $(f + g)(x) \leqslant (\widehat{f + g})(x)$. Thus, for each $x \in D$, $f(x) \leqslant (\widehat{f + g})(x) - g(x)$. Also, $(\widehat{f + g}) - g \in -S$. Thus, by definition of $\hat{f}$, $\hat{f} \leqslant (\widehat{f + g}) - g$ on D.

(ii) From (c) and (d), $\widecheck{f + g} \geqslant \check{f} + \check{g} = \check{f} + g$.

Now to prove $\check{f} \geqslant (\widecheck{f + g}) - g$. Let $x \in D$. From (b), $(f + g)(x) \geqslant (\widecheck{f + g})(x)$. Thus, for each $x \in D$, $f(x) \geqslant (\widecheck{f + g})(x) - g(x)$. Also, $(\widecheck{f + g}) - g \in S$. Thus, by definition of $\check{f}$, $\check{f} \geqslant (\widecheck{f + g}) - g$ on D.

(f). Let $x_0 \in D$. Clearly,

$$\inf\{g(x_0) : g \in -S, f \geqslant f\} \leqslant \inf\{a(x_0) : a \in A_{\mathbb{R}}(D), a \geqslant f\}.$$

Suppose

$$\inf\{g(x_0) : g \in -S, g \geqslant f\} < \inf\{a(x_0) : a \in A_{\mathbb{R}}(D), a \geqslant f\}.$$

Let $m = \inf\{a(x_0) : a \in A_{\mathbb{R}}(D), a \geqslant f\}$. If we can prove that there exists $h \in A_{\mathbb{R}}(D)$ such that $h \geqslant f$ and $h(x_0) < m$, then it would be a contradiction. Now, there exists $g \in -S$ such that $g \geqslant f$ and $g(x_0) < m$. Consider, $A = \{(x,t) \in D \times \mathbb{R} : t \leqslant g(x)\}$ which is, clearly, a convex set and $(x_0, m) \notin A$. Therefore, by Hahn Banach separation theorem, there exists $(\Phi, r) \in (X \times \mathbb{R})^*$ and $\lambda \in \mathbb{R}$ such that

$$(\Phi, r)(x_0, m) > \lambda > \sup_{(x,t)\in A} (\Phi, r)(x,t).$$

Hence, we get the following inequalities, for each $x \in D$:

$$m > \frac{\lambda - \Phi(x_0)}{r} \text{ and } \frac{\lambda - \Phi(x)}{r} > t.$$

Also, we get, $rm > rt$ which implies $r(m-t) > 0$ and hence $r > 0$. Note that $\frac{\lambda-\Phi}{r} \in A_{\mathbb{R}}(D)$ which satisfies our desired properties. This completes the proof. On similar lines, we can show $\check{f}(x) = \sup\{a(x) : a \in A_{\mathbb{R}}(D), a \leqslant f\}$.

(g). Let $f'(x) = \sup\{\mu(f) : \mu \in \mathscr{P}(D), r(\mu) = x\}$, for all $x \in D$. Since $r(\delta_x) = x$, observe that $f'(x) \geqslant f(x)$, for all $x \in D$. Now, let $x, y \in D$ and $\epsilon > 0$. Then there exist measures μ, ν on D with $r(\mu) = x$ and $r(\nu) = y$ such that,

$$\begin{aligned}
\frac{1}{2}f'(x) + \frac{1}{2}f'(y) - \epsilon &= \frac{1}{2}(f'(x) - \epsilon) + \frac{1}{2}(f'(y) - \epsilon) \\
&< \frac{1}{2}(\mu(f) + \nu(f)) \\
&= \frac{\mu+\nu}{2}(f) \leqslant f'(\frac{x+y}{2}), \text{ as } r\left(\frac{\mu+\nu}{2}\right) = \frac{x+y}{2}.
\end{aligned}$$

This implies f' is a concave function on D. Now, we show that f' is usc. Let $x_\alpha \to x_0$. Let $t \leqslant f'(x_\alpha)$, for infinitely many α. Then it remains to prove that $f'(x_0) \geqslant t$. Let $\epsilon > 0$. Then for every infinitely many α, there exists μ_α with $r(\mu_\alpha) = x_\alpha$ such that $\mu_\alpha(f) > t - \epsilon$. Now, (μ_α) has a convergent subsequence in $\mathscr{P}(D)$ say (μ_{α_i}) where $\mu_{\alpha_i} \xrightarrow{w^*} \mu_0$. Without loss of generality, assume that $x_{\alpha_i} \to x_0$ in D. For any $a \in A_{\mathbb{R}}(D)$, $\mu_{\alpha_i}(a) = a(x_{\alpha_i})$ and $a(x_{\alpha_i}) \to a(x_0)$. This implies $\mu_0(a) = a(x_0)$, for all $a \in A_{\mathbb{R}}(D)$. Hence $x_0 = r(\mu_0)$. Now, $\mu_0(f) = \lim_i \mu_{\alpha_i}(f) \geqslant t - \epsilon$. Therefore, $f'(x_0) \geqslant t - \epsilon$, for every $\epsilon > 0$, which implies $f'(x_0) \geqslant t$. So, we get that f' is concave and usc on D and hence by Lemma 4.1.5, $f'(x) = \inf\{a(x) : a \geqslant f', a \in A_{\mathbb{R}}(D)\}$. Now, $\{a \in A_{\mathbb{R}}(D) : a \geqslant f'\} \subseteq \{a \in A_{\mathbb{R}}(D) : a \geqslant f\}$, which implies $f'(x) \geqslant \hat{f}(x)$. On the other hand, if $x \in D$, $h \in A_{\mathbb{R}}(D)$ and $h \geqslant f$, then for any measure μ on D with $r(\mu) = x$, we have $h(x) = \mu(h) \geqslant \mu(f)$. This implies $h(x) \geqslant f'(x)$ and hence (from (f)) $\hat{f}(x) \geqslant f'(x)$. This completes the proof.

(h). It is clear that $\{\mu \in \mathscr{P}(D) : r(\mu) = x\}$ is a w^*-closed subset of $\mathscr{P}(D)$, hence w^*-compact. Hence there exists $\mu_0 \in \mathscr{P}(D)$ with $r(\mu_0) = x$ such that $\mu_0(f) = \hat{f}(x)$. Now $\mu_0(g) \leqslant \hat{g}(x)$. Hence $\hat{f}(x) - \hat{g}(x) \leqslant \mu_0(f) - \mu_0(g) \leqslant \|f - g\|_\infty$. Similarly, $\hat{g}(x) - \hat{f}(x) \leqslant \|f - g\|_\infty$. This completes the proof. □

We can now prove Choquet's theorem for metrizable compact convex sets. Let us consider the below example for motivation:

Consider K to be a compact, T_2 subset of a lctvs X. Then, $(M(K), w^*)$ is a lctvs and $(D = \mathscr{P}(K), w^*)$ is compact convex subset of $M(K)$. Let $\mu \in \mathscr{P}(K)$. Now, clearly, $K \hookrightarrow \{\delta_t : t \in K\} = ext(\mathscr{P}(K))$. Define the same topology on $ext(\mathscr{P}(K))$ as on K. Hence, $ext(\mathscr{P}(K))$ is now topologized. Let B be any Borel subset of $ext(K)$. Define $\tilde{\mu}(B) = \mu(\{t \in K : \delta_t \in B\})$ and $\tilde{\mu} \equiv 0$ on $\mathscr{P}(K) \backslash ext(K)$. Thus, $\tilde{\mu}$ is a measure on $\mathscr{P}(K)$ which we have defined using μ. Now, we claim $\mu = r(\tilde{\mu})$. Let $g \in M(K)^*$. Then, by Riesz representation theorem and definition of w^* topology, g has to be an evaluational functional on $M(K)$ i.e. there exists $f_0 \in C_{\mathbb{R}}(K)$ such that for any $\nu \in M(K)$, $g(\nu) = \nu(f_0)$. Therefore, in particular,

$$g(\mu) = \mu(f_0) = \int_K f_0 d\mu = \int_{\mathscr{P}(K)} \mu(f_0) d\tilde{\mu}$$

Therefore,

$$g(\mu) = \int_{\mathscr{P}(K)} g d\tilde{\mu}.$$

This implies that $\mu = r(\tilde{\mu})$.

Before going to the main result of this chapter let us recall the following well known fact. For the sake of completeness we include the proof here.

Proposition 3.1.10. Let K be a metrizable compact Hausdorff space and $C_{\mathbb{R}}(K)$, the space of all real valued continuous functions on K. Then, $C_{\mathbb{R}}(K)$ is separable.

Proof. Let d be the metric on K. Since K is metrizable, K is separable. Let $\{x_n : n \in \mathbb{N}\}$ be a dense subset of K. Let $m, n \in \mathbb{N}$. Define $f_{mn} : K \to \mathbb{R}$ as

$$f_{mn}(x) = \begin{cases} \frac{1}{m} - d(x, x_n) & \text{if } \frac{1}{m} \geqslant d(x, x_n) \\ 0 & \text{if otherwise} \end{cases} \tag{3.7}$$

It is clear that for all $m, n \in \mathbb{N}$, $f_{mn} \in C_{\mathbb{R}}(K)$ and it vanishes outside the open ball $B(x_n, \frac{1}{m})$. Consider the countable set $\mathscr{A} \subset C_{\mathbb{R}}(K)$ to be collection of all f_{mn}, for every $m, n \in \mathbb{N}$ and the constant function 1. Let $x, y \in K$ such that $x \neq y$, let $m \in \mathbb{N}$ be sufficiently large such that $d(x, y) > \frac{1}{m}$. Since $\{x_n : n \in \mathbb{N}\}$ is dense in K, there exists $n \in \mathbb{N}$ such that $x_n \in B(x, \frac{1}{2m})$. Hence, $f_{mn}(x) \neq 0$ but $f_{mn}(y) = 0$. Therefore, $\mathscr{A}$ separates points on K. We now show that $\mathscr{A}$ forms a separable subalgebra of $C_{\mathbb{R}}(K)$, hence by Stone-Weierstrass Theorem it follows that, $\mathscr{A}$ is dense in $C_{\mathbb{R}}(K)$. Therefore, $C_{\mathbb{R}}(K)$ is separable. Now consider the algebra generated by $\mathscr{A}$, $\{\prod_i f_{m_i n_i} : i$ runs over a finite set$\}$. First observe that for two points x_{n_1} and x_{n_2}, the possible values of m_1 and m_2 for which $f_{m_1 n_1}.f_{m_2 n_2}$ is non zero should satisfy $d(x_{n_1}, x_{n_2}) < \frac{1}{m_1} + \frac{1}{m_2}$. Now cardinality of the set of pair of natural numbers, taken two at a time, is countable and set of all finite product of any such pair is again countable: any function of type $\prod_i f_{m_i n_i}$ is associated to a finite subset of the set of pair of natural numbers, viz. $\{(m_i, n_i)\}$. This leads to that $\mathscr{A}$ generates a separable subalgebra over $C_{\mathbb{R}}(K)$ and this completes the proof. □

The converse of Proposition 3.1.10 is also true and it follows from Proposition B.1.4.

3.2 Main results: Choquet-Bishop-De Leeuw existence Theorem

We now come to the main Theorem of this chapter.

Theorem 3.2.1. Let D be a metrizable compact convex subset of an lctvs X and $x_0 \in D$. Then, there exists $\mu \in \mathscr{P}(D)$ which is supported on $ext(D)$ such that $x_0 = r(\mu)$.

Proof. The metrizability of D implies $C_{\mathbb{R}}(D)$, the space of all real valued continuous functions on D, is separable in the sup norm topology, by earlier remark. We know $A_{\mathbb{R}}(D)$ is uniformly closed in $C_{\mathbb{R}}(D)$ and also separable.

Hence $S_{A_{\mathbb{R}}(D)} = \{f \in A_{\mathbb{R}}(D) : \|f\|_\infty = 1\}$ is separable. Let $\{h_n\}_{n=1}^{\infty}$ be a dense subset of $S_{A_{\mathbb{R}}(D)}$. Define, for each $x \in D$,

$$h(x) = \sum_{n=1}^{\infty} \frac{h_n^2(x)}{2^n}.$$

It is easily seen that $h \in C_{\mathbb{R}}(D)$. We now claim that h is strictly convex on D. Let $x, y \in D$ such that $x \neq y$. Then, there exists h_m such that $h_m(x) \neq h_m(y)$. Let $\lambda \in (0, 1)$.

Consider,

$$\begin{aligned} h(\lambda x + (1-\lambda)y) &= \sum_{n=1}^{\infty} \frac{h_n^2(\lambda x + (1-\lambda)y)}{2^n} \\ &= \frac{h_m^2(\lambda x + (1-\lambda)y)}{2^m} + \sum_{n \neq m} \frac{h_n^2(\lambda x + (1-\lambda)y)}{2^n} \\ &< \frac{(\lambda h_m(x) + (1-\lambda)h_m(y))^2}{2^m} + \sum_{n \neq m} \frac{(\lambda h_n(x) + (1-\lambda)h_n(y))^2}{2^n} \\ &= \sum_{n=1}^{\infty} \frac{(\lambda h_n(x) + (1-\lambda)h_n(y))^2}{2^n}. \end{aligned} \tag{3.8}$$

The strictly inequality in the third step is due to the fact that the map $t \to t^2$ is strictly convex on $\mathbb{R}$. Thus, we showed that h is strictly convex on D. Now, define $Y = span\{A_{\mathbb{R}}(D) \cup \{h\}\} = \{g + \alpha h : g \in A_{\mathbb{R}}(D), \alpha \in \mathbb{R}\}$. Define $p : C_{\mathbb{R}}(D) \to \mathbb{R}$ as $p(g) = \hat{g}(x_0)$. As a result of property (c) of $\hat{g}$ function, p is sublinear on $C_{\mathbb{R}}(D)$. Define $\Phi : Y \to \mathbb{R}$ as $\Phi(g + \alpha h) = g(x_0) + \alpha\hat{h}(x_0)$. Clearly, Φ is linear on Y. If $\alpha > 0$, then by properties (c) and (e) we get,

$$\Phi(g + \alpha h) = g(x_0) + \alpha\hat{h}(x_0) = g(x_0) + \widehat{\alpha h}(x_0) = (\widehat{g + \alpha h})(x_0) = p(g + \alpha h).$$

If $\alpha = 0$, then by property (b),

$$\Phi(g) = g(x_0) \leqslant \hat{g}(x_0) = p(g).$$

Now, assume $\alpha < 0$. Then, by property (b), $\alpha\hat{h}(x_0) \leqslant \alpha h(x_0)$. Since h is strictly convex and continuous on D, αh is strictly concave and also continuous on D. Thus, by property (d), $\widehat{\alpha h}(x_0) = \alpha h(x_0)$. Using all these facts and property (e), we get,

$$\Phi(g + \alpha h) = g(x_0) + \alpha\hat{h}(x_0) \leqslant g(x_0) + \alpha h(x_0) = g(x_0) + \widehat{\alpha h}(x_0) = \widehat{g + \alpha h}(x_0).$$

Hence, $\Phi(g + \alpha h) = p(g + \alpha h)$. This shows that $\Phi \leqslant p$ on D. Therefore, by Hahn Banach extension theorem, we can extend Φ to the whole space $C_{\mathbb{R}}(D)$ such that for each $f \in C_{\mathbb{R}}(D)$, $\Phi(f) \leqslant p(f)$. Now, for any $f \in C_{\mathbb{R}}(D)$ such that $\|f\|_\infty \leqslant 1$,

$$|\Phi(f)| \leqslant |p(f)| = |\hat{f}(x_0)| \leqslant \|f\|_\infty.$$

This implies $\|\Phi\| \leqslant 1$. Since $\Phi(1) = 1$, $\|\Phi\| = 1$

By, Riesz Representation theorem, there exists a regular Borel measure on D such that for each $f \in C_{\mathbb{R}}(D)$, $\Phi(f) = \int_D f dm$. Since $\|\Phi\| = 1$, $\|m\| = 1$. Now,

$$\Phi(1) = \int_D 1 dm = m(D) = \|m\| = 1.$$

This implies

$$\int_D 1 dm = 1.$$

Thus, $m \in \mathscr{P}(D)$. Let $f \in A_{\mathbb{R}}(D)$. Then,

$$f(x_0) = \Phi(f) = \int_D f dm.$$

This implies $x_0 = r(m)$.

We now prove that the measure m is supported on $ext(D)$. Now, by definition of Φ and Riesz Representation theorem, we get

$$\Phi(h) = \hat{h}(x_0) = m(h) = \int_D h dm.$$

Since $h \leqslant \hat{h}$, this implies

$$\int_D h dm \leqslant \int_D \hat{h} dm.$$

This implies $m(h) \leqslant m(\hat{h})$. If $a \in A_{\mathbb{R}}(D)$ such that $a \geqslant h$, then $a \geqslant \hat{h}$, by property (f). Therefore,

$$a(x_0) = \int_D a dm \geqslant \int_D \hat{h} dm.$$

This implies $\hat{h}(x_0) \geqslant \int_D \hat{h} dm$. Hence, $\int_D h dm \geqslant \int_D \hat{h} dm$, which implies $m(h) \geqslant m(\hat{h})$. Thus, $m(h) = m(\hat{h})$. Hence, we get that $\int_D h dm = \int_D \hat{h} dm$, which implies $\int_D (\hat{h} - h) dm = 0$. Therefore, $\hat{h} - h \equiv 0$ $a.e.[m]$. This implies $m(\{x \in D : h(x) \neq \hat{h}(x)\}) = 0$. Thus, $m(D) = m(\{x \in D : h(x) = \hat{h}(x)\})$. We now claim that $\{x \in D : h(x) = \hat{h}(x)\} \subset ext(D)$. Let $x \in D$ such that $h(x) = \hat{h}(x)$. Suppose $x = \frac{x_1 + x_2}{2}$ such that $x_1 \neq x_2$. Consider,

$$\begin{aligned} h(x) = h\left(\frac{x_1 + x_2}{2}\right) &< \frac{h(x_1) + h(x_2)}{2} \\ &\leqslant \frac{\hat{h}(x_1) + \hat{h}(x_2)}{2} \\ &\leqslant \hat{h}\left(\frac{x_1 + x_2}{2}\right) \\ &= \hat{h}(x). \end{aligned} \tag{3.9}$$

The first inequality is due to the fact that h is strictly convex on D. Hence, $h(x) < \hat{h}(x)$, which is a contradiction. This implies $x \in ext(D)$. Therefore, the measure m is supported on $ext(D)$. This completes the proof. □

Let us recall if X is a separable normed linear space then the dual ball is a metrizable compact convex subset of (X^*, w^*)(see Appendix B), hence Theorem 3.2.1 is applicable for such spaces. Now, by Choquet's theorem for metrizable compact convex sets, for every $x^* \in B_{X^*}$, there exists $\mu \in \mathscr{P}(B_{X^*})$ supported on $ext(B_{X^*})$ such that $x^* = r(\mu)$.

We now prove the converse of Proposition 3.1.10. If $C(K)$ is separable, for some compact, Hausdorff space K, then K is metrizable: Let $X = C(K)$, from the above Example it follows that (B_{X^*}, w^*) is metrizable. Now $B_{X^*} = B_{M(K)}$ and hence $ext(B_{X^*}) = \{\pm\delta_t : t \in K\}$. Now as $(\{\delta_t : t \in K\}, w^*) \cong K$ (homeomorphism), the topology on K is metrizable.

We end this Chapter by giving an example of a compact, Hausdorff and also compact, convex subset of a lctvs E which are not metrizable. These examples are relevant in the context of the discussion in the next Chapter.

Let $\beta\mathbb{N}$ be the Stone-Čech compactification of the natural numbers. It is well known that $C(\beta\mathbb{N}) \cong \ell_\infty$ (isometrically isomorphic). Non separability of $C(\beta\mathbb{N})$ ensures that $\beta\mathbb{N}$ is non metrizable. From our earlier arguments it is clear that $\mathscr{P}(\beta\mathbb{N})$ and hence $B_{M(\beta\mathbb{N})}$ are some examples of compact convex subsets which are non metrizable.

Chapter 4

Choquet's theorem for non metrizable compact convex sets

We begin this Chapter with Stone's generalization of Stone-Weierstrass Theorem. The main Theorem in [12] is proved in a more general setup, we only mention the following result which is relevant to our investigation.

Lemma 4.0.1 (Stone, [12])**.** Let K be a compact Hausdorff space and $\mathcal{U}$ be a sublattice of $C_{\mathbb{R}}(K)$ such that, for any $f \in C_{\mathbb{R}}(K)$, for any two points $x, y \in K$ and for any positive number ε, there exists a function $f_{xy} \in \mathcal{U}$ such that $|f(x) - f_{xy}(x)| < \varepsilon$, $|f(y) - f_{xy}(y)| < \varepsilon$. Then $f \in \overline{\mathcal{U}}$.

Proof. It remains to prove that there exists $h \in \mathcal{U}$ such that $\|f - h\|_\infty < \varepsilon$.

Fix $x \in K$ and define $G_y = \{z : |f(z) - f_{xy}(z)| < \varepsilon\}$. By hypothesis $x, y \in G_y$ and hence $\bigcup_{y \in K} G_y = K$. The compactness of K ensures the existence of points $y_1, \ldots, y_n$ such that $\bigcup_{i=1}^n G_{y_i} = K$. Define $g_x = \max\{f_{xy_1}, \ldots, f_{xy_n}\}$. Choose any $z \in K$ and get G_{y_i} such that $z \in G_{y_i}$, then $g_x(z) \geqslant f_{xy_i}(z) > f(z) - \varepsilon$.

On the other hand $f_{xy}(x) < f(x) + \varepsilon$ for all y implies $g_x(x) < f(x) + \varepsilon$. We now continue a similar argument for g_x. Let $H_x = \{z \in K : g_x(z) < f(z) + \varepsilon\}$, then $x \in H_x$ and hence there exist $x_1, x_2, \ldots, x_n$ such that $K = \bigcup_{i=1}^n H_{x_i}$. Let $h(x) = \min\{g_{x_1}, \ldots, g_{x_n}\}$.

Since for any $z \in K$ there exists H_{x_k} such that $z \in H_{x_k}$, hence $h(z) \leqslant g_{x_k}(z) < f(z) + \varepsilon$. On the other hand, the fact that $g_x(z) > f(z) - \varepsilon$ for all z and all x implies that $h(z) > f(z) - \varepsilon$ for all z. Thus we have $|h(z) - f(z)| < \varepsilon$, for all $z \in K$. This completes the proof. $\square$

4.1 A new setup: Boundary measure

We begin this Section with the following important fact.

Proposition 4.1.1. For a compact convex set D of a lctvs E, $S(D)-S(D)$ is uniformly dense in $C_{\mathbb{R}}(D)$.

Proof. First observe that $S(D)-S(D)$ forms a lattice in $C_{\mathbb{R}}(D)$. In other words for any $f, g \in S(D)-S(D)$, $f \vee g, f \wedge g \in S(D)-S(D)$. In fact, for $i=1,2$, $f_i, g_i \in S(D)$, $(f_1-g_1) \vee (f_2-g_2) = (f_1+g_2) \vee (f_2+g_1) - (g_1+g_2)$ and $(f_1-g_1) \wedge (f_2-g_2) = \frac{1}{2}[(f_1+f_2)-(g_2+g_1)-|(f_1+g_2)-(f_2+g_1)|]$. The last identity follows from the fact that, for any two reals a, b $a \wedge b = \frac{1}{2}(a+b-|a-b|)$. Clearly $(f_1+f_2)-(g_2+g_1) \in S(D)-S(D)$ and both $(f_1+g_2)-(f_2+g_1), (f_2+g_1)-(f_1+g_2) \in S(D)-S(D)$. Now for any function h, $|h(x)| = \max\{h(x), -h(x)\}$, hence $|(f_1+g_2)-(f_2+g_1)| \in S(D)-S(D)$; being a subspace $(f_1-g_1) \wedge (f_2-g_2) \in S(D)-S(D)$.

Also the subspace $S(D)-S(D)$ satisfies the condition in Lemma 4.0.1, which is in fact a direct consequence of $A_{\mathbb{R}}(D) \subseteq S(D)$. Hence $S(D)-S(D)$ is uniformly dense in $C_{\mathbb{R}}(D)$. ☐

We now turn our attention to the non-metrizable version of Choquet's theorem. Before proceeding, we need to define few notions.

$S(D)$ is a cone in $C_{\mathbb{R}}(D)$. So, define an ordering $\prec$ on $M^+(D)$, the set of all non-negative regular Borel measures on D, by $\mu \prec \nu$ if and only if $\mu(f) \leqslant \nu(f)$, for all $f \in S(D)$. Clearly, $\prec$ is reflexive and transitive.

Now, we prove that $\prec$ is antisymmetric. Let $\mu, \nu \in M^+(D)$ such that $\mu \prec \nu$ and $\nu \prec \mu$. This implies for all $f, h \in S(D)$,

$$\mu(f) \leqslant \nu(f) \text{ and } \nu(h) \leqslant \mu(h)$$

Thus, for all $g \in S(D)$,

$$\mu(-g) \leqslant \nu(-g)$$

Hence, for all $f, g \in S(D)$,

$$\mu(f-g) \leqslant \nu(f-g)$$

Similarly, we get for all $h, p \in S(D)$,

$$\nu(h-p) \leqslant \mu(h-p)$$

Therefore, for all $f \in S(D) - S(D)$, $\mu(f) = \nu(f)$. By our earlier observation, $S(D) - S(D)$ is dense in $C_{\mathbb{R}}(D)$ and hence, $\mu = \nu$ on $C_{\mathbb{R}}(D)$.

Proposition 4.1.2. (*a*) Let $\mu, \lambda \in M^+(D)$. If $\mu \prec \lambda$, then $r(\mu) = r(\lambda)$.

(*b*) If $\mu \in M^+(D)$ represents $x \in D$, then $\delta_x \prec \mu$.

Proof. (*a*). Let $r(\mu) = x$ and $r(\lambda) = y$. Suppose $x \neq y$. Then there exists $f \in X^* \subset S(D)$ such that $f(x) > f(y)$, without loss of generality(otherwise we will consider $-f$). Now, $f(x) = \mu(f)$ and $f(y) = \lambda(f)$. Thus, $\mu(f) > \lambda(f)$ which is a contradiction since $\mu \prec \lambda$. Therefore, $r(\mu) = r(\lambda)$.

(*b*). Let $f \in S(D)$. Then, $-f$ is concave. Hence, $\widehat{(-f)} = -f$. Since $x = r(\mu)$, we have,

$$\begin{aligned}(-f)(x) &= \inf\{h(x) : h \in A(D), h \geqslant (-f)\} \\ &= \inf\{\mu(h) : h \in A(D), h \geqslant (-f)\} \\ &\geqslant \mu(-f)\end{aligned} \tag{4.1}$$

This implies, $\mu(f) \geqslant f(x)$. Hence, $\mu(f) \geqslant \delta_x(f)$. This is true for any $f \in S(D)$. Hence, $\delta_x \prec \mu$. □

Remark 4.1.3. We can give an intuitive justification for the definition of $\mu \prec \lambda$ as follows:

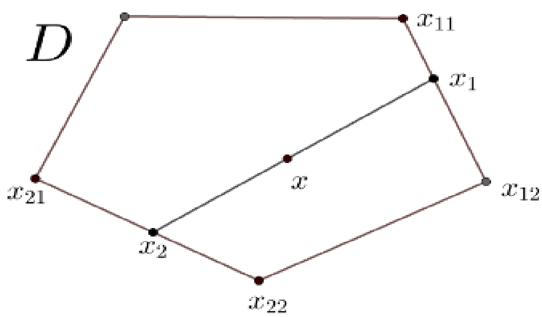

FIGURE 4.1

If D is a polygon and $x \in D$, first write x as the resultant of points on the two faces (as shown in Figure 4.1) and then write these points as the resultants of the vertices of the polygon. Finally, we obtain x as a resultant of the vertices $\{x_{ij}\}$. Roughly speaking, we made x more and more diffused such that the points of support approach the extreme elements.

Thus, $\mu(f) \leqslant \lambda(f)$, for $f \in S(D)$, should mean that the mass of λ is more concentrated in the neighbourhood of extreme points than that of μ. This leads to the hope that, in this ordering, a measure is maximal that is completely diffused when it is concentrated on the extreme points.

Lemma 4.1.4. If $\mu \in M^+(D)$, then there exists a maximal measure $\lambda \in M^+(D)$ such that $\mu \prec \lambda$.

Proof. Let $\mathscr{F} = \{\lambda \in M^+(D) : \mu \prec \lambda\}$. Then, $\mathscr{F}$ is a partially ordered set w.r.t. the ordering $\prec$. It is enough to show that $\mathscr{F}$ has a maximal element in $\mathscr{F}$ since if $\nu \in M^+(D)$ such that $\lambda \prec \nu$, then $\mu \prec \nu$ and hence $\nu \in \mathscr{F}$ and this implies $\nu = \lambda$. Consider $\mathscr{Z}$ to be a chain in $\mathscr{F}$. We now regard $\mathscr{Z}$ as a net in the obvious way (denoted by the elements of $\mathscr{Z}$ themselves) and clearly, using proposition 4.1.2 (a), $\mathscr{Z} \subset \{\lambda \in M^+(D) : \lambda(1) = \mu(1)\}$. Now, the set $\{\lambda \in M^+(D) : \lambda(1) = \mu(1)\} = \mu(1)\mathscr{P}(D)$ is w^*-compact in $C_{\mathbb{R}}(D)^*$. Thus, there exists $\mu_0 \in M^+(D)$ and a subnet $\{\mu_\alpha\}_\alpha$ of $\mathscr{Z}$ which converges to μ_0 in the w^* topology. If $\lambda_1 \in \mathscr{Z}$, then by definition of subnet, $\lambda_1 \prec \mu_\alpha$ eventually and hence $\lambda_1 \prec \mu_0$, which implies that μ_0 is an upper bound for $\mathscr{Z}$. Futhermore, since $\mu \prec \mu_0$, we have $\mu_0 \in \mathscr{Z}$. By Zorn's lemma, $\mathscr{F}$ contains a maximal element. ☐

We now state here a technical result. Let $\check{S}$ or $\check{S}(D)$ denote the collection of lsc convex functions on D.

Lemma 4.1.5. Let D be a compact convex set of a lctvs X. Let $f \in \check{S}(D)$. Then, for every $x \in D$, $f(x) = \sup\{a(x) : a \in A_{\mathbb{R}}(D), a < f\}$.

Proof. Let $f \in \check{S}(D)$. Consider $M = \{(y, \alpha) : y \in D, \alpha \geqslant f(y)\} \subset X \times \mathbb{R}$. Let $\{(y_\beta, \alpha_\beta)\}$ be a net in M converging to (y, α). Thus, $y_\beta \to y$ and $\alpha_\beta \to \alpha$, which implies $f(y_\beta) \to \alpha$. Since f is lsc function, $f(y) \leqslant \liminf_\beta f(y_\beta) \leqslant \liminf_\beta \alpha = \alpha$. This implies $(y, \alpha) \in M$. Hence, M is closed in $X \times \mathbb{R}$. Let $(y_1, \alpha_1), (y_2, \alpha_2) \in M$ and $\lambda \in (0, 1)$. Since D is convex, $\lambda y_1 + (1 - \lambda) y_2 \in D$. Also,

$$\lambda\alpha_1 + (1 - \lambda)\alpha_2 \geqslant \lambda f(y_1) + (1 - \lambda) f(y_2) \geqslant f(\lambda y_1 + (1 - \lambda) y_2).$$

This implies $\lambda(y_1, \alpha_1) + (1 - \lambda)(y_2, \alpha_2) \in M$. Hence, M is convex.

Let $x \in D$ be fixed. Clearly, $f(x) \geqslant \sup\{a(x) : a \in A_{\mathbb{R}}(D), a < f\}$. Let $\epsilon < f(x)$ be arbitrary. We shall show that there exists $a \in A_{\mathbb{R}}(D)$ such that $a < f$ and $\epsilon < a(x)$.

Case 1: $f(x) < \infty$. By Hahn Banach separation theorem, there exists a closed hyperplane H in $X \times \mathbb{R}$ which strictly separates M from the point (x, ϵ). Here $H \subset X \times \mathbb{R}$ is graph of an affine function say $a : X \to \mathbb{R}$ (Note: H cannot be of the form $H_1 \times \mathbb{R}$, where H_1 is a hyperplane in X as H separates (x, ϵ) from $(x, f(x))$. The open half spaces associated with H are $\{(y,\alpha) : \alpha > a(y)\}$ and $\{(y,\alpha) : \alpha < a(y)\}$. By assumptions, one of these contains (x,ϵ) and other all of M. Since $\epsilon < f(x)$, we get $\epsilon < a(x)$ and $a < f$. Also, a is continuous on D, since $a^{-1}(\{0\})$ is closed as $a^{-1}(\{0\}) \times \{0\} = H \cap (X \times \{0\})$.

Case 2: $f(x) = \infty$. Let $\delta > 0$ be an arbitrary number such that $\epsilon < \delta < \infty$ and define $N = \overline{conv}((x,\delta) \cup M)$. We now claim that $(x,\epsilon) \notin N$. Define the sets $A = \{(y,\alpha) : y \in D, \alpha \geqslant \delta\}$ and $B = \{(y,\alpha) : y \in D, f(y) \leqslant \alpha \leqslant \delta\}$.

Claim: $conv((x,\delta) \cup M) \subset A \cup conv((x,\delta) \cup B)$

Proof of the claim: Let $\sum_{i=1}^{n} \delta_i(y_i,\alpha_i) + \beta(x,\delta) \in conv((x,\delta) \cup M)$, where $\sum_{i=1}^{n} \delta_i + \beta = 1$. Since D is convex, $\sum_{i=1}^{n} \delta_i y_i + \beta x \in D$. Suppose $\sum_{i=1}^{n} \delta_i(y_i,\alpha_i) + \beta(x,\delta) \notin conv((x,\delta) \cup B)$. This implies that $(y_i,\alpha_i) \notin B$ and hence $f(y_i) > \alpha_i$ or $\delta < \alpha_i$, for every $i = 1, 2, ...$ Since $(y_i,\alpha_i) \in M$, we have $\delta < \alpha_i$, for every $i = 1, 2, ...$ Thus, $\sum_{i=1}^{n} \delta_i\alpha_i + \beta\delta > \sum_{i=1}^{n} \delta_i\delta + \beta\delta = \delta$. This implies $\sum_{i=1}^{n} \delta_i(y_i,\alpha_i) + \beta(x,\delta) \in A$. This completes the proof.

Therefore, $N \subset A \cup \overline{conv}((x,\delta) \cup B) = A \cup conv((x,\delta) \cup B)$, since B is closed subset of $D \times [m,\delta]$, where $m = \inf\{f(y) : y \in D\}$ which implies it is compact and also, convex hull of finite union of compact sets is compact. Clearly, $(x,\epsilon) \notin A \cup conv((x,\delta) \cup B)$ and so $(x,\epsilon) \notin N$. Now, we apply the same argument as in Case 1 to N. □

Remark 4.1.6. In order to prove the next lemma, we need Dini's lemma for a net of functions which states as follows:

Let D be a compact subset of a lctvs X. Let $\{f_\alpha\}$ be an increasing net of real valued continuous functions where for each α, f_α is continuous on D. Assume that $\{f_\alpha\}$ converges pointwise to a continuous function $f : D \to \mathbb{R}$. Then, $\{f_\alpha\}$ converges to f uniformly on D.

Lemma 4.1.7. $\check{S}(D)$ consists of all pointwise limits of increasing nets of functions of the form $a_1 \vee a_2 \vee ... \vee a_n$, where for all $i = 1, 2, ..., n$, $a_i \in A_{\mathbb{R}}(D)$. Similarly, $S(D)$ consists of all uniform limits of increasing nets of above such functions.

Proof. Let $x \in D$ be arbitrary. By previous lemma,

$$f(x) = \sup\{a(x) : a \in A_{\mathbb{R}}(D), a < f\}.$$

This implies there exists a sequence $\{a_n\}$, where $a_n \in A_{\mathbb{R}}(D)$ and $a_n < f$ such that $f(x) = \sup_n\{a_n(x)\}$. Define $g_n = a_1 \vee ... \vee a_n$ then $g_n \leqslant g_{n+1}$. Hence,

$$f(x) = \sup\{a_n(x) : a_n \in A_{\mathbb{R}}(D), a_n < f\} = \lim_{n\to\infty} g_n(x).$$

Since $x \in D$ is arbitrary, f is pointwise limit of $\{g_n\}$.

Let $f \in S(D)$. Now, $S(D) \subset \check{S}(D)$ and hence f is pointwise limit of increasing nets of functions of the form $a_1 \vee a_2 \vee ... \vee a_n$, where for all $i = 1, 2, ..., n$, $a_i \in A(D)$. By Dini's lemma, we get that f is uniform limit of above such functions. □

Remark 4.1.8. Let $a_1, a_2 \in A(D)$, $\lambda \in (0, 1)$ and $x, y \in D$. Consider,

$$\begin{aligned}(a_1 \vee a_2)(\lambda x + (1-\lambda)y) &= \max\{a_1(\lambda x + (1-\lambda)y), a_2(\lambda x + (1-\lambda)y)\} \\ &\leqslant \max\{\lambda a_1(x) + (1-\lambda)a_2(y), \lambda a_2(x) + (1-\lambda)a_2(y)\} \\ &\leqslant \lambda \max\{a_1(x), a_2(x)\} + (1-\lambda)\max\{a_1(y), a_2(y)\} \\ &= \lambda(a_1 \vee a_2)(x) + (1-\lambda)(a_1 \vee a_2)(y).\end{aligned} \tag{4.2}$$

Thus, $a_1 \vee a_2 \in S(D)$. Therefore, by induction, we can conclude that, for any $n \in \mathbb{N}$, $a_1 \vee ... \vee a_n \in S(D)$, where for all $i = 1, 2, ..., n$, $a_i \in A(D)$.

Proposition 4.1.9. Let $\mu, \nu \in M^+(D)$. Then, the following are equivalent:

(a) $\mu \prec \nu$.

(b) For any $f \in C_{\mathbb{R}}(D)$, $\mu(\check{f}) \leqslant \nu(f)$.

(c) For any $f \in C_{\mathbb{R}}(D)$, $\nu(f) \leqslant \mu(\hat{f})$.

Proof. First we prove (i) implies (ii). Assume $\mu \prec \nu$. We know that $\check{f} \in \check{S}(D)$. Hence, by above Lemma 4.1.7, there exists an increasing sequence say (g_n) such that $g_n \to \check{f}$, where each $g_n = a_1 \vee ... \vee a_n$, for some $a_i \in A(D)$, for all $i = 1, 2, ..., n$. Now, by Remark 4.1.8, for each $n \in \mathbb{N}$, $g_n \in S(D)$. Since $\mu \prec \nu$, $\mu(g_n) \leqslant \nu(g_n)$. Also,

$$\mu(\check{f}) = \lim_{n\to\infty} \mu(g_n) \text{ and } \nu(\check{f}) = \lim_{n\to\infty} \nu(g_n).$$

Hence, $\mu(\check{f}) \leqslant \nu(\check{f})$. Now, since $\nu(\check{f}) \leqslant \nu(f)$, we get $\mu(\check{f}) \leqslant \nu(f)$.

We prove (ii) implies (i). Assume for each $f \in C_{\mathbb{R}}(D)$, $\mu(\check{f}) \leqslant \nu(f)$. Let $f \in S(D)$. Then, $f = \check{f}$. This implies $\mu(f) = \mu(\check{f}) \leqslant \nu(f)$ and hence, for every $f \in S(D)$, $\mu(f) \leqslant \nu(f)$. Therefore, $\mu \prec \nu$.

We now prove (ii) implies (iii). Assume for each $f \in C_{\mathbb{R}}(D)$, $\mu(\check{f}) \leqslant \nu(f)$. Hence, $\mu(\smile{-f}) \leqslant \nu(-f)$. For each $x \in D$, consider,

$$\begin{aligned}\smile{-f}(x) &= \sup\{g(x) : g \leqslant -f, g \in S(D)\} \\ &= \sup\{g(x) : -g \geqslant f, -g \in S(D)\} \\ &= -\inf\{-g(x) : -g \geqslant f, -g \in S(D)\} \\ &= -\hat{f}(x).\end{aligned} \tag{4.3}$$

This implies $\mu(\widecheck{-f}) = \mu(-\hat{f}) = -\mu(\hat{f})$. Hence, $-\mu(\hat{f}) \leqslant \nu(-f)$, which implies $\mu(\hat{f}) \geqslant \nu(f)$. Thus, for each $f \in C_{\mathbb{R}}(D)$, $\nu(f) \leqslant \mu(\hat{f})$.

Finally, we prove (iii) implies (ii). Assume for each $f \in C_{\mathbb{R}}(D)$, $\nu(f) \leqslant \mu(\hat{f})$. Hence, $\nu(-f) \leqslant \mu(\widehat{-f})$. For each $x \in D$, consider,

$$\begin{aligned} \widehat{-f}(x) &= \inf\{g(x) : g \geqslant -f, g \in S(D)\} \\ &= \inf\{g(x) : -g \leqslant f, -g \in -S(D)\} \\ &= -\sup\{-g(x) : -g \leqslant f, -g \in S(D)\} \\ &= -\check{f}(x). \end{aligned} \tag{4.4}$$

This implies $\mu(\widehat{-f}) = \mu(-\check{f}) = -\mu(\check{f})$. Hence, $-\mu(\check{f}) \geqslant \nu(-f)$, which implies $\mu(\check{f}) \leqslant \nu(f)$. Thus, for each $f \in C_{\mathbb{R}}(D)$, $\mu(\check{f}) \leqslant \nu(f)$. □

We will now prove one of the main tools used in this theory.

Proposition 4.1.10. Let $f \in C_{\mathbb{R}}(D)$ and $\mu \in \mathscr{P}(D)$. Then, there exists $\nu \in \mathscr{P}(D)$ such that $\mu \prec \nu$ and $\nu(f) = \mu(\hat{f})$.

Proof. Let $\mu \in \mathscr{P}(D)$. Define for every $g \in C_{\mathbb{R}}(D)$, $\Phi(g) = \mu(\hat{g})$. Then, for any $g, h \in C_{\mathbb{R}}(D)$,

$$\Phi(g+h) = \mu(\widehat{g+h}) \leqslant \mu(\hat{g} + \hat{h}) = \mu(\hat{g}) + \mu(\hat{h}) = \Phi(g) + \Phi(h).$$

Note that the second inequality is due to the fact that the map $g \mapsto \hat{g}$ is sublinear. This implies Φ is sublinear functional on $C_{\mathbb{R}}(D)$. Let $f \in C_{\mathbb{R}}(D)$. Consider the subspace $Y = \{\alpha f : \alpha \in \mathbb{R}\} \subset C_{\mathbb{R}}(D)$. Define the map ν_0 on Y as $\nu_0(\alpha f) = \alpha\mu(\hat{f})$, for all $\alpha \in \mathbb{R}$. Clearly, ν_0 is linear.

Let $\alpha \geqslant 0$. Then, $\nu_0(\alpha f) = \alpha\mu(\hat{f}) = \mu(\widehat{\alpha f}) = \Phi(\alpha f)$. Let $\alpha < 0$. Then, write $\alpha = -\beta$, for some $\beta > 0$. Now,

$$0 = \mu(0) = \mu(\widehat{\beta f - \beta f}) \leqslant \mu(\widehat{\beta f}) + \mu(\widehat{-\beta f}).$$

That is, $-\mu(\widehat{\beta f}) \leqslant \mu(\widehat{-\beta f})$. Hence,

$$\nu_0(\alpha f) = -\beta\mu(\hat{f}) = -\mu(\widehat{\beta f}) \leqslant \mu(\widehat{-\beta f}) = \mu(\widehat{\alpha f}) = \Phi(\alpha f).$$

Therefore, by Hahn Banach theorem, there exists a linear functional ν on $C_{\mathbb{R}}(D)$ which extends ν_0 with $\nu \leqslant \Phi$. Hence, $\nu(f) = \mu(\hat{f})$ and for each $g \in C_{\mathbb{R}}(D)$, $\nu(g) \leqslant \Phi(g) = \mu(\hat{g})$. Hence, by proposition 4.1.9, $\mu \prec \nu$.

Now, it remains to show that $\nu \in \mathscr{P}(D)$. We have, $\nu(1) \leqslant \Phi(1) = \mu(\hat{1}) = \mu(1) = 1$ and since $\mu \prec \nu$, $1 = \mu(1) \leqslant \nu(1)$. This implies $\nu(1) = 1$. Hence, $\nu \in \mathscr{P}(D)$. This completes the proof. □

Our ground work for the Choquet Theorem for non metrizable case is complete. Before going to the main result we prove a very useful characterization of extreme points. We need the following Proposition in this context.

Proposition 4.1.11. Let K be a compact Hausdorff space such that $C_{\mathbb{R}}(K)$ is separable. Let f be an upper semi continuous function on K, then there exists a sequence $(g_n) \subseteq C_{\mathbb{R}}(K)$ such that $g_n \geqslant g_{n+1}$ and $\lim_n g_n(t) = f(t)$, for all t.

Proof. Case 1: When $f > 0$.

Fix $x_0 \in K$ aribitrarily and let $\alpha > f(x_0)$. Let $U = \{t \in K : f(t) < \alpha\}$, then $U \subseteq K$ is open and $x_0 \in U$. By Urysohn's lemma there exists $g \in C_{\mathbb{R}}(K)$ such that $g(U^c) = \{0\}$ and $g(x_0) = \alpha$. Thus $f(x_0) = \inf\{h(x_0) : h \in C_{\mathbb{R}}(K), h > f\}$.

Consider the subset $\mathcal{C}$ in $C_{\mathbb{R}}(K)$ consisting of all positive functions h where $h > f$, as $C_{\mathbb{R}}(K)$ is separable, $\mathcal{C}$ is also separable. Let $(p_n)_{n=1}^{\infty} \subseteq \mathcal{C}$ be a dense subset of $\mathcal{C}$. Then,

$$\begin{aligned} f(x_0) &= \inf\{h(x_0) : h \in C_{\mathbb{R}}(K), h > f\} \\ &= \inf_n \{p_n(x_0)\}. \end{aligned}$$

Define $g_n = \min_{1 \leqslant i \leqslant n} p_i$, then $(g_n) \subseteq C_{\mathbb{R}}(K)$, $g_n \geqslant g_{n+1}$ and finally, $f(x_0) = \inf_n p_n(x_0) = \lim_n g_n(x_0)$. Since $x_0 \in K$ is arbitrary we are through.

Case 2: When $f \neq 0$.

Then $f = f^+ - f^-$, where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$.

For any real scalar r with $f^+(x_0) < r$, it is clear $f(x_0) < r$ and hence there exists an open V containing x_0 such that $V \subseteq \{t \in K : f(t) < r\} \subseteq \{t \in K : f^+(t) < r\}$. On the other hand if $f^-(x_0) > r$ then $-f(x_0) > r$, which is a lower semi continuous, hence there exists an open W containing x_0 such that $W \subseteq \{t \in K : -f(t) > r\} \subseteq \{t \in K : f^-(t) > r\}$.

Hence there exist sequences $(g_n), (h_n) \subseteq C_{\mathbb{R}}(K)$ such that $g_n \downarrow f^+$ and $h_n \uparrow f^-$. Finally, $g_n - h_n \downarrow f$, this completes the proof. □

Theorem 4.1.12. Let x be a point in compact convex set D. Then, the following statements are equivalent.

(a) $x \in ext(D)$.

(*b*) $f(x) = \hat{f}(x)$, for each $f \in C_{\mathbb{R}}(D)$.

(*c*) $f(x) = \hat{f}(x)$, for each usc real function $f : D \to [-\infty, \infty)$.

One may replace $\hat{f}$ by $\check{f}$ in (*b*) and in (*c*) usc by lsc.

Proof. We prove $(a) \Rightarrow (b)$. Let $x \in ext(D)$. Then, δ_x is the unique probability measure which represents x. Let $f \in C_{\mathbb{R}}(D)$. By proposition 4.1.10, there exists $\nu \in \mathscr{P}(D)$ such that $\delta_x \prec \nu$ and $\nu(f) = \delta_x(\hat{f})$. This implies $\delta_x = \nu$ and hence, $\delta_x(f) = \delta_x(\hat{f})$. Therefore, $f(x) = \hat{f}(x)$.

Now, we prove $(b) \Rightarrow (a)$. Assume for each $f \in C_{\mathbb{R}}(D)$, $f(x) = \hat{f}(x)$. We know, $x = r(\delta_x)$. Let $\delta_x \prec \mu$. Hence, by proposition 4.1.9, for each $f \in C_{\mathbb{R}}(D)$,

$$\mu(f) \leqslant \delta_x(\hat{f}) = \hat{f}(x) = f(x) = \delta_x(f).$$

This implies for all $f \in S(D)$, $\mu(f) \leqslant \delta_x(f)$. Hence, $\mu \prec \delta_x$. Therefore, $\mu = \delta_x$. Thus, δ_x is the unique probability measure which represents x. This implies $x \in ext(D)$.

Clearly, $(c) \Rightarrow (b)$ since $C_{\mathbb{R}}(D)$ is contained in the set of all usc real valued functions on D. It remains to prove $(b) \Rightarrow (c)$. Assume for each $f \in C_{\mathbb{R}}(D)$, $f(x) = \hat{f}(x)$. Let f be a bounded usc real function on D. Then, by Proposition 4.1.11 there exists a decreasing net $f_\alpha \subset C_{\mathbb{R}}(D)$ such that $f(x) = \inf_\alpha f_\alpha(x) = \inf_\alpha \hat{f}_\alpha(x)$, using (*b*). Also, $\inf_\alpha \hat{f}_\alpha$ is an usc and concave function. Hence f is usc and concave. Therefore, by property 3.1.9 (*d*), $f = \hat{f}$. This completes the proof. □

We next prove an important characterisation of maximal measures.

Theorem 4.1.13 (Mokobodzki Theorem)**.** Let $\mu \in M^+(D)$. Then, the following statements are equivalent.

(*a*) μ is maximal in $M^+(D)$ w.r.t. ordering $\prec$ defined as above.

(*b*) $\mu(\hat{f}) = \mu(f)$, for each $f \in C_{\mathbb{R}}(D)$.

(*c*) $\mu(\hat{f}) = \mu(f)$, for each $f \in S(D)$.

One may replace $\hat{f}$ by $\check{f}$ in (*b*) and in (*c*) $S(D)$ by $-S(D)$.

Proof. We prove (*a*) implies (*b*). Let $f \in C_{\mathbb{R}}(D)$. Then, by proposition 4.1.10 exists $\nu \in \mathscr{P}(D)$ such that $\mu \prec \nu$ and $\nu(f) = \mu(\hat{f})$. Since μ is maximal measure, $\nu = \mu$ and so $\mu(f) = \mu(\hat{f})$. Now, (*b*) implies (*c*) is obvious. We prove (*c*) implies (*a*). Let $\nu \in M^+(D)$ such that $\mu \prec \nu$.Let $f \in S(D)$, then $f = \check{f}$. By proposition 4.1.9, it follows,

$$\mu(f) = \mu(\check{f}) \leqslant \nu(f) \leqslant \mu(\hat{f}) = \mu(f).$$

This implies $\mu = \nu$ on $S(D)$. Since $S(D) - S(D)$ is uniformly dense in $C_{\mathbb{R}}(D)$ we have $\mu = \nu$ on $C_{\mathbb{R}}(D)$. □

Definition 4.1.14. A measure μ on a compact convex set D is said to be a boundary measure if $|\mu|$, the total variation measure associated with μ, satisfies one of the three equivalent conditions of Mokobodzki Theorem 4.1.13

For each $f \in C_{\mathbb{R}}(D)$, we define the following set,

$$B_f = \{x \in D : f(x) = \hat{f}(x)\}.$$

Clearly,

$$B_f = \bigcap_{n=1}^{\infty} \left\{x \in D : \hat{f}(x) - f(x) < \frac{1}{n}\right\}$$

and it follows that B_f is a G_δ set.

Note that from the characterisation of extreme points given in proposition 4.1.11, it follows that,

$$ext(D) = \bigcap_{f \in C_{\mathbb{R}}(D)} B_f.$$

It is clear from statement (b) of Mokobodzki's theorem that μ is a boundary measure if and only if $|\mu|(D \backslash B_f) = 0$, for all $f \in C_{\mathbb{R}}(D)$. Since on a metrizable compact convex set D, there exists a strictly convex continuous function f and hence $ext(D) = B_f$ and thus μ is a boundary measure on a metrizable compact convex set if and only if $|\mu|(D \backslash ext(D)) = 0$.

We need one more result before establishing Choquet's theorem for non-metrizable compact convex sets.

Lemma 4.1.15. Let D be a compact convex subset of a lctvs X. Let $f = \limsup_n f_n$ where $\{f_n\}$ is a bounded above sequence from $\check{S}(D)$. If $f(x) \leqslant \alpha$ on $ext(D)$, then $f(x) \leqslant \alpha$ on D.

Proof. Let $x \in D$ be arbitrary. By Lemma 4.1.5, there exists a sequence $\{a_n\}$ from $A_{\mathbb{R}}(X)$ such that $a_n \leqslant f_n$ and $f_n(x) < a_n(x) + \frac{1}{n}$, for $n = 1, 2, ...$ Define $\Phi : X \to \mathbb{R}^\infty = \prod_{n=1}^{\infty} \mathbb{R}$ as $y \mapsto \{a_n(y)\}$. Φ is continuous on X, since for any open set U of $\mathbb{R}^\infty$, $\Phi^{-1}(U)$ is intersection of inverse images of open sets of $\mathbb{R}$ under some finite numer of a_n. Also, Φ is affine function on E as for each $n = 1.2...$, a_n is affine on E since $A_{\mathbb{R}}(D)$ is set of all real valued continuous affine functions on E which are restricted to D. Φ maps D onto a metrizable compact convex set say D'. Since Φ is continuous affine function and D is compact, $\Phi(D)$ is compact and convex. Now, since for each $n = 1, 2, ...$, a_n is real valued continuous affine function, $a_n(D)$ is a closed interval say $[a_n, b_n]$. Therefore, $D' = \prod_{n=1}^{\infty}[a_n, b_n] \subset \ell_\infty$. Hence, D' is metrizable.

For every $n = 1, 2, \ldots$, denote the n^{th} canonical projection in $\mathbb{R}^\infty$ as p_n. Then, p_n is continuous linear functional on $\mathbb{R}^\infty$ and $p_n \circ \Phi = a_n$, for $n = 1, 2, \ldots$.

We now claim that

$$\limsup_n p_n(y') \leqslant \alpha$$

for all $y' \in ext(D')$. We first prove that $D \cap \Phi^{-1}(\{y'\})$ is closed face of D. $D \cap \Phi^{-1}(\{y'\})$ is closed since it is intersection of two closed sets in X. Let $y, z \in D$ and $\lambda \in (0, 1)$ such that $\lambda y + (1 - \lambda)z \in D \cap \Phi^{-1}(\{y'\}$. Then, $\Phi(\lambda y + (1 - \lambda)z) = y'$. This implies $\lambda\Phi(y) + (1 - \lambda)\Phi(z) = y'$. Since $y' \in ext(D)$, $\Phi(y) = y'$ and $\Phi(z) = y'$, hence proving that $D \cap \Phi^{-1}(\{y'\})$ is a face of D. Now, by Krein Milman theorem, $ext(D \cap \Phi^{-1}(\{y'\})) \neq \phi$. Let $y \in ext(D \cap \Phi^{-1}(\{y'\}))$. Also $ext(D \cap \Phi^{-1}(\{y'\})) \subset ext(D)$. Therefore, $y' = \Phi(y)$ and so by hypothesis, we get

$$\limsup_n p_n(y') = \limsup_n p_n(\Phi(y)) = \limsup_n a_n(y) \leqslant f(y) \leqslant \alpha.$$

By the metrizable version of Choquet's theorem, there exists $\mu \in M^+(D')$ which represents the point $x' = \Phi(x) \in D'$ and for which $\mu(D' \backslash ext(D')) = 0$.

The sequence $\{p_n\}$ is bounded above on D' since $\{a_n\}$ is bounded above on D. Hence, by Fatou's lemma for superior limits,

$$\begin{aligned} f(x) &= \limsup_n f_n(x) \\ &= \limsup_n a_n(x) \\ &= \limsup_n p_n(x') \\ &= \limsup_n \int_{ext(D')} p_n d\mu \\ &\leqslant \int_{ext(D')} \limsup_n p_n d\mu \\ &\leqslant \alpha. \end{aligned} \tag{4.5}$$

Hence, $f(x) \leqslant \alpha$ on D. □

Theorem 4.1.16 (Choquet-Bishop-deLeeuw)**.** If μ is a boundary measure on D, then $|\mu|(C) = 0$, for every Baire set C disjoint from $ext(D)$.

Proof. We know that if $\lambda \geqslant 0$ is a regular Borel measure, then for any Baire set B, $\lambda(B) = \sup\{\lambda(C) : C \subset B, C \text{ is a compact } G_\delta \text{ set}\}$. Hence, it suffices to assume that C

is a G_δ compact set disjoint from $ext(D)$. Let

$$C = \bigcap_{n=1}^{\infty} U_n$$

where U_n are open sets. By Urysohn's lemma, there exists a bounded sequence $\{f_n\}$ from $C_{\mathbb{R}}(D)$ such that $f_n(x) = 1$, for $x \in C$, for each $n \in \mathbb{N}$ and $\lim_n f_n(x) = 0$, for $x \in D \backslash C$, that is, $\lim_n f_n = \chi_C$. Clearly, $|\mu|(f_n) \geqslant |\mu|(C)$, for $n \in \mathbb{N}$. Now, since for each $n \in \mathbb{N}$, $\check{f}_n \leqslant f_n$, for each $x \in ext(D)$,

$$\limsup_{n\to\infty} \check{f}_n \leqslant 0.$$

Consequently, by Fatou's lemma and Mokobodzki theorem,

$$\begin{aligned} 0 &\leqslant |\mu(C)| \\ &\leqslant \limsup_{n\to\infty} |\mu|(f_n) \\ &= \limsup_{n\to\infty} |\mu|(\check{f}_n) \\ &\leqslant |\mu|(\limsup_{n\to\infty} \check{f}_n) \\ &\leqslant 0. \end{aligned} \tag{4.6}$$

Hence, $|\mu|(C) = 0$. □

4.1.1 Dual of $A_{\mathbb{R}}(D)$

We claim that, $(A_{\mathbb{R}}(D))^* \cong \{\mu \in M(D) : |\mu| \text{ is maximal }\}$. In fact if $F \in (A_{\mathbb{R}}(D))^*$ then extend F to $C(D)^*$ by Hahn-Banach Theorem. Then $\tilde{F} \in C(D)^*$ is a signed measure say μ on D, $\mu = \mu^+ - \mu^-$. Clearly the positive measures μ^+, μ^- are dominated by maximal measures say ν_1, ν_2 with same resultants, by proposition 4.1.2. Hence $\nu_1 - \nu_2$ is a boundary measure on D and for every $a \in A_{\mathbb{R}}(D)$, $\mu(a) = \nu(a)$. $F \mapsto \nu$ is an isometric isomorphism.

4.2 Some applications

We present two non-trivial results which are interesting applications of Choquet-Bishop-deLeeuw theorem.

Theorem 4.2.1 (Rainwater)**.** Let E be a normed linear space and suppose that x, x_n, for $n = 1, 2, ...$ are elements of E. Then, the sequence (x_n) converges to x weakly if and only if (x_n) is bounded and $\lim_n f(x_n) = f(x)$, for all $f \in ext(U)$, where U is closed unit ball of E^*

Proof. Let $J : E \to E^{**}$ be defined as for each $f \in E^*$, $J(x)(f) = f(x)$, which is a linear isometry. Assume that (x_n) converges to x in E weakly. Then for each $f \in E^*$, $\{J(x_n)(f) : n \in \mathbb{N}\}$ is bounded and hence by uniform bounded principle, $\{J(x_n) : n \in \mathbb{N}\}$ is bounded in norm. Since J is an isometry, (x_n) is bounded in norm.

We now prove the converse. Assume that (x_n) is bounded, that is, $(J(x_n))$ is bounded in norm and also for all $f \in ext(U)$,

$$f(x_n) = J(x_n)(f) \to J(x)(f) = f(x).$$

Let g be an arbitrary element of U. It suffices to show that $J(x_n)(g) \to J(x)(g)$. In the weak* topology on E^*, by Banach Alaoglu theorem, U is compact and also, it is convex. Hence, there exists $\mu \in \mathscr{P}(U)$ such that $g = r(\mu)$, support of μ is contained in $ext(D)$ and also for every weak* continuous affine function Φ on U,

$$\Phi(g) = \int_U \Phi d\mu.$$

In particular, for each $n \in \mathbb{N}$,

$$J(x_n)(g) = g(x_n) = \int_U J(x_n) d\mu = \int_{ext(U)} J(x_n) d\mu$$

and

$$J(x)(g) = g(x) = \int_U J(x) d\mu = \int_{ext(U)} J(x) d\mu.$$

Furthermore, by assumption, $J(x_n)$ converges to $J(x)$ on U a.e. $[\mu]$. Hence, by dominated convergence theorem, $\int_{ext(U)} J(x_n) \to \int_{ext(U)} J(x)$.

This implies $g(x_n) \to g(x)$. □

The second application deals with arbitrary Banach spaces. It is clear that if K is a w^*-compact, convex, norm separable subset of E^*, for a real Banach space E, then $ext(K)$ is also norm separable and hence it is natural to ask what are the remaining points in K, not in $\overline{conv}(ext(K))$.

Theorem 4.2.2 (Haydon)**.** Let E be a real Banach space and K be a weak* compact convex subset of E^* such that $ext(K)$ is norm separable. Then K is the norm closed convex hull of its extreme points(and hence is itself norm separable).

Proof. Let $M = \sup\{\|f\| : f \in K\}$. Let $\epsilon > 0$ and $\{f_n\}$ be a norm dense subset of $ext(K)$. For each $n \in \mathbb{N}$, let B_n be the intersection of K and the closed ball of radius $\frac{\epsilon}{3}$ centered at f_n. Since K is weak* compact and by Banach Alaoglu theorem, the closed ball of radius

$\frac{\epsilon}{3}$ centered at f_n is also weak* compact, both the sets are weak* closed and hence their intersection is weak* compact, as it is contained in K. Also, this intersection is convex and

$$\bigcup_n B_n \supset ext(K).$$

Let $f \in K$. Then, there exists $\mu \in \mathscr{P}(K)$ such that μ is maximal and $r(\mu) = f$. Since $\bigcup_n B_n$ is a weak* F_σ set, $\mu(\bigcup_n B_n) = \mu(ext(K)) = 1$. Let $n \in \mathbb{N}$ and $D \stackrel{def}{=} \bigcup_n B_n$. Then, $\mu(D) > 1 - \frac{\epsilon}{3M}$. Then, μ can be decomposed into $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$, where $\lambda = \mu(D)$ and μ_1, μ_2 are probability measures on K defined as follows:

$$\lambda\mu_1 = \mu|_D \text{ and } (1-\lambda)(\mu_2) = \mu|_{K\setminus D}.$$

Note that if $\lambda = 1$ then μ_2 is an arbitrary probability measure on K. Then,

$$f = r(\mu) = \lambda r(\mu_1) + (1-\lambda) r(\mu_2).$$

Since $r(\mu_2) \in K$,

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

Since μ_1 is a probability measure supported on D,

$$r(\mu_1) \in \overline{conv}((B_i)_{i=1}^n) = conv((B_i)_{i=1}^n).$$

Hence, there exists $g_i \in B_i$ and $\lambda_i \geqslant 0$ with $\sum_{i=1}^n \lambda_i = 1$, for $1 \leqslant i \leqslant n$ such that

$$r(\mu_1) = \sum_{i=1}^n \lambda_i g_i.$$

Let $h = \sum_{i=1}^n \lambda_i f_i$. Then, $h \in conv(ext(K))$. Also,

$$\|r(\mu_1) - h\| = \|\sum_{i=1}^n \lambda_i g_i - \sum_{i=1}^n \lambda_i f_i\| \leqslant \sum_{i=1}^n (\lambda_i \|g_i - f_i\|) < 1.\frac{\epsilon}{3}.$$

Consequently,

$$\|f - h\| \leqslant \|f - r(\mu_1)\| + (1-\lambda)\|r(\mu_2)\| + \|r(\mu_2) - h\| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

This implies $h \in \overline{conv}(ext(K))$. Therefore, $conv(ext(K))$ is norm dense in K. □

We give two examples, one where the assumption in the above Theorem is true and the other where it is not true.

Example 4.2.3. Let $K = B_{\ell_1}$, then $ext(K) = \{\pm e_n : n \geqslant 1\}$. It is obvious that $B_{\ell_1} = \overline{conv}^{\|\cdot\|}(ext(K))$.

Example 4.2.4. Let $K = B_{M[0,1]}$. Let λ be the Lebesgue measure on $[0,1]$ and it is clear that $\lambda \notin \overline{conv}^{\|\cdot\|}(ext(K))$. Since (by Mazur's Theorem) $\overline{conv}^{\|\cdot\|}(ext(K)) = \overline{conv}^{w}(ext(K))$, it remains to prove that $\lambda \notin \overline{conv}^{w}(ext(K))$. In fact for any finitely many points $(t_i)_{i=1}^n \subset K$ there exists a continuous $f \in C[0,1]$ such that $\int_{[0,1]} f d\lambda = 1$ but $f(t_i) = 0, 1 \leqslant i \leqslant n$.

Our next result ensures that a net convergence is always possible to any point in K with a very special restriction. Let us recall Remark 4.1.3 and the figure therein. The point x is the resultant of finitely many measures, where the maximal measure is supported only on the extreme points of K.

Proposition 4.2.5. Let E be a lctvs and X be a compact convex subset of E. If $\mu \in \mathscr{P}(X)$, then there exists a net $(\mu_i) \in \mathscr{P}(X)$ with each μ_i discrete($S(\mu_i)$ is discrete), $r(\mu_i) = r(\mu)$ and $\mu_i \xrightarrow{w^*} \mu$.

Proof. Let $\mathscr{U}_0$ be a finite covering of X by closed convex neighbourhoods and let $\{g_k\} \subset C_{\mathbb{R}}(X)$ be a finite subordinate partition of unity, that is, $\mathscr{U}_0 = (U_1, ..., U_n)$ with $X = U_1 \cup ... \cup U_n$; $supp(g_k) \subset U_k$, for each $k = 1, 2, ..., n$ and $1 = \sum_{i=1}^n g_k$. Define

$$d\nu_k = \begin{cases} \frac{g_k}{\mu(g_k)} d\mu & \text{if } \mu(g_k) \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

Then, $\nu_k \in \mathscr{P}(X)$. Let $r(\nu_k) = x_k$. Define

$$\mu_{\mathscr{U}} = \sum_k \mu(g_k)\delta_{x_k}.$$

Clearly, $\mu_{\mathscr{U}} \in \mathscr{P}(X)$, since

$$\sum_{k=1}^n \mu(g_k) = \sum_{k=1}^n \int_X g_k d\mu = \int_X \sum_{k=1}^n g_k = 1.$$

Let P be the set of all finite partitions of X with closed convex subsets of X. Define the partial ordering $\prec$ where for $\mathscr{U}, \mathscr{V} \in P$, if $\mathscr{U} \prec \mathscr{V}$ then $\mathscr{U} \subset \mathscr{V}$, that is, $\mathscr{V}$ is a refinement of $\mathscr{U}$. Then, $(P, \prec)$ is a directed set and $\mathscr{U} \mapsto \mu_{\mathscr{U}}$ is a net. If $supp(g_k) \subset U \in \mathscr{U}_0$, then

$x_k \in U$ since $S(\nu_k) \subset supp(g_k) \subset U$. If $l \in E^*$, then

$$\begin{aligned}
\mu_{\mathscr{U}}(l) &= \sum_{k=1}^{n} \mu(g_k)\delta_{x_k}(l) \\
&= \sum_{k=1}^{n} \mu(g_k) l(x_k) \\
&= \sum_{k=1}^{n} \mu(g_k)\nu_{x_k}(l) \\
&= \sum_{k=1}^{n} \mu(g_k)\frac{1}{\mu(g_k)}\int_X l g_k d\mu \\
&= \sum_{k=1}^{n} \mu(l g_k) \\
&= \mu(l).
\end{aligned} \tag{4.8}$$

This implies $r(\mu_{\mathscr{U}}) = r(\mu)$.

Let $f \in C_{\mathbb{R}}(X)$ and $\epsilon > 0$. Choose a covering $\mathscr{U}_0$ such that for all $U \in \mathscr{U}$, for each $x, y \in U$, $|f(x) - f(y)| < \epsilon$. Then,

$$\begin{aligned}
|\mu_{\mathscr{U}}(f) - \mu(f)| &\leqslant \sum_{k=1}^{n} |\mu(g_k)f(x_k) - \mu(g_k f)| \\
&\leqslant \sum_{k=1}^{n} |\mu((g_k)(f(x_k) - f))| \\
&\leqslant \sum_{k=1}^{n} \mu(g_k)\|f - f(x_k)\|_{supp(g_k)} \\
&< \sum_{k=1}^{n} \epsilon\mu(g_k) \\
&= \epsilon.
\end{aligned} \tag{4.9}$$

This implies $\mu_{\mathscr{U}} \xrightarrow{w^*} \mu$. □

We will now give a proof of Stone Weierstrass theorem using Krein Milman and Choquet theorems.

Lemma 4.2.6. Let T be a compact Hausdorff space and $\mathscr{A}$ be a subalgebra of $C(T)$ with $K = \{\mu \in \mathscr{A}^{\perp} : \|\mu\| \leqslant 1\}$. If $\mu \in ext(K)$ and f is a real valued function in $\mathscr{A}$ such that $0 < f < 1$, then f is constant on $S(\mu)$.

Proof. If $\mu = 0$, $S(\mu) = \phi$, then the assertion is trivially true. If $\mu \neq 0$, then $\|\mu\| = 1$. Define regular complex Borel measure ν and λ by $\nu = f d\mu$ and $\lambda = (1 - f)d\mu$. Since $\mathscr{A}$

is an algebra, it follows that $\nu, \lambda \in \mathscr{A}^{\perp}$. Also, ν and λ are non-zero because $0 < f < 1$. Now,

$$\mu = \|\nu\| \frac{\nu}{\|\nu\|} + \|\lambda\| \frac{\lambda}{\|\lambda\|}$$

is a convex combination of elements of K as

$$\|\nu\| + \|\lambda\| = \int_T f d|\mu| + \int_T (1-f) d|\mu| = |\mu|(T) = \|\mu\| = 1.$$

μ being an extreme point, we must have $\mu = \frac{\nu}{\|\nu\|}$ and so $\nu = \|\nu\|\mu$. Hence, for every Borel set E,

$$\nu(E) = \int_E f d\mu = \int_E \|\nu\| d\mu.$$

This implies $f(t) = \|\nu\|$ a.e. $[\mu]$. Since f is continuous, $f(t) = \|\nu\|$ on $S(\mu)$. □

Theorem 4.2.7 (Stone Weierstrass theorem)**.** Let T be a compact Hausdorff space and $\mathscr{A}$ be a closed (that is supremum closed) subalgebra of $C(T)$ with the property that

- (a) the constant functions are in $\mathscr{A}$.
- (b) $\mathscr{A}$ separates the points of T.
- (c) if $f \in \mathscr{A}$, then $\overline{f} \in \mathscr{A}$.

Then, $\mathscr{A} = C(T)$.

Proof. Let $K = \{\mu \in \mathscr{A}^{\perp} : \|\mu\| \leqslant 1\}$. Then, $K \neq \phi$(in fact $K = \mathscr{A}^{\perp} \cap C(T)^*$). Also, K is convex and weak* compact (as it is the intersection of a weak* closed convex set with the weak* compact set $C(T)^*$). Hence, by Krien-Milman theorem, there exists $\mu \in ext(K)$. Suppose $S(\mu)$ contains two distinct points s and t. From the properties of $\mathscr{A}$, it follows that $\mathscr{A}$ contains a real valued function f such that $0 < f < 1$ and $f(s) \neq f(t)$. But this is impossible by the previous lemma. Hence, $\mu = \alpha\delta_t$, for some $\alpha \in \mathbb{C}$ and $t \in K$ and for every $f \in \mathscr{A}$,

$$\int_T f d\mu = \alpha f(t).$$

But $\mu \in \mathscr{A}^{\perp}$, so $\alpha f(t) = 0$, for $f \in \mathscr{A}$. Since $1 \in \mathscr{A}$, this means $\alpha = 0$ and hence $\mu = 0$, which shows that the only extreme point of K must be the zero measure and hence, by Krein-Milman theorem, $K = \{0\}$ and $\mathscr{A}^{\perp} = \{0\}$. Then, Hahn Banach theorem implies that $\mathscr{A} = C(T)$, as $\mathscr{A}$ is closed by assumption. □

Chapter 5

Choquet Simplex and its Characterizations

In this chapter, we will investigate the characterisations of a simplex as given by Choquet-Meyer. Let E be a real lctvs. Let P be a cone in E. Then we know that P induces a translation invariant partial ordering on E: $x \prec y$ if and only if $y - x \in P$. Examples of cone are $M^+(K)$, where K is a compact Hausdorff space, the set of all positive operators on a Hilbert space H, etc.

Example 5.0.1. Let K be a compact Hausdorff space. Consider $M^+(K)$ which is the set of all non-negative regular Borel measures on K. We know that $M^+(K)$ is a cone.

We now claim that $M^+(K)$ forms a lattice with the ordering, say $\prec$, induced by $M^+(K)$ as a cone. Let $\lambda, \mu \in M^+(K)$. Given any Borel set $A \subset K$, $(\lambda + \mu)(A) = 0$ implies $\lambda(A) + \mu(A) = 0$ and hence $\lambda(A) = \mu(A) = 0$. This implies that λ and μ are absolutely continuous with respect to $\lambda + \mu$ and hence have Radon-Nikodym derivatives say f and g respectively. Let $h = \min(f, g)$, which is defined a.e. $\lambda + \mu$. We claim that $\lambda \wedge \mu = hd(\lambda + \mu)$. Now, $\lambda - hd(\lambda + \mu) = (f - h)d(\lambda + \mu) \in M^+(K)$, since $f \geqslant \min(f, g)$. Similarly, $\mu - hd(\lambda + \mu) \in M^+(K)$. Let $\nu \in M^+(K)$ such that $\nu \prec \lambda$ and $\nu \prec \mu$. Hence, $\lambda - \nu, \mu - \nu \in M^+(K)$. Given any Borel set $A \subset K$, $(\lambda + \mu)(A) = 0$ implies $\lambda(A) = \mu(A) = 0$. Now, $(\lambda - \nu)(A) = -\nu(A)$, which is possible only if $\nu(A) = 0$. It follows that ν is absolutely continuous with respect to $\lambda + \mu$ and hence have Radon-Nikodym derivative say s. We have, therefore, $(f - s)d(\lambda + \mu), (g - s)d(\lambda + \mu) \in M^+(K)$. Since $(f - s) \geqslant 0$ and $(g - s) \geqslant 0$, it follows that $(h - s) \geqslant 0$ and hence $(h - s)d(\lambda + \mu) \in M^+(K)$. Therefore $\nu \prec hd(\lambda + \mu)$. This proves that $M^+(K)$ is a lattice.

Let $0 \neq \mu \in M^+(K)$. Then clearly $\mu/\|\mu\| \in \mathscr{P}(K)$. Also, $0 \notin \mathscr{P}(K)$ and $\mathscr{P}(K)$ is contained in the hyperplane $\{\mu \in M(K) : \mu(1) = 1\}$, which does not contain 0. It follows that $\mathscr{P}(K)$ is a base of $M^+(K)$. Therefore, by Remark D.1.4 (a), $M^+(K) \cap (-M^+(K)) =$

$\{0\}$. By Remark D.1.5 (c), we also get that $M^+(K) - M^+(K)$ is a lattice. Now by Jordan measure decomposition, we have $M(K) = M^+(K) - M^+(K)$. This example serves as a motivation for the definition of a simplex which is as follows.

Definition 5.0.2 (Simplex)**.** Let D be a compact convex subset of a lctvs E, which is contained in a hyperplane H which misses origin. Then D is said to be a simplex if $\widetilde{D} - \widetilde{D}$ is a lattice in E, where $\widetilde{D}$ is the cone generated by D.

Theorem 5.0.3 (Choquet-Meyer Theorem)**.** Let X be a non-empty compact convex set in a lctvs E such that X is contained in a hyperplane which does not contain zero. Then the following are equivalent:

(i) X is a simplex in E.

(ii) For any $f \in C_{\mathbb{R}}(X)$ such that f is convex, $\hat{f}$ is affine on X.

(iii) If μ is maximal measure on X with resultant $x \in X$ and $f \in C_{\mathbb{R}}(X)$, where f is convex, then $\hat{f}(x) = \mu(f)$.

(iv) For each continuous convex functions $f, g \in C_{\mathbb{R}}(X)$, $\widehat{f+g} = \hat{f} + \hat{g}$.

(v) Each $x \in X$ is represented by a unique maximal measure on X.

Proof. (i) implies (ii): Assume X is a simplex in E. This implies there exists a hyperplane $H = \{x \in E : L(x) = r\}$, for some $L \in E^*$, $r > 0$ such that $0 \notin H$, $X \subset H$ and $\widetilde{X}$ is a lattice. Let $f \in C_{\mathbb{R}}(X)$ be a convex function on X. We know that $\hat{f}$ is concave on X. It remains to prove that $\hat{f}$ is convex. Let $x_1, x_2 \in X$ and $\alpha_1, \alpha_2 > 0$ with $\alpha_1 + \alpha_2 = 1$. Let $z = \alpha_1 x_1 + \alpha_2 x_2$. Now, from proposition 3.1.9 and 4.2.5,

$$\hat{f}(z) = \sup\{\mu(f) : \mu \in \mathscr{P}(X), r(\mu) = z\} = \sup\{\mu(f) : \mu \text{ is discrete}, r(\mu) = z\}.$$

Let us assume that μ is a discrete probability measure and $r(\mu) = z$. Then there exists finite sequences $(y_j)_{j\in J}$ and $(\beta_j)_{j\in J}$, with $\beta_j \geqslant 0$, for each $j \in J$, in X such that $\sum_{j\in J} \beta_j = 1$ and $\mu = \sum_{j\in J} \beta_j \delta_{y_j}$. Since for every $g \in E^*$

$$g(z) = \mu(g) = \sum_{j\in J} \beta_j g(y_j),$$

we get

$$z = \alpha_1 x_1 + \alpha_2 x_2 = \sum_{j\in J} \beta_j y_j.$$

Since $\widetilde{X}$ is a lattice, there exists $\{z_{ij} : i = 1, 2; j \in J\} \subset X$ such that for each $i = 1, 2$, $\alpha_i x_i = \sum_{j\in J} z'_{ij}$ and for each $j \in J$, $\beta_j y_j = \sum_{i=1}^{2} z'_{ij}$, where for each i, j, there exists $\gamma_{ij} \geqslant 0$ and $z_{ij} \in X$ such that $z'_{ij} = \gamma_{ij} z_{ij}$. Therefore, for each $i = 1, 2$, $x_i = \alpha_i^{-1} \sum_{j\in J} \gamma_{ij} z_{ij}$. Since

$x_i, z_{ij} \in X \subset H$, $L(x_i) = r = L(z_{ij})$. Thus,

$$r = L(x_i) = \alpha_i^{-1} \sum_{j \in J} \gamma_{ij} L(z_{ij}) = \alpha_i^{-1} \sum_{j \in J} \gamma_{ij} r.$$

This implies $1 = \sum_{j \in J} \alpha_i^{-1} \gamma_{ij}$. Therefore, x_i is a convex combination of elements in X. Define for $i = 1, 2$, $\mu_i = \alpha_i^{-1} \sum_{j \in J} \gamma_{ij} \delta_{z_{ij}}$. Then $r(\mu_i) = x_i$. Thus, $\hat{f}(x_i) \geqslant \mu_i(f) = \sum_{j \in J} \alpha_i^{-1} \gamma_{ij} f(z_{ij})$. Also, $\mu(f) = \sum_{j \in J} \beta_j f(y_j)$. For each $j \in J$,

$$\begin{aligned} f(y_j) &= f(\beta_j^{-1} z_{1j}^{'} + \beta_j^{-1} z_{2j}^{'}) \\ &= f(\beta_j^{-1} \gamma_{1j} z_{1j} + \beta_j^{-1} \gamma_{2j} z_{2j}) \\ &\leqslant \beta_j^{-1} \gamma_{1j} f(z_{1j}) + \beta_j^{-1} \gamma_{2j} f(z_{2j}). \end{aligned} \tag{5.1}$$

Therefore, we have

$$\begin{aligned} \sum_{j \in J} \beta_j f(y_j) &\leqslant \sum_{j \in J} \gamma_{ij} f(z_{1j}) + \gamma_{2j} f(z_{2j}) \\ &= \left(\sum_{j \in J} \gamma_{ij} f(z_{1j}) \right) + \left(\sum_{j \in J} \gamma_{2j} f(z_{2j}) \right) \\ &= \alpha_1 \mu_1(f) + \alpha_2 \mu_2(f). \end{aligned} \tag{5.2}$$

This implies

$$\mu(f) \leqslant \alpha_1 \mu_1(f) + \alpha_2 \mu_2(f) = \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

Thus $\hat{f}(z) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2) \leqslant \alpha_1 \hat{f}(x_1) + \alpha_2 \hat{f}(x_2)$. This implies $\hat{f}$ is convex.

(ii) implies (iii): Suppose μ is a maximal measure such that $r(\mu) = x$, for some $x \in X$ and $f \in C_{\mathbb{R}}(X)$ is a convex function on X. Then by (ii), $\hat{f}$ is affine. Then by the assumption that $r(\mu) = x$ and by Mokobodzki Theorem 4.1.13 we get,

$$\hat{f}(x) = \mu(\hat{f}) = \mu(f).$$

(iii) implies (iv): Let $x \in X$. Then there exists a maximal measure say μ on X and by proposition 4.1.2, $r(\mu) = x$. Let $f, g \in C_{\mathbb{R}}(X)$ such that f, g are convex. Then by (iii),

$$\widehat{f + g}(x) = \mu(f + g) = \mu(f) + \mu(g) = \hat{f}(x) + \hat{g}(x).$$

Therefore, for each $x \in X$, $\widehat{f + g}(x) = \hat{f}(x) + \hat{g}(x)$.

(iv) implies (v): Let $x \in X$. We know that $S := S(X)$ is a cone in $C_{\mathbb{R}}(X)$ and $S - S$ is a subspace of $C_{\mathbb{R}}(X)$. Define $\Lambda : S - S \to \mathbb{R}$ as $\Lambda(f - g) = \hat{f}(x) - \hat{g}(x)$. Let

$f_1 - g_1, f_2 - g_2 \in S - S$ and $\alpha \in \mathbb{R}$. Then by assuming (iv), we get,

$$\begin{aligned} \Lambda((f_1 - g_1) + (f_2 - g_2)) &= \Lambda((f_1 + f_2) - (g_1 + g_2)) \\ &= \widehat{f_1 + f_2} - \widehat{g_1 + g_2} \\ &= \hat{f}_1 + \hat{f}_2 - \hat{g}_1 - \hat{g}_2 \\ &= (\hat{f}_1 - \hat{g}_1) + (\hat{f}_2 - \hat{g}_2) \\ &= \Lambda(f_1 - g_1) + \Lambda(f_2 - g_2). \end{aligned} \tag{5.3}$$

If $\alpha \geqslant 0$, then clearly, $\Lambda(\alpha(f_1 - g_1)) = \alpha\Lambda(f_1 - g_1)$. Now, if $\alpha < 0$, then

$$\begin{aligned} \Lambda(\alpha(f_1 - g_1)) &= \Lambda((-\alpha)g_1 - (-\alpha)f_1) \\ &= (\widehat{-\alpha g_1})(x) - (\widehat{-\alpha f_1})(x) \\ &= -\alpha(\hat{g}_1(x) - \hat{f}_1(x)) \\ &= \alpha(\hat{f}_1(x) - \hat{g}_1(x)), \end{aligned} \tag{5.4}$$

since $S - S$ is a subspace. Therefore, this proves that Λ is linear on $S - S$. Now the set $T = \{\mu \in \mathscr{P}(X) : r(\mu) = x\}$ is w^*-compact in $\mathscr{P}(X)$, since we know that the function $r : \mathscr{P}(X) \to X$ defined by $r(\mu) = x$ is w^*-continuous. Therefore, $\hat{f}(x) = \sup\{\mu(f) : \mu \in T\} = \mu_0(f)$, for some $\mu_0 \in T$. Also, $\hat{g}(x) \geqslant \mu_0(g)$. This implies

$$\hat{f}(x) - \hat{g}(x) \leqslant \mu_0(f - g) \leqslant \|\mu_0\|\|f - g\| = \|f - g\|_\infty.$$

Similarly, we can get

$$\hat{g}(x) - \hat{f}(x) \leqslant \|f - g\|_\infty.$$

Hence

$$|\hat{f}(x) - \hat{g}(x)| \leqslant \|f - g\|_\infty,$$

which implies $\|\Lambda\| \leqslant 1$. Since $1 \in S - S$ and $\Lambda(1) = 1$, $\|\Lambda\| = 1$. Therefore, Λ is continuous linear functional on $S - S$. Since $S - S$ is uniformly dense subspace in $C_{\mathbb{R}}(X)$, Λ has a unique extension to $C_{\mathbb{R}}(X)$. By Riesz Representation theorem, there exists a unique $\mu_x \in M(X)$ such that for each $h \in C_{\mathbb{R}}(X)$, $\Lambda(h) = \mu_x(h)$, $\|\Lambda\| = 1 = \|\mu_x\|$ and $\mu_x(1) = \Lambda(1) = 1$. This implies $\mu_x \in \mathscr{P}(X)$. Also, $r(\mu_x) = x$. Let $f \in S(X)$. Then

$$\mu_x(f) = \Lambda(f) = \hat{f}(x) = \sup\{\nu(f) : \nu \in \mathscr{P}(X); r(\nu) = x\}.$$

Let $\nu \in M^+(X)$ such that $\mu_x \prec \nu$. This implies $r(\mu_x) = r(\nu)$. Hence,

$$\hat{f}(x) = \mu_x(f) \leqslant \nu(f) \leqslant \hat{f}(x).$$

This implies for each $f \in S(X)$, $\mu_x(f) = \nu(f)$. It follows that for $f, g \in S(X)$, then

$\mu_x(f-g) = \nu(f-g)$. Again since $S - S$ is dense in $C_{\mathbb{R}}(X)$, for each $f \in C_{\mathbb{R}}(X)$, $\mu_x(f) = \nu(f)$. This implies $\mu_x = \nu$. Hence μ_x is a maximal measure representing x. Suppose λ is another maximal probability measure such that $r(\lambda) = x$. For each $f \in S$, $\mu_x(f) = \hat{f}(x) \geqslant \lambda(f)$. This implies $\lambda \prec \mu_x$. Since λ is maximal, $\lambda = \mu_x$. This proves the uniqueness of such a maximal measure.

To prove (v) implies (i), we need the following lemmas.

Definition 5.0.4 (Hereditary subcone)**.** Let P_1, P_2 be two cones in a lctvs E such that $P_1 \subset P_2$. Then P_1 is said to be Hereditary subcone of P_2 if for $x \in P_1$, $y \in P_2$ such that $y \prec x$ (where $\prec$ is the ordering induced by P_2), then $y \in P_1$.

Lemma 5.0.5. Let P_1 be a hereditary subcone of cone P_2. If P_2 is a lattice, then P_1 is a lattice.

Proof. We assume that P_1 and P_2 are cones with vertex 0. Let $\prec_1$ and $\prec_2$ be the orderings induced by P_1 and P_2 respectively. Let $x, y \in P_1 \subset P_2$. Since P_2 is a lattice, $x \wedge y \in P_2$. Since $x \wedge y \prec_2 x$, $x \wedge y \in P_1$. Let $z \in P_1 \subset P_2$ such that $z \prec_1 x$, $z \prec_1 y$. This implies $z \prec_2 x \wedge y$. Since $0 \prec_2 z$, $x \wedge y - z \prec_2 x \wedge y$. This implies $x \wedge y - z \in P_1$. Hence $z \prec_1 x \wedge y$. This completes the proof. □

Remark 5.0.6. Let P_1 be a hereditary subcone of cone P_2. Let $x, y \in P_1$. If $x \prec_2 y$, then $x \prec_1 y$.

Lemma 5.0.7. Let X be a compact convex subset of a lctvs E. Let $\mathscr{M}$ be the set of all maximal positive measures on X. Then $\mathscr{M}$ is a hereditary subcone of $M^+(X)$ (which is a cone with vertex 0), $\mathscr{P}(X)$ is a base of $\mathscr{M}$ and $\mathscr{P}(X)$ is a simplex.

Proof. Clearly, $\mathscr{M} \subset M^+(X)$. Let $\lambda \in \mathscr{M}$ and $r > 0$. Let $\mu \in M^+(X)$ such that $r\lambda \prec \mu$. This implies for each $f \in S$, $r\lambda(f) \leqslant \mu(f)$. Since λ is maximal measure on X, it follows that $\lambda = \mu/r$ and hence $r\lambda = \mu$. This proves that $\mathscr{M}$ is a cone. In order to prove that $\mathscr{M}$ is convex, it suffices to show that for each $\lambda, \mu \in \mathscr{M}$, $\lambda + \mu \in \mathscr{M}$. Let $\lambda, \mu \in \mathscr{M}$. Using Mokobodzki Theorem, for each $f \in C_{\mathbb{R}}(X)$,

$$(\lambda + \mu)(\hat{f}) = \lambda(\hat{f}) + \mu(\hat{f}) = \lambda(f) + \mu(f) = (\lambda + \mu)(f).$$

Thus by Mokobodzki theorem, $\lambda + \mu \in \mathscr{M}$. Now, $0 \notin \{\mu \in M(X) : \mu(1) = 1\}$ is a w^*-closed hyperplane in $M(X)$ and

$$\mathscr{P}(X) = \{\mu \in M(X) : \mu(1) = 1\} \cap \mathscr{M}.$$

This implies $\mathscr{P}(X)$ is a base for $\mathscr{M}$. It remains to prove that $\mathscr{M}$ is a hereditary subcone of $M^+(X)$. Let $\lambda \in \mathscr{M}$ and $\mu \in M^+(X)$ such that $\mu \preceq \lambda$ (where $\preceq$ is the ordering induced by $M^+(X)$). Let $\mu_1 \in \mathscr{M}$ such that $\mu \prec \mu_1$. This implies for each $f \in S(X)$, $\mu(f) \leqslant \mu_1(f)$. $\mu \preceq \lambda$ implies $\lambda - \mu \in M^+(X)$. Now, $\lambda = \mu + (\lambda - \mu) \preceq \mu_1 + (\lambda - \mu)$. Since $\lambda \in \mathscr{M}$, $\lambda = \mu_1 + \lambda - \mu$ which implies $\mu_1 = \mu$. Therefore, $\mu \in \mathscr{M}$. Hence, $\mathscr{M}$ is a hereditary subcone of $M^+(X)$. □

We will now prove (v) implies (i). The function $r : \mathscr{P}(X) \to X$ defined by $r(\mu) = x$ is w^*-continuous. By assumption (v), $r^{-1} : X \to \mathscr{P}(X)$ is one-one and onto map. Also, r^{-1} is affine function. Therefore, $\mathscr{P}(X)$ and X are affinely homeomorphic. By Lemma 5.0.7, $\mathscr{P}(X)$ is a simplex. We can continuously extend r to $\widetilde{\mathscr{P}(X)}$. Let us denote the extension also as r for simplicity. Then r is one-one and $r(\widetilde{\mathscr{P}(X)}) = \widetilde{X}$. Let $tx_1, sx_2 \in \widetilde{X}$. This implies $tr^{-1}(x_1) \wedge sr^{-1}(x_2) \in \widetilde{\mathscr{P}(X)}$ and hence $r(tr^{-1}(x_1) \wedge sr^{-1}(x_2)) \in \widetilde{X}$. We now claim that $tx_1 \wedge sx_2 = r(tr^{-1}(x_1) \wedge sr^{-1}(x_2))$. Consider

$$r^{-1}(tx_1 - r(tr^{-1}(x_1) \wedge sr^{-1}(x_2))) = tr^{-1}(x_1) - tr^{-1}(x_1) \wedge sr^{-1}(x_2) \in \widetilde{\mathscr{P}(X)}.$$

Therefore,

$$tx_1 - r(tr^{-1}(x_1) \wedge sr^{-1}(x_2)) \in r(\widetilde{\mathscr{P}(X)}) = \widetilde{X}.$$

Similarly, we can prove that,

$$tx_2 - r(tr^{-1}(x_1) \wedge sr^{-1}(x_2)) \in \widetilde{X}.$$

Let $w \in \widetilde{X}$ such that $w \preceq tx_1$ and $w \preceq sx_2$. Note that the ordering $\preceq$ on $\widetilde{X}$ is induced by r and $\widetilde{\mathscr{P}(X)}$. Then

$$r^{-1}(w) \prec r^{-1}(tx_1) = tr^{-1}(x_1) \text{ and } r^{-1}(w) \prec r^{-1}(sx_2) = sr^{-1}(x_2).$$

Therefore

$$r^{-1}(w) \prec tr^{-1}(x_1) \wedge sr^{-1}(x_2).$$

and hence

$$w \preceq r(tr^{-1}(x_1) \wedge sr^{-1}(x_2)).$$

This proves that $\widetilde{X}$ is a lattice. □

We conclude this chapter with a proof that the definition of simplex coincides with the usual one for finite dimensional spaces.

Result 5.0.8. Let X be a non-empty compact convex set in a lctvs E such that X is contained in a hyperplane which does not contain zero. Suppose $\widetilde{X} - \widetilde{X}$ is of dimension n. X is the convex hull of n linearly independent vectors of E if and only if X has exactly n extreme points.

Proof. Assume $X = conv(\{x_i\}_{i=1}^n)$, where $\{x_i\}_{i=1}^n$ is a linearly independent set in E. By the converse of Krein-Milman Theorem, $ext(X) \subset \{x_i\}_{i=1}^n$. Suppose without loss of generality, $x_n \notin ext(X)$. This implies there exists $\alpha_i > 0$, for each $1 \leqslant i \leqslant n-1$, with $\sum_{i=1}^{n-1} \alpha_i = 1$ such that $x_n = \sum_{i=1}^{n-1} \alpha_i x_i$. This is a contradiction to our assumption that $\{x_i\}_{i=1}^n$ is linearly independent. Hence, X has exactly n extreme points.

Conversely, assume X has exactly n extreme points in E say $x_1, x_2, \ldots, x_n$. Therefore, $X = conv(\{x_i\}_{i=1}^n)$. Suppose $\{x_i\}_{i=1}^n$ is not linearly independent set. Without loss of generality, let $\{x_i\}_{i=1}^m$ be linearly independent set where $1 \leqslant m < n$. This implies $\widetilde{X} - \widetilde{X}$ is a linear span of $\{x_i\}_{i=1}^m$, which is a contradiction to the fact that dimension of $\widetilde{X} - \widetilde{X}$ is n. This implies that $\{x_i\}_{i=1}^n$ is linearly independent set in E. □

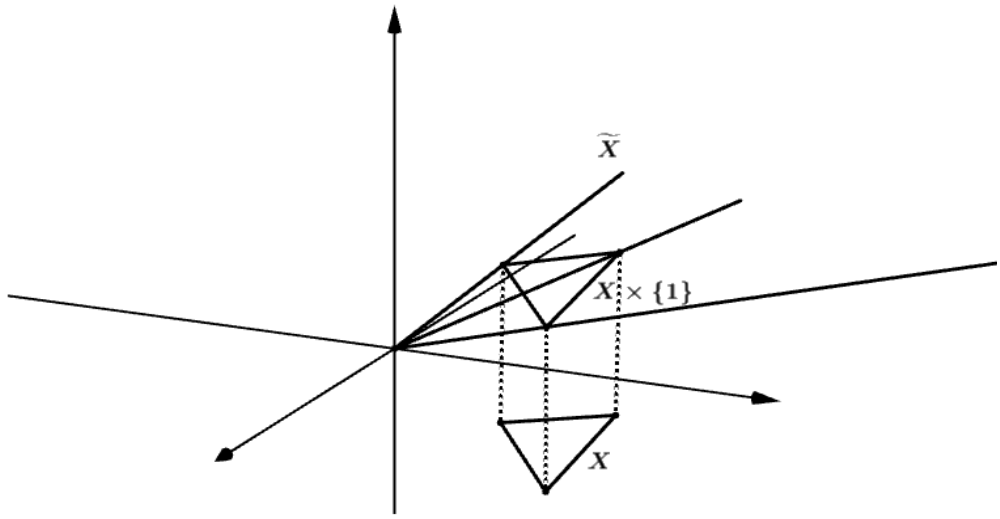

FIGURE 5.1

Example 5.0.9. Consider a triangle X in $\mathbb{R}^2$ as shown in the above figure. We embed X as $X \times \{1\}$ in the hyperplane $\mathbb{R}^2 \times \{1\}$. Clearly, it follows that $\widetilde{X}$ forms a lattice. Hence, X is a simplex. Also, observe that X is the convex hull of its three linearly independent vertices. This example leads us to the following characterization of a simplex in finite dimensional spaces.

Theorem 5.0.10. Let X be a non-empty compact convex set in a lctvs E such that X is contained in a hyperplane which does not contain zero. Suppose $\widetilde{X} - \widetilde{X}$ is of dimension n. Then X is a simplex if and only if X is a convex hull of n linearly independent vectors of E, equivalently X has exactly n extreme points.

Proof. Assume that X is a simplex. Since X is a compact convex subset of E, $X = conv(ext(X))$. X cannot have less than n extreme points, otherwise $\widetilde{X} - \widetilde{X}$ will be a span of less than n vectors which is a contradiction to the fact that $\widetilde{X} - \widetilde{X}$ is of dimension n. Suppose X has $n+1$ extreme points say $y_1, y_2, \ldots, y_{n+1}$. Then there exists α_i, $1 \leqslant i \leqslant n+1$, atleast one of α_i is non-zero, such that

$$\sum_{i=1}^{n+1} \alpha_i y_i = 0.$$

Consider the sets $N = \{\alpha : \alpha = \alpha_i$, for some i such that $\alpha_i < 0\}$ and $P = \{\alpha : \alpha = \alpha_i$ for some i such that $\alpha_i \geqslant 0\}$. Let $\alpha = \sum_{\alpha_i \in P} \alpha_i > 0$. Otherwise if $\alpha = 0$, $f(\sum_{i=1}^{n+1} \alpha_i y_i) = 0$, for a $f \in (\widetilde{X} - \widetilde{X})^*$ such that $f(y_i) = 1$, for each $1 \leqslant i \leqslant n+1$. This implies

$$\sum_{i=1}^{n+1} \alpha_i f(y_i) = \sum_{i=1}^{n+1} \alpha_i = 0.$$

Now the above sum can be written as

$$0 = \alpha + \sum_{\alpha_i \in N} \alpha_i = 0 + \sum_{\alpha_i \in N} \alpha_i = \sum_{\alpha_i \in N} \alpha_i < 0,$$

which is a contradiction. Also, $\alpha = -\sum_{\alpha_i \in N} \alpha_i > 0$. Let $x = \sum_{\alpha_i \in P} \alpha^{-1} \alpha_i y_i$, where $\sum_{\alpha_i \in P} \alpha^{-1} \alpha_i = 1$. Also, $x = \sum_{\alpha_i \in N} (-\alpha^{-1}) \alpha_i y_i$, where $\sum_{\alpha_i \in N} (-\alpha^{-1}) \alpha_i = 1$. Therefore,

$$x = r\left(\sum_{\alpha_i \in P} \alpha^{-1} \alpha_i \delta_{y_i}\right) = r\left(\sum_{\alpha_i \in N} (-\alpha^{-1}) \alpha_i \delta_{y_i}\right).$$

This implies x has two distinct representing measures which is a contradiction since every point of the simplex X is represented by a unique maximal measure on X. Therefore X has exactly n extreme points.

Conversely, let X have exactly n extreme points say $x_1, x_2, \ldots, x_n$. Then $X = conv(\{x_i\}_{i=1}^n)$. Also, $\{x_i\}_{i=1}^n$ is a linearly independent set of vectors in E. Otherwise, without loss of generality, we can remove vectors from the set $\{x_i\}_{i=1}^n$ such that $\{x_i\}_{i=1}^m$, where $1 \leqslant m < n$ is linearly independent set. Now, $\widetilde{X} - \widetilde{X}$ is a linear span of $\{x_i\}_{i=1}^m$, which is a contradiction to the fact that dimension of $\widetilde{X} - \widetilde{X}$ is n. Let $E_1 = \widetilde{X} - \widetilde{X}$ and $\{x_i\}_{i=1}^n$ forms a basis of E_1. Choose $f_1, f_2, \ldots, f_n \in E^*$ such that $f_i(x_j) = \delta_{ij}$, for each $1 \leqslant j \leqslant n$ and $1 \leqslant i \leqslant n$. Define $T : E_1 \to \mathbb{R}^n$ as $Tx = (f_1(x), f_2(x), \ldots, f_n(x))$. Clearly, T is linear and continuous on E_1. Consider $Tx = 0$. This implies $f_i(x) = 0$, for each $1 \leqslant i \leqslant n$. Now, $x = \sum_{i=1}^n \alpha_i x_i$, for some $\alpha_i \in \mathbb{R}$, for each $1 \leqslant i \leqslant n$. Hence, $f_i(x) = \alpha_i = 0$, for each i. Therefore, $x = 0$. This proves that T is one-one. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$. Then $x = \sum_{i=1}^n \alpha_i x_i \in E_1$. Evidently, $Tx = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, which proves that T is onto. Let $\{e_i\}_{i=1}^n$ be the standard basis of $\mathbb{R}^n$. Then $Tx_i = e_i$, for each $1 \leqslant i \leqslant n$. Therefore, $TX = conv(\{e_i\}_{i=1}^n)$ which is a compact convex subset of $\mathbb{R}^n$. Hence, $\widetilde{TX} = \{(\alpha_1, \alpha_2, \ldots, \alpha_n) : \alpha_i \geqslant 0, 1 \leqslant i \leqslant n\}$ and is clearly a lattice in $\mathbb{R}^n$. Therefore, TX is a simplex. It follows that X itself is a simplex. □

Chapter 6

Choquet Boundary

6.1 Definitions and basic properties

Let K be a compact Hausdorff space and $C_{\mathbb{C}}(K)$ denote the space of all complex valued continuous functions on K with the supremum norm. Let us recall some basic facts about any complex Banach space E. The dual of (E^*, w^*) is E itself, with each $x \in E$ defining a continuous linear functional $f \mapsto f(x)$ on E^*.

Definition 6.1.1 (State space)**.** Let M be a linear subspace(not necessarily closed) of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) such that $1 \in M$. Then the state space of M, denoted by $K(M)$, is defined as $K(M) = \{L \in M^* : L(1) = 1, \|L\| = 1\}$.

Consider (M^*, w^*). Then clearly $K(M)$ is non-empty(since for any $t \in K$, $\delta_t \in K(M)$) and convex set. Also, $K(M)$ is a w^*-closed subset of the w^*-compact unit ball of M^*(by Banach Alaoglu theorem) and hence $K(M)$ is w^*-compact set. When $M = C_{\mathbb{R}}(K)$, $K(M) = \mathscr{P}(K)$. Then by Krein-Milman theorem, we know that $\mathscr{P}(K) = \overline{conv}^{w^*}(ext(\mathscr{P}(K)))$. In order to use the results in previous chapters, it is necessary to know the description of extreme points of $K(M)$.

We shall now define a map ϕ from K into $K(M)$ as $k \mapsto \delta_k$, where $\delta_k(f) = f(k)$, for all $f \in M$. Let $k_\alpha \to k$ be any net in K. Then for any $f \in M$, $f(k_\alpha) \to f(k)$, which implies that $\delta_{k_\alpha} \xrightarrow{w^*} \delta_k$. Hence, ϕ is continuous map from K into $(K(M), w^*)$.

If M separates points of K, then ϕ is one-one and hence ϕ is a homeomorphism, embedding K as a compact subset of $K(M)$. If $L \in K(M)$ and μ is a measure on K such that for every $f \in M$, $L(f) = \mu(f)$, then μ can be carried to a measure μ' on $K(M)$ in the obvious way: $\mu' = \mu \circ \phi^{-1}$. Since the dual of (M^*, w^*) is M itself, it follows that μ' represents L.

Lemma 6.1. *Let M be a subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) such that $1 \in M$. Then $K(M) = \overline{conv}^{w^*}(\phi(K))$.*

Proof. Clearly, $\overline{conv}^{w^*}(\phi(K)) \subset K(M)$. Suppose there exists $L \in K(M)$ such that $L \notin \overline{conv}^{w^*}(\phi(K))$. Then by Hahn Banach separation theorem, there exists $f \in M$ such that,

$$\sup Ref(\phi(K)) = \sup\{Ref(y) : y \in K\} < Re(L(f)) \leqslant \|Ref\|\|L\| = \|Ref\| \quad (1).$$

Now, by adding a large positive constant M to both sides of the above inequality such that $g = f + M \geqslant 0$, we get

$$\|Re(g)\| = \sup\{Re(g)(y) : y \in K\} < \|Re(g)\|,$$

which is a contradiction. □

Definition 6.1.2 (Choquet Boundary)**.** Let M be a linear subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) such that $1 \in M$. Then the Choquet boundary of M, denoted by $B(M)$, is defined as $B(M) = \{k \in K : \phi(k) \in ext(K(M))\}$.

Remark 6.1.3. Let $L \in K(M)$. L is an extreme point of $K(M)$ if and only if $L = \phi(y)$, for some $y \in B(M)$.

Proof. Now, by previous lemma, $K(M) = \overline{conv}^{w^*}(\phi(K))$. Hence, by partial converse of Krein-Milman theorem, $ext(K(M)) \subset \phi(K)$. Assume $L = \phi(y)$, for some $y \in B(M)$, then by definition of $B(M)$, L is an extreme point of $K(M)$. □

Let us look at the following characterization of $B(M)$ in terms of measures on K for subspaces M that separates points of K.

Theorem 6.1.4. Let M be a subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) which separates points of K and contains the constant functions. Then $k \in B(M)$ if and only if $\mu = \delta_k$ is the only probability measure on K such that $f(k) = \int_K f d\mu$, for all $f \in M$.

Proof. Assume that $k \in B(M)$ and suppose μ is a measure on K such that $f(k) = \int_K f d\mu$, for all $f \in M$. Hence, we can carry μ to a measure μ' on $K(M)$ such that $\mu'|_{\phi(K)} = \mu \circ \phi^{-1}$. Since $\phi(k) \in ext(K(M))$, $\delta_{\phi(k)}$ is the only probability measure that represents $\phi(k)$. Using the fact that the dual of (M^*, w^*) is M itself, from the above relation, for all $f \in M$, we get the following

$$\delta_k(f) = \int_K f d\mu = \int_{K(M)} f d\mu' = \mu'(f).$$

This implies

$$\mu \circ \phi^{-1} = \mu' = \delta_{\phi(k)} = \delta_k \circ \phi^{-1}.$$

Since ϕ is a homeomorphism, $\mu = \delta_k$.

Conversely, suppose $k \notin B(M)$. This implies $\phi(k) \notin ext(K(M))$. Thus, there exists two distinct measures μ_1 and μ_2 on K which represent $\phi(k)$ such that for every $f \in M$, $\phi(k)(f) = \frac{\mu_1(f)}{2} + \frac{\mu_2(f)}{2}$. Let $\mu = \frac{\mu_1}{2} + \frac{\mu_2}{2}$. By assumption, $\mu = \delta_k$. Since μ_1 and μ_2 are distinct, $\mu_1 \neq \delta_k$ implies $\mu_1(\{k\}) < 1$. Hence $\mu(\{k\}) < 1$. Therefore $\mu \neq \delta_k$, which is a contradiction. This implies $k \in B(M)$. □

The above theorem tells us that when $M = C_{\mathbb{R}}(K)$ or $C_{\mathbb{C}}(K)$, $B(M) = K$. An example where $B(M) \neq K$ can be constructed as follows:

Consider $K = [0,1]$ and let $M = \{f \in C_{\mathbb{R}}(K) : f\left(\frac{1}{2}\right) = \frac{f(0)}{2} + \frac{f(1)}{2}\}$. Then $B(M) = K \setminus \left\{\frac{1}{2}\right\}$. Clearly, from the definition of M, for each $f \in M$,

$$\delta_{\frac{1}{2}}(f) = \frac{\delta_0(f) + \delta_1(f)}{2}.$$

This implies $\delta_{\frac{1}{2}} \notin ext(K(M))$ and hence $\frac{1}{2} \notin B(M)$. Let $x \in K \setminus \left\{\frac{1}{2}\right\}$. Suppose there exists a probability measure μ on K such that for each $f \in M$, $f(x) = \int_K f d\mu$. In order to show that $\mu = \delta_x$, it is enough to show that $\mu(K \setminus \{x\}) = 0$ Now, we can choose a function $g \in M$ such that for any $y \neq x$, $|g(y)| < |g(x)| = \|g\|_\infty$. Consider

$$\|g\|_\infty = |g(x)| \leqslant \int_K |g| d\mu = \int_{\{x\}} |g| d\mu + \int_{\{x\}^c} |g| d\mu \leqslant \|g\|_\infty \mu(\{x\}) + \int_{\{x\}^c} |g| d\mu.$$

If $\mu(K \setminus \{x\}) > 0$, then the above inequality will become,

$$\|g\|_\infty < \|g\|_\infty \mu(\{x\}) + \int_{\{x\}^c} |g| d\mu < \|g\|_\infty \mu(\{x\}) + \|g\|_\infty \mu(K \setminus \{x\}) = \|g\|_\infty,$$

which is a contradiction and hence $\mu(K \setminus \{x\}) = 0$. Therefore, $B(M) = K \setminus \left\{\frac{1}{2}\right\}$.

Definition 6.1.5 (Boundary and Šilov Boundary for M)**.** Let M be a subspace of $C_{\mathbb{R}}(K)$ or of $C_{\mathbb{C}}(K)$ and that $1 \in M$. A subset B of K is said to be a boundary for M if for each $f \in M$, there exists $y \in B$ such that $|f(y)| = \|f\|_\infty$. The smallest closed boundary for M (*i.e.* it is the closed boundary which is contained in every other closed boundary) is called the Šilov boundary for M.

Proposition 6.1.6. Let M be a subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) with $1 \in M$. Then $B(M)$ is a boundary for M.

Proof. Let $f \in M$. Since the dual of (M^*, w^*) is M itself, f attains its supremum on $K(M) \cup -K(M)$. Let $L_0 \in K(M) \cup -K(M)$ such that

$$L_0(f) = \sup f(K(M) \cup -K(M)).$$

Suppose $L_0 \in K(M)$. Then the set $F = \{L \in K(M) : L(f) = L_0(f)\}$ is a non-empty w^*-compact convex subset of $K(M)$ and also clearly, a face of $K(M)$. Therefore, $ext(F) \subset ext(K(M)) \subset \phi(K)$. Let $k \in K$ such that $\phi(k) \in ext(F)$. This implies $\phi(k)(f) = L_0(f) = \|f\|_\infty$ and hence $f(k) = \|f\|_\infty$(since $\sup f(K(M)) = \sup f(\phi(K))$). If $L_0 \in -K(M)$, we do a similar argument as above and get the required conclusion by considering $F = \{L \in K(M) : L(f) = -L_0(f)\}$, a non-empty w^*-compact convex face of $K(M)$. □

Proposition 6.1.7. Let M be a subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) which contains constant functions and separates points of K. Then the closure of $B(M)$ is the Šilov boundary of M.

Proof. By previous proposition, $B(M)$ is a boundary for M. Hence, $\overline{B(M)}$ is a closed boundary for M. Let B be any closed boundary for M. It suffices to show that $B(M) \subset B$. Suppose there exists $y \in B(M) \backslash B$. Since $y \notin B$, there exists a neighbourhood U of y such that $y \in U \subset K \backslash B$. If we show that there exists $f \in M$ such that $\sup |f(K \backslash U)| < \sup |f(U)|$, then we get a contradiction to the fact that B is a boundary for M. Now $\delta_y \in ext(K(M))$ and $\phi(U)$ is a weak* neighbourhood of δ_y in $\phi(K)$. By definition of weak* topology and the fact that the dual of (M^*, w^*) is M itself, there exists $g_1, g_2, \ldots, g_n \in M$ such that

$$\delta_y \in \bigcap_{i=1}^{n} (\{L \in M^* : |L(g_i) - g_i(y)| < \epsilon\} \cap \phi(K)) \subset \phi(U).$$

Now

$$\bigcap_{i=1}^{n} (\{L \in M^* : |L(g_i) - g_i(y)| < \epsilon\} \cap \phi(K)) = \bigcap_{i=1}^{n} \{\delta_z \in \phi(K) : |g_i(z) - g_i(y)| < \epsilon\}.$$

Let for any $1 \leqslant i \leqslant n$, $f_i = g_i - g_i(y) \in M$. It follows that,

$$\begin{aligned} & \bigcap_{i=1}^{n} \{\delta_z \in \phi(K) : |g_i(z) - g_i(y)| < \epsilon\} \\ & = \bigcap_{i=1}^{n} \{\delta_z \in \phi(K) : |f_i(z)| < \epsilon\} \\ & = \bigcap_{i=1}^{n} (\{\delta_z \in \phi(K) : f_i(z) < \epsilon\} \cap \{\delta_z \in \phi(K) : -f_i(z) < \epsilon\}) \\ & \subseteq \bigcap_{i=1}^{n} \{\delta_z \in \phi(K) : f_i(z) < \epsilon\}. \end{aligned} \tag{6.1}$$

Therefore,

$$\delta_y \in \bigcap_{i=1}^{n} \{\delta_z \in \phi(K) : f_i(z) < \epsilon\} \subset \phi(U).$$

For each $i = 1, 2, \ldots, n$, let $K_i = \{L \in M^* : L(f_i) \geqslant \epsilon\} \cap K(M)$. Then K_i is w^*-compact convex sets for each $i = 1, 2, \ldots, n$. Let $J = conv(\bigcup_{i=1}^{n} K_i)$, which is also a w^*-compact convex subset of $K(M)$. Now $\delta_y \notin J$. By Hahn Banach separation theorem, there exists $f \in M$ such that

$$\sup f(J) < \delta_y(f) = f(y).$$

Since

$$\phi(K) \backslash \phi(U) \subset \bigcup_{i=1}^{n} K_i \subset J,$$

we have

$$\sup f(\phi(K) \backslash \phi(U)) < f(y).$$

This implies

$$\sup f(K \backslash U) < f(y) \leqslant \sup f(U).$$

Now there exists a large $C > 0$ such that $g = f + C \geqslant 0$. It follows from the above inequality that

$$\sup |g(K \backslash U)| < \sup |g(U)|.$$

Therefore the closure of $B(M)$ is the Šilov boundary for M. □

We will now show that every non-empty compact convex subset of a lctvs is of the form $K(M)$, for a suitable K and M.

Proposition 6.1.8. If X is a compact convex subset of a lctvs E, then there exists a separating subspace M of $C_{\mathbb{R}}(X)$ with $1 \in M$ such that X is affinely homeomorphic with $K(M)$.

Proof. Let M be set of those functions in $C_{\mathbb{R}}(X)$ of the form $g = f + r$, for some $f \in E^*$ and $r \in \mathbb{R}$. Then, M separates points of X, since E^* separates points of X and clearly $1 \in M$. Define ϕ from X to $K(M)$ as we had earlier. Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha + \beta = 1$ and $x, y \in X$. Then for each $f \in M$,

$$\phi(\alpha x + \beta y)(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha\phi(x)(f) + \beta\phi(y)(f).$$

This implies ϕ is affine. Let $L \in K(M)$. Then by Hahn Banach extension theorem, there exists an extension of L, $\tilde{L} \in C_{\mathbb{R}}(X)^*$ such that $\|\tilde{L}\| = 1$. Now, by Riesz Representation theorem, there exists $\mu \in \mathscr{P}(X)$ such that $L(f) = \mu(f)$, for each $f \in M$. Then by Theorem 2.2.2, μ has a unique resultant in X, which implies $\phi(x) = L$. Hence, $\phi(X) = K(M)$. □

We conclude this chapter with a form of the representation theorem which is due to Bishop-de Leeuw.

Theorem 6.1.9. Let M be a subspace of $C_{\mathbb{R}}(K)$(or of $C_{\mathbb{C}}(K)$) such that it separates points of K and contains constant functions. If $L \in M^*$, then there exists a real measure μ on K such that $L(f) = \int_K f d\mu$, for each $f \in M$ and $\mu(S) = 0$, for any Baire set S in K which is disjoint from the Choquet boundary $B(M)$ for M.

Proof. Let $L \in M^*$. By Hahn Banach extension theorem, L can be extended to continuous linear functional $\tilde{L}$ on $C_{\mathbb{R}}(K)$ such that $\|\tilde{L}\| = \|L\|$. By Riesz Representation theorem, there exists $\mu \in M(K)$ such that $\tilde{L}(g) = \int_K g d\mu$, for each $g \in C_{\mathbb{R}}(K)$. Now there exists positive measures on K say μ_1 and μ_2 such that $\mu = \mu_1 - \mu_2$. We can carry these two measures to measures μ_1' and μ_2' on $K(M)$. Therefore, by lemma 4.1.4 there exists maximal measures ν_1', ν_2' on $K(M)$ such that $\mu_1' \prec \nu_1'$ and $\mu_2' \prec \nu_2'$. By Choquet-Bishop-de Leeuw theorem 4.1.16, for $i = 1, 2$, $\nu_i'(S) = 0$, for any Baire set S which is disjoint from $ext(K(M))$. Therefore, $\nu_1 = \nu_1' \circ \phi$ and $\nu_2 = \nu_2' \circ \phi$ are measures on K and $\nu = \nu_1 - \nu_2$ is a boundary measure on K such that $\mu \prec \nu$ and $\nu(S) = 0$, for any Baire set S in K which is disjoint from $B(M) = \phi^{-1}(ext K(M))$. Also, by proposition 4.1.2, $L(g) = \int_K g d\nu$, for each $g \in M$. □

6.2 Choquet Boundary for uniform algebras

Let K be a compact Hausdorff space.

Definition 6.2.1 (Uniform algebra or function algebra)**.** A uniform algebra $\mathscr{A}$ of $C_{\mathbb{C}}(K)$ is defined as a uniformly closed subalgebra of $C_{\mathbb{C}}(K)$ such that $\mathscr{A}$ separates points of K and contains constant functions.

For a metrizable K, Bishop and de Leeuw prescribe a simple description of the Choquet boundary for a uniform algebra $\mathscr{A}$ *i.e* it consists precisely of all the peak points for $\mathscr{A}$.

Definition 6.2.2 (Peak point)**.** A point $x \in K$ is said to be a peak point for a subalgebra M of $C_{\mathbb{C}}(K)$ if there exists a $f \in M$ such that for each $y \neq x$, $|f(y)| < |f(x)|$.

We will now move to the main theorem of this chapter which is due to Bishop and de Leeuw. A special case of this theorem will yield the above claim regarding the Choquet boundary for uniform algebras.

Definition 6.2.3. Let $\mathscr{A}$ be a uniform algebra of $C_{\mathbb{C}}(K)$ and that $y \in K$. We say that y satisfies:

Condition (I) - if for any open neighbourhood U of y and $\epsilon > 0$, there exists $f \in \mathscr{A}$ such that $\|f\|_\infty \leqslant 1$, $|f(y)| > 1 - \epsilon$ and $|f| \leqslant \epsilon$ in $K\backslash U$.

Condition (II) - if whenever S is a G_δ set containing y, there exists $f \in \mathscr{A}$ such that $|f(y)| = \|f\|_\infty$ and $\{x \in K : |f(x)| = \|f\|_\infty\} \subset S$.

Theorem 6.2.4 (Bishop-de Leeuw)**.** Let $\mathscr{A}$ be a uniform algebra of $C_{\mathbb{C}}(K)$ and $y \in K$. Then the following are equivalent:

(i) The point y satisfies Condition (I).

(ii) For each open set U containing y, there exists $f \in \mathscr{A}$ such that $|f(y)| = \|f\|_\infty$ and $|f| < \|f\|_\infty$ in $K\backslash U$.

(iii) For each $x \in K$ with $y \neq x$, there exists $f \in \mathscr{A}$ such that $|f(x)| < |f(y)| = \|f\|_\infty$.

(iv) The point y satisfies Condition (II).

(v) The point y is in the Choquet boundary $B(\mathscr{A})$ for $\mathscr{A}$.

Proof. We prove first (i) implies (ii). Suppose y satisfies Condition I and let U be any open set containing y. We will construct a sequence $\{g_n\}$ in $\mathscr{A}$ with the following properties:

(a) $\|g_{n+1} - g_n\|_\infty \leqslant 2^{-n+1}$

(b) $\|g_n\|_\infty \leqslant 3(1 - 2^{-n-1})$

(c) $g_n(y) = 3(1 - 2^n)$

(d) $|g_{n+1} - g_n| < 2^{-n-1}$ in $K\backslash U$

Suppose, let us assume, that we have the above sequence. Then (a) implies that $\{g_n\}$ is Cauchy in $\mathscr{A}$ and since $\mathscr{A}$ is complete, there exists $f \in \mathscr{A}$ such that $g_n \to f$ uniformly. (b) implies that $\|f\|_\infty \leqslant 3$ but (c) gives us that $f(y) = 3$, hence $\|f\|_\infty = 3 = f(y)$. If $x \in K\backslash U$, then writing $f = g_n + \sum_{k=n}^{\infty}(g_{k+1} - g_k)$, we get

$$|f(x)| \leqslant \|g_n\|_\infty + \sum_{k=n}^{\infty} |g_{k+1}(x) - g_k(x)| < 3(1 - 2^{-n+1}) + \sum_{k=n}^{\infty} 2^{-k-1} < 3.$$

We apply induction to contruct the sequence $\{g_n\}$. Since $y \in U$, by Condition I, there exists $f \in \mathscr{A}$ such that $\|f\|_\infty \leqslant 1$, $|f(y)| > \frac{3}{4}$ and $|f| \leqslant \frac{1}{4}$ in $K\backslash U$. Define $g_1 = \frac{3}{2}\frac{f}{f(y)}$. Since $|f(y)| > \frac{3}{4}$,

$$\|g_1\|_\infty = \frac{3}{2}.\frac{\|f\|_\infty}{|f(y)|} \leqslant \frac{3}{2}.\frac{4}{3} = 2 < 3(1 - 2^{-2}).$$

Also

$$g_1(y) = \frac{3}{2}.\frac{f(y)}{f(y)} = \frac{3}{2} = 3(1 - 2^{-1}).$$

Therefore g_1 satisfies the relevant conditions.

Suppose we have chosen functions $g_1, g_2, \ldots, g_k$ to satisfy the above four conditions. Since g_k is continuous at y, there exists a neighbourhood of y say $V \subset U$ such that in V,

$$|g_k| < |g_k(y)| + 2^{-k-2} < 3(1-2^{-k}) + 2^{-k-2}.$$

Now again by Condition I, we get another $f \in \mathscr{A}$ such that $\|f\|_\infty \leqslant 1, |f(y)| > \frac{3}{4}$ and $|f| \leqslant \frac{1}{4}$ in $K\backslash V$. Define $h = (3.2^{-k-1})\frac{f}{f(y)}$. Then $\|h\|_\infty \leqslant (3.2^{-k-1})\frac{4}{3} = 2^{-k+1}$, $h(y) = 3.2^{-k-1}$ and in $K\backslash V$, $|h| < (3.2^{-k-1}).\frac{1}{4}.\frac{4}{3} = 2^{-k-1}$. Define $g_{k+1} = g_k + h$. Then it follows easily that g_{k+1} satisfies (a), (c) and (d). To check (b); let $x \in V$, then

$$|g_{k+1}(x)| \leqslant |g_k(x)| + |h(x)| \leqslant 3(1-2^{-k}) + 2^{-k-2} + 2^{-k+1} = 3(1-2^{-k-2})$$

and let $x \in K\backslash V$, then

$$|g_{k+1}(x)| \leqslant \|g_k\|_\infty + |h(x)| \leqslant 3(1-2^{-k-1}) + 2^{-k-1} = 3 - 2^k < 3(1-2^{-k-2}).$$

Thus, $\|g_{k+1}\|_\infty \leqslant 3(1-2^{-k-2})$. This completes the induction and hence the proof of (i) implies (ii).

(ii) implies (iii) is immediate; if $x \neq y$, we consider the open set $K\backslash\{x\}$ which contains y and (iii) follows easily.

To prove (iii) implies (iv); consider S to be any G_δ set containing y. Let $\{U_n\}$ be a decreasing sequence of open sets in K such that $S = \bigcap_{n=1}^{\infty} U_n$. For each $n \in \mathbb{N}$, we will find $f_n \in \mathscr{A}$ with the properties $\|f_n\|_\infty = 1 = f_n(y)$ and $|f_n| < 1$ in $K\backslash U_n$. Once we have this function, we define $f = \sum_{n=1}^{\infty} 2^{-n} f_n$. Then

$$|f(y)| = \sum_{n=1}^{\infty} 2^{-n} f_n(y) = \sum_{n=1}^{\infty} 2^{-n} = 1$$

and

$$\|f\|_\infty \leqslant \sum_{n=1}^{\infty} 2^{-n} \|f_n\|_\infty = \sum_{n=1}^{\infty} 2^{-n} = 1.$$

Hence, $\|f\|_\infty = 1 = f(y)$. Moreover, if $x \in K$ such that $|f(x)| = \|f\|_\infty = 1$; suppose $x \notin U_n$, for some $N \in \mathbb{N}$, then

$$1 = |f(x)| \leqslant \sum_{n=1}^{\infty} 2^{-n} |f_n(x)| < \sum_{n=1}^{\infty} 2^{-n} = 1,$$

which is a contradiction. This implies $x \in S$. Hence, f satisfies the properties of Condition II. Let $n \in \mathbb{N}$ and $x \in K\backslash U_n$. (iii) implies there $f_x \in \mathscr{A}$ such that $\|f_x\|_\infty = 1 = |f_x(y)|$

and $|f_x| < 1$ in the neighbourhood V_x of x. Since $K\backslash U_n$ is compact, we can choose finite number $f_{x_1}, f_{x_2}, \ldots, f_{x_k}$ of such functions for which $V_{x_1}, V_{x_2}, \ldots, V_{x_k}$ cover $K\backslash U_n$. Define $f_n = k^{-1}\sum_{i=1}^{k} f_{x_i}$. Then

$$\|f_n\|_\infty \leqslant k^{-1}\sum_{i=1}^{k}\|f_{x_i}\| = k^{-1}\sum_{i=1}^{k} 1 = 1$$

and

$$f_n(y) = k^{-1}\sum_{i=1}^{k} f_{x_i}(y) = k^{-1}\sum_{i=1}^{k} 1 = 1.$$

Hence $\|f_n\|_\infty = 1 = f_n(y)$. For $x \in K\backslash U_n$, there exists $j \in \{1, 2, \ldots, k\}$ such that $x \in V_{x_j}$, hence

$$|f_n(x)| \leqslant k^{-1}\sum_{i=1}^{k}|f_{x_i}(x)| < k^{-1}\sum_{i=1}^{k} 1 = 1.$$

This completes the proof of (iii) implies (iv).

We now prove (iv) implies (v). Suppose y satisfies Condition II. By theorem 6.1.4, it is enough to show that $\mu = \delta_y$ is the only probability measure on K such that $f(y) = \mu(f)$ for each $f \in \mathscr{A}$. Suppose that μ is a measure on K such that $f(y) = \mu(f)$ for each $f \in \mathscr{A}$. It suffices to show $\mu(S) = 1$ for each S is a G_δ containing y. Suppose we have this, then $K\backslash\{x\}$ is a G_δ set and hence $\mu(K\backslash\{x\}) = 1$, for each $x \neq y$. This implies $S(\mu) = \{y\}$. Let S be a G_δ set containing y. (iv) implies there $f \in \mathscr{A}$ such that $y \in \{x : f(x) = \|f\|_\infty\} \subset S$; then

$$\|f\|_\infty = |f(y)| = |\mu(f)| \leqslant \int_S |f|d\mu + \int_{K\backslash S}|f|d\mu \leqslant \|f\|_\infty\mu(S) + \int_{K\backslash S}|f|d\mu.$$

If $\mu(K\backslash S) > 0$, then $\int_{K\backslash S}|f|d\mu < \|f\|_\infty\mu(K\backslash S)$, which is a contradiction. This implies $\mu(K\backslash S) = 0$. Hence, $\mu(S)$=1. This completes the proof.

To prove (v) implies (i), we need a simple lemma.

Lemma 6.2.5. If M is a separating subspace $C_{\mathbb{C}}(K)$ with $1 \in M$, then ReM(the space of real parts of functions in M) is also a separating subspace of $C_{\mathbb{R}}(K)$ and $B(M) = B(ReM)$.

Proof. Let $x, y \in K$ be such that $x \neq y$. Then there exists $f \in M$ such that $f(x) \neq f(y)$. Suppose $f = u + iv$. This implies either $u(x) \neq u(y)$ or $v(x) \neq v(y)$. If $u(x) \neq u(y)$, then we are done. Otherwise, $if \in M$ and $if(x) \neq if(y)$. This implies

$$-v(x) + iu(x) \neq -v(y) + iu(y).$$

Hence $-v \in ReM$ serves our purpose.

Using theorem 6.1.4 and the fact that for a real measure μ on K, $\mu(Ref) = (Ref)(y)$, for each $f \in M$ if and only if $\mu(f) = f(y)$, for each $f \in M$, we get $B(M) = B(ReM)$. $\square$

Let us now return to the proof of (v) implies (i). Suppose $y \in B(\mathscr{A}) = B(Re\mathscr{A})$ and that U is a open neighbourhood of y and $0 < \epsilon < 1$. We can choose a function $g \in C_{\mathbb{R}}(K)$ such that $0 \leqslant g \leqslant 1$ and $g(y) = 1$ and $g \equiv 0$ in $K \backslash U$. Let us denote the w^* compact convex set $K(Re\mathscr{A}) \subset (Re\mathscr{A})^*$ by X. By Tietze extension theorem, we can obtain $f \in C_{\mathbb{R}}(X)$ such that $f = g \circ \phi$ on $\phi(K) \subset X$. Since $\phi(y) \in ext(X)$, by theorem 4.1.12, $\widehat{-f}(\phi(y)) = (-f)(\phi(y)) = -g(y) = -1$. By proposition 3.1.5, the space of continuous functions on X is isomorphic to the uniform closure of $Re\mathscr{A}$ and hence

$$\begin{aligned}\widehat{-f}(y) &= \inf\{h(y) : h \in Re\mathscr{A}, h \geqslant -f\} \\ &= -\sup\{-h(y) : -h \in \mathscr{A}, -h \leqslant f\} \\ &= 1.\end{aligned} \tag{6.2}$$

It follows that there exists $h_0 \in Re\mathscr{A}$ such that $h_0 \leqslant f$ and $h_0(y) > \frac{\log(\delta-1)}{\log\delta}$, where $\delta = 1/\epsilon$. Define $h = (\log\delta)(h_0 - 1)$, then $h \in \mathscr{A}$ and there exists $k \in Re\mathscr{A}$ such that $h + ik \in \mathscr{A}$. Since $\mathscr{A}$ is closed in $C_{\mathbb{C}}(K)$, $f_0 = e^{h+ik} \in \mathscr{A}$. Now $|f_0| = e^h$. Since $h_0 \leqslant g \leqslant 1$, it implies $h_0 - 1 \leqslant 0$ and hence $h \leqslant 0$. Therefore $e^h \leqslant 1$ *i.e.* $|f_0| \leqslant 1$. Since

$$(\log\delta)(h_0(y) - 1) > (\log\delta)\left(\frac{\log(\delta-1)}{\log\delta} - 1\right) = \log(1-\epsilon),$$

therefore

$$|f_0(y)| = e^{(\log\delta)(h_0(y)-1)} > e^{\log(1-\epsilon)} = 1 - \epsilon.$$

In $K \backslash U$, $g \equiv 0$ and hence $h_0 \leqslant 0$ and $(\log\delta)(h_0 - 1) \leqslant -\log\delta = \log\epsilon$. This implies $|f_0| = e^{(\log\delta)(h_0(y)-1)} \leqslant e^{\log\epsilon} = \epsilon$ in $K \backslash U$. $\square$

Corollary 6.2.6. If K is metrizable compact Hausdorff space and $\mathscr{A}$ is a function algebra in $C_{\mathbb{C}}(K)$, then the Choquet boundary $B(\mathscr{A})$ coincides with the set of peak points for $\mathscr{A}$.

Proof. Since K is metrizable, $\{y\}$ is a G_δ for all $y \in K$. From the above theorem, assertion (iv) implies that y is a peak point for $\mathscr{A}$. Hence, (iv) implies (v) proves the required. $\square$

Example 6.2.7. This example is the motivation for the term *boundary*. Let $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leqslant 1\}$ and consider the disk algebra $A(\mathbb{D}) = \{f \in C_{\mathbb{C}}(\overline{\mathbb{D}}) : f \text{ is analytic in } \mathbb{D}\}$. We claim that the Choquet boundary for $A(\mathbb{D})$ coincides with its Šilov boundary and these equal the boundary $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ of $\overline{\mathbb{D}}$. Let $z_0 \in \mathbb{C}$ such that $|z_0| = 1$. Define $f(z) = z + z_0$, for all $z \in \overline{\mathbb{D}}$. Clearly, $f \in A(\mathbb{D})$. Now $|f(z_0)| = 2$. For any $z \in \mathbb{T}$, $|f(z)| = |z + z_0| < 2$, since $\mathbb{T}$ is strictly convex and for any $z \in \mathbb{C}$ with $|z| < 1$, clearly

$|f(z)| < 2$. This implies z_0 is a peak point for $A(\mathbb{D})$. Also, it follows from the maximum modulus principle for analytic functions that no point in $\mathbb{D}$ can be a peak point for $A(\mathbb{D})$. Since $\overline{\mathbb{D}}$ is metrizable, by above corollary, the claim is proved.

Example 6.2.8. This is another example for the case when Choquet and Šilov boundaries are different. Let $M = \{f \in A(\mathbb{D}) : f(0) = f(1)\}$, a closed subalgebra of $A(\mathbb{D})$. It follows from the Maximum Modulus Principle that the Choquet boundary $B(M)$ does not contain 1, as $B(M)$ and the set of peak points are same for this case. We claim that $B(M) = \mathbb{T}\backslash\{1\}$. It remains to prove that for any $z_0 \in \mathbb{T}\backslash\{1\}$, z_0 is a peak point for M. Let us recall the Condition (I) in Definition 6.2.3. Now choose $\varepsilon > 0$ and a neighborhood U of z_0 in $\mathbb{T}$.

CASE 1: When $|z_0 - 1| \geqslant 1$.

Get $f \in A(\mathbb{D})$ such that $|f(z_0)| > 1 - \frac{\varepsilon}{2}$ and $|f||_{U^c} < \frac{\varepsilon}{2}$.

CASE 2: When $|z_0 - 1| < 1$.

Get $f \in A(\mathbb{D})$ such that $|f(z_0)| > 1 - \frac{\varepsilon}{2}|z_0 - 1|$ and $|f||_{U^c} < \frac{\varepsilon}{2}|z_0 - 1|$.

Now define $g(z) = \frac{z(z-1)}{|z_0-1|}.f(z)$. Clearly $g \in M$ and $|g(z_0)| > 1 - \varepsilon$, $|g||_{U^c} < \varepsilon$.

This shows z_0 is a peak point for M and hence $B(M) = \mathbb{T}\backslash\{1\}$.

Remark 6.2.9. (i) It is not generally true that the peak points and the Choquet boundary coincide(in the metrizable case) for linear subspaces M of $C_{\mathbb{C}}(K)$ which are not algebras.

(ii) One inclusion does hold; in the above Bishop-deLeeuw theorem, the proof of (iii) implies (iv) and (iv) implies (v) did not use the fact that $\mathscr{A}$ is an algebra. For any separating linear subspace, every peak point is in the Choquet boundary.

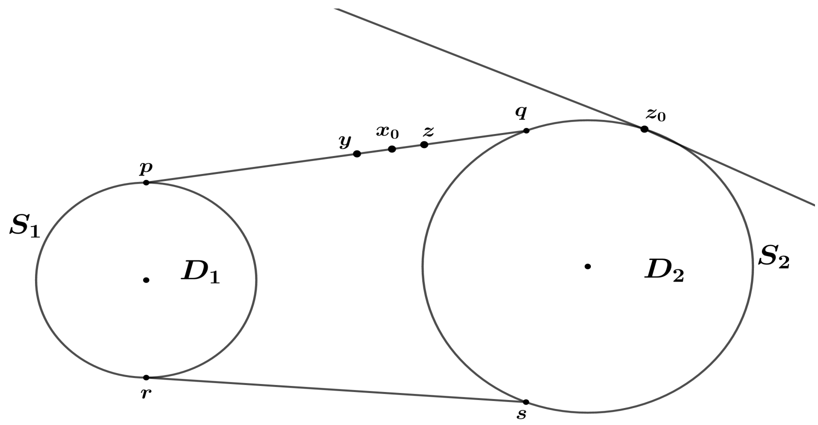

FIGURE 6.1

Now consider the following example:

Example 6.2.10. Let K be the subset of the plane consisting of the convex hull of two disjoint circles(figure 6.1) and let M be the complex-valued affine functions on K. Let the four tangent points be p, q, r, s, denote the open arcs $\overset{\frown}{pr}$ and $\overset{\frown}{qs}$ as S_1 and S_2 and denote the smaller circle as D_1 and the bigger circle as D_2.

Let x_0 be any point on the open line segment joining p and q. Then there exists distinct points y and z on the line segment such that $x_0 = \frac{y+z}{2}$. For any $f \in M$, $f(x_0) = \frac{f(y)+f(z)}{2}$; this implies $f(x_0) = \int_K f d(\frac{\delta_y+\delta_z}{2})$, but $\delta_{x_0} \neq \frac{\delta_y+\delta_z}{2}$. Hence, $x_0 \notin B(M)$ and hence is not a peak point for M.

Consider any $z_0 \in S_1 \cup S_2$ and the tangent line say $ax + by = c$ for some $a, b, c \in \mathbb{R}$ to the point z_0 as shown in the above figure. Then $K \subset \{x + iy \in \mathbb{C} : ax + by \leqslant c\}$ and $g : \mathbb{C} \to \mathbb{R}$ defined as $g(x + iy) = ax + by$ is a real linear functional. Without loss of generality assume that $c > 0$. Hence, $\tilde{g}(z) = g(z) - ig(iz)$, is a complex linear functional. Now $\sup_{z\in K} |\tilde{g}(z)| = \sup_{z\in K} |g(z)|$. Hence the maximum of $\tilde{g}$ is same as that of g. Now $\tilde{g}(z_0) = |\tilde{g}(z_0)|e^{i\theta}$, ie. $e^{-i\theta}\tilde{g}(z_0) = |\tilde{g}(z_0)| \geqslant |g(z_0)| = \sup_K g(z) = c$. Consider the linear functional $e^{-i\theta}\tilde{g}(z)$, which is affine on K and $|e^{-i\theta}\tilde{g}(z)| = |\tilde{g}(z)| \leqslant c$. Hence, z_0 is a peak point for M.

Now suppose p is not in the Choquet boundary for M. This implies $\delta_p \notin ext(K(M))$. Therefore there exists $z_1, z_2 \in S_1 \cup S_2$ such that $\delta_p = \frac{\delta_{z_1}+\delta_{z_2}}{2}$. This implies $p = \frac{z_1+z_2}{2}$ which is a contradiction since p lies on the boundary of the circle. This implies p is in $B(M)$. Similarly, the remaining tangent points also lie in the Choquet boundary for M.

Suppose p is a peak point for M. Then we will get $f \in M$ such that $|f(x)| < |f(p)|$, for $x \in K\backslash\{p\}$. Now $f(p) = |f(p)|e^{i\theta}$, define $h(z) = Re(e^{-i\theta}f(z))$, then $|h(z)| \leqslant |f(z)|$, for all $z \in \mathbb{C}$ and $h(p) = |f(p)| = sup_K|f(z)|$. That is h is a real linear functional on $\mathbb{C}$ which attains its supremum over a smooth surface K at p. It shows that h must be the tangent to the surface. Hence h must coincide with the line segment pq. That is $|f(p)| = h(p) = h(q) \leqslant |f(q)| < |f(p)|$, a contradiction. Hence such an f does not exist. Therefore, p is not a peak point for M.

Therefore, we get that $B(M)$ is precisely $S_1 \cup S_2 \cup \{p, q, r, s\}$ and p, q, r, s are not the peak points for M. Observe that in this example, the peak points are dense in $B(M)$. This is true in general and is a corollary to the following classical result concerning Banach spaces.

Definition 6.2.11 (Smooth point)**.** Let E be a Banach space and S denote the unit sphere of E. A point $x \in S$ is said to be a smooth point of the unit sphere of E if there exists unique $f \in E^*$ such that $f(x) = 1 = \|f\|_\infty$.

Let us recall the smoothness in a normed linear space. See Appendix A.

Proposition 6.2.12 (S. Mazur)**.** Let E be a separable real(or complex) Banach space and let $S = \{x \in E : \|x\| = 1\}$ denote the unit sphere of E. Then the smooth points of S form a dense G_δ subset of S.

Proof. Let us denote the set of smooth points of S by smS. We know that the dual of E as a real space is isometrically isomorphic to the dual of E, where E is considered as a real space and hence in the case of a complex space, we will consider it as a real space and smS is unchanged. We will show that smS is a countable intersection of dense open subsets of S; since S is a complete metric space, the Baire category theorem wil give the desired conclusion.

Let $\{x_n\}$ be a dense sequence in S. Let $m, n \in \mathbb{N}$. Define $D_{mn} = \{x \in S :$ whenever $f, g \in E^*$ satisfy $\|f\| = f(x) = 1 = g(x) = \|g\|, f(x_n) - g(x_n) < m^{-1}\}$. We first prove that $smS = \bigcap_{m,n\in\mathbb{N}} D_{mn}$. Clearly, $smS \subset \bigcap_{m,n\in\mathbb{N}} D_{mn}$. Let $x \notin smS$. Then there exists two distinct $f, g \in E^*$ such that $\|f\| = f(x) = 1 = g(x) = \|g\|$. Since $f \neq g$, for some large enough $n, m \in \mathbb{N}$, $f(x_n) - g(x_n) \geqslant m^{-1}$. This implies $x \notin D_{mn}$. This proves $smS = \bigcap_{m,n\in\mathbb{N}} D_{mn}$.

Next we prove $S \backslash D_{mn}$ is closed in S. Suppose that $(y_k) \subset S\backslash D_{mn}$ and $y_k \to y$. For each $k \in \mathbb{N}$, choose functions $f_k, g_k \in E^*$ such that $\|f_k\| = f_k(y_k) = 1 = g_k(y_k) = \|g_k\|$ and $f_k(x_n) - g_k(x_n) \geqslant m^{-1}$. By Banach Alaoglu theorem, there exists convergent subnets of $\{f_k\}$ and $\{g_k\}$ say $\{f_\alpha\}$ and $\{g_\alpha\}$ respectively such that $f_\alpha \xrightarrow{w^*} f$ and $g_\alpha \xrightarrow{w^*} g$. Clearly, $f(x_n) - g(x_n) \geqslant m^{-1}$. Given $\epsilon > 0$, for all α greater than(w.r.t the ordering) some α_0,

$$|f(y) - 1| \leqslant |f(y) - f_\alpha(y)| + |f_\alpha(y) - f_\alpha(y_\alpha)| < |f(y) - f_\alpha(y)| + \|f_\alpha\|\|y - y_\alpha\| < \epsilon.$$

This implies $f(y) = 1$. Similarly, we can prove $g(y) = 1$. Since w^* convergence implies norm convergence and $|\|f_\alpha\| - \|f\|| \leqslant \|f_\alpha - f\|$, we get $\|f\| = 1 = \|g\|$. This implies $y \in S\backslash D_{mn}$. Hence $S\backslash D_{mn}$ is closed in S.

It remains to show that each set D_{mn} is dense in S. Suppose not; then for some $m, n \in \mathbb{N}$, we can choose $y \in S$ and $\delta > 0$ such that $\|x - y\| < \delta$ and $\|x\| = 1$ imply $x \notin D_{mn}$. Let $y_1 = y$. Then choose $f_1, g_1 \in E^*$ such that $f_1(y_1) = \|f_1\| = 1 = \|g_1\| = g_1(y_1)$ and $f_1(x_n) \geqslant m^{-1} + g_1(x_n)$. We will proceed by induction to define a sequence $\{y_k\}$ in S and corresponding functionals $f_k, g_k \in E^*$ such that $\|y_1 - y_k\| < (1 - 2^{-k})\delta$, $\|f_k\| = f_k(y_k) = 1 = g_k(y_k) = \|g_k\|$ and $f_k(x_n) \geqslant km^{-1} + g_1(x_n)$. Since $f_k(x_n) \leqslant 1$, this implies for each $k \in \mathbb{N}$, $1 \geqslant km^{-1} + g_1(x_n)$, which gives us a contradiction. Suppose we have chosen y_k which has the above properties. We define $y_{k+1} = y_k + \alpha x_n / \|y_k + \alpha x_n\|$, where $\alpha > 0$ is chosen to be small enough so that $\|y_k - y_{k+1}\| < 2^{-k-1}\delta$. Clearly $\|y_{k+1}\| = 1$. Also

$$\|y_1 - y_{k+1}\| < (1 - 2^{-k})\delta + \|y_k - y_{k+1}\| < (1 - 2^{-k-1})\delta < \delta.$$

It follows that $y_{k+1} \notin D_{mn}$. Therefore there exists $f_{k+1}, g_{k+1} \in E^*$ such that $\|f_{k+1}\| = f_{k+1}(y_{k+1}) = 1 = \|g_{k+1}\| = g_{k+1}(y_{k+1})$ and $f_{k+1}(x_n) \geqslant m^{-1} + g_{k+1}(x_n)$. Now

$$1 = \|y_{k+1}\| \geqslant f_k(y_{k+1}) = [1 + \alpha f_k(x_n)]/\|y_k + \alpha x_n\|.$$

Since $g_{k+1}(y_{k+1}) = 1 \geqslant g_{k+1}(y_k)$, we have

$$\|y_k + \alpha x_n\| = g_{k+1}(y_k + \alpha x_n) \leqslant 1 + \alpha g_{k+1}(x_n).$$

Combining the above facts we get the following,

$$f_{k+1}(x_n) \geqslant m^{-1} + g_{k+1}(x_n) \geqslant m^{-1} + f_k(x_n) \geqslant (k+1)m^{-1} + g_1(x_n).$$

This completes the proof. □

Corollary 6.2.13. Suppose K is a compact metrizable space and M is a uniformly closed separating subspace of $C_{\mathbb{C}}(K)$(or of $C_{\mathbb{R}}(K)$) which contains the constant functions. Then the peak points for M are dense in the Choquet boundary for M.

Proof. Let P be the set of $y \in K$ such that there exists a smooth point f of the unit sphere of M such that $f(y) = \|f\|_\infty$. Let $y \in P$. Then there exists a smooth point f of the unit sphere of M such that $f(y) = \|f\|_\infty$. Suppose $f(x) = f(y)$, for some $x \neq y$. Then we have $\delta_x(f) = \delta_y(f) = 1 = \|f\|_\infty$, which is a contradiction to the fact that f is a smooth point of the unit sphere of M. This implies every point of P is a peak point for M. Now P will be dense in $B(M)$ if $\phi(P)$ is a w^* dense in $ext(K(M))$. Suppose $K(M)$ is the w^*-closed convex hull of $\phi(P)$. Then by the partial converse of Krein-Milman theorem, $ext(K(M)) \subset \overline{\phi(P)}^{w^*}$ which implies $B(M) \subset \overline{P}$ and hence the claim. Suppose $K(M)$ is not the w^*-closed convex hull of $\phi(P)$ then there exists $L \in K(M) \backslash \overline{conv}^{w^*}(\phi(P))$. Then by Hahn Banach theorem, there exists $g \in M$ with $\|g\| = 1$ such that (Re g)$(L) >$ sup(Re g)(P). By above theorem, the smooth points are uniformly dense in the unit sphere of M and hence there would exist a smooth point f satisfying the above same inequality. Then we get

$$\|f\|_\infty \geqslant L(f) > \sup(f)(P) = \|f\|_\infty,$$

which is a contradiction. This completes the proof. □

6.3 Choquet Boundary and approximation theory

Many of the classical approximation theory can be formulated in terms of convergence of a sequence of linear operators to the identity operator. To illustrate this, let us consider the Bernstein polynomials. For each $n \in \mathbb{N}$, define the operator B_n on $C[0,1]$. For $f \in C[0,1]$,

$$B_n f = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k},\ x \in [0,1].$$

Bernstein proved that $B_n f$ converges uniformly to f, which gives a constructive proof of the Weierstrass approximation theorem. Now note that for each $n \in \mathbb{N}$. whenever $f \geqslant 0$, $B_n f \geqslant 0$ which implies B_n is a positive operator on $C[0,1]$, for each $n \in \mathbb{N}$. Observing this fact, P. Korovkin proved the following interesting result.

Theorem 6.3.1 (Korovkin)**.** Suppose that $(T_n)_{n\in\mathbb{N}}$ is a sequence of positive operators from $C[0,1]$ into itself with the property that $(T_n f)$ converges uniformly to f for the three functions $f(x) = x^k$, for $k = 0, 1, 2$. Then $(T_n f)$ converges uniformly to f for every $f \in C[0,1]$.

A new proof through inequalities of the above Theorem can be found in [13].

We mention two sets of positive linear operators on $C_{\mathbb{R}}[0,1]$ which approximate identity. A routine verification ensures that these two sets of operators satisfy the stated conditions in Theorem 6.3.1.

Example 6.3.2. (a) The Bernstein operators: $(B_n)_{n=1}^{\infty}$ where $B_n : C_{\mathbb{R}}[0,1] \to C_{\mathbb{R}}[0,1]$ be defined by $B_n(f) = \sum_{k=0}^{n} \binom{n}{k} f(\frac{k}{n}) x^k (1-x)^{n-k}$. A routine verification of the Binomial Theorem guarantees that $(B_n f)$ converges to f uniformly on $[0,1]$ where $f = 1, x, x^2$.

(b) The operators induced by the Schauder basis of $C_{\mathbb{R}}[0,1]$: Define a sequence $(s_n)_{n=0}^{\infty} \subset C_{\mathbb{R}}[0,1]$ as follows: For each $n \in \mathbb{N}$, there exists $m \in \mathbb{N}$ such that $2^{m-1} < n \leqslant 2^m$, then define-

$$s_n(t) = \begin{cases} 2^m \left(t - \frac{2n-2}{2^m} - 1\right) & \text{if } \frac{2n-2}{2^m} - 1 \leqslant t \leqslant \frac{2n-1}{2^m} - 1 \\ (1-2^m)\left(t - \left(\frac{2n-1}{2^m} - 1\right)\right) & \text{if } \frac{2n-1}{2^m} - 1 \leqslant t \leqslant \frac{2n}{2^m} - 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

The sequence (s_n) are called Schauder basis of $C_{\mathbb{R}}[0,1]$. For $f \in C[0,1]$, define another sequence $(p_n)_{n=0}^{\infty}$ as follows:

$p_0 = f(0)s_0$

$p_1 = p_0 + (f(1) - p_0(1))s_1$

$p_2 = p_1 + (f(1/2) - p_1(1/2))s_2$

$p_3 = p_2 + (f(1/4) - p_2(1/4))s_3$

$p_4 = p_3 + (f(3/4) - p_3(3/4))s_4$ and so on.

Hence for each $n \geqslant 0$, there exists $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ such that $p_n = \sum_{i=0}^{n} \alpha_i s_i$. Define $P_n : C[0,1] \to C[0,1]$ as for each $f \in C[0,1]$, $P_n f = p_n$. Clearly each P_n is a positive operator on $C[0,1]$. It is easy to verify that $(P_n f)$ converges to f uniformly on $[0,1]$ when $f = 1, x, x^2$.

Definition 6.3.3 (Korovkin set)**.** Let K be a compact Hausdorff space. Let M be a subset of $C_{\mathbb{R}}(K)$ such that for any countable family of positive operators $(T_n)_{n\in\mathbb{N}}$, whenever for each $g \in M$, $T_n g \to g$(*i.e.* $T_n g$ converges to g uniformly), we have $T_n f \to f$, for each $f \in C_{\mathbb{R}}(K)$. Then M is said to be Korovkin set of $C_{\mathbb{R}}(K)$.

Remark 6.3.4. Let $M \subset C_{\mathbb{R}}(K)$. M is a Korovkin set of $C_{\mathbb{R}}(K)$ if and only if the linear span of M is a Korovkin set of $C_{\mathbb{R}}(K)$.

Remark 6.3.5. Let $K = [0,1]$. Let x_0 be an arbitrary point in $[0,1]$. Now, by Korovkin's theorem and earlier remark, $M = span\{1, x, x^2\}$ is a Korovkin set of $C[0,1]$. Consider $f(x) = 1 - (x - x_0)^2$, for all $x \in [0,1]$. Clearly $f \in M$ and attains its supremum only at x_0 and hence x_0 is a peak point for M. This implies $B(M) = [0,1]$. This leads us to the main theorem of this Chapter.

Theorem 6.3.6 (Šaškin)**.** Let K be a metrizable compact space and M be a linear subspace of $C_{\mathbb{R}}(K)$ which contains 1 and separates points of K. Then M is a Korovkin set of $C_{\mathbb{R}}(K)$ if and only if the Choquet boundary $B(M)$ for M is all of K.

Proof. Assume that M is a Korovkin set of $C_{\mathbb{R}}(K)$. To prove $B(M) = K$, using Theorem 6.1.4, it is enough to prove that for any $t \in K$, $\mu = \delta_t$ is the only probability measure on K such that $\mu(g) = g(t)$, for all $g \in M$. Let $t \in K$ and μ be a probability measure on K such that $\mu(g) = g(t)$, for all $g \in M$. Since K is metrizable space, we can choose a decreasing sequence of open sets $(U_n)_{n\in\mathbb{N}}$ such that $\bigcap_{n\in\mathbb{N}} U_n = \{t\}$. By Urysohn's lemma, for each $n \in \mathbb{N}$, there exists $g_n \in C_{\mathbb{R}}(K)$ such that $0 \leqslant g_n \leqslant 1$, $g_n(t) = 1$ and $g_n = 0$ in U_n^c. For each $n \in \mathbb{N}$, define T_n as $T_n f = \mu(f) g_n + (1 - g_n) f$, for each $f \in C_{\mathbb{R}}(K)$. Clearly T_n is a linear operator on $C_{\mathbb{R}}(K)$ and $T_n(1) = 1$. Since μ is a probability measure on K and $0 \leqslant g_n \leqslant 1$, whenever $f \geqslant 0$, $T_n f \geqslant 0$, which implies T_n is a positive operator on $C_{\mathbb{R}}(K)$. Also $\|T_n f\| \leqslant \|\mu\| \|f\|_\infty \|g_n\|_\infty + (1 + \|g_n\|_\infty)\|f\|_\infty \leqslant \|f\|_\infty + 2\|f\|_\infty = 3\|f\|_\infty$, which implies $\|T_n\| \leqslant 3$. Hence, T_n is a linear bounded operator on $C_{\mathbb{R}}(K)$, for each $n \in \mathbb{N}$. Let $g \in M$ and $\epsilon > 0$, then there exists $N \in \mathbb{N}$ such that $|g(t) - g(y)| < \epsilon$, for all $y \in U_N$. This is because $g \in C_{\mathbb{R}}(K)$ and hence there exists $N \in \mathbb{N}$ such that $t \in U_N \subset g^{-1}((g(t) - \epsilon, g(t) + \epsilon))$. Let $x \in K$, then for all $n > N$, $|T_n g(x) - g(x)| = |\mu(g) - g(x)||g_n(x)| = |g(t) - g(x)||g_n(x)|$. Now if $x \in U_N$, then $|T_n g(x) - g(x)| = |g(t) - g(x)||g_n(x)| < \epsilon$ and if $x \notin U_N$, $|T_n g(x) - g(x)| = 0$. This implies for each $g \in M$, $\|T_n g - g\|_\infty \to 0$ as $n \to \infty$. Since M is Korovkin set of $C_{\mathbb{R}}(K)$, for each $f \in C_{\mathbb{R}}(K)$, $\|T_n f - f\|_\infty \to 0$ as $n \to \infty$. This implies $T_n f(t) \to f(t)$ as $n \to \infty$.

Now, for each $n \in \mathbb{N}$, $T_n f(t) = \mu(f)$. Hence $\mu(f) = f(t)$, for all $f \in C_{\mathbb{R}}(K)$. This implies $\mu = \delta_t$.

Conversely, let $\{T_n\}_{n\in\mathbb{N}}$ be a family of positive operators on $C_{\mathbb{R}}(K)$ such that for each $g \in M$, $T_n g \to g$ as $n \to \infty$. Let $f \in C_{\mathbb{R}}(K)$. In order to show that $\|T_n f - f\|_\infty \to 0$, it suffices to show that every subsequence of $\{\|T_n f - f\|_\infty\}$ itself has a subsequence which converges to 0. For simplicity of notation, assume that $\{\|T_n f - f\|_\infty\}$ is the initial subsequence and for each $n \in \mathbb{N}$, choose $x_n \in K$ such that $\|T_n f - f\|_\infty = |T_n f(x_n) - f(x_n)|$. Since K is compact, by taking a further subsequence, we can assume that $x_n \to x$, for some $x \in K$. For each $n \in \mathbb{N}$, define $L_n : C_{\mathbb{R}}(K) \to \mathbb{R}$ as $L_n(g) = T_n g(x_n)$. Clearly L_n is a linear functional on $C_{\mathbb{R}}(K)$ and also $|L_n(g)| \leqslant \|T_n\|\|g\|_\infty$, which implies L_n is bounded. Also whenever $g \geqslant 0$, since T_n is a positive operator on $C_{\mathbb{R}}(K)$, $L_n(g) \geqslant 0$. Hence, $L_n \in M^+(K)$. As $1 \in M$, $|T_n 1(x_n) - 1(x_n)| \leqslant \|T_n(1) - 1\|_\infty \to 0$. Hence, $L_n 1 \to 1$. Without loss of generality assume that $L_n 1 > 0$, for all $n \in \mathbb{N}$. For any $f \in C_{\mathbb{R}}(K)$ such that $\|f\|_\infty \leqslant 1$, $1 - f \geqslant 0$ and hence $L_n 1 \geqslant L_n f$. This implies $\|\frac{L_n}{L_n 1}\| \leqslant 1$. Therefore, $\mu_n = \frac{L_n}{L_n 1} \in \mathscr{P}(K)$. Since K is metrizable, $C_{\mathbb{R}}(K)$ is separable. This implies $(B_{M(K)}, w^*)$ and hence $(\mathscr{P}(K), w^*)$ is metrizable compact space. Therefore $\{\mu_n\}$ has a convergent subsequence say $\mu_{n_k} \to \mu$, for some $\mu \in \mathscr{P}(K)$. Let $g \in M$. Then

$$|T_{n_k} g(x_{n_k}) - g(x)| \leqslant |T_{n_k} g(x_{n_k}) - g(x_{n_k})| + |g(x_{n_k}) - g(x)| \leqslant \|T_{n_k} g - g\|_\infty + |g(x_{n_k}) - g(x)| \to 0,$$

since $\|T_{n_k} g - g\|_\infty \to 0$, by assumption and $|g(x_{n_k}) - g(x)| \to 0$ because $g \in C_{\mathbb{R}}(K)$ and $x_n \to x$. Therefore we get that $|L_{n_k} g - g(x)| \to 0$. Now, consider

$$|\mu_{n_k}(g) - g(x)| = \left|\frac{L_{n_k} g}{L_{n_k} 1} - g(x)\right| = \left(\frac{1}{L_{n_k} 1}\right) |L_{n_k} g - g(x)| + \left|\frac{g(x)}{L_{n_k} 1} - g(x)\right|.$$

As $k \to \infty$, $|L_{n_k} g - g(x)| \to 0$ and $L_{n_k} 1 \to 1$, hence $|\mu_{n_k}(g) - g(x)| \to 0$, which implies $\mu_{n_k}(g) \to g(x)$ but $\mu_{n_k}(g) \to \mu(g)$. Therefore, $\mu(g) = g(x)$, for all $g \in M$. Since $x \in B(M) = K$, $\mu = \delta_x$, *i.e.* $\frac{L_{n_k}}{L_{n_k} 1} \xrightarrow{w^*} \delta_x$. Consider

$$\begin{aligned} \|T_{n_k} f - f\|_\infty &\leqslant |T_{n_k} f(x_{n_k}) - f(x)| + |f(x) - f(x_{n_k})| \\ &< L_{n_k} 1 \left|\frac{T_{n_k} f(x_{n_k})}{L_{n_k} 1} - f(x)\right| + |L_{n_k} 1.f(x) - f(x)| + |f(x) - f(x_{n_k})|. \end{aligned} \tag{6.4}$$

As $k \to \infty$, $\frac{L_{n_k} f}{L_{n_k} 1} \to f(x)$ and $L_{n_k} 1 \to 1$, $\|T_{n_k} f - f\| \to 0$. This completes the proof. □

Appendix A

Convexity and Smoothness

A.1 Convexity in a topological vector space

Definition A.1.1. (Extreme points)

(a) Let K be a compact convex subset of a topological vector space (tvs for short) E. A point $x \in K$ is said to be an extreme point of K if there does not exist $y, z \in K$ such that $x = \lambda y + (1 - \lambda)z$, for some $\lambda \in (0, 1)$.

(b) A normed linear space X is said to be strictly convex if for every $x \in S_X$, x is an extreme point of B_X.

A.1.1 Characterizations

Proposition A.1.2. Let K be a closed, bounded and convex subset of a tvs X. Then the following are equivalent.

(a) $x \in K$ is an extreme point.

(b) $K \backslash \{x\}$ is convex set in X.

(c) there does not exist $y, z \in K \backslash \{x\}$ such that $x = \frac{y+z}{2}$.

A.1.2 Extreme set(or Face)

Definition A.1.3. (Extreme set) Let X be a tvs, $K \subseteq X$ be a compact convex subset and $F \subseteq K$ be closed and convex. F is said to be an extreme set or face of K if for every $x, y \in K$, $\lambda x + (1 - \lambda)y \in X$, for some $\lambda \in (0, 1)$ would imply $x, y \in K$.

Remark A.1.4. (a) If Y is a subspace of a normed linear space X and $y^* \in S_{Y^*}$, then $HB(y^*) = \{x^* \in S_{X^*} : x^*|_Y = y^*\}$ is an extreme subset of B_{X^*}.

(b) If F is an extreme subset of B_X and $x \in F$ is an extreme point of F, then x is an extreme point of B_X.

A.2 Smoothness in a Normed Linear Space

Let us recall the following definition of differentiability in $\mathbb{R}^n$.

Definition A.2.1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $x_0 \in \mathbb{R}^n$ be nonzero. f is said to be differentiable at x_0 if there exists a linear $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $f(x_0 + y) = f(x_0) + T(y) + \varepsilon(y)$, where $y \in \mathbb{R}^n$ and $\frac{\|\varepsilon(y)\|}{\|y\|} \to 0$ as $\|y\| \to 0$.

The linear map T is called the derivative of f at x_0 and is denoted by $f'(x_0)$. Let us recall that a norm function on a finite dimensional vector space can be considered as a continuous convex function and the smoothness of this function at a non-zero x_0 can be interpreted as the existence of a unique tangent plane passing through the point x_0 at the level surface $\{x \in \mathbb{R}^n : \|x\| = \|x_0\|\}$.

The following result motivates us to define the notion of differentiability in a normed space X.

Proposition A.2.2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as $f(x) = \|x\|$ and $x_0 \in \mathbb{R}^n \backslash \{0\}$. f is differentiable at x_0 if and only if there exists a unique linear functional $\Lambda : \mathbb{R}^n \to \mathbb{R}$ such that $\Lambda(x_0) = \|x_0\|$ and $\|\Lambda\| = 1$.

Proof. Suppose f is differentiable at x_0. We know that $f^{'}(x_0)$ is a linear map on $\mathbb{R}^n$. We will show that $f^{'}(x_0)(x_0) = \|x_0\|$ and $\|f^{'}(x_0)\| = 1$. We know that

$$f(x_0 + y) = f(x_0) + f^{'}(x_0)(y) + \varepsilon(y),$$

where $\varepsilon(y) \in \mathbb{R}$ such that $\frac{|\varepsilon(y)|}{\|y\|} \to 0$ as $\|y\| \to 0$. Now choose $y = tx_0$ for $t > 0$, then as $t \searrow 0$, $\|y\| \to 0$. Hence,

$$f(x_0 + tx_0) = f(x_0) + f^{'}(x_0)(tx_0) + \varepsilon(tx_0),$$

that is,

$$\|(x_0(1 + t)\| = \|(x_0)\| + tf^{'}(x_0)(x_0) + \varepsilon(tx_0).$$

Simplifying the above equation we get,

$$\|x_0\| - f'(x_0)(x_0) = \frac{\varepsilon(tx_0)}{t} = \frac{\varepsilon(tx_0)}{\|tx_0\|}\|x_0\|.$$

Therefore, $\big|\|x_0\| - f'(x_0)(x_0)\big| = \frac{|\varepsilon(tx_0)|}{\|tx_0\|}.\|x_0\| \to 0$ as $t \to 0$. This implies $f'(x_0)(x_0) = \|x_0\|$.

Now for any $y \in \mathbb{R}^n$ and $t > 0$, $\|x_0 + ty\| = \|x_0\| + tf'(x_0)(y) + \varepsilon(ty)$. Then by triangle inequality,

$$\|x_0\| + t\|y\| \geqslant \|x_0\| + tf'(x_0)(y) + \varepsilon(ty).$$

Thus,

$$t\|y\| \geqslant tf'(x_0)(y) + \varepsilon(ty),$$

that is,

$$\|y\| \geqslant f'(x_0)(y) + \frac{\varepsilon(ty)}{\|ty\|}\|y\| = f'(x_0)(y) + \frac{|\varepsilon(ty)|}{\|ty\|}\|y\|.$$

Since $\frac{|\varepsilon(ty)|}{\|ty\|}$ can be made arbitrarily small, we have $\|y\| \geqslant f'(x_0)(y)$. Hence, $\|f'(x_0)\| \leqslant 1$ and also we proved $f'(x_0)(x_0) = \|x_0\|$, which implies $\|f'(x_0)\| = 1$.

Conversely, assume that there exists a unique linear functional Λ on $\mathbb{R}^n$ such that $\Lambda(x_0) = \|x_0\|$ and $\|\Lambda\| = 1$. We will show that for every $y \in \mathbb{R}^n$, $\lim_{t\to 0} \frac{\|x_0+ty\|-\|x_0\|}{t}$ exists. Since $\|.\|$ is a convex function, both

$$\lim_{t\to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t} \text{ and } \lim_{s\to 0^-} \frac{\|x_0 + sy\| - \|x_0\|}{s} \text{ exist.}$$

Also, define for every $y \in \mathbb{R}^n$,

$$p(y) = \lim_{t\to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t} \geqslant \lim_{s\to 0^-} \frac{\|x_0 + sy\| - \|x_0\|}{s} = q(y).$$

We now claim that $p : \mathbb{R}^n \to \mathbb{R}$ and $q : \mathbb{R}^n \to \mathbb{R}$ are sublinear and superlinear functionals respectively. Let $y_1, y_2 \in \mathbb{R}^n$ and $\alpha > 0$. Then we have,

$$\begin{aligned}
p(y_1 + y_2) &= \lim_{t\to 0^+} \frac{\|x_0 + t(y_1 + y_2)\| - \|x_0\|}{t} \\
&= \lim_{t\to 0^+} \frac{\|2x_0 + 2t(y_1 + y_2)\| - \|2x_0\|}{2t} \\
&\leqslant \lim_{t\to 0^+} \frac{\|x_0 + 2ty_1\| - \|x_0\|}{2t} + \lim_{t\to 0^+} \frac{\|x_0 + 2ty_2\| - \|x_0\|}{2t} \\
&= p(y_1) + p(y_2).
\end{aligned} \tag{A.1}$$

and

$$\begin{aligned} p(\alpha y_1) &= \lim_{t\to 0^+} \frac{\|x_0 + t(\alpha y_1)\| - \|x_0\|}{t} \\ &= \lim_{s\to 0^+} \frac{\alpha(\|x_0 + s y_1\| - \|x_0\|)}{s} \\ &= \alpha p(y_1). \end{aligned} \tag{A.2}$$

Hence, p is a sublinear functional on $\mathbb{R}^n$. Now, clearly, $p(-y) = -q(y)$, for any $y \in \mathbb{R}^n$ and hence q is superlinear functional on $\mathbb{R}^n$.

Let $y \in \mathbb{R}^n$. Now for any $p(y) \geqslant t \geqslant q(y)$, we can define $\Phi : span\{y\} \to \mathbb{R}$ as $\Phi(\alpha y) = t\alpha$. Then, clearly Φ is linear functional on $span\{y\}$. Now whenever $\alpha \geqslant 0$, $\Phi(\alpha y) \leqslant p(\alpha y)$, for any $y \in \mathbb{R}^n$. For any $y \in \mathbb{R}^n$, $\Phi(-y) = -t < -q(y) = p(-y)$ and hence whenever $\alpha < 0$, $\Phi(\alpha y) \leqslant p(\alpha y)$. This implies Φ is a linear functional on $span\{y\}$ dominated by the sublinear functional p. Therefore by Hahn Banach theorem, Φ can be extended to $\mathbb{R}^n$, which is also dominated by p. Since $p(x_0) = \|x_0\| = q(x_0)$, $\Phi(x_0) = \|x_0\|$. Now for some $y \in \mathbb{R}^n$, suppose $p(y) > t_1 \neq t_2 > q(y)$, then there exists two distinct linear functionals on $span\{y\}$ and finally by Hahn Banach theorem, two distinct linear functionals on $\mathbb{R}^n$, which contradicts the uniqueness of the given Λ. □

Definition A.2.3. A normed linear space X is said to be smooth at $x_0 \in \mathbb{R}^n$ if there exists unique $x^* \in S_{X^*}$ such that $x^*(x_0) = \|x_0\|$.

Remark A.2.4. (*a*) Let X be a normed linear space and $f : X \to \mathbb{R}$ be a convex function. If the partial derivatives exist in all direction $y \in S_X$ at x_0 then f is differentiable at x_0.

(*b*) Many other notions of differentiability are also available in the literature. The notion defined in Definition A.2.3 is called Gâteaux differentiability. The other notions like Fréchet differentiability, Uniformly Gâteaux differentiability, Uniformly Fréchet differentiability, etc are all strengthenings of Definition A.2.3. All these notions are equivalent under the assumption that the unit sphere is norm-compact.

Appendix B

Weak and Weak* Topologies

Definition B.0.1. Let E be any tvs over a field $\mathbb{F}$. Let $\Sigma = (f_i)_{i\in I}$ be a collection of linear functionals on E. For any finite sub-collection $(f_{i_j})_{j=1}^n$ of Σ and $\epsilon > 0$, let

$$W(f_{i_1}, f_{i_2}, ..., f_{i_n}, \epsilon) := \bigcap_{j=1}^{n} \{x \in E : |f_{i_j}(x)| < \epsilon\}.$$

Then, the collection,

$$\mathscr{B} = \{W(f_{i_1}, f_{i_2}, ..., f_{i_n}, \epsilon) : (f_{i_j})_{j=1}^n \subset \Sigma \text{ and } \epsilon > 0\}$$

forms a basis of E.

For any $x_0 \in E$, one can define the neighbourhood system at x_0 as follows,

$$\{W(f_{i_1}, f_{i_2}, \ldots, f_{i_n}, \epsilon) + x_0 : W(f_{i_1}, f_{i_2}, \ldots, f_{i_n}, \epsilon) \in \mathscr{B}\}.$$

Then there exists a unique topology τ on E for which $\mathscr{B}$ is a neighborhood base of 0 and τ is the smallest such topology on E such that the functionals in Σ are continuous.

Suppose X is a normed linear space. With these notations defined above, the topology defined on X^{**} is said to be weak* topology if $\Sigma = X^*$ and $E = X^{**}$ and is denoted by $\sigma(X^{**}, X^*)$. The corresponding subspace topology on X viz. $\sigma(X^{**}, X^*)|_X$ is called the weak topology on X. If there is no chance of confusion, we denote the topology $\sigma(X^{**}, X^*)$ on X^{**} by (X^{**}, w^*) and $\sigma(X^{**}, X^*)|_X$ by (X, w).

B.1 Basic Properties

Proposition B.1.1. Let $(X, ||.||)$ be an infinite dimensional NLS. Then, the weak topology on X is not first countable.

Proof. Suppose (X, w) is first countable. Let $f \in X^*$. Consider,

$$0 \in \{x \in X : |f(x)| < \epsilon\} \in w.$$

Since X is first countable, there exists $f_1, f_2, ..., f_n \in X^*$ such that

$$W(f_1, f_2, ..., f_n, \epsilon) \subseteq \{x \in X : |f(x)| < \epsilon\}.$$

This implies that,

$$\bigcap_{i=1}^{n}\{x \in X : |f_i(x)| < \epsilon\} \subseteq \{x \in X : |f(x)| < \epsilon\}. \tag{B.1}$$

Claim : $\bigcap_{i=1}^{n} ker(f_i) \subseteq ker(f)$.

Let $x \in \bigcap_{i=1}^{n} ker(f_i)$. Since $f_i's$ are linear functionals on X, for any $n \in \mathbb{N}$, $nx \in \bigcap_{i=1}^{n} ker(f_i)$. This implies that $nx \in \bigcap_{i=1}^{n}\{x \in X : |f_i(x)| < \epsilon\}$. Hence, from (B.1), $nx \in \{x \in X : |f(x)| < \epsilon\}$. Thus, for all $n \in \mathbb{N}$, $|f(nx)| < \epsilon$. This implies for all $n \in \mathbb{N}$, $|f(x)| < \frac{\epsilon}{n}$. Therefore, $f(x) = 0$. This implies that $x \in ker(f)$.

Claim : There exists $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ such that $f = \sum_{i=1}^{n} \alpha_i f_i$.

Define $\psi : X \to \mathbb{F}^n$ as $\psi(x) = (f_1(x), f_2(x), ..., f_n(x))$. Clearly, ψ is a linear map, since $f_1, f_2, ..., f_n$ are linear functionals on X. Hence, $\psi(X)$ is a subspace of $\mathbb{F}^n$. Now, define $\rho : \psi(X) \to \mathbb{F}$ as $\rho((f_1(x), f_2(x), ..., f_n(x))) = f(x)$.

Then, clearly, ρ is well-defined and a linear functional on $\mathbb{F}^n$, since from above claim $\bigcap_{i=1}^{n} ker(f_i) \subseteq ker(f)$.

Since $\psi(X)$ is a subspace of $\mathbb{F}^n$, ρ can be extended linearly to $\mathbb{F}^n$. Let the extension be $\widetilde{\rho} : \mathbb{F}^n \to \mathbb{F}$. This implies that there exists $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{F}$ such that for any $(x_1, x_2, ..., x_n) \in \mathbb{F}^n$,

$$\widetilde{\rho}((x_1, x_2, ..., x_n)) = \sum_{i=1}^{n} \alpha_i x_i.$$

Thus, for any $x \in X$,

$$\rho((f_1(x), f_2(x), ..., f_n(x))) = \widetilde{\rho}((f_1(x), f_2(x), ..., f_n(x))) = \sum_{i=1}^{n} \alpha_i f_i(x).$$

Hence, by definition of ρ, for any $x \in X$, $f(x) = \sum_{i=1}^{n} \alpha_i f_i(x)$. Hence the claim follows.

Therefore, the above claims are true for any $f \in X^*$. This implies that we will get a collection of linear functionals such that $(f_i)_{i\in\mathbb{N}}$.

Let $Z_n =$ span$\{f_i : i = 1, 2, ..., n\} \subseteq X^*$. Since each Z_n is a proper subspace of X^*, each Z_n is closed in X^* and hence, $int(Z_n) = \phi$. Also, $X^* = \bigcup_{n=1}^{\infty} Z_n$. However, this is a contradiction to the Baire Category theorem, since X^* is complete.

Hence, our assumption was wrong. This implies that the weak topology on X is not first countable. □

Remark B.1.2. (i) If X is an infinite dimensional NLS, then neither $\sigma(X^{**}, X^*)$ nor $\sigma(X^{**}, X^*)|_X$ are first countable.

(ii) If X is infinite dimensional then both $\|.\| : (X^{**}, \sigma(X^{**}, X^*)) \to \mathbb{R}$ and $\|.\| : (X, \sigma(X^{**}, X^*)|_X) \to \mathbb{R}$ are lower semi continuous.

(iii) With these notations defined above, the set of continuous linear functionals on $(X, \sigma(X^{**}, X^*)|_X)$ is X^* and the set of continuous linear functionals on $(X^*, \sigma(X^*, X))$ is X.

B.1.1 Banach-Alaoglu Theorem

It is well known in infinite dimensional case that the norm topology cannot allow the closed unit ball to be compact but the situation can occur in a weaker topology.

Theorem B.1.3. Let $(X, ||.||)$ be any normed linear space over a field $\mathbb{F}$. Then, (B_{X^*}, w^*) is compact, where $B_{X^*} = \{f \in X^* : ||f|| \leqslant 1\}$.

Proposition B.1.4. Let X be any separable infinite dimensional normed linear space. Then (B_{X^*}, w^*) is metrizable.

Proof. Let (x_n) be a countable dense subset of S_X. Define $d : B_{X^*} \times B_{X^*} \to [0, \infty)$ such that

$$d(x^*, y^*) = \sum_{n=1}^{\infty} 2^{-n} \frac{|(x^* - y^*)(x_n)|}{1 + |(x^* - y^*)(x_n)|}.$$

Clearly, d is well-defined map since the series in RHS is uniformly convergent by Weierstrass-M test. It is also clear that for any $x^*, y^* \in X^*$, $d(x^*, y^*) \geqslant 0$. Let $x^*, y^* \in X^*$. Then,

$$d(x^*, y^*) = \sum_{n=1}^{\infty} 2^{-n} \frac{|(x^* - y^*)(x_n)|}{1 + |(x^* - y^*)(x_n)|} = \sum_{n=1}^{\infty} 2^{-n} \frac{|(y^* - x^*)(x_n)|}{1 + |(y^* - x^*)(x_n)|} = d(y^*, x^*).$$

Let $x^*, y^*, z^* \in X^*$ and $k \in \mathbb{N}$. We know that $([0,\infty), |.|)$ is a metric space and hence $([0,\infty), \frac{|.|}{1+|.|})$ is also a metric space. Then,

$$\sum_{n=1}^{k} 2^{-n}\frac{|(x^*-y^*)(x_n)|}{1+|(x^*-y^*)(x_n)|} \leqslant \sum_{n=1}^{k} 2^{-n}\frac{|(x^*-z^*)(x_n)|}{1+|(x^*-z^*)(x_n)|} + \sum_{n=1}^{k} 2^{-n}\frac{|(z^*-y^*)(x_n)|}{1+|(z^*-y^*)(x_n)|}.$$

This implies for all $k \in \mathbb{N}$,

$$\sum_{n=1}^{k} 2^{-n}\frac{|(x^*-y^*)(x_n)|}{1+|(x^*-y^*)(x_n)|} \leqslant d(x^*, z^*) + d(z^*, y^*).$$

Hence,

$$d(x^*, y^*) \leqslant d(x^*, z^*) + d(z^*, y^*).$$

Therefore, d is a metric.

Now consider the identity map $I : (B_{X^*}, w^*) \to (B_{X^*}, d)$. Let (x^*_α) be a net in (B_{X^*}, w^*) such that $x^*_\alpha \xrightarrow{w^*} x^*_0$. Then, for every $x \in X$, $x^*_\alpha(x) \to x^*_0(x)$ in $(\mathbb{R}, |.|)$. In particular, for any $n \in \mathbb{N}$, $x^*_\alpha(x_n) \to x^*_0(x_n)$. This implies,

$$\lim_\alpha \frac{|(x^*_\alpha - x^*_0)(x_n)|}{1+|(x^*_\alpha - x^*_0)(x_n)|} = 0.$$

Therefore, since $\sum_{n=1}^{\infty} 2^{-n}\frac{|(x^*_\alpha - x^*_0)(x_n)|}{1+|(x^*_\alpha - x^*_0)(x_n)|}$ is uniformly convergent in $[0,\infty)$,

$$\lim_\alpha \sum_{n=1}^{\infty} 2^{-n}\frac{|(x^*_\alpha - x^*_0)(x_n)|}{1+|(x^*_\alpha - x^*_0)(x_n)|} = \sum_{n=1}^{\infty} 2^{-n} \lim_\alpha \frac{|(x^*_\alpha - x^*_0)(x_n)|}{1+|(x^*_\alpha - x^*_0)(x_n)|} = 0.$$

Hence, $\lim_\alpha d(x^*_\alpha, x^*_0) = 0$. This implies in (B_{X^*}, d), $Id(x^*_\alpha) \to Id(x^*_0)$.

Therefore, I is continuous on (B_{X^*}, w^*). Moreover, (B_{X^*}, w^*) is compact and (B_{X^*}, d) is Hausdorff. As a result, I is homeomorphism. Therefore, (B_{X^*}, w^*) is a metrizable compact convex subset of a non-metrizable lctvs (X^*, w^*). □

Appendix C

On some basic results in Measure Theory

Let (X, τ) be any topological space and $\mathfrak{M}$ be the Borel σ-algebra over X. Let μ be a measure on $(X, \mathfrak{M})$ and $E \in \mathfrak{M}$. Then μ is said to be,

$$\begin{aligned} \textit{Outer Regular} \text{ if } \mu(E) &= \inf\{\mu(U) : U \text{ is open and } E \subseteq U\}, \forall\ E. \\ \textit{Inner Regular} \text{ if } \mu(E) &= \sup\{\mu(K) : K \text{ is compact and } K \subseteq E\}, \forall\ E \text{ with } \mu(E) < \infty. \end{aligned}$$

The measure μ is said to be *Regular* if it is both Inner and Outer Regular. A complex measure μ on K is said to be Regular if the positive measure $|\mu|$ is Regular. Here $|\mu|(E) := \sup_\Gamma \sum_{n=1}^{\infty} |\mu(E_n)|$, where $\Gamma = \{\{E_n\}_{n=1}^{\infty} : \sqcup_{n=1}^{\infty} E_n = E\}$.

Let K be any compact Hausdorff space and $M(K)$ be the space of all regular Borel Complex measure on K. For any $\mu \in M(K)$, define $||\mu|| := |\mu|(K)$. Then, $(M(K), ||.||)$ forms a normed linear space over $\mathbb{C}$. $\|\mu\|$ is called the *total variation* norm of μ.

Analogous structure can be obtained if the underlying scalar field is $\mathbb{R}$. $M(K)$ is the space of all finite *Signed measure* in this case.

C.1 Riesz Representation Theorem

Let $C(K)$ be the space of all Complex(or Real) valued continuous functions on K. The underlying scalar field would be understood from the context.

Theorem C.1.1 (Riesz Representation Theorem)**.** Let K be a compact Hausdorff space and $L : (C(K), ||.||_\infty) \to \mathbb{C}(\mathbb{R})$ be a bounded linear functional. Then, there exists a unique Complex (Signed) measure μ in $(K, \mathfrak{M})$ such that for all $f \in C(K)$,

$$L(f) = \int_K f d\mu.$$

Remark C.1.2. (a) The association $L \mapsto \mu$ is an isometric isomorphism.

(b) For $t \in K$, the *Dirac measure* $(\mu =)\delta_t \in M(K)$ and $\int_K f d\mu = f(t)$.

C.2 Probability measure

Definition C.2.1. (Probability measure) Let K be a compact Hausdroff space. A measure $\mu \in M(K)$ is called probability measure if μ is a positive measure and $||\mu|| = 1$.

Notation:

(a) Let $\mathscr{P}(K)$ denotes the set of all probability measures on K.

(b) Let $M^+(K)$ denotes the class of all positive measure in $(M(K), ||.||)$.

Proposition C.2.2. (a) $\mathscr{P}(K) = \{\nu \in M(K) : ||\nu|| = 1 \text{ and } \int_K 1 d\nu = 1\}$.

(b) $\mathscr{P}(K)$ is a w^*-compact convex subset of $M(K)$. $\mathscr{P}(K)$ is also a face of $B_{M(K)}$.

(c) For any $t \in K$, the Dirac measure δ_t is an extreme point of $\mathscr{P}(K)$.

(d) $ext(B_{M(K)}) = \{\alpha\delta_t : |\alpha| = 1,\ t \in K\}$.

Proof. (a). If $\nu \in \mathscr{P}(K)$, then nothing to prove.

Conversely, assume $\int_K 1 d\nu = 1$. Let $d\nu = hd|\nu|$ be the polar decomposition of ν, where h be a measurable function on K such that $|h| \equiv 1$ a.e $[\nu]$. This implies $\int_K 1 d\nu = \int_K hd|\nu|$. Thus,

$$\int_K h^+ d|\nu| - \int_K h^- d|\nu| = 1.$$

Now,

$$\left|\int_K h^+ d|\nu|\right| \leqslant \int_K |h^+| d|\nu| \leqslant \int_K 1 d|\nu| = |\nu|(K) = 1.$$

Since $\int_K h^+ d|\nu| > 0$, we have $\int_K h^+ d|\nu| \leqslant 1$. Also, $\int_K h^- d|\nu| \geqslant 0$. Therefore,

$$\int_K hd|\nu| = \int_K h^+ d|\nu| = 1 \text{ and } \int_K h^- d|\nu| = 0.$$

This implies $h^+ \equiv 1$ a.e. $[|\nu|]$ and $h^- \equiv 0$ a.e. $[|\nu|]$. Hence, $h = h^+ - h^- \equiv 1$ a.e. $[|\nu|]$. This implies $\nu = |\nu|$ Hence, ν is a positive measure and $||\nu|| = 1$.

(b). It follows from (a) that $\mathscr{P}(K)$ is w^*-compact. Let $\mu_1, \mu_2 \in B_{M(K)}$ such that $\lambda\mu_1 + (1-\lambda)\mu_2 \in \mathscr{P}(K)$, for some $\lambda \in (0,1)$. Since $||\lambda\mu_1 + (1-\lambda)\mu_2|| = 1$, it follows that $||\mu_{1]}|| = ||\mu_2|| = 1$. From result in (a),

$$\lambda \int_K d\mu_1 + (1-\lambda) \int_K d\mu_2 = 1.$$

Now,

$$\left| \int_K d\mu_1 \right| \leqslant \int_K d|\mu_1| = |\mu_1|(K) = ||\mu_1|| = 1.$$

Similarly, we will get, $\left|\int_K d\mu_2\right| \leqslant 1$. Hence,

$$\int_K d\mu_1 = \int_K d\mu_2 = 1.$$

This implies $\mu_1, \mu_2 \in \mathscr{P}(K)$. Therefore, $\mathscr{P}(K)$ is a face of $B_{M(K)}$.

(c). Let $\mu_1, \mu_2 \in \mathscr{P}(K)$ such that

$$\delta_t = \frac{\mu_1 + \mu_2}{2}.$$

Suppose μ_1 and μ_2 are not Dirac measures. This implies $\mathrm{S}(\mu_1)$ is not singleton. Hence, we can choose $s \in S(\mu_1)$ such that $s \neq t$. Then by Urysohn's lemma, there exists a continuous function $f : K \longrightarrow [0,1]$ such that $f(t) = 0$ and $f(s) = 1$. Now, since $s \in S(\mu_1)$ and $f(s) = 1$, $\mu_1(f) = \int_K f d\mu_1 > 0$. Also, since $\mu_2 \in \mathscr{P}(K)$, μ_2 is a positive measure. This implies $\int_K f d\mu_2 \geqslant 0$. Now, $\delta_t(f) = (\mu_1(f) + \mu_2(f))/2$. Hence,

$$0 = f(t) = \frac{1}{2}\left(\int_K f d\mu_1 + \int_K f d\mu_2 \right) > 0.$$

which is a contradiction. This implies there exists $x, y \in K$ such that $\mu_1 = \delta_x$ and $\mu_2 = \delta_y$.

Suppose $x \neq t$. Now,

$$\delta_t(\{t\}) = \frac{\delta_x(\{t\}) + \delta_y(\{t\})}{2}.$$

Hence, $\delta_y(\{t\}) = 2$, which is a contradiction. Similarly, we will arrive at a contradiction if $y \neq t$. Hence, $x = y = t$. Therefore, $\mu_1 = \delta_t$ and $\mu_2 = \delta_t$. Hence, the Dirac measure δ_t is an extreme point of $\mathscr{P}(K)$.

As $\mathscr{P}(K)$ is a face of $B_{M(K)}$, we have, $ext(\mathscr{P}(K)) \subseteq ext(B_{M(K)})$.

(d). First we prove that for any $\alpha \in \mathbb{F}$ such that $|\alpha| = 1$ and $t \in K$, $\mu := \alpha\delta_t$ is an extreme point of $B_{M(K)}$. Let $B := B_{M(K)}$. Suppose there exists $\nu_1, \nu_2 \in B$ such that for some $\lambda \in (0,1)$, $\mu = \lambda\nu_1 + (1-\lambda)\nu_2$. We want to prove that $\nu_1 = \nu_2 = \mu$.

Now, $||\nu_{1]}|| = ||\nu_2|| = 1$. By definition of μ,

$$\delta_t = |\mu| = |\lambda\nu_1 + (1-\lambda)\nu_2| \leqslant \lambda|\nu_1| + (1-\lambda)|\nu_2|.$$

Now, let $\nu := \lambda|\nu_1| + (1-\lambda)|\nu_2|$. Hence, $\delta_t \leqslant \nu$. In fact, $\delta_t = \nu$ because if E is any measurable set containing t, then $1 = \delta_t(E) \leqslant \nu(E) \leqslant 1$. On the other hand, if t is not in E, then $\nu(E) = \nu(K) - \nu(K\backslash E) = 0$. Hence, $\nu = \delta_t$, that is, $\delta_t = \lambda|\nu_1| + (1-\lambda)|\nu_2|$. Now, $|\nu_1|$ and $|\nu_2|$ are probability measures and δ_t is extreme point of $\mathscr{P}(K)$ (from result in (c)). Hence, $|\nu_1| = |\nu_2| = \delta_t$.

Now, for $i = 1, 2$, if E is any measurable set not containing t, then $|\nu_i(E)| \leqslant |\nu_i|(E) = \delta_t(E) = 0$. Hence, $\nu_i(E) = 0$. On the other hand, if E contains t, then $\nu_i(E) = \nu_i(K) - \nu_i(K\backslash E) = \nu_i(K)$. Therefore, we have $\nu_i = b_i\delta_t$, where $b_i = \nu_i(K)$ and $|\nu_i|(K) = |b_i|\delta_t(K)$ implies $|b_i| = 1$.

Hence, by our assumption, we get, $\alpha\delta_t = (\lambda b_1 + (1-\lambda)b_2)\delta_t$, which gives us $\alpha = \lambda b_1 + (1-\lambda)b_2$. Since $|b_1| = |b_2| = 1$ and $\lambda \in (0,1)$, this implies, $\alpha = b_1 = b_2$. Hence, we can conclude that $\nu_1 = \nu_2 = \alpha\delta_t$.

Conversely, let μ be any extreme point of $B_{M(K)}$. It is enough to show that S(μ) is singleton. Suppose it is not the case. Let $x, y \in S(\mu)$ such that $x \neq y$. Then, there exists open sets U and V such that $x \in U$ and $y \in V$ and $\overline{U} \cap \overline{V} = \phi$. By Urysohn's lemma, there exists a continuous function $f : K \longrightarrow [0,1]$ such that $f|_{\overline{U}} = 1$ and $f|_{\overline{V}} = 0$. Consider, $fd\mu$ and $(1-f)d\mu \in M(K)$. Let $\alpha = ||fd\mu||$. Then,

$$\alpha = \int_K |f|d|\mu| = \int_K fd|\mu| \leqslant \int_K 1d|\mu| = |\mu|(K) = ||\mu|| \leqslant 1.$$

Now, consider, $1 - \alpha = 1 - \int_K fd\mu = \int_K (1-f)d\mu = ||(1-f)d\mu||$. Now,

$$\alpha = \int_K fd|\mu| \geqslant \int_U fd|\mu| = \int_U d|\mu| = |\mu|(U) > 0.$$

Also, $1 - \alpha = \int_K (1-f)d|\mu| \geqslant \int_V (1-f)d|\mu| = \int_V 1d|\mu| = |\mu|(V) > 0$. Thus, $\alpha < 1$. Hence, $0 < \alpha < 1$. Therefore, $\frac{fd\mu}{\alpha}, \frac{(1-f)d\mu}{\alpha} \in B_{M(K)}$. Now,

$$\mu = \alpha\left(\frac{fd\mu}{\alpha}\right) + (1-\alpha)\left(\frac{(1-f)d\mu}{\alpha}\right).$$

Since μ is an extreme point of $B_{M(K)}$, we have, $\mu = \frac{fd\mu}{\alpha} = \frac{(1-f)d\mu}{\alpha}$. Thus, we get, $\alpha d\mu = fd\mu$. This implies $f \equiv \alpha < 1$ a.e. $[\mu]$. But, $f \equiv 1$ on U and $|\mu|(U) > 0$, which is not possible. Hence, $S(\mu)$ is singleton. Thus, $\mu = \alpha\delta_t$, for some $\alpha \in \mathbb{F}$ such that $|\alpha| = 1$ and $t \in K$. □

Appendix D

Preliminaries on Vector Lattice

Definition D.0.1. (*a*) A vector space V over $\mathbb{R}$ is said to be an ordered vector space if for any $x, y, z \in V$ and $x \leqslant y$ imply $x + z \leqslant y + z$. If $\lambda \geqslant 0$ then $x \leqslant y$ implies $\lambda x \leqslant \lambda y$.

(*b*) An ordered set $(L, \leqslant)$ is called a *lattice* if for each pair $(x, y) \in L \times L$, there exist elements $x \vee y := lub\{x, y\}$ and $x \wedge y := glb\{x, y\}$ exist in L. If, in addition, the distributive law

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z) \tag{D.1}$$

is satisfied for all $x, y \in L$, L is called a distributive lattice.

A vector lattice is an ordered vector space which satisfies the condition (*b*). It is due to Birkhoff that the condition D.1 is equivalent to $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$.

If $(L, \leqslant)$ is a lattice, the mappings $(x, y) \mapsto x \vee y$ and $(x, y) \mapsto x \wedge y$ are usually called the lattice operations. As laws of composition they are idempotent, associative, commutative, and satisfy $x \wedge (x \vee y) = x$ as well as $x \vee (x \wedge y) = x$. On the other hand it is not difficult to verify that if a non-void set L is endowed with two laws of composition having these properties, then $x \geqslant y$ if and only if $x \vee y = y$ defines an ordering under which L is a lattice in the sense of Definition D.0.1. Recall also that a lattice L is called (countably) complete if every (countable) subset of L possesses a least upper bound and a greatest lower bound.

Let L be a lattice, a subset L_0 closed under the lattice operations is called a sublattice of L. L_0 is called a (countably) complete sublattice of L if L_0 is closed under the formation of arbitrary (countable) infima and suprema. However, infima or suprema of a sublattice may not be the same as those of the lattice.

Example D.0.2. If M is a non-void set and $(L, \leqslant)$ is a lattice. The set L^M of all mappings $f : M \to L$ is a lattice under the canonical ordering defined by:$f \leqslant g$ if and only if $f(t) \leqslant g(t)$ for all $t \in M$.

Definition D.0.3. Let E be a vector lattice. For all $x \in E$, we define $x^+ := x \vee 0$, $x^- := (-x) \vee 0$, $|x| := x + (-x)$. x^+, x^- and $|x|$ are called the positive part, the negative part, and the modulus (or absolute value,) of x, respectively.

D.1 Basic Properties

Theorem D.1.1. Let E be a vector lattice and $x, y \in E$ and $\lambda \in \mathbb{R}$. Then the following are true.

(*a*) $x = x^+ - x^-$.

(*b*) $|x| = x^+ + x^-$.

(*c*) $|\lambda x| = |\lambda||x|$, $|x + y| \leqslant |x| + |y|$.

(*d*) $x + y = (x \vee y) + (x \wedge y)$, $|x - y| = (x \vee y) - (x \wedge y)$.

Definition D.1.2 (Base of a cone)**.** Let X be a compact convex subset of a cone P in a lctvs E. X is said to be a base for the cone P if $P = \{tx : t \geqslant 0, x \in X\}$.

Note D.1.3. If P is a cone with base X, then we denote P as $\widetilde{X}$.

In the rest of this Section, we will assume that P is a cone with vertex at the origin and a base X in a lctvs E such that X is contained in some hyperplane $H = \{x \in E : L(x) = r\}$, for some $L \in E^*, r \in \mathbb{R}$, which misses the origin.

More precisely, $P = \{\alpha x : x \in X, \alpha \geqslant 0\}$. Note that there is no generality lost in making this assumption since we may embed E as a hyperplane $E \times \{1\}$ in $E \times \mathbb{R}$(with product topology); the image of $X \times \{1\}$ is affinely homeomorphic with X. Due to this assumption, we have uniqueness as follows: $y \in P$ if and only if there exists a unique $\alpha \geqslant 0$ and $x \in X$ such that $y = \alpha x$.

Proposition D.1.4. Let P be a cone in a lctvs E and X be its base.

(*a*) $P \cap (-P) = \{0\}$

(*b*) $P \cap (-P) = \{0\}$ if and only if given $x, y \in E$ such that $x \prec y$ and $y \prec x$ implies $x = y$, in other words, $\prec$ is antisymmetric.

Proof. (*a*) Suppose $0 \neq z \in P \cap (-P)$. Then there exists $t, s > 0$ and $x, y \in X$ such that $z = tx = -sy$. This can be written as

$$\frac{t}{t+s}x + \frac{s}{t+s}y = 0 \in X.$$

which is a contradiction. Therefore, $P \cap (-P) = \{0\}$.

(*b*) Suppose $P \cap (-P) = \{0\}$. Let $x, y \in E$ such that $x \prec y$ and $y \prec x$. This implies $y - x, x - y \in P$ and hence $y - x \in P \cap (-P) = \{0\}$, which implies $x = y$. Now suppose $\prec$ is antisymmetric. Let $z \in P \cap (-P)$. This implies $z = tx = -sy$, for some $t, s \geqslant 0; x, y \in X$. Hence, $tx + sy = 0$. Now, $z - 0 = tx - (tx + sy) \in P$ and $0 - z = (tx + sy) - sy \in P$. This implies $z \prec 0$ and $0 \prec z$ and hence $z = 0$.

□

Let us now look at few elementary properties of a lattice.

Proposition D.1.5. .

(*a*) Let $(E, \prec)$ be a lattice where $\prec$ is the ordering induced by the cone P. Suppose $x, y \in E$, then $-(-x \wedge -y)$ is the least upper bound (lub for short) of x and y.

(*b*) Let P be a cone in a real lctvs E. Then $P - P = \{x - y : x, y \in P\}$ is a subspace of E.

(*c*) Let P be a cone in E such that $P \cap (-P) = \{0\}$. $(P, \prec)$ is a lattice w.r.t. the partial ordering induced by P if and only if $P - P$ is a lattice subspace of E.

Proof. (*a*) Clearly, $(-x \wedge -y) \prec -x$ and $(-x \wedge -y) \prec -y$. This implies $-x - (-x \wedge -y) \in P$, which can be written as $-(-x \wedge -y) - x \in P$. Therefore, $x \prec -(-x \wedge -y)$. Similarly, $y \prec -(-x \wedge -y)$. Let $z \in P$ such that $x \prec z$ and $y \prec z$. It follows that $-z \prec -x$ and $-z \prec -y$ and hence, $-z \prec (-x \wedge -y)$. This implies $z - [-(-x \wedge -y)] = (-x \wedge -y) - (-z) \in P$ and thus $-(-x \wedge -y) \prec z$. Therefore, $-(-x \wedge -y)$ is the lub of x and y.

(*b*) Let $x, y \in P - P$. Then there exists $x_1, x_2, y_1, y_2 \in P$ such that $x = x_1 - x_2$ and $y = y_1 - y_2$. Now, since $x_1 + y_1, x_2 + y_2 \in P$, $x + y = x_1 - x_2 + y_1 - y_2 = (x_1 + y_1) - (x_2 + y_2) \in P - P$. Let $\alpha \geqslant 0$. Then $\alpha x = \alpha(x_1 - x_2) = \alpha x_1 - \alpha x_2 \in P - P$. Let $\alpha < 0$. Hence, $\alpha = -\beta$, for some $\beta > 0$. We then have $\alpha x = -\beta(x_1 - x_2) = \beta x_2 - \beta x_1 \in P - P$. Therefore, $P - P$ is a subspace of E.

(*c*) Assume P is a lattice. Let $x = x_1 - x_2, y = y_1 - y_2 \in P - P$, where $x_1, x_2, y_1, y_2 \in P$. Let $z = (x_1 + y_2) \vee (y_1 + x_2) - (x_2 + y_2)$. Then $z - x = (x_1 + y_2) \vee (y_1 + x_2) - (x_1 + y_2) \in P$, which implies $x \prec z$. Similarly, $y \prec z$. Let $w = w_1 - w_2 \in P - P$ such that $x \prec w$

and $y \prec w$. Now, $x_1 - x_2 \prec w_1 - w_2$ implies $x_1 + w_2 + y_2 \prec w_1 + x_2 + y_2$ and $y_1 - y_2 \prec w_1 - w_2$ implies $y_1 + w_2 + x_2 \prec w_1 + y_2 + x_2$. Then $w - z = (w_1 + x_2 + y_2) - [(w_2 + x_1 + y_2) \vee (w_2 + y_1 + x_2)] \in P$. Therefore, $z \prec w$. Hence, z is the lub for x, y in $P - P$. $P - P$ is a subspace of E by the earlier remark. Conversely, $P - \{0\} \subset P - P$. Since $P - P$ is a lattice subspace of E, so is $P - \{0\}$. Hence P is a lattice.

□

Proposition D.1.6. Let $(E, \leqslant)$ be a lattice, where $\leqslant$ is induced by a cone P. Then,

(i) For any $a, b, c \in E$, $(a \wedge b) + c = (a + c) \wedge (b + c)$

(ii) For $0 \leqslant a, b, c \in E$, $(a + b) \wedge c \leqslant (a \wedge c) + (b \wedge c)$.

(iii) If $\{a_i : i \in I\}$ and $\{b_j : j \in J\}$ are finite sequences of non-negative elements of E, and if $\sum_{i \in I} a_i = \sum_{j \in J} b_j$ then there exists $z_{ij} \geqslant 0$, $(i, j) \in I \times J$ such that $a_i = \sum_{j \in J} z_{ij}$ $(i \in I)$ and $b_j = \sum_{i \in I} z_{ij}$.

Proof. (i). Let $a, b, c \in E$. Then $(a + c) - (a \wedge b + c) = a - (a \wedge b) \in P$. Similarly, we get $(b + c) - (a \wedge b + c) \in P$. Therefore, $(a \wedge b) + c \leqslant (a + c) \wedge (b + c)$. We now prove the other inequality. We know $a \geqslant (a + c) \wedge (b + c) - c$ and $b \geqslant (a + c) \wedge (b + c) - c$. Therefore $a \wedge b \geqslant (a + c) \wedge (b + c) - c$. This implies $a \wedge b + c \geqslant (a + c) \wedge (b + c)$.

(ii). Let $a, b, c \geqslant 0$. Let $u = (a + b) \wedge c$. Since $a \geqslant 0$, $u \leqslant a + c$ ($a + c \geqslant c \geqslant u$ implies $(a + c - c) + (c - u) \in P$). Thus, $u \leqslant (a + b) \wedge (a + c) = b \wedge c + a$, since $u \leqslant a + b$ and from (i). Also, $u \leqslant c + (b \wedge c)$ (since $b \wedge c \geqslant 0$). Therefore, $u \leqslant [c + (b \wedge c)] \wedge [a + (b \wedge c)] = (a \wedge c) + (b \wedge c)$. This implies $(a + b) \wedge c \leqslant (a \wedge c) + (b \wedge c)$.

(iii). We can easily use induction on the number of elements in I and in J and reduce the proof to the case $I = J = \{1, 2\}$. Assume $a_1 + a_2 = b_1 + b_2$. Define $z_{11} = a_1 \wedge b_1 \geqslant 0$, $z_{12} = a_1 - z_{11} \geqslant 0$ and $z_{21} = b_1 - z_{11} \geqslant 0$. We will now choose $z_{22} \geqslant 0$ such that $z_{22} = a_2 - z_{21} = b_2 - z_{12}$. We now claim that $a_2 - z_{21} = b_2 - z_{12}$, that is, $b_2 - a_1 + z_{11} = a_2 - b_1 + z_{11}$. This is clearly true because of the assumption that $a_1 + a_2 = b_1 + b_2$. Clearly, $a_1 \geqslant a_1 - b_2$ and $b_1 \geqslant a_1 - b_2$. This implies $z_{11} \geqslant a_1 - b_2$ and hence $b_2 - a_1 + z_{11} \geqslant 0$. Therefore, $z_{22} \geqslant 0$. This completes the proof. □

Bibliography

[1] Erik M. Alfsen, *Compact convex sets and boundary integrals*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 57. Springer-Verlag, New York-Heidelberg, 1971. x+210 pp.

[2] L. Asimow and A. J. Ellis, *Convexity Theory and its applications in Functional Analysis*, London Mathematical Society Monographs, 16. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1980. x+266.

[3] H. G. Eggleston, *Convexity*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 47 Cambridge University Press, New York 1958 viii+136 pp.

[4] Michel Grabisch, Ivan Kojadinovic, Patrick Meyer, *A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package* European J. Operational Research, vol 186, **2** 766-785.

[5] Michel Grabisch, Marc Roubens, *Application of the Choquet Integral in Multicriteria Decision Making.* http://metronu.ulb.ac.be/npauly/art_2014_2015/grabisch_2000.pdf

[6] Michel Grabisch and Christophe Labreuche, *A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid*, Annals of Operations Research, **175** 1 247–286.

[7] Steven G. Krantz, *Convex Analysis*, CRC Press, Boca Raton, FL, 2015. xiv+161.

[8] S. R. Lay, *Convex sets and their applications*, Wiley, New Yark 1958.

[9] R. R. Phelps, *Lectures on Choquet's Theorems* Second edition. Lecture Notes in Mathematics, 1757. Springer-Verlag, Berlin, 2001. viii+124.

[10] R. T. Rockefeller, *Convex Analysis*, Princeton University Press,1970.

[11] J. Stoer and C. Witzgall, *Convexity and Optimization on Finite Dimensions*, Springer-Verlag, Heidelberg, New York 1970.

[12] M. Stone, *The Generalized Weistrass Approximation Theorem*, Mathematics Magazine, **21** 4 (Mar.–Apr., 1948), 167–184.

[13] Mitsuru Uchiyama, *Proofs of Korovkin's Theorems via Inequalities*, Amer. Math. Monthly, 110 **4**, 334–336.

[14] F. A. Valentine, *Convex Sets*, Mcgraw Hill,1964.