STUDIA MATHEMATICA 203 (1) (2011)

Universal bounds for matrix semigroups

by

LEO LIVSHITS (Waterville, ME), GORDON MACDONALD (Charlottetown), and HEYDAR RADJAVI (Waterloo)

Abstract. We show that any compact semigroup of $n \times n$ matrices is similar to a semigroup bounded by $\sqrt{n}$. We give examples to show that this bound is best possible and consider the effect of the minimal rank of matrices in the semigroup on this bound.

1. Introduction. It is well-known that any compact group $\mathcal{G}$ contained in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), the $n \times n$ real (resp. complex) matrices, is similar to a group of orthogonal (resp. unitary) matrices. Expressed another way, we have the following (see [RR]).

THEOREM 1.1. *If $\mathcal{G}$ is a compact group in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), then there exists an invertible matrix X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) so that*

$$\sup\{\|X^{-1}GX\| : G \in \mathcal{G}\} = 1.$$

The norm we are using here (and throughout this paper) is the usual operator norm for matrices acting on a finite-dimensional Hilbert space. So we consider the norm of $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) to be $\|x\| = (\sum_{i=1}^n |x_i|^2)^{1/2}$, and then for A in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$),

$$\|A\| = \max\{\|Ax\| : \|x\| = 1\}.$$

We are interested in related problems for matrix semigroups, sets of matrices closed under matrix multiplication:

(1) Given a compact semigroup $\mathcal{S}$ in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), for what values of $K_S > 0$ does there exist an invertible X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) such that

$$\sup\{\|X^{-1}SX\| : S \in \mathcal{S}\} \le K_S?$$

(2) Do there exist universal constants K_n (independent of the semigroup), for each $n = 1, 2, \ldots$, such that for each compact semigroup

2010 *Mathematics Subject Classification*: Primary 15A30; Secondary 15A60.
Key words and phrases: semigroup, matrix, similarity, norm bounds.

DOI: 10.4064/sm203-1-4

$\mathcal{S}$ in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), we have an invertible matrix X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) with

$$\sup\{\|X^{-1}SX\| : S \in \mathcal{S}\} \le K_n?$$

Also, if such universal constants do exist, what is the best value of K_n, $n = 1, 2, \ldots$?

From now on, we shall motivate our theorems by considering only the real case, which most people find easier to visualize. When we state our theorems we shall mention both the real and complex cases.

We must stress that the problem we are are considering is not equivalent to the problem of renorming $\mathbb{R}^n$, and considering the operator norm on $M_n(\mathbb{R})$ induced by that norm. In that case the universal constant would be 1, since for a given compact semigroup $\mathcal{S}$, we can define a new norm on $\mathbb{R}^n$ by

$$|x|_{\mathcal{S}} = \sup\{\|Sx\| : S \in \mathcal{S} \cup \{I\}\},$$

which clearly has the property that for any $S \in \mathcal{S}$, $|Sx|_{\mathcal{S}} \le |x|_{\mathcal{S}}$ for all $x \in \mathbb{R}^n$. This norm is not a Hilbert space norm, in other words, its unit ball $B_{|\cdot|_{\mathcal{S}}}(1) = \{x \in \mathbb{R}^n : |x|_{\mathcal{S}} \le 1\}$ is not an ellipsoid. However, it is a symmetric convex body (i.e. a convex set with interior and with the property that if $x \in B_{|\cdot|_{\mathcal{S}}}(1)$ then $-x \in B_{|\cdot|_{\mathcal{S}}}(1)$). Fritz John [J] has a remarkable theorem relating such sets to ellipsoids.

THEOREM 1.2 (Fritz John [J]). *Let $K \in \mathbb{R}^n$ be a symmetric convex body. Then there is a unique ellipsoid $E \subseteq K$ of maximal volume, and for this ellipsoid, $K \subseteq \sqrt{n}E$.*

There is also a complex version of the Fritz John Theorem (see [CM]), where the definition of a symmetric set K is now taken to be that if $x \in K$ and $\alpha \in \mathbb{T}$ (the complex numbers of modulus 1), then $\alpha x \in K$, and the conclusion is the same.

2. Main theorems. Our first theorem shows that there is a universal bound, which can be obtained from the Fritz John Theorem.

THEOREM 2.1. *If $\mathcal{S}$ is a compact semigroup in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), then there is an invertible matrix X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) so that $X^{-1}\mathcal{S}X$ is bounded by $\sqrt{n}$.*

Proof. We shall begin with the real case. By adjoining the identity to $\mathcal{S}$ if necessary, we may assume that $\mathcal{S}$ has no common kernel, and so for $x \in \mathbb{R}^n$, $|x|_{\mathcal{S}} = \sup\{\|Sx\| : S \in \mathcal{S}\}$ defines a norm on $\mathbb{R}^n$ with respect to which each $S \in \mathcal{S}$ is a contraction. That is, for each $S \in \mathcal{S}$, $|Sx|_{\mathcal{S}} \le |x|_{\mathcal{S}}$ for all $x \in \mathbb{R}^n$. As mentioned above, the unit ball in this norm, $B_{|\cdot|_{\mathcal{S}}}(1) = \{x \in \mathbb{R}^n : |x|_{\mathcal{S}} \le 1\}$, is a compact, convex, symmetric set (i.e. if $x \in B_{|\cdot|_{\mathcal{S}}}(1)$ then $-x \in B_{|\cdot|_{\mathcal{S}}}(1)$).

The set $B_{|\cdot|_{\mathcal{S}}}(1)$ also has interior since if $\mathcal{S}$ is bounded by M, then any vector $x \in \mathbb{R}^n$ with $\|x\| \le 1/M$ is in B. Thus, by the Fritz John Theorem there exists an ellipsoid E contained in $B_{|\cdot|_{\mathcal{S}}}(1)$ so that $B_{|\cdot|_{\mathcal{S}}}(1)$ is contained in $\sqrt{n}E$. Take E to be the unit ball of a new Hilbert space norm. Then

$$S(E) \subseteq S(B_{|\cdot|_{\mathcal{S}}}(1)) \subseteq B_{|\cdot|_{\mathcal{S}}}(1) \subseteq \sqrt{n}E$$

so for this new Hilbert space norm we have $\|S\| \le \sqrt{n}$. The X mentioned in the theorem is any invertible matrix which maps $\{x \in \mathbb{R}^n : \|x\| \le 1\}$ onto E. The proof in the complex case is identical, except we use the complex version of the Fritz John Theorem [CM]. ∎

If we are searching for a universal bound, this result is best possible, as the following examples show.

We let $\{e_1, \ldots, e_n\}$ denote the standard basis for $\mathbb{C}^n$, that is, e_i is the vector in $\mathbb{C}^n$ with a 1 in the ith entry and zeroes in all other entries. Let $\{f_1, \ldots, f_n\}$ denote the Fourier basis, that is, for ρ a primitive nth root of unity, $f_j = \sum_{k=1}^n \rho^{jk} e_k$ for $j = 1, \ldots, n$.

We follow the usual convention of considering vectors in $\mathbb{C}^n$ to be columns of numbers and for x in $\mathbb{C}^n$, x^* will denote the conjugate transpose, the row vector whose entries are the complex conjugates of the entries of x.

Example 2.2 (Complex example). For $n \in \mathbb{N}$,

$$\mathcal{S}_n = \left\{ z e_i f_j^* : i, j = 1, \ldots, n,\ z \in \mathbb{T} \right\}$$

is an irreducible compact semigroup in $M_n(\mathbb{C})$ with the property that, for any invertible matrix X in $M_n(\mathbb{C})$,

$$\max\left\{ \|X^{-1} S X\| : S \in \mathcal{S} \right\} \ge \sqrt{n}.$$

It is easy to verify that $\mathcal{S}$ is a semigroup and is compact and irreducible (i.e. has no non-trivial proper invariant subspaces). Let X be an invertible matrix X in $M_n(\mathbb{C})$. By writing the polar decomposition of $X = UP$ where U is unitary and P is positive definite, we can see that it suffices to consider the case where X is positive definite. So there exists an orthonormal basis $\{w_i\}_{i=1}^n$ and positive numbers $\{\alpha_i\}_{i=1}^n$ so that $X = \sum_{i=1}^n \alpha_i w_i w_i^*$. Then

$$\|X^{-1} e_j\|^2 = \sum_{i=1}^n \frac{1}{\alpha_i^2} |w_i^* e_j|^2$$

and so

$$\sum_{j=1}^n \|X^{-1} e_j\|^2 = \sum_{i=1}^n \frac{1}{\alpha_i^2}$$

(this is the square of the Hilbert–Schmidt norm).

Similarly,

$$\sum_{j=1}^{n} \|Xf_j\|^2 = \sum_{i=1}^{n} \alpha_i^2 n.$$

So there exist j_0 and j_1 so that

$$\|X^{-1}e_{j_0}\|^2 \geq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha_i^2} \quad \text{and} \quad \|Xf_{j_1}\|^2 \geq \sum_{i=1}^{n}\alpha_i^2.$$

Thus

$$\begin{aligned}\|X^{-1}(e_{j_0}f_{j_1}^T)X\|^2 = \|X^{-1}e_{j_0}\|^2\|Xf_{j_1}\|^2 &\geq \frac{1}{n}\Big(\sum_{i=1}^{n}\frac{1}{\alpha_i^2}\Big)\Big(\sum_{i=1}^{n}\alpha_i^2\Big)\\ &\geq \frac{1}{n}\left\langle \begin{bmatrix}1/\alpha_1\\ \vdots \\ 1/\alpha_n\end{bmatrix}, \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{bmatrix}\right\rangle^2 \qquad \text{(Cauchy–Schwarz inequality)}\\ &= \frac{1}{n}n^2 = n,\end{aligned}$$

so $\|X^{-1}(e_{j_0}f_{j_1}^T)X\| \geq \sqrt{n}$.

This example is very complex in nature, but by taking the tensor product of the 2×2 version of the above example, we can obtain an example in $M_{2^n}(\mathbb{R})$. To see that Theorem 2.1 is best possible in all dimensions, even in the real case, we have the following construction.

Let $\mathcal{Y}$ denote the set of vertices of the n-cube

$$[-1,1]^n = \{y \in \mathbb{R}^n : |y_i| \leq 1 \text{ for all } i = 1,\dots,n\}.$$

Example 2.3 (Real example). For $n \in \mathbb{N}$,

$$\mathcal{S}_n = \{ye_i^T : y \in \mathcal{Y},\ i = 1,\dots,n\}$$

is an irreducible compact semigroup in $M_n(\mathbb{R})$ with the property that, for any invertible matrix X in $M_n(\mathbb{R})$,

$$\max\{\|X^{-1}SX\| : S \in \mathcal{S}\} \geq \sqrt{n}.$$

To see this, let X be an invertible matrix in $M_n(\mathbb{R})$ implementing a similarity. By scaling if necessary, with no loss of generality we may assume that the determinant of X is 1. Then, just as $\mathcal{S}_n$ contains rank-one matrices which map the centers of faces of the n-cube $[-1,1]^n$ to the corners of $[-1,1]^n$, $X^{-1}\mathcal{S}_nX$ contains rank-one matrices which map the centers of the faces of the n-parallelotope $X^{-1}[-1,1]^n$ to the corners of $X^{-1}[-1,1]^n$.

Our first claim is that the center of at least one face has norm greater than or equal to 1. To see this, use the fact that for an $n \times n$ matrix M with columns $(m_1, \ldots, m_n)$,

$$|\det M| \le \|m_1\| \cdots \|m_n\|,$$

applied to $(X^{-1})^T$. Since $|\det X| = 1$ we must have $\|e_{i_0}^T X^{-1}\| \ge 1$ for some i_0.

Our second claim is that under the map X, at least one corner of the n-cube gets mapped to a corner of the n-parallelotope $X[-1,1]^n$ of greater or equal norm. If this were not the case then $X[-1,1]^n$ would be contained in the ball of radius $\sqrt{n}$ and would have the same volume as the n-cube. But the Fritz John Theorem tells us that the unique n-parallelotope of largest volume in a ball is a cube, so X must map the n-cube to another n-cube in this case and thus X is orthogonal. Thus there exists $y_0 \in \mathcal{Y}$ such that $\|Xy_0\| \ge \sqrt{n}$.

Now we have $\|Xy_0 e_{i_0}^T X^{-1}\| \ge \sqrt{n}$.

If we are going to consider properties internal to each semigroup, we can obtain some refinements of Theorem 2.1.

Definition 2.4. A subspace M of $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) is *semiinvariant* for a set $\mathcal{T}$ in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), if there exist subspaces $M_1 \subseteq M_2$ in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) which are both invariant for $\mathcal{T}$ and satisfy $M_2 = M_1 \oplus M$.

Letting P_M denote the orthogonal projection onto M, we define the *maximal diagonal dimension* of $\mathcal{T}$, written $\operatorname{mdd}(\mathcal{T})$, to be the maximum of the dimensions of all semiinvariant subspaces M of $\mathcal{T}$ such that $P_M \mathcal{T} P_M|_M$ is irreducible (i.e. has no non-trivial proper invariant subspaces).

Definition 2.5. A *maximal block triangularization* of a set $\mathcal{T}$ in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) is a chain of subspaces of maximal length among all chains $\{0\} = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_k = \mathbb{R}^n$ (resp. $\mathbb{C}^n$) with the property that each M_i is invariant under each T in $\mathcal{T}$.

If we have a maximal block triangularization of a set $\mathcal{T}$, then by choosing a basis for M_1 and extending it to a basis for M_2, then to a basis for M_3 and so on until we have a basis for $\mathbb{R}^n$, and then letting X be the transition matrix from this basis to the standard basis, we see that

$$X^{-1}\mathcal{T}X = \{X^{-1}TX : T \in \mathcal{T}\}$$

is block upper-triangular and maximality implies that the diagonal blocks are irreducible. Thus in any maximal block triangularization, the size of each diagonal block of $X^{-1}\mathcal{T}X$ will be less than or equal to $\operatorname{mdd}(\mathcal{T})$.

The next theorem shows that the smaller the diagonal blocks are in a triangularization, the more we can reduce the norm.

Theorem 2.6. *If $\mathcal{S}$ is a compact semigroup in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$), then for all $\epsilon > 0$ there is an invertible matrix X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) so that $X^{-1}\mathcal{S}X$ is bounded by $\sqrt{\operatorname{mdd}(\mathcal{S})} + \epsilon$.*

Proof. By choosing a maximal triangularization for $\mathcal{S}$, and applying the similarity mentioned above, we may assume that $\mathcal{S}$ is block upper-triangular and the largest size of a diagonal block is $\mathrm{mdd}(\mathcal{S})$. Applying Theorem 2.1 to each diagonal block, and taking the direct sum of the invertible matrices X_i we get from that theorem, we obtain an invertible matrix $X = \bigoplus_i X_i$ so that $X^{-1}\mathcal{S}X$ has diagonal blocks bounded by $\sqrt{\mathrm{mdd}(\mathcal{S})}$. Now apply a similarity induced by an invertible matrix of the form

$$Y = \begin{bmatrix} I & & & \\ & \delta I & & \\ & & \ddots & \\ & & & \delta^k I \end{bmatrix}$$

(with $0 \le \delta < 1$) to $X^{-1}\mathcal{S}X$, where Y is block diagonal with diagonal blocks of the same size as those of $\mathcal{S}$. This leaves the diagonal blocks unchanged and scales the upper-triangular, non-diagonal blocks by scalars less than or equal to δ. By choosing $\delta > 0$ small enough we can ensure that the norm of $(XY)^{-1}\mathcal{S}(XY)$ is less than $\sqrt{\mathrm{mdd}(\mathcal{S})} + \epsilon$. ∎

There is still room for improvement. In the case where we have a compact group $\mathcal{G}$ in $M_n(\mathbb{R})$, we know we can achieve norm 1. This is usually proved using the existence of Haar measure (see [RR] or [A]), but can also be shown by using the uniqueness of the maximal volume ellipsoid in the Fritz John Theorem (again see [RR] or [DS]). In brief, the argument is as follows. Using compactness and the Spectral Mapping Theorem, we can see that any invertible G in $\mathcal{G}$ must have all eigenvalues on the unit circle, so have determinant 1, and so it is volume preserving. But this means that the image of the maximal-volume ellipsoid E in $B_{|\cdot|_\mathcal{G}}(1)$ under any G in $\mathcal{G}$ is another maximal-volume ellipsoid $G(E)$ in $B_{|\cdot|_\mathcal{G}}(1)$. By uniqueness, $G(E) = E$, and so taking the Hilbert space norm associated to this ellipsoid we obtain the norm 1 bound.

Is there a fundamental difference between the invertible case and all other cases? Or, as the minimal rank of matrices in our compact semigroup increases, does the actual bound after a similarity decrease? The following theorem gives some evidence for the second conjecture.

THEOREM 2.7. *If $\mathcal{S}$ is a compact semigroup in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) and* $\mathrm{rank}(S) = n-1$ *for all $S \in \mathcal{S}$ then there exists an invertible matrix X in $M_n(\mathbb{R})$ (resp. $M_n(\mathbb{C})$) such that $\|X^{-1}\mathcal{S}X\| \le \sqrt{2}$.*

Proof. We shall prove the theorem in the complex case, as there is then a slight complication. The real case follows similarly.

Each matrix in $\mathcal{S}$ must have spectral radius equal to one, since if $\rho(S) > 1$ for some $S \in \mathcal{S}$, then $\{S^j\}_{j=1}^\infty$ is unbounded, contradicting compactness; and if $\rho(S) < 1$ then $\{S^j\}_{j=1}^{\infty}$ converges to 0, which contradicts constant rank $n-1$. Also, since rank is constant, it must be that the Jordan form of any S in $\mathcal{S}$ has no nilpotent blocks. Since $\mathcal{S}$ is bounded, all Jordan blocks must be 1×1. Also, by compactness, given any $S \in \mathcal{S}$ there exists a increasing sequence of natural numbers n_j so that S^{n_j} converges. The limit must also be of rank $n-1$, which means that for each S in $\mathcal{S}$, except for the eigenvalue 0 of multiplicity 1, the rest of the spectrum lies on the unit circle. Thus, each S in $\mathcal{S}$ is similar to a matrix of the form $U \oplus 0$ where U is an $(n-1)\times(n-1)$ unitary matrix.

Assume that such a similarity has been applied to our semigroup and so with no loss of generality we may assume that some matrix of the form $U \oplus 0$, where U is an $(n-1) \times (n-1)$ unitary matrix, is in our semigroup. By taking a limit of a sequence of powers of this matrix we may assume that the projection $E = I_{n-1} \oplus 0$ is in our semigroup. Now consider $E\mathcal{S}E|_{\mathrm{Ran}(E)}$. This is a compact group of invertibles and hence by Theorem 1.1 is similar (via an invertible of the form $X \oplus I_1$) to a group of unitaries. Again we shall assume such a similarity has been applied to our semigroup, and so we deduce that there exists a unitary group $\mathcal{G}$ in $M_{n-1}(\mathbb{C})$ such that

$$\mathcal{S} \subseteq \left\{ \begin{bmatrix} U & y \\ x^* & \alpha \end{bmatrix} : U \in \mathcal{G},\ x, y \in \mathbb{C}^{n-1},\ \alpha \in \mathbb{C} \right\},$$

and rank restrictions force our semigroup to be of the form

$$\mathcal{S} \subseteq \left\{ \begin{bmatrix} U & Uy \\ x^*U & x^*Uy \end{bmatrix} = \begin{bmatrix} I_{n-1} \\ x^* \end{bmatrix} U \begin{bmatrix} I_{n-1} & y \end{bmatrix} : U \in \mathcal{G},\ x, y \in \mathbb{C}^{n-1} \right\}.$$

So if we choose two elements in $\mathcal{S}$,

$$A = \begin{bmatrix} I_{n-1} \\ x_A^* \end{bmatrix} U_A \begin{bmatrix} I_{n-1} & y_A \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} I_{n-1} \\ x_B^* \end{bmatrix} U_B \begin{bmatrix} I_{n-1} & y_B \end{bmatrix},$$

then

$$AB = \begin{bmatrix} I_{n-1} \\ x_A^* \end{bmatrix} U_A \left(I_{n-1} + y_A x_B^* \right) U_B \begin{bmatrix} I_{n-1} & y_B \end{bmatrix}$$

and so $I_{n-1} + y_A x_B^*$ must be a unitary. In that case $y_A x_B^*$ is normal. Thus $y_A x_B^* x_B y_A^* = x_B y_A^* y_A x_B^*$ and so $\|x_B\|^2 y_A y_A^* = \|y_A\|^2 x_B x_B^*$.

There are three possibilities at this stage: (1) all x_B and y_A are zero, in which case the universal norm bound is 1; (2) all the x_B are zero or all the y_A are zero, in which case $\mathcal{S}$ is triangularized with a unitary diagonal block and a zero diagonal block, and so we can apply a diagonal similarity and get a universal bound of $1+\epsilon$ for arbitrarily small ϵ; or (3) there is at least one non-zero x_A and one non-zero y_B. In case (3) the equation

$\|x_B\|^2 y_A y_A^* = \|y_A\|^2 x_B x_B^*$ immediately implies that there exists a single unit vector $z \in \mathbb{C}^{n-1}$ such that for each x_A we have a scalar α_A such that $x_A = \alpha_A z$ and for each y_B we have a scalar β_B so that $y_B = \beta_B z$. Therefore, any two A and B in $\mathcal{S}$ are of the form

$$A = \begin{bmatrix} I_{n-1} \\ \alpha_A z^* \end{bmatrix} U_A \begin{bmatrix} I_{n-1} & \beta_A z \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} I_{n-1} \\ \alpha_B z^* \end{bmatrix} U_B \begin{bmatrix} I_{n-1} & \beta_B z \end{bmatrix},$$

and as above, when we consider the product we see that the scalars must satisfy the condition $|1 + \alpha_A \beta_B| = 1$.

If we let $\mathcal{A} = \{\alpha_A : A \in \mathcal{S}\}$ and $\mathcal{B} = \{\beta_B : B \in \mathcal{S}\}$ then we know both these sets contain non-zero elements (otherwise we would be in case (1) or (2)). Suppose one of these sets, say $\mathcal{A}$, contains at least two non-zero elements. Then by applying a diagonal similarity we may assume that $1 \in \mathcal{A}$ and $\alpha_0 \in \mathcal{A}$ and α_0 is not 0 or 1. Then every β_B lies on the intersection of the circles $|1 + t| = 1$ and $|1 + \alpha_0 t| = 1$. But these two circles intersect at zero and hence can intersect in at most one other location. By symmetry, this shows that at least one of the sets $\mathcal{A}$ and $\mathcal{B}$ contains at most one non-zero element. By transposing if necessary and applying a diagonal similarity, we may assume that $1 \in \mathcal{A} \subseteq \{0, 1\}$. Then $|1 + \beta_B| = 1$ for all $\beta_B \in \mathcal{B}$.

Now we see that there is a unitary group $\mathcal{G}$ in $M_{n-1}(\mathbb{C})$ and a unit vector $z \in \mathbb{C}^{n-1}$ so that

$$\mathcal{S} \subseteq \left\{ \begin{bmatrix} I_{n-1} \\ \alpha z^* \end{bmatrix} U \begin{bmatrix} I_{n-1} & \beta z \end{bmatrix} : U \in \mathcal{G},\ \alpha \in \{0, 1\},\ |1 + \beta| = 1 \right\}.$$

Finally, applying the similarity induced by

$$T = \begin{bmatrix} I_{n-1} & -z \\ 0 & 1 \end{bmatrix} \quad \left(\text{so } T^{-1} = \begin{bmatrix} I_{n-1} & z \\ 0 & 1 \end{bmatrix} \right),$$

we obtain

$$T\mathcal{S}T^{-1} \subseteq \left\{ \begin{bmatrix} I_{n-1} - \alpha z z^* \\ \alpha z^* \end{bmatrix} U \begin{bmatrix} I_{n-1} & (1 + \beta) z \end{bmatrix} : U \in \mathcal{G},\ \alpha \in \{0, 1\},\ |1 + \beta| = 1 \right\}.$$

The first and second matrices in the above product have norm 1 and the third has norm $\sqrt{2}$, and so the result is proven. ∎

Acknowledgements. The first author's research was supported by Colby College Natural Sciences Division Grant. The second and third authors' research was supported by NSERC (Canada) research grants.

References

[A] H. Auerbach, *Sur les groupes bornés de substitutions linéaires*, C. R. Acad. Sci. Paris 195 (1932), 1367–1369.

[CM] B. D. Craven and B. Mond, *Real and complex Fritz John theorems*, J. Math. Anal. Appl. 44 (1973), 773–778.

[DS] E. Deutsch and H. Schneider, *Bounded groups and norm-Hermitian matrices*, Linear Algebra Appl. 9 (1974), 9–27.

[J] F. John, *Extremum problems with inequalities as subsidiary conditions*, in: Studies and Essays Presented to R. Courant on his 60th Birthday, Interscience, New York, NY, 1948, 187–204.

[RR] H. Radjavi and P. Rosenthal, *Simultaneous Triangularization*, Universitext, Springer, New York, 2000.

Leo Livshits
Department of Mathematics
Colby College
Waterville, ME 04901, U.S.A.
E-mail: llivshi@colby.edu

Gordon MacDonald
Department of Mathematics and Statistics
University of Prince Edward Island
Charlottetown, PE, Canada C1A 4P3
E-mail: gmacdonald@upei.ca

Heydar Radjavi
Department of Pure Mathematics
University of Waterloo
Waterloo, ON, Canada N2L 3G1
E-mail: hradjavi@uwaterloo.ca

Received October 27, 2010
Revised version January 19, 2011 (7018)