An Infinite Dimensional Version of the Schur–Horn Convexity Theorem

Andreas Neumann

Fachbereich Mathematik AG 5, Technische Universität Darmstadt, Schloßgartenstraße 7, 64289 Darmstadt, Germany
E-mail: neumann.mathematik.tu-darmstadt.de

Received March 4, 1998; revised July 10, 1998; accepted July 20, 1998

The Schur–Horn Convexity Theorem states that for \( a \in \mathbb{R}^n \)

\[
p(\{ U^* \text{diag}(a) U : U \in \mathcal{U}(n) \}) = \text{conv}(\mathcal{S}_n a),
\]

where \( p \) denotes the projection on the diagonal. In this paper we generalize this result to the setting of arbitrary separable Hilbert spaces. It turns out that the theorem still holds, if we take the \( l^\infty \)-closure on both sides. We will also give a description of the left-hand side for nondiagonalizable hermitian operators. In the last section we use this result to get an extension theorem for invariant closed convex subsets of the diagonal operators.

1. INTRODUCTION

In [Ho54] Horn proved that for every \( x \in \mathbb{R}^n \) we have that

\[
p(\{ U^* \text{diag}(x) U : U \in \mathcal{U}(n) \}) = \text{conv}(\mathcal{S}_n x),
\]

where \( p \) denotes the projection on the diagonal, \( \mathcal{U}(n) \) is the group of unitary \( n \times n \) matrices, and \( \mathcal{S}_n \) is the group of permutations of \( \{1, \ldots, n\} \), which acts on \( \mathbb{R}^n \) by permutation of the entries. The inclusion \( \subseteq \) had already been proven by Schur 31 years earlier in [Sch23]. This result is therefore called the Schur–Horn Convexity Theorem.

This theorem was the first in a long series of convexity theorems. Among them is the Kostant Convexity Theorem for semisimple Lie algebras [Ko73], Ólafsson's Convexity Theorem on symmetric spaces [Ol90], and several versions for Hamiltonian torus actions on symplectic manifolds [HN94]. It is an interesting question, which of these convexity theorems generalize to an infinite dimensional setting and which new phenomena arise in this context. A particular convexity theorem of this type has been obtained in [BFR93], where an analog of Kostant's theorem was shown.
for the group of area preserving diffeomorphisms of the annulus. This result was motivated by the study of certain nonlinear PDEs.

In our paper we start from the original theorem by Schur and Horn. We prove an operator-theoretic generalization of this theorem to infinite dimensional separable Hilbert spaces. This provides insight into new phenomena occurring in the infinite dimensional situation and should be considered as a first step towards an appropriate version of Kostant’s theorem for infinite dimensional Lie groups.

Our main object of study is the algebra of bounded linear operators on $l^2(\mathbb{N})$. We replace $\mathfrak{U}(n)$ by $\mathfrak{U}$, the group of unitary isomorphisms, and $\mathfrak{S}_n$ by $\mathfrak{S} := \mathfrak{S}(\mathbb{N})$, the group of bijections on $\mathbb{N}$, our “Weyl group”. It turns out that with these changes the theorem still holds if we take the closure on both sides, that is we get

$$p_\infty(\mathfrak{U}.a) = \text{conv } \mathfrak{S}.a^\infty,$$

where $\mathfrak{U}.a := \{ U^* \text{diag}(a) U : U \in \mathfrak{U} \}$. The natural choice here is the $l^\infty$-closure, since $l^\infty$ is canonically isomorphic to the space of diagonal matrices equipped with the operator norm.

The infinite dimensional Schur–Horn Theorem will no longer describe the projection of the adjoint orbit of an arbitrary hermitian operator $A$ under $\mathfrak{U}$, since in infinite dimensions a hermitian operator in general is nondiagonalizable. We have to use other methods to describe $p_\infty(\mathfrak{U}.A)^\infty$, where $\mathfrak{U}.A := \{ U^*AU : U \in \mathfrak{U} \}$.

With the help of the spectral theorem we can decompose $A$ into a sum $A = A^- + A' + A^+$, where $A'$, $A^-$ are compact, $A^+$ has positive spectrum, $A^-$ has negative spectrum and each spectral value of $A'$ lies between the upper bound $x_1$ and the lower bound $x_0$ of its essential spectrum. Then

$$p_\infty(\mathfrak{U}.a) = p_\infty(\mathfrak{U}.A^-)^\infty + [x_0, x_1]^\mathbb{N} + p_\infty(\mathfrak{U}.A'^\infty)$$

is the desired description.

With these theorems we are able to prove that the projection on the diagonal gives us a bijection between the closed convex $\mathfrak{U}$-invariant subsets of the hermitian operators and the closed convex $\mathfrak{S}$-invariant subsets of the real diagonal operators.

In Section 2 we describe the set $\text{conv } \mathfrak{S}.a^\infty$ for $a \in l^p$ and $c_0$. We denote by $c_0$ the Banach space of all sequences converging to 0 equipped with the $l^\infty$ norm. We denote by $\mathfrak{C}_k$ the set of $k$-elementary subsets of $\mathbb{N}$ and define

$$L_k((a_j)_j) := \sup_{E \in \mathfrak{C}_k} \sum_{j \in E} a_j$$
for \( a = (a_i) \in l^\infty \). For \( p > 1 \) we obtain for \( a, b \in l^p \) that \( b \in \text{conv } \Xi \mathcal{a}^\infty \) if and only if \( L_k(b) \leq L_k(a) \) and \( L_k(-b) \leq L_k(-a) \) for all \( k \in \mathbb{N} \). This condition also describes the elements of \( \text{conv } \Xi \mathcal{a}^\infty \) whenever \( a \in c_0 \). It turns out in particular that \( \text{conv } \Xi \mathcal{a}^\infty = \text{conv } \Xi \mathcal{a}^\infty \) for \( p, q > 1 \), as long as \( a \in l^p \cap l^q \).

The only distinct situation is \( p = 1 \), where we need the additional condition \( \sum_i a_i = \sum_i b_i \).

For general sequences we need different techniques. We write \( a \in l^\infty \) as

\[
\mathcal{a} = \mathcal{a} + \mathcal{\tilde{a}}, \quad \text{where } g, \tilde{a} \in c_0, \quad g < 0, \quad \tilde{a} > 0 \quad \text{and} \quad \lim \inf \mathcal{a}' \leq \mathcal{a}' \leq \lim \sup \mathcal{a}' \quad \text{for all } i.
\]

Then

\[
\text{conv } \Xi \mathcal{a}^\infty = \text{conv } \Xi \mathcal{\tilde{a}}^\infty + [\lim \inf \mathcal{a}', \lim \sup \mathcal{a}']^\mathbb{N} + \text{conv } \Xi \mathcal{\tilde{a}}^\infty .
\]

Closer examination reveals that again \( \text{conv } \Xi \mathcal{a}^\infty \) is the set of all \( b \in l^\infty \) satisfying \( L_k(b) \leq L_k(a) \) and \( L_k(-b) \leq L_k(-a) \) for all \( k \in \mathbb{N} \).

In Section 3 we prove the generalization of the Schur–Horn Theorem. There we use the fact that the map \( a \mapsto p(U^*\text{diag}(a)U) \) is a linear map given by a doubly stochastic matrix. Using the description of \( \text{conv } \Xi \mathcal{a}^\infty \) the inclusion \( \subseteq \) is then straightforward. To show that \( p(U, a) \) is dense in \( \text{conv } \Xi \mathcal{a}^\infty \) we permute \( a \) in a useful way and apply the finite dimensional version of the Schur–Horn Theorem “piecewise”. As a byproduct we can prove a similar result for the algebra of Hilbert–Schmidt operators.

In Section 4 we consider arbitrary hermitian operators and describe \( p(U, A) \). To do this we use the spectral theorem to approximate \( A \) with diagonalizable operators.

In the last section we use these results to show that the projection on the diagonal gives a bijection of closed convex invariant subsets of the hermitian and the diagonal operators. We will also prove that a continuous \( U \)-invariant function on a closed convex subset of the hermitian operators is convex if its restriction to the diagonal operators is convex.

2. THE CONVEX HULL

We consider only real valued sequences throughout this section and denote by \( l^p, c_0 \) etc. real Banach spaces of real valued sequences.

Let \( \Xi := \Xi(\mathbb{N}) \) denote the set of all bijections of \( \mathbb{N} \). Then \( \Xi \) acts on \( l^\infty \) by permutation of the entries. For \( a \in l^\infty \) let \( \Xi a \) denote the orbit of \( \Xi \). In this section we analyze the sets \( \text{conv } \Xi a^\infty \), where \( \bar{A}^p \) denotes the closure of \( A \subseteq l^p \) in the \( l^p \)-topology, \( 1 \leq p \leq \infty \). If we omit \( p \) we mean the \( l^\infty \)-topology.

In the finite dimensional case we have the following lemma:

**Lemma 2.1.** Let \( x_1, ..., x_n \) and \( y_1, ..., y_n \) be in \( \mathbb{R} \). Then \( (y_1, ..., y_n) \) lies in \( \text{conv } \Xi (x_1, ..., x_n) \) if and only if:
1. \[
\sum_i x_i = \sum_j y_j
\]
2. \[
\sum_{i=1}^k y_{(\sigma(i))} \leq \sup \{ \sum_{i=1}^k x_{(\tau(i))} : \tau \in \mathfrak{S}_n \} \text{ for all } k \leq n, \sigma \in \mathfrak{S}_n.
\]

**Proof.** This is [HLP52, pp. 49, 89].

We define \( \mathfrak{S}_k \) as the set of subsets of \( \mathbb{N} \) of order \( k \). Then \( \mathfrak{S} := \bigcup_{k=1}^\infty \mathfrak{S}_k \) is the set of finite subsets of \( \mathbb{N} \). For \( E \in \mathfrak{S} \) we define the continuous linear functional \( L_E \) on \( l^\infty \) by

\[
L_E(a) := \sum_{j \in E} a_j, \quad a = (a_j), \in l^\infty.
\]

For \( k \in \mathbb{N} \) we define

\[
L_k(a) := \sup_{E \in \mathfrak{S}_k} L_E(a).
\]

If \( a \in c_0 \), then the positive entries of \( a \) can be ordered by size. If \( a_1, \ldots, a_k \geq 0 \) are the \( k \) largest entries of \( a \) then \( L_k(a) = a_1 + \cdots + a_k \). In the infinite dimensional case we have the following lemma:

**Lemma 2.2.** Let \( a \in l^\infty(\mathbb{N}) \) and \( b \in \text{conv } \mathfrak{S}a \). Then for every \( k \in \mathbb{N} \):

1. \( L_k(b) \leq L_k(a) \);
2. \( L_k(-b) \leq L_k(-a) \).

Furthermore the set of all elements \( b \) fulfilling (1) and (2) is closed in the product-topology. If \( a \in l^p \), \( 1 \leq p \leq \infty \) then this set is in particular closed in the \( l^p \)-topology.

**Proof.** The set satisfying (1) and (2) is closed in the product topology, because for all \( A \geq 0 \) we have

\[
\{ b : \sup_{E \in \mathfrak{S}_k} L_E(b) \leq A \} = \bigcap_{E \in \mathfrak{S}_k} \{ b : L_E(b) \leq A \}
\]

and the sets on the right hand side are obviously closed in the product topology.

For \( E \in \mathfrak{S}_k \) and \( b \in \text{conv } \mathfrak{S}a \) we have

\[
L_E(b) \leq \sup \{ L_E(\text{conv } \mathfrak{S}a) \} = \sup \{ L_E(\mathfrak{S}a) \}
\]

and (2) follows exactly the same way.

For \( a \in \mathbb{R} \) we let \( a^+ := \max \{a, 0\} \) and \( a^- := \max \{-a, 0\} \). For a sequence \( b = (b_i) \), we set \( b^+ := (b_i^+) \), and \( b^- := (b_i^-) \).
Lemma 2.3. Let \( a = (a_i) \in c_d(\mathbb{N}) \) be arbitrary. Then for every \( k \in \mathbb{N} \) we have:

1. \( L_k(a) = L_k(a^+) \)
2. \( L_k(-a) = L_k(a^-) \)

In particular \( L_k(a) \) and \( L_k(-a) \) are both nonnegative.

Proof. (1) Since \( a \in c_d(\mathbb{N}) \), it is possible to find \( \{i_1, ..., i_k\} \) such that \( a^+_{i_j} \leq a^-_{i_j} \) for all \( j \notin \{i_1, ..., i_k\} \) and \( 1 \leq m \leq k \). Then \( L_k(a^-) = \sum_{m=1}^{k} a^-_{i_m} \). This shows \( \leq \). The other direction is obvious in case \( a \) has at least \( k \) positive entries. But if \( \{a_{i_1}, ..., a_{i_k}\} \) are all the positive entries and \( n < k \), then \( L_k(a^+) = \sum_{j=1}^{n} a^+_{i_j} \). In view of \( a \in c_d(\mathbb{N}) \), we find for every \( \varepsilon > 0 \) numbers \( a_{i_{k+1}}, ..., a_{i_{k'}} \) such that \( \sum_{j=1}^{m} a^+_{i_j} \geq \varepsilon \). Then

\[
\sup_{E \in \mathcal{E}_k} L_k(a) \geq \sum_{j=1}^{n} a^+_{i_j} + \sum_{j=n+1}^{k} a^-_{i_j} \geq \sum_{j=1}^{n} a^+_{i_j} - \varepsilon.
\]

and thus both sides are equal.

Assertion (2) follows by applying (1) to \( -a \).

Lemma 2.4. Let \( \lambda_1, ..., \lambda_n, \lambda_1', ..., \lambda_m' \in [0, 1] \) with \( \sum_{i=1}^{n} \lambda_i = \sum_{j=1}^{m} \lambda_j' = 1 \). Let \( E_1 \) and \( E_2 \) be subsets of \( \mathbb{N} \) and \( \sigma_1, ..., \sigma_n \in \mathcal{E}(E_1) \), \( \sigma_1', ..., \sigma_m' \in \mathcal{E}(E_2) \). Then there exist \( \mu_1, ..., \mu_s \) and \( i_1, ..., i_s \), \( j_1, ..., j_s \), such that

\[
\sum_{i=1}^{n} \lambda_i \sigma_i = \sum_{k=1}^{s} \mu_k \sigma_k \quad \text{and} \quad \sum_{j=1}^{m} \lambda_j' \sigma_j' = \sum_{k=1}^{s} \mu_k \sigma_k'.
\]

Proof. We set \( \mu_{i,j} := \lambda_{i,j} \lambda_{i,j}' \). Then \( \sum_{i,j} \mu_{i,j} = 1 \) and the assertion obviously holds if we consider

\[
\sum_{i=1}^{n} \lambda_i \sigma_i = \sum_{i,j} \mu_{i,j} \sigma_i \quad \text{and} \quad \sum_{j=1}^{m} \lambda_j' \sigma_j' = \sum_{i,j} \mu_{i,j} \sigma_j'.
\]

and change the index set of the \( \mu \) from \( \{1, ..., l\} \times \{1, ..., j\} \) to \( \{1, ..., s\} \). \( \blacksquare \)

Now we can prove the following lemma.

Lemma 2.5. Let \( a = (a_i) \in l^1 \). Then \( b = (b_j) \in \text{conv} \supseteq a^1 \) if and only if

1. \( \sum_i a_i = \sum_j b_j \),
2. \( L_k(b) \leq L_k(a) \) for all \( k \in \mathbb{N} \),
3. \( L_k(-b) \leq L_k(-a) \) for all \( k \in \mathbb{N} \).
Proof. First let \( b \in \text{conv} \mathcal{Z}a \), i.e., \( b = \sum_{i=1}^{n} \lambda_i \sigma_i a \) for \( \lambda_i \geq 0 \) such that \( \sum_{i=1}^{n} \lambda_i = 1 \). Then

\[
\sum_{j=1}^{\infty} \left( \sum_{i=1}^{n} \lambda_i \sigma_i a_{\sigma_{(j)}} \right) = \sum_{j=1}^{\infty} \lambda_i \sum_{i=1}^{n} a_{\sigma_{(j)}} = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} a_j
\]

establishes (1). Lemma 2.2 shows that (2) and (3) hold for \( b \in \text{conv} \mathcal{Z}a \).

Let us assume that (1) to (3) hold for \( b \). Our main idea is to truncate \( a \) and \( b \) and then apply Lemma 2.1 to their positive and negative entries separately. However we have to make sure that the truncated sequences will still fulfill condition (1). Note also that the sequences \( a + \) and \( b + \) no longer fulfill condition (1).

Let \( x_i^+ \geq x_i^+ \geq \cdots \) denote the positive and \( x_i^- \leq x_i^- \leq \cdots \) denote the negative entries of \( a \). Similarly let \( \beta_i^+ \geq \beta_i^+ \geq \cdots \) denote the positive and \( \beta_i^- \leq \beta_i^- \leq \cdots \) denote the negative entries of \( b \). Continue any of these sequences with 0 if \( a \) or \( b \) has only finitely many positive or negative entries.

Now let \( \epsilon > 0 \) be arbitrary. Then there are integers \( r, s \) and positive numbers \( B^+, B^- \) such that

\[
B^+ := \sum_{i=r+1}^{\infty} \beta_i^+ \leq \epsilon, \quad -B^- := \sum_{i=s+1}^{\infty} \beta_i^- \geq -\epsilon.
\]

Further there are integers \( p \) and \( q \) such that

\[
\sum_{i=p}^{\infty} x_i^+ > B^+ \geq \sum_{i=p+1}^{\infty} x_i^- +, \quad \sum_{i=q}^{\infty} x_i^- < -B^- \leq \sum_{i=q+1}^{\infty} x_i^-.
\]

Therefore we can find \( x_p^* \in [0, x_p^+ \text{]} \) and \( x_q^- \in [x_q^-, 0[ \) such that

\[
(x_p^+ - x_p^*) + x_p^+ + \cdots = B^+, \quad (x_q^-- x_q^-*) + x_q^- + \cdots = -B^-.
\]

Let \( m := \max \{ p, r \} \) and \( n := \max \{ q, s \} \). We define the sequences

\[
a' := (x_1^+, \ldots, x_{m-p}^+, x_{m-p}^*, 0, \ldots, 0, x_1^-, \ldots, x_{n-q-1}^-, x_{n-q}^*, 0, \ldots, 0, 0, \ldots),
\]

\[
b' := (\beta_1^+, \ldots, \beta_{m-r}^+, 0, \ldots, 0, \beta_1^-, \ldots, \beta_{n-s}^-, 0, \ldots, 0, 0, \ldots).
\]

Then after permuting \( a \) and \( b \), which will not violate the condition \( b \in \text{conv} \mathcal{Z}a \), we have \( \|a - a'\|_1 = \|b - b'\|_1 = B^+ + B^- \leq 2\epsilon \).
We also have

\[
0 = \sum_{i=1}^{m} a_i - \sum_{i=1}^{m} b_i = \sum_{i=1}^{m} x_i^+ - \sum_{i=1}^{m} x_i^- - \sum_{i=1}^{m} \beta_i^+ - \sum_{i=1}^{m} \beta_i^- \\
= \sum_{i=1}^{m} a_i' + B^+ + \sum_{i=m+1}^{m+n} a_i' - B^- - \sum_{i=1}^{m} b_i' - B^+ - \sum_{i=m+1}^{m+n} b_i' + B^- \\
= \left( \sum_{i=1}^{m} a_i' - \sum_{i=1}^{m} b_i' \right) + \left( \sum_{i=m+1}^{m+n} a_i' - \sum_{i=m+1}^{m+n} b_i' \right).
\]

So there exists a \( \delta > 0 \) such that

\[
\delta = \sum_{i=1}^{m} a_i' - \sum_{i=1}^{m} b_i' = \sum_{i=m+1}^{m+n} b_i' - \sum_{i=m+1}^{m+n} a_i'.
\]

We can find a positive \( \epsilon \) smaller than any absolute value of a nonzero entry of \( a' \) or \( b' \) and a \( k \in \mathbb{N} \) such that \( k\epsilon = \delta \). The entries of \( a' \) are the largest and smallest entries of \( a \). So we can use Lemma 2.1 to find numbers \( \mu_j, \mu'_j \in [0,1] \) satisfying \( \sum_j \mu_j = \sum_j \mu'_j = 1 \) and \( \tau_j \in \mathbb{Z}_{m+k}, \tau'_j \in \mathbb{Z}_{n+k} \) such that

\[
\sum_j \mu_j \tau_j (x_1^+, \ldots, x_{p-1}^+, x_p^{*}, 0, \ldots, 0) = (\beta_1^+, \ldots, \beta_s^+, c, \ldots, c, 0, \ldots, 0)_{m-p+k}
\]

and

\[
\sum_j \mu'_j \tau'_j (x_1^-, \ldots, x_{q-1}^-, x_q^{*}, 0, \ldots, 0) = (\beta_1^-, \ldots, \beta_s^-, -c, \ldots, -c, 0, \ldots, 0)_{n-q+k}
\]

With Lemma 2.4 we obtain \( \lambda_j, \lambda'_j \) and \( \sigma_j, \sigma'_j \in \Xi \) such that

\[
\sum_j \lambda_j \sigma_j (x_1^+, \ldots, x_{p-1}^+, x_p^{*}, 0, \ldots, 0, x_1^-, \ldots, x_{q-1}^-, x_q^{*}, 0, \ldots, 0, 0, \ldots, 0) = (\beta_1^+, \ldots, \beta_s^+, c, \ldots, c, 0, \ldots, 0, 0, \ldots, 0)_{m-r}
\]

and

\[
\sum_j \lambda'_j \sigma'_j (x_1^+, \ldots, x_{p-1}^+, x_p^{*}, 0, \ldots, 0, x_1^-, \ldots, x_{q-1}^-, x_q^{*}, 0, \ldots, 0, 0, \ldots, 0) = (\beta_1^+, \ldots, \beta_s^+, -c, \ldots, -c, 0, \ldots, 0, 0, \ldots, 0)_{n-s}
\]

So we have \( \frac{1}{2} \sum \lambda_j \sigma_j + \frac{1}{2} \sum \lambda'_j \sigma'_j \) \( a' = b' \). Note that \( \frac{1}{2} \| \sum \lambda_j \sigma_j + \sum \lambda'_j \sigma'_j \|_{ap,1} \leq 1 \). Now
\[ b - \frac{1}{2} \left( \sum_i \lambda_i \sigma_i + \frac{1}{2} \sum_i \lambda_i' \sigma_i' \right) a \]
\[ \leq \| b - b' \|_1 + \left\| b' - \frac{1}{2} \left( \sum_i \lambda_i \sigma_i + \frac{1}{2} \sum_i \lambda_i' \sigma_i' \right) a \right\|_1 + \frac{1}{2} \left( \sum_i \lambda_i \sigma_i + \sum_i \lambda_i' \sigma_i' \right) (a' - a) \]
\[ \| a \|_1 \leq 4 \epsilon. \]

Therefore \( b \) lies in \( \text{conv} \mathcal{S} a \).

We can describe \( \text{conv} \mathcal{S} a \) for \( 1 < p \leq \infty \) in a similar way.

**Lemma 2.6.** Let \( a \) be an \( l^1 \)-sequence. Then \( b \) lies in \( \text{conv} \mathcal{S} a \) if and only if for all \( k \in \mathbb{N} \):

1. \( L_k(b) \leq L_k(a) \);
2. \( L_k(-b) \leq L_k(-a) \).

In particular \( \| b \|_1 \leq \| a \|_1 \).

**Proof.** Lemma 2.2 tells us that every \( b \in \text{conv} \mathcal{S} a \) fulfills (1) and (2). With the help of Lemma 2.3 we observe that

\[ \sum_{i=1}^{\infty} |b_i| = \sum_{i=1}^{\infty} b_i^+ + \sum_{i=1}^{\infty} b_i^- = \lim_{k \to \infty} L_k(b^+) + \lim_{k \to \infty} L_k(b^-) \]
\[ \leq \lim_{k \to \infty} L_k(a^+) + \lim_{k \to \infty} L_k(a^-) = \sum_{i=1}^{\infty} |a_i| = \| a \|_1. \]

Let \( \delta > 0 \). Suppose we have found \( b' \) satisfying (1) and (2), such that \( \| b - b' \|_p < \delta \) and \( \sum_{i=1}^{\infty} b_i = \sum_{i=1}^{\infty} a_i \). Then \( b' \in \text{conv} \mathcal{S} a \) fulfills (1) and (2) (Lemma 2.5) and therefore \( \text{dist}(b, \text{conv} \mathcal{S} a) \leq \delta \) for every \( \delta > 0 \), which means \( b \in \text{conv} \mathcal{S} a \).

To construct \( b' \) we define \( \delta := \sum_{i=1}^{\infty} (a_i - b_i) \). Then \( \delta \) is well defined and finite since \( b \in \ell^1 \). If \( \delta = 0 \), then we are done. So we can assume that \( \delta > 0 \). Otherwise we apply the following proof to \( -a \) and \( -b \).

We know from Lemma 2.3 that \( \sum_{i=1}^{\infty} b_i^- \leq \sum_{i=1}^{\infty} a_i^- \). Thus

\[ \delta = \sum_{i=1}^{\infty} a_i - \sum_{i=1}^{\infty} b_i = \sum_{i=1}^{\infty} a_i^+ - \sum_{i=1}^{\infty} b_i^- \leq \sum_{i=1}^{\infty} a_i^+ - \sum_{i=1}^{\infty} b_i^+, \]

and therefore \( \sum_{i=1}^{\infty} b_i^+ + \delta \leq \sum_{i=1}^{\infty} a_i^+ \).
Let \( x_1 \geq x_2 \geq \cdots \) denote the positive entries of \( a \) and \( \beta_1 \geq \beta_2 \geq \cdots \) those of \( b \). Then there exists an \( N \in \mathbb{N} \) such that for all \( n \geq N \)
\[
\sum_{i=1}^{n} \beta_i + \frac{\delta}{2} \leq \sum_{i=1}^{n} x_i
\]
and \( \beta_{N+1} < \beta_N \). We observe that for every \( \lambda \in \mathbb{R} \)
\[
\left( \frac{\lambda}{N}, \ldots, \frac{\lambda}{N}, 0, \ldots \right) \xrightarrow{n \to \infty} 0,
\]
the convergence to 0 holding in the \( l^p \)-norm for \( 1 < p \leq \infty \). So we find \( k \in \mathbb{N} \) with
\[
\left\| \left( \frac{\delta}{2k}, \ldots, \frac{\delta}{2k}, 0, \ldots \right) \right\|_p \leq \frac{\epsilon}{2}.
\]
By enlarging \( k \), which will not violate the above condition, we get \( \beta(2k) \leq \beta_N - \beta_{N+1} \). We can find \( i_1, \ldots, i_k \in \mathbb{N} \) such that \( b_{i_j} < \beta_N \) for \( j = 1, \ldots, k \). We define
\[
b^1 := (b^1_i)_i, \quad b^1_i := \begin{cases} b_i + \frac{\delta}{2k} & i \in \{i_1, \ldots, i_k\} \\ b_i & \text{otherwise} \end{cases}.
\]
The \( \beta_1, \ldots, \beta_N \) are the \( N \) largest entries of \( b^1 \). Let \( E \in \mathcal{E} \), \( l > N \), then
\[
L_E(b^1) \leq L_E(b^+) + k \frac{\delta}{2k} \leq \beta_1 + \cdots + \beta_l + \frac{\delta}{2} \leq x_1 + \cdots + x_l,
\]
so \( b^1 \) fulfills condition (1). It also fulfills condition (2) because \( b^{1-} \leq b^* \). Furthermore \( \sum_{i=1}^{\infty} (a_i - b^1_i) = \frac{\delta}{2} \) and \( \|b - b^1\| \leq \frac{\delta}{2} \).
We can apply the same procedure to \( b^1 \) to get \( b^2 \) with \( \sum_{i=1}^{\infty} (a_i - b^2_i) = \frac{\delta}{4} \) and \( \|b^1 - b^2\| \leq \frac{\delta}{4} \). This \( b^2 \) will still satisfy conditions (1) and (2). Continuing this way we obtain a sequence \( (b^n)_{n \in \mathbb{N}} \) satisfying the conditions (1) and (2), \( \sum_{i=1}^{\infty} (a_i - b^n_i) = \delta 2^{n+1} \) and \( \|b^{n-1} - b^n\|_p \leq \epsilon / 2^n \). This is a Cauchy sequence and therefore converges to some \( b^* \). Since conditions (1) and (2) define a closed set, \( b^* \) will satisfy them and for continuity reasons \( \sum_{i=1}^{\infty} (a_i - b_i^*) = 0 \). Thus we obtain \( b^* \in \text{conv } \mathcal{E} : a^* \). Further \( \|b - b^*\|_p \leq \sum_{n=1}^{\infty} \|b^{n-1} - b^n\|_p \leq \sum_{n=1}^{\infty} \epsilon / 2^n = \epsilon \), so \( b^* \) is as desired.
**Corollary 2.7.** For $a \in l^1$ we have $\text{conv} \, \mathcal{S}a \subseteq l^1$ and $\text{conv} \, \mathcal{S}a^p = \text{conv} \, \mathcal{S}a^q$ for $1 < p, q \leq \infty$.

**Proposition 2.8.** (1) Let $a \in l^p(\mathbb{N})$. Then $b \in \text{conv} \, \mathcal{S}a^p$ if and only if for all $k \in \mathbb{N}$ we have:

1. $L_k(b) \leq L_k(a)$;
2. $L_k(-b) \leq L_k(-a)$.

(2) Let $a \in c_0(\mathbb{N})$. Then $b \in \text{conv} \, \mathcal{S}a$ if and only if for all $k \in \mathbb{N}$ we have:

1. $L_k(b) \leq L_k(a)$;
2. $L_k(-b) \leq L_k(-a)$.

**Proof.** The proof is exactly the same in both cases. That each $b \in \text{conv} \, \mathcal{S}a^p$ fulfills (1) and (2) follows from Lemma 2.2.

Conversely assume $b$ satisfies (1) and (2). Note that $a, b \in c_0(\mathbb{N})$ in both cases. Let $\beta_i^+, \beta_i^- \geq \cdots$ denote the positive and $\beta_i^-, \beta_i^+ \leq \cdots$ denote the negative entries of $b$ and let $\alpha_i^+, \alpha_i^- \geq \cdots$ denote the positive and $\alpha_i^- \leq \alpha_i^+ \leq \cdots$ denote the negative entries of $a$. Let $\epsilon > 0$. Then in Case (1) there exist $r$ and $s$ such that

$$
\left( \sum_{i=r}^{s} \beta_i^+ \right)^{1/p} \leq \epsilon, \quad \left( \sum_{i=r}^{s} \beta_i^- \right)^{1/p} \leq \epsilon.
$$

In Case (2) we choose $r$ such that $\alpha_i^+, \alpha_i^- \leq \epsilon/2$ for $i \geq r$ and $\alpha_i^-, \alpha_i^+ \geq -\epsilon/2$ for $i \geq s$. We define new sequences $b'$ and $a'$ by setting the entries $\beta_i^+, \alpha_i^-$ to 0 for $i \geq r$ and $\beta_i^-, \alpha_i^+$ to 0 for $i \geq s$.

Then $\|b - b'\|_p \leq \epsilon$, $\|a - a'\|_p \leq \epsilon$. We have $b', a' \in l^1$ satisfying the conditions (1) and (2) since the largest and smallest entries of $a$ and $b$ have not been eliminated. With Lemma 2.6 we get $\lambda_i, \sigma_i$ with $\|\sum \lambda_i, \sigma_i a' - b'\|_p \leq \epsilon$. We note that the $l^1$-operator norm of $\sum \lambda_i, \sigma_i$ is less than or equal to 1. Therefore

$$
\left\| \sum \lambda_i, \sigma_i a - b \right\|_p \leq \left\| \sum \lambda_i, \sigma_i (a - a') \right\|_p + \left\| \sum \lambda_i, \sigma_i a' - b' \right\|_p + \|b' - b\|_p < 3\epsilon,
$$
which proves the lemma.

**Remark 2.9.** (1) With the aid of Lemma 2.3 the conditions (1) and (2) in Proposition 2.8 can be formulated in another way: (1) just says that the sum over the $k$ largest positive entries of $b$ is smaller than the sum over
the \( k \) largest positive entries of \( a \) for all \( k \). If there are less than \( k \) positive entries in \( a \) or \( b \), just look at the sum over all of them. Condition (2) says the same thing for the smallest entries of \( a \) and \( b \).

(2) It is interesting to note, that the condition on the positive entries of \( b \) only uses the positive entries of \( a \) and the same holds for the negative entries.

(3) Lemma 2.3 implies that if \( b \) lies in \( \operatorname{conv} \mathcal{S} a \), then so does every \( b' \) fulfilling \( b'^+ \leq b^+ \) and \( b'^- \leq b^- \), where the order relation is taken componentwise.

**Corollary 2.10.** For \( a \in c_0(\mathbb{N}) \) \( (a \in l^p(\mathbb{N}), p > 1) \) we have \( 0 \in \operatorname{conv} \mathcal{S} a \) \( (0 \in \operatorname{conv} \mathcal{S} a'^p) \).

**Remark 2.11.** (1) From the explicit description in Proposition 2.8 it follows in particular that \( \operatorname{conv} \mathcal{S} a \) is closed even in the weak-*-topology of \( l^\infty \) as the dual of \( l^1 \).

(2) It is obvious that for \( a \in l^p \) (resp. \( a \in c_0 \) for \( p = \infty \)) we have

\[
\operatorname{conv} \mathcal{S} a'^p = \operatorname{conv} \mathcal{S} a'^p,
\]

where \( \mathcal{S} \_\infty = \bigcup_n \mathcal{S} _n \) and \( \mathcal{S} _n \) acts on a sequence by acting on the first \( n \) entries.

Now we take a closer look at the case when the sequence is not \( c_0 \).

**Definition 2.12.** We define

\[
\mathcal{F}_n := \{(a_i)_i \in [0, 1]^\mathbb{N} : (\forall i) n a_i \text{ integer}\}
\]

and

\[
\mathcal{F} := \bigcup_{n \in \mathbb{N}} \mathcal{F}_n.
\]

Note that \( \mathcal{F} \) is a dense subset of \([0, 1]^\mathbb{N}\).

We start with a special case:

**Lemma 2.13.** Let \( a = (1, 0, 1, 0, 1, 0, ...) \). Then

\[
\operatorname{conv} \mathcal{S} a = [0, 1]^\mathbb{N}.
\]

**Proof.** The inclusion \( \subseteq \) is obvious. To show \( \supseteq \) it is sufficient to prove that \( \mathcal{F} \supseteq \operatorname{conv} \mathcal{S} a \). It is even enough to consider only those sequences with entries other than 0 or 1 infinitely many times, since these form a
dense subset of $\mathcal{F}$. Let $f \in \mathcal{F}_n$ be such a sequence. After a permutation we can write $f$ as

$$f = \frac{1}{n} (a_1, \ldots, a_s, m_1, \ldots, m_k, m_1, \ldots, m_s, \ldots)$$

with $0 \leq a_i, m_i \leq n$ integers. Then $\sum_{i=1}^s a_i = pn + q$, where $0 \leq p \leq s$, $0 \leq q < n$ are integers. We permute $a = (1, 0, 1, 0, \ldots)$ to the sequences

$$a'_1 = (1, \ldots, 1, 0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0, \ldots)$$

with $\kappa \in \{0, 1\}$. This can be done, because at least one of the $m_i$ is neither $n$ nor $0$, and therefore we have infinitely many places to put the 1’s and 0’s. So the sequence

$$a' := \frac{n-q}{n} a_0 + \frac{q}{n} a_1$$

lies in $\text{conv } \mathcal{G}_a$.

Consider the permutation

$$\sigma^* : \mathbb{R}^n \to \mathbb{R}^n \ (x_1, \ldots, x_n) \mapsto (x_n, x_1, \ldots, x_{n-1}).$$

Then

$$\left(\sum_{i=0}^{n-1} \frac{1}{n} (\sigma^*)^i \right) \left(1, \ldots, 1, 0, \ldots, 0, \right) = \left(\frac{m}{n}, \ldots, \frac{m}{n}\right).$$

By Lemma 2.1 there exist numbers $\mu_i$ and $\sigma_i \in \mathcal{G}_a$ such that

$$\sum_{i=1}^n \mu_i \sigma_i \left(1, \ldots, 1, \frac{q}{n}, 0, \ldots, 0, \right) = \left(\frac{d_1}{n}, \ldots, \frac{d_s}{n}\right).$$
From Lemma 2.1 there exist $\lambda_j$ and integers $i_j$ and $k_j$ such that for the following block diagonal matrices we have

$$\sum_j \lambda_j \begin{pmatrix}
\sigma_j & 0 & 0 & \ldots \\
0 & (\sigma^+)^{k_j} & 0 & \ldots \\
0 & 0 & (\sigma^+)^{k_j} & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix} a' = f$$

and thus $f \in \text{conv } \Xi a' \subseteq \text{conv } \Xi a.$

**Corollary 2.14.** Let $a = (a_i) \in [0,1]^N$ such that $\limsup a_i = 1$ and $\liminf a_i = 0$. Then

$$\text{conv } \Xi a = [0,1]^N$$

**Proof.** Again, “$\subseteq$” holds trivially. First we assume that $a \notin \Xi$. Then $a$ has infinitely many entries 0 and 1. Given an arbitrary $N > 0$ there exists $n > N$ with $a \notin \Xi_n$. Then we can permute $a$ to

$$a' = \left(1, \frac{m_1}{n}, 0, \ldots, 0, 1, \frac{m_2}{n}, 0, \ldots, 0, \ldots \right)$$

for $0 \leq m_i \leq n$ integer. Let $m_1, \ldots, m_s$ be all the different nominators occurring. We denote by $(1, k)$ the transposition permuting the 1st and k-th entry ($(1, 1) := \text{id}$). We observe that

$$\sum_{i=1}^m \frac{1}{m} (1, i) \left(\frac{m}{n}, 0, \ldots, 0, \frac{m_i}{n_{m-1}}\right) = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right).$$

By Lemma 2.4 there are $\lambda_i$ and $\sigma_{i,j}$ $j = 1, \ldots, s$ such that

$$\sum_i \lambda_i \sigma_{i,j} (m_j/n, 0, \ldots, 0) = (1/n, \ldots, 1/n)$$

for all $j$. Then

$$\sum_i \lambda_i \begin{pmatrix}
1 & 0 & 0 & 0 & \ldots \\
\sigma_{i,1} & 0 & 0 & \ldots \\
0 & 1 & 0 & \ldots \\
0 & 0 & \sigma_{i,2} & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix} a' = \left(1, \frac{1}{n}, \ldots, 1, \frac{1}{n}, \ldots, 1, \ldots \right)$$
and thus by Lemma 2.13 \([1/n, 1]^N \subseteq \text{conv } S\). So we have
\[
[0, 1]^N = \bigcup_{N \in \mathbb{N}} \left[ \frac{1}{N}, 1 \right]^N \subseteq \text{conv } S.
\]

If now \(a\) is arbitrary, then an approximation argument proves the corollary: Given \(b \in [0, 1]^N\) and \(\varepsilon > 0\), there exists an \(f \in \mathbb{R}\) with infinitely many entries 0 and 1, such that \(\|a - f\| < \varepsilon\). The above consideration gives \(\lambda_i \geq 0, \sum_i \lambda_i = 1, \text{ and } \sigma_i \in \mathbb{R}\) such that \(\sum_i \lambda_i \sigma_i f - b\). Then
\[
\sum_i \lambda_i \sigma_i a - b \leq \sum_i \lambda_i \sigma_i (a - f) + \sum_i \lambda_i \sigma_i f - b \leq 2\varepsilon
\]
because \(\|\sum_i \lambda_i \sigma_i\| \leq 1\).

Now we can put all the pieces together.

**Theorem 2.15.** Let \(a = (a_i) \in l^\infty\). Define 
\[A^+ := \lim \sup a_i \text{ and } A^- := \lim \inf a_i.\]
Define \(\bar{a} := \max\{a_i, A^+\} - A^- \text{ and } \underline{a} := \min\{a_i, A^-\} - A^\cdot.\)
Let \(\bar{a} = (\bar{a}_i)\) and \(a = (a_i)\). Then
\[
\text{conv } S - a = \text{conv } S - a + [A^-, A^+] + \text{conv } S.a.
\]

**Proof.** The inclusion “\(\subseteq\)” follows immediately from Proposition 2.8 and Corollary 2.14. To see that the right hand side is closed, just note that each summand is closed in the weak-* topology (Remark 2.11.1 and obvious for \([A^-, A^+]^N\)). Because they are also bounded they are weak-* compact and so is their sum. Then in particular the sum is closed in the norm topology.

“\(\supseteq\)” If \(A^+ = A^-\) then this is just Proposition 2.8. Otherwise we can assume that \(A^+ = 1\) and \(A^- = 0\). First let us assume that there are infinitely many entries in \(a\) equal to 1 and to 0 and that \(a\) has only finitely many entries not in \([0, 1]\). Let \(x_i^+ + 1 \geq \ldots \geq x_i^+ + 1\) denote the entries greater than 1 and \(x_i^- \leq \ldots \leq x_i^-\) denote those smaller than 0.

By definition we have \(a - a - \bar{a} \in [0, 1]^N\). Pick \(b \in \text{conv } S - a, b \in \text{conv } S - a\) and \(d \in [0, 1]^N\) and let \(b := b + d + b\). It is sufficient to look at those \(b\) for which \(b\) and \(b\) have only finitely many nonzero entries and \(d\) has at least one rational entry in \([0, 1]\.\) The elements with these properties lie densely in \(\text{conv } S - a + [A^-, A^+] + \text{conv } S.a\). Let \(b_i := \max\{1, b_i, -1, b_i := \min\{b_i, 0\}\) and \(b := (b_i), b := (b_i)\). Then \(0 \leq b \leq b\) and therefore by Remark 2.9.3 we have \(b \in \text{conv } S - a\). Analogously \(b \in \text{conv } S - a\) and obviously \(d := b - b \in [0, 1]^N\). If \(b_i = 0\) then \(b_i = 0\). So \(b\) has only finitely many nonzero entries. The same holds for \(b\). The sequence \(d\) differs from \(d\) only where \(b\) or \(b\) are nonzero. These are only
finitely many entries, so $d'$ will have the same rational entries infinitely often. Therefore we have to look only at the case that:

1. $d$ has at least one rational entry not in $\{0, 1\}$ infinitely often;
2. $\bar{b}$ and $b$ have only finitely many nonzero entries;
3. For every $i$ at most one of $b_i$ and $\bar{b}_i$ is not zero;
4. If $\bar{b}_i \neq 0$ then $d_i = 0$. If $\bar{b}_i \neq 0$ then $d_i = 1$.

We let $b := \bar{b} + d + \bar{b}$.

Let $\beta_1^+ \geq \ldots \beta_p^+$ denote the nonzero entries of $\bar{b}$ and $\beta_1^- \leq \ldots \leq \beta_q^-$ those of $\bar{b}$. Then we know

$$(\beta_1^+, \ldots, \beta_p^+, 0, \ldots) \in \operatorname{conv} \mathcal{Z}(\pi_1^+, \ldots, \pi_p^+, 0, \ldots).$$

By Proposition 2.8 we deduce that $0 \leq \delta := \pi_1^+ + \ldots + \pi_n^+ - \beta_1^+ - \ldots - \beta_p^+$. Analogously we get $0 \geq \delta^* := \pi_1^- + \ldots + \pi_n^- - \beta_1^- - \ldots - \beta_q^-$. Let $\delta_0 \in \mathbb{N}$ denote a rational entry of $d$ occurring infinitely often. We pick $\varepsilon > 0$. Then there exists an $n \in \mathbb{N}$ with $n\delta_0$ an integer, $n\delta_0 \geq r - \delta^* \delta^*/n \leq \varepsilon$ and $\delta/n \leq 1 - \delta_0$. Now we look at the sequences

$$a_i = (\pi_1^+ + 1, \ldots, \pi_n^+ + 1, 1, 1, 0, \ldots, 0),$$

$$b_1 = (\beta_1^+ + 1, \ldots, \beta_p^+ + 1, d_0 + \frac{\delta}{n}, \ldots, d_0 + \frac{\delta}{n}),$$

$$b_1 = (\beta_1^+ + 1, \ldots, \beta_p^+ + 1, d_0, \ldots, d_0).$$

We have that

$$\sum_i (a_i) = \sum_{i=1}^r (\pi_1^+ + 1) + (p - r + nd_0) = \sum_{i=1}^r \pi_i^+ + p + nd_0$$

$$= \sum_{i=1}^r \beta_i^+ + \delta + p + nd_0 = \sum_{i=1}^n (\beta_i^+ + 1) + n \left(\frac{\delta}{n} + d_0\right)$$

$$= \sum_i (b_i).$$

Since $b \in \operatorname{conv} \mathcal{Z}$ we have that $L_k(b'_1) \leq L_k(a_1)$ for every $k$. By Lemma 2.1 there are $\lambda_i^+, \sigma_i^-$ such that $b'_1 = \sum \lambda_i^+ \sigma_i^- a_1$, and therefore $\|b_1 - \sum \lambda_i^+ \sigma_i^- a_1\| = \delta/n \leq \varepsilon$.

In the same way we can find $n' \in \mathbb{N}$ with $n'd_0$ an integer, $n'(1 - d_0) \geq s - q$, $\delta' / n' \leq \varepsilon$ and $|\delta'/n'| \leq \delta_0$. We set
\(a_2 = (x_1^-, \ldots, x_n^-, \frac{0, \ldots, 0}{\kappa(1-d_0+q-\gamma)}, \frac{1, \ldots, 1}{\sigma d_0}),\)

\(b_2 = \left( \beta_1^-, \ldots, \beta_q^-, \frac{d_0 + \delta}{\mu}, \ldots, \frac{d_0 + \delta}{\nu} \right),\)

\(b_1 = (\beta_1^-, \ldots, \beta_q^-, \frac{d_0, \ldots, d_0}{\nu}).\)

By Lemma 2.1 we can find \(\lambda_1^-, \sigma_1^-\) satisfying \(b_2' = \sum \lambda_1^- \sigma_1^- a_2\) and therefore \(\|b_2 - \sum \lambda_1^- \sigma_1^- a_2\| \leq \varepsilon.\)

From our previous considerations we know that after a permutation we can write

\(a = (a_1, a_2, a_3), \quad b = (b_1, b_2, b_3)\)

with \(a_3, b_3 \in [0, 1]^N, \limsup a_1 = 1\) and \(\liminf a_3 = 0.\) By Corollary 2.14 there are \(\lambda_1^+, \sigma_1^+\) such that \(\|b_3 - \sum \lambda_1^+ \sigma_1^+ a_3\| \leq \varepsilon.\)

Now we use Lemma 2.4 to find \(\lambda_i\) and simultaneously change the indices of the \(\sigma_i^+, \sigma_i^-, \sigma_i'\) so that

\(\left\| \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} - \sum \lambda_i \begin{bmatrix} \sigma_i^+ & a_1 \\ \sigma_i^- & a_2 \\ \sigma_i' & a_3 \end{bmatrix} \right\| \leq \varepsilon.\)

This proves the assertion.

If \(a\) does not have infinitely many entries 0 and 1 and only finitely many entries not in \([0, 1]\), then for \(\varepsilon > 0\) we define \(a'\) by

\[
\begin{cases}
1: a_i \in [1 - \varepsilon, 1 + \varepsilon], \\
0: a_i \in [-\varepsilon, \varepsilon], \\
a_i: \text{otherwise.}
\end{cases}
\]

Then \(\|a - a'\| \leq \varepsilon.\) In particular we conclude \(\|a - a'\| \leq \varepsilon\) and \(\|a - a'\| \leq \varepsilon.\)

We pick \(b := b + d + \hat{b}\) in \(\text{conv} \mathcal{S}_a + [A^-, A^+]^N + \text{conv} \mathcal{S}.\) That is, we have \(b = \sum \lambda_i \sigma_i a_i, \quad b = \sum \mu_j \tau_j a_i, \quad d \in [0, 1]^N.\) As we have shown above, \(d + \sum \lambda_i \sigma_i a_i + \sum \mu_j \tau_j a_i \in \text{conv} \mathcal{S}.\) So we can find \(\lambda_1^+, \sigma_1^+\) such that

\[
\left\| \left( d + \sum \lambda_i \sigma_i a_i + \sum \mu_j \tau_j a_i \right) - \sum \lambda_i \sigma_i a_i \right\| \leq \varepsilon.
\]
Hence we have
\[
\left\| b - \sum_i \lambda_i' \sigma_i' a \right\| = \left\| \left( d + \sum_i \lambda_i \sigma_i a + \sum \mu_j \tau_j \overline{a} \right) - \sum_i \lambda_i' \sigma_i' a \right\|
\leq \left\| \left( d + \sum_i \lambda_i \sigma_i a + \sum \mu_j \tau_j \overline{a} \right) - \left( d + \sum_i \lambda_i \sigma_i' a' + \sum \mu_j \tau_j \overline{a'} \right) \right\|
+ \left\| \left( d + \sum_i \lambda_i \sigma_i a + \sum \mu_j \tau_j \overline{a} \right) - \sum_i \lambda_i' \sigma_i' a' \right\|
+ \left\| \sum_i \lambda_i' \sigma_i' a' - \sum_i \lambda_i' \sigma_i' a \right\|
\leq \left\| \sum_i \lambda_i \sigma_i \left( a - a' \right) \right\| + \left\| \sum_j \mu_j \tau_j \left( \overline{a} - \overline{a'} \right) \right\| + \varepsilon
+ \left\| \sum_i \lambda_i' \sigma_i' \left( a - a' \right) \right\| \leq 4 \varepsilon
\]
and the theorem follows from the closedness of \( \text{conv} \mathcal{S} a \).

Remark 2.16. As we have seen in the proof of Theorem 2.15 the decomposition \( b = b + d + \overline{b} \in \text{conv} \mathcal{S} a + [A^-, A^+]^\infty + \text{conv} \mathcal{S} a \) is not unique. However we can pick the \( b \) and \( \overline{b} \) such that \( d_i = 0 \) if \( b_i \neq 0 \) and \( d_i = 1 \) if \( b_i = 0 \). Remark 2.9.3 tells us that still \( \overline{b} \in \text{conv} \mathcal{S} a \) and \( b \in \text{conv} \mathcal{S} a \). These special \( b \) and \( \overline{b} \) will be minimal, that is they are the sequences with the smallest absolute value for each entry that still give a valid decomposition for \( b \).

Lemma 2.17. Let \( a \in l^\infty \) be arbitrary. We use the notation of Theorem 2.15. Then we have for every \( k \in \mathbb{N} \):

1. \( L_k(a) = L_k(\overline{a}) + k A^+; \)
2. \( L_k(-a) = L_k(-\overline{a}) - k A^- \).

We also have that
\[
\lim_{k \to \infty} \frac{1}{k} L_k(a) = \limsup a \quad \text{and} \quad \lim_{k \to \infty} -\frac{1}{k} L_k(-a) = \liminf a.
\]

Proof. The proof of (1) and (2) is almost the same as for Lemma 2.3. For (1) we distinguish two cases. If \( \overline{a} \) has at least \( k \) nonzero entries the assertion follows immediately, since the entries where \( a \) is nonzero are the largest entries of \( a \). So let us assume \( \overline{a} \) has only \( l < k \) nonzero entries. Then \( L_{k}(\overline{a}) = L_{l}(\overline{a}) \). Let these be the first \( l \) entries. Then \( a_1, ..., a_l \) are the \( l \) largest entries of \( a \) and any other entries of \( a \) are no greater than \( A^+ \). So for any
E ∈ C_k we have \( L_k(a) \leq \sum_{i=1}^{k-l} a_i + (k-l) A^+ = \sum_{i=1}^{k-l} \bar{a}_i + kA^+ = L_k(\bar{a}) + kA^+ \). But for every \( \varepsilon > 0 \) we can find entries \( \alpha_1, \ldots, \alpha_{k-l} \) of \( a \) that lie in \([A^+ - \varepsilon/k, A^+]\). Then

\[
L_k(a) \geq \sum_{i=1}^{k-l} \bar{a}_i + A^+ + (k-l) \left( A^+ - \frac{\varepsilon}{k-l} \right)
\]

\[
= \sum_{i=1}^{k-l} \bar{a}_i + kA^+ - \varepsilon
\]

So (1) follows immediately. We obtain (2) by applying (1) to \(-a\).

Now we prove

\[
\lim_{k \to \infty} \frac{1}{k} L_k(a) = \lim \sup \, a.
\]

“\( \geq \)”: For every \( \varepsilon > 0 \), \( k \in \mathbb{N} \), we have \( L_k(a) \geq k \lim \sup \, a - \varepsilon \).

“\( \leq \)”: We note that for any given \( \varepsilon > 0 \) there exist only \( m \) entries of \( a \) greater than \( \lim \sup \, a + \varepsilon \). So for \( k > m \) we have

\[
\frac{1}{k} L_k(a) \leq \frac{1}{k} \left( (a_1 + \cdots + a_m) + (\lim \sup \, a + \varepsilon)(k-m) \right)
\]

\[
= \frac{1}{k} \left( a_1 + \cdots + a_m \right) \left( 1 - \frac{m}{k} \right) \left( \lim \sup \, a + \varepsilon \right)
\]

\[
\lim_{k \to \infty} \lim \sup \, a + \varepsilon,
\]

which proves the assertion. Using \( \lim \inf \, a = -\lim \sup \, (-a) \), we obtain the last remaining claim. □

**Corollary 2.18.** Let \( a, b \in l^\infty \). Then \( b \in \text{conv} \, \overline{za} \) if and only if for all \( k \in \mathbb{N} \):

1. \( L_k(b) \leq L_k(a) \);
2. \( L_k(-b) \leq L_k(-a) \).

**Proof.** One inclusion follows directly from Lemma 2.2. Now let us assume \( b \) satisfies (1) and (2).

By Lemma 2.17 we have that \( \lim \sup \, b \leq \lim \sup \, a \) and \( \lim \inf \, b \geq \lim \inf \, a \). We use the notation of Theorem 2.15. Then we can write \( b = b' + \bar{b} \) where \( b' \in [A^-, A^+]^\infty \), \( \bar{b} \leq 0 \) and \( \bar{b} \geq 0 \). According to Remark 2.9.3 we can choose the decomposition such that \( b_i \neq 0 \) only if
\( b_i = A^- \) and \( \tilde{b}_i \neq 0 \) only if \( b_i = A^+ \). Lemma 2.17 tells us that \( L_k(a) = L_k(\tilde{a}) + kA^+ \). If \( \tilde{b} \) has at least \( k \) nonzero entries we have \( L_k(b) = L_k(\tilde{b}) + kA^+ \) and thus
\[
L_k(\tilde{b}) = L_k(b) - kA^+ \leq L_k(a) - kA^+ = L_k(\tilde{a}).
\]
If \( \tilde{b} \) has only \( m \) nonzero entries and \( m < k \) then \( L_k(\tilde{b}) = L_m(\tilde{b}) \leq L_m(\tilde{a}) \leq L_k(\tilde{a}) \). In this way we get for every \( k \) that \( L_k(\tilde{b}) \leq L_k(\tilde{a}) \).

Because \( b \) and \( \tilde{a} \) both are positive we conclude \( L_k(-b) = 0 = L_k(-\tilde{a}) \). By Proposition 2.8 we have \( \tilde{b} \in \text{conv } \tilde{a} \). In the same way we can show \( \tilde{b} \in \text{conv } \tilde{a} \). Theorem 2.15 then finishes the proof.

**Remark 2.19.** Let us have a look at the inverse of the problem, that is, we try to determine which sequences give the same set \( \text{conv } \tilde{a} \). Theorem 2.15, Lemma 2.17 and Corollary 2.18 tell us that \( \text{conv } \tilde{a}_1 = \text{conv } \tilde{a}_2 \) if and only if

1. \( \liminf a_1 = \liminf a_2 \),
2. \( \limsup a_1 = \limsup a_2 \),
3. \( L_k(\tilde{a}_1) = L_k(\tilde{a}_2) \) and \( L_k(a_1) = L_k(a_2) \) for all \( k \in \mathbb{N} \).

The third condition only affects entries that do not lie in \([A^-, A^+]\). Each such entry only occurs finitely many times in \( a_i \) or \( a_2 \). Condition (3) then tells us that each such entry must occur equally often in \( a_i \) and \( a_2 \). It is also interesting to notice that the entries that lie in \([A^-, A^+]\) do not affect the shape of \( \text{conv } \tilde{a} \).

The following two examples illustrate this:

Let \((a_1, a_2, a_3, ...) \in c_0\) be arbitrary. Then the sequences
\[
(a_1, a_2, a_3, ...)
\]
\[
(0, ..., 0, a_1, a_2, a_3, ...)
\]
\[
(0, a_1, 0, a_2, 0, a_3, 0, ...)
\]
and each permutation of one of these sequences will generate the same set \( \text{conv } \tilde{a} \). If \((a_1, a_2, ...) \) has entries that are 0 we can omit them as well.

Let \((a_1, a_2, a_3, ...) \in [0, 1]^\mathbb{N}\) be arbitrary. Then the sequences
\[
(3, 2, 2, -1, -2, 1, 0, 1, 0, 1, 0, ...)
\]
\[
(3, 2, 2, -1, -2, 1, a_1, 0, a_2, 1, a_3, 0, a_4, 1, ...)
\]
\[
(3, 2, 2, -1, -2, \frac{3}{1}, \frac{7}{1}, \frac{1}{1}, \frac{15}{1}, \frac{1}{1}, ...)
\]
and each permutation of one of these sequences will generate the same set \( \text{conv } \tilde{a} \).
3. THE SCHUR HORN THEOREM

Consider the Hilbert space $l^2 = l^2(\mathbb{N})$ (which is unitarily isomorphic to every separable Hilbert space). Then every bounded linear operator can be written as an $\mathbb{N} \times \mathbb{N}$ matrix

$$
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots \\
a_{21} & a_{22} & a_{23} & \cdots \\
a_{31} & a_{32} & a_{33} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}.
$$

Let $D$ denote the space of all bounded diagonal matrices. Then $D$ is isomorphic to $l^\infty(\mathbb{N})$, where $D$ is equipped with the operator norm. Let $p$ denote the projection on the diagonal:

$$
p\begin{pmatrix}a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33} \\
\vdots & \vdots & \vdots
\end{pmatrix} = (a_{11}, a_{22}, a_{33}, \ldots).
$$

We consider $p$ as a continuous linear map from the bounded linear operators to $l^\infty$. For $(a_1, a_2, \ldots) \in l^\infty$ let $\text{diag}(a_1, a_2, \ldots)$ denote the diagonal matrix with the corresponding entries.

Let $\mathcal{U}(n)$ denote the group of unitary isomorphisms of $\mathbb{C}^n$ and $\mathcal{U}$ the group of unitary isomorphisms of $l^2$. We write $U.a := U^* \text{diag}(a) U$ and $\mathcal{U}.a := \{ U.a : U \in \mathcal{U} \}$. In finite dimensions we have the Schur-Horn Theorem:

**Theorem 3.1.** Let $(a_1, \ldots, a_n) \in \mathbb{R}^n$. Then

$$
p(\{ U^* \text{diag}(a_1, \ldots, a_n) U : U \in \mathcal{U}(n) \}) = \text{conv} \ \Xi_n(a_1, \ldots, a_n)
$$

**Proof.** This is [Ho54, Theorem 5].

We seek a similar result in the infinite dimensional case. We generalize the Schur–Horn Theorem for two algebras. Therefore we need some technical details.

**Definition 3.2.** An $\mathbb{N} \times \mathbb{N}$ matrix $(a_{ij})_{i,j}$ is called *doubly stochastic* (or d.s.), if and only if $0 \leq a_{ij}$ for all $i, j$ and

$$
\sum_{j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} a_{ij} = 1 \quad \forall i, j \in \mathbb{N}.
$$

In particular the $l^1$- and $l^\infty$-operator norm of each d.s. matrix equals 1.
Remark 3.3. Let \( d \in l^\infty \). Then for \( U = (u_{ij}) \in \mathfrak{U} \) we have \( p(U^* \text{ diag}(d) U) = Cd \) for \( C = |u_{ij}|^2 \) a doubly stochastic matrix.

**Proposition 3.4.** Let \( a \in l^\infty \) be an arbitrary sequence and \( D \) be a d.s. matrix. Then

\[
D \cdot a \in \text{conv} \mathfrak{Z}a
\]

**Proof.** As seen above, we can split \( a \) up into \( a = q + \tilde{a} + c \) where \( q \) and \( -c \) are positive sequences converging to 0 and \( \lim \inf q = \lim \inf c \leq c_j \leq \lim \sup c = \lim \sup a \) for all \( j \). Let \( D = (d_{ij}) \). Choose a \( k > 0 \). We can assume that \( \tilde{a}_1, \ldots, \tilde{a}_k \) are the \( k \) largest entries of \( \tilde{a} \). Then for \( E \in \mathfrak{E}_k \)

\[
L_k(D \cdot \tilde{a}) = \sum_{i \in E} (D \cdot \tilde{a})_i = \sum_{i \in E} \sum_{j=1}^\infty d_{ij} \tilde{a}_j \leq \sum_{i \in E} \sum_{j=1}^k d_{ij} \tilde{a}_j + \sum_{i \in E} \sum_{j=k+1}^\infty d_{ij} \tilde{a}_k
\]

\[
= \sum_{j=1}^k \sum_{i \in E} d_{ij} \tilde{a}_j + \sum_{j=1}^k \sum_{i \in E} d_{ij} \tilde{a}_k - \sum_{i \in E} \sum_{j=1}^k d_{ij} \tilde{a}_k
\]

\[
= \sum_{j=1}^k \left( \sum_{i \in E} d_{ij} \tilde{a}_j + \left( 1 - \sum_{i \in E} d_{ij} \right) \tilde{a}_k \right) \leq \sum_{j=1}^k \tilde{a}_j = L_k(\tilde{a}).
\]

Therefore \( D \cdot \tilde{a} \in \text{conv} \mathfrak{Z}a \), since \( \tilde{a} \) has no negative entries. The same proof works for \( a \). On the other hand it is obvious, that

\[
D \cdot c \in [\lim \inf c_j, \lim \sup c_j]^{\mathfrak{Z}a}
\]

So Theorem 2.15 finishes the proof.

The first algebra we will look at is \( B_2 \), the algebra of Hilbert–Schmidt operators on \( l^2(\mathbb{N}) \). Every Hilbert–Schmidt operator is in particular compact, and therefore every hermitian Hilbert–Schmidt operator is diagonalizable. The diagonal Hilbert–Schmidt operators equipped with the Hilbert–Schmidt norm are canonically isomorphic to \( l^2(\mathbb{N}) \). The corresponding unitary group is \( \mathfrak{U}_2 := \mathfrak{U} \cap (1 + B_2) \). We have

**Theorem 3.5.** Let \( a \in l^2 \). Then

\[
p(\mathfrak{U}_2, a) = \text{conv} \mathfrak{Z}a
\]

**Proof.** Proposition 2.8 and Remark 2.11.2 tell us that \( \text{conv} \mathfrak{Z}a = \text{conv} \mathfrak{Z}a \). Then “\( \subseteq \)” follows directly from Remark 3.3 and Proposition 3.4, and “\( \supseteq \)” is a direct consequence of Theorem 3.1.
Now we will look at the algebra of all bounded endomorphisms of $l^2$. In the finite dimensional case there is at least one diagonal matrix in each $U(n)$-orbit of a hermitian operator. So Theorem 3.1 gives us the projection of the $U(n)$-orbit of any hermitian operator. In the infinite dimensional case we have to assume that the considered matrix is diagonalizable. We treat the other case in the next section. Because all diagonal entries of hermitian matrices are real, we only have to consider real valued sequences $a$. It will turn out that

$$p(U(a_1, a_2, ...)) = \text{conv } \mathcal{E}(a_1, a_2, ...)$$

but we cannot omit the closure on either side. First we prove that $p(U.a) \subseteq \text{conv } \mathcal{E}a$ and then that the left hand side is dense.

**Theorem 3.6.** Let $a \in l^\infty$. Then

$$p(U.a) \subseteq \text{conv } \mathcal{E}a.$$ 

**Proof.** This follows immediately from Remark 3.3 and Proposition 3.4. □

Remark 3.7. It can be shown quite easily that the left hand side is not closed in general. We consider the sequence $a = (1, 0, 1, 0, 1, 0, ...)$. There exists a basis $\{v_i, w_i\}$ of $l^2$ such that the $v_i$ are eigenvectors of $A := \text{diag}(a)$ with eigenvalue $1$ and $w_i$ those with eigenvalue $0$. Now we switch to another o.n.b. To do so we look at $V_1 = \text{span}\{v_1, \ldots, v_{2^n-1}, w_1\}$. Choose a $2^n \times 2^n$-matrix $(b_{ij})$ as in Lemma 3.8. Then the vectors $z_i = (\sqrt{2^n})^{-1} (\sum_{j=1}^{2^n-1} b_{ij} v_j + b_{i2^n} w_1)$ form another orthonormal basis of $V_1$. The diagonal entries of $A_{|V_1}$ with respect to this basis are

$$A_i = \langle A z_i, z_i \rangle = \frac{1}{2^n} \left( \sum_{j=1}^{2^n-1} b_{ij} v_j, \sum_{j=1}^{2^n-1} b_{ij} v_j + b_{i2^n} w_1 \right) = \frac{2^n - 1}{2^n} = 1 - \frac{1}{2^n}.$$ 

We do the same with all $V_k = \text{span}\{v_k(2^n + 1), \ldots, v_k(2^n - 1), w_k\}$ and obviously $l^2$ is the orthogonal direct sum of the $V_k$. Therefore in this o.n.b. the diagonal entries of $A$ are all $1 - 1/2^n$. Since this basis change is unitary

$$\left( 1 - \frac{1}{2^n}, 1 - \frac{1}{2^n}, \ldots \right) \in p(U.(1, 0, 1, 0, \ldots))$$

and obviously $(1, 1, 1, 1, \ldots) \notin p(U.(1, 0, 1, 0, \ldots))$. So $p(U.(1, 0, 1, 0, \ldots))$ is not closed.

On the other hand we have to take the $l^\infty$ closure on the right hand side, because with a construction similar to the one above we get a $U$ such that $p(U.a)$ has infinitely many entries infinitely often. This sequence lies in no $l^p$ closure of $\text{conv } \mathcal{E}a$ for any $p < \infty$. 

THE SCHUR-HORN CONVEXITY THEOREM
Lemma 3.8. For every $n \in \mathbb{N}$ there exist a $2^n \times 2^n$ matrix $B$ with entries only 1 or $-1$ and pairwise orthogonal row-vectors.

Proof. We prove the statement by induction: $n = 0$ is obvious. If $B$ is a $2^{n-1} \times 2^{n-1}$ matrix with the above properties, then

$$
\begin{pmatrix}
B & B \\
B & -B
\end{pmatrix} \in M_{2^n}
$$

is as desired.

Now we show, that $p(\mathbb{U} \cdot a)$ is dense in $\text{conv} \, \overline{\mathbb{S}}$. We do this in several steps.

Lemma 3.9. Let $a \in c_0(\mathbb{N})$. Then $p(\mathbb{U} \cdot a)$ is dense in $\text{conv} \, \overline{\mathbb{S}}$.

Proof. This follows directly from the fact that for $a \in c_0(\mathbb{N})$ we have $\text{conv} \, \overline{\mathbb{S}} = \text{conv} \, \overline{\mathbb{S}}_a$ (see Remark 2.11) and Theorem 3.1.

Remark 3.10. Before we consider the other cases, let us note the following fact: Let $A$ denote an arbitrary subset of $l^\infty(\mathbb{N})$ and let us assume we know for every $a \in A$ that $p(\mathbb{U} \cdot a)$ is dense in $\text{conv} \, \overline{\mathbb{S}}$. Then the same holds for every $a_0 \in A$.

Proof. Pick an $a \in A$, $b \in \text{conv} \, \overline{\mathbb{S}}$ and $\varepsilon > 0$. Then there exists a $b' = \sum \lambda_j \sigma_j a$ with $\|b - b'\| \leq \varepsilon$. We can find an $a' \in A$ with $\|a - a'\| \leq \varepsilon$. Since $a' \in A$ there exists a $U_0 \in \mathbb{U}$ such that $\|p(\mathbb{U} \cdot \text{diag}(a') U) - \sum \lambda_j \sigma_j a'\| \leq \varepsilon$. Now we have

$$
\|p(\mathbb{U} \cdot \text{diag}(a) U) - b\| \leq \|p(\mathbb{U} \cdot \text{diag}(a') U) - p(\mathbb{U} \cdot \text{diag}(a) U)\|
$$

$$
\quad + \|p(\mathbb{U} \cdot \text{diag}(a') U - \sum \lambda_j \sigma_j a'\|
$$

$$
\quad + \|\sum \lambda_j \sigma_j a' - \sum \lambda_j \sigma_j a\| + \|\sum \lambda_j \sigma_j a' - b\|
$$

$$
\leq \|p(\mathbb{U} \cdot \text{diag}(a - a') U)\| + \varepsilon + \|\sum \lambda_j \sigma_j (a' - a)\| + \varepsilon
$$

$$
\leq 4\varepsilon,
$$

which establishes the remark.

Lemma 3.11. Let $a \in l^\infty(\mathbb{N})$ such that $A^- := \liminf a \subseteq a \subseteq A^+ := \limsup a$ for all $j \in \mathbb{N}$. Then

$$
p(\mathbb{U} \cdot a) = [A^-, A^+]^\mathbb{N}.
$$
Proof. If $A^{-} = A^{+}$, then $a$ is constant and we are done. So we assume that $A^{-} = 0$ and $A^{+} = 1$.

In view of Remark 3.10, we may assume that $a$ is in $\mathbb{R}$ and has infinitely many entries 0 and 1. After a permutation we can write

$$a := (1, 0, x_1, 1, 0, x_2, 1, 0, x_3, ...)$$

where $x_i \in \{0, 1\}$.

Now we pick $b \in \mathbb{R}$ that has no entry 0 and has at least one entry different from 1 infinitely often. These sequences form a dense subset of $\{0, 1\}^{\mathbb{N}}$, so if we can show that each of them lies in $p(U.a)$ we are done. Then

$$b := (\beta_1, ..., \beta_p, n_1, ..., n_q, n_1, ..., n_q, ...)$$

We pick $\varepsilon > 0$. Since at least one entry different from 0 and 1 appears infinitely often in $b$ we can assume that $n_1 \in [0, 1[$. There exists a $k \in \mathbb{N}$ such that $\beta_1 + \cdots + \beta_p \leq k\varepsilon$. We let $\delta := (\beta_1 + \cdots + \beta_p)/k \leq \varepsilon$. Since $b \in \mathbb{R}$ we can assume (by enlarging $k$) that $kn_1$ is an integer and $\delta \leq n_1$. Now we permute $a$ and $b$ to

$$a' = (1, 1, 0, 0, 1, 0, x_1, 1, 0, x_2, 1, 0, x_3, ...)$$

and

$$b' = (\beta_1, ..., \beta_p, n_1, ..., n_q, n_1, ..., n_q, ...)$$

By Lemma 2.1 and Theorem 3.1 there is a $U' \in U(k + p)$ such that

$$p(U'^* \text{ diag}(1, ..., 1, 0, ..., 0) U') = (\beta_1, ..., \beta_p, n_1 - \delta, ..., n_1 - \delta).$$

We consider the sequences $a'' = (1, 0, x_1, 1, 0, x_2, ...)$ and $b'' = (n_1, ..., n_q, n_1, ...)$. We know that $b'' \in \mathbb{R}$ for some $l$. We can assume that $1/l \leq \varepsilon$. There are integers $0 < m_2 \leq l$ such that $n_1 = m_2/l$. We permute $a''$ and $b''$ to

$$a'' = (1, ..., 1, x_1, 0, ..., 0, 1, ..., 1, x_2, 0, ..., 0, 1, ..., 1, x_3, +, ..., 0, ...)$$

and

$$b'' = (n_1, ..., n_q, n_1, ..., n_q).$$
and 

\[ b^* = (n_1, ..., n_1, n_2, ..., n_2, ..., n_q, ..., n_q, n_1, ..., n_1, ...). \]

By Lemma 2.1 and Theorem 3.1 we can find \( U_{s,t} \in \mathcal{H}(m) \) for every \( s,t \in \mathbb{N} \) such that 

\[
p(U_{s,t}^* \text{ diag}(1, ..., 1, \frac{1}{s}, ..., 0, ..., 0) U_{s,t}) = \left( \begin{array}{cc}
\frac{1 - \frac{s}{l}}{l} & \vdots \\
\vdots & \frac{1 - \frac{s}{l}}{l}
\end{array} \right).
\]

Since \( b \in \mathcal{R}_l \) we have that \( n_s \geq 1/l \), so none of the entries will be negative. Forming 

\[
U := \left( \begin{array}{ccc}
U' & U_{1,1} & \cdots \\
& U_{2,1} & \\
& & \ddots & U_{q,1} \\
& & & U_{1,2} & \cdots \\
& & & & \ddots
\end{array} \right)
\]

we obtain 

\[
\| p(U^* \text{ diag}(a) U) - b \| \leq \sup_{s,t} \frac{1 - \frac{s}{l}}{l} \leq \frac{1}{l} \leq \varepsilon.
\]

This proves the lemma.

**Proposition 3.12.** Let \( a \in l^{\infty}([n]) \). Then \( p(\mathcal{H}(a)) \) is dense in \( \text{conv} \mathcal{Z}a \).

**Proof.** If \( a \) converges this is Lemma 3.9. So we can assume that \( \lim \sup a_i = 1 \) and \( \lim \inf a_i = 0 \). We use Remark 3.10 and restrict ourselves to the case that \( a \) has only finitely many \( a_i \) with \( a_i > 1 \) or \( a_i < 0 \), but infinitely many entries 1 and 0.

Now we pick an arbitrary \( b \in \text{conv} \mathcal{Z}a \) and an \( \varepsilon > 0 \). Theorem 2.15 tells us that \( b = b + d + b \) with \( b \in \text{conv} \mathcal{Z}a, b \in \text{conv} \mathcal{Z}a \) and \( d \in [0, 1]^N \). In particular we have \( \lim \sup b \leq 1 \) and \( \lim \inf b \geq 0 \). We pick \( b \) and \( b \) minimal in the sense of Remark 2.16. Further \( b \) will have at least one subsequence
(\(b_n\))_n converging to some \(B' \in [0, 1]\). Then there exists a rational \(B \in ]0, 1[\) with \(|B - B'| \leq \frac{\varepsilon}{2}\). We define \(b'\) by

\[
\begin{align*}
  b'_i := \\
  \begin{cases} \\
    1 & b_i \in [1, 1 + \varepsilon], \\
    0 & b_i \in [-\varepsilon, 0], \\
    B & b_i \in [B' - \varepsilon/2, B' + \varepsilon/2], \\
    b_i & \text{otherwise}. \\
  \end{cases}
\end{align*}
\]

By Proposition 2.8 and Theorem 2.15 \(b' \in \text{conv} \, \mathcal{S}a\), \(b'\) has only finitely many entries not in \([0, 1]\), one rational entry \(B\) infinitely often and \(|b' - b| \leq \varepsilon\). So we only have to show \(b \in \rho(U.a)\) for \(b\) with only finitely many entries not in \([0, 1]\) and a rational entry in \([0, 1]\) occurring infinitely often. This will finish the proof since the sequences with these properties lie densely in \(\text{conv} \, \mathcal{S}a\).

Let \(x^+_1 \geq \cdots \geq x^+_p\) denote the nonzero entries of \(a\) and \(x^-_1 \leq \cdots \leq x^-_q\) those of \(a\). We denote by \(\beta^+_1 \geq \cdots \geq \beta^+_p\) and \(\beta^-_1 \leq \cdots \leq \beta^-_q\) the nonzero entries of \(b\) and \(b\). We define \(\delta := x^+_1 + \cdots + x^+_p - \beta^+_p - \cdots - \beta^-_q\). By Proposition 2.8 and Theorem 2.15 we have \(\delta \geq 0\). We know there exists a \(c \in \mathbb{N}\) such that \(cB\) an integer, \(\delta/c \leq \varepsilon\), \(\delta/c \leq 1 - B\) and \(cB \geq r - p\).

Now we look at

\[
\begin{align*}
  a^+ & := (x^+_1 + 1, \ldots, x^+_p + 1, 1, \ldots, 1, 0, \ldots, 0), \\
  b^+ & := (\beta^+_1 + 1, \ldots, \beta^+_p + 1, B + \frac{\delta}{c}, \ldots, B + \frac{\delta}{c}), \\
  a^- & := (x^-_1, \ldots, x^-_q, 0, \ldots, 0), \\
  b^- & := (\beta^-_1, \ldots, \beta^-_q, B + \frac{\delta}{c}, \ldots, B + \frac{\delta}{c}).
\end{align*}
\]

By Lemma 2.1 and Theorem 3.1 there exists \(U_+ \in \mathcal{U}(c + p)\) such that \(p(U_+ \cdot \text{diag}(a^+)) U_+ \cdot \mathcal{S}a = b^+\). In the same way we find \(c' \in \mathbb{N}\) such that \(c'B\) is an integer and for \(\delta' := x^-_1 + \cdots + x^-_q - \beta^-_1 - \cdots - \beta^-_q \leq 0\) we have \(|\delta'/c' \leq \varepsilon\), \(|\delta'/c' \leq B\) and \(c'(1 - B) \geq s - q\). Then for

\[
\begin{align*}
  a^- & := (x^-_1, \ldots, x^-_q, 0, \ldots, 0, 1, \ldots, 1), \\
  b^- & := (\beta^-_1, \ldots, \beta^-_q, B + \frac{\delta'}{c'}, \ldots, B + \frac{\delta'}{c'}). \\
\end{align*}
\]

there exists a \(U_- \in \mathcal{U}(c' + q)\) with \(p(U_- \cdot \text{diag}(a^-)) U_- \cdot \mathcal{S}a = b^-\).

By our choice of \(a^+\) and \(b^+\) we can permute \(a\) to \(a' := (a^+, a^-, a'^+)'\) and \(b\) to \(b' := (b^+, b^-, b'^+)'\), where \(a', b' \in [0, 1]^n\) and \(\limsup a' = 1\), \(\liminf a' = 0\).
By Lemma 3.11 we can find a $U \in \mathcal{U}$ such that $\|p(U^* \text{diag}(a') U^{-1}) - b^*\| \leq \varepsilon$.

Now

$$U := \begin{pmatrix} U_+ & \text{ } \\ U_- & U \end{pmatrix}$$

satisfies $\|p(U^* \text{diag}(a) U) - b\| \leq \varepsilon$. This proves the theorem.

**Theorem 3.13.** Let $a \in l^\infty$. Then

$$\overline{\rho(U, a)} = \text{conv } \Xi a.$$  

**Proof.** This follows immediately from Theorem 3.6 and Proposition 3.12.

4. NONDIAGONALIZABLE OPERATORS

In this section we look at the case that a given hermitian operator $A$ is nondiagonalizable. We write $U.A := U^*AU$ for $U \in \mathcal{U}$ and $\mathcal{U}.A := \{U.A: U \in \mathcal{U}\}$. Our aim is to describe the set $\overline{\rho(\mathcal{U}.A)}$. We recall first the Spectral Theorem:

**Theorem 4.1.** Let $A \in B(\mathfrak{H})$ be a normal operator and let $\sigma(A)$ denote the spectrum of $A$. Then there exists a spectral measure $P$ on $\sigma(A)$ equipped with the Borel $\sigma$-algebra such that

$$A := \int_{\sigma(A)} x \, dP(x)$$

and the mapping

$$\varphi: L^\infty(\sigma(A)) \to B(\mathfrak{H}); \quad f \mapsto \int_{\sigma(A)} f(x) \, dP(x)$$

is an embedding.

Now let us fix a hermitian operator $A$. Then $\sigma(A) \subseteq \mathbb{R}$. By Theorem 4.1 we can write $A = \int_{\sigma(A)} x \, dP(x)$. Define

$$\sigma_p(A) := \{ x \in \sigma(A): P(\{x\}) \neq \{0\} \},$$

and

$$\sigma_p(A) := \{ x \in \sigma(A): P(\{x\}) \neq \{0\} \}.$$
the point spectrum of \( \mathcal{A} \). We set

\[
\mathcal{H}_p := \bigoplus_{x \in \sigma_p(\mathcal{A})} P(\{x\}) \mathcal{H} \quad \text{and} \quad \mathcal{H}_c := \mathcal{H}_p^\perp.
\]

The subspaces \( \mathcal{H}_p \) and \( \mathcal{H}_c \) are both \( \mathcal{A} \)-invariant. Further \( \mathcal{A} \) is diagonalizable on \( \mathcal{H}_p \) and has no eigenvectors in \( \mathcal{H}_c \). From now on we will assume that \( \mathcal{H}_c \neq \{0\} \), as the other case was already done in Theorem 3.13. Then \( \mathcal{A}' := \mathcal{A}|_{\mathcal{H}_c} \) is also hermitian. Its spectrum is real and has a minimum \( \lambda_0 \) and a maximum \( \lambda_1 \). Since \( \mathcal{H}_c \) is nontrivial, we have \( \lambda_0 < \lambda_1 \).

**Lemma 4.2.** Let \( \mathcal{A} \) be a hermitian operator with empty point spectrum, i.e., \( \mathcal{H}_p = \mathcal{H}_c \). Denote by \( \lambda_0 \) the minimum and by \( \lambda_1 \) the maximum of \( \sigma(\mathcal{A}) \). Then for every \( \varepsilon > 0 \) there exists a diagonalizable operator \( \mathcal{A}_\varepsilon \) satisfying

1. \( \mathcal{A}_\varepsilon \) has only finitely many eigenvalues, and each eigenvalue has an infinite dimensional eigenspace.
2. The smallest eigenvalue of \( \mathcal{A}_\varepsilon \) is \( \lambda_0 \) and the largest is \( \lambda_1 \).
3. \( \| \mathcal{A} - \mathcal{A}_\varepsilon \| \leq \varepsilon \).

**Proof.** Let \( \mathcal{A} = \int_{\sigma(\mathcal{A})} x dP(x) \). Since \( \mathcal{H}_p = \mathcal{H}_c \) no \( x \in \sigma(\mathcal{A}) \) has nonzero measure. Now we pick a step-function \( s_\varepsilon \) on \( [\lambda_0, \lambda_1] \) such that \( s_\varepsilon = \lambda_0 \) in a neighborhood of \( \lambda_0 \), \( s_\varepsilon = \lambda_1 \) in a neighborhood of \( \lambda_1 \) and \( \|s_\varepsilon - \text{id}\|_\infty \leq \varepsilon \). Then

\[
\mathcal{A}_\varepsilon := \int_{\sigma(\mathcal{A})} s_\varepsilon(x) dP(x)
\]

fulfills (1)–(3). Since \( \sigma_\varepsilon \) is a step-function \( \mathcal{A}_\varepsilon \) is obviously diagonalizable and has only finitely many eigenvalues. For every open \( V \subseteq \sigma(\mathcal{A}) \) we have \( \dim(P(V), \mathcal{H}) = \infty \), because otherwise we would have \( P(V), \mathcal{H} \cong \mathcal{H}_p \). So each eigenspace will be infinite dimensional. The condition (2) follows directly from the special shape of \( s_\varepsilon \) at the borders of the interval \( [\lambda_0, \lambda_1] \) and condition (3) is a direct consequence of \( \|s_\varepsilon - \text{id}\|_\infty \leq \varepsilon \). –

**Theorem 4.3.** Let \( \mathcal{A} \) be an arbitrary hermitian operator on \( \mathcal{H} \). Let \( \mathcal{H} = \mathcal{H}_p \oplus \mathcal{H}_c \). Write \( \mathcal{A}' := \mathcal{A}|_{\mathcal{H}_c} \) and denote by \( \lambda_0 \) and \( \lambda_1 \) the minimum and maximum of \( \sigma(\mathcal{A}') \). Then

\[
\overline{p(U, \mathcal{A})} = \overline{p(\mathcal{H}, \mathcal{A})}
\]

where \( \mathcal{A}' = \mathcal{A} \) on \( \mathcal{H}_p \) and \( \mathcal{A}' = \text{diag}(\lambda_0, \lambda_1, \lambda_1, ...) \) on \( \mathcal{H}_c \) in some orthonormal basis.
Proof. First we note that the choice of basis in $\mathcal{S}_p$ does not change the set $\mathbb{U}$. $\mathcal{A}$.

On $\mathcal{S}_p$ we find a basis such that $A|_{\mathcal{S}_p} = \text{diag}(a_p)$ for some $a_p \in l^\infty(\mathbb{N})$. Let $\epsilon > 0$. By Lemma 4.2 we can find a diagonalizable operator $A'_p$ on $\mathcal{S}_p$ such that $\|A' - A'_p\| \leq \epsilon$ and $A'_p = \text{diag}(x_0, x_1, x_2, ...)$ with $\lambda_i \in [x_0, x_1]$. We define operators $A_p$ by

$$A_p v := \left\{ \begin{array}{ll}
A v & v \in \mathcal{S}_p^+ \\
A'_p v & v \in \mathcal{S}_p^-
\end{array} \right.$$  

Then $A_p$ is diagonalizable. By Theorem 2.15 and Theorem 3.13 we have

$$p(\mathbb{U} A_p) = p(\mathbb{U} A)$$

since for any operator $\text{diag}(a)$ for $a \in l^\infty(\mathbb{N})$ the set $\overline{\text{conv}}(\overline{a})$ depends only on $\lim \sup a_i, \lim \inf a_i$ and the entries in $a$ that do not lie between these two values. It does not matter whether or not $A'_p$ and $A_p$ are diagonal in different bases, since obviously for any $U \in \mathbb{U}$ we have $p(\mathbb{U} (U^{-1} A U)) = p(\mathbb{U} A)$.

"$\leq$": Choose $U$ in $\mathbb{U}$. Pick $\epsilon > 0$. Then $\|p(U^* A U) - p(U^* A U')\| = \|p(U^* (A - A'_p) U)\| \leq \|A - A'_p\| \leq \epsilon$. Here we make use of the fact that the mapping $A \mapsto p(U^* A U)$ has norm 1 for every unitary matrix $U$. The distance between $p(U^* A U)$ and $p(\mathbb{U} A_p)$ is smaller than $\epsilon$. Since $\epsilon$ was arbitrary the distance is zero. So $p(U^* A U) \in p(\mathbb{U} A)$ for any $U$ and therefore $p(\mathbb{U} A) \subseteq p(\mathbb{U} A)$.

"$\geq$": Pick $U$ in $\mathbb{U}$ arbitrary and $\epsilon > 0$. Then there exists a $U'$ such that $\|p(U^* A U') - p(U^* A U)\| \leq \epsilon$. We have $\|p(U^* A U') - p(U^* A U)\| \leq \|p(U^* (A - A'_p) U')\| + \|p(U^* A U') - p(U^* A U)\| \leq 2\epsilon$. With the same argumentation as in the last paragraph we get the assertion. \]

Lemma 4.4. For a given hermitian operator $A$ let $\tilde{A}$ be defined as in Theorem 4.3. Then $\tilde{A}$ lies in $\text{conv}(\mathbb{U} \tilde{A})$.

Proof. Since $\tilde{A}$ differs from $A$ only on $\mathcal{S}_p$, we can assume $\mathcal{S} = \mathcal{S}_p$. For the matrix representation of the operators we pick a basis on which $\tilde{A}$ is diagonal, i.e., $\tilde{A} = \text{diag}(x_0, x_1, x_2, ...)$ and $\tilde{A}_p$ as in Lemma 4.2. Then there is $U \in \mathbb{U}$ such that $U^{-1} \tilde{A} U = \text{diag}(x_0, x_1, x_2, ...)$ with $\lambda_i \in [x_0, x_1]$. By Corollary 2.14 there are $\mu_i \in [0, 1], \sum_i \mu_i = 1$ and $\sigma_i \in \mathcal{S}$ such that $\sum_i \mu_i \sigma_i (x_0, x_1, \lambda_1, x_0, x_1, \lambda_2, ...) = (x_0, x_1, x_0, x_1, ...)$ with $\|\sum_i \mu_i \sigma_i (x_0, x_1, \lambda_1, x_0, x_1, \lambda_2, ...) - (x_0, x_1, x_0, x_1, ...)\| \leq \epsilon$. But every permutation $\sigma$ of the entries of a diagonal matrix can be achieved by conjugating with a permutation matrix $S$, which in particular is unitary. Now
\[ \sum_i \mu_i (US_\varepsilon)^{-1} AUS_\varepsilon - \bar{A} \leq \left\| \sum_i \mu_i ((US_\varepsilon)^{-1} AUS_\varepsilon - (US_\varepsilon)^{-1} A_x US_\varepsilon) \right\| \\
+ \left\| \sum_i \mu_i S_\varepsilon^{-1} U^{-1} A_x US_\varepsilon - \bar{A} \right\| \\
\leq \sum_i \mu_i \| A - \bar{A} \| + \left\| \sum_i \mu_i \sigma_i (x_0, x_1, \lambda_1, \sigma_0, x_1, \lambda_2, \ldots) \\
- (x_0, x_1, \sigma_0, x_1, \ldots) \right\| \\
\leq 2\varepsilon. \]

Since \( \varepsilon \) was arbitrary this proves the lemma. \[ \square \]

**Remark 4.5.** We can formulate the result of Theorem 4.3 in another way. For a given hermitian operator \( A = \int_{\sigma(A)} x \, dP(x) \) we define the essential spectrum of \( A \) to be

\[ \sigma_e(A) := \{ x \in \sigma(A) : (\forall V \text{ neighborhood of } x) \dim(P(V) \cap \sigma(A)) = \infty \}. \]

Then \( \sigma_e(A) \) is a closed subset of \( \sigma(A) \) with minimum \( \underline{\sigma}_e \) and maximum \( \overline{\sigma}_e \). We set \( \sigma^+ := \{ x \in \sigma(A) : x > \overline{\sigma}_e \} \) and \( \sigma^- := \{ x \in \sigma(A) : x < \underline{\sigma}_e \} \). We define

\[ A^+ := \int_{\sigma^+} (x - \overline{\sigma}_e) \, dP(x) \]
\[ A^- := \int_{\sigma^-} (x - \underline{\sigma}_e) \, dP(x) \]
\[ A_0 := \int_{\sigma^-} \underline{\sigma}_e \, dP(x) + \int_{\sigma(A) \cap \{ x \in [\underline{\sigma}_e, \overline{\sigma}_e] \}} x \, dP(x) + \int_{\sigma^+} \overline{\sigma}_e \, dP(x). \]

Then \( A = A^- + A_0 + A^+ \) and \( A^- \) and \( A^+ \) are compact hermitian operators and therefore can be diagonalized. Let \( a^+ (a^-) \) denote the diagonal of \( A^+ (A^-) \) with respect to some orthonormal basis. Then

\[ p(\Pi, A) = \text{conv} \, \Xi a^+ + [\underline{\sigma}_e, \overline{\sigma}_e] \cap \text{conv} \, \Xi a^- \].

It might be possible to get more detailed information in special cases, for example for trace class or Hilbert Schmidt operators. In these cases it might be true in general that \( p(\Pi, A) \) is closed. It would also be interesting...
to know whether \( Da \) lies in \( p(H, a) \) for every d.s. matrix \( D \). In the finite dimensional case this is true. Another approach might be to look at

\[
\text{conv}_\sigma Z a := \left\{ \sum_{i=1}^\infty \lambda_i \sigma_i a : 0 \leq \lambda_i \leq 1, \sum_{i=1}^\infty \lambda_i = 1 \right\}
\]

instead of \( \text{conv} Z a \)

5. APPLICATIONS

In this section we use the convexity theorem to study invariant convex sets and functions.

**Lemma 5.1.** Let \( V \) be a topological vector space, \( C \subseteq V \) convex and \( C^* \neq \emptyset \). Then \( C \subseteq C^* \).

**Proof.** Let us assume that the convex subset \( C \) of the topological vector space \( V \) has nonempty interior. We pick an \( x \in C \) arbitrary. We show that \( x \in C^* \). Pick \( y \in C^* \). Then there exists an open neighborhood \( U \) of 0 in \( V \) such that \( y + U \subseteq C^* \). Since \( C \) is convex we know that for every \( \lambda \in [0, 1] \)

\[
(1 - \lambda) x + \lambda (y + U) = (1 - \lambda) x + \lambda y + \lambda U \subseteq C.
\]

Thus for \( \lambda > 0 \) we have \( (1 - \lambda) x + \lambda y \in C^* \). So \( x \in C^* \). \[\square\]

In this section we will denote by \( \mathfrak{h} \) the set of all hermitian operators on \( l^2(\mathbb{N}) \). We recall that \( \mathfrak{D} \) denotes the space of diagonal operators. We let \( t := \mathfrak{D} \cap \mathfrak{h} \). Then \( \mathfrak{Z} \) acts on \( t \) as it does on \( \mathfrak{D} \), by permutation of entries. We identify \( t \) with \( l^\infty(\mathbb{N}, \mathbb{R}) \).

**Lemma 5.2.** Let \( C_1 \) be a closed convex \( \mathfrak{Z} \)-invariant subset of \( \mathfrak{Z} \). Then one of the two mutually exclusive cases occurs:

1. Every element in \( C_1 \) is a convergent sequence, that is viewing them as diagonal operators each element can be written in the form \( \lambda \text{id} + K \), where \( K \) is a compact operator. In this case the interior of \( C_1 \) is empty.

2. \( C_1 \) has nonempty interior and contains an interior point invariant under \( \mathfrak{Z} \), that is, its interior contains a real multiple of the identity.

**Proof.** Let \( C \) be a closed convex \( \mathfrak{Z} \)-invariant subset of \( t \). First we prove that \( C_1 \) cannot consist entirely of convergent sequences if it has nonempty interior. We pick a sequence \( \varepsilon \in C_1^* \) converging to \( \gamma \). Since \( \varepsilon \) lies in the interior we find an \( \varepsilon > 0 \) such that \( \{ \varepsilon' \in t : \| \varepsilon - \varepsilon' \| < \varepsilon \} \subseteq C \). It is then clear that there exists a sequence \( \varepsilon' \in C_1 \) with cluster points \( \gamma - \varepsilon/2 \) and \( \gamma + \varepsilon/2 \).
So if \( C_1 \) has nonempty interior we can find a non-convergent \( c \) in \( C_1 \). Then we have \( \text{conv} \mathcal{E} C \subseteq C_1 \). By Theorem 2.15 we get \( [\lim \inf c, \lim \sup c] \subseteq C_1 \). So for any \( \gamma \in ]\lim \inf c, \lim \sup c[ \) we have \( (\gamma, \gamma, \gamma, \ldots) \in C_1^0 \). This finishes the proof.

**Lemma 5.3.** Let \( C_1 \) be a closed convex \( \mathcal{E} \)-invariant subset of \( \mathfrak{t} \). If \( C_1^0 \neq \emptyset \) then \( (\mathbb{H}, C_1^0) \neq \emptyset \).

**Proof.** By Lemma 5.2 there is a \( \lambda \) such that \( \lambda \text{ id} \in C_1^0 \). Then for some \( \varepsilon > 0 \) we have that \( \{ c \in \mathfrak{t} : ||c - \lambda \text{ id}|| < \varepsilon \} \subseteq C_1 \). Now we claim that \( B := \{ X \in \mathfrak{h} : ||X - \lambda \text{ id}|| < \varepsilon \} \subseteq \mathbb{H} C_1 \). This proves the lemma.

Pick \( X \in B \) diagonalizable. Then we can find \( U \in \mathbb{H} \) such that \( U^{-1} X U \in \mathfrak{t} \). Further we get \( ||U^{-1} X U - \lambda \text{ id}|| = ||U^{-1} (X - \lambda \text{ id}) U|| = ||X - \lambda \text{ id}|| < \varepsilon \).

Therefore \( U^{-1} X U \in C_1 \) and \( X \in \mathbb{H} C_1 \). But the diagonalizable operators are dense in \( B \) and that proves our claim.

**Theorem 5.4.** Let \( p : \mathfrak{h} \to \mathfrak{t} \) be the projection on the diagonal as in Section 2.

1. For every closed convex \( \mathbb{H} \)-invariant subset \( C \) of \( \mathfrak{h} \) we have \( p(C) = C \cap \mathfrak{t} \).

2. The mapping \( P : C \mapsto p(C) \) is a bijection between the closed convex \( \mathbb{H} \)-invariant subsets of \( \mathfrak{h} \) and the closed convex \( \mathcal{E} \)-invariant subsets of \( \mathfrak{t} \). The set \( C \) has nonempty interior if and only if \( p(C) \) has nonempty interior. The inverse of \( P \) is given by \( P^{-1} : C_1 \mapsto \mathbb{H} C_1 \).

**Proof.** (1) For \( C \subseteq \mathfrak{h} \) a \( \mathbb{H} \)-invariant closed convex set let \( C_1 := C \cap \mathfrak{t} \). First we will show that \( \mathbb{H} C_1 \) is a dense subset of \( C \). If every element in \( \mathbb{H} C_1 \) is of the form \( \lambda \text{ id} + K \), \( K \) a compact operator, then every element in \( C \) is unitary diagonalizable. This implies \( C = \mathbb{H} C_1 \).

Now let us look at the case where we can find an \( A \in C \) not of the above form, that is the upper and lower bound of its essential spectrum \( \sigma_e(A) \) are different. Then either \( A \) is diagonalizable or at least \( C \) contains a diagonalizable element \( \tilde{A} \) for which \( \sigma_e(\tilde{A}) \) has different upper and lower bounds (Lemma 4.4). The \( \mathbb{H} \)-invariance of \( C \) gives us a \( X \in \tilde{C}_1 := \mathcal{C} \cap \mathfrak{t} \) that has a non-convergent diagonal. Since \( \text{conv} \mathcal{E} \tilde{C} \subseteq \tilde{C}_1 \) we see with Theorem 2.15 that \( C_1^0 \neq \emptyset \). By Lemma 5.3 \( C_1^0 \neq \emptyset \). Hence by Lemma 5.1 \( C \) has dense interior and since the diagonalizable operators form a dense subset of \( \mathfrak{h} \) we have \( \mathbb{H} \tilde{C}_1 = C \). Now we apply Theorem 2.15 and get \( p(C) = p(\mathbb{H} \tilde{C}_1) \subseteq p(\mathbb{H} C_1) = \text{conv} (\mathcal{E} C_1) = C_1 \).

(2) For every closed convex \( \mathcal{E} \)-invariant set \( C_1 \subseteq \mathfrak{t} \) define

\[
\hat{C} := \text{conv}(\mathbb{H} C_1)
\]
and
\[ \tilde{C} := \bigcap_{U \in \mathcal{U}} U^{-1}p^{-1}(C_t) U. \]

Then with (1) obviously \( \tilde{C} \) is the minimal and \( \hat{C} \) is the maximal closed convex \( \mathcal{U} \)-invariant subset of \( \mathfrak{h} \) with \( p(C) = C_t \). But in the proof of (1) we have seen that \( \mathcal{U}(C \cap t) \) is dense in \( C \). Thus \( \hat{C} = \tilde{C} \). So \( P \) is a bijection. If \( C_t^0 \neq \emptyset \), then by Lemma 5.3 \( \hat{C}^0 \neq \emptyset \). On the other hand if \( C_t^0 = \emptyset \) then by Lemma 5.2 we know that \( C_t \) and therefore \( \text{conv}(\mathcal{U}.C_t) \) consists of operators of the form \( \lambda \text{id} + K \) where \( K \) is compact. Therefore \( \text{conv}(\mathcal{U}.C_t) \) cannot have interior points.

Remark 5.5. This theorem gives rise to an interesting geometric observation. Let us consider an arbitrary hermitian operator \( A \). By Theorem 5.4 we know that \( \overline{p(\mathcal{U}.A)} = \text{conv}(\mathcal{U}.A) \cap t \). Theorem 2.15 and Theorem 4.3 together give us a very good description of this set. It depends only on the lower bound \( x_0 \) and the upper bound \( x_1 \) of the essential spectrum and the eigenspace-dimensions of the spectral values not in the interval \([x_0, x_1]\). But since there is a 1-1-correspondence between closed convex \( \mathcal{U} \)-invariant sets and their projections the same is true for \( \text{conv}(\mathcal{U}.A) \). In particular it will always contain a multiple of the identity, more precisely it will contain exactly all \( \lambda \text{id} \) for \( \lambda \in [x_0, x_1] \).

Theorem 5.6. A \( \mathcal{U} \)-invariant continuous function \( f \) on a \( \mathcal{U} \)-invariant closed convex set \( C = \mathcal{U} \subseteq \mathfrak{h} \) with nonempty interior is convex if and only if \( f|_{C \cap t} \) is convex.

Proof. One implication is obvious. Now let us assume \( f \) is convex on \( C \cap t \).

We let \( h' = \mathfrak{h} \oplus \mathbb{R} \) and \( t' = t \oplus \mathbb{R} \). On \( C' := C \oplus \mathbb{R} \) we look at
\[ E := \{ (X, \lambda) \in C \oplus \mathbb{R} : f(X) \leq \lambda \}, \]
the epigraph of \( f \). It is well known that \( f \) is convex if and only if \( E \) is convex. We know that \( E_t := E \cap t' \) is convex, since \( f \) is convex on \( C \cap t \). Let \( \mathcal{U} \) and \( \mathcal{S} \) act on \( C' \) resp. \( C' \cap t' \) by operating on the first component. Since \( f \) is \( \mathcal{U} \)-invariant so is \( E \), and thus \( E_t \) is \( \mathcal{S} \)-invariant.

The set \( E \) is obviously closed and since \( f \) is continuous it has dense interior in \( C \oplus \mathbb{R} \). Let \( pr := p \oplus \text{id} : \mathfrak{h} \oplus \mathbb{R} \to t \oplus \mathbb{R} \). First we show that \( E_t = pr(E) \). Theorem 5.4.2 tells us that \( \mathcal{U}(C \cap t) \) lies dense in \( C \) and therefore using the fact that \( E \) is \( \mathcal{U} \)-invariant, we see that \( \mathcal{U}.E_t \) lies dense in \( E \). Therefore \( pr(E) = pr(\mathcal{U}.E_t) \subseteq pr(\mathcal{U}.E_t) \subseteq \text{conv} \mathcal{S}E_t = E_t \). Now we prove that
\[ E = \tilde{E} := \bigcap_{U \in \mathcal{U}} U.pr^{-1}(E_t). \]
This will prove the theorem, since $\tilde{E}$ as an intersection of convex sets again is convex. We observe that $E_t = \text{pr}(E) = \text{pr}(U^{-1}E)$ and therefore $\tilde{E} = U(U^{-1}E) \subseteq U \cdot \text{pr}^{-1}(E_t)$ for every $U \in \mathcal{U}$. So we get $E \subseteq \tilde{E}$ and $\text{pr}(E) \subseteq \text{pr}(\tilde{E})$. But by definition $\text{pr}(\tilde{E}) \subseteq \text{pr}(E)$. So we have $\text{pr}(E) = \text{pr}(\tilde{E})$. This tells us that $\tilde{E} \cap U^{-1} \subseteq \text{pr}(\tilde{E}) = \text{pr}(E) = E_t$. In $E$ and $\tilde{E}$ the elements with diagonalizable operators in the first component form a dense subset so we get

$$\tilde{E} = \bigcup (\tilde{E} \cap U^{-1}) \subseteq \bigcup E_t = E,$$

which proves the claim.

**Remark 5.7.** In Theorem 5.6 it is necessary to assume a priori that $f$ is continuous. Even if $f |_{C^{\infty}}$ is continuous and convex $f$ does not need to be either as can be seen by looking at the characteristic function of the set of diagonalizable operators.

**REFERENCES**


