Available online at www.sciencedirect.com
SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 399 (2005) 17–34
www.elsevier.com/locate/laa

On the geometry of numerical ranges in spaces with an indefinite inner product

N. Bebiano [a], R. Lemos [b], J. da Providência [c], G. Soares [d,*]

[a] *Mathematics Department, University of Coimbra, P 3001-454 Coimbra, Portugal*
[b] *Mathematics Department, University of Aveiro, P 3810-193 Aveiro, Portugal*
[c] *Physics Department, University of Coimbra, P 3004-516 Coimbra, Portugal*
[d] *Mathematics Department, University of Trás-os-Montes e Alto Douro, P 5000-911 Vila Real, Portugal*

Received 24 December 2003; accepted 22 April 2004
Available online 2 July 2004
Submitted by F. Zhang

Abstract

Geometric properties of the numerical ranges of operators on an indefinite inner product space are investigated. In particular, classes of matrices are presented such that the boundary generating curves of the J-numerical range are hyperbolical. The curvature of the J-numerical range at a boundary point is studied, generalizing results of Fiedler on the classical numerical range.

Keywords: Indefinite inner product; Numerical range; Generalized Levinger curve

1. Introduction

Throughout the paper, $M_{n\times m}$ denotes the set of $n \times m$ complex matrices, simply M_n, if $n = m$, denoting H_n the set of $n \times n$ Hermitian matrices. For $H \in H_n$ and $A \in M_n$, consider the subsets of the complex plane

$$V_H(A) = \left\{ \frac{x^*Ax}{x^*Hx} : x \in \mathbb{C}^n,\ x^*Hx \neq 0 \right\},$$

* Corresponding author.
E-mail addresses: bebiano@ci.uc.pt (N. Bebiano), rute@mat.ua.pt (R. Lemos), providencia@teor.fis.uc.pt (J. da Providência), gsoares@utad.pt (G. Soares).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.04.021

and the H-*numerical range* of A denoted and defined by

$$W_H(A) = V_H(HA).$$

If H is the identity matrix $I_n \in M_n$, then $V_H(A)$ and $W_H(A)$ reduce to the *classical numerical range*, usually denoted by $W(A)$. If H is a non-singular indefinite Hermitian matrix, the sets $W_H(A)$ and $V_H(A)$ can be understood as natural generalizations of the numerical range with respect to the Krein structure defined by the indefinite inner product $\langle x, y\rangle_H = y^*Hx,\ x, y \in \mathbb{C}^n$ [12].

For convenience, we consider the related sets

$$V_H^{\pm}(A) = \left\{x^*Ax :\ x \in \mathbb{C}^n,\ x^*Hx = \pm 1\right\} \quad \text{and} \quad W_H^{\pm}(A) = V_H^{\pm}(HA).$$

Evidently, we have

$$W_{-H}^{+}(A) = -W_H^{-}(A), \quad W_H(A) = W_H^{+}(A) \cup W_{-H}^{+}(A).$$

If $H = I_n$, then $V_H(A) = V_H^{+}(A) = W_H(A) = W_H^{+}(A) = W(A)$ and $V_H^{-}(A) = W_H^{-}(A) = \emptyset$.

For any $A \in M_n$, $W(A)$ contains $\sigma(A)$, the spectrum of A. For the H-numerical range, we have the following inclusion property: $\sigma_H(A) \subset W_H(A)$, $\sigma_H(A)$ denoting the set of the eigenvalues of A that have H-*anisotropic* eigenvectors, that is, vectors x for which $x^*Hx \neq 0$. Compactness and convexity are basic properties of the classical numerical range. Actually, $W(A)$ is always a compact and convex set for $A \in M_n$ [9]. In contrast with the classical case, the set $W_H(A)$ may not be closed and is either unbounded or a singleton [12,13]. (For $\lambda \in \mathbb{C}$, $W_H(A) = \{\lambda\}$ if and only if $HA = \lambda H$.) On the other hand, $W_H(A)$ may not be convex. Nevertheless, $W_H(A)$ is *pseudo-convex* [12]; that is, for any pair of distinct points $x, y \in W_H(A)$, either $W_H(A)$ contains the closed line segment joining x and y, or $W_H(A)$ contains the line defined by x and y, except the open line segment joining x and y.

Let $A = \operatorname{Re} A + \mathrm{i} \operatorname{Im} A$, where $\operatorname{Re} A = (A + A^*)/2$ and $\operatorname{Im} A = (A - A^*)/2\mathrm{i}$, be the *cartesian decomposition* of $A \in M_n$. For $H \in H_n$, the H-*shell* of A is the subset of $\mathbb{R}^3$ denoted and defined by

$$\mathscr{S}_H(A) = \left\{\left(\frac{x^* \operatorname{Re}(HA)x}{x^*Hx}, \frac{x^* \operatorname{Im}(HA)x}{x^*Hx}, \frac{x^*A^*HAx}{x^*Hx}\right) : x \in \mathbb{C}^n, x^*Hx \neq 0\right\}.$$

This concept motivated some investigation [2]. If $H = I_n$, then $\mathscr{S}_H(A)$ reduces to the *Davis–Wiedlant shell* $\mathscr{S}(A)$ studied by Davis [4,5]. The sets

$$\mathscr{S}_H^{\pm}(A) = \left\{\left(x^* \operatorname{Re}(HA)x, x^* \operatorname{Im}(HA)x, x^*A^*HAx\right) : x \in \mathbb{C}^n, x^*Hx = \pm 1\right\}$$

are closely related to the H-shell of A. Obviously,

$$\mathscr{S}_{-H}^{+}(A) = -\mathscr{S}_H^{-}(A), \quad \mathscr{S}_H(A) = \mathscr{S}_H^{+}(A) \cup \mathscr{S}_{-H}^{+}(A).$$

Since $W_H(A)$ ($W_H^{\pm}(A)$) is the image of $\mathscr{S}_H(A)$ ($\mathscr{S}_H^{\pm}(A)$) under the projection $(x, y, r) \mapsto x + \mathrm{i}y$, more information on the matrix A can be obtained from $\mathscr{S}_H(A)$ ($\mathscr{S}_H^{\pm}(A)$). Moreover, there is an interesting interplay between the algebraic properties of the matrix A and the geometrical properties of $\mathscr{S}_H(A)$.

In this paper, we assume H non-singular. Without loss of generality, we can consider in the definitions of $W_H(A)$ and $\mathscr{S}_H(A)$, instead of H, the matrix $J = P(I_r \oplus -I_{n-r})P^{\mathrm{T}}$, where P is a permutation matrix. In fact, using Sylvester's law of inertia [8], we can easily check that $W_H(A) = W_J(R^{-1}AR)$ and $\mathscr{S}_H(A) = \mathscr{S}_J(R^{-1}AR)$, R being a non-singular matrix such that $R^*HR = J$ is the inertia matrix of H.

We recall that a matrix $U \in M_n$ is *pseudo-unitary* of signature $(r, n-r)$, $0 \leqslant r \leqslant n$, if and only if $U^*JU = J$. This matrix is also called J*-unitary*. The pseudo-unitary matrices of signature $(r, n-r)$ form a group denoted by $U_{r,n-r}$. For any $U \in U_{r,n-r}$, $V_J(A) = V_J(U^*AU)$.

For simplicity of notation, we consider throughout

$$H_1 = \mathrm{Re}(JA) \quad \text{and} \quad H_2 = \mathrm{Im}(JA).$$

Let $\kappa \in W_J(A)$ be a boundary point of $W_J(A)$. A line containing κ and defining two half planes, such that one of them does not contain $W_J^+(A)$ $(-W_J^-(A))$ but it contains $-W_J^-(A)(W_J^+(A))$ will be called a *supporting line* of $W_J(A)$. Supporting lines may not exist, and they may not be unique. As proved in [1, Theorem 2.2], if $ux + vy + w = 0$ is the equation of a supporting line L of $W_J(A)$, then

$$\det(uH_1 + vH_2 + wJ) = 0, \tag{1}$$

and $-w$ is the maximum or the minimum eigenvalue of the matrix $uJH_1 + vJH_2$, according to $ux + vy + w \leqslant 0$, for all points in $W_J^+(A)$ $(W_J^-(A))$, or $ux + vy + w \geqslant 0$, for these points. Conversely, the intersection $L \cap W_J^{\pm}(A)$ consists of all points $\langle Az, z\rangle_J$ for which z is an eigenvector of $uJH_1 + vJH_2$ corresponding to $-w$ such that $z^*Jz = \pm 1$. Since $\det(uH_1 + vH_2 + wJ)$ is a homogeneous polynomial of degree n, (1) can be considered the dual (line) equation of an algebraic curve (for details on algebraic curves, see e.g. [16,18]). Its real part, throughout denoted by $C_J(A)$ (simply $C(A)$, if $J = I_n$), is called the *boundary generating curve* of $W_J(A)$. Kippenhahn [10] proved that the curve $C(A)$ generates $W(A)$ as its convex hull. The algebraic curve $C_J(A)$ generates $W_J(A)$ as its *pseudo-convex hull*, which is obtained in the following way: for any two points x_1, x_2 in the boundary generating curve, let $z_i \in \mathbb{C}^n$ be such that

$$\frac{z_i^*JAz_i}{z_i^*Jz_i} = x_i, \quad i = 1, 2;$$

take the closed line segment defined by x_1, x_2, if $(z_1^*Jz_1)(z_2^*Jz_2) > 0$; and take the two rays $\{\alpha x_1 + (1-\alpha)x_2 : \alpha \leqslant 0 \text{ or } \alpha \geqslant 1\}$, if $(z_1^*Jz_1)(z_2^*Jz_2) < 0$. The boundary generating curve $C_J(A)$ has class n, that is, through a general point in the plane there are n lines (may be complex) tangent to $C_J(A)$, and it has the eigenvalues of A as its real foci.

This paper is organized as follows. In Section 2, $W_J(A)$ is described for A a J-normal matrix and A a J-unitary matrix with simple eigenvalues, answering affirmatively questions posed by Li et al. in [12]. It is also proved that if A is a J-normal matrix and JH_1 has simple eigenvalues, then $\mathscr{S}_J^+(A)$ is a polyhedron, proving, in a particular case, a conjecture formulated in [2]. In Section 3, classes of matrices are presented such that the boundary generating curves of $W_J(A)$ are hyperbolical. In Section 4, a formula for the curvature of $W_J(A)$ at a boundary point is obtained, and the connection between the curvature of the boundary of $W_J(A)$ at $\lambda_{\max}(H_1)$ and of the local J-generalized Levinger curve at 1/2 is investigated. These results generalize results of Fiedler [6,7] on the classical numerical range.

2. Results for *J*-Hermitian, *J*-normal and *J*-unitary matrices

It is known that $W_J(A) \subseteq \mathbb{R}$ if and only if A is *J-Hermitian*, that is, $JA \in H_n$. Now, we prove the following.

Proposition 2.1. *If A is a J-Hermitian matrix such that the eigenvalues of A are not all real, then $W_J(A)$ is the whole real line.*

Proof. Since $JA \in H_n$, then it is obvious that $W_J(A) \subseteq \mathbb{R}$. Suppose that the eigenvalues $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_r, \bar{\lambda}_r, \lambda_{2r+1}, \ldots, \lambda_n$ of A are all distinct, and suppose that only $\lambda_{2r+1}, \ldots, \lambda_n$ are real. It is clear that the corresponding eigenvectors $u_1, v_1, \ldots, u_r, v_r, u_{2r+1}, \ldots, u_n$ are linearly independent. Clearly, $u_i^* J u_i = v_i^* J v_i = 0$, $i = 1, \ldots, r$. Let $D \in M_n$ be the matrix whose columns are the vectors of this eigenbasis. The matrix $D^* J D$ is a block diagonal matrix with 2×2 and 1×1 blocks, corresponding to complex and to real eigenvalues, respectively. Moreover, the 2×2 blocks have zero diagonal entries and the off diagonal entries non-zero. Hence, $\det(D^* J D) \neq 0$, which implies $u_i^* J v_i = \gamma_i \neq 0$, and so $u_i^* J A v_i = \bar{\lambda}_i \gamma_i$, $i = 1, \ldots, r$.

Consider the subset $R_{u_i,v_i}(A)$ of $W_J(A)$ defined by

$$R_{u_i,v_i}(A) = \left\{ \frac{(u_i + av_i)^* J A (u_i + av_i)}{(u_i + av_i)^* J (u_i + av_i)} : a \in \mathbb{C},\ (u_i + av_i)^* J (u_i + av_i) \neq 0 \right\}.$$

Since $\lambda_i \neq \bar{\lambda}_i$, by straightforward computation, we can prove that

$$\frac{\operatorname{Re}(a\bar{\lambda}_i\gamma_i)}{\operatorname{Re}(a\gamma_i)} = |\lambda_i|(\cos\phi_i - \sin\phi_i \tan\xi_i)$$

are elements of $R_{u_i,v_i}(A)$, for $\phi_i = -\arg\lambda_i$ and $\xi_i = \arg a + \arg\gamma_i$. Therefore, $R_{u_i,v_i}(A)$ is the whole real line, as well as $W_J(A)$. By a perturbation, we can take the eigenvalues of A simple, and the result follows by a continuity argument, as it can be easily seen, for instance, by a contradiction argument. □

Remark 2.1. The set $W_J(A)$ may be the whole real line even if the eigenvalues of A are all real. Indeed, consider $A = \operatorname{diag}(1, 2, 3)$ and $J = \operatorname{diag}(1, -1, 1)$. It can be easily shown that $W_J(A)$ is the real line.

A matrix $A \in M_n$ is called *essentially J-Hermitian* if $\mu A + \upsilon I_n$ is J-Hermitian, for some $0 \neq \mu \in \mathbb{C}$ and $\upsilon \in \mathbb{C}$.

Corollary 2.1. *If A is essentially J-Hermitian and the eigenvalues of the J-Hermitian matrix $B = \mu A + \upsilon I_n$, $0 \neq \mu \in \mathbb{C}$, $\upsilon \in \mathbb{C}$, are not all real, then $W_J(A)$ is the whole straight line passing through $-\upsilon/\mu$ and with direction* $-\arg \mu$.

Proof. By the hypothesis and using the previous theorem, we can conclude that $W_J(B) = \mathbb{R}$. Since $W_J(\mu A + \upsilon I_n) = \mu W_J(A) + \upsilon$, the result easily follows. □

Let $A \in M_n$. In [14], it was proved that if a boundary point w in $W_J(A)$ is a *corner* of $W_J(A)$, that is, it is on more than one supporting line of $W_J(A)$, then w is an eigenvalue of A and there exists an eigenvector x associated to w, such that $Ax = wx$, $A^{[*]}x = \bar{w}x$, $x^*Jx = \pm 1$.

Clearly, $\operatorname{Re} W_J(A) = W_J(JH_1)$ and $\operatorname{Im} W_J(A) = W_J(JH_2)$.

Lemma 2.1. *Let $A \in M_n$ and $x \in \mathbb{C}^n$ such that $x^*Jx = \pm 1$. If $W_J^{\pm}(JH_1)$ is a closed half line then the following conditions are equivalent*:

(a) $\operatorname{Re}(x^*JAx)$ *is the extreme point of* $\pm\operatorname{Re} W_J^{\pm}(A)$;
(b) x^*H_1x *is the extreme point in* $\pm W_J^{\pm}(JH_1)$;
(c) $JH_1x = \lambda_M x$, *where λ_M is the maximum or the minimum eigenvalue in* $\sigma_J(JH_1)$.

Proof. The equivalence of (a) and (b) is obvious. Suppose that w is the extreme point of the closed half-line $W_J^{\pm}(JH_1)$. Then w is a corner of $W_J^{\pm}(JH_1)$, and so it is the maximum or the minimum eigenvalue in $\sigma_J(JH_1)$, that is, $w = \lambda_M$.

(c) ⇒ (b) If x is an eigenvector of JH_1 associated with λ_M, then $\pm(x^*H_1x) = \lambda_M$ is the extremum of $W_J^{\pm}(JH_1)$.

(b) ⇒ (c) (By contradiction.) Suppose that x is not an eigenvector of JH_1 associated with the eigenvalue λ_M, that is, $H_1x \neq \lambda_M Jx$. Then $\pm(x^*H_1x) \neq \lambda_M$, and λ_M would not be the extreme point of $W_J^{\pm}(JH_1)$, a contradiction. □

In [1], an algorithm to describe the boundary generating curve of $V_J(A)$ was presented. In this spirit, we prove the following proposition.

Proposition 2.2. *Let $A \in M_n$ and let θ belong to a subset Ω of $[0, 2\pi)$ such that the n eigenvalues $\lambda_1(\theta), \ldots, \lambda_n(\theta)$ of $J\operatorname{Re}(e^{-i\theta}JA)$ are real and simple. Let $u_k(\theta)$ be an eigenvector of $J\operatorname{Re}(e^{-i\theta}JA)$ associated with $\lambda_k(\theta)$, $k = 1, \ldots, n$. Then the boundary generating curve $C_J(A)$ of $W_J(A)$ is given by*

$$\left\{ z_k(\theta) = \frac{u_k^*(\theta)JAu_k(\theta)}{u_k^*(\theta)Ju_k(\theta)} : \theta \in \Omega, k = 1, \ldots, n \right\}$$

and $W_J(A)$ is the pseudo-convex hull of $C_J(A)$.

Proof. Let $p_A(u, v, w) = \det(uH_1 + vH_2 + wJ)$. The curve $C_J(A)$ is the dual curve to the algebraic curve determined by $p_A(u, v, w) = 0$ in the complex projective plane $\mathbb{CP}^2$, that is, consists of all points $[x, y, z]$ in $\mathbb{CP}^2$ such that $ux + vy + wz = 0$ is a tangent line to $p_A(u, v, w) = 0$. As usual, we identify the point (x, y) in $\mathbb{C}^2$ with $[x, y, 1]$ in $\mathbb{CP}^2$, and identify any point $[x, y, z]$ in $\mathbb{CP}^2$ such that $z \neq 0$ with $(x/z, y/z)$ in $\mathbb{C}^2$. Hence, in particular, the plane $\mathbb{R}^2$ (identified with $\mathbb{C}$) sits in $\mathbb{CP}^2$ by the way of the identification of the point (a, b) of $\mathbb{R}^2$ with $[a, b, 1]$ in $\mathbb{CP}^2$. In [1, Theorem 2.2], it was proved that if $ux + vy + w = 0$ is the equation of a supporting line of $W_J(A)$, then $p_A(u, v, w) = 0$. Since the dual curve of $C_J(A)$ is the original curve $p_A(u, v, w) = 0$, we infer, in particular, that every supporting line of $W_J(A)$ is tangent to $C_J(A)$. Consider the real direction $(\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$, for which (1) provides n real eigenvalues for $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta} JA) = \cos\theta JH_1 + \sin\theta JH_2$, namely, $\lambda_1(\theta), \dots, \lambda_n(\theta)$. Let $u_k(\theta)$ be an eigenvector of $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta} JA)$ associated with $\lambda_k(\theta)$, $k = 1, \dots, n$. If the eigenvalues corresponding to eigenvectors with positive J-norm are all smaller (or larger) than all the eigenvalues corresponding to eigenvectors with negative J-norm, then supporting lines of $W_J(A)$ exist. Let $-w(\theta)$ be the maximum (or the minimum) eigenvalue of $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta} JA)$, so that $\cos\theta x + \sin\theta y = w(\theta)$ is a supporting line L of $W_J(A)$. Since $\lambda_1(\theta), \dots, \lambda_n(\theta)$ are distinct, then $u_1(\theta), \dots, u_n(\theta)$ form an eigenbasis and $u_k^*(\theta) J u_k(\theta) \neq 0$, $k = 1, \dots, n$. Applying Lemma 2.1 to the matrix $\mathrm{e}^{-\mathrm{i}\theta} JA$, we conclude that the intersection $L \cap W_J^{\pm}(A)$ consists of all points $\langle Az, z\rangle_J$ for which z is an eigenvector of $\cos\theta JH_1 + \sin\theta JH_2$ corresponding to $-w(\theta)$ such that $z^* J z = \pm 1$. These points belong to $C_J(A)$ which is given by

$$\frac{u_k(\theta)^* JA u_k(\theta)}{u_k(\theta)^* J u_k(\theta)} = z_k(\theta), \quad \theta \in \Omega,\ k = 1, \dots, n.$$

By the pseudo-convexity of $W_J(A)$, we conclude the result.

If there does not exist supporting lines of $W_J(A)$, then in any real direction in $\mathbb{R}^2$, say $(\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$, there are n real lines tangent to $C_J(A)$ with this direction, namely the n real eigenvalues $\lambda_1(\theta), \dots, \lambda_n(\theta)$ of $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta} JA)$ and the result easily follows. □

Remark 2.2. The hypothesis on the simplicity of the eigenvalues in Proposition 2.2 ensures that all eigenvectors are anisotropic. If $\lambda_k(\theta)$, $k = 1, \dots, n$, are not all simple, but have anisotropic associated eigenvectors $u_k(\theta)$, $k = 1, \dots, n$, the result is still valid.

The J-*adjoint* of $A \in M_n$, denoted by $A^{[*]}$, is defined by $\langle Ax, y\rangle_J = \langle x, A^{[*]} y\rangle_J$, $x, y \in \mathbb{C}^n$, and it can be expressed explicitly in terms of A and J by $A^{[*]} = JA^*J$.

A matrix $A \in M_n$ is said to be J-*normal* if and only if $AA^{[*]} = A^{[*]}A$. It is well-known that if $A \in M_n$ is normal, then $W(A)$ is the convex hull of the spectrum of A. An analogous result for $W_J(A)$ and a J-normal matrix A, with simple eigenvalues, is obtained in the next proposition. This answers, in a particular case, a question proposed in [12].

Proposition 2.3. *Let $A \in M_n$ be a J-normal matrix with simple eigenvalues. If the eigenvalues of $A + A^{[*]}$ are all real, then $W_J(A)$ is the pseudo-convex hull of these eigenvalues.*

Proof. Denote by $\lambda_1, \ldots, \lambda_n$ the distinct eigenvalues of A. For $k = 1, \ldots, n$, let v_k be an eigenvector of A associated with the eigenvalue λ_k. Since by hypothesis A is J-normal, then $Aw_k = \lambda_k w_k$, for $w_k = A^{[*]}v_k$. Because the eigenvalues of A are all distinct, the corresponding eigenvectors are linearly independent. Hence, there exist complex numbers c_k such that $w_k = c_k v_k$. Moreover, $(A + A^{[*]})v_k = (\lambda_k + c_k)v_k$, and since the eigenvalues of $A + A^{[*]}$ are all real, we have $c_k = \bar{\lambda}_k$. Therefore, $A^{[*]}v_k = \bar{\lambda}_k v_k$. Easy calculations show that $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta} JA)v_k = \mathrm{Re}(\mathrm{e}^{-\mathrm{i}\theta}\lambda_k)v_k$, $[0, 2\pi]$. We notice that the eigenvector v_k does not depend on θ and v_k is an eigenvector of A associated with the eigenvalue λ_k, $k = 1, \ldots, n$. By Proposition 2.2, $(v_k^* JAv_k)/(v_k^* Jv_k) = \lambda_k,\ \ k = 1, \ldots, n$, give the boundary generating curve of $W_J(A)$. It is straightforward to show that $W_J(A)$ is the pseudo-convex hull of the eigenvalues of A. $\square$

Remark 2.3. If $J = I_n$, then $A^{[*]} = A^*$ and $A + A^{[*]}$ is Hermitian, and so its eigenvalues are all real. Hence, in the particular case $J = I_n$, Proposition 2.3 yields the well-known property that $W(A)$ is the convex hull of the spectrum of A, when A is a normal matrix (valid even if the eigenvalues of A are not all simple).

Remark 2.4. The hypothesis on the simplicity of the eigenvalues in Proposition 2.3 ensures that all eigenvectors are anisotropic. The existence of multiple eigenvalues may lead to the existence of isotropic eigenvectors, that is, vectors x for which $x^*Jx = 0$. Let $J = -I_1 \oplus I_1$ and consider the J-Hermitian matrix

$$A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

The double eigenvalue of A is 2, the eigenvectors associated with 2 are isotropic and $W_J(A)$ is $\mathbb{R}\backslash\{2\}$. In this example, the concept of pseudo-convex hull is meaningless.

We say that $W_S^{\pm}(A)$ is *polygonal* if it is the intersection of a finite number of closed half-planes or the whole complex plane. The set $W_S(A)$ is said to be *polygonal* if the closures of the both convex components $W_S^{\pm}(A)$ are polygonal. In [17], it was proved that if A is a J-normal matrix, then $W_J(A)$ is polygonal, the converse not being true in general.

It is well-known that $A \in M_n$ is unitary if and only if $W(A)$ is a polygon inscribed in the unit disc $\mathscr{D}$ and $\sigma(A) \subset \partial\mathscr{D}$, the boundary of $\mathscr{D}$, that is, the unit circle. Whether an analogous result is valid for $W_J(U)$, when U is a pseudo-unitary matrix, is a question proposed in [12]. In this vein, we have the following result. We observe that $U \in U_{r,n-r}$ if and only if $U^{[*]}U = I_n$.

Corollary 2.2. *If the eigenvalues of* $U \in U_{r,n-r}$ *are all simple, then* $W_J(U)$ *is the pseudo-convex hull of these eigenvalues.*

Proof. Let v_k be an eigenvector of U associated with the eigenvalue $\lambda_k, k = 1, \ldots, n$. Hence $(Uv_k)^* = \bar{\lambda}_k v_k^*$, and it can be easily seen that

$$v_k^* U^* J v_k = \bar{\lambda}_k v_k^* J v_k. \tag{2}$$

Since $U^{[*]}U = I_n$ and $Uv_k = \lambda_k v_k$, we easily get $\lambda_k U^{[*]} v_k = v_k$, and so

$$v_k^* J v_k = \lambda_k v_k^* U^* J v_k. \tag{3}$$

From (2) and (3), we can conclude that $|\lambda_k| = 1$, $k = 1, \ldots, n$, that is, the spectrum of U is on the unit circle. We may also conclude that $U^{[*]} v_k = \bar{\lambda}_k v_k$, and v_k is an eigenvector of $U + U^{[*]}$ corresponding to the eigenvalue $\lambda_k + \bar{\lambda}_k$, $k = 1, \ldots, n$. Since the eigenvalues of $U + U^{[*]}$ are all real, applying Proposition 2.3 to the J-normal matrix U, the result follows. □

A *polyhedral set* is the intersection of finitely many closed half-planes or the whole $\mathbb{R}^3$. The Davis–Wiedlant shell $\mathscr{S}(A)$ is a *polyhedron*, that is, the convex hull of a finite number of points in $\mathbb{R}^3$, if and only if A is a normal matrix [4]. A polyhedron is clearly a bounded polyhedral set.

In [2], it was conjectured that, for H-normal operators, $\mathscr{S}_H^+(A)$ is convex and its closure is a polyhedral set. The conjecture was proved for indefinite inner product spaces of dimension at most three, and for H invertible with only one positive eigenvalue. In this vein we have the following result.

Proposition 2.4. *If* $A \in M_n$ *is a* J*-normal matrix and the eigenvalues of* JH_1 *are all simple, then* $\mathscr{S}_J^+(A)$ $(\mathscr{S}_J^-(A))$ *is a polyhedron.*

Proof. Let $H_1 = \mathrm{Re}(JA)$ and $H_2 = \mathrm{Im}(JA)$. By the hypothesis, A is J-normal, that is, $AJA^*J = JA^*JA$. This implies that

$$H_1 J H_2 = H_2 J H_1 \tag{4}$$

and so

$$A^{[*]}A = H_1 J H_1 + H_2 J H_2. \tag{5}$$

If $\alpha_1, \ldots, \alpha_n$ are the distinct eigenvalues of JH_1, then $JH_1 v_k = \alpha_k v_k$, where the eigenvectors v_k of JH_1 associated with α_k are linearly independent vectors such that $v_k^* J v_k = 1$, $k = 1, \ldots, n$. We show that the eigenvectors v_k of JH_1 are also eigenvectors of JH_2 and of $A^{[*]}A$. Indeed, from $JH_1 v_k = \alpha_k v_k$ and using (4), we have $JH_1 JH_2 v_k = JH_2 JH_1 v_k = \alpha_k JH_2 v_k$. It follows that $JH_2 v_k$ is an eigenvector of JH_1 associated with α_k. Thus, there exist non-zero real numbers β_k such that

$JH_2v_k = \beta_k v_k$. In fact, since $\beta_k = v_k^* H_2 v_k$, it is a real number. Moreover, from (5), we have $A^{[*]}Av_k = J(H_1JH_1 + H_2JH_2)v_k = (\alpha_k^2 + \beta_k^2)v_k$, that is, v_k is an eigenvector of $A^{[*]}A$ associated with the eigenvalue $\alpha_k^2+\beta_k^2$. Now, let $z=(x^*H_1x, x^*H_2x, x^*A^*JAx)$ be an arbitrary element in $\mathscr{S}_J^+(A)$. Writing $x = \sum_{i=1}^n c_k v_k$, $c_k \in \mathbb{C}$, we get $z = \sum_{k=1}^n |c_k|^2(\alpha_k, \beta_k, \alpha_k^2 + \beta_k^2)$. An analogous result holds for $\mathscr{S}_J^-(A)$. □

Remark 2.5. The present approach does not allow to relax the condition on the simplicity of the eigenvalues of JH_1.

A matrix $A \in M_n$ is called *J-decomposable* if there exists a pseudo-unitary matrix $U \in M_n$ such that $U^*JAU = A_1 \oplus A_2$, where A_1, A_2 are square matrices. Otherwise, A is called *J-indecomposable*. We notice that the next result was previously obtained by Li and Rodman in [14]. Here we present an alternative proof.

Theorem 2.1. *Let $A \in M_n$ be J-indecomposable. If $z \in \sigma_J(A)$, then z is in the interior of $W_J(A)$.*

Proof. (By contradiction.) Suppose that $z = a + \mathrm{i}b \in \sigma_J(A)$ is in the boundary of $W_J(A)$. Then there exists a supporting line passing through z and there exists $\phi_0 \in \mathbb{R}$ such that $a\cos\phi_0 + b\sin\phi_0 = \lambda_M(\mathrm{e}^{-\mathrm{i}\phi_0}JA)$, where $\lambda_M(\mathrm{e}^{-\mathrm{i}\phi_0}JA)$ is the maximum (or minimum) eigenvalue of $J\,\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\phi_0}JA)$. Let x be an eigenvector of A associated with the eigenvalue z, such that $x^*Jx = \pm 1$. From $\mathrm{e}^{-\mathrm{i}\phi_0}JAx = \mathrm{e}^{-\mathrm{i}\phi_0}(a+ib)Jx$, we easily get

$$\frac{x^*\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\phi_0}JA)x}{x^*Jx} = a\cos\phi_0 + b\sin\phi_0. \tag{6}$$

Clearly $\mathrm{Re}(\mathrm{e}^{-\mathrm{i}\phi_0}JA) = H_1\cos\phi_0 + H_2\sin\phi_0$, and so it follows from (6) that

$$(H_1\cos\phi_0 + H_2\sin\phi_0)x = (a\cos\phi_0 + b\sin\phi_0)Jx.$$

This in conjunction with the condition $Ax = zx$ gives

$$(H_1\sin\phi_0 - H_2\cos\phi_0)x = (a\sin\phi_0 - b\cos\phi_0)Jx.$$

Hence, $H_1x = aJx$ and $H_2x = bJx$. Now, it can be easily shown that $x^*JA = zx^*J$. Taking the pseudo-unitary matrix U whose first column is x, we have $U^*JAU = [z] \oplus A_1$, a contradiction. □

3. Hyperbolical boundary generating curves

The classical numerical range of a 2×2 matrix A is an elliptical disc, possibly degenerate, with the eigenvalues λ_1 and λ_2 of A as foci and minor axis of length $\sqrt{\mathrm{Tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$. For $A \in M_2$ and $J = I_1 \oplus -I_1$, the Hyperbolical

Range Theorem [1] asserts that $W_J(A)$ is bounded by a non-degenerate hyperbola, with the eigenvalues λ_1 and λ_2 of A as foci and non-transverse axis of length $\sqrt{|\lambda_1|^2 + |\lambda_2|^2 - \mathrm{Tr}(A^{[*]}A)}$, if $2\,\mathrm{Re}(\bar{\lambda}_1\lambda_2) < \mathrm{Tr}(A^{[*]}A) < |\lambda_1|^2 + |\lambda_2|^2$. For the degenerate cases, $W_J(A)$ is a singleton, a line, a subset of a line, the whole complex plane, or the complex plane except a line. In this section, we present a hyperbolical range theorem for a wide class of matrices.

Consider the block matrix

$$A = \begin{bmatrix} aI_r & X \\ Y & bI_{n-r} \end{bmatrix}, \quad a, b \in \mathbb{C}, \tag{7}$$

where X, Y^* are $r \times (n-r)$ complex matrices, such that XY and YX are normal matrices, and let $p = \min(r, n-r)$. In [3], it was proved that the numerical range of A is the convex hull of at most p ellipses, all of which centered at $(a+b)/2$. We prove that, for A of type (7) and $J = I_r \oplus -I_{n-r}$, $0 < r < n$, the set $W_J(A)$ is the pseudo-convex hull of at most p hyperbolas, all centered at $(a+b)/2$.

Lemma 3.1. *Let $J = I_r \oplus -I_{n-r}$, $0 < r < n$, and $p = \min(r, n-r)$. Let A be a block matrix of type* (7), *such that XY and YX are normal matrices, $\sigma_1, \ldots, \sigma_p$ and $\delta_1, \ldots, \delta_p$ being the singular values of X and Y, respectively. Then, there exists a pseudo-unitary matrix U such that*

$$U^*JAU = \begin{bmatrix} aI_r & \Sigma \\ \Delta & -bI_{n-r} \end{bmatrix}, \tag{8}$$

where $\Sigma, \Delta^* \in M_{r,n-r}$, the diagonal entries of Σ and Δ are $\sigma_1 e^{i\phi_1}, \ldots, \sigma_p e^{i\phi_p}$ and $\delta_1 e^{i\phi_1}, \ldots, \delta_p e^{i\phi_p}$, respectively, for some $\phi_1, \ldots, \phi_p \in \mathbb{R}$, all the other entries being zero.

Proof. By the singular value decomposition, there exist unitary matrices $U_1 \in M_r$, $U_2 \in M_{n-r}$ such that the diagonal entries of $U_1^* X U_2$ are the singular values $\sigma_1, \ldots, \sigma_p$ of X, all the other entries being zero. If XY and YX are normal matrices, then the diagonal entries of $U_1^* Y^* U_2$ are $-\delta_1 e^{2i\phi_1}, \ldots, -\delta_p e^{2i\phi_p}$, with $\delta_1, \ldots, \delta_p$ the singular values of Y, and $\phi_1, \ldots, \phi_p \in \mathbb{R}$, all the other entries being zero [8, p. 426]. Moreover, let $D_1 = \mathrm{diag}(e^{i\eta_1}, \ldots, e^{i\eta_p}) \oplus I_{r-p}$, $D_2 = \mathrm{diag}(e^{i\mu_1}, \ldots, e^{i\mu_p}) \oplus I_{n-r-p}$ such that $\mu_l - \eta_l = \phi_l$, $l = 1, \ldots, p$. It can be easily seen that the block matrix $U = (U_1 D_1) \oplus (U_2 D_2)$ satisfies $U^* J U = J$ and has the asserted property. □

Before the main result of this section, we recall that if $T \in M_n$ is a block matrix

$$T = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix},$$

such that X, Y, Z, W are square matrices and all commute, then $\det(T) = \det(XW - ZY)$ (see, e.g. [11]).

Theorem 3.1. *Let A and J be matrices under the conditions of Lemma* 3.1, *and let*

$$2\beta_{l\pm} = a + b \pm \sqrt{(a-b)^2 - 4\sigma_l\delta_l e^{2i\phi_l}}, \quad l = 1, \ldots, p. \tag{9}$$

If

$$2\,\mathrm{Re}(\overline{\beta_{l+}}\beta_{l-}) < |a|^2 + |b|^2 - \sigma_l^2 - \delta_l^2 < |\beta_{l+}|^2 + |\beta_{l-}|^2, \quad l = 1, \ldots, p, \tag{10}$$

then the boundary generating curve of $W_J(A)$ is given by the p hyperbolas (*some possibly coincident*), *all centered at $(a+b)/2$, with foci at $\beta_{l\pm}$, with non-transverse axis of length $\sqrt{|\beta_{l+}|^2 + |\beta_{l-}|^2 - |a|^2 - |b|^2 + \sigma_l^2 + \delta_l^2}$, $l = 1, \ldots, p$, and possibly a point, a if $n < 2r$, and b if $n > 2r$. The set $W_J(A)$ is the pseudo-convex hull of these p hyperbolas.*

Proof. Since $W_J(A) = V_J(U^*JAU)$, without loss of generality, we may concentrate on the study of $V_J(B)$, where $B = U^*JAU$ is given by (8). The characteristic polynomial of the matrix $J\,\mathrm{Re}(e^{i\theta}B)$ is

$$P_n(t) = (t - c_{a\theta})^{r-p}(t - c_{b\theta})^{n-r-p} \det\big((t - c_{a\theta})(t - c_{b\theta})I_p + D_\theta\big),$$

where $c_{a\theta} = |a|\cos(\theta + \arg a)$, $c_{b\theta} = |b|\cos(\theta + \arg b)$ and D_θ is the diagonal matrix whose lth diagonal entry is

$$d_l(\theta) = \frac{\sigma_l\delta_l}{2}\cos(2\theta + 2\phi_l) + \frac{\sigma_l^2}{4} + \frac{\delta_l^2}{4}, \quad l = 1, \ldots, p.$$

The roots of $P_n(t)$ satisfy

$$(t - c_{a\theta})^{r-p}(t - c_{b\theta})^{n-r-p}\prod_{l=1}^{p}\big(t^2 - (c_{a\theta} + c_{b\theta})t + c_{a\theta}c_{b\theta} + d_l(\theta)\big) = 0,$$

and so the eigenvalues of $J\,\mathrm{Re}(e^{i\theta}B)$ are

$$\lambda_{l\pm}(\theta) = \tfrac{1}{2}(c_{a\theta} + c_{b\theta}) \pm \tfrac{1}{2}\sqrt{(c_{a\theta} - c_{b\theta})^2 - 4d_l(\theta)}, \quad l = 1, \ldots, p, \tag{11}$$

and $c_{a\theta}$ ($c_{b\theta}$), if $n < 2r$ ($n > 2r$). Since (10) holds, there exist directions for which all the characteristic roots of $J\,\mathrm{Re}(e^{i\theta}B)$ are non-zero real numbers. We observe that σ_l, δ_l are not simultaneously zero, otherwise the second inequality in (10) would not hold. The eigenvectors of $J\,\mathrm{Re}(e^{i\theta}B)$ associated with the eigenvalue $\lambda_{l\pm}(\theta)$ are the vectors $u_{l\pm}(\theta)$ whose lth entry is

$$\frac{2(c_{b\theta} + \lambda_{l\pm}(\theta))}{\sigma_l e^{-i(\theta+\phi_l)} + \delta_l e^{i(\theta+\phi_l)}},$$

the $(r+l)$th entry is 1 and all the others are 0, $l = 1, \ldots, p$. Hence, it can be easily seen that the boundary generating curve of $V_J(B)$ coincides with the boundary generating curve of $V_{J_2}(B_l)$, $J_2 = I_1 \oplus -I_1$ and

$$B_l = \begin{bmatrix} a & \sigma_l e^{i\phi_l} \\ \delta_l e^{i\phi_l} & -b \end{bmatrix}, \quad l = 1, \ldots, p, \tag{12}$$

and possibly a point. The eigenvalues of $J_2 B_l$ are $\beta_{l\pm}$ defined in (9) and $\mathrm{Tr}(B_l^{[*]} B_l) = |a|^2 + |b|^2 - \sigma_l^2 - \delta_l^2$. By the Hyperbolical Range Theorem for 2×2 matrices, we conclude that $V_{J_2}(B_l)$ is bounded by a hyperbola, with foci at $\beta_{l\pm}$ and non-transverse axis of length $\sqrt{|\beta_{l+}|^2 + |\beta_{l-}|^2 - |a|^2 - |b|^2 + \sigma_l^2 + \delta_l^2}$. If $n = 2r$, the boundary generating curve of $V_J(B)$ is given precisely by these p hyperbolas (some of them possibly coincident). If $n < 2r(n > 2r)$, the eigenvectors associated with the eigenvalue $c_{a\theta}(c_{b\theta})$ are the vectors e_j (e_m) of the standard basis of $\mathbb{C}^n$, and $e_j^* B e_j = a e_j^* J e_j$, $j = n - r + 1, \ldots, r (e_m^* B e_m = b e_m^* J e_m$, $m = 2r + 1, \ldots, n)$. By Proposition 2.2, the boundary generating curve of $V_J(B)$ is given by the previous p hyperbolas and a point, a if $n < 2r$, and b if $n > 2r$.

We observe that $a \in V_{J_2}^+(B_l)$ and $b \in -V_{J_2}^-(B_l)$, $l = 1, \ldots, p$. Therefore, even in the case of rectangular matrices X and Y, the set $V_J(B)$ is the pseudo-convex hull of these p hyperbolas. $\square$

Example 1. Let $J = I_2 \oplus -I_2$ and A be a block matrix of type (7), with $n = 4$, $r = 2$, $a = 1$, $b = -1$, $X = \mathrm{diag}(1/2, 3/2)$, $Y = \mathrm{diag}(1, i)$. The boundary generating curve of $W_J(A)$ is shown in Fig. 1.

Now, we consider $Y^* = kX$ in (7), for some constant $k \in \mathbb{C}$. In this special case, the block matrix A has an elliptical classical numerical range [3]. In the following corollary, we prove that $W_J(A)$ is bounded by a hyperbola. Our proof follows the steps of the proof of Corollary 2.3 in [3].

Corollary 3.1. *Let $J = I_r \oplus -I_{n-r}$, $0 < r < n$, and let $p = \min(r, n - r)$. Let A be a block matrix of type* (7), *with $Y^* = kX$, $k \in \mathbb{C}$. Moreover, let $2\beta_{l\pm} = a + b \pm$*

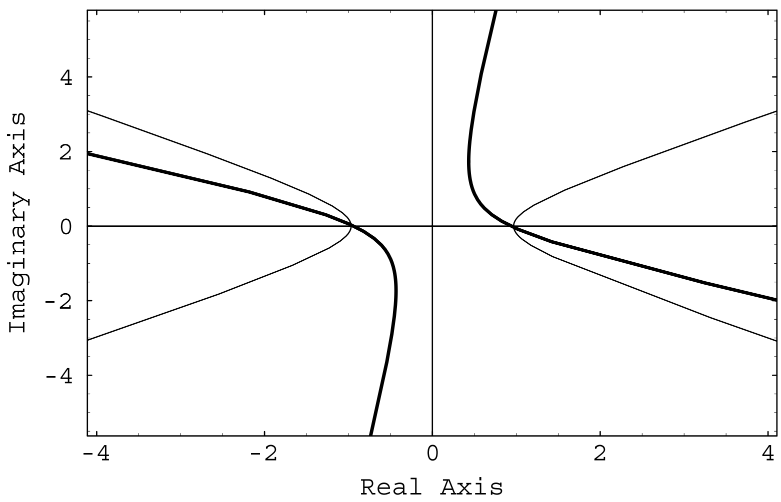

Fig. 1. The boundary generating curve of $W_J(A)$ for the matrix A in Example 1.

$\sqrt{(a-b)^2+4k\sigma_l^2}$, *$l=1,\ldots,p$, where $\sigma_1 \geqslant \ldots \geqslant \sigma_p$ are the singular values of X. If*

$$2\,\mathrm{Re}(\overline{\beta_{l+}}\beta_{l-}) < |a|^2+|b|^2-\sigma_l^2(1+|k|^2) < |\beta_{l+}|^2+|\beta_{l-}|^2, \quad l=1,\ldots,p,$$

then $W_J(A)$ is bounded by a hyperbola, with foci at $\beta_{1\pm}$ and with non-transverse axis of length $\sqrt{|\beta_{1+}|^2+|\beta_{1-}|^2-|a|^2-|b|^2+\sigma_1^2(1+|k|^2)}$.

Proof. From the proof of Theorem 3.1, we may conclude that $W_J^+(A)$ is the convex hull of the sets $V_{J_2}^+(B_1)$, where $J_2=I_1\oplus -I_1$ and B_l are the 2×2 matrices in (12), $l=1,\ldots,p$. Since $Y^*=kX$, then $\delta_l=|k|\sigma_l$ and $2\phi_l=\arg k+\pi$, $l=1,\ldots,p$. We shall prove that

$$V_{J_2}^+(B_l)\subset V_{J_2}^+(B_1), \quad l=2,\ldots,p. \tag{13}$$

Recalling the definition of $V_{J_2}^+(B_l)$, we have

$$V_{J_2}^+(B_l)=\left\{a|x_1|^2-b|x_2|^2+\sigma_l\mathrm{e}^{\mathrm{i}\phi_l}(\bar{x}_1x_2+|k|\bar{x}_2x_1) : |x_1|^2-|x_2|^2=1\right\}. \tag{14}$$

Taking in (14) $r=|x_1|^2$, $\gamma=\arg x_2-\arg x_1$, $\gamma\in[0,2\pi)$, and denoting by $\mathscr{E}$ the curve

$$\left\{(1+|k|)\cos\gamma+\mathrm{i}(1-|k|)\sin\gamma : \gamma\in[0,2\pi)\right\},$$

we may write

$$V_{J_2}^+(B_l)=\bigcup_{r\geqslant 1}\left\{(a-b)r+b+\sigma_l\mathrm{e}^{\mathrm{i}\phi_l}\sqrt{r(r-1)}\mathscr{E}\right\}.$$

Any $z\in V_{J_2}^+(B_l)$ lies on the curve $\Gamma_l=(a-b)r+b+\sigma_l\mathrm{e}^{\mathrm{i}\phi_l}\sqrt{r(r-1)}\mathscr{E}$, for a certain choice of r. Depending on the value of k, the curve $\mathscr{E}$ is either the boundary of an ellipse or a line segment, in both cases centered at the origin. Since $\sigma_1\geqslant\sigma_l$, then Γ_l lies in the domain bounded by Γ_1. Since $\Gamma_1\subset V_{J_2}^+(B_1)$ and the later set is convex, it follows that $z\in\Gamma_l\subset V_{J_2}^+(B_1)$ and (13) holds. Hence, $W_J^+(A)=V_{J_2}^+(B_1)$. An analogous result holds for $W_J^-(A)$ and the corollary follows directly from Theorem 3.1. □

If $k=0$ in Corollary 3.1, then we have the following result concerning matrices of type:

$$A=\begin{bmatrix} aI_r & X\\ 0 & bI_{n-r}\end{bmatrix}, \quad a,b\in\mathbb{C},\ X\in M_{r,n-r}. \tag{15}$$

Corollary 3.2. *Let $J=I_r\oplus -I_{n-r}$, $0<r<n$, and let A be a block matrix of type* (15). *If $0<\sigma_1<|a-b|$, where σ_1 is the largest singular value of X, then $W_J(A)$ is bounded by a hyperbola, with a and b as foci, and with non-transverse axis of length σ_1.*

4. Curvature of $W_J(A)$ at a boundary point

We will state our results for $W_J^+(A)$, with the understanding that analogous results hold for $W_J^-(A)$, because $W_J^-(A) = -W_{-J}^+(A)$. Suppose that the boundary point under consideration belongs to $W_J^+(A)$. For any $\alpha, \beta \in \mathbb{C}$, $W_J^+(\alpha I + \beta A) = \alpha + \beta W_J^+(A)$. Thus, by an appropriate rotation and translation, we can assume that the boundary point stays at the origin and the supporting line coincides with the y-axis, $W_J^+(A)$ being in the half-plane $x \leqslant 0$. For

$$H_1 x_i = \lambda_i J x_i, \quad i = 0, \dots, n-1, \tag{16}$$

this situation obviously corresponds to the case that $\lambda_0 = 0$,

$$H_1 x_0 = 0, \quad x_0^* H_2 x_0 = 0, \quad \text{for some vector } x_0 \in \mathbb{C}^n, \quad x_0^* J x_0 = 1, \tag{17}$$

and

$$\lambda_i x_i^* J x_i \leqslant 0, \quad i = 1, \dots, n-1. \tag{18}$$

It is known that if a boundary point of $W_J(A)$ is a corner point of $W_J(A)$, then it is an eigenvalue of A and there exists a pseudo-unitary matrix $U \in U_{r,n-r}$ such that $U^* J A U = A_1 \oplus A_2$, for square matrices A_1, A_2 of order at least one [14]. Therefore, if we restrict our study to the case that the curve $C_J(A)$ given by (1) is irreducible and $n \geqslant 2$, the boundary of $W_J(A)$ is a smooth curve. As will be proved in Theorem 4.1, this boundary can then contain a line segment only if the corresponding line is a *singular tangent* to $C_J(A)$, that is, its coordinates satisfy (1) and also all the three partial derivatives of the left-hand side with respect to u, v and w vanish. A line is called a *non-singular tangent* to $C_J(A)$ if its coordinates satisfy (1) and at most one of the partial derivatives of the left-hand side of (1) with respect to u, v and w vanish.

The proofs in this section are inspired on the parallel results on the classical numerical range due to Fiedler [6,7].

Lemma 4.1. *Let A be a $n \times n$ complex matrix satisfying* (16)–(18). *Then the line $x = 0$ is a non-singular tangent of $C_J(A)$ at the origin if and only if* 0 *is a simple eigenvalue of H_1.*

Proof. The left-hand side of (1) may be written as

$$c_0 u^n + c_1(v, w) u^{n-1} + \cdots + c_n(v, w), \tag{19}$$

with $c_k(v, w)$ being homogeneous polynomials in v, w of degree k (possibly zero). By simple algebraic-geometrical considerations, if $x = 0$ is a non-singular tangent of $C_J(A)$ then $c_0 = 0$ and $c_1(v, w) = kw$, $k \neq 0$. (Then $w = 0$ is the dual tangent at the dual point $(1, 0, 0)$, which means that the origin is the tangent point of the tangent $x = 0$.) If U is a unitary matrix for which $U^* H_1 U$ is diagonal with the first entry zero, it follows that H_1 has to have rank $n - 1$, since, otherwise, the term kwu^{n-1} with $k \neq 0$ would not appear in (19). The converse is easily checked. □

Theorem 4.1. *Let $A \in M_n$. The boundary of $W_J^+(A)$ contains a line segment only if the corresponding supporting line is a singular tangent to $C_J(A)$.*

Proof. Without loss of generality, suppose that the supporting line referred to in the theorem is $x = 0$, with $W_J^+(A)$ lying in the half-plane $x \leqslant 0$. Let $(0, y_1)$ and $(0, y_2)$ be the (distinct) extreme points of the line segment. It can be easily seen that

$$x_1^* J A x_1 = y_1 i \quad \text{and} \quad x_2^* J A x_2 = y_2 i, \tag{20}$$

for $x_1, x_2 \in \mathbb{C}^n$ linearly independent vectors (since $y_1 \neq y_2$) such that $x_1^* J x_1 = x_2^* J x_2 = 1$. From (20), we have that $x_1^* H_1 x_1 = x_2^* H_1 x_2 = 0$. Since $x = 0$ is a supporting line of $W_J^+(A)$ and $W_J^+(A)$ is on the half-plane $x \leqslant 0$, we conclude that $x_1^* H_1 x_1 = x_2^* H_1 x_2 = 0$ are extreme points of the quadratic form $x^* H_1 x$, with x satisfying $x^* J x = 1$. The boundary generating curve of $W_J^+(A)$ is $\det J \det(J H_1 + wI) = 0$, therefore 0 is not a simple eigenvalue of $J H_1$. By the previous Lemma, $x = 0$ is a singular tangent to $C_J(A)$. □

Through the paper, we denote by $A^\dagger$ the *Moore–Penrose generalized inverse* of A.

We state a lemma of the perturbation theory for matrices, which will be used in the study of the curvature of $W_J(A)$ at a boundary point.

Lemma 4.2. *Let 0 be a simple eigenvalue of $L \in H_n$, with a corresponding eigenvector u such that $u^* J u = 1$. Then for any $K \in M_n$ and $\epsilon \in \mathbb{C}$, $|\epsilon| \to 0$, the expansion of the eigenvalue $\lambda(\epsilon)$ of the matrix $J(L + \epsilon K)$ in the neighborhood of 0 is*

$$\lambda(\epsilon) = \epsilon\lambda^{(1)} + \epsilon^2\lambda^{(2)} + \mathrm{O}(\epsilon^3),$$

where $\lambda^{(1)} = u^* K u$ and $\lambda^{(2)} = -u^*(K - \lambda^{(1)} J) L^\dagger (K - \lambda^{(1)} J) u$, and the expansion of the eigenvector $u(\epsilon)$ associated with $\lambda(\epsilon)$ in the same neighborhood of 0 is

$$u(\epsilon) = u - \epsilon L^\dagger (K - \lambda^{(1)} J) u + \mathrm{O}(\epsilon^2).$$

Proof. Let $(L + \epsilon K) u(\epsilon) = \lambda(\epsilon) J u(\epsilon)$. By the simplicity of the eigenvalue 0, $\lambda(\epsilon)$ and $u(\epsilon)$ are analytic functions of ϵ in the neighborhood of 0. Consider the power series for $\lambda(\epsilon)$ and $u(\epsilon)$, respectively,

$$\lambda(\epsilon) = \epsilon\lambda^{(1)} + \epsilon^2\lambda^{(2)} + \mathrm{O}(\epsilon^3) \quad \text{and} \quad u(\epsilon) = u + \epsilon u^{(1)} + \epsilon^2 u^{(2)} + \mathrm{O}(\epsilon^3).$$

It can easily be seen that $Ku + Lu^{(1)} = \lambda^{(1)} J u$ and

$$K u^{(1)} + L u^{(2)} = \lambda^{(1)} J u^{(1)} + \lambda^{(2)} J u. \tag{21}$$

Thus, $\lambda^{(1)} = u^* K u$ and $u^{(1)} = -L^\dagger (K - \lambda^{(1)} J) u$. Multiplying (21) on the left by u^*, we get $\lambda^{(2)} = u^*(K - \lambda^{(1)} J) u^{(1)}$, and the result follows. □

Theorem 4.2. *Let $x_0 \in \mathbb{C}^n$ and $A \in M_n$ satisfying* (16)–(18). *If $x = 0$ is a non-singular tangent of $C_J(A)$, then $C_J(A)$ has a positive finite curvature at the origin and the radius of curvature at* 0 *is given by $r = -2x_0^* H_2 H_1^\dagger H_2 x_0$.*

Proof. To evaluate the radius r of curvature at 0, we determine the coordinates $x(v)$ and $y(v)$ of the intersection of the supporting line $x + vy + w = 0$ (in the "neighborhood" of $x = 0$) with $C_J(A)$. We can conclude that $\det(JH_1 + vJH_2 + wI) = 0$, and so $-w$ is an eigenvalue of the matrix $J(H_1 + vH_2)$. Since $x = 0$ is a non-singular tangent of $C_J(A)$, by Lemma 4.1, $w = 0$ is a simple eigenvalue of H_1. Therefore, we can apply Lemma 4.2 to the left-hand side of (1), for H_1 as L, H_2 as K, x_0 as u, and v as ϵ. Since $x_0^* H_2 x_0 = 0$, we obtain

$$w = w(v) = -v^2 x_0^* H_2 H_1^\dagger H_2 x_0 + \mathrm{O}(v^3). \tag{22}$$

An eigenvector $\nu(v)$ associated with the eigenvalue $-w(v)$ of $J(H_1 + vH_2)$ has the following expansion $\nu(v) = x_0 + vH_1^\dagger H_2 x_0 + \mathrm{O}(v^2)$. The parametric equations of $C_J(A)$ in the "neighborhood" of $x = 0$ are

$$x(v) = \frac{\nu^*(v) H_1 \nu(v)}{\nu^*(v) J \nu(v)}, \quad y(v) = \frac{\nu^*(v) H_2 \nu(v)}{\nu^*(v) J \nu(v)},$$

with $|v| < \epsilon$ sufficiently small. From (22), it follows that

$$x(v) = v^2 x_0^* H_2 H_1^\dagger H_2 x_0 + \mathrm{O}(v^3), \quad y(v) = 2v x_0^* H_2 H_1^\dagger H_2 x_0 + \mathrm{O}(v^2).$$

Since $r = -\lim_{v \to 0} y^2(v)/(2x(v))$ and $x(0) = y(0) = x'(0) = 0$, the theorem is proved. □

Considering $J = I_n$ in Lemma 4.1 and in Theorem 4.2, we obtain Theorem 3.3 of Fiedler in [6].

The *J-generalized Levinger transformation* of $A \in M_n$ is defined by

$$\mathscr{L}_J(A, \alpha) = (1 - \alpha)A + \alpha A^{[*]}, \quad \alpha \in [0, 1]$$

(denoted simply by $\mathscr{L}(A, \alpha)$, if $J = I_n$). Obviously,

$$\mathscr{L}_J(A, \alpha) = JH_1 + \mathrm{i}(1 - 2\alpha)JH_2 \tag{23}$$

and $\mathscr{L}_J(A^{[*]}, \alpha) = \mathscr{L}_J(A, 1 - \alpha)$, $\alpha \in [0, 1]$.

There is a relation between $W_J(A)$ and $W_J(\mathscr{L}_J(A, \alpha))$, in case the sets are hyperbolical. A parallel result for the classical case ($J = I_n$) was presented by Maroulas et al. [15]. In fact, due to (23), we may write

$$W_J(\mathscr{L}_J(A, \alpha)) = \{x + \mathrm{i}(1 - 2\alpha)y : x, y \in \mathbb{R}, x + \mathrm{i}y \in W_J(A)\}.$$

Supposing that the boundary of $W_J(A)$ in the plane (u, v) has equation

$$\frac{u^2}{M^2} - \frac{v^2}{N^2} = 1, \quad M, N > 0,$$

via the change of variables $x = u$ and $y = (1 - 2\alpha)v$, then the boundary of $W_J(\mathscr{L}_J(A, \alpha))$ has equation

$$\frac{x^2}{M^2} - \frac{y^2}{(1-2\alpha)^2N^2} = 1.$$

If $\alpha \neq 1/2$, then $W_J(A)$ is bounded by a non-degenerate hyperbola with transverse and non-transverse axis of length M and N, respectively, if and only if $W_J(\mathscr{L}_J(A, \alpha))$ is bounded by a non-degenerate hyperbola with transverse and non-transverse axis of length M and $|1 - 2\alpha|N$, respectively. If $\alpha = 1/2$, then $W_J(\mathscr{L}_J(A, \alpha)) = \operatorname{Re} W_J(A)$ is a subset of a line.

In [7], Fiedler investigated the connection between the curvature of the boundary of $W(A)$, $A \in M_n$, and the curvature of the *generalized Levinger curve*, which is the graph of the function $\phi_A : [0, 1] \to \mathbb{R}$ such that

$$\phi_A(\alpha) = \max\left\{\operatorname{Re}\lambda : \lambda \in \sigma(\mathscr{L}(A, \alpha))\right\}.$$

Now, we consider the *J-generalized Levinger function* $\phi_{A,J} : [0, 1] \to \mathbb{R}$ defined by

$$\phi_{A,J}(\alpha) = \max\left\{\operatorname{Re}\lambda : \lambda \in \sigma(\mathscr{L}_J(A, \alpha))\right\}.$$

The following theorem reduces to [7, Theorem 3.2], considering $J = I_n$.

Theorem 4.3. *Let $x_0 \in \mathbb{C}^n$ and $A \in M_n$ satisfy* (16)–(18), *and let the maximum eigenvalue* 0 *of H_1 be simple.*

(i) *If x_0 is not an eigenvector of A associated with* 0, *then there is an open interval $\mathscr{I}$ with midpoint* 1/2, *such that the J-generalized Levinger function $\phi_{A,J}$ is increasing in the left half of $\mathscr{I}$, has zero derivative at the point* 1/2 *and is concave in $\mathscr{I}$. Moreover, the radius of curvature R of $\phi_{A,J}$ at the point* 1/2 *is related to the radius of curvature r of the boundary of $W_J^+(A)$ at the point* 0 *by* $Rr = 1/4$.

(ii) *If x_0 is an eigenvector of A associated with* 0, *then $\phi_{A,J}$ is constant.*

Proof. (i) Since 0 is a simple eigenvalue of H_1, there is a unique eigenvalue $\lambda(\epsilon)$ of $\mathscr{L}_J(A, 1/2 + \epsilon) = J(H_1 - i2\epsilon H_2)$, in some neighborhood of 0. By Lemma 4.2, we have $\lambda(\epsilon) = \epsilon\lambda^{(1)} + \epsilon^2\lambda^{(2)} + O(\epsilon^3)$, where $\lambda^{(1)} = x_0^* H_2 x_0 = 0$ and $\lambda^{(2)} = 4x_0^* H_2 H_1^\dagger H_2 x_0$. For ϵ real, $\operatorname{Re}\lambda(\epsilon) = \epsilon^2 \operatorname{Re}\lambda^{(2)} + O(\epsilon^3)$. So, there is an open interval $\mathscr{I}$ containing 0 in which $\phi_{A,J}(1/2 + \epsilon) = \operatorname{Re}\lambda(\epsilon)$. For the derivatives, we have $\phi'_{A,J}(1/2) = 0$, $\phi''_{A,J}(1/2) = 2\operatorname{Re}\lambda^{(2)} = 8x_0^* H_2 H_1^\dagger H_2 x_0$. Since x_0 is not an eigenvector of A, and therefore of JA, associated to 0, $H_2 x_0$ is not a multiple of x_0. By the hypothesis, 0 is the maximum eigenvalue of the matrix H_1, therefore $-H_1$ is positive semidefinite. Under our assumptions, the eigenspace associated with 0 is one-dimensional and H_1 has only multiples of x_0 as annihilating vectors, and the same holds for its Moore-Penrose inverse. We can conclude that the second derivative of the J-generalized Levinger function at the point 1/2 is positive and remains positive in some neighborhood of 1/2. Hence, the J-generalized Levinger function is concave in $\mathscr{I}$ and the radius of curvature R of the J-generalized Levinger function at the point 1/2 is $R = -1/\phi''_{A,J}(1/2)$. By Theorem 4.2, the radius of the curvature r of the boundary of $W_J^+(A)$ at the origin is $r = -2x_0^* H_2 H_1^\dagger H_2 x_0$. Therefore, $Rr = 1/4$.

(ii) If $Ax_0 = 0$, then $JAx_0 = 0$. Thus, 0 is a common eigenvalue of all matrices $\mathscr{L}_J(A, \alpha)$, $\alpha \in [0, 1]$. □

Remark 4.1. In Theorem 4.3, if x_0 is an eigenvector of A, then 0 is a corner of $W_J(A)$, so that the radius of curvature r can be considered as 0 and R as infinity. Moreover, $W_J(A)$ has a flat point at 0 if the maximum eigenvalue of H_1 is multiple.

References

[1] N. Bebiano, R. Lemos, J. da Providência, G. Soares, On generalized numerical ranges of operators on an indefinite inner product space, Linear and Multilinear Algebra 52 (2004) 203–233.
[2] V. Bolotnikov, C.-K. Li, P.R. Meade, C. Mehl, L. Rodman, Shells of matrices in indefinite inner product spaces, Electr. J. Linear Algebra 9 (2002) 67–92.
[3] E. Brown, I. Spitkovsky, On matrices with elliptical numerical ranges, Linear and Multilinear Algebra 52 (2004) 177–193.
[4] C. Davis, The shell of a Hilbert-space operator, Acta Sci. Math. (Szeged) 29 (1968) 69–86.
[5] C. Davis, The shell of a Hilbert-space operator II, Acta Sci. Math. (Szeged) 31 (1970) 301–318.
[6] M. Fiedler, Geometry of the numerical range of matrices, Linear Algebra Appl. 37 (1981) 81–96.
[7] M. Fiedler, Numerical range of matrices and Levinger's theorem, Linear Algebra Appl. 220 (1995) 171–180.
[8] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[9] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[10] R. Kippenhahn, Uber der wertevorrat einer matrix, Math. Nach. 6 (1951) 193–228.
[11] P. Lancaster, M. Tismenetsky, The Theory of Matrices: Second Edition with Applications, Academic Press, New York, 1985.
[12] C.-K. Li, N.K. Tsing, F. Uhlig, Numerical ranges of an operator in an indefinite inner product space, Electr. J. Linear Algebra 1 (1996) 1–17.
[13] C.-K. Li, L. Rodman, Shapes and computer generation of numerical ranges of Krein space operators, Electr. J. Linear Algebra 3 (1998) 31–47.
[14] C.-K. Li, L. Rodman, Remarks of numerical ranges of operators in spaces with an indefinite inner product, Proc. Amer. Math. Soc. 126 (1998) 973–982.
[15] J. Maroulas, P.J. Psarrakos, M.J. Tsatsomeros, Perron–Frobenius type results on the numerical range, Linear Algebra Appl. 348 (2002) 49–62.
[16] H. Shapiro, A conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices. II, Linear Algebra Appl. 45 (1982) 97–108.
[17] C.K. Sze, S-normality and polygonal S-numerical ranges, M.Phil. Thesis, University of Hong Kong, 1998.
[18] R.J. Walker, Algebraic Curves, Princeton University Press, Princeton, 1950.