Canad. J. Math. Vol. **56** (1), 2004 pp. 134–167

Linear Operators on Matrix Algebras that Preserve the Numerical Range, Numerical Radius or the States

Chi-Kwong Li and Ahmed Ramzi Sourour

Abstract. Every norm ν on $\mathbf{C}^n$ induces two norm numerical ranges on the algebra M_n of all $n \times n$ complex matrices, the spatial numerical range

$$W(A) = \{x^*Ay : x, y \in \mathbf{C}^n, \nu^D(x) = \nu(y) = x^*y = 1\},$$

where ν^D is the norm dual to ν, and the algebra numerical range

$$V(A) = \{f(A) : f \in \mathcal{S}\},$$

where $\mathcal{S}$ is the set of states on the normed algebra M_n under the operator norm induced by ν. For a symmetric norm ν, we identify all linear maps on M_n that preserve either one of the two norm numerical ranges or the set of states or vector states. We also identify the numerical radius isometries, *i.e.*, linear maps that preserve the (one) numerical radius induced by either numerical range. In particular, it is shown that if ν is not the ℓ_1, ℓ_2, or ℓ_∞ norms, then the linear maps that preserve either numerical range or either set of states are "inner", *i.e.*, of the form $A \mapsto Q^*AQ$, where Q is a product of a diagonal unitary matrix and a permutation matrix and the numerical radius isometries are unimodular scalar multiples of such inner maps. For the ℓ_1 and the ℓ_∞ norms, the results are quite different.

1 Introduction

Let M_n be the algebra of $n \times n$ complex matrices, and let $\|\cdot\|$ be the operator norm on M_n induced by a norm ν on $\mathbf{C}^n$, *i.e.*,

$$\|A\| = \max\{\nu(Ax) : x \in \mathbf{C}^n, \nu(x) \le 1\}.$$

Suppose X and Y are complex matrices or vectors of the same size. Denote by $(X, Y) = \operatorname{tr}(XY^*)$ the usual inner product on matrices, and denote by ν^D and $\|\cdot\|^D$ the dual norms of ν and $\|\cdot\|$, respectively, *i.e.*,

$$\nu^D(y) = \max\{|(y, x)| : \nu(x) \le 1\} \quad \text{and} \quad \|B\|^D = \{|(B, A)| : \|A\| \le 1\}.$$

Received by the editors April 30, 2002; revised September 20, 2002.

The first author was supported by an NSF (USA) grant, and the second author was supported by an NSERC (Canada) grant.

AMS subject classification: 15A60, 15A04, 47A12, 47A30.

Keywords: Numerical range, numerical radius, state, isometry.

We note that in our version of duality, the identification of $(\mathbf{C}^n, \nu^D)$ with the dual space of $(\mathbf{C}^n, \nu)$, is conjugate-linear rather than linear, *i.e.*, the linear functional f_y on $(\mathbf{C}^n, \nu)$ induced by a vector y is given by $f_y(x) = \sum_{j=1}^n x_j \overline{y_j}$. This is at variance with standard notation in classical Banach space theory, but consistent with Hilbert space duality. Similarly the duality (X, Y) between $(M_n, \|\cdot\|)$ and $(M_n, \|\cdot\|^D)$ is variably given in the literature as $\operatorname{tr}(XY^*)$, as we do, and also $\operatorname{tr}(XY^t)$ or $\operatorname{tr}(XY)$.

A *state* on a normed algebra $\mathcal{A}$ (with identity I of norm one) is a linear functional f on $\mathcal{A}$ such that $f(I) = \|f\| = 1$. With the usual identification of M_n with its dual, the set Σ of states is then identified with the set $\mathcal{S} \subset M_n$ described below, which we shall also refer to as the set of states on M_n with respect to the norm $\|\cdot\|$.

$$\mathcal{S} = \{B \in M_n : \operatorname{tr} B = \|B\|^D = 1\}. \tag{1.1}$$

The following subset

$$\mathcal{R} = \{xy^* : x, y \in \mathbf{C}^n, \nu^D(x) = \nu(y) = x^*y = 1\} \tag{1.2}$$

is called the set of *vector states*. There are two *norm numerical ranges* of $A \in M_n$ associated with ν. The *spatial numerical range*

$$\begin{aligned} W_\nu(A) &= \{(A, Z) : Z \in \mathcal{R}\} \\ &= \{x^*Ay : x, y \in \mathbf{C}^n, \nu^D(x) = \nu(y) = x^*y = 1\}, \end{aligned} \tag{1.3}$$

and the *algebra numerical range*

$$V_\nu(A) = \{(A, Z) : Z \in \mathcal{S}\}. \tag{1.4}$$

When there is no ambiguity about the norm ν, we simply write

$$W(A) = W_\nu(A) \quad \text{and} \quad V(A) = V_\nu(A).$$

It is known (see [3, p. 84] and Corollary 2.2 below) that $V(A)$ is the convex hull of $W(A)$, *i.e.*,

$$V(A) = \operatorname{conv} W(A). \tag{1.5}$$

In view of this, there is only one *norm numerical radius* associated with the numerical ranges, namely,

$$r_\nu(A) = \max\{|\mu| : \mu \in W(A)\} = \max\{|\mu| : \mu \in V(A)\}.$$

Again, when there is no ambiguity about ν, we simplify write

$$r(A) = r_\nu(A).$$

The numerical radius is a norm on M_n and (see [3, p. 34])

$$\mathbf{e}^{-1}\|A\| \le r(A) \le \|A\|, \tag{1.6}$$

where **e** is the Euler constant. It is obvious that $W(A)$ includes every eigenvalue of A and hence

$$\rho(A) \le r(A), \tag{1.7}$$

where $\rho(A)$ is the spectral radius of A.

The standard reference for norm numerical ranges is [3]. See also [2, 17].

When ν is the ℓ_2 norm, $\mathcal{S}$ is the set of positive semidefinite matrices with trace 1 and $W(A) = V(A)$ is the classical numerical range of A acting on the n-dimensional Hilbert space $\mathbf{C}^n$, which has been studied extensively; see [8, 9] for background. In this case, the linear preservers of the numerical range are known. A result of Pellegrini [18] implies that a linear operator ϕ on M_n satisfies $V\big(\phi(A)\big) = V(A)$ for all $A \in M_n$ if and only if its dual transformation ϕ^* satisfies $\phi^*(\mathcal{S}) = \mathcal{S}$. If ν is the ℓ_2 norm, then one may use a result of Kadison [10] on state preserving maps to deduce that there is a unitary $U \in M_n$ such that ϕ has the form

$$A \mapsto U^*AU \quad \text{or} \quad A \mapsto U^*A^tU.$$

In this article, we consider linear maps on M_n that preserve numerical ranges and radii induced by *symmetric norms* ν on $\mathbf{C}^n$, *i.e.*, norms ν that satisfy $\nu(Px) = \nu(x)$ for every P that is either a permutation matrix or a diagonal unitary matrix. (Some authors refer to these norms as *symmetric and absolute*.) We give a complete characterization of those linear operators ϕ on M_n satisfying

$$F\big(\phi(A)\big) = F(A) \quad \text{for all } A \in M_n, \tag{1.8}$$

where $F(A) = W(A)$, $V(A)$ or $r(A)$. A linear operator ϕ on M_n satisfying (1.8) is called a *linear preserver* of the function F.

It is evident that if ν is any norm on $\mathbf{C}^n$ and if T is a linear isometry of $(\mathbf{C}^n, \nu)$, then the "inner" map $A \mapsto T^{-1}AT$ preserves each of the two numerical ranges. The main result of Section 3 is that the converse is also true when ν is a symmetric norm other than multiples of the ℓ_1, ℓ_2 or ℓ_∞ norms. *i.e.*, the numerical range preservers are all inner. These results differ from the ℓ_2 results in as much as the transpose map is no longer present and that the group of isometries of the underlying space $(\mathbf{C}^n, \nu)$ is much smaller. We also give in Section 5 a complete description for the preservers of the numerical ranges when ν is the ℓ_1 or the ℓ_∞ norm. The preservers of the spatial numerical range are also inner, but there are more linear preservers of the algebra numerical range.

In Section 4, we identify the numerical radius isometries, *i.e.*, the numerical radius preservers. When ν is a symmetric norm other than multiples of the ℓ_1 or the ℓ_∞ norm, such maps are unimodular scalar multiples of the numerical range preservers. This is again not true for the ℓ_1 or the ℓ_∞ norms and we characterize the numerical radius preservers in these exceptional cases in Section 5.

We end this section by fixing notation and terminology.

By a *complex unit*, we mean a complex number of modulus one. Also the term "*unimodular complex number*" is used synonymously. Vectors in $\mathbf{C}^n$ are always assumed to be column vectors so that xy^* is an $n \times n$ matrix, while y^*x is the inner product of x and y. We always assume that $\|\cdot\|$ is the operator norm on M_n induced by a symmetric norm ν on $\mathbf{C}^n$ unless specified otherwise. Since one may replace ν by $\gamma\nu$ for any $\gamma > 0$ without changing $\|\cdot\|$, $W(A)$ and $V(A)$, we often assume that $\nu(e_1) = 1$.

Furthermore, we use the following notation and terminology.

- $\{e_1, \dots, e_n\}$: the standard basis for $\mathbf{C}^n$,
- $e = e_1 + \cdots + e_n$,
- $(x, y) = y^*x$: the usual inner product on $\mathbf{C}^n$,
- $\ell_p(x)$: The ℓ_p norm of $x \in \mathbf{C}^n$; $(1 \le p \le \infty)$,
- $\{E_{11}, E_{12}, \dots, E_{nn}\}$: the standard basis for M_n,
- $(X, Y) = \operatorname{tr}(XY^*)$: the usual inner product in M_n,
- conv S: the convex hull of a given set S,
- Ext S: the set of extreme points of a compact convex set S,
- $\mathcal{E} = \operatorname{Ext} \mathcal{B}$, where $\mathcal{B} = \{x \in \mathbf{C}^n : \nu(x) \le 1\}$,
- $\mathcal{E}^D = \operatorname{Ext} \mathcal{B}^D$, where $\mathcal{B} = \{x \in \mathbf{C}^n : \nu^D(x) \le 1\}$,
- $\mathcal{E}_{\|\cdot\|^D} = \operatorname{Ext} \mathcal{B}_{\|\cdot\|^D}$, where $\mathcal{B}_{\|\cdot\|^D} = \{X \in M_n : \|X\|^D \le 1\}$.
- $\mathrm{GP}(n)$: the group of generalized permutation matrices in M_n, *i.e.*, the group generated by permutation matrices and diagonal unitary matrices.

For $A \in M_n$, the sets

- $\mathcal{D}(A) = \{DAD^* : D \text{ is a diagonal unitary}\}$,
- $\mathcal{P}(A) = \{PAP^t : P \text{ is a permutation}\}$, and
- $\mathcal{GP}(A) = \{QAQ^* : Q \in \mathrm{GP}(n)\}$

will be called the *diagonal-unitary orbit* of A, *the permutation orbit* of A and the *generalized permutation orbit* of A (or the GP-*orbit* of A) respectively.

2 Auxiliary Results

We begin with some general results on operator norms not necessarily induced by symmetric norms on $\mathbf{C}^n$. Some of them are well known. We will mention some convenient references or give short proofs for completeness. We start with the following proposition, whose proof may be found in [15, Proposition 4.1].

Proposition 2.1 *Let ν be a norm on $\mathbf{C}^n$. Then*

$$\mathcal{E}_{\|\cdot\|^D} = \{xy^* : x \in \mathcal{E}^D, y \in \mathcal{E}\},$$

and

$$\mathcal{B}_{\|\cdot\|^D} = \operatorname{conv}\{xy^* : x \in \mathcal{E}^D, y \in \mathcal{E}\} = \operatorname{conv}\{xy^* : \nu^D(x) = \nu(y) = 1\}.$$

We use Proposition 2.1 to give a proof for the following known result. We mention, in passing, that the extreme points of the set $\mathcal{S}$ of states are called *pure states*.

Corollary 2.2 *Let ν be a norm on $\mathbf{C}^n$ and let $\mathcal{S}$ and $\mathcal{R}$ be the corresponding set of states and set of vector states respectively. Then*

$$\operatorname{conv}\mathcal{R} = \mathcal{S}, \quad V(A) = \operatorname{conv} W(A), \quad \textit{and} \quad \operatorname{Ext}\mathcal{S} = \mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D}. \tag{2.1}$$

Proof It is easy to verify the well-known fact that $\mathcal{S}$ is convex. Indeed, it is the intersection of the unit ball $\mathcal{B}_{\|\cdot\|^D}$ and the set $\mathcal{T}$ of matrices of trace 1 since $1 = \operatorname{tr} S = (S, I)$ implies that $\|S\|^D \geq 1$. The convexity of each of $\mathcal{B}_{\|\cdot\|^D}$ and $\mathcal{T}$ is quite easy to see.

It is obvious that $\mathcal{R} \subseteq \mathcal{S}$ and hence $\operatorname{conv}\mathcal{R} \subseteq \mathcal{S}$ To prove the reverse inclusion, let $B \in \mathcal{S}$. Since $\mathcal{S} \subseteq \mathcal{B}_{\|\cdot\|^D}$, then by Proposition 2.1, B is a convex combination of elements of the form xy^* with $\nu^D(x) = \nu(y) = 1$. Since $|\operatorname{tr}(xy^*)| = |y^*x| \leq \nu^D(x)\nu(y) = 1$, each matrix xy^* in the convex combination must satisfy $1 = \operatorname{tr}(xy^*) = y^*x$ to ensure that $\operatorname{tr} B = 1$. We conclude that $B \in \operatorname{conv}\mathcal{R}$.

For the second equality in the corollary, we have

$$\begin{aligned}
\operatorname{conv} W(A) &= \operatorname{conv}\{(A, Z) : Z \in \mathcal{R}\} \\
&= \{(A, Z) : Z \in \operatorname{conv}\mathcal{R}\} \\
&= \{(A, Z) : Z \in \mathcal{S}\} \\
&= V(A).
\end{aligned}$$

Finally, we consider the third equality. Since $\mathcal{S} = \operatorname{conv}\mathcal{R}$ and $\mathcal{S} \subseteq \mathcal{B}_{\|\cdot\|^D}$, then $\operatorname{Ext}\mathcal{S} \subseteq \mathcal{R}$ and $\operatorname{Ext}\mathcal{S} \subseteq \mathcal{E}_{\|\cdot\|^D}$. Thus, $\operatorname{Ext}\mathcal{S} \subseteq \mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D}$. For the reverse inclusion, if $X \in \operatorname{Ext}\mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D}$, then $X \in \mathcal{S}$ and X cannot be written as the convex combination of two different matrices in $\mathcal{S} \subseteq \mathcal{E}_{\|\cdot\|^D}$. Hence, $X \in \operatorname{Ext}\mathcal{S}$. ■

In the following proposition, we use the notation $\bar{S} = \{\bar{\mu} : \mu \in S\}$ for $S \subseteq \mathbf{C}$. The first assertion of the proposition can be found (with different notation) in [3, p. 85].

Proposition 2.3 *Let $W_\nu(A)$ or $V_\nu(A)$ be the numerical ranges associated with a norm ν on $\mathbf{C}^n$. For $F_\nu(A) = W_\nu(A)$ or $V_\nu(A)$, we have*

$$\overline{F_\nu(A)} = F_{\nu^D}(A^*) \quad \textit{and} \quad r_\nu(A) = r_{\nu^D}(A^*).$$

If, in addition, $\nu(x) = \nu(\bar{x})$ for all $x \in \mathbf{C}^n$, which is true for a symmetric norm ν, then

$$F_\nu(A) = F_{\nu^D}(A^t) \quad \textit{and} \quad r_\nu(A) = r_{\nu^D}(A^t).$$

Proof Observe that

$$\begin{aligned}
\overline{W_\nu(A)} &= \{\overline{x^*Ay} : \nu^D(x) = \nu(y) = x^*y = 1\} \\
&= \{(x^*Ay)^* : \nu^D(x) = \nu(y) = x^*y = 1\} \\
&= \{y^*A^*x : \nu^D(x) = \nu(y) = x^*y = 1\} \\
&= W_{\nu^D}(A^*).
\end{aligned}$$

The first assertion on F_ν and r_ν follows. If $\nu(y) = \nu(\bar{y})$ for all $y \in \mathbf{C}^n$, then $\nu^D(x) = \nu^D(\bar{x})$ for all $x \in \mathbf{C}^n$. It follows that

$$\begin{aligned} W_\nu(A) &= \{(x^*Ay)^t : \nu^D(x) = \nu(y) = x^*y = 1\} \\ &= \{\bar{y}^*A^t\bar{x} : \nu^D(x) = \nu(y) = x^*y = 1\} \\ &= W_{\nu^D}(A^t). \end{aligned}$$

The second assertion follows. ∎

Corollary 2.4 *Suppose $F_\nu(A) = W_\nu(A)$, $V_\nu(A)$ or $r_\nu(A)$. A mapping $\phi\colon M_n \to M_n$ satisfies*

$$F_\nu\big(\phi(A)\big) = F_\nu(A) \quad \text{for all } A \in M_n \tag{2.2}$$

if and only if the mapping $\tilde{\phi} : M_n \to M_n$ defined by $\tilde{\phi}(A) = \phi(A^*)^*$ satisfies

$$F_{\nu^D}\big(\tilde{\phi}(A)\big) = F_{\nu^D}(A) \quad \text{for all } A \in M_n. \tag{2.3}$$

If, in addition, $\nu(x) = \nu(\bar{x})$ for all $x \in \mathbf{C}^n$, in particular if ν is a symmetric, then (2.2) holds if and only if the mapping $\tilde{\phi}\colon M_n \to M_n$ defined by $\tilde{\phi}(A) = \phi(A^t)^t$ satisfies (2.3).

Here is another well known result needed in our discussion.

Lemma 2.5 *Let W_ν and V_ν be the numerical ranges associated with a norm ν on $\mathbf{C}^n$. Then a matrix $A \in M_n$ is such that any one (or both) of the sets $W_\nu(A)$ or $V_\nu(A)$ equals $\{\mu\}$ if and only if $A = \mu I$.*

The rest of the results in this section concern operator norms on M_n induced by symmetric norms ν on $\mathbf{C}^n$.

Lemma 2.6 *Suppose $\|\cdot\|$ is the operator norm on M_n induced by a symmetric norm ν on $\mathbf{C}^n$. Let $\mathcal{S}$ and $\mathcal{R}$ be the set of states and the set of vector states corresponding to ν, respectively. If $xy^* \in \mathcal{R}$, then there exists a generalized permutation $Q \in \mathrm{GP}(n)$ such that both Qx and Qy have nonnegative entries in descending order. Consequently, all matrices in $\mathcal{R}$ and $\mathcal{S}$ have nonnegative diagonal entries.*

Proof Suppose Q is the generalized permutation matrix satisfying $Qx = (x_1, \dots, x_n)^t$ with $x_1 \geq \cdots \geq x_n \geq 0$. If $Qy = (y_1, \dots, y_n)^t$ is not a nonnegative vector, then there exists a diagonal unitary matrix D such that $DQxy^*Q^*D^* \in \mathcal{R}$ has positive trace larger than $\operatorname{tr}(xy^*) = 1$. Thus $\nu(y)\nu^D(x) = \nu(DQy)\nu^D(DQx) \geq \operatorname{tr}(DQxy^*Q^*D^*) > 1$, a contradiction. This establishes that Qy is a nonnegative vector. Furthermore, if $y_j < y_{j+1}$, then we claim that $x_j = x_{j+1}$; otherwise, we can let P be the permutation matrix obtained from I be interchanging the j-th and $(j+1)$-st rows so that $PQxy^*Q^*P^* \in \mathcal{R}$ has trace larger than $\operatorname{tr}(xy^*) = 1$. We may replace Q by PQ. After at most $n-1$ of such modifications, the resulting matrix Q will satisfy the asserted property. ∎

The following is a key lemma in this paper and will be extensively used in the sequel.

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

Lemma 2.7 *Let ν be a symmetric norm on $\mathbf{C}^n$, and $\|\cdot\|$ be the corresponding operator norm on M_n.*

(a) *If $c = \nu(e_1)$ then $c\ell_\infty(x) \le \nu(x) \le c\,\ell_1(x)$ for every $x \in \mathbf{C}^n$.*
(b) *At least one of the vectors $e/\nu(e)$ and $e/\nu^D(e)$ belongs to the set of extreme points $\mathcal{E}$ or $\mathcal{E}^D$ of the corresponding unit ball respectively.*
(c) *$\nu(e)\nu^D(e) = n$.*
(d) *If $e/\nu(e) \in \mathcal{E}$ then*
 (i) *for every nonnegative $x \in \mathcal{B}^D$, the unit ball of ν^D, we have $e^t x \le \nu(e)$;*
 (ii) *for every nonnegative $y \in \mathcal{B}$, the unit ball of ν, we have $e^t y \le n/\nu(e)$ and equality holds if and only if $y = e/\nu(e)$;*
 (iii) *there exists $u = (u_1, \dots, u_n)^t$ with $u_1 \ge \cdots \ge u_n \ge 0$ and $u_1 + \cdots + u_n = 1$ such that $ue^t \in \operatorname{Ext} \mathcal{S}$, and so the QP-orbit $\mathcal{D}_1$ of ue^t satisfies*

$$\mathcal{D}_1 = \{Q^* ue^t Q : Q \in \mathrm{GP}(n)\} \subseteq \operatorname{Ext} \mathcal{S}. \tag{2.4}$$

 Furthermore $u = e_1$ if and only if ν is a multiple of the ℓ_∞ norm.
(e) *If $e/\nu^D(e) \in \mathcal{E}^D$ then*
 (i) *for every nonnegative $y \in \mathcal{B}$, we have $e^t y \le \nu^D(e)$;*
 (ii) *for every nonnegative $x \in \mathcal{B}^D$, we have $e^t x \le n/\nu^D(e)$ and equality holds if and only if $x = e/\nu^D(e)$;*
 (iii) *there exists $v = (v_1, \dots, v_n)^t$ with $v_1 \ge \cdots \ge v_n \ge 0$ and $v_1 + \cdots + v_n = 1$ such that $ev^t \in \operatorname{Ext} \mathcal{S}$, and so the QP-orbit $\mathcal{D}_2$ of ev^t satisfies*

$$\mathcal{D}_2 = \{Q^* ev^t Q : Q \in \mathrm{GP}(n)\} \subseteq \operatorname{Ext} \mathcal{S}. \tag{2.5}$$

 Furthermore $v = e_1$ if and only if ν is a multiple of the ℓ_1 norm.
(f) *The following conditions are equivalent to each other:*
 (i) *there exist vectors u and v in $\mathbf{C}^n$ such that both ue^t and ev^t belong to* Ext $\mathcal{S}$*;*
 (ii) *$e/\nu(e) \in \mathcal{E}$ and $e/\nu^D(e) \in \mathcal{E}^D$;*
 (iii) *$ee^t/n \in$* Ext $\mathcal{S}$.

Proof The inequality $\nu(x) \le c\,\ell_1(x)$ follows easily from the triangle inequality. The other inequality in part (a) follows by duality.

To prove part (b), let $\gamma = \nu(e)$. If $e/\gamma \notin \mathcal{E}$, then it is a convex combination of a finite set $\mathcal{Z} \subseteq \mathcal{E}$. Let w be a vector in $\mathcal{E}^D$ such that $w^t e/\gamma = 1$. We shall show that w is scalar multiple of e. By Lemma 2.6, we have that w is a nonnegative vector. For every $z \in \mathcal{Z}$ we have $(w, |z|) \le 1$, but $1 = (w, e/\gamma)$ is a convex combination of the numbers (w, z) for $z \in \mathcal{Z}$. It follows that $(w, z) = (w, |z|) = 1$. Since we may replace w by Pw for any permutation P, we get that $(Pw, z) = (Pw, |z|) = 1$ for every permutation P. This easily implies that z is nonnegative and that either z or w is a multiple of e. By assumption z cannot be, so w is. In other words $e/\nu^D(e) \in \mathcal{E}^D$. This proves part (b).

Part (c) will be proved after parts (d) and (e).

To prove part (d), we again let $\gamma = \nu(e)$. There exists a vector $w = (w_1, \dots, w_n)^t \in \mathcal{E}^D$ such that $w^*e = \gamma$. By Lemma 2.6, the vector w is nonnegative. We may, with no loss of generality, assume that $w_1 \geq \cdots \geq w_n$. Now let $u = w/\gamma$. Then $ue^t = we^t/\gamma \in \mathcal{E}_{\|\cdot\|^D}$ by Proposition 2.1. By Corollary 2.2, we conclude that $\mathcal{D}_1 \subseteq \mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D} = \operatorname{Ext} \mathcal{S}$. This proves the set inclusion part of (d).

If x is a nonnegative vector in $\mathcal{B}^D$, then

$$\sum_{j=1}^{n} x_j = (x, e) \leq \nu^D(x)\nu(e) = \nu(e).$$

If y is a nonnegative vector in $\mathcal{B}$, let z be the average of Py as P runs through all permutations. Therefore $\nu(z) \leq 1$. But $z = (e^t y)e/n$. It follows that $e^t y \leq n/\nu(e)$. If equality holds, then $z = e/\nu(e)$ is an extreme point of the unit ball of ν and is at the same time a convex combinations of the vectors Py in the unit ball. It follows that $y = e/\nu(e)$.

Next, we turn our attention to the last assertion of the last part of (d). It is obvious that if ν is a multiple of the ℓ_∞ norm, then $Q^*e_1e^tQ \in \operatorname{Ext} \mathcal{S}$ for every $Q \in \mathrm{GP}(n)$. Conversely, if $e_1e^t \in \operatorname{Ext} \mathcal{S}$, then by Proposition 2.1 and Corollary 2.2, we have that (after normalisation) $e_1 \in \mathcal{E}^D$ and $e \in \mathcal{E}$. In particular $\nu(e) = 1$. The set of extreme points of the unit ball of the ℓ_∞ norm is precisely the set $\{Qe : Q \in \mathrm{GP}(n)\}$. Thus every vector in the unit ball of the ℓ_∞ norm belongs to $\operatorname{conv}\{Qe : Q \in \mathrm{GP}(n)\} \subseteq \mathcal{B}$. This shows that $\nu(x) \leq \ell_\infty(x)$ for every $x \in \mathbf{C}^n$. The reverse inequality follows from part (a) and so the normalised ν is the ℓ_∞ norm.

The proof of part (e) follows by duality.

Finally, we consider part (f). The implication (iii) $\Rightarrow$ (i) is clear and the implication (i) $\Rightarrow$ (ii) follows from Proposition 2.1 and Corollary 2.2. If (ii) is satisfied, then using the same two results together with part (c), we get that

$$\frac{ee^t}{n} = \frac{e}{\nu(e)} \frac{e^t}{\nu^D(e)} \in \mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D} = \operatorname{Ext} \mathcal{S}. \qquad \blacksquare$$

Lemma 2.8 *Suppose ν is a symmetric norm on $\mathbf{C}^n$ not equal to a multiple of the ℓ_2-norm.*

(a) *For every $k \in \{1, \dots, n\}$, and for every $Q \in \mathrm{GP}(n)$, we have*

$$Q(e_1 + \cdots + e_k)(e_1 + \cdots + e_k)^t Q^*/k \in \mathcal{S},$$

in particular $ee^t/n \in \mathcal{S}$ and $E_{jj} \in \mathcal{S}$ for every j.

(b) *There exists an element of the form $xy^* \in \operatorname{Ext} \mathcal{S}$ such that $x, y \in \mathbf{C}^n$ are not multiples of each other.*

(c) *$\mathcal{S}^t \neq \mathcal{S} \neq \mathcal{S}^*$ and $\mathcal{R}^t \neq \mathcal{R} \neq \mathcal{R}^*$.*

(d) *There exists a matrix of the form $S = aE_{12} + bE_{21} + \sum_{j=1}^{n} d_jE_{jj}$ in $\mathcal{S}$ with $a > b \geq 0$.*

Proof To prove (a), suppose $y = (e_1 + \cdots + e_k)$ and $\nu(y) = \gamma$. Then there exists $x = (x_1, \dots, x_n)^t \in \mathcal{E}^D$ such that $(x, y)/\gamma = 1$, and hence $x_1 + \cdots + x_k = \gamma$. Let $\tilde{x} = (x_1, \dots, x_k, 0, \dots, 0)^t$, where the zeros are absent if $k = n$. Note that

$$\tilde{x} = \big(x + (I_k \oplus -I_{n-k})x\big) \big/ 2 \in \mathcal{B}^D.$$

Let P be the permutation matrix $E_{12} + \cdots + E_{k-1,k} + E_{k,1} + \sum_{j>k} E_{jj}$. Then

$$u = \Big(\sum_{j=1}^{k} P^j \tilde{x}\Big) \Big/ k = \gamma(e_1 + \cdots + e_k)/k \in \mathcal{B}^D.$$

Thus,

$$uy^t/\gamma = (e_1 + \cdots + e_k)(e_1 + \cdots + e_k)^t/k \in \operatorname{conv} \mathcal{R} = \mathcal{S}.$$

Clearly, for any $Q \in \mathrm{GP}(n)$, we have $Q(uy^t/\gamma)Q^* \in \mathcal{S}$.

To prove part (b) assume, to the contrary, that every element in $\operatorname{Ext} \mathcal{S}$ has the form xx^* for some vector $x \in \mathbf{C}^n$. It follows that every element of $\mathcal{S}$ is self-adjoint. Therefore every element of $\mathcal{R}$ has the form xx^* for some vector $x \in \mathbf{C}^n$. If $x \in \mathbf{C}^n$ satisfies $\nu(x) = 1$, then there exists $y \in \mathbf{C}^n$ such that $1 = \nu^D(y) = (x, y)$. By our assumption, we have $y = x$ and hence $1 = (x, x) = \ell_2(x)^2$. It follows that ν is the ℓ_2 norm, a contradiction.

For part (c) we will only prove the first inequality; the second inequality may be proved by a similar argument and the last two follow from the first two. So assume that $\mathcal{S} = \mathcal{S}^t$. It follows that $\operatorname{Ext} \mathcal{S} = (\operatorname{Ext} \mathcal{S})^t$. By part (b) there exist linearly independent vectors $x \in \mathcal{E}^D$ and $y \in \mathcal{E}$ such that the matrix $A = xy^* \in \operatorname{Ext} \mathcal{S}$. By Lemma 2.6, we may assume that x and y have nonnegative entries. Now, $yx^t = A^t \in \operatorname{Ext} \mathcal{S}$. Hence there is a positive number r such that $ry \in \mathcal{E}^D$ and $x/r \in \mathcal{E}$. But then

$$1 \geq (x, x/r)(y, ry) = (x, x)(y, y) \geq (x, y)(x, y) = (\operatorname{tr} A)^2 = 1$$

implies that x and y are multiples of each other, a contradiction.

For part (d), we again choose linearly independent nonnegative vectors x and y such that $xy^t \in \mathcal{S}$. Since the matrix xy^t is not symmetric, there exists a permutation matrix P such that Pxy^tP^t has its $(1, 2)$ and $(2, 1)$ entries equal to a and b with $a > b \geq 0$. Let $A = Pxy^tP^t$ and $D = -I_2 \oplus I_{n-2}$. By the convexity of $\mathcal{S}$, the matrix $A_0 = (DAD + A)/2$ belongs to $\mathcal{S}$. But $A_0 = B \oplus C$ where $B \in M_2$ is the top left 2×2 corner of A. Now, set

$$S = \sum_D (DA_0D)/2^{n-2} \in \mathcal{S},$$

where D ranges over all diagonal orthogonal matrices whose $(1, 1)$ and $(2, 2)$ entries are equal to 1. Then S has the form described in (d). ■

Let ν be a norm on $\mathbf{C}^n$. Then $A \in M_n$ is said to be *ν-hermitian* if $W_\nu(A)$ (or equivalently $V_\nu(A)$) is contained in $\mathbf{R}$. If ν is the ℓ_2 norm, this reduces to the usual notion of hermitian (or self-adjoint) matrices. (ν-positive definite or semi-definite matrices may be defined analogously.) The ν-hermitian matrices associated with an absolute norm have been characterized in [21]. Specializing their results to symmetric norms, we have the following corollary, which can also be derived from our previous lemmas as we presently show.

Corollary 2.9 *Let ν be a symmetric norm on $\mathbf{C}^n$ not equal to a multiple of the ℓ_2-norm. Then a matrix $A \in M_n$ is ν-hermitian if and only if A is a diagonal matrix and the diagonal entries are real. If D is such a diagonal hermitian with diagonal entries d_{jj}, then $V_\nu(D) = W_\nu(D) = \operatorname{conv}\{d_{jj} : 1 \le j \le n\}$*

Proof Note that every state S in $\mathcal{S}$ satisfies $\operatorname{tr} S = 1$ and, by Lemma 2.6, has nonnegative diagonal entries. If A is a diagonal matrix with real diagonal entries, then (A, S) is real. Thus $V(A) = \{(A, S) : S \in \mathcal{S}\}$ is a subset of $\mathbf{R}$.

Conversely, suppose $V(A)$ is a subset of $\mathbf{R}$. Since $E_{jj} \in \mathcal{S}$, we have (A, E_{jj}) is real for all $j = 1, \dots, n$. Thus A has real diagonal entries. Suppose A has a nonzero (j, k) entry a_{jk} for some $j \ne k$. By Lemma 2.8(d), there exists a matrix $B = \sum_{j=1}^n d_j E_{jj} + aE_{12} + bE_{21}$ with $a > b \ge 0$ in $\mathcal{S}$. For every $s \in [0, 2\pi)$ there exists a $Q_s \in \mathrm{GP}(n)$ such that $Q_s B Q_s^*$ has (j, k) entry $e^{is}a$ and (k, j) entry $e^{-is}b$. Let D_A and D_s denote the diagonal part of A an $Q_s B Q_s^*$, respectively. Then both of them are real matrices and thus $\operatorname{tr}(D_A D_s) \in \mathbf{R}$. But then for all $s \in [0, 2\pi)$, we have

$$(A, Q_s B Q_s^t) = \operatorname{tr}(D_A D_s) + a_{jk} a e^{-is} + a_{kj} b e^{is} \in \mathbf{R}.$$

Thus $a_{jk}a$ and $a_{kj}b$ are complex conjugates. It follows that $|a_{jk}| < |a_{kj}|$.

On the other hand for every permutation P, we have that $P^t \mathcal{S} P = \mathcal{S}$, which implies that $P^t A P$ is also a ν-hermitian matrix. If we take P to be the transposition that transposes e_j and e_k, we conclude also that $|a_{jk}| > |a_{kj}|$, which is impossible. Hence A cannot have non-zero off-diagonal entries.

If D is a diagonal hermitian and S is a state, then it is obvious that (D, S) is a convex combination of the diagonal entries of D. On the other hand, every eigenvalue d_{jj} of D evidently belongs to the spatial numerical range and hence to the algebra numerical range of D. The convexity of the algebra numerical range now implies that $V_\nu(D) = \operatorname{conv}\{d_{jj} : 1 \le j \le n\}$. Since the spatial numerical range is always connected [3, p. 102], we also conclude that $W_\nu(D) = \operatorname{conv}\{d_{jj} : 1 \le j \le n\}$. ■

The isometries of the space $(\mathbf{C}^n, \nu)$ for a symmetric norm ν are known (see [20]). As we now have all the ingredients needed for a short proof, we take this opportunity to present it.

Corollary 2.10 *Let ν be a symmetric norm on $\mathbf{C}^n$ not equal to a multiple of the ℓ_2-norm. Then a matrix $U \in M_n$ is an isometry of $(\mathbf{C}^n, \nu)$ if and only if U is a generalized permutation.*

Proof The fact that a generalized permutation is an isometry is nothing more than the definition of a symmetric norm. To prove the converse, assume that U is an isometry. It follows that UHU^{-1} is ν-hermitian for every ν-hermitian matrix H. In particular $UE_{jj}U^{-1}$ is a rank-one ν-hermitian matrix for every j. Therefore $UE_{jj}U^{-1}$ is a scalar multiple of E_{kk} for an index k depending on j. This implies that $Ue_j = \lambda_j e_{\pi(j)}$, for unimodular complex numbers λ_j, and a permutation π of the set $\{1, \dots, n\}$, *i.e.*, U is a product of a permutation matrix and a diagonal unitary matrix as required. ■

3 States and Numerical Range Preservers

In this section, we characterize linear operators on M_n that preserve the states or the vector states or any of the two norm numerical ranges arising from a symmetric norm which is not a multiple of the ℓ_1, ℓ_2 or ℓ_∞ norm.

Theorem 3.1 *Let ν be a symmetric norm not equal to a multiple of the ℓ_q norm with $q \in \{1, 2, \infty\}$. The following conditions are equivalent for a linear operator ϕ on M_n.*

(a) *ϕ preserves the spatial numerical range, i.e., $W\big(\phi(A)\big) = W(A)$ for all $A \in M_n$.*
(b) *ϕ preserves the algebra numerical range, i.e., $V\big(\phi(A)\big) = V(A)$ for all $A \in M_n$.*
(c) *There exists a generalized permutation $Q \in \mathrm{GP}(n)$ such that*

$$\phi(A) = Q^*AQ \quad \forall A \in M_n,$$

or equivalently,

$$\phi^*(A) = QAQ^* \quad \forall A \in M_n.$$

(d) $\phi^*(\mathcal{R}) = \mathcal{R}$.
(e) $\phi^*(\mathcal{S}) = \mathcal{S}$.

Proof The equivalence of the form for ϕ and the form for ϕ^* in part (c) follows from

$$(Q^*AQ, B) = \mathrm{tr}(Q^*AQB^*) = \mathrm{tr}(AQB^*Q^*) = (A, QBQ^*).$$

Next, we consider the following chain of implications: (c) $\Rightarrow$ (d) $\Rightarrow$ (a) $\Rightarrow$ (b) $\Rightarrow$ (e)

The implications (c) $\Rightarrow$ (d) and (d) $\Rightarrow$ (a) are readily verified. The implication (a) $\Rightarrow$ (b) follows the fact that $V(A) = \mathrm{conv}\, W(A)$ and the implication (b) $\Rightarrow$ (e) follows from the result in [18].

It remains to establish (e) $\Rightarrow$ (c). We divide the proof into several lemmas. In the rest of this section, *we always assume that $\mathcal{S} \subseteq M_n$ is the set of states arising from a symmetric norm ν on $\mathbf{C}^n$, and $\psi = \phi^*$ is a linear operator on M_n satisfying $\psi(\mathcal{S}) = \mathcal{S}$.*

Lemma 3.2 *Let ψ be a linear map on M_n such that $\psi(\mathcal{S}) = \mathcal{S}$, where $\mathcal{S}$ is the set of states induced by a symmetric norm ν which is not a multiple of the ℓ_2 norm. Then*

(a) *ψ preserves the algebra numerical range;*
(b) *ψ maps the set of diagonal matrices onto itself;*
(c) *If $\tilde{\mathcal{E}} = \{E_{jj} : 1 \le j \le n\}$ then $\psi(\tilde{\mathcal{E}}) = \tilde{\mathcal{E}}$.*

Furthermore, there exists a permutation matrix P such that the operator $\tilde{\psi}$ defined by $X \mapsto P^*\psi(X)P$

(i) *preserves the usual inner product on M_n,*
(ii) *fixes every diagonal matrix in M_n, and*
(iii) *maps the set of matrices with zero diagonal onto itself.*

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

Proof Let $\mathcal{G}$ be the set of all linear operators on M_n that map $\mathcal{S}$ onto itself. Since $\mathcal{S}$ is a compact set that spans M_n, it is easy to see that $\mathcal{G}$ is a compact group of invertible linear operators on M_n. Viewing M_n as a Hilbert space under the Frobenius (or Hilbert-Schmidt) norm, then using a result in [1] (see also [7]), we get that there exists a positive definite linear operator ξ on M_n such that $\xi\mathcal{G}\xi^{-1}$ is a subgroup of the group of unitary operators on M_n.

Note that for any $Q \in \mathrm{GP}(n)$ the linear operator ψ_Q defined by $A \mapsto Q^*AQ$ is a member of $\mathcal{G}$. Thus $\xi\psi_Q\xi^{-1}$ is a unitary operator, *i.e.*, $(\xi\psi_Q\xi^{-1})^*(\xi\psi_Q\xi^{-1})$ is the identity operator. It follows that $\xi^2\psi_Q = \psi_Q\xi^2$ for all $Q \in \mathrm{GP}(n)$.

Let $\mathcal{G}_0$ be the group of operators ψ_Q, where $Q \in \mathrm{GP}(n)$. Then $\mathcal{G}_0$ has three irreducible subspaces on M_n, namely, the span of I, the space of trace zero diagonal matrices, and the spaces of matrices with zero diagonal. If $n > 2$, the three subspaces have different dimensions, namely, 1, $n-1$, n^2-n. By Schur's lemma, (see, *e.g.* [6, p. 182]), these are also reducing subspaces of ξ^2, and ξ^2 acts as a scalar operator on each of them. If $n = 2$, the dimensions of the first two irreducible subspaces are equal and so ξ^2 may apparently interchange them. But then the matrix of ξ^2 with respect to the orthonormal basis $\{I_2, E_{11}-E_{22}, E_{12}, E_{21}\}$ will then be of the form

$$\begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix} \oplus \gamma I_2,$$

which contradicts the fact that ξ^2 is the square of a positive definite operator on M_2. Thus, we conclude that ξ^2 acts as scalar operators on the three irreducible subspaces of $\mathcal{G}_0$, and so does ξ. Therefore $\xi(I) = \lambda I$ for a scalar λ and $\xi(\mathcal{D}) = \mathcal{D}$, where $\mathcal{D}$ is the space of diagonal matrices.

For every matrix $A \in M_n$, we have

$$\begin{aligned} V\big(\xi^{-2}\psi\xi^2(A)\big) &= \{(\xi^{-2}\psi\xi^2(A), Z) : Z \in \mathcal{S}\} \\ &= \big\{\big(\xi^{-2}\psi\xi^2(A), \psi(Z)\big) : Z \in \mathcal{S}\big\} \quad \text{as } \psi(\mathcal{S}) = \mathcal{S} \\ &= \big\{\big(\xi^{-1}\psi\xi^2(A), \xi^{-1}\psi(Z)\big) : Z \in \mathcal{S}\big\} \quad \text{as } \xi \text{ is self-adjoint} \\ &= \big\{\big(\xi(A), \xi^{-1}(Z)\big) : Z \in \mathcal{S}\big\} \quad \text{as } \xi^{-1}\psi\xi \text{ is unitary on } M_n \\ &= \{(A, Z) : Z \in \mathcal{S}\} = V(A). \end{aligned}$$

That is, $\xi^{-2}\psi\xi^2$ preserves the algebra numerical range. From this, we conclude that:

1. $\xi^{-2}\psi\xi^2(I) = (I)$
2. $\xi^{-2}\psi\xi^2(\mathcal{H}) = \mathcal{H}$, where $\mathcal{H}$ is the set of ν-hermitian elements.
3. i $\xi^{-2}\psi\xi^2(\mathcal{D}) = \mathcal{D}$, where $\mathcal{D}$ is the set of all diagonal matrices, since $\mathcal{D} = \operatorname{span}\mathcal{H}$.

However $\xi(I) = \lambda I$ for a scalar λ and $\xi(\mathcal{D}) = \mathcal{D}$, and so items 1 and 3 above imply that:

4. $\psi(I) = (I)$, and $\psi(\mathcal{D}) = \mathcal{D}$,
5. $\xi^{-1}\psi\xi(I) = I$, and $\xi^{-1}\psi\xi(\mathcal{D}) = \mathcal{D}$.

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

Since $\xi^{-1}\psi\xi$ is a unitary operator on M_n, we also conclude that

6. $\xi^{-1}\psi\xi$ maps the set of zero-diagonal matrices onto itself since the set of zero-diagonal matrices is the orthogonal complement of $\mathcal{D}$.
7. $\xi^{-1}\psi\xi$ maps the set of trace zero diagonal matrices onto itself since this set is the orthogonal complement of $\{I\}$ in $\mathcal{D}$.

Thus $\xi^{-1}\psi\xi$ leaves invariant every one of the three eigenspaces of ξ, and so it commutes with ξ, *i.e.*, $\xi^{-2}\psi\xi^2 = \xi^{-1}\psi\xi = \psi$. We conclude that ψ itself preserves the algebra numerical range, leaves invariant the identity, the set of trace zero diagonal matrices and set of zero-diagonal matrices. Consequently, ψ also preserves the trace. In particular if $\mathcal{H}_1$ is the set of matrices whose algebra numerical range is in the interval $[0, 1]$ and whose trace is 1, then $\psi(\mathcal{H}_1) = \mathcal{H}_1$. By Corollary 2.9, $\mathcal{H}_1$ consists of all diagonal matrices with trace 1 and whose entries are in $[0, 1]$. It is evident that $\tilde{\mathcal{E}}$ is the set of extreme points of $\mathcal{H}_1$ and hence $\psi(\tilde{\mathcal{E}}) = \tilde{\mathcal{E}}$. This proves assertions (a)–(c) of the lemma.

Since $\psi(\tilde{\mathcal{E}}) = \tilde{\mathcal{E}}$, there exists a permutation matrix P such that the map $\tilde{\psi}$ defined by $X \mapsto P^*\psi(X)P$ satisfies $\tilde{\psi}(E_{jj}) = E_{jj}$ for $j = 1, \dots, n$. It then follows that $\tilde{\psi}$ satisfies the asserted properties (i)–(iii). ∎

Lemma 3.3 *Let ψ be as in the preceding lemma. Assume also that ν is not a multiple of the ℓ_1 or the ℓ_∞ norms and let $\mathcal{F} = \{Dee^*D^* : D \text{ is diagonal unitary}\}$. Then $\psi(\mathcal{F}) = \mathcal{F}$.*

Proof By Lemma 3.2, we may assume that ψ acts as the identity on diagonal matrices, preserves the usual inner product on M_n, and maps the set of matrices with zero diagonal onto itself.

Recall that $\mathcal{E}$ and $\mathcal{E}^D$ denote the sets of extreme points of the unit ball of ν and ν^D, respectively and that $\mathcal{E}_{\|\cdot\|^D}$ denotes the set of extreme points of the unit ball of $\|\cdot\|^D$. By Lemma 2.7, at least one of $\mathcal{E}$ and $\mathcal{E}^D$ contains a multiple of e. We shall assume the former as the latter may be treated by a similar argument. So by Lemma 2.7, there is a vector $u = (u_1, \dots, u_n)^t$ with $u_1 \geq \dots \geq u_n = 1$ and $u_1 + \dots + u_n = 1$ such that $\operatorname{Ext} \mathcal{S}$ contains a subset of the form $\mathcal{D}_1$ defined in (2.4). We use this fact to show that $\psi(\mathcal{F}) = \mathcal{F}$.

First, we claim that

$$\psi(\mathcal{D}_1) = \mathcal{D}_1.$$

Note that elements in $\mathcal{D}_1$ are extreme points of $\mathcal{S}$. Thus, $\psi(\mathcal{D}_1) \subseteq \operatorname{Ext} \mathcal{S} = \mathcal{R} \cap \mathcal{E}_{\|\cdot\|^D}$. Suppose $\psi(ue^*) = xy^*$, where $x = (x_1, \dots, x_n)^t$ and $y = (y_1, \dots, y_n)^t$. Since ψ fixes the diagonal entries, we have

$$u_j = x_j y_j \quad \text{for } j = 1, \dots, n. \tag{3.1}$$

By Lemma 2.6, there exists $S \in \mathrm{GP}(n)$ such that Sx and Sy has nonnegative entries arranged in descending order. Since the diagonal entries $u_1, \dots, u_n$ of the matrix xy^* are already in descending order, the matrix S must be a diagonal unitary. Hence

$$|x_1| \geq \dots \geq |x_n| \quad \text{and} \quad |y_1| \geq \dots \geq |y_n|. \tag{3.2}$$

Let $Q_0 = E_{n1} + \sum_{j=1}^{n-1} E_{j,j+1}$ be the basic circulant matrix. By (3.1) and (3.2), we have

$$\sum_{j=1}^{n} u_j^2 \geq \big(|x_1|^2, \ldots, |x_n|^2\big)\, Q_0^k \big(|y_1|^2, \ldots, |y_n|^2\big)^t, \quad k = 1, \ldots, n. \tag{3.3}$$

Now,

$$\begin{aligned} n \sum_{j=1}^{n} u_j^2 &= (ue^*, ue^*) \\ &= \big(\psi(ue^*), \psi(ue^*)\big) \\ &= (xy^*, xy^*) \\ &= \Big(\sum_{j=1}^{n} |x_j|^2\Big)\Big(\sum_{j=1}^{n} |y_j|^2\Big) \\ &= \sum_{j=1}^{n} \big(|x_1|^2, \ldots, |x_n|^2\big)\, Q_0^j \big(|y_1|^2, \ldots, |y_n|^2\big)^t. \end{aligned}$$

Hence, all the inequalities in (3.3) become equalities. It follows that $|x_1| = |x_n|$ or $|y_1| = |y_n|$. If $|y_1| = |y_n|$ then $xy^* = P^*ue^*P \in \mathcal{D}_1$. If $|x_1| = |x_n|$, by Lemma 2.7(f), we see that $ue^* = xy^* = ee^*/n$. Again we have $\psi(ue^*) \in \mathcal{D}_1$. Similarly, we can show that for any $R \in \mathrm{GP}(n)$, $\psi(Rue^*R^*) \in \mathcal{D}_1$. So, we have $\psi(\mathcal{D}_1) \subseteq \mathcal{D}_1$. Applying the same argument to ψ^{-1}, we conclude that $\psi(\mathcal{D}_1) = \mathcal{D}_1$ as asserted.

Next, we shall prove that ψ leaves invariant certain subsets of $\mathcal{D}_1$. These are the sets $\mathcal{D}_{1j}$, $(1 \leq j \leq n)$ defined as

$$\mathcal{D}_{1j} = \{DQ_0^j ue^t (Q_0^j)^t D^* : D \text{ is diagonal unitary}\}$$

where Q_0 is the basic circulant matrix $E_{n1} + \sum_{j=1}^{n-1} E_{j,j+1}$. In other words $\mathcal{D}_{1j}$ is the diagonal-unitary orbit of the matrix $(u_{j+1}, \ldots, u_n, u_1, \ldots, u_j)^t e^t$, while $\mathcal{D}_1$ is the unions of the diagonal-unitary orbits of all the matrices $(u_{\pi(1)}, \ldots, u_{\pi(n)})^t e^t$ for all permutations π of the set of indices $\{1, 2, \ldots, n\}$. It is clear that a matrix X belongs to $\mathcal{D}_{1j}$ if and only if $X \in \mathcal{D}_1$ and the diagonal entries of X are $u_{j+1}, \ldots, u_n, u_1, \ldots, u_j$. Now ψ maps $\mathcal{D}_1$ onto itself and fixes the diagonal of every matrix. Thus is $X \in \mathcal{D}_{1j}$, then $X \in \mathcal{D}_1$ and so $\psi(X) \in \mathcal{D}_1$. Also

$$\mathrm{diag}\big(\psi(X)\big) = \mathrm{diag}(X) = (u_{j+1}, \ldots, u_n, u_1, \ldots, u_j).$$

Therefore $\psi(X) \in \mathcal{D}_{1j}$. Applying the same reasoning to ψ^{-1}, we conclude that

$$\psi(\mathcal{D}_{1j}) = \mathcal{D}_{1j}, \quad j = 1, \ldots, n. \tag{3.4}$$

If u is a scalar multiple of e, then $ue^* = ee^*/n$, and $\mathcal{D}_{1j} = \mathcal{D}_1 = \mathcal{F}$ for all $j = 1, \ldots, n$, and the assertion of the lemma is already established.

We now assume that u is not a scalar multiple of e, *i.e.*, $u_1 > u_n$. In this case, the sets $\mathcal{D}_{1j}$ are all distinct. We claim that $A \in \mathcal{F}$ if and only if $(A, A) = n^2$ and

$$(3.5) \qquad A = A_1 + \cdots + A_n \text{ with } A_j \in \mathcal{D}_{1j} \quad \text{for } j = 1, \dots, n.$$

Suppose $A = Dee^*D^* \in \mathcal{F}$, where D is a diagonal unitary matrix. Then $(A, A) = n^2$ and (3.5) holds with $A_j = DQ_0^j ue^*(Q_0^j)^*D^* \in \mathcal{D}_{1j}$ for $j = 1, \dots, n$. Conversely, suppose $A = (a_{ij}) \in M_n$ is such that

$$A = \sum_{j=1}^{n} D_j Q_0^j ue^*(Q_0^j)^* D_j^*,$$

where D_j is a diagonal unitary matrix for $j = 1, \dots, n$. Let $d_{j1}, \dots, d_{jn}$ be the diagonal entries of D_j. Then $a_{pp} = \sum_{j=1}^{n} u_j = 1$ for all $j = 1, \dots, n$, while if $p \neq q$, then $a_{pq} = \sum_{j=1}^{n} d_{jp}\overline{d_{jq}}u_{p+j}$, with addition $p + j$ taken to be addition modulo n. Thus a_{pq} is a linear combination of $u_1, u_2, \dots, u_n$, with unimodular coefficients and hence $|a_{ij}| \leq 1$ and $(A, A) \leq n^2$. If we also assume that $(A, A) = n^2$, then we must have $|a_{pq}| = 1$, for every p, q. Since ν is not the ℓ_∞, we have $u_1 \geq u_2 > 0$ by Lemma 2.7(a). It follows that $|a_{pq}| = 1$, for every p, q if and only if all D_j are the same up to multiplication by a complex unit. Thus

$$A = D_1 \Big(\sum_{j=1}^{n} Q_0^j ue^*(Q_0^j)^* \Big) D_1^* = D_1 ee^t D_1^* \in \mathcal{F}.$$

Since ψ satisfies (3.4), preserves the inner product on M_n, and the set $\mathcal{D}_{1j}$ for each j, we see that ψ maps the set of matrices A satisfying (3.5) onto itself. So, we have $\psi(\mathcal{F}) = \mathcal{F}$ as asserted. ■

We are now ready to finish the proof of our theorem by establishing the following.

Lemma 3.4 *The implication* (e) $\Rightarrow$ (c) *in Theorem 3.1 holds.*

Proof By the result in Lemma 3.2, there exists a permutation matrix P_1 such that the map ψ_1 defined by $\psi_1(X) = P_1^*\psi(X)P_1$ fixes every diagonal matrix. By Lemma 3.3, $\psi_1(\mathcal{F}) = \mathcal{F}$. By the result in [14], there exists $Q \in \mathrm{GP}(n)$ such that

(i) $\psi_1(X) = QXQ^*$ for all X with zero diagonal, or
(ii) $\psi_1(X) = QX^tQ^*$ for all X with zero diagonal.

Define ψ_2 by $\psi_2(X) = Q^*\psi_1(X)Q$, then

(iii) $\psi_2(X) = X$ for all X with zero diagonal, or
(iv) $\psi_2(X) = X^t$ for all X with zero diagonal.

Since $\psi_1(D) = D$ for every diagonal matrix D, there exists a permutation matrix P such that

$$\psi_2(D) = PDP^t \quad \text{for every diagonal matrix } D. \tag{3.6}$$

We will show that ψ_2 is the identity or (in the case $n = 2$ only) a mapping of the form of Theorem 3.1(c). This would then imply that ψ itself is of the form of Theorem 3.1(c).

For $n = 2$, the forms (i) and (ii) coincide since if $P_0 = E_{12} + E_{21}$, then the off-diagonal entries of X^t are the same as the off-diagonal entries of $P_0XP_0^*$. So we may assume that ψ_2 satisfies (iii). If $P \neq I$ then $P = E_{12} + E_{21}$ and $\psi_2(X) = PX^tP$ for all $X \in M_2$. It follows that $\mathcal{S}^t = \mathcal{S}$, which contradicts Lemma 2.8. Therefore $P = I$ and ψ_2 is the identity mapping.

Suppose $n \geq 3$. Then every rank one matrix is completely determined by its off-diagonal entries with the exception of those matrices that are "essentially 2×2", *i.e.*, matrices of the form uv^t where $u, v \in \operatorname{span}\{e_i, e_j\}$ for some indices i and j. In particular a vector state S is completely determined by its off-diagonal entries except when S takes the following form:

$$S = tE_{ii} + s_{ij}E_{ij} + s_{ji}E_{ji} + (1-t)E_{jj}. \tag{3.7}$$

where $i \neq j$, $0 \leq t \leq 1$ and $s_{ij}s_{ji} = t(1-t)$. Furthermore in this exceptional case the only other vector state that has the same off-diagonal entries as S is

$$T = (1-t)E_{ii} + s_{ij}E_{ij} + s_{ji}E_{ji} + tE_{jj}, \tag{3.8}$$

unless $S = E_{ii}$, in which case, T can be any of the matrices E_{jj}. (The case $S = E_{ii}$ will be excluded presently, as we consider only nonsymmetric states.) We use the above to prove that (iv) is not possible and that $P = I$.

Since $\psi_2(\mathcal{S}) = \mathcal{S}$, we have $\psi_2(\operatorname{Ext}\mathcal{S}) = \operatorname{Ext}\mathcal{S}$, which is a set of rank one matrices by Corollary 2.2. By Lemma 2.8(c), there exist nonnegative vector $x, y \in \mathbf{C}^n$ such that $Z = xy^t \in \operatorname{Ext}(\mathcal{S})$ but $Z^t = yx^t \notin \operatorname{Ext}(\mathcal{S})$. If (iv) holds and if xy^t is not one of the exceptional states of (3.7) then the off-diagonal entries of $\psi_2(xy^t) \in \operatorname{Ext}(\mathcal{S})$ are the same as those of yx^t, and thus $\psi(xy^t) = yx^t \in \operatorname{Ext}(\mathcal{S})$, which is a contradiction. If xy^t is one of the exceptional states of (3.7), then either we reach a contradiction as before or the permutation P in (3.6) satisfies $Pe_i = e_j$ or $Pe_j = e_i$ or both according as $t = 1$ or $t = 0$ or neither. Since this may be applied to Uxy^tU^t for any permutation U, we get that $Pe_1 = e_2 = e_3$ which is absurd. This proves that (iii) is satisfied.

The diagonal of the state $Z = xy^t$ considered in the previous paragraph is not constant. Indeed it follows from Lemma 2.8 that the only possible nonnegative vector state with constant diagonal is ee^t/n, which is symmetric. We now consider the image $\psi_2(UZU^t)$ for every permutation matrix U. If Z is not one of the exceptional states of (3.7), then the permutation P in (3.6) leaves $\operatorname{diag}(Z)$ and every permutation of it invariant. This implies that $P = I$. If Z is one of the exceptional states in (3.7) and if $0 < t < 1$, then by (3.7) and (3.8), we see that P leaves the span of $\{e_i, e_j\}$ invariant. This is true for every pair of distinct indices and again we get $P = I$. The only remaining cases are when (after possible relabelling of indices) $Z = E_{ii} + sE_{ij}$ or $Z = E_{ii} + sE_{ij}$; $s \neq 0$. In the former case, we get that if $Pe_i \neq e_i$, then $Pe_i = e_j$ and that would be true for every $j \neq i$, which is absurd. Again we get $P = I$. The remaining case follow similarly. Thus ψ_2 is the identity and hence $\psi(A) = QAQ^*$ for a generalized permutation Q. ■

The extremal cases where ν is the ℓ_1 or the ℓ_∞ norm will be considered in Section 5.

4 Numerical Radius Isometries

The main result of this section is the following theorem. The result is known when ν is a multiple of the ℓ_2 norm [12] (see also [4, 5])

Theorem 4.1 *Suppose $r = r_\nu$ is the norm numerical radius associated with a symmetric norm ν on $\mathbf{C}^n$, where ν is not a multiple of the ℓ_1 norm or the ℓ_∞ norm. Then a linear operator ϕ on M_n is a ν-numerical radius isometry, i.e., satisfies*

$$r\big(\phi(A)\big) = r(A) \quad \text{for all } A \in M_n$$

if and only if there is a complex number μ of modulus 1 such that

$$W\big(\mu\phi(A)\big) = W(A) \quad \text{for all } A \in M_n.$$

In other words, ϕ preserves the numerical radius if and only if it is a unit multiple of a numerical range preserver.

If we exclude the ℓ_2 norm and use Theorem 3.1, we obtain the following.

Corollary 4.2 *Suppose ν is a symmetric norm on $\mathbf{C}^n$ and that ν is not a multiple of the ℓ_1, ℓ_2 or the ℓ_∞ norm. Then a linear operator ϕ on M_n is a ν-numerical radius isometry if and only if there exists a complex number λ of modulus 1 and a generalized permutation $Q \in \mathrm{GP}(n)$ such that*

$$\phi(A) = \lambda Q^* AQ \quad \text{for all } A \in M_n,$$

To prove Theorem 4.1, we need the following characterization of scalar matrices in terms of the numerical radius, which may be of independent interest.

Proposition 4.3 *Suppose ν is a symmetric norm on $\mathbf{C}^n$ not equal to a multiple of ℓ_q norm with $q \in \{1, \infty\}$, and $r = r_\nu$ is the corresponding norm numerical radius. Let $\mathcal{L}$ be the set of matrices $A \in M_n$ such that for every $Y \in M_n$ there exists a complex unit η such that*

$$r(\eta A + Y) = 1 + r(Y). \tag{4.1}$$

Then $A \in \mathcal{L}$ if and only if $A = \mu I$ for some $\mu \in \mathbf{C}$ with $|\mu| = 1$.

The "if" part is clear. The converse will be proved by establishing a sequence of lemmas. In all that follows, the set $\mathcal{L}$ will be the set of matrices defined by (4.1) for initially an arbitrary symmetric norm. Restrictions on the norm (to exclude the extremal norms) will be required when first needed.

Lemma 4.4 *If $A \in \mathcal{L}$, then $r(A) = 1$.*

Proof Take $Y = 0$ in (4.1).

Lemma 4.5 *If $A \in \mathcal{L}$, then so does $\delta Q^* AQ$ for every generalized permutation Q and every complex unit δ.*

Proof Let $Y \in M_n$. Then there exists a complex unit η such that

$$1 + r(Y) = 1 + r(PYP^*) = r(\eta A + PYP^*) = r\big((\eta\bar{\delta})\delta P^* AP + Y\big).$$

This proves the lemma. ■

Lemma 4.6 *Let $B \in M_n$, let m be a positive integer and $A_1, A_2, \dots, A_m \in \mathcal{L}$. Then there exists a state $S \in \operatorname{Ext} \mathcal{S}$ such that*

(4.2) $$|(B, S)| = r(B) \text{ and } |(A_j, S)| = 1 \quad \text{for } j = 1, 2, \dots, n.$$

Proof By applying equation (4.1) repeatedly, or by induction, we see that there exist complex numbers η_j; $(1 \le j \le n)$ of modulus one such that

$$r(\eta_1 A_1 + \cdots + \eta_m A_m + B) = m + r(b).$$

Therefore there exists a state $S \in \mathcal{S}$, such that $|(\eta_1 A_1 + \cdots + \eta_m A_m + B, S)| = m + r(B)$. But $|(A_j, S)| \le 1$ and $|(B, S)| \le r(B)$. This implies that S satisfies (4.2). If S is not an extreme point of $\mathcal{S}$, then S is a convex combination of states in $\operatorname{Ext} \mathcal{S}$, each of which must also satisfy (4.2). ■

We next prove a generalization of the above lemma.

Lemma 4.7 *Let $B \in M_n$. Then there exists a state $S \in \operatorname{Ext} \mathcal{S}$ such that*

(4.3) $$|(B, S)| = r(B) \text{ and } |(A, S)| = 1 \quad \text{for every } A \in \mathcal{L}.$$

Proof For every $A \in \mathcal{L}$, let

$$\mathcal{S}_A = \{X \in \mathcal{S} : |(A, X)| = 1 \text{ and } |(B, X)| = r(B)\}.$$

Each $\mathcal{S}_A$ is evidently a closed subset of $\mathcal{S}$. Furthermore, by Lemma 4.6, the intersection of any finite number of the sets $\mathcal{S}_A$ is nonempty. Since $\mathcal{S}$ is compact, it follows that the intersection

$$\bigcap_{A \in \mathcal{L}} \mathcal{S}_A$$

is nonempty, *i.e.*, there exists a state X satisfying (4.3). As in the proof of Lemma 4.6, if X is not in Ext $\mathcal{S}$, then there exists an $S \in$ Ext $\mathcal{S}$ that satisfy the same equations (4.3). ■

To effectively exploit the above lemma, we shall choose one particular matrix B for which the corresponding set of *numerical radius norming states*, *i.e.*, the states S that satisfy $|(B, S)| = r(B)$, can be determined.

Recall that by Lemma 2.7, there exists a nonnegative vector u such that ue^t or eu^t belong to Ext $\mathcal{S}$. We shall henceforth assume the latter case as the former case would then follow by duality. In other words, we are assuming that $e/\nu^D(e)$ belongs to $\mathcal{E}^D$, the set of extreme points of the unit ball of ν^D. We recall also that we must have that $u_1 + \cdots + u_n = 1$.

Let k be the smallest positive integer such that the set

$$\mathbf{U} = \{e(u_1, \ldots, u_k, 0, \ldots, 0) \in \mathcal{S} : u_1, \ldots, u_k > 0, u_1 + \cdots + u_k = 1\} \tag{4.4}$$

is non-empty. Now, we set

$$B = e(e_1 + \cdots + e_k)^t. \tag{4.5}$$

The set $\mathbf{U}$ is convex. Indeed if S is a convex combination of elements of $\mathbf{U}$, then S is evidently a state of the form $e(v_1, \ldots, v_k, 0, \ldots, 0)$. Furthermore, by the minimality of k, we must have $v_1, \ldots, v_k > 0$, that is, S is a member of $\mathbf{U}$. We also note that by Lemma 2.7, $k = 1$ if and only if ν is the ℓ_1 norm.

Lemma 4.8 *Let B be the matrix defined by equation* (4.5) *and let* $\mathbf{U}$ *be the subset of* $\mathcal{S}$ *described in* (4.4). *A state $S \in \mathcal{S}$ is numerical radius norming for the matrix B, i.e.,* $|(B, S)| = r(B)$, *if and only if* $S \in \mathbf{U}$.

Proof Let $\gamma = \nu^D(e)$. First we consider extremal states $vw^* \in$ Ext $\mathcal{S}$. By Lemma 2.7(e), for every $vw^* \in$ Ext$(\mathcal{S})$ such that $v = (v_1, \ldots, v_n)^t \in \mathcal{E}^D$ and $w = (w_1, \ldots, w_n)^t \in \mathcal{E}$, we have

$$\sum_{j=1}^{n} |v_j| \le n/\gamma \quad \text{and} \quad \sum_{j=1}^{n} |w_k| \le \gamma,$$

and the first equality holds if and only if $(|v_1|, \ldots, |v_n|)^t = e/\gamma$. Therefore, for every $vw^* \in$ Ext $\mathcal{S}$, we have

$$|(B, vw^*)| = |(e, v)(e_1 + \cdots + e_k, w)| \le (n/\gamma)\gamma = n.$$

Thus $r(B) = n$, and if $|(B, uv^*)| = n$, then all the inequalities above become equalities and this occurs only if $v_1 = v_2 = \cdots = v_n$ and $w_j = 0$ for $j > k$. This means that $vw^* = eu^*$ for some $u \in \mathbf{C}^n$. Since eu^* is a state, the vector u must be nonnegative and since $u = (u_1, \ldots, u_k, 0, \ldots, 0)^t$ then by the minimality of k in (4.4), we must have that $u_j > 0$ for $j = 1, \ldots, k$. Thus $xw^* = eu^* \in \mathbf{U}$.

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

Now let S be any state that satisfies $|(B,S)| = r(B) = n$. The state S is a convex combination of extremal states $X_1, X_2, \dots, X_m$. Each of these extremal states X_j must then satisfy $|(B, X_j)| = n$. Therefore every $X_j \in \mathbf{U}$ and it then follows that $S \in \mathbf{U}$ since $\mathbf{U}$ is convex.

The converse is easily verified as direct calculation shows that $(B,U) = n = r(B)$ for every $U \in \mathbf{U}$. ■

Corollary 4.9 *There exists a state $U \in \mathbf{U}$ such that $|(A,U)| = 1$ for every $A \in \mathcal{L}$.*

Proof This follows immediately from Lemma 4.7 and Lemma 4.8. ■

In the following, we need to partition matrices. If $A = (a_{ij})$, we also write

$$A = \begin{pmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

So A_{12} is a $1 \times (n-1)$ matrix, A_{21} is an $(n-1) \times 1$ matrix and A_{22} is an $(n-1) \times (n-1)$ matrix.

Lemma 4.10 *Let*

$$U = \begin{pmatrix} u_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \in \mathcal{S} \quad \text{and} \quad A = \begin{pmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

and assume that $|(DAD^*, U)| = 1$ for every diagonal unitary matrix D. Assume further that $(A_{21}, U_{21}) \neq 0$. Then

$$1 = |(A_{21}, U_{21})| \quad \text{and} \quad 0 = (A_{12}, U_{12}) = a_{11}\bar{u}_{11} + (A_{22}, U_{22}).$$

Proof Let $D_\theta = e^{i\theta} \oplus I_{n-1}$, and let $A_\theta = D_\theta A D_\theta^*$. By assumption, we have $|(A_\theta, U)| = 1$, *i.e.*,

$$1 = |a_{11}\bar{u}_{11} + (A_{22}, U_{22}) + e^{i\theta}(A_{12}, U_{12}) + e^{-i\theta}(A_{21}, U_{21})| \tag{4.6}$$

for every θ. It is not hard to see that if a, b, and c are complex numbers, such that $|a + be^{i\theta} + ce^{-i\theta}| = 1$ for every θ, then two of a, b, or c must be 0. Indeed if $f(\theta) = a + be^{i\theta} + ce^{-i\theta}$, then we have $f(\theta)\overline{f(\theta)} - 1 = 0$. Upon equating the coefficients of every power of $e^{i\theta}$ to zero, we reach the above conclusion.

Since $(A_{21}, U_{21}) \neq 0$, then the other coefficients in equation (4.6) must be zero, *i.e.*, $a_{11}\bar{u}_{11} + (A_{22}, U_{22}) = 0$ and $(A_{12}, U_{12}) = 0$. It then follows that $|(A_{21}, U_{21})| = 1$. ■

Lemma 4.11 *Let $A \in \mathcal{L}$. Then each column of A has at most one nonzero off-diagonal entry.*

Proof If the assertion is not true, then there exists $Q \in \mathrm{GP}(n)$ and a complex unit δ such that δQAQ^* has nonnegative first column and such that the $(2,1)$ and $(3,1)$ entries are positive. We may now replace A by δQAQ^*, *i.e.*, we assume that A itself has a nonnegative first column and that $a_{21} > 0$ and $a_{31} > 0$. By Corollary 4.9, there exists a state $U \in \mathbf{U}$ such that $|(C, U)| = 1$ for every $C \in \mathcal{L}$. We conclude that the results of Lemma 4.10 hold for the matrix A as well as a certain perturbation of A introduced below.

First applying Lemma 4.10 to A, we get that $|(A_{21}, U_{21})| = 1$, *i.e.*,

$$u_1 \sum_{j=2}^{n} a_{j1} = 1 \tag{4.7}$$

Next we describe the perturbation $\tilde{A}$ of A alluded to above. Towards this, let

$$R = \mathrm{diag}(1, e^{i\varepsilon}, e^{i2\varepsilon}, \dots, e^{i(n-1)\varepsilon}), \tag{4.8}$$

and $\tilde{A} = RAR^*$. Then

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$$

where $\tilde{A}_{21} = (a_{21}e^{i\varepsilon}, a_{31}e^{2i\varepsilon}, \dots, a_{n1}e^{i(n-1)\varepsilon})^t$. This matrix $\tilde{A}$ is in the diagonal-unitary orbit of A and for sufficiently small $\varepsilon > 0$, the inner product $(\tilde{A}_{21}, U)$ is close enough to (A_{21}, U) so that it is not zero. But by assumption $|(\tilde{A}, U)| = 1$, so, by Lemma 4.10, $|(\tilde{A}_{21}, U)| = 1$, *i.e.*,

$$u_1 \Big| \sum_{j=2}^{n} e^{(j-1)i\varepsilon} a_{j1} \Big| = 1. \tag{4.9}$$

However, equations (4.7) and (4.9) are not simultaneously possible as $e^{i\varepsilon}a_{21}u_1$ and $e^{2i\varepsilon}a_{31}u_1$ are nonzero and have different arguments. ■

We note that all of the preceding lemmas in Section 4 are still true when ν is the ℓ_1 norm. (Recall that the ℓ_∞ norm is excluded by our assumption that eu^t is a state in Ext $\mathcal{S}$.) For most of the remainder of this Section, we must exclude the ℓ_1 norm as well. But we pause momentarily to state the characterization of the set $\mathcal{L}$ for these extremal norms.

Lemma 4.12 *When ν is the ℓ_1 norm (respectively, ℓ_∞ norm), $A \in \mathcal{L}$ if and only if every column (respectively, row) of A has exactly one nonzero entry, and every such entry has modulus* 1.

Proof For the ℓ_1 norm, we have already seen that if $A \in \mathcal{L}$ then every column of A has at most one off-diagonal nonzero entry. If $a_{21} \neq 0$, then using Lemma 4.10 and noting that $k = 1$, we get that $a_{11} = 0$ and $|a_{21}| = 1$. If there are no off-diagonal nonzero entries in the first column, then by Corollary 4.9 and the fact that $\mathbf{U} = \{ee_1^t\}$, we get that $|a_{11}| = 1$. The same argument applies to any column, thus very column of A has exactly one nonzero entry, and every such entry is of modulus 1. For the converse assume every column of A has exactly one nonzero entry and every such entry is of modulus 1. Let $Y = (y_{ij}) \in M_n$. The numerical radius of Y is assumed at an extremal state $S = Qee_1^t Q^*$ for some $Q \in \mathrm{GP}(N)$, *i.e.*, $r(Y) = |(Y, S)|$. But from the structure of A we easily see that $|(A, S)| = 1 = r(A)$ for every such S. If η is a complex unit chosen so that the complex numbers (Y, S) and $\eta(A, S)$ have the same argument then $|(\eta A + Y, S)| = |(A, S)| + |(Y, S)| = 1 + r(Y)$, proving that $A \in \mathcal{L}$.

The ℓ_∞ result follows by duality. ∎

For the remainder of this Section we shall assume that ν is not the ℓ_1 or the ℓ_∞ norm.

Lemma 4.13 *If $A \in \mathcal{L}$, then each row of A has at most one nonzero off-diagonal entry.*

Proof The proof is similar to the proof of Lemma 4.11. We must however make sure that U_{12} has at least two nonzero entries. Recall that $U = eu^t$. Since ν is not the ℓ_1 norm, by Lemma 2.7(e)(iii), the vector u has at least two nonzero entries. Also since the set $\mathcal{L}$ is invariant under permutations, *i.e.*, $P\mathcal{L}P^t = \mathcal{L}$, for every permutation P, then the same holds for the states that are simultaneously numerical radius norming for $\mathcal{L}$. We choose a permutation P so that if $v = Pu$, then v_2 and v_3 are both nonzero. By Corollary 4.9 and the observations above, the state $V = ev^t = Peu^tP^t$ satisfies $|(C, V)| = 1$ for ever $C \in \mathcal{L}$. We now follow the same argument as in the proof of Lemma 4.11, first reducing to the case A having nonnegative first row and finally reaching

$$\sum_{j=2}^{n} a_{1j} v_j = 1 \tag{4.10}$$

and

$$\Bigl|\sum_{j=2}^{n} e^{-(j-1)i\varepsilon} a_{1j} v_j\Bigr| = 1. \tag{4.11}$$

and the result follows as before. ∎

Lemma 4.14 *If $A \in \mathcal{L}$ has a nonzero off-diagonal entry in the j-th column, then all the off-diagonal entries in the j-th row are all zero.*

Proof We may replace A by a matrix in its GP-orbit, and so we may assume that $j = 1$, *i.e.*, $A_{21} \neq 0$ and we may also assume by Lemma 4.13 that $a_{1j} = 0$ for $j \geq 3$. We must then show that $a_{12} = 0$ and hence $A_{12} = 0$. By Corollary 4.9, there exists a state $U \in \mathbf{U}$ such that $|(C, U)| = 1$ for every C in the GP-orbit of A. By Lemma 4.10, we get that $a_{12}u_2 = (A_{12}, U_{12}) = 0$. Since ν is not the ℓ_1 norm, $u_2 \neq 0$ and so $a_{12} = 0$. ∎

Lemma 4.15 *If $A \in \mathcal{L}$, then A has at most one nonzero off-diagonal entry.*

Proof Assume to the contrary that A has two nonzero off-diagonal entries a_{ij} and a_{pq}. By the above assertion, the indices i, j, p, q must all be different from each other. We may replace A by a matrix in its permutation orbit and so we may assume that a_{31} and a_{42} are nonzero. By Lemma 4.14, the matrix A must now be of the form

$$\begin{pmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ * & 0 & * & 0 \\ 0 & * & 0 & * \end{pmatrix} \oplus C.$$

As before we get a state $U \in \mathbf{U}$ that is numerical radius norming for the GP-orbit of A, and we conclude by Lemma 4.10 that

$$a_{11}u_1 + (A_{22}, U_{22}) = 0. \tag{4.12}$$

This equation is also satisfied when A is replaced by any matrix $\tilde{A}$ in its diagonal-unitary orbit, *i.e.*,

$$a_{11}u_{11} + (\tilde{A}_{22}, U_{22}) = 0. \tag{4.13}$$

In particular, if we take $D = \operatorname{diag}(1, -1, 1, \dots, 1)$ and $\tilde{A} = DAD^*$, then the entries $(\tilde{a}_{ij})$ of $\tilde{A}$ are the same as a_{ij} except for the $(4, 2)$ entry, where $\tilde{a}_{4,2} = -a_{4,2}$. Now subtracting equation (4.13) form equation (4.12), we get $2a_{42}u_2 = 0$. But $a_{42} \neq 0$ by assumption and $u_2 \neq 0$ since ν is not the ℓ_1 norm. The contradiction establishes the lemma. ■

Lemma 4.16 *If $A = (a_{ij}) \in \mathcal{L}$ and $a_{21} \neq 0$, then*

(a) *The integer k in* (4.4) *equals* 2 *and the set* $\mathbf{U} = \{e(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0)\}$.
(b) $|a_{21}| = 2$ *and* $|a_{jj}| = 1$ *for every j.*
(c) $a_{11} = -a_{jj}$ *for every $j \geq 2$.*

Proof Let $U = eu^t \in \mathbf{U}$ be a numerical radius norming state for all matrices in $\mathcal{L}$. Using $|(A, U)| = 1$ and Lemma 4.10, we have $u_1|a_{21}| = 1$. Using other matrices in the permutation orbit of A, we get that $u_j|a_{21}| = 1$ for $j = 1, 2, \dots, k$. Therefore $u_1 = u_2 = \cdots = u_k = \frac{1}{k}$ and $|a_{21}| = k$. We recall that $\|A\| \leq \mathbf{e}r(A)$ where $\mathbf{e}$ is the Euler constant. Thus

$$k = |a_{21}| \leq \|A\| \leq \mathbf{e}r(A) = \mathbf{e} < 3.$$

Also $k \geq 2$ as ν is not the ℓ_1 norm. Thus $k = 2$, $u_1 = u_2 = \frac{1}{2}$ and $|a_{21}| = 2$. This proves (a) and the first assertion of (b). Furthermore, we get from Lemma 4.10 that $a_{11}u_1 + a_{22}u_2 = 0$, which implies that $a_{11} + a_{22} = 0$. Since we may also replace A

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

by P^tAP for any permutation P that fixes e_1, we also conclude that $a_{11} + a_{jj} = 0$ for $2 \leq j \leq n$. This proves (c). In particular, we have

$$a_{22} = a_{33} = \cdots = a_{nn}.$$

It remains only to show that every diagonal entry of A has modulus 1.

If $n \geq 3$, we use a permutation P so that the first column of PUP^t is zero and the second and third columns are not zero. Each of these columns must then equal $e/2$ and we get

$$1 = |(P^tAP, U)| = |(A, PAP^t)| = |(a_{22} + a_{33})|/2 = |a_{22}|.$$

We then have $|a_{11}| = 1$ since $a_{11} + a_{22} = 0$, and consequently $|a_{jj}| = 1$ for other indices j since $a_{jj} = -a_{11}$.

If $n = 2$, we apply Lemma 4.7 with $B = E_{11}$. Any extremal state X that is numerical radius norming for E_{11} is either of the form $X = e_1x^t$ with $x = (1, \xi)^t$ or of the form $X = ye_1^t$ with $y = (1, \eta)^t$. In the former case, we get $|a_{11}| = |(A, X)| = 1$. In the latter case, we take P to be the permutation $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, then $|a_{22}| = |(PAP^t, X)| = 1$. Since $a_{11} = -a_{22}$, then in each case we get $|a_{11}| = |a_{22}| = 1$. ■

Lemma 4.17 *Every matrix in $\mathcal{L}$ is a diagonal unitary matrix.*

Proof If $A \in \mathcal{L}$ is not diagonal, then by Lemma 4.15, it has only one off-diagonal entry, which we may assume, without loss of generality, to be a_{21}. By Lemma 4.16 and Lemma 4.5, we must also have the following matrix A_1 in $\mathcal{L}$.

$$A_1 = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} \oplus I_{n-2}.$$

It follows also that the matrix

$$A_2 = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \oplus I_{n-2}$$

also belongs to $\mathcal{L}$ since $A_2 = PDA_1D^*P^*$ for $D = (-1) \oplus I_{n-1}$ and $P = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \oplus I_{n-2}$. Therefore

$$r(A_1 - A_2) \leq r(A_1) + r(A_2) = 2.$$

On the other hand,

$$A_1 - A_2 = \begin{pmatrix} -2 & 2 \\ 2 & 2 \end{pmatrix} \oplus 0$$

has eigenvalues $\pm 2\sqrt{2}$ and $n - 2$ zeros. Therefore its spectral radius is larger than its numerical radius, which is impossible. This proves that A is diagonal.

Next, we use Lemma 4.7 with $B = E_{mm}$. Any state $S = (s_{ij})$ that satisfies $|(E_{mm}, S)| = 1 = r(E_{mm})$ must have $s_{mm} = 1$ and all other diagonal entries zero. Therefore $|d_m| = |(A, S)| = 1$. This proves that A is a diagonal unitary. ■

Finally, we are ready to prove Proposition 4.3.

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

Lemma 4.18 *Every matrix in $\mathcal{L}$ is a unimodular scalar multiple of the identity.*

Proof Let $A \in \mathcal{L}$. We have now shown that $A = \operatorname{diag}(d_1, d_2, \dots, d_n)$ with $|d_j| = 1$ for every j. By Corollary 4.9, there exits $U \in \mathbf{U}$ such that $|(PAP^t, U)| = 1$ for every permutation P. Thus $1 = |(d_1, \dots, d_n)P(u_1, \dots, u_k, 0, \dots, 0)^t|$ for every permutation matrix P. Since ν is not the ℓ_1 norm, $k \geq 2$, and we conclude that $d_1 = \cdots = d_n$. ■

We remark that our proof of Proposition 4.3 is computational and quite long. It would be of interest to have a short conceptual proof. Since the result is not valid for r_ν if ν is a multiple of the ℓ_1 or ℓ_∞ norm, any proof must use the fact that these two norms behave differently from other symmetric norms. In our proof, the set $\mathbf{U}$ defined in (4.4) allows us to make the distinction.

We establish another general result which is useful in proving Theorem 4.1.

Proposition 4.19 *Let $\mathcal{K}(\mathbf{C})$ be the set of all compact convex subsets of $\mathbf{C}$ and let $\mathcal{M}$ be a linear subspace of M_n such that $I \in \mathcal{M}$. Suppose $F \colon \mathcal{M} \to \mathcal{K}(\mathbf{C})$ is a function that satisfies*

$$F(A + \beta I) = F(A) + \beta, \quad \text{for every } \beta \in \mathbf{C},$$

and define $f \colon \mathcal{M} \to \mathbf{R}$ by $f(A) = \max\{|z| : z \in F(A)\}$. If ϕ is a linear operator from $\mathcal{M}$ into $\mathcal{M}$ satisfying $\phi(I) = I$ and $f\big(\phi(A)\big) = f(A)$ for all $A \in \mathcal{M}$, then $F\big(\phi(A)\big) = F(A)$ for all $A \in \mathcal{M}$.

Proof Assume that $F\big(\phi(A)\big) \neq F(A)$ for some $A \in \mathcal{M}$. If there is $\mu \in F\big(\phi(A)\big) \setminus F(A)$, then by a standard separation theorem for convex sets, there exists $\eta \in \mathbf{C}$ such that

$$f\big(\phi(A - \eta I)\big) = f\big(\phi(A) - \eta I\big) \geq |\mu - \eta| > \max_{z \in F(A)} |z - \eta| = f(A - \eta I),$$

which is a contradiction. Similarly, if there is $\mu \in F(A) \setminus F\big(\phi(A)\big)$, then there exists $\eta \in \mathbf{C}$ such that

$$f\big(\phi(A - \eta I)\big) = f\big(\phi(A) - \eta I\big) = \max_{z \in F(\phi(A))} |z - \eta| < |\mu - \eta| \leq f(A - \eta I),$$

which is a contradiction. ■

Now, we are ready to complete the proof of the main theorem in this section.

Proof of Theorem 4.1 First, we show that a linear numerical radius isometry ϕ satisfies $\phi(I) = \mu I$ for some complex unit μ. Suppose $\phi(I) = C$. Since the numerical radius is a norm, ϕ is invertible. Then for every $Y \in M_n$ there exists $X \in M_n$ such that $\phi(X) = Y$. By Proposition 4.3 there exists a complex unit η such that

$$1 + r(Y) = 1 + r\big(\phi(X)\big) = 1 + r(X) = r(\eta I + X) = r\big(\eta\phi(I) + \phi(X)\big) = r(\eta C + Y).$$

Since this is true for every $Y \in M_n$, by Proposition 4.3 we conclude that $C = \mu I$ for some complex unit μ.

We may now replace ϕ by the mapping $A \mapsto \bar{\mu}\phi(A)$ and assume that $\phi(I) = I$. Applying Proposition 4.19 with $\mathcal{M} = M_n$ and $F(A) = V(A)$, we see that $V\big(\phi(A)\big) = V(A)$ for all $A \in M_n$, and the conclusion follows. ■

5 The Extremal ℓ_1 and ℓ_∞ Norms

In this section, we assume that the norm ν on $\mathbf{C}^n$ is the ℓ_1 or the ℓ_∞ norm and we characterize the linear operators on M_n that preserve the corresponding states or vector states or any of the two norm numerical ranges or the induced numerical radius. For the most part, we shall state and prove the results only for the ℓ_1 norm as the ℓ_∞ norm may then be treated using duality. First, the preservers of the spatial numerical range are the same as those of the other (non Hilbert space) symmetric norms. The labels of the parts of the following theorems conform with the labels in Theorem 3.1.

Theorem 5.1 *Let ν be the ℓ_1 or the ℓ_∞ norm on $\mathbf{C}^n$, and let $W(A)$ denote the induced spatial numerical range on M_n. Let ϕ be a linear operator on M_n. Then the following conditions are equivalent.*

(a) *ϕ preserves the spatial numerical range, i.e., $W\big(\phi(A)\big) = W(A)$ for all $A \in M_n$.*
(c) *There exists a generalized permutation $Q \in \mathrm{GP}(n)$ such that*

$$\phi(A) = Q^*AQ \quad \forall A \in M_n,$$

or equivalently,

$$\phi^*(A) = QAQ^* \quad \forall A \in M_n.$$

(d) $\phi^*(\mathcal{R}) = \mathcal{R}$.

The next result shows that the group of algebra numerical range preservers is much larger. When we write $A = [A_1|\cdots|A_n]$, we mean that the columns of A are the column vectors $A_1, \ldots, A_n$.

Theorem 5.2 *Let ν be the ℓ_1 norm on $\mathbf{C}^n$, and let $V(A)$ denote the induced algebra numerical range on M_n. Let ϕ be a linear operator on M_n. Then the following conditions are equivalent.*

(b) *ϕ preserves the algebra numerical range, i.e., $V\big(\phi(A)\big) = V(A)$ for all $A \in M_n$.*
(c)† *There exist a permutation matrix P and generalized permutations $Q_1, \ldots, Q_n \in \mathrm{GP}(n)$ with $Q_je_j = e_j$ for $j \in \{1, \ldots, n\}$ such that*

$$\phi([A_1|\cdots|A_n]) = P^*[Q_1A_1|\cdots|Q_nA_n]P \quad \text{for all } A = [A_1|\cdots|A_n] \in M_n,$$

or equivalently

$$\phi^*([A_1|\cdots|A_n]) = P[\tilde{Q}_1A_1|\cdots|\tilde{Q}_nA_n]P^* \quad \text{for all } A = [A_1|\cdots|A_n] \in M_n,$$

with $\tilde{Q}_j = P^*Q_{\pi(j)}P$ where π is the permutation determined by $Pe_j = e_{\pi(j)}$

(e) $\phi^*(\mathcal{S}) = \mathcal{S}$.

The analogous result for ℓ_∞ is similar but with columns replaced by rows.

Remarks 1. We note that P in Theorem 5.2 is taken to be a permutation rather than a generalized permutation since the action of any diagonal unitary may be absorbed in the action of the generalized permutations Q_j.

2. The group of operators appearing in Theorems 3.1 and 5.1 is ubiquitous in the theory of linear preservers. This is the group $\mathrm{GP}(n)/Z\big(\mathrm{GP}(n)\big)$, where $Z\big(\mathrm{GP}(n)\big)$ is the centre of $\mathrm{GP}(n)$, *i. e.*, the group of scalar matrices. On the other hand, the group appearing in Theorem 5.2 does not seem to appear anywhere else. This is the semidirect product of S_n and $\big(\mathrm{GP}(n-1)\big)^n$.

The next result shows that the group of numerical radius isometries are much larger than the unit multiples of the numerical range preservers. This is a deviation from all the known results on linear preservers of generalized numerical ranges and radii; see [13].

Theorem 5.3 *Suppose $r = r_1$ is the norm numerical radius associated with the ℓ_1 norm on $\mathbf{C}^n$. Then a linear operator ϕ on M_n is an ℓ_1-numerical radius isometry, i.e., satisfies*

$$r\big(\phi(A)\big) = r(A) \quad \text{for all } A \in M_n$$

if and only if there exists a permutation P and generalized permutations $Q_1, \dots, Q_n \in$ $\mathrm{GP}(n)$ *such that ϕ has the form*

$$[A_1|\cdots|A_n] \mapsto [Q_1A_1|\cdots|Q_nA_n]P;$$

Again, a similar result holds for the ℓ_∞ norm with columns replaced by rows.

Remarks 1. We note that the action of the permutation P is a (one-sided) right multiplication and not the usual "conjugation" $A \mapsto P^tAP$. This is due to the fact that a left multiplication by a permutation may be absorbed in the action of the generalized permutations Q_j.

2. The group of operators in Theorem 5.3, *i.e.*, the semidirect product of S_n and $\mathrm{GP}(n)^n$, is also a rare group among the groups of linear preservers or the groups of isometries of a normed space.

Before proving our theorems, we present some general observations, some of which are well known, about norms and numerical ranges induced by these extremal norms.

Proposition 5.4 *Let $A = (a_{ij}) \in M_n$. The operator norm $\|A\|$, its dual $\|A\|^D$ and the algebra numerical range $V(A)$ induced by the ℓ_1 norm satisfy the following.*

(i) $\|A\| = \max_j(\sum_{i=1}^n |a_{ij}|)$, *i.e., the maximum of the ℓ_1 norms of the columns of A.*

(ii) $\|A\|^D = \sum_{j=1}^n(\max_i |a_{ij}|)$, *i.e., the sum of the ℓ_∞ norms of the columns of A.*

(iii) *A matrix* $S = (s_{ij})$ *is a state if and only if* $s_{jj} \geq 0$ *for every* j, $\operatorname{tr}(S) = 1$ *and*

$$s_{jj} = \max_{1 \leq i \leq n} |s_{ij}| \quad \text{for every } j$$

(iv) $V(A) = \operatorname{conv}(\bigcup_{j=1}^{n} \Omega_j)$, *where* Ω_j *is the disk with centre* a_{jj} *and radius* $\rho_j = \sum_{i \neq j} |a_{ij}|$, *in other words,* $V(A)$ *is the convex hull of the (column-)Gershgorin disks of* A.

(v) $W(A)$ *includes the union of the Gershgorin disks* Ω_j.

(vi) $r(A) = \|A\|$ *for every* A.

Proof The first assertion is well known in both the finite and infinite dimensional cases. The second assertion follow easily. Also assertion (iii) is an easy consequence of (ii).

To prove (v), consider the following subset of vector states.

$$\mathcal{S}_j = \{xe_j : \ell_\infty(x) \leq 1, x_j = 1\}$$

It is not too hard to see that $\{(A, Z) : Z \in \mathcal{S}_k\} = \Omega_j$ and hence $\Omega_j \subseteq W(A)$.

By the convexity of $V(A)$, we get $\operatorname{conv}\{\Omega_j : 1 \leq j \leq n\} \subseteq V(A)$. As to the reverse inclusion, we notice that the extreme points of the set of states are precisely the states in $\bigcup_j \mathcal{S}_j$, so that $(A, S) \in \operatorname{conv}\{\Omega_j : 1 \leq j \leq n\}$ for every state S. This proves (iv).

Finally (vi) follows easily from (iv) and (ii). ■

In the following we use the notation $[z_1, z_2]$ for the closed line segment joining the two complex numbers z_1 and z_2, *i.e.*, $[z_1, z_2] = \operatorname{conv}\{z_1, z_2\}$.

Lemma 5.5 *Let* $m < n$, *and let* $\tilde{A} = A \oplus 0_{m-n}$ *where* $A \in M_m$. *The spatial numerical ranges* $W(A)$ *and* $W(\tilde{A})$ *induced by the* ℓ_1 *norm are related by*

$$W(\tilde{A}) = \bigcup_{0 \leq t \leq 1} tW(A)$$

i.e., $W(\tilde{A})$ *is the union of the line segments* $[0, w]$ *for* $w \in W(A)$.

Proof From the description of the states in Lemma 5.4, we see that if $R = \left(\begin{smallmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{smallmatrix}\right) \in \mathcal{R}$ with $R_{11} \in M_m$ and if $R_{11} \neq 0$ then $R_{11}/\operatorname{tr} R_{11}$ is a vector state in M_m. Therefore, $(\tilde{A}, R) = (A, R_{11}) \in tW(A)$, where $t = \operatorname{tr} R_{11}$. Conversely, if $z \in W(A)$, and $1 \leq t \leq 1$, then $z = (A, Z)$ for a vector state $Z = yx^*$ in M_m with $x, y \in \mathbf{C}^m$ such that $\ell_1(x) = \ell_\infty(y) = x^*y = 1$. We define $\tilde{x}$ and $\tilde{y} \in \mathbf{C}^n$ by

$$\tilde{x} = (tx_1, \dots, tx_m, 1 - t, 0, \dots, 0)^t, \quad \tilde{y} = (y_1, \dots, y_m, 1, 1, \dots, 1)^t.$$

Then $\tilde{y}\tilde{x}^*$ is a vector state in M_n, and $(\tilde{A}, \tilde{y}\tilde{x}^*) = t(A, yx^*) \in tW(A)$. ■

Next we give some examples of spatial numerical range calculations. In addition to being illuminating, these examples will be used to show that certain maps on M_n,

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

which preserve the algebra numerical range, nevertheless fail to preserve the spatial numerical range.

Example 1 Let

$$A_\theta = \begin{pmatrix} 2 & -1 \\ e^{i\theta} & -2 \end{pmatrix} \oplus 0_{n-2}, \quad \theta \in [0, 2\pi);$$

in particular

$$A = A_0 = \begin{pmatrix} 2 & -1 \\ 1 & -2 \end{pmatrix} \oplus 0_{n-2}, \quad B = A_\pi = \begin{pmatrix} 2 & -1 \\ -1 & -2 \end{pmatrix} \oplus 0_{n-2}.$$

The numerical ranges of these matrices with respect to the ℓ_1 norm satisfy the following.

(i) $W(A) = V(A) = \operatorname{conv}\big(D(2;1) \cup D(-2;1)\big)$, where $D(z;\rho)$ denotes the disk with centre z and radius ρ
(ii) If $\theta \neq 0$, then $W(A_\theta)$ contains the endpoints but none of the interior points of the line segment $[-2+i, 2+i]$.
(iii) $W(A_\theta)$ is convex for $\theta = 0$ but not convex for $0 < \theta < 2\pi$.

First the assertion about $V(A)$ is a direct consequence of Proposition 5.4. For the spatial numerical range, we first calculate $W(C_\theta)$ for the 2×2 matrices $C_\theta = \left(\begin{smallmatrix} 2 & -1 \\ e^{i\theta} & -2 \end{smallmatrix}\right)$ which are the top left 2×2 compressions of A_θ. The vector states $U_z = (1, z)^t(1, 0)$ for $|z| \leq 1$ give rise to points (C_θ, U_z) in $W(C_\theta)$. These points are precisely the points in the disk $D(2;1)$. Similarly the states $(z, 1)^t(0, 1)$ give us the disk $D(-2, 1) \subseteq W(B_\theta)$. The remaining vector states are

$$\begin{pmatrix} s & (1-s)e^{i\alpha} \\ se^{-i\alpha} & 1-s \end{pmatrix}; \quad 0 < s < 1,\ 0 \leq \alpha < 2\pi.$$

Such a state is a convex combination of states $R_1 = (1, e^{-i\alpha})^t e_1^t$ and $R_2 = (e^{i\alpha}, 1)^t e_2^t$. Furthermore $(C_0, R_1) = 2 + e^{i\alpha}$ and $(C_0, R_2) = -2 - e^{-i\alpha}$ and so $W(C_0)$ include the horizontal line segment $[-2 - e^{-i\alpha}, 2 + e^{i\alpha}]$. The union of all these segments is exactly the convex hull of $D(2;1) \cup D(-2;1)$. Using Lemma 5.5, we see that $W(A) = \operatorname{conv}\big(D(1;1) \cup D(-1;1)\big)$. This proves (i).

We note that the only convex combinations of extreme points in $V(A_\theta)$ that gives an interior point in the line segment $[-2+i, 2+i]$ are just the convex combinations of $-2+i$ and $2+i$. Furthermore if X and Y are states that satisfy $(A_\theta, X) = 2+i$ and $(A_\theta, Y) = -2+i$, then the top left 2×2 compressions of X and Y must be

$$X_0 = \begin{pmatrix} 1 & 0 \\ -ie^{i\theta} & 0 \end{pmatrix} \quad \text{and} \quad Y_0 = \begin{pmatrix} 0 & i \\ 0 & 1 \end{pmatrix}.$$

But then any proper convex combination $sX + (1-s)Y$, $(0 < s < 1)$, has rank one (*i.e.*, is a vector state) if and only if $\theta = 0$. This proves (ii) and (iii). ■

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

The matrices in the next example are obtained from the matrices in the previous example by applying generalized permutations to the columns.

Example 2 Let

$$F = \begin{pmatrix} 2 & \mu & 0 \\ 0 & -2 & 0 \\ \lambda & 0 & 0 \end{pmatrix} \oplus 0_{n-3}, \quad G = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ \lambda & \mu & 0 \end{pmatrix} \oplus 0_{n-3},$$

$$H = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ \lambda & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \end{pmatrix} \oplus 0_{n-4},$$

where λ and μ are arbitrary complex numbers of modulus 1. Then

$$W(F) = W(G) = W(H) = \operatorname{conv}\big(D(2;1) \cup D(-2;1)\big),$$

To verify this, we note that the two indicated disks together with $\{0\}$ are the Gershgorin disks for each of the given matrices. By Proposition 5.4, each of $W(F)$, $W(G)$ and $W(H)$ includes each of the two disks and is included in their convex hull. Every point in the convex hull of these two disks which is not already in their union is evidently a convex combination of a point z_1 in $\partial D(2;1)$, the boundary of $D(2;1)$ and a point z_2 in $\partial D(-2;1)$, the boundary of $D(-2;1)$. Now $z_1 = (F, Z_1)$ and $z_2 = (F, Z_2)$ where $Z_1 = ue_1^t$, $Z_2 = ve_2^t$,

$$u = \big(1, *, \lambda(\overline{z_1} - 2), 0, \dots, 0\big)^t \quad \text{and} \quad v = \big(\mu(\overline{z_2} + 2), 1, *, 0, \dots, 0\big)^t,$$

and where $*$ indicates an arbitrary complex number of modulus at most 1. It clear that those arbitrary entries may be chosen so that the vectors u and v are linearly dependent and so every convex combination of the vector states Z_1 and Z_2 is thus a vector state. This shows that every convex combination of z_1 and z_2 belongs to $W(F)$. A similar calculation establishes the same for G and H.

Example 3 Let

$$A_\theta = \begin{pmatrix} 4 & 1 & 0 \\ 1 & -4 & 0 \\ 1 & ie^{i\theta} & 0 \end{pmatrix} \oplus 0_{n-3}, \quad \theta \in [0, 2\pi).$$

Then $2i \in W(A_\theta)$ if and only if $\theta = 0$.

This is established by similar methods to the methods used in the previous two examples. The Gershgorin disks are $D(4;2)$ and $D(-4;2)$ together with $\{0\}$. For $2i$ to belong to the spatial numerical range, we must have a vector state S such that $(A_\theta, S) = 2i$. But this occurs if and only if the vector state S satisfies $S = (X+Y)/2$ where X and Y are states that satisfy $(A_\theta, X) = 4 + 2i$ and $(A_\theta, Y) = -4 + 2i$. The states that satisfy the later two equations are

$$X = (1, i, i, *, \dots, *)^t e_1^t \quad \text{and} \quad Y = (-i, 1, e^{i\theta}, *, \dots, *)e_2^t.$$

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

But if $\theta \neq 0$, then $(X+Y)/2$ has rank two and so is not a vector state. While if $\theta = 0$, then by taking all the undetermined entries in X and Y to be zero, we get a vector state $(X+Y)/2$.

We are now ready to prove Theorems 5.1, 5.2, and 5.3. We start with second one.

Proof of Theorem 5.2 From the description of the algebra numerical range in Proposition 5.4, it is clear that (c)† implies (b). The implication (b) $\Rightarrow$ (e) follows, as before, from [18]. It remains to prove that (e) $\Rightarrow$ (c)†. As in Section 3, we let $\psi = \phi^*$ and so assume that ψ is a linear operator on M_n satisfying $\psi(\mathcal{S}) = \mathcal{S}$. By Lemma 3.2, there exists $Q \in \mathrm{GP}(n)$ such that the mapping $\tilde{\psi}(X) = Q\psi(X)Q^*$ preserves the usual inner product on M_n, fixes all diagonal matrices, and maps the set of matrices with zero diagonal onto itself. Furthermore, the set of extreme points of $\mathcal{S}$ is the generalized permutation orbit of ee_1^t. This set must also be mapped onto itself by $\tilde{\psi}$. The set

$$\mathcal{C}_k = \{D^* ee_k^t D : D \text{ is diagonal unitary}\}, \quad k = 1, \dots, n,$$

is the intersection of $\operatorname{Ext} \mathcal{S}$ and the set of matrices with diagonal equal $\operatorname{diag} E_{kk}$. Consequently, $\tilde{\psi}(\mathcal{C}_k) = \mathcal{C}_k$. One easily checks that the restriction of $\tilde{\psi}$ on the span of $\mathcal{C}_k$ (identified with $\mathbf{C}^n$) is a linear operator preserving the dual norm ball of the ℓ_1 norm and maps e_k to itself; thus, it is of the form $v \mapsto \tilde{Q}_k v$ for some $\tilde{Q}_k \in \mathrm{GP}(n)$ satisfying $\tilde{Q}_k e_k = e_k$. This proves the form of ϕ^* in (c)†. It is straightforward to establish the equivalence of the form for ϕ and the form of ϕ^* given in (c)†. ∎

Proof of Theorem 5.1 It is clear that (c) $\Rightarrow$ (d). To show that (d) $\Rightarrow$ (a), assume that $\phi^*(\mathcal{R}) = \mathcal{R}$. Then

$$\begin{aligned} W\big(\phi(A)\big) &= \big\{ \big(\phi(A), R\big) : R \in \mathcal{R} \big\} = \big\{ \big(A, \phi^*(R)\big) : R \in \mathcal{R} \big\} \\ &= \{(A,S) : S \in \phi^*(\mathcal{R})\} = \{(A,S) : S \in \mathcal{R}\} \\ &= W(A). \end{aligned}$$

It remains to prove the implication (a) $\Rightarrow$ (c). Assume that ϕ satisfies (a). Since $V(A) = \operatorname{conv} W(A)$, then ϕ preserves the algebra numerical range and so is of the form (c)† of Theorem 5.2. Replacing ϕ by the mapping $A \mapsto P\phi(A)P^t$ for a permutation P, we may assume that

$$\phi([A_1|\cdots|A_n]) = [Q_1A_1|\cdots|Q_nA_n] \quad \text{for all } A = [A_1|\cdots|A_n] \in M_n, \tag{5.1}$$

where $Q_1, \dots, Q_n \in \mathrm{GP}(n)$ are generalized permutations in GP satisfying $Q_j e_j = e_j$ for $j \in \{1, \dots, n\}$. We must then show that $\phi(A) = DAD^*$ for a diagonal unitary D.

We start by showing that Q_1 is a diagonal unitary. If not, then there exist $j \neq k$ such that $Q_1 e_j = -\lambda e_k$ for a complex unit λ. Without loss of generality, we may assume that $j = 2$ and $k = 3$. Furthermore $Q_2(e_1) = \mu e_p$ for a complex unit μ and an index p which may be 1 or 3 or $p \geq 4$. The latter case may be reduced to $p = 4$ as follows. Let P be the permutation matrix obtained from the identity by interchanging the 4-th and the p-th rows, and consider $\hat{\phi}_1$ defined by $A \mapsto P\phi(P^tAP)P^t$. This new map also preserves the spatial numerical range and is of the form (5.1) with generalized permutations $\hat{Q}_j$ acting on the columns such that $\hat{Q}_1(e_2) = -\lambda e_3$ and $\hat{Q}_2(e_1) = e_4$.

Now consider the matrix

$$B = \begin{pmatrix} 2 & -1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \oplus 0_{n-3}$$

of Example 1. Then $\phi(B)$ is one of the matrices F, G or H of Example 2 according as $Q_2(e_1)$ is μe_1, μe_3 or μe_4 respectively. From the calculations in Examples 1 and 2, we have that $W\big(\phi(B)\big)$ is convex while $W(B)$ is not. This is a contradiction, proving that Q_1 is a diagonal unitary. This argument may be used on any column and so we have that Q_j is a diagonal unitary for every j.

Next, we replace ϕ by the map $A \mapsto D^*\phi(A)D$, where D is the matrix Q_1. Thus we may assume that ϕ is of the form (5.1) with $Q_1 = I$. We will then show that every $Q_j = I$ and thus ϕ is the identity. We shall prove this only for $j = 2$ as it can be seen that the same argument applies to any index j. Let $Q_2 = \operatorname{diag}(d_1, 1, d_3, \dots, d_n)$. First consider the matrix $A = \left(\begin{smallmatrix} 2 & -1 \\ 1 & -2 \end{smallmatrix}\right) \oplus 0_{n-2}$, of Example 1. Then $\phi(A) = \left(\begin{smallmatrix} 2 & -d_1 \\ 1 & -2 \end{smallmatrix}\right) \oplus 0_{n-2}$. This is diagonal-unitary equivalent to the matrix $C = \left(\begin{smallmatrix} 2 & -1 \\ d_1^* & -2 \end{smallmatrix}\right) \oplus 0_{n-2}$. By Example 1, we see that $W\big(\phi(A)\big) = W(A)$ if and only if $d_1 = 1$.

Next we consider the matrix

$$T = \begin{pmatrix} 4 & 1 & 0 \\ 1 & -4 & 0 \\ 1 & i & 0 \end{pmatrix} \oplus 0_{n-3},$$

of Example 3. Then

$$\phi(T) = \begin{pmatrix} 4 & 1 & 0 \\ 1 & -4 & 0 \\ 1 & id_3 & 0 \end{pmatrix} \oplus 0_{n-3}.$$

By Example 3, we see that $W(\phi(A) = W(A)$ if and only if $d_3 = 1$. Using PTP^t for a permutation P that fixes e_1 and e_2, we get in the same way that $d_j = 1$ for $j \geq 3$. This proves that $Q_2 = I$ and ends the proof. ■

Proof of Theorem 5.3 By Proposition 5.4, the norm numerical radius coincides with the operator norm. Thus, numerical radius isometries are just the isometries of the operator norm, whose structure is known to be as given in the statement of Theorem 5.3; see [11, 22].

We present another proof. Assume ν is the ℓ_1 norm. Let use denote by $\mathcal{G}$ the group of operators ϕ on M_n of the form

$$[A_1|\cdots|A_n] \mapsto [Q_1A_1|\cdots|Q_nA_n]P;$$

for a permutation P and generalized permutations $Q_1, \dots, Q_n \in \mathrm{GP}(n)$. From the description of the norm and numerical radius given in Proposition 5.4, it follows easily that every $\phi \in \mathcal{G}$ is indeed a numerical radius isometry. For the converse, we have

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press

already seen that A belongs to the set $\mathcal{L}$ defined by (4.1) if and only if every column has exactly one nonzero entry and that this entry has modulus 1, *i.e.*, A belongs to the orbit of I under the action of the group $\mathcal{G}$. Now if ϕ is a spectral radius isometry, then $\phi(I) \in \mathcal{L}$. Therefore, we may compose ϕ with a member of the group $\mathcal{G}$ to get a map $\hat{\phi}$ which preserves the numerical radius and maps I to I. By Proposition 4.19, $\hat{\phi}$ is a unimodular scalar multiple of a map that preserve the algebra numerical range. By Theorem 5.2, every such map belong to $\mathcal{G}$, so the original operator $\phi \in \mathcal{G}$. ∎

Acknowledgments We thank the referee for numerous questions and suggestions which helped improve the presentation.

The research was started when the two authors attended the Linear Algebra Workshop in Bled (Slovenia) in May 1999. The authors would like to thank the organizers for making this collaboration possible.

References

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

[2] F. L. Bauer, *On the field of values subordinate to a norm*. Numer. Math. **4**(1962), 103–113.

[3] F. F. Bonsall and J. Duncan, *Numerical ranges of operators on normed spaces and elements of normed algebras*. London Math. Soc., Lecture Note Series **2**, Cambridge University Press, Cambridge, 1971.

[4] J. T. Chan, *Numerical radius preserving operators on $B(H)$*. Proc. Amer. Math. Soc. **123**(1995), 1437–1439.

[5] ———, *Numerical radius preserving operators on C^*-algebras*. Arch. Math. (Basel) **70**(1998), 486–488.

[6] C. Chevalley, *Theory of Lie groups*. Princeton University Press, Princeton, 1946.

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

[8] P. R. Halmos, *A Hilbert Space Problem Book*. Springer-Verlag, New York, 1974.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, New York, 1985.

[10] R. V. Kadison, *Isometries of operator algebras*. Ann. of Math. **54**(1951), 325–338.

[11] A.-L. S. Klaus and C. K. Li, *Isometries for the vector (p, q) norm and the induced (p, q) norm*. Linear and Multilinear Algebra **38**(1995), 315–332.

[12] C. K. Li, *Linear operators preserving the numerical radius of matrices*. Proc. Amer. Math. Soc. (4) **99**(1987), 601–608.

[13] ———, *A survey on linear preservers of numerical ranges and radii*. Taiwanese J. Math. (3) **5**(2001), 477–496.

[14] C. K. Li and S. Pierce, *Linear operators preserving correlation matrices*. Proc. Amer. Math. Soc., to appear.

[15] C. K. Li and H. Schneider, *Orthogonality of Matrices*. Linear Algebra Appl. **347**(2002), 115–122.

[16] C. K. Li, P. Šemrl and G. Soares, *Linear operators preserving the numerical range (radius) on triangular matrices*. Linear and Multilinear Algebra **48**(2001), 281–292.

[17] N. Nirschl and H. Schneider, *The Bauer fields of values of a matrix*. Numer. Math. **6**(1964), 355–365.

[18] V. Pellegrini, *Numerical range preserving operators on a Banach algebra*. Studia Math. **54**(1975), 143–147.

[19] R. T. Rockafeller, *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[20] S. Rolewicz, *Metric Linear Spaces*. Monografie Matematyczne, Warsaw, 1972.

[21] H. Schneider and R.E.L. Turner, *Matrices hermitian for an absolute norm*. Linear and Multilinear Algebra **1**(1973), 9–31.
[22] A. R. Sourour, *Isometries of* $L^p(\Omega, X)$. J. Funct. Anal. **30**(1978), 276–285.

Department of Mathematics
College of William and Mary
P.O. Box 8795
Williamsburg, Virginia 23187-8795
USA
e-mail: ckli@math.wm.edu.

Department of Mathematics and Statistics
University of Victoria
Victoria, British Columbia
V8W 3P4
e-mail: sourour@math.uvic.ca.

https://doi.org/10.4153/CJM-2004-007-4 Published online by Cambridge University Press