Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

SOME SUBPOLYTOPES OF THE BIRKHOFF POLYTOPE*

EDUARDO MARQUES DE SÁ†

Abstract. Some special subsets of the set of uniformly tapered doubly stochastic matrices are considered. It is proved that each such subset is a convex polytope and its extreme points are determined. A minimality result for the whole set of uniformly tapered doubly stochastic matrices is also given. It is well known that if x and y are nonnegative vectors of $\mathbb{R}^n$ and x is weakly majorized by y, there exists a doubly substochastic matrix S such that $x = Sy$. A special choice for such S is exhibited, as a product of doubly stochastic and diagonal substochastic matrices of a particularly simple structure.

Key words. Doubly-stochastic matrices, Inequalities, Polytopes, Majorization.

AMS subject classifications. 15A39, 52B11.

1. Introduction. A square, nonnegative matrix with row and column sums equal to 1 is called *doubly stochastic*. There is an extensive literature on Ω_n, the set of doubly stochastic matrices of order n. The name Birkhoff polytope given to Ω_n comes from a famous theorem of G. Birkhoff [1] who showed that Ω_n is a polytope whose vertices are the $n \times n$ permutation matrices.

For any *interval* F of $\{1, \dots, n\}$, of cardinality q, *i.e.*, a set of the form $F = \{r+1, \dots, r+q\}$ (for some r, $0 \leqslant r < n$) let E_F be the $n \times n$ matrix

$$E_F := I_r \oplus J_q \oplus I_{n-r-q},$$

where J_q is the $q \times q$ matrix with all entries $= 1/q$. An *interval partition of* $\{1, \dots, n\}$, is a partition $\mathscr{P} = \{P_1, \dots, P_s\}$ of $\{1, \dots, n\}$ into disjoint, nonempty intervals P_i. For such $\mathscr{P}$, we let

$$E_{\mathscr{P}} := E_{P_1} E_{P_2} \cdots E_{P_s}. \tag{1.1}$$

The set $\mathfrak{U}_n$ of the so-called *uniformly tapered doubly stochastic matrices* was introduced in [7, 11] by means of a set of linear inequalities. Theorem 1 of [9] asserts that $\mathfrak{U}_n$ is the convex hull of all matrices $E_{\mathscr{P}}$. We shall prove that all $E_{\mathscr{P}}$ are vertices of $\mathfrak{U}_n$, and settle a minimality property of $\mathfrak{U}_n$. Note that $E_{\mathscr{P}}$ is the barycenter of the face of Ω_n consisting of all doubly stochastic matrices whose (i,j)-entry is 0 if the (i,j)-entry of $E_{\mathscr{P}}$ is 0. The facial structure of Ω_n has been thoroughly studied in [2, 3, 4, 5], however, the sub-polytopes of Ω_n we shall consider are not faces of Ω_n.

A *nested family* of intervals of $\{1, \dots, n\}$ is a set $\mathscr{F} = \{F_1, \dots, F_t\}$ of intervals of $\{1, \dots, n\}$, such that any two intervals in the family either have an empty intersection,

* Received by the editors 22 December 2005. Accepted for publication 4 January 2006. Handling Editor: João Filipe Queiró.

†Departamento de Matemática, Universidade de Coimbra, 3000 Coimbra, Portugal (emsa@mat.uc.pt). Member of the research unit Centro de Matemática da Universidade de Coimbra/Fundação para a Ciência e a Tecnologia (CMUC/FCT).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

or one of them is contained in the other. Note that, in these conditions, the matrices $E_{F_1}, \ldots, E_{F_t}$ commute. We define $\mathfrak{U}(\mathscr{F})$ as the set of all $n \times n$ matrices of the form

$$\prod_{i=1}^{t} \left[\alpha_i I + (1-\alpha_i) E_{F_i}\right], \tag{1.2}$$

where $\alpha_1, \ldots, \alpha_t$ run over $[0, 1]$, independently of each other. We shall prove that $\mathfrak{U}(\mathscr{F})$ is a subpolytope of $\mathfrak{U}_n$, and determine its vertices.

We denote by $\mathfrak{D}(n)$ the set of all $x \in \mathbb{R}^n$, such that $x_1 \geqslant \cdots \geqslant x_n$, and $\mathfrak{D}_+(n)$ is the set of all nonnegative vectors of $\mathfrak{D}(n)$. We adopt the following *majorization symbols*: for $x, y \in \mathbb{R}^n$, we write $x \preccurlyeq_w y$ whenever

$$x'_1 + \cdots + x'_k \leqslant y'_1 + \cdots + y'_k\,, \quad \text{for all } k \in \{1, \ldots, n\}, \tag{1.3}$$

where $z'_1, \ldots, z'_k$ denotes the non-increasing rearrangement of $z \in \mathbb{R}^n$; and we write $x \preccurlyeq y$ if (1.3) holds with equality for $k = n$. In [9], the reader may find the following refinement of a well-known theorem of Hardy, Littlewood and Pólya [8]: *if $x, y \in \mathfrak{D}(n)$ satisfy $x \preccurlyeq y$, there exists $R \in \mathfrak{U}_n$ such that* $x = Ry$, together with three proofs of this result. In section 2, we show that the third of these proofs, due to D.Z. Djokovic (see [9, p. 325]) may be conveniently adapted to give a little bit more than the referred refinement. Then we extend that result to the case of weak majorization.

2. Nested Families and Majorization.

PROPOSITION 2.1. *For any $\mathscr{F}$, a nested family of intervals of $\{1, \ldots, n\}$, $\mathfrak{U}(\mathscr{F})$ is a subset of $\mathfrak{U}_n$.*

Proof. Let us expand the polynomial

$$f(u_1, \ldots, u_t) := \prod_{i=1}^{t} \left[\alpha_i + (1-\alpha_i) u_i\right],$$

where the α_i are real numbers and the u_i are commutative variables, as a sum of monomials. The sum of all coefficients of f's monomials is $f(1, \ldots, 1)$, which obviously equals 1. So (1.2) is a convex combination of the products $E_{X_1} \cdots E_{X_s}$, for $0 \leqslant s \leqslant t$ and $X_1, \ldots, X_s \in \mathscr{F}$. Note that, if $X \supseteq Y$, then $E_X E_Y = E_Y E_X = E_X$. Thus we only have to consider products $E_{X_1} \cdots E_{X_s}$ for pairwise disjoint sets $X_1, \ldots, X_s$. Therefore (1.2) lies in $\mathfrak{U}_n$, and so $\mathfrak{U}(\mathscr{F}) \subseteq \mathfrak{U}_n$. □

The proof of the following theorem is essentially due to D. Djokovic [9, p. 325].

THEOREM 2.2. *Let $x, y \in \mathfrak{D}(n)$ satisfy $x \preccurlyeq y$. There exists a nested family of intervals of $\{1, \ldots, n\}$, and a matrix $R \in \mathfrak{U}(\mathscr{F})$, such that* $x = Ry$.

Proof. We consider the two cases of D.Z. Djokovic's proof [9, p. 325]. In Case 1, it is assumed there is $k < n$ such that $x_1 + \cdots + x_k = y_1 + \cdots + y_k$. By induction, there exist a nested family $\mathscr{F}'$ of intervals of $\{1, \ldots, k\}$, a nested family $\mathscr{F}''$ of intervals of $\{1, \ldots, n-k\}$, and there exist $R' \in \mathfrak{U}(\mathscr{F}')$ and $R'' \in \mathfrak{U}(\mathscr{F}'')$ such that $x = Ry$, with $R := R' \oplus R''$. Define

$$\mathscr{F} := \mathscr{F}' \cup (\mathscr{F}'' + k)\,,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

where $\mathscr{F}''+k$ is the family of all sets $\{i+k:\, i\in X\}$, for X running over $\mathscr{F}''$. Clearly, $\mathscr{F}$ is a nested family of intervals of $\{1,\dots,n\}$. On the other hand, it is also clear that $R'\oplus I_{n-k}$ and $I_k\oplus R''$ both lie in $\mathfrak{U}(\mathscr{F})$; therefore, R lies in $\mathfrak{U}(\mathscr{F})$ as well. So we are done with Case 1. In Case 2, D.Z. Djokovic proves that $x=R\big[\beta I+(1-\beta)E_{\{1,\dots,n\}}\big]\,y$, where R is obtained as in Case 1. In our situation, this means R lies in $\mathfrak{U}(\mathscr{F})$ for some nested family $\mathscr{F}$ of intervals. Note that $\mathscr{F}\cup\{\{1,\dots,n\}\}$ is also a nested family of intervals. So the theorem holds in this case as well. □

Theorem 2.2 gives us a representation of matrix R as a product of type (1.2), of t doubly stochastic matrices of simple structure, where t is the cardinality of $\mathscr{F}$. On the other hand, the only sets $F_i\in\mathscr{F}$ which are relevant in (1.2) are those having cardinality at least 2. A straightforward argument, left to the reader, shows that *any maximal nested family of intervals of* $\{1,\dots,n\}$ *has precisely* $n-1$ *elements of cardinality at least* 2. So, $n-1$ is an upper bound to the number of relevant factors in R's factorization (1.2).

It is well known [10, p. 27] that if $x,y\in\mathfrak{D}_+(n)$ satisfy $x\preccurlyeq_w y$, then $x=Sy$ for some doubly sub-stochastic matrix S. In the following theorem we give a factorization for a special choice of S, in the spirit of Theorem 2.2.

We shall use the following notation: for each $p\in\{1,\dots,n\}$, Δ_p is the $n\times n$ diagonal matrix

$$\Delta_p\;:=\;\mathrm{Diag}(\underbrace{1,1,\dots,1}_{p},0,0,\dots,0)\,.$$

THEOREM 2.3. *Let* $x\in\mathfrak{D}_+(n)$ *be a vector whose distinct coordinates are* $\chi_1>\cdots>\chi_s$. *Suppose* m_i *is the number of times* χ_i *occurs in* x. *If* $y\in\mathfrak{D}_+(n)$ *satisfies* $x\preccurlyeq_w y$, *then the following conditions hold:*

(I) *There exist real numbers* $\theta_1,\dots,\theta_s$ *in the interval* $[0,1]$, *a nested family* $\mathscr{F}$ *on* $\{1,\dots,n\}$ *and a matrix* R *in* $\mathfrak{U}(\mathscr{F})$, *such that* $x=DRy$, *where* D *is the diagonal matrix*

$$D:=\prod_{i=1}^{s}\,[\theta_i I+(1-\theta_i)\Delta_{m_1+\cdots+m_i}]\,. \tag{2.1}$$

(II) *The following entities exist: a positive integer* p, *real numbers* $\sigma_1,\dots,\sigma_p$ *in the interval* $[0,1]$, *nested families,* $\mathscr{F}_1,\dots,\mathscr{F}_p$, *of intervals of* $\{1,\dots,n\}$, *and matrices* $R_1\in\mathfrak{U}(\mathscr{F}_1),\dots,R_p\in\mathfrak{U}(\mathscr{F}_p)$, *such that* $x=[D_pR_p\cdots D_2R_2D_1R_1]\,y$, *where*

$$D_i\;:=\;\sigma_i I+(1-\sigma_i)\Delta_{n-m_s}\,,\ \textit{for } i=1,\dots,s. \tag{2.2}$$

Proof. For each $z\in\mathbb{R}^n$ let $\Sigma(z):=z_1+\cdots+z_n$. For each $t\in\mathbb{R}$ let $x(t)\in\mathfrak{D}(n)$ be the vector with i-th entry $\max\{x_i,t\}$. Clearly $x(t)\geqslant x$ for all t, with equality iff $t\leqslant x_n$. $\Sigma(x(t))$ is a continuous function, and it is strictly increasing with t, for $t\geqslant x_n$. As $x\preccurlyeq_w y$, we have $\Sigma(x)=\Sigma(x(x_n))\leqslant\Sigma(y)\leqslant\Sigma(x(y_1))$. So there is a

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

unique $\tau \geqslant x_n$ such that $\Sigma(x(\tau)) = \Sigma(y)$. We prove

$$\sum_{i=1}^{k} [y_i - x(\tau)_i] \geqslant 0\,, \tag{2.3}$$

for $k = 1, \ldots, n-1$. If $\tau \geqslant x_1$, then $x(\tau) = (\tau, \ldots, \tau)$ and (2.3) is obvious. Now assume $\tau < x_1$, and let $v := \sup\{i : x_i > \tau\}$. Note that $1 \leqslant v < n$. As $x_i(\tau) = x_i$ for $i \in \{1, \ldots, v\}$, (2.3) is true for $k \in \{1, \ldots, v\}$. So we are left with the case $v < k < n$. Clearly

$$\sum_{i=1}^{k} [y_i - x(\tau)_i] = \sum_{i=k+1}^{n} (\tau - y_i)\,. \tag{2.4}$$

On the other hand, as $x \preccurlyeq_w y$ and $(y_i - \tau)_{i=1}^n$ in non-increasing, we have

$$\begin{aligned} 0 = \Sigma(y) - \Sigma(x(\tau)) &= \sum_{i=1}^{v} (y_i - x_i) + \sum_{i=v+1}^{n} (y_i - \tau) \\ &\geqslant \sum_{i=v+1}^{n} (y_i - \tau) \geqslant \frac{n-v}{n-k} \sum_{i=k+1}^{n} (y_i - \tau)\,. \end{aligned} \tag{2.5}$$

So (2.4) is nonnegative. This proves (2.3). Therefore $x \leqslant x(\tau) \preccurlyeq y$. By Theorem 2.2 we know that

$$x(\tau) = Ry\,, \tag{2.6}$$

where $R \in \mathfrak{U}(\mathscr{F})$ for some nested family of intervals, $\mathscr{F}$. From now on we assume that x and y lie in $\mathfrak{D}_+(n)$.

Proof of (I). If $x = x(\tau)$, then (I) holds with $D := I$, i.e. with $\theta_i := 1$ for $i = 1, \ldots, s$. Now assume $x \neq x(\tau)$. Let $u := \min\{i : x_i < \tau\}$. Then define $\theta_i := 1$ for $i = 1, \ldots, u-1$, $\theta_u := \chi_u/\tau$ and $\theta_j := \chi_j/\chi_{j-1}$ for $j = u+1, \ldots, s$. We clearly have $x = Dx(\tau)$, for D as given in (2.1). So (I) holds.

Proof of (II). The proof is easy when $s = 1$, i.e. when all entries of x are equal. For, we define $p := 1$, $\sigma_1 := x_n/\tau$ if $\tau > 0$ and $\sigma_1 := 0$ if $\tau = 0$ (note that in this case $x = x(\tau)$). Then put $R_1 := R$, the matrix of (2.6). With these definitions (II) holds. We now work out the case $s \geqslant 2$. For any $z \in \mathbb{R}^n$, let $\kappa(z)$ be the smallest integer greater than $[\Sigma(z) - \Sigma(x)]/(m_s \chi_{s-1})$. In particular

$$\kappa(z) m_s \chi_{s-1} \;\geqslant\; \Sigma(z) - \Sigma(x)\,. \tag{2.7}$$

The proof goes by induction on $\kappa(y)$. Note that $\kappa(y) = \kappa(x(\tau))$. We have two cases.

Case 1: *when* $m_s\tau \geqslant \Sigma(y) - \Sigma(x)$. Define $p := 2$,

$$\sigma_1 := \frac{m_s\tau - \Sigma(y) + \Sigma(x)}{m_s\tau}\,,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

$\sigma_2 := 0$ and $R_1 := R$, the matrix of (2.6). Moreover, let D_i be as given in (2.2) and let $y' := D_1 x(\tau)$. As $\Sigma(y') = \Sigma(x(\tau)) - m_s\tau(1-\sigma_1)$, some easy computations show $\Sigma(y') = \Sigma(x)$. This identity may be written as:

$$\sum_{i=1}^{n-m_s} x(\tau)_i + m_s\tau\sigma_1 = \sum_{i=1}^{n-m_s} x_i + m_s\tau \,. \tag{2.8}$$

As $\sigma_1 \leqslant 1$, this implies, for each $k \in \{1, \ldots, m_s\}$:

$$\sum_{i=1}^{n-m_s} x(\tau)_i + k\tau\sigma_1 \geqslant \sum_{i=1}^{n-m_s} x_i + k\tau \,. \tag{2.9}$$

Taking into account that $x(\tau) \geqslant x$, (2.8)-(2.9) show that $x \preccurlyeq y'$. So, for some nested family of intervals $\mathscr{F}_2$, there exists $R_2 \in \mathfrak{U}(\mathscr{F}_2)$ such that $x = R_2 y'$. Therefore $x = [D_2R_2D_1R_1]\, y$ and (II) holds. Case 2: *when* $m_s\tau < \Sigma(y) - \Sigma(x)$. Here, we let $\sigma_1 := 0$ and D_1 be as in (2.2). The vector $y' := D_1 x(\tau)$ clearly satisfies $\Sigma(y') = \Sigma(y) - m_s\tau > \Sigma(x)$. It is now easy to show that

$$x \preccurlyeq_w y' \,. \tag{2.10}$$

On the other hand,

$$\begin{aligned} 0 < \Sigma(x(\tau)) - \Sigma(x) - m_s\tau &= \sum_{i=1}^{s} m_i \cdot \max\{0, \tau - \chi_i\} - m_s\tau \\ &\leqslant n \cdot \max\{0, \tau - \chi_{s-1}\} \,. \end{aligned}$$

Therefore $\tau > \chi_{s-1}$. Taking (2.7) into account we obtain:

$$\begin{aligned} \Sigma(y') - \Sigma(x) &= \Sigma(y) - \Sigma(x) - m_s\tau \\ &\leqslant \kappa(y)m_s\chi_{s-1} - m_s\tau < [\kappa(y) - 1]m_s\chi_{s-1} \,. \end{aligned}$$

This yields $\kappa(y') \leqslant \kappa(y) - 1$, and this, taken together with (2.10), allows us to use induction: there exist nested families of intervals, $\mathscr{F}'_1, \ldots, \mathscr{F}'_q$, matrices $R'_1 \in \mathfrak{U}(\mathscr{F}'_1), \ldots, R'_q \in \mathfrak{U}\mathscr{F}'_q$ and diagonal matrices, $D'_1, \ldots, D'_q$, of the type of (2.2), such that $x = [D'_qR'_q \cdots D'_1R'_1]\, y'$. Therefore

$$x = [D'_qR'_q \cdots D'_1R'_1D_1R]\, y$$

and the proof is done. □

Incidentally, in the course of proof, we showed the existence of a z such that $x \leqslant z \preccurlyeq y$. This is a result of [6] (see also [10, p. 123] and references therein). However, we got a little bit more: that we may choose z of the form $x(\tau)$. We point out that our inductive proof of Theorem 2.3(II) also yields an upper bound for the number, p, of factors D_iR_i, namely $p \leqslant \kappa(y) + 1$. This gives an indication on the complexity of the procedure given by the proof.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

3. Extreme Points. There exist 2^{n-1} distinct interval partitions of $\{1,\ldots,n\}$, and so this is the cardinality of the set $\{E_{\mathscr{P}}\}$ of all matrices defined in (1.1). Theorem 1 of [9] says that $\{E_{\mathscr{P}}\}$ contains the set of all extreme points of $\mathfrak{U}_n$. Our aim now is to prove that any $E_{\mathscr{P}}$ is an extreme point of $\mathfrak{U}_n$.

Lemma 3.1. *Let $w \in \mathbb{R}^n$ be a vector satisfying $w_1 > \cdots > w_n$, R an element of $\mathfrak{U}_n$ and $\mathscr{G}$ an interval partition of $\{1,\ldots,n\}$. The identity $Rw = E_{\mathscr{G}}w$ implies $R = E_{\mathscr{G}}$.*

Proof. By Theorem 1 of [9], R is a convex combination of the $E_{\mathscr{P}}$, for all partitions $\mathscr{P}$, *i.e.*, $R = \sum \lambda_{\mathscr{P}} E_{\mathscr{P}}$, for some nonnegative coefficients $\lambda_{\mathscr{P}}$ which sum up 1. As $Rw = E_{\mathscr{G}}w$,

$$E_{\mathscr{G}}w \;=\; \sum \lambda_{\mathscr{P}} E_{\mathscr{P}} w\,. \tag{3.1}$$

The second proof of Theorem 2 of [9] shows that the 2^{n-1} vectors $E_{\mathscr{P}}w$ are pairwise distinct, and are the extreme points of $\{x \in \mathfrak{D}(n) : x \preccurlyeq w\}$. Therefore (3.1) implies that all $\lambda_{\mathscr{P}}$ are 0, except $\lambda_{\mathscr{G}}$ that equals 1. Thus $R = E_{\mathscr{G}}$ as required. □

Theorem 3.2. *For any interval partition $\mathscr{G}$, $E_{\mathscr{G}}$ is an extreme point of $\mathfrak{U}_n$.*

Proof. Pick any $E_{\mathscr{G}}$ and write it as a convex combination of the $E_{\mathscr{P}}$. Then an equation like (3.1) arises. The argument under (3.1) now proves that $E_{\mathscr{G}}$ is not a convex combination of the *other* generators $E_{\mathscr{P}}$ of $\mathfrak{U}_n$. This means $E_{\mathscr{G}}$ is an extreme point of $\mathfrak{U}_n$. □

Theorem 3.3. *$\mathfrak{U}_n$ is minimal among all sets $\mathfrak{M}$ of $n \times n$ matrices satisfying the conditions: $\mathfrak{M}$ is convex, and, if $x, y \in \mathfrak{D}(n)$ satisfy $x \preccurlyeq y$, there exists $M \in \mathfrak{M}$ such that $x = My$.*

Proof. Assume $\mathfrak{M} \subseteq \mathfrak{U}_n$ satisfies the given conditions. With w as in Lemma 3.1 we have, for any interval partition $\mathscr{P}$: $E_{\mathscr{P}}w \in \mathfrak{D}(n)$ and $E_{\mathscr{P}}w \preccurlyeq w$. So $E_{\mathscr{P}}w = M_{\mathscr{P}}w$, for some $M_{\mathscr{P}} \in \mathfrak{M}$. Lemma 3.1 implies $E_{\mathscr{P}} = M_{\mathscr{P}}$, and so $E_{\mathscr{P}} \in \mathfrak{M}$. Therefore $\mathfrak{M} = \mathfrak{U}_n$. □

We now prove the convexity of the set $\mathfrak{U}(\mathscr{F})$, whose members are matrix products as (1.2), and determine the set of its extreme points.

Theorem 3.4. *Given a nested family $\mathscr{F}$ of intervals of $\{1,\ldots,n\}$, the set $\mathfrak{U}(\mathscr{F})$ is convex, and $\{E_{\mathscr{X}} : \mathscr{X} \subseteq \mathscr{F}\}$ is the set of $\mathfrak{U}(\mathscr{F})$'s extreme points.*

Proof. By Theorem 3.2 we only need to prove that $\mathfrak{U}(\mathscr{F})$ is the convex hull of the $E_{\mathscr{X}}$, for $\mathscr{X} \subseteq \mathscr{F}$. We argue by induction on $t = |\mathscr{F}|$. Let $M_1,\ldots,M_r$ be the elements of $\mathscr{F}$ which are maximal for inclusion. Without loss of generality, assume $M_1 = F_1,\ldots,M_r = F_r$. Define $\mathscr{F}_i := \{X \in \mathscr{F} : X \subseteq F_i\}$, for $i = 1,\ldots,r$. Clearly, $\mathscr{F} = \mathscr{F}_1 \cup \cdots \cup \mathscr{F}_r$, and this union is disjoint. In the first place suppose $r = 1$, that is $F_1 \supseteq [F_1 \cup \cdots \cup F_t]$. By induction, $\mathfrak{U}(\{F_2,\ldots,F_t\}) = \operatorname{conv}\{E_{\mathscr{X}} : \mathscr{X} \subseteq \{F_2,\ldots,F_t\}\}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 1-7, January 2006

ELA
http://math.technion.ac.il/iic/ela

We therefore have

$$\begin{aligned}\mathfrak{U}(\mathscr{F}) &= \bigcup_{\alpha\in[0,1]} \left[\alpha I + (1-\alpha)E_{F_1}\right]\cdot \mathfrak{U}(\{F_2,\ldots,F_t\})\\ &= \bigcup_{\alpha\in[0,1]} \left[\alpha\, \mathfrak{U}(\{F_2,\ldots,F_t\}) + (1-\alpha)E_{F_1}\right]\\ &= \operatorname{conv}\left(\{E_{F_1}\}\cup\left\{E_{\mathscr{X}} : \mathscr{X}\subseteq\{F_2,\ldots,F_t\}\right\}\right)\\ &= \operatorname{conv}\{E_{\mathscr{Y}} : \mathscr{Y}\subseteq\mathscr{F}\}.\end{aligned}$$

This settles the case $r = 1$. We now assume $r \geqslant 2$. By induction, $\mathfrak{U}(\mathscr{F}_i) = \operatorname{conv}\{E_{\mathscr{X}_i} : \mathscr{X}_i \subseteq \mathscr{F}_i\}$. The proof is finished in the following two lines:

$$\begin{aligned}\mathfrak{U}(\mathscr{F}) &= \bigoplus_{i=1}^{r}\mathfrak{U}(\mathscr{F}_i) = \bigoplus_{i=1}^{r}\operatorname{conv}\{E_{\mathscr{X}_i} : \mathscr{X}_i\subseteq\mathscr{F}_1\}\\ &= \operatorname{conv}\bigoplus_{i=1}^{r}\{E_{\mathscr{X}_i} : \mathscr{X}_i\subseteq\mathscr{F}_i\} = \operatorname{conv}\{E_{\mathscr{X}} : \mathscr{X}\subseteq\mathscr{F}\}. \qquad\square\end{aligned}$$

REFERENCES

[1] G. Birkhoff. Tres Observaciones sobre el Algebra Lineal. *Univ. Nac. de Tucumán Rev. Sér. A*, 5:147–151, 1946.
[2] R. Brualdi and P. Gibson. Convex Polyhedra of Doubly Stochastic Matrices, I. Applications of the Permanent Function. *J. Combinatorial Theory Ser. A*, 22:194–230, 1977.
[3] ———. Convex Polyhedra of Doubly Stochastic Matrices, II. Graph of Ω_n. *J. Combinatorial Theory Ser. B*, 22:175–198, 1977.
[4] ———. Convex Polyhedra of Doubly Stochastic Matrices, III. Affine and Combinatorial Properties of Ω_n. *J. Combinatorial Theory Ser. A*, 22:338–351, 1977.
[5] ———. Convex Polyhedra of Doubly Stochastic Matrices, IV. *Linear Algebra Appl.*, 15:153–172, 1976.
[6] Ky Fan. Maximum Properties and Inequalities for the Eigenvalues of Completely Continuous Operators. *Proc. Nat. Acad. Sci. USA*, 37:760–766, 1951.
[7] E.A. Guillemin. On the Analysis and Synthesis of Single-Element-Kind Networks. *IRE Trans.*, vol. CT-7:303–312, 1960.
[8] G. Hardy, F. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, London, 1952.
[9] A.F. Hoffman. A Special Class of Doubly Stochastic Matrices. *Aequationes Math.*, 2:319–326, 1969.
[10] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications*. Academic Press, New York, 1979.
[11] L. Weinberg. A Survey of Linear Graphs: Fundamentals and Applications to Network Theory. *Matrix Tensor Quart.*, 14:81–89 and 14:103–115, 1963/64.