UiO : Department of Mathematics
University of Oslo

On Markov operators and cones

Stefan Ivkovic

Master's Thesis, Autumn 2015

Acknowledgment

First of all I wish to thank to my parents for all their love and support and I dedicate this thesis to them. Also I wish to thank prof. Ivan Milic for technical consultation regarding LaTex.

Contents

1 Introduction

In this thesis we will consider Markov operators on cones. More precisely, we let $(X, ||\cdot||)$ be real Banach space, $K \subseteq X$ be a closed and normal cone with nonempty interior and $e \in \text{Int}K$ an order unit. (The notations and notions used here will be detailed in section 1).

A bounded, linear operator $T : X \to X$ is a Markov operator w.r.t. K and e if $T(K) \subseteq K$ and $T(e) = e$. We consider then the adjoint of T and a homogeneous discrete time Markov system given by

$$\pi_{k+1} = T^*(\pi_k),\ \ k = 0, 1, 2, ...$$

where $\pi_0 \in X^*$, is s.t. $\pi_0(x) \geq 0$ for all $x \in K$ and $\pi_0(e) = 1$. The final goal of the theoretical part of this thesis which includes the first 6 sections is to give a proper conditions on T that will guarantee the convergence of the Markov system given above to a unique invariant measure. These conditions are given in the theorem 6.1 in section 6, but before that, we will need to develop certain theory and introduce some new concepts, tools and definitions. For instance, we consider the quotient space $X/\mathbb{R}e$, define a norm $|||\cdot|||_H$ on this space and let

$$\tilde{T} : X/\mathbb{R}e \to X/\mathbb{R}e$$

be the induced linear map given by

$$\tilde{T}(x + \mathbb{R}e)) = T(x) + \mathbb{R}e$$

for all $x \in X$. Furthermore, we consider the annihilator of $\mathbb{R}e$ in X^* denoted by $M(e)$ and define a norm $||\cdot||_H^*$ on $M(e)$. We also show that the dual of $((X/\mathbb{R}e), |||\cdot|||_H)$, is isometrically isomorphic to $(M(e), ||\cdot||_H^*)$. The theorem 6.1 states then that if $|||\tilde{T}|||_H \leq 1$, then there exists $\pi \in P(e)$ s.t.

$$||(T^*)^n(\mu) - \pi||_H^* \leq |||\tilde{T}|||_H^n$$

for all n and all $\mu \in P(e)$ where

$$P(e) = \{\mu \in X^* \mid \mu(e) = 1 \text{ and } \mu(x) \geq 0\ \forall x \in K\}.$$

As $|||\tilde{T}|||_H < 1$, this in particular means that $((T^*)^n(\mu) - \pi)$ converges to 0, as $n \to \infty$, since $|||\tilde{T}|||_H^n$, converges to 0, as $n \to \infty$. This is a very important result and in sections 7,8 and 9, we will apply this result to the different concrete cases when X is equal to $\mathbb{R}^n$, $C_{\mathbb{R}}(\Omega)$ and S_n respectively.

Now, in order to apply theorem 6.1 to different examples, we need that $|||\tilde{T}|||_H < 1$, hence we must be able to somehow calculate an estimate on $|||\tilde{T}|||_H$. The theorem 6.2 is section 6 gives an expression for $|||\tilde{T}|||_H$ in terms of disjoint, extreme points of $P(e)$, and we are going to use this expression in sections 7,8 and 9 to calculate\estimate $|||\tilde{T}|||_H$ in concrete examples. ("disjointness" of elements in $P(e)$ will be defined in section 4).

It is well known that a Markov chain with n-states can be described by an $n \times n$ column stochastic matrix P where $p_{i,j}$ denotes the probability to move from the state j to the state i in one step. Such matrix P is then called the transition matrix for the Markov chain. The most standard example of a Markov operator is therefore the operator $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = A(x)$ for all $x \in \mathbb{R}^n$ where A is an $n \times n$ row stochastic matrix. Its adjoint is then $T^n : \mathbb{R}^n \to \mathbb{R}^n$ given by $T^*(x) = A^t x$ for all $x \in \mathbb{R}$ and, as A is $n \times n$ row stochastic matrix, A^t is as an $n \times n$ column stochastic matrix.Thus A^t is a transition matrix for a Markov chain.We are interested in convergence of a Markov system $x_{k+1} = A^t x_k$ to some unique stochastic vector $\pi \in \mathbb{R}^n$ where the initial vector x_0 is a stochastic vector in $\mathbb{R}$ and this is what we are going to study in section 7. From Perron Frobenius theorem we know that the Markov system given above will be convergent if A^t is regular. The less known fact is that the system will be convergent if and only if A^t semiregular, that is if there exists some $k \in \mathbb{N}$ s.t. $A^k(A^k)^t$ has only positive coefficients. This is for instance stated in [ABS] in the theorem 1.1. In section 7 we will show that this theorem is a direct consequence of the theorem 6.1 given in this thesis.
More precisely, in section 7 we let $X = \mathbb{R}^n$, $K = \mathbb{R}^n_+$ and $e = \vec{1} = (1, 1, ..., 1)$. Then, the operator $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = Ax$ for all $x \in \mathbb{R}^n$ will be a Markov operator w.r.t. K and e when A is an $n \times n$ row stochastic matrix. In lemma 7.1 we will first show, using the theorem 6.2, that

$$|||\tilde{T}|||_H \ = \ 1 - \min_{i<j} \sum_{k=1}^{n} \min\{a_{i,k}, a_{i,j}\}.$$

in this case. The expression

$$1 - \min_{i<j} \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}$$

is, for instance knows as Dobrushin ergodicity coefficient. We will then use this expression to show that $|||\tilde{T}|||_H < 1$ if and only if AA^T has only positive coefficients. Thus, for A^T being semiregular is equivalent to that there exists

$k \in \mathbb{N}$ s.t. $|||\tilde{T}^k|||_H < 1$. Relating this to the theorem 1.1. in [ABS], we see that the Markov system given by $T^*(x_k) = x_{k+1}$, converges some unique stochastic vector in $\mathbb{R}^n$ if and only if there exists $k \in \mathbb{N}$ s.t $|||\tilde{T}^k|||_H < 1$. This is stated in theorem 7.2 in section 7 and the proof of this theorem is just an application of theorem 6.1, as mentioned. At the end of section 7, in proposition 7.3, we show that Doeblin contraction coefficient is equal to Dobrushin ergodicity coefficient, that is

$$\frac{1}{2}\max_{i<j}\sum_{1\leq k\leq n}|a_{i,k}-a_{j,k}| = 1-\min_{i<j}\sum_{k=1}^{n}\min\{a_{i,k},a_{j,k}\}.$$

Another interesting application of the theory in this thesis is in the case when $X = C_{\mathbb{R}}(\Omega)$. Here $C_{\mathbb{R}}(\Omega)$ denotes the space of all continuous real valued functions, on Ω, where Ω is a compact, Hausdorff topological space, and $C_{\mathbb{R}}(\Omega)$ is equipped with supremums norm, $||\cdot||_\infty$. In remark 2.4, section 2, we show that the dual of $(C_{\mathbb{R}}(\Omega), ||\cdot||_\infty)$ is the space of all signed Radon measures on Ω equipped with the total variation norm, $(M_r(\Omega), ||\cdot||)$. We let in this case K be the cone in $C_{\mathbb{R}}(\Omega)$ consisting of all nonnegative functions on Ω and the constant function 1 on Ω be the order unit. In section 8 we consider then Markov operators $C_{\mathbb{R}}(\Omega)$ w.r.t 1 and K.There are 2 such examples in section 8, example 8.1 and example 8.2. In example 8.1, we let μ be a nonzero Radon measure on Ω and we choose a nonnegative continuous functions $\tilde{k}$ on $\Omega \times \Omega$ such that

$$\int_\Omega \tilde{k}(x,y)\ d\mu(y) = 1$$

for all $x \in \Omega$. Then we let

$$T_{\tilde{k}} : C_{\mathbb{R}}(\Omega) \to C_{\mathbb{R}}(\Omega)$$

be s.t. for all $f \in C_{\mathbb{R}}(\Omega)$, $T_{\tilde{k}}(f)$ is the function given by

$$T_{\tilde{k}}(f)(x) = \int_\Omega \tilde{k}(x,y)\ f(y)\ d\mu(y)$$

for all $x \in \Omega$. The operator $T_{\tilde{k}}$ is hence a Markov operator on $C_{\mathbb{R}}(\Omega)$ w.r.t. 1 and K. We show then that if

$$||\tilde{k}||_\infty < \frac{2}{\mu(\Omega)},$$

then $|||\tilde{T_{\tilde{k}}}|||_H < 1$, so that we can apply theorem 6.1. In example 8.2 we choose $w : \Omega \to \Omega$ to be a continuons map and we let

$$T_w : C_{\mathbb{R}}(\Omega) \to C_{\mathbb{R}}(\Omega)$$

be given by $T_w(f) = f \circ w$. Also in this case, T_w is a Markov operator on $C_{\mathbb{R}}(\Omega)$ w.r.t. 1 and K. However, we show that $|||\tilde{T}_w||| \geq 1$, so we can not apply theorem 6.1. in this example.

In section 9, which is the last section in this thesis, we let $X = S_n$ be the space of all Hermitian $n \times n$ matrices equipped with the operator norm, $K = S_n^+$ be the cone in X consisting of all positive semidefinite $n \times n$ matrices and I_n be the order unit. We consider then the Kraus map $\Phi : S_n \to S_n$ given by

$$\Phi(A) = \sum_{i=1}^{m} V_i^* A V_i$$

for all $A \in S_n$ where $\sum_{i=1}^{m} V_i^* V_i = I_n$. The operator Φ is then the Markov operator on S_n w.r.t. I_n and K and its adjoint $\Psi : S_n \to S_n$ is given by

$$\Psi(A) = \sum_{i=1}^{m} V_i A V_i^*$$

for all $A \in S_n$. We will show that $|||\tilde{\Phi}|||_H \leq 1$, if and only if there are no nonzero vectors $\mu, v \in \mathbb{C}^n$ with the property that $< V_i u, V_j v > = 0$ for all $i, j \in \{1, ..., m\}$. In that case, we can hence apply theorem 6.1. Furthermore, in theorem 9.7 at the end of section 9, we show that the Markov system

$$\Pi_{k+1} = \Psi(\Pi)_k)$$

converges to a unique invariant measure if and only if there exists $k_0 \in \mathbb{N}$ s.t. $|||\tilde{\Phi}^{k_0}|||_H < 1$. This is actually quite similar to theorem 7.2 mentioned before in the introduction, which states that the Markov system

$$x_{k+1} = T^*(x_k)$$

converges to a unique stochastic vector in $\mathbb{R}^n$ if and only if there exists $k_0 \in \mathbb{N}$ s.t.

$$|||\tilde{T^{k_0}}||| < 1.$$

The proof of theorem 9.7 also applies theorem 6.1, but is somewhat different than the proof of theorem 7.2.
Another important topic of this thesis is to study the convergence of the system

$$x_{k+1} = T(x_k), \quad k = 0, 1, ...$$

We will show in theorem 6.1 that if $|||\tilde{T}|||_H < 1$, then there exist $\pi \in P(e)$ s.t.

$$||T^n(X) - <\pi, x> e||_H \leq |||\tilde{T}|||_H^n \; ||x||_H$$

for all $x \in X$, where $|| \cdot ||$ is a certain seminorm on X that depends on e.
(We will define this seminorm in section 1.)
Since $|||\tilde{T}|||_H < 1$ in this case, it follows that $T^n(x)$ converges to $<\pi, x> e$, w.r.t. $|| \cdot ||_H$ as $n \to \infty$ for all $x \in X$.
In others words, the system

$$x_{k+1} = T(x_k), \quad k = 0, 1, ...$$

converges w.r.t $|| \cdot ||_H$ to a scalar multiple of e, so called "consensus state". This is also a very important result and we will apply this in theorem 9.6 to show that every orbit of the system given by

$$X_{k+1} = \Phi(X_k), k = 0, 1, ...$$

converges to on equibrilium co- linear to I_n if and only if there exists $k_0 \in \mathbb{N}$ s.t. $|||\tilde{\Phi}^{k_0}||| < 1$, where $\Phi : S_n \to S_n$ is a Kraus map as described earlier in this introduction.
This thesis builds on the 3. version of the article "Dobrushin ergodicity coefficient for Markov operators on cones" writen by Stephane Guobert and Zheng Qu. As it was written by my supervisor, prof. Erik C. Bedos , in the project description for this thesis, the main aim of the thesis is to give on improved and detailed presentation of the above mentioned article of Gaubert and Qu.
The 3. version of the above mentioned article by Gaubert and Qu will be denoted by [GQ] throughout this thesis.

2 Preliminaries - Cones on vector spaces, order units, Thompson's norm and Hilbert's seminorm

In this section we will give some definitions and introduce concepts like cones on vector spaces, order units, Thompson's norm and Hilbert's seminorm. These concepts will be fundamental for the theory which will be developed and discussed later in the thesis. After introducing these concepts, we will give some examples where these concepts occur. The examples 1.4 and 1.5 in this section are the same as examples 2.2 and 2.3 in [GQ] respectively. However we give here detailed proofs and explanations for all the statements used in those examples. In addition, we give an example that is not in [GQ]. It is example 1.6 it this section and deals with the applications of the above mentioned concepts on the space $\mathbb{C}_{\mathbb{R}}(\Omega)$, which is the space of all continuous real valued functions on Ω where Ω is compact, Hausdorff topological space.

Let X be a real vector space, let $K \subseteq X,\ K \neq \emptyset$. We say that K is a cone in X if K satisfies the following properties:

1. $K + K \subseteq K$
2. $\lambda K \subseteq K$ for all $\lambda > 0$
3. $K \cap (-K) = \{0\}$.

Let $\leq$ denote the associated order on X, so $x \leq y \Leftrightarrow (y - x) \in K$. For $u \in K$, set

$$I_u = \{x \in X \mid -u \leq x \leq u\}.$$

Recall that $A \subset X$ is called absorbing if

$$\{t > 0 \mid x \in tA\} \neq \emptyset$$

for every $x \in X$. An element $u \in K$ is called an order unit for (X, K) if I_u is absorbing.

Furthermore, for $x \leq y$, we define the order interval

$$[x, y] := \{z \in X \mid x \leq z \leq y\}.$$

If Int$K \neq \emptyset$ for $x \in X$ and $y \in \text{Int}(K)$ we define then

$$M(x/y) := \inf\{t \in \mathbb{R} \ : \ x \leq ty\}$$

$$\text{and } m(x/y) := \sup\{t \in \mathbb{R} \ : \ x \geq ty\}.$$

For an order unit $u \in$ K, we define a Thompson's norm w.r.t. u to be given by $||x||_T := \max\{M(x/u), -m(x/u)\}$ for all $x \in X$.

We also define a Hilbert's seminorm w.r.t. u to be given by

$$||x||_H = M(x/u) - m(x/u).$$

We wish to prove that if $(X, ||\cdot||)$ is a real Banach space and K satisfies certain properties, then $||\cdot||$ and $||\cdot||_T$ are equivalent. To prove this, we will first state and prove 2 observations, before we give and prove the main proposition, the proposition 1.2, where for instance this result is stated.

Observation 1 Let $u \in K \backslash \{0\}$.

Then u is an order unit if and only if u is an "internal point" of K (that is $\forall x \in X \ \exists \delta > 0$ s.t. $(u + \lambda x) \in K \ \forall x \in [-\delta, \delta]$)

Observation 2 For $u \in K \setminus \{0\}$, set

$$X_u = \bigcup_{t \geq 0} tI_u \ , \ K_u = K \cap X_u.$$

Then X_u is a subspace of X, I_u is an absorbing, absolutely convex subset of X_u, K_u is a cone of X_u and u is an order unit for (X_u, K_u).
Let $|\cdot|_u$ denote the Minkowski functional on X_u associated with I_u, so $|x|_u = \inf\{t > 0 \mid x \in tI_u\}$. Then $|\cdot|_u$ is a seminorm on X_u s.t. $|u|_u = |-u| = 1$.

Definition 1.1 Let $(X, ||\cdot||)$ be a real normed space, let K be a cone in X. We say that K in normal if there is some constant $M > 0$ s.t $||x|| \leq M||y||$ for all $x, y \in K$ satisfying $x \leq y$.

Proposition 1.2 *Assume that $(X, ||\cdot||)$ is a real normed space, K is closed and normal cone in X.Let $u \in K \setminus \{0\}$. Then we have the following*

1. $|\cdot|_u$ *is a norm on* X_u *and* $||\cdot|| \leq \alpha|\cdot|_u$ *on* X_u *for some* $\alpha > 0$

2. *If* $(X, ||\cdot||)$ *is a Banach space, then* $(X_u, |\cdot|_u)$ *is a Banach space.*

3. *If* u *is an order-unit in* (X, K)*, then* $X = X_u$.

4. *If* $u \in \text{Int}(K)$*, where* $\text{Int}(K)$ *is the interior of* K *in topological sense, then* u *is order-unit in* (X, K) *and* $||\cdot||$, $|\cdot|_u$ *are equivalent norms on* X.

Moreover $|\cdot|_u$ *is the same as Thomson's norm on* X *that we defined.*

Proof of observation 1. We observe first that K is convex:
If $x, y \in K$, and $\delta \in [0, 1]$, then $\delta x \in K$ and $(1-\delta)y \in K$ since $tK \subseteq K$ for all $t > 0$. Then $(\delta x + (1-\delta)y) \in K$. since $K + K \subseteq K$.

Let now $x \in X$ and assume that I_u is absorbing. Since I_u is absorbing, there exists $t > 0$ s.t. $x \in tI_u$. This means that $-tu \leq x \leq tu$, so we have that $tu - x \in K$ and $x + tu \in K$. It gives that $u - \frac{1}{t}x \in K$ and $u + \frac{1}{t}x \in K$.

Let now $\lambda \in \left[0, \frac{1}{t}\right]$.

Since $u + \lambda x$ lies in the segment between u and $u + \frac{1}{t}x$, we have
$(u + \lambda x) \in K$.

More precisely, as $0 < \lambda t < 1$, we have $(1 - \lambda t)u \in K$ because $1 - \lambda t > 0$ and $u \in K$.

Hence

$$(1 - \lambda t)u + \lambda t(u + \tfrac{1}{t}x) \in K.$$

Thus $(u + \lambda x) = (1 - \lambda t)u + \lambda t(u + \frac{1}{t}x) \quad \in K$.

Similarly if $-\frac{1}{t} \leq \lambda \leq 0$, then $u + \lambda x$ lies is the segment between $u - \frac{1}{t}x$ and u. Hence $(u + \lambda x) \in K$ since K is convex.

Since x was and arbitrary, this shows that u is an internal point of K.

Assume now that u is an internal point of K, that is,

$$\forall x \in X \ \exists \delta_x > 0$$

$$\text{s.t. } (u + \lambda x) \in K \ \ \forall x \in [-\delta_x, \ \delta_x]$$

Then in particular $u + \delta_x x \in K$ and $u - \delta_x x$ are in K.

Let $t_x = \frac{1}{\delta_x}$ Then $-u \leq \frac{1}{t_x} x \leq u$, so $x \in t_x I_u$.

This shows that for all $x \in X$, we have $\{t > 0 \mid x \in tI_u\} \neq \emptyset$.

Hence I_u is absorbing. This complets the proof of observation 1.

Proof of observation 2. X_u is a subspace of X:

Let $x, y \in X_u$. Then there exist r, s >0 s.t. $x \in rI_u, y \in sI_u$.

Hence

$$-u \leq \tfrac{1}{r} x \leq u \text{ and } -u \leq \tfrac{1}{s} y \leq u.$$

This means that

$$u - \tfrac{1}{r} x \in K,$$

$$u + \tfrac{1}{r} x \in K,$$

$$u - \tfrac{1}{s} y \in K$$

$$\text{and } u + \tfrac{1}{s} y \in K.$$

Since K is a cone, we must have

$$\tfrac{r}{r+s}(u - \tfrac{1}{r} x) \in K$$

$$\text{and } \tfrac{s}{r+s}(u - \tfrac{1}{s} y) \in K.$$

As $K + K \subseteq K$, we get that

$$\tfrac{r}{r+s}(u - \tfrac{1}{r} x) + \tfrac{s}{r+s}(u - \tfrac{1}{s} y) \in K,$$

$$\text{so } (u - \tfrac{1}{r+s}(x + y)) \in K.$$

In the same way , we can show that

$$(u + \tfrac{1}{r+s}(x + y)) \in K.$$

Hence

$$-u \leq \frac{1}{(r+s)}(x+y) \leq u,$$

which gives that $(x+y) \in (r+s)I_u$.

Thus $(x+y) \in X_u$.

Let $x \in X_u$ and consider cx for some $c \in \mathbb{R}$.

a) If $c > 0$:

Since $x \in X_u$ then $x \in tI_u$ for some $t > 0$.

Thus $(u - \frac{1}{t}x) \in K$

and $(u + \frac{1}{t}x) \in K$.

Hence

$$(u - \frac{1}{ct}cx) \in K$$

and

$$(u + \frac{1}{ct}cx) \in K.$$

This gives that

$$-u \leq \frac{1}{ct}cx \leq u.$$

So

$$cx \in ctI_u.$$

Hence

$cx \in X_u$ since $ct > 0$.

b)If $c < 0$:

Since $(u + \frac{1}{t}x) \in K$, we have

$$(u + \frac{1}{|c|t}|c|x) \in K,$$

that is,

$(u - \frac{cx}{|c|t}) \in K$ because $|c| = -c$.

Also, since

$$(u - \tfrac{1}{t}x) \in K, \text{ we get that}$$

$$(u - \tfrac{1}{|c|t}|c|x) \in K$$

$$\text{so } (u + \tfrac{cx}{|c|t}) \in K \text{ (because } c = -|c| \text{).}$$

Hence

$$-u \leq \tfrac{1}{|c|t}cx \leq u$$

which means that

$$cx \in |c|tI_u.$$

So

$$cx \in X_u.$$

Finally, since

$$-u \leq 0 \leq u$$

we have $0 \in I_u$, so $0 \in X_u$.

Thus X_u is a subspace of X.

Consider $K_u = K \cap X_u$.

We claim that K_u is a cone in X_u:

If $x, y \in K_u$, then $(x + y) \in K$ since $x, y \in K$ and $K + K \subseteq K$. Also $(x + y) \in X_u$ since $x, y \in X_u$ and X_u is a subspace of X.

Hence $(x + y) \in K \cap X_u = K_u$. This gives that $K_u + K_u \subseteq K_u$ since $x, y \in K_u$ were arbitrary.

If $x \in K_u$ and $t > 0$, then $tx \in K$ since $x \in K$ and $tK \subseteq K$. Also $tx \in X_u$ since $x \in X_u$ and X_u is a subspace.

Thus $tx \in K_u$ This gives that $tK_u \subseteq K_u$ for all $t \geq 0$ since $x \in K_u$ and $t > 0$ were arbitrary.

Also $K_u \cap (-K_u) = (K \cap (-K)) \cap X_u = \{0\}$, so K_u is a cone in X_u.
Next we claim that I_u is an absorbing, absolutely convex subset of X_u:
Clearly, I_u is subset of X_u by definition of X_u. Let now $x \in X_u$.

Then $x \in tI_u$ for some $t > 0$,

$$\text{so } \{t > 0 \mid x \in tI_u\} \neq \emptyset.$$

Hence I_u is absorbing in X_u. Thus u is an order unit in (X_u, K_u). In order to show that I_u is absolutely convex, we have to show that I_u is convex and balanced.

a) I_u is convex:

We have by definition

$$I_u = \{x \in X | -u \leq x \leq u\}$$

Let $x, y \in I_u$ and $\lambda \in (0,1)$. Then $u-(\lambda x+(1-\lambda)y) = \lambda(u-x)+(1-\lambda)(u-y)$.

Now, $u - x$ and $u - y$ are in K since $x, y \in I_u$, hence

$$(\lambda(u - x) + (1 - \lambda)(u - y)) \in K$$

because K is cone. Thus

$$(u - (\lambda x + (1 - \lambda)y)) \in K,$$

so

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

Similarly , since

$$u + (\lambda x + (1 - \lambda)y) = \lambda(u + x) + (1 - \lambda)(u + y)$$

and

$$(u + x), (u + y) \in K,$$

we have that

$$(u + (\lambda x + (1 - \lambda)y)) \in K.$$

Hence

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

Thus we get

$$-u \leq \lambda x + (1-\lambda)y \leq u,$$

which means that

$$\lambda x + (1-\lambda)y \in I_u.$$

Since $\lambda \in (0,1)$ was arbitrary, it follows that I_u is convex.

b) I_u is balanced:

Let $x \in I_u$ and let $\lambda \in \mathbb{R}, |\lambda| \leq 1$.

If $\lambda \leq 0$, then $-1 \leq \lambda \leq 0$.

Furthermore $\lambda x + u = u - |\lambda| x = |\lambda|(u-x) + (1-|\lambda|)u \in K$,

since $(u-x)$ and u are in K, $|\lambda| \geq 0$, $1 - |\lambda| \geq 0$ and K is a cone.

Hence $-u \leq \lambda x$.

We have also

$$u - \lambda x = u + |\lambda| x = |\lambda|(u+x) + (1-|\lambda|)u.$$

Again, since $(u+x)$ and u are in K, we get that

$$u - \lambda x = |\lambda|(u+x) + (1-|\lambda|)u \in K, \text{ so } \lambda x \leq u.$$

Thus $-u \leq \lambda x \leq u$, so $\lambda x \in I_u$. This shows that I_u is balanced. Since I_u is convex and balanced, it is absolutely convex.

We know from proposition 14.8 on page 525 in[MW] that the Minkowski functional $|\cdot|_u$ is a seminorm.

Finally we show that $|u|_u = 1$:
Clearly $u \in I_u$ since $u \leq u$ and $u \geq -u$ because $2u \in K$.
Hence

$$\inf\{t > 0 |\ u \in tI_u\} \leq 1.$$

Assume now that $u \in \lambda I_u$ for some $0 < \lambda < 1$

Then

$$-u \leq \frac{1}{\lambda} u \leq u.$$

$$\text{In particular } \frac{1}{\lambda} \leq u, \text{ so } u - \frac{1}{\lambda}u \in K.$$

$$\text{Thus } (1 - \frac{1}{\lambda})u \in K.$$

Since we have that $\frac{1}{\lambda} > 1$, $u \in K$ and $tK \subset K$ for all $t > 0$, we get that $(\frac{1}{\lambda} - 1)u \in K$.

Hence $(\frac{1}{\lambda} - 1)u \in K \cap (-K) = \{0\}$.

But $\frac{1}{\lambda} - 1 > 0$ and $u \neq 0$, so $(\frac{1}{\lambda} - 1)u \neq 0$.

That is a contradiction,

Hence $\inf\{t > 0 | u \in tI_u\} = 1$, so $|u|_u = 1$. This completes the proof of observation 2.

Proof of proposition 1.2

1)

Assume that $x \in X_u$.

Then there exists $t > 0$ s.t. $x \in tI_u$.

As we have seen, this means that

$$-u \leq \frac{1}{t}x \leq u,$$

that is

$$(u - \frac{1}{t}x) \in K$$

$$\text{and } (u + \frac{1}{t}x) \in K.$$

$$\text{Hence } \frac{1}{2}(u - \frac{1}{t}x) \in K$$

$$\text{and } \frac{1}{2}(u + \frac{1}{t}x) \in K.$$

Thus

$$u - \frac{1}{2}(u - \frac{1}{t}x) = \frac{1}{2}(u + \frac{1}{t}x) \in K.$$

Hence

$$u \geq \frac{1}{2}(u - \frac{1}{t}x).$$

Since u and $\frac{1}{2}(u - \frac{1}{t}x)$ are in K $\frac{1}{2}(u - \frac{1}{t}x) \leq \mu$ and K is normal by assumption, we get that

$$M\|u\| \geq \tfrac{1}{2}\|u - \tfrac{1}{t}x\|.$$

Hence

$$2M \parallel u \parallel \geq \parallel u - \tfrac{1}{t}x \parallel \geq \tfrac{1}{t} \parallel x \parallel - \parallel u \parallel .$$

So

$$(2M + 1) \parallel u \parallel \geq \tfrac{1}{t} \parallel x \parallel,$$

$$\text{which means that } (2M + 1) \parallel u \parallel t \geq \parallel x \parallel .$$

Since this is true for all $t > 0$ with $x \in tI_u$, taking inf over such $t's$

we get that $(2M + 1) \parallel u \parallel |x|_u \geq \parallel x \parallel$. So $||x|| \leq \alpha |x|_u$ with $\alpha = (2M + 1)||u||$.

We have already proved that $| \cdot |_u$ is a seminorm.

Assume now that $|x|_u = 0$

Then since $(2M + 1) \parallel u \parallel |x|_u \geq \parallel x \parallel$, it follows that $\parallel x \parallel = 0$.

Hence $x = 0$ since $\| \cdot \|$ is a norm.

This shows that $| \cdot |_u$ is a norm.

2)

Let $\{x_n\}_{n \in N}$ be Cauchy in $(X_u, | \cdot |_u)$. We observe that since $\{x_n\}_n$ is Cauchy w.r.t. $| \cdot |_u$, then $\{x_n\}_n$ is Cauchy w.r.t. $| \cdot |_u$ as $|| \cdot || \leq \alpha | \cdot |_u$. Since $(X, || \cdot ||)$ is Banach space, there is an $x \in X$ s.t. $x_n \to x$ w.r.t. $|| \cdot ||$.

Given $\varepsilon > 0$, choose N s.t. $|x_n - x_m|_u < \varepsilon \ \forall n, m \geq N$.

Fix $n \geq N$ and consider $y_m^{(n)} = |x_n - x_m|_u, m \geq n$.

Then $\{y_m^{(n)}\}_m$ is bounded , hence there exists a convergent subsequence

$$\{y_{m_k}^{(n)}\}_k \subseteq \{y_m^{(n)}\}_m.$$

Assume first that $\lim_{k\to\infty} y_{m_k}^{(n)} > 0$.

As $y_{m_k}^{(n)} = |x_n - x_{m_k}|_u$, per definition of $|\cdot|_u$ we can find some $t_k^{(n)} > 0$ s.t.

$$0 < t_k^{(n)} - y_{m_k}^{(n)} < \tfrac{1}{k} \text{ and}$$

$$-u \leq \tfrac{1}{t_k^{(n)}}(x_n - x_{m_k}) \leq u \text{ for all } k.$$

That is

$$u - \tfrac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K \text{ and}$$

$$u + \tfrac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K \text{ for all } k.$$

As K is closed w.r.t $\|\cdot\|$, and

$$u - \tfrac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K$$

$$\text{and } u + \tfrac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K \text{ for all } k,$$

we get that

$$\lim_{k\to\infty}(u - \frac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K$$

$$\text{and } \lim_{k\to\infty}(u - \frac{1}{t_k^{(n)}}(x_n - x_{m_k}) \in K,$$

since limits exist because

$$t^{(n)} = \lim_{k\to\infty} t_k^{(n)} = \lim_{k\to\infty} y_{m_k}^{(n)} > 0.$$

(here we use that

$$|t_k^{(n)} - y_{m_k}^{(n)}| \ < \ \frac{1}{k}$$

for all k so

$$\lim_{k\to\infty} t_k^n = \lim_{k\to\infty} y_{m_k}^{(n)}. \text{)}$$

Then we get that

$$(u - \tfrac{1}{t^{(n)}}(x_n - x)) \in K \text{ and } (u + \tfrac{1}{t^{(n)}}(x_n - x)) \in K,$$

$$\text{so } \; -u \leq \tfrac{1}{t^{(n)}}(x_n - x) \leq u.$$

Hence

$$(x_n - x) \in t^{(n)} I_u.$$

Now, as

$$y^{(n)}_{m_k} = |x_n - x_{m_k}|_u < \varepsilon \text{ for all } k \;,$$

it follows that

$$t^{(n)} = \lim_{k \to \infty} y^{(n)}_{m_k} \leq \varepsilon.$$

$$\text{As } (x_n - x) \in t^{(n)} I_u \text{ we must have } |x_n - x|_u \leq t^{(n)} \leq \varepsilon.$$

This is true for all $n \geq N$.
(Furthermore since $(x_n - x) \in t^{(n)} I_u$ for all $n \geq N$, it follows that

$$(x_n - x) \in X_u$$

for all $n \geq \mathbb{N}$.
Hence

$$(x - x_n) = -(x_n - x) \in X_u$$

for all $n \geq N$, because X_u is a subspace. Then $x = (x - x_n) + x_n$ is in X_u , since $x - x_n$ and x_u are in X_u for all $n \geq N$ and X_u is a subspace).

Assume now that

$$\lim_{k \to \infty} y^{(n)}_{m_k} = 0$$

This means that

$$\lim_{k \to \infty} |x_n - x_{m_k}|_u = 0$$

$$\text{As } \| \cdot \| \leq \alpha | \cdot |_u \text{ , we get that } \lim_{k \to \infty} \| x_n - x_{m_k} \| = 0$$

But

$$\lim_{k\to\infty} \| x_n - x_{m_k} \| = \| x_n - x \| \text{ as } \lim_{k\to\infty} x_{m_k} = x \text{ w.r.t } \| \cdot \| .$$

Hence

$$\| x_n - x \| = 0, \text{ so } x = x_n \text{ which is in } X_u$$

Since $|\cdot|_u$ is a norm, we get that

$$|x_n - x| = 0 < \varepsilon \text{ (this is again true for all } n \geq \mathbb{N})$$

So we have proved that given $\varepsilon > 0$, if we choose N

$$\text{s.t. } |x_n - x_m|_u < \varepsilon \ \forall n, m \geq N,$$

$$\text{then } |x_n - x|_u < \varepsilon \ \forall n \geq N.$$

$$\text{Thus } \lim_{n\to\infty} |x_n - x|_u = 0.$$

This shows that $(X_u, |\cdot|_u)$ is a Banach space since $\{x_n\}_{n\in\mathbb{N}}$ was an arbitrary Cauchy sequence in X_u and $x \in X_u$.

3)

If u is an order unit for (X, K), then I_u is absorbing per definition.

That is, for all $x \in X$, we have $\{t > 0 | x \in tI_u\} \neq \emptyset$

Thus, for all $x \in X$, there exists $t_x > 0$ s.t. $x \in t_x I_u$.
Hence

$$x \in \bigcup_{t>0} tI_u \ \forall x \in X, \text{ so } X = X_u.$$

4)

If $u \in$Int(K) then there exist $\delta > 0$ s.t. $B(u, \delta) \subset \text{Int}(K)$.
(Here $B(u, \delta) := \{x \in X | \ ||x - u|| < \delta\}$)

Hence, given $x \in X$ with $x \neq u$, we have $u + \lambda x \subset B(u, \delta) \subset \text{Int}(K)$

$$\text{for all } \lambda \in \left(-\tfrac{\delta}{\|x\|}, \tfrac{\delta}{\|x\|}\right).$$

This means that u is an internal point of K From observation 1 it follows then that u is an order unit. From part 1) in proposition 1.2 it follows then that there is an $\alpha > 0$ s.t. $||\cdot|| \leq \alpha|\cdot|_u$. In order to prove that there is a $\beta > 0$ s.t $\|\cdot|_u \leq \beta||\cdot||$, we first observe that part 3) of proposition 1.2 gives that $X = X_u$ since u is an order unit. Hence $|\cdot|_u$ is defined on the whole X, so we may consider $(X, |\cdot|_u)$. and the map $\phi : (X, |\cdot|_u) \to (X, ||\cdot||)$ given by $\phi(x) = x$ for all $x \in X$. As we already know, there is an $\alpha > 0$ s.t. $||\cdot|| \leq \alpha|\cdot|_u$, so this gives that ϕ is bounded. Also, ϕ is bijective by definition. Moreover, since $(X, ||\cdot||)$ is a Banach space, we have from part 2) of proposition 1.2 that $(X_u, |\cdot|_u)$ is a Banach space. Thus $(X, |\cdot|_u)$ is a Banach space as $X = X_u$. Hence, we can apply the open mapping theorem and deduce that ϕ^{-1} is bounded. But this means that there exists a $\beta > 0$ s.t. $|\cdot|_u \leq \beta||\cdot||$, so $|\cdot|_u$ and $||\cdot||$ are equivalent.

Now, even if $(X, ||\cdot||)$ is not a Banach space, we still have the equivalence of the norms as long we assume that u is the interior point of K in the topological sense:
Since $u \in \text{Int}(K)$, there exists a $\delta > 0$ s.t. $u \pm \frac{\delta}{||x||}x \in K$ for all $x \in X \setminus \{0\}$.

Hence

$$-u \leq \tfrac{\delta}{||x||}x \leq u, \forall x \in X \setminus \{0\}.$$

This means that

$$\tfrac{\delta}{||x||}x \in I_u, \forall x \in X \setminus \{0\},$$

so

$$x \in \tfrac{||x||}{\delta} I_u, \forall x \in X \setminus \{0\}.$$

Hence

$$|x|_u = \inf\{t > u | x \in tI_u\} \leq \tfrac{||x||}{\delta} \quad \forall x \in X \setminus \{0\}.$$

Thus we have that $|\cdot|_u \leq \frac{1}{\delta} \parallel \cdot \parallel$.

From 1) we know that $\parallel \cdot \parallel \leq \alpha|\cdot|_u$.

Hence $|\cdot|_u$ and $\|\cdot\|$ are equivalent also in the case when $(X, ||\cdot||)$ is not a Banach space.

Now we wish to show that $|\cdot|_u = ||\cdot||_T$ where $||\cdot||_T$ is the Thomson's norm on X defined in the begining of this section, that is

$$||x||_T = \max\{\text{Inf}\{t \in \mathbb{R} \mid x \leq tu\}, \; -\sup\{t \in \mathbb{R} \mid x \geq tu\}\}.$$

We have

$$-\sup\{t \in \mathbb{R} | x \geq tu\}$$

$$= \inf\{-t \in \mathbb{R} | x \geq tu\}$$

$$= \inf\{s \in \mathbb{R} | x \geq -su\}.$$

(change of variables $s = -t$).

If we consider $t > 0$, we see that

$$x \leq tu \Leftrightarrow tu - x \in K \Leftrightarrow u - \tfrac{1}{t}x \in K.$$

Since $u \in \mathrm{Int}(K)$, there exists a $\delta > 0$ s.t. $B(u, \delta) \subseteq K$.

Hence

$$(\tfrac{\|x\|}{\delta}, \infty) \subseteq \{t > 0 | x \leq tu\}$$

since

$$(\mu - \frac{1}{t}x) \in B(\mu, \delta) \subseteq K$$

for all $t > \frac{||x||}{\delta}$.

Assume now that

$$0 < a < \tfrac{\|x\|}{\delta} \text{ and } (u - \tfrac{1}{a}x) \in K$$

Since K is convex and u , $(u - \frac{1}{a}x) \in K$ we must have that $v \in K$ for all v that lies in the segment between $u - \frac{1}{a}x$ and u.

This means that $(u - \frac{1}{t}x) \in K$ for any $t \geq a$.

Hence

$$[a, \infty) \subseteq \{t > 0 | x \leq tu\}.$$

Thus either

$$\{t > 0 | x \leq tu\} = [a, \infty) \text{ for some } a > 0 \text{ or}$$

$$\{t > 0 | x \leq tu\} = (0, \infty).$$

Similarly, if we consider

$$\{t > 0 | x \geq -tu\},$$

since we have

$$x \geq -tu \Leftrightarrow x + tu \in K \Leftrightarrow u + \frac{1}{t}x \in K$$

we deduce that either

$$\{t > 0 | x \geq -tu\} = [b, \infty) \text{ for some } b > 0 \text{ or}$$

$$\{t > 0 | x \geq -tu\} = (0, \infty).$$

Hence we have 4 possible situations:

i) $\{t > 0 | x \leq tu\} = [a, \infty)$ and $\{t > 0 | x \geq -tu\} = [b, \infty)$ for some a $b > 0$.

Then

$$\max\{\inf\{[a, \infty)\}, \inf\{[b, \infty)\}\}$$

$$= \max\{a, b\} = \inf\{[a, \infty) \cap [b, \infty)\}$$

$$= \inf\{\{t > 0 | x \leq tu\} \cap \{t > 0 | x \geq -tu\}\}$$

$$= \inf\{t > 0 | -tu \leq x \leq tu\}$$

$$= \inf\{t > 0 | x \in tI_u\} = |x|_u.$$

ii) $\{t > 0 | x \leq tu\} = (0, \infty)$ and $\{t > 0 | x \geq -tu\} = [b, \infty)$ for some $b > 0$.

Then

$$\max\{\inf\{(0, \infty)\}, \inf\{(b, \infty)\}\}$$

$$= b = \inf\{[b, \infty)\} = inf\{(0, \infty) \cap [b, \infty)\}$$

$$= \inf\{\{t > 0 | x \leq tu\} \cap \{t > 0 | x \geq -tu\}\}$$

$$= \inf\{t > 0 | -tu \leq x \leq tu\}$$

$$= |x|_u.$$

iii) $\{t > 0 | x \geq -tu\} = [a, \infty)$ for some $a > 0$ and $\{t > 0 | x \geq -tu\} = (0, \infty)$. This case can be treated similarly as ii) .

iv) $\{t > 0 | x \leq tu\} = \{t > 0 | x \geq -tu\} = (0, \infty)$.
Then

$$= max\{\inf\{t > 0 | x \leq tu\}, \inf\{t > 0 | x \geq -tu\}\}$$
$$= 0 = inf\{(0, \infty)\} = \inf\{\{t > 0 | x \leq tu\} \cap \{t > 0 | x \geq -tu\}\}$$
$$= \inf\{t > 0 | -tu \leq x \leq tu\} = |x|_u.$$

So, in all cases, we have

$$\max\{\inf\{t > 0 | x \leq tu\}, \inf\{t > 0 | x \geq -tu\}\} = |x|_u.$$

Now we have to show that

$$\max\{\inf\{t \in \mathbb{R} | x \leq tu\}, \inf\{t \in \mathbb{R} | x \geq -tu\}\}$$

$$= \max\{\inf\{t > 0 | x \leq tu\}, \inf\{t > 0 | x \geq -tu\}\}$$

The idea is to show that for all x, we have that

$$\max\{\inf\{t \in \mathbb{R} | x \leq tu\}, \inf\{t \in \mathbb{R} | x > -tu\}\}$$

is greater or equal to zero. Then it would follow that we can replace $t \in \mathbb{R}$ by $t \geq 0$.

Now, if we assume that we can replace $t \in \mathbb{R}$ by $t \geq 0$, we observe further that, if $0 \in \{t \geq 0 \mid x \leq tu\}$, then $(0 - x) \in K$. so $-x \in K$.

Hence $tu - x \in K$ for all $t > 0$ since K is a cone.

Then

$$0 = \inf\{t \geq 0 | x \leq tu\} = \inf\{t > 0 | t \in \mathbb{R}\} = \inf\{t > 0 | x \leq tu\}$$

since $\{t > 0 \mid x \leq tu\}$ contains all positive t's.

A similar argument gives that

$$\inf\{t \geq 0 | x \geq -tu\} = \inf\{t > 0 | x \geq -tu\}.$$

Hence if we can replace $t \in \mathbb{R}$ by $t \geq 0$, it follows that we can replace further $t \geq 0$ by $t > 0$.

Let us now show that we can replace $t \in \mathbb{R}$ by $t \geq 0$ that is

$$\max\{\inf\{t \in \mathbb{R} | x \leq tu\}, \inf\{t \in \mathbb{R} | x \geq -tu\}\} \geq 0, \ \forall x \in X_u :$$

Assume first that there is an $s < 0$ s.t. $x \geq -su$ and that there exist $t < 0$ s.t. $x \leq tu$.

Then

$$-su \leq x \leq tu$$

Since $s, t < 0$, we get that $|s|u \leq x \leq -|t|u$.

This means that $-|t|u - x \in K$

and $x - |s|u \in K$.

Since $K + K \subseteq K$, we get that:

$$(-|t| - |s|)u = (-|t|u - x) + (x - |s|u) \in K.$$

Since $u \in K$ and $|t| + |s| > 0$ because $s, t < 0$ we have

$$(|t| + |s|)u \in K$$

Thus $(|t| + |s|)u \in K \cap (-K)$.

That is a contradiction since $(|t| + |s|)u \neq 0$, as $|t| + |s| > 0$ and $u \neq 0$.

Hence, if there exists an $s > 0$ s.t. $s \in \{t \in \mathbb{R} | x \geq -tu\}$,

then the set $\{t \in \mathbb{R} | x \leq tu\}$ contains only nonnegative elements and vice versa.

This shows that

$$\max\{\inf\{t \in \mathbb{R} | x \leq tu\}, \inf\{t \in \mathbb{R} | x \geq -tu\}\} \geq 0, \text{ for all } x \in X_u$$

and the proof of proposition 1.2 is completed.

Comment: In [GQ] it is just stated without proof that $|| \cdot ||$ and $|| \cdot ||_T$ are equivalent when K is closed and normal. For the proof they refer to the article of Nussbaum. Here we have given our own proof of this result by proving the proposition 1.2 .

Let now $(X, ||\cdot||)$ be a normed space, $u \in K - \{0\}$, $u \in Int(K)$ and K be a normal, closed cone in X.

As we mentioned, one can also define the Hilbert seminorm at $x \in X$

by $\| x \|_H = M(x/u) - m(x/u)$

where

$$M(x/u) = \inf\{t \in \mathbb{R} | x \leq tu\}$$

$$m(x/u) = \sup\{t \in \mathbb{R} | x \geq tu\}.$$

Let $L_u = \mathbb{R}u$ and the quotient space X/L_u be equiped with the quotient norm (ass. with the Thomson norm) i.e.

$$\| x + L_u \|_T = \inf\{\| x - y \|_T | y \in L_u\}$$

$$= \inf\{\| x + \lambda u \|_T | \lambda \in \mathbb{R}\}$$

We will show the following lemma:

Lemma 1.3 $\| x \|_H = 2 \| x + L_u \|_T$ *for all* $x \in X$

Proof:

Claim 1) $\| x + \lambda u \|_T = \max(M(x/u) + \lambda, -m(x/u) - \lambda)$:

By definition

$$\| x + \lambda u \|_T = \max(M((x + \lambda u)/u), -m((x + \lambda u)/u))$$

We have that:

$$x + \lambda u \leq tu \Leftrightarrow tu - (x + \lambda u) \in K$$

$$\Leftrightarrow (t - \lambda)u - x \in K$$

$$\Leftrightarrow x \leq (t - \lambda)u.$$

Hence $\tilde{t} \in \{t \in \mathbb{R} | (x + \lambda u) \leq tu\}$ if and only if

$$(\tilde{t} - \lambda) \in \{t \in \mathbb{R} | x \leq tu\}.$$

This gives that

$$\inf\{t \in \mathbb{R} | (x + \lambda u) \leq tu\}$$

$$\inf\{t \in \mathbb{R} | x \leq tu\} + \lambda.$$

Thus $M((x + \lambda u)/u) = M(x/u) + \lambda$.

Similarly it can be shown that

$$m((x + \lambda u) = m(x/u) + \lambda$$

Claim 2) Given $a, b \in \mathbb{R}$, the expression $\max\{a + \lambda, b - \lambda\}$ is minimal when $a + \lambda = b - \lambda$, that is when $\lambda = \frac{1}{2}(b - a)$:

Consider the functions f, g :$\mathbb{R} \to \mathbb{R}$ given by $f(x) = x + a,\ g(x) = -x + b$

Let $\tilde{x}$ be s.t. $f(\tilde{x}) = g(\tilde{x})$, more precisely $\tilde{x} = \frac{1}{2}(b - a)$.

As $f^{'}(x) = 1$ and $g^{'}(x) = -1 \forall x \in \mathbb{R}$,

$$\text{we get that } f(x) < f(\tilde{x}) = g(\tilde{x}) < g(x)\ \forall x < \tilde{x} \text{ and}$$

$$f(x) > f(\tilde{x}) = g(\tilde{x}) > g(x)\ \forall x > \tilde{x}$$

(since f is strictly increasing and g is strictly decreasing)

Hence

$$\max\{f(x), g(x)\} = g(x) > g(\tilde{x}) = f(\tilde{x}) = \max\{f(\tilde{x}), g(\tilde{x})\} \forall x < \tilde{x} \text{ and}$$

$$\max\{f(x), g(x)\} = f(x) > f(\tilde{x} = g(\tilde{x}) = \max\{f(\tilde{x}), g(\tilde{x})\} \forall x > \tilde{x}$$

This proves the claim 2).

Claim 1 and claim 2 give that

$$\inf \parallel x + \lambda u \parallel_T = \inf_{\lambda \in \mathbb{R}} \max\{M(x/u) + \lambda, -m(x/u - \lambda)\}$$

$$= M(x/u) + \tfrac{1}{2}(-m(x/u) - M(x/u))$$

$$= \tfrac{1}{2}(M(x/u) - m(x/u)) = \tfrac{1}{2} \parallel x \parallel_H$$

which proves the lemma 1.3 .Here we have used that the minimum of the

$$\max\{M(x/\mu)+\lambda, -m(x/\mu)-\lambda\}$$

wil be attained at

$$\lambda = \frac{1}{2}(-m(x/\mu) - M(x/\mu)$$

by claim 2.

Now we will consider the examples that we mentioned in the introduction of this section.

In examples 1.4 and 1.5 we will denote the respective order unit by e, as done in [GQ]. In example 1.6, we will denote the respective order unit by u as we have done so far in this section.

Example 1.4 We consider the finite dimensional vector space $X = \mathbb{R}^n$ with its Euclidian norm, the standard positive cone $K = \mathbb{R}^n_+$ and the order unit vector $e = \overrightarrow{1} := (1, ..., 1)^T$. We claim that Thompson's norm with respect to $\overrightarrow{1}$ is nothing but the sup norm

$$||x||_T = \max_i |x_i| = ||x||_\infty,$$

whereas Hilbert's seminorm with respect to $\overrightarrow{1}$ is the so called *diameter*:

$$||x||_H = \max_{i \leq i,j \leq n} (x_i - x_j) = \Delta(x).$$

Proof: We have:

$$x \leq te \Leftrightarrow$$

$$(te - x) \in K \Leftrightarrow$$

$$t \geq x_i \quad \forall i \quad 1 \leq i \leq n.$$

Clearly then

$$\inf\{t \in \mathbb{R} | x \leq te\} = \max_i x_i.$$

Similarly we have:

$$te \leq x \Leftrightarrow$$

$$(x - te) \in K \Leftrightarrow$$

$$t \leq x_i \ \forall i \ 1 \leq i \leq n.$$

Hence

$$\sup\{t \in \mathbb{R} | te \leq x\} = \min_i x_i.$$

Thus

$$M(x/e) = \max_i x_i$$

and

$$m(x/e) = \min_i x_i.$$

Then we get that

$$\| x \|_T = \max(M(x/e), -m(x/e))$$

$$= \max(\max_i x_i \ , \ -\min_i x_i).$$

Choose j and k s.t.

$$x_j = \max_i x_i$$

and

$$x_k = \min_i x_i.$$

Then

$$\| x \|_T = \begin{cases} x_j & \text{if } x_j \geq -x_k; \\ -x_k & \text{if } -x_k \geq x_j. \end{cases}$$

1) If $x_j \geq -x_k$, then $\| x \|_T = x_j$.

Since

$$x_j = \max_i x_i,$$

we also have $x_j \geq x_k$. Hence $x_j \geq |x_k|$.

Since

$$x_j = \max_i x_i \text{ and } x_k = \min_i x_i$$

we must have

$$\max_i |x_i| = \max\ \{|x_j|, |x_k|\}.$$

But since $x_j \geq |x_k|$, we get then that $\max\ \{|x_j|, |x_k|\} = x_j$.

Hence

$$\max_i |x_i| = x_j = \parallel x \parallel_T .$$

2) If $-x_k \geq x_j$, then $\parallel x \parallel_T = -x_k$.

Since

$$x_k = \min_i x_i,$$

we have $x_k \leq x_j$.

Hence $-x_k \geq -x_j$. Combining these 2 inequalities together, $(-x_k \geq x_j)$ and $-x_k \geq -x_j$, we get that $-x_k \geq |x_j|$.

Thus

$$|x_k| = -x_k \geq |x_j|,$$

so

$$\max_i |x_i| = \max\{|x_j|, |x_k|\} = |x_k| = -x_k = \parallel x \parallel_T .$$

So in any case

$$\max_i |x_i| = \parallel x \parallel_T .$$

Furthermore

$$\parallel x \parallel_H = M(x/e) - m(x/e) = \max_i x_i - \min_i x_j = x_j - x_k = \max_{1 \leq i,l \leq n} (x_i - x_l)$$

Example 1.5 Let $X = S_n$ be the space of all Hermitian matrices of dimension n equipped with the operator norm and $K = S_n^+$ be the cone of positive semi-definite matrices. Let the identity matrix I_n be the order unit: $e = I_n$. Then we claim that ThompsonÂt's norm with respect to I_n is nothing but the spectral radius of X, i.e..

$$||X||_T = \max_{1 \le i \le n} |\lambda_i(X)| = ||\lambda(X)||_\infty,$$

where $\lambda(X) := (\lambda_1(X), ..., \lambda_n(X))$, is the vector of ordered eigenvalues of X counted with multiplicities, whereas HilbertÂťs seminorm with respect to I_n is the diameter of the spectrum:

$$||X||_H = \max_{1 \le i,j \le n} (\lambda_i(X) - \lambda_j(X)) = \Delta(\lambda(X)).$$

Proof: First we want to show that $K = S_n^+$ indeed is a cone in X:

It is clear that $aK \subset K$ for all $a \ge 0$ and $K + K \subseteq K$. Assume that $A \in K \cap (-K)$. As $A \in K$, all eigenvalues of A are nonnegative and since $A \in (-K)$, then $-A \in K$, so all eigenvalues of A are less or equal to zero. We must then have that all eigenvalues of A are 0. But since A is Hermitian it is unitary diagonalisable.

Hence $A = UDU^*$ where U is a unitary matrix and D is a diagonal matrix having the eigenvalues of A on its diagonal.

Since 0 is the only eigenvalue of A, we get that D is the zero matrix, hence $A = 0$. So $K \cap (-K) = 0$.

To simplify notation, from now on we let $I = I_n$. If $X \le tI$, then

$$tI - X \in K,$$

so $tI - X$ has only non-negative eigenvalues . Again, since X is Hermitian, it is unitary diagonalisable, so there exists an orthonormal basis of eigenvectors for X.

Let $\lambda = (\lambda_1, ..., \lambda_n)$ be the eigenvalues counting multiplicities and $\{v_1, ..., v_n\}$ corresponding eigenvectors. Then they are eigenvectors for $tI - X$ with corresponding eigenvalues $\{t - \lambda_1, ..., t - \lambda_n\}$.

So, $\{v_1, ..., v_n\}$ is an orthonormal basis of eigenvectors for $(tI - X)$ also.

Hence all the eigenvalues for $tI - X$ are of the form $t - \lambda_i$, $i \in \{1, ..., n\}$.

This gives that $(tI - X) \in K$ if and only $t - \lambda_i \ge 0$, for all i with $i \in \{1, ..., n\}$.

We clearly then have

$$\inf\{t \in \mathbb{R} | (tI - X) \in K\} = \max_i \lambda_i.$$

That is

$$M(X/I) = \max_i \ \lambda_i.$$

A similar argument gives that

$$(X - tI) \in K \Leftrightarrow \lambda_i - t \geq 0 \ \forall i.$$

since all the eigenvalues of $X - tI$ are of the form $\lambda_i - t$ where $i \in \{1, ..., n\}$.

Then

$$\sup\{t \in \mathbb{R} | (X - tI) \in K\} = \min \ \lambda_i.$$

Thus

$$m(X/I) = \min \ \lambda_i.$$

Hence

$$\| X \|_T = \max \ \{\max \ \lambda_i, -\min \ \lambda_i\} = \max \ |\lambda_i|$$

by the same argument as in example 1.4.

Furthermore

$$\| X \|_H = M(X/I) - m(X/I) = \max \ \lambda_i - \min \ \lambda_i$$

$$= \max_{1 \leq i,j, \leq n} (\lambda_i - \lambda_j),$$

which proves the statement.

We should also show that I indeed is in Int(K), so that I actually is an order unit:

Let

$$A \in S_n \text{ s.t. } \parallel I - A \parallel < 1.$$

If λ is an eigenvalue for A, then $1 - \lambda$ is an eigenvalue for $I - A$ so

$$|1 - \lambda| \leq \parallel I - A \parallel < 1.$$

Hence

$$|1 - \lambda| < 1.$$

But then we must clearly have

$$0 < \lambda < 2.$$

In particular $\lambda > 0$.

Since λ was a general eigenvalue for A, we get that A is positive definite, so $A \in K$.

Hence

$$I \in \text{Int}(K).$$

At the end, we show that $K = S_n^+$ is closed and normal.

K is closed:

Suppose for a contradiction that K is not closed in X. Then there exists a sequence $\{A_n\} \subseteq K$ s.t. A_n converges to some A in $X \setminus K$ as $n \to \infty$. Since A is in $X \setminus K$, A is then Hermitian, but not a positive semidefinite, hence there exists an $x \in \mathbb{C}^n$ s.t. $< Ax, x > < 0$. Furthermore, since A_n is in K for all n, we have $< A_n x, x > \geq 0$ for all n.

Hence

$$|\langle A_n x, x \rangle - \langle Ax, x \rangle|$$

$$= |\langle A_n x, x \rangle + \langle Ax, x \rangle|$$

$$= \langle A_n x, x \rangle + |\langle Ax, x \rangle| \geq |\langle Ax, x \rangle| > 0$$

for all $n \in \mathbb{N}$.

On the other hand,

$$|\langle A_n x, x\rangle - \langle Ax, x\rangle|$$

$$= |\langle (A_n - A)x, x\rangle| \leq \| (A_n - A)x \| \| x \|$$

$$\leq \| (A_n - A) \| \| x \|^2, \text{ and } ||(A_n - A)||^2 ||x||^2 \text{ goes to } 0 \text{ as } n \to \infty$$

$$\text{since } \| A_n - A \| \to 0 \text{ as } n \to \infty$$

by assumption.

Hence we get a contradiction, so we deduce that K is closed.

K is normal:

Assume that $A, B \in K$ and $A \leq B$, that is $(B - A) \in K$.

Then

$$0 \leq < (B - A)x, x > = < Bx, x > - < Ax, x > \text{ for all } x \in \mathbb{C}^n.$$

Since A and B are positive semidefinite , we have

$$\| A \| = \sup_{\|x\| \leq 1} < Ax, x >$$

$$\text{and } \| B \| = \sup_{\|x\| \leq 1} < Bx, x >$$

Combining this and the fact that $< Bx, x > \geq < Ax, x >$ for all $x \in \mathbb{C}^n$ as we proved, we get that $\| B \| \geq \| A \|$. Since $A, B \in K$ were arbitrary with $A \leq B$, we deduce that K is normal.

The example 1.4 and 1.5 are also given in [GQ], denoted by examples 2.1 and 2.2, but the proofs are omitted in [GQ]. The next example, example 1.6 is not given in [GQ].

Example 1.6 We now let Ω be a compact Hausdorff space and we let $X = C_{\mathbb{R}}(\Omega)$, that is X is the space of all continuous real valued functions on Ω.

X is then a Banach space with the norm $\| \cdot \|_\infty$, where

$$\| f \|_\infty = \sup_{w\in\Omega} |f(w)|.$$

We let

$$K = \{f \in X | f(w) \geq 0 \ \forall w \in \Omega\}$$

and we let the order unit $u = 1$, that is u is the constant function 1 in this case.on Ω. It is obvious that K closed and normal cone w.r.t $||\cdot||_\infty$ and that $1 \in \mathrm{Int}(K)$.

Now we claim that

$$\| f \|_T = \| f \|_\infty \ \ \forall f \in X \ :$$

We have

$$\| f \|_T = \inf\{t > 0 | \ f \in tI_u\}$$

since $\| \cdot \|_T$ is equal to the Minkowski functional w.r.t I_u as we have proved.

So

$$\| f \|_T = \inf\{t > 0 | \ f \in tI_u\}$$

$$= \inf\{t > 0 | \ -t \leq f(w) \leq t \ \forall w \in \Omega\}$$

$$= \inf\{t > 0 | \ |f(w)| \leq t \ \forall w \in \Omega\}.$$

If t is s.t. $t \geq |f(w)| \ \forall w \in \Omega$,

then clearly

$$t \geq \sup_{w\in\Omega} |f(w)| = \| f \|_\infty .$$

Hence

$$\inf\{t > 0 | \ t \geq |f(w)| \ \forall w \in \Omega\} \geq \| f \|_\infty .$$

On the other hand

$$\| f \|_\infty = \sup_{w\in\Omega} |f(w)| \geq |f(\tilde{w})| \ \forall \tilde{w} \in \Omega.$$

Hence

$$\| f \|_\infty \in \{t > 0 | \ t \geq |f(w)| \ \forall w \in \Omega\},$$

$$\text{so } \| f \|_\infty \geq \inf\{t > 0 | \ t \geq |f(w)| \ \forall w \in \Omega\}.$$

Thus

$$\| f \|_\infty = \inf\{t > 0 | \ t \geq |f(w)| \ \forall w \in \Omega\} = \| f \|_T .$$

which proves the claim.

Next we claim that

$$M(f/u) = \max_{w\in\Omega} f(w) \text{ and } m(f/w) = \min_{w\in\Omega} f(w) :$$

(Comment: Since Ω is compact and f is continuous, real valued, we know that f will attain a maximal and minimal value at some points w_1 and w_2 in Ω, so it makes sense to write

$$\max_{w\in\Omega} f(w) \text{ and } \min_{w\in\Omega} f(w).)$$

By definition,

$M(f/u) = \inf\{t \in \mathbb{R} | \ f(w) \leq t \ \ \forall w \in \Omega\}$, since the order unit u is the constant function 1 in this case.

Clearly, if $t \in \mathbb{R}$ is s.t. $t \geq f(w) \ \ \forall w \in \Omega$, then

$$t \geq \max_{w\in\Omega} f(w)$$

Taking inf over such t's, we get that

$$M(f/u) \geq \max_{w\in\Omega} f(w).$$

On the other hand,

$$\max_{w\in\Omega} f(w) \geq f(\tilde{w}) \ \ \forall \tilde{w} \in \Omega ,$$

so

$$\max_{w\in\Omega} f(w) \in \{t \in \mathbb{R} \mid f(w) \leq t \ \ \forall \ w \in \Omega\}.$$

Hence

$$\max_{w\in\Omega} f(w) \geq \inf\{t \in \mathbb{R} \mid f(w) \leq t \ \ \forall\, w \in \Omega\} = M(f/u).$$

We conclude that

$$M(f/u) = \max_{w\in\Omega} f(w).$$

In the similar way, we can show that

$$\min_{w\in\Omega} f(w) = m(f/u).$$

Hence we get that

$$\| \, f \, \|_H = (\max_{w\in\Omega} f(w)) - (\min_{w\in\Omega} f(w))$$

$$= \max_{w,w^{'}\in\Omega} (f(w) - f(w^{'})) = \triangle(f).$$

3 Thompson's dual norm, the dual cone and the abstract simlex

From now on in this thesis, unless else is specified, we let $(X, ||\cdot||)$ be a real Banach space, $K \subseteq X$ be a normal, closed cone with $\text{Int}(K) \neq \emptyset$, $e \in \text{Int}K$ be an order unit and $||\cdot||_T$ be a Thompson' norm w.r.t. e.
In this section we consider the dual of $(X, ||\cdot||_T)$ and define a norm $||\cdot||_T^*$ on this space. We also define the dual cone K^* and the abstract simplex P(e) in $(X, ||\cdot||_T)^*$. (Those definitions can also be found in [GQ], section 2.) After defining these new concepts, we introduce remarks 2.1, 2.2 and 2.3, where we give the concrete expressions and descriptions for $||\cdot||_T^*, K^*$ and $P(e)$ in the cases when X is equal to $\mathbb{R}^n$, S_n and $C_{\mathbb{R}}(\Omega)$. The remarks 2.1 and 2.2 are also given in [GQ] section 2, denoted by remark 3.1 and 3.2 respectively. However, here we give a detailed proof of all the statements in those remarks. In addition we introduce the remark 2.3 not given in [GQ], that deals with the dual of $C_{\mathbb{R}}(\Omega)$ which turns out to be the space of all signed Radon measures on Ω equipped with the total variation norm.

We define

$$||\cdot||_T^* \; : \; (X, ||\cdot||_T)^* \to \mathbb{R}^+ \text{ by}$$

$$||z||_T^* = \sup_{||x||_T \leq 1} | < z, x > | \; \forall z \in (X, ||\cdot||_T)^*$$

(here $< z, x >$ means $z(x)$)

(Comment; Since $||\cdot||$ and $||\cdot||_T$ are equivalent because K is closed and normal, we have $(X, ||\cdot||)^* = (X, ||\cdot||_T)^*$. This space will be denoted by X^* from now on.)

Furthermore, we define the dual cone K^* in X^* by

$$K^* = \{x \in X^* \mid \; < z, x > \; \geq 0 \; \forall x \in K\}$$

and the abstract simplex by

$$P(e) = \{\mu \in K^* \mid \; < \mu, e > = 1\}$$

Remark 2.1*For the standard positive cone (Example 1.4 $X = \mathbb{R}^n$, $K = \mathbb{R}^n_+$ and $e = \overrightarrow{1}$) the dual space X^* is $X = \mathbb{R}^n$ itself and the dual norm $||\cdot||_T^*$ is the l_1 norm:*

$$||x||_T^* = \sum_i |x_i| = ||x||_1.$$

Furthermore the dual cone K^* is $\mathbb{R}^n_+$ in this case.

The abstract simplex $P(\overrightarrow{1})$ *is the standard simplex in* $\mathbb{R}^n$:

$$P(\overrightarrow{1}) = \{v \in \mathbb{R}^n_+ \ : \ \sum_i v_i = 1\},$$

i.e., the set of probability measures on the discret space $\{1, ..., n\}$.

Proof: For each $\phi \in X^*$, if we let

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $y_k = \phi(e_k)$, we see that $\phi(x) = x \cdot y$ for all $x \in \mathbb{R}^n = X$. Indeed, $e_k \cdot y = y_k = \phi(e_k)$ for all $k \in \{1, ..., n\}$, $\{e_k\}_{1 \le k \le n}$ is a basis for $\mathbb{R}^n$ and ϕ is linear, so it is completely determined by its values on $e_k's$. Hence $\phi(x) = x \cdot y$ for all $x \in \mathbb{R}^n$, so we can identify X^* with $\mathbb{R}^n$.
If $||x||_T \le 1$, then $||x||_\infty \le 1$ since $|| \cdot ||_T = || \cdot ||_\infty$ by example 1.4, so we get that

$$|\phi(x)| = |x \cdot y| = |\sum_{k=1}^{n} x_k y_k| \le \sum_{k=1}^{n} |x_k||y_k| \le ||x||_\infty \sum_{k=1}^{n} |y_k| \le \sum_{k=1}^{n} |y_k|.$$

On the other hand, if $x \in \mathbb{R}^n$ is given by

$$x_k = \begin{cases} \frac{y_k}{|y_k|} & \text{when } y_k \neq 0 \\ 0 & \text{when } y_k = 0 \end{cases}$$

for all $k \in \{1, ..., n\}$, then $||x||_\infty = 1$ and

$$\phi(x) = x \cdot y = \sum_{k=1}^{n} |y_k|.$$

Hence

$$||\phi||_T^* = ||y||_T^* = \sum_{k=1}^{n} |y_k| = ||y||_1$$

Next we show that $K^* = \mathbb{R}^n_+$:

Assume that $\phi \in K^*$ and let $y \in \mathbb{R}^n$ be s.t. $\phi(x) = x\dot{y}$ for all $x \in \mathbb{R}^n_+$. (Such $y \in \mathbb{R}^n$ exists by what we have proved above.) Then, in particular $\phi(e_j) = y_j \geq 0$ for all $j \in \{1, ..., n\}$ since $e_j \in K$ for all j. Thus $y \in \mathbb{R}^n_+$. On the other hand, if $y \in \mathbb{R}^n_+$ and $x \in K$, then $x \cdot y = \sum_{j=1}^{n} x_j y_j \geq 0$ since $x_j y_j \geq 0$ for all $j \in \{1, ..., n\}$. Hence ϕ given by $\phi(x) = x \cdot y$ for all $x \in X$ is in K^*, and this shows that $K^* = \mathbb{R}^n_+$.

Furthermore, for $\phi \in X^*$ given by $\phi(x) = x \cdot y$ for all $x \in \mathbb{R}$, we have

$$\phi(\vec{1}) = 1 \Leftrightarrow y \cdot \vec{1} = 1 \Leftrightarrow \sum_{k=1}^{n} y_k = 1.$$

Combining this and the the fact that $K^* = \mathbb{R}^n_+$, we get

$$P(\vec{1}) = \{\phi \in K^* \mid \phi(\overrightarrow{1}) = 1\} = \{y \in \mathbb{R}^n_+ | \sum_{k=1}^{n} y_k = 1\}.$$

This completes the proof of remark 2.1.

Remark 2.2 *For the cone of semidefinite matrices (Example 1,5 $X = S_n$, $K = S_n^+$ and $e = I_n$), $X^* = S_n$ itself and the dual norm $||\cdot||^*_T$ it is the trace norm:*

$$||X||^*_T = \sum_{1 \leq i \leq n} |\lambda_i(X)| = ||X||_1, \ X \in S_n$$

The dual cone K^* is equal to S_n^*, the set of all positive semidefinite matrices. The simplex $P(I_n)$ is the set of positive semidefinite matrices with trace 1:

$$P(I_n) = \{\rho \in S_n^+ \ : \mathrm{trace}(\rho) = 1\}.$$

Proof: We want to show first that for each $\phi \in X^*$, there is a unique $B \in X$ s.t. $\phi(A) = Tr(AB)$ for all $A \in X$:

As ϕ is linear, it is completely determined by its values on a basis.

Let $V_{k,l}$ be a matrix in X having 1 as its $l, k - th$ component and its $k, l - th$ component and 0 otherwise.

F.ex. if n=2, then

$$V_{2,1} = \left\{ \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right\}$$

Let $\tilde{V_{k,l}}$ be a matrix in X having i as its k, l-th component, -i as its l, k-th components and 0 otherwise.

F.ex. if n=2, then

$$\tilde{V}_{2,1} = \left\{ \begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right\}$$

In this notation,we always assume that $k > l$. Furthermore, let $V_{i,i}$ be a matrix having 1 as its i, i-th component and 0 otherwise.

Again if n=2 then

$$V_{1,1} = \left\{ \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right\}$$

$$V_{2,2} = \left\{ \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right\}.$$

Clearly, we get that

$\beta = \{\{V_{i,i}\}_{1 \leq i \leq n}, \{V_{k,l}\}_{1 \leq l < k \leq n}, \{\tilde{V_{k,l}}\}_{1 \leq l < k \leq n}\}$ is a basis for X.

Given (k, l) with $1 \leq l < k \leq n$, it is easy to see that

$$Tr(V_{k,l}B) = b_{l,k} + b_{k,l}.$$

$$Tr(\tilde{V}_{k,l}\ B) = i(b_{l,k} - b_{k,l}) \text{ where } B = [b_{i,j}]$$

If $B \in X$, we know that $Re(b_{l,k}) = Re(b_{k,l})$ $Im((b_{l,k}) = -Im(b_{k,l})$

Hence, for any pair (k,l) with $1 \leq l < k \leq n$, we get a unique solution of $b_{l,k}$ and $b_{k,l}$ by considering the equations:

$$Tr((V_{k,l}B) = \phi(V_{k,l}) = b_{l,k} + b_{k,l} = 2Re(b_{l,k}),$$

$$Tr((\tilde{V}_{k,l}B) = \phi(\tilde{V}_{k,l}) = i(b_{l,k} - b_{k,l}) = 2Im(b_{l,k}).$$

Furthermore, since $Tr((V_{i,i}B) = b_{i,i}$ we also get a unique solution for $b_{i,i}$, namely $b_{i,i} = \phi(V_{i,i})$.

Hence there is a unique $B \in X$ s.t. $\phi(A) = Tr(AB)$ for all $A \in \beta$.

But since β is a basis and ϕ is linear, we have then $\phi(A) = Tr(AB)$ for all $A \in X$.

This shows that $\psi : X \to X^*$ given by $\psi(B) = \phi_B$ where $\phi_B(A) = Tr(AB)$ for all $A \in X$,

is an isomorphism, so we can identify X with X^*.

By definition,

$$P(I_n) = \{\phi \in K^* | \phi(I_n) = 1\}.$$

But, as we have shown, for each $\phi \in X^*$, there is a $B \in X$ s.t.

$$\phi(A) = Tr(AB) \text{ for all } A \in X.$$

Hence, if $\phi(I_n) = 1$, it follows that $Tr((I_nB) = Tr((B) = 1$.

So we can rewrite the expression for $P(I_n)$ as

$$P(I_n) = \{B \in S_n^+ | Tr(B) = 1\}$$

by identifying ϕ_B with B for all $\in X$. (It remains to show that S_n^+ can be identified with K^* and it will be done later.)

Let $B \in X$ and consider $\phi_B \in X^*$ given by $\phi_B(M) = Tr(MB)$ for all $M \in X$. We want to show that $||\phi_B||^* = ||B||_1$

Let $M \in X$ and assume that $||M||_T \leq 1$, that is $|\lambda_i(M)| \leq 1$ for all $i \in \{1, ..., n\}$. (Recall the example 1.5 where it was shown that $||M||_T = ||\lambda(M)||_\infty$).

Since M and B are Hermitian, hence unitary diagonalisable, there are unitaries U_1, U_2 and diagonal matrices D_1, D_2 s.t.

$$M = U_1D_1U_1^* \ , \ B = U_2D_2U_2^*.$$

Then

$$Tr(MB) = Tr(U_1D_1U_1^* \ U_2D_2U_2^*)$$

$$= Tr(D_1 U_1^* U_2 \ D_2 U_2^* U_1)$$

$$= Tr(D_1 (U_1^* U_2 D_2 (U_1^* U_2)^*))$$

$$= Tr(D_1 (V D_2 V^*)) = Tr(D_1 A)$$

where $V = U_1^* U_2$ and $A = V D_2 V^*$. We have

$$Tr(D_1 A) = \sum_{j=1}^{n} \lambda_j a_{j,j}$$

where $A = [a_{i,j}]$, and λ_j s are the eigenvalues of M so that

$$D_1 = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

Let $V = [v_{i,j}]$ and write D_2 as

$$D_2 = \begin{bmatrix} \eta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \eta_n \end{bmatrix}.$$

Since $A = V D_2 V^*$

$$= \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix} \begin{bmatrix} \eta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \eta_n \end{bmatrix} \begin{bmatrix} \bar{v}_{11} & \cdots & \bar{v}_{n1} \\ \vdots & \ddots & \vdots \\ \bar{v}_{1n} & \cdots & \bar{v}_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix} \begin{bmatrix} \eta_1 \bar{v}_{11} & \cdots & \eta_1 \bar{v}_{n1} \\ \vdots & \ddots & \vdots \\ \eta_n \bar{v}_{1n} & \cdots & \eta_n \bar{v}_{nn} \end{bmatrix},$$

we see that for each $j \in \{1, ..., n\}$, we have

$$a_{j,j} = \sum_{k=1}^{n} \eta_k |v_{j,k}|^2.$$

Now, since

$$Tr(MB) = Tr(D_1 A) = \sum_{j=1}^{n} \lambda_j a_{j,j}$$

and $|\lambda_j| \leq 1$ for all j by assumption ($||M||_T \ \leq 1$), we get that

$$|\phi(M)| = |Tr(MB)| = |\sum_{j=1}^{n} \lambda_j a_{j,j}|$$

$$\leq \sum_{j=1}^{n} |\lambda_j||a_{j,j}| \leq \sum_{j=1}^{n} |a_{j,j}|$$

$$= \sum_{j=1}^{n} |\sum_{k=1}^{n} \eta_k |v_{j,k}|^2| \leq \sum_{j=1}^{n}\sum_{k=1}^{n} |\eta_k||v_{j,k}|^2 = \sum_{k=1}^{n}\sum_{j=1}^{n} |\eta_k||v_{j,k}|^2 = \sum_{k=1}^{n} |\eta_k|$$

(since $\sum_{j=1}^{n} |v_{j,k}|^2 = 1$ for all $k \in \{1, ..., n\}$ because V is unitary).

But since $D_2 = \begin{bmatrix} \eta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \eta_n \end{bmatrix}$

and $B = U_2 D_2 U_2^*$, we have that η_k's are the eigenvalues of B.

Since M was arbitrary with $||M||_T \leq 1$, we obtain that

$$||\phi_B||_T^* \leq \sum_{k=1}^{n} |\eta_k| = ||B||_1$$

On the other hand, if we let $M = U_2 \tilde{D} U_2^*$ where

$$\tilde{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

and the λ_k's are defined by

$$\lambda_k = \begin{cases} \frac{\eta_k}{|\eta_k|} & \text{when } \eta_k \neq 0 \\ 0 & \text{otherwise} \end{cases},$$

we get that

$$\phi_B(M) = Tr(MB) = Tr(U_2 \tilde{D} U_2^* U_2 D U_2^*) = Tr(U_2 \tilde{D} D_2 U_2^*) = Tr(\tilde{D} D) = \sum_{k=1}^{n} |\eta_k|.$$

As $|\lambda_k| \leq 1$ for all $k \in \{1, ..., n\}$,

we have $||M||_T \leq 1$.

Since

$$\phi_B(M) = \sum_{k=1}^{n} |\eta_k|,$$

we get that

$$||\phi_B||_T^* \geq \sum_{k=1}^{n} |\eta_k|.$$

All this together gives that

$$||\phi_B||_T^* = \sum_{k=1}^{n} |\eta_k| = ||B||_1,$$

for all $B \in S_n$.

Now we have to show that $K^* = S_+^n$:

Let

$$E_{j,j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & 1_{j,j} & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

If B=U $\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ U^* where U is unitary,

let then $A_{j,j} = UE_{j,j}U^*$. Hence $A_{j,j} \in K$, since the eigenvalues of $A_{j,j}$ are 0 and 1.

Then

$$\phi_B(A_{j,j}) = Tr(A_{j,j}B) = Tr(UE_{j,j}U^*U \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} U^*)$$

$$= Tr(UE_{j,j} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} U^*)$$

$$= Tr(U^*UE_{j,j} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix})$$

$$= Tr(E_{j,j} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}) = \lambda_j.$$

If $\phi_B \in K^*$, (where ϕ_B is given by $\phi_B(M) = Tr(MB)$ for all $M \in S_n$) we must then have $\lambda_j \geq 0$, since $A_{j,j} \in K$.

This is true for all j, hence B must be then a positive semidefinite. Thus $K^* \subseteq S^n_+$. On the other hand, if B is positive semidefinite and $M \in K$, let again $M = U_1 D_1 U_1^*$ and $B = U_2 D_2 U_2^*$ where

$$D_1 = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

$$D_2 = \begin{bmatrix} \eta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \eta_n \end{bmatrix}$$

As we saw, $Tr(MB) = Tr(D_1 A)$ where $A = U_1^* U_2 D_2 U_2^* U_1$

We calculated that

$$Tr(D_1 A) = \sum_{j=1}^{n} \lambda_j a_{j,j} = \sum_{j=1}^{n} \lambda_j (\sum_{k=1}^{n} \eta_k |v_{j,k}|^2) = \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda_j \eta_k |v_{j,k}|^2$$

where $V = [v_{j,k}] = U_1^* U_2$.

But since B was assumed to be positive semidefinite and the η_k's are the eigenvalues of B, it follows that $\eta_k \geq 0$ for all k.

Hence

$$\phi_B(M) = Tr(MB) = Tr(D_1 A) = \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda_j \eta_k |v_{j,k}|^2 \geq 0$$

(here we use that $\lambda_j \geq 0$ for all j since $M \in K$).

This is true for any $M \in K$, hence $\phi_B \in K^*$. Thus $S_+^n \subseteq K^*$. We conclude that $K^* = S_+^n$ and this completes the proof of the remark 2.2.

The remark 2.1 and 2.2 are also given in [GQ] denoted by remark 3.2 and 3.2 respectively, but the proofs are omitted in [GQ]. The next remark, remark 2.3 is not given in [GQ].

Remark 2.3 *Let Ω be a compact Hausdorff space and let $M_r(\Omega)$ denote the space of all signed Radon measures on Ω with the norm $\| v \| = |v|(\Omega)$. For $v \in M_r(\Omega)$, let $\phi_v : C_{\mathbb{R}}(\Omega) \to \mathbb{R}$ be defined by*

$$\phi_v(f) = \int_\Omega f dv.$$

Then $v \to \phi_v$ is an isometric isomorphism of $(M_r, (\Omega), \| \cdot \|)$ onto $((C_{\mathbb{R}}(\Omega))^*, \| \cdot \|)$.

Proof: Consider $\phi \in (C_{\mathbb{R}}(\Omega))^*$. Since ϕ is real, by theorem 13.13 in [MW], ϕ can be written as $\phi = \phi_+ - \phi_-$ where ϕ_+ , $\phi_- \in (C_{\mathbb{R}}(\Omega))^*$ and ϕ_+ , ϕ_- are positive linear functionals .

Since Ω is compact, by Riesz - Markov theorem, there are μ_1 and μ_2 positive Radon measures

s.t.

$$\phi_+(f) = \int_\Omega f d\mu_1 \; \forall f \in C_{\mathbb{R}}(\Omega),$$

$$\phi_-(f) = \int_\Omega f d\mu_2 \; \forall f \in C_{\mathbb{R}}(\Omega).$$

Let $v = \mu_1 - \mu_2$. Then v is a signed Radon measure. (This will be proved at the end of this section under " comment ")

Thus, if $\phi \in (C_{\mathbb{R}}(\Omega))^*$, then there is a signed Radon measure v s.t.

$$\phi(f) = \int_\Omega f dv \; \forall f \in C_{\mathbb{R}}(\Omega).$$

On the other hand, if $v \in M_r(\Omega)$ and

$$\phi_v(f) = \int_\Omega f dv \ \forall f \in C_{\mathbb{R}}(\Omega),$$

then ϕ_v is clearly a real valued linear functional on $C_{\mathbb{R}}(\Omega)$ since v is a signed measure.

Furthermore, ϕ_v is bounded:

$$|\phi_v(f)| = |\int_\Omega f dv| = |\int_\Omega f d(v_+ - v_-|$$

$$= |\int_\Omega f dv_+ - \int_\Omega f dv_-| \leq |\int_\Omega f dv_+| + |\int_\Omega f dv_-|$$

$$\leq \int_\Omega |f| dv_+ + \int_\Omega |f| dv_- = \int_\Omega |f| d(v_+ + v_-)$$

$$= \int_\Omega |f| d|v| \leq ||f||_\infty |v|(\Omega) \ \forall f \in C_{\mathbb{R}}(\Omega).$$

Thus ϕ_v is bounded and $||\phi_v|| \leq |v|(\Omega)$.

Hence $\phi_v \in (C_{\mathbb{R}}(\Omega))^*$

So we have an isomorphism between $M_r(\Omega)$ and $(C_{\mathbb{R}}(\Omega))^*$ via the map $v \to \phi_v$ where $v \in M_r(\Omega)$ and $\phi_v(f) = \int_\Omega f dv \ \forall f \in C_{\mathbb{R}}(\Omega)$.

Now we will prove that this isomorphism is an isometry. Then we have to prove that for all $\phi_v \in (C_{\mathbb{R}}(\Omega))^*$ we have that $||\phi_v|| \geq |v|(\Omega)$ when

$$\phi_v(f) = \int_\Omega f dv \ \forall f \in C_{\mathbb{R}}(\Omega), v \in M_r(\Omega).$$

(We have already proved the opposite inequality.)

Let $\{P, N\}$ be the Hahn decomposition for v, so $v_+(N) = v_-(P) = 0$

Let $\varepsilon > 0$.

Choose K,L compact subsets of Ω s.t. $K \subseteq P, L \subseteq N$ and $v_+(P \setminus K) < \frac{\varepsilon}{4}$ and $v_-(N \setminus L) < \frac{\varepsilon}{4}$ (This is possible since v_+, v_- are regular Borel measures.)

Since $P \bigcap N = \emptyset$ and $K \subseteq P, L \subseteq N$, it follows that $K \bigcap L = \emptyset$.

Furthermore, since K,L are compact and Ω is Hansdorff, it follows that K and L are closed.

Also, since Ω is compact Hausddorff, it is normal.

Hence, by Urisohn's lemma

$$\exists f_1 \in (C_{\mathbb{R}}(\Omega)) \text{ s.t. } f_{1|_K} = 1,\ f_{1|_L} = 0 \text{ and } 0 \leq f_1 \leq 1.$$

Similarly there exists $f_2 \in C_{\mathbb{R}}(\Omega)$ s.t. $f_{2|_K} = 0,\ f_{2|_L} = 1$ and $0 \leq f_2 \leq 1$.

Let $f = f_1 - f_2$.

Then $f \in C_{\mathbb{R}}(\Omega),\ ||f||_\infty \leq 1$ and $f_{|_K} = 1,\ f_{|_L} = -1$.

We claim that $|\int_\Omega f dv| \geq |v|(\Omega) - \varepsilon$:

We have $K \subseteq P, L \subseteq N$ and $v_+(P \setminus K), v_-(N \setminus L) < \frac{\varepsilon}{4}$ and

$f : \Omega \to [-1, 1],\ f(K) = \{1,\ \}\ f(L) = \{-1\}$.

$$\text{Then we get } |\int_\Omega f dv| = |\int_\Omega f dv_+ - \int_\Omega f dv_-|$$

$$= |\int_P f dv_+ - \int_N f dv_-|$$

$$= |\int_K f dv_+ + \int_{P\setminus K} f dv_+ - \int_L f dv_- - \int_{N\setminus L} f dv_-|$$

$$= |\int_K 1 dv_+ + \int_L 1 dv_- + \int_{P\setminus K} f dv_+ - \int_{N\setminus L} f dv_-|$$

$$= |v_+(K) + v_-(L) - (\int_{N\setminus L} f dv_- - \int_{P\setminus K} f dv_+)|$$

$$\geq v_+(K) + v_-(L) - |(\int_{N\setminus L} f dv_- - \int_{P\setminus K} f dv_+|$$

$$\geq v_+(K) + v_-(L) - |\int_{N\setminus L} f dv_-| - |\int_{P\setminus K} f dv_+|$$

$$\geq v_+(K) + v_-(L) - \int_{N\setminus L} |f| dv_- - \int_{P\setminus K} |f| dv_+$$

$$\geq v_+(K) + v_-(L) - \int_{N\setminus L} 1 dv_- - \int_{P\setminus K} 1 dv_+$$

$$= v_+(K) + v_-(L) - v_-(N \setminus L) - v_+(P \setminus K)$$

$$> (v_+(P) - \tfrac{\epsilon}{4}) + (v_-(N) - \tfrac{\epsilon}{4}) - \tfrac{\epsilon}{4} - \tfrac{\epsilon}{4}$$

$$> v_+(P) + v_-(N) - \epsilon = v_+(\Omega) + v_-(\Omega) - \epsilon = |v|(\Omega) - \epsilon.$$

Since $||f||_\infty \leq 1$, we get that

$$||\phi_v|| \geq |\phi_v(f)| = |\int_\Omega f dv| \geq |v|(\Omega) - \epsilon$$

Since $\epsilon > 0$ was arbitrary, we get that

$$||\phi_v|| \geq |v|(\Omega),$$

hence

$$||\phi_v|| = |v|(\Omega).$$

Thus $v \to \phi_v$ is an isometric isomorphism.

Comment:

It was stated that if μ_1 and μ_2 are Radon measures on Ω, then $v = \mu_1 - \mu_2$ is signed Radon measure. We will prove this now:

Let $A \subseteq \Omega$, A Borel. Given $\epsilon > 0$, choose U_1, K_1 and U_2, K_2 s.t.
$K_1 \subseteq A \subseteq U_1$ $K_2 \subseteq A \subseteq U_2$, K_1, K_2 compact, U_1, U_2 open and

$$\mu_1(U_1 \setminus K_1) < \tfrac{\epsilon}{2} \text{ , } \mu_2(U_2 \setminus K_2) < \tfrac{\epsilon}{2}.$$

Let $U = U_1 \cap U_2$, $K = K_1 \cup K_2$.

Then U is open, K is compact, $K \subset A \subset U$,

$$\mu_1(U \setminus K) \leq \mu_1(U_1 \setminus K_1) < \tfrac{\epsilon}{2},$$

$$\text{and } \mu_2(U \setminus K) \leq \mu_2(U_2 \setminus K_2) < \tfrac{\epsilon}{2}.$$

Hence

$$v_+(U \setminus K) = v((U \setminus K) \cap P)$$

$$= \mu_1((U \setminus K) \cap P) - \mu_2((U \setminus K) \cap P)$$

$$\leq \mu_1((U \setminus K) \cap P) + \mu_2((U \setminus K) \cap P)$$
$$\leq \mu_1((U \setminus K) + \mu_2((U \setminus K) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

Similarly $v_-(U \setminus K) < \epsilon$, so v_+ and v_- are regular.

Furthermore , if $K \subseteq \Omega$ is compact, then

$$v_+(K) = v(K \cap P) = \mu_1(K \cap P) - \mu_2(K \cap P))$$
$$\leq \mu_1(K) + \mu_2(K) < \infty.$$

Similarly $v_-(K) < \infty$.

Hence v_+ and v_- are Radon measures.

Since Ω is compact and v_+, v_- are Radon measures, it follows that

$$v_+(\Omega), v_-(\Omega) < \infty.$$

Let now again $A \subseteq \Omega$, A Borel.

Then we have:

$$|v(A)| \leq |v|(A) = v_+(A) + v_-(A) \leq v_+(\Omega) + v_-(\Omega) < \infty.$$

This shows that v is finite signed Radon measure and complets the proof.

We have shown in example 1.6 that $||\cdot||_T$w.r.t the constant function 1 in the space $C_{\mathbb{R}}(\Omega)$ is exactly $||\cdot||_\infty$ when

$$K = \{f \in C_{\mathbb{R}}(\Omega) \mid f(x) \geq 0 \ \forall x \in \Omega\}.$$

Hence the dual of $(C_{\mathbb{R}}(\Omega), ||\cdot||_T)$ is actually $(C_{\mathbb{R}}(\Omega), ||\cdot||_\infty)^*$.

Since $(C_{\mathbb{R}}(\Omega), ||\cdot||_\infty)^*$ is isometrically isomorphic to $M_r(\Omega)$ equipped with the total variation norm (as we proved), it follows that

$$||\mu||_T^* = |\mu|(\Omega) \ \forall \mu \in M_r(\Omega).$$

Also, it is clear that the dual cone and simplex in this case are $K^* = M_+(\Omega)$,

$$P(1) = \{\mu \in M_+(\Omega) \mid \mu(\Omega) = 1\} = P(\Omega),$$

so the simplex is the set of all probability Borel measures on Ω. To see that $K^* = M_+(\Omega)$,, we refer to the exercise 13.75 in [MW] which states that a linear functional ϕ_μ on Ω given by

$$\phi(f) = \int_\Omega f \ d\mu$$

is nonnegative if and only if $\mu \in M_+(\Omega)$.

4 The dual unit ball and its relation to the abstract simplex

In this section we refer to the 2. version of the article by Gaubert and Qu denoted by [GQ2]. We consider again the space $(X^*, ||\cdot||_T^*)$ defined in section 2 and introduce lemma 3.2 where we give a description of the dual unit ball $B_T^*(e)$ of the space $(X^*, ||\cdot||_T^*)$ in terms of the abstract simplex $P(e)$ and we also give a complete proof of this lemma. In order to prove lemma 3.2 , we first need to introduce lemma 3.1 which we will use later in the proof of lemma 3.2. Lemma 3.1 is denoted by (6) in [GQ2] and is given without proof. We give here a detailed proof of this lemma. After proving lemma 3.1 we prove the observation 3 which states that $P(e)$ is a w^* - compact subset of $B_T^*(e)$. This observation will also be used in the proof of lemma 3.2. Finally we state and prove lemma 3.2 . Lemma 3.2 is denoted by " lemma 3.1 " in [GQ2]. In our proof of this lemma, we will mainly follow the proof of lemma 3.1 given in [GQ2], but we supply most of the statements that are used in this proof with further, detailed explanations.

Lemma 3.1 *Let K^* be a dual cone in X^* as defined in the begining of the section 2. Then, for all $z \in K^*$ with $z \neq 0$, we have $< z, e > > 0$. Furthermore, for all $x \in X$, we have*

$$\sup_{z\in K^*,\ z\neq 0} \frac{|< z, x >|}{< z, e >} = ||x||_T$$

Proof: We prove first that if $z \in K^*$ and $z \neq 0$, then $< z, e > > 0$:
Since $e \in \text{Int } K$, given $x \in X$ there exists $t > 0$ s.t. $(e + \frac{1}{t}x) \in K$ and $(e - \frac{1}{t}x) \in K$
If $z \in K^*$, then $< z, e > + \frac{1}{t} < z, x > \geq 0$ and $< z, e > - \frac{1}{t} < z, x > \geq 0$
Hence, if $< z, e > = 0$, we must then have that $\frac{1}{t} < z, x > \geq 0$ and $\frac{1}{t} < z, x > \leq 0$. Thus we get that $< z, x > = 0$. Since x was arbitrary, we deduce that $z = 0$.
So, if $z \neq 0$ and $z \in K^*$, then $< z, e > > 0$. This proves the first statement of the lemma.
Now we prove the second statement.

Let $x \in X,\ x \neq 0$, let $n \in N$ be s.t. $||x||_T - \frac{1}{n} > 0$ and let $s_n = ||x||_T - \frac{1}{n}$. It follows that $x \notin s_n I_e$,
since $||x||_T = \inf\{t > 0 \mid x \in tI_e\}$.

Hence, either

$$(e - \frac{1}{s_n}x) \notin K \text{ or } (e + \frac{1}{s_n}x) \notin K.$$

Suppose $(e + \frac{1}{s_n}x) \notin K$.

Since K is convex and closed, by corollay 14.7 in [MW], (X is LCS since it is a normed space)

$$\exists z \in X^* \text{ s.t. } < z, e + \frac{1}{s_n}x > < \inf z(K).$$

It follows then that $\inf z(K) > -\infty$.
Assume now that there exists an $y \in K$ s.t. $< z, y > < 0$. Since K is a cone, then $\lambda y \in K$ for all $\lambda > 0$.
Hence $< z, \lambda y > = \ \lambda < z, y > \in z(K)$ for all $\lambda > 0$. Since $< z, y > < 0$, it follows that

$$\lim_{\lambda \to \infty} \lambda < z, y > = -\infty.$$

Then we get that $\inf z(K) = -\infty$ and this is a contradiction. So there is no $y \in K$ s.t. $< z, y > < 0$, hence we have $< z, y > \geq 0$ for all $y \in K$. Thus $z \in K^*$.
Since $z \in K^*$, we get that $\inf z(K) \geq 0$. As $0 \in K$, we deduce that $\inf z(K) = 0$.
Thus

$$< z, e + \frac{1}{s_n}x > < \inf z(K) = 0.$$

This gives that $\frac{1}{s_n} < z, x > < - < z, e > \leq 0$.
Hence $\frac{1}{s_n} | < z, x > | \ > < z, e >$. Now, since $z \in K^*$ and $z \neq 0$, we have $< z, e > > \ 0$, by the first statement of the lemma. So we can divide with $\frac{<z,e>}{S_n}$ on the both sides of the inequality and obtain

$$\frac{| < z, x > |}{< z, e >} > s_n = ||x||_T - \frac{1}{n}.$$

Since $z \in K^*$, we get that

$$\sup_{\tilde{z} \in K^*, \tilde{z} \neq 0} \frac{| < \tilde{z}, x > |}{< \tilde{z}, e >} \geq \frac{| < z, x > |}{< z, e >} > s_n = ||x||_T - \frac{1}{n}.$$

Since this is true for all n with $||x||_T - \frac{1}{n} > 0$, letting $n \to \infty$,we obtain

$$\sup_{\tilde{z} \in K^*, z \neq 0} \frac{| < \tilde{z}, x > |}{< \tilde{z}, e >} \geq ||x||_T.$$

Now, if $(e - \frac{1}{s_n}x) \notin K$, in the same way

$$\exists w \in X^* s.t. < w, e - \frac{1}{s_n}x > < \inf w(K).$$

Then, using, similar arguments we can deduce that $\frac{<w,x>}{<w,e>} > s_n = ||x||_T - \frac{1}{n}$ and that $w \in K^*$.

Hence

$$\sup_{\tilde{w} \in K^*, \tilde{w} \neq 0} \frac{| < \tilde{w}, x > |}{< \tilde{w}, e >} > s_n = ||x||_T - \frac{1}{n}.$$

Again, letting $n \to \infty$, we get that

$$\sup_{w \in K^*, w \neq 0} \frac{| < w, x > |}{< w, e >} \geq ||x||_T.$$

Now we have to prove the opposite inequality, that is

$$\sup_{z \in K^*, z \neq 0} \frac{| < z, x > |}{< z, e >} \leq ||x||_T :$$

We have

$$||x||_T = \inf\{t > 0 | x \in tI_e\}.$$

Assume that $x \in tI_e$ for some $t > 0$ and let $z \in K^*,\ z \neq 0$.

Then $e - \frac{1}{t}x \in K$ and $e + \frac{1}{t}x \in K$.

Hence

$$< z, e > -\frac{1}{t} < z, x > \geq 0 \text{ and } < z, e > +\frac{1}{t} < z, x > \geq 0.$$

Thus

$$- < z, e > \leq \frac{1}{t} < z, x > \leq < z, e > \text{, which gives that}$$

$$\frac{1}{t}| < z, x > | \leq | < z, e > | = < z, e >$$

$$(\text{since } z \in K^* \ < z, e > > 0, \text{ so } < z, e > = | < z, e > |)$$

Hence

$$\frac{1}{t}\frac{| < z, x > |}{< z, e >} \leq 1,$$

$$\text{so} \quad \frac{|< z, x >|}{< z, e >} \leq t.$$

Since this is true for any t s.t. $x \in tI_e$, taking inf over all such t's, we get that

$$\frac{|< z, x >|}{< z, e >} \leq ||x||_T.$$

Now, since $z \in K^*$, $z \neq 0$ was arbitrary, we get that

$$\sup_{\substack{z \in K^* \\ z \neq 0}} \frac{|< z, x >|}{< z, e >} \leq ||x||_T.$$

Combining these 2 inequalities together, we obtain that

$$\sup_{z \in K^*, z \neq 0} \frac{|< z, x >|}{< z, e >} = ||x||_T.$$

This complets the proof of lemma 3.1
Next we have the following observation:

Observation 3 $P(e)$ is a w^*- compact subset of $B_T^*(e)$:

Proof: First we observe that $P(e) \subseteq B_T^*(e)$:

Let $\tilde{\mu} \in P(e)$. Then

$$|< \tilde{\mu}, x >| \ \leq \ \sup_{\mu \in P(e)} |< \mu, x >|$$

$$= \sup_{\mu \in P(e)} \frac{|< \mu, x >|}{< \mu, x >} \leq \sup_{\substack{z \in K^* \\ z \neq 0}} \frac{|< z, x >|}{< z, e >} = ||x||_T$$

(Here we have used that $< \mu, e > = 1$ for all $\mu \in P(e)$, $P(e) \subseteq K^* \setminus \{0\}$ and for the last equality, we have used lemma 3.1)

This is true for all $x \in X$, hence $||\tilde{\mu}||_T^* \leq 1$. Since $\tilde{\mu} \in P(e)$ was arbitrary, we get that $P(e) \subseteq B_T^*(e)$.
Next we show that $P(e)$ is w^*- closed:
Let $\{v_\alpha\}$ be a net in $P(e)$ and assume that $v_\alpha \to v$ in the w^* topology. This simply means that $< v_\alpha, x > \to < v, x >$ for all $x \in X$. Since $< v_\alpha, x > \geq 0$ for all $x \in K$ and for all α (as $v_\alpha \in P(e) \subseteq K^*$ for all α), we get that $< v, x > \geq 0$ for all $x \in K$. Thus $v \in K^*$.Also since $< v_\alpha, e > = 1$ for all α as $v_\alpha \in P(e)$ for all α, we get that $< v, e > = 1$ Thus $v \in P(e)$. Hence $P(e)$ is w^* - closed. Since $B_T^*(e)$ is w^* - compact by Alaoglu's theorem and $P(e)$ is a w^* closed subset of $B_T^*(e)$, it follows that $P(e)$ is w^* compact. This proves observation 3.

Lemma 3.2 *The unit ball $B_T^*(e)$ of the space $(X^*, ||\cdot||_T^*)$ satisfies*

$$B_T^*(e) = \text{conv}(P(e) \cup (-P(e))).$$

Proof: To simplify notation, throughout this proof we let $P = P(e)$.

We will prove this lemma by proving the following:
a) $||x||_T = \sup_{\mu \in P} | < \mu, x > | = \sup_{\mu \in P \cup (-P)} < \mu, x >$
b) $B_T^*(e) = \overline{\text{conv}}^{w^*}(P \cup (-P))$
c) $\overline{\text{conv}}^{w^*}(P \cup (-P)) = \text{conv}(P \cup (-P))$

Proof of a): Let $z \in K^*$, $z \neq 0$. Then $< z, e > > 0$ as we have proved in lemma 3.1. Also, if we let $\mu = \frac{z}{<z,e>}$, then clearly $\mu \in P$. From lemma 3.1 we have

$$||x||_T = \sup_{z \in K^*, z \neq 0} \frac{| < z, x > |}{< z, e >}.$$

Using all this together, we get that

$$||x||_T = \sup_{z \in K^*, z \neq 0} \frac{| < z, x > |}{< z, e >} = \sup_{z \in K^*, z \neq 0} | < \frac{z}{< z, e >}, x > |$$

$$\leq \sup_{\mu \in P} | < \mu, x > |.$$

On the other hand, if $\mu \in P$ then $< \mu, e > = 1$ and $\mu \in K^*$ by definition of P, so

$$\sup_{\mu \in P} | < \mu, x > | = \sup_{\mu \in P} \frac{| < \mu, x > |}{< \mu, e >} \leq \sup_{z \in K^*, z \neq 0} \frac{| < z, x > |}{< z, e >} = ||x||_T.$$

Hence

$$||x||_T = \sup_{\mu \in P} | < \mu, x > |.$$

Now

$$||x||_T = \sup_{\mu \in P} | < \mu, x > | = \sup_{\mu \in P} \{\max\{< \mu, x >, - < \mu, x >\}\}$$

$$= \sup_{\mu \in P} \{\max\{< \mu, x >, < -\mu, x >\}\} = \sup_{\mu \in P \cup (-P)} < \mu, x > .$$

Thus

$$||x||_T = \sup_{\mu \in P} | < \mu, x > | = \sup_{\mu \in P \cup (-P)} < \mu, x > .$$

and this proves the part a).

Let now $z \in X^*$.

We then have

$$||z||_T^* \leq 1 \ \Leftrightarrow | < z, x > | \leq ||x||_T \ \ \forall x \in X.$$

Since

$$||x||_T = \sup_{\mu \in P \cup (-P)} < \mu, x > ,$$

we get that $||z||_T^* \leq 1 \Leftrightarrow \sup_{\mu \in P \cup (-P)} < \mu, x > \geq | < z, x > | \;\; \forall x \in X.$ (1)

We wish to use this to prove the part b), that is

$$B_T^*(e) = \overline{\text{conv}}^{w^*}(P \cup (-P)).$$

Proof of b) First we prove that $B_T^*(e) \subseteq \overline{\text{conv}}(P \cup (-P))^*$.

Suppose, this is not the case.

Then there exists a $z \in B_T^*(e) \setminus \overline{\text{conv}}(P \cup (-P))$

(Here we always consider the closure w.r.t the w^* - topology.)

Since $z \notin \overline{\text{conv}}(P \cup (-P))$, by prop. 14.9. part e) in [MW], there exists $\phi \in (X^*)^*$

$$\text{s.t. } \phi(z) > \sup \; \phi(P \cup (-P)).$$

(Here we use that X^* with w^* -topology is locally convex topological space.)

Thus there exists scalar $\gamma \in \mathbb{R}$

$$s.t. \; \phi(z) > \gamma \geq \sup \; \phi(P \cup (-P)) \geq \phi(\mu) \; \forall \mu \in P \cup (-P).$$

Now, by theorem 1.3. in [C], we have $(X^*, w^*)^* \cong X$. Thus there exists an $x \in X$ $s.t.$ $\phi(\mu) = < \mu, x >$ for all $\mu \in X^*$

So $< z, x > = \phi(z) > \gamma \geq \phi(\mu) = < \mu, x >$ for all $\mu \in P \cup (-P)$.

Hence

$$< z, x > > \gamma \geq \sup_{\mu \in P \cup (-P)} < \mu, x > \quad as \;\; \gamma \geq < \mu, x > \forall \mu \in P \cup (-P).$$

On the other hand since $z \in B_T^*(e)$, we have $||z||_T^* < 1$. By (1), it follows then that

$$< z, x > \leq | < z, x > | \leq \sup_{\mu \in P \cup (-P)} < \mu, x >$$

Thus we get a contradiction.
We can conclude then that there is no

$$z \in B_T^*(e) \setminus \overline{\text{conv}}(P \cup (-P)),$$

hence we must have

$$B_T^*(e) \subseteq \overline{\text{conv}}(P \cup (-P)).$$

Now, we prove that

$$\overline{\text{conv}}(P \cup (-P)) \subseteq B_T^*(e) :$$

By observation 3, $P \subseteq B_T^*(e)$ and then clearly also $(-P) \subseteq B_T^*(e)$, so

$$P \cup (-P) \subseteq B_T^*(e).$$

Since $B_T^*(e)$ is w^* - compact by Alaoglu's theorem and the w^* - topology is Hausdorff, we get that $B_T^*(e)$ is w^* - closed. Also $B_T^*(e)$ is obviously convex. Hence, by prop 14.9. part d) in [MW], $\overline{\text{conv}}(P \cup (-P)) \subseteq B_T^*(e)$. Combining these 2 inclusions, we deduce that $B_T^*(e) = \overline{\text{conv}}(P \cup (-P))$. This proves the part b).

Proof of c): We observe first that if $y \in \text{conv}(P \cup (-P))$, then

$$y = \sum_{k=1}^{n} a_k v_k + \sum_{j=1}^{m} b_j w_j,$$

$$\text{where } a_k \geq 0, \; b_j \geq 0 \; \forall k, j$$

$$v_k \in P, \; w_j \in (-P) \; \forall k, j$$

$$\text{and } \sum_{k=1}^{n} a_k + \sum_{j=1}^{m} b_j = 1.$$

Since $w_j \in (-P)$ for all j, it follows that for all j there is some $\pi_j \in P$ s.t. $w_j = -\pi_j$.

Hence

$$y = \sum_{k=1}^{n} a_k v_k - \sum_{j=1}^{m} b_j \pi_j$$

where $v_k, \pi_j \in P \ \ \forall k, j$.

Let

$$s = \sum_{k=1}^{n} a_k \quad t = \sum_{j=1}^{m} b_j.$$

Then $s, t \geq 0$ and $s + t = 1$.

Furthermore, if we let

$$v = \sum_{k=1}^{n} \frac{a_k}{s} v_k, \quad \pi = \sum_{j=1}^{m} \frac{b_j}{t} \pi_j \text{ when } s, t > 0,$$

$$\text{then } < v, e > = \sum_{k=1}^{n} \frac{a_k}{s} < v_k, e > = \sum_{k=1}^{n} \frac{a_k}{s} = 1 \text{ (since } v_k \in P \ \forall k).$$

Similarly $< \pi, e > = 1$ since $< \pi_j, e > = 1$ for all j. Since $a_k, b_j \geq 0$ for all k, j and $v_k, \pi_j \in P \subseteq K^*$, we have that v and π are in K^*. Thus we get that $v, \pi \in P$.

This shows that if $y \in$ conv $P \cup (-P)$, then $y = sv - t\pi$ where $v, \pi \in P$ $s, t \geq 0$ and $s + t = 1$. Since $B_T^* = \overline{conv}(P \cup (-P))$, if $\mu \in B_T^*(e)$, then there exists a net $\{\mu_\alpha\}_{\alpha \in A}$ in conv$(P \cup (-P))$, s.t. $\mu_\alpha \to \mu$ in w^* - topology. But as we have shown, any such μ_α can be written as $s_\alpha v_\alpha - t_\alpha \pi_\alpha$ where $v_\alpha, \pi_\alpha \in P$, $s_\alpha, t_\alpha \geq 0$ and $s_\alpha + t_\alpha = 1$. Since $\{v_\alpha\}_{\alpha \in A}$, $\{\pi_\alpha\}_{\alpha \in A}$ are nets in $P, \{s_\alpha\}_{\alpha \in A}, \{t_\alpha\}_{\alpha \in A}$ are nets in [0,1] and P is w^* - compact by observation 3, by passing to a subnet if necessary, we may assume that $v_\alpha \underset{w^*}{\to} v$,$\pi_\alpha \underset{w^*}{\to} \pi$ for some $v, \pi \in P$ and $s_\alpha \to s$, $t_\alpha \to t$ for some $s, t \in [0, 1]$.

Then the net $\{s_\alpha v_\alpha - t_\alpha \pi_\alpha\}_{\alpha \in A}$ converges to $(sv - t\pi)$ in the w^* -topology.

Moreover, since $s_\alpha + t_\alpha = 1$ for all α and $(s_\alpha + t_\alpha) \to (s + t)$ we get that $s + t = 1$.

But we have $(s_\alpha v_\alpha - t_\alpha \pi_\alpha) \underset{w^*}{\to} \mu$, so we must have $\mu = sv - t\pi$.

Since $v, \pi \in P$, $s, t \in [0, 1]$ and $s + t = 1$,
we have $\mu = sv - t\pi \in$ conv$(P \cup (-P)$. Since $\mu \in \overline{\text{conv}}(P \cup (-P))$ was arbitrary, we get that

$$\overline{\text{conv}}(P \cup (-P)) = \text{conv}(P \cup (-P)).$$

Thus

$$B_T^*(e) = \overline{\text{conv}}(P \cup (-P)) = \text{conv}(P \cup (-P))$$

and this completes the proof of lemma 3.2.

5 Hilbert's quotient norm and dual norm. Characterisation of disjoint, extreme points of the simplex.

In this section, we return back to the 3. version of the article of Gaubert and Qu denoted by [GQ]. We will first consider the quotient space $X/\mathbb{R}e$ and define a norm $|||\cdot|||_H$ on this space. In lemma 4.2 we will prove that the dual of this space $((X/\mathbb{R}e)^*, |||\cdot|||_H^*)$ is isometrically isomorphic to $(M(e), ||\cdot||_H^*)$ where $M(e)$ is the anihilator of $\mathbb{R}e$ in X^* and $||\cdot||_H^* = \frac{1}{2}||\cdot||_T^*$. (Recall that $||\cdot||_T^*$ was already defined in section 2). After that we will consider the simplex $P(e)$, define "disjointness" of elements in $P(e)$ and then give a concrete description of disjoint, extreme points of $P(e)$ in different examples\ remarks. Throughout section 4 we will mainly follow section 4 in [GQ]. However, the lemma 4.2 in [GQ] is slightly modified and reformulated here. In addition, all remarks and examples in [GQ] are given here with complete proofs.

Given $(X, e, ||\cdot||_T)$, consider now the quotient space $X/\mathbb{R}e$ and define Hilbert's quotient norm $|||\cdot|||_H : X/\mathbb{R}e \to \mathbb{R}^+$ by

$$|||x + \mathbb{R}e|||_H = 2 \inf_{\lambda\in\mathbb{R}} ||x + \lambda e||_T.$$

Then $|||\cdot|||_H$ is a norm on $X/\mathbb{R}e$:
Let $|||\cdot|||_T : X/\mathbb{R}e \to \mathbb{R}^+$ be given by

$$|||x + \mathbb{R}e|||_T = 2 \inf_{\lambda\in\mathbb{R}} ||x + \lambda e||_T.$$

Since $(X, ||\cdot||_T)$ is a Banach space and $\mathbb{R}e$ is a closed subspace of X, by theorem 4.2 in chapter 3 in [C], we have that $|||\cdot|||_T$ is a norm on $X/\mathbb{R}e$. Now, since $|||\cdot|||_H = \frac{1}{2}|||\cdot|||_T$ by definition, it follows that $|||\cdot|||_H$ is also a norm on $X/\mathbb{R}e$.

Comment: In [GQ] they consider $((X/\mathbb{R}e, ||\cdot||_H)$ instead of

$$((X/\mathbb{R}e, |||\cdot|||_H),$$

where $||\cdot||_H$ is a Hilbert seminorm on X w.r.t. e that is defined is section 1 ($||x||_H = M(x/e) - m(X/e)$). It turns out that $||x||_H = |||x + \mathbb{R}e|||_H$ for all $x \in X$: By our definition,

$$|||x + \mathbb{R}e|||_H = 2 \inf_{\lambda\in\mathbb{R}} ||x + \lambda e||_T.$$

Now, by lemma 1.4 is section 1 (lemma 4.1 in [GQ])

$$||x||_H = 2 \inf_{\lambda \in \mathbb{R}} ||x + \lambda e||_T.$$

Hence $||x||_H = |||x + \mathbb{R}e|||_H$ for all $x \in X$, so in [GQ] they actually identify $||\cdot||_H$ with $|||\cdot|||_H$. However, it is not quite precise to write "$((X/\mathbb{R}e, ||\cdot||_H)$" as they do in [GQ] , since, by definition, $||\cdot||_H$ acts on X whereas $|||\cdot|||_H$ acts on the quotient space $X/\mathbb{R}e$.

Furthermore, let $M(e) = \{\mu \in X^* | \ < \mu, e > \ = 0\}$ and define Hilbert's dual norm

$$|| \cdot ||_H^* : M(e) \to \mathbb{R}^+ \text{by}$$

$$||\mu||_H^* = \frac{1}{2} ||\mu||_T^* \ \ \forall \mu \in M(e).$$

We have then the following lemma:
Lemma 4.2 $(X/\mathbb{R}e^*, |||\cdot|||_H^*)$ *is isometrically isomorphic to* $(M(e), ||\cdot||_H^*)$.

To prove this, we will first prove the following lemma:

Lemma 4.1 *Let Z be a normed space with the norm* $\| \cdot \|_1$. *Let* $\| \cdot \|_2$ *be another norm on Z given by* $\| z \|_2 = C \| z \|_1$ *for all* $z \in Z$ *where* $C > 0$ *is a constant. Then*

$$\| \varphi \|_2^* = \frac{1}{C} \| \varphi \|_1^* \ \ \textit{for all } \varphi \in Z^*$$

Proof: Assume that $\| z \|_2 \leq 1$.

Then $\| z \|_1 \leq \frac{1}{C}$, hence $\| Cz \|_1 \leq 1$ which gives that

$$C|\phi(z)| = |\phi(Cz)| \leq \| \phi \|_1^* \quad \forall \phi \in Z^*.$$

So

$$|\phi(z)| \leq \frac{1}{C} \| \phi \|_1^* \ \ \forall \phi \in Z^*.$$

Thus

$$\sup_{\|z\|_2 \leq 1} |\phi(z)| = \| \phi \|_2^* \leq \frac{1}{C} \| \phi \|_1^* \quad \forall \phi \in Z^*.$$

Now, since we have

$$\| \cdot \|_1 = \frac{1}{C} \| \cdot \|_2$$

on Z, by the same argument we can deduce that

$$\| \cdot \|_1^* \leq \frac{1}{\frac{1}{C}} \| \cdot \|_2^* = C \| \cdot \|_2^* .$$

So

$$\frac{1}{C} \parallel \cdot \parallel_1^* \leq \parallel \cdot \parallel_2^* .$$

This gives that

$$\parallel \cdot \parallel_2^* = \tfrac{1}{C} \parallel \cdot \parallel_1^*,$$

which proves the lemma 4.1. Now we prove the lemma 4.2:
By theorem 10.2 in chapter 3 in [C], we have $((X/\mathbb{R}_e)^*, ||| \cdot |||_T^*)$ is isometrically isomorphic to $(M(e), \| \cdot \|_T^*)$.

Since $||| \cdot |||_H$ is the norm on $X/\mathbb{R}e$ given by

$$||| \cdot |||_H = 2||| \cdot |||_T.$$

the lemma 4.1 gives

$$||| \cdot |||_H^* = \tfrac{1}{2}||| \cdot |||_T^*.$$

Since $((X/\mathbb{R}_e)^*, ||| \cdot |||_T^*)$ is isometrically isomorphic to $(M(e), \| \cdot \|_T^*)$ and $||| \cdot |||_H^* = \frac{1}{2}||| \cdot |||_T^*$, it follows that $((X/\mathbb{R}_e)^*, ||| \cdot |||_H^*)$ is isometrically isomorphic to $(M(e), \frac{1}{2} \parallel \cdot \parallel_T^*)$. This proves lemma 4.2. since $|| \cdot ||_H^* = \frac{1}{2}|| \cdot ||_T^*$ by definition.

The lemma 4.2 implies that the unit ball of the space $(M(e), || \cdot ||_H^*)$, denoted by $B_H^*(e)$, satisfies:

$$B_H^*(e) = 2B_T^*(e) \cap M(e).$$

Remark 4.3 *In the case of the standard positive cone (example 1.4, $X = \mathbb{R}^n$, $K = \mathbb{R}^n$ and $e = \vec{1}$) we claim that implies that for any two probability $\mu, v \in P(\vec{1})$. the dual norm $||\mu - v||_H^*$ is the total variation distance between μ and v:*

$$||\mu - v||_H^* = \frac{1}{2}||\mu - v||_1 = ||\mu - v||_{TV}$$

Proof We have already proved in remark 3.1 that in this case we have $|| \cdot ||_T^* = || \cdot ||_1$.

Hence

$$||\cdot||_H^* = \frac{1}{2}||\cdot||_T^* = \frac{1}{2}||\cdot||_1$$

Let now

$$w \in M(e), \text{ that is } \sum_{k=1}^{n} w_i = 0.$$

Set

$$J = \{i \mid 1 \leq i \leq n,\ w_i \geq 0\},$$
$$J^c = \{i \mid 1 \leq i \leq n,\ w_i < 0\}.$$

If $L \subseteq \{1, ..., n\}$, then by definition of J and J^c, we have

$$|\sum_{i \in L} w_i| = |\sum_{i \in L \cap J} w_i + \sum_{i \in L \cap J^c} w_i|$$

$$= |\sum_{i \in L \cap J} w_i| - |\sum_{i \in L \cap J^c} w_i| \leq |\sum_{i \in L \cap J} w_i|$$

$$= \sum_{i \in L \cap J} w_i \leq \sum_{i \in J} w_i.$$

Since $L \subseteq \{1, ..., n\}$ was arbitrary, we get that

$$\sum_{i \in J} w_i = \sup_{L \subseteq \{1,...,n\}} |\sum_{i \in L} w_i| = ||w||_{TV}.$$

Now, since

$$0 = \sum_{i=1}^{n} w_i = \sum_{i \in J} w_i + \sum_{i \in J^c} w_i = \sum_{i \in J} |w_i| - \sum_{i \in J^c} |w_i|,$$

we get that

$$\sum_{i \in J} |w_i| = \sum_{i \in J^c} |w_i| = \sum_{i \in J} w_i \quad (\text{as } \ w_i \geq 0 \ \ \forall i \in J).$$

Hence

$$||w||_1 = \sum_{i=1}^{n} |w_i| = \sum_{i \in J} |w_i| + \sum_{i \in J^c} |w_i|$$

$$= 2\sum_{i \in J} w_i = 2||w||_{TV}.$$

Thus $\frac{1}{2}||w||_1 = ||w||_{TV}$, so in general $\frac{1}{2}||\cdot||_1 = ||\cdot||_{TV}$ on $M(\vec{1})$ as $w \in M(\vec{1})$ was arbitrary. This proves the statement in remark 4.3. The remark 4.3 is also given in [GQ] , but without proof.

Definition 4.4 For all $v, \pi \in P(e)$, we say that v and π are disjoint, denoted by $v \perp \pi$, if

$$\mu = \frac{v + \pi}{2}$$

for all $\mu \in P(e)$ such that $\mu \geq \frac{v}{2}$ and $\mu \geq \frac{\pi}{2}$. (This definition is also given in [GQ]).The notation $\mu \geq \frac{v}{2}$ and $\mu \geq \frac{\pi}{2}$, means that $\mu - \frac{v}{2}$ and $\mu - \frac{\pi}{2}$ are in K^*.

Example 4.5 In the case of the standard positive cone ($X = \mathbb{R}^n$, $K = \mathbb{R}^n_+$ and $e = \vec{1}$), we claim that two points v, π in $P(\vec{1})$ are disjoint if and only if for all $i \in \{1, ..., n\}$, $v_i = 0$ or $\pi_i = 0$ holds, meaning that v and π, thought of as discrete probability measures, have disjoint supports:
Proof: We observe first that if $v,\ \pi \in P(\vec{1})$ then $\mu \geq \frac{\pi}{2}$ and $\mu \geq \frac{v}{2}$ if and only if $\mu_i - \frac{\pi_i}{2} \geq 0$ and $\mu_i - \frac{v_i}{2} \geq 0$ for all i, since $K^* = \mathbb{R}^*_+$ by remark 2.1.

Hence

$$\mu_i \geq \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\}\ \forall i.$$

Let now

$$I = \{i|\ i \in \{1, ..., n\} \text{ and } \pi_i \geq v_i\}.$$

Then

$$\sum_{i=1}^{n} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} = \sum_{i=1}^{n} \frac{v_i}{2} + \sum_{i \in I} (\frac{\pi_i}{2} - \frac{v_i}{2}).$$

Since $v_i, \pi_i \geq 0$ for all i, because $v, \pi \in P(\vec{1})$, we have:

$$\sum_{i \in I} (\frac{\pi_i}{2} - \frac{v_i}{2}) \leq \sum_{i \in I} \frac{\pi_i}{2} \leq \sum_{i=1}^{n} \frac{\pi_i}{2} = \frac{1}{2} \text{ (since } \pi \in P(e)).$$

This gives that:

$$\sum_{i=1}^{n} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} = \sum_{i=1}^{n} \frac{v_i}{2} + \sum_{i \in I} (\frac{\pi_i}{2} - \frac{v_i}{2}) =$$

$$= \frac{1}{2} + \sum_{i \in I} (\frac{\pi_i}{2} - \frac{v_i}{2}) \leq \frac{1}{2} + \frac{1}{2} = 1.$$

Assume now that v and π in $P(\vec{1})$ do not have disjoint support. This means that we have $v_j \neq 0$ and $\pi_j \neq 0$ for some $j \in \{1, ..., n\}$.

Set

$$M = 1 - \sum_{i=1}^{n} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\}.$$

Then $M \geq 0$, as we have shown .

Let $\mu \in \mathbb{R}^n$ be given by

$$\mu_i = \begin{cases} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} + \frac{M}{n-1} & \text{for } i \neq j; \\ \max\{\frac{v_j}{2}, \frac{\pi_j}{2}\} & \text{for } i = j. \end{cases}$$

Then $\mu_i - \frac{\pi_i}{2} \geq 0$ and $\mu_i - \frac{v_i}{2} \geq 0$ for all $i \in \{1, ..., n\}$, so $\mu \geq \frac{\pi}{2}$ and $\mu \geq \frac{v}{2}$.

Furthermore

$$\sum_{i=1}^{n} \mu_i = (\sum_{i \neq j, 1 \leq i \leq n} (\max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} + \frac{M}{n-1})) + \max\{\frac{v_j}{2}, \frac{\pi_j}{2}\}$$

$$= M + \sum_{i=1}^{n} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} = 1,$$

so $\mu \in P(\vec{1})$. But $\mu \neq \frac{\pi+v}{2}$ since $\mu_j = \max\{\frac{v_j}{2}, \frac{\pi_j}{2}\} < \frac{v_j+\pi_j}{2}$ because $v_j > 0$ and $\pi_j > 0$.

Hence, if $v, \pi \in P(\vec{1})$ are s.t. whenever $\mu \in P(\vec{1})$ and $\mu \geq \frac{\pi}{2}, \mu \geq \frac{v}{2}$, implies that $\mu = \frac{v+\pi}{2}$, then v and π must have disjoint support.

Assume now that $v, \pi \in P(\vec{1})$ have disjoint support. Then for each i we have that max $\{\frac{v_i}{2}, \frac{\pi_i}{2}\} = \frac{v_i}{2} + \frac{\pi_i}{2}$ since either $v_i = 0$ or $\pi_i = 0$. Then, if μ is s.t. $\mu \geq \frac{v}{2}$ and $\mu \geq \frac{\pi}{2}$, that is $\mu_i \geq \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\}$ for all i, we get that

$$\sum_{i=1}^{n} \mu_i \geq \sum_{i=1}^{n} \max\{\frac{v_i}{2}, \frac{\pi_i}{2}\} = \sum_{i=1}^{n} \frac{v_i + \pi_i}{2} = \sum_{i=1}^{n} \frac{v_i}{2} + \sum_{i=1}^{n} \frac{\pi_i}{2} = \frac{1}{2} + \frac{1}{2} = 1.$$

So, if we in addition want that $\mu \in P(\vec{1})$, that is

$$\sum_{i=1}^{n} \mu_i = 1,$$

then we must have

$$\mu_i = \max\{\tfrac{v_i}{2}, \tfrac{\pi_i}{2}\} = \tfrac{v_i+\pi_i}{2} \ \forall i.$$

But then $\mu = \frac{v+\pi}{2}$. This complets the proof.
The example 4.5 is also given in [GQ] but without proof.

We have the following characterization of the disjointness property.

Lemma 4.6 *Let $v, \pi \in P(e)$. The following assertions are equivalent:*

a) $v \perp \pi$

b) The only elements $\rho, \sigma \in P(e)$ satisfying $v - \pi = \rho - \sigma$ are $\rho = v$ and $\sigma = \pi$.

Proof: a) $\Longrightarrow$ b) : Let any $\rho, \sigma \in P(e)$ be such that $v - \pi = \rho - \sigma$. Then is it immediate that $v + \sigma = \pi + \rho$. Let $\mu = \frac{v+\sigma}{2} = \frac{\pi+\rho}{2}$. Then $\mu \in P(e), \mu \geq \frac{v}{2}$ and $\mu \geq \frac{\pi}{2}$. Since $v \perp \pi$, we obtain that $\mu = \frac{v+\pi}{2}$. It follows that $\rho = v$ and $\sigma = \pi$.
b) $\Longrightarrow$ a): Let $\mu \in P(e)$ be such that $\mu \geq \frac{v}{2}$ and $\mu \geq \frac{\pi}{2}$. Then

$$v - \pi = (2\mu - \pi) - (2\mu - v).$$

From b) we know that $2\mu - \pi = v$.

We denote by extr$(\cdot)$ the set of extreme points of a convex set.

Proposition 4.7 *The set of extreme points of $B^*_H(e)$, denoted by extr $(B^*_H(e))$, is characterized by:*

$$extr\ (B^*_H(e)) = \{v - \pi \mid v, \pi \in \ extr(P(e)),\ v \perp \pi\}.$$

Proof: It follows from lemma 3.2 that every $\mu \in B^*_T(e)$ can be written as

$$\mu = sv - t\pi$$

with $s + t = 1,\ s, t \geq 0,\ v, \pi \in P(e)$. Moreover, if $\mu \in M(e)$, then

$$0 = < \mu, e > = s < v, e > -t < \pi, e > = s - t.$$

Thus $s = t = \frac{1}{2}$. Therefore every $\mu \in B^*_T(e) \cap M(e)$ can be written as

$$\mu = \frac{v - \pi}{2},\ v, \pi \in P(e).$$

Therefore by (13), we have proved that

$$B_H^*(e) = \{v - \pi \ : \ v, \pi \in P(e)\}.$$

Now let $v, \pi \in P(e)$ and $v \perp \pi$. We are going to prove that $v - \pi \in \text{extr} B_H^*(e)$. Let $v_1, \pi_1, v_2, \pi_2 \in P(e)$ be such that

$$v - \pi = \frac{v_1 - \pi_1}{2} + \frac{v_2 - \pi_2}{2}$$

Then

$$v - \pi = \frac{v_1 + v_2}{2} - \frac{\pi_1 + \pi_2}{2}$$

By lemma 4.6, the only possibility is $2v = v_1 + v_2$ and $2\pi = \pi_1 + \pi_2$. Since $v, \pi \in \text{extr} P(e)$ we obtain that $v_1 = v_2 = v$ and $\pi_1 = \pi_2 = \pi$. Therefore $v - \pi \in \text{extr} B_H^*(e)$

Now let $v, \pi \in P(e)$ such that $v - \pi \in \text{extr}(B_H^*(e))$. Assume for contradiction that v is not extreme in $P(e)$(the case where π is not extreme can be dealt with similarly). Then, we can find $v_1, v_2 \in P(e),\ v_1 \neq v_2$, such that $v = \frac{v_1 + v_2}{2}$

It follows that

$$\mu = \frac{v_1 - \pi}{2} + \frac{v_2 - \pi}{2}$$

where $v_1 - \pi, v_2 - \pi$ are distinct elements of $B_H^*(e)$, which is a contradiction. Next we show that $v \perp \pi$. To this end let $\rho, \sigma \in P(e)$ be such that

$$v - \pi = \rho - \sigma.$$

Then

$$v - \pi = \frac{v - \pi + \rho - \sigma}{2} = \frac{v - \sigma}{2} + \frac{\rho - \pi}{2}.$$

If $\sigma \neq \pi$, then $v - \sigma \neq v - \pi$ and this contradicts the fact that $v - \pi$ is extremal. Therefore $\sigma = \pi$ and $\rho = v$ From Lemma 4.6, we deduce that $v \perp \pi$. This completes the proof of lemma 4.6.

The lemma 4.6 and proposition 4.7 with proofs are alreeady given in [GQ]. We will give now a remark on the proof of proposition 4.7 that is not given in [GQ].

Remark on the proof of prop 4.7

In this proof it is used the general assumption that if C is a convex set, $v \in C$ and v is not an extreme point of C, then

$$\exists v_1, v_2 \in C,\ v_1 \neq v_2, \text{ s.t. } v = \tfrac{v_1+v_2}{2}.$$

We are going to prove this:

According to the definition 14.9. in [MW], if v is not an extreme point, then

$$\exists v_1, ..., v_n \in C,\ \alpha_1, ..., \alpha_n \geq 0$$

$$\text{s.t. } \alpha_1 + ... + \alpha_n = 1,$$

$$v = \sum_{k=1}^{n} \alpha_k v_k,$$

$\alpha_i, \alpha_j > 0$ for some i and $j, i \neq j$ and either $v \neq v_i$ or $v \neq v_j$.

Assume that $v \neq v_i$.

Write v as

$$v = \alpha_i v_i + \sum_{k \neq i, 1 \leq k \leq n} \alpha_k v_k.$$

Since $\alpha_j > 0$, then

$$\alpha_i = 1 - \sum_{k \neq i, 1 \leq k \leq n} \alpha_k \leq 1 - \alpha_j < 1,$$

so we get that $1 - \alpha_i \neq 0$.

Then

$$v = \alpha_i v_i + (1 - \alpha_i) w$$

where

$$w = \frac{1}{1 - \alpha_i} \sum_{k \neq i, 1 \leq k \leq n} \alpha_k v_k.$$

Assume now that $\alpha_i \leq 1 - \alpha_i$.

This implies that $1 - 2\alpha_i \geq 0$. Also $2\alpha_i \geq 0$ as $\alpha_i \in (0, 1)$.

Set

$$\mu = 2\alpha_i v_i + (1 - 2\alpha_i) w.$$

Then $\mu \in [v_i, w]$.

Also we have

$$\frac{1}{2}(w+\mu) = \frac{1}{2}w + \alpha_i v_i + (\frac{1}{2} - \alpha_i)w = \alpha_i v_i + (1-\alpha_i)w = v.$$

Now we wish to show that $w \neq \mu$. Assume first that $w = \mu$.

Then

$$2\alpha_i w = w - (1-2\alpha_i)w = \mu - (1-2\alpha_i)w = 2\alpha_i v_i.$$

Since $\alpha_i > 0$, we must have $w = v_i$.

But then $v = \alpha_i v_i + (1-\alpha_i)w = \alpha_i v_i + (1-\alpha_i)v_i = v_i$.

This contradicts the assumption in the begining that $v \neq v_i$. Hence we must have $w \neq \mu$,, as we wanted to show.

Next, we show that $w \in C$:
We have

$$w = \frac{1}{1-\alpha_i} \sum_{k\neq i, 1\leq k\leq n} \alpha_k v_k.$$

Clearly $\frac{\alpha_k}{1-\alpha_i} \geq 0$ for all $k \in \{1, ..., n\}$.

Furthermore

$$\sum_{k=1}^{n} \alpha_k = 1 \text{ gives that } 1-\alpha_i = \sum_{k\neq i, 1\leq k\leq n} \alpha_k.$$

Thus

$$\frac{1}{1-\alpha_i} \sum_{k\neq i, 1\leq k\leq n} \alpha_k = 1.$$

So w is a convex combination of elements in C, hence $w \in C$ since C is convex. But then $\mu \in C$, as $v_i, w \in C$ and $[\mu \in v_i, w]$. So we have shown that $v = \frac{1}{2}(w+\mu)$ where $\mu, w \in C$ and $\mu \neq w$ as desired.

Now, if $\alpha_i \geq 1-\alpha_i$ then $1-2\beta_i \geq 0$ where $\beta_i = 1-\alpha_i$.
Let $\tilde{\mu} = 2\beta_i w + (1-2\beta_i)v_i$. Then $\tilde{\mu} \in [v_i, w] \subseteq C$.
Furthermore

$$v = \frac{1}{2}(v_i + \tilde{\mu}).$$

If $\mu = v_i$, it would follow that $w = v_i$ since $\beta_i > 0$. But then $v_i = v$ which is a contradiction. Hence $\mu \neq v_i$, $\tilde{\mu}, v_i \in C$ and $v = \frac{1}{2}(\mu + v_i)$. This completes the proof.

Remark 4.8 *In the case of the standard positive cone ($X = \mathbb{R}, K = \mathbb{R}_+$ and $e = \vec{1}$) the set of extreme points is the set of standard basis vectors $\{e_i\}_i = 1, ..., n$. The extreme points are pairwise disjoint.*

Proof: Let $\mu \in P(\overrightarrow{1})$ and write

$$\mu = \sum_{i=1}^{n} \mu_i e_i$$

If there is $k, j \in \{1, ..., n\}$, s.t. $\mu_k > 0$ and $\mu_j > 0$, then

$$\mu = \mu_j e_j + \sum_{i=1, i\neq j}^{n} \mu_i e_i = \mu_j e_j + (1-\mu_j) \sum_{i=1, i\neq j}^{n} \frac{\mu_i}{1-\mu_j} e_j.$$

(Observe that since $\mu \in P(\overrightarrow{1})$, then $1 = \sum_{i=1}^{n} \mu_i$ and $\mu_i \geq 0 \quad$ for all i. Hence

$$1 - \mu_j = \sum_{i=1, i\neq j}^{n} \mu_i \geq \mu_k > 0,$$

so we can divide with $1 - \mu_j$.)

Furthermore

$$\sum_{i=1, i\neq j}^{n} \frac{\mu_i}{1-\mu_j} = \frac{1}{(1-\mu_j)} \sum_{i=1, i\neq j}^{n} \mu_i = \frac{1}{1-\mu_j}(1-\mu_j) = 1$$

Hence, since obviously $e_i \in P(\overrightarrow{1})$ for all $i \in \{1, ..., n\}$ and $P(\overrightarrow{1})$ is convex, we get that

$$v = \sum_{i=1, i\neq j}^{n} \frac{\mu_i}{1-\mu_j} e_i \in P(\overrightarrow{1}),$$

as v is a convex combination of elements in $P(\overrightarrow{1})$. Since $\mu = \mu_j e_j + (1-\mu_j)v$, $\mu_j, (1-\mu_j) > 0$ and $e_j, v \in P(\overrightarrow{1})$, we get that $\mu \notin \text{extr}P(\overrightarrow{1})$.

On the other hand, given $e_i \in P(\overrightarrow{1})$, if $e_i = \lambda\mu + (1-\lambda)v$ for some $\mu, v \in P(\overrightarrow{1})$ and some $\lambda \in (0,1)$, then $\lambda\mu_i + (1-\lambda)v_i = 1$.

Since $u, v \in P(\overrightarrow{1})$, we have that $\mu_i, v_i \in [0,1]$. Hence, since $\lambda \in (0,1)$, $\mu, v_i \in [0,1]$ and $\lambda\mu_i + (1-\lambda)v_i = 1$, we deduce that $\mu_i = v_i = 1$.

But, since $\mu, v \in P(\overrightarrow{1})$, then

$$\sum_{j=1}^{n} \mu_j = \sum_{j=1}^{n} v_j = 1$$

and $\mu_j, v_j \geq 0$ for all j, $1 \leq j \leq n$. As $\mu_i = v_i = 1$, we then get that $\mu_j = v_j = 0$ whenever $j \in \{1, ..., n\}$ with $j \neq i$. Thus $\mu = v = e_i$, so $e_i \in \text{extr}P(\overrightarrow{1})$. We conclude that

$$\text{extr}(P(\overrightarrow{1})) = \{e_j \mid 1 \leq j \leq n\}.$$

Let now $k, j \in \{1, ..., n\}$ s.t $k \neq j$.
If $\mu \in P(\overrightarrow{1})$ and $\mu \geq \frac{e_j}{2}$, $\mu \geq \frac{e_k}{2}$, that is $\mu_j \geq \frac{1}{2}$ and $\mu_k \geq \frac{1}{2}$, then we must have $\mu_j = \mu_k = \frac{1}{2}$, and $\mu = 0$ whenever $i \notin \{j, k\}$. Because, if not, then since $\mu_i \geq 0$ for all $i \in \{1, ...n\}$, (as $\mu \in P(\overrightarrow{1}) \subseteq K = \mathbb{R}^n_+$), we get that

$$\sum_{i=1}^{n} \mu_i > \mu_k + \mu_j \geq \frac{1}{2} + \frac{1}{2} = 1.$$

That is a contradiction since $\mu \in P(\overrightarrow{1})$. Thus $\mu_j = \mu_k = \frac{1}{2}$, $\mu_i = 0$ whenever $i \notin \{j, k\}$, which gives that $\mu = \frac{e_j}{2} + \frac{e_i}{2}$. This shows that $e_i^{\prime}s$ are pairwise disjoint and completes the proof. The remark 4.8 is also given in [GQ] but without proof.

Remark 4.9 *The set of extreme points in* $P(I_n)$ *is* $\{xx^* \mid x \in \mathbb{C}^n , x^*x = 1\}$

Furthermore, two extreme points xx^* *and* yy^* *are disjoint if and only if* $x^*y = 0$.

Proof: Let $x \in \mathbb{C}^n$.

Assume that $x^*x = 1$ and construct then an orthonormal basis for $\mathbb{C}^n$ that contains x . Denote this basis by β and write $\beta = \{x, v_1, ..., v_{n-1}\}$. Then $(xx^*)x = x$ and $(xx^*)v_j = 0$ for all j with $1 \leq j \leq n-1$.
Assume now that there exist $A, B \in P(I_n)$ and $\alpha \in (0, 1)$ s.t.

$$\alpha A + (1 - \alpha)B = xx^*.$$

Since $A, B \in P(I_n)$, then $A, B \in S_n^+$, so we must have $< Ay, y > \geq 0$ and $< By, y > \geq 0$ for all $y \in \mathbb{C}^n$.

Since

$$0 =< (xx^*)v_j, v_j > = < \alpha A + (1 - \alpha)B)v_j, v_j >$$
$$= \alpha < Av_j, v_j > +(1 - \alpha) < Bv_j, v_j >, \quad \forall j \in \{1, ..., n-1\},$$

we must have $< Av_j, v_j > = < Bv_j, v_j > = 0$ for all $j \in \{1, ..., k-1\}$, as α and $1 - \alpha$ are strictly positive. Hence both for A and for B, all $v_j's$ are the eigenvectors with the corresponding eigenvalue 0. Inded, since $A \geq 0$ and $B \geq 0$, there exist $A^{\frac{1}{2}}$, $B^{\frac{1}{2}}$. Hence

$$||A^{\frac{1}{2}} v_j||^2 \ = < Av_j, v_j > = \ 0,$$

so $A^{\frac{1}{2}} v_j = 0$ for all $j \in \{1, ..., k-1\}$.
Thus

$$Av_j = A^{\frac{1}{2}}(A^{\frac{1}{2}} v_j) = 0$$

for all $j \in \{1, ..., k-1\}$. The same argument applies for B.

Thus we have that $\{x, v_1, ..., v_{n-1}\}$ is an orthonormal basis consisting of eigenvectors both for A and for B. Since both A and B are in $P(I_n)$, we must have trace(A) = trace$(B) = 1$. So the sum of eigenvalues for A and the sum of the eigenvalues for B must be equal to 1. Since 0 is the corresponding eigenvalue for all $v_j's$, both for A and for B, then the corresponding eigenvalue for x must be 1 (both for A and for B).
But this means that $A = B = xx^*$ and this shows that xx^* is an extreme point of $P(I_n)$. Thus $\{xx^* | x \in \mathbb{C}^n, x^* x = 1\} \subseteq (\mathrm{P(I_n)})$ since x was arbitrary.

Let now $M \in P(I_n)$.

If

$$M = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^*$$

where $\{\mu_1, ..., \mu_n\}$ is an orthonormal basis for $\mathbb{C}^n$, then $M = \mu_1 \mu_1^*$. If not, then using that $M \in S_n^+$ and that $\mathrm{tr}(M) = 1$ we see that we can write M as

$$M = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^*$$

where at least λ_1 and λ_2 are strictly greater than 0 and $\sum_{k=1}^{n} \lambda_k = 1$

(and $\lambda_k \geq 0$ for $k \geq 3$).

Let

$$A = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^*$$

and

$$B = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\lambda_2}{1-\lambda_1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_n}{1-\lambda_1} \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^*$$

(this is well defined since

$$\lambda_1 = 1 - \sum_{k=2}^{n} \lambda_k \leq 1 - \lambda_2 < 1$$

since $\lambda_2 > 0$, so $1 - \lambda_1 \neq 0$).

Then

$$\sum_{k=2}^{n} \frac{\lambda_k}{1-\lambda_1} = \frac{1}{1-\lambda_1} \sum_{k=2}^{n} \lambda_k = \frac{1}{1-\lambda_1}(1-\lambda_1) = 1,$$

so $A, B \in P(I_n)$ and $A \neq B$ as $A\mu =_1 = \mu_1,\ B\mu_1 = 0$ and $\mu \neq 0$.

Furthermore,

$$\lambda_1 A + (1-\lambda_1)B$$

$$= \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \left(\lambda_1 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + (1-\lambda_1) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\lambda_2}{1-\lambda_1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_n}{1-\lambda_1} \end{bmatrix} \right) \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^*$$

$$= \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^* = M.$$

Hence M is a convex combination of A and B which are the elements of $P(I_n)$ and $A \neq B$, so $M \notin \text{extr}P(I_n)$.

This shows that $\text{extr}P(I_n) = \{xx^* \mid x \in \mathbb{C}^n \ \ x^*x = 1\}$.

Thus we have proved the first part of remark 4.9 . This proof is omitted in [GQ] . Now we will prove the second part of remark 4.9 . Here we will mainly follow the proof given in [GQ], but we will supply most of the statements used in this proof with further, detailed explanations.

Suppose first that $x^*y = 0,\ x^*x = y^*y = 1$. Then we have the following result: "If $X \geq xx^*$, $X \geq yy^*$, $x^*x = y^*y = 1$ and $\text{tr}(X) = 2$, then necessarily $X = xx^* + yy^*$." . The proof of this result and some comments about it are given in "comments" at the end of the proof of remark 4.9. Now, if $X \in P(I_n)$, then $\text{tr}(X) = 1$. If, in addition, $X \geq \frac{1}{2}xx^*$ and $X \geq \frac{1}{2}yy^*$, then the result given above implies that $X = \frac{1}{2}(xx^* + yy^*)$. Hence xx^* and yy^* are disjoint by definition .

Suppose now that xx^* and yy^* are disjoint extreme point of $P(e)$. Observe that this implies that x and y are linearly independent. Because, if not, then there is some $\alpha \in \mathbb{C}$ s.t. $x = \alpha y$. Since xx^* and yy^* are assumed to be extreme points of $P(e)$, then we must have $x^*x = 1$ and $y^*y = 1$ by the first statement of the remark 4.9 which we already proved . But then

$$1 = x^*x = |\alpha|^2 y^*y = |\alpha|^2,$$

so $1 = |\alpha|^2$.
Hence

$$xx^* = |\alpha|^2 yy^* = yy^*.$$

As $xx^* = yy^*$, they can clearly not be disjoint, so we get a contradiction. Thus we must have that x and y are linearly independent, so

$$\dim(\text{Span}\{x, y\}) = 2.$$

Let $W = \dim(\text{Span}\{x, y\})$.
Then $\mathbb{C}^n = W \oplus W^\perp$ and if $v \in W^\perp$, then v is an eigenvector of the matrix $xx^* - yy^*$ with corresponding eigenvalue 0. Hence we can find an orthonormal basis consisting of eigenvectors of $xx^* - yy^* = \{\mu, v.w_1, ..., w_{n-2}\}$ where $w_j \in W^\perp$ for all $j \in \{1, ..., n-2\}$ Since this basis is orthonormal, it follows then that

$$u, v \in (W^\perp)^\perp = W = \text{Span}\{x, y\}.$$

(here we use that $\dim W = 2$.)

Then $xx^* - yy^*$

$$= \begin{bmatrix} \mu & v & w_1 \cdots & w_{n-2} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \cdots & 0 & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu & v & w_1 \cdots & w_{n-2} \end{bmatrix}^*.$$

Next since

$$\mathrm{Tr}(xx^* - yy^*) = \mathrm{Tr}(xx^*) - \mathrm{Tr}(yy^*) = 1 - 1 = 0,$$

it follows that $\lambda_1 + \lambda_2 = 0$, so $\lambda_2 = -\lambda_1$ Hence

$$xx^* - yy^* = \lambda_1(\mu\mu^* - vv^*)$$

(Here we choose $\lambda_2 \leq 0$ so that $\lambda_1 \geq 0$).

Now we want to show that $\lambda_1 \leq 1$.

Let

$$x = c_1\mu + c_2 v,$$

$$y = d_1\mu + d_2 v.$$

Since $||x|| = ||y|| = 1$ (because $xx^* = yy^* = 1$,) by Pythagoras',

$$|c_1|^2 + |c_2|^2 = 1,$$

$$|d_1|^2 + |d_2|^2 = 1.$$

We also have:

$$(xx^* - yy^*)\mu = xx^*\mu - yy^*\mu$$

$$= [(c_1\mu + c_2 v)(\overline{c}_1\mu^* + \overline{c}_2 v^*) - (d_1\mu + d_2 v)(\overline{d}_1\mu^* + \overline{d}_2 v^*)]\mu$$

$$= [(|c_1|^2\mu\mu^* + c_1\overline{c}_2\mu v^* + c_2\overline{c}_1 v\mu^* + |c_2|^2 vv^*) - (|d_1|^2\mu\mu^* + d_1\overline{d}_2\mu v^* + d_2\overline{d}_1 v\mu^* + |d_2|^2 vv^*)]\mu$$

$$= |c_1|^2\mu + c_2\overline{c_1}v - |d_1|^2\mu - d_2\overline{d_1}v$$

since $\mu^*\mu = v^*v = 1$ and $v^*\mu = 0$.

As $(xx^* - yy^*)\mu = \lambda_1\mu$ because $xx^* - yy^* = \lambda_1(\mu\mu^* - vv^*)$ and $v * \mu = 0$, we get that

$$(|c_1|^2 - |d_1|^2)\mu + (c_2\overline{c_1} - d_2\overline{d_1})v = \lambda_1\mu$$

Taking inner product on both sides with μ and again using that $\mu^*\mu = 1$ and $v^*\mu = 0$, we get that $|c_1|^2 - |d_1|^2 = \lambda_1$. Since $|c_1|^2, |d_1|^2 \in [0,1]$, $\lambda_1 \geq 0$ and $\lambda_1 = |c_1|^2 - |d_1|^2$, we get that $\lambda_1 \in [0,1]$ so $\lambda_1 \leq 1$ as we wanted to show. Furthermore:

$$xx^* - yy^* = \lambda_1(\mu\mu^* - vv^*) = \mu\mu^* - (1-\lambda_1)\mu\mu^* - \lambda_1 vv^*$$

$$= \mu\mu^* - (1-\lambda_1)\mu\mu^* + \lambda_1 vv^*.$$

Since $\mu\mu^*,\ vv^* \in P(I_n)$, $P(I_n)$ is convex and $\lambda_1 \in [0,1]$, we have that $(1-\lambda_1)\mu\mu^* + \lambda_1 vv^* \in P(I_n)$. Lemma 4.6, gives then that

$$xx^* = \mu\mu^*, \quad yy^* = (1-\lambda_1)\mu\mu^* + \lambda_1 vv^*$$

since xx^* and yy^* are disjoint by assumption.
Then $(xx^*)\mu = \mu\mu^*\mu = \mu$, so μ is the eigenvector of xx^* with the corresponding eigenvalue 1. Hence $\mu = \alpha_1 x$ where $|\alpha_1| = 1$ since $E_1^{xx^*} = \text{Span}\{x\}$ and $||x|| = ||\mu|| = 1$. Since $yy^* = (1-\lambda_1)uu^* + \lambda_1 vv^*$, we get that $(yy^*)v = \lambda_1 v$, so v is the eigenvector of yy^* with the eigenvalue λ_1. Since $\lambda_1 > 0$, then λ_1 must be equal to 1 as 0 and 1 are the only eigenvalues of yy^*, and $v = \alpha_2 y$ where $|\alpha_2| = 1$.
Hence

$$x^*y = \frac{1}{\overline{\alpha_1}\alpha_2}\mu v = 0.$$

This complets the proof of the second statement in remark 4.9 .

Comments
We have the following observation:
Observation 4

If $X \geq xx^*,\ X \geq yy^*, x, y$ are unitary vectors, $x^*y = 0$, and trace(X) $= 2$ then necessarily

$$X = xx^* + yy^*.$$

Proof W.l.o.g let us assume that $x = e_1$ and $y = e_2$. Let

$$A = X - e_1e_1^* - e_2e_2^*.$$

Then

$$\text{trace}(A) = 0,$$

and

$$A + e_1e_1^* \geq 0,$$
$$A + e_2e_2^* \geq 0.$$

The only difference between A and $A + e_1e_1^*$ is on the first diagonal element. We then deduce from (1) that the second to the last diagonal elements of A must be nonnegative. Similarly, it follows from (2) that the first to the penultimate diagonal elements are nonnegative. Hence, all the diagonal elements are nonnegative, and since trace$(A) = 0$, all these diagonal elements are zero.

Now, note that if B is a positive semidefinite matrice, all 2×2 principal submatrices of B are positive semidefinite, and so $B_{ij}^2 \leq B_{ij}B_{ij}$, for all $i \neq j$. By applying this to the matrix $B := A + e_1 e_1^* \geq 0$, we deduce that $A_{ij}^2 = B_{ij}^2 \leq B_{ii}Bjj = 0$ for all $i \neq j$, and so, A is the zero matrix.

This observation and proof were given by Stephane Gaubert and Zheng Qu.

The next remark is not given in [GQ]. **Remark 4.10** We consider now the $C_{\mathbb{R}}(\Omega)$ where Ω is a compact, Hausdorff topological space Recall that $(C_{\mathbb{R}}(\Omega), || \cdot ||_\infty)^*$ is isometrically isomorphic to $M_r(\Omega)$ where $M_r(\Omega)$ is the space of all signed Radon measures on Ω. The simplex here is

$$P(\Omega) = \{\mu \in M_+(\Omega) \ : \ \mu(\Omega) = 1\}$$

By exercise 14.86 on page 547 in [MW], we have that

$$P(\Omega) = \{\delta_x : \quad x \in \Omega\}.$$

It is clear that $\delta_x's$ are pairwise disjoint, since $K^* = M_+(\Omega)$ in this case.

6 The operator norm induced by Hilbert's quotient norm

In this section we consider two real Banach spaces

$$(X_1, ||\cdot||) \text{ and } (X_2, ||\cdot||).$$

We let $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ be respectively two closed, normal cones with nonempty interiors K_1° and K_2°. Furthermore we let $e_1 \in K_1^\circ$ and $e_2 \in K_2^\circ$ be order units. From section 4 lemma 4.2 we know that the duals spaces of the quotient spaces $(X_1/\mathbb{R}e_1, |||\cdot|||_H)$ and $(X_2/\mathbb{R}e_2, |||\cdot|||_H)$ are isometrically isomorphic to $(M(e_1), ||\cdot||_H^*)$ and $(M(e_2), ||\cdot||_H^*)$ respectively.

We will state and prove here one of the 2 main theorems in this thesis, the theorem 5.1 [GQ] However it was not written in the correct way in [GQ] so it had to be reformulated. In [GQ] they let $T : X_1 \to X_2$ be a bounded, linear map satisfying $T(e_1) \in \mathbb{R}e_2$ and they consider "$||T||_H$", where $||\cdot||_H$, is a seminorm on X_1 defined in section 1. It is not correct to defined an operator norm w.r.t. seminorm, so this had to be reformulated. We define instead the induced linear map

$$\tilde{T} : X/\mathbb{R}e_1 \to X_2/\mathbb{R}e_2$$

by

$$\tilde{T}(x + \mathbb{R}e_1) = T(x) + \mathbb{R}e_2.$$

Then we show that $\tilde{T}$ is a well defined, bounded linear map w.r.t. $|||\cdot|||_H$. Hence we may consider $|||\tilde{T}|||_H$, that is the operator norm of $\tilde{T}$ w.r.t. $|||\cdot|||_H$. We also define $S^* : M(e_2) \to M(e_1)$ by letting $S^* = T_{|M(e_2)}$ and we show that $||S^*||_H^* = |||\tilde{T}|||_H$. First after introducing all these concepts, definitions and relations between them, we finally state and prove the reformulated version of the theorem 5.1.
This theorem gives then an expression for $|||\tilde{T}|||_H$ in terms of disjoint, extreme points of $P(e_2)$. In the proof of this theorem we will mainly follow the proof given in [GQ], but we will supply most of the statements used in [GQ] with detailed proofs and explanations.
At the end this section, we will introduce the remark 5.2 which states that $[0, e_1]$is the closed, convex hull of the set of its extreme points, when X_1 of finite dimension. This remark is also given in [GQ] but without proof. However, here we give a complete proof of this remark. The remark 5.2 will be

used later in section 7 and 9 which deal with applications of the theory from first 6 sections.

Let $T : X_1 \to X_2$ be a bounded, linear map satisfying $T(e_1) \in \mathbb{R}e_2$.

We define the induced map

$$\tilde{T} : (X_1/\mathbb{R}e_1, |||\cdot|||_H) \to (X_2/\mathbb{R}e_2, |||\cdot|||_H) \text{ by}$$

$$\tilde{T}(x + \mathbb{R}e_1) = T(x) + \mathbb{R}e_2$$

Then $\tilde{T}$ is well defined:

Assume that $x + \mathbb{R}e_1 = y + \mathbb{R}e_1$, for some $x, y \in X_1$.

Then $x = y + ke_1$, for some $k \in \mathbb{R}$. Furthermore $T(e_1) = ce_2$ for some $c \in \mathbb{R}$.

Hence

$$T(x) = T(y) + kT(e_1) = T(y) + kce_2.$$

Then

$$T(x) + \mathbb{R}e_2 = (T(y) + kce_2 + \mathbb{R}e_2) = T(y) + \mathbb{R}e_2,$$

so $\tilde{T}$ is well defined.

Furthermore $\tilde{T}$ is bounded:

Assume that

$$|||x + \mathbb{R}e_1|||_H = 1.$$

This means that $\inf_{\lambda \in \mathbb{R}} ||x + \lambda e_1||_T = \dfrac{1}{2}$

Hence, there exists a sequence and $\{\lambda_n\}_n \subset \mathbb{R}$ s.t.

$$\tfrac{1}{2} \leq ||x + \lambda_n e_1||_T \leq \tfrac{1}{2} + \tfrac{1}{n} \quad \forall n \in \mathbb{N}.$$

We have

$$\tilde{T}(x + \mathbb{R}e_1) = T(x) + \mathbb{R}e_2.$$

But since $T(e_1) \in \mathbb{R}e_2$, we get that

$$\tilde{T}(x+\mathbb{R}e_1) = T(x)+\mathbb{R}e_2 = T(x)+\lambda_n T(e_1)+\mathbb{R}e_2 = T(x+\lambda_n e_1)+\mathbb{R}e_2 \ \forall n \in \mathbb{N}.$$

Thus

$$|||\tilde{T}(x+\mathbb{R}e_1)|||_H = |||T(x+\lambda_n e_1) + +\mathbb{R}e_2|||_H$$

$$= 2\inf_{\eta\in\mathbb{R}} ||T(x+\lambda_n e_1)+\eta e_2||_T \leq 2||T(x+\lambda_n e_1)||_T \quad \forall n.$$

Hence, for all n, we have:

$$|||\tilde{T}(x+\mathbb{R}e_1)|||_H \leq 2||T(x+\lambda_n e_1)||_T \leq 2||x+\lambda_n e_1||_T \ ||T||_T$$

Since

$$\lim_{n\to\infty} ||x+\lambda_n e_1||_T = \frac{1}{2},$$

we get that $|||\tilde{T}(x+\mathbb{R}e_1)|||_H \leq 2\frac{1}{2}||T||_T = ||T||_T$ so $\tilde{T}$ is bounded.

Also $\tilde{T}$ is obviously linear, since T is linear.

We can then define the map

$$(\tilde{T})^* : ((X_2/\mathbb{R}e_2)^*, |||\cdot|||_H^*) \to ((X_1/\mathbb{R}e_1)^*, |||\cdot|||_H^*) \text{ given by}$$

$$< \tilde{T}^*(l), x+\mathbb{R}e_1 > = < l, \tilde{T}(x+\mathbb{R}e_1) > \quad \forall l \in (X_2/\mathbb{R}e_2)^* \ \ \forall x \in X_1$$

From [P] (2.3.20.),we have

$$|||(\tilde{T})^*|||_H^* = |||\tilde{T}|||_H.$$

From lemma 4.2, we have that the maps

$$\phi_1 : (M(e_1), \tfrac{1}{2}||\cdot||_T^*) \to ((X_1/\mathbb{R}e_1)^*, |||\cdot|||_H^*) \text{ and}$$

$$\phi_2 : (M(e_2), \tfrac{1}{2}||\cdot||_T^*) \to ((X_2/\mathbb{R}e_2)^*, |||\cdot|||_H^*) \text{ given by}$$

$$< \phi_1(v), x+\mathbb{R}e_1 > = < v, x > \quad \forall v \in M(e_1) \ \ x \in X_1 \text{ and}$$

$$< \phi_2(\mu), y+\mathbb{R}e_2 > = < \mu, y > \quad \forall \mu \in M(e_2) \ \ y \in X_2$$

are isometric isomorphisms.

We then define the map

$$S^* : (M(e_2), ||\cdot||_H^*) \to (M(e_1), ||\cdot||_H^*) \text{ by } S^* = \phi_1^{-1} \circ (\tilde{T})^* \circ \phi_2.$$

So we have the following diagram:

$$\begin{array}{ccc}((X_1/\mathbb{R}e_1)^*, |||\cdot|||_H^*) & \overset{\longleftarrow}{(\tilde{T})^*} & ((X_2/\mathbb{R}e_2)^*, |||\cdot|||_H^*) \\ \uparrow \phi_1 & & \uparrow \phi_2 \\ (M(e_1), ||\cdot||_H^*) & \overset{\longleftarrow}{S^*} & (M(e_2), ||\cdot||_H^*)\end{array}$$

Claim: *We have*

$$< S^*(\mu), x > = < \mu, Tx > = < T^*(\mu), x > \ \textit{for all } \mu \in M(e_2) \textit{ and } x \in X,$$

that is

$$S^* = T^*_{|_{M(e_2)}} :$$

Proof:

$$\begin{aligned} < S^*(\mu), x > &= < (\phi_1^{-1}(\tilde{T})^*\phi_2)(\mu), x > = < ((\tilde{T})^*\phi_2)(\mu), x + \mathbb{R}(e_1) > \\ &= < \phi_2(\mu), \tilde{T}(x + \mathbb{R}e_1) > = < \phi_2(\mu), Tx + \mathbb{R}(e_2) \\ &= < \mu, Tx > = < T^*(\mu), x > \blacksquare \end{aligned}$$

Moreover we have

$||S^*||_H^* = |||(\tilde{T})^*|||_H^*$ since ϕ_1 and ϕ_2 are isometric isomorphisms.

Since

$|||(\tilde{T})^*|||_H^* = |||\tilde{T}|||_H$, we get that $||S^*||_H^* = |||\tilde{T}|||_H$.

Now we are ready to state and prove the reformulated version of the theorem 5.1.

Theorem 5.1 *Let $T : X_1 \to X_2$ be a bounded, linear map s.t.*

$$T(e_1) \in \mathbb{R}e_2$$

Then

$$\begin{aligned} |||\tilde{T}|||_H = ||S^*||_H^* &= \frac{1}{2} \sup_{v,\pi \in P(e_2)} ||T^*(v) - T^*(\pi)||_T^* \\ &= \sup_{v,\pi \in P(e_2)} \sup_{x \in [o,e_1]} < v - \pi, T(x) > \end{aligned}$$

Moreover, the supremum can be restricted to the set of mutually disjoint extreme points of $P(e_2)$:

$$|||\tilde{T}|||_H = ||S^*||_H^* = \frac{1}{2} \sup_{v,\pi \in extP(e_2),\ v \perp \pi} ||T^*(v) - T^*(\pi)||_T^*$$

$$= \sup_{v,\pi \in extP(e_2), v\perp\pi} \ \sup_{x \in [o,e_1]} < v - \pi, T(x) >$$

Proof We have already proved that $|||\tilde{T}|||_H = ||S^*||_H^*$.

Moreover
$$||S^*||_H^* = \sup_{\mu \in B_H^*(e_2)} ||S^*(\mu)||_H^*.$$

By the characterisation of $B_H^*(e_2)$ obtained earlier, that is

$B_H^*(e_2) = \{v - \pi : v, \pi \in P(e_2)\}$, we get that

$$\sup_{\mu \in B_H^*(e2)} ||S^*(\mu)||_H^* = \sup_{v,\pi \in P(e_2)} ||S^*(v - \pi)||_H^*.$$

Using that $S^* = T_{|_{M(e_2)}}$ and that T^* is linear, we obtain

$$\sup_{\mu \in B_H^*(e2)} ||S^*(\mu)||_H^* = \sup_{v,\pi \in P(e_2)} ||S^*(v - \pi)||_H^*$$

$$= \sup_{v,\pi \in P(e_2)} ||T^*(v - \pi)||_H^* = \sup_{v,\pi \in P(e_2)} ||T^*(v) - T^*(\pi))||_H^*$$

$$= \frac{1}{2} \sup_{v,\pi \in P(e_2)} ||T^*(v) - T^*(\pi))||_T^*$$

since $|| \cdot ||_H^* = \frac{1}{2} || \cdot ||_T^*$ by definition.

Now we wish to show that

$$\frac{1}{2} \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} ||T^*(v) - T^*(\pi)||_T^*$$

$$= \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} \ \sup_{x \in [0,e_1]} < v - \pi, T(x) >$$

Thus we have to show that

$$||T^*(v) - T^*(\pi)||_T^* = 2 \sup_{x \in [0,e_1]} < v - \pi, T(x) >$$

for all $v, \pi \in P(e_2)$.
We will show this by proving that $B_T(e_1)$ can be written as

$$B_T(e_1) = 2[0, e_1] - e_1.$$

Then it would follow that whenever $v, \pi \in P(e_2)$, we have

$$||T^*(v) - T^*(\pi)||_T^* = \sup_{||x||_T \leq 1} | < T^*(v) - T^*(\pi), x > |$$

$$= \sup_{\tilde{x} \in [0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} - e_1 > |.$$

When we have established this equality, we can proceed further by observing that if $v \in P(e_2)$, then

$$< T^*(v), e_1 > = < v, T(e_1) > = < v, ke_2 > = k < v, e_2 > = k$$

for some $k \in \mathbb{R}$, since $T(e_1) \in \mathbb{R}e_2$ and $< v, e_2 > = 1$ as $v \in P(e_2)$. Similarly, if $\pi \in P(e_2)$, we get that $< T^*(\pi), e_1 > = k$. Hence

$$< T^*(v) - T^*(\pi), e_1 > = 0.$$

Thus

$$\sup_{\tilde{x} \in [0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} - e_1 > | = \sup_{\tilde{x} \in [0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} > |$$

$$= 2 \sup_{\tilde{x} \in [0,e_1]} | < T^*(v) - T^*(\pi), \tilde{x} > |$$

whenever $v, \pi \in P(e_2)$.
Hence, if we can show that

$$\sup_{||x||_T \leq 1} | < T^*(v) - T^*(\pi), x > |$$

$$= \sup_{\tilde{x} \in [0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} - e_1 > |,$$

then we would get

$$||T^*(v) - T^*(\pi)||_T^* = \sup_{||x||_T \leq 1} | < T^*(v) - T^*(\pi), x > |$$

$$= \sup_{\tilde{x}\in[0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} - e_1 > |$$

$$= 2 \sup_{\tilde{x}\in[0,e_1]} | < T^*(v) - T^*(\pi), \tilde{x}|$$

whenever $v, \pi \in P(e_2)$.
Now, in order to prove that

$$\sup_{||x||_T\leq 1} | < T^*(v) - T^*(\pi), x > | = \sup_{\tilde{x}\in[0,e_1]} | < T^*(v) - T^*(\pi), 2\tilde{x} - e_1 > |$$

it suffices to show that $B_T(e_1) = 2[0, e_1] - e_1$, as mentioned before.

Claim: $B_T(e_2) = 2[0, e_1] - e_1$

Proof: Assume that $x \in B_T(e_1)$.
If $||x||_T = 1$,

then $\inf\{t > 0 | x \in tI_{e_1}\} = 1$, as

$$||x||_T = \inf\{t > 0 \mid x \in tI_{e_1}\}.$$

Since

$$||x||_T = 1, \text{ for all } n \text{ there exists a } t_n \in [1, 1 + \tfrac{1}{n}]$$

$$\text{s.t. } x \in t_n I_{e_1} \text{ which means that}$$

$$(e_1 - \tfrac{1}{t_n}x) \in K \text{ and}$$

$$(e_1 + \tfrac{1}{t_n}x) \in K.$$

Since K is closed, we get that

$$\lim_{n\to\infty} (e_1 + \frac{1}{t_n}x) = (e_1 + x) \in K \text{ and}$$

$$\lim_{n\to\infty} (e_1 - \frac{1}{t_n}x) = (e_1 - x) \in K.$$

Hence

$$\tfrac{1}{2}(e_1 + x) \in K$$

$$\text{and } e_1 - \tfrac{1}{2}(e_1 + x) = \tfrac{1}{2}(e_1 - x) \in K$$

Thus $\frac{1}{2}(e_1 + x) \in [0, e_1]$.

If $||x||_T < 1$, then

$$\inf\{t > 0 | x \in tI_{e_1}\} < 1.$$

Hence, by prop:14.8 part c) in [MW] we have that $x \in I_{e_1}$.

Thus $(e_1 - x) \in K$ and $(e_1 + x) \in K$ that is $\frac{1}{2}(e_1 - x)$ and $\frac{1}{2}(e_1 + x)$ are in K.

Then we can use the similar arguments as above to deduce that

$$\frac{1}{2}(x + e_1) \in [0, e_1].$$

Since $x \in B_T(e_1)$ was arbitrary, we get $\frac{1}{2}B_T(e_1) + e_1 \subseteq [0, e_1]$, or equivalently $B_T(e_1) \subseteq 2[0, e_1] - e_1$.

Assume now that $\tilde{x} \in [0, e_1]$. Then $\tilde{x} \in K$ and $e_1 - \tilde{x} \in K$.

Hence

$$e_1 - (2\tilde{x} - e_1) = 2(e_1 - \tilde{x}) \in K \text{ and } e_1 + (2\tilde{x} - e_1) = 2\tilde{x} \in K$$

If we let $x = 2\tilde{x} - e_1$, we see then that $(e_1 + x),\ (e_1 - x) \in K$.

Hence

$$1 \geq \inf\{t > 0 | (e_1 - \tfrac{1}{t}x) \in K \text{ and } (e_1 + \tfrac{1}{t}x) \in K\} =$$

$$= \inf\{t > 0 | x \in tIe_1\} = ||x||_T.$$

Thus

$$2\tilde{x} - e_1 = x$$

is in $B_T(e_1)$. Since

$$\tilde{x} \in [0, e_1]$$

was arbitrary, we conclude that

$$2[0, e_1] - e_1 \subseteq B_T(e_1).$$

Combining these 2 inclusions, we get that

$$B_T(e_1) = 2[0, e_1] - e_1$$

and this proves the claim. Hence we have proved the first part of the theorem 5.1.

Next, we will show that the supremum can be restricted to the set of extreme points.
We show this by proving the following:
a) $M(e_2)$ is a locally convex topological space in its relative w^* topology, and $B^*_H(e_2)$ is a w^* compact subset of $M(e_2)$.
Furthermore

$$B^*_H(e_2) = \overline{\text{conv}(\text{extr}(B^*_H(e_2)))}^{w^*}$$

Hence every $\rho \in B^*_H(e_2)$ is a limit of a met

$$\{\rho_\alpha\}_{\alpha \in A} \subseteq \text{conv}(\text{extr}(B^*_H(e_2)))$$

in the w^*-topology.
b) Let $\phi : M(e_2) \to [0, \infty)$ be given by

$$\phi(\mu) = ||S^*(\mu)||^*_H$$

Then

$$\phi(\mu) = \sup_{(x+\mathbb{R}e_1) \in B_H(e_1)} | < \mu, T(x) > |$$

c) ϕ is w^* lower semicontinuous
d) If $\rho \in B^*_H(e_2)$, $\{\rho_\alpha\}_{\alpha \in A}$ is a net in $\text{conv}(\text{extr}B^*_H(e_2))$ and $\rho_\alpha \to \rho$ in w^*-topology , then $\phi(\rho) \leq \liminf_\alpha \phi(\rho_\alpha)$. (Here $\liminf_\alpha \phi(\rho_\alpha)$ denotes the limit of a net $\{\beta_\alpha\}_{\alpha \in A}$ in $\mathbb{R}^+$ given by

$$\beta_\alpha = \inf_{\alpha', \alpha \leq \alpha} \{\phi(\rho_\alpha)\} \text{)}$$

e)

$$\sup\{\phi(\rho) : \mu \in \text{conv}(\text{extr}(B^*_H(e_2)))\}$$
$$= \sup\{\phi(\rho) : \mu \in (\text{extr}(B^*_H(e_2))\}$$

f)

$$\sup_{\mu \in B^*_H(e_2)} \phi(\mu) = \sup_{\mu \in \text{extr}(B^*_H(e_2))} \phi(\mu))$$

g)

$$\sup_{\mu \in \text{extr}(B^*_H(e_2))} ||S^(\mu)||^*_H$$

$$= \frac{1}{2} \sup_{v, \pi \in (\text{extr}(P(e_2))} \sup_{x \in [0, e_1]} < v - \pi, T(x) >$$

Proof of a) By prop. 14.5 in [MW], a vector space with the topology induced by a separating family of seminorms is locally convex topological vector space. Hence $M(e_2)$ with the relative w^* - topology is a locally convex topological space (since the w^* - topology on X_2^* is induced by the family of seminorms $\{\rho_x | x \in X_2\}$ where $\rho_x(\varphi) = |\varphi(x)|$ for all $\varphi \in X_2^*$).

We have also that $B_H^*(e_2)$ is w^* - compact:
Since the convergence of a net in X_2^* in the w^* - topology is the same as the pointwise convergence of this net on X_2 and

$$M(e_2) = \{\mu \in X_2^* | < \mu, e_2 > = 0\}$$

it is obvious that $M(e_2)$ is w^* closed in X_2^*.
Now, Banach - Alaoglu's theorem gives that $B_T^*(e_2)$ is w^* - compact in X_2^*
Hence $2B_T^*(e_2)$ is also w^* -compact.

By definition, $B_H^*(e_2) = 2B_T^*(e_2) \cap M(e_2)$. Since $M(e_2)$ is w^* - closed, it follows that $B_H^*(e_2)$ is w^*- closed subset of $2B_T^*(e_2)$. Hence $B_H^*(e_2)$ is w^*- compact in $M(e_2)$.
Since $M(e_2)$ with the relative w^* - topology is LCS and $B_H^*(e_2)$ is w^*- compact and convex subset of $M(e_2)$, the Krein - Milman theorem, gives

$$B_H^*(e_2) = \overline{\text{conv}(\text{extr } B_H^*(e_2))}^{w^*}.$$

Hence every $\rho \in B_H^*(e_2))$ is a limit in w^* - topology of a net

$$\{\rho_\alpha\}_{\alpha \in A} \in \text{conv}(\text{extr } B_H^*(e_2)).$$

This proves the part a).

Proof of b) Consider now the function: $\phi : \mu \to ||S^*(\mu)|||_H^*$ from $M(e_2)$ into $[0, \infty)$.

Since the map

$$\varphi : (M(e_1), || \cdot ||_H^*) \to ((X_1/\mathbb{R}e_1)^*, ||| \cdot |||_H^*)$$

given by

$$< \varphi(v), x + \mathbb{R}e_1 > = < v, x > \ \ \forall v \in M(e_1), x \in X_1$$

is an isometry, we get that

$$||S^*(\mu)||_H^* = |||\varphi(S^*(\mu))|||_H^*$$

$$= \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} |< \varphi(S^*(\mu)), x+\mathbb{R}e_1 >|$$

$$= \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} |< S^*(\mu), x >| = \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} |< \mu, Tx >|.$$

Hence

$$\phi(\mu) = \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} |< \mu, Tx >|.$$

This proves part b).

Proof of c) Since the w^* - topology on X_2^* is the topology of pointwise convergence, the map: $\phi_x : \mu \to |< \mu, T(x) >|$ is w^* - continuous for each $x \in X_1$.

If we let

$$\phi : \mu \to \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} |< \mu, T(x) >| = \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x(\mu),$$

we see that ϕ is a supremum of a family of weak star continuous maps. We claim that this implies that ϕ is w^* lower semicontinuous.

By definition on page 410 in [MW] a function $f : \Omega \to (-\infty, \infty]$ is weak star lower semi continuous if $f^{-1}((r,\infty])$ is open in Ω for all $r \in \mathbb{R}$. Since

$$\phi(\mu) = \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x(\mu) \quad \forall \mu \in M(e_2),$$

we get that

$$\phi^{-1}((r,\infty]) = \bigcup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x^{-1}((r,\infty])$$

(Namely, if

$$\phi(\mu) = \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x(\mu) > r$$

for some $r \in \mathbb{R}$, then $\exists x \in X_1$ s.t. $(x + \mathbb{R}e_1) \in B_H(e_1)$ and $(\phi_x(\mu)) > r$.
Hence

$$\mu \in \bigcup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x^{-1}((r,\infty])$$

$$so \ \ \phi^{-1}((r,\infty]) \subseteq \bigcup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x^{-1}((r,\infty])$$

The other inclusion is trivial)

Since each ϕ_x is w^*- continuous for each $x \in X_1$, we get that $\phi_x^{-1}((r,\infty])$ is w^*- open for all x with $x + \mathbb{R}e_1 \in B_H(e_1)$, so

$$\phi^{-1}((r,\infty]) = \bigcup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \phi_x^{-1}((r,\infty])$$

is w^*- open.

Hence ϕ is w^*- lower semicontinuous, since r was arbitrary. This complets the proof of part c).

Proof of d)

Since ϕ lower semicontinuous and $\rho_\alpha \longrightarrow \rho$ in the w^* topology (where

$$\rho_\alpha \in \mathrm{conv}(\mathrm{extr}(B_H^*(e_2))$$

for all α and

$$\rho \in B_H^*(e_2) = \overline{\mathrm{conv}(\mathrm{extr}(B_H^*(e_2)))}^{w^*} \ ) \ )$$

we claim that

$$\phi(\rho) \leq \liminf_\alpha \phi(\rho_\alpha)$$

where $\liminf_\alpha \phi(\rho_\alpha)$ the limit of the net $\{\beta_\alpha\}_\alpha$ given by $\beta_\alpha = \inf_{\alpha',\alpha\leq\alpha'} \phi(\rho_{\alpha'})$: Note first that $\{\beta_\alpha\}$ is a nondecreasing net, hence the limit of this net is well defined.

Let $c \in \mathbb{R}, c < \phi(\rho)$. Then $\rho \in \phi^{-1}((c,\infty])$ and $\phi^{-1}((c,\infty])$ is weak star open (since ϕ is weak star lower semicontinoous).

Since $\rho_\alpha \to \rho$ in w^* topology, there is an $\alpha_0 \in A$ s.t. $\rho_\alpha \in \phi^{-1}((c,\infty])$ whenever $\alpha_0 \leq \alpha$, as $\phi^{-1}((c,\infty])$ is w^*-open.

This gives that $\phi(\rho_\alpha) \in (c,\infty]$ which means that $\phi(\rho_\alpha) > c$, whenever $\alpha_0 \leq \alpha$.

Hence

$$\inf_{\alpha_0 \leq \alpha} \{\phi(\rho_\alpha)\} \geq c.$$

Thus

$$\liminf_{\alpha} \phi(\rho_\alpha) \geq \inf_{\alpha_0 \leq \alpha} \{\phi(\rho_\alpha)\} \geq c.$$

Since c was arbitrary with $c < \phi(\rho)$, we get that

$$\liminf_{\alpha} \phi(\rho_\alpha) \geq \phi(\rho)$$

and this proves the part d).

Proof of e): Since $\text{extr}(B_H^*(e_2)) \subseteq \text{conv}(\text{extr}(B_H^*(e_2)))$, then clearly

$$\sup\{\phi(\mu) : \mu \in \text{extr}(B_H^*(e_2))\} \leq \sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}(B_H^*(e_2)))\}$$

so we have to prove the opposite inequality:

Let $\mu \in \text{conv}(\text{extr}(B_H^*(e_2)))$. Then

$$\exists v_1, ..., v_n \in \text{extr} B_H^*(e_2) \text{ and}$$

$$\alpha_1, ..., \alpha_n \geq 0 \ s.t. \ \sum_{k=1}^{n} \alpha_k = 1 \ \text{ and}$$

$$\sum_{k=1}^{n} \alpha_k v_k = \mu.$$

We then get that:

$$\phi(\mu) = \sup_{(x+\mathbb{R}e_1) \in B_H(e_1)} | < \mu, T(x) > | = \sup_{(x+\mathbb{R}e_1) \in B_H(e_1)} | < \sum_{k=1}^{n} \alpha_k v_k, T(x) > |$$

$$= \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \{|\sum_{k=1}^{n} \alpha_k < v_k, T(x) > |\}$$

$$\leq \sum_{k=1}^{n} \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} \{|\alpha_k < v_k, T(x) > |\}$$

$$= \sum_{k=1}^{n} \alpha_k \sup_{(x+\mathbb{R}e_1)\in B_H(e_1)} | < v_k, T(x) > | = \sum_{k=1}^{n} \alpha_k \phi(v_k)$$

Furthermore, since $v_1, ..., v_n \in \text{extr}B_H(e_2)$, we get that

$$\phi(\mu) \leq \sum_{k=1}^{n} \alpha_k \phi(v_k) \leq \sum_{k=1}^{n} \alpha_k \sup\{\phi(v) : v \in \text{extr}B_H^*(e_2)\}$$

$$= (\sup\{\phi(v) : v \in \text{extr}B_H^*(e_2)\}) \sum_{k=1}^{n} \alpha_k = \sup\{\phi(v) : v \in \text{extr}B_H^*(e_2)\}$$

Since this is true for any $\mu \in \text{conv}(\text{extr}B_H^*(e_2))$, we get that sup

$$\sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}B_H^*(e_2))\}$$

$$\leq \sup\{\phi(v) : v \in \text{extr}B_H^*(e_2)\} \ .$$

Combining these 2 inequalities together, we obtain the equality and this proves the part e)

Proof of f) Given $\rho \in B_H^*(e_2)$ by part a) there exists a net $\{\rho_\alpha\}_{\alpha\in A}$ in $\text{conv}(\text{extr}(B_H^*(e_2))$ s.t $\rho_\alpha \to \rho$ in the w^* - topology. By part d), we must then have that $\phi(\rho) \leq \liminf_\alpha \phi(\rho_\alpha)$.

Now, since

$$\rho_\alpha \in \text{conv}(\text{extr}(B_H^*(e_2))$$

for all α, then

$$\phi(\rho_\alpha) \leq \sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}(B_H^*(e_2)))\}$$

for all α.

Hence

$$\liminf_\alpha \phi(\rho_\alpha) \leq \ \sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}(B_H^*(e_2)))\}$$

Thus

$$\phi(\rho) \leq \liminf_{\alpha} \phi(\rho_\alpha)$$
$$\leq \sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}(B_H^*(e_2)))\}.$$

By part e)

$$\sup\{\phi(\mu) : \mu \in \text{conv}(\text{extr}(B_H^*(e_2)))\}$$
$$= \sup\{\phi(\mu) : \mu \in (\text{extr}(B_H^*(e_2)))\},$$

so we deduce that

$$\phi(\rho) \leq \sup\{\phi(\mu) : \mu \in (\text{extr}(B_H^*(e_2)))\}.$$

Since $\rho \in B_H^*(e_2)$ was arbitrary, we get

$$\sup_{\mu \in B_H^*(e_2)} \phi(\mu) \leq \sup_{\mu \in \text{extr}(B_H^*(e_2))} \phi(\mu).$$

On the other hand, since $\text{extr}B_H^*(e_2) \subseteq B_H^*(e_2)$, we have

$$\sup_{\mu \in \text{extr}(B_H^*(e_2))} \phi(\mu) \leq \sup_{\mu \in B_H^*(e_2)} \phi(\mu).$$

Thus

$$\sup_{\mu \in B_H^*(e_2)} ||S^*(\mu)||_H^* = \sup_{\mu \in B_H^*(e_2)} \phi(\mu) = \sup_{\mu \in \text{extr}(B_H^*(e_2))} \phi(\mu) = \sup_{\mu \in \text{extr}(B_H^*(e_2))} ||S^*(\mu)||_H^*.$$

Proof of g) Since $\text{extr}(B_H^*(e_2) = \{v - \pi |\ v, \pi \text{extr} P(e_2), v \perp \pi\}$ by prop. 4.7 in section 4, we have

$$\sup_{\mu \in \text{extr}(B_H^*(e_2))} ||S^*(\mu)||_H^* = \sup_{v,\pi \in \text{extr}(P(e_2)), v \perp \pi} ||S^*(v - \pi)||_H^*$$

Since $S^* = T^*_{|_{M(e_2)}}$ and T^* is linear, we get that:

$$\sup_{v,\pi \in \text{extr}(P(e_2)), v \perp \pi} ||S^*(v - \pi)||_H^* = \sup_{v,\pi \in \text{extr}(P(e_2)), v \perp \pi} ||T^*(v - \pi)||_H^*$$

$$= \sup_{v,\pi \in \text{extr}P(e_2), v \perp \pi} ||T^*(v) - T^*(\pi)||_H^* = \frac{1}{2} \sup_{v,\pi \in \text{extr}P(e_2), v \perp \pi} ||T^*(v) - T^*(\pi)||_T^*$$

Now, in the proof of the first part of theorem 5.1. we have shown that

$$||T^*(v) - T^*(\pi)||_T^* = \sup_{x \in [0,e_1]} 2 < v - \pi, T(x) >$$

Hence

$$\frac{1}{2} \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} ||T^*(v) - T^*(\pi)||_T^*$$

$$= \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} \sup_{x \in [0,e_1]} < v - \pi, T(x) >$$

Thus

$$||S^*||_H^* = \sup_{\mu \in B_H^*(e_2)} ||S^*(\mu)||_H^*$$

$$= \sup_{\mu \in \text{extr}(B_H^*(e_2))} ||S^*(\mu)||_H^*$$

$$= \frac{1}{2} \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} ||T^*(v) - T^*(\pi)||_T^*$$

$$= \sup_{\substack{v,\pi \in P(e_2) \\ v \perp \pi}} \sup_{x \in \text{extr}[0,e_1]} < v - \pi, T(x) >$$

and this completes the proof of the part g). Hence we have proved theorem 5.1 .

Remark 5.2 *Assume that X_1 is finite dimensional. Then $[0,e] = \overline{\text{conv}}(\text{ex}([0,1]))$. Furthermore*

$$\sup_{x \in [0,e_1]} < v - \pi, T(x) > = \sup_{x \in ex([o,e_1])} < v - \pi, T(x) >$$

for all $v, \pi \in P(e_2)$.
Proof: Let

$$B_T(e_1) = \{x \in X_1 \mid ||x||_T \leq 1\}$$

We claim that $[0,e_1]$ is a closed, convex subset of $B_T(e_1)$:

If $x \in [0,e_1]$, by definition of $[0,e_1]$ we have $x \in K$ and $e_1 - x \in K$.

Since $x \in K$ and $e_1 \in K$ we have also $(x + e_1) \in K$. Thus $(x + e_1) \in K$ and $(e_1 - x) \in K$, so we have $-e_1 \leq x \leq e_1$.

Hence

$$||x||_T = \inf\{t > 0 \mid -e_1 \leq \tfrac{1}{t}x \leq e_1\} \leq 1, \text{ so } x \in B_T(e_1).$$

This gives that $[0, e_1] \subseteq B_T(e_1)$

Next we show that $[0, e_1]$ is closed :

Let $\{z_n\}_{n\in N} \subseteq [0, e_1]$ and assume that $||z_n - z||_T \to 0$ as $n \to \infty$ for some $z \in X_1$.

Hence $||(e_1 - z_n) - (e_1 - z)||_T \to 0$ as $n \to \infty$

Since z_n and $e_1 - z_n$ are in K for all n and K is closed, we get that z and $e_1 - z$ are in K. Thus $z \in [0, e_1]$, so it follows that $[0, e_1]$ is closed.

Furthermore $[0, e_1]$ is convex:

If $x, y \in [0, e_1]$ and $\lambda \in (0, 1)$, then $\lambda x + (1 - \lambda)y \in K$ since $x, y \in K$ and K is a convex.

Also

$$e_1 - (\lambda x + (1 - \lambda)y) = \lambda(e_1 - x) + (1 - \lambda)(e_1 - y)) \in K$$

again since $(e_1 - x), (e_1 - y) \in K$ and K is convex.

Hence

$$0 \leq e_1 - (\lambda x + (1 - \lambda)y) \leq e_1 \text{ so } (\lambda x + (1 - \lambda)y) \in [0, e_1]$$

so it follows that $[0, e_1]$ is convex since $x, y \in [0, e_1]$ and $\lambda \in (0, 1)$ were arbitrary.
Since dim $X_1 < \infty$, we have that $B_T(e_1)$ is compact and since $[0, e_1]$ is closed subset of $B_T(e_1)$, it follows that $[0, e_1]$ is compact. Since $[0, e_1]$ also is convex, the Krein - Milman theorem gives that

$$[0, e_1] = \overline{\text{conv}}(ex[0, e_1]).$$

This proves the first statement of the remark 5.2. Now we will prove the second statement:
Suppose that $v, \pi \in P(e_2)$. Then v, π are continuous, hence $v \circ T$ and $\pi \circ T$ are continuous linear functionals on X_1.
Therefore

$$\sup_{x \in \overline{\text{conv}}(\text{ex}([0,e_1]))} < v - \pi, T(x) > = \sup_{x \in \text{conv}(\text{ex}([0,e_1]))} < v - \pi, T(x) >$$

Now, by the first statement of remark 5.2, $[0, e_1] = \overline{\text{conv}}(\text{ex}([0, e_1]))$.
Hence, we have

$$\sup_{x\in[0,e_1]} < v - \pi, T(x) > = \sup_{x\in\overline{\text{conv}}(\text{ex}([0,e_1]))} < v - \pi, T(x) >$$

$$= \sup_{x\in\text{conv}(\text{ex}([0,e_1]))} < v - \pi, T(x) > .$$

Next, let $x \in \text{conv}(ex[0, e_1])$.
Then

$$x = \sum_{k=1}^{n} \alpha_k w_k,$$

where $\alpha_k \geq 0$ for all $k \in \{1, ..., n\}$,

$$\sum_{k=1}^{n} \alpha_k = 1$$

and $w_k \in ex[0, e_1]$ for all $k \in \{1, ..., n\}$.

Hence

$$< v - \pi, T(x) > = < v - \pi, T(\sum_{k=1}^{n} \alpha_k w_k) > = < v - \pi, \sum_{k=1}^{n} \alpha_k T(w_k) >$$

$$= \sum_{k=1}^{n} \alpha_k < v-\pi, T(w_k) > \leq \sum_{k=1}^{n} \alpha_k \sup_{w\in ex([0,e_1])} < v-\pi, T(w) > = \sup_{w\in ex([0,e_1])} < v-\pi, T(w) > .$$

Since $x \in \text{conv}(ex([0, e_1]))$ was arbitrary, we get that

$$\sup_{x\in\text{conv}(ex([0,e_1]))} < v - \pi, T(x) > \leq \sup_{x\in ex([0,e_1])} < v - \pi, Tx > .$$

On the other hand, since $ex([0, e_1]) \subseteq \text{conv}(ex([0, e_1]))$, we have that

$$\sup_{x\in ex([0,e_1])} < v - \pi, Tx > \ \leq \sup_{x\in\text{conv}(ex([0,e_1]))} < v - \pi, Tx > .$$

Then we deduce that

$$\sup_{x\in ex([0,e_1])} < v - \pi, T(x) > \ = \sup_{x\in\text{conv}ex([0,e_1])} < v - \pi, T(x) > \ \forall v, \pi \in P(e_2)$$

Since

$$\sup_{x\in\text{conv}(\text{ex}([0,e_1]))} < v - \pi, T(x) >$$

$$= \sup_{x \in [0,e_1]} < v - \pi, T(x) >$$

for all $v, \pi \in P(e_2)$ as we have shown, we get that

$$\sup_{x \in \mathrm{ex}[0,e_1]} < v - \pi, T(x) > = \sup_{x \in [0,e_1]} < v - \pi, T(x) > .$$

This proves the second statement of the remark 5.2.

7 The convergence of homogeneous discrete time Markov systems

In this section we let again $(X, || \cdot ||)$ be a real Banach space, $K \subseteq X$ be a closed, normal cone with nonempty interior, $e \in \text{Int } K$ be an order unit. All the norms and seminorms that depends on an order unit like $|| \cdot ||_T$, $|| \cdot ||_H$, $||| \cdot |||_H$ etc, are assumed to be given w.r.t. e in this section.
We will state and prove here the most important theorem in this thesis, the theorem 6.1, which considers a Markov operator $T : X \to X$ w.r.t. K and e. Again, this theorem is also given in [GQ] but we have somewhat reformulated it since we are considering $|||\tilde{T}|||_H$ and $||S^*||_H^*$ instead of $||T||_H$ and $||T^*||_H^*$. As we will see, this theorem gives a sufficient condition for the convergence of homogeneous discrete time Markov system given by

$$\mu_{n+1} = (T^*)(\mu_k), \ k = 0, 1, ...$$

where $\mu_0 \in P(e)$.
In fact, the theorem states for instance that if $|||\tilde{T}|||_H \leq 1$, then there exists a $\pi \in P(e)$ s.t.

$$||(T^*)^n(\mu) - \pi||_H^* \leq |||\tilde{T}|||_H^n$$

for all $\mu \in P(e)$ and all $n \in \mathbb{N}$.
In other words, if $|||\tilde{T}|||_H \ \leq 1$, then there exists a unique invariant measure s.t. the homogenous discrete time Markov system given above converges to this measure regardless of the initial distribution.
The proof of the theorem given in this thesis mainly follows the proof given in [GQ], however all the statements used in the proof in [GQ] are given here with detailed proofs and explanations. At the end of the section we state and prove the theorem 6.2 which is also given in [GQ].This theorem applies the theorem 5.1 on the case when $T : X \to X$ is Markov operator w.r.t. K and e and gives thus the expression of $|||\tilde{T}|||_H$ in terms of disjoint extreme point of $P(e)$ is this particular case. The theorem 6.2 will be frequently applied later in sections 7,8 and 9.

Theorem 6.1 *Let $T : X \to X$ be a Markov operator with respect K and e. If $|||\tilde{T}|||_H < 1$ or equivalently $\| S^* \|_H^* < 1$ then there is $\pi \in P(e)$ s.t. for all $x \in X$ and $n \in \mathbb{N}$, we have*

$$\| T^n(x) - < \pi, x > e \|_T \leq (|||\tilde{T}|||_H)^n \| x \|_H$$

$$\text{and } \ \| (T^*)^n(\mu) - \pi \|_H^* \leq (|||\tilde{T}|||_H)^n$$

for all $\mu \in P(e)$, *and all* $n \in \mathbb{N}$.

Proof We will prove this theorem by proving the following :
a) Set

$$I_n(x) = [m(T^n(x)/e), M(T^n(x)/e)]$$

Then $I_{n+1}(x) \subseteq I_n(x)$ *for all* $x \in X$.
Furthermore

$$|I_n(x)| \ \leq \ ||\tilde{T}|||_H^n \ ||x||_H$$

for all n *and there is a real number* $c(x)$ *depending on* x *s.t.*

$$\{c(x)\} = \bigcap_{n \in \mathbb{N}} I_n(x).$$

b) Define $\omega : X \times (Int\ K) \to \mathbb{R}$ *by* $\omega(x/y) = M(x/y) - m(x,y)$.
Then we have

$$-\omega(T^n(x)/e)e \leq (T^n(x) - c(x)e) \leq \omega(T^n(x)/e)e$$

for all $x \in X$ *and all* $n \in \mathbb{N}$.
Hence

$$||(T^n(x) - c(x)e)||_T \leq \omega(T^n(x)/e)) \leq |||\tilde{T}|||_H^n \ ||x||_H$$

for all $x \in X$ *and all* $n \in \mathbb{N}$.

c) Define $\pi : X \to \mathbb{C}$ *by* $< \pi, x > = c(x)$ *where*

$$\{c(x)\} = \bigcap_{n \in \mathbb{N}} I_n(x)$$

as given in part a). Then $\pi \in P(e)$.
d) For all $\mu \in P(x)$, *we have* $||\mu||_T^* = 1$.
Furthermore,

$$||(T^*)^n(\mu) - \pi||_H^* \leq |||T|\tilde{|}|_H^n$$

for all $n \in \mathbb{N}$ *and all* $\mu \in P(e)$,
(Here π *is the functional on* X *defined in c), that is* $< \pi, x > = c(x)$ *for all* $x \in X$.*)*

Proof of a) We want to show that

$$[m(T^{n+1}(x)/e), M(T^{n+1}(x)/e)]$$

$$\subseteq [m(T^n(x)/e), M(T^n(x)/e)] \ \forall n :$$

Assume that $te - T^n x \in K$ for some $t \in \mathbb{R}$. Since $T(K) \subseteq K$, then

$$T(te - T^n(x)) \in K,$$

so we have

$$T(te - T^n(x)) = tT(e) - T^{n+1}(x) = te - T^{n+1}(x) \in K.$$

Hence

$$\{t \in \mathbb{R}|\ T^n(x) \leq te\} \subseteq \{t \in \mathbb{R}|\ T^{n+1}(x) \leq te\},$$

$$\text{so } M(T^n(x)/e) = \inf\{t \in \mathbb{R}|T^n(x) \leq te\} \geq \inf\{t \in \mathbb{R}|T^{n+1}(x) \leq te\}$$

$$= M(T^{n+1}(x)/e).$$

Similarly, if

$$T^n(x) - te \in K, \text{ then}$$

$$T(T^n x - te) = T^{n+1}x - tT(e) = T^{n+1}x - te \in K,$$

$$\text{so } \{t \in \mathbb{R}|T^n(x) \geq te\} \subseteq \{t \in \mathbb{R}|T^{n+1}(x) \geq te\}.$$

Hence

$$m(T^n(x)/e) = \sup\{t \in \mathbb{R}|T^n x \geq te\}$$

$$\leq \sup\{t \in \mathbb{R}|T^{n+1}(x) \geq te\} = m(T^{n+1}(x)/e).$$

This shows that

$$I_{n+1}(x) = [m(T^{n+1}(x)/e), M(T^{n+1}(x)/e)] \subseteq [m(T^n(x)/e), M(T^n(x))/e)] = I_n(x)$$

for all n, that is $I_{n+1}(x) \subseteq I_n(x)$ for all n as desired.

Furthermore observe that each $I_n(x)$ is included in $I_1(x)$ which is compact in $\mathbb{R}$, and each $I_n(x)$ is closed.

Since the family of closed sets $\{I_n(x)\}_{n \in N}$ has finite intersection property,

(because given $I_{n_1}(x), ..., I_{n_k}(x)$ with $n_1 \leq n_2 \leq ... \leq n_k$ then

$$I_{n_1}(x) \supseteq I_{n_2}(x) \supseteq ... \supseteq I_{n_k}(x),$$

so

$$\bigcap_{j=1}^{k} I_{n_j}(x) = I_{n_k}(x) \neq \emptyset.)$$

theorem 11.5 in [MW] part b) gives that

$$\bigcap_{n=1}^{\infty} I_n(x) \neq \emptyset.$$

We observed in section 4 that $|||x + \mathbb{R}e|||_H = ||x||_H$ for all $x \in X$.

Furthermore, we have (by definition) that

$$\tilde{T}(x + \mathbb{R}e) = T(x) + \mathbb{R}\ e \text{ which gives that}$$

$$\tilde{T}^2(x + \mathbb{R}e) = \tilde{T}(\tilde{T}(x + \mathbb{R}e) = \tilde{T}(Tx + \mathbb{R}e) = T(T(x)) + \mathbb{R}e = T^2(x) + \mathbb{R}e.$$

By induction, we get

$$\tilde{T}^n(x + \mathbb{R}e) = T^n(x) + \mathbb{R}e \text{ for all } n \in \mathbb{N}.$$

Hence,

$$\| T^n(x) \|_H = |||T^n(x) + \mathbb{R}e|||_H = |||\tilde{T}^n(x + \mathbb{R}e)|||_H \leq |||\tilde{T}^n|||_H|||(x + \mathbb{R}e)|||_H$$

$$\leq |||\tilde{T}|||_H^n|||(x + \mathbb{R}e)|||_H = |||\tilde{T}|||_H^n \| x \|_H .$$

But

$$||(T^n(x))||_H = M(T^n(x)/e) - m(T^n(x)/e) = |I_n(x)|$$

Hence

$$|I_n(x)| \;=\; ||(T^n(x))||_H \;\leq\; |||\tilde{T}|||_H^n \;\| x \|_H,$$

which goes to 0 as $n \to \infty$, because $|||\tilde{T}|||_H < 1$ by assumption.

Since the length of $I_n(x)$ - s gets arbitrary small when $n \to \infty$,and

$$\bigcap_{n \in N} I_n(x) \neq \emptyset$$

we must have that the intersection is reduced to a real number $c(x) \in \mathbb{R}$ (which certainly depends on x since each $I_n(x)$ depends on x).

So

$$\{c(x)\} = \bigcap_n [m(T^n(x)/e), M(T^n(x)/e)]$$

This proves the part a).

Proof of b): Since

$$c(x) \in [m(T^n(x)/e), M(T^n(x)/e)] \ \forall n,$$

it follows that

$$m(T^n(x)/e) \leq c(x) \leq M(T^n(x)/e) \ \forall n.$$

Hence

$$M(T^n(x)/e) - c(x) \geq 0 \text{ and } c(x) - m(T^n(x)/e) \geq 0 \ \forall n.$$

Furthermore

$$(M(T^n(x)/e - m(T^n(x)/e))e - (T^n(x) - c(x)e)$$

$$= M(T^n(x)/e)e - T^n(x) + (c(x) - m(T^n(x)/e))e)$$

Now,

$$M(T^n(x)/e) = \inf\{t \in \mathbb{R} \mid (te - T^n(x)) \in K\}.$$

Thus there exists $\{t_j\}_{j \in N} \subseteq \mathbb{R}$ s.t.

$$\lim_{j \to \infty} t_j = M(T^n(x)/e)$$

$$\text{and } (t_j e - T^n(x)) \in K$$

for all $j \in \mathbb{N}$. So, we get that

$$M(T^n(x)/e)e - T^n(x) = \lim_{j \to \infty} (t_j e - T^n(x)) \in K,$$

which is in K since K is closed.

Since

$$\lambda K \subseteq K \ \forall \lambda \geq 0, \text{ and } c(x) - m(T^n(x)/e) \geq 0$$

for all $n \in \mathbb{N}$, we have

$$(c(x) - m(T^n(x)/e))e \in K$$

for all $n \in \mathbb{N}$.

Finally, since $K + K \subseteq K$, we have

$$(M(T^n(x)/e)e - T^n(x)) + (c(x) - m(T^n(x)/e)e \in K.$$

that is

$$(M(T^n(x)/e) - m(T^n(x)/e))e - (T^n(x) - c(x)e) \in K$$

for all $n \in \mathbb{N}$.

Thus

$$T^n(x) - c(x)e \leq (M(T^n(x)/e) - m(T^n(x)/e))e = \omega(T^n(x)/e)e \ \forall n.$$

Since

$$m(T^n(x)/e) = \sup\{t \in \mathbb{R} \mid (T^n(x) - te) \in K\}$$

and using again that K is closed, by similar arguments as before we deduce that

$$T^n(x) - m(T^n(x)/e)e \in K$$

Furthermore $M(T^n(x)/e) \geq c(x)$ for all n, so $(M(T^n(x)/e) - c(x))e \in K$ and

$$T^n(x) - c(x)e + \omega(T^n(x)/e)e$$

$$= T^n(x) - c(x)e + M(T^n(x)/e)e - m(T^n(x)/e)e$$
$$= (T^n(x) - m(T^n(x)/e)e) + (M(T^n(x)/e) - c(x))e \quad (*)$$

Since

$$(T^n(x) - m(T^n(x)/e)e), (M(T^n(x)/e) - c(x))e$$

are in K and $K + K \subseteq K$, we get by (*) that

$$(T^n(x) - c(x))e + \omega(T^n(x)/e)e) \in K.$$

Hence

$$-\omega(T^n(x)/e)e \leq T^n(x) - c(x)e.$$

So we have shown that

$$-\omega(T^n(x)/e)e \leq (T^n(x) - c(x))e) \leq \omega(T^n(x)/e)e \text{ for all n.}$$

But this means that for all n,

$$(T^n(x) - c(x))e) \in w(T^n(x)/e)I_e.$$

Since

$$\| T^n(x) - c(x)e \|_T = \inf\{t > 0 | (T^n(x) - c(x)e) \in tI_e\},$$

we get immediately that

$$\| T^n(x) - c(x)e \|_T \leq \omega(T^n(x)/e) \ \forall n.$$

But we have already shown that

$$\omega(T^n(x)/e) \leq |||\tilde{T}|||_H^n \| x \|_H \ \ \forall n.$$

Hence,

$$\| T^n(x) - c(x)e \|_T \leq |||\tilde{T}|||_H^n ||x||_H$$

for all n and this proves the part b).

Proof of c) As $|||\tilde{T}|||_H < 1$, we get that from part b) that

$$\lim_{n\to\infty} \| T^n(x) - c(x)e \|_T = 0,$$

for all $x \in X$, that is

$$\lim_{n\to\infty} T^n(x) = c(x)e \ w.r.t. \ \| \cdot \|_T$$

for all $x \in X$.

Define L: $X \to X$ by

$$L(x) = \lim_{n\to\infty} T^n(x) = c(x)e.$$

Then L is well-defined since the limit exists (as we have proved) and it is linear since T^n's are linear.

Furthermore, L is bounded:

Since $\lim_{n\to\infty} T^n(x)$ exist for all n and all $x \in X$, we have that

$$\sup_n ||T^n(x)||_T < \infty \quad \forall x \in X.$$

By the uniform boundedness principle,

$$\sup_n ||T^n(x)||_T < \infty$$

(Here we have used that $(X, ||\cdot||_T)$ is a Banach space. This follows, since $(X, ||\cdot||)$ is a Banach space and $||\cdot||$ and $||\cdot||_T$ are equivalent because K is closed and normal.)

Now, for all $x \in X$ with $||x||_T \leq 1$,

set

$$||L(x)||_T = \lim_{n\to\infty} ||T^n(x)||_T \leq \sup_n ||T^n(x)||_T \leq \sup_n ||T^n||_T.$$

Hence,

$$\sup_{||x||_T\leq 1} ||L(x)||_T \leq \sup_n ||T^n||_T < \infty.$$

Thus L is bounded with

$$||L(x)||_T \leq \sup_n ||T^n||_T.$$

From this we can deduce that π as defined in part c) is a continuous linear functional:

First observe that $L(x) = c(x)e \ = <\pi, x> e$ for all $x \in X$. Since L is linear and $e \neq 0$ it follows that π is linear.
Since $||e||_T = 1$, we get

$$| <\pi, x> | = |c(x)| = |c(x)| \; ||e||_T = ||c(x)e||_T = ||L(x)||_T \leq ||L||_T ||x||_T$$

Hence π is bounded and $||\pi||_T^* \leq ||L||_T$ so $\pi \in X^*$.

As $T(e) = e$, we have $T^n(e) = e$ for all $n \in \mathbb{N}$.

Hence

$$<\pi, e> e = c(e)e = \lim_{n\to\infty} T^n(e) = e,$$

so it follows that $<\pi, e> = 1$. Furthermore, if $x \in K$, then $T^n(x) \in K$ for all n since $T(K) \subseteq K$, so, we get

$$c(x)e = \lim_{n\to\infty} T^n(x) \in K$$

when $x \in K$.

Now, we have $|c(x)|e \in K$, since $e \in K$ and $|c(x)| \geq 0$.

Since $K \cap (-K) = \{0\}$, and $c(x)e$, $|c(x)|e \in K$ it follows that $c(x) = |c(x)|$ as $e \neq 0$. So we have shown that $< \pi, x > = c(x) \geq 0 \ \ \forall x \in K$.

Hence $\pi \in P(e)$. This proves part c).

Proof of part d) Since $\mu \in P(e)$, by the consequence of lemma 3.1, we have

$$||x||_T = \sup_{v \in P(e)} | < v, x > | \geq | < \mu, x > |$$

for all $x \in X$.

So, if $||x||_T \leq 1$, then $| < \mu, x > | \leq 1$.

Hence

$$||\mu||_T^* = \sup_{||x||_T \leq 1} | < \mu, x > | \leq 1$$

On the other hand , $< \mu, e > = 1$ and $||e||_T = 1$, so

$$\sup_{||x||_T \leq 1} | < \mu, x > | \geq | < \mu, e > | = 1$$

Thus

$$||\mu||_T^* = \sup_{||x||_T \leq 1} | < \mu, x > | = 1$$

which proves the first statement of part d).

Now we prove the second statement of the part d):

Since $(T^*)^n = (T^n)^*$ for all $n \in \mathbb{N}$, then for all $\mu \in P(e)$ we have:

$$< (T^*)^n(\mu) - \pi, x > = < (T^n)^*(\mu) - \pi, x > = < (T^n)^*(\mu), x > - < \pi, x >$$

$$= < \mu, T^n(x) > - < \pi, x > = < \mu, T^n(x) > - < \pi, x > < \mu, e >$$

(since $< \mu, e > = 1$ when $\mu \in P(e)$.)

But $< \pi, x > < \mu, e > = < \mu, < \pi, x > e > = < \mu, c(x)e >$ since μ is linear.

Hence

$$< (T^*)^n(\mu) - \pi, x > = < \mu, T^n(x) > - < \pi, x > < \mu, e >$$

$$= < \mu, T^n(x) > - < \mu, c(x)e >$$

$$= < \mu, T^n(x) - c(x)e > \leq \ ||\mu||_T^* ||T^n(x) - c(x)e||_T$$

$$= ||\mu||_T^* \ ||T^n(x) - c(x)e||_T \ \leq \ ||\mu||_T^* \ |||\tilde{T}|||_H^n \ ||x||_H,$$

for all $n \in \mathbb{N}$, $x \in X$ since we proved that

$$||T^n(x) - c(x)e||_T \leq |||\tilde{T}|||_H^n \ ||x||_H.$$

for all $n \in \mathbb{N}$, $x \in X$. Since $\mu \in P(e)$, we have $||\mu||_T^* = 1$.
Hence

$$< (T^*)^n(\mu) - \pi, x > \leq ||\mu||_T^* \ |||\tilde{T}|||_H^n \ ||x||_H \ = \ |||\tilde{T}|||_H^n \ ||x||_H$$

Let now

$$\phi : M(e) \to (X/\mathbb{R}e)^*$$

be isometric isomorphism (which we considered in the begining of the section 5) given by:

$$< \phi(v), x + \mathbb{R}e > = < v, x > \ \ \forall v \in M(e), \ x \in X.$$

We observe that

$$< (T^*)^n(\mu), e > = < (T^n)^*(\mu), e > = < \mu, T^n(e) > = < \mu, e > = 1$$

for all $n \in \mathbb{N}$ and all $\mu \in M(e)$.

Hence

$$< (T^*)^n(\mu) - \pi, e > = 1 - < \pi, e > = 0$$

for all $n \in \mathbb{N}$ and all $\mu \in M(e)$.

So

$$((T^*)^n(\mu) - \pi) \in M(e).$$

for all n and all $\mu \in M(e)$.

Furthermore, we have

$$< \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > = < (T^*)^n(\mu) - \pi, x > \leq |||\tilde{T}|||_H^n \ ||x||_H$$

for all $x \in X$ and all $\mu \in M(e)$.

Since $||x||_H = |||x + \mathbb{R}e|||_H$ for all $x \in X$,

we get

$$< \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > \leq \ |||\tilde{T}|||_H^n \ |||x + \mathbb{R}e|||_H$$

for all $n \in \mathbb{N}$ and all $\mu \in M(e)$.

This shows that

$$|||\phi((T^*)^n(\mu) - \pi)|||_H \leq |||\tilde{T}|||_H^n.$$

for all $n \in \mathbb{N}$ and all $\mu \in M(e)$.

(Indeed,

$$\text{if} < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > < 0, \text{ then}$$

$$| < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > |$$

$$= - < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e >$$

$$= < \phi((T^*)^n(\mu) - \pi), -x + \mathbb{R}e > \leq \ |||\tilde{T}|||_H^n \ ||| - x + \mathbb{R}e|||_H$$

$$= |||\tilde{T}|||_H^n \ |||x + \mathbb{R}e|||_H.$$

If

$$< \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > \geq 0,$$

then

$$| \ < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > \ |$$
$$= < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > \leq \ |||\tilde{T}|||_H^* \ |||x + \mathbb{R}e|||_H$$

Hence, in any case

$$| < \phi((T^*)^n(\mu) - \pi), x + \mathbb{R}e > | \leq |||\tilde{T}|||_H^n \ |||x + \mathbb{R}e|||_H)$$

so

$$|||\phi((T^*)^n\mu) - \pi|||_H \ \leq \ |||\tilde{T}|||_H^n. \ )$$

Since ϕ is an isometry, obtain

$$||(T^*)^n(\mu) - \pi||_H^* = |||\phi((T^*)^n(\mu) - \pi)|||_H \leq |||\tilde{T}|||_H^n$$

and this proves the second statement of the part d).

This completes the proof of theorem 6.1. As mentioned before, now we will introduce theorem 6.2 which is an application of theorem 5.1 to the particular case when $T : X \to X$ is a Markov operator w.e.t. K and e.

Theorem 6.2 *(Abstract Dolrushin's ergodicity coefficient)*

Let $T : X \to X$ be a Markov operator w.r.t. K and e.

Then

$$|||\tilde{T}|||_H = ||S^*||_H^* = 1 - \inf_{v,\pi \in \text{extr}P(e), v \perp \pi} \inf_{x \in [0,e]} < \pi, T(x) > + < v, T(e-x) >$$

Proof: By theorem 5.1, we have that

$$|||\tilde{T}|||_H = ||S^*||_H^* = \sup_{v,\pi \in \text{extr}P(e), v \perp \pi} \sup_{x \in [0,e]} < v - \pi, T(x) >$$

Now, since $T(e) = e$, we get that

$$< v - \pi, T(x) > = 1- < \pi, T(x) > -(1- < v, T(x) >)$$

$$= 1- < \pi, T(x) > -(< v, e > - < v, T(x) >)$$

$$= 1- < \pi, T(x) > -(< v, T(e) > - < v, T(x) >)$$

$$= 1- < \pi, T(x) > -(< v, T(e) - T(x) >)$$

$$= 1- < \pi, T(x) > -(< v, T(e-x))$$

whenever $v, \pi \in P(e)$

Hence

$$|||\tilde{T}|||_H = ||S^*||_H^* = \sup_{v,\pi \in \text{extr}P(e), v \perp \pi} \sup_{x \in [0,e]} < v - \pi, T(x) >$$

$$= \sup_{v,\pi \in \text{extr}P(e), v \perp \pi} \sup_{x \in [0,e]} (1- < \pi, T(x) > - < v, T(e-x) >$$

$$= 1 - \inf_{v,\pi \in \text{extr}P(e), v \perp \pi} \inf_{x \in [0,e]} (1- < \pi, T(x) > - < v, T(e-x) >$$

8 Applications to stochastic matrices

In this section we let $X = \mathbb{R}^n$ equipped with its usual Euclidian norm, $K = \mathbb{R}^n_+$ be the standard positive cone, $e = \vec{1}$. As mentioned in the introduction of the thesis, throughout this section we will consider $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = Ax$ for all $x \in \mathbb{R}^n$ where A is an $n \times n$ row stochastic matrix. The first part of lemma 7.1. in this section which states that

$$|||\tilde{T}|||_H = \delta(A) = 1 - \min_{i<j} \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}$$

is also given in [GQ] in example 6.3 but without proof. However, we give here the complete proof of lemma 7.1. Similarly, the proposition 7.3 which states that

$$1 - \min_{i<j} \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\} = \frac{1}{2} \max_{i<j} \sum_{k=1}^{n} |a_{i,k} - a_{j,k}|$$

is given in the introduction in [GQ] without proof, whereas we give here a detailed proof of this equality. The rest of the material in this section is not given in [GQ].

Consider the case of the standard positive cone from example 1.4 that is $X = \mathbb{R}^n,\ K = \mathbb{R}^n_+$ and $e = \overrightarrow{1}$.
Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be given by $T(x) = Ax$ for all $x \in \mathbb{R}^n$ where A is an $n \times n$ row stochastic matrix.
Then T is a Markov operator w.r.t K and $\vec{1}$. Denote $|||\tilde{T}|||_H$ by $\delta(A)$, that is we let $|||\tilde{T}|||_H = \delta(A)$.
We have then the following lemma:
Lemma 7.1 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Markov operator defined above. Then we have*

$$\begin{aligned}\delta(A) &= 1 - \min_{i<j} \min_{I \subseteq \{1,\ldots,n\}} \left[\sum_{k \in I} a_{i,k} + \sum_{k \notin I} a_{j,k} \right] \\ &= 1 - \min_{i<j} \sum_{s=1}^{n} \min\{a_{i,s} a_{j,s}\}\end{aligned}$$

Furthermore, $\delta(A) < 1$ if and only if AA^t has only positive coefficients.

Proof: We have

$$\text{extr } [\overrightarrow{0}, \overrightarrow{1}] = \{v_j : J \subseteq \{1, ..., n\}\}$$

where

$$v_j(i) = \begin{cases} 1 & \text{if } i \in J \\ 0 & \text{if } i \notin J. \end{cases}$$

Let $J^c := \{1, ..., n\} \setminus J$

By theorem 6.2 and remark 4.8, we get that

$$|||\tilde{T}|||_H = \delta(A) = 1 - \min_{v,\pi \in extr(P(1)), v \perp \pi} \min_{x \in [\overrightarrow{0}, \overrightarrow{1}]} (< \pi, Ax > + < v, A(\overrightarrow{1} - x) >).$$

$$= 1 - \min_{i<j} \min_{x \in [\overrightarrow{0}, \overrightarrow{1}]} (e_i^t Ax + e_j^t A(\overrightarrow{1} - x))$$

By remark 5.2, since $\mathbb{R}^n$ is finite dimensional, the minimum will be attained at an extreme point of $[\overrightarrow{0}, \overrightarrow{1}]$, hence

$$\delta(A) = 1 - \min_{i<j} \min_{[\overrightarrow{0}, \overrightarrow{1}]} (e_i^t Ax + e_j^t A(\overrightarrow{1} - x))$$

$$= 1 - \min_{i<j} \min_{x \in extr[\overrightarrow{0}, \overrightarrow{1}]} (e_i^t Ax + e_j^t A(\overrightarrow{1} - x))$$

$$= 1 - \min_{i<j} \min_{J \subseteq \{1,...,n\}} (e_i^t A v_J + e_j^t A(\overrightarrow{1} - v_J))$$

$$= 1 - \min_{i<j} \min_{J \subseteq \{1,...,n\}} (e_i^t A v_J + e_j^t A v_{J^c)})$$

$$= 1 - \min_{i,j} \min_{J \subseteq \{1,...,n\}} (\sum_{k \in J} a_{i,k} + \sum_{k \notin J} a_{j,k})$$

where we have used that

$$Av_J = \sum_{k \in J} a_{j,k}$$

$$\text{and } Av_{J^C} = \sum_{k \notin J} a_{j,k}.$$

Fix now $i, j \in \{1, ..., n\}$ and let $\tilde{J} = \{k \mid 1 \leq k \leq n \;\; a_{i,k} \leq a_{j,k}\}$.

Then

$$\sum_{k \in \tilde{J}} a_{i,k} + \sum_{k \notin \tilde{J}} a_{j,k} = \min_{J \subseteq \{1,...,n\}} (\sum_{k \in J} a_{i,k} + \sum_{k \notin J} a_{j,k}) \quad (*)$$

(The explanation of (*) will be given at the end of this section.) By definition of $\tilde{J}$,

$$\sum_{k \in \tilde{J}} a_{i,k} + \sum_{k \notin \tilde{J}} a_{j,k} = \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\},$$

so

$$\sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\} = \min_{J \subseteq \{1,...,n\}} (\sum_{k \in J} a_{i,k} + \sum_{k \notin J} a_{j,k}).$$

Hence, using this in the formula for $\delta(A)$ obtained earlier in this proof, we get

$$\delta(A) = 1 - \min_{i,j} \min_{J \subseteq \{1,...,n\}} (\sum_{k \in J} a_{i,k} + \sum_{k \notin J} a_{j,k}) = 1 - \min_{i,j} (\sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}).$$

which proves the first part of the lemma.

Thus

$$\delta(A) < 1 \;\Leftrightarrow\; \min_{i,j}(\sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}) > 0$$

$$\Leftrightarrow \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}) > 0 \quad \text{for all } i, j \in \{1, ..., n\}$$

This now holds if and only if for all $i, j \in \{1, ..., n\}$, there exists an $k \in \{1, ..., n\}$ s.t $a_{i,k} > 0$ and $a_{j,k} > 0$, which again is true if and only if

$$\sum_{k=1}^{n} (a_{i,k})(a_{j,k}) > 0 \;\text{ for all } i, j \in \{1, ..., n\}$$

If we now let $S = AA^t$, we observe that

$$s_{i,j} = \sum_{k=1}^{n} (a_{i,k})(a_{j,k}).$$

Hence we get that $\sum_{k=0}^{n} (a_{i,k})(a_{j,k}) > 0$ for all $i, j \in \{1, ..., n\}$ if and only if AA^T has only positive coefficients. Combining all these equivalences, we conclude $\delta(A) < 1$ if and only if AA^T has only positive coefficients. This complets the proof of lemma 7.1.

Now we have the following theorem:
Theorem 7.2 *Let A be an $n \times n$ row stochastic matrix, let $Q = A^t$. Then the following 3 conditions are equivalent:*
a) There exists some $k \in \mathbb{N}$ s.t. $\delta(A^k) < 1$.
b) Q has an attractor, that is there exists a stochastic vector $\mu \in \mathbb{R}^n$ s.t. $Q^n v \to \mu$ as $n \to \infty$ for all stochastic vectors $v \in \mathbb{R}^n$.
c) There exists some $k \in \mathbb{N}$ s.t. A^k has (at least) one column with only positive coefficients.

Proof: a)$\Rightarrow$ b): Assume that a) holds.

First observe that A^k is also a row stochastic matrix.

Consider $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = A^k x$.

Since A^k is row stochastic, T is a Markov operator w.r.t. K and $\vec{1}$

$||S^*||_H^* = |||\tilde{T}|||_H = \delta(A^k) < 1$, so S^* is a contraction w.r.t. $||\cdot||_H^*$.

Recall that $S^* = T^*_{|M(\overrightarrow{1})}$.

Now we observe that

$$T^*(x) = (A^k)^t x = (A^t)^k x = Q^k x \quad \forall x \in \mathbb{R}^n \ (**)$$

We also observe that if $\mu, v \in P(\overrightarrow{1})$, then $(\mu - v) \in M(\overrightarrow{1})$.

Hence

$$||T^*(\mu) - T^*(v)||_H^* = ||T^*(\mu - v)||_H^* = ||S^*(\mu - v)||_H^* \leq ||S^*||_H^* \ ||(\mu - v)||_H^*$$

for all $\mu, v \in P(\overrightarrow{1})$.

Recall also that $P(\overrightarrow{1})$ in this case is the set of all stochastic vectors in $\mathbb{R}^n$ by remark 3.1. and that $||\cdot||_H^* = \frac{1}{2}||\cdot||_1$ by remark 4.3.

Using all this together with (**), we get that

$$||Q^k\mu - Q^k v||_1 = 2||T^*(\mu) - T^*(v)||_H^* \leq ||S^*||_H^* 2||\mu - v||_H^*$$

$$= ||S^*||_H^* \ ||\mu - v||_1 \quad \forall \mu, v \in P(\overrightarrow{1}).$$

Thus, there is $c < 1$ $(c = ||S^* \ ||_H^*)$ s.t.

$$||Q^k\mu - Q^k v||_1 \leq c||\mu - v||_1$$

for all stochastic vectors $\mu, v \in \mathbb{R}^n$.

Now, by theorem 6.1, there exists a $\pi \in P(\overrightarrow{1})$ s.t.

$$||(T^*)^n(\mu) - \pi||_H^* \leq |||\tilde{T}|||_H^n$$

for all $\mu \in P(\overrightarrow{1})$ and all $n \in \mathbb{N}$. Then clearly $||(T^*)^n(\mu) - \pi||_H^* \to 0$ as $n \to \infty$, because $|||\tilde{T}|||_H < 1$. Hence it follows that

$$||Q^{kn}\mu - \pi||_1 = 2||(T^*)^n(\mu) - \pi||_H^* \to 0$$

as $n \to \infty$ for all stochastic vectors $\mu \in \mathbb{R}^n$. This shows that π is an attractor for Q^k.
Now we will prove that π is an attractor for Q.

First we prove that $Q\pi = \pi$:

Since π is an attractor for Q^k, it follows that it $Q^k\pi = \pi$. But this gives that $Q\pi = Q(Q^k\pi) = Q^k(Q\pi)$, so $Q\pi$ is also a stochastic vector that is fixed by Q^k.

Now, since
$$||Q^k\mu - Q^k v||_1 \leq c||\mu - v||_1$$
for all stochastic vectors $\mu, v \in \mathbb{R}^n$ and $c = ||S^*||_H^* < 1$, it follows that a stochastic vector fixed by Q^k must be unique. Hence we get $\pi = Q\pi$.
Next we prove that π is an attractor for Q:

Let $\mu \in \mathbb{R}^n$ be an arbitrary stochastic vector. For $n \in \mathbb{N}$, there is an r in $\mathbb{N}$ and $m < k$ s.t. $n = rk + m$.

Since Q is an $n \times n$ column stochastic, matrix, it is not difficult to see that $||Q^s x||_1 \leq ||x||_1$ for all $s \in \mathbb{N}$ and all $x \in \mathbb{R}^n$.
Hence we have:

$$||Q^n(\mu - \pi)||_1 = ||Q^{rk+m}(\mu - \pi)||_1 = ||Q^m(Q^{rk}(\mu - \pi)||_1 \leq ||Q^{rk}(\mu - \pi)||_1.$$

We also have

$$||Q^{rk}(\mu - \pi)||_1 \leq ||S^*||_H^* ||Q^{(r-1)k}(\mu - \pi)||_1 = c||Q^{(r-1)k}(\mu - \pi)||_1$$

where $c = ||S^*||_H^* < 1$. It follows by induction that

$$||Q^{rk}(\mu - \pi)||_1 \leq c^r||(\mu - \pi)||_1 = c^{(\frac{n}{k} - \frac{m}{k})}||(\mu - \pi)||_1 \leq c^{\frac{n}{k}-1}||(\mu - \pi)||_1$$

since $\frac{m}{k} < 1$ and $0 < c < 1$, so $\frac{1}{c^{\frac{m}{k}}} < \frac{1}{c}$

Hence

$$||Q^n\mu - \pi||_1 \leq c^{\frac{n}{k}-1}||(\mu - \pi)||_1$$

which gives that for any stochastic vector $\mu \in \mathbb{R}^n$, $Q^n\mu \to \pi$ as $n \to \infty$. This means that π is an attractor for Q.

b) $\Rightarrow$ c)

Since Q has an attractor, there exists $\pi \in P(\overrightarrow{1})$ s.t. for all $\mu \in P(\overrightarrow{1})$, we have $Q^n\mu \to \pi$ as $n \to \infty$. Let e_j denote the j-th unit vector of $\mathbb{R}^n$. Since $Q^n e_j \to \pi$ as $n \to \infty$ for all $j \in \{1, ..., n\}$, we must have that $q_j^n \to \pi$ as $n \to \infty$ for all $j \in \{1, ..., n\}$ where q_j^n denotes j-th column of Q^n.

Since π is stochastic, there exists an entry π_s, of π s.t. $\pi_s > 0$. Thus, for each j with $1 \leq j \leq n$, there is $k_j \in \mathbb{N}$ s.t. the s-th entry of $Q^m e_j$ is strictly larger than 0 for all $m \geq k_j$. Put $k = \max\{k_j \mid 1 \leq j \leq n \}$. It follows then that s-the entry of $Q^k e_j$ is strictly larger than 0 for all $j \in \{1, ..., n\}$. But since $Q^k e_j$ is exactly the j-th column of Q^k, it follows that the s-th row of Q^k has only strictly positive coefficients. This gives that A^k has r-th column with only positive coefficients because $A^t = Q$.

c) $\Longrightarrow$ a)

Assume that c) holds. As we just saw, there exists $k \in \mathbb{N}$ s.t. for some $r \in \{1, ...n\}$, the r-th row of Q^k has only positive coefficients. This gives that:

$$((Q^k)^t Q^k)_{ij} = \sum_{s=1}^{n} ((Q^k)^t_{is} (Q^k)_{sj}) = \sum_{s=1}^{n} ((Q^k)^t_{si} (Q^k)_{sj})$$

$$= (Q^k)_{ri} (Q^k)_{rj} + \sum_{s=1, s \neq r} (Q^k)_{si} (Q^k)_{sj}$$

$$\geq (Q^k)_{ri} (Q^k)_{rj} > 0$$

for all $i, j \in \{1, ..., n\}$.
Hence $(Q^k)^t Q^k > 0$, which means that $A^k (A^k)^t > 0$. But by lemma 7.1, this is equivalent to

$$\delta(A^k) < 1.$$

Thus, there exists $k \in \mathbb{N}$ s.t. $\delta(A^k) < 1$ and this complets the proof of theorem 7.2. Comment: A similar version of theorem 7.2 is given in [ABS], denoted by "theorem 1.1". It is given in the following way:
Theorem 1.1 *Let P be an $n \times n$ stochastic matrix. Then the following conditions are equivalent:*
(1) P has an attractor.
(2) P is semiregular, i.e. there exists some $s \in \mathbb{N}$ such that $(P^s)^T P^s$ has only positive coefficients.
(3) There exist some $s \in \mathbb{N}$ such that P^s has (at least) one row with only positive coefficients.

Comparing to the theorem 7.2 in this thesis, we observe first that since A is assumed to be an $n \times n$ row stochastic matrix in the theorem 7.2, then $Q = A^t$ is an $n \times n$ column stochastic matrix. Furthermore , by lemma 7.1, we have that $\delta(A^k) < 1$ if and only if $A^k (A^k)^t$ has only positive coefficients. Again, as $Q = A^t$, it follows that $\delta(A^k) < 1$ if and only if $(Q^k)^T Q^k$ has only positive coefficients.So the condition a) in theorem 7.2 is equivalent to that Q is semiregular. Thus the theorem 7.2 in this thesis is actually a reformulated version of the theorem 1.1 in [ABS]. However, we have proved this theorem

by applying directly the theorem 6.1 from section 6, whereas in [ABS] they first need to prove the completness of $\mathbb{R}$ w.r.t. $||x||_1$ and a lemma which they denote by lemma 4.1 before they can prove the main theorem.
The rest of the proof theorem 7.2 in this thesis uses the similar techniques as in the proof of theorem 1.1 in [ABS].

At the end of this section, we introduce the proposition 7.3 which gives an alternative expression for $\delta(A)$ known as Doeblim contraction coefficient:
Proposition 7.3: *Let A be $n \times n$ row stochastic matrix. Then we have*

$$\delta(A) = \frac{1}{2} \max_{i<j} \sum_{1 \leq j \leq n} |a_{i,s} - a_{j,s}|.$$

Proof: From lemma 7.1, we already know that

$$\delta(A) = 1 - \min_{i<j} \sum_{j=1}^{n} \{a_{i,s}, a_{j,s}\}.$$

Hence, it suffices to prove that

$$1 - min_{i<j} \sum_{s=1}^{n} \{aa_{i,s}, a_{j,s}\} = \frac{1}{2} \max_{i<j} \sum_{1 \leq j \leq n} |a_{i,s} - a_{j,s}|.$$

Given $i, j \in \{1, ..., n\}$ with $i < j$, let $K = \{s \in \{1, ..., n\} | a_{i,s} \geq a_{j,s}\}$.

Then

$$\min\{a_{i,s}, a_{j,s}\} = \begin{cases} a_{i,s} & \text{if } s \in K^c \\ a_{j,s} & \text{if } s \in K. \end{cases}$$

where $K^c = \{1, ..., n\} \setminus K$.

Also

$$|a_{i,s} - a_{j,s}|\} = \begin{cases} a_{i,s} - a_{j,s} & \text{if } s \in K \\ a_{j,s} - a_{i,s} & \text{if } s \in K^c. \end{cases}$$

Hence we get

$$\sum_{1 \leq s \leq n} |a_{i,s} - a_{j,s}| = \sum_{s \in K} (a_{i,s} - a_{j,s}) + \sum_{s \in K^c} (a_{j,s} - a_{i,s})$$

$$= \sum_{s \in K} (a_{i,s} + a_{j,s} - 2a_{j,s}) + \sum_{s \in K^c} (a_{i,s} + a_{j,s} - 2a_{i,s})$$

$$= \sum_{1 \leq s \leq n} (a_{i,s} + a_{j,s}) - 2 \left[\sum_{s \in K} a_{j,s} + \sum_{s \in K^c} a_{i,s} \right]$$

$$= 2 - 2 \left[\sum_{s \in K} \min\{a_{i,s}, a_{j,s}\} + \sum_{s \in K^c} \min\{a_{i,s}, a_{j,s}\} \right]$$

$$= 2 - 2 \sum_{1 \leq s \leq n} \min\{a_{i,s}, a_{j,s}\}.$$

This gives that

$$\frac{1}{2} \sum_{1 \leq s \leq n} |a_{i,s} - a_{j,s}| = 1 - \sum_{1 \leq s \leq n} \min\{a_{i,s}, a_{j,s}\}.$$

So we get that

$$\frac{1}{2} \max_{i \leq j} \sum_{1 \leq s \leq n} |a_{i,s} - a_{j,s}| = \max_{i \leq j} (1 - \sum_{1 \leq s \leq n} \min\{a_{i,s}, a_{j,s}\})$$

$$= 1 - \min_{i \leq j} \sum_{1 \leq s \leq n} \min\{a_{i,s}, a_{j,s}\}$$

and this proves the proposition 7.3.

Additional comments:

Explanation of (*)

Let $J \subseteq \{1, ..., n\}$ and $1 \leq i, j \leq n$.

Then

$$\sum_{k \in J} a_{i,k} + \sum_{k \notin J} a_{j,k}$$

$$\geq \sum_{k\in J} \min\{a_{i,k}, a_{j,k}\} + \sum_{k\notin J} \min\{a_{i,k}, a_{j,k}\} = \sum_{k=1}^{n} \min\{a_{i,k}, a_{j,k}\}$$

$$= \sum_{k\in \tilde{J}} a_{i,k} + \sum_{k\notin \tilde{J}} a_{i,k}$$

(Remark on notation: Here $k \notin J$ and $k \notin \tilde{J}$ means $k \in \{1, ..., n\} \setminus J$ and $k \in \{1, ..., n\} \setminus \tilde{J}$ respectively.)

Since we have this inequality for all $J \subseteq \{1, ..., n\}$, we get that

$$\min_{J\subseteq\{1,...,n\}} (\sum_{k\in J} a_{i,k} + \sum_{k\notin J} a_{j,k}) \geq \sum_{k\in \tilde{J}} a_{i,k} + \sum_{k\notin \tilde{J}} a_{j,k}$$

The inequality the other way is obvious since $\tilde{J}$ is a subset of $\{1, ..., n\}$
Hence we must have the equality.

9 Application to Markov operators on $C_{\mathbb{R}}(\Omega)$

As mentioned in the introduction of the thesis, in this section we let $X = C_{\mathbb{R}}(\Omega)$ be equipped with the supremums norm, $K \subseteq C_{\mathbb{R}}(\Omega)$ be the cone consisting of all nonnegative functions on Ω, the order unit $u \in \text{Int } K$ be equal to the constant function 1. The material in this section is not given in [GQ].

Example 8.1: Consider now the space $C_{\mathbb{R}}(\Omega)$ with the sup norm, where Ω is compact metric space.

Let $k : \Omega \times \Omega \to \mathbb{R}$ be a continuous non - negative function and let μ be a positive Radon measure on Ω s.t.

$$\int_{\Omega} k(x, y) d\mu(y) > 0 \quad \forall x \in \Omega$$

(Here we assume that we have such k and μ). Let then $\tilde{k} : \Omega * \Omega \to \mathbb{R}$ be defined as

$$\tilde{k}(x, y) = \frac{k(x, y)}{\int_{\Omega} k(x, y) d\mu(y)}.$$

Then $\tilde{k}$ is continuous, nonnegative and $\int_{\Omega} k(x, y) d\mu(y) = 1$.

Let now $F_{\mathbb{R}}(\Omega)$ denote the space of all real valued functions on Ω and consider the integral operator $T_k : C_{\mathbb{R}}(\Omega) \to F_{\mathbb{R}}(\Omega)$. given by

$$T_k(f)(x) = \int_{\Omega} \tilde{k}(x, y) f(y) d\mu(y) \quad \forall x \in \Omega.$$

Then T_k is clearly linear.

We want to show that

$$T_k(C_{\mathbb{R}}(\Omega)) \subseteq C_{\mathbb{R}}(\Omega):$$

Let $f \in C_{\mathbb{R}}(\Omega)$

Consider the function $T_k(f)$.

If $f \equiv 0$ then $T_k(f) \equiv 0$ so $T_k(f)$ is continuous.

If $f \neq 0$, then $||f||_\infty \neq 0$

Since $\tilde{k}$ is uniformly continuous on $\Omega \times \Omega$, given $\epsilon > 0$, there exists $\delta > 0$ s.t. if $x, x_0 \in \Omega$ and $d(x, x_0) < \delta$, then

$$|\tilde{k}(x, y) - \tilde{k}(x_0, y)| < \frac{\epsilon}{\mu(\Omega)||f||_\infty}$$

for all $y \in \Omega$.

Then, if $d(x, x_0) < \delta$, we get that

$$|T_k(f)(x) - T_k(f)(x_0)| = |\int_\Omega (\tilde{k}(x, y) - \tilde{k}(x_0, y))f(y)d\mu(y)|$$

$$\leq ||f||_\infty \int_\Omega |\tilde{k}(x, y) - \tilde{k}(x_0, y)|d\mu(y) < ||f||_\infty \int_\Omega \frac{\epsilon}{\mu(\Omega)||f||_\infty} d\mu(y) = \epsilon$$

Since $x_0 \in \Omega$ was arbitrary, this shows that $T_k(f)$ is continuous . Thus $T_k : C_{\mathbb{R}}(\Omega) \to C_{\mathbb{R}}(\Omega)$.

T_k is then bounded, since

$$|T_k(f)(x)| = |\int_\Omega \tilde{k}(x, y)f(y)d\mu(y)| \leq \int_\Omega \tilde{k}(x, y)|f(y)|d\mu(y)$$

$$\leq ||f||_\infty \int_\Omega \tilde{k}(x, y)d\mu(y) = ||f||_\infty$$

for all $x \in \Omega$, hence

$$||T_k(f)||_\infty \leq ||f||_\infty$$

for all $f \in C_{\mathbb{R}}(\Omega)$).

We observe that if $f \in K$, then since $\tilde{k}$ is non - negative and μ is positive Radon measure, we get that

$$T_k(f)(x) = \int_\Omega \tilde{k}(x, y)f(y)d\mu(y) \geq 0 \quad \forall x \in \Omega, \text{ so } T_k(f) \in K.$$

Thus $T_k(K) \subseteq K$.

Furthermore

$$T_k(1)(x) = \int_\Omega \tilde{k}(x, y)d\mu(y) = 1 \quad \forall x \in \Omega \text{ so } T_k(1) = 1.$$

Hence T_k is a Markov operator, w.r.t. K and the constant function 1.

Consider now the adjoint of T_k, T_k^*. We have shown before that $(C_{\mathbb{R}}(\Omega))^* = M_r(\Omega)$, where $M_r(\Omega)$ is the space of all signed Radon measures on Ω.

Hence $T_k^* : M_r(\Omega) \to M_r(\Omega)$, and for given $f \in C_{\mathbb{R}}(\Omega)$ and $v \in M_r(\Omega)$ we have

$$< T_k^*(v), f > = < v, T_k(f) >$$

Let $w = T_k^*(v)$.

Then

$$\begin{aligned} \textstyle\int_\Omega f(y)dw(y) = < w, f > = < v, T_k(f) > = &\textstyle\int_\Omega T_k(f)dv(x) \\ = &\textstyle\int_\Omega(\int_\Omega \tilde{k}(x,y)f(y)d\mu(y))dv(x) \end{aligned}$$

Let $v = v_+ - v_-$ be Jordan decomposition of v. Then we have:

$$\int_\Omega \Big(\int_\Omega |\tilde{k}(x,y)f(y)|d\mu(y) \Big) dv_+(x)$$

$$\leq ||\tilde{k}||_\infty \; ||f||_\infty \; \mu(\Omega)v_+(\Omega) < \infty.$$

since u, v_+ are Radon measures and Ω is compact.
Similarly

$$\int_\Omega \Big(\int_\Omega |\tilde{k}(x,y)f(y)|d\mu(y) \Big) dv_-(x) < \infty.$$

By Fubini's theorem, we may change the order of the integration and get

$$\int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)d\mu(y) \Big) dv(x)$$

$$= \int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)d\mu(y)) \Big) dv_+(x) - \int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)|d\mu(y) \Big) dv_-(x)$$

$$= \int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)dv_+(x) \Big) d\mu(y) - \int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)|dv_-(x) \Big) d\mu(y)$$

$$= \int_\Omega \Big(\int_\Omega \tilde{k}(x,y)f(y)dv(x) \Big))d\mu(y).$$

Hence we get

$$\textstyle\int_\Omega f(y)dw(y) = \int_\Omega(\int_\Omega \tilde{k}(x,y)f(y)d\mu(y))dv(x)$$

$$= \int_\Omega (\int_\Omega \tilde{k}(x,y) f(y) dv(x)) d\mu(y)$$

$$= \int_\Omega f(y)(\int_\Omega \tilde{k}(x,y) dv(x)) d\mu(y)$$

$$= \int_\Omega f(y)\rho(y) d\mu(y) \text{ where } \rho(y) = \int_\Omega \tilde{k}(x,y) dv(x))$$

Thus $\int_\Omega f(y) dw(y) = \int_\Omega f(y)\rho(y) d\mu(y)$ for all $f \in C_\mathbb{R}(\Omega)$.

Now, given $y, y_0 \in \Omega$ we have

$$|\rho(y) - \rho(y_0)| \leq \int_\Omega |(\tilde{k}(x,y) - \tilde{k}(x,y_0)| \ d|v|(x)$$

Again, since $\tilde{k}$ is uniformly continuous, given $\epsilon > 0$ we can find $\delta > 0$ s.t if $d(y, y_0) < \delta$, then $|(\tilde{k}(x,y) - \tilde{k}(x,y_0)| \leq \frac{\epsilon}{|v|(\Omega)}$ for all $x \in \Omega$.
Hence, given $y_0 \in \Omega$ and $\epsilon > 0$ if $d(y, y_0) < \delta$, then

$$|\rho(y) - \rho(y_0)| \leq \int_\Omega |(\tilde{k}(x,y) - \tilde{k}(x,y_0)| \ d|v|(x) \leq \frac{\epsilon}{|v|(\Omega)} |v|(\Omega) = \epsilon.$$

This shows that ρ is continuous, hence Borel measurable. (Of course, we assume here that $|v|(\Omega) \neq 0$..If not, then v is a 0 measure, hence ρ is a constant function equal to 0. But then, ρ is continuous hence Borel measurable.)
Since ρ is Borel measurable, the integral

$$\int_E \rho(y) d\mu(y)$$

is well-defined for all $E \subseteq B(\Omega)$.
Let $\eta : B(\Omega) \to \mathbb{R}$ be given by $\eta(E) = \int_E \rho(y) d\mu(y)$.
Then η is a signed Radon measure on Ω since μ is a Radon measure on Ω and ρ is continous, hence bounded on Ω. By exercise 9.51 in [MW] we have

$$\int_\Omega g d\eta = \int_\Omega g p d\mu$$

for all $g \in L^1(|\eta|)$ that is $d\eta = \rho d\mu$. Since η is signed Radon measure on Ω and Ω is compact we have that $|\eta|(\Omega)$ is finite. Hence $C_\mathbb{R}(\Omega) \subseteq L^1(|\eta)|)$, as all f in $C_\mathbb{R}(\Omega)$ are bounded on Ω. So in particular, for all $f \in C_\mathbb{R}(\Omega)$, we have

$$\int_\Omega f \ d\eta = \int_\Omega f\rho \ d\mu$$

by exercise 9.51 in [MW]
But

$$\int_\Omega f\ d\eta = \int_\Omega f\rho\ d\mu$$

for all $f \in C_{\mathbb{R}}(\Omega)$ as we have shown, hence

$$\int_\Omega f\ d\eta = \int_\Omega f\ dw$$

for all $f \in C_{\mathbb{R}}(\Omega)$. So $\phi_\eta = \phi_w$, where $\phi_\eta,\ \phi_w$ are in $(C_{\mathbb{R}}(\Omega))^*$ given by

$$\phi_\eta(f) = \int_\Omega f\ d\eta \quad \forall f \in C_{\mathbb{R}}(\Omega),$$

$$\phi_w(f) = \int_\Omega f\ dw \quad \forall f \in C_{\mathbb{R}}(\Omega).$$

Since we have isometric isomorphism between the space of all signed Radon measures on Ω and $(C_{\mathbb{R}}(\Omega))^*$ via the map $v :\rightarrow \phi_v$ (as we proved in the remark 2.3), it follows that $\eta = w.$. Hence $\rho d\mu = d\eta = dw$

So we have

$$dw = \rho d\mu$$

$$\text{where } w = T_k^*(v)$$

$$\text{and } \rho(y) = \textstyle\int_\Omega \tilde{k}(x,y)dv(x))$$

Then we get

$$d|w| = |\rho|d\mu$$

since μ is a positive measure.

So

$$|T_k^*(v)|(\Omega) = |w|(\Omega) = \textstyle\int_\Omega |\rho(y)|d\mu(y) = \int_\Omega |\int_\Omega \tilde{k}(x,y)dv(x)|d\mu(y)$$

$$\textstyle\int_\Omega |\int_\Omega \tilde{k}(x,y)dv_+(x) - \int_\Omega \tilde{k}(x,y)dv_-(x)|d\mu(y)$$

From section 5 we have

$$|||\tilde{T}_k|||_H = ||S_k^*||_H^* \text{ where } S_k^* = T^*_{k\,|M(1)}.$$

Also we have

$$v \in M(1) \Leftrightarrow \int_\Omega 1 dv = 0 \Leftrightarrow v(\Omega) = 0.$$

Hence, using that

$$|| \cdot ||_H^* = \frac{1}{2} || \cdot ||_T^*$$

by definition and that $||v||_T^* = |v|(\Omega)$ for all $v \in M_r(\Omega)$ by remark 2.3, we obtain:

$$||S_k^*||_H^* = ||T^*_{k_{|M(1)}}||_H^*$$

$$= \sup_{v \in M(1),\ v \neq 0} \frac{||T_K^*(v)||_H^*}{||v||_H^*}$$

$$= \sup_{v \in M(1),\ v \neq 0} \frac{\frac{1}{2}||T_K^*(v)||_T^*}{\frac{1}{2}||v||_T^*}$$

$$= \sup_{v \in M(1),\ v \neq 0} \frac{||T_K^*(v)||_T^*}{||v||_T^*}$$

$$= \sup_{v \in M(1),\ v \neq 0} \frac{|T_K^*(v)|(\Omega)}{|v|(\Omega)}$$

$$= \sup_{v \in M(1),\ v \neq 0} \frac{\int_\Omega | \int_\Omega \tilde{k}(x,y) dv_+(x) - \int_\Omega \tilde{k}(x,y) dv_-(x) | d\mu(y)}{v_+(\Omega) + v_-(\Omega)}$$

Now, if $v \in M(1)$, then $0 = v(\Omega) = v_+(\Omega) - v_-(\Omega)$.
Also

$$|v|(\Omega) = v_+(\Omega) + v_-(\Omega).$$

Hence,

$$v_+(\Omega) = v_-(\Omega) = \tfrac{1}{2}|v|(\Omega).$$

Since $\tilde{k} \geq 0$, we have

$$\int_\Omega \tilde{k}(x,y) dv_+(x),\ \int_\Omega \tilde{k}(x,y) dv_-(x) \ \geq 0 \ \ \forall y \in \Omega.$$

So we get

$$|\int_\Omega \tilde{k}(x,y)dv_+(x) - \int_\Omega \tilde{k}(x,y)dv_-(x)|$$

$$\leq \max\{\int_\Omega \tilde{k}(x,y)dv_+(x),\ \int_\Omega \tilde{k}(x,y)dv_-(x)\}$$

$$\leq \max\{\int_\Omega ||\tilde{k}||_\infty dv_+,\ \int_\Omega ||\tilde{k}||_\infty dv_-\}$$

$$= \max\{||\tilde{k}||_\infty v_+(\Omega),\ ||\tilde{k}||_\infty v_-(\Omega)\}$$

$$= \tfrac{1}{2}||\tilde{k}||_\infty |v|(\Omega) \text{ for all } y \in \Omega.$$

Hence

$$||T^*_{k_{|M(1)}}||^*_H = \sup_{v\in M(1),\ v\neq 0} \frac{|T^*_K(v)|(\Omega)}{|v|(\Omega)}$$

$$\leq \sup_{v\in M(1),\ v\neq 0} \frac{\int_\Omega \frac{1}{2}||\tilde{k}||_\infty |v|(\Omega)d\mu}{|v|(\Omega)}$$

$$= \tfrac{1}{2}||\tilde{k}||_\infty \mu(\Omega).$$

So, if $\tilde{k}$ is s.t. $||\tilde{k}||_\infty < \frac{2}{\mu(\Omega}$, then $|||\tilde{T}_k|||_H < 1$.

As a concrete example, let now $\Omega = [0, 2\pi],\ \mu$ be the Lebesgue measure on $[0, 2\pi]$ and $k : \Omega\ \times\ \Omega \to \mathbb{R}$ be given as $k(x,y) = \frac{1}{4}sin(\frac{1}{4}(x+y))$.

Then k is continuous.

If $x, y \in [0, 2\pi]$, then $\frac{1}{4}(x+y) \in [0, \pi]$, hence $k(x+y) = \frac{1}{4}sin(\frac{1}{4}(x+y)) \geq 0$ for all $x, y \in [0, 2\pi]$.

Furthermore

$$\int_\Omega k(x,y)d\mu(y) = \int_0^{2\pi} k(x,y)dy = \int_0^{2\pi} \frac{1}{4}sin(\frac{1}{4}(x+y))dy$$

$$= [-cos\tfrac{1}{4}(x+y)]_0^{2\pi} = cos(\tfrac{1}{4}x) - cos\tfrac{1}{4}(x+2\pi)$$

$$= \cos(\tfrac{1}{4}x) - cos(\tfrac{1}{4}x + \tfrac{\pi}{2}) = cos(\tfrac{1}{4}x) + sin(\tfrac{1}{4}x).$$

Using elementary calculus, it can be easily checked that

$$\cos(\frac{1}{4}x) + \sin(\frac{1}{4}x) \geq 1$$

for all $x \in [0, 2\pi]$.
In particular,

$$\cos(\frac{1}{4}x) + \sin(\frac{1}{4}x) > 0$$

for all $x \in [0, 2\pi]$, so we can then define $\tilde{k} : \Omega \times \Omega \to \mathbb{R}$ as described in the begining of the section by betting

$$\tilde{k}(x, y) = \frac{k(x, y)}{\int_\Omega k(x, y) d\mu(y)}$$

$$= \frac{k(x, y)}{\cos(\frac{1}{4}x) + \sin(\frac{1}{4}y)}$$

Since $\cos(\frac{1}{4}x) + \sin(\frac{1}{4}x \geq 1$ for all $x \in [0, 2\pi]$, it follows that

$$||\tilde{k}||_\infty = \sup_{x,y \in [0,2\pi]} \frac{\frac{1}{4}\sin(\frac{1}{4}(x + y))}{(\cos(\frac{1}{4}x) + \sin(\frac{1}{4}x))} \leq \frac{1}{4},$$

so

$$||\tilde{k}||_\infty \leq \frac{1}{4} < \frac{1}{\pi} = \frac{2}{2\pi} = \frac{2}{\mu(\Omega)}$$

since $\Omega = [0, 2\pi]$ and μ is the Lebesgue measure.
Hence, it follows that $|||\tilde{T}_k|||_H < 1$. The theorem 6.1 gives then that there exists a unique invariant measure $\tilde{v} \in P([0, 2\pi])$ s.t.

$$|(T_k^*)^n(v) - \tilde{v}|\ ([0, 2\pi]) \to 0 \text{ as } n \to \infty$$

for all $v \in [0, 2\pi]$.

Example 8.2: Again let Ω be a compact Hausdorff topological space, consider $C_\mathbb{R}(\Omega)$.

Let $w : \Omega \to \Omega$, be continuous and let $T_w : C_\mathbb{R}(\Omega) \to C_\mathbb{R}(\Omega)$ be given as $T_w(f) = f \circ w$. We then have:

$$T_w(f)(x) = f(w(x)) \geq 0$$

for all $x \in \Omega$ whenever $f \geq 0$.
Thus $T_w(f) \geq 0$ whenever $f \geq 0$, which means that $T_w(K) \subseteq K$. Furthermore $T(1) = 1 \circ w = 1$, so T_w is a Markov operator w.r.t K and 1.

By theorem 5.1

$$|||\tilde{T}_w|||_H = \sup_{v,\pi \in extrP(1), v\perp\pi} \quad \sup_{0\leq f\leq 1} < v-\pi, T_w(f) >$$

$$= \sup_{x,y\in\Omega, x\neq y} \quad \sup_{0\leq f\leq 1} < \delta_x - \delta_y, T_w(f) >$$

$$= \sup_{x,y\in\Omega, x\neq y} \quad \sup_{0\leq f\leq 1} (\int_\Omega f\circ w \ d_{\delta_x} - \int_\Omega f\circ w \ d_{\delta_y})$$

$$= \sup_{x,y\in\Omega, x\neq y} \quad \sup_{0\leq f\leq 1} (f(w(x)) - f(w(y)))$$

Now, if there exist $\tilde{x}, \tilde{y} \in \Omega, \tilde{x} \neq \tilde{y}$ s.t. $w(\tilde{x}) \neq w(\tilde{y})$, by Urisohn's lemma,

there exists $\tilde{f} \in C_{\mathbb{R}}(\Omega)$ s.t. $\tilde{f}(w(\tilde{x})) = 1 \quad \tilde{f}(w(\tilde{y})) = 0$ and $0 \leq \tilde{f} \leq 1$.

Thus

$$1 = \tilde{f}(w(\tilde{x})) - \tilde{f}(w(\tilde{y}))$$
$$\leq \sup_{x,y\in\Omega, x\neq y} \quad \sup_{0\leq f\leq 1} (f(w(x)) - f(w(y))) = |||\tilde{T}_w|||_H,$$

so $|||\tilde{T}_w|||_H \geq 1$. This is a situation where we can not apply the theorem 6.1
.

10 Application to Kraus maps

As mentioned in the introduction, of the thesis in this section we let $X = S_n$ be the space of all Hermitian matrices in $\mathbb{C}^{n\times n}$ equipped with the operator norm, we let $K = S_n^+$ be the cone in S_n consisting of all positive semidefinite $n \times n$ matrices, and I_n be the order unit.We will mainly follow section 7 in [GQ] but most of the statements used in [GQ] will be supplied with detailed proofs and explanations. In addition, the examples 9.1, 9.2 and 9.8 and the proof of theorem 9.7 in this section are not given in [GQ].
Corollary 9.3 in this section is denoted by corollary 7.1 in [GQ], corollary 9.4 is denoted by corollary 7.3 in [GQ], lemma 9.5 is denoted by lemma 7.6 in [GQ], theorem 9.6 is denoted by theorem 7.7 in [GQ] and theorem 9.7 is denoted by theorem 7.8 in [GQ].

Non commutative Markov operators - Kraus maps:
Let

$$\Phi : S_n \to S_n$$

be defined by

$$\Phi(X) = \sum_{i=1}^{m} V_i^* X V_i$$

for all $X \in S_n$ where $v_1, ..., v_m$ are in $\mathbb{C}^{n\times n}$ and

$$\sum_{i=1}^{m} V_i^* V_i = I_n.$$

Such operator Φ is called a Kraus map.
Then

$$\Phi(I_n) = \sum_{i=1}^{m} V_i^* V_i = I_n$$

and if $X \in K$, then for all $x \in \mathbb{C}^n$, we get

$$< \Phi(X)x, x > = \sum_{i=1}^{m} < V_i^* X V_i x, x > = \sum_{i=1}^{m} < X V_i x, V_i x > \geq 0$$

since X is a positive operator.
Thus

$$\Phi(X) \in K, \text{ so } \Phi(K) \subseteq K \text{ since } X \in K \text{ was arbitrary}$$

Hence Φ is a Markov operator w.r.t I_n and K.

Example 9.1 Let $U \in \mathbb{C}^{n\times n}$ be unitary. For $k \in \{1, ..., m\}$, let

$$D_k = \frac{1}{\sqrt{m}} \begin{bmatrix} e^{i\theta_1^{(k)}} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n^{(k)}} \end{bmatrix}$$

where $\theta_j^{(k)} \in [0, 2\pi]$ for all $j \in \{1, ..., n\}$.
Let $V_k = UD_kU^*$.
Then

$$\begin{aligned} \sum_{k=1}^{m} V_k^* V_k &= \sum_{k=1}^{m} ((UD_k^*U^*)(UD_kU^*)) \\ &= \sum_{k=1}^{m} UD_k^*D_kU^* = \sum_{k=1}^{m} U(\frac{1}{m} I_n)U^* \\ &= \frac{1}{m}\sum_{k=1}^{m} UU^* = I_n, \end{aligned}$$

so the operator Φ given by

$$\Phi(X) = \sum_{k=1}^{m} V_k^* X V_k$$

for all $X \in S_n$ is a Kraus map.

Example 9.2 Let

$$V_1 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

$$V_2 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then

$$V_1^*V_1 + V_2^*V_2 = \tfrac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = I_2,$$

so Φ given by $\Phi(X) = V_1^* X V_1 + V_2^* X V_2$ for all $X \in S_n$ is a Kraus map.

By remark 3.2. we can identify the dual of S_n with S_n itself via the map $B \rightarrow < \cdot, B >$ where $< X, B > = \operatorname{tr}(XB) \ \ \forall X \in S_n$.
The map $B \rightarrow < \cdot, B >$ is an isometric bijection as it was proved in remark 2.2.
If we define $\Psi : S_n \rightarrow S_n$ by

$$\Psi(X) = \sum_{i=1}^{m} V_i X V_i^*,$$

for all $X \in S_n$, we see that

$$< X, \Psi(B) > = \operatorname{tr}(X\Psi(B)) = \operatorname{tr}(\sum_{i=1}^{m} X V_i B V_i^*) = \sum_{i=1}^{m} \operatorname{tr}(V_i^* X V_i B)$$

$$= \operatorname{tr}((\sum_{i=1}^{m} V_i^* X V_i) B) = \operatorname{tr}(\Phi(X)B) = < \Phi(X), B > .$$

for all $X \in S_n$. Hence Ψ is the dual operator of Φ.
Observe that Ψ is trace preserving, since given $X \in S_n$, we have

$$\operatorname{tr}(\sum_{i=1}^{m}(V_i X V_i^*) = \sum_{i=1}^{m} \operatorname{tr}(V_i X V_i^*) = \sum_{i=1}^{m} \operatorname{tr}(X V_i^* V_i) = \operatorname{tr}(\sum_{i=1}^{m} X V_i^* V_i)$$

$$= \operatorname{tr}(X(\sum_{i=1}^{m} V_i^* V_i)) = \operatorname{tr}(X I_n) = \operatorname{tr}(X).$$

As we have shown in section 5, since $\Phi(I_n) = I_n$, we get the induced map $\tilde{\Phi} : S_n/\mathbb{R}I_n \rightarrow S_n/\mathbb{R}I_n$ given by $\tilde{\Phi}(X + \mathbb{R}I_n) = \Phi(X) + \mathbb{R}I_n$ for all $X \in S_n$
Then

$$|||\tilde{\Phi}|||_H = \sup_{\substack{X \in S_n \\ X + \mathbb{R}I_n \neq 0}} \frac{|||\tilde{\Phi}(X + \mathbb{R}I_n)|||_H}{|||X + \mathbb{R}I_n|||_H}.$$

Now, as we have shown in section 4, the lemma 4.1. gives us

$$|||\Phi(X) + \mathbb{R}I_n|||_H = ||\Phi(X)||_H$$

and

$$|||X + \mathbb{R}I_n|||_H = ||X||_H,$$

for all $X \in S_n$, hence

$$|||\tilde{\Phi}|||_H = \sup_{\substack{X \in S_n \\ ||X||_H \neq 0}} \frac{||\Phi(X)||_H}{||X||_H}.$$

By example 1.5, $||X||_H = \lambda_{\max}(X) - \lambda_{\min}(X)$ for all $X \in S_n$, hence

$$|||\tilde{\Phi}|||_H = \sup_{\substack{X \in S_n \\ ||X||_H \neq 0}} \frac{\lambda_{\max}(\Phi(X)) - \lambda_{\min}(\Phi(X))}{\lambda_{\max}(X) - \lambda_{\min}(X)}.$$

As we have proved before, the adjoint map of $\tilde{\Phi}$ is the dual operator of Φ restricted to $M(I_n)$, that is $\Psi_{|_{M(I_n)}} = \tilde{\Psi}$

$$\text{and } ||\tilde{\Psi}||_H^* = \sup_{\mu \in B_H^*(I_n)} ||\tilde{\Psi}(\mu)||_H^*$$

$$= \sup_{\substack{\mu \in B_H^*(I_n) \\ ||\mu||_H^* = 1}} ||\tilde{\Psi}(\mu)||_H^* = \sup_{\substack{\mu \in B_H^*(I_n) \\ ||\mu||_H^* \neq 0}} \frac{||\tilde{\Psi}(\mu)||_H^*}{||\mu||_H^*}.$$

By an earlier result in section 4, $B_H^*(I_n) = \{\rho_1 - \rho_2 : \ \rho_1, \rho_2 \in P(I_n)\}$, hence

$$||\tilde{\Psi}||_H^* = \sup_{\substack{\mu \in B_H^*(I_n) \\ ||\mu||_H^* \neq 0}} \frac{||\tilde{\Psi}(\mu)||_H^*}{||\mu||_H^*}$$

$$= \sup_{\substack{\rho_1, \rho_2 \in P(I_n) \\ \rho_1 \neq \rho_2}} \frac{||\tilde{\Psi}(\rho_1 - \rho_2)||_H^*}{||\rho_1 - \rho_2||_H^*}$$

$$= \sup_{\substack{\rho_1, \rho_2 \in P(I_n) \\ \rho_1 \neq \rho_2}} \frac{||\Psi(\rho_1 - \rho_2)||_H^*}{||\rho_1 - \rho_2||_H^*}$$

(here we use that $\tilde{\Psi} = \Psi_{|_{M(I_n)}}$)
By definition, $||\mu||_H^* = \frac{1}{2}||\mu||_T^* \ \ \forall \mu \in M(I_n)$ and $||X||_T^* = ||X||_1 \ \ \forall X \in S_n$ by remark 2.2 . Now, $M(I_n) \subseteq S_n^* = S_n$, so $||\mu||_H^* = \frac{1}{2}||\mu||_T^* = \frac{1}{2}||\mu||_1$.
Hence

$$||\tilde{\Psi}||_H^* = \sup_{\substack{\rho_1, \rho_2 \in P(I_n) \\ \rho_1 \neq \rho_2}} \frac{||\Psi(\rho_1) - \Psi(\rho_2)||_H^*}{||\rho_1 - \rho_2||_H^*}$$

$$= \sup_{\substack{\rho_1, \rho_2 \in P(I_n) \\ \rho_1 \neq \rho_2}} \frac{\frac{1}{2}||\Psi(\rho_1) - \Psi(\rho_2)||_1}{\frac{1}{2}||\rho_1 - \rho_2||_1}$$

$$= \sup_{\substack{\rho_1, \rho_2 \in P(I_n) \\ \rho_1 \neq \rho_2}} \frac{||\Psi(\rho_1) - \Psi(\rho_2)||_1}{||\rho_1 - \rho_2||_1}.$$

In the next corollary, we will apply theorem 6.2 to derive a concrete expresion for $|||\tilde{\Phi}|||_H$.

Corollary 9.3 (Noncommutative Dobrushin's ergodicity coefficient). *Let Φ be a Kraus map defined in the begining of this section. Then,*

$$|||\tilde{\Phi}|||_H = ||\tilde{\Psi}||_H^* = 1 - \inf_{\substack{u,v:u^*v=0 \\ u^*u=v^*v=1}} \inf_{J\subseteq\{1,\dots,n\}} \inf_{\substack{X=[x_1,\dots,x_n] \\ XX^*=I_n}} (\sum_{i\in J} u^*\Phi(x_ix_i^*)u + \sum_{i\notin J} v^*\Phi(x_ix_i^*)v)$$

Proof First we prove the following claim.
Claim 1: $\mathrm{extr}[0, I_n] = \{P \in S_n : P^2 = P\}$.
If $A \in [0, I_n]$, then A is a positive semidefinite and all its eigenvalues are between 0 and 1. By lemma 3.3.7 in [P], a compact normal operator T has an eigenvalue λ s.t $|\lambda| = ||T||$. Since A is a finite rank operator, it is compact and it is normal since it is self adjoint, so there is an eigenvalue λ for A s.t. $|\lambda| = ||A||$. Since λ is an eigenvalue for A then $\lambda \in [0, 1]$as we observed above so it follows that $0 \leq ||A|| \leq 1$.
By prop 3.2.27 in [P], if T is normal then

$$||T|| = \sup_{||x||\leq 1} \{| < Tx, x > |\}.$$

Since $A = A^*$, we get that

$$||A|| = \sup_{||x||\leq 1} \{| < Ax, x > |\}.$$

Consider now some $P \in S_n$, P is a projection.
Then there is an orthonormal basis of eigenvectors for P $\{x_1, ..., x_k, x_{k+1}, ..., x_n\}$ s.t.

$$P(x_j) = x, \quad \forall j \;\; 1 \leq j \leq k$$
$$P(x_j) = 0, \quad \forall j \;\; k+1 \leq j \leq n.$$

So if $A, B \in [0, I_n]$, $0 < \lambda < 1$ and $\lambda A + (1-\lambda)B = P$, we get that

$$1 = ||x_j||^2 = < Px_j, x_j > = \lambda < Ax_j, x_j > +(1-\lambda) < Bx_j, x_j >$$
$$\leq \lambda||Ax_j||\;||x_j|| + (1-\lambda)||Bx_j||\;||x_j||$$
$$\leq \lambda||A||\;||x_j||^2 + (1-\lambda)||B||\;||x_j||^2$$
$$= \lambda||A|| + (1-\lambda)||B|| \leq \lambda + (1-\lambda) = 1 \text{ for all } j \in \{1, ..., k\}.$$

($A, B \in [0, I_n]$ implies that $0 \leq ||A||, ||B|| \leq 1$ as we proved above .) Thus we must have equality all the way.
In particular, we must have the equality between 1. and 2. line, that is:

$$\lambda < Ax_j, x_j > +(1-\lambda) < Bx_j, x_j > = \lambda||Ax_j||\;||x_j|| + (1-\lambda)||Bx_j||\;||x_j||$$

Since

$$< Ax_j, x_j > \leq ||Ax_j||\ ||x_j||,$$
$$< Bx_j, x_j > \leq ||Bx_j||\ ||x_j||,$$

and $\lambda, 1-\lambda > 0$, we get that

$$< Ax_j, x_j > = ||Ax_j||\ ||x_j|| \text{ and } < Bx_j, x_j > = ||Bx_j||\ ||x_j||$$

Hence there exists $\eta_1^{(j)}$ and $\eta_2^{(j)}$ in $\mathbb{R}$ s.t. $Ax_j = \eta_1^{(j)} x_j$ and $Bx_j = \eta_2^{(j)} x_j$. But we also have the equality between 2. and 3. line, that is

$$\lambda ||Ax_j||\ ||x_j|| + (1-\lambda)||Bx_j||\ ||x_j||$$

$$= \lambda ||A||\ ||x_j||^2 + (1-\lambda)||B||\ ||x_j||^2.$$

Again, since

$$||A||\ ||x_j|| \geq ||Ax_j||$$
$$||B||\ ||x_j|| \geq ||Bx_j||$$

and $\lambda, 1-\lambda > 0$ we must have that

$$||Ax_j|| = ||A||\ ||x_j|| = ||A||$$

$$\text{and } ||Bx_j|| = ||B||\ ||x_j|| = ||B||.$$

But we have that

$$||Ax_j|| = ||\eta_1^{(j)} x_j|| = |\eta_1^{(j)}|\ ||x_j|| = |\eta_1^{(j)}|$$

$$\text{and } ||Bx_j|| = ||\eta_2^{(j)} x_j|| = |\eta_2^{(j)}|\ ||x_j|| = |\eta_2^{(j)}|$$

hence

$$|\eta_1^{(j)}| = ||A|| \text{ and } |\eta_2^{(j)}| = ||B||.$$

Now, by the last equality, that is

$$\lambda ||A|| + (1-\lambda)||B|| = \lambda + (1-\lambda) = 1,$$

it follows that $||A|| = ||B|| = 1$, since $||A||, ||B|| \in [0,1]$ as we proved before.

Hence $|\eta_1^{(j)}| = 1$ and $|\eta_2^{(j)}| = 1$
Since $A, B \in [0, I_n]$, in particular $0 \leq A$ and $0 \leq B$, so $A, B \in K = S_n^+$. Then A and B have nonnegative eigenvalues. Since $\eta_1^{(j)}$, $\eta_2^{(j)}$ are eigenvalues

of A and B respectively (because $Ax_j = \eta_1^{(j)} x_j \quad Bx_j = \eta_2^{(j)} x_j$), they are nonnegative, so we have that $\eta_1^{(j)} = \eta_2^{(j)} = 1$, because $|\eta_1^{(j)}| = |\eta_2^{(j)}| = 1$.
Thus $Ax_j = \eta_1^{(j)} x_j = x_j$ and $Bx_j = \eta_2^{(j)} x_j = x_j$.

If $k+1 \leq j \leq n$, then $Px_j = 0$.
Hence $0 = \lambda < Ax_j, x_j > +(1-\lambda) < Bx_j, x_j >$.
Since $< Ax_j, x_j > \geq 0$ and $< Bx_j, x_j > \geq 0$ we get that

$$< Ax_j, x_j > = < Bx_j, x_j > = 0.$$

Since $0 \leq A$ and $0 \leq B$, there exist $A^{\frac{1}{2}}$ and $B^{\frac{1}{2}}$ so

$$< Ax_j, x_j > = ||A^{\frac{1}{2}} x_j||^2 = 0$$

and

$$< Bx_j, x_j > = ||B^{\frac{1}{2}} x_j||^2 = 0.$$

Thus $A^{\frac{1}{2}} x_j = B^{\frac{1}{2}} x_j = 0$.
Hence

$$Ax_j = A^{\frac{1}{2}}(A^{\frac{1}{2}} x_j) = 0$$

$$\text{and } Bx_j = B^{\frac{1}{2}}(B^{\frac{1}{2}} x_j) = 0.$$

We conclude that A and B have same eigenvectors with same eigenvalues as P, so $A = B = P$. Hence $P \in \text{extr}[0, I_n]$.
If $A \in [0, I_n]$ and A is not a projection , then there exists $\lambda \in (0,1)$ s.t. λ is an eigenvalue of A.
Write A as

$$A = U \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & n_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_{n-1} \end{bmatrix} U^*$$

where U is unitary. Then

$$A = \lambda U \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & n_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_{n-1} \end{bmatrix} U^* + (1-\lambda) U \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & n_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_{n-1} \end{bmatrix} U^*$$

Since $A \in [0, I_n]$, we have that all eigenvalues of are in $[0, 1]$ as we have shown before.

We wrote A as

$$A = U \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \eta_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_{n-1} \end{bmatrix} U^*$$

hence $\eta_k' \ s$ are eigenvalues of A for $1 \leq k \leq n-1$. Thus $\eta_k \in [0, 1]$ for all $k \in \{1, ..., n-1\}$.

Let

$$B = U \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \eta_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_{n-1} \end{bmatrix} U^*$$

$$\text{and } C = U \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \eta_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_{n-1} \end{bmatrix} U^*$$

Then we have $A = \lambda B + (1-\lambda)C$. The set of eigenvalues of B is $\{1, \eta_1, ..., \eta_{n-1}\}$ and the set of eigenvalues for C is $\{0, \eta_1, ..., \eta_{n-1}\}$.
Since $\eta_k \ \in [0, 1]$ for all $k \in \{1, ..., n-1\}$, we have that all eigenvalues of B and C are in $[0, 1]$.
Then it is easz to see that $0 \leq B \leq I_n$ and $0 \leq C \leq I_n$, so $B, C \in [0, I_n]$.
Since $A \neq B$, $A \neq C$, $A = \lambda B + (1-\lambda)C$, and $B, C \in [0, I_n]$ it follows that $A \notin \text{extr}[0, I_n]$.
Thus $\text{extr}[0, I_n] = \{P \in S_n : P^2 = P\}$ and this proves the claim.

By theorem 6.2, $|||\tilde{\Phi}|||_H = ||\tilde{\Psi}||_H^*$

$$= 1 - \inf_{\substack{v,\pi \in \text{extr}P(I_n) \\ v \perp \pi}} \inf_{X \in [0, I_n]} (< \pi, \Phi(X) > + < v, \Phi(I_n - X) >)$$

$$= 1 - \inf_{\substack{v,\pi \in \text{extr}P(I_n) \\ v \perp \pi}} \inf_{X \in [0, I_n]} (\text{tr}(\Phi(X)\pi) + \text{tr}(\Phi(I_n - X)v))$$

By remark 4.9, $\text{extr}P(I_n) = \{xx^* \mid x \in \mathbb{C}^n \; xx^* = 1\}$ and $xx^* \perp yy^*$ if and only if $x^*y = 0$.
Hence

$$1 - \inf_{\substack{v,\pi \in \text{extr}P(I_n) \\ v \perp \pi}} \inf_{X \in [0, In]} (\text{tr}(\Phi(X)\pi) + \text{tr}(\Phi(I_n - X)v))$$

$$= 1 - \inf_{\substack{\mu, v : \mu^* v = 0 \\ \mu^* \mu = v^* v = 1}} \inf_{X \in [0, I_n]} (\text{tr}(\Phi(X)\mu\mu^*) + \text{tr}(\Phi(I_n - X)vv^*)).$$

Now, if

$$A = [a_{i,j}] = \Phi(X)\mu\mu^* = (\Phi(X)\mu)\mu^*,$$

we see that

$$a_{i,j} = (\Phi(X)\mu)_i \overline{\mu}_j.$$

Hence

$$\text{tr}A = \text{tr}(\Phi(X)\mu\mu^*) = \sum_{j=1}^{n} (\Phi(X)u)_j \overline{u}_j.$$

Thus

$$\text{tr}A = \text{tr}(\Phi(X)\mu u^*) = \sum_{j=1}^{n} (\Phi(X)u)_j \overline{u}_j = u^* \Phi(X)\mu.$$

Similarly,

$$\text{tr } \Phi(I_n - X)vv^*) = v^* \Phi(I_n - X)v.$$

Also, by remark 5.2 , since S_n is of finite dimension, we have

$$[0, I_n] = \overline{\text{cov}}(\text{extr}[0, I_n]),$$

so by linearity and continuity of Φ it suffices to consider the infimum over extr $[0, I_n]$.
So we get:

$$|||\tilde{\Phi}|||_H = ||\tilde{\Psi}||_H^*$$

$$= 1 - \inf_{\substack{\mu, v : \mu^* v = 0 \\ \mu * \mu = v^* v = 1}} \inf_{X \in [0, I_n]} (\text{tr}(\Phi(X)\mu\mu^*) + \text{tr}(\Phi(I_n - X)vv^*))$$

$$= 1 - \inf_{\substack{\mu, v : \mu^* v = 0 \\ \mu^{\mu} * = v^* v = 1}} \inf_{X \in extr[0, I_n]} (\text{tr}(\Phi(X)\mu\mu^*) + \text{tr}(\Phi(I_n - X)vv^*))$$

$$= 1 - \inf_{\substack{\mu, v : \mu^* v = 0 \\ \mu^* \mu = v^* v = 1}} \inf_{X \in extr[0, I_n]} (\mu^* \Phi(X)\mu + v^* \Phi(I_n - X)v)$$

In the previous claim, we have proved that

$$X \in [0, I_n] = \{P \in S_n : \; P^2 = P\}.$$

If P is a projection, there exists an orthonormal basis $\{x_1, ..., x_n\}$ for $\mathbb{C}^n$ and a $J \subseteq \{1, 2, ..., n\}$ s.t. $P(x_k) = x_k \;\; \forall k \in J, \;\; P(x_k) = 0 \;\; \forall k \notin J$.
Thus $\{x_k | k \in J\}$ is an orthonormal basis for $P(\mathbb{C}^n)$.
Then, we have

$$P = \sum_{i \in J} x_i x_i^*.$$

Furthermore, since $\{x_1, ..., x_n\}$ is an o.n.b for $\mathbb{C}^n$, we get that $X^*X = I_n$ where $X^* = [x_1, ..., x_n]$. On the other hand, if $X = [x_1, ..., x_n]$ and $X^*X = I_n$ then $\{x_1, ..., x_n\}$ is an o.n.b for $\mathbb{C}^n$ since

$$x_i^* x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Hence if $J \subseteq \{1, ..., n\}$ and

$$A = \sum_{i \in J} x_i x_i^*$$

then A is the projection onto span$\{x_i : i \in J\}$.
So we obtain that

$$\inf_{\text{P:P is a projection}} (\mu^* \Phi(X) \mu + v^* \Phi(I_n - X) v)$$

$$\inf_{J \subseteq \{1,...,n\}} \inf_{\substack{X=[x_1,...,x_n] \\ X^*X=I_n}} (\sum_{i \in J} \mu^* \Phi(x_i x_i^*) \mu + \sum_{i \neq J} v^* \Phi(x_i x_i^*) v)$$

for all $\mu, v \in \mathbb{C}^n$.
Combining all these facts, we get:

$$|||\tilde{\Phi}_H = ||\tilde{\Psi}||_H^* = 1 - \inf_{\substack{\mu,v:\mu^*v=0 \\ \mu^*\mu=v^*v=1}} \inf_{X \in extr[0,I_n]} (\mu^* \Phi(X) \mu + v^* \Phi(I_n - X) v)$$

$$= 1 - \inf_{\substack{\mu,v:\mu^*v=0 \\ \mu^*\mu=v^*v=1}} \inf_{P:P \; is \; a \; projection} (\mu^* \Phi(X) \mu + v^* \Phi(I_n - X) v)$$

$$= 1 - \inf_{\substack{\mu,v:\mu^*v=0 \\ \mu^*\mu=v^*v=1}} \inf_{J \in \{1,...,n\}} \inf_{\substack{X=[x_1,...,x_n] \\ X^*X=I_n}} (\sum_{i \in J} \mu^* \Phi(x_i x_i^*) \mu + \sum_{i \notin J} v^* \Phi(x_i x_i^*) v).$$

This completes the proof of corollary 9.3. Next corollary gives necessary and sufficient condition for the operator norm induced by $||| \cdot |||_H$ to be 1.

Corollary 9.4 *The following conditions are equivalent:*

1. $|||\tilde{\Phi}|||_H = ||\tilde{\Psi}||^*_H = 1$.

2. There are nonzero vectors $u, v \in \mathbb{C}^n$ *such that*

$$< V_i u, V_j v > = 0, \quad \forall i, j \in \{1, ..., m\}.$$

3. There is a rank one matrix $Y \subset \mathbb{C}^{n\times n}$*such that*

$$\mathrm{trace}(V_i^* V_j Y) = 0, \quad \forall i, j \in \{1, ..., m\}.$$

Proof: 1) $\Rightarrow$ 2) :
From corollary 7.1, it follows that $|||\tilde{\Phi}|||_H = 1$ if and only if there is an o.n.b. $\{x_1, ..., x_n\}$ and $u, v \in \mathbb{C}^n$ with $\mu^* v = 0, \mu^*\mu = v^* v = 1$ s.t.

$$\min_{J\subset\{1,...,n\}} (\sum_{i\in J} \mu^*\Phi(x_i x_i^*)\mu + \sum_{i\notin J} v^*\Phi(x_i x_i^*)v) = 0.$$

If we let $\tilde{J} = \{i|\ 1 \le i \le n, \mu^*\Phi(x_i x_i^*)\mu \le v^*\Phi(x_i x_i^*)v\}$,
then

$$\sum_{i\in\tilde{J}} \mu^*\Phi(x_i x_i^*)\mu + \sum_{i\notin\tilde{J}} v^*\Phi(x_i x_i^*)v$$

$$= \min_{J\subset\{1,...,n\}} (\sum_{i\in J} \mu^*\Phi(x_i x_i^*)\mu + \sum_{i\notin J} v^*\Phi(x_i x_i^*)v)$$

So, $|||\tilde{\Phi}|||_H = 1$, if and only if there exists an o.n.b. $\{x_1, ..., x_n\}$ for $\mathbb{C}^n$ and 2 orthonormal vectors $\mu, v \in \mathbb{C}^n$ s.t.

$$\sum_{i\in\tilde{J}} \mu^*\Phi(x_i x_i^*)\mu + \sum_{i\notin\tilde{J}} v^*\Phi(x_i x_i^*)v \ = \ 0$$

where $\tilde{J}$ is as defined above. ($\tilde{J}$ depends on $\{x, ..., x_n\}$ and μ, v)
But, by definition of $\tilde{J}$, we have

$$\sum_{i\in\tilde{J}} \mu^*\Phi(x_i x_i^*)\mu + \sum_{i\notin\tilde{J}} v^*\Phi(x_i x_i^*)v = \sum_{i=1}^{n} \min\{\mu^*\Phi(x_i x_i^*)\mu, v^*\Phi(x_i x_i^*)v\}$$

$$= \sum_{i=1}^{n} \min\{\sum_{k=1}^{m} u^* V_k^* x_i x_i^* V_k \mu, \sum_{j=1}^{m} v^* V_j^* x_i x_i^* V_j v\}$$

Now, for all i, k and j, we have

$$\mu^* V_k^* x_i x_i^* V_j \mu = ||x_i^* V_k \mu||^2 \ge 0,$$

$$v^*V_j^*x_ix_i^*V_jv = ||x_i^*V_jv||^2 \geq 0,$$

so

$$\min\{\sum_{k=1}^{m} \mu^*V_k^*x_ix_i^*V_k\mu, \sum_{j=1}^{m} v^*V_j^*x_ix_i^*V_jv\} \geq 0$$

for all i.
Thus

$$\sum_{i=1}^{n} \min\{\sum_{k=1}^{m} \mu^*V_k^*x_ix_i^*V_k\mu, \sum_{j=1}^{m} v^*V_j^*x_ix_i^*V_jv\} = 0 \text{ for all } i \;\Leftrightarrow$$

$$\Leftrightarrow \min\{\sum_{k=1}^{m} \mu^*V_k^*x_ix_i^*V_k\mu, \sum_{j=1}^{m} v^*V_j^*x_ix_i^*V_jv\} = 0 \text{ for all } i \;\Leftrightarrow$$

$$\Leftrightarrow \min\{\sum_{k=1}^{m} ||x_i^*V_k\mu||^2, \sum_{j=1}^{m} ||x_i^*V_jv||^2\} = 0$$

for all i which is true if and only if for each i, either

$$\sum_{k=1}^{m} ||x_i^*V_k\mu||^2 = 0 \; or \; \sum_{j=1}^{m} ||x_i^*V_jv||^2 = 0$$

This again holds if and only if for each i, either $x_i^*V_k\mu = 0$ for all $k \in \{1, ..., m\}$ or $x_i^*V_jv = 0$ for all $j \in \{1, ..., m\}$.
So

$$\sum_{i \in \tilde{J}} \mu^*\Phi(x_ix_i^*)\mu + \sum_{i \neq \tilde{J}} v^*\Phi(x_ix_i^*)v \;=\; 0$$

if and only if for each i, either $x_i^*V_k\mu = 0$ for all $k \in \{1, ..., m\}$ or $x^*V_jv = 0$ for all $j \in \{1, ..., m\}$.

Thus we have proved so far that $|||\tilde{\Phi}|||_H \; < 1$, if and only if there exist an orthonormal basis $\{x_1, ..., x_n\}$ for $\mathbb{C}^n$ and 2 orthonormal vectors u, v in $\mathbb{C}$ s.t. for each $i \in \; \{1, ..., n\}$ either $x_i^*V_ku = 0$ for all $k \in \{1, ..., m\}$ or $x_iVjv = 0$ for all $j \in \{1, ..., m\}$.

Let now $X = [x_1, ..., x_n]$. Then, since $\{x_1, ..., x_n\}$ is an o.n.b. for $\mathbb{C}^n$, we have that X is unitary, so $XX^* = I_n$.
Hence, if $k, j \in \{1, ..., m\}$, we get that

$$< V_k\mu, V_jv > = < V_k\mu, XX^*V_jv > = < X^*V_k\mu, X^*V_jv >$$

$$= \; \sum_{i=1}^{n} (X^*V_k\mu)_i(\overline{X^*V_jv})_i \; = \; \sum_{i=1}^{n} (x_i^*V_k\mu)(\overline{x_i^*V_jv}).$$

Then, if for each i we have that either $x_i^* V_k \mu = 0$ for all $k \in \{1, ...m\}$ or $x_i^* V_j v = 0$ for all $j \in \{1, ..., m\}$, it follows that

$$< V_k\mu, V_j v > = \sum_{i=1}^{n} (x_i^* V_k \mu)(\overline{x_i^* V_j v}) = 0 \ \ \forall k, j \in \{1, ..., n\}.$$

2) $\Rightarrow$ 1)
Assume that $< V_k\mu, V_j v > = 0$ for all $k, j \in \{1, ..., m\}$.
Let

$$U = \text{span}\{V_k\mu | \ 1 \leq k \leq m\}$$
$$V = \text{span}\{V_j v | \ 1 \leq j \leq m\}.$$

Then $V \subseteq U^{\perp}$. Now, we first find an orthonormal basis for U:

$$\beta_U = \{x_1, ..., x_r\}, \ r \leq n.$$

Then we extend it further to an o.n.b. for whole $\mathbb{C}^n$:

$$\beta = \{x_1, ..., x_r, x_{r+1}, ..., x_n\}.$$

So, if $1 \leq i \leq r$, we have that $x_i \in U$, hence $x_i^* V_j v = 0$ for all $j \in \{1, ..., m\}$ since $V \subseteq U^{\perp}$. If $r + 1 \leq i \leq n$, then $x_i^* V_k \mu = 0$ for all $k \in \{1, ..., m\}$ since $x_i \in U^{\perp}$ for $r + 1 \leq i \leq n$. Thus for each i, either $x_i^* V_k \mu = 0$ for all $k \in \{1, ..., m\}$ or $x_i^* V_j v = 0$ for all $j \in \{1, ..., m\}$. As we have shown, this is equivalent to

$$\sum_{i=1}^{n} \min(\sum_{j=1}^{m} \mu^* V_j^* x_i x_i^* V_j \mu, \sum_{i=1}^{m} v^* V_j^* x_i x_i^* V_j v\} = 0$$

which again is equivalent to $|||\tilde{\Phi}|||_H = 1$.
2) $\Leftrightarrow$ 3)
We have

$$< V_k\mu, V_j v > = \text{tr}((V_k\mu)(V_j v)^*) = \text{tr}(V_k \mu v^* V_j^*) = \text{tr}(V_j^* V_k \mu v^*)$$

for all $k, j \in \{1, ..., m\}$. Hence the equivalence follows if we let $Y = \mu v^*$ This completes the proof of corollary 9.4.

We observe that if Φ is of the form given in example 9.1, then $||\tilde{\Phi}|||_H = 1$: Let u_i and u_j be any 2 column vectors of U s.t. $i \neq j$ (where U is the unitary matrix in example 9.1). Since U is unitary, then u_i, u_j are non zero and $u_i \perp u_j$. Since $V_k = UD_kU^*$ for all $k \in \{1, ..., m\}$, and D_k is a diagonal matrix, it follows that both u_i and u_j are eigenvectors for V_k for all $k \in \{1, ..., m\}$. Hence $< V_{k_1}u_i, V_{k_2}u_j > = 0$ for all $k_1, k_2 \in \{1, ..., m\}$.

Also we observe, that if Φ is of the form given in example 9.2, then

$$|||\tilde{\Phi}|||_H = 1 :$$

Let $u \in \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right],\ v \in \left[\begin{smallmatrix}-1\\1\end{smallmatrix}\right]$ Then u and v satisfy the condition 2) in corollary 9.4.

Consider now a sequence of matrix subspaces of $\mathbb{C}^{n\times n}$ defined as follows: $K_u = \text{span}\{I_n\}$,

$$K_{k+1} = \text{span}\{V_i^* X V_j \mid X \in K_k,\ i, j = 1, ..., m \quad k = 0, 1, ...\}$$

where V_j s are s.t

$$\sum_{j=1}^{m} V_j^* V_j = I_n.$$

Then we have the following lemma:
Lemma 9.5 *There is* $k_0 \leq n^2 - 1$ *s.t.* $K_{k_0+s} = K_{k_0}, \quad \forall s \in \mathbb{N}$

Proof We claim that $K_{k+1} \supseteq K_k\ \forall k \in \mathbb{N}$:
Since

$$K_0 = \text{span}\{I_n\} = \text{span}\{\sum_{j=1}^{m} V_j^* V_j\}$$

it follows that $K_0 \subseteq \text{span}\{V_i^* V_j | i, j = 1, ..., m\}$
If $X \in K_0$, by definition of K_0 , $X = \alpha I_n$ for some $\alpha \in \mathbb{C}$, hence

$$K_1 = \{V_i^* X V_j \mid X \in K_0\ \ i, j = 1, ..., m\} = \text{span}\{\alpha V_i^* V_j :\ \alpha \in \mathbb{C}\ i, j = 1, ..., m\}$$
$$= \text{span}\{V_i^* V_j :\quad i, j = 1, ..., m\}. \text{ Thus } K_0 \subseteq K_1.$$

Assume now that $K_{k+1} \supseteq K_k$ for all $k \in \{1, ..., r\}$ and consider

$$K_{r+1} = \text{span}\{V_i^* X V_j :\ \ X \in K_r\ i, j = 1, ..., m\ \}.$$

If $X \in K_{r-1}$, then $X \in K_r$ by hypothesis, so

$$K_r = \text{span}\{V_i^* X V_j :\ \ X \in K_{r-1}\ i, j = 1, ..., m\ \}$$

$$\subseteq \text{span}\{V_i^* X V_j :\ \ X \in K_r\ i, j = 1, ..., m\ \} = K_{r+1}.$$

By induction, it follows that $K_k \subseteq K_{k+1}$ $\forall k \in \mathbb{N}$.

Also, if for some $k_0 \in \mathbb{N}$ we have that $K_{k_0} = K_{k_0+1}$, then we claim that

$$K_{k_0+s} = K_{k_0} \ \forall s \in \mathbb{N} :$$

Assume that this is true for all $s \in \{1, ..., r\}$. Consider K_{k_0+r+1}.
Since $K_{k_0+r-1} = K_{k_0} = K_{k_0+r}$ by hypothesis, we get that

$$K_{k_0} = K_{k_0+r} = \text{span}\{V_i^* X V_j : \ X \in K_{k_0+r-1}, \quad i, j = 1, ..., m\}$$

$$= \text{span}\{V_i^* X V_j : \ X \in K_{k_0+r}, \quad i, j = 1, ..., m\} = K_{k_0+r+1}.$$

By induction, it follows that

$$K_{k_0+s} = K_{k_0} \quad \forall s \in \mathbb{N}.$$

Since the dimension of $\mathbb{C}^{nxn}$ is n^2 and $K_k \subseteq K_{k+1}$ for all k, as we proved we get that the inequality $K_{k_0+1} \neq K_{k_0}$ can not happen more than n^2 times. Hence, the exists $k_0 \leq n^2 - 1$ s.t. $K_{k_0+1} = K_{k_0}$. By what we have proved, it follows that

$$K_{k_0+s} = K_{k_0} \quad \forall s \in \mathbb{N}.$$

This completes the proof of lemma 9.5.

For all $k \in \mathbb{N}$ let G_k be the orthogonal complement of K_k, Observe that it follows from lemma 9.5 that there is $k_0 \leq n^2 - 1$ s.t $G_{k_0} = G_{k_0+s}$ for all $s \in \mathbb{N}$.
Now we will use the results which we have obtained so far in this section to prove main theorems in this section, theorem 9.6 and theorem 9.7.

Theorem 9.6 *The following conditions are equivalent:*

1. *There exists $k \in \mathbb{N}$ s.t $|||\tilde{\Phi}^k|||_H < 1$.*
2. *Every orbit of the system $X_{k+1} = \Phi(X_k)$ converges to an equilibrium co linear to I_n.*
3. *The subspace $\cap_k G_k$ dos not contain a rank one matrix.*
4. *There exists $k_0 \leq n^2 - 1$ st $|||\tilde{\Phi}^{k_0}|||_H < 1$.*

Proof: (2) $\Rightarrow$ (1) Let $W = S_n/\mathbb{R}I_n$ be equipped with the norm $|||\cdot|||_H$. As we have proved, since $\Phi(I_n) = I_n$, we get the induced bounded linear map $\tilde{\Phi} : W \to W$ given by

$$\tilde{\Phi}(A + \mathbb{R}I_n) = \Phi(A) + \mathbb{R}I_n.$$

We observe now that $|||\tilde{\Phi}|||_H \leq 1$:
By corollary 9.3, we have

$$= 1 - \inf_{\substack{\mu,v:\mu^*v=0\\ \mu^*\mu=v^*v=1}} \inf_{J\in\{1,\dots,n\}} \inf_{\substack{X=[x_1,\dots,x_n]\\ X^*X=I_n}} (\sum_{i\in J} \mu^*\Phi(x_ix_i^*)\mu + \sum_{i\notin J} v^*\Phi(x_ix_i^*)v).$$

Now, given any o.n.b. $\{x_1, ..., x_n\}$ for $\mathbb{C}^n$ and any 2 orthonormal vectors u, v in $\mathbb{C}^n$, we have shown in the begining of the proof of corollary 9.4 part "1) $\Rightarrow$ 2)" that

$$\min_{J\in\{1,\dots,n\}} (\sum_{i\in J} \mu^*\Phi(x_ix_i^*)\mu + \sum_{i\notin J} v^*\Phi(x_ix_i^*)v) \ = \ \sum_{i=1}^{n} \min\{\mu^*\Phi(x_ix_i^*)\mu,\ v^*\Phi(x_ix_i^*)v)$$

Hence, we obtain:

$$|||\tilde{\Phi}|||_H = 1 - \min_{\substack{u,v:u^*v=0\\ u^*u=v^*v=1}} \min_{\substack{X=(x_1,\dots,x_n)\\ XX^*=I_n}} \sum_{i=1}^{n} \min\{u^*\Phi(x_ix_i^*)u, v^*\Phi(x_ix_i^*)v\} \quad (*)$$

Since $\Phi(K) \subset K$, it follows that $\Phi(A)$ is positive semidefinite whenever A is positive semidefinite. Hence, since $x_ix_i^*$ is positive semidefinite for all i (being the orthogonal projection onto Span $\{x_i\}$) we get that $\Phi(x_ix_i^*)$ is positive semidefinite. Hence $u^*\Phi(x_ix_i^*)u = < \Phi(x_ix_i^*)u, u > \geq 0$ for all $u \in \mathbb{C}^n$, so

$$\min\{u^*\Phi(x_ix_i^*)u,\ v^*\Phi(x_ix_i^*)v\} \geq 0 \ \ \forall u, v \in \mathbb{C}^n.$$

Thus

$$\min_{\substack{u,v:u^*v=0\\ u^*u=v^*v=1}} \min_{\substack{X=(x_1,\dots,x_n)\\ XX^*=I_n}} \sum_{i=1}^{n} \min\{u^*\Phi(x_ix_i^*)u, v^*\Phi(x_ix_i^*)v\} \geq 0$$

Using this together with the formula (*), we get that $|||\tilde{\Phi}|||_H \leq 1$. It follows that $|||\tilde{\Phi}^k|||_H \leq |||\tilde{\Phi}|||_H^k \leq 1 \ \ \forall k \geq 1$. Hence, if (1) is not true, then

$$|||\tilde{\Phi}^k|||_H = 1$$

for all k, so

$$1 = \lim_{k\to\infty} |||\tilde{\Phi}^k|||_H^{\frac{1}{k}}$$

Let now $(V, ||\cdot||_v)$ be a real normed vector space and let $V_{\mathbb{C}} = V + iV$. To be more precise, by $V_{\mathbb{C}}$ we mean $V \otimes V$ equipped with usual addition and with scalar multiplication given by

$$(a + ib)(v_1, v_2) = (av_1 - bv_2, bv_1 + av_2).$$

Then we can write $V_{\mathbb{C}}$ as

$$V_{\mathbb{C}} = (V \otimes \{0\}) + i(V \otimes \{0\})$$

Since $V_{\mathbb{C}}$ can be identified with V, we see that we can write $V_{\mathbb{C}}$ as $V_{\mathbb{C}} = V + iV$. We claim that $||\cdot||_{V_{\mathbb{C}}} : V_{\mathbb{C}} \to \mathbb{R}_+$ given by

$$||x + iy||_{V_{\mathbb{C}}} = \sup_{0 \le \theta \le 2\pi} ||x \cos\theta - y \sin\theta||_V$$

is a norm on $V_{\mathbb{C}}$: Clearly $||x + iy||_{V_{\mathbb{C}}} \ge 0$ for all $x, y \in V$ since $||\cdot||_V$ is a norm on V. Assume now that $||x + iy||_{V_{\mathbb{C}}} = 0$ for some $x, y \in V$.
Then

$$\sup_{0 \le \theta \le 2\pi} ||x \cos\theta - y \sin\theta||_V = 0,$$

so

$$||x \cos(0) - y \sin(0)||_V = ||x||_V = 0$$

and

$$||x \cos(\frac{\pi}{2}) - y \sin(\frac{\pi}{2})||_V = ||-y||_V = 0.$$

Since $||\cdot||_V$ is a norm, we get that $x = y = 0$ Furthermore given

$$x_1, x_2, y_1, y_2 \in V$$

we have

$$||(x_1 + iy_1) + (x_2 + iy_2)||_{V_{\mathbb{C}}} = ||(x_1 + x_2) + i(y_1 + y_2)||_{V_{\mathbb{C}}}$$

$$= \sup_{0 \le \theta \le 2\pi} ||(x_1 + x_2)\cos\theta - (y_1 + y_2)\sin\theta||_V$$

$$\le \sup_{0 \le \theta \le 2\pi} (||x_1 \cos\theta - y_1 \sin\theta||_V + ||x_2 \cos\theta - y_2 \sin\theta||_V$$

$$\le \sup_{0 \le \theta \le 2\pi} ||x_1 \cos\theta - y_1 \sin\theta||_V + \sup_{0 \le \theta \le 2\pi} ||x_2 \cos\theta - y_2 \sin\theta||_V$$

$$= ||x_1 + iy_1||_{V_{\mathbb{C}}} + ||x_2 + iy_2||_{V_{\mathbb{C}}}.$$

Let $\alpha \in \mathbb{C}$. Then $\alpha = r(\cos\psi + i \sin\psi)$ where $r \ge 0,\ \psi \in [0, 2\pi]$.

Hence

$$||r(\cos\psi + i\sin\psi)(x+iy)||_{V_{\mathbb{C}}}$$

$$= ||r[(x\cos\psi - y\sin\psi) + i(x\sin\psi + y\cos\psi)]||_{V_{\mathbb{C}}}$$

$$= \sup_{0\le\theta\le 2\pi} ||r[(x\cos\psi - y\sin\psi)\cos\theta - (x\sin\psi + y\cos\psi)\sin\theta||_V$$

$$= \sup_{0\le\theta\le 2\pi} r||x(\cos\psi\cos\theta - \sin\psi\sin\theta) - y(\sin\psi\cos\theta + \cos\psi\sin\theta)||_V$$

$$= \sup_{0\le\theta\le 2\pi} r||x(\cos(\theta+\psi) - y\sin(\theta+\psi)||_V$$

$$\sup_{\psi\le u\le(2\pi+\psi)} r||x\cos(u) - y\sin(u)||_V$$

$$= \sup_{0\le u\le 2\pi} r||x\cos(u) - y\sin(u)||_V = r||x+iy||_{V_{\mathbb{C}}}$$

Thus $||\cdot||_{V_{\mathbb{C}}}$ is a norm on the complex vector space $V_{\mathbb{C}}$. We can then define the linear map $T_{\mathbb{C}}$ on $V_{\mathbb{C}}$ by $T_{\mathbb{C}}(x+iy) = T(x) + iT(y)$ for all $x, y \in V$. We claim that $||T_{\mathbb{C}}||_{V_{\mathbb{C}}} = ||T||_V$ (where $||\cdot||_{V_{\mathbb{C}}}$ and $||\cdot||_V$ are the operator norms induced by $||\cdot||_{V_{\mathbb{C}}}$ and $||\cdot||_V$ respectively):

We have

$$||T_{\mathbb{C}}(x+iy)||_{V_{\mathbb{C}}} = ||T(x) + iT(y)||_{V_{\mathbb{C}}}$$

$$= \sup_{0\le\theta\le 2\pi} ||T(x)\cos\theta - T(y)\sin\theta||_V$$

$$= \sup_{0\le\theta\le 2\pi} ||T(x\cos\theta - y\sin\theta)||_V$$

$$\le \sup_{0\le\theta\le 2\pi} ||T||_V||(x\cos\theta - y\sin\theta)||_V$$

$$= ||T||_V \sup_{0\le\theta\le 2\pi} ||(x\cos\theta - y\sin\theta)||_V$$

$$= ||T||_V||x+iy||_{V_{\mathbb{C}}} \quad \forall x, y \in V.$$

Hence $||T||_{V_{\mathbb{C}}} \le ||T||_V$.
On the other hand

$$||T(x)||_V = \sup_{0\le\theta\le 2\pi} ||T(x)\cos\theta||$$

$$= \sup_{0\le\theta\le 2\pi} ||T(x)\cos\theta - 0\sin\theta||$$

$$= ||T(x) + i0||_{V_{\mathbb{C}}} = ||T(x) + iT(0)||_{V_{\mathbb{C}}}$$

$$= ||T_{\mathbb{C}}(x+i0)||_{V_{\mathbb{C}}} \le ||T_{\mathbb{C}}||_{V_{\mathbb{C}}}||(x+i0)||_{V_{\mathbb{C}}}$$

$$= ||T_{\mathbb{C}}||_{V_{\mathbb{C}}} \sup_{0\le\theta\le 2\pi} ||x\cos\theta||_V$$

$$= ||T_{\mathbb{C}}||_{V_{\mathbb{C}}}||x||_V \quad \forall x \in V$$

Thus

$$||T||_V \leq ||T_{\mathbb{C}}||_{V_{\mathbb{C}}}.$$

Hence we deduce that $||T||_V = ||T_{\mathbb{C}}||_{V_{\mathbb{C}}}$. Now we consider $(S_n)_{\mathbb{C}}$ and claim that $(S_n)_{\mathbb{C}} \simeq \mathbb{C}^{n\times n}$:

By definition, $(S_n)_{\mathbb{C}} = S_n + iS_n$. We define the map $\iota : (S_n + iS_n) \to \mathbb{C}^{n\times n}$ simply by $\iota(A + iB) = A + iB$ Then ι is obviosly linear. Assume that $\iota(A + iB) = 0$ for some $A, B \in S_n$. Then $A + iB = 0$ so $A = -iB$. Since $A, B \in S_n$, we have $A = A^*,\ B = B^*$. Hence

$$A^* = (-iB)^* = iB^* = iB = -A = -A^*.$$

Thus $A = A^* = 0$ and hence also $B = 0$, so ι is injective. (**)

It is also surjective: Let $A \in \mathbb{C}^{n\times n}$. Then clearly $\frac{1}{2}(A + A^*) \in S_n$ and $\frac{1}{2i}(A - A^*) \in S_n$. $(***)$

Furthermore

$$A = \frac{1}{2}(A + A^*) + i[\frac{1}{2i}(A - A^*)] \in S_n + iS_n$$

Hence ι is an isomorphism, so

$$(S_n)_{\mathbb{C}} = (S_n + iS_n) \cong \mathbb{C}^{n\times n}.$$

Next we claim that

$$(S_n/\mathbb{R}I_n)_{\mathbb{C}} \simeq \mathbb{C}^{n\times n}/\mathbb{C}I_n :$$

Let

$$\tilde{\iota} : (S_n/\mathbb{R}I_n)_{\mathbb{C}} \to \mathbb{C}^{n\times n}/\mathbb{C}I_n$$

be given by

$$\tilde{\iota}((A + \mathbb{R}I_n)) + i(B + \mathbb{R}I_n)) = A + iB + \mathbb{C}I_n$$

Then $\tilde{\iota}$ is well defined: Assume that

$$(A_1 + \mathbb{R}I_n) + i(B_1 + \mathbb{R}I_n) = (A_2 + \mathbb{R}I_n) + i(B_2 + \mathbb{R}I_n)$$

Then we must have $A_1 + \mathbb{R}I_n = A_2 + \mathbb{R}I_n$ and $B_1 + \mathbb{R}I_n = B_2 + \mathbb{R}I_n$ Hence there exists a,b $\in \mathbb{R}$ s.t. $A_1 = A_2 + aI_n,\ B_1 = B_2 + bI_n$. But then

$$\tilde{\iota}((A_1 + \mathbb{R}I_n) + i(B_1 + \mathbb{R}I_n))$$

$$= A_1 + iB_1 + \mathbb{C}I_n$$

$$= A_2 + iB_2 + (a+ib)I_n + \mathbb{C}I_n$$
$$= A_2 + iB_2 + \mathbb{C}I_n$$
$$= \tilde{\imath}((A_2 + \mathbb{R}I_n) + i(B_2 + \mathbb{R}I_n)).$$

Thus $\tilde{\imath}$ is well defined. Also, $\tilde{\imath}$ is linear since for instance

$$\tilde{\imath}([(A+\mathbb{R}I_n) + i(B+\mathbb{R}I_n)] + [(C+\mathbb{R}I_n) + i(D+\mathbb{R}I_n)]$$
$$= \tilde{\imath}((A+\mathbb{R}I_n) + (C+\mathbb{R}I_n) + i[(B+\mathbb{R}I_n) + (D+\mathbb{R}I_n)])$$
$$= \tilde{\imath}(A+C+\mathbb{R}I_n + i(B+D+\mathbb{R}I_n))$$
$$= (A+C) + i(B+D) + \mathbb{C}I_n = (A+iB) + (C+iD) + \mathbb{C}I_n$$
$$(A+iB+\mathbb{C}I_n) + (C+id+\mathbb{C}I_n)$$
$$\tilde{\imath}((A+\mathbb{R}I_n) + i(B+\mathbb{R}I_n) + \tilde{\imath}(C+\mathbb{R}I_n + i(D+\mathbb{R}I_n)).$$

If $A+iB+\mathbb{C}I_n = \mathbb{C}I_n$ for some $A, B \in S_n$, then $A+iB = (a+ib)I_n = aI+ibI_n$ for some $a, b \in \mathbb{R}$ Hence $(A-aI_n) = -i(B-bI_n)$. Since $(A-aI)$ and $(B-bI)$ are in S_n, by (**) on previous page , it follows that $A - aI_n = B - bI_n = 0$. Hence $A = aI_n$ and $B = bI_n$, so $A, B \in \mathbb{R}I_n$ which gives that

$$A + \mathbb{R}I_n = B + \mathbb{R}I_n = \mathbb{R}I_n$$

Thus $\tilde{\imath}$ is injective . Also $\tilde{\imath}$ is surjective since each $A \in \mathbb{C}^{n\times n}$ can be written as $A = H + iK$ where $H, K \in S_n$ (by (***) on previous page). Thus $\tilde{\imath}$ is isomorphism, so

$$(S_n/\mathbb{R}I_n)_{\mathbb{C}} \simeq \mathbb{C}^{n\times n}/\mathbb{C}I_n$$

Recall now that $W := S_n/\mathbb{R}I_n$, so that

$$W_{\mathbb{C}} = (S_n/\mathbb{R}I_n)_{\mathbb{C}} \simeq \mathbb{C}^{n\times n}/\mathbb{C}I_n.$$

Then, we have

$$||\tilde{\Phi}^k_{\mathbb{C}}||_{W_{\mathbb{C}}} = ||\tilde{\Phi}^k||_W = ||\tilde{\Phi}^k_W||_H.$$

Hence

$$\lim_{k\to\infty} ||\tilde{\Phi_{\mathbb{C}}}^k||^{\frac{1}{k}}_{W_{\mathbb{C}}} = \lim_{k\to\infty} |||\tilde{\Phi}^k|||^{\frac{1}{k}}_W = 1.$$

Now, we have that $\tilde{\Phi}_{\mathbb{C}} \in B(W_{\mathbb{C}}, ||\cdot||_{W_{\mathbb{C}}})$ and that $B(W_{\mathbb{C}}, ||\cdot||_{W_{\mathbb{C}}})$ is a Banach algebra, so by spectral radius theorem (theorem 4.1.13 on page 131 in [P]), we have that $\lim_{k\to\infty} ||\tilde{\Phi_{\mathbb{C}}}^k||^{\frac{1}{k}}_{W_{\mathbb{C}}}$ is the spectral radius of $\tilde{\Phi_{\mathbb{C}}}$. Thus $\sup\{|\lambda| \mid \lambda I_{\mathbb{C}} - \tilde{\Phi}_{\mathbb{C}}$ is not invertible $\} = 1$. Since $W_{\mathbb{C}}$ is finite dimensional, we have that the spectrum of $\tilde{\Phi}_{\mathbb{C}}$ consists of eigenvalues for $\tilde{\Phi}_{\mathbb{C}}$. Thus there is s $\lambda \in \Pi$ and $X, Y \in S_n$ s.t.

$$\tilde{\Phi}_{\mathbb{C}}((X + \mathbb{R}I_n) + i(Y + \mathbb{R}I_n)$$

$$= \lambda((X + \mathbb{R}I_n) + i(Y + \mathbb{R}I_n))$$

(here Π is the unit circle in $\mathbb{C}$). Let now

$$\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}} = \tilde{\iota} \circ \tilde{\Phi}_{\mathbb{C}} \circ \tilde{\iota}^{-1}$$

where $\tilde{\iota} : W_{\mathbb{C}} \to \mathbb{C}^{n\times n}/\mathbb{C}I_n$ is the isomorphism we considered, that is

$$\tilde{\iota}((A + \mathbb{R}I_n) + i(B + \mathbb{R}I_n) = A + iB + \mathbb{C}I_n \quad \forall A, B \in S_n.$$

Then

$$\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}} : \mathbb{C}^{n\times n}/\mathbb{C}I_n \to \mathbb{C}^{n\times n}/\mathbb{C}I_n.$$

Since λ is an eigenvalue for $\tilde{\Phi}_{\mathbb{C}}$ with a coresponding eigenvector

$$(X + \mathbb{R}I_n) + i(Y + \mathbb{R}I_n)$$

it follows that λ is an eigenvalue for $\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}}$ with a corresponding eigenvector

$$\tilde{\iota}((X + \mathbb{R}I_n) + i(Y + \mathbb{R}I_n)) = X + iY + \mathbb{C}I_n.$$

Thus we have

$$\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}}(X + iY + \mathbb{C}I_n) = \lambda(X + iY) + \mathbb{C}I_n.$$

Hence

$$(\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}})^k(X + iY + \mathbb{C}I_n) = \lambda^k(X + iY) + \mathbb{C}I_n.$$

We observe that

$$\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}}(X + iY + \mathbb{C}I_n) = \tilde{\iota}(\tilde{\Phi}_{\mathbb{C}}(\tilde{\iota}^{-1}(X + iY + \mathbb{C}I_n)))$$

$$= \iota(\tilde{\Phi}_{\mathbb{C}}((X + \mathbb{R}I_n) + i(Y + \mathbb{R}I_n)))$$

$$= \tilde{\iota}(\tilde{\Phi}(X + \mathbb{R}I_n) + i\tilde{\Phi}(Y + \mathbb{R}I_n)))$$

$$= \tilde{\iota}((\Phi(X) + \mathbb{R}I_n) + i(\Phi(Y) + \mathbb{R}I_n))$$

$$= \Phi(X) + i\Phi(Y) + \mathbb{C}I_n$$

By induction, it is easy to see that

$$(\tilde{\Phi}_{\mathbb{C}}^{\tilde{\iota}})^k(X + iY + \mathbb{C}I_n) = \Phi^k(X) + i\Phi^k(Y) + \mathbb{C}I_n.$$

So we get that

$$\Phi^k(X) + i\Phi^k(Y)\mathbb{C}I_n = \lambda^k(X + iY) + \mathbb{C}I_n \quad \forall k \in \mathbb{N}.$$

Now, since $\lambda \in \Pi$, there exist θ in $[0, 2\pi]$ s.t. $\lambda = e^{i\theta}$.
Hence we have

$$\Phi^k(X) + i\Phi^k(Y) + \mathbb{C}I_n = e^{ik\theta}(X + iY)+) + \mathbb{C}I_n$$

for all $k \in \mathbb{N}$, that is

$$\Phi^k(X) + i\Phi^k(Y) - e^{ik\theta}(X + iY) \in \mathbb{C}I_n$$

for all $k \in \mathbb{N}$. We also observe that since $X+iY+\mathbb{C}I_n$ is an eigenvector for $\tilde{\Phi}^{\tilde{\iota}}_{\mathbb{C}}$ it is then non - zero vector. Thus $X+iY+\mathbb{C}I_n \neq \mathbb{C}I_n$ that is $(X+iY) \notin \mathbb{C}I_n$.
Since

$$\Phi^k(X) + i\Phi^k(Y) - e^{ik\theta}(X + iY) \in \mathbb{C}I_n \quad \forall k \geq 1$$

we have that for all $k \geq 1$ there exist $a_k, b_k \in \mathbb{R}$ s.t.

$$\Phi^k(X) + i\Phi^k(Y) - e^{ik\theta}(X + iY) = (a_k + ib_k)I_n$$

Hence

$$\Phi^k(X) - \cos(k\theta)X + \sin(k\theta)Y - a_kI_n + i(\Phi^k(Y) - \sin(k\theta)X - \cos(k\theta)Y - b_kI_n) = 0 \ \forall k \geq 1.$$

Let

$$A_k = \Phi^k(X) - \cos(k\theta)X + \sin(k\theta)Y - a_kI_n,$$
$$B_k = \Phi^k(Y) - \sin(k\theta)X - \cos(k\theta)Y - b_kI_n.$$

So we have that $A_k + iB_k = 0$ for all $k \geq 1$ which gives that $A_k = -iB_k$ for all $k \geq 1$. Since $A_k, B_k \in S_n$ for all $k \geq 1$, we have $A_k^* = A_k$, $B_k^* = B_k$, so

$$A_k^* = (-iB_k)^* = iB_k^* = iB_k = -A_k = -A_k^*$$

where we have used twice that $A_k = -iB_k$ for all $k \geq 1$.
Hence $A_k^* = A_k = 0$ and also $B_k = iA_k = 0$.
Thus

$$\Phi^k(X) = \cos(k\theta)X - \sin(k\theta)Y + a_kI_n$$

and

$$\Phi^k(Y) = \sin(k\theta)X + \cos(k\theta)Y + b_kI_n$$

since $A_k, B_k = 0$ for all $k \geq 1$.
Since $(X + iY) \in \mathbb{C}I_n$, it also follows that we can not have that both X and

Y are in $\mathbb{R}I_n$.
If we have that $\Phi^k(X) \to \alpha I_n$ and $\Phi^k(Y) \to \beta I_n$ for some constants α, β , then we must have

$$(\Phi^k(X))_{i,j} \to (\alpha I_n)_{i,j})$$

and

$$(\Phi^k(Y))_{i,j} \to (\beta I_n)_{i,j})$$

ak $k \to \infty$ for all $i, j \in \{1, ..., n\}$. If $i \neq j$ then $(\alpha I_n)_{i,j}) = (\beta I_n)_{i,j}) = 0$, so $(\Phi^k(X))_{i,j}$ and $(\Phi^k(Y))_{i,j}$ converges to 0 as $k \to \infty$. Using the expressions that we have for $(\Phi^k(X))$ and $(\Phi^k(Y))$ above we get that

$$\cos(k\theta)x_{i,j} - \sin(k\theta)y_{i,j} \to 0$$

and

$$\sin(k\theta)x_{i,j} + \cos(k\theta)y_{i,j} \to 0 \quad as\ k\ \to \infty$$

(since $(a_k I_n)_{i,j} = (b_k I_n)_{i,j} = 0$ when $i \neq j$).
Thus

$$\begin{bmatrix} \cos(k\theta) & -\sin(k\theta) \\ \sin(k\theta) & \cos(k\theta) \end{bmatrix} \begin{bmatrix} x_{i,j} \\ y_{i,j} \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad as\ k \to \infty.$$

But

$$\begin{bmatrix} \cos(k\theta) & -\sin(k\theta) \\ \sin(k\theta) & \cos(k\theta) \end{bmatrix} = (R_\theta)^k,$$

where R_θ is the rotation matrix given by

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

so

$$(R_\theta)^k = \begin{bmatrix} x_{i,j} \\ y_{i,j} \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{as } k \to \infty.$$

Since (R_θ) is the rotation matrix, it preserves the length of the vector, so we have the convergence towards the zero vector if and only if for all
$x_{i,j} \in \{1, ..., n\}$ with $i \neq j$, we have $x_{i,j} = y_{i,j} = 0$ Since $\Phi^k(X) \to \alpha I_n$ and $\Phi^k(Y) \to \beta I_n$ we must also have that $(\Phi^k(X))_{i,i}$ and $(\Phi^k(Y))_{i,i}$ converges to α and β respectively as $k \to \infty$ for all $i \in \{1, ..., n\}$. Using again the expresions for $\Phi^k(X)$ and $\Phi^k(Y)$, we obtain then that

$$(\cos(k\theta)x_{i,i} - \sin(k\theta)y_{i,i} + a_k) \to \alpha$$

and

$$(\sin(k\theta)x_{i,i} + \cos(k\theta)y_{i,i} + b_k) \to \beta \;\; as \; k \to \infty.$$

Since this should hold for all $i \in \{1, ..., n\}$, we get that

$$(\cos(k\theta)(x_{i,i} - x_{j,j}) - \sin(k\theta)(y_{i,i} - y_{j,j})) \to 0$$

and

$$\sin(k\theta)(x_{i,i} - x_{j,j}) - + \cos(k\theta)(y_{i,i} - y_{j,j}) \to 0 \;\text{ as } k \to \infty.$$

for all $i, j \in \{1, ..., n\}$. By the same arguments as before, we deduce that $x_{i,i} - x_{j,j} = y_{i,i} - y_{j,j} = 0$ for all $i, j \in \{1, ..., n\}$.

Hence $x_{i,i} = x_{j,j}$, $y_{i,i} = y_{j,j} = 0$ whenever $i \neq j$. This means that X and Y must be a (real) multiple of I_n which is a contradiction. Thus we can not have that both $\Phi^k(X)$ and $\Phi^k(Y)$ converges to real multiple of I_n. Hence if 1) does not hold, then 2) does not hold.

1) $\Longrightarrow$ 2) : We observe first that if $(X, ||\cdot||)$ is a real Banach space, $K \subseteq X$ is a normal, closed cone and $e \subseteq \text{Int } \; K$ is an order unit, then $||T(x)||_T \leq ||x||_T$ for all $x \in X$ where $||\cdot||_T$ is Thompson norm w.r.t. e:
In section 1 we have proved that

$$||x||_T = \inf\{t > 0 \mid x \in tI_e\}$$

for all $x \in X$.
Now, $x \in tI_e$ with $t > 0$ if and only if $-e \leq \frac{1}{t}x \leq e$, that is if and only if $(e - \frac{1}{t}x), (e + \frac{1}{t}x) \in K$. But, if $(e - \frac{1}{t}x), (e + \frac{1}{t}x) \in K$, we get that

$$(e - \frac{1}{t}T(x)), (e + \frac{1}{t}T(x)) \in K,$$

since $T(K) \subseteq K$, $T(e) = e$ and T is linear. Then $T(x) \subseteq tI_e$.
This means that if $x \in tI_e$, then $T(x) \subseteq tI_e$ whenever $t > 0$. Hence

$$||x||_T = \inf\{t > 0 \; | x \in tI_e\} \geq \inf\}t : 0 \mid T(x) \in tI_e\} = ||T(x)||_T.$$

This proves the observation.

Now, if 1) in theorem 9.6 holds, then by applying theorem 6.1 to the Markov operator Φ^k, we deduce that there exists $Q \in P(I_n)$ s.t.

$$||\Phi^{km}(X) - \text{tr}(XQ)I_n||_T \;\leq\; |||\tilde{\Phi}^k|||_H^m \, ||X||_H$$

for all $X \in P(I_n)$ and all $m \in \mathbb{N}$. This means that $\Phi^{km}(X) \to \text{tr}(XQ)I_n$ as $m \to \infty$ for all $X \in P(I_n)$ since $|||\tilde{\Phi}^k|||_H \; < 1$ by assumption.
Let $s \in \mathbb{N}$. Then there is $j \in \mathbb{N}$ and $r \in \{1, ..., k-1\}$ s.t. $s = jk + r$. Since

$$||\Phi^r(A)||_T \;\leq\; ||A||_T,$$

for all $A \in S_n$ by the observation above, as Φ^r is a Markov operator w.r.t. K and I_n we get that

$$\begin{aligned} ||\Phi^s(X) - \mathrm{tr}(XQ)I_n||_T &= ||\ \Phi^s(X - \mathrm{tr}(XQ)I_n)||_T \\ &= ||\Phi^{jk+r}(X - \mathrm{tr}(XQ)I_n)||_T \\ = ||\Phi^r(\Phi^{jk}(X - \mathrm{tr}(XQ)I_n))||_T &\leq ||\Phi^{jk}(X - \mathrm{tr}(XQ)I_n)||_T \\ = ||\Phi^{jk}(X) - \mathrm{tr}(XQ)I_n)||_T &\leq |||\tilde{\Phi}^k|||_H^j \ ||X||_H. \end{aligned}$$

But

$$j = \frac{s-r}{k} > \frac{s-k}{k} = \frac{s}{k} - 1.$$

Since $|||\tilde{\Phi}^k|||_H < 1$ by assumption in 1), we get,

$$|||\tilde{\Phi}^k|||_H^j < |||\tilde{\Phi}^k|||_H^{(\frac{s}{k}-1)},$$

hence

$$||\Phi^s(X) - \mathrm{tr}(XQ)I_n||_T \leq |||\tilde{\Phi}^k|||_H^{(\frac{s}{k}-1)} \ ||X||_H.$$

Letting $s \to \infty$, we obtain that

$$\Phi^s(X) \to \mathrm{tr}(XQ)I_n$$

as

$$|||\tilde{\Phi}^k|||_H^{(\frac{s}{k}-1)} \to 0$$

This proves 1) $\to$ 2)
The proof of "1) $\to$ 2)" given here is omitted in the proof of theorem 7.7 in [GQ]. In [GQ] it just stated that the theorem 6.1 is applied to the application Φ^k without any further details.

3) $\Leftrightarrow$ 1)
First observe that

$$\Phi^k(x) = \sum_{i_1,\dots,i_k} V_{i_k}^* \dots V_{i_1}^* X V_{i_1} \dots V_{i_k}$$

for all $k \in \mathbb{N}$ where $i_1, \dots, i_k \in \{1, \dots, m\}$. This follows easily from the definition of Φ and an obvious inductive argument. We have also that Φ^k is a Kraus map for all $k \in \mathbb{N}$, since for

$$\Phi^k(I_n) = \Phi^{k-1}(\Phi(I_n)) = \Phi^{k-1}(I_n) = \dots = I_n$$

Then we can apply corollary 9.4 to Φ^k to deduce that $|||\tilde{\Phi^k}|||_H = 1$ if and only if there exists a rank one matrix $Y \in \mathbb{C}^{n\times n}$ s.t. $\mathrm{tr}(V_{i_k}^* \dots V_{i_1}^* V_{j_i} \dots V_{j_k} Y) = 0$ for all $i_1, ..., i_k, j_1, ..., j_k \in \{1, ..., m\}$. Next we observe that for each k, the matrix subspace H_k defined in lemma 9.5, can be described as

$$H_k = \text{span}\{V_{i_k}^* ... V_{i_1}^* V_{j_i} ... V_{j_k} \left| \begin{array}{l} 1 \leq i_1, ..., i_k \leq m \\ 1 \leq i_1, ..., i_k \leq m \end{array} \right\}]$$

This follows easily by an induction argument. Since the trace is linear it follows that tr(XY)=0 for all $X \in H_k$, since

$$\text{tr}(V_{i_k}^* ... V_{i_1}^* V_{j_i} ... V_{j_k} Y) = 0$$

for all $i_1, ..., i_k, j_1, ..., j_k \in \{1, ..., m\}$. Then we have that $Y \in H_k^{\perp} = G_k$. Conversely, if $Y \in G_k$, it follows that

$$\text{tr}(V_{i_k}^* ... V_{i_1}^* V_{j_i ...} V_{j_k} Y) = 0$$

for all $i_1, ..., i_k, j_1, ..., j_k \in \{1, ..., m\}$ as $V_{i_k}^* ... V_{i_1}^* V_{j_i} ... V_{j_k} \in H_k$ for all

$$i_1, ..., i_k, j_1, ..., j_k \in \{1, ..., m\}.$$

Combing all these equivalences, we get the following:

$$|||\tilde{\Phi}^k|||_H = 1$$

if and only if there is a rank one matrix

$$Y \in \mathbb{C}^{n \times n} \text{s.t.tr}(V_{i_k}^* ... V_{i_1}^* V_{j_i} ... V_{j_k} Y) = 0$$

for all $i_1, ..., i_k, j_1, ..., j_k \in \{1, ..., m\}$, which again is true if and only if there is a rank one matrix $Y \in G_k$.
Thus, if

$$|||\tilde{\Phi^k}|||_H = 1 \quad \forall k,$$

we have that for all k there is a rank one matrix

$$Y_k \in \mathbb{C}^{n \times n} \text{ s.t. } Y_k \in G_k.$$

Now, by the consequence of lemma 9.5 there exists $k_0 \leq n^2 - 1$ s.t. $G_{k_0} = G_{k_0 + s}$ for all $s \in \mathbb{N}$. As $G_k \supseteq G_{k+1}$ for all $k \in \mathbb{N}$, we obtain that

$$\bigcap_k G_k = G_{k_0}.$$

If $|||\tilde{\Phi^k}|||_H = 1$ for all k, then in particular $|||\tilde{\Phi^{k_0}}|||_H = 1$. Thus there is a rank one matrix Y_{k_0}

$$\text{s.t. } Y_{k_0} \in G_{k_0} = \bigcap_k G_k.$$

On the other hand, if there is a rank one matrix $Y \in \bigcap_k G_k$, then $Y \in G_k$ for all k and by what we proved, we get that $|||\tilde{\Phi}^k|||_H = 1$ for all k.

3) $\rightarrow$ 4) Here we refer to the proof of "3) $\rightarrow$ 4)" in theorem 7.7 in [GQ].

Recall now that $\Phi^* = \Psi$. The next theorem gives the necesary and sufficient conditions for the convergence of the Markov chain given by

$$\Pi_{k+1} = \Psi(\Pi_k) \quad k = 1, 2, 3, \dots, \Pi_1 \in P(I_n).$$

Theorem 9.7: *The following conditions are equivalent:*

1. *There exists $k \in \mathbb{N}$ s.t. $||\tilde{\Psi}^k||_H^* < 1$.*
2. *The Markov chain given by $\Pi_{k+1} = \Psi(\Pi_k)$ converges to a unique invariant measure regardless of initial distribution.*
3. *The subspace $\cap_k G_k$ does not contain a rank one matrix.*
4. *There exists $k_0 \leq n^2 - 1$ s.t. $||\tilde{\Psi}^{k_0}||_H^* < 1$.*

(2) $\Rightarrow$ (1) Assume that there is an $A \in P(I_n)$ s.t. whenever $\sqcap \in P(I_n)$, then $||\Psi^k(\sqcap) - A||_H^* \to 0$ when $k \to \infty$.

Let $\tilde{\Psi} = \Psi_{|_{M(I_n)}}$. Then $\tilde{\Psi} : M(I_n) \to M(I_n)$. (clearly $\tilde{\Psi}(M(I_n)) \subseteq M(I_n)$ since $\tilde{\Psi}$ is trace preserving).

We have $|||\tilde{\Phi}|||_H = ||\tilde{\Psi}||_H^*$ and since $|||\tilde{\Phi}|||_H \leq 1$ (as we have shown in the proof of theorem 9.6), it follows that $||\tilde{\Psi}||_H^* \leq 1$, so $||\tilde{\Psi}^k||_H^* \leq 1$ for all $k \geq 1$. Hence, if (1) is not true, then $||\tilde{\Psi}^k||_H^* = 1$ for all k, so

$$\lim_{k\to\infty} ||\tilde{\Psi}^k|||_H^{*\frac{1}{k}} = 1.$$

Consider now $M_{\mathbb{C}} = M(I_n) + iM(I_n)$ equipped with $|| \cdot ||_{M_{\mathbb{C}}}^*$ where

$$||R + iQ||_{M_{\mathbb{C}}}^* = \sup_{0 \leq \theta \leq 2\pi} ||R\cos\theta - Q\sin\theta||_H^*$$

for all $R, Q \in M(I_n)$ and let $\tilde{\Psi}_{\mathbb{C}} : M_{\mathbb{C}} \to M_{\mathbb{C}}$ be given by

$$\tilde{\Psi}_{\mathbb{C}}(R + iQ) = \tilde{\Psi}(R) + i\tilde{\Psi}(Q)$$

(as done before in the proof of theorem 9.6).
As we proved, then $||\tilde{\Psi}_{\mathbb{C}}^k||_{M_{\mathbb{C}}}^* = ||\tilde{\Psi}^k||_H^*$ for all k, so

$$\lim_{k\to\infty} ||\tilde{\Psi}_{\mathbb{C}}^k||_{M_{\mathbb{C}}}^{*\frac{1}{k}} = \lim_{k=\infty} ||\tilde{\Psi}^k||_H^{*\frac{1}{k}} = 1.$$

By the spectral radius theorem and again since $M_{\mathbb{C}}$ is finite dimensional, we deduce as before that there exists a $\lambda \in \Pi$ s.t. λ is an eigenvalue for $\tilde{\Psi}_{\mathbb{C}}$. Thus there exists $B \in M_{\mathbb{C}}$, $\;B \neq 0$ and $\theta \in [0, 2\pi)$ s.t

$$\tilde{\Psi_{\mathbb{C}}}(B) = e^{i\theta}B.$$

Hence
$$\tilde{\Psi_{\mathbb{C}}}^k(B) = e^{ik\theta}B \ \ \forall k \geq 1.$$
Write B as $B = C + iD$ where $C, D \in M(I_n)$.
Then we get that $\tilde{\Psi_{\mathbb{C}}}(B) = \tilde{\Psi_{\mathbb{C}}}(C + iD) = \tilde{\Psi}(C) + i\tilde{\Psi}(D) = \Psi(C) + i\Psi(D)$ (since $\tilde{\Psi} = \Psi_{|_{M(I_n)}}$). It is easy to see then that

$$\tilde{\Psi_{\mathbb{C}}}^k(B) = \Psi^k(C) + i\Psi^k(D) \ \ \forall k \geq 1.$$

Hence
$$\tilde{\Psi_{\mathbb{C}}}^k(B) = \Psi^k(C) + i\Psi^k(D) = e^{ik\theta}(C + iD) \ \ \forall k.$$
Identifying the real and imaginary parts in the same way as in the proof of theorem 9.6 (observe that $C, D, \Psi^k(C), \Psi^k(D) \in S_n \ \ \forall k$) we get that

$$\Psi^k(C) = \cos(k\theta)C - \sin(k\theta)D,$$

$$\Psi^k(D) = \sin(k\theta)C + \cos(k\theta)D.$$

Now we recall a result from section 4 which gives that the dual unit ball satisfies
$$B^*_H(I_n) = \{R - Q : R, Q \in P(I_n)\}.$$
Hence, if $L \in M(I_n)$, $L \neq 0$, then there exist $R, Q \in P(I_n)$ s.t. $\frac{1}{||L||^*_H}L = R - Q$. Thus $L = ||L||^*_H(R - Q)$. Also, if $L = 0$, then $||L||^*_H = 0$, so $L = ||L||^*_H(R - Q)$ for any $R, Q \in P(I_n)$. Hence since $\tilde{\Psi} = \Psi_{|_{M(I_n)}}$, we get

$$||\tilde{\Psi}^k(L)||^*_H = ||\tilde{\Psi}^k(||L||^*_H(R - Q)||^*_H = ||L||^*_H||\tilde{\Psi}^k(R - Q)||^*_H$$
$$= ||L||^*_H \ ||\Psi^k(R) - \Psi^k(Q)||^*_H$$
$$\leq ||L||^*_H(||\Psi^k(R) - A||^*_H + ||A - \Psi^k(Q)||^*_H) \to_{k\to\infty} 0$$
since $||\Psi^k(\Pi) - A||^*_H \to 0$ for all $\Pi \in P(I_n)$ by assumption in 2).
So we get that
$$\Psi^k(C) \to 0 \ \ (as\ k \to \infty)$$
and
$$\Psi^k(D) \to 0 \ \ (as\ k \to \infty)$$
as $C, D \in M(I_n)$, and $\Psi_{|_{M(I_n)}} = \tilde{\Psi}^k$.
Thus
$$\Psi^k(C) = (\cos(k\theta)C - \sin(k\theta)D) \to 0 \ \ (as\ k \to \infty)$$
and
$$\Psi^k(D) = (\sin(k\theta)C + \cos(k\theta)D) \to 0 \ \ (as\ k \to \infty)$$

Hence, for each i, j with $1 \leq i, j \leq n$, we have

$$(\cos(k\theta)C_{i,j} - \sin(k\theta)D_{i,j}) \to 0 \ \ (as\ k \to \infty)$$

and

$$(\sin(k\theta)C_{i,j} + \cos(k\theta)D_{i,j}) \to 0 \ \ (as\ k \to \infty).$$

This gives that

$$R_{k\theta} \begin{bmatrix} C_{i,j} \\ D_{i,j} \end{bmatrix} = (R_\theta)^k \begin{bmatrix} C_{i,j} \\ D_{i,j} \end{bmatrix} \to \overrightarrow{0} \ \ (as\ k \to \infty)$$

where

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Since R_θ is a rotation matrix, it follows that $C_{i,j} = D_{i,j} = 0$.
Thus $C = D = 0$, so $B = C + iD = 0$. But we had that $B \neq 0.$, so we get a contradiction.

1) $\Rightarrow$ 2)

Assume that 1) holds. Then there is a $k \in \mathbb{N}$ s.t. $||\tilde{\Psi}^k||_H < 1$. Now we recall the theorem 6.1, which says that if $|||\tilde{T}|||_H = ||S^*||_H^*$ is strictly less than 1, then there is a $\pi \in P(e)$ s.t. $||(T^*)^n(\mu) - \pi||_H^* \leq |||\tilde{T}|||_H^n$ for all n. We apply the theorem 6.1 to the operator $T = \Phi^k$. This is possible since $|||\tilde{\Phi}^k|||_H = ||\tilde{\Psi}||_H^* < 1$. Then $T^* = (\Phi^k)^* = \Psi^k$ in this case and

$$S^* = T^*_{|_{M(I_n)}} = \Psi^k_{|_{M(I_n)}} = \tilde{\Psi}^k,$$

so by theorem 6.1, there is a $Q \in P(I_n)$ s.t. $||\Psi^{kn}(R) - Q||_H^* \leq |||\tilde{\Phi}^k|||_H^n$ for all n and all $R \in P(I_n)$. Since $|||\tilde{\Phi}^k|||_H < 1$, we get that $||\Psi^{kn}(R) - Q||_H^* \to 0$ as $n \to \infty$ for all $R \in P(I_n)$. Furthermore, we have

$$||\Psi^k(R) - \Psi^k(R')||_H^* = \frac{1}{2}||\Psi^k(R) - \Psi^k(R')||_T^*$$

$$= \frac{1}{2}||\Psi^k(R - R')||_T^* = \frac{1}{2}||\tilde{\Psi}^k(R - R')||_1$$

$$\leq ||\tilde{\Psi}^k||_H^* \ ||(R - R')||_H^* = \frac{1}{2}||\tilde{\Psi}^k||_H^* \ ||(R - R')||_1 \ \ \forall R, R' \in P(I_n)$$

where we have used that $(R - R') \in M(I_n)$ for all $R, R' \in P(I_n)$ and

$$\Psi^k_{|_{M(I_n)}} = \tilde{\Psi}^k$$

and that
$$||\cdot||_H^* = \frac{1}{2}||\cdot||_T^* = \frac{1}{2}||\cdot||_1$$
on $M(I_n)$.
So,
$$||\tilde{\Psi}^k(R-R')||_1 \leq ||\tilde{\Psi}^k||_H^*||(R-R')||_1$$
for all $R, R' \in P(I_n)$ and
$$||\tilde{\Psi}^k||_H^* < 1.$$
Also observe that $||\Phi||_T \leq 1$ since $||\Phi(X)||_T \leq ||X||_T$ for all $X = S_n$. Then $||\Psi||_T^* \leq 1$ as $\Psi = \Phi^*$, so $||\Psi(X)||_T^* \leq ||X||_T^*$ for all $X \in S_n$. Hence $||\Psi(X)||_1 \leq ||X||_1$ for all $X \in S_n$ as $||\cdot||_T^* = ||\cdot||_1$
Then we can apply similar arguments as in the proof theorem 7.2 in section 7 part "a)$\Longrightarrow$ b)" to deduce that $\Psi^n(R) \to Q$ as $n \to \infty$ for all $R \in P(I_n)$.

3) $\Leftrightarrow$ 1)

Since $||\tilde{\Psi}^k||_H^* = |||\tilde{\Phi}^k|||_H$ for all k, this equivalence follows from the equivalence 3 $\Leftrightarrow$ 1) in theorem 9.6.

3) $\Leftrightarrow$ 4)

Again, since $||\tilde{\Psi}^{k_0}||_H^* = |||\tilde{\Phi}^{k_0}|||_H$, this equivalence follows from the equivalence 3) $\Leftrightarrow$ 1) in theorem 9.6.
Theorem 9.7 is denoted by theorem 7.8 in [GQ], but the proof is omitted in [GQ].

Example 9.8

Let

$$V_1 = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, V_2 = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$V_3 = \tfrac{1}{2\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, V_4 = \tfrac{1}{2\sqrt{2}}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

Then $\sum_{j=1}^{4} V_j^* V_j = I_2$. Assume now that $u, v \in \mathbb{C}^2$ are s.t.

$$< V_i u, V_j v > = 0$$

for all $i, j \in \{1, ..., 4\}$.
Then, in particular

$$< V_1 u, V_3 v > = \ 0$$

and

$$< V_1 u, V_4 v > = \ 0$$

This gives that

$$u_1(\overline{v}_1 + \overline{v}_2) = 0$$

and

$$u_1(\overline{v}_1 - \overline{v}_2) = 0$$

where

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \ v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

If $u \neq 0$, we must then have

$$\overline{v}_1 + \overline{v}_2 = 0$$

and

$$\overline{v}_1 - \overline{v}_2 = 0$$

Thus we get $v_1 = v_2 = 0$, so v is a zero vector. Hence, if v is not a zero vector, we must have $u_1 = 0$.
By assumption in the beginning, we also have

$$< V_2 u, V_3 v > = 0$$

and

$$< V_2 u, V_4 v > = 0.$$

This gives

$$u_2(\overline{v}_1 + \overline{v}_2) = 0$$

and

$$u_2(\overline{v}_1 - \overline{v}_2) = 0.$$

Hence, again either v is a zero vector on $u_2 = 0$. So, if $< V_i u, V_j v > = \ 0$ for all $i, j \in \{1, ..., 4\}$, then either v is a zero vector or $u_1 = u_2 = 0$, that is u is a zero vector.
We conclude that there are no nonzero vectors
$u, v \in \mathbb{C}^2$ s.t. $< V_i u, V_j v > = \ 0$ for all $i, j \in \{1, ..., 4\}$. If we let $\Phi : S_2 \to S_2$ be given by

$$\Phi(X) = \sum_{j=1}^{4} V_j^* X V_j,$$

by corollary 9.4 it follows that $|||\tilde{\Phi}|||_H \neq 1$. Now, as we have shown in the proof of theorem 9.6, we always have $|||\tilde{\Phi}|||_H \leq 1$. Hence we conclude that in this case $|||\tilde{\Phi}|||_H < 1$, so theorem 6.1 applies.

11 Further research

So far in this thesis, we have applied the theorem 6.1 to stochastic matrices, Markov operators on $C_{\mathbb{R}}(\Omega)$ and Kraus maps acting on S_n. There are many other examples of Markov operators acting on different real Banach spaces, so there are many other cases where we could investigate whether the condition of theorem 6.1 is satisfied, that is whether $|||\tilde{T}|||_H < 1$, for the respective Markov operator T. Theorem 6.1 gives a sufficient condition on $\tilde{T}$ that ensures the convergence of the Markov system given by

$$\mu_{k+1} = T^*\mu_k \quad k = 1, 2, ...$$

to some unique invariant measure. The interesting question is whether there is some weaker condition on T or $\tilde{T}$ that still ensures the convergence of the Markov system to unique invariant measure. In other words, the question is whether it is possible to give a necessary and sufficient condition on T or $\tilde{T}$ that guarantees ergodic property of the Markov system. In the particular case when $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = Ax$ for all $x \in \mathbb{R}^n$ and A is $n \times n$ row stochastic matrix, we proved in section 7 that the corresponding Markov system $x_{k+1} = A^t x_k$, is ergodic if and only if there exists $k_0 \in \mathbb{N}$ s.t.

$$|||\tilde{T}^{k_0}|||_H = \delta(A^{k_0}) < 1.$$

Similarly, in section 9, when$\Phi : S_n \to S_n$ is a Kraus map, we proved that the corresponding Markov system

$$\Pi_{k+1} = \Psi(\Pi_k), \ k = 0, 1, ...$$

(where Ψ is the adjoint of Φ) is ergodic if and only if there exists $k_0 \in \mathbb{N}$ s.t. $|||\tilde{\Phi}^{k_0}|||_H < 1$. One might ask whether this is true in general that the Markov system given by $\mu_{k+1} = T^*\mu_k, \quad k = 0, 1, ...$, is ergodic if and only if there exists $k_0 \in \mathbb{N}$ s.t $|||\tilde{\Phi}^{k_0}|||_H < 1.$. The implication in one direction holds:

Recall that we observed in the proof of theorem 9.6 part "1) $\to$ 2)" that $||T(x)||_T \leq ||x||_T$ for all $x \in X$.
Hence

$$|||\tilde{T}(x+\mathbb{R}e)|||_H = |||T(x)+\mathbb{R}e|||_H = 2\inf_{\lambda\in\mathbb{R}} ||T(x)+\lambda e||_T = 2\inf_{\lambda\in\mathbb{R}} ||T(x+\lambda e)||_T$$

$$\leq 2\inf_{\lambda\in\mathbb{R}} ||(x+\lambda e)||_T = |||x+\mathbb{R}e|||_H$$

for all $x \in X$.
Thus $|||\tilde{\Phi}|||_H \leq 1$ and consequently $||S^*||_H^* \leq 1$ as $|||\tilde{\Phi}|||_H = ||S^*||_H^*$.
Suppose now that there exists $k_0 \in \mathbb{N}$ s.t. $|||\tilde{T}^{k_0}|||H < 1$. Then

$$||(S^*)^{k_0}||H^* = |||\tilde{T}^{k_0}|||_H < 1.$$

Hence

$$||(T^*)^{k_0}(\mu) - (T^*)^{k_0}(v)||_H^* \;=\; ||(S^*)^{k_0}(\mu - v)||_H^* \;\leq\; ||(S^*)^{k_0}||_H^* \; ||(\mu - v)||_H^*$$

for all $\mu, v \in P(e)$.
Furthermore, as $|||\tilde{T}^{k_0}|||_H \; < \; 1$, by theorem 6.1 there exist $\pi \in P(e)$ s.t.

$$||(T^*)^{nk_0}(\mu) - \pi||_H^* \;\leq\; |||\tilde{T}^{k_0}|||_H^n$$

for all $\mu \in P(e)$ and all $n \in \mathbb{N}$. Now, using this, we can show that $(T^*)(\pi) = \pi$ in the same way as we have shown that $Q\pi = \pi$ in the proof of theorem 7.2 part "$a) \; \rightarrow \; b)$.
Next, given $n \in \mathbb{N}$, write n as $n = rk_0 + m$, as we have done in the proof of theorem 7.2 "$a) \; \rightarrow \; b)$".
For any $\mu \in P(e)$, we then have

$$||(T^*)^n(\mu) - \pi||_H^* \;=\; ||(T^*)^n(\mu - \pi)||_H^* \;=\; ||S^*(\mu - \pi)||_H^*.$$

$$= \; ||(S^*)^{rk_0+m}(\mu - \pi)||_H^* \;\leq\; ||(S^*)^m||_H^* \; ||(S^*)^{rk_0}(\mu - \pi)||_H^*$$

$$||(S^*)^*||_H^m \; ||(S^*)^{rk_0}(\mu - \pi)||_H^* \;\leq\; ||(S^*)^{rk_0}(\mu - \pi)||_H^*$$

where we have used that $||S^*||_H^* \;\leq\; 1$, which we observed before.
Then we can proceed as in the proof of theorem 7.2 part "$a) \; \rightarrow \; b)$" to deduce that $(T^*)^n(\mu) \rightarrow \pi$ as $n \rightarrow \infty$.
Hence, this is true in general for any Markov operator T that if there exists $k_0 \in \mathbb{N}$ s.t. $|||\tilde{T}^{k_0}|||_H \; < 1$, then the Markov system given by $\mu_{n+1} = (T^*)^n \mu_n$, $n = 0, 1, ...$ converges to some unique invariant measure regardless of the initial distribution. This is a weaker condition than the condition in theorem 6.1, which is the requirement that $|||\tilde{T}|||_H \; < 1$. However, we are looking for the weakest possible condition that ensures the convergence of the Markov systems, so therefore we wish to find a necessary and sufficient condition. We do not know whether the condition given above is necessary since we didn't prove in general that implication the other way, that is that if the Markov system converges to some unique invariant measure regardless of initial distribution then there exists some $k_0 \in \mathbb{N}$ s.t. $|||\tilde{T}^{k_0}|||_H \; < 1$. We have proved this only in the particular cases in section 7 and section 9.
The interesting question that arises is whether it is possible to give an upper bound for such k_0 to occur. In section 9 we have proved that wen $T = \Phi$ where $\Phi : S_n \rightarrow S_n$ it suffices to consider all k's satisfying $k \leq n^2 - 1$. This fact has followed from the proof of lemma 9.6 where we have shown that the inequality $K_m = K_{m+1}$ can not happen more than n^2 times since the dimension of $\mathbb{C}^{n \times n}$ is equal to n^2 and $K_m \; \subseteq K_{m+1}$ for all m. So we have

used that $\mathbb{C}^{n\times n}$ finite dimensional. In the case when $T : X \to X$ and X is infinite dimensional it is not obvious that we can give an upper bound on such k_0 to occur.
As mentioned in the introduction of the thesis, we are dependent of being able to calculate or give an estimate on $|||\tilde{T}|||_H$ in order to apply theorem 6.1. In this thesis and in [GQ], the theorem 6.2 gives a general expression of $|||\tilde{T}|||_H$, in terms of disjoint, extreme points of $P(e)$. However, using the expression for $|||\tilde{T}|||_H$ from theorem 6.2 is not always the easisest way of calculating $|||\tilde{T}|||_H$. We observe that in example 8.2 we have used the definition of $||S^*||_H^*$ instead of the alternative formula from theorem 6.2 to give an estimate on $||S^*||_H^*$. In section 9, using the formula from theorem 6.2, we have shown that $|||\tilde{\Phi}|||_H \ < 1$ if and only if there are no nonzero vectors $u, v \in \mathbb{C}^n$ s.t. $< V_i, u, V_j v > = \ 0$ for all i, j.
In the examples 9.1, 9.2 and 9.8, it was straightforward to check whether this criteria is satisfied, but on other more complicated examples with higher dimensions, this could be difficult to check. So one may look for some other formulas \ expressions for $|||\tilde{T}|||_H$ that will be more efficient for calculating in concrete examples.
Finally, we observe that most of the concepts and tools we have been used so far in the thesis depends on the choice of K and e. It would be interesting to investigate in concrete, different examples how the change of choice of K and e would reflect on the results.

12 Bibliography

References

[1] [GQ] S. Gaubert, Z. Qu : “Dobrushin‘s ergodicity coefficient for Markov operators on cones“ arXiv:1307.4649v3 [math.FA] 1. Nov 2014.,

[2] [GQ2] : S. Gaubert, Z. Qu : “ Dobrushin‘s ergodicity coefficient for Markov operators on cones“ arXiv:1307.4649v2 [math.FA] 7. Nov 2013.

[3] [ABS] A. B. Svinland : “On stochastic matrices and attractors “ Project for MAT2000, UiO 2012.

[4] [MW] J. N. McDonald, N. A. Weiss : “A course in real analysis“ Second Edition, Academic Press 2013.

[5] [P] G. Pedersen : “Analysis Now“ Revised printing, Springer 1995.

[6] [C] J. B. Conway : “A Course in Functional Analysis “ Second Edition, Springer 1989.