BULL. AUSTRAL. MATH. SOC. 15A63, 15A60
VOL. 66 (2002) [105–117]

H-JOINT NUMERICAL RANGES

CHI-KWONG LI AND LEIBA RODMAN

The notion of the joint numerical range of several linear operators with respect to a sesquilinear form is introduced. Geometrical properties of the joint numerical range are studied, in particular, convexity and angle points, in connection with the algebraic properties of the operators. The main focus is on the finite dimensional case.

1. INTRODUCTION AND ELEMENTARY PROPERTIES

The *joint numerical range* of a k-tuple of $n \times n$ Hermitian matrices $A = (A_1, \dots, A_k)$ is defined by

$$W(A_1, \dots, A_k) = \{(x^*A_1x, \dots, x^*A_kx) : x \in \mathbb{C}^n, x^*x = 1\} \subseteq \mathbb{R}^{1\times k}.$$

We denote by $\mathbb{C}^n$ the n-dimensional complex linear space of column vectors, and by $\mathbb{R}^{1\times k}$ the k-dimensional real linear space of row vectors. The joint numerical range has been studied by many researchers (for example, see [**1, 2, 10**] and their references), and is useful in various theoretical and applied subjects, (for example, see [**7, 12**] and their references), in particular, control systems (structured singular value) [**4, 14**].

By analogy with the joint numerical range of several Hermitian matrices, we introduce the notion of joint numerical range with respect to a Hermitian matrix.

To set up notation, let $\mathcal{H}_n$ be the set of $n \times n$ Hermitian matrices. Suppose $H \in \mathcal{H}_n$ and $A = (A_1, \dots, A_k) \in \mathcal{H}_n^k$. Then the *H-joint numerical range* of $(A_1, \dots, A_k)$, or in short the *H-numerical range* of A is defined by

$$W_H(A) = W_H(A_1, \dots, A_k) = \{(x^*A_1x, \dots, x^*A_kx)/(x^*Hx) : x \in \mathbb{C}^n, x^*Hx \neq 0\} \subseteq \mathbb{R}^{1\times k}.$$

We also consider the sets

$$W_H^{\pm}(A) = \{\pm(x^*A_1x, \dots, x^*A_kx) : x \in \mathbb{C}^n, x^*Hx = \pm 1\}.$$

It will be implicitly assumed for the definition of $W_H(A)$ that $x^*Hx \neq 0$ for some $x \in \mathbb{C}^n$, and for the definitions of $W_H^{\pm}(A)$ we assume that $x^*Hx = \pm 1$ for some $x \in \mathbb{C}^n$.

When $H = I$, $W_H(A) = W_H^+(A)$ reduces to the usual joint numerical range $W(A)$. The following properties can be easily proved.

Received 24th January, 2002
Research for the authors was supported by NSF grants DMS–0071944 and DMS–9988579.

(A) $W_H(A) = W_H^+(A) \cup W_H^-(A)$.

(B) $W_H^-(A) = W_{(-H)}^+(-A)$.

By Property (B), we can focus our study on $W_H^+(A)$ and translate the results on $W_H^+(A)$ to $W_H^-(A)$ and $W_H(A)$.

In this paper we study geometric properties, such as convexity, closedness, boundedness, *et cetera* of the sets $W_H(A)$ and $W_H^\pm(A)$ in connection with the algebraic properties of $A_1, \ldots, A_k$. These problems were raised in [**10**].

In studies of the numerical range and its generalisations, a useful technique is reducing the problems to the 2×2 case. For instance, convexity results are proven using such a reduction (in clever ways), see [**1, 5, 9, 13**]. So, in Section 2 we give a complete description of the 2×2 case in our study of H-joint numerical ranges, and use this description in Section 3, as well as known results on convexity of joint numerical ranges of Hermitian matrices, to derive convexity results for H-joint numerical ranges. Furthermore, the careful analysis of the 2×2 case often leads to interesting results, such as characterisation of non-differentiable boundary points of the (generalised) numerical range, and of the conditions under which the (generalised) numerical range is a polygon (with interior), a line segment, or a singleton, see [**3, 6, 11**]. In fact, we address these problems for H-joint numerical ranges in Sections 3 and 4. In the last section we consider H-joint numerical ranges in infinite dimensional spaces.

In the remaining part of the introduction, we list several elementary and useful properties of H-joint numerical ranges.

(C) For any $n \times m$ matrix S , let $S^*AS = (S^*A_1S, \ldots, S^*A_kS)$. Then

$$W_{S^*HS}(S^*AS) \subseteq W_H(A) \quad \text{and} \quad W_{S^*HS}^+(S^*AS) \subseteq W_H^+(A).$$

The set inclusions become equalities if S is $n \times n$ and invertible.

(D) Suppose $T = (t_{ij})$ is a $k \times q$ real matrix and $B = (B_1, \ldots, B_q) \in \mathcal{H}_n^q$ is such that $B_j = \sum_{i=1}^{k} t_{ij} A_i$, $j = 1, \ldots, q$. Then

$$W_H(B) = \{vT : v \in W_H(A)\} \quad \text{and} \quad W_H^+(B) = \{vT : v \in W_H^+(A)\}.$$

Consequently, if $W_H^+(A)$ is convex, or bounded, or compact, then $W_H^+(B)$ has the corresponding property. In particular, if T is square size and invertible, then $W_H^+(A)$ and $W_H^+(B)$ share the same geometrical properties. Thus, to study the geometrical properties of $W_H^+(A)$, we may replace $A = (A_1, \ldots, A_k)$ by $B = (B_1, \ldots, B_m, 0, \ldots, 0)$, or simply by $B = (B_1, \ldots, B_m)$, where $\{B_1, \ldots, B_m\}$ is a basis for $\operatorname{span}\{A_1, \ldots, A_k\}$.

The next observation was made in [**10**]. It shows that one can study $W_H^+(A)$ and $W_H(A)$ via the cone generated by the usual joint numerical range. Define

$$K(H, A_1, \ldots, A_k) = \{(x^*Hx, x^*A_1x, \ldots, x^*A_kx) : x \in \mathbb{C}^n\}. \tag{1}$$

(E) Let $(a_1, \dots, a_k)$ be a real vector. Then $(a_1, \dots, a_k) \in W_H^+(A_1, \dots, A_k)$ if and only if $(1, a_1, \dots, a_k) \in K(H, A_1, \dots, A_k)$; $(a_1, \dots, a_k) \in W_H(A_1, \dots, A_k)$ if and only if

$$(1, a_1, \dots, a_k) \in K(H, A_1, \dots, A_k) \text{ or } -(1, a_1, \dots, a_k) \in K(H, A_1, \dots, A_k).$$

(F) If H is positive definite, then $W_H^+(A_1, \dots, A_k)$ is compact; otherwise, $W_H^+(A_1, \dots, A_k)$ need not be closed or bounded.

The next observation is an extension of an idea in [7]. We denote by $\operatorname{tr} X$ the trace of a matrix X.

(G) Let

$$V_H = \{xx^* : x \in \mathbb{C}^n,\ x^*Hx = 1\}.$$

Then $W_H^+(A)$ is the image of the set V_H under the linear map $\phi_A : \mathcal{H}_n \to \mathbb{R}^{1\times k}$ defined by

$$\phi_A(X) = (\operatorname{tr} A_1X, \dots, \operatorname{tr} A_kX), \quad X \in \mathcal{H}_n. \tag{2}$$

We have:

$$\operatorname{rank} \phi_A = \text{dimension of span}\,\{A_1, \dots, A_k\}. \tag{3}$$

To verify (3), observe that by Property (D) we may assume that $A_1, \dots, A_k$ are linearly independent. Then (3) amounts to the assertion that ϕ_A is onto. If ϕ_A were not onto, then there would exist $y_1, \dots, y_k \in \mathbb{R}$, not all zeros, such that $\sum_{j=1}^{k} y_j \operatorname{tr} A_jX = 0$ for all $X \in \mathcal{H}_n$. Then $\operatorname{tr}\left(\sum_{j=1}^{k} y_jA_j\right)X = 0$ for all $X \in \mathcal{H}_n$, which implies $\sum_{j=1}^{k} y_jA_j = 0$, a contradiction with the linear independence of $A_1, \dots, A_k$.

(H) The set $W_H(A)$ is always connected. Indeed, it is easy to see that the set

$$\{x \in \mathbb{C}^n \ :\ x^*Hx = 1\}$$

is connected (assuming without loss of generality that H is diagonal). Therefore, the set V_H is connected. The assertion now follows readily from Property (G).

2. Two by two case

We first consider the two by two case. In view of Property (G) and equality (3), the shape of $W_H^+(A_1, \dots, A_k)$ can be determined via the description of V_H. Recall that a subset $\mathcal{S}$ in $\mathbb{R}^{1\times k}$ has affine dimension m if there exists $v \in \mathcal{S}$ such that the linear span of $\mathcal{S} - v$ has dimension m. Since $V_H \subseteq \mathcal{H}_2$ has real affine dimension 4, we see by (G) that $W_H^+(A)$ has affine dimension at most 4. We have the following theorem.

https://doi.org/10.1017/S0004972700020724 Published online by Cambridge University Press

THEOREM 2.1. *Suppose $H, A_1, \ldots, A_k \in \mathcal{H}_2$, $A = (A_1, \ldots, A_k)$, and $\operatorname{span}\{H, A_1, \ldots, A_k\}$ has dimension $m \leqslant 4$. Then $W_H^+(A)$ has affine dimension $m-1$. In particular, if $m = 1$, then $W_H^+(A)$ is a singleton. Assume $m > 1$ in* (a), (b), *and* (c) *below.*

(a) *Suppose H is positive definite. Then*

$$V_H = \left\{ H^{1/2} \begin{pmatrix} 1/2 + a & b + ic \\ b - ic & 1/2 - a \end{pmatrix} H^{1/2} : a^2 + b^2 + c^2 = 1/4 \right\}. \tag{4}$$

Moreover, $W_H^+(A)$ is a closed line segment, a closed ellipse with interior, or a closed ellipsoid without interior, depending on $m = 2, 3$, or 4, respectively.

(b) *Suppose there is an invertible matrix S such that $S^*HS = \operatorname{diag}(1,0)$. Then*

$$V_H = \left\{ S \begin{pmatrix} 1 & b + ic \\ b - ic & a \end{pmatrix} S^* : a = b^2 + c^2 \right\}. \tag{5}$$

Moreover, $W_H^+(A)$ is a closed half line or a straight line if $m = 2$; $W_H^+(A)$ is a closed parabola with interior or a two-dimensional plane if $m = 3$; $W_H^+(A)$ is a closed paraboloid without interior if $m = 4$.

(c) *Suppose there is an invertible matrix S such that $S^*HS = \operatorname{diag}(1,-1)$. Then*

$$V_H = \left\{ S \begin{pmatrix} a + 1/2 & b + ic \\ b - ic & a - 1/2 \end{pmatrix} S^* : a^2 = b^2 + c^2 + 1/4 \text{ and } a \geqslant 1/2 \right\}. \tag{6}$$

Moreover, $W_H^+(A)$ a closed or open half line or a straight line if $m = 2$; $W_H^+(A)$ is a closed one-component hyperbola with interior, an open two-dimensional half plane, or a two-dimensional plane if $m = 3$; a closed one-component hyperboloid without interior if $m = 4$.

PROOF: If $m = 1$, then all A_j's are multiples of H. Thus, $W_H^+(A)$ is a singleton.

Suppose $m > 1$. Note that

$$W_H^+(H, A_1, \ldots, A_k) = \{(1, a_1, \ldots, a_k) : (a_1, \ldots, a_k) \in W_H^+(A)\}.$$

Thus, to prove (a) – (c), we can always assume that $H \in \operatorname{span}\{A_1, \ldots, A_k\}$. Replacing each A_j by a suitable linear combination of $A_1, \ldots, A_k$ and using Property (D), we may assume without loss of generality that $A_1 = H$ and $k = m$.

If $k = 4$, then by (3) the map ϕ_A defined in (2) is invertible, and therefore $W_H^+(A) = \phi_A(V_H)$ has the same geometric shape as V_H, and the result follows by inspection of the shape of V_H. Thus, we shall focus on the cases $1 < k < 4$ in the following discussion once V_H is determined.

(a) By Property (C), without loss of generality we may assume $H = I$. Note that $\begin{pmatrix} 1/2 + a & b + ic \\ b - ic & 1/2 - a \end{pmatrix}$, where $a^2 + b^2 + c^2 = 1/4$, is a general form of a 2×2 rank 1 orthogonal projection. Thus, V_H indeed is given by (4).

Assume $1 < k < 4$. Since $A_1 = I_2$, we may replace each A_j by a suitable linear combinations of $A_1, \ldots, A_k$ so that

$$A_2 = \begin{pmatrix} d_2 & q_2 \\ q_2^* & -d_2 \end{pmatrix}, \quad \text{and} \quad A_3 = \begin{pmatrix} 0 & q_3 \\ q_3^* & 0 \end{pmatrix} \quad \text{if } k = 3.$$

By (4) and (G), we have

$$W_H^+(A) = \left\{ \left(1, 2\left[d_2 a + \mathrm{Re}\big(q_2(b - ic)\big)\right]\right) : a^2 + b^2 + c^2 = 1/4 \right\} \quad \text{if } k = 2;$$

and

$$W_H^+(A) = \left\{ \left(1, 2\left[d_2 a + \mathrm{Re}\big(q_2(b - ic)\big)\right], 2\,\mathrm{Re}\big(q_3(b - ic)\big)\right) : a^2 + b^2 + c^2 = 1/4 \right\} \quad \text{if } k = 3.$$

By elementary considerations, we see that $W_H^+(A_1, \ldots, A_k)$ satisfies condition (a).

(b) Let S be invertible such that $S^*HS = \mathrm{diag}(1, 0)$. Then $x^*Hx = 1$ if and only if $x = Sy$ such that $y^*\mathrm{diag}(1, 0)y = 1$, equivalently, $xx^* = Syy^*S^*$ with

$$yy^* = \begin{pmatrix} 1 & b + ic \\ b - ic & a \end{pmatrix}$$

for some $a, b, c \in \mathbb{R}$ satisfying $a = b^2 + c^2$. Thus, without loss of generality, we may assume that $H = \mathrm{diag}(1, 0)$ and $S = I$. Then

$$V_H = \left\{ \begin{pmatrix} 1 & b + ic \\ b - ic & a \end{pmatrix} : a = b^2 + c^2 \right\}. \tag{7}$$

Assume first $k = 2$. Recalling our standing assumption $A_1 = H$, and replacing if necessary A_2 by a linear combination of A_1 and A_2, we may assume that A_2 has the form $A_2 = \begin{pmatrix} 0 & q \\ q^* & s \end{pmatrix}$, for some $q \in \mathbb{C}$ and $s \in \mathbb{R}$, $s \geqslant 0$, where s and q are not both zeros. Then, for $X \in V_H$ given in the form (7), we have

$$\mathrm{tr}\,(A_2 X) = 2\mathrm{Re}\,(q\sqrt{a}e^{it}) + sa, \tag{8}$$

where $\sqrt{a}e^{it} = b - ic$. So, as $a \geqslant 0$ and $t \in \mathbb{R}$ vary, the range of ϕ_A on V_H is a straight line if $s = 0$, and is a closed half line with end point

$$\min\left\{2\mathrm{Re}\,(q\sqrt{a}e^{it}) + sa : a \geqslant 0,\ t \in \mathbb{R}\right\} = -|q|^2/s$$

if $s > 0$.

Assume now $k = 3$. Again, $A_1 = H$, and replacing A_2 and A_3 by their suitable linear combinations with H, we can assume that

$$A_2 = \begin{pmatrix} 0 & q_2 \\ q_2^* & s_2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & q_3 \\ q_3^* & 0 \end{pmatrix}, \tag{9}$$

https://doi.org/10.1017/S0004972700020724 Published online by Cambridge University Press

where $q_2, q_3 \in \mathbb{C}$, and either $s_2 = 1$ or $s_2 = 0$. In the latter case ($s_2 = 0$), the function $(\operatorname{tr}(A_2X), \operatorname{tr}(A_3X))$ maps V_H onto $\mathbb{R}^{1\times 2}$, hence W_H^+ is a two-dimensional plane. In the former case ($s_2 = 1$) the function $(\operatorname{tr}(A_2X), \operatorname{tr}(A_3X))$ maps V_H onto a closed parabola with interior in $\mathbb{R}^{1\times 2}$.

(c) Let S be invertible such that $S^*HS = \operatorname{diag}(1,-1)$. Then $x^*Hx = 1$ if and only if $x = Sy$ such that $y^*\operatorname{diag}(1,-1)y = 1$, equivalently, $xx^* = Syy^*S^*$ with

$$yy^* = \begin{pmatrix} a+1/2 & b+ic \\ b-ic & a-1/2 \end{pmatrix}$$

for some $a, b, c \in \mathbb{R}$ satisfying $a^2 = b^2 + c^2 + 1/4$ and $a \geqslant 1/2$. We may assume therefore that $H = \operatorname{diag}(1,-1)$ and

$$(10)\quad V_H = \left\{ \begin{pmatrix} a+1/2 & b+ic \\ b-ic & a-1/2 \end{pmatrix} \ : \ a, b, c \in \mathbb{R} \text{ such that } a^2 = b^2 + c^2 + 1/4, a \geqslant 1/2 \right\}.$$

Assume $k = 2$. Since $A_1 = H = \operatorname{diag}(1,-1)$, we may assume that $A_2 = \begin{pmatrix} 0 & q \\ q^* & s \end{pmatrix}$ for some $q \in \mathbb{C}$, $s \geqslant 0$, where q and s are not both zero. Then for $X \in V_H$ given by (10), we have

$$\operatorname{tr}(A_2X) = 2\operatorname{Re}\left(q(b-ic)\right) + s(a-1/2).$$

So, if $s = 0$, then $W_H^+(A)$ is a straight line, and if $q = 0$, then $W_H^+(A)$ is a closed half line. Otherwise, applying simultaneous congruence $A_1 = T^*A_1T$, $A_2 \to T^*A_2T$, for a suitable unitary diagonal matrix T, and scaling A_2, we can further assume that $s = 1$ and $q > 0$. Then for $X \in V_H$, where X is given by the matrix in the right hand side of (10), we have $\operatorname{tr}(A_2X) = 2qb + (a-1/2)$, where $|b| \leqslant \sqrt{a^2 - (1/4)}$. Denoting $y = a - 1/2$, it is easy to see that the range of $\operatorname{tr}(A_2X)$ coincides with the range of the function

$$f_q(y) = y - 2q\sqrt{y(y+1)}, \qquad y \geqslant 0.$$

Elementary analysis shows that the function $f_q(y)$ has only one minimum if $0 < q < 1/2$, which is attained at $y = \left(-1 + (1-4q^2)^{-1/2}\right)/2$. Also, if $q > 1/2$, then $f_q(y) \to -\infty$ as $y \to \infty$, and if $q = 1/2$, then $f_q(y)$ is a decreasing function with $f_q(y) \to -1/2$ as $y \to \infty$. Thus, the range of $\operatorname{tr}(A_2X)$, $X \in V_H$, is a closed half line if $0 < q < 1/2$, an open half line if $q = 1/2$, and the whole of $\mathbb{R}$ if $q > 1/2$.

Assume now $k = 3$. We may further assume that A_2 and A_3 have the form (9), where either $s_2 = 0$ or $s_2 = 1$. In the former case,

$$W_H^+(A) = \left\{ \left(1, 2\operatorname{Re}\left(q_2(b-ic)\right), 2\operatorname{Re}\left(q_3(b-ic)\right)\right) : b, c \in \mathbb{R} \right\}.$$

Since q_2 and q_3 are linearly independent over $\mathbb{R}$, we see that $W_H^+(A)$ is a two-dimensional plane.

In the latter case, that is, when $s_2 = 1$, we may assume that $q_3 > 0$. Otherwise, let D be a diagonal unitary matrix such that the $(1,2)$ and $(2,1)$ entry of D^*A_3D equal $|d_3|$, and replace (A_1, A_2, A_3) by $(D^*A_1D, D^*A_2D, D^*A_3D)$. We may further replace A_2 by $2A_2 + A_1 - 2(\operatorname{Re} q_2/q_3)A_3$ so that

$$A_2 = \begin{pmatrix} 1 & iy \\ -iy & 1 \end{pmatrix}, \quad y \in \mathbb{R}.$$

Then

$$W_H^+(A_1, A_2, A_3) = \left\{\left(1,\ 2\sqrt{b^2 + c^2 + 1/4} + 2yc,\ 2q_3b\right) : b, c \in \mathbb{R}\right\}.$$

For each $b \in \mathbb{R}$, we can apply elementary calculus to the function

$$f(c) = \sqrt{b^2 + c^2 + 1/4} + yc, \qquad c \in \mathbb{R},$$

to conclude that

(i) if $|y| = 1$, then $f(c) > 0$ for all $c \in \mathbb{R}$, and $\inf_{c\in\mathbb{R}} f(c) = 0$;

(ii) if $|y| < 1$, then

$$\min_{c\in\mathbb{R}} f(c) = \sqrt{(1-y^2)(b^2+1/4)} \quad \text{occurs at} \quad c = -y\sqrt{(b^2+1/4)/(1-y^2)}.$$

As a result,

$$\{f(c) : c \in \mathbb{R}\} = \begin{cases} \left[\sqrt{(1-y^2)(b^2+1/4)}, \infty\right) & \text{if } |y| < 1, \\ (0, \infty) & \text{if } |y| = 1, \\ \mathbb{R} & \text{if } |y| > 1. \end{cases}$$

Hence, $W_H^+(A)$ is a closed one-component hyperbola with interior, an open two-dimensional half plane, or a two-dimensional plane. ∎

All the shapes of $W_H^+(A)$ of k-tuples A of 2×2 Hermitian matrices as asserted in Theorem 2.1 actually appear in examples. We present here only one example (taken from [**10**]) in which the joint numerical range is an open half plane.

EXAMPLE 2.2. Let

$$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Letting $x = \begin{pmatrix} y \\ z \end{pmatrix} \in \mathbb{C}^2$, $y, z \in \mathbb{C}$, a computation shows that

$$W_H^+(A) = \left\{(2y^*y, iz^*y - iy^*z) \in \mathbb{R}^{1\times 2} : z^*y + yz^* = 1\right\}.$$

Thus, $W_H^+(A) = \left\{(a, b) \in \mathbb{R} : a > 0\right\}$ is indeed an open half plane.

The proof of Theorem 2.1 shows that under given hypotheses some shapes of $W_H^+(A)$ are *generic*, that is, for fixed H they appear for all k-tuples of linearly independent 2×2 Hermitian matrices $A = (A_1, \dots, A_k)$, excepting a proper algebraic set, and other shapes are *special*, that is, they appear for A belonging in a proper algebraic set, as follows:

COROLLARY 2.3. *Suppose $H, A_1, \dots, A_k \in \mathcal{H}_2$, $A = (A_1, \dots, A_k)$, and* $\operatorname{span}\{H, A_1, \dots, A_k\}$ *has dimension $m \leqslant 4$.*

(a) *Assume that H is positive semidefinite and singular. Then for $m = 2$ the closed half line is a generic shape of $W_H^+(A)$, whereas the straight line is a special shape; for $m = 3$, the closed parabola with interior is a generic shape of $W_H^+(A)$, whereas the two-dimensional plane is a special shape.*

(b) *Assume that H is indefinite. Then for $m = 2$, the shape of $W_H^+(A)$ which is either the closed half line or the straight line is generic, whereas the open half line is a special shape; for $m = 3$, the closed one-component hyperbola with interior and the two-dimensional plane are generic shapes of $W_H^+(A)$, whereas the open two-dimensional half plane is a special shape.*

3. Convexity and affine dimension

By Theorem 2.1, we have the following result concerning the convexity of $W_H^+(A)$ for matrices in $\mathcal{H}_2$.

PROPOSITION 3.1. *Let $H, A_1, \dots, A_k \in \mathcal{H}_2$ and $A = (A_1, \dots, A_k)$. Then $W_H^+(A)$ is convex if and only if* $\operatorname{span}\{H, A_1, A_2, \dots, A_k\}$ *has dimension less than 4.*

In general, we only have sufficient conditions for the convexity of $W_H^+(A)$ as shown in the following.

THEOREM 3.2. *Let $H \in \mathcal{H}_n$, and $A = (A_1, \dots, A_k) \in \mathcal{H}_n^k$. Then $W_H^+(A)$ is convex if one of the following holds.*

(a) $\operatorname{span}\{H, A_1, \dots, A_k\}$ *has dimension less than or equal to three.*

(b) *$n \geqslant 3$ and* $\operatorname{span}\{H, A_1, \dots, A_k\}$ *has dimension four and contains a (positive or negative) definite matrix.*

(c) *There exists an invertible S such that* $\operatorname{span}\{S^*HS, S^*A_1S, \dots, S^*A_kS\} \subseteq \mathcal{S}$, *where $\mathcal{S}$ is one of the following subspaces:*

$$\mathcal{S}_1 = \left\{ aI_n + \begin{pmatrix} 0 & v^* \\ v & 0_{n-1} \end{pmatrix} : a \in \mathbb{R},\ v \in \mathbb{C}^{n-1} \right\}$$

or

$$\mathcal{S}_2 = \left\{ A \in \mathcal{H}_n : A \text{ is a real tridiagonal matrix } \right\}.$$

PROOF: Suppose (a) holds. By Property (D), we may assume that $k \leqslant 2$. Then for any $x, y \in \mathbb{C}^n$ satisfying $x^*Hx = y^*Hy = 1$, we can consider $S = [x|y]$. By Theorem

2.1, $W^+_{S^*HS}(S^*A_1S, \dots, S^*A_kS)$ is a convex set containing the points $(x^*A_1x, \dots, x^*A_kx)$ and $(y^*A_1y, \dots, y^*A_ky)$, and thus containing the whole line segment joining them. By Property (C), we see that the line segment lies in $W^+_H(A_1, \dots, A_k)$.

Suppose (b) holds. By Property (D), we may assume that $k = 3$ and A_3 is positive definite. We may further assume that $A_3 = I_n$; otherwise, replace X by $A_3^{-1/2}XA_3^{-1/2}$ for $X = H, A_1, A_2, A_3$. Then $W(H, A_1, A_2, I) = \{(a, b, c, 1) : (a, b, c) \in W(H, A_1, A_2)\}$ is convex and so is $K(H, A_1, A_2, I)$. By Property (E), $W^+_H(A_1, A_2, A_3)$ is convex.

Suppose condition (c) holds. By Theorems 2.1 and 3.2 of [**7**], $W(H, A_1, \dots, A_k)$ is convex. By Property (E), the result follows. $\square$

By Theorem 3.2 (c), for $n \geqslant 3$, there exist linearly independent $H, A_1, A_2, A_3 \in \mathcal{S}_1$ such that $\operatorname{span}\{H, A_1, A_2, A_3\}$ does not contain a definite matrix, but $W^+_H(A_1, A_2, A_3)$ is convex. However, in general, there exist A_1, A_2, A_3 such that $W^+_H(A_1, A_2, A_3)$ is not convex. We have the following stronger non-convexity result.

Theorem 3.3. *Suppose $H, A_1, A_2 \in \mathcal{H}_n$ are linearly independent, and $\operatorname{span}\{H, A_1, A_2\}$ does not contain a positive definite matrix. Then there exists $A_3 \in \mathcal{H}_n$ such that $W^+_H(A_1, A_2, A_3)$ is not convex.*

Proof: Suppose $H, A_1, A_2 \in \mathcal{H}_n$ satisfy the hypotheses. Then the set $\{I, H, A_1, A_2\}$ is linearly independent. By Theorem 4.1 and its proof in [**7**, pp. 673–674], we see that there exists an $n \times 2$ matrix P such that $P^*P = I_2$ and the three matrices P^*HP, P^*B_1P, P^*B_2P, are linearly independent and indefinite, where $B_1, B_2 \in \operatorname{span}\{I, A_1, A_2\}$. Let $A_3 = PP^*$. Then

$$W^+_H(B_1, B_2, A_3) \cap \{(a, b, 1) : a, b \in \mathbb{R}\}$$

is not convex. Thus, $W^+_H(B_1, B_2, A_3)$ is not convex. It follows that $W^+_H(A_1, A_2, A_3)$ is not convex. $\square$

Theorem 2.1 contains information about the affine dimension of $W^+_H(A)$ for matrices in $\mathcal{H}_2$. We show that the result is also valid for higher dimensions.

Theorem 3.4. *Let $H, A_1, \dots, A_k \in \mathcal{H}_n$. Then $W^+_H(A_1, \dots, A_k)$ has affine dimension m if and only if $\operatorname{span}\{H, A_1, \dots, A_k\}$ has dimension $m + 1$. Consequently,*

(a) *$W^+_H(A_1, \dots, A_k)$ is a point if and only if A_j is a multiple of H for each $j \in \{1, \dots, k\}$.*

(b) *$W^+_H(A_1, \dots, A_k)$ is a subset of a line if and only if $A_j = a_jH + b_jK$ for some $K \in \mathcal{H}_n$ for each $j \in \{1, \dots, k\}$.*

Proof: Note that

$$W^+_H(H, A_1, \dots, A_k) = \{(1, a_1, \dots, a_k) : (a_1, \dots, a_k) \in W^+_H(A)\}.$$

Thus, the sets $W^+_H(H, A_1, \dots, A_k)$ and $W^+_H(A)$ have the same affine dimension. Furthermore, using Property (D), we may assume without loss of generality that $H, A_1, A_2, \dots, A_k$

https://doi.org/10.1017/S0004972700020724 Published online by Cambridge University Press

are linearly independent, that is, $m = k$. Then for any $v \in W_H^+(H, A_1, \dots, A_k)$, we have

$$W_H^+(H, A_1, \dots, A_k) - v \subseteq \{(0, a_1, \dots, a_k) : a_j \in \mathbb{R},\ j = 1, \dots, k\}.$$

We see that the affine dimension of $W_H^+(A)$ is at most k. Using arguments in the proof of (3), we see that the affine dimension of $W_H^+(A)$ is indeed $k = m$. Statements (a) and (b) are now clear. ∎

4. Polyhedral properties and angle points

We describe a sufficient condition on $H, A_1, \dots, A_k \in \mathcal{H}_n$ so that $W_H^+(A_1, \dots, A_k)$ is polyhedral, that is, the set is the intersection of finitely many half spaces.

Theorem 4.1. *Suppose $H, A_1, \dots, A_k \in \mathcal{H}_n$ and there exists an invertible matrix S such that $S^*HS = D_0 \oplus B_0$ and $S^*A_jS = D_j \oplus B_j$ for $j = 1, \dots, k$, where $D_0, \dots, D_k$ are diagonal matrices and $W(B_0, \dots, B_k) \subseteq W(H, A_1, \dots, A_k)$. Then $W_H^+(A_1, \dots, A_k)$ is a polyhedral set.*

Proof: By the assumption, we have (see (1) and (C))

$$K(H, A_1, \dots, A_k) = K(D_0, D_1, \dots, D_k),$$

which is a polyhedral cone; see [**2**]. By Property (E), the result follows. ∎

A point $q \in W_H^+(A)$ is called an *angle point* of $W_H^+(A)$ if there exist a nonzero $p \in \mathbb{R}^{1\times k}$ and a positive number α such that

$$p(v - q)^T \geqslant \alpha \|v - q\| \tag{11}$$

for all $v \in W_H^+(A)$ sufficiently close to q.

Theorem 4.2. *Let $H, A_1, \dots, A_k \in \mathcal{H}_n$. If $q \in W_H^+(A_1, \dots, A_k)$, $q = (x^*A_1x, \dots, x^*A_jx)$ for some x such that $x^*Hx = 1$ is an angle point of $W_H^+(A_1, \dots A_k)$, then*

$$A_jx = (x^*A_jx)Hx, \quad j = 1, \dots, k. \tag{12}$$

As we shall see later, equation (12) when suitably interpreted under some additional hypotheses, shows existence of a joint eigenvector. Thus, Theorem 4.2 can be thought of as a generalisation of Theorem 3.1 in [**8**].

Proof: For simplicity of notation we assume that the nonzero vector $p \in \mathbb{R}^{1\times k}$, whose existence is assured by q being an angle point of $W_H^+(A) = W_H^+(A_1, \dots A_k)$, is equal to $(1, 0, \dots, 0)$. Fix an arbitrary vector $y \in \mathbb{C}^n$. Then for real t sufficiently close to zero,

$$v := \left(\frac{(x+ty)^*A_1(x+ty)}{(x+ty)^*H(x+ty)}, \dots, \frac{(x+ty)^*A_k(x+ty)}{(x+ty)^*H(x+ty)} \right) \in W_H^+(A)$$

is close to q. Write $v = (v_1, \dots, v_k)$, $q = (q_1, \dots, q_k)$. Then

$$v_j - q_j = t\big(y^*A_jx + x^*A_jy - (x^*A_jx)(y^*Hx + x^*Hy)\big) + O(t^2)$$

as t approaches zero. Thus, equation (11) gives

$$(13)\quad t\big(y^*A_1x + x^*A_1y - (x^*A_1x)(y^*Hx + x^*Hy)\big) + O(t^2) \\ \geqslant \alpha|t|\left(\sum_{j=1}^{k}\big(y^*A_jx + x^*A_jy - (x^*A_jx)(y^*Hx + x^*Hy)\big)^2\right)^{1/2}.$$

If $y^*A_1x + x^*A_1y - (x^*A_1x)(y^*Hx + x^*Hy)$ were positive, we would obtain a contradiction with (13) when $t < 0$, and if $y^*A_1x + x^*A_1y - (x^*A_1x)(y^*Hx + x^*Hy)$ were negative, we would obtain a contradiction with (13) when $t > 0$. Thus,

$$y^*A_1x + x^*A_1y - (x^*A_1x)(y^*Hx + x^*Hy) = 0.$$

Therefore, since y is arbitrary, we obtain

$$A_1x - (x^*A_1x)Hx = 0.$$

Now (13) gives

$$y^*\big(A_jx - (x^*A_jx)Hx\big) + \big(A_jx - (x^*A_jx)Hx\big)^*y = 0, \quad j = 2, 3, \dots,$$

for every vector y. Therefore,

$$A_jx - (x^*A_jx)Hx = 0, \quad j = 2, 3, \dots.$$

□

5. A DIFFERENT FORMULATION AND THE INFINITE DIMENSIONAL CASE

We now offer a slightly different interpretation of $W_H^+(A)$ and $W_H(A)$. Consider a fixed sesquilinear form $[x, y]$, $x, y \in \mathbb{C}^n$. We introduce the joint numerical range with respect to $[\cdot,\cdot]$. Namely, let $X = (X_1, \dots, X_k)$ be an ordered k-tuple of linear transformations on $\mathbb{C}^n$ such that each X_j is $[\cdot,\cdot]$-selfadjoint, that is, $[X_jx, y] = [x, X_jy]$ for all $x, y \in \mathbb{C}^n$. Define the *joint numerical range* of X with respect to $[\cdot,\cdot]$ by

$$W_{[\cdot,\cdot]}(X) = \Big\{\big([X_1x, x], \dots, [X_kx, x]\big)/[x, x] \in \mathbb{R}^{1\times k} : x \in \mathbb{C}^n, [x, x] \neq 0\Big\}.$$

Define also the related sets

$$W_{[\cdot,\cdot]}^+(X) = \Big\{\big([X_1x, x], \dots, [X_kx, x]\big) \in \mathbb{R}^{1\times k} : x \in \mathbb{C}^n, [x, x] = 1\Big\}$$

and

$$W_{[\cdot,\cdot]}^-(X) = \Big\{-\big([X_1x, x], \dots, [X_kx, x]\big) \in \mathbb{R}^{1\times k} : x \in \mathbb{C}^n, [x, x] = -1\Big\}.$$

https://doi.org/10.1017/S0004972700020724 Published online by Cambridge University Press

If $k = 1$, then we obtain the notion of numerical range of X with respect to a sesquilinear form that was studied in [**8**, **10**].

It will be convenient to work with matrices representing linear transformations with respect to the standard basis in $\mathbb{C}^n$. Thus, let H be an $n \times n$ Hermitian matrix such that $[x, y] = y^* H x$ for all $x, y \in \mathbb{C}^n$. A matrix X_j is $[\cdot,\cdot]$-selfadjoint if and only if HX_j is Hermitian. Hence

$$W_{[\cdot,\cdot]}(X) = \{(x^* H X_1 x, \ldots, x^* H X_k x)/(x^* H x) : x \in \mathbb{C}^n, x^* H x \neq 0\}$$

and

$$\begin{aligned} W^{\pm}_{[\cdot,\cdot]}(X) &= \{\pm(x^* H X_1 x, \ldots, x^* H X_k x) : x \in \mathbb{C}^n, x^* H x = \pm 1\} \\ &= W^{\pm}_H(HX_1, \ldots HX_k) \end{aligned} \tag{14}$$

If an invertible matrix S is such that $S^* H S = \begin{pmatrix} H_0 & 0 \\ 0 & 0 \end{pmatrix}$, where H_0 is $p \times p$ and invertible, it is easy to see that

$$W^{\pm}_{[\cdot,\cdot]}(X) = \left\{\pm\big(x^* H_0 (S^{-1} X_1 S)_0 x, \ldots, x^* H_0 (S^{-1} X_k S)_0 x\big) : x \in \mathbb{C}^p, x^* H_0 x = \pm 1\right\},$$

where $(\cdot)_0$ indicates projection on the component that corresponds to H_0. Similar formula holds for $W_{[\cdot,\cdot]}(X)$. This allows us to assume in the study of joint numerical ranges with respect to a sesquilinear form that the sesquilinear form is non-degenerate, that is, the corresponding matrix H is invertible. One can easily translate the results in the previous sections in this context. Some of them have interesting forms. For example, assuming the sesquilinear form is nondegenerate, Theorem 4.2 gives:

THEOREM 5.1. *Suppose that the sesquilinear form is nondegenerate, that is, $[x_0, y] = 0$ for all $y \in \mathbb{C}^n$ implies that $x_0 = 0$. If $q = \big([X_1 x, x], \ldots, [X_k x, x]\big)$, $x \in \mathbb{C}^n$, $[x, x] = 1$ is an angle point of $W^{+}_{[\cdot,\cdot]}(X_1, \ldots, X_k)$, then*

$$X_j x = (x^* H X_j x) x = [X_j x, x] x, \quad j = 1, 2, \ldots, k.$$

Here H is the Hermitian matrix that determines $[\cdot,\cdot]$.

The proof is obvious from Theorem 4.2 and equation (14).

Many statements and results of previous sections can be extended verbatim to the infinite dimensional case. Thus, let $\mathcal{G}$ be a Hilbert space with the inner product $\langle\cdot,\cdot\rangle$, H a (bounded) selfadjoint operator on $\mathcal{G}$, and $A = (A_1, \ldots, A_k)$ a k-tuple of selfadjoint operators on $\mathcal{G}$. Define

$$W^{\pm}_H(A) = \left\{\pm\big(\langle x, A_1 x\rangle, \ldots, \langle x, A_k x\rangle\big) : x \in \mathcal{G}, \langle x, Hx\rangle = \pm 1\right\}.$$

When using $W^{\pm}_H(A)$ it will be implicitly assumed that the set of vectors $x \in \mathcal{G}$ such that $\langle x, Hx\rangle = \pm 1$ is not empty.

Statements (A) through (E), and (H) remain valid. (The connectedness of the set $Q := \{x \in \mathcal{G} : \langle x, Hx\rangle = 1\}$ is easily seen upon restricting H to the two dimensional subspace generated by given $x \in Q$ and $y \in Q$).

Theorem 3.2, parts (a) and (b), and Theorems 5.1 and 4.2 are valid in the infinite dimensional case as well.

References

[1] Y.H. Au-Yeung and N.K. Tsing, 'An extension of the Hausdorff-Toeplitz theorem on the numerical range', *Proc. Amer. Math. Soc.* **89** (1983), 215–218.

[2] P. Binding and C.K. Li, 'Joint numerical ranges of Hermitian matrices and simultaneous diagonalization', *Linear Algebra Appl.* **151** (1991), 157–168.

[3] W.F. Donoghue, Jr., 'On the numerical range of a bounded operator', *Michigan Math. J.* **4** (1957), 261–263.

[4] M.K.H. Fan and A.L. Tits, 'm-form numerical range and the computation of the structured singular value', *IEEE Trans. Automat. Control* **33** (1988), 284–289.

[5] C.K. Li, 'A simple proof of the elliptical range theorem', *Proc. Amer. Math. Soc.* **124** (1996), 1985–1986.

[6] C.K. Li, 'The c-spectral, c-radial and c-convex matrices', *Linear and Multilinear Algebra* **20** (1986), 5–15.

[7] C.K. Li and Y.T. Poon, 'Convexity of the joint numerical range', *SIAM J. Matrix Anal. Appl.* **21** (2000), 668–678.

[8] C.K. Li and L. Rodman, 'Remarks on numerical ranges of operators in spaces with an indefinite metric', *Proc. Amer. Math. Soc.* **126** (1998), 973–982.

[9] C.K. Li and T.Y. Tam, 'Numerical ranges arising from simple Lie algebras', *Canad. J. Math.* **52** (2000), 141–171.

[10] C.K. Li, N.K. Tsing and F. Uhlig, 'Numerical range of an operator on an indefinite inner product space', *Electron. J. Linear Algebra* **1** (1996), 1–17.

[11] M. Marcus and I. Filippenko, 'Nondifferentiable boundary points of higher numerical range', *Linear Algebra Appl.* **21** (1978), 217–232.

[12] Y.T. Poon, 'On the convex hull of the multiform numerical range', *Linear and Multilinear Algebra* **37** (1994), 221–223.

[13] Y.T. Poon, 'Generalized numerical ranges, joint positive definiteness and multiple eigenvalues', *Proc. Amer. Math. Soc.* **125** (1997), 1625–1634.

[14] J.-C. Wang, M.K.H. Fan and A.L. Tits, 'Structured singular value and geometry of the m-form numerical range', in *Linear circuits, systems and signal processing: theory and application (Phoenix, AZ, 1987),* (North-Holland, Amsterdam, 1988), pp. 609–615.

Department of Mathematics
College of William and Mary
Williamsburg, VA 23187-8795
United States of America
e-mail: ckli@wm.edu
lxrodm@math.wm.edu

https://doi.org/10.1017/S0004972700020724 Published online by Cambridge University Press