BIRKHOFF'S THEOREM AND MULTIDIMENSIONAL NUMERICAL RANGE

YU. SAFAROV

Abstract. We show that, under certain conditions, Birkhoff's theorem on doubly stochastic matrices remains valid for countable families of discrete probability spaces which have nonempty intersections. Using this result, we study the relation between the spectrum of a self-adjoint operator A and its multidimensional numerical range. It turns out that the multidimensional numerical range is a convex set whose extreme points are sequences of eigenvalues of the operator A. Every collection of eigenvalues which can be obtained by the Rayleigh–Ritz formula generates an extreme point of the multidimensional numerical range. However, it may also have other extreme points.

Recall that a (possibly infinite) matrix is said to be *doubly stochastic* if all its entries are non-negative and the sum of entries in every row and every column is equal to one. Birkhoff's theorem [B] says that

(i) the extreme points of the convex set of doubly stochastic matrices are permutation matrices and

(ii) the set of doubly stochastic matrices coincides with the closed convex hull of the set of permutation matrices.

The first aim of this paper is to show that, under certain conditions, Birkhoff's theorem remains valid for a countable family of discrete probability spaces which have nonempty intersections (see Remark 2.1). We join every two points lying in the same probability space by an edge and reformulate the problem in terms of weighted graphs. It turns out that **(i)** and **(ii)** hold true whenever the underlying graph satisfies the conditions ($\mathbf{g}_1$)–($\mathbf{g}_3$) introduced in Section 2. The conditions ($\mathbf{g}_1$) and ($\mathbf{g}_3$) are purely technical and can probably be removed or weakened. The geometric condition ($\mathbf{g}_2$) is necessary (see Remark 2.5).

The second aim of the paper is to study the relation between the spectrum of a self-adjoint operator A and its m-dimensional numerical range $\Sigma(m,A)$. The latter is defined as the set of all m-dimensional vectors of the form $\{Q_A[u_1], Q_A[u_2], \dots\}$, where Q_A is the corresponding quadratic form, $\{u_1, u_2, \dots\} \subset \mathcal{D}(Q_A)$ is an arbitrary orthonormal set containing m elements

Date: August 2004.

1991 *Mathematics Subject Classification.* 47A12, 05C50.

Key words and phrases. Stochastic matrices, weighted graphs, Birkhoff's theorem, numerical range, extreme points, variational principle.

and $m = 1, 2, \dots, \infty$. Using an infinite dimensional version of Birkhoff's theorem, we prove that

- **(1)** the m-dimensional numerical range $\Sigma(m, A)$ is a convex set,
- **(2)** the extreme points of $\Sigma(m, A)$ belong to the corresponding m-dimensional point spectrum $\sigma_p(m, A)$,
- **(3)** every collection of m lowest or highest eigenvalues which can be found with the use of the Rayleigh–Ritz formula generates an extreme point of $\Sigma(m, A)$,
- **(4)** the extreme points of the closure $\overline{\Sigma(m, A)}$ belong to the m-dimensional spectrum $\sigma(m, A)$,
- **(5)** the closed convex hull of $\sigma(m, A)$ coincides with $\overline{\Sigma(m, A)}$

(see Section 4 for precise statements and definitions). The item **(3)** can be regarded as a geometric version of the variational principle. The set $\Sigma(m, A)$ may also have other extreme points (see Remark 4.12). Therefore one can obtain more information about the point spectrum by studying the extreme points of $\Sigma(m, A)$ than by applying the standard variational formulae.

The paper is organised as follows. For the sake of convenience, in Section 1 we give definitions and results on sequence spaces and locally convex topologies, which are used throughout other sections. Almost all these results are well known; most of them can be found in [K], Sections 20.9, 21.2, 30 and [Ru], Section 2.4.

Section 2 is devoted to Birkhoff's theorem. Many proofs of this theorem are known for finite matrices (see, for example, [MO] or [BP]). The problem of extending **(i)** and **(ii)** to infinite matrices is known as Birkhoff's problem 111. It has been studied in [Gr], [Is], [Ke], [Le], [Mu] and [RP]. However, their results are not sufficient for our purposes because

- (i) in order to deal with unbounded operators, we need **(i)** not only for the whole set of stochastic matrices but also for some its subsets which were not considered in these papers,
- (ii) we need **(ii)** with respect to a finer topology than the topology introduced in [Ke] or [RP], whereas [Is] deals with a too strong topology such that **(ii)** does not hold true.

Our proof of **(i)** and **(ii)** is based on the well known idea of shifting weights along edges of the underlying graph. It is almost purely combinatorial and works equally well for finite and infinite weighted graphs or matrices. Formally speaking, in Sections 3 and 4 we consider only infinite matrices. However, in the proof of Theorem 3.15 we apply results related to more general weighted graphs. For infinite graphs and matrices **(ii)** depends upon the choice of an appropriate topology. We give an explicit description of the strong and Mackey topologies on the set of (sub)stochastic weights (Corollaries 2.11 and 2.12), and show that **(ii)** holds true with respect to the Mackey topology (Theorem 2.15), but not necessarily with respect to the strong topology (Example 2.19).

In Section 3 we consider operators generated by stochastic matrices and derive a number of corollaries from Birkhoff's theorem. Many of these results seem almost obvious. However, our proofs of the key Theorems 3.10 and 3.15 are surprisingly long and complicated. It is not clear whether they can be essentially simplified.

Section 4 is about multidimensional spectra and numerical ranges. Here we give precise statements and proofs of **(1)**–**(5)** for a self-adjoint operator A (see Corollaries 4.7, 4.11 and Lemma 4.10). The corresponding results for finite matrices A are well known and rather elementary (see, for example, [AU] or [MO]). If A is compact, one can probably obtain **(1)**–**(5)** by considering its finite dimensional approximations (in [Ma1] and [Si] similar ideas have been used for studying s-numbers of compact operators). However, the general case is much more complex as the operator A may have continuous spectrum or (and) several accumulation points of its discrete spectrum, which makes it impossible to find an effective approximation procedure. In the end of Section 4 we prove two variational formulae (Corollaries 4.16 and 4.17) and show that $\sigma(m,A)$ is a subset of the closed convex hull of $\bigcup_\theta \sigma(m,A_\theta)$ whenever the self-adjoint operator A belongs to the closed convex hull of the family of self-adjoint operators A_θ (Corollary 4.21); all these results are simple consequences of **(1)**–**(5)**.

There are many other concepts of multidimensional numerical range [BD], [H], [LMMT]. We briefly discuss some of them in Subsection 4.1.

Acknowledgements. I would like to thank E.B. Davies, A. Markus, Yu. Netrusov and M. Solomyak for their valuable comments and constructive criticism. I am also grateful to L. Landau for his useful remarks on Birkhoff's theorem.

1. Sequence spaces

1.1. **Notation and definitions.** Let

$\hat{\mathbb{R}} := [-\infty, +\infty]$,

$\mathbb{R}^\infty$ be the linear spaces of all real sequences $\mathbf{x} = \{x_1, x_2, \dots\}$,

$\mathbb{R}_0^\infty$ be the subspace of sequences which converge to zero and

$\mathbb{R}_{00}^\infty$ be the subspace of sequences with finitely many nonzero entries.

We shall often consider the Euclidean space $\mathbb{R}^m$ as a finite dimensional subspace of $\mathbb{R}_{00}^\infty$, so that the m-dimensional real vector $(x_1, x_2, \dots, x_m)$ is identified with the sequence $(x_1, x_2, \dots, x_m, 0, 0, \dots)$. If $\mathbf{x} \in \mathbb{R}^\infty$, let

$$\begin{aligned} |\mathbf{x}| &:= \{|x_1|, |x_2|, \dots\}, \\ \mathbf{x}^{(m)} &:= \{x_1, x_2, \dots, x_m, 0, 0, \dots\}, \qquad m = 1, 2, \dots, \\ \mathbf{x}^{(\infty)} &:= \mathbf{x}. \end{aligned} \tag{1.1}$$

Throughout the paper X denotes a real linear subspace of $\mathbb{R}^\infty$ endowed with a locally convex topology $\mathfrak{T}$ and X^* is its dual space. We shall always be assuming that $\mathfrak{T}$ is finer (that is, not weaker) than the topology of element-wise convergence.

If Ω is a subset of X then $\operatorname{ex}\Omega$, $\operatorname{conv}\Omega$, $\overline{\operatorname{conv}}\,\Omega$ denote the set of extreme points of Ω, the convex hull of Ω and its $\mathfrak{T}$-closure respectively. Recall that $\mathbf{x}\in\Omega$ is called an *extreme point* of Ω if $\mathbf{x}$ cannot be represented as a convex linear combination of two other elements of Ω. If the set Ω is $\mathfrak{T}$-compact then, according to the Krein–Milman theorem, $\overline{\operatorname{conv}}\,\Omega=\overline{\operatorname{conv}}\,(\operatorname{ex}\Omega)$. An element $\mathbf{x}\in\Omega$ is said to be $\mathfrak{T}$-*exposed* if there exists a linear $\mathfrak{T}$-continuous functional $\mathbf{x}^*\in X^*$ such that $\langle\mathbf{x},\mathbf{x}^*\rangle>\langle\mathbf{y},\mathbf{x}^*\rangle$ for all $\mathbf{y}\in\Omega$. Every exposed point of Ω belongs to $\operatorname{ex}\Omega$ but an extreme point is not necessarily exposed.

Denote by X' the linear space of all real sequences $\mathbf{x}'=\{x'_1,x'_2,\dots\}\in\mathbb{R}^\infty$ such that $\sum_{i=1}^\infty|x_i\,x'_i|<\infty$ for all $\mathbf{x}\in X$. If $X''=X$ then the space X is said to be *perfect*. We have $X\subseteq X''$ and $\mathbb{R}^\infty_{00}\subseteq X'=X'''$; in particular, X' is perfect. The intersection of an arbitrary collection of perfect spaces is perfect. However, the linear span of a collection of perfect spaces may not be perfect. For example, if X is a one dimensional subspace of $\mathbb{R}^\infty_0$ then $X''\subset\mathbb{R}^\infty_0$ but $(\mathbb{R}^\infty_0)''=l^\infty$.

The set of sequences $\tilde{\mathbf{x}}=\{\tilde x_1,\tilde x_2,\dots\}$ such that $|\tilde x_j|\le|x_j|$ for some $\mathbf{x}\in\Omega$ is said to be the *normal cover* of the set Ω. A set (or subspace) of $\mathbb{R}^\infty$ is said to be *normal* if it coincides with its normal cover. We have $X'=(\tilde X)'$, where $\tilde X$ is a normal cover of X. Therefore a perfect space is normal.

1.2. **Topologies on sequence spaces.** Every sequence $\mathbf{x}'\in X'$ defines the linear functional $\langle\mathbf{x},\mathbf{x}'\rangle:=\sum_{j=1}^\infty x_j\,x'_j$ on the space X. Further on we shall always be assuming that $\mathbb{R}^\infty_{00}\subseteq X$. Then every nonzero element of X' defines a nonzero functional and therefore we can introduce the weak* topology $\mathfrak{T}_{\rm w}(X',X)$ on X'. If $\mathfrak{S}$ is an arbitrary family of weak* bounded sets $\Omega'\in X'$ then the family of seminorms

$$p_{\Omega'}(\mathbf{x})\;:=\;\sup_{\mathbf{x}'\in\Omega'}|\langle\mathbf{x},\mathbf{x}'\rangle|\,,\qquad\Omega'\in\mathfrak{S}\,,\tag{1.2}$$

defines a locally convex topology on the space X, which is usually called the $\mathfrak{S}$-topology. We shall deal with the following $\mathfrak{S}$-topologies on X:

(1) the topology of element-wise convergence $\mathfrak{T}_0$, generated by the family $\mathfrak{S}$ of all finite subsets of $\mathbb{R}^\infty_{00}$;
(2) the weak topology $\mathfrak{T}_{\rm w}(X,X')$, generated by the family $\mathfrak{S}$ of all finite subsets of X';
(3) the Mackey topology $\mathfrak{T}_{\rm m}(X,X')$, generated by the family $\mathfrak{S}$ of all absolutely convex $\mathfrak{T}_{\rm w}(X',X)$-compact subsets of X';
(4) the strong topology $\mathfrak{T}_{\rm b}(X,X')$, generated by the family $\mathfrak{S}$ of all $\mathfrak{T}_{\rm w}(X',X)$-bounded subsets of X'.

Every next topology in this list is finer than the previous one. Each of them is equivalent to the usual Euclidean topology whenever $\dim X<\infty$.

The strong topology $\mathfrak{T}_{\rm b}(X,X')$ is generated by all lower $\mathfrak{T}_{\rm w}(X,X')$-semicontinuous seminorms on X and the Mackey topology $\mathfrak{T}_{\rm m}(X,X')$ is defined by all lower $\mathfrak{T}_{\rm w}(X,X')$-semicontinuous seminorms p on X such that

$$p(\mathbf{x}-\mathbf{x}^{(m)}) \underset{m\to\infty}{\to} 0\,,\qquad \forall\mathbf{x}\in X\,. \tag{1.3}$$

The perfect space X'' is obtained from X by adding all $\mathfrak{T}_0$-limits of $\mathfrak{T}_{\rm w}(X,X')$-Cauchy sequences in X. A perfect space X is $\mathfrak{T}_{\rm b}(X,X')$-complete, $\mathfrak{T}_{\rm m}(X,X')$-complete and sequentially $\mathfrak{T}_{\rm w}(X,X')$-complete but is not necessarily $\mathfrak{T}_{\rm w}(X,X')$-complete. By the Mackey–Arens theorem, $\mathfrak{T}_{\rm m}(X,X')$ is the finest locally convex topology on the space X such that its topological dual X^* coincides with X'. If X' is $\mathfrak{T}_{\rm w}(X',X)$-complete then the $\mathfrak{T}_{\rm b}(X,X')$-dual of X also coincides with X'.

By Mackey's theorem, a subset of a locally convex space is weakly bounded if and only if it is bounded in any topology generating the same dual space. For a sequence space X, we have the following stronger result which implies that $\Omega\subset X$ is $\mathfrak{T}_{\rm w}(X,X')$-bounded if and only if it is $\mathfrak{T}_{\rm b}(X,X')$-bounded.

Theorem 1.1. *Assume that $\Omega\subset X$ is $\mathfrak{T}_{\rm w}(X,X')$-bounded and $\Omega'\subset X'$ is $\mathfrak{T}_{\rm w}(X',X)$-bounded. Then the set of sequences $\{x_1 x'_1, x_2 x'_2,\dots\}$, where $\mathbf{x}=\{x_1,x_2,\dots\}\in\Omega$ and $\mathbf{x}'=\{x'_1,x'_2,\dots\}\in\Omega'$, is bounded in l^1.*

Proof. See [Ru], Chapter 2, Proposition 1.4. □

The following theorem can be proved in the same way as Theorem 2.4 in [Ru], Chapter 2, where the author assumed that X is perfect.

Theorem 1.2. *If X is a normal space and $\Omega'\subset X'$ then the following two conditions are equivalent:*

(1) *Ω' is $\mathfrak{T}_{\rm w}(X',X)$-compact,*

(2) *Ω' is $\mathfrak{T}_0$-compact and $\lim\limits_{n\to\infty}\sup\limits_{\mathbf{x}'\in\Omega'}\sum\limits_{i=n}^{\infty}|x_i\,x'_i|=0$ for each $\mathbf{x}\in X$.*

Remark 1.3. If $\{\mathbf{x}_n\}\subset X$ is a $\mathfrak{T}_{\rm w}(X,X')$-Cauchy sequence which converges to $\mathbf{x}\in X''$ in the topology $\mathfrak{T}_0$, then by Fatou's lemma

$$\sup_{\mathbf{x}'\in\Omega'}|\langle\mathbf{x},\mathbf{x}'\rangle| \;\le\; \sup_{\mathbf{x}'\in\Omega'}\langle|\mathbf{x}|,|\mathbf{x}'|\rangle \;\le\; \sup_{\mathbf{x}'\in\Omega'}\sup_n\langle|\mathbf{x}_n|,|\mathbf{x}'|\rangle\,.$$

Since the Cauchy sequence $\{\mathbf{x}_n\}$ is $\mathfrak{T}_{\rm w}(X,X')$-bounded, Theorem 1.1 and the above inequality imply that the set $\Omega'\subset X'$ is $\mathfrak{T}_{\rm w}(X',X'')$-bounded if and only if it is $\mathfrak{T}_{\rm w}(X',X)$-bounded. Therefore the strong topology $\mathfrak{T}_{\rm b}(X,X')$ coincides with the restriction of $\mathfrak{T}_{\rm b}(X'',X')$ to X. However, this is not necessarily the case with the Mackey topologies.

Example 1.4. If $X=\mathbb{R}^\infty_0$ then $X'=l^1$, $X''=l^\infty$ and $\mathfrak{T}_{\rm b}(l^\infty,l^1)$ is the l^∞-topology. Theorem 1.2 implies that the closed unit ball in the space l^1 is $\mathfrak{T}_{\rm w}(l^1,\mathbb{R}^\infty_0)$-compact. Therefore $\mathfrak{T}_{\rm m}(\mathbb{R}^\infty_0,l^1)=\mathfrak{T}_{\rm b}(l^\infty,l^1)|_{\mathbb{R}^\infty_0}$. The Mackey topology $\mathfrak{T}_{\rm m}(l^\infty,l^1)$ on $\mathbb{R}^\infty_0$ is strictly coarser than $\mathfrak{T}_{\rm m}(\mathbb{R}^\infty_0,l^1)$. Indeed, if

$\mathbf{x} = \{1, 1, \ldots\}$ and $\tilde{\mathbf{x}}_m := \mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}$ then $\tilde{\mathbf{x}}_m \in \mathbb{R}^\infty_0$, $\|\tilde{\mathbf{x}}_m\|_{l^\infty} = 1$ but, by Theorem 1.2, $\tilde{\mathbf{x}}_m \to 0$ as $m \to \infty$ in the topology $\mathfrak{T}_{\mathrm{m}}(l^\infty, l^1)$.

Remark 1.5. Let $\tilde{\Omega}'$ be the normal cover of the set $\Omega' \subset X'$. Theorem 1.1 implies that $\tilde{\Omega}'$ is $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded whenever Ω' is $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded. If X is normal then, by Theorem 1.2, $\tilde{\Omega}'$ is $\mathfrak{T}_{\mathrm{w}}(X', X)$-compact whenever Ω' is $\mathfrak{T}_{\mathrm{w}}(X', X)$-compact. Obviously,

$$p_{\Omega'}(\mathbf{x}) = \sup_{\mathbf{x}' \in \Omega'} |\langle \mathbf{x}, \mathbf{x}' \rangle| \le \sup_{\mathbf{x}' \in \tilde{\Omega}'} |\langle \mathbf{x}, \mathbf{x}' \rangle| = \sup_{\mathbf{x}' \in \tilde{\Omega}'} \sum_{j=1}^{\infty} |x_j| \, |x'_j| = p_{\tilde{\Omega}'}(\mathbf{x})$$

and the seminorms $p_{\tilde{\Omega}'}$ are lower $\mathfrak{T}_0$-semicontinuous. Therefore the strong topology $\mathfrak{T}_{\mathrm{b}}(X, X')$ on an arbitrary space X is generated by all lower $\mathfrak{T}_0$-semicontinuous seminorms and the Mackey topology $\mathfrak{T}_{\mathrm{m}}(X, X')$ on a normal space X is generated by all lower $\mathfrak{T}_0$-semicontinuous seminorms satisfying (1.3).

1.3. **Symmetric sequence spaces.** Our choice of notation in the following definition will become clear in Section 3.

Definition 1.6. If $\mathbf{x} \in \mathbb{R}^\infty$, let

- $P_{\mathbf{x}}$ be the set of all sequences $\mathbf{y} \in \mathbb{R}^\infty$ obtained from the sequence $\mathbf{x}$ by permutations of its entries,
- $P^r_{\mathbf{x}}$ be the set of all sequences $\tilde{\mathbf{y}} \in \mathbb{R}^\infty$ whose entries form a subsequence of a sequence $\mathbf{y} \in P_{\mathbf{x}}$ and
- $P^\emptyset_{\mathbf{x}}$ be the set of all sequences obtained from sequences $\tilde{\mathbf{y}} \in P^r_{\mathbf{x}}$ by adding an arbitrary collection of zero entries.

We shall say that a sequence space X is *symmetric* if $P_{\mathbf{x}} \subset X$ for every $\mathbf{x} \in X$. A seminorm p on a symmetric space X is said to be symmetric if $p(\mathbf{y}) = p(\mathbf{x})$ whenever $\mathbf{y} \in P_{\mathbf{x}}$.

If X is symmetric then $P^\emptyset_{\mathbf{x}'} \subset X'$ for every $\mathbf{x}' \in X'$. The seminorm $p_{\Omega'}$ defined by (1.2) is symmetric if and only if $\Omega' = \bigcup_{\mathbf{x}' \in \Omega'} P_{\mathbf{x}'}$. The following result is a consequence of Theorems 1.1 and 1.2 (see Remark 3.2).

Corollary 1.7. *Let X be a symmetric space such that $X \not\subset \mathbb{R}^\infty_{00}$, Ω' be a subset of X' and $\Omega'_{\mathrm{sym}} := \bigcup_{\mathbf{x}' \in \Omega'} P^\emptyset_{\mathbf{x}'}$. If $X \subseteq l^\infty$ then Ω'_{sym} is $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded whenever Ω' is $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded. If $X \subseteq \mathbb{R}^\infty_0$ and X is normal then Ω'_{sym} is $\mathfrak{T}_{\mathrm{w}}(X', X)$-compact whenever Ω' is $\mathfrak{T}_{\mathrm{w}}(X', X)$-compact.*

By Corollary 1.7, if X is a symmetric subspace of l^∞ and $X \not\subset \mathbb{R}^\infty_{00}$ then the strong topology $\mathfrak{T}_{\mathrm{b}}(X, X')$ is generated by a family of symmetric $\mathfrak{T}_0$-semicontinuous seminorms p such that

$$p(\mathbf{y}) \le p(\mathbf{x}), \qquad \forall \mathbf{y} \in X \bigcap P^\emptyset_{\mathbf{x}}, \quad \forall \mathbf{x} \in X. \tag{1.4}$$

If X is a normal symmetric subspace of $\mathbb{R}^\infty_0$ and $X \not\subset \mathbb{R}^\infty_{00}$ then the Mackey topology $\mathfrak{T}_{\mathrm{m}}(X, X')$ is generated by a family of symmetric $\mathfrak{T}_0$-semicontinuous seminorms p satisfying (1.3) and (1.4).

Example 1.8. If $X = \mathbb{R}^\infty$ then $X' = \mathbb{R}^\infty_{00}$ and $\mathfrak{T}_0 = \mathfrak{T}_{\mathrm{m}}(\mathbb{R}^\infty, \mathbb{R}^\infty_{00}) = \mathfrak{T}_{\mathrm{b}}(\mathbb{R}^\infty, \mathbb{R}^\infty_{00})$. This topology cannot be defined with the use of symmetric seminorms. If $X = l^p$ with $1 \le p \le \infty$ then $X' = l^{p'}$ and $\mathfrak{T}_{\mathrm{b}}(l^p, l^{p'})$ is the usual l^p-topology. If $p < \infty$ then $\mathfrak{T}_{\mathrm{b}}(l^p, l^{p'}) = \mathfrak{T}_{\mathrm{m}}(l^p, l^{p'})$, but the Mackey topology $\mathfrak{T}_{\mathrm{m}}(l^\infty, l^1)$ is strictly coarser than the l^∞-topology and is not generated by a family of symmetric seminorms.

Example 1.9. Let Φ be a symmetric lower $\mathfrak{T}_0$-semicontinuous Schatten norm on $\mathbb{R}^\infty_0$ and $s^{(0)}_\Phi \subseteq s_\Phi \subseteq l^\infty$ be the corresponding linear subspaces of sequences (see, for example, [Si] or [Ma1]; in the latter paper Φ is called a symmetric gauge function and the corresponding subspaces are denoted by l_Φ and $l^{(0)}_\Phi$). Then the norm topology on a subspace $X \subset s_\Phi$ is always coarser than $\mathfrak{T}_{\mathrm{b}}(X, X')$ and is coarser than $\mathfrak{T}_{\mathrm{m}}(X, X')$ whenever $X \subset s^{(0)}_\Phi$.

Example 1.10. Let $\mathbf{x} \in \mathbb{R}^\infty_0$, $\mathbf{x} \notin l^1$ and X be the subspace spanned by the normal cover $\tilde{P}_{\mathbf{x}}$ of the set $P_{\mathbf{x}}$. Then X' consists of all sequences $\mathbf{x}' \in \mathbb{R}^\infty_0$ such that

$$\|\mathbf{x}'\|_L \;:=\; \sup_{\mathbf{y}\in\tilde{P}_{\mathbf{x}}} |\langle \mathbf{y}, \mathbf{x}'\rangle| \;<\; \infty\,. \tag{1.5}$$

The space X' provided with the norm (1.5) is called the *Lorentz* space associated with the weight sequence $\mathbf{x}$ (see, for example, [LT], Section 4.e). We have

$$\sum_{k=1}^{\infty} |y_j|\,|x'_j|^* \;=\; \sum_{m=1}^{\infty} (|x'_m|^* - |x'_{m+1}|^*) \sum_{j=1}^{m} |y_j|\,, \qquad \forall \mathbf{x}', \mathbf{y} \in \mathbb{R}^\infty_0\,, \tag{1.6}$$

where $\{|z_1|^*, |z_2|^* \ldots\}$ denotes either the non-increasing rearrangement of the sequence $|\mathbf{z}|$ or (if $|\mathbf{z}|$ contains infinitely many nonzero entries and at least one zero entry) the non-increasing rearrangement of its nonzero entries. Using this identity, one can easily show that $\mathbf{y} \in X''$ if and only if

$$\|\mathbf{y}\|_M \;:=\; \sup_{m\ge 1} R_m(|\mathbf{y}|)\,(R_m(|\mathbf{x}|))^{-1} \;=\; \sup_{\|\mathbf{x}'\|_L<1} |\langle \mathbf{y}, \mathbf{x}'\rangle| \;<\; \infty\,, \tag{1.7}$$

where $R_m(|\mathbf{z}|) := \sum_{j=1}^m |z_j|^*$. The space X'' provided with the norm (1.7) is called the *Marcinkiewicz* space associated with $\mathbf{x}$. Since the set $\tilde{P}_{\mathbf{x}}$ is $\mathfrak{T}_{\mathrm{w}}(X, X')$-bounded, Theorem 1.1 implies that the $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded set $\{\mathbf{x}' \in X' : \|\mathbf{x}'\|_L < 1\}$ absorbs any other $\mathfrak{T}_{\mathrm{w}}(X', X)$-bounded subset of X'. Therefore, in view of Remark 1.3, the strong topology $\mathfrak{T}_{\mathrm{b}}(X'', X')$ is generated by the norm $\|\cdot\|_M$. The Mackey topology $\mathfrak{T}_{\mathrm{m}}(X'', X')$ is strictly coarser than $\mathfrak{T}_{\mathrm{b}}(X'', X')$ as $\|\mathbf{y} - \mathbf{y}^{(m)}\|_M$ may be equal to $\|\mathbf{y}\|_M$ for all m.

Remark 1.11. Let $X_{P_{\mathbf{x}}}$ be the linear space spanned by $P_{\mathbf{x}}$. Then $X''_{P_{\mathbf{x}}}$ is the minimal symmetric perfect space which contains $\mathbf{x}$. Obviously,

(1) if $\mathbf{x}$ is unbounded then $X'_{P_{\mathbf{x}}} = \mathbb{R}^\infty_{00}$ and $X''_{P_{\mathbf{x}}} = \mathbb{R}^\infty$;
(2) if $\mathbf{x} \in l^\infty$ but $\mathbf{x} \notin \mathbb{R}^\infty_0$ then $X'_{P_{\mathbf{x}}} = l^1$ and $X''_{P_{\mathbf{x}}} = l^\infty$;

(3) if $\mathbf{x} \in \mathbb{R}_0^\infty$ but $\mathbf{x} \notin l^1$ then $X'_{P_\mathbf{x}}$ is the Lorentz space and $X''_{P_\mathbf{x}}$ is the Marcinkiewicz space associated with $\mathbf{x}$ (see Example 1.10);
(4) if $\mathbf{x} \in l^1$ but $\mathbf{x} \notin \mathbb{R}_{00}^\infty$ then $X'_{P_\mathbf{x}} = l^\infty$ and $X''_{P_\mathbf{x}} = l^1$;
(5) if $\mathbf{x} \in \mathbb{R}_{00}^\infty$ then $X'_{P_\mathbf{x}} = \mathbb{R}^\infty$ and $X''_{P_\mathbf{x}} = \mathbb{R}_{00}^\infty$.

Remark 1.12. If $\mathbf{x} \notin \mathbb{R}_{00}^\infty$ and $P_\mathbf{x} \subset X$ then $X' \subseteq l^\infty$ and $l^1 \subseteq X''$. Therefore for every $\mathfrak{T}_\mathrm{b}(X, X')$-continuous seminorm p on X there exists a constant C_p such that $p(\mathbf{x}) \le C_p \|\mathbf{x}\|_{l^1}$ for all $\mathbf{x} \in X \bigcap l^1$.

2. Birkhoff's theorem

2.1. **Notation and definitions.** Let $\mathbf{G} = \{G_1, G_2, \dots\}$ be a family of countable sets G_k which may have non-empty intersections. Define a simple graph G as follows: the set of vertices of G coincides with $\bigcup_k G_k$ and two vertices are joined by an edge in G if and only if they belong to the same set G_k. Then G_k become complete subgraphs of G. Throughout this section we denote by g (with or without indices) the vertices of G or, in other words, the elements of $\bigcup_k G_k$. Let

$\mathcal{W}$ be the linear space of real-valued functions $\mathbf{w}$ on G,
$\mathcal{W}_+$ be the cone of non-negative functions $\mathbf{w} \in \mathcal{W}$ and
$\mathcal{W}_0$ be the set of functions $\mathbf{w} \in \mathcal{W}$ which take only finitely many non-zero values.

We shall call $\mathbf{w} \in \mathcal{W}_+$ *weights* over G and denote by $\mathbf{w}(g)$ the weight assigned to $g \in G$ (that is, the value of $\mathbf{w}$ at g). If $\mathbf{w} \in \mathcal{W}$, let

$G_\mathbf{w}$ be the subgraph of G which includes all vertices $g \in G$ such that $\mathbf{w}(g) \neq 0$ and all edges joining these vertices.

Let $\mathbf{G}_1$ be an arbitrary subset of $\mathbf{G}$. We shall say that a weight $\mathbf{w} \in \mathcal{W}_+$ is $\mathbf{G}_1$-*stochastic* if $\sum_{g \in G_k} \mathbf{w}(g) \le 1$ for every $G_k \in \mathbf{G}$ and $\sum_{g \in G_k} \mathbf{w}(g) = 1$ for every $G_k \in \mathbf{G}_1$. Denote by $\mathcal{S}^{\mathbf{G}_1}$ the convex set of all $\mathbf{G}_1$-stochastic weights and let $\mathcal{P}^{\mathbf{G}_1}$ be the set of $\mathbf{G}_1$-stochastic weights taking only the values 0 and 1. Clearly, $\mathbf{w} \in \mathcal{P}^{\mathbf{G}_1}$ if and only if the restriction of $\mathbf{w}$ to every subset G_k takes at most one value 1, all other values being 0, and $\mathbf{w}$ does take the value 1 at some vertex $g \in G_k$ whenever $G_k \in \mathbf{G}_1$. If $\mathbf{G}_1 \subseteq \mathbf{G}'_1 \subseteq \mathbf{G}$ then $\mathcal{P}^{\mathbf{G}'_1} \subseteq \mathcal{P}^{\mathbf{G}_1} \subseteq \mathcal{S}^{\mathbf{G}_1}$ and $\mathcal{S}^{\mathbf{G}'_1} \subseteq \mathcal{S}^{\mathbf{G}_1}$.

Remark 2.1. The weights $\mathbf{w} \in \mathcal{S}^{\mathbf{G}}$ and $\mathbf{w} \in \mathcal{S}^{\emptyset}$ are said to be *stochastic* and, respectively, *sub-stochastic*. A stochastic weight $\mathbf{w}$ can be considered as a family of probability measures $\mathbf{w}_{(k)} := \mathbf{w}|_{G_k}$ on the sets G_k such that $\mathbf{w}_{(k)} = \mathbf{w}_{(j)}$ on $G_k \bigcap G_j$.

Since the set of vertices is countable, $\mathcal{W}$ can be identified with the sequence space $\mathbb{R}^\infty$ (or with its subspace if G is finite). Further on we use definitions and notation introduced in Section 1.

2.2. **Extreme points.** We shall say that a path $g_0 \to g_1 \to \cdots \to g_l$ in G is

admissible if no three adjacent vertices in this path belong to the same set $G_k \in \mathbf{G}$;

a *cycle* if $g_0 = g_l$ and the number of distinct vertices g_j is not smaller than 3 (that is, $g_0 \to g_1 \to g_2 = g_0$ is not a cycle).

Proposition 2.2. *Every two vertices lying in the same connected component of G can be joined by an admissible path. If there are no admissible cycles and, in addition,*

(c_1) *the intersection $G_k \bigcap G_l$ of two distinct sets $G_k, G_l \in \mathbf{G}$ contains at most one vertex of G*

then this admissible path is unique.

Proof. Let g_0 and g_m belong to the same connected component of G. Then a path $g_0 \to g_1 \to \cdots \to g_m$ with the minimal possible number of vertices is admissible (otherwise we could obtain a shorter path from g_0 to g_m replacing $g_j \to g_{j+1} \to \cdots \to g_{j+i}$ with $g_j \to g_{j+i}$). This proves the first statement.

Let $g_1 \to g_1 \to \cdots \to g_m$ and $g_1 \to g_{n+m} \to \cdots \to g_{m+1} \to g_m$ be two distinct admissible paths from g_1 to g_m. Without loss of generality we may assume that these paths have only two common vertices g_1 and g_m. Then the vertices $g_1, \ldots, g_{m+n}$ are distinct and do not belong to the same set G_k. Consider the graph $\mathcal{G}$ formed by all these vertices and all joining them edges. Let $\tilde{g}_1 \to \tilde{g}_2 \to \cdots \to \tilde{g}_{l+1} = \tilde{g}_1$ be a cycle in $\mathcal{G}$ with the minimal possible number of vertices which do not belong to the same set G_k (since $\mathcal{G}$ contains at least one cycle $g_1 \to g_2 \to \cdots \to g_{m+n}$ with this property, such a 'minimal' cycle exists). The condition (c_1) implies that this cycle is admissible. Indeed, if two non-adjacent vertices $\tilde{g}_i$ and $\tilde{g}_{i+j}$ in this path are joined by an edge then all vertices of the cycle $\tilde{g}_i \to \tilde{g}_{i+1} \to \cdots \to \tilde{g}_{i+j} \to \tilde{g}_i$ belong to some set $G_k \in \mathbf{G}$ and all vertices of the cycle $\tilde{g}_{i+j} \to \tilde{g}_{i+j+1} \to \cdots \to \tilde{g}_l \to \tilde{g}_1 \to \cdots \to \tilde{g}_i \to \tilde{g}_{i+j}$ belong to a distinct set G_l, in which case the intersection $G_k \bigcap G_l$ contains at least two elements $\tilde{g}_i$ and $\tilde{g}_{i+j}$. This proves the second statement. □

Further on we shall be assuming that

($\mathbf{g}_1$) every vertex of G belongs to at most two sets G_k,

($\mathbf{g}_2$) every admissible cycle in G has an even number of vertices.

If the conditions ($\mathbf{g}_1$) and ($\mathbf{g}_2$) are fulfilled then $\mathbf{G}$ can be split into two groups $\mathbf{G}^+ = \{G_1^+, G_2^+, \ldots\}$ and $\mathbf{G}^- = \{G_1^-, G_2^-, \ldots\}$ in such a way that any two sets from the same group do not have common elements (two sets G_k and G_j belong to the same group if every admissible path $G_k \ni g_0 \to g_1 \to \cdots \to g_{l-1} \to g_l \in G_j$ in G with $g_1 \notin G_k$ and $g_{l-1} \notin G_j$ has an even number of vertices). The intersection $G_k^+ \bigcap G_j^-$ may consist of several elements or be empty, and every set $G_k^\pm$ may contain a 'tail' subset $\tilde{G}_k^\pm$ which does not have common elements with any other set G_j.

In view of the following example, all results of this section are valid for finite and infinite matrices which we shall discuss in more detail in Section 3.

Example 2.3. Let $\mathbf{G}$ satisfy $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$ and $\mathbf{G}^\pm$ be defined as above. Denote by $m_\pm$ the number of sets $G_k^\pm$ lying in $\mathbf{G}^\pm$; we allow $m_+ = \infty$ and (or) $m_- = \infty$. If every intersection $G_k^+ \bigcap G_j^-$ consists of one element and all the tail subsets $\tilde{G}_k^\pm$ are empty then $\mathcal{W}$ is isomorphic to the linear space of $m_+ \times m_-$-matrices. Indeed, the value of $\mathbf{w} \in \mathcal{W}$ at the vertex $g \in G_k^+ \bigcap G_j^-$ can be considered as the entry of an $m_+ \times m_-$-matrix at the intersection of its jth row and kth column. In this case $\mathcal{S}^{\mathbf{G}}$, $\mathcal{S}^\emptyset$ and $\mathcal{P}^{\mathbf{G}}$ are the sets of doubly stochastic, sub-stochastic and permutation matrices respectively.

If G is a general family of sets satisfying $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$ then one can think of $\mathcal{W}$ as a space of matrices which may have 'multiple' or 'forbidden' entries and 'tails' $\tilde{G}_k^\pm$ attached to their rows and columns.

Theorem 2.4. *Let the conditions* $(\mathbf{g}_1)$ *and* $(\mathbf{g}_2)$ *be fulfilled and let* $\mathcal{V}$ *be a normal conic subset of* $\mathcal{W}$. *Then* $\operatorname{ex}(\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}) = \mathcal{P}^{\mathbf{G}_1} \bigcap \mathcal{V}$.

Proof. Obviously, $\mathcal{P}^{\mathbf{G}_1} \bigcap \mathcal{V} \subset \operatorname{ex}(\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V})$. In order to prove the converse, let us consider a weight $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$ such that $\mathbf{w}(g') \in (0,1)$ for some $g' \in G$ and show that $\mathbf{w} \notin \operatorname{ex}(\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V})$. Let G' be the connected component of $G_{\mathbf{w}}$ containing the vertex g'. Then $\mathbf{w}(g) \in (0,1)$ at every vertex $g \in G'$.

(**1**) Assume that, for some $k \neq l$, the intersection $G' \bigcap G_k \bigcap G_l$ contains two distinct vertices g_1 and g_2. Let $\mathbf{w}_\varepsilon^\pm(g_j) = \mathbf{w}(g_j) \pm (-1)^j \varepsilon$ and $\mathbf{w}_\varepsilon^\pm(g) = \mathbf{w}(g)$ whenever $g \neq g_j$, $j = 1, 2$. Then $\mathbf{w} = \frac{1}{2}(\mathbf{w}_\varepsilon^+ + \mathbf{w}_\varepsilon^-)$ and, in view of $(\mathbf{g}_1)$, $\mathbf{w}_\varepsilon^\pm \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$ provided that $\varepsilon > 0$ is sufficiently small. Therefore without loss of generality we can assume that G' satisfies (c_1).

(**2**) Similarly, if G' contains an admissible cycle $\mathcal{G} = g_0 \to g_2 \to \cdots \to g_n = g_0$, let $\mathbf{w}_\varepsilon^\pm(g_j) = \mathbf{w}(g_j) \pm (-1)^j \varepsilon$ and $\mathbf{w}_\varepsilon^\pm(g) = \mathbf{w}(g)$ whenever $g \notin \mathcal{G}$. The condition $(\mathbf{g}_2)$ implies that $\mathbf{w}_\varepsilon^+$ and $\mathbf{w}_\varepsilon^-$ are correctly defined weights over G. We have $\mathbf{w} = \frac{1}{2}(\mathbf{w}_\varepsilon^+ + \mathbf{w}_\varepsilon^-)$ and $\mathbf{w}_\varepsilon^\pm \in \mathcal{W}_+ \bigcap \mathcal{V}$ provided that ε is sufficiently small. In view of $(\mathbf{g}_1)$, if $g_j \in G_k$ then one of the adjacent vertices g_{j-1}, g_{j+1} belongs to G_k and the other does not. This implies that $\sum_{g \in G_k} \mathbf{w}_\varepsilon^\pm(g) = \sum_{g \in G_k} \mathbf{w}(g)$ for every k. Therefore $\mathbf{w}_\varepsilon^\pm \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$.

(**3**) Finally, let us assume that G' does not contain admissible cycles and satisfies (c_1). Then, by Proposition 2.2, every two vertices $g_0, g_l \in G'$ are joined by a unique admissible path. Let us fix $g_0 \in G'$ and denote by $\mathcal{G}_n$ the set of vertices in G' obtained from g_0 by moving along all admissible paths with n edges. Then for each $k = 1, 2, \ldots$ there exists $n \geq 0$ such that $G_k \subseteq \mathcal{G}_n \bigcup \mathcal{G}_{n+1}$. Moreover, if $G_k \subseteq \mathcal{G}_n \bigcup \mathcal{G}_{n+1}$ then the intersection $G_k \bigcap \mathcal{G}_n$ consists of one element $g_{k,n}$. Indeed, if there are two distinct admissible paths $g_0 \to g_1 \to \cdots \to g_{k,n}$ and $g_0 \to g_1' \to \cdots \to g_{k,n}' \in G_k \bigcap \mathcal{G}_n$ then g_0 and $g_{k,n}$ can be joined by the two distinct admissible paths $g_0 \to g_1 \to \cdots \to g_{k,n}$ and $g_0 \to g_1' \to \cdots \to g_{k,n}' \to g_{k,n}$.

If $g_0 \in G_k$ then $G_k \subseteq \mathcal{G}_0 \bigcup \mathcal{G}_1$ and $g_{k,0} = g_0$. Let us denote

$$\varepsilon_{k,0} := \min\left\{\frac{1}{2}\,,\ \frac{1-\mathbf{w}(g_0)}{2\,\mathbf{w}(g_0)}\right\}\,,\qquad \varepsilon_{k,n+1} := \frac{\varepsilon_{k,n}\,\mathbf{w}(g_{k,n})}{1-\mathbf{w}(g_{k,n})}\,, \tag{2.1}$$

where $n = 0,1,2,\dots$ and k is such that $G_k \subseteq \mathcal{G}_n \bigcup \mathcal{G}_{n+1}$. Since $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1}$, we have $\mathbf{w}(g_{k,n}) + \mathbf{w}(g_{k,n+1}) \le 1$ and, consequently,

$$\frac{\mathbf{w}(g_{k,n})}{1-\mathbf{w}(g_{k,n})} \ \le\ \frac{1-\mathbf{w}(g_{k,n+1})}{\mathbf{w}(g_{k,n+1})}\,.$$

Using these inequalities, one can easily prove by induction in n that

$$\varepsilon_{k,n} \ \le\ \min\left\{\frac{1}{2}\,,\ \frac{1-\mathbf{w}(g_{k,n})}{2\,\mathbf{w}(g_{k,n})}\right\}\,. \tag{2.2}$$

Consider two sequences of weights $\mathbf{w}^+_{\varepsilon,n}$ and $\mathbf{w}^-_{\varepsilon,n}$ such that

$\mathbf{w}^\pm_{\varepsilon,0}(g_0) := (1 \pm \varepsilon_{k,0})\,\mathbf{w}(g_0)$ and $\mathbf{w}^\pm_{\varepsilon,0}(g) := \mathbf{w}(g)$ for all $g \ne g_0$,
$\mathbf{w}^\pm_{\varepsilon,n+1}(g) := \mathbf{w}^\pm_{\varepsilon,n}(g)$ for all $g \in \bigcup_{j\le n} \mathcal{G}_j$,
$\mathbf{w}^\pm_{\varepsilon,n+1}(g) := \mathbf{w}(g)$ whenever $g \notin \bigcup_{j\le n+1} \mathcal{G}_j$,
if $G_k \subseteq \mathcal{G}_n \bigcup \mathcal{G}_{n+1}$ then $\mathbf{w}^\pm_{\varepsilon,n}(g_{k,n}) := (1 \pm \varepsilon_{k,n})\,\mathbf{w}(g_{k,n})$ and
$\mathbf{w}^\pm_{\varepsilon,n+1}(g) := (1 \mp \varepsilon_{k,n+1})\mathbf{w}(g)$ whenever $g \in G_k \bigcap \mathcal{G}_{n+1}$ and $g \ne g_{k,n}$.

Obviously, $\mathbf{w}(g) = \frac{1}{2}(\mathbf{w}^+_{\varepsilon,n}(g) + \mathbf{w}^-_{\varepsilon,n}(g))$. The estimates (2.2) imply that $\mathbf{w}^\pm_{\varepsilon,n} \in \mathcal{W}_+ \bigcap \mathcal{V}$. Finally, if $G_k \subseteq \mathcal{G}_n \bigcup \mathcal{G}_{n+1}$ and $\sum_{g\in G_k} w(g) = t$ then

$$\sum_{g\in G_k} \mathbf{w}^\pm_{\varepsilon,n}(g) = (1\pm\varepsilon_{k,n})\,\mathbf{w}(g_{k,n}) + (1\mp\varepsilon_{k,n+1})\,(t-\mathbf{w}(g_{k,n})) \ =\ t\mp\varepsilon_{k,n+1}(t-1)\,.$$

This identity and the estimates $\varepsilon_{k,n}, \varepsilon_{k,n+1} \le 1/2$ imply that

$$\frac{t+1}{2} \ -\ \frac{1-t}{2\,(1-w(g_{k,n}))} \ \le\ \sum_{g\in G_k} \mathbf{w}^\pm_{\varepsilon,n}(g) \ \le\ \frac{t+1}{2}\,. \tag{2.3}$$

Let $\mathbf{w}^\pm_\varepsilon(g) := \lim_{n\to\infty} \mathbf{w}^\pm_{\varepsilon,n}(g)$. Then $\mathbf{w}_0 = \frac{1}{2}(\mathbf{w}^+_\varepsilon + \mathbf{w}^-_\varepsilon)$ and, in view of (2.2) and (2.3), $\mathbf{w}^\pm_\varepsilon(g) \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$.

Thus, under conditions of the theorem, a weight $\mathbf{w} \in (\mathcal{S}^{\mathbf{G}_1} \setminus \mathcal{P}^{\mathbf{G}_1}) \bigcap \mathcal{V}$ can always be represented as a convex combination of two other weights from $\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$ and therefore is not an extreme point. □

Remark 2.5. If the condition $(\mathbf{g}_2)$ is not fulfilled then an extreme point of $\mathcal{S}^{\mathbf{G}_1}$ does not necessarily belong to $\mathcal{P}^{\mathbf{G}_1}$. The simplest example is $G_1 = \{g_1, g_2\}$, $G_2 = \{g_2, g_3\}$, $G_3 = \{g_3, g_1\}$ and $\mathbf{G} = \{G_1, G_2, G_3\}$. In this case $\mathcal{S}^{\mathbf{G}}$ consists of one weight which takes the value $\frac{1}{2}$ at each vertex.

Remark 2.6. The sets $\mathcal{S}^{\mathbf{G}_1}$ and $\mathcal{P}^{\mathbf{G}_1}$ may well be very poor or even empty. However, even in this situation Theorem 2.4 may be useful. In particular, by the Krein–Milman theorem, under conditions of Theorem 2.4 we have

$$\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{W}_0 \ =\ \operatorname{conv} \mathcal{P}^{\mathbf{G}_1} \bigcap \mathcal{W}_0\,.$$

Therefore $\mathcal{S}^{\mathbf{G}_1}\bigcap\mathcal{W}_0=\emptyset$ whenever $\mathcal{P}^{\mathbf{G}_1}\bigcap\mathcal{W}_0=\emptyset$.

Remark 2.7. If the conditions $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$ are fulfilled and $\mathcal{V}$ is a normal linear subspace of $\mathcal{W}$ then every extreme point $\mathbf{w}\in\operatorname{ex}(\mathcal{S}^{\mathbf{G}_1}\bigcap\mathcal{V})=\mathcal{P}^{\mathbf{G}_1}\bigcap\mathcal{V}$ is $\mathfrak{T}_{\mathrm{m}}(\mathcal{V},\mathcal{V}')$-exposed. Indeed, if $\mathbf{w}'(g)>0$ whenever $\mathbf{w}(g)=1$, $\mathbf{w}'(g)<0$ whenever $\mathbf{w}(g)=0$ and $\mathbf{w}'\in\mathcal{V}'$ then we have $\langle\mathbf{w},\mathbf{w}'\rangle>\langle\tilde{\mathbf{w}},\mathbf{w}'\rangle$ for all $\tilde{\mathbf{w}}\in\mathcal{S}^{\mathbf{G}_1}\bigcap\mathcal{V}$.

2.3. **Topologies on the space of stochastic weights.** The aim of this subsection is to describe locally convex topologies $\mathfrak{T}$ on a linear subspace $\mathcal{V}\supset\mathcal{P}^{\mathbf{G}_1}$ such that the $\mathfrak{T}$-closure of $\operatorname{conv}\mathcal{P}^{\mathbf{G}_1}$ coincides with $\mathcal{S}^{\mathbf{G}_1}\bigcap\mathcal{V}$. By Fatou's lemma we always have $\overline{\operatorname{conv}}\,\mathcal{P}^{\emptyset}\subset\mathcal{S}^{\emptyset}$ (as $\mathfrak{T}$ is finer than $\mathfrak{T}_0$). Tychonoff's theorem and Fatou's lemma also imply that the set $\mathcal{S}^{\emptyset}$ is $\mathfrak{T}_0$-compact. Therefore, in view of Theorem 2.4 and the Krein–Milman theorem, under the conditions $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$ we have $\mathcal{S}^{\emptyset}=\overline{\operatorname{conv}}\,\mathcal{P}^{\emptyset}$, where the closure is taken in the topology of element-wise convergence $\mathfrak{T}_0$. However, if $\mathbf{G}_1$ contains an infinite set G_k then the set $\mathcal{S}^{\mathbf{G}_1}$ is not $\mathfrak{T}_0$-closed and, by Theorem 1.2, is not $\mathfrak{T}$-compact whenever the functional $\mathbf{w}\to\sum_{g\in G_k}\mathbf{w}(g)$ is $\mathfrak{T}$-continuous. In this case **(ii)** does not directly follow from **(i)** and the Krein–Milman theorem.

Definition 2.8. Denote by $\mathcal{V}_{\mathcal{P}}$ and $\mathcal{V}_{\mathcal{S}}$ the normal covers of the subspaces spanned by $\mathcal{P}^{\emptyset}$ and $\mathcal{S}^{\emptyset}$ respectively. If $\mathbf{w}\in\mathcal{V}_{\mathcal{S}}$, let $\mathbf{w}_{(k)}$ be the restriction of $\mathbf{w}$ to G_k and $p_k(\mathbf{w}):=\|\mathbf{w}_{(k)}\|_{l^1}$.

Lemma 2.9. *Let us enumerate the sets G_k in an arbitrary way and define $F_n:=\bigcup_{k=1}^n G_k$. If D' is a $\mathfrak{T}_{\mathrm{w}}(\mathcal{V}_{\mathcal{P}}',\mathcal{V}_{\mathcal{P}})$-compact subset of $\mathcal{V}_{\mathcal{P}}'$ then*

$$\sup_{\mathbf{w}\in\mathcal{P}^{\emptyset},\,\mathbf{w}'\in D'}\ \sum_{g\in G\setminus F_n}|\mathbf{w}(g)\,\mathbf{w}'(g)|\ \to\ 0\,,\qquad n\to\infty\,, \tag{2.4}$$

whenever $\mathbf{G}$ satisfies $(\mathbf{g}_1)$ and

$$\sup_{\mathbf{w}\in\mathcal{S}^{\emptyset},\,\mathbf{w}'\in D'}\ \sum_{g\in G\setminus F_n}|\mathbf{w}(g)\,\mathbf{w}'(g)|\ \to\ 0\,,\qquad n\to\infty\,, \tag{2.5}$$

whenever $\mathbf{G}$ satisfies $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$.

Proof. If the conditions $(\mathbf{g}_1)$ and $(\mathbf{g}_2)$ are fulfilled then $\mathcal{S}^{\emptyset}$ coincides with the $\mathfrak{T}_0$-closure of $\operatorname{conv}\mathcal{P}^{\emptyset}$. Therefore, in view of Fatou's lemma, it is sufficient to prove only the first statement.

If (2.4) is not true then there exists $\delta>0$ and two sequences of weights $\{\mathbf{w}_n\}\subset\mathcal{P}^{\emptyset}$ and $\{\mathbf{w}_n'\}\subset D'$ such that $\sum_{g\in G\setminus F_n}|\mathbf{w}_n(g)\,\mathbf{w}_n'(g)|\geq\delta>0$ for all $n=1,2,\dots$ Let $n_\star$ be the minimal positive integer satisfying the estimate $\sum_{g\in F_{n_\star}\setminus F_n}|\mathbf{w}_n(g)\,\mathbf{w}_n'(g)|\geq\delta/2$ and $\mathbf{w}_n^\star\in\mathcal{W}_0\bigcap\mathcal{P}^{\emptyset}$ be the weight which takes the same values as $\mathbf{w}_n$ on $F_{n_\star}\setminus F_n$ and vanishes outside $F_{n_\star}\setminus F_n$. In view of $(\mathbf{g}_1)$, there exists a positive integer $n^\star>n_\star$ such that $\mathbf{w}_n^\star|_{G_k}\equiv0$ for all $k\geq n^\star$. Let us take an arbitrary n_1 and define $n_{j+1}:=n_j^\star$, where $j=1,2,\dots$

Then for each $g \in G$ the sum $\mathbf{w}^\star(g) := \sum_j \mathbf{w}^\star_{n_j}(g)$ is equal either to 0 or to 1 and $\sum_{g\in G_k} \mathbf{w}^\star(g) \le 1$, $\forall k = 1,2,\ldots$ Therefore the corresponding weight $\mathbf{w}^\star$ belongs to $\mathcal{P}^\emptyset$. On the other hand, $n_j \to \infty$ and

$$\sum_{g\in F_{n_j^\star}\setminus F_{n_j}} |\mathbf{w}^\star(g)\,\mathbf{w}'_{n_j}(g)| \;=\; \sum_{g\in F_{n_j^\star}\setminus F_{n_j}} |\mathbf{w}_{n_j}(g)\,\mathbf{w}'_{n_j}(g)| \;\ge\; \delta/2\,,$$

which contradicts to Theorem 1.2. □

We do not assume in Lemma 2.9 that $D' \subset \mathcal{V}_\mathcal{S}$. Therefore, for each fixed n, the supremum in (2.5) may well be $+\infty$. However, under conditions ($\mathbf{g}_1$) and ($\mathbf{g}_2$), it eventually becomes finite and converges to zero as $n\to\infty$.

Lemma 2.10. *If the condition* ($\mathbf{g}_1$) *is fulfilled and* D' *is a* $\mathfrak{T}_\mathrm{w}(\mathcal{V}'_\mathcal{S},\mathcal{V}_\mathcal{S})$-*bounded* $\mathfrak{T}_\mathrm{w}(\mathcal{V}'_\mathcal{P},\mathcal{V}_\mathcal{P})$-*compact subset of* $\mathcal{V}'_\mathcal{S}$ *then the weights* $\mathbf{w}'\in D'$ *are uniformly bounded.*

Proof. Let F_n be defined as in Lemma 2.9. If the restrictions of weights $\mathbf{w}'\in D'$ to F_n are not uniformly bounded then, for some $k\le n$, their restrictions to G_k form an unbounded subset of l^∞. This implies that the set D' is not $\mathfrak{T}_\mathrm{w}(\mathcal{V}'_\mathcal{S},\mathcal{V}_\mathcal{S})$-bounded.

Assume that there exist sequences $\{g_j\}_{j=1,2,\ldots}\in G$ and $\{\mathbf{w}'_j\}_{j=1,2,\ldots}\in D'$ such that $\mathbf{w}'_j(g_j)\to\infty$ as $j\to\infty$ and $\{g_j\}\not\subset F_n$ for any finite n. Since ($\mathbf{g}_1$) holds true, every vertex g belongs only to finitely many sets G_k and we can find a subsequence $\{g_{j_i}\}_{i=1,2,\ldots}$ with at most one entry at each set G_k. If $\mathbf{w}(g_{j_i})=1$ and $\mathbf{w}(g)=0$ whenever $g\notin\{g_{j_i}\}$ then $\mathbf{w}\in\mathcal{P}^\emptyset$ and $\sum_i|\mathbf{w}(g_{j_i})\,\mathbf{w}'_{j_i}(g_{j_i})| = \infty$. Therefore, by Theorem 1.2, the set D' is not $\mathfrak{T}_\mathrm{w}(\mathcal{V}'_\mathcal{P},\mathcal{V}_\mathcal{P})$-compact. □

Corollary 2.11. *If the conditions* ($\mathbf{g}_1$) *and* ($\mathbf{g}_2$) *are fulfilled then the strong topology* $\mathfrak{T}_\mathrm{b}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$ *is generated by the norm*

$$\|\mathbf{w}\|_\mathcal{S} \;:=\; \sup_k p_k(\mathbf{w}) \tag{2.6}$$

Proof. Since the norm (2.6) is lower $\mathfrak{T}_0$-semicontinuous, it is $\mathfrak{T}_\mathrm{b}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-continuous (see Remark 1.5). The set

$$\mathcal{S}^\emptyset \;=\; \{\mathbf{w}\in\mathcal{V}_\mathcal{S} : \|\mathbf{w}\|_\mathcal{S} < 1\}$$

is absorbing and, in view of (2.5) and Lemma 2.10, is $\mathfrak{T}_\mathrm{m}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-bounded. By Theorem 1.1, this set is $\mathfrak{T}_\mathrm{b}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-bounded, which implies that every $\mathfrak{T}_\mathrm{b}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-continuous seminorm is continuous in the norm topology. □

Corollary 2.12. *Let* $\mathfrak{T}$ *be the locally convex topology on* $\mathcal{V}_\mathcal{S}$ *generated by the seminorms* p_k, $k=1,2,\ldots$ *If the conditions* ($\mathbf{g}_1$) *and* ($\mathbf{g}_2$) *are fulfilled then the Mackey topology* $\mathfrak{T}_\mathrm{m}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$ *is finer than* $\mathfrak{T}$ *and coincides with* $\mathfrak{T}$ *on every* $\mathfrak{T}_\mathrm{b}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-*bounded subset of* $\mathcal{V}_\mathcal{S}$.

Proof. The seminorms p_k are lower $\mathfrak{T}_0$-semicontinuous and satisfy (1.3). Therefore, by Remark 1.5, $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$ is finer than $\mathfrak{T}$. On the other hand, if Ω is a bounded subset of $\mathcal{V}_\mathcal{S}$ then, in view of (2.5) and Lemma 2.10, for every Mackey seminorm p on $\mathcal{V}_\mathcal{S}$, every $\mathbf{x}\in\Omega$ and every $\varepsilon>0$ there exist a positive integer m and $\delta>0$ such that

$$\{\mathbf{y}\in\Omega : p_k(\mathbf{x}-\mathbf{y})<\delta\,,\forall k=1,2,\dots,m\}\ \subseteq\ \{\mathbf{y}\in\Omega : p(\mathbf{x}-\mathbf{y})<\varepsilon\}\,.$$

This implies that every $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-neighbourhood of $\mathbf{x}$ in Ω contains a $\mathfrak{T}$-neighbourhood. □

Remark 2.13. If the conditions $(\mathbf{g}_1)$, $(\mathbf{g}_2)$ are fulfilled and G does not coincide with the union of a finite collection of the sets G_k then the topology $\mathfrak{T}$ generated by the seminorms p_k is strictly coarser than $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$. Indeed, in this case there exists a sequence of weights $\mathbf{w}_n\in\mathcal{V}_\mathcal{P}$ such that $p_k(\mathbf{w}_n)=0$ for all $k<n$ and $p_n(\mathbf{w}_n)\to\infty$ as $n\to\infty$. This sequences converges to the zero weight in the topology $\mathfrak{T}$ but is not $\mathfrak{T}_{\mathrm{b}}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-bounded and, consequently, is not $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}_\mathcal{S},\mathcal{V}'_\mathcal{S})$-convergent.

In the rest of this section we shall be assuming that

$(\mathbf{g}_3)$ one can enumerate the sets G_j in such a way that either $G=F_n$ or $G_{n+1}\not\subset F_n$ for all sufficiently large n, where $F_n:=\bigcup_{k\le n}G_k$.

Every finite collection $\mathbf{G}=\{G_1,G_2,\dots,G_n\}$ satisfies $(\mathbf{g}_3)$. More generally, the condition $(\mathbf{g}_3)$ is fulfilled whenever the number of finite sets G_k is finite and the intersections of every two sets $G_j,G_k\in\mathbf{G}$ is finite. In particular, $(\mathbf{g}_3)$ is fulfilled for finite and infinite matrices (see Example 2.3).

Lemma 2.14. *Let the conditions* $(\mathbf{g}_1)$ *and* $(\mathbf{g}_3)$ *be fulfilled,* G_k *be enumerated as in* $(\mathbf{g}_3)$*,* $F_n:=\bigcup_{k\le n}G_k$ *and* $\mathbf{G}_{1,n}$ *be the collection of all sets* $G_k\in\mathbf{G}_1$ *with* $k\le n$*. Then there exists a positive integer* n_0 *such that for every* $n\ge n_0$ *and every weight* $\mathbf{w}\in\mathcal{P}^\emptyset$ *satisfying*

$$\sum_{g\in G_k}\mathbf{w}(g)=1\,,\qquad\forall G_k\in\mathbf{G}_{1,n}\,,\tag{2.7}$$

one can find a weight $\tilde{\mathbf{w}}\in\mathcal{P}^{\mathbf{G}_1}$ *whose restriction to* F_n *coincides with* $\mathbf{w}|_{F_n}$.

Proof. If for some positive integer n_1 there are no weights $\mathbf{w}\in\mathcal{P}^\emptyset$ satisfying (2.7) with $n=n_1$ then the lemma automatically holds true for $n_0=n_1$. Therefore we can assume without loss of generality that for each $n=1,2,\dots$ there exists a weight $\mathbf{w}_n\in\mathcal{P}^\emptyset$ satisfying (2.7).

If $G=F_n$ for all $n\ge n_1$ then, in view of $(\mathbf{g}_1)$, $\sum_k\sum_{g\in G_k}\mathbf{w}_n(g)\le 2n_1$. This estimate and (2.7) imply that the set $\mathbf{G}_{1,n}$ contains at most $2n_1$ elements for each $n=1,2,\dots$. Therefore there exists a positive integer n_0 such that $\mathbf{G}_{1,n}=\mathbf{G}_{1,n_0}$ for all $n\ge n_0$. In this case the inclusion $\mathbf{w}\in\mathcal{P}^\emptyset$ and (2.7) with $n\ge n_0$ imply that $\mathbf{w}\in\mathcal{P}^{\mathbf{G}_1}$.

If $G_{n+1}\not\subset F_n$ for all $n\ge n_1$ then we take $n_0=n_1$. Given $n\ge n_0$ and a weight $\mathbf{w}\in\mathcal{P}^\emptyset$ satisfying (2.7), we choose arbitrary vertices $g_{n+j}\in$

$G_{n+j} \setminus F_{n+j-1}$ and define $\tilde{\mathbf{w}}$ as follows: $\tilde{\mathbf{w}}(g) := \mathbf{w}(g)$ whenever $g \in F_n$, $\tilde{\mathbf{w}}(g_{n+j}) := 1$ for all $j = 1, 2, \ldots$ and $\tilde{\mathbf{w}}(g) := 0$ otherwise. Then $\mathbf{w} = \tilde{\mathbf{w}}$ on F_n and $\tilde{\mathbf{w}} \in \mathcal{P}^{\mathbf{G}_1}$ because $\sum_{g \in G_k} \tilde{\mathbf{w}}(g) = 1$ for all $k > n$. □

Theorem 2.15. *Let the conditions* $(\mathbf{g}_1)$–$(\mathbf{g}_3)$ *be fulfilled and* $\mathcal{V}$ *be a normal subspace of* $\mathcal{W}$ *such that* $\mathcal{V}_{\mathcal{P}} \subseteq \mathcal{V} \subseteq \mathcal{V}_{\mathcal{S}}$. *Then*

$$\mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V} \ = \ \overline{\operatorname{conv}}\, \mathcal{P}^{\mathbf{G}_1}, \qquad \forall \mathbf{G}_1 \subseteq \mathbf{G}, \tag{2.8}$$

where the closure is taken in the Mackey topology $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}, \mathcal{V}')$.

Proof. Since $\mathcal{V} \subseteq \mathcal{V}_{\mathcal{S}}$, the functionals $\mathbf{w} \to \sum_{g \in G_k} \mathbf{w}(g)$ are $\mathfrak{T}_{\mathrm{w}}(\mathcal{V}, \mathcal{V}')$-continuous and, consequently, $\overline{\operatorname{conv}}\, \mathcal{P}^{\mathbf{G}_1} \subseteq \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$. If $\mathbf{w} \notin \overline{\operatorname{conv}}\, \mathcal{P}^{\mathbf{G}_1}$ then, by the separation theorem (see, for example, [K], Section 20.7), there exist $\mathbf{w}' \in \mathcal{V}'$ and $\varepsilon > 0$ such that $\langle \mathbf{w}, \mathbf{w}' \rangle - \langle \tilde{\mathbf{w}}, \mathbf{w}' \rangle > \varepsilon$ for all $\tilde{\mathbf{w}} \in \operatorname{conv} \mathcal{P}^{\mathbf{G}_1}$. Therefore, in order to prove (2.8), it is sufficient to show that for each fixed $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$, $\mathbf{w}' \in \mathcal{V}'$ and $\varepsilon > 0$ one can find $\tilde{\mathbf{w}} \in \operatorname{conv} \mathcal{P}^{\mathbf{G}_1}$ such that $\langle \mathbf{w}, \mathbf{w}' \rangle - \langle \tilde{\mathbf{w}}, \mathbf{w}' \rangle \leq \varepsilon$.

Assume that the intersection $G_k \bigcap G_l \bigcap G_{\mathbf{w}}$ contains more than one vertex so that $G_k \bigcap G_l \bigcap G_{\mathbf{w}} = \{g_1, g_2, \ldots\}$. Since $\sum_j |\mathbf{w}(g_j)\, \mathbf{w}'(g_j)| \leq \infty$ and $\sum_j \mathbf{w}(g_j) \leq 1$, we have $\mathbf{w}'(g_i) \geq \sum_j \mathbf{w}(g_j)\, \mathbf{w}'(g_j)$ for some positive integer i. If $\mathbf{w}^\star(g) := \mathbf{w}(g)$ whenever $g \notin G_k \bigcap G_l \bigcap G_{\mathbf{w}}$, $\mathbf{w}^\star(g_i) := \sum_j \mathbf{w}(g_j)$ and $\mathbf{w}^\star(g) := 0$ whenever $g \in G_k \bigcap G_l \bigcap G_{\mathbf{w}}$ but $g \neq g_i$ then $\mathbf{w}^\star \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}$ and $\langle \mathbf{w}^\star, \mathbf{w}' \rangle \geq \langle \mathbf{w}, \mathbf{w}' \rangle$. Therefore we can assume without loss of generality that $G_{\mathbf{w}}$ satisfies the condition (c_1) of Proposition 2.2.

Let us enumerate the sets G_k and define F_n and n_0 as in Lemma 2.14. Let $n \geq n_0$ and $G_k \bigcap G_{\mathbf{w}} = \{g_1^k, g_2^k, ...\}$, where $k = 1, 2, \ldots, n$. By $(\mathbf{g}_1)$, for every g_j^k there exists at most one positive integer $l \neq k$ such that $g_j^k \in G_l$. Denote

$$v_n(g_j^k) \ := \ \begin{cases} \mathbf{w}(g_j^k), & \text{if } g_j^k \notin \bigcup_{l=n+1}^{\infty} G_l, \\ \sum_{g \in G_l \bigcap F_n} \mathbf{w}(g), & \text{if } g_j^k \in G_l \text{ for some } l > n. \end{cases}$$

In view of $(\mathbf{g}_1)$ and (c_1), we have $\sum_j v_n(g_j^k) \leq n$. Therefore $v_n(g_j^k) \to 0$ as $j \to \infty$ whenever the set $G_k \bigcap G_{\mathbf{w}}$ is infinite. If $G_k \bigcap G_{\mathbf{w}}$ is finite, denote by j_k the number of elements of $G_k \bigcap G_{\mathbf{w}}$. If $G_k \bigcap G_{\mathbf{w}}$ is infinite, denote by j_k the minimal positive integer such that

$$v_n(g_{j_k}^k) + \sum_{j > j_k} \mathbf{w}(g_j^k) \ \leq \ 1 \qquad \text{and} \qquad g_j^k \notin F_n \setminus G_k, \quad \forall j \geq j_k \tag{2.9}$$

(since $v_n(g_j^k) \to 0$, $\sum_j \mathbf{w}(g_j^k) \leq 1$ and $G_{\mathbf{w}}$ satisfies (c_1), such a minimal integer exists).

Let $\mathbf{w}_n(g) := 0$ whenever $g \notin F_n$ and $\mathbf{w}_n(g) := \mathbf{w}(g)$ for all $g \in F_n$. Then $\langle \mathbf{w} - \mathbf{w}_n, \mathbf{w}' \rangle \to 0$ as $n \to \infty$ because the series $\sum_{g \in G} \mathbf{w}(g)\, \mathbf{w}'(g)$ is absolutely convergent. Let $m \geq \max\{j_1, j_2, \ldots, j_n\}$, $\mathbf{w}_{n,m}(g) := 0$ whenever

$\mathbf{w}_n(g) = 0$ and

$$\mathbf{w}_{n,m}(g^k_j) \ := \ \begin{cases} 0 & \text{if } j > m, \\ \mathbf{w}(g^k_{j_k}) + \sum_{j>m} \mathbf{w}(g^k_j) & \text{if } j = j_k \le m, \\ \mathbf{w}(g) & \text{if } j \le m \text{ and } j \ne j_k\,. \end{cases}$$

Then $\langle \mathbf{w}_n - \mathbf{w}_{n,m}, \mathbf{w}'\rangle \to 0$ as $m \to \infty$ for each fixed n because the series $\sum_j \mathbf{w}(g^k_j)\,\mathbf{w}'(g^k_j)$ are absolutely convergent and $\sum_{j>m} \mathbf{w}(g^k_j) \to 0$.

The weight $\mathbf{w}_{n,m}$ vanishes outside a finite subset of G and, in view of (2.9), belongs to $\mathcal{S}^\emptyset$ and satisfies the condition (2.7). Applying Theorem 2.4 to the family of sets $\{G_1 \bigcap G_{\mathbf{w}_{n,m}}, \dots, G_n \bigcap G_{\mathbf{w}_{n,m}}\}$ and then the Krein–Milman theorem, we see that $\mathbf{w}_{n,m}$ can be represented as a finite convex combination $\sum_i \alpha_i\, \mathbf{w}^{(i)}_{n,m}$ of some weights $\mathbf{w}^{(i)}_{n,m} \in \mathcal{P}^\emptyset$. Obviously, each weight $\mathbf{w}^{(i)}_{n,m}$ also satisfies (2.7). By Lemma 2.14, we can find $\tilde{\mathbf{w}}^{(i)}_{n,m} \in \mathcal{P}^{\mathbf{G}_1}$ such that $\tilde{\mathbf{w}}^{(i)}_{n,m} = \tilde{\mathbf{w}}^{(i)}_{n,m}$ on the set F_n. If $\tilde{\mathbf{w}}_{n,m} := \sum_i \alpha_i\, \tilde{\mathbf{w}}^{(i)}_{n,m}$ then $\tilde{\mathbf{w}}_{n,m} \in \operatorname{conv} \mathcal{P}^{\mathbf{G}_1}$ and, in view of (2.4), we have $\langle \mathbf{w}_{n,m} - \tilde{\mathbf{w}}_{n,m}, \mathbf{w}'\rangle < \varepsilon/3$ for all $m \ge \max\{j_1, j_2, \dots, j_n\}$ provided that n is sufficiently large. Therefore, choosing a sufficiently large $n \ge n_0$ and then a sufficiently large $m \ge \max\{j_1, j_2, \dots, j_n\}$, we can make the right hand side of the identity

$$\langle \mathbf{w}, \mathbf{w}'\rangle - \langle \tilde{\mathbf{w}}_{n,m}, \mathbf{w}'\rangle \ = \ \langle \mathbf{w} - \mathbf{w}_n, \mathbf{w}'\rangle + \langle \mathbf{w}_n - \mathbf{w}_{n,m}, \mathbf{w}'\rangle + \langle \mathbf{w}_{n,m} - \tilde{\mathbf{w}}_{n,m}, \mathbf{w}'\rangle$$

smaller than ε. □

Remark 2.16. Theorem 2.15 implies that $\mathcal{P}^{\mathbf{G}_1} \ne \emptyset$ whenever $\mathcal{S}^{\mathbf{G}_1} \ne \emptyset$ and $\mathbf{G}$ satisfies ($\mathbf{g}_1$)–($\mathbf{g}_3$). If $G_{n+1} \not\subset \bigcup_{k=1}^n G_j$ for all $n = 1, 2, \dots$ then, using the same procedure as in the proof of Lemma 2.14, one can show that $\mathcal{P}^{\mathbf{G}} \ne \emptyset$.

Remark 2.17. If $\mathcal{V}$ is a proper normal subspace of $\mathcal{V}_1$ then $\mathcal{V}_1'$ is a proper subspace of $\mathcal{V}'$ and the Mackey topology $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}, \mathcal{V}')$ is strictly finer than $\mathfrak{T}_{\mathrm{m}}(\mathcal{V}_1, \mathcal{V}_1')$. Therefore choosing a smaller space $\mathcal{V}$ in Theorem 2.15 we obtain a stronger result which is valid for a narrower class of $\mathbf{G}_1$-stochastic weights.

Remark 2.18. Taking $\mathcal{V} = \mathcal{V}_{\mathcal{S}}$ in Theorem 2.15 and applying Corollary 2.12, we obtain $\mathcal{S}^{\mathbf{G}_1} = \overline{\operatorname{conv}}\, \mathcal{P}^{\mathbf{G}_1}$, where the closure is taken in the topology generated by the seminorms p_k. This topology is metrizable. Therefore, under conditions ($\mathbf{g}_1$)–($\mathbf{g}_3$), for every $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1}$ there exists a sequence of weights $\mathbf{w}_n \in \operatorname{conv} \mathcal{P}^{\mathbf{G}_1}$ such that $p_k(\mathbf{w} - \mathbf{w}_n) \to 0$ as $n \to \infty$ for all $k = 1, 2, \dots$

The following simple example shows that, generally speaking, $\overline{\operatorname{conv}}\, \mathcal{P}^\emptyset$ does not contain $\mathcal{S}^{\mathbf{G}} \bigcap \mathcal{V}$ if we take the closure in the strong topology $\mathfrak{T}_{\mathrm{b}}(\mathcal{V}, \mathcal{V}')$.

Example 2.19. Let $\mathbf{G}$ be an infinite collection of mutually disjoint sets G_k such that G_k contains k elements. Then the weight $\mathbf{w}$ which takes the values k^{-1} on G_k belongs to $\mathcal{S}^{\mathbf{G}}$. On the other hand, for every weight $\tilde{\mathbf{w}} \in \operatorname{conv} \mathcal{P}^\emptyset$ there exists a positive integer n such that the number of nonzero entries in $\tilde{\mathbf{w}}|_{G_k}$ does not exceed n for every k. Therefore $\|\mathbf{w} - \tilde{\mathbf{w}}\|_{\mathcal{S}} = 1$ for all $\tilde{\mathbf{w}} \in \operatorname{conv} \mathcal{P}^\emptyset$, where $\|\cdot\|_{\mathcal{S}}$ is defined by (2.6).

The strong closure of the convex hull of the set of permutation matrices is also strictly smaller then the set of doubly stochastic matrices [Is].

3. Operators generated by stochastic matrices

3.1. **Notation and definitions.** In the rest of the paper (with the exception of the proof of Theorem 3.15) we shall be assuming that $\mathcal{W}$ is the space of real matrices $\mathbf{w} = \{w_{ij}\}_{i,j=1,2,\dots}$ and the sets G_k are the rows and columns (see Example 2.3). Recall that in this case $\mathbf{G}$ satisfies the conditions $(\mathbf{g}_1)$–$(\mathbf{g}_3)$, $\mathcal{S}^{\mathbf{G}}$ and $\mathcal{S}^{\emptyset}$ are the sets of doubly stochastic and sub-stochastic matrices respectively, $\mathcal{P}^{\mathbf{G}}$ is the set of permutation matrices and $\mathcal{P}^{\emptyset}$ is the set of sub-stochastic matrices whose entries are equal either to 0 or to 1. For the sake of definiteness we shall consider only infinite matrices; the corresponding results for finite matrices are much simpler and can be proved in a similar manner.

Every matrix $\mathbf{w} \in \mathcal{W}$ generates the linear operator

$$\mathbb{R}^\infty \ni \mathbf{x} \;\to\; \{\sum_{j=1}^{\infty} w_{1j}\, x_j\,,\; \sum_{j=1}^{\infty} w_{2j}\, x_j\,,\dots\} \;\in\; \mathbb{R}^\infty \tag{3.1}$$

with domain $\mathcal{D}(\mathbf{w}) \;=\{\mathbf{x} \in \mathbb{R}^\infty\;:\sum_{j=1}^{\infty} |\, w_{ij}\, x_j\,| < \infty\,,\ \forall i = 1,2,\dots\}$. We shall denote this operator by the same letter $\mathbf{w}$. Obviously, $l^\infty \subseteq \mathcal{D}(\mathbf{w})$ for all $\mathbf{w} \in \mathcal{S}^{\emptyset}$ and $\mathcal{D}(\mathbf{w}) = \mathbb{R}^\infty$ for all $\mathbf{w} \in \mathcal{P}^{\emptyset}$, but $\mathcal{D}(\mathbf{w}) \neq \mathbb{R}^\infty$ whenever $\mathbf{w}$ has a row with infinitely many nonzero entries.

Lemma 3.1. *If $X \subseteq l^\infty$ is a symmetric perfect space and $X \neq \mathbb{R}^\infty_{00}$ then the operator generated by a matrix $\mathbf{w} \in \mathcal{S}^{\emptyset}$ maps X into X and is continuous in the topologies $\mathfrak{T}_{\mathrm{w}}(X,X')$, $\mathfrak{T}_{\mathrm{m}}(X,X')$ and $\mathfrak{T}_{\mathrm{b}}(X,X')$.*

Proof. Since $X \neq \mathbb{R}^\infty_{00}$, by Remark 1.11 we have $X' \subseteq l^\infty$. The inclusions $X \subseteq l^\infty$, $X' \subseteq l^\infty$ and (2.5) imply that $\mathbf{x}' \otimes \mathbf{x} \in \mathcal{V}'_{\mathcal{S}}$ for all $\mathbf{x} \in X$ and $\mathbf{x}' \in X'$, which means that $\mathbf{w}$ maps the perfect space X into itself. Similarly, the transposed operator $\mathbf{w}^T$ maps the perfect space X' into itself. Therefore $|\langle \mathbf{w}\mathbf{x}, \mathbf{x}'\rangle| = |\langle \mathbf{x}, \mathbf{w}^T\mathbf{x}'\rangle|$ is a $\mathfrak{T}_{\mathrm{w}}(X,X')$-continuous seminorm on X for each $\mathbf{x}' \in X'$ and is a $\mathfrak{T}_{\mathrm{w}}(X',X)$-continuous seminorm on X' for each $\mathbf{x} \in X$. This implies that $\mathbf{w}$ is $\mathfrak{T}_{\mathrm{w}}(X,X')$-continuous and $\mathbf{w}^T$ is $\mathfrak{T}_{\mathrm{w}}(X',X)$-continuous. Since the continuous operator $\mathbf{w}^T$ maps compact sets into compact sets and bounded sets into bounded sets, the operator $\mathbf{w}$ is $\mathfrak{T}_{\mathrm{m}}(X,X')$-continuous and $\mathfrak{T}_{\mathrm{b}}(X,X')$-continuous. □

Remark 3.2. Let $X \subseteq l^\infty$ be a symmetric space, $X \not\subset \mathbb{R}^\infty_{00}$ and $\{\mathbf{x}' \otimes \mathbf{x}\}$ be the set which contains one element $\mathbf{x}' \otimes \mathbf{x}$, where $\mathbf{x} \in X$ and $\mathbf{x}' \in X'$. Applying (2.4) to $D' = \{\mathbf{x}' \otimes \mathbf{x}\}$, we see that the set $P^{\emptyset}_{\mathbf{x}}$ is $\mathfrak{T}_{\mathrm{w}}(X,X')$-bounded. Therefore Theorem 1.1 implies the first statement of Corollary 1.7. If Ω' is a $\mathfrak{T}_{\mathrm{w}}(X',X)$-compact subset of X' then, by Theorem 1.2, the set $D' := \bigcup_{\mathbf{x}' \in \Omega'} \{\mathbf{x}' \otimes \mathbf{x}\}$ is $\mathfrak{T}_{\mathrm{w}}(\mathcal{V}'_{\mathcal{P}}, \mathcal{V}_{\mathcal{P}})$-compact. Therefore Theorem 1.2 and (2.4) imply the second statement of Corollary 1.7.

Definition 3.3. Let $\mathbf{G}_r$ be the set of all rows, $\mathcal{S}^r := \mathcal{S}^{\mathbf{G}_r}$ and $\mathcal{US}^r$ be the set of matrices $\mathbf{w} = \{w_{ij}\}_{i,j=1,2,\dots} \in \mathcal{S}^r$ such that $w_{ij} = |(u_i, e_j)_H|^2$, where $\{e_1, e_2, \dots\}$ is a complete orthonormal subset of a separable complex Hilbert space H, $\{u_1, u_2, \dots\}$ is an orthonormal subset of the same Hilbert space H and $(\cdot,\cdot)_H$ is the inner product in H.

If the set $\{u_1, u_2, \dots\}$ is also complete then the inner products $(u_i, e_j)_H$ coincide with entries of a unitary matrix. In this case the corresponding matrix $\mathbf{w} \in \mathcal{US}^r$ is doubly stochastic and is said to be *unistochastic*. In the finite dimensional case every matrix $\mathbf{w} \in \mathcal{US}^r$ is unistochastic.

Definition 3.4. If $\mathbf{x} = \{x_1, x_2, \dots\} \in \mathbb{R}^\infty$, let

$$R_m^+(\mathbf{x}) := \sup_{\{x_{j_1},\dots,x_{j_m}\}} \sum_{n=1}^m x_{j_n} \quad \text{and} \quad R_m^-(\mathbf{x}) := \inf_{\{x_{j_1},\dots,x_{j_m}\}} \sum_{n=1}^m x_{j_n}, \tag{3.2}$$

where $m = 1, 2, \dots$ and the supremum and infimum are taken over all subsets of $\mathbf{x}$ containing m elements. Denote by $Q_\mathbf{x}$ the set of all sequences $\mathbf{y} = \{y_1, y_2, \dots\} \in \mathbb{R}^\infty$ such that

$$R_m^-(\mathbf{x}) \;\le\; \sum_{n=1}^m y_{i_n} \;\le\; R_m^+(\mathbf{x}) \tag{3.3}$$

for each $m = 1, 2, \dots, p$ and each collection of m distinct positive integers $i_1, \dots, i_m$. Finally, let $X_{Q_\mathbf{x}}$ be the subspace of $\mathbb{R}^\infty$ spanned by $Q_\mathbf{x}$.

By Remark 1.11, $X_{Q_\mathbf{x}}$ is the minimal symmetric perfect space containing $\mathbf{x}$ whenever $\mathbf{x} \notin \mathbb{R}^\infty_{00}$ and $X_{Q_\mathbf{x}} = l^1$ whenever $\mathbf{x} \in l^1 \setminus \{0\}$.

Definition 3.5. If $\mathbf{x} = \{x_1, x_2, \dots\} \in \mathbb{R}^\infty$, let

$\mathcal{V}_\mathbf{x}$ be the linear space of matrices $\mathbf{w}$ such that $\mathbf{x} \in \mathcal{D}(\mathbf{w})$;
$P_\mathbf{x}^{\mathbf{G}_1}$, $S_\mathbf{x}^{\mathbf{G}_1}$ and $US_\mathbf{x}^r$ be the sets of all sequences $\mathbf{y} \in \mathbb{R}^\infty$ such that $\mathbf{y} = \mathbf{w}\mathbf{x}$ for some $\mathbf{w} \in \mathcal{P}^{\mathbf{G}_1}$, $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}_\mathbf{x}$ and $\mathbf{w} \in \mathcal{US}^r \bigcap \mathcal{V}_\mathbf{x}$ respectively and $S_\mathbf{x}^r := S_\mathbf{x}^{\mathbf{G}_r}$.

Obviously, the sets $S_\mathbf{x}^{\mathbf{G}_1}$, $P_\mathbf{x}^{\mathbf{G}_1}$ $Q_\mathbf{x}$ do not depend on the order of entries in the sequence $\mathbf{x}$. We have $S_\mathbf{x}^{\mathbf{G}} \subset S_\mathbf{x}^r \subset S_\mathbf{x}^\emptyset$, $P_\mathbf{x}^{\mathbf{G}_r} = P_\mathbf{x}^r \subset US_\mathbf{x}^r \subseteq S_\mathbf{x}^r$ and $P_\mathbf{x}^{\mathbf{G}} = P_\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^\infty$ (see Definition 1.6).

Lemma 3.6. *Let $\mathbf{x} := \{x_1, x_2, \dots\} \in \mathbb{R}^\infty$, $\{e_1, e_2, \dots\}$ be a complete orthonormal subset of a separable complex Hilbert space H and A be the self-adjoint operator in H such that $Ae_j = x_j\, e_j$. Then $\mathbf{y} \in US_\mathbf{x}^r$ if and only if there exists an orthonormal set $\{u_i\} \subset \mathcal{D}(|A|^{1/2})$ such that $y_i := (Au_i, u_i)_H$.*

Proof. A sequence $\mathbf{y}$ belongs to $US_\mathbf{x}^r$ if and only if $y_i = \sum_j |(u_i, e_j)_H|^2 x_j$, where $\{u_i\}$ is an orthonormal set such that $\sum_j |(u_i, e_j)_H|^2 |x_j| < \infty$ for each $i = 1, 2, \dots$ These estimates are equivalent to the inclusion $\{u_i\} \subset \mathcal{D}(|A|^{1/2})$. Since $u_i = \sum_j (u_i, e_j)_H\, e_j$, we have $y_i = \sum_j |(u_i, e_j)_H|^2 x_j = (Au_i, u_i)$. $\square$

3.2. **The sets $P^r_{\mathbf{x}}$ $US^r_{\mathbf{x}}$, $S^r_{\mathbf{x}}$ and $Q_{\mathbf{x}}$.** The main result of this subsection is Theorem 3.10 which clarifies the relation between these sets. Given a sequence $\mathbf{x}$ and a set $\Lambda\subset\hat{\mathbb{R}}$, we shall denote by $\mathbf{x}\bigcap\Lambda$ the sequence obtained from $\mathbf{x}$ by removing all its entries lying outside Λ.

Lemma 3.7. *Assume that the sequence $\mathbf{x}\in\mathbb{R}^\infty$ has one accumulation point $\lambda\in\hat{\mathbb{R}}$, $\mathbf{y}\in Q_{\mathbf{x}}$ and $\mathbf{y}\bigcap(-\infty,\lambda)=\emptyset$. Then $\mathbf{y}\in US^r_{\mathbf{x}}$ provided that*

(a) *either $\mathbf{x}\bigcap[\lambda,+\infty)$ is infinite and $\sharp\{i: y_i=\lambda\}\le\sharp\{j: x_j=\lambda\}$*

(b) *or $\mathbf{x}\bigcap[\lambda,+\infty)$ is finite and $\sum_j(x_j-\lambda)_+-\sum_i(y_i-\lambda)_+=\varepsilon>0$.*

Proof. Let A be defined as in Lemma 3.6. In order to prove the inclusion $\mathbf{y}\in US^r_{\mathbf{x}}$, we have to find an orthonormal set $\{u_1,u_2,\dots\}\subset\mathcal{D}(|A|^{1/2})$ such that $y_i:=(Au_i,u_i)_H$.

Assume first that (a) holds true. Then there are two entries $x_{j_1},x_{k_1}\in\mathbf{x}\bigcap[\lambda,+\infty)$ such that $y_1\in[x_{j_1},x_{k_1}]$ and $\mathbf{x}\bigcap(x_{j_1},x_{k_1})=\emptyset$. If $y_1=\alpha x_{i_1}+(1-\alpha)x_{k_1}$ and $u_1:=\alpha^{1/2}e_{j_1}+(1-\alpha)^{1/2}e_{k_1}$ then $\|u_1\|_H=1$ and $y_1=(Au_1,u_1)_H$. Let $\mathbf{x}^{(1)}$ be the sequence obtained from $\mathbf{x}$ by replacing the two entries x_{i_1} and x_{k_1} with one entry $x_{i_1}+x_{k_1}-y_1$ and $\mathbf{y}^{(1)}$ be the sequence obtained from $\mathbf{y}$ by removing the entry y_1. The entries of $\mathbf{x}^{(1)}$ coincide with the eigenvalues of the self-adjoint operator $A_1:=\Pi_1A|_{H_1}$ in the Hilbert space $H_1:=\Pi_1H$, where Π_1 is the orthogonal projection onto the annihilator of u_1.

If $y_1=\lambda$ then at least one of the entries x_{j_1},x_{k_1} coincides with λ, which implies that $\mathbf{x}^{(1)}$ and $\mathbf{y}^{(1)}$ are obtained from $\mathbf{x}$ and $\mathbf{y}$ by removing one entry λ. Therefore the sequences $\mathbf{x}^{(1)}$ and $\mathbf{y}^{(1)}$ satisfy the condition (a). We also have $\mathbf{y}^{(1)}\in Q_{\mathbf{x}^{(1)}}$. Indeed, if the number of entries in $\mathbf{x}$ lying in the interval $(x_{k_1},+\infty)$ is equal to p then $R^+_m(\mathbf{x}^{(1)})=R^+_m(\mathbf{x})$ whenever $m<p$. If $m\ge p$ then $R^+_m(\mathbf{x}^{(1)})=R^+_{m+1}(\mathbf{x})-y_1\ge\sum_{k=1}^m y_{l_k}$ for each subset $\{y_{l_1},\dots,y_{l_m}\}\subset\mathbf{y}^{(1)}$.

Applying the same procedure to $\mathbf{x}^{(i-1)}$, $\mathbf{y}^{(i-1)}$ and A_{i-1} with $i=2,3,\dots$, we can find $x_{j_i},x_{k_i}\in\mathbf{x}^{(i-1)}$, $u_i\in H$ and $\mathbf{x}^{(i)}$ such that $y_i\in[x_{j_i},x_{k_i}]$, $\mathbf{x}^{(i-1)}\bigcap(x_{j_i},x_{k_i})=\emptyset$, $\Pi_{i-1}u_i=0$, $\|u_i\|_H=1$, $y_i=(A_{i-1}u_i,u_i)_H$ and $R^+_m(\mathbf{y}^{(i)})\le R^+_m(\mathbf{x}^{(i)})$. The entries of $\mathbf{x}^{(i)}$ coincide with the eigenvalues of $A_i:=\Pi_iA_{i-1}|_{\Pi_iH}$, where Π_i is the orthogonal projection onto the annihilator H_i of the set $\{u_1,\dots,u_i\}$. The set $\{u_1,u_2,\dots\}$, obtained by induction in i, is orthonormal and every its element u_i is a finite linear combination of the eigenvectors $e_1,e_2,\dots$ The latter implies that $u_i\in\mathcal{D}(A)\subset\mathcal{D}(|A|^{1/2})$ and $y_i=(A_{i-1}u_i,u_i)_H=(Au_i,u_i)_H$ for all $i=1,2,3,\dots$

If (b) holds true then λ is an accumulation point of $\mathbf{x}\bigcap(-\infty,\lambda)$. Without loss of generality we may assume that the sequence $\mathbf{x}\bigcap(-\infty,\lambda)$ converges to λ and that $\sum_j(\lambda-x_j)_+<\varepsilon/2$ (this can always be achieved by removing a collection of entries from $\mathbf{x}$). Let us denote $\mathbf{x}^{(0)}:=\mathbf{x}$, $\mathbf{y}^{(0)}:=\mathbf{y}$ and apply the same procedure as above with x_{j_i},x_{k_i} defined as follows:

x_{k_i} is the smallest entry of $\mathbf{x}^{(i-1)}$ lying in the interval $[y_i,+\infty)$,

x_{j_i} is either the largest entry of $\mathbf{x}^{(i-1)}$ lying in (λ, y_i) or, if such an entry does not exists, x_{j_i} is an arbitrary entry of $\mathbf{x}\bigcap(-\infty,\lambda)$.

The inequality $\sum_j(\lambda - x_j)_+ < \varepsilon/2$ implies that $R_m^+(\mathbf{y}^{(i)}) < R_m^+(\mathbf{x}^{(i)}) - \varepsilon/2$ for all $i, m = 1, 2, \ldots$. Therefore, by induction in i, we can find the required representation for all entries y_i. □

Lemma 3.8. *Assume that the sequence* $\mathbf{x} \in \mathbb{R}^\infty$ *has two accumulation points* $\lambda, \mu \in \hat{\mathbb{R}}$ *and* $\lambda < \mu$. *If* $\mathbf{y} = \mathbf{y}\bigcap[\lambda,\mu]$, $\sharp\{i : y_i = \lambda\} \le \sharp\{j : x_j \le \lambda\}$ *and* $\sharp\{i : y_i = \mu\} \le \sharp\{j : x_j \ge \mu\}$ *then* $\mathbf{y} \in US^r_{\mathbf{x}}$.

Proof. Under the conditions of the lemma, there exists a set of distinct positive integers $\{j_1, j_2, \ldots, k_1, k_2, \ldots\}$ such that $y_i \in [x_{j_i}, x_{k_i}]$ for all $i = 1, 2, \ldots$ If $y_i = \alpha_i x_{j_i} + (1-\alpha_i)x_{k_i}$ and A is defined as in Lemma 3.6 then $y_i = (Au_i, u_i)_H$, where $u_i := \alpha_i^{1/2} e_{j_i} + (1-\alpha_i)^{1/2} e_{k_i}$. □

Definition 3.9. If $\mathbf{x} = \{x_1, x_2, \ldots\} \in \mathbb{R}^\infty$, let $x^- := \liminf_{j\to\infty} x_j \in \hat{\mathbb{R}}$, $x^+ := \limsup_{j\to\infty} x_j \in \hat{\mathbb{R}}$ and $\hat{\mathbf{x}}$ be the sequence obtained from $\mathbf{x}$ by adding infinitely many entries x^- whenever $x^- > -\infty$ and infinitely many entries x^+ whenever $x^+ < +\infty$.

Theorem 3.10. *For every* $\mathbf{x} = \{x_1, x_2, \ldots\} \in \mathbb{R}^\infty$ *we have*

$$US^r_{\mathbf{x}} \;=\; S^r_{\mathbf{x}} \;\subseteq\; Q_{\mathbf{x}} = Q_{\hat{\mathbf{x}}} \;=\; US^r_{\hat{\mathbf{x}}}. \tag{3.4}$$

Proof. The equality $Q_{\mathbf{x}} = Q_{\hat{\mathbf{x}}}$ immediately follows from the definition of $Q_{\mathbf{x}}$. If $\mathbf{y} \in S^r_{\mathbf{x}}$ then for every collection of m distinct positive integers $i_1, \ldots, i_m$ we have $\sum_{k=1}^m y_{i_k} = \sum_j \alpha_j x_j$, where $\alpha_j \in [0,1]$ and $\sum_j \alpha_j = m$. This implies (3.3). Therefore $S^r_{\mathbf{x}} \subseteq Q_{\mathbf{x}}$.

It remains to prove that $\mathbf{y} \in US^r_{\mathbf{x}}$ provided that either $\mathbf{y} \in S^r_{\mathbf{x}}$ or $\mathbf{x} = \hat{\mathbf{x}}$ and $\mathbf{y} \in Q_{\mathbf{x}}$. We are going to show that there exist countable families of disjoint subsequences $\mathbf{x}_n \subset \mathbf{x}$ and $\mathbf{y}_n \subset \mathbf{y}$ such that $\bigcup_n \mathbf{x}_n = \mathbf{x}$, $\bigcup_n \mathbf{y}_n = \mathbf{y}$ and $\mathbf{y}_n \in US^r_{\mathbf{x}_n}$. Obviously, this implies that $\mathbf{y} \in US^r_{\mathbf{x}}$. Given a sequence $\mathbf{z}$, in the rest of the proof we shall denote $\mathbf{z}^+ := \mathbf{z}\bigcap(x^+, +\infty)$, $\mathbf{z}^- := \mathbf{z}\bigcap(-\infty, x^-)$, $\mathbf{z}_+ := \mathbf{z}\bigcap[x^+, +\infty)$ and $\mathbf{z}_- := \mathbf{z}\bigcap(-\infty, x^-]$.

Assume first that $\mathbf{x} = \hat{\mathbf{x}}$. Then we can split $\mathbf{x}$ into the union of three disjoint subsequences $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ such that $\mathbf{x}_1 = \hat{\mathbf{x}}^+$, $\mathbf{x}_2 = \hat{\mathbf{x}}^-$, $\mathbf{x}_3$ does not have any entries lying outside $[x^-, x^+]$ and $\mathbf{x}_3$ has infinitely many entries $x^\pm$ whenever $x^\pm$ is finite. If $\mathbf{y}_1 := \mathbf{y}_+$, $\mathbf{y}_2 := \mathbf{y}_-$ and $\mathbf{y}_0 := \mathbf{y}\bigcap(x^-, x^+)$ then, by Lemmas 3.7 and 3.8, we have $\mathbf{y}_n \in Q_{\mathbf{x}_n}$ whenever $\mathbf{y} \in Q_{\mathbf{x}}$. Therefore $Q_{\hat{\mathbf{x}}} \subset US^r_{\hat{\mathbf{x}}}$.

Assume that $\mathbf{y} \in S^r_{\mathbf{x}}$. We have to consider the following possibilities:

(1_+) $\mathbf{y}^+ \ne \emptyset$, $\mathbf{x}^+$ is infinite and

$$\liminf_{m\to\infty} \left(R_m^+(\mathbf{x}^+) - R_m^+(\mathbf{y}^+)\right) \;=\; 0\,; \tag{3.5}$$

(2_+) $\mathbf{y}^+ \ne \emptyset$, $\mathbf{x}^+$ is finite and (3.5) holds true;

(3_+) $\mathbf{y}^+ \neq \emptyset$, $\mathbf{x}^+$ is infinite and

$$R_m^+(\mathbf{x}^+) - R_m^+(\mathbf{y}^+) \geq \varepsilon > 0\,, \qquad \forall m = 1, 2, \dots\,; \tag{3.6}$$

(4_+) $\mathbf{x}^+ \neq \emptyset$ is finite and (3.6) holds true;
(5_+) $\mathbf{y}^+ = \emptyset$ and $\mathbf{x}^+$ is infinite;
($6'_+$) $\mathbf{y}^+ = \emptyset$ and $\mathbf{x}^+ = \emptyset$.

Note that ($6'_+$) and the inclusion $\mathbf{y} \in S^r_{\mathbf{x}}$ imply

(6_+) $\mathbf{y}^+ = \emptyset$, $\mathbf{x}^+ = \emptyset$ and $\sharp\{i : y_i = x^+\} \leq \sharp\{j : x_j = x^+\}$.

We shall say that $\mathbf{x}$ and $\mathbf{y}$ satisfy (n_-) if the corresponding condition (n_+) is fulfilled for $-\mathbf{x}$ and $-\mathbf{y}$.

Assume first that (1_+) holds true. By Lemma 3.7, we have $\mathbf{y}^+ \in SU^r_{\mathbf{x}^+}$. Let $\tilde{\mathbf{y}} := \mathbf{y} \setminus \mathbf{y}^+$ and $\tilde{\mathbf{x}} := \mathbf{x} \setminus \mathbf{x}^+$ be the sequences obtained from $\mathbf{y}$ and $\mathbf{x}$ by removing all the entries $y_i \in \mathbf{y}^+$ and $x_j \in \mathbf{x}^+$ respectively. If $\mathbf{y} = \mathbf{w}\mathbf{x}$ and $\mathbf{w} \in \mathcal{S}^r$ then, in view of (3.5), the entry w_{ij} of the matrix $\mathbf{w}$ is equal to zero whenever $x_j > x^+$ and $y_i \leq x^+$. Therefore $\tilde{\mathbf{y}} = \tilde{\mathbf{w}}\tilde{\mathbf{x}}$, where $\tilde{\mathbf{w}} \in \mathcal{S}^r$ is the matrix obtained from $\mathbf{w}$ by crossing out all the ith rows corresponding to $y_i \in \mathbf{y}^+$. If $\limsup_j \tilde{x}_j = \tilde{x}^+ < x^+$ and $\tilde{\mathbf{x}}, \tilde{\mathbf{y}}$ satisfy (1_+) then, in a similar manner, we remove the subsequences $\tilde{\mathbf{x}}^+ := \tilde{\mathbf{x}} \bigcap (\tilde{x}^+, +\infty)$ and $\tilde{\mathbf{y}}^+ := \tilde{\mathbf{y}} \bigcap (\tilde{x}^+, +\infty)$. After sufficiently (possibly, infinitely) many iterations we either obtain two required families of disjoint subsequences $\mathbf{x}_n$ and $\mathbf{y}_n$ or end up with two remaining sequences satisfying one of the conditions (2_+)–(6_+). If (1_-) holds true then we can apply the same procedure to the sequences $-\mathbf{x}$ and $-\mathbf{y}$. Therefore it is sufficient to consider the sequences $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{y} \in S^r_{\mathbf{x}}$ and one of the conditions ($2_\pm$)–($6_\pm$) is fulfilled.

Assume that (2_+) is fulfilled and $\mathbf{y} = \mathbf{w}\mathbf{x}$, where $\mathbf{w} \in \mathcal{S}^r$. If $\mathbf{x}$ has finitely many entries x^+, we define $\mathbf{x}_\star := \mathbf{x}_+$. The condition (3.5) imply that the entry w_{ij} of the matrix $\mathbf{w}$ is equal to zero whenever $x_j < x^+$ and $y_i \geq x^+$. Therefore the number of entries in $\mathbf{y}_+$ does not exceed the number of entries in $\mathbf{x}_\star$ and $\mathbf{y}_+ = \mathbf{w}_\star \mathbf{x}_\star$ for some finite matrix $\mathbf{w}_\star \in \mathcal{S}^r$. In the same way as in the proof of Lemma 3.7 one can show that $\mathbf{y}_+ \in S^r_{\mathbf{x}_\star}$. If $\mathbf{x}$ has infinitely many entries x^+ then we represent $\mathbf{x}$ as the union of two disjoint subsequences $\tilde{\mathbf{x}}$ and $\mathbf{x}_\star$ such that $\mathbf{x}_\star = \hat{\mathbf{x}}^+$, $\tilde{\mathbf{x}} \bigcap (x^+, +\infty) = \emptyset$ and $\tilde{\mathbf{x}}$ contains infinitely many entries x^+. By Lemma 3.7, we have $\mathbf{y}_+ \in US^r_{\mathbf{x}_\star}$.

In both cases the sequences $\mathbf{x} \setminus \mathbf{x}_\star$ and $\mathbf{y} \setminus \mathbf{y}_+$ satisfy (6_+). If (2_-) holds true then, in a similar way, we can remove all the entries lying below x^-. Therefore it is sufficient to prove the inclusion $\mathbf{y} \in US^r_{\mathbf{x}}$ assuming that $\mathbf{x}$ and $\mathbf{y}$ satisfy (3.3) and one of the conditions ($3_\pm$)–($6_\pm$).

If (3_+) is fulfilled then we choose a subsequence $\mathbf{x}^\star$ of the sequence $\mathbf{x}^+$ in such a way that the remaining sequence $\mathbf{x}^+ \setminus \mathbf{x}^\star$ contains infinitely many entries and $R_m^+(\mathbf{y}^+) \leq R_m^+(\mathbf{x}^\star)$ for all $m = 1, 2, \dots$ By Lemma 3.7, $\mathbf{y}_+ \in US^r_{\mathbf{x}_\star}$ If we remove all entries $x_j \in \mathbf{x}^\star$ and $y_i \in \mathbf{y}^+$ then the remaining sequences $\mathbf{x} \setminus \mathbf{x}^\star$ and $\mathbf{y} \setminus \mathbf{y}^+$ satisfy (5_+). Similarly, if (3_-) holds true then, after applying this procedure to $-\mathbf{x}$ and $-\mathbf{y}$, we arrive at (5_-). Therefore we can assume without loss of generality that $\mathbf{x}$ and $\mathbf{y}$ satisfy (3.3) and one of the conditions ($4_\pm$)–($6_\pm$).

Let (4_+) be fulfilled. If $\mathbf{y}_+ = \emptyset$ then we simply remove all the entries $x_j^+ \geq x^+$ and arrive at (6_+). Otherwise we choose a subsequence $\mathbf{x}_\star$ of the sequence $\mathbf{x}$ in such a way that $\mathbf{x}_+ \subset \mathbf{x}_\star$ and x^+ is an accumulation point of both sequences $\mathbf{x}_\star$ and $\mathbf{x} \setminus \mathbf{x}_\star$. Lemma 3.7 implies that $\mathbf{y}_+ \in US^r_{\mathbf{x}_\star}$. Removing the subsequences $\mathbf{y}^+$, $\mathbf{x}_\star$ and all remaining entries $x_j > x^+$, we arrive at (6_+). If (4_-) is fulfilled then, in a similar manner, we can remove the entries $x_j \in (-\infty, x^-)$ and the entries $y_i \in (-\infty, x^-]$ so that (6_-) holds true.

Finally, under conditions ($5_\pm$) or ($6_\pm$) the inclusion $\mathbf{y} \in US^r_{\mathbf{x}}$ follows from Lemma 3.8. □

Theorem 3.10 implies, in particular, that the set $US^r_{\mathbf{x}}$ is convex. Note that the set of matrices $\mathcal{US}^r$ is not convex even in the finite dimensional case (see Example 4.3). Since the set $Q_{\mathbf{x}}$ is $\mathfrak{T}_0$-closed, Theorem 3.10 also implies that

$$\overline{\operatorname{conv}}\, P^r_{\mathbf{x}} \subseteq \overline{S^r_{\mathbf{x}}} \subseteq Q_{\mathbf{x}}, \qquad \forall \mathbf{x} \in \mathbb{R}^\infty, \tag{3.7}$$

where the closure is taken in any topology which is finer than $\mathfrak{T}_0$.

Corollary 3.11. *Let $\mathfrak{T}$ be an arbitrary topology on $X_{Q_{\mathbf{x}}}$, which is finer than $\mathfrak{T}_0$ and coarser than the Mackey topology $\mathfrak{T}_{\mathrm{m}}(X_{Q_{\mathbf{x}}}, X'_{Q_{\mathbf{x}}})$. Then*

$$\overline{\operatorname{conv}}\, P^r_{\mathbf{x}} = \overline{S^r_{\mathbf{x}}} = Q_{\mathbf{x}}, \qquad \forall \mathbf{x} \in \mathbb{R}^\infty, \tag{3.8}$$

where $\overline{\operatorname{conv}}\, P^r_{\mathbf{x}}$ and $\overline{S^r_{\mathbf{x}}}$ are the sequential $\mathfrak{T}$-closures of the sets $\operatorname{conv} P^r_{\mathbf{x}}$ and $S^r_{\mathbf{x}}$ respectively.

Proof. In view of (3.7), it is sufficient to prove (3.8) for $\mathfrak{T} = \mathfrak{T}_{\mathrm{m}}(X_{Q_{\mathbf{x}}}, X'_{Q_{\mathbf{x}}})$. In the rest of the prove $\bar{\Omega}$ denotes the sequential $\mathfrak{T}_{\mathrm{m}}(X_{Q_{\mathbf{x}}}, X'_{Q_{\mathbf{x}}})$-closure of the set $\Omega \in X_{Q_{\mathbf{x}}}$ and $\overline{\operatorname{conv}}\,\Omega$ is the sequential $\mathfrak{T}_{\mathrm{m}}(X_{Q_{\mathbf{x}}}, X'_{Q_{\mathbf{x}}})$-closure of its convex hull.

Let $\mathcal{V}_{\mathcal{S},\mathbf{x}} := \mathcal{V}_{\mathcal{S}} \bigcap \mathcal{V}_{\mathbf{x}}$, where $\mathcal{V}_{\mathcal{S}}$ is the subspace introduced in Definition 2.8. By Lemma 3.1, we have $\mathbf{w}\mathbf{x} \in X'_{Q_{\mathbf{x}}}$ for all $\mathbf{w} \in \mathcal{V}_{\mathcal{S},\mathbf{x}}$ and, consequently, $\mathbf{x}' \otimes \mathbf{x} \in \mathcal{V}'_{\mathcal{S},\mathbf{x}}$ for all $\mathbf{x}' \in X'_{Q_{\mathbf{x}}}$. If $\mathbf{x} \notin l^\infty$ then $X_{Q_{\mathbf{x}}} = \mathbb{R}^\infty$ and $\mathfrak{T}_{\mathrm{m}}(X_{Q_{\mathbf{x}}}, X'_{Q_{\mathbf{x}}}) = \mathfrak{T}_0$ is a metrizable topology. If $\mathbf{x} \in l^\infty$ then $\mathcal{V}_{\mathcal{S},\mathbf{x}} = \mathcal{V}_{\mathcal{S}}$. Therefore Theorem 2.15 and Remark 2.18 imply that $S^r_{\mathbf{x}} \subseteq \overline{\operatorname{conv}}\, P^r_{\mathbf{x}} = \overline{S^r_{\mathbf{x}}}$.

Note that

- (*) for each $\varepsilon > 0$ there exists $\mathbf{x}_\varepsilon \in P^r_{\mathbf{x}}$ such that $\hat{\mathbf{x}} - \mathbf{x}_\varepsilon \in l^1$ and $\|\hat{\mathbf{x}} - \mathbf{x}_\varepsilon\|_{l^1} < \varepsilon$.

Indeed, if $x^+ < +\infty$ then we can always find a subsequence $\{x_{j_k}\}_{k=1,2,\dots}$ of $\mathbf{x}$ such that the l^1-norm of the sequence $\{x^+ - x_{j_k}\}_{k=1,2,\dots}$ is smaller than $\varepsilon/6$. Similarly, if $x^- > -\infty$ then there exists a subsequence $\{x_{i_n}\}_{n=1,2,\dots}$ such that $i_n \neq j_k$ for all k, n and the l^1-norm of the sequence $\{x^- - x_{i_n}\}_{n=1,2,\dots}$ is smaller than $\varepsilon/6$. The required sequence $\mathbf{x}_\varepsilon$ is obtained from $\hat{\mathbf{x}}$ by replacing the entries x^+ and x^- with $x_{j_{2k-1}}$ and $x_{i_{2n-1}}$ and changing the entries x_{j_k} and x_{i_n} of the sequence $\hat{\mathbf{x}}$ to $x_{j_{2k}}$ and $x_{i_{2n}}$ respectively.

In view of Remark 1.12, (*) implies that $\overline{\operatorname{conv}}\, P^r_{\hat{\mathbf{x}}} \subseteq \overline{\operatorname{conv}}\, P^r_{\mathbf{x}}$. Since $Q_{\mathbf{x}}$ is sequentially closed, applying Theorem 3.10 and taking into account the identity $\overline{\operatorname{conv}}\, P^r_{\mathbf{x}} = \overline{S^r_{\mathbf{x}}}$, we obtain $Q_{\mathbf{x}} = S^r_{\hat{\mathbf{x}}} = \overline{\operatorname{conv}}\, P^r_{\hat{\mathbf{x}}} \subseteq \overline{\operatorname{conv}}\, P^r_{\mathbf{x}} = \overline{S^r_{\mathbf{x}}} \subseteq Q_{\mathbf{x}}$. □

Remark 3.12. If $\mathbf{x}' \in X'$ contains a subsequence which converges to zero and $\tilde{\mathbf{x}} \in P^r_{\mathbf{x}}$ then one can find $\mathbf{x}_n \in P_{\mathbf{x}}$ such that $\langle \tilde{\mathbf{x}} - \mathbf{x}_n, \mathbf{x}' \rangle \to 0$ as $n \to \infty$. This observation and the separation theorem immediately imply that, under the conditions of Corollary 3.11,

1. $\overline{\operatorname{conv}}\, P_{\mathbf{x}} = Q_{\mathbf{x}}$ whenever $\mathbf{x} \notin l^1$,
2. $\overline{\operatorname{conv}}\, P_{\mathbf{x}} = Q^\star_{\mathbf{x}} := \{\mathbf{y} \in Q_{\mathbf{x}} : y_1 + y_2 + \ldots = x_1 + x_2 + \ldots\}$ whenever $\mathbf{x} \in l^1$ and $\mathfrak{T}$ is the l^1-topology (indeed, if $\mathbf{x}' \in l^\infty$ separates $P_{\mathbf{x}}$ and $\mathbf{x}^\star \in Q^\star_{\mathbf{x}}$ and c' is an accumulation point of the sequence $\mathbf{x}'$ then, by the above, $\tilde{\mathbf{x}}' := \{x'_1 - c', x'_2 - c', \ldots\}$ separates $P^r_{\mathbf{x}}$ and $\mathbf{x}^\star$, which contradicts to Corollary 3.11).

The latter result is well known (see, for example, [Ma1], Theorem 4.2), the former was proved in [Ma1] for the topology $\mathfrak{T}$ generated by a symmetric norm which satisfies (1.3).

Remark 3.13. By Corollary 3.11, (3.8) holds true in the Mackey topology $\mathfrak{T}_{\mathrm{m}}(l^\infty, l^1)$ whenever $\mathbf{x} \in l^\infty$. If, in addition, $x_j \to c \neq 0$ as $j \to \infty$ then, applying Corollary 3.11 to the sequence $\tilde{\mathbf{x}} := \{x_1 - c, x_2 - c, \ldots\}$, one can show that (3.8) remains valid with respect to a stronger topology.

3.3. **Extreme points.** Theorem 2.4 suggests that $\operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x}} \subset P^{\mathbf{G}_1}_{\mathbf{x}}$. In the next theorem we prove this inclusion only under some additional conditions.

Definition 3.14. Denote $S^{\mathbf{G}_1}_{\mathbf{x},(m)} := \{\mathbf{y}^{(m)} \in \mathbb{R}^m : \mathbf{y} \in S^{\mathbf{G}_1}_{\mathbf{x}}\}$, $P^{\mathbf{G}_1}_{\mathbf{x},(m)} := \{\mathbf{y}^{(m)} \in \mathbb{R}^m : \mathbf{y} \in P^{\mathbf{G}_1}_{\mathbf{x}}\}$, $S^r_{\mathbf{x},(m)} := S^{\mathbf{G}_r}_{\mathbf{x},(m)}$ and $P^r_{\mathbf{x},(m)} := P^{\mathbf{G}_r}_{\mathbf{x},(m)}$, where $\mathbf{y}^{(m)}$ is defined as in (1.1).

Clearly, $S^{\mathbf{G}_1}_{\mathbf{x},(\infty)} = S^{\mathbf{G}_1}_{\mathbf{x}}$ and $P^{\mathbf{G}_1}_{\mathbf{x},(\infty)} = P^{\mathbf{G}_1}_{\mathbf{x}}$.

Theorem 3.15. *Let $\mathbf{x} = \{x_1, x_2, \ldots\} \in \mathbb{R}^\infty$ and $\mathbf{G}_1$ be a set of rows and columns. Assume that at least one of the following conditions is fulfilled:*

1. *$m < \infty$,*
2. *$m = \infty$ and either $\mathbf{G}_1 \subseteq \mathbf{G}_r$ or $\mathbf{G}_1$ contains all columns,*
3. *$m = \infty$ and $x_j \neq x_k$ whenever $j \neq k$.*

Then $\operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x},(m)} \subset P^{\mathbf{G}_1}_{\mathbf{x},(m)}$. If (3) holds true and $\mathbf{x}$ does not contain zero entries then $\mathbf{w}\mathbf{x} \notin \operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x}}$ whenever $\mathbf{w} \in \left(\mathcal{S}^{\mathbf{G}_1} \setminus \mathcal{P}^{\mathbf{G}_1}\right) \bigcap \mathcal{V}_{\mathbf{x}}$.

Proof. Let $\mathbf{w} \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}_{\mathbf{x}}$ and $(\mathbf{w}\mathbf{x})^{(m)} \in \operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x},(m)}$. The proof consists of two parts. In the first part we shall construct a special matrix $\tilde{\mathbf{w}} \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}_{\mathbf{x}}$

such that $(\tilde{\mathbf{w}}\mathbf{x})^{(m)} = (\mathbf{w}\mathbf{x})^{(m)}$. Then we shall show that $(\tilde{\mathbf{w}}\mathbf{x})^{(m)} = (\mathbf{w}_0\mathbf{x})^{(m)}$ with some $\mathbf{w}_0 \in \mathcal{P}^{\mathbf{G}_1}$ and that $\mathbf{w} = \tilde{\mathbf{w}} \in \mathcal{P}^{\mathbf{G}_1}$ whenever (3) holds true and $\mathbf{x}$ does not have zero entries.

Let Λ be the countable set of all distinct values taken by the entries of $\mathbf{x}$, $\mathbf{J}_\lambda = \{j_1, j_2, \ldots\}$ be the ordered set of all indices $j_1 < j_2 < \ldots$ such that $x_{j_k} = \lambda$ and $v_{i,\lambda} := \sum_{j_k \in \mathbf{J}_\lambda} w_{ij_k}$.

If (3) is fulfilled then we take $\tilde{\mathbf{w}} := \mathbf{w}$. Otherwise, given $m \le +\infty$ and an ordered set $\mathbf{J}_\lambda = \{j_1, j_2, \ldots\}$, we define

(1) $\tilde{w}_{1j_1}^{(m;\lambda)} := v_{1,\lambda}$ and $\tilde{w}_{1j_k}^{(m;\lambda)} := 0$ for all $k > 1$;
(2) if $1 < i \le m$ and $j_l \in \mathbf{J}_\lambda$ is the maximal positive integer such that $\tilde{w}_{(i-1)j_l} > 0$, then
$\tilde{w}_{ij_k}^{(m;\lambda)} := 0$ for all $k < l$ and $k > l+1$,
$\tilde{w}_{ij_l}^{(m;\lambda)} := \min\{\, v_{i,\lambda}\,, 1 - \sum_{n=1}^{i-1} \tilde{w}_{nj_l}^{(m;\lambda)}\,\}$ and
$\tilde{w}_{ij_{l+1}}^{(m;\lambda)} := v_{i,\lambda} - w_{ij_l}^{(m;\lambda)}$.

Let $\tilde{\mathbf{w}}^{(m)}$ be the $m \times \infty$-matrix whose entries $\tilde{w}_{ij}^{(m)}$ coincide with $\tilde{w}_{ij_k}^{(m;\lambda)}$ for all $j = j_k \in \mathbf{J}_\lambda$. Obviously, we have $\tilde{\mathbf{w}}^{(m)} \in \mathcal{V}_{\mathbf{x}}$ and $\tilde{\mathbf{w}}^{(m)}\mathbf{x} = (\mathbf{w}\mathbf{x})^{(m)}$. For each $\lambda \in \Lambda$ the matrix $\tilde{\mathbf{w}}^{(m)}$ has at most two nonzero entries lying at the intersections of a given ith row and the $\mathbf{J}_\lambda$-columns. The minimal column-number of such a nonzero entry in the $(i+1)$th row is not smaller than the maximal column-number of a nonzero entry in the ith row; in other words, the set of nonzero entries lying in the $\mathbf{J}_\lambda$-columns is ladder-shaped.

If (2) is fulfilled then we take $\tilde{\mathbf{w}} := \tilde{\mathbf{w}}^{(m)}$. The matrix $\tilde{\mathbf{w}}^{(m)}$ has the same row-sums as $\mathbf{w}$ and its column-sums $u_j := \sum_{i=1}^m \tilde{w}_{ij}^{(m)}$ are not greater than 1. If all column-sums of $\mathbf{w}$ are equal to 1 then each column-sum of $\tilde{\mathbf{w}}^{(m)}$ is also equal to 1. Therefore $\tilde{\mathbf{w}} \in \mathcal{S}^{\mathbf{G}_1}$.

If (1) is fulfilled, let us consider the ordered set $\mathbf{J} = \{j_1, j_2, \ldots\}$ of all indices $j_1 < j_2 < \ldots$ such that $u_{j_k} < 1$. Note that every set $\mathbf{J}_\lambda$ contains at most one element of $\mathbf{J}$. Let $\tilde{\mathbf{w}} = \{\tilde{w}_{ij}\}$ be the $\infty \times \infty$-extension of the $m \times \infty$-matrix $\tilde{\mathbf{w}}^{(m)}$ defined as follows:

(1) $\tilde{w}_{ij} = 0$ for all $j \notin \mathbf{J}$ and $i > m$;
(2) $\tilde{w}_{(m+1)j_1} := 1 - u_{j_1}$ and $\tilde{w}_{(m+1)j_k} := \min\{1 - u_{j_k}\,,\, 1 - \sum_{n=1}^{k-1} \tilde{w}_{(m+1)j_n}\}$ for all $j_k \in \mathbf{J}$ with $k = 2, 3, \ldots$;
(3) if $i > m+1$ and $j_l \in \mathbf{J}$ is the maximal positive integer such that $\tilde{w}_{(i-1)j_l} > 0$ then
$\tilde{w}_{ij_k} := 0$ for all $j_k \in \mathbf{J}$ with $k < l$,
$\tilde{w}_{ij_l} := 1 - u_{j_l} - \tilde{w}_{(i-1)j_l}$ and
$\tilde{w}_{ij_k} := \min\{1 - u_{j_k}\,,\, 1 - \sum_{n=1}^{k-1} \tilde{w}_{ij_n}\}$ for all $j_k \in \mathbf{J}$ with $k > l$.

We have $\sum_{j \in \mathbf{J}} u_j \le m$ and, consequently, $\sum_{j \in \mathbf{J}} (1 - u_j) = +\infty$. Therefore, for each $i > m$, the set of nonzero entries in the ith row of the matrix $\tilde{\mathbf{w}}$ is finite and non-empty. Since $\tilde{\mathbf{w}}^{(m)} \in \mathcal{V}_{\mathbf{x}}$, this implies that $\tilde{\mathbf{w}} \in \mathcal{V}_{\mathbf{x}}$. All column-sums of the matrix $\tilde{\mathbf{w}}$ are equal to 1. Its ith row-sum coincides with the ith row sum of $\mathbf{w}$ whenever $i \le m$ and is equal to 1 whenever $i > m$. Therefore $\tilde{\mathbf{w}} \in \mathcal{S}^{\mathbf{G}_1}$. The set of nonzero entries $\tilde{w}_{ij}$ with $i > m$ is also ladder-shaped. More precisely, the jth column contains at most two such nonzero entries (if it does then these entries lie in adjacent rows) and the minimal column-number of a nonzero entry in the $(i+1)$th row is not smaller than the maximal column-number of a nonzero entry in the ith row.

Let $\tilde{G}$ the subgraph of G, which contains all the vertices g_{ij} (that is, the intersections of ith rows and jth columns) such that $\tilde{w}(g_{ij}) := \tilde{w}_{ij} \in (0,1)$. Denote $\tilde{G}_\lambda := \{g_{ij} \in \tilde{G} : j \in \mathbf{J}_\lambda\}$ and $\tilde{G}' := \{g_{ij} \in \tilde{G} : i > m\}$.

Assume first that $\tilde{G}$ contains an admissible cycle $g_{i_1 j_1} \to g_{i_2 j_1} \to g_{i_2 j_2} \to \cdots \to g_{i_1 j_1}$. Replacing $g_{i_k j_k} \to g_{i_{k+1} j_k} \to g_{i_{k+1} j_{k+1}} \to \cdots \to g_{i_{k+l} j_{k+l}}$ with $g_{i_k j_k} \to g_{i_{k+l} j_{k+l}}$ whenever $i_k = i_{k+l}$, we obtain an admissible cycle $g_{i'_1 j'_1} \to g_{i'_2 j'_1} \to g_{i'_2 j'_2} \to \cdots \to g_{i'_1 j'_1}$ which has at most two vertices in every row. By our construction, the subgraphs $\tilde{G}_\lambda$ and $\tilde{G}'$ are ladder-shaped and, for every $\lambda \in \Lambda$, the intersection $\tilde{G}_\lambda \bigcap \tilde{G}'$ contains at most one element. Therefore this admissible cycle has at least two vertices lying in the same ith row with $i \le m$ but in distinct sets $\tilde{G}_\lambda$. If $\mathbf{w}_\varepsilon^\pm$ are defined as in the part (**2**) of the proof of Theorem 2.4 then $\mathbf{w}_\varepsilon^\pm \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}_\mathbf{x}$, $\tilde{\mathbf{w}} = \frac{1}{2}(\mathbf{w}_\varepsilon^+ + \mathbf{w}_\varepsilon^-)$ and $(\mathbf{w}_\varepsilon^+ \mathbf{x})^{(m)} \ne (\mathbf{w}_\varepsilon^- \mathbf{x})^{(m)}$. Therefore $(\tilde{\mathbf{w}}\mathbf{x})^{(m)} \notin \operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x},(m)}$.

Thus, the graph $\tilde{G}$ does not have any admissible cycles. Let us take an arbitrary vertex $g_{ij_0} = g_0 \in \tilde{G}$ with $i \le m$, define $\mathbf{w}_\varepsilon^\pm$ as in the part (**3**) of the proof of Theorem 2.4 and denote $\mathbf{w}^* := \frac{1}{2}(\mathbf{w}_\varepsilon^+ - \mathbf{w}_\varepsilon^-)$. Then $\mathbf{w}_\varepsilon^\pm \in \mathcal{S}^{\mathbf{G}_1} \bigcap \mathcal{V}_\mathbf{x}$, $\tilde{\mathbf{w}} = \frac{1}{2}(\mathbf{w}_\varepsilon^+ + \mathbf{w}_\varepsilon^-)$ and

$$w^\star_{ij_0} = \pm\varepsilon\, \tilde{w}_{ij_0}\,, \qquad w^\star_{ij} = \mp\varepsilon\, \tilde{w}_{ij_0}\,(1-\tilde{w}_{ij_0})^{-1}\,\tilde{w}_{ij}\,, \quad \forall j \ne j_0\,.$$

Since $(\tilde{\mathbf{w}}\mathbf{x})^{(m)} \in \operatorname{ex} S^{\mathbf{G}_1}_{\mathbf{x},(m)}$, we have $\mathbf{w}^\star \mathbf{x} = 0$ which implies that $\tilde{w}_{ij_0} x_{j_0} = \tilde{w}_{ij_0}(1-\tilde{w}_{ij_0})^{-1} \sum_{j \ne j_0} \tilde{w}_{ij} x_j$ and $x_{j_0} = \sum_{j=1}^\infty \tilde{w}_{ij} x_j$. The integer j_0 can be chosen in an arbitrary way. Therefore for each $i \le m$ we have either $\tilde{w}_{ij} = 0$ or $j \in \mathbf{J}_{\lambda_i}$, where $\lambda_i := \sum_{j=1}^\infty \tilde{w}_{ij} x_j \in \Lambda$. The first row of $\tilde{\mathbf{w}}$ may contain only one nonzero entry w_{1j} with $j \in \mathbf{J}_{\lambda_1}$. If it does then $x_j = w_{1j} x_j$ and, consequently, either $w_{1j} = 1$ or $x_j = \lambda_1 = 0$. By induction in i, the same is true for all $i = 1, 2, \dots, m$; namely, each of the first m rows either contains one entry 1 in a $\mathbf{J}_{\lambda_i}$-column corresponding to some $\lambda_i \ne 0$ or has nonzero entries only in the $\mathbf{J}_0$-columns. If $\mathbf{J}_0 = \emptyset$, this implies that $\tilde{\mathbf{w}} \in \mathcal{P}^{\mathbf{G}_1}$. If $\mathbf{J}_0 \ne \emptyset$ then, by Remark 2.16, there exists a matrix $\mathbf{w}_0 \in \mathcal{P}^{\mathbf{G}_1}$ whose entries at the intersections of the first m rows and the $\mathbf{J}_{\lambda_i}$-columns with $\lambda_i \ne 0$ coincide with the corresponding entries of $\tilde{\mathbf{w}}$. Since $(\mathbf{w}\mathbf{x})^{(m)} = (\tilde{\mathbf{w}}\mathbf{x})^{(m)} = (\mathbf{w}_0\mathbf{x})^{(m)}$, this completes the proof. □

4. Applications to spectral theory

4.1. **Notation and definitions.** Let H be a separable complex Hilbert space with the inner product $(\cdot,\cdot)_H$ and norm $\|\cdot\|_H$. For the sake of definiteness, we shall be assuming that $\dim H = \infty$; the finite dimensional versions of our results are either well known or can be proved in a similar manner.

Consider a linear operator A in H and denote by $Q_A[\cdot]$ its quadratic form defined on the domain $\mathcal{D}(Q_A) := \mathcal{D}(|A|^{1/2})$. We shall always be assuming that the operator A is self-adjoint. Let $\sigma(A)$, $\sigma_{\rm c}(A)$, and $\sigma_{\rm ess}(A)$ be its spectrum, continuous spectrum and essential spectrum respectively and let $\sigma_{\rm p}(A) = \{\lambda_1, \lambda_2, \dots\}$ be the set of its eigenvalues. As usual, we enumerate the eigenvalues λ_j taking into account their multiplicities. If $\Lambda \in \hat{\mathbb{R}}$ and $\mathbb{R} \bigcap \Lambda$ is a Borel set, we shall denote by Π_Λ and A_Λ the spectral projection of A corresponding to $\mathbb{R} \bigcap \Lambda$ and the restriction of A to the subspace $\Pi_\Lambda H$ respectively.

Definition 4.1. Let $\hat\sigma^{\pm}_{\rm ess}(A)$ and $\hat\sigma_{\rm ess}(A)$ be the subsets of $\hat{\mathbb{R}}$ such that

$\lambda \in \hat\sigma^{+}_{\rm ess}(A)$ if and only if $\dim \Pi_{[\lambda,\mu)} H = \infty$ for all $\mu > \lambda$,

$\lambda \in \hat\sigma^{-}_{\rm ess}(A)$ if and only if $\dim \Pi_{(\mu,\lambda]} H = \infty$ for all $\mu < \lambda$,

and $\hat\sigma_{\rm ess}(A) := \hat\sigma^{-}_{\rm ess}(A) \bigcup \hat\sigma^{+}_{\rm ess}(A)$

Obviously, $\sigma_{\rm ess}(A) = \mathbb{R} \bigcap \hat\sigma_{\rm ess}(A)$, $+\infty \notin \sigma^{+}_{\rm ess}(A)$ and $-\infty \notin \sigma^{-}_{\rm ess}(A)$. We have $\pm\infty \in \hat\sigma_{\rm ess}(A)$ if and only if $\pm A$ is not bounded from above.

Definition 4.2. If m is a positive integer or $m = \infty$, let

(1) $\sigma(m, A)$ be the set of vectors $\mathbf{x} = (x_1, x_2, \dots) \in \mathbb{R}^m$ such that $x_j \in \sigma(A)$ for each j and the number of entries $x_j = \lambda \notin \sigma_{\rm ess}(A)$ does not exceed the multiplicity of the eigenvalue λ;

(2) $\sigma_{\rm p}(m, A)$ be the set of vectors $\mathbf{x} = (x_1, x_2, \dots) \in \mathbb{R}^m$ such that $x_j \in \sigma_{\rm p}(A)$ for each j and the number of entries $x_j = \lambda$ does not exceed the multiplicity of the eigenvalue λ.

If $\mathbf{u} = \{u_1, u_2, \dots\}$ is an orthonormal subset of $\mathcal{D}(Q_A)$ which contains m elements u_k, denote $Q_A[\mathbf{u}] := \{Q_A[u_1], Q_A[u_2], \dots\} \in \mathbb{R}^m$ and define

(3) $\Sigma(m, A) := \{\, \mathbf{y} \in \mathbb{R}^m : \mathbf{y} = Q_A[\mathbf{u}]$ for some $\mathbf{u} \subset \mathcal{D}(Q_A) \,\}$.

The sets $\sigma(m, A)$, $\sigma_{\rm p}(m, A)$ and $\Sigma(m, A)$ will be called the *m-spectrum*, *point m-spectrum*, and *m-numerical range* of A respectively.

The m-spectra and m-numerical range are symmetric with respect to permutations of the coordinates x_k. The ∞-spectra and ∞-numerical range are subsets of $\mathbb{R}^\infty$, whose projections onto the subspace spanned by any m coordinate vectors coincide with the m-spectra and m-numerical range. In particular, $\sigma(1, A) = \sigma(A)$, $\sigma_{\rm p}(1, A) = \sigma_{\rm p}(A)$ and $\Sigma(1, A)$ is the numerical range of the operator A. Since $\sigma(A)$ is a closed set, the m-spectrum $\sigma(m, A)$ is closed in the topology of element-wise convergence and, consequently, in any finer topology.

Definition 4.2 can be extended to an arbitrary linear operator A acting in the separable Hilbert space H. In [BD], Section 36, the authors defined a matrix m-numerical range as the set of all $m \times m$-matrices of the form $\Pi A \Pi$, where Π is an orthogonal projection of rank $m < \infty$. Halmos defined an m-numerical range as the set of traces of such matrices (see [H], Chapter 17). Our definition lies in between: we consider the sets of diagonal elements of the matrices $\Pi A\Pi$ instead of their traces. Yet another concept of multidimensional numerical range, related to a given block representation of the operator A, was introduced in [LMMT]. Halmos' m-numerical range is always convex (in the self-adjoint case this immediately follows from Corollary 4.7). The m-numerical range $\Sigma(m,A)$ is convex if A is self-adjoint. The matrix m-numerical range considered in [BD] and the multidimensional numerical range introduced in [LMMT] are not necessarily convex. The latter depends on the choice of block representation and is not unitary invariant.

If $A \neq A^*$ then $\Sigma(m,A)$ does not have to be convex, even if the operator A is normal and $\dim H < \infty$. The following simple example was suggested by A. Markus [Ma2].

Example 4.3. Let $A = \{a_{ij}\}$ be the diagonal 3×3-matrix with $a_{11} = i$, $a_{22} = 1$ and $a_{33} = 0$. Then $\{i,1,0\} \in \Sigma(3,A)$ and $\{0,i,1\} \in \Sigma(3,A)$. However, the half-sum $\{\frac{i}{2}, \frac{1}{2}+\frac{i}{2}, \frac{1}{2}\}$ does not belong to $\Sigma(3,A)$. In the same way as in Lemma 3.6, one can show that $\Sigma(3,A) = \bigcup_{\mathbf{w}} \mathbf{w}\,\mathbf{z}$, where $\mathbf{z}$ is the three dimensional complex vector $\{0,i,1\}$ and the union is taken over all unistochastic 3×3-matrices $\mathbf{w}$. This implies that the set of unistochastic matrices is not convex.

In [FW] the authors proved that $\overline{\operatorname{conv}}\,\sigma_{\mathrm{p}}(m,A) = \overline{\operatorname{conv}}\,\Sigma(m,A)$ whenever A is a normal $m\times m$-matrix. There are also some results on the so-called c-numerical range of a finite matrix A, which is defined as the image of $\Sigma(m,A)$ under the map $\mathbf{x} \to \langle \mathbf{x}, c\rangle \in \mathbb{C}$ where c is a fixed m-dimensional complex vector (see [GR], [MMF], [MS]).

4.2. **Extreme points of the multidimensional numerical range.** We shall need the following simple lemma.

Lemma 4.4. *If $\sigma(A) \subset [\lambda_-, \lambda_+]$ and $\lambda_\pm \in \hat\sigma_{\mathrm{ess}}(A)$ then $\Sigma(\infty,A)$ coincides with the set of all sequences $\mathbf{z} = \{z_1, z_2, \ldots\}$ such that $z_i \in [\lambda_-,\lambda_+]$ for all i and the number of entries $z_i = \lambda_\pm$ does not exceed the multiplicity of the eigenvalue $\lambda_\pm$ (we assume that the multiplicity is zero whenever $\lambda_\pm$ is not an eigenvalue).*

Proof. The spectral theorem implies that every sequence $\mathbf{z} \in \Sigma(\infty,A)$ satisfies the above two conditions. On the other hand, if $z_1 \in [\lambda_-,\lambda_+]$ then, using the spectral theorem, one can easily find $u_1 \in \mathcal{D}(A)$ such that $\|u\|_H = 1$ and $z_1 = Q_A[u_1]$. Clearly, u_1 is an eigenvector whenever $z_1 = \lambda_-$ or $z_1 = \lambda_+$. If Π_1 is the orthogonal projection onto the annihilator of u_1 and $A_1 := \Pi_1 A \Pi_1$ then $\mathcal{D}(A_1) = \mathcal{D}(A)$ and $A - A_1$ is a finite rank operator. Since a finite rank perturbation does not change the essential spectrum, by induction in i we can construct an orthonormal set $\mathbf{u} = \{u_i\} \subset \mathcal{D}(A)$ such that $z_i = Q_A[u_i]$ for all $i = 1,2,\ldots$ $\square$

Definition 4.5. We shall say that $\mathbf{x} = \{x_1, x_2, \dots\} \in \sigma(\infty, A)$ is a *generating* sequence of the self-adjoint operator A if

1. either $\sigma_{\mathrm{c}}(A) = \emptyset$, $\mathbf{x} \subset \sigma_{\mathrm{p}}(\infty, A)$ and $\mathbf{x}$ contains all the eigenvalues λ_j of A according to their multiplicities;
2. or $\sigma_{\mathrm{c}}(A) \neq \emptyset$ and $\mathbf{x}$ can be represented as the union of three disjoint subsequences, one of which is defined as above and the other two lie in the open interval $(\inf \sigma_{\mathrm{c}}(A), \sup \sigma_{\mathrm{c}}(A))$ and converge to $\inf \sigma_{\mathrm{c}}(A)$ and $\sup \sigma_{\mathrm{c}}(A)$ respectively.

Theorem 4.6. *If $\mathbf{x}$ is a generating sequence of A then $\Sigma(m, A) = S^r_{\mathbf{x},(m)}$.*

Proof. Since $\Sigma(m, A)$ and $S^r_{\mathbf{x},(m)}$ coincide with the projections of $\Sigma(\infty, A)$ and $S^r_{\mathbf{x}}$ onto the subspace spanned by the first m coordinate vectors, it is sufficient to prove that $\Sigma(\infty, A) = S^r_{\mathbf{x}}$. If $\tilde{\mathbf{x}}$ is another generating sequence then, by Lemma 3.8, $\tilde{\mathbf{x}} \in S^r_{\mathbf{x}}$. Therefore $S^r_{\mathbf{x}}$ does not depend on the choice of generating sequence $\mathbf{x}$.

Let $\lambda^- := \inf \sigma_{\mathrm{c}}(A)$, $\lambda^+ := \sup \sigma_{\mathrm{c}}(A)$, $\Lambda := (\inf \sigma_{\mathrm{c}}(A), \sup \sigma_{\mathrm{c}}(A))$, λ_j be the eigenvalues of A lying outside Λ and $\{e_j\}$ be the orthonormal set of eigenvectors corresponding to λ_j.

Assume first that $\mathbf{y} = Q_A[\mathbf{u}]$, where $\mathbf{u} \subset \mathcal{D}(Q_A)$ is an orthonormal set. Let $d_i := \|\Pi_\Lambda u_i\|_H$ and $\{z_i\} \subset \Lambda$ be a sequence with two accumulation points $\lambda^\pm$, such that $Q_A[\Pi_\Lambda u_i] = d_i^2(\alpha_i\, z_{2i-1} + (1-\alpha_i)\, z_{2i})$ with some $\alpha_i \in [0,1]$. Then

$$Q_A[u_i] = Q_A[\Pi_\Lambda u_i] + Q_A[\Pi_{\mathbb{R}\setminus\Lambda} u_i] = d_i^2\, \alpha_i\, z_{2i-1} + d_i^2\,(1-\alpha_i)\, z_{2i} + \sum_j w_{ij}\lambda_j\,,$$

where $w_{ij} := |(u_i, e_j)_H|^2$. Since $\sum_j w_{ij} = \|\Pi_{\mathbb{R}\setminus\Lambda} u_i\|_H^2 = 1 - d_i^2$ and $\sum_i w_{ij} \le \|e_j\|_H^2 \le 1$, this implies that $\mathbf{y} \in S^r_{\mathbf{x}}$, where $\mathbf{x}$ is an arbitrary generating sequence containing all the eigenvalues λ_j and the subsequence $\{z_i\}$.

Assume now that $\mathbf{y} = \{y_1, y_2, \dots\} \in S^r_{\mathbf{x}}$ for some generating sequence $\mathbf{x}$. By Lemma 4.4, there exists an orthonormal set $\{v_n\} \subset \Pi_\Lambda H \bigcap \mathcal{D}(Q_A)$ such that $x_n = Q_A[v_n]$ for all $x_n \in \Lambda$. Let $\tilde{A}$ be the self-adjoint operator in the space H such that $\tilde{A} e_j = \lambda_j e_j$ and $\tilde{A} v_n = x_n\, v_n$ for all $x_n \in \Lambda$. In view of Lemma 3.6 and Theorem 3.10, we have $Q_{\tilde{A}}[\tilde{\mathbf{u}}] = \mathbf{y}$ for some orthonormal set $\tilde{\mathbf{u}} = \{\tilde{u}_1, \tilde{u}_2, \dots\} \subset \mathcal{D}(Q_{\tilde{A}})$. If $\tilde{d}_i := \|\Pi_\Lambda \tilde{u}_i\|_H$ and $\tilde{z}_i := \tilde{d}_i^{-2} Q_{\tilde{A}}[\Pi_\Lambda \tilde{u}_i]$ then the sequence $\{\tilde{z}_i\}$ satisfies conditions of Lemma 4.4. Therefore $\tilde{z}_i = Q_A[u'_i]$ for some orthonormal set $\{u'_i\} \subset \Pi_\Lambda H \bigcap \mathcal{D}(Q_A)$. Since $A_{\mathbb{R}\setminus\Lambda} = \tilde{A}_{\mathbb{R}\setminus\Lambda}$, the orthonormal set $\mathbf{u} := \{d_i u'_i + \Pi_{\mathbb{R}\setminus\Lambda} \tilde{u}_i\}$ satisfies $Q_A[\mathbf{u}] = \mathbf{y}$. □

Corollary 4.7. *For each $m = 1, 2, \dots, \infty$ the set $\Sigma(m, A)$ is convex and $\operatorname{ex}\Sigma(m, A) \subset \sigma_{\mathrm{p}}(m, A)$. A sequence $\mathbf{y} \in \sigma_{\mathrm{p}}(m, A)$ belongs to $\operatorname{ex}\Sigma(m, A)$ if and only if there is a (possibly, degenerate) interval $[\mu^-, \mu^+] \subset \hat{\mathbb{R}}$ such that*

1. $\sigma_{\mathrm{c}}(A) \subset [\mu^-, \mu^+]$, $\hat{\sigma}^+_{\mathrm{ess}}(A) \bigcap [-\infty, \mu^-) = \emptyset$, $\hat{\sigma}^-_{\mathrm{ess}}(A) \bigcap (\mu^+, +\infty] = \emptyset$;
2. $\mathbf{y} \bigcap (\mu^-, \mu^+) = \emptyset$ *and* $\mathbf{y}$ *contains all the eigenvalues* $\lambda_j \notin [\mu^-, \mu^+]$ *according to their multiplicities.*

Proof. Let $\mathbf{x}$ be a generating sequence. Theorems 3.15 and 4.6 imply that the set $\Sigma(m,A)=S^r_{\mathbf{x},(m)}$ is convex and $\operatorname{ex}\Sigma(m,A)=\operatorname{ex}S^r_{\mathbf{x},(m)}\subset P^r_{\mathbf{x},(m)}$.

Let $\mathbf{y}\in\sigma_{\mathrm{p}}(m,A)$ and $\mu^\pm\in\hat{\mathbb{R}}$ satisfy (1) and (2). Then $\mathbf{y}\bigcap(-\infty,\mu^-]$ either is empty or coincides with the union of disjoint nondecreasing subsequences $\mathbf{y}_n$ such that $\sup\mathbf{y}_n\le\inf\mathbf{y}_{n+1}$ and $\sup\mathbf{y}_n\notin\mathbf{y}_n$ whenever $\mathbf{y}_n$ is infinite (in the latter case A is bounded from below). Using this observation, one can easily show by induction in n that the sequence $\mathbf{y}\bigcap(-\infty,\mu^-]$ cannot be represented as a convex combination of two distinct sequences from $S^r_{\mathbf{x},(k)}$. Similarly, $\mathbf{y}\bigcap[\mu^+,+\infty)$ is not a convex combination of two distinct sequences from $S^r_{\mathbf{x},(k)}$. Therefore every sequence $\mathbf{y}\in\sigma_{\mathrm{p}}(m,A)$ satisfying the conditions of the corollary belongs to $\operatorname{ex}S^r_{\mathbf{x},(m)}$.

Assume now that $\mathbf{y}\in\operatorname{ex}S^r_{\mathbf{x},(m)}$ and denote $\sigma_{\mathbf{y}}:=\{\lambda\in\sigma(A):\lambda\notin\mathbf{y}\}$. If $\hat\sigma^-_{\mathrm{ess}}(A)\ne\emptyset$, $\hat\sigma^+_{\mathrm{ess}}(A)\ne\emptyset$, $\inf\hat\sigma^+_{\mathrm{ess}}(A)<\sup\hat\sigma^-_{\mathrm{ess}}(A)$ and

$$\mathbf{y}_\star\ :=\ \mathbf{y}\bigcap(\inf\hat\sigma^+_{\mathrm{ess}}(A),\sup\hat\sigma^-_{\mathrm{ess}}(A))\ \ne\ \emptyset$$

then $\mathbf{y}_\star$ coincides with a convex combination of two distinct sequences $\mathbf{y}^\pm_\star$ whose entries lie in the open interval $(\inf\hat\sigma^+_{\mathrm{ess}}(A),\sup\hat\sigma^-_{\mathrm{ess}}(A))$. By Lemma 3.8, we have $\mathbf{y}^\pm_\star\in S^r_{\mathbf{x}_\star,(k)}$, where k is the number of entries in $\mathbf{y}_0$ and $\mathbf{x}_\star:=\mathbf{x}\bigcap(\inf\hat\sigma^+_{\mathrm{ess}}(A),\sup\hat\sigma^-_{\mathrm{ess}}(A))$. Therefore the sequence $\mathbf{y}\in\operatorname{ex}S^r_{\mathbf{x},(m)}$ does not have entries which are greater than $\inf\hat\sigma^+_{\mathrm{ess}}(A)$ and smaller than $\sup\hat\sigma^-_{\mathrm{ess}}(A)$. In particular, $\mathbf{y}\bigcap(\inf\sigma_{\mathrm{c}}(A),\sup\sigma_{\mathrm{c}}(A))=\emptyset$. Since the number of entries $\inf\sigma_{\mathrm{c}}(A)$ and $\sup\sigma_{\mathrm{c}}(A)$ in the generating sequence $\mathbf{x}$ does not exceed the multiplicity of the corresponding eigenvalue and $\mathbf{y}\in P^r_{\mathbf{x},(m)}$, this implies that $\mathbf{y}\in\sigma_{\mathrm{p}}(m,A)$.

Let

$\mu^-=\mu^+:=\inf\hat\sigma^+_{\mathrm{ess}}(A)$ if $\sigma_{\mathbf{y}}=\emptyset$ and $\hat\sigma^-_{\mathrm{ess}}(A)=\emptyset$;
$\mu^-=\mu^+:=\sup\hat\sigma^-_{\mathrm{ess}}(A)$ if $\sigma_{\mathbf{y}}=\emptyset$ and $\hat\sigma^+_{\mathrm{ess}}(A)=\emptyset$;
$\mu^-=\mu^+:=\mu$ if $\sigma_{\mathbf{y}}=\emptyset$ and $\inf\hat\sigma^+_{\mathrm{ess}}(A)\ge\sup\hat\sigma^-_{\mathrm{ess}}(A)$, where μ is an arbitrary number from the closed interval $[\sup\hat\sigma^-_{\mathrm{ess}}(A),\inf\hat\sigma^+_{\mathrm{ess}}(A)]$;
$\mu^-:=\inf\sigma_{\mathbf{y}}$ and $\mu^+:=\sup\sigma_{\mathbf{y}}$ if $\sigma_{\mathbf{y}}\ne\emptyset$.

Obviously, in the first three cases (1) and (2) hold true. It remains to prove that $\mathbf{y}\bigcap(\mu^-,\mu^+)=\emptyset$, $\hat\sigma^+_{\mathrm{ess}}(A)\bigcap(-\infty,\mu^-)=\emptyset$ and $\hat\sigma^-_{\mathrm{ess}}(A)\bigcap(\mu^+,+\infty)=\emptyset$ in the last case.

Let $\sigma_{\mathbf{y}}\ne\emptyset$ and $\mu^\pm$ be defined as above. If $\sigma_{\mathbf{y}}$ contains two distinct entries λ and μ and $\mathbf{y}$ has an entry $y_i\in(\lambda,\mu)$ then $\mathbf{y}$ coincides with a convex combination of two distinct sequences obtained by replacing y_i with λ and μ respectively. Both these sequences belong to $P^r_{\mathbf{x},(m)}$ for some generating sequence $\mathbf{x}$. Therefore the inclusion $\mathbf{y}\in\operatorname{ex}S^r_{\mathbf{x},(m)}$ implies that $\mathbf{y}\bigcap(\mu^-,\mu^+)=\emptyset$.

If $m<\infty$ then $\hat\sigma^+_{\mathrm{ess}}(A)\bigcap(-\infty,\mu^-)=\emptyset$ as the number of eigenvalues lying below μ^- is finite. Assume that $m=\infty$ and that there exists $\hat\lambda\in\hat\sigma^+_{\mathrm{ess}}(A)$ such that $\lambda<\mu^-$. Let $\mathbf{y}^\star$ be a decreasing subsequence of $\mathbf{y}\bigcap(-\infty,\mu^-)$, which converges to λ, and $\mathbf{x}^\star \in P^r_{\mathbf{x}}$ be the sequence obtained from $\mathbf{y}^\star$ by adding an entry $\mu \in \sigma_{\mathbf{y}}$. By Lemma 3.7, the sequences $\mathbf{y}^\star_\pm$ obtained from $\mathbf{y}^\star$ by replacing an arbitrary entry $y_i \in \mathbf{y}^\star$ with $y_i - \varepsilon > \lambda$ and $y_i + \varepsilon < \mu^-$ respectively belong to $S^r_{\mathbf{x}^*}$. Therefore $\mathbf{y}_\pm \in S^r_{\mathbf{x}}$, where $\mathbf{y}_\pm$ are the sequences obtained from $\mathbf{y}$ by replacing the entry y_i with $y_i \pm \varepsilon$. Since $\mathbf{y} = \frac{1}{2}(\mathbf{y}_- + \mathbf{y}_+)$, this contradicts to the inclusion $\mathbf{y} \in \operatorname{ex} S^r_{\mathbf{x}}$.

In a similar way one can show that $\hat\sigma^-_{\rm ess}(A) \bigcap (\mu^+, +\infty) = \emptyset$. □

Remark 4.8. Let $\Lambda_{\rm e}(A) \subset \hat{\mathbb{R}}$ be the intersection of all intervals $[\mu^-, \mu^+]$ satisfying the condition (1) of Corollary 4.7. If the number of eigenvalues lying outside $\Lambda_{\rm e}(A)$ is smaller then m then, by Corollary 4.7, the set $\Sigma(m, A)$ does not have any extreme points.

Definition 4.9. If $\mathbf{x}$ is a generating sequence of A, let $Q(\infty, A) := Q_{\mathbf{x}}$, $X_A := X_{Q_{\mathbf{x}}}$ (see Definition 3.5), $Q(m, A)$ be the projection of $Q(\infty, A)$ on the subspace of X_A spanned by the first m coordinate vectors and $\mathfrak{T}^{(m)}_A$ be the topology on $Q^{(m)}_A$ induced by $\mathfrak{T}_{\rm m}(X_A, X'_A)$.

Obviously, the symmetric perfect space X_A and its subset $Q(\infty, A)$ do not depend on the choice of generating sequence $\mathbf{x}$. By Theorems 3.10 and 4.6, we have

$$S^r_{\mathbf{x}} \subseteq \Sigma(\infty, A) \subseteq Q(\infty, A) \tag{4.1}$$

for each generating sequence $\mathbf{x}$.

Lemma 4.10. *For every $\mathbf{x} \in \sigma(\infty, A)$ and every sequence of strictly positive numbers ε_k there exists $\mathbf{y} \in \Sigma(\infty, A)$ such that $|y_k - x_k| \le \varepsilon_k$. For every $\mathbf{y} \in \Sigma(\infty, A)$ there exists a sequence of vectors $\mathbf{y}_n \in \operatorname{conv} \sigma(\infty, A)$ which converges to $\mathbf{y}$ in the Mackey topology $\mathfrak{T}_{\rm m}(X_A, X'_A)$.*

Proof. Let $\Lambda_k := (x_k - \varepsilon_k, x_k + \varepsilon_k)$. If $x_j \in \sigma_{\rm ess}(A)$ then $\dim P_{\Lambda_j} H = \infty$. Since a finite dimensional perturbation does not change the essential spectrum, by induction in k one can find an orthonormal sequence $\{u_1, u_2, \ldots\}$ such that $Au_k = x_k u_k$ whenever $x_k \notin \sigma_{\rm ess}(A)$ and $u_k \in P_{\Lambda_j} H$ otherwise. If $y_k = Q_A[u_k]$ then $\mathbf{y} \in \Sigma(\infty, A)$ and $|y_k - x_k| \le \varepsilon_k$.

The second statement of the lemma follows from Corollary 3.11 and the second inclusion (4.1). □

Lemma 4.10 immediately implies that

$$\overline{\operatorname{conv}}\, \sigma(m, A) = \overline{\Sigma(m, A)} = Q(m, A), \qquad \forall m = 1, 2, \ldots, \infty, \tag{4.2}$$

where the bar denotes the sequential closure taken in any topology which is finer than the topology of element-wise convergence $\mathfrak{T}_0$ and coarser than $\mathfrak{T}^{(m)}_A$. Since $\mathfrak{T}_0$ is a metrizable topology, (4.2) remains valid if we take the usual closure.

Corollary 4.11. *For each* $m = 1, 2, \ldots, \infty$ *the set* $Q(m, A)$ *is convex and* $\operatorname{ex} Q(m, A) \subset \sigma(m, A)$. *A sequence* $\mathbf{y} \in \sigma(m, A)$ *belongs to* $\operatorname{ex} Q(m, A)$ *if and only if there is a (possibly, degenerate) interval* $[\mu^-, \mu^+] \subset \hat{\mathbb{R}}$ *such that*

(1) $\hat{\sigma}_{\mathrm{ess}}(A) \subset [\mu^-, \mu^+]$,

(2) $\mathbf{y} \bigcap (\mu^-, \mu^+) = \emptyset$ *and* $\mathbf{y}$ *contains all the eigenvalues* $\lambda_j \notin [\mu^-, \mu^+]$ *according to their multiplicities.*

Proof. Let $\mathbf{x}$ be a generating sequence, $x^+ = \limsup \mathbf{x}$ and $x^- = \liminf \mathbf{x}$. In view of (3.4), (3.8) and (4.2), we have $Q(m, A) = S^r_{\hat{\mathbf{x}},(m)}$. Therefore the corollary is obtained by applying Corollary 4.7 to the operator $A \oplus A_+ \oplus A_-$ acting in the orthogonal sum $H \oplus H_+ \oplus H_-$, where $A_\pm$ is multiplication by $x^\pm$ in $H_\pm$, $\dim H_\pm = \infty$ whenever $|x^\pm| < \infty$ and $H_\pm = \emptyset$ otherwise. □

Remark 4.12. By Corollary 4.11, each sequence $\mathbf{y} \in \operatorname{ex} Q(m, A)$ consists of eigenvalues $\lambda_j \notin \overline{\operatorname{conv}}\, \sigma_{\mathrm{ess}}(A)$ and, possibly, a collection of entries $\inf \sigma_{\mathrm{ess}}(A)$ and $\sup \sigma_{\mathrm{ess}}(A)$. All these eigenvalues can be found with the use of the Rayleigh–Ritz variational formula. The interval $\Lambda_{\mathrm{e}}(A)$ defined in Remark 4.8 is a subset of $\hat{\sigma}_{\mathrm{ess}}(A)$ and may be strictly smaller. Therefore a sequence $\mathbf{y} \in \operatorname{ex} \Sigma(\infty, A)$ may contain eigenvalues lying inside $\overline{\operatorname{conv}}\, \sigma_{\mathrm{ess}}(A)$.

Example 4.13. Assume that the continuous spectrum of A is empty and that the eigenvalues of A form a sequence $\mathbf{x}$ which has two accumulation points $\lambda^\pm$ such that $\lambda^+ > \lambda_-$. Then $\overline{\operatorname{conv}}\, \sigma_{\mathrm{ess}}(A) = [\lambda^-, \lambda^+]$. However, if λ^- or λ^+ is not an accumulation point of the sequence $\mathbf{x} \bigcap [\lambda^-, \lambda^+]$ then $\Lambda_{\mathrm{e}}(A) = \emptyset$ and $\mathbf{x}$ is an extreme point of $\Sigma(\infty, A)$.

Example 4.14. If $\hat{\sigma}_{\mathrm{ess}}(A) = [-\infty, +\infty]$ then $\Sigma(\infty, A) = Q(\infty, A) = \mathbb{R}^\infty$ and $\operatorname{ex} \Sigma(\infty, A) = \emptyset$. If $\hat{\sigma}_{\mathrm{ess}}(A) = \{+\infty\}$ then $\Sigma(\infty, A) = Q(\infty, A)$ and the extreme points of $\Sigma(\infty, A)$ are the sequences formed by all the eigenvalues λ_j. If $\hat{\sigma}_{\mathrm{ess}}(A) = [\mu, +\infty]$, $\mu \in \hat{\sigma}^+_{\mathrm{ess}}(A) \bigcap \mathbb{R}$ and $\mathbf{x}$ is the sequence formed by all the eigenvalues $\lambda_j < \mu$ then every extreme point of $Q(\infty, A)$ is obtained from $\mathbf{x}$ by adding an arbitrary collection of entries μ and every extreme point of $\Sigma(\infty, A)$ is obtained from $\mathbf{x}$ by adding a collection of entries μ whose number does not exceed the multiplicity of the eigenvalue μ (we assume that the multiplicity is zero if μ is not an eigenvalue).

Remark 4.15. Let $\mathbb{R}^\infty_\lambda$ be the set of all real sequences with entries in the interval $(-\infty, \lambda]$. Theorem 4.6 implies that $\Sigma(\infty, A) \bigcap \mathbb{R}^\infty_\lambda = \Sigma(\infty, A_{(-\infty,\lambda]})$ whenever $\operatorname{rank} A_{(\lambda,+\infty)} = \infty$. This observation allows one to extend Theorem 4.6 and Corollaries 4.7, 4.11 to the sets $\Sigma(\infty, A) \bigcap \mathbb{R}^\infty_\lambda$ and $\sigma(\infty, A) \bigcap \mathbb{R}^\infty_\lambda$. Note that the linear space $X_{A_\lambda - \lambda \mathrm{I}}$ may well be smaller than X_A. In this case one can refine Lemma 4.10 and related results by considering the operator $A_\lambda - \lambda \mathrm{I}$ instead of A.

4.3. **Variational formulae and exposed points.** Recall that a function $\psi : \Omega \to \hat{\mathbb{R}}$ defined on a convex set Ω is called *quasi-concave* if

$$\psi(\alpha\, \mathbf{x} + (1-\alpha)\, \mathbf{y}) \ \geq\ \min\{\psi(\mathbf{x}), \psi(\mathbf{y})\}, \qquad \forall \mathbf{x}, \mathbf{y} \in \Omega,\ \forall \alpha \in (0,1), \tag{4.3}$$

and *strictly quasi-concave* if the left hand side of (4.3) is strictly greater than the right hand side. The function ψ is quasi-concave if and only if the sets $\{\mathbf{x} \in X : \psi(\mathbf{x}) \geq \lambda\}$ are convex for all $\lambda \in \hat{\mathbb{R}}$. The function ψ is said to be sequentially upper $\mathfrak{T}$-semicontinuous if these sets are sequentially closed in the topology $\mathfrak{T}$. The identity (4.2) and Corollary 4.7 immediately imply the following two variational results.

Corollary 4.16. *If ψ is a quasi-concave sequentially upper $\mathfrak{T}_A^{(m)}$-semicontinuous function on $Q(m,A)$ then*

$$\inf_{\mathbf{x}\in\sigma(m,A)} \psi(\mathbf{x}) \;=\; \inf_{\mathbf{x}\in\Sigma(m,A)} \psi(\mathbf{x})\,. \tag{4.4}$$

For each finite m the functions $\psi(\mathbf{x}) = x_1 + x_2 \cdots + x_m$ and $\psi(\mathbf{x}) = x_1\, x_2 \ldots x_m = \exp(\ln x_1 + \ldots \ln x_m)$ defined on the set of positive sequences are quasi-concave and $\mathfrak{T}_0$-upper semicontinuous. Therefore the variational formulae for the sum and product of the first m eigenvalues of a positive self-adjoint operator are particular cases of (4.4).

Corollary 4.17. *Let ψ be a real-valued function defined on $\Sigma(m,A)$. If*

(a) *either ψ is quasi-concave and $\psi(\mathbf{y}) < \psi(\tilde{\mathbf{y}})$ for all $\tilde{\mathbf{y}} \neq \mathbf{y}$*
(b) *or ψ is strictly quasi-concave and $\psi(\mathbf{y}) \leq \psi(\tilde{\mathbf{y}})$ for all $\tilde{\mathbf{y}}$*

then $\mathbf{y} \in \sigma_{\mathrm{p}}(m,A)$.

Note that $\mathbf{y}$ is a $\mathfrak{T}$-exposed point of the set $\Sigma(m,A)$ if and only if there exists a linear $\mathfrak{T}$-continuous function ψ satisfying the condition (a).

Example 4.18. If $m < \infty$ then $Q(m,A)$ is a closed convex polytope, $\Sigma(m,A)$ is a convex dense subset of $Q(m,A)$ and, by Corollaries 4.7 and 4.11, we have $\operatorname{ex}\Sigma(m,A) \subset \operatorname{ex}Q(m,A)$. In this case the extreme points of $\Sigma(m,A)$ and $Q(m,A)$ are exposed in the standard Euclidean topology.

The sets $\Sigma(\infty,A)$ and $Q(\infty,A)$ may contain extreme points which are not $\mathfrak{T}_{\mathrm{m}}(X_A,X'_A)$-exposed.

Example 4.19. If A is not bounded then $X_A = \mathbb{R}^\infty$ and $X'_A = \mathbb{R}^\infty_{00}$. For every $\mathbf{y} \in \mathbb{R}^\infty$ and $\mathbf{x}' \in \mathbb{R}^\infty_{00}$ there exists $\tilde{\mathbf{y}} \in P_{\mathbf{y}}$ such that $\tilde{\mathbf{y}} \neq \mathbf{y}$ and $\langle \mathbf{y},\mathbf{x}'\rangle = \langle \tilde{\mathbf{y}},\mathbf{x}'\rangle$. Therefore the sets $\Sigma(\infty,A)$ and $Q(\infty,A)$ do not contain $\mathfrak{T}_{\mathrm{m}}(X_A,X'_A)$-exposed points whenever A is unbounded.

If $\mathbf{y} \in \operatorname{ex}\Sigma(\infty,A)$ or $\mathbf{y} \in \operatorname{ex}Q(\infty,A)$, let $[\mu^-,\mu^+]$ be the interval introduced in Corollary 4.7 or 4.11 respectively, $\mathbf{y}_{(+)} := \mathbf{y}\bigcap[\mu^+,+\infty)$, $\mathbf{y}_{(-)} := \mathbf{y}\bigcap(-\infty,\mu^-]$ and $\lambda^\pm$ be defined as follows:

$\lambda^+_{\mathbf{y}} := \limsup \mathbf{y}_{(+)}$ whenever $\mathbf{y}_{(+)}$ is infinite, $\lambda^+_{\mathbf{y}} := \inf \mathbf{y}_{(+)}$ whenever $\mathbf{y}_{(+)}$ is finite and nonempty, and $\lambda^+_{\mathbf{y}} := \mu^+$ whenever $\mathbf{y}_{(+)} = \emptyset$;
$\lambda^-_{\mathbf{y}} := \liminf \mathbf{y}_{(-)}$ whenever $\mathbf{y}_{(-)}$ is infinite, $\lambda^-_{\mathbf{y}} := \sup \mathbf{y}_{(-)}$ whenever $\mathbf{y}_{(-)}$ is finite and nonempty, and $\lambda^-_{\mathbf{y}} := \mu^-$ whenever $\mathbf{y}_{(-)} = \emptyset$.

If $\lambda_{\mathbf{y}}^{-} < \lambda_{\mathbf{y}}^{+}$, denote by $\Lambda_{\mathbf{y}}$ the interval with end points $\lambda_{\mathbf{y}}^{-}$ and $\lambda_{\mathbf{y}}^{+}$ such that $\lambda_{\mathbf{y}}^{\pm} \in \Lambda_{\mathbf{y}}$ if and only if $\lambda_{\mathbf{y}}^{\pm}$ is an accumulation point of the sequence obtained from $\mathbf{y}$ by removing all the entries $y_j \in [\lambda_{\mathbf{y}}^{-}, \lambda_{\mathbf{y}}^{+}]$. If $\lambda_{\mathbf{y}}^{-} = \lambda_{\mathbf{y}}^{+}$, let $\Lambda_{\mathbf{y}} := [\lambda_{\mathbf{y}}^{-}, \lambda_{\mathbf{y}}^{-}]$.

Obviously, $\sigma_{\mathrm{ess}}(A) \subset \bar{\Lambda}_{\mathbf{y}}$ and $\mathbf{y}$ contains all the eigenvalues lying outside the closure $\bar{\Lambda}_{\mathbf{y}}$ of the interval $\Lambda_{\mathbf{y}}$. The entries of $\mathbf{y}$ lying below and above $\Lambda_{\mathbf{y}}$ form a nondecreasing sequence $\mathbf{y}^{(-)}$ and a nonincreasing sequence $\mathbf{y}^{(+)}$ respectively (either of these sequences may be empty).

Theorem 4.20. *If A belongs to the trace class then every extreme point $\mathbf{y} \in \operatorname{ex} Q(\infty, A)$ or $\mathbf{y} \in \operatorname{ex} \Sigma(\infty, A)$ is $\mathfrak{T}_{\mathrm{m}}(X_A, X'_A)$-exposed. If A is bounded but does not belong to the trace class then*

> *$\mathbf{y} \in \operatorname{ex} Q(\infty, A)$ is a $\mathfrak{T}_{\mathrm{m}}(X_A, X'_A)$-exposed point of $Q(\infty, A)$ if and only if either $\mathbf{y} \bigcap \Lambda_{\mathbf{y}} = \emptyset$ or $\Lambda_{\mathbf{y}}$ consists of one point;*
> *$\mathbf{y} \in \operatorname{ex} \Sigma(\infty, A)$ is a $\mathfrak{T}_{\mathrm{m}}(X_A, X'_A)$-exposed point of $\Sigma(\infty, A)$ if and only if either $\mathbf{y} \bigcap \Lambda_{\mathbf{y}} = \emptyset$ or $\Lambda_{\mathbf{y}}$ is closed and the spectrum of the truncation $A_{\Lambda_{\mathbf{y}}}$ consists of one point.*

Proof. Assume that $\tilde{\mathbf{y}} \in Q(\infty, A)$ or $\mathbf{y} \in \operatorname{ex} \Sigma(\infty, A)$ and $\tilde{\mathbf{y}} \in \Sigma(\infty, A) \subset Q(\infty, A)$. Let $y_{j_1} \le y_{j_2} \le \dots$ be the entries of $\mathbf{y}^{(-)}$, $y_{k_1} \ge y_{k_2} \ge \dots$ be the entries of $\mathbf{y}^{(+)}$ and $y_{n_1}, y_{n_2}, \dots$ be the entries of $\mathbf{y}$ lying in $\Lambda_{\mathbf{y}}$. Consider an arbitrary sequence $\mathbf{x}' \in X'_A$ such that

$$x'_{j_1} < x'_{j_2} < \dots < 0\,, \quad x'_{k_1} > x'_{k_2} > \dots > 0 \quad \text{and} \quad x'_{n_1} = x'_{n_2} = \dots = 0\,.$$

The identity (1.6) implies that $\sum_i y_{j_i}\, x'_{j_i} \ge \sum_i \tilde{y}_{j_i}\, x'_{j_i}$ and these two sums coincide only if $\tilde{y}_{j_i} = y_{j_i}$ for all i. Similarly, $\sum_i y_{k_i}\, x'_{j_i} \ge \sum_i \tilde{y}_{k_i}\, x'_{k_i}$ and the sums coincide only if $\tilde{y}_{k_i} = y_{k_i}$ for all i. If $\Lambda_{\mathbf{y}}$ satisfies the conditions of the theorem, $\tilde{y}_{j_i} = y_{j_i}$ for all i and $\tilde{y}_{k_i} = y_{k_i}$ for all i then, in view of Theorem 4.6, we have $\tilde{\mathbf{y}} = \mathbf{y}$. Therefore the sequence $\mathbf{y}$ is $\mathfrak{T}_{\mathrm{m}}(X_A, X'_A)$-exposed.

If $\mathbf{y} \in l^1$ and

$$x'_{j_1} < x'_{j_2} < \dots < -2\,, \quad x'_{k_1} > x'_{k_2} > \dots > 2\,, \quad x'_{n_1} = x'_{n_2} = \dots = 1$$

then the same arguments show that $\langle \mathbf{y}, \mathbf{x}' \rangle > \langle \tilde{\mathbf{y}}, \mathbf{x}' \rangle$ for all $\tilde{\mathbf{y}} \in Q(\infty, A)$. This proves the first statement of the theorem.

Assume now that A does not belong to the trace class and that $\mathbf{y}$ is $\mathfrak{T}_{\mathrm{m}}(X_A, X'_A)$-exposed. Then there exists a sequence $\mathbf{x}' \in X'_A \subset \mathbb{R}_0^\infty$ such that $\langle \mathbf{y}, \mathbf{x}' \rangle > \langle \tilde{\mathbf{y}}, \mathbf{x}' \rangle$ whenever $\tilde{\mathbf{y}} \in P_{\mathbf{y}}$ and $\tilde{\mathbf{y}} \ne \mathbf{y}$. If $y_i > y_j$ but $x'_i \le x'_j$ then $\langle \mathbf{y}, \mathbf{x}' \rangle \le \langle \tilde{\mathbf{y}}, \mathbf{x}' \rangle$, where $\tilde{\mathbf{y}} \in P_{\mathbf{y}}$ is the sequence obtained from $\mathbf{y}$ by interchanging the entries y_i and y_j. Therefore

(c$_2$) $x'_i > x'_j$ whenever $y_i > y_j$.

If $\lambda_{\mathbf{y}}^{-} = \lambda_{\mathbf{y}}^{+}$ then $\Lambda_{\mathbf{y}}$ satisfies the conditions of the theorem. Assume that $\lambda_{\mathbf{y}}^{-} < \lambda_{\mathbf{y}}^{+}$. Then $\lambda_{\mathbf{y}}^{\pm}$ are accumulation points of $\mathbf{y}$ and Λ is not closed if and only if $\mathbf{y}$ contains infinitely many entries $\lambda_{\mathbf{y}}^{-}$ or $\lambda_{\mathbf{y}}^{+}$. The inclusion $\mathbf{x}' \in \mathbb{R}_0^\infty$ and (c$_2$) imply that $x_i = 0$ whenever $y_i \in \Lambda_{\mathbf{y}}$. If $\mathbf{y}$ has two distinct entries in $\Lambda_{\mathbf{y}}$ then $\langle\mathbf{y},\mathbf{x}'\rangle = \langle\tilde{\mathbf{y}},\mathbf{x}'\rangle$, where $\tilde{\mathbf{y}} \neq \mathbf{y}$ is the sequence obtained by interchanging these entries. Therefore either $\mathbf{y}\bigcap\Lambda_{\mathbf{y}} = \emptyset$ or there exists λ such that $y_i = \lambda$ whenever $y_i \in \Lambda_{\mathbf{y}}$. If $\mathbf{y}\bigcap\Lambda_{\mathbf{y}} \neq \emptyset$ and $\sigma(A_{\Lambda_{\mathbf{y}}})$ contains another point $\mu \neq \lambda$ then we can find $u \in \Pi_{[\lambda,\mu]}H$ such that $\tilde{\lambda} := Q_A[u] \neq \lambda$ and the sequence $\tilde{\mathbf{y}}$ obtained by replacing λ with $\tilde{\lambda}$ belongs to $\Sigma(\infty, A)$. Since $\langle\mathbf{y},\mathbf{x}'\rangle = \langle\tilde{\mathbf{y}},\mathbf{x}'\rangle$, we see that $\sigma(A_{\Lambda_{\mathbf{y}}}) = \{\lambda\}$ whenever $\mathbf{y}\bigcap\Lambda_{\mathbf{y}} \neq \emptyset$. Finally, if $\mathbf{y} \in Q(\infty, A)$ or $\Lambda_{\mathbf{y}}$ is not closed then $\langle\mathbf{y},\mathbf{x}'\rangle = \langle\tilde{\mathbf{y}},\mathbf{x}'\rangle$ for the sequence $\tilde{\mathbf{y}}$ obtained by replacing λ with $\lambda_{\mathbf{y}}^-$ or $\lambda_{\mathbf{y}}^+$. Therefore in either case $\mathbf{y}\bigcap\Lambda_{\mathbf{y}} = \emptyset$. □

4.4. **Family of operators.** Finally, let us consider a family of self-adjoint operators $\{A_\theta\}_{\theta\in\Theta}$ acting in H, where Θ is an arbitrary index set. The following corollary implies that

$$\sigma(\infty, A) \subset \overline{\operatorname{conv}} \bigcup_{\theta\in\Theta} \sigma(\infty, A_\theta) \tag{4.5}$$

whenever $A \in \overline{\operatorname{conv}}\,\{A_\theta\}$, provided that the closures are taken in appropriate topologies.

Corollary 4.21. *Let X be a subspace of $\mathbb{R}^\infty$ and A be a self-adjoint operator in H such that $X_A \subset X$ and $X_{A_\theta} \subset X$ for all $\theta \in \Theta$. Assume that for every orthonormal set $\mathbf{u} \subset \mathcal{D}(Q_A)$, every $\mathbf{x}' \in X'$ and every $\varepsilon > 0$ there exist an operator A_θ and an orthonormal set $\tilde{\mathbf{u}} \subset \mathcal{D}(Q_{A_\theta})$ such that $\langle Q_A[\mathbf{u}],\mathbf{x}'\rangle \le \langle Q_{A_\theta}[\tilde{\mathbf{u}}],\mathbf{x}'\rangle + \varepsilon$. Then we have* (4.5)*, where the closure is taken in the Mackey topology $\mathfrak{T}_{\mathrm{m}}(X, X')$.*

Proof. By the separation theorem, under conditions of the corollary we have $\Sigma(\infty, A) \subset \overline{\operatorname{conv}} \bigcup_{\theta\in\Theta} \Sigma(\infty, A_\theta)$. Therefore (4.5) follows from (4.2). □

In Corollary 4.21 we can always take $X = \mathbb{R}^\infty$, in which case $X' = \mathbb{R}^\infty_{00}$ and $\mathfrak{T}_{\mathrm{m}}(X, X')$ coincides with the topology of element-wise convergence $\mathfrak{T}_0$. If A and A_θ satisfy the conditions of Corollary 4.21 and are compact then we can take $X = \mathbb{R}^\infty_0$, which implies (4.5) with the closure taken in the l^∞-topology.

References

[AU] P.M. Alberti and A. Uhlmann. *Stochasticity and Partial Order*, Reidel Publishing Company, 1982.

[B] G. Birkhoff. Three observations on linear algebra. (Spanish) *Univ. Nac. Tucumán. Revista A.* **5** (1946), 147–151.

[BD] F.F. Bonsall and J. Duncan. *Numerical Ranges* II, LMS Lecture Note Series 10, Cambridge University Press, 1973.

[BP] A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, 1979.

[FW] P.A. Fillmore & J.P. Williams. *Some convexity theorems for matrices,* Glasgow Math. J. **12** (1971), 110–117.

[Gr] R. Grzaślewicz. On extreme infinite doubly stochastic matrices. *Illinois J. Math.* **31**, no. 4 (1987), 529–543.

[GR] K.E. Gustafson & D.K.M. Rao. *Numerical range: the field of values of linear operators and matrices.* Springer-Verlag, New York, 1996.

[H] P.R. Halmos. *A Hilbert Space Problem Book*, D. van Nostrand Company, 1967.

[Is] J.R. Isbell. Birkhoff's problem 111. *Proc. Amer. Math. Soc.* **6** (1955), 217–218.

[K] G. Köthe. *Topological Vector Spaces* I, Springer-Verlag, 1969.

[Ke] D.G. Kendall. On infinite doubly-stochastic matrices and Birkhoff's problem, *J. London Math. Soc.* **35** (1960), 81–84.

[Le] G. Letac. Represéntation des mesures de probabilité sur le produit de deux espaces dénombrables, de marqes données. (French) *Illinois J. Math.* **10** (1966), 497–507.

[LMMT] H. Langer, A. Markus, V. Matsaev, C. Tretter. A new concept for block operator matrices: The quadratic numerical range. *Linear Algebra Appl.* **330** (2001), 89–112.

[LT] J. Lindenstrauss and L. Tzafriri. *Classical Banach Spaces* I, Springer-Verlag, 1977.

[Ma1] A.S. Markus. Eigenvalues and singular values of the sum and product of linear operators. *Uspehi Mat. Nauk* **19** (1964), 93–123 (Russian), English transl. in *Russian Math. Surveys* **19** (1964), 91–120.

[Ma2] A.S. Markus. Private communication, 2001.

[MMF] M. Markus, B.N. Moyls & I. Filippenko. *Normality and the higher numerical range,* Canad. J. Math. **30** (1978), no. 2, 419–430.

[MS] M. Markus & M. Sandy. *Symmetry properties of higher numerical ranges,* Linear Algebra Appl. **104** (1988), 141–164.

[MO] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[Mu] H.G. Mukerjee. Supports of extremal measures with given marginals. (French) *Illinois J. Math.* **29**, no. 2 (1985), 248–260.

[RP] B.A. Rattray and J.E.L. Peck. Infinite stochastic matrices. *Trans. Roy. Soc. Canada Sect. III (3)* **49** (1955), 55–57.

[Ru] W.H. Ruckle, *Sequence Spaces.* Research Notes in Mathematics **35**, Pitman Advanced Publishing Program, 1981.

[Si] B. Simon. *Trace Ideals and Their Applications*, LMS lecture Notes Series **35**, Cambridge University Press, 1979.

Department of Mathematics, King's College, Strand, London WC2R 2LS
E-mail address: ysafarov@mth.kcl.ac.uk