ON THE FREE SEPARATION THEOREM

Master Thesis

by

Manuel Astner

submitted to the Faculty of Mathematics, Computer Science and Physics of the University of Innsbruck

for the academic degree of Diplom-Ingenieur in Technical Mathematics

Supervisor: Univ.-Prof. Dr. Tim Netzer

Innsbruck, November 2020

Abstract

This thesis shows a non commutative analogue of the Hahn-Banach separation theorem for matrix convex sets in infinite and finite dimensional vector spaces. The former case will only be discussed briefly, giving an insight into the separation method of Effros and Winkler. We then investigate the latter case and formulate and prove the statement for matrix convex sets or matrix cones as simple as possible. The results to separate matrix convex sets from a single point are used to give statements regarding free spectrahedral cones or about the separation of two disjoint matrix convex sets.

Contents

1 Introduction

In classical convexity theory the Hahn-Banach separation theorem for closed convex sets in locally convex vector spaces is in itself an interesting fact and forms a useful proof argument for important statements in various fields of mathematics. In recent years free analogues of algebraic geometry and functional analysis were developed, accompanied by non-commutative notions of convexity, such as matrix convexity. Matrix convex sets are collections $K = (K_n)_n$ of subsets of Hermitian $n \times n-$matrices with entries in a locally convex vector space V for all $n \in \mathbb{N}$, conserving a refined convexity property. They occur in different areas, especially in connection with free spectrahedra as solution sets of free linear matrix inequalities [[1], [2], [3]]. However, this non-commutative convexity theory was further extended with the aim to generalize basic results of classical convexity theory to the setting of matrix convex sets [see for example [4], [5]]. For this Edward G. Effros and Soren Winkler showed and used a non-commutative analogue of the Hahn-Banach separation theorem, which also proves to be convenient in the non-commutative theory. While in classical theory the separation of a closed convex set from a single point is based on a continuous linear functional, in the non-commutative setting the separation of a closed matrix convex set $K = (K_n)_n$ from a point $v_0 \in \text{Her}_n(V) \backslash K_n$ becomes possible by a continuous *-linear mapping $\phi : V \to M_n$ with associated mapping $\phi_r : M_r(V) \to M_r(M_n)$ and corresponding Hermitian matrix $a \in \text{Her}_n(V)$, in the sense that $\phi_r|_{K_r} \leq I_r \otimes a$ for all $r \in \mathbb{N}$ but $\phi_n(v_0) \nleq I_n \otimes a$.
We investigate Effros' and Winkler's separation method for closed matrix convex sets and start with the general case of infinite dimensional locally convex vector spaces in chapter 2. Here we introduce some basic concepts and notations we are working with throughout this thesis and give an overview of the proof. In chapter 3 we then go into the case of finite dimensional real vector spaces and simplify the statement and proof starting with matrix cones. Subsequently we shortly present free spectrahedral cones and apply the result to show that every closed matrix cone equals the intersection of all free spectrahedral cones containing it. Then we prove the separation statement for general closed matrix convex sets by reducing it to the cone case. In chapter 4 we also examine criteria for the separation of two disjoint matrix cones or matrix convex sets.

2 Preliminaries

For a detailed insight into the case of infinite dimensional vector spaces we refer to [3], [4] and especially [5]. Here we only want to give a short review in order to get the idea of the proof. So for now let V be a complex vector space with involution $*$ and $V_h := \{v \in V : v^* = v\}$ the real subspace of fixpoints. We write $M_n = M_{n,n} = M_{n,n}(\mathbb{C})$ for the $\mathbb{C}$-algebra of complex $n\times n$-matrices and $M_n(V) = M_{n,n}(V)$ for the corresponding vector space of $n \times n$-matrices with entries in V and canonical involution $[m_{i,j}]^*_{i,j} := [m^*_{j,i}]_{i,j}$. In order to define a matrix convex set in V, we consider for each $n \in \mathbb{N}$ the real subspace of Hermitian matrices $\mathrm{Her}_n(V) = M_n(V)_h$ and subsets $K_n \subseteq \mathrm{Her}_n(V)$, such that the collection $K = (K_n)_n$ is closed under unitary conjugation and forming of direct sums:

Definition 2.0.1.
The collection $K = (K_n)_n$ is a matrix convex set in V, if:

- $\forall v \in K_r, \alpha \in M_{r,n}$ with $\alpha^*\alpha = I_n :\ \alpha^* v\alpha \in K_n$,
- $\forall v \in K_n, w \in K_m :\ v \oplus w \in K_{n+m}$.

If in addition $\alpha^* v\alpha \in K_n$ for any $\alpha \in M_{r,n}$, we say that K is a matrix cone in V.

For the separation theorem 2.0.2 below to apply, the matrix convex set K in V has to be closed. So let V further be a locally convex vector space, or more generally, a vector space with distinguished dual space. Then V is equipped with a topology or weak topology that comes from the dual pairing. For simplicity we assume V to be locally convex and say that $K = (K_n)_n$ in V is closed, if $K_n \subseteq \mathrm{Her}_n(V)$ is closed in the corresponding product topology on $M_n(V)$ for every $n \in \mathbb{N}$. We shall also write $0 \in K$ if $0 \in K_n$ for all $n \in \mathbb{N}$ (or equivalently $0 \in K_r$ for some $r \in \mathbb{N}$).
Given two vector spaces V, W and a linear mapping $\varphi : V \to W$, we obtain for all $n \in \mathbb{N}$ a linear mapping $\varphi_n : M_n(V) \to M_n(W)$ by $\varphi_n([v_{i,j}]_{i,j}) := [\varphi(v_{i,j})]_{i,j}$. Now we can formulate the separation statement:

Theorem 2.0.2.
Let K be a closed matrix convex set in a locally convex vector space V with $0 \in K$. Then for any $v_0 \in Her_n(V)\backslash K_n$ there exists a continuous *-linear mapping $\phi : V \to M_n$, such that:

$$\forall r \in \mathbb{N} : \phi_r|_{K_r} \leq I_{r\times n} \qquad \wedge \qquad \phi_n(v_0) \nleq I_{n\times n}.$$

In particular, for a closed matrix cone K in V and $v_0 \in Her_n(V) \backslash K_n$ there exists such a mapping ϕ with:

$$\forall r \in \mathbb{N} : \phi_r|_{K_r} \leq 0 \qquad \wedge \qquad \phi_n(v_0) \nleq 0.$$

For the proof we shall need an auxiliary lemma providing a suitable state on M_n, and the GNS-method to construct a corresponding *-representation π of M_n on a Hilbert space. For that purpose we remember a linear functional $p : A \to \mathbb{C}$ to be a state of a C^*-algebra A, if it is positive and $p(1) = 1$. In the following we formulate these statements without proof, which can be found in [5] or [6].

Lemma 2.0.3.
Given a matrix convex set K in V with $0 \in K$ and a *-linear function $f : M_n(V) \to \mathbb{C}$ with $f|_{K_n} \leq 1$, there exists a state $p : M_n \to \mathbb{C}$ such that:

$$\forall v \in K_r, \alpha \in M_{r,n} : \ f(\alpha^* v \alpha) \leq p(\alpha^* \alpha).$$

Theorem 2.0.4.
For a state p of a C^*-algebra A there is a *-representation $\pi : A \to \mathcal{B}(H)$ on a Hilbert space H with $e \in H$, such that for all $a \in A$:

$$p(a) = \langle \pi(a)e, e \rangle. \tag{1}$$

Proof of theorem 2.0.2.
Since K is a closed matrix convex set containing the origin in V, we have that $K_n \subseteq \mathrm{Her}_n(V)$ is closed convex and contains 0. The classical Bipolar theorem guarantees the existence of a continuous linear function $f : \mathrm{Her}_n(V) \to \mathbb{R}$ with:

$$f|_{K_n} \leq 1 < f(v_0).$$

Now any matrix $M \in M_n(V)$ may be written as sum $M = A + iB$ of matrices $A, B \in \mathrm{Her}_n(V)$. This determines a unique *-linear continuation $F : M_n(V) \to \mathbb{C}$ since $F(M) = F(A) + iF(B)$. The state p on M_n resulting from lemma 2.0.3 and the normalized trace $\tau : M_n \to \mathbb{C} : \tau(m) := \mathrm{tr}(m) \cdot n^{-1}$ are now used to define a faithful state $q := (1 - \epsilon)p + \epsilon\tau$ on M_n. Here $\epsilon \in (0, 1)$ can be chosen small enough to ensure the following inequalities for $G := (1-\epsilon)F$ and all $v \in K_r, \alpha \in M_{r,n}$:

$$G(\alpha^* v \alpha) \le q(\alpha^* \alpha) \quad \wedge \quad G|_{K_n} \le 1 < G(v_0). \tag{2}$$

These are later used to show the separation properties of ϕ. First, the GNS-construction from q provides a *-representation $\pi : M_n \to \mathcal{B}(H)$ of M_n on a finite dimensional Hilbert space H with separating vector $e \in H$ and property (1). Now for any $a = (a_1, \dots, a_n) \in M_{1,n}$ we define the $n \times n$ matrix:

$$\tilde{a} := \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in M_n.$$

The corresponding subspace $\tilde{M}_{1,n} := \{\tilde{a} : a \in M_{1,n}\} \subseteq M_n$ determines the subspace $\pi_{\tilde{M}_{1,n}} := \{\pi_{\tilde{a}} := \pi(\tilde{a}) : \tilde{a} \in \tilde{M}_{1,n}\} \subseteq \mathcal{B}(H)$, which again determines the n-dimensional subspace $H_0 := \{\pi_{\tilde{a}}(e) : \pi_{\tilde{a}} \in \pi_{\tilde{M}_{1,n}}\} \subseteq H$. By means of G for each $v \in V$ a unique sesquilinearform $\langle \cdot | \cdot \rangle_v : H_0 \times H_0 \to \mathbb{C}$ is defined on H_0:

$$\langle \pi_{\tilde{a}}(e) | \pi_{\tilde{b}}(e) \rangle_v := G(b^* v a).$$

This determines a unique *-linear mapping $\phi_v : H_0 \to H_0$ by the condition:

$$\langle \pi_{\tilde{a}}(e) | \pi_{\tilde{b}}(e) \rangle_v = \langle \phi_v(\pi_{\tilde{a}}(e)), \pi_{\tilde{b}}(e) \rangle.$$

We obtain a continuous *-linear mapping $\phi : V \to \mathcal{B}(H_0) : v \mapsto \phi_v$, where $\mathcal{B}(H_0)$ can be identified with M_n. For arbitrary $v \in \mathrm{Her}_r(V)$ and $\alpha = (\alpha_1^\top, \dots \alpha_r^\top)^\top \in M_{r,n}$ we just receive the suitable representation:

$$G(\alpha^* v \alpha) = \langle \phi_r(v)\eta, \eta \rangle,$$

where $\eta := (\pi_{\tilde{\alpha_1}}(e)^\top, \dots, \pi_{\tilde{\alpha_r}}(e)^\top)^\top \in \mathbb{C}^{rn}$ satisfies $\langle \eta, \eta \rangle = q(\alpha^* \alpha)$. Applying the estimates (2) then yields for all $v \in K_r$ the desired separation property of ϕ:

$$\forall \eta \in \mathbb{C}^{nr} : \quad \langle \phi_r(v)\eta, \eta \rangle = G(\alpha^* v \alpha) \le q(\alpha^* \alpha) = \langle \eta, \eta \rangle.$$

For v_0 we have instead that $\langle \phi_n(v_0)\eta_0, \eta_0 \rangle = G(v_0) > 1 = \langle \eta_0, \eta_0 \rangle$ and the proof is complete.

Given a matrix cone K in V, we have that $\alpha^* v \alpha \in K_n$ for any $v \in K_r, \alpha \in M_{r,n}$ and take advantage of the fact that K_n is a cone. Thus the corresponding linear functional f on $\mathrm{Her}_n(V)$ resulting from the Bipolar theorem satisfies $f|_{K_n} \leq 0 < f(v_0)$. So for sufficient small $\epsilon \in (0,1)$ and the representation for $G = (1-\epsilon)F$ we obtain:

$$\langle \phi_r(v)\eta, \eta \rangle = G(\alpha^* v \alpha) \leq 0 \quad \wedge \quad \langle \phi_r(v_0)\eta_0, \eta_0 \rangle = G(v_0) > 0.$$

□

The theorem also applies to a general closed matrix convex set K that does not contain 0. If we fix an element $\hat{v} \in K_1$, the translation $\hat{K} := (K_n - \hat{v} I_n)_n$ is again a closed matrix convex set in V containing 0. Theorem 2.0.2 provides a contiunous *-linear mapping ϕ such that:

$$\phi_r(v - \hat{v} I_r) = \phi_r(v) - \phi_r(\hat{v} I_r) \leq I_r \otimes I_n \;\Leftrightarrow\; \phi_r(v) \leq I_r \otimes a,$$

with self-adjoint matrix $a := (I_n + \phi(\overline{v})) \in \mathrm{Her}_n(V)$. In the next chapter we consider finite dimensional vector spaces and simplify the proof. The proof steps we make are similar to those above, so some arguments will be clarified later.

3 Separation in finite dimensional vector spaces

Throughout the rest of this thesis, let V be a finite dimensional vector space ($\dim(V) = s \in \mathbb{N}$), which we assume to be real. The case of complex vector spaces can be shown analogously. We want to state and prove theorem 2.0.2 as simple as possible avoiding the use of lemma 2.0.3 and theorem 2.0.4. Therefore we start with some remarks and simplifications that become possible through this special setup and consider matrix cones first. Later we use our result to show the statement for general matrix convex sets by extending them to matrix cones.

3.1 Separation for matrix cones

After fixing a basis in V we may identify it with the s-dimensional euclidean space $\mathbb{R}^s$. Now writing $M_n = M_{n,n}(\mathbb{R})$, we have that $M_n(V) = M_n(\mathbb{R}^s) = M_n^s$. Again we say that a matrix convex set $K = (K_n)_n$ in V is closed, if each K_n is closed in the euclidean topology on $M_n^s = \mathbb{R}^{sn^2}$. However, K is now a collection of subsets $K_n \subseteq M_n(V)_h = \mathrm{Sym}_n^s$ satisfying $\alpha^\top K_r \alpha \subseteq K_n$ and $K_n \oplus K_m \subseteq K_{n+m}$ for every $\alpha \in M_{r,n}$ with $\alpha^\top \alpha = I_n$ and $r, n, m \in \mathbb{N}$. As in the infinite dimensional case the convexity of all sets K_n follows directly from these properties, which can easily seen by writing an arbitrary convex combination with elements $v, w \in K_n$ and $\lambda \in [0,1]$ as:

$$\lambda v + (1-\lambda)w = (\sqrt{\lambda}\cdot I_n, \sqrt{1-\lambda}\cdot I_n)\overbrace{(v \oplus w)}^{\in K_{2n}}\overbrace{(\sqrt{\lambda}\cdot I_n, \sqrt{1-\lambda}\cdot I_n)^\top}^{\in M_{2n,n}} \in K_n.$$

For a matrix cone $K = (K_n)_n$ in V follows that all sets K_n are cones, since for any $v \in K_n$ and $\lambda \geq 0$ we have in addition that:

$$\lambda v = (\sqrt{\lambda}\cdot I_n)^\top v\,\overbrace{(\sqrt{\lambda}\cdot I_n)}^{\in M_n} \in K_n.$$

So given a closed matrix cone $K = (K_n)_n$ in V and $v_0 \in \mathrm{Sym}_n^s \backslash K_n$, all requirements for K_n are met in order to apply the Bipolar theorem 3.1.2, which is a simple consequence of the classical Hahn-Banach separation theorem [7]. In this regard we recall the definition of the bipolar of a set:

Definition 3.1.1.
For a real vector space V with algebraic dual space V' and subsets $A \subseteq V, B \subseteq V'$ we define:

- the polar $A^\circ := \{f \in V' : f|_A \leq 1\} \subseteq V'$ of A,
- the prepolar ${}^\circ B := \{v \in V : f(v) \leq 1, \forall f \in B\} \subseteq V$ of B,
- the bipolar ${}^\circ(A^\circ) := \{v \in V : f(v) \leq 1, \forall f \in A^\circ\} \subseteq V$ of A.

Theorem 3.1.2 (classical Bipolar theorem)**.**
For any subset $M \subseteq \mathbb{R}^n$ containing 0, the bipolar ${}^\circ(M^\circ)$ is given by:

$${}^\circ(M^\circ) = \overline{\mathrm{conv}(M)}.$$

Theorem 3.1.3.
For a closed matrix cone K in V and $v_0 \in Sym_n^s \backslash K_n$, there exists a linear mapping $\phi : V \to Sym_n$ such that:

$$\forall r \in \mathbb{N} : \phi_r|_{K_r} \leq 0 \qquad \wedge \qquad \phi_n(v_0) \nleq 0.$$

Proof.
We have already seen that given a closed matrix cone K in V, the set K_n is a closed cone containing 0. Theorem 3.1.2 states that $K_n = {}^\circ(K_n^\circ)$ and thus $v_0 \notin {}^\circ(K_n^\circ)$. So there must be a linear functional $F : Sym_n^s \to \mathbb{R} \in K_n^\circ \subseteq (\mathrm{Sym}_n^s)'$ satisfying $F|_{K_n} \leq 1 < F(v_0)$. Since K_n is a cone, we even have that for all $v \in K_n$ and $\lambda \geq 0$:

$$F(\lambda v) = \lambda F(v) \leq 1 \Leftrightarrow F(v) \leq 0. \tag{3}$$

With the identifications above, elements $v \in V$ or $A \in M_n(V)_h$ correspond to s−tuples $(v_1, \dots, v_s)^\top \in \mathbb{R}^s$ or $(A_1, \dots, A_s)^\top \in \mathrm{Sym}_n^s$. Further the euclidean vector space Sym_n^s with standard scalar product $\langle \cdot, \cdot \rangle$ can be identified with its dual space $(\mathrm{Sym}_n^s)'$, using the mapping $\mathrm{Sym}_n^s \to (\mathrm{Sym}_n^s)' : A \mapsto \langle \cdot, A \rangle$. This means for F, that there exists $N \in \mathrm{Sym}_n^s$ with:

$$F(M) = \langle M, N \rangle = \langle (M_1, \dots, M_s)^\top, (N_1, \dots, N_s)^\top \rangle = \sum_{k=1}^{s} \langle M_k, N_k \rangle = \sum_{k=1}^{s} \mathrm{tr}(M_k N_k).$$

Now we use N to define an obviously linear mapping $\phi : V \to \mathrm{Sym}_n$ by:

$$\phi(v) = \phi((v_1, \dots, v_s)^\top) := \sum_{k=1}^{s} v_k N_k.$$

Any matrix $\alpha \in M_{r,n}$ may be written as $\alpha = (\alpha_1^\top, \dots, \alpha_r^\top)^\top$ with matrices $\alpha_i \in M_{1,n}$ for all $i \in \{1, \dots, r\}$. So given $v = [v_{i,j}]_{i,j} \in \mathrm{Sym}_r^s$, the matrix $\alpha^\top v \alpha \in \mathrm{Sym}_n^s$ corresponds to the r^2-fold sum:

$$\alpha^\top v \alpha = \sum_{i=1}^{r} \sum_{j=1}^{r} \alpha_i^\top v_{i,j} \alpha_j = \sum_{i,j}^{r} \alpha_i^\top \alpha_j v_{i,j}.$$

With this expression and using the mapping ϕ, we obtain for F:

$$
\begin{aligned}
F(\alpha^\top v\alpha) &= F(\sum_{i,j}^{r} \alpha_i^\top \alpha_j v_{i,j}) = \sum_{i,j}^{r} F(\alpha_i^\top \alpha_j v_{i,j}) = \sum_{i,j}^{r}\sum_{k=1}^{s} \operatorname{tr}((\alpha_i^\top \alpha_j v_{i,j})_k N_k) = \\
&= \sum_{i,j}^{r}\sum_{k=1}^{s} \operatorname{tr}((\alpha_i^\top \alpha_j)(v_{i,j})_k N_k) = \sum_{i,j}^{r} \operatorname{tr}((\alpha_i^\top \alpha_j)\sum_{k=1}^{s}(v_{i,j})_k N_k) = \\
&= \sum_{i,j}^{r} \operatorname{tr}((\alpha_i^\top \alpha_j)\phi(v_{i,j})) = \sum_{i,j}^{r} \operatorname{tr}(\phi(v_{i,j})(\alpha_i^\top \alpha_j)^\top) = \sum_{i,j}^{r} \operatorname{tr}(\phi(v_{i,j})\alpha_j^\top \alpha_i).
\end{aligned}
$$

One can easily check that the trace of $\phi(v_{i,j})\alpha_j^\top \alpha_i$ equals the product $\alpha_i \phi(v_{i,j})\alpha_j^\top$. This further yields the following representation for F:

$$
\begin{aligned}
F(\alpha^\top v\alpha) &= \sum_{i=1}^{r}\sum_{j=1}^{r} \alpha_i \phi(v_{i,j})\alpha_j^\top = \sum_{j=1}^{r} \alpha_1 \phi(v_{1,j})\alpha_j^\top + \ldots + \sum_{j=1}^{r} \alpha_r \phi(v_{r,j})\alpha_j^\top = \\
&= (\alpha_1, \ldots, \alpha_r)\begin{pmatrix} \sum_{j=1}^{r} \phi(v_{1,j})\alpha_j^\top \\ \vdots \\ \sum_{j=1}^{r} \phi(v_{r,j})\alpha_j^\top \end{pmatrix} = (\alpha_1, \ldots, \alpha_r)[\phi(v_{i,j})]_{i,j}\begin{pmatrix} \alpha_1^\top \\ \vdots \\ \alpha_r^\top \end{pmatrix} = \langle \phi_r(v)\eta, \eta\rangle,
\end{aligned}
$$

where the vector $\eta := (\alpha_1, \ldots, \alpha_r)^\top \in \mathbb{R}^{rn}$ is dependent of α. Because of $\alpha^\top v\alpha \in K_n$ for any $v \in K_r, \alpha \in M_{r,n}$ due to the property of the matrix cone K, the following inequality, given by (3), holds for all $\eta \in \mathbb{R}^{rn}$:

$$
\langle \phi_r(v)\eta, \eta\rangle = F(\alpha^\top v\alpha) \leq 0 \;\Leftrightarrow\; \eta^\top \phi_r(v)\eta \leq 0.
$$

This just means that $\phi_r(v)$ is negative semidefinite and since $v \in K_r$ was arbitrary, we have that $\phi_r|_{K_r} \leq 0$ for all $r \in \mathbb{R}$. In contrast, for $v_0 \in \mathrm{Sym}_n^s \backslash K_n$ and $\alpha = I_n$ with $\alpha_i = e_i^\top$ for $i \in \{1, \ldots, n\}$, we see that there is at least one vector $\eta_0 := (e_1^\top, \ldots, e_n^\top)^\top \in \mathbb{R}^{nr}$ such that:

$$
\langle \phi_n(v_0)\eta_0, \eta_0\rangle = F(I_n^\top\, v_0\, I_n) > 0 \;\Leftrightarrow\; \eta_0^\top\, \phi_n(v_0)\eta_0 > 0.
$$

Accordingly, $\phi_n(v_0)$ can not be negative semidefinite. □

Now we take a look at free spectrahedra, or more precisely, free spectrahedral cones. These are important examples of closed matrix cones and we can apply theorem 3.1.3. For a closer look into the theory of free spectrahedra, their properties and applications see for example [1], [2] or [3].

Definition 3.1.4.
Given $n \in \mathbb{N}$ and $M = (M_1, \ldots, M_s)^\top \in \mathrm{Sym}_n^s$ we define for every $r \in \mathbb{N}$ the set:

$$S(M)_r := \{(A_1, \ldots, A_s)^\top \in \mathrm{Sym}_r^s : \sum_{k=1}^{s} (M_k \otimes A_k) \geq 0\} \subseteq \mathrm{Sym}_r^s,$$

and call the collection $S(M) = (S(M)_r)_r$ the free spectrahedral cone defined by M.

It can easily be seen that free spectrahedral cones are closed matrix cones. ϕ constructed as above satisfies the following for $v \in \mathrm{Sym}_r^s$ and all $r \in \mathbb{N}$:

$$\phi_r(v) = [\phi(v_{i,j})]_{i,j} = [\sum_{k=1}^{s} (v_{i,j})_k N_k]_{i,j} = \sum_{k=1}^{s} [(v_{i,j})_k N_k]_{i,j} = \sum_{k=1}^{s} v_k \otimes N_k. \tag{4}$$

So with the separation result above one can see that every closed matrix cone is contained in a free spectrahedral cone, which results in the following corollary. Here we write $K \subseteq L$ (or L contains K) for two collections $K = (K_n)_n, L = (L_n)_n$ again if $K_n \subseteq L_n$ stepwise for all $n \in \mathbb{N}$.

Corollary 3.1.5.
Every closed matrix cone is the intersection of all free spectrahedral cones containing it.

Proof.
K is obviously subset of $S := \bigcap_{K \subseteq S(M), M \text{ arbitrary}} S(M)$. In contrast if we assume that there is a point $v_0 \in S_n$ with $v_0 \notin K_n$, theorem 3.1.3 yields a linear mapping $\phi : V \to \mathrm{Sym}_n : v \mapsto \sum_{k=1}^{s} v_k N_k$ such that:

$$\forall r \in \mathbb{N} : \phi_r|_{K_r} \geq 0 \quad \wedge \quad \phi_n(v_0) \ngeq 0.$$

By (4) this means that $\sum_{k=1}^{s} N_k \otimes v_k \geq 0$ for all $v \in K_r, r \in \mathbb{N}$ and $\sum_{k=1}^{s} N_k \otimes (v_0)_k \ngeq 0$. This shows $K \subseteq S(N)$ and $v_0 \notin S(N)$. Since $v_0 \in S_n$ by assumption, it must be contained in $S(N)$ as we have $K \subseteq S(N)$, a contradiction. □

3.2 Separation for matrix convex sets

Now, given a general closed matrix convex set $K = (K_n)_n$ in V, the idea is to extend elements $(v_1, \dots, v_s)^\top \in K_n$ to $s+1$–tuples $(v_1, \dots, v_s, -I_n)^\top \in \mathbb{R}^{s+1}$ and the corresponding matrix convex set to a matrix cone in $\mathbb{R}^{s+1}$. We can then apply theorem 3.1.3 in order to obtain the following statement:

Theorem 3.2.1.
For a closed matrix convex set K in V and $v_0 \in Sym_n^s \backslash K_n$, there exists a linear mapping $\phi : V \to Sym_n$ and $a \in Sym_n$ such that:

$$\forall r \in \mathbb{N} : \phi_r|_{K_r} \leq I_r \otimes a \qquad \wedge \qquad \phi_n(v_0) \nleq I_n \otimes a.$$

Proof.
For elements $v = (v_1, \dots, v_s)^\top \in K_r$ we consider s+1-tuples $\hat{v} := (v, -I_r)^\top = (v_1, \dots, v_s, -I_r)^\top$ and define the set $\hat{K}_r := \{\hat{v} : v \in K_r\} \subseteq \mathrm{Sym}_r^{s+1}$. In order to extend the matrix convex set $\hat{K} = (\hat{K}_n)_n$ to a matrix cone in $\mathbb{R}^{s+1}$ we define $\tilde{K}_r := \{\alpha^\top \hat{v} \alpha : \hat{v} \in \hat{K}_m, \alpha \in M_{m,r}, m \in \mathbb{N}\} \subseteq \mathrm{Sym}_r^{s+1}$. Thus the collection $\tilde{K} = (\tilde{K}_n)_n$ satisfies the following for any $\alpha^\top \hat{v} \alpha \in \tilde{K}_r, \gamma^\top \hat{w} \gamma \in \tilde{K}_t$ and $\beta \in M_{r,n}$:

- $\beta^\top (\alpha^\top \hat{v} \alpha) \beta = (\alpha\beta)^\top \hat{v} \overbrace{\alpha\beta}^{\in M_{m,n}} \in \tilde{K}_n$.
- $(\alpha^\top \hat{v} \alpha) \oplus (\gamma^\top \hat{w} \gamma) = (\alpha \oplus \gamma)^\top (\hat{v} \oplus \hat{w}) \overbrace{(\alpha \oplus \gamma)}^{\in M_{m+p,r+t}} \in \tilde{K}_{r+t}$.

In particular the point $\hat{v}_0$ cannot be contained in the closure of $\tilde{K}_n$. To see this consider a sequence $(\tilde{v}_k)_k$ in $\tilde{K}_n$ with limit $\lim_{k \to \infty} \tilde{v}_k = \hat{v}_0 = (v_0, -I_n)^\top$. Since every $\tilde{v}_k$ is of the form $(\alpha_k^\top v_k \alpha_k, -\alpha_k^\top \alpha_k)^\top$ we see that $\alpha_k^\top \alpha_k$ converges to I_n, which means that the columns of α_k converge to an orthonormal basis. Now the Gram-Schmidt orthonormalization applied to the columns of α_k yields an orthorormal basis and corresponding unitary matrix $\dot{\alpha}_k$ (whose columns form this orthonormal basis). With increasing k the difference between α_k and $\dot{\alpha}_k$ decreases and we obtain the sequence $(\dot{\alpha}_k^\top \hat{v}_k \dot{\alpha}_k)_k$ converging towards $\hat{v}_0$. By the matrix convexity property $(\dot{\alpha}_k^\top v_k \dot{\alpha}_k)_k$ is a sequence in K_n with limit v_0. But K_n is closed, which leads to a contradiction since $v_0 \notin K_n$.
So applying theorem 3.1.3 to the closure of $\tilde{K}$ yields a linear functional $F : \mathrm{Sym}_n^{s+1} \to \mathbb{R}$ and a linear mapping $\tilde{\phi} : \mathbb{R}^{s+1} \to \mathrm{Sym}_n$ with the following properties for arbitrary $\tilde{v} \in$

$\tilde{K}_r, \beta \in M_{r,n}$ and corresponding $\eta := (\beta_1, \dots, \beta_r)^\top \in \mathbb{R}^{rn}$:

$$F(\beta^\top \tilde{v} \beta) = \langle \tilde{\phi}_r(\tilde{v})\eta, \eta\rangle \leq 0 \qquad \wedge \qquad F(\hat{v}_0) > 0. \tag{5}$$

$\tilde{\phi}$ is given by $\tilde{\phi}(v) := \sum_{k=1}^{s+1} v_k N_k$, where $N = (N_1, \dots, N_{s+1})^\top \in \mathrm{Sym}_n^{s+1}$ stems from the identification $F(\cdot) = \langle \cdot, N\rangle$. We define $\phi : V \to \mathrm{Sym}_n : \phi(v) := \sum_{k=1}^{s} v_k N_k$ and obtain for any $\hat{v} \in \hat{K}_r \subseteq \tilde{K}_r$:

$$\begin{aligned} F(\beta^\top \hat{v} \beta) &= \sum_{i,j}^{r} F((\beta_i^\top \beta_j v_{i,j}, \beta_i^\top \beta_j (-I_r)_{i,j})) = \sum_{i,j}^{r} \Big(\sum_{k=1}^{s} \langle \beta_i^\top \beta_j (v_{i,j})_k, N_k\rangle - \langle \beta_i^\top \beta_j (I_r)_{i,j}, N_{s+1}\rangle\Big) \\ &= \langle \phi_r(v)\eta, \eta\rangle - \sum_{i=1}^{r} \langle \beta_i^\top \beta_i, N_{s+1}\rangle = \langle \phi_r(v)\eta, \eta\rangle - \mathrm{tr}(\beta^\top \beta \cdot N_{s+1}). \end{aligned}$$

One can verify that the trace of the matrix $\beta^\top \beta \cdot N_{s+1}$ corresponds to $\langle (I_r \otimes N_{s+1})\eta, \eta\rangle$. Togehther with (5) this results in:

$$F(\beta^\top \hat{v} \beta) = \langle \phi_r(v)\eta, \eta\rangle - \langle (I_r \otimes N_{s+1})\eta, \eta\rangle \leq 0 \;\Leftrightarrow\; \langle (\phi_r(v) - I_r \otimes N_{s+1})\eta, \eta\rangle \leq 0.$$

Since $\hat{v} \in \hat{K}_r$ and $\beta \in M_{r,n}$ was arbitrary, the inequality holds for all $v \in K_r$ and $\eta \in \mathbb{R}^{rn}$, which just means that $\phi_r|_{K_r} \leq I_r \otimes N_{s+1}$ for all $r \in \mathbb{N}$. For the point $\hat{v}_0$ there is again a vector $\eta_0 := (e_1^\top, \dots, e_n^\top)^\top \in \mathbb{R}^{nr}$ with:

$$F(\hat{v}_0) = \langle \phi_n(v_0)\eta_0, \eta_0\rangle - \langle (I_n \otimes N_{s+1})\eta_0, \eta_0\rangle > 0,$$

and thus $\phi_n(v_0) \nleq I_n \otimes N_{s+1}$. □

4 Separation of two matrix convex sets

In this section we concern ourself with the separation of two disjoint matrix convex sets $K = (K_n)_n$ and $L = (L_n)_n$ in V, which means that $K_r \cap L_r = \varnothing$ for some $r \in \mathbb{N}$, or equivalently $K_n \cap L_n = \varnothing$ for all $n \in \mathbb{N}$. To see this, just assume $K_r \cap L_r = \varnothing$ and $v \in K_n \cap L_n$ for some $n \in \mathbb{N}$. Given $\alpha \in M_{n,r}$ satisfying $\alpha^\top \alpha = I_r$, we have that $\alpha^\top v \alpha \in K_r$ and $\alpha^\top v \alpha \in L_r$ due to the property of the matrix convex sets K and L, which leads to a contradiction. Again we start with the simpler case of disjoint matrix cones in V that in particular only intersect at the origin, so $K_n \cap L_n = \{0\}$ for all $n \in \mathbb{N}$. A corresponding separation is given by the following well known hyperplane separation theorem (for a proof see [8]):

Corollary 4.0.2.
For two disjoint convex subsets $A, B \subseteq \mathbb{R}^s$ there exists a linear functional $f : \mathbb{R}^s \to \mathbb{R}$, such that $f|_A \leq f|_B$.

Theorem 4.0.3.
Given two disjoint matrix cones $K = (K_n)_n$ and $L = (L_n)_n$ in V, there exists a linear mapping $\phi : V \to Sym_n$ such that:

$$\forall r \in \mathbb{N}: \ \phi_r|_{K_r} \leq 0 \leq \phi_r|_{L_r}.$$

Proof.
Since both subsets $K_n, L_n \subseteq \mathrm{Sym}_n^s$ are convex and $K_n \cap L_n = \{0\}$, we can apply corollary 4.0.2 to obtain a linear functional $F : \mathrm{Sym}_n^s \to \mathbb{R}$ with:

$$F|_{K_n} \leq 0 \leq F|_{L_n}. \tag{6}$$

The matrix $N \in \mathrm{Sym}_n^s$ satisfying $F(\cdot) = \langle \cdot, N \rangle$ is used again to define $\phi : V \to \mathrm{Sym}_n : v \mapsto \sum_{i=1}^{s} v_i N_i$ and to provide the following representation for all $v \in \mathrm{Sym}_r^s$ and $\alpha \in M_{r,n}$ with corresponding vector $\eta \in \mathbb{R}^{rn}$:

$$F(\alpha^\top v \alpha) = \langle \phi_r(v)\eta, \eta \rangle.$$

With (5) and due to the property of the matrix cones K and L, for arbitrary elements $v \in K_r, w \in L_r$ and $\alpha \in M_{r,n}$ we have that:

$$F(\overbrace{\alpha^\top v \alpha}^{\in K_n}) = \langle \phi_r(v)\eta, \eta \rangle \leq \langle \phi_r(w)\eta, \eta \rangle = F(\overbrace{\alpha^\top w \alpha}^{\in L_n}),$$

which overall proves the statement. □

In order to obtain a separation statement also for general disjoint matrix convex sets $K = (K_n)_n$ and $L = (L_n)_n$ in V, we have to tightening the requirements for K and L.

One of them has to be closed and the other compact (which is again to be understood step wise for all sets K_n). The idea is to consider the difference $\overline{K} \coloneqq K - L$ and apply theorem 3.2.1, which yields the following statement:

Theorem 4.0.4.
Given two disjoint matrix convex sets $K = (K_n)_n$ and $L = (L_n)_n$ in V, where K is closed and L is compact, there exists a linear mapping $\phi : V \to Sym_n$ and $a \in Sym_n$ with $a \ngeq 0$, such that:

$$\forall r \in \mathbb{N} : \ \phi_r|_{K_r} \leq I_r \otimes a + \phi_r|_{L_r}.$$

Proof.
Consider the difference $\overline{K} \coloneqq K - L = (\overline{K}_n)_n$, where $\overline{K}_n \coloneqq K_n - L_n = \{v - w : v \in K_n, w \in L_n\}$. This is again a matrix convex set in V since for arbitrary $\overline{v} = (v - w) \in \overline{K}_r, \overline{z} = (x - y) \in \overline{K}_s$ and $\alpha \in M_{r,n}$ with $\alpha^\top \alpha = I_n$ we have that:

$$\alpha^\top \overline{v} \alpha = \alpha^\top v \alpha - \alpha^\top w \alpha \in \overline{K}_n \qquad \wedge \qquad \overline{v} \oplus \overline{z} = v \oplus x - w \oplus y \in \overline{K}_{r+s}.$$

The assumption $K_n \cap L_n = \varnothing$ implies that $v \neq w$ for all $v \in K_n, w \in L_n$ and thus $0 \notin \overline{K}_n$. Further the difference $\overline{K}_n$ of a compact set K_n and a closed set L_n is again closed for all $n \in \mathbb{N}$. Applying theorem 3.2.1 to the closed matrix convex set $\overline{K}$ in V and the point $0 \notin \overline{K}_n$ then guarantees a linear mapping $\phi : V \to \mathrm{Sym}_n$ and a matrix $a \in \mathrm{Sym}_n$ satisfying:

$$\forall r \in \mathbb{N} : \ \phi_r|_{\overline{K}_r} \leq I_r \otimes a \quad \wedge \quad \phi_n(0) = 0 \nleq I_n \otimes a.$$

$I_n \otimes a \ngeq 0$ implies $a \ngeq 0$ and for arbitrary $\overline{v} \in K_r$ we have that $\phi_r(\overline{v}) = \phi_r(v) - \phi_r(w) \leq I_r \otimes a$ and thus $\phi_r|_{K_r} \leq I_r \otimes a + \phi_r|_{L_r}$ for all $r \in \mathbb{N}$.

□

5 Conclusion

In both cases of infinite and finite dimensional vector spaces the method remains the same: The separation of a closed matrix convex set $K = (K_n)_n$ in V and $v_0 \in \mathrm{Her}_n(V) \setminus K_n$ is traced back to the classical Hahn-Banach separation theorem, which provides a continuous linear functional f on $\mathrm{Her}_n(V)$ separating K_n and v_0. This (or more generally its *-linear continuation F) is used to construct a continuous *-linear mapping $\phi : V \to M_n$, whose separation properties result from the matrix convexity properties of K and the representation of f (or $G := (1-\epsilon)F$) via ϕ together with the inequalities holding for f (or G).
In infinite dimensional vector spaces the case of general closed matrix convex sets follows by an easy translation argument from the case of closed matrix convex sets containing 0. Here f can be extended to a unique *-linear functional F on $M_n(V)$, which comes with a state p on M_n such that overall for all $v \in K_r, \alpha \in M_{r,n}$ the following applies:

$$F(\alpha^* v \alpha) \leq p(\alpha^* \alpha) \qquad \wedge \qquad F|_{K_n} \leq 1 < F(v_0). \tag{7}$$

Together with the normalized trace τ one can define the faithful state $q := (1-\epsilon)p + \epsilon\tau$ on M_n, such that for sufficiently small $\epsilon \in (0,1)$ the estimates also apply for $G := (1-\epsilon)F$ and q. The GNS-construction yields a corresponding *–representation π of M_n on a Hilbert space H with separating vector $e \in H$. These are used to determine a n-dimensional subspace $H_0 \subseteq H$ on which a sesquilinearform is defined for every $v \in V$. This way one obtains *-linear mappings $\phi_v : H_0 \to H_0 \in \mathcal{B}(H_0) \cong M_n$ with associated *-linear mapping $\phi : V \to M_n : v \mapsto \phi_v$ and representations $G(\alpha^* v \alpha) = \langle \phi_r(v)\eta, \eta\rangle$ and $q(\alpha^*\alpha) = \langle \eta, \eta \rangle$. By the matrix convexity properties of K, (7) holds for every vector $\eta \in \mathbb{C}^{nr}$, which yields the corresponding separation properties of ϕ.
The finite dimensional real case simplifies vastly by identifying V with $\mathbb{R}^s$ and $\mathrm{Her}_n(V) = \mathrm{Sym}_n^s$ with its dual space. Thus f corresponds to a matrix $N \in \mathrm{Sym}_n^s$, which can be used to define the linear mapping $\phi : \mathbb{R}^s \to \mathrm{Sym}_n : v \mapsto \sum_{k=1}^{s} v_k N_k$. The corresponding representation $f(\alpha^\top v \alpha) = \langle \phi_r(v)\eta, \eta \rangle$ can be verified relatively easy and the claim for closed matrix cones follows directly due to their special cone properties. Almost all further evidence is based on this separation theorem 3.1.3 for closed matrix cones. By looking at free spectrahedral cones for example, it shows that every closed matrix cone K in $\mathbb{R}^s$ is contained in such a solution set of a free matrix inequality, which can be seen by the identification $\phi_r(v) = \sum_{k=1}^{s} v_k \otimes N_k$. More precisely, by a simple contradiction argument, theorem 3.1.3 implies that K equals the intersection of all free spectrahedral cones over-

lying it. Also the separation statement for a general closed matrix convex set K in $\mathbb{R}^s$ and $v_0 \notin K_n$ goes back to theorem 3.1.3. One can extend elements $v \in \mathrm{Sym}_n^s$ with $-I_r$ to $(v_1, \ldots, v_s, -I_r)^\top$ and consider the corresponding homogenization $\hat{K}$ in $\mathbb{R}^{s+1}$. Forming its matrix cone hull yields a matrix cone $\tilde{K}$ in $\mathbb{R}^{s+1}$, whose closure cannot contain $\hat{v}_0$. Theorem 3.1.3 guarantees the separation of $\tilde{K}$ and v_0 by a linear mapping $\tilde{\phi} : \mathbb{R}^{s+1} \to \mathrm{Sym}_n$. From this one obtains the linear mapping ϕ as above with corresponding separation properties on K and v_0.
The separation of two disjoint matrix cones K, L in V underlies the fact that any two disjoint convex subsets $K_n, L_n \subseteq \mathrm{Her}_n(V)$ may be separated by a linear functional f on $\mathrm{Her}_n(V)$. In the same manner as above it can be used to construct ϕ, separating K and L due to their matrix cone properties. In the general case of two disjoint matrix convex sets one can consider the difference $\overline{K} := K - L$, which can be guaranteed to be a closed matrix convex set in V not containing the origin, if K is closed and L compact. The separation statement for closed matrix convex sets yields separating ϕ with corresponding separation properties on K and L.
Altogether the thesis shows a useful analogue of the classical Hahn-Banach separation theorem with significant simplifications in the finite dimensional case. It appears in several works dealing with non-commutative versions of interesting results from classical convexity theory, which rely on this Effros-Winkler separation theorem. Thus it plays a central role in rising non-commutative algebra or functional analysis and motivates to further research in this area.

Acknowledgments

Primary i wish to thank my supervisor Univ.-Prof. Dr. Tim Netzer, who supported me throughout this master's thesis and was always a great help with suggestions for improvement or approaches when facing problems. I would also like to thank my fellow students and friends who accompanied me during my studies. Special thanks go to my family and especially to my parents who made my studies possible in the first place.

References

[1] J. William Helton, Igor Klep, and Scott McCullough. The tracial hahn–banach theorem, polar duals, matrix convex sets, and projections of free spectrahedra. 19:1845–1897, 2017.

[2] Tom-Lukas Kriel. An introduction to matrix convex sets and free spectrahedra. *Complex Analysis and Operator Theory*, 13(7):3251–3335, 2019.

[3] Tobias Fritz, Tim Netzer, and Andreas Thom. Spectrahedral containment and operator systems with finite-dimensional realization. *arXiv e-prints*, page arXiv:1609.07908, September 2016.

[4] Webster Corran and Winkler Soren. The krein-milman theorem in operator convexity, January 1999.

[5] Edward G. Effros and Soren Winkler. Matrix convexity: Operator analogues of the bipolar and hahn–banach theorems. 144:117–152, 1997.

[6] Martin Bordemann and Stefan Waldmann. Formal gns construction and states in deformation quantization. July 1996.

[7] Frank Deutsch, Hein Hundal, and Ludmil Zikatanov. Some applications of the hahn-banach separation theorem. December 2017.

[8] Robert Peng. The hahn-banach separation theorem and other separation results. 2014.

universität innsbruck

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

Datum	Unterschrift