行政院國家科學委員會補助專題研究計畫成果報告

擁有固定列和及行和的非負矩陣的強線性保持問題

Linear maps strongly preserving nonnegative matrices with given row sums and column sums

計畫類別: ☑ 個別型計畫　□ 整合型計畫
計畫編號: NSC 91−2115−M−032−008
執行期間: 91 年 8 月 1 日　至　92 年 7 月 31 日

計畫主持人: 譚必信
共同主持人:
計畫參與人員: 國科會助理　謝旻宜
國科會助理　洪劍能
國科會助理　林淑慧

成果報告類型 (依經費核定清單規定繳交): ☑ 精簡報告 □ 完整報告

本成果報告包括以下應繳交之附件:
□ 赴國外出差或研習心得報告一份
□ 赴大陸地區出差或研習心得報告一份
☑ 出席國際學術會議心得報告及發表之論文各一份
□ 國際合作研究計畫國外研究報告書一份

處理方式: 除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外, 得立即公開查詢
□ 涉及專利或其他智慧財產權, □ 一年 □ 二年後可公開查詢

執行單位: 淡江大學數學系

中　華　民　國　92 年 7 月 31 日

摘要

若 C 是某一有限維向量空間的非空集合, T 爲 $\operatorname{span} C$ 上的線性變換且 $T(C)=C$, 則稱 T 爲 C 的強線性保持者。在本計劃我們刻劃下列的對稱非負矩陣多面體的強線性保持者: $\mathbf{SDS}(n)$, n 階對稱雙隨機矩陣多面體, $\mathbf{SDsS}(n)$, n 階對稱子雙隨機矩陣多面體, 及 $\mathbf{U}(\mathbf{r})$, n 階對稱非負矩陣其列和向量爲 $\mathbf{r}$, $n \leq 3$。如預期, $\mathbf{SDS}(n)$ 及 $\mathbf{SDsS}(n)$ 的強線性保持者都是形如 $T(X)=P^tXP$, 其中 P 爲一置換矩陣。從研究 $n=3$ 時的情況, 我們猜想在一般情形 $\mathbf{U}(\mathbf{r})$ 的強線性保持者並沒有良好的表示法。

關鍵詞: 強線性保持者、對稱雙隨機矩陣、對稱子雙隨機矩陣、LOCC 圖、相鄰頂點、列和向量。

Abstract

If C is a nonempty subset of a finite-dimensional linear space and T is a linear map on $\operatorname{span} C$, then we say T is a strong linear preserver of C if $T(C) = C$. In this project we characterize the strong linear preservers of the following polytopes of symmetric nonnegative matrices: $\mathbf{SDS}(n)$, the polytope of $n \times n$ symmetric doubly stochastic matrices, $\mathbf{SDsS}(n)$, the polytope of $n \times n$ symmetric doubly substochastic matrices, and $\mathbf{U}(\mathbf{r})$, the polytope of $n \times n$ symmetric nonnegative matrices with a fixed row sums vector $\mathbf{r}$ for $n \leq 3$. We prove that strong linear preservers T of $\mathbf{SDS}(n)$ and $\mathbf{SDsS}(n)$ are of the expected form, namely, $T(X) = P^t X P$ for some $n \times n$ permutation matrix P. However, by examining the case $n = 3$, we suspect that there is no nice characterization for the strong linear preservers of $\mathbf{U}(\mathbf{r})$ for a general n.

Key words: Strong linear preserver, symmetric doubly stochastic matrix, symmetric doubly substochastic matrices, LOCC graph, neighborly extreme point, row sums vector.

1. Motivation and Aims

In 1999, Professors Chi-Kwong Li and Nam-Kiu Tsing, my academic brothers, invited me to join in their study of the strong linear preservers of the polytope of doubly stochastic matrices. After some hard work, we finally resolved the problem and obtained the following:

Theorem A. *Let T be a linear map on* $\operatorname{span}(\mathbf{DS}(n))$. *The following conditions are equivalent*:
(a) $T(\mathbf{DS}(n)) = \mathbf{DS}(n)$.
(b) $T(\mathbf{P}(n)) = \mathbf{P}(n)$.
(c) *T is given by $T(X) = PXQ$ or $T(X) = PX^tQ$ for some $P, Q \in \mathbf{P}(n)$.*

In the above, $\mathbf{DS}(n)$ and $\mathbf{P}(n)$ denote respectively the set of $n \times n$ doubly stochastic matrices and the set of $n \times n$ permutation matrices.

Subsequently, we also resolved the strong linear preservers problem for $\mathbf{DsS}(m,n)$, the polytope of $m \times n$ doubly substochastic matrices, $\mathbf{CS}(m,n)$, the polytope of $m \times n$ column stochastic matrices, and $\mathbf{CsS}(m,n)$, the polytope of $m \times n$ column substochastic matrices. The work formed the contents of the paper [3] in the reference list. It appears that [3] is the first paper in the literature that deals with the strong linear preservers of polytopes of nonnegative matrices. Later, Professor Li and his student H. Chiang [1] also characterized the strong linear preservers of $\mathbf{A}(n)$, the set of all $n \times n$ even permutation matrices. In the table below we give a summary of the results on the strong linear preserver problems as done in [1] and [3]. We will denote by P, Q, $P_1, \ldots, P_n$ permutation matrices of appropriate sizes and use X^j to denote the jth column of X.

Table 1.

Polytopes	Strong Linear Preservers
$\mathbf{DS}(n)$	$T(X) = PXQ$ or PX^tQ
$\mathbf{CS}(m,n)$	$T[X^1 \cdots X^n] = [P_1X^1 \cdots P_nX^n]Q$
$\mathbf{DsS}(m,n)$	$T(X) = PXQ$ or (PX^tQ and $m = n$)
$\mathbf{CsS}(m,n)$	$T[X^1 \cdots X^n] = [P_1X^1 \cdots P_nX^n]Q$
$\operatorname{conv}\mathbf{A}(n)$	$T(X) = PXQ$ or PX^tQ with $PQ \in \mathbf{A}(n)$, $n \geq 5$

Professor Li is a well-known expert in linear preserver problems. I had to supervise two Ph.D. students at Tamkang University, so I asked Professor Li for suitable related problems on this topic. Li suggested to me the problem of characterizing the strong linear preserves of $U(\mathbf{r}, \mathbf{c})$, the polytope of nonnegative matrices with fixed row sums vector $\mathbf{r}$ and column sums vector $\mathbf{c}$.

Here is a list of the conjectures given in this project in related to the strong linear preserver problems of $U(\mathbf{r}, \mathbf{c})$:

Conjecture 1. Let $\mathbf{r} = (r_1, \ldots, r_m)$ and $\mathbf{c} = (c_1, \ldots, c_n)$ be nonnegative vectors such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$. Let T be a linear map on $\operatorname{span} U(\mathbf{r}, \mathbf{c})$. In almost all cases the following are equivalent:
(a) $T(U(\mathbf{r}, \mathbf{c})) = U(r, c)$.
(b) T is given by $T(X) = P_\pi^t X Q_\tau$ or $\mathbf{r} = \mathbf{c}$ and $T(X) = P_\pi^t X^t Q_\tau$, for some $\pi \in S_m$, $\tau \in S_n$ such that $r_{\pi(i)} = r_i$ for $i = 1, \ldots, m$, and $c_{\tau(j)} = c_j$ for $j = 1, \ldots, n$, where P_π denotes the $m \times m$ permutation matrix $[e_{\pi(1)} \cdots e_{\pi(m)}]$ (e_j being the jth standard unit vector) and Q_τ is the $n \times n$ permutation matrix defined in a similar way.

Conjecture 2. Let $\mathbf{r} = (r_1, \ldots, r_m)$ be a nonnegative vector and let T be a linear map on $\operatorname{span} U(\mathbf{r})$. In almost all cases the following are equivalent:
(a) $T(U(\mathbf{r}, \mathbf{c})) = U(\mathbf{r}, \mathbf{c})$.
(b) T is of the form $T(X) = P_\pi^t X P_\pi$, where $\pi \in S_n$ satisfies $r_{\pi(i)} = r_i$ for $i = 1, \ldots, n$.

Conjecture 3. Conjecture 1 still holds if we replace $U(\mathbf{r}, \mathbf{c})$ by $sU(\mathbf{r}, \mathbf{c})$, the polytope of all $m \times n$ nonnegative matrices with row sums (respectively, column sums) vector less than or equal to $\mathbf{r}$ (respectively, $\mathbf{c}$).

Conjecture 4. Conjecture 2 still holds if we replace $U(\mathbf{r})$ by $sU(\mathbf{r})$.

Conjecture 5. A linear map T on $\operatorname{span} \mathbf{SDS}(n)$ is a strong linear preserver of $\mathbf{SDS}(n)$ if and only if there exists a permutation matrix P such that $T(X) = P^t X P$.

Together with my Ph.D. student Shwu-Huey Lin I embarked on the project. It did not take us too long to realize that just mimicking the arguments of the papers [1] and [3] would not work. The situation for the $U(\mathbf{r}, \mathbf{c})$ case is considerably more intricate. Moreover, the extreme matrices of the polytope $U(\mathbf{r}, \mathbf{c})$ are determined by their associated bipartite graphs and it is difficult to deal with bipartite graphs. So, instead of treating this general problem, we switched to consider the $U(\mathbf{r})$ case, in which the extreme matrices are determined by their associated (undirected) graphs. As a start we began with the case $\mathbf{SDS}(n)$, i.e., $U(e)$, where e denotes the vector of all 1's of $\mathbb{R}^n$.

2. Results and Discussions

After many months of hard work, we finally obtained the following result, which confirms Conjecture 5:

Theorem 1. *Let T be a linear map on* $\operatorname{span}\mathbf{SDS}(n)$, $n \geq 3$. *The following conditions are equivalent*:

(a) $T(\mathbf{SDS}(n)) = \mathbf{SDS}(n)$.

(b) *T is given by $T(X) = P^t X P$ for some $P \in \mathbf{P}(n)$.*

The case $n = 2$ is straightforward. The polytope $\mathbf{SDS}(2)$ is a line segment with endpoints I_2 and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. So there are exactly two strong linear preservers, namely, the identity map and the one which interchanges I_2 and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Note that the latter map is not of the form as given by Theorem 1(b) (for $n = 2$).

The most difficult part of the proof of Theorem 1 is to show that every strong linear preserver of $\mathbf{SDS}(n)$ satisfies $T(I_n) = I_n$. For the purpose, we need to make use of the concept of neighborly extreme points of a polytope. Two extreme points of a polytope are said to be neighborly if the line segment joining them is a face of the polytope. Clearly, if T is strong linear preserver of $\mathbf{SDS}(n)$, then T maps $\mathcal{E}(\mathbf{SDS}(n))$, the set of extreme points of $\mathbf{SDS}(n)$, onto itself. Moreover, for any $A \in \mathcal{E}(\mathbf{SDS}(n))$, $N(A)$ and $N(T(A))$ have the same cardinality, where we use $N(A)$ to denote the set of extreme points of $\mathbf{SDS}(n)$ that are neighborly to A. So, a relevant question to ask is, given $A \in \mathcal{E}(\mathbf{SDS}(n))$, what is the cardinality of $N(A)$? We translate this into a problem on graphs as follows.

By a result due to M. Katz, $\mathcal{E}(\mathbf{SDS}(n))$ consists of those matrices whose graphs have line segments or odd cycles as connected components. We call a graph with the latter property an LOCC graph. We call two LOCC graphs G, H on the same vertex set neighborly if their union $G \cup H$ contains G and H as its only spanning LOCC subgraphs. It is straightforward to check that for any $A, B \in \mathcal{E}(\mathbf{SDS}(n))$, A and B are neighborly extreme points if and only if their graphs $G(A)$ and $G(B)$ are neighborly (as LOCC graphs). With a little work, one can also show that if G and H are two neighborly LOCC graphs, then $G \cup H$ has a unique connected component which is not a line segment or an odd cycle and moreover this unique connected component is itself the union of two LOCC graphs. We refer to it as the distinguished component of $G \cup H$. It turns out that for any LOCC graph G, the cardinality of $N(G)$ is equal to the cardinality of the collection of distinguished components of

$G \cup K$ as K runs through the set $N(G)$. So, a relevant question is, when is a connected graph the union of two neighborly LOCC graphs ? We answer it in the following theorem:

Theorem 2. *A connected graph is the union of two neighborly LOCC graphs if and only if it is one of the following*:

(a) *a path of length* ≥ 1 *with odd cycles attached at its two ends* (*such that the path and the two odd cycles are pairwise internally disjoint*);

(b) *an odd cycle of length* ≥ 3 *with an odd simple walk* (*open or closed, internally disjoint from the cycle*) *joining two* (*not necessarily distinct*) *vertices of the cycle*; *or*

(c) *an even cycle of length* ≥ 4.

Next, we found that for any LOCC graph G, the cardinality of $N(G)$ is determined by the number of line segment connected components of G. More precisely, we have the following result:

Theorem 3. *Let* $n \geq 3$ *be a given positive integer. For any* i, $i = 0, \ldots, [n/2]$, *let* G_i^n *denote the collection of all LOCC graphs with vertex set* $\langle n \rangle$ *which have precisely* i *line segments among their connected components. Then the cardinality of* $N(G)$ *is independent of the choice of* G *from* G_i^n. *If* N_i^n *denotes the common value of* $|N(G)|$ *for* $G \in G_i^n$, *then we have*

$$N_0^n < N_1^n < \cdots < N_{[n/2]}^n.$$

The proofs of Theorems 2 and 3 are the hardest part of this project.

Perhaps, it is also worth mentioning that originally we tried to prove the relation $T(I_n) = I_n$ (for a strong linear preserver T of $\mathbf{SDS}(n)$) without using the concept of LOCC graphs and we reduced it to the problem of proving the following:

Conjecture. For each positive integer $n \geq 3$, $(n-1)h(n-1)$ is not divisible by $h(n) - h(n-1)$, where $h(n) = |\mathcal{E}(\mathbf{SDS}(n))|$.

By running a computer, we verified the above conjecture for $n = 3, \ldots, 171$. However, for $n \geq 172$, the numbers involved are too large (larger than 10^{305}) and our data overflow.

After we completed our proof of Theorem 1, we learned that Professor Li and his student Chiang [2] had found a shorter and different proof for the same result. In fact, they also showed that, except for certain low dimensional cases, the strong linear preservers of the set of symmetric permutation matrices are also of the form $X \mapsto P^t XP$. However, our approach has the extra bonus of increasing our understanding of the structure of the polytope $\mathbf{SDS}(n)$ — we are able to characterize completely the neighborly relation between the extreme points of this polytope.

Until today we are unable to establish the above Conjecture. But that is perhaps something of minor interest now.

For the strong linear preserver of $\mathbf{SDsS}(n)$, the polytope of $n \times n$ symmetric doubly substochastic matrices, we obtained the following:

Theorem 4. *Let T be a linear map on the space of $n \times n$ real symmetric matrices, $n \geq 1$. The following conditions are equivalent*:
(a) $T(\mathbf{SDsS}(n)) = \mathbf{SDsS}(n)$.
(b) *T is given by $T(X) = P^t XP$ for some $P \in \mathbf{P}(n)$.*

Again by a result of Katz, the extreme points of $\mathbf{SDsS}(n)$ are those matrices whose connected components are each a line segment, an odd cycle, or an isolated point. So, it should be more difficult to characterize the neighborly relation between the extreme points of $\mathbf{SDsS}(n)$. Fortunately, we need only partial information about this neighborly relation. The proof of Theorem 4 is easier than that of Theorem 1, though far from being trivial.

For the polytope of $U(\mathbf{r})$, it is natural to ask the following:

Question. Are the strong linear preservers of $U(\mathbf{r})$ always of the form $X \mapsto P_\pi^t XP_\pi$ for some $\pi \in S_n$ that satisfies $r_j = r_{\pi(j)}$ for $j = 1, \ldots, n$?

By examining the case $n = 3$, we showed that when the components of $\mathbf{r}$ are distinct, $U(\mathbf{r})$ has many strong linear preservers other than the identity map. This implies that the answer to the above question is in the negative, and we suspect that, in general, there is no nice characterization for the strong linear preservers of $U(\mathbf{r})$.

For the purpose, we need to make use of the known result (discovered independently by Converse and Katz, Lewin, and Brualdi) that a matrix $A \in U(\mathbf{r})$ is an extreme point of $U(\mathbf{r})$ if and only if the connected components of $G(A)$ are trees or simple odd cacti. Notice, however, that the preceding result does not characterize the extreme points of $U(\mathbf{r})$. Indeed, with a given $n \times 1$ nonnegative vector $\mathbf{r}$, it is not true that every graph on $\langle n \rangle$ with trees or simple odd cacti as connected components can be realized as the graph of some extreme matrix in $U(\mathbf{r})$.

Clearly, a strong linear preserver of $U(\mathbf{r})$ induces a permutation on $\mathcal{E}(U(\mathbf{r}))$, but not conversely. In order that a permutation on $\mathcal{E}(U(\mathbf{r}))$ gives rise to a strong linear preserver of $U(\mathbf{r})$, it is necessary and sufficient that the permutation preserves all the linear relations between the elements of $\mathcal{E}(U(\mathbf{r}))$. For the case $n = 3$ (and also $n = 2$), we determine completely the set $\mathcal{E}(U(\mathbf{r}))$ and the linear relations between its elements. It turns out that the answers depend on the relations between r_1, r_2, r_3, the components of the vector $\mathbf{r}$; there are eight different cases if we assume $r_3 > r_2 > r_1 > 0$. Then we determine completely the strong linear preservers (of $U(\mathbf{r})$) for the individual cases. The number of strong linear preservers varies between 1 and 8, as given by the following table:

Table 2.

Cases	Number of strong linear preservers
$r_1 + r_2 < r_3$, $r_2 > r_1$	4
$r_1 + r_2 > r_3 > r_2 > r_1$	1
$r_1 + r_2 < r_3$, $r_2 = r_1$	8
$r_1 + r_2 = r_3$, $r_2 > r_1$	4
$r_1 + r_2 = r_3$, $r_2 = r_1$	8
$r_1 + r_2 > r_3 = r_2 > r_1$	2
$r_1 + r_2 > r_3 > r_2 = r_1$	2
$r_1 = r_2 = r_3$	6

Because of the limit of time, we have not considered the strong linear preservers of $U(\mathbf{r})$ for $n \geq 4$. The problem seems to be intractable for a general n. When n is large, we cannot think of a feasible way to determine all the extreme matrices of $U(\mathbf{r})$, as the number of linearly independent linear relations between the extreme matrices may be large. So for large n, the determination of all the strong linear preservers of $U(\mathbf{r})$ seems impossible. But we hope it is possible to answer some interesting questions, like the following:

Question 1. Determine when $U(\mathbf{r})$ has only one strong linear preserver (namely, the identity operator).

Question 2. For a fixed n, what is the maximum number of strong linear preservers of $U(\mathbf{r})$? Is 2^n an upper bound ?

They are for future work.

3. Self-evaluation of Performance

For the five conjectures raised in our project, we only answered Conjecture 5 (and in the affirmative). We have only treated Conjecture 2 partially. The case $n = 2$ or 3 seems to suggest that there is no nice characterization of the strong linear preservers of $U(\mathbf{r})$ for general n, and so the answer to Conjecture 2 should be in the negative. But at this stage, we are not completely sure. Existing results on strong linear preservers (for instance, of $\mathbf{DS}(2)$ or $A(4)$) also indicate that the lower-dimensional cases are often the atypical cases. So it is also possible that when n is large enough, the strong linear preservers of $U(\mathbf{r})$ are all of the desirable forms. We have also not treated Conjectures 1, 3 and 4 (for $U(\mathbf{r}, \mathbf{c})$, $sU(\mathbf{r}, \mathbf{c})$ and $sU(\mathbf{r})$ respectively). At present, we incline to believe that the answers to these conjectures are all in the negative.

The project has been carried out pretty well, inspite of the fact that only one out of the five conjectures answered definitely. As a consequence of working for this project, my Ph.D. student Shwu-Huey Lin has finished her Ph.D. thesis "Strong linear preservers of some polytopes of symmetric nonnegative matrices". The first half of her thesis, which is about the strong linear preservers of $\mathbf{SDS}(n)$ and $\mathbf{SDsS}(n)$, forms the contents of paper [4], which is accepted for publication in Linear Algebra and Its Applications. I expect that with more follow-up works on the strong linear preservers of $U(\mathbf{r})$, another paper will come out.

REFERENCES

[1] H. Chiang and C.K. Li, Linear maps leaving the alternating group invariant, *Linear Algebra Appl.* **340** (2002), 69–80.

[2] H. Chiang and C.K. Li, Linear maps leaving invariant subsets of nonnegative symmetric matrices, preprint.

[3] C.K. Li, B.S. Tam and N.K. Tsing, Linear maps preserving permutation and stochastic matrices, *Linear Algebra Appl.* **341** (2002), 5–22.

[4] S.H. Lin and B.S. Tam, Strong linear preservers of symmetric doubly stochastic or doubly substochastic matrices, to appear in *Linear Algebra Appl.*

附

件

Strong linear preservers of symmetric doubly stochastic or doubly substochastic matrices

Shwu-Huey Lin[a], **Bit-Shun Tam**[b,*,1]

[a,b]Department of Mathematics, Tamkang University, Tamsui, Taiwan 251, ROC

January 12, 2003

Abstract.

Let $\mathcal{S}$ denote either the set of $n \times n$ symmetric doubly stochastic matrices or the set of $n \times n$ symmetric doubly substochastic matrices and let T be a linear map on span $\mathcal{S}$. We prove that $T(\mathcal{S}) = \mathcal{S}$ if and only if there exists an $n \times n$ permutation matrix P such that $T(X) = P^t XP$ for all $X \in \operatorname{span} \mathcal{S}$. Our proofs make use of the concept of neighborly extreme points of a polytope and depend on some intricate graph theory.

AMS classification: 15A04; 15A51; 05C50.
Keywords: Linear preserver; Symmetric doubly stochastic matrix; Symmetric doubly substochastic matrix; LOCC graph; Neighborly extreme point.

[*] Corresponding author
E-mail addresses: bsm01@mail.tku.edu.tw (B.-S. Tam),
shlin@email.au.edu.tw (S.-H. Lin).

[1] Supported by the National Science Council of the Republic of China

1. Introduction

Let C be a nonempty subset of a finite-dimensional linear space, and let T be a linear map on span C. We say that T *preserves* (respectively, *strongly preserves*) C if $T(C) \subseteq C$ (respectively, $T(C) = C$). It is clear that T strongly preserves C if and only if T is bijective and T, T^{-1} both preserve C. Also, for a compact convex set C, T strongly preserves C if and only if T strongly preserves $\mathcal{E}(C)$, where we use $\mathcal{E}(C)$ to denote the set of extreme points of C.

Recently, Li, Tam and Tsing [L–T–T] obtained, besides other results, the following characterizations of the strong linear preservers of $\mathbf{DS}(n)$, the polytope of doubly stochastic matrices, and of $\mathbf{DsS}(m, n)$, the polytope of $m \times n$ doubly substochastic matrices:

Theorem A. *Let T be a linear map on* $\operatorname{span}(\mathbf{DS}(n))$. *The following conditions are equivalent*:
(a) $T(\mathbf{DS}(n)) = \mathbf{DS}(n)$.
(b) $T(\mathbf{P}(n)) = \mathbf{P}(n)$.
(c) *T is given by $T(X) = PXQ$ or $T(X) = PX^tQ$ for some $P,\ Q \in \mathbf{P}(n)$.*

Theorem B. *Let T be a linear map on $\mathbb{R}^{m\times n}$. The following conditions are equivalent*:

(a) $T(\mathbf{DsS}(m, n)) = \mathbf{DsS}(m, n)$.

(b) $T(\mathbf{sP}(m, n)) = \mathbf{sP}(m, n)$.

(c) *There exist $P \in \mathbf{P}(m)$ and $Q \in \mathbf{P}(n)$ such that T is given by*:
(i) $T(X) = PXQ$, *or*
(ii) $T(X) = PX^tQ$ (*and* $m = n$).

In the above, $\mathbf{P}(n)$ denotes the set of $n \times n$ permutation matrices, and $\mathbf{sP}(m, n)$ denotes the set of $m \times n$ subpermutation matrices.

Subsequently, Chiang and Li [C–L1] also characterized the strong linear preservers of $\mathbf{A}(n)$, the set of $n \times n$ even permutation matrices. The case $n = 2$ or 3 is straightforward, the case $n = 4$ is atypical (see [C–L1, Theorem 1.2] for the detail), and for $n \geq 5$, the answer is the expected one:

Theorem C. *Let $n \geq 5$. A linear map T on* span $\mathbf{A}(n)$ *satisfies $T(\mathbf{A}(n)) = \mathbf{A}(n)$ if and only if there exist $P,\ Q \in \mathbf{P}(n)$ with $PQ \in \mathbf{A}(n)$ such that T is given by*:

$$T(X) = PXQ \text{ or } T(X) = PX^tQ.$$

The purpose of this paper is to establish the following characterizations of the strong linear preservers of $\mathbf{SDS}(n)$, the polytope of $n \times n$ symmetric doubly stochastic matrices, and of $\mathbf{SDsS}(n)$, the polytope of $n \times n$ symmetric doubly substochastic matrices:

Theorem 1. *Let T be a linear map on* $\operatorname{span}\mathbf{SDS}(n)$, $n \geq 3$. *The following conditions are equivalent*:

(a) $T(\mathbf{SDS}(n)) = \mathbf{SDS}(n)$.

(b) *T is given by $T(X) = P^tXP$ for some $P \in \mathbf{P}(n)$.*

Theorem 2. *Let T be a linear map on the space of $n \times n$ real symmetric matrices, $n \geq 1$. The following conditions are equivalent*:

(a) $T(\mathbf{SDsS}(n)) = \mathbf{SDsS}(n)$.

(b) *T is given by $T(X) = P^tXP$ for some $P \in \mathbf{P}(n)$.*

One may think that, for a compact convex set C, it is more natural to study the (strong) affine preservers of C instead of (strong) linear preservers. In this respect, we would like to make some relevant remarks. First of all, if the affine hull of C does not contain the origin (for instance, if C is the polytope $\mathbf{DS}(n)$ or $\mathbf{SDS}(n)$), then the problem of studying affine preservers of C (with domain and codomain both equal to aff C) is equivalent to the problem of studying linear preservers of C (with domain and codomain both equal to span C). This is because, when $O \notin \operatorname{aff} C$, any affine map $T : \operatorname{aff} C \to \operatorname{aff} C$ can be extended in a unique way to a linear map $\widetilde{T} : \operatorname{span} C \to \operatorname{span} C$, and moreover the association $T \mapsto \widetilde{T}$ is a linear isomorphism. On the other hand, if $O \in \operatorname{aff} C$, then there may exist strong affine preservers of C, which are not linear. (For instance, take C to be an equilateral triangle in the plane with the origin as one of the vertices.) However, when C is the polytope $\mathbf{SDsS}(n)$, the strong linear preserver problem and the strong affine preserver problem have the same answer. In other words, in Theorem 2, if we replace "linear map" by "affine map", then the result is still valid. This is because, as we shall explain at the end of Section 4, every strong affine preserver of $\mathbf{SDsS}(n)$ necessarily fixes O_n and hence is linear.

We shall need the following characterizations of the extreme points of $\mathbf{SDS}(n)$ and $\mathbf{SDsS}(n)$ due to M. Katz [K1, K2] (or see [B–P, Chapter 4, Section 3]).

Theorem D. *The extreme points of the polytope* $\mathbf{SDS}(n)$ *are those matrices which are permutationally similar to direct sums of* (*some of*) *the following three types of matrices*:

(i) $[1]$, 1×1 *matrix,*

(ii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, *$2 \times 2$ matrix, and*

(iii) *The $k \times k$ symmetric matrix with $1/2$ at its $(1,k)$, $(k,1)$ and $(i,i+1)$, $(i+1,i)$ entries for $i = 1, \ldots, k-1$ and zero elsewhere, where k is an odd integer ≥ 3.*

Theorem E. *The extreme points of the polytope $\mathbf{SDsS}(n)$ are those matrices that are permutationally similar to matrices of the form $B \oplus O_{n-k}$, where B is an extreme point of $\mathbf{SDS}(k)$ for some nonnegative integer $k \leq n$.*

Our proof of Theorem 1 resembles that for Theorem A as done in [L–T–T] or Theorem C as done in [C–L1]. However, there are also enough differences that deserves mentioning. In [L–T–T] (respectively, [C–L1]), in order to show that a strong linear preserver of $\mathbf{DS}(n)$ (respectively, of $\mathbf{A}(n)$) is of the desired form, the problem is reduced to one in which the strong linear preserver under consideration fixes the identity matrix I_n. This reduction can be carried out easily, because the polytopes $\mathbf{DS}(n)$ and conv $\mathbf{A}(n)$ (or, more correctly, their groups of strong linear preservers) are transitive (on their sets of extreme points) in the sense that for any pair of extreme points there is a strong linear preserver which takes one extreme point to the other. In contrast, the polytope $\mathbf{SDS}(n)$ is not transitive (as can be readily seen from Theorem 1). A difficult part of our proof is to show that every strong linear preserver of $\mathbf{SDS}(n)$ fixes I_n. For the purpose, we shall make use of the concept of neighborly extreme points of a polytope. The polytope $\mathbf{DsS}(m,n)$ (also $\mathbf{SDsS}(n)$) is also not transitive. Nonetheless, the proof for the strong linear preservers of $\mathbf{DsS}(m,n)$ (respectively, $\mathbf{SDsS}(n)$) is easier than that for the strong linear preservers of $\mathbf{DS}(n)$ (respectively, of $\mathbf{SDS}(n)$). In our proof of $\mathbf{SDsS}(n)$ we shall again make use of the concept of neighborly extreme points. Conceivably, the idea of neighborly extreme points may also be applied to other strong linear preserver problems on polytopes that are not transitive. We elaborate on this in what follows.

We shall assume elementary properties of a convex set (see, for instance, [R]).

Let C be a polytope. Two extreme points x, y of C are said to be *neighborly* if $\{(1-\lambda)x + \lambda y : 0 \leq \lambda \leq 1\}$, the line segment joining x and y, is a face of C (of dimension 1); equivalently, the face $\Phi(\frac{x+y}{2})$ contains precisely two extreme points (namely, x and y), where we use $\Phi(w)$ to denote the face of C generated by w, i.e., the set $\{y \in C : w + \mu(w-y) \in C$ for some $\mu > 0\}$. If T is a strong linear preserver of C, then for any $w \in C$, we have $T\Phi(w) = \Phi(Tw)$, and hence $\dim \Phi(Tw) = \dim \Phi(w)$ as T is bijective. So it is clear that a strong linear preserver maps neighborly extreme points to neighborly extreme points. Denote by $N(x)$ the set of extreme points of C neighborly to an extreme point x of C. Then, for any strong linear preserver T of C, we have $TN(x) = N(Tx)$ for all $x \in \mathcal{E}(C)$. Also, for any nonnegative integer k, T maps the set $\mathcal{E}_k := \{x \in \mathcal{E}(C) : |N(x)| = k\}$ onto itself, where we use $|S|$ to denote the cardinality of the set S. Moreover, for any $x \in \mathcal{E}(C)$, we have $|\mathcal{E}_k \cap N(x)| = |\mathcal{E}_k \cap N(Tx)|$ for all positive integers k.

To treat the strong linear preserver problem of $\mathbf{SDS}(n)$, a relevant question to ask is, when two extreme points of $\mathbf{SDS}(n)$ are neighborly. In the light of Theorem D, we shall translate this into a problem on graphs.

After this work was completed, we learned that Chiang and Li had found a shorter and different proof for the characterization of the strong linear preservers of the set of symmetric doubly stochastic matrices. In fact, they also showed that, except for certain low dimensional cases, the strong linear preservers of the set of symmetric permutation matrices are also of the form $X \mapsto P^T XP$. We would like to thank them for showing us the preprint of their paper [C–L2]. We would also like to mention that our approach has the extra bonus of increasing our understanding of the structure of the polytope $\mathbf{SDS}(n)$ — now we are able to characterize completely the neighborly relation between the extreme points of this polytope.

2. LOCC graphs

All graphs considered in this paper are finite, have no multiple edges, and may contain loops. By a *path* (respectively, *cycle*) we mean a simple open (respectively, closed) walk. A path (or cycle) is said to be *odd* (or *even*) if it has odd (or even) number of edges. A loop is treated as an odd cycle of length 1. We call an edge which is not a loop a *line segment.* (It should be clear from the context whether we are dealing with a line segment of a graph or a line segment that joins two points in a linear space.)

By the *graph* of an $n \times n$ real symmetric matrix A, denoted by $G(A)$, we mean, as usual, the graph with vertex set $\langle n \rangle := \{1, \ldots, n\}$, where $\{i, j\}$ is an edge if and only if $a_{ij} \neq 0$.

We call a graph an *LOCC graph* if its connected components are each either a line segment or an odd cycle. By Theorem D, for any $A \in \mathcal{E}(\mathbf{SDS}(n))$, $G(A)$ is an LOCC graph. Indeed, there is a one-to-one correspondence between the set of extreme points of $\mathbf{SDS}(n)$ and the set of LOCC graphs with vertex set $\langle n \rangle$.

For our purposes, we shall adopt the following special definitions of union and join of graphs. Let G, H be two graphs. If G and H have the same vertex set, then we use $G \cup H$ to denote the graph whose vertex set is the common vertex set of G and H and whose edge set is the union of those of G and H, and refer to it as the *union of* G *and* H. If G and H have disjoint vertex sets, then we use $G \vee H$ to denote the graph whose vertex set and edge set are respectively the union of those of G and H, and refer to it as the *join of* G *and* H. Of course, we can also define the union and join of more than two graphs in a similar way. Evidently, the join of LOCC graphs is still an LOCC graph.

We call H a *spanning LOCC subgraph* of a graph G if H is an LOCC graph which is also a spanning subgraph of G. Two LOCC graphs G, H on the same vertex set are said to be *neighborly* if their union $G \cup H$ contains G and H as its only spanning LOCC subgraphs. For any LOCC graph G, we use $N(G)$ to denote the set of LOCC graphs which are neighborly to G.

Note that for any $A,\ B \in \mathbf{SDS}(n)$, $A \in \Phi(B)$ if and only if for all $i,\ j \in \langle n \rangle$, $a_{ij} = 0$ whenever $b_{ij} = 0$, or equivalently, $G(A)$ is a (spanning) subgraph of $G(B)$. Also, we have $G(\frac{A+B}{2}) = G(A) \cup G(B)$. Consequently, two extreme points A, B of $\mathbf{SDS}(n)$ are neighborly if and only if the LOCC graphs $G(A)$ and $G(B)$ are neighborly.

Remark 1. Let G, H_1, H_2 be LOCC graphs such that $H_1,\ H_2 \in N(G)$. If $G \cup H_1 = G \cup H_2$, then $H_1 = H_2$.

This is because, H_1 and H_2 are both spanning LOCC subgraphs of $G \cup H_1$, different from G, and $G \cup H_1$ has only two spanning LOCC subgraphs.

Consider a connected component C of $G \cup H$, where $G,\ H$ are LOCC graphs on the same vertex set. Clearly, there is no edge in G (or H) joining a vertex of C to a vertex that lies outside C. So the subgraph of G (respectively, of H) induced by the vertex set of C, which we call G_1 (respectively, H_1), must be the join of some connected components of G (respectively, of H) and hence is an LOCC graph. The LOCC graph G_1 (also H_1) is always a spanning subgraph of C. If C is a line segment or an odd cycle, then G_1 (also H_1) must be C itself, and in this case C is clearly a common connected component of G and H. If C is not a line segment or an odd cycle, then from the above, C is the union of two different LOCC subgraphs, namely, G_1 and H_1. In general, the LOCC graphs G_1 and H_1 need not be neighborly. (For instance, take G and H to be respectively the odd cycles $1 \to 2 \to 3 \to 4 \to 5 \to 1$ and $1 \to 4 \to 2 \to 5 \to 3 \to 1$. Then we have $C = G \cup H$, $G_1 = G$ and $H_1 = H$, and $G_1,\ H_1$ are not neighborly.) However, they must be neighborly if G and H are. This is clearly so if C equals $G \cup H$. If $C \neq G \cup H$, then we can argue by way of contradiction as follows. Suppose C contains a spanning LOCC subgraph, say K, different from G_1 and H_1. Take the join of K and the connected components of G, (vertex-)disjoint from C. The resulting graph is a spanning LOCC subgraph of $G \cup H$, which is different from G and H, in contradiction with the assumption that G and H are neighborly. Indeed, the preceding argument can be adapted to show that, in case G and H are neighborly, $G \cup H$ cannot contain more than one connected components which are not line segments or odd cycles.

From the above discussion, we see that if G and H are two neighborly LOCC graphs, then $G \cup H$ has a unique connected component which is not a line segment or an odd cycle. We shall refer to it as the *distinguished component* of $G \cup H$. Note that the distinguished component is itself the union of two neighborly LOCC graphs, and also that $G \cup H$ is equal to the join of its distinguished component and the common connected components of G and H, which are disjoint from the distinguished component. In other words, the union of two neighborly LOCC graphs is completely determined by its distinguished component. And, in view of Remark 1, we have

Remark 2. For any LOCC graph G, the cardinality of $N(G)$ is equal to the cardinality of the collection of distinguished components of $G \cup K$ as K runs through the set $N(G)$.

An example is in order.

Example. Let G denote the odd cycle $1 \to 2 \to 3 \to 4 \to 5 \to 1$. We are going to determine the LOCC graphs neighborly to G and also the distinguished components of $G \cup K$ as K runs through all LOCC graphs neighborly to G.

First, consider the LOCC graph H on $\langle 5 \rangle$ which is composed of a loop at the vertex 1 together with the line segments $\{2, 3\}$ and $\{4, 5\}$. Clearly $G \cup H$ is equal to the odd cycle G together with the loop at the vertex 1. Suppose K is a spanning LOCC subgraph of the latter graph. The possible candidates for the connected component of K that contains the vertex 1 are: the line segments $\{1, 2\}$, $\{1, 5\}$, the loop and the odd cycle G. If it is the line segment $\{1, 2\}$, then the connected component (of K) that contains the vertex 3 must be the line segment $\{3, 4\}$, and so the line segment $\{1, 5\}$ must be the connected component that contains the vertex 5 (and the vertex 1), which is a contradiction. Similarly, the connected component that contains the vertex 1 cannot be the line segment $\{1, 5\}$. If the connected component is the loop, then K equals H, and if it is G, then K equals G. So G and H are neighborly LOCC graphs. In this case, since $G \cup H$ is connected, $G \cup H$ is its own distinguished component.

Next, let $\widetilde{H}$ denote the LOCC graph on $\langle 5 \rangle$ which is composed of the odd cycle $1 \to 2 \to 3 \to 1$ and the line segment $\{4.5\}$. Then $G \cup \widetilde{H}$ is equal to the odd cycle G together with the line segment $\{1,3\}$ (joining the nonconsecutive vertices 1 and 3). Let K be a spanning LOCC subgraph of $G \cup \widetilde{H}$. One can show that the connected component of K that contains the vertex 1 cannot be the line segment $\{1,2\}$, $\{1,3\}$ or $\{1,5\}$. So the said connected component must be the odd cycle G or the odd cycle $1 \to 2 \to 3 \to 1$; hence K must be G or $\widetilde{H}$. This shows that G and $\widetilde{H}$ are neighborly LOCC graphs. In this case, $G \cup \widetilde{H}$ is also its own distinguished component.

By the above, we see that if we add one new edge (which is either a loop or a line segment joining two nonconsecutive vertices) to the odd cycle G, we obtain the distinguished component of the union of G and some LOCC graph neighborly to G. But if we add more than one edges, then clearly the resulting graph must contain more than two spanning LOCC subgraphs. So we have captured all LOCC graphs neighborly to G and also the corresponding distinguished components, and the cardinality of $N(G)$ is equal to $\binom{5}{2}$.

For our purpose, it is important to characterize connected graphs which are the union of two neighborly LOCC graphs. In this respect, we have the following result:

Theorem 3. *A connected graph is the union of two neighborly LOCC graphs if and only if it is one of the following*:

(a) *a path of length* ≥ 1 *with odd cycles attached at its two ends* (*such that the path and the two odd cycles are pairwise internally disjoint*);

(b) *an odd cycle of length* ≥ 3 *with an odd simple walk* (*open or closed, internally disjoint from the cycle*) *joining two* (*not necessarily distinct*) *vertices of the cycle*; *or*

(c) *an even cycle of length* ≥ 4.

Proof of Theorem 3. "If" part: One can readily show that if P is a graph of the form (a), (b) or (c), then P can be expressed as the union of two neighborly LOCC graphs. The less trivial part is to show that the two involved LOCC graphs are in fact neighborly. We demonstrate how this can be done when P is a graph of the form as given by (a) and the path involved is of odd length.

Suppose P is composed of the path $w_0 \to w_1 \to \cdots \to w_{2r-2} \to w_{2r-1}$, where $r \geq 1$, together with the odd cycles $\Gamma_1 : u_0 \to \cdots \to u_{2p-2} \to u_0$ and $\Gamma_2 : v_0 \to \cdots \to v_{2q-2} \to v_0$, where $w_0 = u_0$ and $w_{2r-1} = v_0$. Then the graph P is the union of the LOCC graphs G and H, where G is the join of the odd cycles Γ_1, Γ_2 and the line segments $\{w_{2l-1}, w_{2l}\}$, $l = 1, \ldots, r-1$, and

H is the join of the line segments $\{w_{2i}, w_{2i+1}\}$, $i = 0, \ldots, r-1$, $\{u_{2j-1}, u_{2j}\}$, $j = 1, \ldots, p-1$, and $\{v_{2k-1}, v_{2k}\}$, $k = 1, \ldots, q-1$. To show that the LOCC graphs G and H are neighborly, let K be a spanning LOCC subgraph of P. The possible candidates for the connected component of K that contains the vertex u_0 are: the odd cycle Γ_1 and the line segments $\{u_0, u_{2p-2}\}$, $\{u_0, u_1\}$ and $\{w_0, w_1\}$. One can show that the said connected component cannot be the line segment $\{u_0, u_{2p-2}\}$ or $\{u_0, u_1\}$ (cf. our Example). If it is the odd cycle Γ_1, then, arguing one by one, one can show that the following line segments are each connected components of $K : \{w_1, w_2\}, \{w_3, w_4\}, \ldots, \{w_{2r-3}, w_{2r-2}\}$. But one can also show that the connected component of K containing v_0 cannot be the line segment $\{v_0, v_1\}$ or $\{v_0, v_{2q-2}\}$. So, in this case, the connected component of K containing v_0 must be the cycle Γ_2. Hence, K is G. If the connected component of K containing u_0 is the line segment $\{w_0, w_1\}$, then by a similar argument one can also show that K is H. So G and H are the only spanning LOCC subgraphs of P, i.e., G and H are neighborly.

"Only if" part: We will depend on the following useful observation:

Assertion. *Let G, H be neighborly LOCC graphs such that $G \cup H$ is connected. If $G \cup H$ contains a subgraph K of the form* (a), (b) *or* (c), *and K, in turn, contains the join of some connected components of G as a spanning subgraph, then $G \cup H$ is equal to K and consequently is of the form* (a), (b) *or* (c).

Proof of Assertion. First, observe that if we have two graphs each of which is the union of two neighborly LOCC graphs such that one of the graph is a spanning subgraph of the other, then the two graphs are the same. Let $\widetilde{G}$ denote the spanning subgraph of K which is the join of some connected components of G. According to our assumption or the proved "if" part, the graphs $G \cup H$ and K are each the union of two neighborly LOCC graphs. So it suffices to show that $G \cup H$ and K, or equivalently, G and $\widetilde{G}$, has the same vertex set. Suppose not. Since K is the union of two neighborly LOCC graphs and $\widetilde{G}$ is a spanning LOCC subgraph of K, K must be the union of $\widetilde{G}$ and another LOCC graph, say $\widetilde{H}$. Let P be the join of $\widetilde{H}$ and the connected components of G disjoint from $\widetilde{G}$. Clearly, P is a spanning LOCC subgraph of $G \cup H$. Note that G and H cannot have a common connected component, as $G \cup H$ is connected and G, H are different. Since P shares at least one common connected component with G but H does not, P must be different from H. Also, P is different from G, because the subgraphs of P and G induced by the vertex set of K are respectively $\widetilde{H}$ and $\widetilde{G}$ and are different. So P is a spanning LOCC subgraph of $G \cup H$, different from G and H, which contradicts the assumption that G and H are neighborly. This completes the proof. ■

First, consider the case when G and H both have only line segment components. If $G \cup H$ has a vertex of degree 1, then G and H must share a common line segment component, which is a contradiction. So the degree of each vertex of $G \cup H$ must be 2 and $G \cup H$ contains at least one cycle K. It is clear that the edges of K are alternately connected components of G or H, i.e., K is a cycle of even length ≥ 4. Now K is of the form (c) and clearly it contains as a spanning subgraph the join of certain line segment components of G. So by the Assertion, $G \cup H$ equals K and hence is of the form (c).

Now, consider the case when G or H has an odd cycle component, say G. If G has only one connected component, then G must be an odd cycle of length ≥ 3 and we readily show that $G \cup H$ is of the form (b) (cf. our Example). So, without loss of generality, we may assume that G has an odd cycle C_G (possibly a loop) and there is an edge, say, $\{v_0, v_1\}$ of H such that v_0 lies on C_G but v_1 does not. We want to prove that $G \cup H$ is of the form (a) or (b). Assume to the contrary that this is not true. If the connected component of G that contains v_1 is an odd cycle, say C, then the graph which is composed of C, C_G and the edge $\{v_0, v_1\}$ is of the form (a), and moreover it contains as a spanning subgraph the join of the connected components C_G and C of G. Then, by the Assertion, $G \cup H$ is of the form (a), which contradicts our assumption. So the connected component of G that contains v_1 is a line segment, say, $\{v_1, v_2\}$.

Proceeding inductively, suppose that for $t \geq 1$, we have already constructed distinct vertices $v_1, \ldots, v_{2t}$, all lying outside C_G, such that for $j = 1, \ldots, t$, $\{v_{2j-2}, v_{2j-1}\}$ is an edge of H and $\{v_{2j-1}, v_{2j}\}$ is a connected component of G. If H has no edge incident with v_{2t} other than $\{v_{2t-1}, v_{2t}\}$, then the line segment $\{v_{2t-1}, v_{2t}\}$ is a common connected component of G and H, which is a contradiction, as $G \cup H$ is connected and G, H are different. So H must have an edge incident with v_{2t} other than $\{v_{2t-1}, v_{2t}\}$. If the edge joins v_{2t} to v_{2s}, where $1 \leq s \leq t$, (or, to a vertex of C_G), then by applying the Assertion, we can conclude that $G \cup H$ is of the form (a) (or, of the form (b)), which contradicts our assumption. Note that the case when the edge joins v_{2t} to v_{2s-1}, where $1 \leq s < t$, cannot happen; because, then the join of C_G, the line segments $\{v_{2j-1}, v_{2j}\}$, $1 \leq j \leq s-1$, $\{v_{2k}, v_{2k+1}\}$, $s \leq k \leq t-1$, and $\{v_{2t}, v_{2s-1}\}$, and the connected components of G not incident with $v_0, v_1, \ldots, v_{2t}$ (if any) is a spanning LOCC subgraph of $G \cup H$, different from G and H, which contradicts the hypothesis that G and H are neighborly LOCC graphs. So H must have an edge, say, $\{v_{2t}, v_{2t+1}\}$ such that v_{2t+1} does not lie on C_G and is different from $v_1, \ldots, v_{2t}$. If the connected component of G that contains v_{2t+1} is an odd cycle, then again by applying the Assertion, we arrive at a contradiction. So, v_{2t+1} is contained in a line segment component of G, say, $\{v_{2t+1}, v_{2t+2}\}$. Continuing in this way, we construct an infinite sequence $(v_k)_{k\in\mathbb{N}}$ of distinct vertices (all lying outside C_G such that for $j = 1, 2, \ldots, \{v_{2j-2}, v_{2j-1}\}$ is an edge of H and $\{v_{2j-1}, v_{2j}\}$ is a connected component of G), which is a contradiction, as our graphs are finite.

The proof is complete. ■

As noted in the discussions preceding Remark 2, if G and H are neighborly LOCC graphs, then the distinguished component of $G \cup H$ can be expressed as $G_1 \cup H_1$, where G_1, H_1 are neighborly LOCC graphs such that G_1 (respectively, H_1) is equal to the join of certain connected components of G (respectively, of H), and moreover G_1 and H_1 share no common connected components. So, for a given LOCC graph G, a graph D is equal to the distinguished component of $G \cup H$ for some $H \in N(G)$ if and only if D is of one of the forms (a), (b) or (c) as given by Theorem 3 and moreover D contains as a spanning subgraph the join of certain connected components of G, among which at most two are odd cycles (as can be seen from Theorem 3). To obtain such D, we choose some of the connected components of G (at most two of which are odd cycles) and add edges (but not vertices) so that the resulting graph is connected and satisfies (a), (b) or (c) of Theorem 3. In view of Remark 2, the number of ways this can be done is equal to the cardinality of $N(G)$.

Let n be a given positive integer. For $i = 0, 1, \ldots, [n/2]$, let G_i^n denote the collection of all LOCC graphs with vertex set $\langle n\rangle$ which have precisely i line segments among their connected components. The following result is crucial to our treatment of strong linear preservers of $\mathbf{SDS}(n)$.

Theorem 4. *Let $n \geq 3$ be a given positive integer. For any i, $i = 0, \ldots, [n/2]$, the cardinality of $N(G)$ is independent of the choice of G from G_i^n. If N_i^n denotes the common value of $|N(G)|$ for $G \in G_i^n$, then we have $N_0^n < N_1^n < \cdots < N_{[n/2]}^n$.*

Proof. Note that every member of $G_{[n/2]}^n$ is a graph composed of $[n/2]$ line segment components or $[n/2]$ line segment components together with a loop component, depending on whether n is even or odd. So any two graphs G, H in $G_{[n/2]}^n$ are isomorphic, and it is clear that the isomorphism between G and H induces a one-to-one correspondence between $N(G)$ and $N(H)$, hence we have $|N(G)| = |N(H)|$. To complete the proof of the first half, it remains to consider the case when $0 \leq i < [n/2]$.

We are going to prove the following:

Assertion. *Let $G \in G_i^n$, where $n \geq 3$ and $0 \leq i < [n/2]$. If G has at least one odd cycle component with more than one vertex and if we replace one* *such odd cycle component by loops at each of its vertices, then the resulting LOCC graph and the graph G have the same number of neighborly graphs.*

It is clear that, once the above assertion is proved, it will follow that $|N(G)| = |N(H)|$ whenever G, H belong to the same G_i^n.

Proof of Assertion. Let H be the LOCC graph obtained from G by replacing the odd cycle $C : u_1 \to \cdots \to u_k \to u_1$, where $k \geq 3$, by k loops $R_1, \ldots, R_k$ attached at the vertices $u_1, \ldots, u_k$ respectively. Let $\mathcal{G}$ (respectively, $\mathcal{H}$) denote the collection of distinguished components of $G \cup K$ (respectively, of $H \cup K$) as K runs through all LOCC graphs neighborly to G (respectively, to H). In view of Remark 2, it suffices to show that $\mathcal{G}$ and $\mathcal{H}$ have the same cardinality.

To obtain an element of $\mathcal{G}$ (respectively, of $\mathcal{H}$), we choose certain connected components of G (respectively, of H) and add edges so that the resulting graph is connected and satisfies (a), (b) or (c) of Theorem 3. For our choice, we may take C alone (respectively, precisely two loops, both from $R_1, \ldots, R_k$), or take only connected components of G (respectively, of H) other than C (respectively, $R_1, \ldots, R_k$), or take C (respectively, at least one of the loops $R_1, \ldots, R_k$) together with at least one connected component of G (respectively, of H) other than C (respectively, $R_1, \ldots, R_k$). So the elements of $\mathcal{G}$ (respectively, of $\mathcal{H}$) can be classified into three kinds according to the above choices.

If P is an element of $\mathcal{G}$ of the first kind, then P equals either the cycle C with a loop attached at one of its vertices or the cycle C with an edge joining two nonconsecutive vertices (cf. our Example). There are altogether $\binom{k}{2}$ such P. Similarly, if P is an element of $\mathcal{H}$ of the first kind, then P equals one of the line segments $\{u_r, u_s\}$, $r, s \in \langle k \rangle$, $r \neq s$, together with loops at its ends. Again, there are altogether $\binom{k}{2}$ such P. So $\mathcal{G}$ and $\mathcal{H}$ have the same number of elements of the first kind.

$\mathcal{G}$ and $\mathcal{H}$ also have the same number of elements of the second kind and, in fact, the same set of elements, because the connected components of G other than C and those of H other than $R_1, \ldots, R_k$ are the same.

It remains to compare the elements of $\mathcal{G}$ and those of $\mathcal{H}$ of the third kind.

By examining Theorem 3 and its proof carefully, one can see that an element of $\mathcal{G}$ (respectively, of $\mathcal{H}$) of the third kind must be one of the following:

(i) an odd path of length ≥ 1 with the odd cycle C (respectively, with one of the loops $R_1, \ldots, R_k$) attached at one end and an odd cycle of G other C (respectively, of H other than the loops $R_1, \ldots, R_k$) attached at the other end;

(ii) an even path of length ≥ 2 with the odd cycle C (respectively, with one of the loops $R_1, \ldots, R_k$) attached at one end and another odd cycle, which is not a cycle of G (respectively, of H), attached at the other end; or

(iii) the odd cycle C with an odd simple walk of length ≥ 3 joining two (not necessarily distinct) vertices of C (respectively, an odd path of length ≥ 3 with loops, both chosen from $R_1, \ldots, R_k$, attached at its two ends [accounting for graphs of the form (a) in Theorem 3 when the path is odd and the odd cycles at the two ends are both chosen from $R_1, \ldots, R_k$], or an odd cycle of length ≥ 3 with one of the loops $R_1, \ldots, R_k$ attached).

Let P be an element of $\mathcal{G}$ of the third kind. If P is of the form (i) or (ii), we obtain a graph $\widetilde{P}$ from P by replacing the cycle C by a loop at the vertex at which the path is attached to C. If it is of the form (iii), we obtain a graph $\widetilde{P}$ from P by replacing the cycle C by two loops at the two ends of the walk (or by one loop, in case we have a closed walk). In each case, $\widetilde{P}$ is of the form (a) or (b) of Theorem 3 and, in addition, it contains as a spanning subgraph the join of at least one of the loops $R_1, \ldots, R_k$ and certain connected components of H other than $R_1, \ldots, R_k$. So $\widetilde{P}$ is an element of $\mathcal{H}$ of the third kind. Furthermore, it is not difficult to see that the association $P \mapsto \widetilde{P}$ gives a one-to-one correspondence between the elements of $\mathcal{G}$ and those of $\mathcal{H}$ of the third kind. This completes the proof of our Assertion and hence the first half of the theorem.

To establish the last half of the theorem, it suffices to prove the following:

If G is an LOCC graph on $\langle n \rangle$ which is composed of i line segments and $n-2i(\geq 2)$ loops, where $n \geq 3$ and $0 \leq i \leq [n/2]-1$, and if H is the LOCC graph obtained from G by replacing two of its loops by a line segment joining their vertices, then $|N(G)| < |N(H)|$.

Let $\mathcal{G}$ and $\mathcal{H}$ have the same meanings as before. We want to show that for any $P \in \mathcal{G}$, we can associate with it some $\widetilde{P} \in \mathcal{H}$, and moreover the association is one-to-one but not onto.

We may assume that H is obtained from G by replacing the loops R_1, R_2 at the vertices u_1, u_2 by the line segment $\{u_1, u_2\}$. Consider any $P \in \mathcal{G}$. If P does not contain the vertex u_1 or u_2 or if P is equal to the line segment $\{u_1, u_2\}$ together with the loops R_1, R_2, then we take $\widetilde{P}$ to be P. If P contains exactly one of the vertices u_1, u_2, say u_1, then, in view of Theorem 3, P is either a path of length ≥ 1 with the loop R_1 attached at one end and an odd cycle attached at the other end, or is an odd cycle with the loop R_1 attached. In either case, we obtain $\widetilde{P}$ from P by adding the edge $\{u_1, u_2\}$ and the loop R_2 and deleting the loop R_1. If P contains the vertices u_1, u_2 and also other vertices, then P must be an odd path of length ≥ 3, with the loops R_1, R_2 attached at its two ends. In this case, we obtain $\widetilde{P}$ from P by adding the line segment $\{u_1, u_2\}$ and deleting the loops R_1 and R_2 (giving rise to an even cycle). Using Theorem 3, one can show that the association

$P \mapsto \widetilde{P}$ provides a well-defined one-to-one mapping from $\mathcal{G}$ into $\mathcal{H}$.

To complete the proof, it remains to show that there exists $\widehat{P} \in \mathcal{H}$ which is not an image of the above mapping. If n is odd or n is even and $i \leq [n/2]-2$, then G (and hence also H) has at least one loop other than R_1 and R_2, say, R_3 at the vertex u_3. In this case, take $\widehat{P}$ to be the graph obtained from $\{u_1, u_2\} \vee R_3$ by adding the edges $\{u_1, u_3\}$ and $\{u_2, u_3\}$. Since $\widehat{P}$ is of the form (b) as given by Theorem 3 and contains the join of the connected components $\{u_1, u_2\}$ and R_3 of H as a spanning subgraph, $\widehat{P} \in \mathcal{H}$. Note that any image of the above mapping has to be the line segment $\{u_1, u_2\}$ with loops attached at both ends, or it does not involve the vertices u_1 or u_2, or it is a path with the loops R_1 or R_2 attached at one end and another odd cycle attached at the other end and in addition $\{u_1, u_2\}$ is an edge of the path, or it is an even cycle. But $\widehat{P}$ satisfies none of the above, so $\widehat{P}$ has no pre-image. In the remaining case, we have, n is an even integer ≥ 4 and $i = [n/2]-1$. Then G is the join of $[n/2]-1$ line segments and the two loops R_1, R_2 (at the vertices u_1, u_2 respectively). Take any line segment of G, say, $L_1 = \{u_3, u_4\}$ and let $\widehat{P}$ be the connected graph obtained from $L_1 \vee \{u_1, u_2\}$ by adding the edges $\{u_1, u_3\}$, $\{u_2, u_3\}$ and a loop at u_4. Again, one can check that $\widehat{P} \in \mathcal{H}$ but $\widehat{P}$ is not an image of the said mapping. The proof is complete. ■

We shall need also the following technical lemma:

Lemma 1. *Let G and H be LOCC graphs on $\langle n\rangle$, where n is an odd integer ≥ 5. Suppose that G is composed of a cycle of length 3 and $[\frac{n}{2}]-1$ line segments and H is composed of three loops and $[\frac{n}{2}]-1$ line segments. Let*

$$\mathcal{N}_1 = \{K \in N(G) \cap G_0^n : G \cup K \textit{ is connected}\}$$

and

$$\mathcal{N}_2 = \{K \in N(H) \cap G_0^n : \textit{the distinguished component of } H \cup K \textit{ contains at least one of the three loops of } H\}.$$

Then $|\mathcal{N}_1| = 2|\mathcal{N}_2| > 0$.

Proof. It suffices to show that the cardinality of the collection of distinguished components of $G \cup K$ as K runs through all elements of $\mathcal{N}_1$ is twice that of the collection of distinguished components of $H \cup K$ as K runs through all elements of $\mathcal{N}_2$. There is no loss of generality in assuming that G and H have the same line segments. We are going to make use of the proof of the Assertion in the proof of Theorem 4.

Let $\mathcal{G}$ and $\mathcal{H}$ have the same meanings as before. First, we want to identify the elements of $\mathcal{N}_1$ and $\mathcal{N}_2$. The elements of $\mathcal{G}$ (respectively, of $\mathcal{H}$) can be classified into three kinds as in the proof of the Assertion. Let $K \in N(G)$ (respectively, $N(H)$) and let P denote the distinguished component of $G \cup K$ (respectively, of $H \cup K$). It is easy to see that if P is an element of $\mathcal{G}$ (respectively, of $\mathcal{H}$) of the first kind, then $G \cup K$ (respectively, $H \cup K$), and hence K, has at least one line segment component [as $G \cup K$ (respectively, $H \cup K$) and K share common line segment and odd cycle components if they exist]; if P is an element of $\mathcal{G}$ (respectively, of $\mathcal{H}$) of the second kind, then P does not contain the 3-cycle of G (respectively, any one of the three loops of H) as a subgraph. So, for such P, the corresponding K does not belong to $\mathcal{N}_i$ for $i = 1$ or 2. If P is an element of $\mathcal{G}$ (respectively, of $\mathcal{H}$) of the third kind, then it must be of the form (i), (ii) or (iii) as described in the proof of the Assertion (where the odd cycle C mentioned there becomes the 3-cycle of G in this lemma). Since G (respectively, H) has no odd cycles other than C (respectively, the three loops), P cannot be of the form (i). It is ready to check that when P (in $\mathcal{G}$ or $\mathcal{H}$) is of the form (ii), we have $K \notin G_0^n$ and so $K \notin \mathcal{N}_i$ for $i = 1$ or 2. Moreover, when $P \in \mathcal{G}$ is of the form (iii), we have $K \notin \mathcal{N}_1$ if the odd simple walk of P is closed; and if the walk is open, then we have, $K \in \mathcal{N}_1$ if and only if K has no line segment component, i.e., P and $G \cup K$ are the same and is equal to the 3-cycle C together with a path of length $2m + 1$ joining two distinct vertices of C, where, for convenience, we have introduced m to stand for $[\frac{n}{2}] - 1$. Denote the line segments of G (and H) by $L_1, \ldots, L_m$. Summarizing and rephrasing, for any $P \in \mathcal{G}$, we have $K \in \mathcal{N}_1$ if and only if P can be obtained as follows: Link up the line segments $L_1, \ldots, L_m$ to form a path of length $2m - 1$ and then join the two ends of the path to different vertices of C. On the other hand, when $P \in \mathcal{H}$ is of the form (iii), we have $K \notin \mathcal{N}_2$ if the odd simple walk of P is open; and if the walk is closed, then $K \in \mathcal{N}_2$ if and only if P equals an odd cycle of length of $2m+1$ that contains all the line segments $L_1, \ldots, L_m$ and with one of the three loops of H attached. Putting it differently, for any $P \in \mathcal{H}$, we have $K \in \mathcal{N}_2$ if and only if P can be obtained in the following way: Link up the line segments $L_1, \ldots, L_m$ to form a path of length $2m - 1$ and then join the two ends of the path to one of the three loops of H. But for each way of linking up the line segments $L_1, \ldots, L_m$ to form a path of length $2m - 1$ and then joining the two ends to one of the three loops, there are two ways of joining the two ends of the same path to two different vertices of C. It follows that we have $|\mathcal{N}_1| = 2|\mathcal{N}_2| > 0$. ■

One can readily check that in Lemma 1 if $n \leq 3$, then the sets $\mathcal{N}_1$ and $\mathcal{N}_2$ are both empty.

3. Strong preservers of symmetric doubly stochastic matrices

We would like point out that Theorem 1 does not extend to the case $n = 2$. The polytope $\mathbf{SDS}(2)$ has exactly two strong linear preservers, namely, the identity operator and the one which interchanges I_2 and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, as $\mathbf{SDS}(2)$ is a line segment with endpoints I_2 and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Evidently, the latter operator is not of the form $T(X) = P^tXP$, where $P \in \mathbf{P}(2)$.

This section is devoted to proving Theorem 1. The implication (b) $\Longrightarrow$ (a) is clear. It remains to show (a) $\Longrightarrow$ (b). Before we do that, we need two more lemmas.

We shall denote by $P(r, s)$ the $n \times n$ transposition (permutation) matrix with 1 at its (r, s), (s, r) and (t, t) positions for $t \in \langle n\rangle\backslash\{r, s\}$ and 0 elsewhere.

Lemma 2. *Let $\mathcal{C}_2$ denote the collection of all transposition matrices of $\mathbf{P}(n)$. Then $\mathcal{C}_2 \cup \{I_n\}$ is a basis for* span $\mathbf{SDS}(n)$.

Proof. Clearly, $\mathcal{C}_2 \cup \{I_n\}$ is a linearly independent subset of $\mathbf{SDS}(n)$. It remains to show that each extreme matrix of $\mathbf{SDS}(n)$ can be written as a linear combination of matrices in $\mathcal{C}_2 \cup \{I_n\}$. Consider any $A \in \mathcal{E}(\mathbf{SDS}(n))$. By Theorem D and the fact that $\mathcal{C}_2 \cup \{I_n\}$ is closed under taking permutation similarity, we may assume that A is already of the form $A_1 \oplus \cdots \oplus A_m$, where each A_j is of one of the three types (i), (ii) or (iii) as given by Theorem D. For $j = 1, \ldots, m$, let k_j denote the size of A_j. Also let

$$\widetilde{A}_j = I_{k_1} \oplus \cdots \oplus I_{k_{j-1}} \oplus A_j \oplus I_{k_{j+1}} \oplus \cdots \oplus I_{k_m}.$$

It is easy to check that $A = \widetilde{A}_1 + \widetilde{A}_2 + \cdots + \widetilde{A}_m - (m-1)I_n$. Also, for each j, if A_j is of type (i) or (ii), then clearly $\widetilde{A}_j \in \mathcal{C}_2 \cup \{I_n\}$; if A_j is of type (iii), then we also have $\widetilde{A}_j \in \operatorname{span}(\mathcal{C}_2 \cup \{I_n\})$ as

$$\begin{aligned} \widetilde{A}_j \;&=\; \frac{1}{2}[P(l_j+1, l_j+2) + P(l_j+2, l_j+3) + \cdots \\ &\quad + P(l_j+k_j-1, l_j+k_j) + P(l_j+k_j, l_j+1) - (k_j-2)I_n], \end{aligned}$$

where $l_j = k_1 + \cdots + k_{j-1}$ for $2 \le j \le m$ and $l_1 = 0$. This shows that $\mathcal{E}(\mathbf{SDS}(n)) \subseteq \operatorname{span}(\mathcal{C}_2 \cup \{I_n\})$, as desired. ■

For any integer i, $0 \le i \le [n/2]$, we use E_i^n to denote the set of all matrices $A \in \mathcal{E}(\mathbf{SDS}(n))$ that satisfy $G(A) \in G_i^n$. We use J_n to denote the $n \times n$

matrix all of whose entries equal 1. We also write $\widetilde{J}_n$ for $J_n - I_n$.

Proof of Theorem 1, (a) $\Longrightarrow$ (b).

Hereafter, we use T to denote a strong linear preserver of $\mathbf{SDS}(n)$ for $n \geq 3$.

Assertion 1. $T(I_n) = I_n$.

Proof of Assertion 1. For each $i = 0, \ldots, [n/2]$, by definition, E_i^n equals the set of matrices $A \in \mathcal{E}(\mathbf{SDS}(n))$ that satisfy $G(A) \in G_i^n$. By Theorem 4 the latter set, in turn, is equal to the set of all $A \in \mathcal{E}(\mathbf{SDS}(n))$ for which $|N(A)| = N_i^n$. So by the discussion near the end of Section 1, T maps the set E_i^n onto itself.

We first treat the case when n is even. By symmetry, clearly $\sum_{A \in E^n_{[n/2]}} A$ equals $\alpha_n \widetilde{J}_n$ for some $\alpha_n > 0$. But $T(E^n_{[n/2]}) = E^n_{[n/2]}$, so we have

$$T(\widetilde{J}_n) = \alpha_n^{-1} \left(\sum_{A \in E^n_{[n/2]}} T(A) \right) = \alpha_n^{-1} \left(\sum_{A \in E^n_{[n/2]}} A \right) = \widetilde{J}_n.$$

On the other hand, T also fixes the matrix $\sum_{A \in \mathcal{E}(\mathbf{SDS}(n))} A$, which is clearly of the form $\beta_n I_n + \gamma_n \widetilde{J}_n$ for some $\beta_n,\ \gamma_n > 0$. It follows that T fixes I_n.

Now consider the case when n is odd. When $n = 3$, by direct calculation, we have

$$\sum_{A \in \mathcal{E}(\mathbf{SDS}(3))} A = \frac{1}{2} I_3 + \frac{3}{2} J_3,$$

and so

$$\frac{1}{2} I_3 + \frac{3}{2} J_3 = \frac{1}{2} T(I_3) + \frac{3}{2} T(J_3).$$

On the other hand, $\sum_{A \in E_1^3} A = J_3$, so $J_3 = T(J_3)$. It follows that T fixes I_3.

For odd n, $n \geq 5$, we consider the class $E^n_{[n/2]-1}$. Each matrix in this class has graph made up of $[n/2] - 1$ (disjoint) line segments, together with three loops or one 3-cycle. So we can partition $E^n_{[n/2]-1}$ as $\mathcal{L} \cup \mathcal{T}$, where

$$\mathcal{L} = \{A \in E^n_{[n/2]-1} : G(A) \text{ contains three loops}\}$$

and

$$\mathcal{T} = \{A \in E^n_{[n/2]-1} : G(A) \text{ contains a 3-cycle}\}.$$

As explained before, T preserves the class $E^n_{[n/2]-1}$. We contend that T also preserves $\mathcal{L}$ and $\mathcal{T}$. Suppose not. Then there must exist $A \in \mathcal{T}$ such that

$T(A) = B$ for some $B \in \mathcal{L}$. An element R of $N(A) \cap E_0^n$ (respectively, of $N(B) \cap E_0^n$) can be classified as of the first or second kind according to whether or not the distinguished component of $G(A) \cup G(R)$ (respectively, of $G(B) \cup G(R)$) contains the join of the $[n/2] - 1$ line segments of $G(A)$ (respectively, of $G(B)$) as a spanning subgraph (noting that the distinguished component has to contain each of the $[n/2]-1$ line segments, as $G(R) \in G_0^n$). It is clear that $N(A) \cap E_0^n$ and $N(B) \cap E_0^n$ have the same number of elements of the first kind. Now we apply Lemma 1 with $G = G(A)$ and $H = G(B)$. Recall how one can obtain the distinguished components of the unions of an LOCC graph and its neighborly graphs (as given in the paragraph following Theorem 3). Note that an element R of $N(A) \cap E_0^n$ is of the second kind if and only if the distinguished component of $G(A) \cup G(R)$ contains the 3-cycle of $G(A)$ (besides all of the $[n/2] - 1$ line segments), i.e., if and only if $G(R) \in \mathcal{N}_1$. Similarly, an element R of $N(B) \cap E_0^n$ is of the second kind if and only if $G(R) \in \mathcal{N}_2$. So, by Lemma 1, the number of elements in $N(A) \cap E_0^n$ of the second kind is twice of that of $N(B) \cap E_0^n$. Hence, we have

$$|N(A) \cap E_0^n| > |N(B) \cap E_0^n| = |T(N(A)) \cap T(E_0^n)| = |N(A) \cap E_0^n|,$$

where the first equality holds as we have $T(N(A)) = N(T(A)) = N(B)$ and $T(E_0^n) = E_0^n$ and the second equality holds as T is bijective. So we arrive at a contradiction. This shows that T preserves the sets $\mathcal{L}$ and $\mathcal{T}$. But $\sum_{A \in \mathcal{T}} A = \omega \widetilde{J}_n$ for some $\omega > 0$, so T fixes $\widetilde{J}_n$. On the other hand, T also fixes the matrix $\sum_{A \in \mathcal{E}(\mathbf{SDS}(n))} A$, which is of the form $\beta_n I_n + \gamma_n \widetilde{J}_n$ for some β_n, $\gamma_n > 0$. It follows that T fixes I_n. This completes the proof of Assertion 1.

The next assertion can be proved by modifying the argument used for [L–T–T, Assertion 2 in the proof of Theorem 2.2], noting that in the course of our proof of Assertion 1 we also established $T(\widetilde{J}_n) = \widetilde{J}_n$. Here we give an alternative proof.

Assertion 2. *$T(\mathcal{C}_2) = \mathcal{C}_2$, where $\mathcal{C}_2$ has the same meaning as in Lemma 2.*

Proof of Assertion 2. Since the LOCC graphs on $\langle n \rangle$ neighborly to $G(I_n)$ are precisely the LOCC graphs on $\langle n \rangle$ which are joins of a line segment and $n - 2$ loops, $N(I_n)$ equals $\mathcal{C}_2$. So we have

$$T(\mathcal{C}_2) = T(N(I_n)) = N(T(I_n)) = N(I_n) = \mathcal{C}_2.$$

Assertion 3. *Suppose that $T(P(i, j)) = P(p, q)$ and $T(P(k, l)) = P(r, s)$. If $\{i, j\} \cap \{k, l\}$ is a singleton, then so is $\{p, q\} \cap \{r, s\}$.*

The proof of Assertion 3 is the same as that for [L–T–T, Assertion 3 in the proof of Theorem 2.2]. The proof of the next assertion is also a modification of that for [L–T–T, Assertion 4 in the proof of Theorem 2.2].

Assertion 4. *There exists $P \in P(n)$ such that $T(X) = P^tXP$ for all $X \in \mathcal{C}_2$.*

Proof of Assertion 4. First, we may assume that $T(P(1,2)) = P(1,2)$. Otherwise, since $T(\mathcal{C}_2) = \mathcal{C}_2$ by Assertion 2, we can choose a permutation $\sigma \in S_n$ that satisfies $T(P(1,2)) = P(\sigma(1), \sigma(2))$ and replace T by $\widetilde{T}$ defined by $\widetilde{T}(X) = P_\sigma^t T(X) P_\sigma$, where P_σ denotes the $n \times n$ permutation matrix whose jth column is the standard unit vector $e_{\sigma(j)}$. [Here and in what follows we use implicitly the formula $P_\sigma^t P(i_1, \ldots, i_k) P_\sigma = P(\sigma^{-1}(i_1), \ldots, \sigma^{-1}(i_k))$, where $P(i_1, \ldots, i_k)$ denotes P_τ for the cyclic permutation τ in S_n given by $\tau(j) = j$ for $j \notin \{i_1, \ldots, i_k\}$, $\tau(i_r) = i_{r+1}$ for $r = 1, \ldots, k-1$, and $\tau(i_k) = i_1$.]

By Assertion 3, $T(P(1,3)) = P(1,s)$ or $P(2,s)$ for some $s \geq 3$. We may assume that T also fixes $P(1,3)$. Otherwise, replace T by $\widetilde{T}$ defined by

$$\widetilde{T}(X) = P(3,s)T(X)P(3,s) \text{ or } P(3,s)P(1,2)T(X)P(1,2)P(3,s),$$

depending on whether $T(P(1,3)) = P(1,s)$ or $P(2,s)$. (If $s = 3$, $P(3,s)$ is treated as I_n.)

Consider any $l \geq 4$. By Assertion 3 with $(k,l) = (1,l)$ and $(i,j) = (1,2)$, (1,3) in turn, we infer that $T(P(1,l))$ is either $P(2,3)$ or $P(1,s_l)$ for some $s_l \geq 4$. Suppose that the former happens. Consider the element $A = \frac{1}{3}(P(1,2) + P(1,3) + P(1,l))$ of $\mathbf{SDS}(n)$. Notice that the face $\Phi(A)$ of $\mathbf{SDS}(n)$ contains precisely three extreme matrices, namely, $P(1,2)$, $P(1,3)$ and $P(1,l)$. On the other hand, $\Phi(T(A))$, which is $\frac{1}{3}(P(1,2) + P(1,3) + P(2,3))$, contains precisely five extreme matrices, namely, $P(1,2)$, $P(1,3)$, $P(2,3)$, I_n and $\begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix} \oplus I_{n-3}$. Since T is a strong linear preserver of $\mathbf{SDS}(n)$, $\Phi(A)$ and $\Phi(T(A))$ should have the same number of extreme matrices. So we arrive at a contradiction. This shows that for each $l \geq 4$, we have $T(P(1,l)) = P(1,s_l)$ for some $s_l \geq 4$. Let $\tau \in S_n$ be given by $\tau(i)$ equals i for $i = 1,2,3$ and equals s_i for $i = 4, \ldots, n$, and replace T by $\widetilde{T}$ where $\widetilde{T}(X) = P_\tau^t T(X) P_\tau$. Then we may assume that T fixes $P(1,i)$ for all $i \geq 2$.

Consider any distinct $r,\ s \in \{2, \ldots, n\}$. Since T fixes $P(1,r)$ and $P(1,s)$, by Assertion 3 and the fact that T fixes $P(1,i)$ for all $i \geq 2$, we readily infer that T also fixes $P(r,s)$. This completes the proof of Assertion 4.

Now by Lemma 2, Assertions 1 and 4, the implication (a) $\Longrightarrow$ (b) of Theorem 1 clearly follows. ■

4. Strong preservers of symmetric doubly substochastic matrices

In this section we treat the strong linear preserver problem for the polytope $\mathbf{SDsS}(n)$.

We shall denote by E_{ij} the $n \times n$ matrix with 1 at its (i, j) position and 0 elsewhere. Clearly, $\mathbf{SDsS}(n)$ contains all E_{ii} and $E_{ij} + E_{ji}$ for $i,\ j \in \langle n\rangle$, $i \neq j$. So span $\mathbf{SDsS}(n)$ equals the space of all $n \times n$ real symmetric matrices.

By Theorem E (and D), if $A \in \mathcal{E}(\mathbf{SDsS}(n))$, then the connected components of $G(A)$ are each a line segment, an odd cycle or an isolated vertex. In fact, it is easy to see that there is a one-to-one correspondence between the set $\mathcal{E}(\mathbf{SDsS}(n))$ and the collection of graphs on $\langle n\rangle$ whose connected components are each a line segment, an odd cycle or an isolated vertex.

One can show that for any $A,\ B \in \mathbf{SDsS}(n)$, $A \in \Phi(B)$ if and only if for all $i,\ j \in \langle n\rangle$, $a_{ij} = 0$ whenever $b_{ij} = 0$ (or, equivalently, $G(A)$ is a subgraph of $G(B)$) and moreover the ith row sum of A equals 1 whenever the corresponding row sum of B equals 1. (See [L–T–T, Proposition 1.1] for a more general result.) One would expect that the problem of determining when two extreme elements of $\mathbf{SDsS}(n)$ are neighborly is more difficult than the corresponding problem for the polytope $\mathbf{SDS}(n)$, which is already nontrivial. Fortunately, for our purposes, we need not resolve the said problem completely.

Lemma 3. *For any $A \in \mathcal{E}(\mathbf{SDsS}(n))$, $\Phi(A/2)$ contains exactly two extreme elements* (*namely, O_n and A*) *if and only if A equals E_{ii} or $E_{ij} + E_{ji}$ for some $i,\ j \in \langle n\rangle$, $i \neq j$.*

Proof. "If" part: For any $i,\ j \in \langle n\rangle$, $i \neq j$, it is readily checked that the face $\Phi((E_{ij} + E_{ji})/2)$ contains exactly two extreme elements, namely, $E_{ij} + E_{ji}$ and the zero matrix O_n. A similar assertion also holds for E_{ii}.

"Only if" part: For convenience, denote by A_k the $k \times k$ symmetric matrix as given in Theorem D, (iii). Consider any nonzero extreme element A of $\mathcal{E}(\mathbf{SDsS}(n))$ which is not of the form E_{ii} or $E_{ij} + E_{ji}$, where $i,\ j \in \langle n\rangle$, $i \neq j$. Since the class of matrices of the said form is invariant under permutation similarity, by Theorems E and D we may assume that A is already of one of the following forms:

(i) $(1) \oplus (1) \oplus B$, where $B \in \mathcal{E}(\mathbf{SDsS}(n-2))$,

(ii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus B$, where $B \in \mathcal{E}(\mathbf{SDsS}(n-4))$,

(iii) $(1) \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus B$, where $B \in \mathcal{E}(\mathbf{SDs}(n-3))$, and

(iv) $A_k \oplus B$, where $B \in \mathcal{E}(\mathbf{SDsS}(n-k))$ for some odd integer $k \geq 3$.

If A is of the form (i), then $\Phi(A/2)$ contains at least four extreme elements, namely, $A,\ O_n,\ (1) \oplus (0) \oplus B$ and $(0) \oplus (1) \oplus (B)$.

If A is of the form (ii), then $\Phi(A/2)$ contains at least four extreme elements, namely, $A,\ O_n,\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus O_2 \oplus B$ and $O_2 \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus B$.

If A is of the form (iii), then $\Phi(A/2)$ contains at least four extreme elements, namely, A, O_n, $(1) \oplus O_2 \oplus B$ and $(0) \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus B$.

If A is of the form (iv), then $\Phi(A/2)$ contains the extreme elements A, O_n, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus O_{k-2} \oplus B$ and many more.

This proves that for any $A \in \mathcal{E}(\mathbf{SDsS}(n))$, if $\Phi(A/2)$ contains exactly two extreme elements, then necessarily A equals E_{ii} or $E_{ij} + E_{ji}$ for some $i, j \in \langle n \rangle$, $i \neq j$. ■

It would be helpful to keep in mind the following observation, though we do not need it in our proofs.

Remark 3. Let A, B be extreme elements of $\mathbf{SDsS}(n)$, both different from the zero matrix O_n. If A, B are neighborly extreme points, then there exist an edge e of $G(A)$ and also an edge f of $G(B)$ such that e and f meet at a common vertex (and possibly, $e = f$).

To see this, suppose that the edges of $G(A)$ and those of $G(B)$ do not meet at a common vertex. Then all row (column) sums of $(A+B)/2$ is less than or equal to $1/2$. So, besides A and B, $\Phi((A+B)/2)$ also contains O_n as an extreme element. Hence, A and B are not neighborly extreme points.

Proof of Theorem 2.

It suffices to consider the implication (a) $\Longrightarrow$ (b). In what follows, we denote by T a strong linear preserver of $\mathbf{SDsS}(n)$.

Assertion 5. *T permutes the elements of the set*

$$\mathcal{B} = \{E_{ii} : i \in \langle n \rangle\} \cup \{E_{ij} + E_{ji} : i, j \in \langle n \rangle,\ i \neq j\}.$$

Proof of Assertion 5. Since T is a strong linear preserver of $\mathbf{SDsS}(n)$, T maps the elements of $\mathcal{E}(\mathbf{SDsS}(n))$, and in particular the elements of $\mathcal{B}$, into $\mathcal{E}(\mathbf{SDsS}(n))$. In view of Lemma 3, for any $A \in \mathcal{E}(\mathbf{SDsS}(n))$, we have, $A \in \mathcal{B}$ if and only if $\Phi(A/2)$ has exactly two extreme elements if and only if $\Phi(T(A)/2)$ has exactly two extreme elements if and only if $T(A) \in \mathcal{B}$. So we have $T(\mathcal{B}) = \mathcal{B}$.

Assertion 6. *For any $i, j \in \langle n \rangle$, $i \neq j$, we have*

$$|N(E_{ii})| = \frac{n^2}{2} + \frac{n}{2} \quad \textit{and} \quad |N(E_{ij} + E_{ji})| = \frac{3}{2}n^2 - \frac{3}{2}n + 1.$$

Proof of Assertion 6. It is clear that we need only consider the case when $i = 1$ and $j = 2$.

The first equality of Assertion 6 follows readily once we establish the following claim:

$$N(E_{11}) = \{O_n, E_{11} + E_{ii}, E_{11} + E_{ij} + E_{ji}, E_{1i} + E_{i1} : i, j \in \langle n \rangle \backslash \{1\},\ i \neq j\}.$$

To show that $E_{1i} + E_{i1} \in N(E_{11})$, where $i \in \langle n \rangle \backslash \{1\}$, consider any $B \in \mathcal{E}(\Phi((E_{11} + E_{1i} + E_{i1})/2))$. Note that the connected component of B containing the vertex 1 cannot be an isolated vertex, because the first row sum of B has to be 1, as the first row sum of $(E_{11} + E_{1i} + E_{i1})/2$ equals 1 and $B \in \Phi((E_{11} + E_{1i} + E_{i1})/2)$. But $G(B)$ is a spanning subgraph of $G((E_{11}+E_{1i}+E_{i1})/2)$, and $G((E_{11}+E_{1i}+E_{i1})/2)$ consists of the line segment $\{1, i\}$, a loop at the vertex 1, together with isolated vertices, so $G(B)$ must consist of isolated vertices together with either a loop at the vertex 1 or the line segment $\{1, i\}$. It follows that B equals E_{11} or $E_{1i} + E_{i1}$. This shows that $|\mathcal{E}(\Phi((E_{11} + E_{1i} + E_{i1})/2))| = 2$, hence $E_{1i} + E_{i1} \in N(E_{11})$. In a similar way, one can also show that each of the other elements in the set on the right side of our claim belongs to $N(E_{11})$.

To prove the reverse inclusion, let $O_n \neq B \in N(E_{11})$. First, consider the case when $G(B)$ has an edge e (possibly a loop) which is not incident with the vertex 1. Let H denote the graph on $\langle n \rangle$ consisting of the edge e and the loop at the vertex 1, together with other isolated vertices. Clearly, H is the graph of some $C \in \mathcal{E}(\mathbf{SDsS}(n))$. Note that the 1st row is the only possible row of $(E_{11} + B)/2$ with row sum equal to 1, that H is a subgraph of $G((E_{11} + B)/2)$, and also that the 1st row sum of C is equal to 1; hence $C \in \Phi((E_{11} + B)/2)$. Clearly, $C \neq E_{11}$. Since B and E_{11} are neighborly extreme points, this implies that we must have $C = B$. So, in this case, B equals $E_{11} + E_{jj}$ or $E_{11} + E_{ij} + E_{ji}$, where $i,\ j \in \langle n \rangle \backslash \{1\}$, $i \neq j$. In the remaining case, the edges of $G(B)$ are all incident with the vertex 1. Then, necessarily, $G(B)$ is a line segment containing the vertex 1, together with isolated vertices, and so B equals $E_{1i} + E_{i1}$ for some $i \in \langle n\rangle\backslash\{1\}$.

To prove the second equality of Assertion 6, it suffices to show the following:

$N(E_{12}+E_{21}) = \{O_n, E_{12}+E_{21}+E_{ii}, E_{12}+E_{21}+E_{ij}+E_{ji}, E_{11}, E_{22}, E_{1i}+E_{i1}, E_{2i}+E_{i2}, E_{11}+E_{22}, E_{1i}+E_{i1}+E_{22}, E_{2i}+E_{i2}+E_{11}, E_{1i}+E_{i1}+E_{2j}+E_{j2}, \frac{1}{2}(E_{12}+E_{21}+E_{1i}+E_{i1}+E_{2i}+E_{i2}) : i, j \in \langle n\rangle\backslash\{1,2\}, i \neq j\}$.

Case by case, one can show that each element in the set on the right side belongs to $N(E_{12}+E_{21})$.

To prove the reverse inclusion, let $O_n \neq B \in N(E_{12}+E_{21})$. First, consider the case when $G(B)$ has an edge e (possibly a loop) which is not adjacent to the edge $\{1,2\}$. Let H denote the graph on $\langle n\rangle$ which is composed of the edges e and $\{1,2\}$, together with isolated vertices. Clearly, $H = G(C)$ for some $C \in \mathcal{E}(\mathbf{SDsS}(n))$. Also, $C \neq E_{12}+E_{21}$ and H is a subgraph of $G((E_{12}+E_{21}+B)/2)$. Note that the first and the second row are the only possible rows of $(E_{12}+E_{21}+B)/2$ with row sum equal to 1, and also that the first and the second row sums of C are both equal to 1. Hence, $C \in \mathcal{E}(\Phi((E_{12}+E_{21}+B)/2)$. But $E_{12}+E_{21}$, B are neighborly extreme points, so we must have $C = B$. So, in this case, B must be of the form $E_{12}+E_{21}+E_{ii}$ or $E_{12}+E_{21}+E_{ij}+E_{ji}$, where $i,\ j \in \langle n\rangle\backslash\{1,2\}$, $i \neq j$. In the remaining case, all edges of $G(B)$ are adjacent to the edge $\{1,2\}$. Then $G(B)$ must consist of isolated vertices together with one of the following: one loop at the vertex 1 or 2; a line segment of the form $\{1,i\}$ or $\{2,i\}$, where $i \in \langle n\rangle\backslash\{1,2\}$; two loops, one at each of the vertices 1 and 2; a line segment of the form $\{j_1,i\}$ together with a loop at the vertex j_2, where $i \in \langle n\rangle\backslash\{1,2\}$ and the sets $\{j_1,j_2\}$, $\{1,2\}$ are equal; two line segments of the form $\{1,i\}$, $\{2,j\}$, where $i,j \in \langle n\rangle\backslash\{1,2\}$, $i \neq j$; a 3-cycle of the form $1 \to 2 \to i \to 1$, where $i \in \langle n\rangle\backslash\{1,2\}$. So, in this case, B must be one of the following: $E_{11}, E_{22}, E_{1i}+E_{i1}, E_{2i}+E_{i2}, E_{11}+E_{22}, E_{1i}+E_{i1}+E_{22}, E_{2i}+E_{i2}+E_{11}, E_{1i}+E_{i1}+E_{2j}+E_{j2}, \frac{1}{2}(E_{12}+E_{21}+E_{1i}+E_{i1}+E_{2i}+E_{i2})$, where $i,\ j \in \langle n\rangle\backslash\{1,2\}$, $i \neq j$.

Assertion 7. *For $n \geq 2$, T maps the sets $\{E_{ii} : i \in \langle n\rangle\}$ and $\{E_{ij}+E_{ji} : i,\ j \in \langle n\rangle, i \neq j\}$ each onto themselves.*

Proof of Assertion 7. By Assertion 6, the elements of $\{E_{ii} : i \in \langle n\rangle\}$ have the same number of neighborly extreme points, namely $\frac{n^2}{2} + \frac{n}{2}$, and the elements of $\{E_{ij}+E_{ji} : i,j \in \langle n\rangle,\ i \neq j\}$ also have the same number of neighborly extreme points, namely, $\frac{3}{2}n^2 - \frac{3}{2}n + 1$. As can be readily checked, $\frac{n^2}{2} + \frac{n}{2} = \frac{3}{2}n^2 - \frac{3}{2}n + 1$ if and only if $n = 1$. So, for $n \geq 2$, T cannot map some E_{ii} to some $E_{rs}+E_{sr}$ $(r \neq s)$ or conversely. Now, by Assertion

5, T maps $\mathcal{B}$ onto itself. But $\mathcal{B}$ is the union of the sets $\{E_{ii} : i \in \langle n \rangle\}$ and $\{E_{ij}+E_{ji} : i,j \in \langle n \rangle,\ i \neq j\}$, so T must map these sets each onto themselves.

Clearly, our theorem is true for the case $n = 1$. So consider $n \geq 2$. By Assertion 7, there exists a permutation $\sigma \in S_n$ such that $T(E_{ii}) = E_{\sigma(i)\sigma(i)}$ for each $i \in \langle n \rangle$. Consider any $i, j \in \langle n \rangle$, $i \neq j$. By Assertion 7 again, $T(E_{ij}+E_{ji})$ is of the form $E_{rs}+E_{sr}$ for some $r,\ s \in \langle n \rangle$, $r \neq s$. By the claim (but with $N(E_{ii})$ in place of $N(E_{11})$) given in the proof of the first equality of Assertion 6, we have, $E_{ij}+E_{ji} \in N(E_{ii})$; so $T(E_{ij}+E_{ji}) \in N(E_{\sigma(i)\sigma(i)})$ and hence, by the claim again, $T(E_{ij}+E_{ji})$ must be of the form $E_{\sigma(i)r}+E_{r\sigma(i)}$ for some $r \in \langle n \rangle \backslash \{\sigma(i)\}$. Similarly, from $E_{ij}+E_{ji} \in N(E_{jj})$, we also infer that $T(E_{ij}+E_{ji})$ is of the form $E_{\sigma(j)s}+E_{s\sigma(j)}$ for some $s \in \langle n \rangle \backslash \{\sigma(j)\}$. Hence, we must have $T(E_{ij}+E_{ji}) = E_{\sigma(i)\sigma(j)}+E_{\sigma(j)\sigma(i)}$. But $\mathcal{B}$ forms a basis for the space of $n \times n$ real symmetric matrices, it follows that T is given by $T(X) = P^t X P$, where P is the permutation matrix $[e_{\sigma(1)} \cdots e_{\sigma(n)}]$. The proof is complete. ■

In the introductory section, we have pointed out that Theorem 2 is still valid if “linear map” is replaced by “affine map”, and also mentioned the reason that every strong affine preserver of $\mathbf{SDsS}(n)$ necessarily fixes O_n and hence is linear. Now we elaborate. First, note that every strong affine preserver T of a polytope C shares with a strong linear preserver the properties that $T(\mathcal{E}(C)) = \mathcal{E}(C)$, $T(\Phi(x)) = \Phi(Tx)$ and $|\mathcal{E}(\Phi(x))| = |\mathcal{E}(\Phi(Tx))|$ for all $x \in C$. But by Lemma 3 (and its proof), we readily see that for any $A \in \mathcal{E}(\mathbf{SDsS}(n))$, we have, $A = O_n$ if and only if $\Phi(A/2)$ contains exactly one extreme element; so every strong affine preserver of $\mathbf{SDsS}(n)$ fixes O_n.

References

[B–P] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Revised reprint of the 1979 original, Classics in Applied Mathematics, **9**, SIAM, Philadelphia, 1994.

[C–L1] H. Chiang and C.K. Li, Linear maps leaving the alternating group invariant, *Linear Algebra Appl.* **340** (2002), 69–80.

[C–L2] H. Chiang and C.K. Li, Linear maps leaving invariant subsets of nonnegative symmetric matrices, preprint.

[L–T–T] C.K. Li, B.S. Tam and N.K. Tsing, Linear maps preserving permutation and stochastic matrices, *Linear Algebra Appl.* **341** (2002), 5–22.

[K1] M. Katz, On the extreme points of a certain convex polytope, *J. Combinatorial Theory* **8** (1970), 417–423.

[K2] M. Katz, On the extreme points of a set of substochastic and symmetric matrices, *J. Math. Anal. Appl.* **37** (1972), 576–579.

[R] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

出席國際會議報告

會議名稱: 第八屆 SIAM 應用線性代數會議

會議地點: 美國威廉斯堡威廉瑪琍大學

會議時間: 92年7月16日–19日

報　告　人: 淡江大學數學系　譚必信

撰寫日期: 92年8月12日

《SIAM 應用線性代數會議》每三年舉辦一次, 今年是第八屆, 輪到在美國威廉瑪琍大學舉行, 參加者來自世界各地, 共約243人。

在四個全天的緊湊議程中總共安排了211場演講, 其中包括10場一小時的大會演講, 89場半小時的迷你會議演講及112個15分鐘的分組報告。會議的特色之一爲爲迷你會議特別多, 總共26個。粗略統計發現約一半迷你會議屬數值線性代數, 三成爲線性代數的應用, 剩下兩成則爲核心 (純) 線性代數。數值線性代數的主題包括: 大型高速計算 (涉及線性系統、特徵值、奇異值、偏微分方程、積分方程等)、多項式及整數矩陣的準確計算、結構矩陣、稀疏矩陣及線性代數的組合學等等。線性代數的應用主題包括動力系統、計算生物醫學、影像回復、資料整理及回饋等。核心線性代數方面則含蓋不定內積、組合線性代數、矩陣不等式、逆特徵值問題及矩陣完成問題等。

這屆的 SIAM Activity Group on Linear Algebra Prize 是由肯刹斯州大學的 Karen S. Braman 及 Ralph Byers 與威廉瑪琍大學的 Roy C. Mathias 共同奪得。他們的得獎論文爲“The multishift QR algorithm, Part II: early deflation", 該論文已於2001年在 SIAM Matrix Analysis and Applications 刊登。得獎主之一的 Karen S. Braman 剛完成她的博士學位, 眞是前途無量。

這屆的 ILAS speakers (由國際線性代數學會支助) 爲華盛頓大學的 Judith J. McDonald 及懷奧明大學的 Bryan L. Shaders。他們的大會演講題目分別爲"Combinatorial Matrix Theory"及 "Non-negative matrix pairs, 2-D dynamical systems and road-colorings"。

除了大會報告以外, 會議的演講都是分四場平行進行。本人主要選擇聆聽主題跟自己興趣相關的迷你會議及分組報告, 其中包括: indefinite inner product and applications、inverse eigenvalue problems、matrix inequalities and applications、matrix completion problems 等等。在這次會議我聽到的精彩 (或感興趣的) 演講有10多個。特別欣賞的包括: R. Bhatia 的 "Convexity of some matrix functions"、C.K. Johnson 的 "Matrix completion problems"、Chandler Davis 的 "Explicit computation of some polynomial hulls of matrices"、Michael Neumann 的 "On functions that preserve M-matrices and inverse M-matrices"、Leslie Hogben 的 "Relationship between the completion problems for various classes of matrices" 及 Chi-Kwong Li 的 "Unitarily invariant metrics on subspaces of $\mathbb{C}^n$" 等等。

本人的演講是安排在會議第二天晚上的分組 CP8, 講題爲"Equilibria of pairs of nonlinear maps associated with cones", 是報告本人最近與 H. Schneider、G.P. Barker、M. Takane 及 M. Neumann-Coto 等人合作在非線性 Perron-Frobenius 理論所做的工作。我的演講吸引了不少聽衆。

在開會期間某天用餐時我無意中認識了現職於賓州大學的查宏遠教授, 獲悉他有興趣做非負矩陣的近似非負秩分解, 因爲這方面的結果在

影像處理問題上有重要的實用價值, 交談間我才記起二十多年前我曾獲得非負矩陣擁有非負秩分解的一個幾何等價條件。說不定將來我們在這方面可以合作做些工作。

會議結束後, 我飛往亞特蘭大再轉乘車至美國阿拉巴馬州的奧本市, 訪問奧本大學的譚天祐教授, 與他討論了一些涉及李代數 (或李群) 及錐體 (或數值域) 的數學問題。8月9日我起程回國。

攜回資料: 會議議程及摘要一本。

Equilibria of Pairs of Nonlinear Maps Associated with Cones

George Phillip Barker[1], Max Neumann-Coto[2],
Hans Schneider[3,*], Martha Takane[4,*], Bit-Shun Tam[5,†]

April 10, 2003

[1]Department of Mathematics, University of Missouri-Kansas City, Kansas City, MO 64110-2499, U.S.A; [2,4]Instituto de Matemáticas, UNAM, Área de la Investigación Científica, Circuito Exterior, C.U., 04510 México, D.F., E-Mail: max@matem.unam.mx, takane@matem.unam.mx; [3]Department of Mathematics, University of Wisconsin, Madison, WI 53706, U.S.A, E-Mail: hans@math.wisc.edu; [5]Department of Mathematics, Tamkang University, Tamsui, Taiwan 251, R.O.C, E-Mail: bsm01@mail.tku.edu.tw

Let K_1, K_2 be closed, full, pointed convex cones in finite-dimensional real vector spaces of the same dimension, and let $F : K_1 \to \operatorname{span} K_2$ be a homogeneous, continuous, K_2-convex map that satisfies $F(\partial K_1) \cap \operatorname{int} K_2 = \emptyset$ and $FK_1 \cap \operatorname{int} K_2 \neq \emptyset$. Using an equivalent formulation of the Borsuk-Ulam theorem in algebraic topology, we show that we have $F(K_1 \backslash \{0\}) \cap (-K_2) = \emptyset$ and $K_2 \subseteq FK_1$. We also prove that if, in addition, $G : K_1 \to \operatorname{span} K_2$ is any homogeneous, continuous map which is (K_1, K_2)-positive and K_2-concave, then there exist a unique real scalar ω_0 and a (up to scalar multiples) unique nonzero vector $x_0 \in K_1$ such that $Gx_0 = \omega_0 F x_0$, and moreover we have $\omega_0 > 0$ and $x_0 \in \operatorname{int} K_1$ and we also have a characterization of the scalar ω_0. Then, we reformulate the above result in the setting when K_1 is replaced by a compact convex set and recapture a classical result of Ky

[*]Partially supported by UNAM-DGAPA IN102399
[†]Supported by the National Science Council of the Republic of China

Fan on the equilibrium value of a finite system of convex and concave functions.

AMS Classification: 47H15, 47B55, 15A48, 65J05.
Keywords: Proper cone, compact convex set, homogeneous map, convex map, concave map, equilibrium point, Ky Fan's theorem, Borsuk-Ulam theorem, Perron-Frobenius.

1. Introduction

In this paper we prove equilibrium theorems of Perron-Frobenius type for a pair of nonlinear maps F and G from a proper cone K_1 in a finite dimensional real space to another finite dimensional real space ordered by another proper cone K_2; namely, we determine conditions under which there is a unique positive scalar ω_0 and a unique fixed vector x_0 (up to scalar multiples) in K_1 such that $Gx_0 = \omega_0 Fx_0$, see Theorem 3. We also show that ω_0 can be obtained as infimum or supremum of analogs of the Collatz-Wielandt sets further discussed in our last section. In Theorem 4 we derive a version of our equilibrium theorem with a compact convex set as the domain space.

Our motivation is [F, Theorem 1] due to Ky Fan on the equilibrium value of a finite system of convex and concave functions which we state at the beginning of next section. However, we do not use this theorem in deriving our main results, Theorems 3 and 4, which may be considered as its extensions. Instead, we use the Borsuk-Ulam theorem to establish a geometric result about a nonlinear map (see Theorem 1) and then use it to deduce our main results. Ky Fan's theorem can be recovered from our extension by means of Sperner's Lemma [Spe].

Our paper continues a long tradition of generalizations of the Perron-Frobenius theorem. While the setting of our work is strictly finite dimensional (which is natural in view of our use of the Borsuk-Ulam theorem and Sperner's Lemma), many generalizations are to operators in a Banach space which leave a cone invariant. We point to recent linear and nonlinear generalizations in [N1], [N2] and [NVL], and to the recent surveys [Do],[Z], [T] and books [A2], [KLS] and [HIR] for different aspects of the theory and many further references.

2. Statements of Main Results

In [F, Theorem 1] Ky Fan obtained the following result discussed in our introduction.

Ky Fan's Theorem. *Let S denote the standard $(n-1)$-simplex of $\mathbb{R}^n$, i.e., $S = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n_+ : \sum_{j=1}^n \xi_j = 1\}$, and let $S_i = \{(\xi_1, \ldots, \xi_n) \in S : \xi_i = 0\}$ for $i = 1, \ldots, n$. For $i = 1, \ldots, n$, also let $f_1, \ldots, f_n$, $g_1, \ldots, g_n$ be $2n$ real-valued functions defined on S that satisfy the following*:

(a) *Each f_i is continuous and convex on S*;

(b) *$f_i(x) \leq 0$ for each $x \in S_i$*;

(c) *For each $x \in S$ there is an index i for which $f_i(x) > 0$*; *and*

(d) *Each g_i is continuous, concave and positive on S.*

Then there exist a unique real number λ and a unique point $\hat{x} \in S$ such that for every i, $g_i(\hat{x}) = \lambda f_i(\hat{x})$. Moreover, we have $\lambda > 0$, $\hat{x}$ has positive components, and

$$\frac{1}{\lambda} = \min_{x \in S} \max_{1 \leq i \leq n} \frac{f_i(x)}{g_i(x)} = \max_{x \in S} \min_{1 \leq i \leq n} \frac{f_i(x)}{g_i(x)}.$$

Notice that under the hypotheses of Ky Fan's theorem, if we define a map $f : S \to \mathbb{R}^n$ by $f(x) = (f_1(x), \ldots, f_n(x))$, then f is a convex map in the sense that, for any scalar λ, $0 < \lambda < 1$, and x, $y \in S$, we have $f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x)+\lambda f(y)$, where the ordering is componentwise. Similarly, if we define $g : S \to \mathbb{R}^n$ by $g(x) = (g_1(x), \ldots, g_n(x))$, then g is a concave map (i.e., $-g$ is a convex map). The conclusion of Ky Fan's theorem can now be restated as: $g(\hat{x}) = \lambda f(\hat{x})$ for some real number λ and $\hat{x} \in S$. In this case, we say that λ is an *equilibrium value* and $\hat{x}$ is an *equilibrium point* for the system (g, f). The concepts of equilibrium value and equilibrium point come from economic models (see, for instance, [A2]).

As already noted in [F], if $A = (a_{ij})$ is an $n \times n$ (entrywise) positive matrix, and if we define f_i, g_i $(1 \leq i \leq n)$ on S by $f_i(x) = \xi_i$ and $g_i(x) = \sum_{i=1}^n a_{ij}\xi_j$ for $x = (\xi_1, \ldots, \xi_n) \in S$, then conditions (a), (b), (c) and (d) of Ky Fan's theorem are satisfied. In this case, the first part of Ky Fan's theorem becomes the classical Perron's theorem on positive matrices (with λ being the spectral radius and $\hat{x}$ the Perron vector of A). The last part of Ky Fan's theorem becomes Wielandt's extremal characterization of the spectral radius.

In Aubin [A1] one may find extensions or variants of Ky Fan's theorem in the setting of a pair of multi-valued maps. In [Sim, Theorem 4.1] Simons generalized the first part of Ky Fan's theorem in such a way that the finite systems of functions are replaced by two (single-valued) maps whose common range space is a real vector space with a given sublinear (i.e., positively homogeneous, convex) function, referred to as a sublineared space, and which are convex or concave in a certain generalized sense defined with respect to the sublinear structure, and moreover the domain space is not restricted to an $(n-1)$-simplex. In fact, Simons obtained first a continuity result about a pair of multi-valued maps that involve a sublineared space and used it to deduce the aforementioned result and also to obtain a result that generalizes [A1, Theorem 2], and hence the last part of Ky Fan's theorem, in the setting of a pair of multi-valued maps. In this paper, we give a generalization in a different direction. We first examine conditions (a)–(c) of Ky Fan's theorem in the setting of a homogeneous map on a proper cone.

We call a nonempty subset K in a finite-dimensional real vector space V a *proper cone* of V if K is a convex cone (i.e. $\alpha K + \beta K \subseteq K$ for all $\alpha, \beta \geq 0$), which is pointed (i.e. $K \cap (-K) = \{0\}$), closed (with respect to the usual topology of V) and has nonempty interior (or equivalently, $\operatorname{span} K$, the linear span of K, is V). We use $\geq^K$ to denote the partial ordering on $\operatorname{span} K$ induced by the proper cone K, i.e. $x \geq^K y$ if and only if $x - y \in K$. For convenience, we also adopt the following notation:

$$\begin{aligned} & x >^K 0 \quad \text{if and only if } x \geq^K 0 \text{ and } x \neq 0, \\ \text{and} \quad & x \gg^K 0 \quad \text{if and only if } x \in \operatorname{int} K. \end{aligned}$$

Sometimes we also use $\geq$, $>$ and $\gg$ in place of $\geq^K$, $>^K$ and $\gg^K$, when there is no danger of confusion.

We obtain the following result:

Theorem A. *Let $K_1,\ K_2$ be proper cones. Let $F : K_1 \to \operatorname{span} K_2$ be a homogeneous map that satisfies each of the following conditions*:

(a) *For any $x,\ y \in K_1$, there exist positive constants $\alpha,\ \beta$ (depending on x and y) such that $\alpha Fx + \beta Fy \geq^{K_2} F(x+y)$;*

(b) *$F(\partial K_1) \cap \operatorname{int} K_2 = \emptyset$; and*

(c) *$FK_1 \cap \operatorname{int} K_2 \neq \emptyset$.*

Then $F(K_1 \backslash \{0\}) \cap (-K_2) = \emptyset$. If, in addition, $\dim K_1 = \dim K_2$ and F is continuous, then $K_2 \subseteq FK_1$.

Here we use $\operatorname{int} S$ (respectively, ∂S) to denote the interior (respectively, boundary) of S. A map $T : D \subseteq V_1 \to V_2$, where $V_1,\ V_2$ are real vector spaces and D satisfies $\lambda D \subseteq D$ for all $\lambda > 0$, is said to be *homogeneous* (of degree one) if $T(\lambda x) = \lambda Tx$ for all $\lambda > 0$ and $x \in D$.

To avoid trivialities, we assume that the cones $K_1,\ K_2$ considered in Theorem 1 are nonzero. The same remark also applies (sometimes to K) in the remaining parts of the paper.

Note that, when $K_1 = K_2 = K$, condition (b) of Theorem 1 is weaker than the following natural extension of condition (b) of Ky Fan's theorem: For any $x \in \partial K$, $p \in \partial K^*$, where K^* denotes the dual cone of K, we have $p(Fx) \leq 0$ whenever $p(x) = 0$.

The proof of Theorem 1 relies on the use of an equivalent formulation of the Borsuk-Ulam theorem in algebraic topology. A modification of the argument used in the proof also leads to the following unexpected side-product:

Theorem B. *Let K_1, K_2 be proper cones such that $\dim K_1 > \dim K_2$. Let $F : K_1 \rightarrow \operatorname{span} K_2$ be a homogeneous, continuous map with the property that for any x, $y \in K_1$, there exist α, $\beta > 0$ such that $\alpha Fx + \beta Fy \geq^{K_2} F(x+y)$. If $FK_1 \cap \operatorname{int} K_2 \neq \emptyset$, then $F(\partial K_1) \cap \operatorname{int} K_2 \neq \emptyset$ and moreover we have either $F(K_1 \backslash \{0\}) \cap (-K_2) \neq \emptyset$ or $\operatorname{int} K_2 \cap F(\operatorname{int} K_1) \subseteq F(\partial K_1)$.*

Theorems 1 and 2 can be restated as results about solvability of nonlinear systems.

Using Theorem 1, we derive the following result which extends Ky Fan's theorem and also [F, Corollaries 1 and 2] in the setting of homogeneous maps on proper cones.

A map $F\colon K_1 \rightarrow \operatorname{span} K_2$ is said to be K_2-*convex* (respectively, K_2-concave) if $(1-\lambda)Fx + \lambda Fy \geq^{K_2} F((1-\lambda)x + \lambda y)$ (respectively, $(1-\lambda)Fx + \lambda Fy \;{}^{K_2}\!\!\leq F((1-\lambda)x + \lambda y)$) for all real scalar λ, $0 < \lambda < 1$, and x, $y \in K_1$; F is (K_1, K_2)-*nonnegative* (respectively, (K_1, K_2)-*positive*) if $FK_1 \subseteq K_2$ (respectively, $F(K_1 \backslash \{0\}) \subseteq \operatorname{int} K_2$); F is (K_1, K_2)-*monotone* (or, *order-preserving*, according to some authors) if $y \geq^{K_1} x$ implies $Fy \geq^{K_2} Fx$. Clearly, if F is homogeneous, K_2-convex, then F possesses the property that for any x, $y \in K_1$, there exist α, $\beta > 0$ such that $\alpha F(x) + \beta F(y) \geq^{K_2} F(x+y)$.

Theorem C. *Let K_1, K_2 be proper cones such that $\dim K_1 = \dim K_2$. Let $F : K_1 \rightarrow \operatorname{span} K_2$ be a homogeneous, continuous map that satisfies each of the following conditions*:

(a) *F is K_2-convex*;

(b) $F(\partial K_1) \cap \operatorname{int} K_2 = \emptyset$; *and*

(c) $FK_1 \cap \operatorname{int} K_2 \neq \emptyset$.

Then, for any homogeneous, continuous, K_2-concave and (K_1, K_2)-positive map $G : K_1 \rightarrow \operatorname{span} K_2$, there exist a unique scalar ω_0 and a (up to scalar multiples) unique nonzero vector x_0 of K_1 such that $Gx_0 = \omega_0 Fx_0$. We have, $\omega_0 > 0$, $x_0 \in \operatorname{int} K_1$ and $\sup \Omega = \inf \Sigma_1 = \omega_0$, where

$$\begin{array}{rcl} \Omega & = & \{\omega \geq 0 : \exists x >^{K_1} 0,\ Gx \geq^{K_2} \omega Fx\} \\ \textit{and} \quad \Sigma_1 & = & \{\sigma \geq 0 : \exists x \gg^{K_1} 0,\ Gx \;{}^{K_2}\!\!\leq \sigma Fx\}. \end{array}$$

Moreover, for any $x >^{K_1} 0$ and $\omega, \sigma > 0$, we have

$$\begin{aligned} &\omega < \omega_0 \text{ whenever } Gx \geq^{K_2} \omega Fx \text{ and } x \text{ is not a multiple of } x_0 \\ \text{and} \quad &\sigma > \omega_0 \text{ whenever } Gx \ {}^{K_2}\!\!\leq \sigma Fx \text{ and } x \text{ is not a multiple of } x_0. \end{aligned}$$

In Theorem 4 below we give a reformulation of Theorem 3 in the setting when the common domain K_1 of F and G is replaced by a compact convex set.

For a convex set C, we use $\operatorname{ri} C$ and $\operatorname{rbd} C$ to denote respectively the relative interior and the relative boundary of C. A map $g : C \to W$ from a convex set C to a real vector space W ordered by a proper cone K is said to be (C, K)-*nonnegative* (respectively, (C, K)-*positive*) if $g(C) \subseteq K$ (respectively, $g(C) \subseteq \operatorname{int} K$); K-convexity and K-concavity of g are defined in the same way as in the case when C is a proper cone.

Theorem D. *Let C be a compact convex set in a finite-dimensional real vector space, and let $f : C \to W$ be a continuous map from C to a finite-dimensional real vector space W ordered by a proper cone K such that* $\dim W = \dim C + 1$. *Suppose that f satisfies each of the following conditions*:

(a) *f is K-convex*;

(b) $f(\operatorname{rbd} C) \cap \operatorname{int} K = \emptyset$; *and*

(c) $f(C) \cap \operatorname{int} K \neq \emptyset$.

Then, for any continuous, K-concave and (C, K)-positive map $g : C \to W$, there exist a unique real scalar ω_0 and a unique point x_0 of C such that $g(x_0) = \omega_0 f(x_0)$. We have, $\omega_0 > 0$, $x_0 \in \operatorname{ri} C$ and $\sup \Omega = \inf \Sigma_1 = \omega_0$, where

$$\begin{aligned} \Omega &= \{\omega \geq 0 : \exists x \in C : g(x) \geq^{K} \omega f(x)\} \\ \text{and} \ \Sigma_1 &= \{\sigma \geq 0 : \exists x \in \operatorname{ri} C : g(x) \ {}^{K}\!\!\leq \sigma f(x)\}. \end{aligned}$$

Moreover, for any $x \in C$ and $\omega, \sigma > 0$, we have

$$\begin{aligned} &\omega < \omega_0 \quad \text{whenever } g(x) \geq^{K} \omega f(x) \text{ and } x \neq x_0 \\ \text{and} \quad &\sigma > \omega_0 \quad \text{whenever } g(x) \ {}^{K}\!\!\leq \sigma f(x) \text{ and } x \neq x_0. \end{aligned}$$

3. Nonlinear Solvability Theorems

In this section we shall prove Theorems 1, 2 and make relevant remarks and illustrative examples. Before we begin, we recall some facts from topology, which we shall need.

We shall identify finite-dimensional real vector spaces with euclidean spaces. Let B^n, S^{n-1} denote respectively the euclidean unit ball and unit sphere of $\mathbb{R}^n$.

For a proper cone K in $\mathbb{R}^n$, $n \geq 2$, we define a map π_K from the set $\{(z,v) : z \in \operatorname{int} K \cap S^{n-1},\ v \in S^{n-1},\ v \neq z, -z\}$ to $\partial K \cap S^{n-1}$ as follows: Let $z, v \in S^{n-1}$ with $z \in \operatorname{int} K$ and $v \neq z, -z$. Then $\operatorname{span}\{z,v\} \cap S^{n-1}$ is a circle, and $\operatorname{span}\{z,v\} \cap K \cap S^{n-1}$ is a closed circular arc whose endpoints belong to opposite semicircles determined by z and $-z$ and constitute the set $\operatorname{span}\{z,v\} \cap (\partial K \cap S^{n-1})$. We denote by $\pi_K(z,v)$ the endpoint in the semicircle that contains v. Observe that the point $\pi_K(z,v)$ is uniquely determined by the property that it belongs to ∂K and can be expressed in the form $\frac{\lambda z + v}{\|\lambda z + v\|}$ for some $\lambda \in \mathbb{R}$. That π_K is a continuous map is probably known. We give a proof below, as we have not been able to find any suitable reference.

Assume to the contrary that π_K is not continuous at (z,v) for some $z \in \operatorname{int} K \cap S^{n-1}$ and $v \in S^{n-1}$, $v \neq z, -z$. Then there exist a sequence $(z_k)_{k\in\mathbb{N}}$ in $\operatorname{int} K \cap S^{n-1}$ converging to z and a sequence $(v_k)_{k\in\mathbb{N}}$ in S^{n-1} converging to v such that, for some fixed $\delta > 0$, we have $\|\pi_K(z_k, v_k) - \pi_K(z,v)\| \geq \delta$ for all k. Now, for each k, we have, $\pi_K(z_k, v_k) = \frac{\lambda_k z_k + v_k}{\|\lambda_k z_k + v_k\|}$ for some real scalar λ_k. Note that the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is bounded; otherwise, $(z_k + \lambda_k^{-1} v_k)_{k\in\mathbb{N}}$ is a sequence in ∂K with a subsequence converging to z, which is a contradiction, as $z \in \operatorname{int} K$. Replacing by a subsequence, if necessary, we may assume that $(\lambda_k)_{k\in\mathbb{N}}$ converges to λ. Then we have $\lim_{k\to\infty} \pi_K(z_k, v_k) = \frac{\lambda z + v}{\|\lambda z + v\|}$. But $\lim_{k\to\infty} \pi_K(z_k, v_k)$ belongs to ∂K, so it is, in fact, equal to $\pi_K(z,v)$, which is a contradiction.

If $z \in \operatorname{int} K$, and x, $\bar{x} \in \partial K \cap S^{n-1}$ are such that z can be expressed as a linear combination of x and $\bar{x}$ with positive coefficients, then we say that x and $\bar{x}$ form *a pair of antipodal points of $\partial K \cap S^{n-1}$ relative to z*. Notice that the map $\pi_K(z, \cdot)$ takes each pair of antipodal points of the sphere $(\operatorname{span}\{z\})^{\perp} \cap S^{n-1}$ (which can be identified with S^{n-2}) to a pair of antipodal points of $\partial K \cap S^{n-1}$ relative to z.

Recall that two continuous maps f_0, $f_1 : X \to Y$ between topological spaces X, Y are said to be *homotopic* if one can be deformed continuously to the other, i.e., f_0 and f_1 belong to a family of continuous maps $f_t : X \to Y$, $t \in [0,1]$, so that $\Phi : X \times [0,1] \to Y$ given by $\Phi(x,t) = f_t(x)$ is continuous.

We shall make use of the following known results from algebraic topology:

Lemma A *A continuous map $f : S^{n-1} \to Y$, where Y is a topological space and $n \geq 1$, is homotopic to a constant map if and only if f can be extended to a continuous map from B^n to Y.*

Theorem A *If $f : S^n \to S^n$, $n \geq 0$, is a continuous map which is homotopic to a constant map, then there exists $x \in S^n$ such that $f(x) = f(-x)$.*

Corollary A *If $f : S^n \to S^m$, where $0 \leq m < n$, is a continuous map, then there exists $x \in S^n$ such that $f(x) = f(-x)$.*

Lemma A is elementary and can be found in many textbooks of topology; see, for instance, [Du, p.316, **1.2**(2)]. Theorem A is equivalent to the Borsuk-Ulam theorem, which asserts that every continuous map $f : S^n \to \mathbb{R}^n$, $n \geq 1$, sends at least one pair of antipodal points to the same points, and, in fact, equivalent to them are also several other geometric results about the n-sphere, such as the Borsuk antipodal theorem, the Lusternik-Schnirelman-Borsuk theorem, etc. (see, for instance, [DG, Theorems 5.2 and 6.1]). Corollary A can be deduced from Theorem A as follows: If $m < n$, we may regard S^m as lying in the equator of S^n and consider the map $\hat{f} : S^n \to S^n$ which is obtained from f by enlarging its range space to S^n. Since the image set $\hat{f}(S^n)$ is included in the upper hemisphere S^{n+} and S^{n+}, being homeomorphic to B^n, is a contractible space (i.e., one whose identity map is homotopic to a constant map), the map $\hat{f}$ is homotopic to a constant map. By Theorem A, it follows that there exists a pair of antipodal points of S^n with the same image under $\hat{f}$. Since $f(x) = \hat{f}(x)$ for all $x \in S^n$, we also have two antipodal points with the same image under f.

Proof of Theorem 1. Assume to the contrary that there exists $x > 0$ such that $Fx \in -K_2$. By conditions (c) and (b), there exists $u \gg 0$ such that $Fu \gg 0$. Since $u \in \operatorname{int} K_1$ and $-x \notin K_1$, there exists $\varepsilon > 0$ such that $u - \varepsilon x \in \partial K_1$. By the homogeneity of F and condition (a), we have

$$0 \ll Fu \leq \alpha F(u - \varepsilon x) + \beta \varepsilon Fx$$

for some $\alpha,\ \beta > 0$. Thus, $F(u - \varepsilon x) \geq \alpha^{-1} Fu - \alpha^{-1} \beta \varepsilon Fx \gg 0$, as $-Fx \geq 0$ and $Fu \gg 0$. This contradicts condition (b).

Now suppose, in addition, that F is continuous and K_1, K_2 have the same dimension. There is no loss of generality in assuming that $\mathbb{R}^n = \operatorname{span} K_1 = \operatorname{span} K_2$. The case $n = 1$ is trivial. Hereafter, we assume that $n \geq 2$. Let $f : K_1 \cap S^{n-1} \to S^{n-1}$ be the map given by: $f(x) = Fx/\|Fx\|$, where $\|\cdot\|$ denotes the euclidean norm of $\mathbb{R}^n$. Note that f is well-defined, as $Fx \neq 0$ for all $x \in K_1 \backslash \{0\}$, and is also continuous. Since F is homogeneous, it suffices to show that $K_2 \cap S^{n-1} \subseteq f(K_1 \cap S^{n-1})$.

Assume to the contrary that there exists $y \in K_2 \cap S^{n-1}$ such that $y \notin f(K_1 \cap S^{n-1})$. Since the set $f(K_1 \cap S^{n-1})$ is compact and hence closed, we may choose y so that $y \in \operatorname{int} K_2$. Let $\theta_y : \partial K_1 \cap S^{n-1} \to \partial K_2 \cap S^{n-1}$ be the map defined by: $\theta_y(v) = \pi_{K_2}(y, fv)$, where $\pi_{K_2} : \{(z,v) : z \in \operatorname{int} K_2 \cap S^{n-1},\ v \in S^{n-1},\ v \neq z, -z\} \to \partial K_2 \cap S^{n-1}$ is the continuous map that we have introduced at the beginning of this section. Since $y, -y \notin f(K_1 \cap S^{n-1})$, θ_y is a well-defined map. Indeed, for the same reason, we can extend the domain of θ_y to $K_1 \cap S^{n-1}$, using the same formula for definition. Of course, θ_y and its extension are continuous maps. But there is a homeomorphism from $K_1 \cap S^{n-1}$ onto B^{n-1} which takes $\partial K_1 \cap S^{n-1}$ onto S^{n-2}, so by Lemma A, it follows that the map θ_y is homotopic to a constant map. Now we are going to obtain another map from $\partial K_1 \cap S^{n-1}$ to $\partial K_2 \cap S^{n-1}$, which is homotopic to θ_y, as follows. By conditions (c) and (b), there exists a vector $u \in \operatorname{int} K_1 \cap S^{n-1}$ such that $fu \in \operatorname{int} K_2 \cap S^{n-1}$. Denote fu by z and define the desired map θ_z by $\theta_z(v) = \pi_{K_2}(z, fv)$. Clearly, θ_z is well-defined and continuous. Moreover, the continuous map $\Phi : (\partial K_1 \cap S^{n-1}) \times [0,1] \to \partial K_2 \cap S^{n-1}$ given by $\Phi(v,t) = \pi_{K_2}(y(t), fv)$, where $y(t) = \frac{(1-t)y+tz}{\|(1-ty)+tz\|}$, establishes a homotopy of θ_y to θ_z. Since θ_y is homotopic to a constant map, so is θ_z. On the other hand, the continuous map $\pi_{K_1}(u, \cdot)$ takes the compact set $(\operatorname{span}\{u\})^{\perp} \cap S^{n-1}$ one-to-one, and hence homeomorphically, onto $\partial K_1 \cap S^{n-1}$ and moreover it sends each pair of antipodal points of the sphere $(\operatorname{span}\{u\})^{\perp} \cap S^{n-1}$ (which can be identified with S^{n-2}) to a pair of antipodal points of $\partial K_1 \cap S^{n-1}$ relative to u. Also, $\partial K_2 \cap S^{n-1}$ is homeomorphic with S^{n-2}. In view of Theorem A, it follows that there exists a pair of antipodal points x, $\bar{x}$ of $\partial K_1 \cap S^{n-1}$ relative to u such that $\theta_z(x) = \theta_z(\bar{x})$. The fact that x, $\bar{x}$ are antipodes clearly implies that there exist $\nu,\ \eta > 0$ such that $u = \nu x + \eta \bar{x}$. By the homogeneity of F and condition (a), we have

$$\alpha\nu F(x) + \beta\eta F(\bar{x}) \geq F(\nu x + \eta \bar{x}) = F(u) \gg 0$$

for some $\alpha,\ \beta > 0$. On the other hand, the condition $\theta_z(x) = \theta_z(\bar{x})$, which amounts to $\pi_{K_2}(z, fx) = \pi_{K_2}(z, f\bar{x})$, together with the fact that $Fx, F\bar{x} \notin \operatorname{int} K_2$, clearly implies that $\lambda F(x) + \mu F(\bar{x}) \notin \operatorname{int} K_2$ for any $\lambda,\ \mu > 0$. So we arrive at a contradiction. ■

It can be readily checked that in Theorem 1 if we assume that F is homogeneous of degree p, where p is a positive number possibly different from 1, then the result is still valid.

The following examples illustrate the irredundancy of condition (a) of Theorem 1.

Example 1. Let K be the proper convex cone in $\mathbb{R}^2$ given by:

$$K = \{\lambda(\cos\theta, \sin\theta) : \lambda \geq 0, -\pi/4 \leq \theta \leq \pi/4\},$$

and let $F\colon K \to \mathbb{R}^2$ be the map defined by: $F(\lambda(\cos\theta, \sin\theta)) = \lambda(\cos 3\theta, \sin 3\theta)$. Then F is homogeneous, continuous and satisfies conditions (b) and (c) of Theorem 1 (with $K_1 = K_2 = K$). However, $F(K\backslash\{0\}) \cap (-K) \neq \emptyset$, as $F(1, 1) = (1, -1) \in -K$. (But we do have $FK \supseteq K$ in this case.)

Example 2. Let g be any real-valued concave continuous function defined on the closed interval [0,1] such that $g(0) = g(1) = 0$ and $g(t) > 0$ for all $t \in (0, 1)$. Let $F\colon \mathbb{R}^2_+ \to \mathbb{R}^2$ be the homogeneous map determined by: $F(1 - t, t) = g(t)(\frac{1}{2}, \frac{1}{2})$ for all $t \in [0, 1]$. Then F is continuous, $\mathbb{R}^2_+$-concave (but not $\mathbb{R}^2_+$-convex). Also, conditions (b) and (c) of Theorem 1 are satisfied. However, we have $F(\mathbb{R}^2_+\backslash\{0\}) \cap (-\mathbb{R}^2_+) = \{0\} \neq \emptyset$ and $\mathbb{R}^2_+ \not\subseteq F\mathbb{R}^2_+$.

Example 3. Let $F : \mathbb{R}^2_+ \to \mathbb{R}^2$ be defined by: $F(\xi_1, \xi_2)$ equals (ξ_1, ξ_2) if $\xi_1 \geq \xi_2$ and equals (ξ_2, ξ_1) if $\xi_1 < \xi_2$. Then F is homogeneous, continuous and we have $F(\partial\mathbb{R}^2_+)\cap\operatorname{int}\mathbb{R}^2_+ = \emptyset$, $F\mathbb{R}^2_+\cap\operatorname{int}\mathbb{R}^2_+ \neq \emptyset$ and $F(\mathbb{R}^2_+\backslash\{0\})\cap(-\mathbb{R}^2_+) = \emptyset$. However, $\mathbb{R}^2_+ \not\subseteq F\mathbb{R}^2_+$. So, in Theorem 1, when F is continuous and $\dim K_1 = \dim K_2$, without condition (a), we cannot infer that $K_2 \subseteq FK_1$, even if we add as an extra assumption the condition that $F(K_1\backslash\{0\}) \cap (-K_2) = \emptyset$.

We would also like to point out that the last part of Theorem 1 is invalid if we assume $\dim K_1 < \dim K_2$ instead of the equality. Indeed, in this case, for any map $F : K_1 \to \operatorname{span} K_2$ which is linear (i.e., $F(\alpha x+\beta y) = \alpha Fx+\beta Fy$ for all $\alpha, \beta \geq 0$ and $x, y \in K_1$) and satisfies conditions (a)–(c) of Theorem 1 (for instance, take $K_2 = \mathbb{R}^3_+$, $K_1 = \operatorname{pos}\{(1, 0, 0), (0, 1, 1)\}$, where we use $\operatorname{pos} S$ to denote the positive hull of S, i.e., the set of all (finite) nonnegative linear combinations of vectors in S, and $F : K_1 \to \operatorname{span} K_2$ to be the canonical injection), it is impossible that the inclusion $K_2 \subseteq FK_1$ holds.

On the other hand, if we have $\dim K_1 > \dim K_2$, then we have Theorem 2 which, rather surprisingly, indicates that for a homogeneous, continuous map $F : K_1 \to \operatorname{span} K_2$ which satisfies condition (a) of Theorem 1, conditions (b) and (c) of Theorem 1 are incompatible !

Proof of Theorem 2. First, assume to the contrary that $F(\partial K_1)\cap\operatorname{int} K_2 = \emptyset$. As done in the proof for the first part of Theorem 1, we have $F(K_1\backslash\{0\})\cap(-K_2) = \emptyset$. Then we borrow part of the arguments used in the proof of the last part of Theorem 1, now assuming instead that $\operatorname{span} K_1 = \mathbb{R}^n$ and $\operatorname{span} K_2 = \mathbb{R}^m$. The continuous map $f : K_1 \cap S^{n-1} \to S^{m-1}$ can be defined in the same way as before, but we do not introduce the map θ_y. We do choose a vector u from $\operatorname{int} K_1 \cap S^{n-1}$ such that $fu \in \operatorname{int} K_2 \cap S^{m-1}$, denote fu by z and define the map $\theta_z : \partial K_1 \cap S^{n-1} \to \partial K_2 \cap S^{m-1}$ by $\theta_z(v) = \pi_{K_2}(z, fv)$. Note that z, $-z \notin f(\partial K_1 \cap S^{n-1})$; so θ_z is well-defined, continuous. Since the sets $\partial K_1 \cap S^{n-1}$ and $\partial K_2 \cap S^{m-1}$ are homeomorphic to S^{n-2} and S^{m-2} respectively and $m < n$ by our assumption, we can now apply Corollary A to conclude that there exists a pair of antipodal points x, $\bar{x}$ of $\partial K_1 \cap S^{n-1}$ relative to u such that $\theta_z(x) = \theta_z(\bar{x})$. Then we can derive a contradiction in the same way as before. So we must have $F(\partial K_1) \cap \operatorname{int} K_2 \neq \emptyset$.

To prove the second half of the theorem, suppose that $F(K_1 \backslash \{0\}) \cap (-K_2) = \emptyset$. Then the map f is well-defined. If, in addition, we have $\operatorname{int} K_2 \cap F(\operatorname{int} K_1) \not\subset F(\partial K_1)$, then we can choose a vector u from $\operatorname{int} K_1 \cap S^{n-1}$ such that $0 \ll f(u) \notin f(\partial K_1 \cap S^{n-1})$. Then we denote $f(u)$ by z, introduce the continuous map $\theta_z : \partial K_1 \cap S^{n-1} \rightarrow \partial K_2 \cap S^{m-1}$, and derive a contradiction in the same way as done above. ■

Below we give some "natural" conditions on a map $F : K_1 \rightarrow \operatorname{span} K_2$, which guarantee that F satisfies condition (a) of Theorem 1. The proof is straightforward.

A subset F of K is called a *face* of K if it is a convex cone and in addition possesses the property that $x \geq^K y \geq^K 0$ and $x \in F$ imply $y \in F$. For any nonempty subset S of a closed, pointed convex cone K, we denote by $\Phi(S)$ the *face of K generated by S*, i.e., the intersection of all faces of K that include S; equivalently, we have, $\Phi(S) = \{y \in K : y \overset{K}{\leq} \alpha x$ for some $\alpha > 0$ and $x \in \operatorname{pos} S\}$, where $\operatorname{pos} S$ denotes the positive hull (i.e., the set of all nonnegative linear combinations of vectors) of S. If $S = \{x\}$, where $x \in K$, we denote $\Phi(S)$ simply by $\Phi(x)$.

Remark 1. Consider the following conditions on a map T: $K_1 \rightarrow \operatorname{span} K_2$, where K_1, K_2 are proper cones in finite-dimensional real vector spaces.

(a) T is K_2-convex.

(b) For any $S \subseteq K_1$, $T(\Phi(S)) \subseteq \Phi(TS)$.

(c) For any $S \subseteq K_1$, $T(\operatorname{pos} S) \subseteq \Phi(TS)$.

(d) For any $x, y \in K_1$ and $\lambda, \mu > 0$, there exist $\alpha, \beta > 0$ such that $\alpha Tx + \beta Ty \geq^{K_2} T(\lambda x + \mu y)$.

(e) For any $x, y \in K_1$, there exist $\alpha, \beta > 0$ such that $\alpha Tx + \beta Ty \geq^{K_2} T(x + y)$.

Conditions (c) and (d) are equivalent, and we always have the implications (b) $\Longrightarrow$ (c) $\Longrightarrow$ (e) and (a) $\Longrightarrow$ (e). When T is homogeneous, (d) and (e) are also equivalent. When T is homogeneous and satisfies the condition that $T(\Phi(x)) \subseteq \Phi(Tx)$ for all $x \in K_1$ (which is the case if T is (K_1, K_2)-monotone), we also have (a) $\Longrightarrow$ (b).

4. Extensions of Ky Fan's Theorem

We need the following, parts of which are undoubtedly known:

Remark 1. Let $T : K_1 \to \operatorname{span} K_2$ be a homogeneous map.
(i) If T is (K_1, K_2)-monotone, then $T(0) = 0$ and T is (K_1, K_2)-nonnegative.
(ii) The following are equivalent statements:

(a) T is (K_1, K_2)-convex.

(b) For any $x,\ y \in K_1$, $Tx + Ty \geq^{K_2} T(x + y)$.

(c) $x \geq^{K_1} y \geq^{K_1} 0$ implies $T(x - y) \geq^{K_2} Tx - Ty$.

A similar assertion also holds for K_2-concavity.

(iii) If T is K_2-concave and (K_1, K_2)-nonnegative, then T is (K_1, K_2)-monotone.

(iv) If T is (K_1, K_2)-monotone, then T is bounded, in the sense that it maps bounded sets to bounded sets, or equivalently, there exists a positive constant M such that $\|Tx\|_2 \leq M\|x\|_1$ for all $x >^{K_1} 0$ and for some (and hence, for all) norms $\|\cdot\|_1$ and $\|\cdot\|_2$ of $\operatorname{span} K_1$ and $\operatorname{span} K_2$ respectively.

Notice that the (K_1, K_2)-monotonicity of T alone does not guarantee (K_1, K_2)-nonnegativity nor $T(0) = 0$. The point is, if T is a (K_1, K_2)-monotone map, then the map S defined by $Sx = Tx + y$, where y is any fixed vector of K_2, is still a (K_1, K_2)-monotone map. However, if T is homogeneous and (K_1, K_2)-monotone, then from $2.0 \geq 0$, we obtain $2T(0) \geq T(0)$ and hence $T(0) \geq 0$. On the other hand, from $\frac{1}{2}.0 \geq 0$, we also obtain $T(0) \leq 0$. Hence, we have, $T(0) = 0$, and then by the (K_1, K_2)-monotonicity of T, the (K_1, K_2)-nonnegativity of T follows. This proves part(i) of Remark 2.

Parts (ii) and (iii) of Remark 2 can be readily proved. To prove (iv), choose any vector $v \in \operatorname{int} K_1$. By definition of interior, there exists $\varepsilon > 0$ such that $v + \varepsilon x \in K_1$ for all $x \in V_1$ with $\|x\|_1 \leq 1$, where $\|\cdot\|_1$ is any norm of $\operatorname{span} K_1$. Now choose a norm $\|\cdot\|_2$ of $\operatorname{span} K_2$ which is monotonic with respect to K_2; that is, $0 \overset{K_2}{\leq} x \overset{K_2}{\leq} y$ implies $\|x\|_2 \leq \|y\|_2$. (For the existence of monotonic norms, see [BP, pp.5–6, Exercise 2.24].) Consider any vector $x \in K_1$ with $\|x\|_1 \leq 1$. Clearly, we have $v - \varepsilon x \in K_1$. Since T is homogeneous and (K_1, K_2)-monotone, we also have $0 \leq \varepsilon Tx \leq Tv$. By the monotonicity of $\|\cdot\|_2$, it follows that $\varepsilon\|Tx\|_2 \leq \|Tv\|_2$ and $\varepsilon^{-1}\|Tv\|_2$ is the desired constant for the boundedness of T.

Proof of Theorem 3. First, we show that the set Ω contains some positive elements. Take any $u >^{K_1} 0$. By the positivity of G, $Gu \gg 0$. So, for $\varepsilon > 0$

sufficiently small, we have $Gu - \varepsilon Fu \geq 0$, i.e., $\varepsilon \in \Omega$. Also, note that Ω is bounded. Otherwise, choose $x_k \geq 0$, $\omega_k > 0$ for $k = 1, 2, \ldots$ such that $\lim_{k\to\infty} \omega_k = \infty$ and $Gx_k - \omega_k Fx_k \geq 0$ for each k. By the homogeneity of G and F, we may assume that each x_k is a unit vector (with respect to some norm of $\operatorname{span} K_1$). Replacing by a subsequence, if necessary, we may also assume that $(x_k)_{k\in\mathbb{N}}$ converges to $\bar{x}$. By Remark 2(iii) and (iv), the sequence $(Gx_k)_{k\in\mathbb{N}}$ is bounded. Rewriting the above inequalities, we have $\omega_k^{-1} Gx_k \geq Fx_k$ for each k. Letting $k \to \infty$ and making use of the continuity of F at $\bar{x}$, we obtain $-F\bar{x} \geq 0$. On the other hand, since F is K_2-convex, by Remark 1, F satisfies condition (a) and hence the assumptions of Theorem 1. So by Theorem 1, we have $F(K_1\backslash\{0\}) \cap (-K_2) = \emptyset$. Hence, we arrive at a contradiction.

Denote $\sup \Omega$ by ω_0. Clearly $\omega_0 > 0$. By a modification of the above argument, it is clear that there exists $x_0 > 0$ such that $Gx_0 - \omega_0 Fx_0 \in \partial K_2$. We are going to show that $Gx_0 = \omega_0 Fx_0$.

In view of the last part of Theorem 1, there exists $z > 0$ such that $Fz = Gx_0$. By the positivity of G and condition (b), clearly $z \gg 0$. Since $-x_0 \notin K_1$, there exists $\lambda > 0$ such that $z - \lambda x_0 \in \partial K_1$. If $\lambda < \omega_0$, then by the convexity and homogeneity of F and the choice of z, we have

$$F(z-\lambda x_0) \geq Fz-\lambda Fx_0 = Gx_0-\lambda Fx_0 = \left(1 - \frac{\lambda}{\omega_0}\right) Gx_0 + \frac{\lambda}{\omega_0}(Gx_0 - \omega Fx_0) \gg 0,$$

which contradicts condition (b). So, we must have $\lambda \geq \omega_0$. Then, since G is concave, positive, and $z - \omega_0 x_0 \geq z - \lambda x_0 \geq 0$, we have

$$Gz - \omega_0 Fz = Gz - \omega_0 Gx_0 \geq G(z - \omega_0 x_0) \geq 0.$$

If $z - \omega_0 x_0 > 0$, then by the positivity of G and the above, we would obtain $Gz - \omega_0 Fz \gg 0$, which clearly contradicts the maximality of ω_0. So we must have $z - \omega_0 x_0 = 0$, and from the above we obtain $\lambda = \omega_0$ and $z = \omega_0 x_0$. Hence,

$$Gx_0 - \omega_0 Fx_0 = Gx_0 - \omega_0 F(\omega_0^{-1} z) = Gx_0 - Fz = 0,$$

which is what we want. Since x_0 is a positive scalar multiple of z, we also have $x_0 \gg 0$.

From the above, clearly $\omega_0 \in \Omega \cap \Sigma_1$. In order to establish the equalities $\sup \Omega = \inf \Sigma_1 = \omega_0$, it suffices to prove that $\sigma \geq \omega$ for any $\sigma \in \Sigma_1$ and $\omega \in \Omega$. We are going to show that the latter assertion is true even if we replace Σ_1 by Σ, which is defined by $\Sigma = \{\sigma \geq 0 : \exists x >^{K_1} 0,\ Gx \ {}^{K_2}\!\!\leq \sigma Fx\}$ (and, in fact, as the proof will show, in this case we have $\Sigma_1 = \Sigma$). Let $x >^{K_1} 0$, $y >^{K_1} 0$ be such that $Gx \ {}^{K_2}\!\!\leq \sigma Fx$ and $Gy \geq^{K_2} \omega Fy$. By the

(K_1, K_2)-positivity of G and condition (b), the first inequality clearly implies that $\sigma > 0$ and $x \in \operatorname{int} K_1$. So there exists $\varepsilon > 0$ such that $x - \varepsilon y \in \partial K_1$. Assume to the contrary that $\sigma < \omega$. Then $x - \varepsilon y \neq 0$ (otherwise, we would have $\sigma = \omega$) and by the given properties of F and G, we have

$$F(x-\varepsilon y) \geq Fx-\varepsilon Fy \geq \sigma^{-1}(Gx-\varepsilon\sigma\omega^{-1}Gy) \geq \sigma^{-1}(Gx-\varepsilon Gy) \geq \sigma^{-1}G(x-\varepsilon y),$$

which is a contradiction, as $G(x - \varepsilon y) \in \operatorname{int} K_2$ and $F(x - \varepsilon y) \notin \operatorname{int} K_2$,

The uniqueness of ω_0 and x_0 (up to positive multiples) will follow once we establish the last part of our result.

Last part. Let $y > 0$ and $\omega \geq 0$ be such that $Gy \geq \omega Fy$. Then $\omega \in \Omega$ and, by what we have proved, $\omega \leq \omega_0$. If the strict inequality does not hold, then from the above argument (with $x = x_0$ and $y = y$), we obtain $F(x_0 - \varepsilon y) \geq \omega_0^{-1}G(x_0 - \varepsilon y)$ and with $x_0 - \varepsilon y \in \partial K_1$ for some $\varepsilon > 0$, which is not possible, unless $x_0 = \varepsilon y$.

Similarly, we can also show that if $\sigma \geq 0$ is such that $Gx \leq \sigma Fx$ for some $x \in K_1 \backslash \{0\}$, which is not a multiple of x_0, then $\sigma > \omega_0$. ■

With some hindsight, we can give a few remarks on the relevance of conditions (a)–(c) of Theorem 3. First, the conclusion of Theorem 3, namely, $Gx_0 = \omega Fx_0$, where $x_0 > 0$, $\omega > 0$, together with the assumption that G is positive, forces the necessity of condition (c). But conditions (a), (b) together do not guarantee condition (c); for instance, if we take $K_1 = K_2 = K$ and F to be a linear map that maps K into ∂K, then F satisfies (a) and (b) but not (c). That is why we impose the condition. Next, according to Theorem 1, conditions (a)–(c) and the assumption that $\dim K_1 = \dim K_2$, together with the continuity and homogeneity of F, guarantee two conditions, namely, $F(K_1 \backslash \{0\}) \cap (-K_2) = \emptyset$ and $FK_1 \supseteq K_2$. In the proof of Theorem 3, the former condition is needed to guarantee the boundedness of Ω. The latter condition is also crucial for our desirable conclusion. For, if $FK_1 \not\supset \operatorname{int} K_2$, then we can choose $z \in \operatorname{int} K_2 \backslash FK_1$ and find a positive linear map G which maps K_1 onto the ray generated by z. For any such G, it is clear that the system (F, G) has no equilibrium point.

Remark 2. Let K be a proper cone. If $F : K \to \operatorname{span} K$ is linear and satisfies conditions (b) and (c) of Theorem 3, then for any homogeneous, continuous, (K, K)-nonnegative map $G : K \to \operatorname{span} K$, there exist a positive scalar ω and a nonzero vector x of K such that $Gx = \omega Fx$. However, the uniqueness of the equilibrium point is not guaranteed, even if we assume, in addition, that G is linear and K-irreducible (i.e. $GK \subseteq K$ and G leaves invariant no faces of K other than $\{0\}$ and K itself).

To show the existence of an equilibrium point for the system (F, G), we first note that F can be readily extended to a linear map on $\operatorname{span} K$. We still use the same symbol to denote its extension map. By Theorem 1, we have, $FK \supseteq K$. Since K is a full cone in $\operatorname{span} K$, this implies that F is nonsingular and we have $F^{-1}K \subseteq K$. Then one can readily verify that the map $F^{-1}G : K \to \operatorname{span} K$ is homogeneous, continuous and (K, K)-nonnegative. But any such map has a (necessarily, nonnegative) eigenvalue and a corresponding eigenvector in K (as can be proved by applying the Brouwer fixed-point theorem to the continuous map $T : C \to C$ given by $Tx = (f(F^{-1}Gx))^{-1}F^{-1}Gx$, where f is any fixed vector chosen from the interior of the dual cone of K and C is the compact convex full cross-section of K given by $C = \{x \in K : fx = 1\}$, assuming that $Gx \neq 0$ for all $x \in K \backslash \{0\}$). If ω is an eigenvalue and $x >^K 0$ is a corresponding eigenvector of $F^{-1}G$, then ω is an equilibrium value and x is an equilibrium point for the original system (F, G).

To see that uniqueness of the equilibrium point is not guaranteed, just take $K = \mathbb{R}^2_+$ and choose F, G to be the same and be the restriction to $\mathbb{R}^2_+$ of the linear map determined by the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

We would also like to add the following, which extends [F, Theorem 3]:

Corollary 1. *Let K_1, K_2 be proper cones such that $\dim K_1 = \dim K_2$. Let $F, G : K_1 \to \operatorname{span} K_2$ be maps that satisfy the hypotheses of Theorem 3. Also let ω_0 denote the positive number which has the same meaning as given in the theorem. Then the following are equivalent conditions on a real number σ :*

(a) $\sigma > \omega_0$;

(b) *For all $y >^{K_2} 0$, there exists $x \in K_1$ (which, necessarily, lies in $\operatorname{int} K_1$) such that $(\sigma F - G)x = y$;*

(c) *For some $y >^{K_2} 0$, there exists $x \in K_1$ (which, necessarily, lies in $\operatorname{int} K_1$) such that $(\sigma F - G)x = y$.*

Proof. (a) ⇒ (b): It is easy to see that, when $\sigma > \omega_0$, the map $\sigma F - G$ is homogeneous, continuous and satisfies conditions (a), (b) of Theorem 1. Since $(\sigma F - G)x_0 = (\sigma - \omega_0)Fx_0 \gg 0$, the map also satisfies condition (c). So, by the last part of Theorem 1, our assertion follows. [Since G is (K_1, K_2)-positive and F satisfies condition (b) of Theorem 3, it is clear that the solution vector x must lie in $\operatorname{int} K_1$.]

(b) ⇒ (c): Obvious.

(c) ⇒ (a): Suppose that condition (c) holds. If x is a multiple of x_0, then we have $0 < (\sigma F - G)x = (\sigma - \omega_0)Fx$, which implies $\sigma > \omega_0$, as

$Fx = \omega_0^{-1} Gx \gg 0$. If x is not a multiple of x_0, then by the last part of Theorem 3 we also obtain $\sigma > \omega_0$. ■

In order to obtain Theorem 4 from Theorem 3, we need to make use of the following lemma (except for its last part, which has interest of its own).

Lemma 1. *Let C be a compact convex set in a finite-dimensional real vector space and let $f : C \to W$ be a map from C to a finite-dimensional real vector space W ordered by a proper cone K. Suppose that 0 is not in the affine hull of C and let $F : \operatorname{pos} C \to W$ be the homogeneous map defined by $F(\lambda x) = \lambda f(x)$ for $x \in C$ and $\lambda \geq 0$. Then f is continuous (respectively, K-convex, K-concave, (C, K)-nonnegative, (C, K)-positive) if and only if F is continuous (respectively, K-convex, K-concave, $(\operatorname{pos} C, K)$-nonnegative, $(\operatorname{pos} C, K)$-positive). Furthermore, F is (C, K)-monotone if and only if for any $x,\ y \in C$ and $t > 1$, $(1-t)x + ty \in C$ implies $f(y) \geq (1 - \frac{1}{t})f(x)$.*

Proof. First, note that since $0 \notin \operatorname{aff} C$, each nonzero vector y of $\operatorname{pos} C$ can be expressed uniquely as λx, where $x \in C$ and $\lambda > 0$. So F is a well-defined map. By definition of F, it is clear that F is always homogeneous. Since f is the restriction of F to C, clearly f is continuous (or, convex, concave, nonnegative, positive), whenever F is. It is also easy to show that if f is continuous (respectively, nonnegative, positive), then so is F. We are going to show that if f is convex, then so is F, the proof for the corresponding concavity part being similar.

Suppose that f is convex. Since F is homogeneous, to establish the convexity of F, it suffices to show that for any $v,\ w \in \operatorname{pos} C \backslash \{0\}$, we have $F(v + w) \leq F(v) + F(w)$. Express v, w and $v + w$ in terms of vectors in C, say, $v = \alpha x$, $w = \beta y$ and $v + w = \gamma z$, where $\alpha,\ \beta,\ \gamma > 0$ and $x,\ y,\ z \in C$. Rewriting, we have $z = ax + by$, where $a = \alpha/\gamma$, $b = \beta/\gamma$ are both positive. Since $\operatorname{aff} C$ does not contain the origin 0, we can choose a nonzero vector e such that the inner product between e and each vector in $\operatorname{aff} C$ equals 1. Taking inner product of e with vectors on opposite side of the relation $\alpha x + \beta y = \gamma z$, we obtain $a + b = 1$. So by definition of F and the convexity of f, we have

$$F(v+w) = F(\gamma z) = \gamma f(z) = \gamma f(ax+by) \leq \gamma a f(x) + \gamma b f(y) = F(v) + F(w).$$

Last Part. Suppose that F is monotone. Let $x,\ y \in C$, $t > 1$ be such that $(1-t)x + ty \in C$. Then $y \geq (1 - \frac{1}{t})x$ and by the homogeneity and monotonicity of F, we have, $F(y) \geq (1 - \frac{1}{t})F(x)$, hence $f(y) \geq (1 - \frac{1}{t})f(x)$.

Conversely, suppose that f possesses the given property. Consider any vectors $u,\ v \in \operatorname{pos} C \backslash \{0\}$ with $v \geq u$. Express v, u and $v - u$ in terms of vectors in C, say, $v = \beta y$, $u = \alpha x$ and $v - u = \gamma z$ where $x,\ y,\ z \in C$ and $\alpha,\ \beta,\ \gamma > 0$. Set $t = \beta/\gamma$. After some manipulations (and again making use of the fact that $\langle x, e\rangle = \langle y, e\rangle = \langle z, e\rangle = 1$, where the vector e has the same meaning as above), we obtain $(1-t)x + ty = z \in C$ and $t > 1$. By the property of f, we have $f(y) \geq (1 - \frac{1}{t})f(x)$. Rewriting the latter inequality in terms of u, v (and α, β, γ) and simplifying, we obtain $F(v) \geq F(u)$. This shows that F is monotone. ■

Proof of Theorem 4. We may assume that $0 \notin \operatorname{aff} C$. Otherwise, choose a one-to-one affine map that takes C onto some compact convex set $\widetilde{C}$ for which $0 \notin \operatorname{aff} \widetilde{C}$, define maps $\tilde{f}$, $\tilde{g}$ corresponding to f, g in the natural way, and work with $\widetilde{C}$, $\tilde{f}$ and $\tilde{g}$ instead.

Let $F : \operatorname{pos} C \to W$ be the map defined by $F(y) = \lambda f(x)$ for $y \in \operatorname{pos} C$, where $y = \lambda x$, $x \in C$ and $\lambda \geq 0$. Since f is continuous, convex on C, by Lemma 1, F is continuous, convex on $\operatorname{pos} C$. In view of (b) and (c) (and the homogeneity of F), it is clear that, we have, $F(\partial(\operatorname{pos} C)) \cap K = \emptyset$ and $F(\operatorname{pos} C) \cap \operatorname{int} K \neq \emptyset$. Now let $G : \operatorname{pos} C \to W$ be the homogeneous map defined in a similar way (in terms of g). By Lemma 1 again, G is a continuous, concave positive map. Since the restriction of F (respectively, G) to C equals f (respectively, g) and $0 \notin \operatorname{aff} C$, we can apply Theorem 3 to the pair (F, G) to draw the desired conclusions. ■

With the aid of Sperner's Lemma (and by adapting the proof of [F, Theorem 1]), one can derive the first part of Ky Fan's theorem from the first part of Theorem 4. The last part of Ky Fan's theorem can also be deduced from the identity $\sup \Omega = \inf \Sigma_1 = \omega_0$ (of Theorem 4) by making use of the following readily-proved facts: $\sup \Omega = \max_{x \in S} r(x)$, $\inf \Sigma = \min_{x \in S} R(x)$, $\Sigma = \Sigma_1$ in this case, and for any $x \in S$, $r(x)^{-1} = \max_{1 \leq i \leq n} f_i(x)/g_i(x)$ and $R(x)^{-1} = \min_{1 \leq i \leq n} f_i(x)/g_i(x)$, where

$$\begin{aligned} r(x) &= \max\{\omega \geq 0 : g(x) \geq \omega f(x)\}, \\ R(x) &= \min\{\sigma \geq 0 : g(x) \leq \sigma f(x)\} \text{ (by convention } \min \emptyset = \infty), \end{aligned}$$

and $\Sigma = \{\sigma \geq 0 : \exists x \in C, g(x) \overset{K}{\leq} \sigma f(x)\}$.

Actually, Theorems 3 and 4 are equivalent. Also, Theorem 1 admits the following equivalent formulation with K_1 replaced by a compact convex set:

Theorem 1′. *Let C be a compact convex set in a finite-dimensional real vector space, and let $f : C \to W$ be a continuous map from C to a finite-dimensional real vector space W ordered by a proper cone K such that $\dim W = \dim C + 1$. Suppose that f satisfies each of the following conditions*:

(a) *f is K-convex*;

(b) $f(\operatorname{rbd} C) \cap K = \emptyset$; *and*

(c) $f(C) \cap \operatorname{int} K \neq \emptyset$.

Then $f(C) \cap (-K) = \emptyset$ *and* $K \subseteq \bigcup_{\lambda \geq 0} \lambda f(C)$.

Note that if C is an $(n-1)$-dimensional compact convex set whose affine hull does not contain the origin, then $\operatorname{pos} C$ is an n-dimensional closed, pointed convex cone. Then C (respectively, $\operatorname{rbd} C$) is homeomorphic with $(\operatorname{pos} C) \cap S^{n-1}$ (respectively, $\partial(\operatorname{pos} C) \cap S^{n-1}$), after identifying $\operatorname{span} C$ (= $\operatorname{span}(\operatorname{pos} C)$) with $\mathbb{R}^n$. Indeed, we could have introduced the concept of a pair of antipodal points of $\operatorname{rbd} C$ relative to a relative interior point of C, and also could have derived Theorem $1'$ directly (using an argument similar to that for Theorem 1) and then used it to prove Theorem 4.

Certainly we can also reformulate Corollary 1 in the setting when the common domain of F and G is a compact convex set.

5. Final Remarks

In Theorem 3, if K_1, K_2 are the same and equal to a proper cone K, F equals the identity map on $\operatorname{span} K$ and G equals a linear map A that preserves K (i.e. $AK \subseteq K$), then the sets Ω and Σ_1 considered in the theorem become two of the four Collatz-Wielandt sets associated with the cone-preserving map A. Collatz-Wielandt sets were first introduced by Barker and Schneider [BS]. The greatest lower bound and the least upper bound of the Collatz-Wielandt sets are studied in [T–W]; in particular, it is proved that, for any linear map A that preserves K, we have $\sup \Omega = \inf \Sigma_1 = \rho(A)$, where $\rho(A)$ denotes the spectral radius of A. For more recent developments of the topic, we refer the reader to the review paper [T]. In the book [A2, Chapter 11], Aubin has also elaborated on the results of [F] in the setting of a pair of maps F, G from the standard simplex of $\mathbb{R}^n$ to $\mathbb{R}^m$ and with the continuity assumptions on F, G replaced respectively by the lower and upper semicontinuity assumptions. The study of the Collatz-Wielandt sets associated with a pair of nonlinear maps (in particular, the determination of when $\sup \Omega$ and $\inf \Sigma_1$ are the same and equal to an equilibrium value, etc.), and also the introduction and study of the concepts of lower or upper semicontinuity of a map with respect to a proper cone seem worthwhile and will form the subject matter of future work.

References

[A1] J.P. Aubin, Propriété de Perron-Frobenius pour des correspondences, *C.R. Acad. Sci. Paris* **286** (1978), 911–914.

[A2] J.P. Aubin, *Optima and Equilibria*, 2nd ed., Springer-Verlag, Berlin, 1998.

[BS] G.P. Barker and H. Schneider, Algebraic Perron-Frobenius theory, *Linear Algebra Appl.* **11** (1975), 219–233.

[Do] P.G. Doods, Positive compact operators, *Quaestiones Math.* **18** (1995), 21–45.

[Du] J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.

[DG] J. Dugundji and A. Granas, *Fixed Point Theory*, Vol.1, PWN–Polish Scientific Publishers, Warszawa, 1982.

[F] Ky Fan, On the equilibrium value of a system of convex and concave functions, *Math. Zeitschr.* **70** (1958), 271–280.

[HIR] D.H. Hyers, G. Isac and T.M. Rassias, *Topics in Nonlinear Analysis and Applications*, World Scientific, Singapore, 1997.

[KLS] M.A. Krasnosel'skij, Je.A. Lifshits and A.V. Sobolev, Positive Linear Systems: The Method of Positive Operators, translated from the Russian by J. Appell, Berlin, Heldermann, 1989.

[N1] R.D. Nussbaum, Convexity and log convexity for the spectral radius, *Linear Algebra Appl.* **73** (1986), 59–122.

[N2] R.D. Nussbaum, Eigenvectors of order-preserving linear operators, *J. London Math. Soc.* (2) **58** (1998), 480–496.

[NVL] R.D. Nussbaum and S.M. Verduyn Lunel, Generalizations of the Perron-Frobenius theorem for nonlinear maps. *Memoirs of the American Mathematical Society*, **659** (1999).

[Sim] S. Simons, The continuity of inf sup, with applications, *Arch. Math.* **48** (1987), 426–437.

[Spe] E. Sperner, Neuer Beweis fur die Invarianz der Dimensionszahl und des Gebietes, *Abh. Math. Sem. Ham. Univ.* **6** (1928), 265–272.

[T] B.S. Tam, A cone-theoretic approach to the spectral theory of positive linear operators: the finite-dimensional case, *Taiwanese J. Math.* **5** (2001), 207–277.

[TW] B.S. Tam and S.F. Wu, On the Collatz-Wielandt sets associated with a cone-preserving map, *Linear Algebra Appl.* **125** (1989), 77–95.

[Z] M. Zerner, Quelques propriétés spectrales des opérateurs positifs, *J. Funct. Anal.* **72** (1987), 381–417.