On the convexity and circularity of the numerical range of nilpotent quaternionic matrices

Luís Carvalho, Cristina Diogo and Sérgio Mendes

ABSTRACT. We provide a sufficient condition for the numerical range of a nilpotent matrix N to be circular in terms of the existence of cycles in an undirected graph associated with N. We prove that if we add to this matrix N a diagonal real matrix D, the matrix $D + N$ has convex numerical range. For 3×3 nilpotent matrices, we strength further our results and obtain necessary and sufficient conditions for circularity and convexity of the numerical range.

CONTENTS

1. Introduction

Let $\mathbb{H}$ denote the Hamilton quaternions and let A be a $n \times n$ matrix with quaternionic entries. The quaternionic numerical range of A, denoted $W(A)$, was introduced in 1951 in Kippenhahn's seminal article [Ki] as an analogue of the long established complex numerical range (see [R] for an account on quaternionic numerical ranges). Specifically, $W(A)$ is the subset of $\mathbb{H}$ whose elements have the form $\boldsymbol{x}^* A \boldsymbol{x}$, where $\boldsymbol{x}$ runs over the unit sphere of $\mathbb{H}^n$. Due to the failure of Toeplitz-Hausdorff theorem in the quaternionic setting, the convexity of $W(A)$ has been studied by several authors. In [Ki], Kippenhahn introduced the Bild of A, denoted $B(A)$, as the intersection of

$W(A)$ with the complex plane and studied its convexity. The Bild is indeed a planar substitute of $W(A)$ in the sense that every element of $W(A)$ is equivalent to an element in $W(A) \cap \mathbb{C}$.

The first remarkable result on convexity is due to Au-Yeung who proved in 1984 that $W(A)$ is convex if, and only if, the projection of $W(A)$ over $\mathbb{R}$ (resp., $\mathbb{C}$) equals the real (resp., the complex) elements in $W(A)$, see [AY1, theorems 2 and 3]. In that same paper, the author gives necessary and sufficient conditions on the eigenvalues of a normal matrix A for $W(A)$ to be convex. The convexity of the Bild, already an issue in [Ki], was established for normal matrices in 1994 by So, Thompson and Zhang. They proved in [STZ, p. 192] that the closed upper half plane part of the Bild (the upper Bild $B^+(A)$) of a normal matrix A is the convex hull of eigenvalues and cone vertices. Later on, the proof was simplified by Au-Yeung [AY2].

The general case was settled by So and Thompson in 1996. In [ST, theorem 15.2] they proved that for any matrix A, the intersection of $W(A)$ with the closed upper half plane is always convex.

Another problem that attracted much attention in the complex setting is the shape of the numerical range. In [ST, theorem 17.1], So and Thompson characterized the numerical range of 2×2 quaternionic matrices, the analogue of the elliptical range theorem. However, compared with the complex case, the shape of the numerical range of quaternionic matrices seems to have been more neglected.

In this article we study the convexity and shape of the numerical range for nilpotent quaternionic matrices. To be more specific, we determine under what conditions $W(A)$ has circular shape or, at least, is convex. In section 2 of this article we recall some definitions and fix notation. In section 3 we deal with the circularity of the numerical range. Theorem 3.1 shows that if the numerical range of a nilpotent matrix is a disk, its center must be located at the origin. This is the quaternionic analogue of [MM, proposition 1]. We conclude this section with theorem 3.5 which says that a sufficient condition for the numerical range of a nilpotent matrix A to be a disk is that the associated graph of A is a tree. This condition is not necessary as example 3.8 shows. In section 4 we extend the results of the previous section. Every matrix $A \in \mathcal{M}_n(\mathbb{H})$ is, up to unitary equivalence, upper triangular and every upper triangular matrix decomposes as a sum of a diagonal with a nilpotent matrix. The main result of this section is theorem 4.2, where it is proved that when the diagonal part is real and the nilpotent part is a tree then the numerical range is a union of disks. To reach this result we apply Berge's maximum theorem, a technique not much seen in the literature. Corollary 4.3 proves that this class of matrices have convex numerical range. We end the section providing an example of one of these matrices where the union of disks that compose its numerical range is in fact an ellipse. In section 5, we focus on 3×3 nilpotent matrices concerning convexity and circularity of the numerical range. Theorem 5.1 says that a necessary and sufficient condition

for the numerical range of $A \in \mathcal{M}_3(\mathbb{H})$ to be a disk with center at the origin is that A is cycle-free. On the other hand, theorem 5.3 gives a necessary and sufficient condition for the same class of matrices to have convex numerical range. Specifically, $W(A)$ is convex if, and only if, $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. The link with our work is now provided by theorem 1 of Chien and Tam [CT] in the complex setting.

2. Preliminaries and notation

In this section we present some well known facts about quaternions and fix some notation. The quaternionic skew-field $\mathbb{H}$ is an algebra of rank 4 over $\mathbb{R}$ with basis $\{1, i, j, k\}$. The product in $\mathbb{H}$ is given by $i^2 = j^2 = k^2 = ijk = -1$. Denote the pure quaternions by $\mathbb{P} = \mathrm{span}_{\mathbb{R}} \{i, j, k\}$. For any $q = a_0 + a_1 i + a_2 j + a_3 k \in \mathbb{H}$ let $\pi_{\mathbb{R}}(q) = a_0$ and $\pi_{\mathbb{P}}(q) = a_1 i + a_2 j + a_3 k$ be the real and imaginary parts of q, respectively. The conjugate of q is given by $q^* = \pi_{\mathbb{R}}(q) - \pi_{\mathbb{P}}(q)$ and the norm is defined by $|q|^2 = qq^*$. Two quaternions $q_1, q_2 \in \mathbb{H}$ are called similar if there exists a unitary quaternion s such that $s^* q_2 s = q_1$. Similarity is an equivalence relation and we denote by $[q]$ the equivalence class containing q. A necessary and sufficient condition for the similarity of q_1 and q_2 is that $\pi_{\mathbb{R}}(q_1) = \pi_{\mathbb{R}}(q_2)$ and $|\pi_{\mathbb{P}}(q_1)| = |\pi_{\mathbb{P}}(q_2)|$ [R, theorem 2.2.6].

Let $\mathbb{F}$ denote $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. Let $\mathbb{F}^n$ be the n-dimensional $\mathbb{F}$-space. For $\boldsymbol{x} \in \mathbb{F}^n$, $\boldsymbol{x}^*$ denote the conjugate transpose of $\boldsymbol{x}$. The disk with center $\boldsymbol{a} \in \mathbb{F}^n$ and radius $r \geq 0$ is the set $\mathbb{D}_{\mathbb{F}^n}(\boldsymbol{a}, r) = \{\boldsymbol{x} \in \mathbb{F}^n : |\boldsymbol{x} - \boldsymbol{a}| \leq r\}$ and its boundary is the sphere $\mathbb{S}_{\mathbb{F}^n}(\boldsymbol{a}, r)$. In particular, if $\boldsymbol{a} = \boldsymbol{0}$ and $r = 1$, we simply write $\mathbb{D}_{\mathbb{F}^n}$ and $\mathbb{S}_{\mathbb{F}^n}$. The group of unitary quaternions is denoted by $\mathbb{S}_{\mathbb{H}}$. Notice that we are considering the singleton $\{\boldsymbol{a}\}$ to be the disk with center a and radius $r = 0$. Any $x \in \mathbb{H}$ can be written as $x = \beta z$, with $\beta = |x|$ and $z \in \mathbb{S}_{\mathbb{H}}$[1]. We introduce the following notation:

$$\mathbb{R}^{n,+} = \{(\beta_1, ..., \beta_n) \in \mathbb{R}^n : \beta_i \geq 0, 1 \leq i \leq n\}$$
$$\mathbb{S}_{\mathbb{R}^n}^+ = \mathbb{S}_{\mathbb{R}^n} \cap \mathbb{R}^{n,+}.$$

Let $\mathcal{M}_n(\mathbb{F})$ be the set of all $n \times n$ matrices with entries over $\mathbb{F}$. Let $A \in \mathcal{M}_n(\mathbb{H})$. The set

$$W(A) = \{\boldsymbol{x}^* A \boldsymbol{x} : \boldsymbol{x} \in \mathbb{S}_{\mathbb{H}^n}\}$$

is called the numerical range of A in $\mathbb{H}$. As usual, the complex numerical range of a complex matrix is defined by

$$W_{\mathbb{C}}(A) = \{\boldsymbol{x}^* A \boldsymbol{x} : \boldsymbol{x} \in \mathbb{S}_{\mathbb{C}^n}\}.$$

It is well known that if $q \in W(A)$ then $[q] \subseteq W(A)$ [R, page 38]. Therefore, it is enough to study the subset of complex elements in each similarity

---

[1]If $x \neq 0$, z is uniquely defined. However, if $x = 0$ that is not the case and we take $z = 1$.

class. This set is known as $B(A)$, the Bild of A

$$B(A) = W(A) \cap \mathbb{C}.$$

Although the Bild may not be convex, the upper Bild $B^+(A) = W(A) \cap \mathbb{C}^+$ is always convex, see [ST].

Taking into account that $\mathbb{R}$ can be seen as a real subspace of $\mathbb{H}$, what we denoted by $\pi_{\mathbb{R}}$ is, in fact, the projection of $\mathbb{H}$ over $\mathbb{R}$, $\pi_{\mathbb{R}} : \mathbb{H} \to \mathbb{R}$. The projection of $W(A)$ over $\mathbb{R}$ is

$$\pi_{\mathbb{R}}(W(A)) = \{\pi_{\mathbb{R}}(w) : \ w \in W(A)\}.$$

In the next section we will define a relation between the circularity of the numerical range of A and the lack of cycles of an associated undirected graph. To be more specific, given a matrix $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{H})$ we may define the underlying undirected graph $\mathcal{G}_A$ with n vertices as the graph with an edge between i and j whenever $a_{ij} \neq 0$ or $a_{ji} \neq 0$. That is, if $\delta : \mathbb{H} \to \{0,1\}$ is the indicator function, $\delta(q) = 1$ if $q \neq 0$ and $\delta(q) = 0$ otherwise, let A_δ be the symmetric matrix given by

$$A_\delta = \Big[\delta(\max\{a_{ij}, a_{ji}\})\Big]_{i,j=1}^n.$$

Then A_δ is precisely the adjacency matrix of the undirected graph $\mathcal{G}_A$. We say that the graph $\mathcal{G}_A$ has a path between the vertices $i, j \in \{1, \ldots, n\}$, if there is a sequence of vertices $(i_1, i_2, \ldots, i_p)$ such that:

$$i_1 = i, \qquad i_p = j, \qquad (A_\delta)_{i_k, i_{k+1}} = 1, \text{ for } k = 1, \ldots, p-1.$$

In terms of the elements a_{km} of the matrix A this condition is equivalent to $a_{i_k i_{k+1}} \neq 0$, for all $k \in \{1, \ldots, p-1\}$ and to $a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{p-1} i_p} \neq 0$. The graph $\mathcal{G}_A$ is connected if there is a path between any pair of vertices $i, j \in \{1, \ldots, n\}$, otherwise it is disconnected. We say that the matrix A is connected (resp., disconnected) whenever $\mathcal{G}_A$ is connected (resp., disconnected) . The graph $\mathcal{G}_A$ has a cycle (or the matrix A has a cycle) if there is a vertex $i \in \{1, \ldots, n\}$ and a path connecting i to itself. Loops are seen as cycles. The graph $\mathcal{G}_A$ is cycle-free (or the matrix A is cycle-free) if there are no cycles in $\mathcal{G}_A$.

If the graph $\mathcal{G}_A$ is connected and cycle-free then it is a tree, and in this case the number of edges is $n-1$ [Di, corollary 1.5.3]. It follows that there exists one vertex with only one edge and, if we eliminate this vertex and its edge, we get a graph with $n-1$ vertices and $n-2$ edges, which is also a tree. If A is a nilpotent matrix and the graph $\mathcal{G}_A$ is a tree, then there exists a permutation matrix P such that $P^\top A P$ is upper triangular. More generally, two matrices $A, A' \in \mathcal{M}_n(\mathbb{H})$ are unitarily equivalent if there exists a unitary $U \in \mathcal{M}_n(\mathbb{H})$ such that $A' = U^* A U$, in which case we write $A' \sim A$. The relation $\sim$ is an equivalence relation. By Schur's triangularization theorem [R, theorem 5.3.6], every matrix $A \in \mathcal{M}_n(\mathbb{H})$ is unitarily equivalent to an upper triangular matrix whose diagonal is complex. By [R, theorem 3.5.4], the numerical range is invariant for the equivalence classes $[A]_\sim$, with

$A \in \mathcal{M}_n(\mathbb{H})$. Therefore, it is enough to consider upper triangular matrices $A \in \mathcal{M}_n(\mathbb{H})$ with complex diagonal entries.

For the rest of this article we will assume that $A \in \mathcal{M}_n(\mathbb{H})$ is upper triangular.

3. Circularity of the numerical range

Our first result shows that the numerical range of a nilpotent matrix is either circular with center at the origin or it is not a disk. In other words, when circular, the disk must be centered at the origin.

Given $A \in \mathcal{M}_n(\mathbb{H})$ there exists an associated complex matrix

$$\chi(A) = \begin{bmatrix} A_1 & A_2 \\ -\bar{A}_2 & \bar{A}_1 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C}),$$

where $A_1, A_2 \in \mathcal{M}_n(\mathbb{C})$ and $A = A_1 + A_2 j$.

Theorem 3.1. *Let $A \in \mathcal{M}_n(\mathbb{H})$ be nilpotent. If $W(A)$ is a disk, then its center is at the origin.*

Proof. Since $\chi(AB) = \chi(A)\chi(B)$,[Zh, theorem 4.2], A is nilpotent if and only if $\chi(A)$ is nilpotent. When $W(A)$ is circular, it is convex, and according to [AY1, theorem 2] and [AY2, p. 280], $W(A) \cap \mathbb{C} = W_{\mathbb{C}}(\chi(A))$. Thus $W_{\mathbb{C}}(\chi(A))$ is a disk in $\mathbb{C}$. According to [MM, proposition 1] a nilpotent complex matrix whose numerical range is a disk must have center at the origin. Thus $W(A) \cap \mathbb{C} = W_{\mathbb{C}}(\chi(A)) = \mathbb{D}_{\mathbb{C}}(0, r)$, where r is the radius. We conclude, rebuilding the numerical range by taking the equivalence classes of the elements of the Bild, that $W(A)$ is a disk centered at the origin, i.e. $W(A) = \mathbb{D}_{\mathbb{H}}(0, r)$. $\qquad \square$

In theorem 3.5 we will prove that if the graph associated with the nilpotent matrix $A \in \mathcal{M}_n(\mathbb{H})$ has no cycles then the numerical range of A is a disk. When the graph $\mathcal{G}_A$ is disconnected, we can partition the set of vertices into connected components, where each component has no edge to the other components. Then, in terms of the original matrix A, we can (through a reordering of the vertices if necessary, *i.e.* through $P^\top A P$ where P is a permutation of $\{1, \ldots, n\}$) write A as a block matrix. Now, each block of matrix A is connected. In addition, if A is cycle-free, then each block is cycle-free.

The fact that the quaternionic numerical range is not always convex motivates the following definition.

Definition 3.2. Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be subsets of $\mathbb{H}$. The inter-convex hull of the $\mathcal{A}_i$'s is the set

$$\operatorname{iconv}\{\mathcal{A}_1, \ldots, \mathcal{A}_n\} = \Big\{ \sum_i \alpha_i^2 a_i : \boldsymbol{\alpha} \in \mathbb{S}_{\mathbb{R}^n}, \ a_i \in \mathcal{A}_i, \ i = 1, \ldots, n \Big\}.$$

We can easily prove that:

Proposition 3.3. *Let $A_1 \oplus \ldots \oplus A_k \in \mathcal{M}_n(\mathbb{H})$. Then*

$$W(A_1 \oplus \ldots \oplus A_k) = \mathrm{iconv}\Big\{ W(A_1), \ldots, W(A_k) \Big\}.$$

In particular, if $W(A_i)$ is convex, for $i = 1, \ldots, k$, then

$$W(A_1 \oplus \ldots \oplus A_k) = \mathrm{conv}\Big\{ W(A_1), \ldots, W(A_k) \Big\},$$

where conv *denotes the convex hull.*

Then, to figure out the numerical range of a nilpotent matrix A without cycles, we just need to consider the numerical range of each block A_i, a nilpotent matrix without cycles and connected (that is, a tree). Thus, to establish the relation between the circularity of the numerical range of A and the existence of cycles, we will focus only on connected matrices. We start by an auxiliary and technical result.

Lemma 3.4. *Let $A = [a_{ij}]_{i,j=1}^n \in \mathcal{M}_n(\mathbb{H})$ be a nilpotent and tree matrix and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{S}_{\mathbb{R}^n}^+$. Then*

$$\bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \left\{ \sum_{i,j=1}^n \beta_i \beta_j z_i^* a_{ij} z_j \right\} = \sum_{i,j}^n \mathbb{S}_{\mathbb{H}}\Big(0, \beta_i \beta_j |a_{ij}|\Big).$$

Proof. The proof is done by induction in n. If $n = 1$, then A is the zero matrix and the result is obvious.

Now assume that

$$\bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n-1}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j \right\} = \sum_{i,j}^{n-1} \mathbb{S}_{\mathbb{H}}\Big(0, \beta_i \beta_j |a_{ij}|\Big).$$

In the tree $\mathcal{G}_A$ associated with A, pick any vertex that has only one edge. If necessary change its label to n. In this case, since $\mathcal{G}_A$ is a tree, we have that $a_{ni} = 0$, for any $i \in \{1, \ldots, n\}$ and, for some $p \in \{1, \ldots, n-1\}$, $a_{pn} \ne 0$, $a_{in} = 0$ if $i \in \{1, \ldots, n-1\}\backslash\{p\}$. We then have

$$\bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \left\{ \sum_{i,j=1}^n \beta_i \beta_j z_i^* a_{ij} z_j \right\}$$

$$= \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j + \sum_{\max\{i,j\}=n} \beta_i \beta_j z_i^* a_{ij} z_j \right\}$$

$$= \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n-1}} \bigcup_{z_n \in \mathbb{S}_{\mathbb{H}}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j + \beta_p \beta_n z_p^* a_{pn} z_n \right\}$$

$$= \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n-1}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j + \mathbb{S}_{\mathbb{H}} \Big(0, \beta_p \beta_n |a_{pn}| \Big) \right\}$$

$$= \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n-1}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j \right\} + \mathbb{S}_{\mathbb{H}} \Big(0, \beta_p \beta_n |a_{pn}| \Big)$$

$$= \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n-1}} \left\{ \sum_{i,j=1}^{n-1} \beta_i \beta_j z_i^* a_{ij} z_j \right\} + \sum_{i=1}^{n-1} \mathbb{S}_{\mathbb{H}} \Big(0, \beta_i \beta_n |a_{in}| \Big)$$

In the second and the last equality we used that the only non-zero a_{in}, for $i \in \{1, \ldots, n\}$, is a_{pn} and that all a_{ni} are zero.

By the induction hypothesis, the last equality can be written as

$$\bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \left\{ \sum_{i,j=1}^{n} \beta_i \beta_j z_i^* a_{ij} z_j \right\} = \sum_{i,j}^{n-1} \mathbb{S}_{\mathbb{H}} \Big(0, \beta_i \beta_j |a_{ij}| \Big) + \sum_{i=1}^{n-1} \mathbb{S}_{\mathbb{H}} \Big(0, \beta_i \beta_n |a_{in}| \Big)$$

$$= \sum_{i,j}^{n} \mathbb{S}_{\mathbb{H}} \Big(0, \beta_i \beta_j |a_{ij}| \Big),$$

again using that $a_{nj} = 0$ for $j \in \{1, \ldots, n\}$.

Notice that the previous calculation was carried out under the assumption that the $(n-1) \times (n-1)$ matrix had no cycles. In fact, since the initial $n \times n$ matrix A is a tree, as we mentioned before, if we eliminate a one edge vertex together with its edge, we end up with a new graph that is also a tree. $\qquad\square$

Theorem 3.5. *Let $A \in \mathcal{M}_n(\mathbb{H})$ be a nilpotent and tree matrix. Then, $W(A)$ is a circular disk with center at the origin and radius $\max_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \sum_{i,j}^{n} \beta_i \beta_j |a_{ij}|$.*

Proof. We have:

$$W(A) = \bigcup_{\boldsymbol{x} \in \mathbb{S}_{\mathbb{H}^n}} \left\{ \sum_{i,j=1}^{n} x_i^* a_{ij} x_j \right\}.$$

Each summand $x_i^* a_{ij} x_j$ may be written as

$$\beta_i \beta_j z_i^* a_{ij} z_j$$

with $\beta_i, \beta_j \ge 0$ and $z_i, z_j \in \mathbb{S}_{\mathbb{H}}$. It follows that

$$W(A) = \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \left\{ \sum_{i,j=1}^{n} \beta_i \beta_j z_i^* a_{ij} z_j \right\}.$$

From lemma 3.4, we have

$$W(A) = \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_{i,j=1}^n \mathbb{S}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|) \right\}. \tag{3.1}$$

Therefore, it remains to prove that

$$\bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_{i,j=1}^n \mathbb{S}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|) \right\} = \mathbb{D}_{\mathbb{H}}\left(0, \max_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \sum_{i,j}^n \beta_i \beta_j |a_{ij}|\right).$$

This is achieved by proving a double inclusion. If $y = \sum_{i,j=1}^n \beta_i \beta_j r_{ij}$ for a given $\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+$, where $r_{ij} \in \mathbb{S}_{\mathbb{H}}(0, |a_{ij}|)$, then

$$|y| \le \sum_{i,j=1}^n \beta_i \beta_j |a_{ij}| \le \max_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \sum_{i,j} \beta_i \beta_j |a_{ij}|.$$

For the converse inclusion, first observe that the function $f : \mathbb{S}_{\mathbb{R}^n}^+ \to \mathbb{R}$ defined by

$$f(\boldsymbol{\beta}) = \sum_{i,j=1}^n \beta_i \beta_j |a_{ij}|$$

is a continuous function on a compact set, so it has a maximum at, say, $\boldsymbol{\beta}^* \in \mathbb{S}_{\mathbb{R}^n}^+$. Since f is continuous, $f(1, 0, \ldots, 0) = 0$ and $\mathbb{S}_{\mathbb{R}^n}^+$ is connected, for every $r_0 \in [0, f(\boldsymbol{\beta}^*)]$ there exists $\boldsymbol{\beta}_0 \in \mathbb{S}_{\mathbb{R}^n}^+$ such that $f(\boldsymbol{\beta}_0) = r_0$, that is, for any path connecting $(1, 0, \ldots, 0)$ to $\boldsymbol{\beta}^*$, the values of f run surjectively over the interval $[0, f(\boldsymbol{\beta}^*)]$. Take $y \in \mathbb{D}_{\mathbb{H}}(0, f(\boldsymbol{\beta}^*))$. Then, $|y| \le f(\boldsymbol{\beta}^*)$ and $|y| = f(\boldsymbol{\beta})$, for some $\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+$. If $y = 0$, then we can take $\boldsymbol{\beta} = (1, 0, \ldots, 0)$ and the inclusion follows. Now, suppose $y \ne 0$. Let $y_{ij} = \beta_i \beta_j |a_{ij}| \dfrac{y}{|y|}$. It follows that

$$\sum_{i,j=1}^n y_{ij} = f(\boldsymbol{\beta}) \frac{y}{|y|} = y$$

and $y_{ij} \in \mathbb{S}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|)$. Therefore,

$$y \in \sum_{i,j=1}^n \mathbb{S}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|).$$

$\square$

This result can be applied to more general matrices as shown in the next example.

Example 3.6. Let $A = \begin{bmatrix} 0 & 2j & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and write A as a direct sum

$$A = A_1 \oplus A_2 = \begin{bmatrix} 0 & 2j \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 1 \end{bmatrix}.$$

By theorem 3.5 we have

$$W(A_1) = \mathbb{D}_{\mathbb{H}}(0,1) \ \text{ and } \ W(A_2) = \{1\}.$$

From proposition 3.3 it follows that

$$W(A) = \text{conv}\,\{W(A_1), W(A_2)\} = \mathbb{D}_{\mathbb{H}}(0,1).$$

The previous result, theorem 3.5, can be extended to disconnected matrices.

Corollary 3.7. *Let $A \in \mathcal{M}_n(\mathbb{H})$ be a nilpotent and cycle-free matrix. Then, $W(A)$ is a disk with center at the origin.*

Proof. There exist a permutation matrix P such that $P^\top A P = A_1 \oplus \ldots \oplus A_k$, where each A_i is a tree, square matrix. By proposition 3.3 we have

$$\begin{aligned}
W(A) = W(P^\top A P) &= W(A_1 \oplus \ldots \oplus A_k) \\
&= \text{iconv}\,\{W(A_1), ..., W(A_k) \\
&= \text{conv}\,\{W(A_1), ..., W(A_k)\}.
\end{aligned}$$

The result follows from the convexity of $W(A_i)$, see theorem 3.5.   □

We will now give an example that shows the implication of the previous result cannot be strengthened to an equivalence. We will provide a nilpotent (real) matrix A with cycles that have a circular numerical range. The existence of such matrix is supported on two results. Firstly, in theorem 3.7 of [CDM] it was shown that the quaternionic numerical range of a real matrix A is the equivalence classes of the complex numerical range of the matrix A, that is

$$W(A) = \left[W_{\mathbb{C}}(A)\right], \text{ for } A \in M_n(\mathbb{R}).$$

And secondly, theorem 1 of [CT] provide necessary and sufficient conditions for a 4×4 nilpotent complex matrix to have a circular complex numerical range.

Example 3.8. Let $A \in M_4(\mathbb{R})$ be

$$A = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This matrix satisfies both conditions of [CT] for the numerical range to be circular, therefore the complex numerical range of A is $W_{\mathbb{C}}(A) = \mathbb{D}_{\mathbb{C}}(0,r)$. Since A is real, by [CDM], $W(A) = \left[\mathbb{D}_{\mathbb{C}}(0,r)\right] = \mathbb{D}_{\mathbb{H}}(0,r)$.

4. A class of matrices with convex numerical range

So far we have dealt with nilpotent matrices. In this section we will enlarge our domain and consider matrices that have real entries in the diagonal. It is know that when $A = \alpha I + N \in \mathcal{M}_n(\mathbb{H})$, for $\alpha \in \mathbb{R}$, the numerical range is $W(\alpha I + N) = \alpha + W(N)$, see [R, proposition 3.5.4]. Therefore, if the numerical range of N is convex, so is the numerical range of A. We expand this result on convexity for the case where A can be written as the sum of real diagonal matrix D and a nilpotent cycle-free matrix N, i.e. $T = N + D$.

Let $D = diag(d_1, \ldots, d_n) \in \mathcal{M}_n(\mathbb{R})$ and define

$$\underline{d} = \min\{d_i : 1 \le i \le n\} \quad \text{and} \quad \overline{d} = \max\{d_i, 1 \le i \le n\}.$$

We will find out in theorem 4.2 that the numerical range of A can be decomposed into a union of disks $\mathbb{D}_\mathbb{H}(d, r(d))$, one for each $d \in [\underline{d}, \overline{d}]$, that is,

$$W(A) = \bigcup_{d \in [\underline{d}, \overline{d}]} \mathbb{D}_\mathbb{H}(d, r(d)).$$

To prove the above decomposition, we will need to show that the radius of each disk $r(d)$ varies continuously with the center d.

Lemma 4.1. *Let $A \in \mathcal{M}_n(\mathbb{H})$ be a matrix with real entries in the diagonal. Let*

$$f : \mathbb{R}^{n,+} \longrightarrow \mathbb{R}^+ \qquad\qquad g : \mathbb{R}^{n,+} \longrightarrow [\underline{d}, \overline{d}]$$
$$\boldsymbol{\beta} \mapsto \sum_{i \ne j} \beta_i \beta_j |a_{ij}| \qquad\qquad \boldsymbol{\beta} \mapsto \sum_i d_i \beta_i^2.$$

Then, the function $r : [\underline{d}, \overline{d}] \to \mathbb{R}^+$ defined by

$$r(d) = \max\{f(\boldsymbol{\beta}) : g(\boldsymbol{\beta}) = d \text{ and } \boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+\} \tag{4.1}$$

is continuous.

Proof. Define a correspondence $\Gamma : [\underline{d}, \overline{d}] \rightrightarrows \mathbb{S}_{\mathbb{R}^n}^+$ to be the intersection of fibers

$$\Gamma(d) = g^{-1}(d) \cap h^{-1}(0),$$

where $h : \mathbb{R}^{n,+} \to \mathbb{R}$ is $h(\boldsymbol{\beta}) = \|\boldsymbol{\beta}\| - 1$. We may rewrite function r using the correspondence Γ as follows:

$$r(d) = \max\{f(\boldsymbol{\beta}) : \boldsymbol{\beta} \in \Gamma(d)\}.$$

According to Berge's maximum theorem, see [Be, p.116], r is continuous provided that f and Γ are continuous and, for each $d \in [\underline{d}, \overline{d}]$, $\Gamma(d) \ne \emptyset$. Clearly, f is continuous, and since for each $d \in [\underline{d}, \overline{d}]$, there is a convex linear combination of $\underline{d}$ and $\overline{d}$ equal to d, $\Gamma(d)$ is nonempty.

We will now prove that Γ is continuous by showing that it is sequentially upper and lower semi-continuous. To prove upper semi-continuity, take any convergent sequence $\{\delta_k\}_k$ such that $\delta_k \to d$. We now prove that every

sequence $\boldsymbol{\beta}_k \in \Gamma(\delta_k)$ has a convergent subsequence $\{\boldsymbol{\beta}_{k_p}\}_p$ where $\boldsymbol{\beta}_{k_p} \to$ $\boldsymbol{\beta} \in \Gamma(d)$. In fact, since $\boldsymbol{\beta}_k$ is a sequence in the compact set $\mathbb{S}_{\mathbb{R}^n}^+$, it has a subsequence that converges to some $\boldsymbol{\beta}$. And since h and g are continuous it follows that $\delta_{k_p} = g(\boldsymbol{\beta}_{k_p}) \to g(\boldsymbol{\beta}) = d$ and $0 = h(\boldsymbol{\beta}_{k_p}) \to h(\boldsymbol{\beta})$. Thus $\boldsymbol{\beta} \in g^{-1}(d) \cap h^{-1}(0) = \Gamma(d)$.

A correspondence is lower semi-continuous if, for any convergent sequence $\{\delta_k\}_k \subseteq [\underline{d}, \overline{d}]$, such that $\delta_k \to d$, and any $\boldsymbol{\beta} \in \Gamma(d)$, there is a convergent sequence $\{\boldsymbol{\beta}_k\}_k$, such that $\boldsymbol{\beta}_k \in \Gamma(\delta_k)$ and $\boldsymbol{\beta}_k \to \boldsymbol{\beta}$. If $D = \alpha I$, for some $\alpha \in \mathbb{R}$, then $\underline{d} = \overline{d}$ and there is nothing to prove since the correspondence's domain is a singleton and the correspondence is trivially lower semi-continuous. When $\underline{d} < \overline{d}$ we need to find for each δ_k a vector $\boldsymbol{\beta}_k$ satisfying $\boldsymbol{\beta}_k \in \Gamma(\delta_k)$, and, the whole sequence $\{\boldsymbol{\beta}_k\}_k$, must be such that $\boldsymbol{\beta}_k \to \boldsymbol{\beta}$. When $\delta_k = d$ we choose $\boldsymbol{\beta}_k = \boldsymbol{\beta}$. When $d < \delta_k$, to find $\boldsymbol{\beta}_k$ we proceed in the following manner. The vector $\boldsymbol{\beta}$ is such that $g(\boldsymbol{\beta}) = d$, and with it define the sets:

$$P(\boldsymbol{\beta}) = \{j \in \{1, \ldots, n\} : \beta_j^2 > 0\},$$
$$D(\boldsymbol{\beta}) = \{j \in \{1, \ldots, n\} : d_j > d\},$$
$$d(\boldsymbol{\beta}) = \{j \in \{1, \ldots, n\} : d_j \leq d\}.$$

Since $d < \delta_k \leq \overline{d}$, $D(\boldsymbol{\beta}) \neq \emptyset$. On the other hand, we also have that $P(\boldsymbol{\beta}) \cap d(\boldsymbol{\beta}) \neq \emptyset$, because $g(\boldsymbol{\beta}) = d$ is a weighted average of the d_i over the indices $i \in P(\boldsymbol{\beta})$, thus we cannot have them all with strictly higher value than d. We will choose one element from each of these sets, without loss of generality, the element 1 from $D(\boldsymbol{\beta})$ and the element 2 from $P(\boldsymbol{\beta}) \cap d(\boldsymbol{\beta})$. Clearly $d_1 > d_2$. Let $0 < r^2 = \beta_1^2 + \beta_2^2 \leq 1$ and let the function $\tilde{g} : [0, 2\pi] \to \mathbb{R}$ be

$$\tilde{g}(\theta) = \sum_{i \geq 3} d_i \beta_i^2 + r^2 \sin^2(\theta) d_1 + r^2 \cos^2(\theta) d_2.$$

When $\theta_0 = \arcsin\left(\frac{\beta_1}{\sqrt{\beta_1^2 + \beta_2^2}}\right)$ then $\beta_1^2 = r^2 \sin^2(\theta_0)$ and $\beta_2^2 = r^2 \cos^2(\theta_0)$, and $\tilde{g}(\theta_0) = g(\boldsymbol{\beta})$. We have, for $d_1 > d_2$,

$$\tilde{g}(0) = \sum_{i \geq 3} d_i \beta_i^2 + r^2 d_2 = d + \beta_1^2(d_2 - d_1) \equiv \tilde{d} \leq d, \ (\beta_1^2 \geq 0),$$

$$\tilde{g}\left(\frac{\pi}{2}\right) = \sum_{i \geq 3} d_i \beta_i^2 + r^2 d_1 = d + \beta_2^2(d_1 - d_2) \equiv \hat{d} > d, \ (\beta_2^2 > 0).$$

The function $\tilde{g}$ is continuous and increasing in the interval $(0, \frac{\pi}{2})$, since $\tilde{g}'(\theta) = 2\sin(\theta)\cos(\theta)r^2(d_1 - d_2) > 0$. Therefore $\tilde{g}$ is a homeomorphism (in fact a diffeomorphism) between $[0, \frac{\pi}{2}]$ and the interval $[\tilde{d}, \hat{d}]$. In order to find a vector $\boldsymbol{\beta}_k$ such that $g(\boldsymbol{\beta}_k) = \delta_k$, k must be sufficiently large for $\delta_k \in [\tilde{d}, \hat{d}]$. Since, on one hand $\delta_k \to d$ and, in this case, $d < \delta_k$ and, on the other hand, $\tilde{d} \leq d < \hat{d}$, there is a $K \in \mathbb{N}$ such that $\delta_k \in [\tilde{d}, \hat{d}]$, for any $k > K$. Then, for $k \leq K$, we can take any $\boldsymbol{\beta}_k \in \Gamma(\delta_k)$. For $k > K$, we start by finding θ_k such that $\tilde{g}(\theta_k) = \delta_k$. Now, let $\beta_{1,k} = r\sin(\theta_k)$, $\beta_{2,k} = r\cos(\theta_k)$ and $\beta_{i,k} = \beta_i$,

for $i \in \{3, \ldots, n\}$, that is, $\beta_{1,k} = r \sin(\widetilde{g}^{-1}(\delta_k))$, $\beta_{2,k} = r \cos(\widetilde{g}^{-1}(\delta_k))$ and the remaining terms constant. We can easily verify that $g(\boldsymbol{\beta_k}) = \widetilde{g}(\theta_k) = \delta_k$. When $\{\delta_k\}_k$ converges to d, clearly $\{\theta_k\}_k$ converges to θ since $\widetilde{g}^{-1}$ is continuous. Therefore $\{\boldsymbol{\beta_k}\}_k$ converges to $\boldsymbol{\beta}$, as trigonometric functions are continuous.

When $\underline{d} \le \delta_k < d$ we proceed in a similar way, and conclude that if the set $\{\delta_k : \delta_k < d\}$ is infinite then for each element of this subsequence there is a $\boldsymbol{\beta_k} \in \Gamma(\delta_k)$ and these $\boldsymbol{\beta_k}$ converge to $\boldsymbol{\beta}$.

In conclusion, for any convergent sequence $\{\delta_k\}_k$ we can find a sequence $\boldsymbol{\beta_k} \in \mathbb{S}_{\mathbb{H}}$ such that $g(\boldsymbol{\beta_k}) = \delta_k \to d = g(\boldsymbol{\beta})$. $\qquad \square$

We are now able to show that $W(A)$ decompose as unions of disks.

Theorem 4.2. *Let $A = D + N \in \mathcal{M}_n(\mathbb{H})$, with D a diagonal matrix of real entries and N a nilpotent and tree matrix. Then, we have*

$$W(A) = \bigcup_{d \in \left[\underline{d}, \bar{d}\right]} \mathbb{D}_{\mathbb{H}}(d, r(d)),$$

where $r(d)$ is given by (4.1).

Proof. We begin by proving that $W(A) \subseteq \bigcup_{d \in \left[\underline{d}, \bar{d}\right]} \mathbb{D}_{\mathbb{H}}(d, r(d))$. We have:

$$
\begin{aligned}
W(A) &= \bigcup_{\boldsymbol{x} \in \mathbb{S}_{\mathbb{H}^n}} \boldsymbol{x}^* A \boldsymbol{x} = \bigcup_{\boldsymbol{x} \in \mathbb{S}_{\mathbb{H}^n}} \left\{ \sum_i d_i |x_i|^2 + \sum_{i<j} x_i^* a_{ij} x_j \right\} \\
&= \bigcup_{\substack{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+ \\ \substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}}} \left\{ \sum_i d_i \beta_i^2 + \sum_{i<j} \beta_i \beta_j z_i^* a_{ij} z_j \right\} \\
&= \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_i d_i \beta_i^2 + \bigcup_{\substack{z_k \in \mathbb{S}_{\mathbb{H}} \\ 1 \le k \le n}} \sum_{i<j} \beta_i \beta_j z_i^* a_{ij} z_j \right\} \\
&= \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_i d_i \beta_i^2 + \sum_{i<j} \beta_i \beta_j \mathbb{S}_{\mathbb{H}}(0, |a_{ij}|) \right\} \quad (4.2) \\
&\subseteq \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_i d_i \beta_i^2 + \sum_{i<j} \beta_i \beta_j \mathbb{D}_{\mathbb{H}}(0, |a_{ij}|) \right\} \\
&= \bigcup_{d \in \left[\underline{d}, \bar{d}\right]} d + \mathbb{D}_{\mathbb{H}}(0, r(d)) \quad (4.3) \\
&= \bigcup_{d \in \left[\underline{d}, \bar{d}\right]} \mathbb{D}_{\mathbb{H}}(d, r(d)).
\end{aligned}
$$

Equality (4.2) follows from lemma 3.4 applied to matrix N. Equality (4.3) follows from dividing the set $\mathbb{S}_{\mathbb{R}^n}^+$ into the fibers of the function $g(\boldsymbol{\beta}) =$

$\sum_i d_i \beta_i^2$, i.e.

$$\mathbb{S}_{\mathbb{R}^n}^+ = \bigcup_d \Gamma(d).$$

For each fiber $g^{-1}(d) \cap \mathbb{S}_{\mathbb{R}^n}^+ = g^{-1}(d) \cap h^{-1}(0) = \Gamma(d)$, with $\underline{d} \le d \le \overline{d}$, h and $\Gamma(d)$ defined as in lemma 4.1, we have

$$\bigcup_{\boldsymbol{\beta} \in \Gamma(d)} \left\{ \sum_i d_i \beta_i^2 + \sum_{i<j} \mathbb{D}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|) \right\} = \bigcup_{\boldsymbol{\beta} \in \Gamma(d)} \left\{ d + \sum_{i<j} \mathbb{D}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|) \right\}$$

$$= d + \bigcup_{\boldsymbol{\beta} \in \Gamma(d)} \sum_{i<j} \mathbb{D}_{\mathbb{H}}(0, \beta_i \beta_j |a_{ij}|)$$

$$= d + \bigcup_{\boldsymbol{\beta} \in \Gamma(d)} \mathbb{D}_{\mathbb{H}}\left(0, \sum_{i<j} \beta_i \beta_j |a_{ij}|\right)$$

$$= d + \mathbb{D}_{\mathbb{H}}\left(0, \max_{\boldsymbol{\beta} \in \Gamma(d)} \sum_{i<j} \beta_i \beta_j |a_{ij}|\right)$$

$$= \mathbb{D}_{\mathbb{H}}(d, r(d)).$$

Now, to prove the converse inclusion we need to consider three different cases.

Case 1: $y \in \mathbb{S}_{\mathbb{H}}(d, r(d))$. We defined in (4.1)

$$r(d) = \max\{f(\boldsymbol{\beta}) : \sum_i \beta_i^2 d_i = d, \boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+\}$$

and, in (4.2), we saw that

$$W(A) = \bigcup_{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+} \left\{ \sum_i \beta_i^2 d_i + \sum_{i<j} \beta_i \beta_j \mathbb{S}_{\mathbb{H}}(0, |a_{ij}|) \right\}.$$

Choose $\boldsymbol{\beta}^* \in g^{-1}(d) \cap \mathbb{S}_{\mathbb{R}^n}^+$ such that $r(d) = f(\boldsymbol{\beta}^*)$. To conclude that $y \in W(A)$ we will find $y_{ij} \in \beta_i^* \beta_j^* \mathbb{S}_{\mathbb{H}}(0, |a_{ij}|)$ such that $y - d = \sum_{i<j} y_{ij}$. This is precisely what we did at the end of theorem 3.5, taking this time $y_{ij} = \beta_i^* \beta_j^* |a_{ij}| \dfrac{y-d}{|y-d|}$. And we just follow the same reasoning we did there, noting that $f(\boldsymbol{\beta}^*) = r(d) = |y - d|$.

Case 2.1: $y \in \mathbb{D}_{\mathbb{H}}(d, r(d)) \backslash \mathbb{S}_{\mathbb{H}}(d, r(d))$, with $d = \underline{d}$ or $d = \overline{d}$.

Suppose $y \in \mathbb{D}_{\mathbb{H}}(\underline{d}, r(\underline{d}))$ and $|y - \underline{d}| < r(\underline{d})$. Notice that for $\boldsymbol{\beta}$ be such that $\sum \beta_i^2 d_i = \underline{d}$, then $\beta_k^2 > 0$ only for those k's for which $d_k = \underline{d}$. Assume, without loss of generality, that those k's are the first p elements in the diagonal of A, that is,

$$\{k \in \{1, \ldots, n\} : d_k = \underline{d}\} = \{1, \ldots, p\}.$$

Define

$$\mathcal{A} = \{\boldsymbol{\beta} \in \mathbb{S}_{\mathbb{R}^n}^+ : \sum \beta_i^2 d_i = \underline{d}\} = \mathbb{S}_{\mathbb{R}^p}^+ \times \{0\}^{n-p}.$$

Then, $\mathcal{A}$ is a connected set and we can take a path form the element where f vanishes (for example some vector $\boldsymbol{\beta}$ of the canonical basis) to the element where f is maximum and equal to $r(\underline{d})$, as we did in theorem 3.5. Since f is continuous, the intermediate value theorem ensures that, in such path, all the values in between 0 and $r(\underline{d})$ are taken by some element in the path. We can conclude that there exists an element $\boldsymbol{\gamma} \in \mathcal{A}$ such that $f(\boldsymbol{\gamma}) = |y - \underline{d}|$ and that $y \in \mathbb{S}(\underline{d}, f(\boldsymbol{\gamma}))$. By (4.2), $y \in W(A)$. For $d = \overline{d}$ the procedure is analogous.

Case 2.2: $y \in \mathbb{D}_{\mathbb{H}}(d, r(d)) \backslash \mathbb{S}_{\mathbb{H}}(d, r(d))$, with $\underline{d} < d < \overline{d}$. Assume that $y \notin \mathbb{D}_{\mathbb{H}}(\underline{d}, r(\underline{d}))$, otherwise recur to the previous cases. Then $|y - \underline{d}| > r(\underline{d})$. Let ρ be a function defined over the interval $[\underline{d}, \overline{d}]$ by $\rho(t) = |y - t| - r(t)$. The function ρ is continuous since the norm is continuous and, by lemma 4.1, r is continuous. Since $\rho(\underline{d}) > 0$ and $\rho(d) < 0$, continuity of ρ implies the existence of an element $\widetilde{d}$ such that $|y - \widetilde{d}| = r(\widetilde{d})$, and so $y \in \mathbb{S}_{\mathbb{H}}(\widetilde{d}, r(\widetilde{d}))$. This concludes the proof since, by (4.2), $y \in W(A)$. $\qquad \square$

The next result identify a class of upper triangular matrices that has convex numerical range.

Corollary 4.3. *Let $A = D + N \in \mathcal{M}_n(\mathbb{H})$, with D a diagonal matrix with real entries, N nilpotent and cycle-free matrix. Then, $W(A)$ is convex.*

Proof. We start assuming that N is a tree. In this case, from theorem 4.2, for any $w \in W(A) = \bigcup_{d \in [\underline{d}, \overline{d}]} \mathbb{D}_{\mathbb{H}}(d, r(d))$, there is a $d \in [\underline{d}, \overline{d}]$ such that $w \in \mathbb{D}_{\mathbb{H}}(d, r(d))$, which implies $|\pi_{\mathbb{R}}(w) - d| \leq r(d)$. Thus $d - r(d) \leq \pi_{\mathbb{R}}(w) \leq d + r(d)$ and we end up concluding that

$$\pi_{\mathbb{R}}(W(A)) \subset W(A) \cap \mathbb{R} = \left[\min_{d \in [\underline{d}, \overline{d}]} \big(d - r(d) \big), \max_{d \in [\underline{d}, \overline{d}]} \big(d - r(d) \big) \right]$$

and since $\pi_{\mathbb{R}}(W(A)) \supset W(A) \cap \mathbb{R}$, we conclude that $\pi_{\mathbb{R}}(W(A)) = W(A) \cap \mathbb{R}$. By [AY1, theorem 3] the numerical range is convex.

In the case where N is not a tree, then there exists a permutation matrix P such that $P^\top N P = N_1 \oplus \ldots \oplus N_k$, where each N_i is a tree, square matrix. Since $P^\top D P$ is still a real diagonal matrix, we have

$$
\begin{aligned}
W(A) &= W(D + N) \\
&= W(P^\top (D + N) P) = W((D_1 + N_1) \oplus \ldots \oplus (D_k + N_k)) \\
&= \text{iconv}\, \{W(D_1 + N_1), ..., W(D_k + N_k)\} \\
&= \text{conv}\, \{W(D_1 + N_1), ..., W(D_k + N_k)\}. \qquad (4.4)
\end{aligned}
$$

The result follows from proposition 3.3 and the first part of this corollary. $\quad \square$

Example 4.4. Let $A = \begin{pmatrix} d_1 & a_{12} & a_{13} \\ 0 & d_2 & 0 \\ 0 & 0 & d_2 \end{pmatrix}$, where $d_1, d_2 \in \mathbb{R}$, $d_1 \neq d_2$, and $a_{12}, a_{13} \in \mathbb{H}$. Since A can be written as $A = d_1 I + (d_2 - d_1)\tilde{A}$, with $\tilde{A} =$

$\begin{pmatrix} 0 & q_{12} & q_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, we have $W(A) = d_1 + (d_2 - d_1)W(\tilde{A})$, which means that it is enough to study $W(\tilde{A})$.

By theorem 4.2, we have that $W(\tilde{A}) = \bigcup_{d \in [0,1]} \mathbb{D}_{\mathbb{H}}(d, r(d))$, with $r(d) = \sqrt{kd(1-d)}$ and $k = |q_{12}|^2 + |q_{13}|^2$. Since $W(\tilde{A}) \cap \mathbb{C} = \bigcup_{d \in [0,1]} \mathbb{D}_{\mathbb{C}}(d, r(d))$, we will prove that this union of disks is an ellipse:

$$\bigcup_{d \in [0,1]} \mathbb{D}_{\mathbb{C}}(d, \sqrt{kd(1-d)}) = \left\{ (x,y) \in \mathbb{R}^2 : \frac{(x - \frac{1}{2})^2}{\frac{k+1}{4}} + \frac{y^2}{\frac{k}{4}} \leq 1 \right\} \equiv \mathcal{E}.$$

To show that $\mathcal{E} \subseteq \bigcup_{d \in [0,1]} \mathbb{D}_{\mathbb{C}}(d, \sqrt{kd(1-d)})$, notice that for $(a,b) \in \mathcal{E}$, $(a,b) \in \mathbb{D}_{\mathbb{C}}(d_0, \sqrt{kd_0(1-d_0)})$ with $d_0 = \frac{2a+k}{2(k+1)} \in [0,1]$.

Conversely, let $(a,b) \in \mathbb{D}_{\mathbb{C}}(d, \sqrt{kd(1-d)})$. We want to show now that $\frac{(a - \frac{1}{2})^2}{\frac{k+1}{4}} + \frac{b^2}{\frac{k}{4}} \leq 1$, i.e, $\frac{k}{k+1}(a - \frac{1}{2})^2 + b^2 - \frac{k}{4} \leq 0$. Since $b^2 = kd(1-d) - (a-d)^2$, we have that

$$\frac{k}{k+1}\left(a - \frac{1}{2}\right)^2 + b^2 - \frac{k}{4} = -\frac{1}{k+1}((k+1)d - a)^2 + \frac{k}{k+1}((k+1)d - a) - \frac{k^2}{4(k+1)}.$$

This is a second degree polynomial in $(k+1)d - a$, with down concavity and always non-positive, so we conclude that $(a,b) \in \mathcal{E}$.

5. Convexity and circularity of 3×3 nilpotent matrices

Our main goal in this section is to establish necessary and sufficient conditions for quaternionic 3×3 nilpotent matrices to have circular or convex numerical range. We start by finding out that this condition for circularity is related to the product of all non-zero elements of the matrix. In particular, it relates the numerical range's circularity with the product $a_{13}^* a_{12} a_{23}$ vanishing or not. Theorem 5.3 gives a condition for the convexity of the numerical range in terms of the values assumed by exactly the same product. Consequently, the numerical range is convex if and only if $a_{13}^* a_{12} a_{23} \in \mathbb{R}$.

Theorem 5.1. *Let $A \in \mathcal{M}_3(\mathbb{H})$ be a nilpotent matrix. Then, $W(A)$ is a disk with center at the origin, if and only if, A is cycle-free.*

Proof. Sufficiency was proved in corollary 3.7. For necessity we will show that if A has a cycle then $W(A)$ is not a disk. We just need to observe that there are some elements $q \in \mathbb{S}_{\mathbb{H}}$ for which $aq \in W(A)$, with

$$a = \max\{\beta_1\beta_2|a_{12}| + \beta_1\beta_3|a_{13}| + \beta_2\beta_3|a_{23}|\},$$

and some others $\tilde{q} \in \mathbb{S}_{\mathbb{H}}$ for which $a\tilde{q} \notin W(A)$. If $w \in W(A)$ then, by the triangle inequality, $|w| \leq a$, and the equality holds if, and only if, all the terms of $\boldsymbol{x}^* A \boldsymbol{x}$ are collinear. We have that $aq \in W(A)$ if, and only if, any

term of $\boldsymbol{x}^* A \boldsymbol{x}$ is in the real span of q, for some $q \in \mathbb{S}_{\mathbb{H}}$, that is, writing $a_{ij} = |a_{ij}| w_{ij}$ with $w_{ij} \in \mathbb{S}_{\mathbb{H}}$,

$$\begin{cases} \beta_1 \beta_2 |a_{12}| z_1^* w_{12} z_2 &= \beta_1 \beta_2 |a_{12}| q \\ \beta_1 \beta_3 |a_{13}| z_1^* w_{13} z_3 &= \beta_1 \beta_3 |a_{13}| q \\ \beta_2 \beta_3 |a_{23}| z_2^* w_{23} z_3 &= \beta_2 \beta_3 |a_{23}| q \end{cases}$$

If $a_{13} a_{12} a_{23} \neq 0$ (i.e, A has cycles) the system has solution only when $q = z_3^* w_{13}^* w_{12} w_{23} z_3$, for some $z_3 \in \mathbb{S}_{\mathbb{H}}$. That is, only for those $q \in [w_{13}^* w_{12} w_{23}]$ can we reach the maximum $aq \in W(A)$. But then $W(A)$ is not circular since for $\tilde{q} \notin [w_{13}^* w_{12} w_{23}]$, $a\tilde{q} \notin W(A)$ but $aq \in W(A)$ for all $q \in [w_{13}^* w_{12} w_{23}]$. We conclude that when A has cycles the numerical range is not circular. $\quad\square$

We will now obtain a similar result for the convexity of the numerical range of a matrix A in $\mathcal{M}_3(\mathbb{H})$, that is, $W(A)$ is convex, if and only if, $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. The argument uses two known results: [AY1, Theorem 3], which says that $W(A)$ is convex if, and only if, $W(A) \cap \mathbb{R} = \pi_{\mathbb{R}}(W(A))$, and that $W(A) \cap \mathbb{R}$ is a closed interval, see [AY1, Corollary 1]. Since the numerical range is connected [R, Theorem 3.10.7] we know that $\pi_{\mathbb{R}}(W(A)) = [m, M]$ with

$$m = \min \pi_{\mathbb{R}}(W(A)),$$
$$M = \max \pi_{\mathbb{R}}(W(A)).$$

Thus we can conclude that the numerical range is convex if, and only if, $W(A) \cap \mathbb{R} = [m, M]$. That is, $W(A)$ is convex if, and only if, there are $\boldsymbol{v}$ and $\hat{\boldsymbol{v}}$ in $\mathbb{S}_{\mathbb{H}^3}$, such that

$$\boldsymbol{v}^* A \boldsymbol{v} = M \text{ and } \hat{\boldsymbol{v}}^* A \hat{\boldsymbol{v}} = m. \tag{5.1}$$

Necessarily, there are $\boldsymbol{y}, \hat{\boldsymbol{y}} \in \mathbb{S}_{\mathbb{H}^3}$, such that $M = \pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y})$ and $m = \pi_{\mathbb{R}}(\hat{\boldsymbol{y}}^* A \hat{\boldsymbol{y}})$. The following lemma is preparatory to reach the conclusion in (5.1).

Lemma 5.2. *Let* $A \in \mathcal{M}_3(\mathbb{H})$ *be a nilpotent matrix satisfying the following condition:*

the two quaternions a_{12}, $a_{13} a_{23}^*$ *are* $\mathbb{R} -$ *linearly independent.* $\tag{5.2}$

Suppose that $\pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y}) = M$. *Then,*

$$y_1^*(a_{12} y_2 + a_{13} y_3), y_2^*(a_{12}^* y_1 + a_{23} y_3), (y_1^* a_{13} + y_2^* a_{23}) y_3 \in \mathbb{R} \setminus \{0\}$$

Proof. To prove that $\omega = y_1^*(a_{12} y_2 + a_{13} y_3) \in \mathbb{R}$ we start by writing

$$M = \pi_{\mathbb{R}}(y_1^*(a_{12} y_2 + a_{13} y_3) + y_2^* a_{23} y_3)$$
$$= \pi_{\mathbb{R}}(\omega + y_2^* a_{23} y_3).$$

There exists $z \in \mathbb{S}_{\mathbb{H}}$ such that

$$z^* \omega = |\omega| \in \mathbb{R}$$

Taking now $\tilde{\boldsymbol{y}} = (y_1 z, y_2, y_3) \in \mathbb{S}_{\mathbb{H}^3}$, we have:

$$M = \pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y}) \geq \pi_{\mathbb{R}}(\tilde{\boldsymbol{y}}^* A \tilde{\boldsymbol{y}}).$$

Then,

$$M = \pi_{\mathbb{R}}(\omega + y_2^* a_{23} y_3) \geq \pi_{\mathbb{R}}(z^* \omega + y_2^* a_{23} y_3)$$

and by $\mathbb{R}$-linearity of $\pi_{\mathbb{R}}$ we have

$$\pi_{\mathbb{R}}(\omega) \geq \pi_{\mathbb{R}}(z^* \omega) = |\omega|.$$

We conclude that $\omega = |\omega|$ and so $\omega \in \mathbb{R}$.

Now we prove that $y_1^*(a_{12} y_2 + a_{13} y_3) \neq 0$ by assuming that $y_1^*(a_{12} y_2 + a_{13} y_3) = 0$ and then finding a vector $\boldsymbol{t} \in \mathbb{S}_{\mathbb{H}^3}$ with $\boldsymbol{t}^* A \boldsymbol{t} > \pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y})$, reaching a contradiction with $\boldsymbol{y}$ being the maximizer of $\pi_{\mathbb{R}}(\boldsymbol{x}^* A \boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{S}_{\mathbb{H}^3}$. Condition (5.2) implies that $a_{23} \neq 0$ and $a \equiv a_{12} + a_{13}\dfrac{a_{23}^*}{|a_{23}|} \neq 0$. If $y_1^*(a_{12} y_2 + a_{13} y_3) = 0$ then

$$\pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y}) = y_2^* a_{23} y_3 = \frac{|a_{23}|}{2}$$

The previous equality comes from the maximum of $f(\alpha_1, \alpha_2) = \alpha_1 \alpha_2$, subject to $\alpha_1^2 + \alpha_2^2 = 1$, being $\dfrac{1}{2}$. Let $t_1 = \beta_1 \in \mathbb{R}$, $t_2 = \beta_2 \dfrac{a^*}{|a|}$ and $t_3 = \dfrac{a_{23}^*}{|a_{23}|} t_2$, with $\beta_2 = \dfrac{\sqrt{2}}{2} - \epsilon$ and $\beta_1^2 + 2\beta_2^2 = 1$.

$$\boldsymbol{t}^* A \boldsymbol{t} - \frac{|a_{23}|}{2} = \beta_1 \beta_2 |a| + \beta_2^2 |a_{23}| - \frac{|a_{23}|}{2}$$

$$= \sqrt{1 - 2\beta_2^2} \beta_2 |a| + \left(\beta_2^2 - \frac{1}{2}\right)|a_{23}|$$

$$= \sqrt{2\sqrt{2}\epsilon - 2\epsilon^2}\left(\frac{\sqrt{2}}{2} - \epsilon\right)|a| - \left(\sqrt{2} - \epsilon\right)\epsilon|a_{23}|$$

$$= \epsilon\left\{\sqrt{\frac{2\sqrt{2}}{\epsilon} - 2}\left(\frac{\sqrt{2}}{2} - \epsilon\right)|a| - \left(\sqrt{2} - \epsilon\right)|a_{23}|\right\}$$

Clearly $\dfrac{1}{\epsilon}\left(\boldsymbol{t}^* A \boldsymbol{t} - \dfrac{|a_{23}|}{2}\right) > 0$ for very small and positive ϵ. Thus, $\boldsymbol{t}^* A \boldsymbol{t} > \pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y})$, and we found a contradiction.

The second case, $y_2^*(a_{12}^* y_1 + a_{23} y_3) \in \mathbb{R} \setminus \{0\}$, follows from writing M as follows

$$\begin{aligned}
M &= \pi_{\mathbb{R}}(y_1^* a_{12} y_2 + y_1^* a_{13} y_3 + y_2^* a_{23} y_3) \\
&= \pi_{\mathbb{R}}(y_1^* a_{12} y_2) + \pi_{\mathbb{R}}(y_1^* a_{13} y_3) + \pi_{\mathbb{R}}(y_2^* a_{23} y_3) \\
&= \pi_{\mathbb{R}}(y_2^* a_{12}^* y_1) + \pi_{\mathbb{R}}(y_1^* a_{13} y_3) + \pi_{\mathbb{R}}(y_2^* a_{23} y_3) \\
&= \pi_{\mathbb{R}}\left(y_2^*(a_{12}^* y_1 + a_{23} y_3)\right) + \pi_{\mathbb{R}}(y_1^* a_{13} y_3).
\end{aligned}$$

Now we proceed as in the first case to find out that $y_2^*(a_{12}^* y_1 + a_{23} y_3) \in \mathbb{R}$. To prove that $y_2^*(a_{12}^* y_1 + a_{23} y_3) \neq 0$ we let this time $a \equiv a_{23} + a_{12}^* \dfrac{a_{13}}{|a_{13}|}$. From (5.2) we have that $a_{13} \neq 0$ and $a \neq 0$. We will find a contradiction, in the same way as in the first case, assuming that $y_2^*(a_{12}^* y_1 + a_{23} y_3) = 0$. In this case it must be that $\pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y}) = \dfrac{|a_{13}|}{2}$, and we will find a $\boldsymbol{t} \in \mathbb{S}_{\mathbb{H}^3}$ with $\boldsymbol{t}^* A \boldsymbol{t} > \dfrac{|a_{13}|}{2}$. Such $\boldsymbol{t}$ has $t_1 = \dfrac{a_{13}}{|a_{13}|} t_3$, $t_2 = \beta_2$, $t_3 = \beta_3 \dfrac{a^*}{|a|}$, with $\beta_3 = \dfrac{\sqrt{2}}{2} - \epsilon$ and $\beta_2^2 + 2\beta_3^2 = 1$. The rest of the proof proceeds just like the first case.

The proof for case 3 mimics the previous two. $\qquad\qquad\qquad\square$

Now we will state and prove a necessary and sufficient condition for a nilpotent 3×3 matrix to have convex numerical range.

Theorem 5.3. *Let $A \in \mathcal{M}_3(\mathbb{H})$ be a nilpotent matrix. Then, $W(A)$ is convex if, and only if, $a_{13}^* a_{12} a_{23} \in \mathbb{R}$.*

Proof. First, consider that $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. The case where $a_{13}^* a_{12} a_{23} = 0$ was dealt in theorem 5.1, the numerical range is circular and therefore convex. For the other cases, the matrix A is unitary equivalent to a real matrix, *i.e* there exists an unitary matrix $U \in \mathcal{M}_n(\mathbb{H})$, such that $U^* A U \in \mathcal{M}_n(\mathbb{R})$. By [CDM, theorem 3.6], we know that the numerical range of any real matrix is convex, thus $W(A) = W(U^* A U)$ is convex. For the unitary matrix U, take the diagonal matrix $diag(\rho, z_{12}^* \rho, z_{13}^* \rho)$, where $\rho \in \mathbb{S}_{\mathbb{H}}$ and $z_{ij} \in \mathbb{S}_{\mathbb{H}}$ are such that $a_{ij} = |a_{ij}| z_{ij}$. Its now a matter of simple calculations, using that $z_{13}^* z_{12} z_{23} = \pm 1$, to check that $U^* A U \in \mathcal{M}_n(\mathbb{R})$.

Now we consider the converse implication, that is, if $W(A)$ is convex then $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. If a_{12}, $a_{13} a_{23}^*$ are $\mathbb{R}$-linearly dependent we easily see that $a_{13} a_{23}^* a_{12}^* \in \mathbb{R}$ and, since $\pi_{\mathbb{R}}(ab) = \pi_{\mathbb{R}}(ba)$, then $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. Therefore, we can assume that a_{12}, $a_{13} a_{23}^*$ are $\mathbb{R}$-linearly independent and, by lemma 5.2, conclude that

$$y_1^*(a_{12} y_2 + a_{13} y_3), \, y_2^*(a_{12}^* y_1 + a_{23} y_3), \, (y_1^* a_{13} + y_2^* a_{23}) y_3 \in \mathbb{R} \setminus \{0\}. \quad (5.3)$$

Hence, for some $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} \setminus \{0\}$ we can write

$$\begin{cases} y_1 = \alpha_1(a_{12} y_2 + a_{13} y_3) \\ y_2 = \alpha_2(a_{12}^* y_1 + a_{23} y_3) \\ y_3 = \alpha_3(a_{13}^* y_1 + a_{23}^* y_2) \end{cases} \quad (5.4)$$

Substituting y_1 in the second equation, we get

$$(1 - \alpha_1 \alpha_2 |a_{12}|^2) y_2 = \alpha_2(\alpha_1 a_{12}^* a_{13} + a_{23}) y_3. \quad (5.5)$$

Suppose $1 - \alpha_1 \alpha_2 |a_{12}|^2 \neq 0$. We have

$$y_2 = r_2(\alpha_1 a_{12}^* a_{13} + a_{23}) y_3, \quad \text{where} \quad r_2 = \frac{\alpha_2}{1 - \alpha_1 \alpha_2 |a_{12}|^2}.$$

Therefore, for $\boldsymbol{y} = (y_1, y_2, y_3) \in \mathbb{S}_{\mathbb{H}^3}$ such that $M = \pi_{\mathbb{R}}(\boldsymbol{y}^* A \boldsymbol{y})$,

$$\boldsymbol{y}^* A \boldsymbol{y} = y_1^*(a_{12} y_2 + a_{13} y_3) + y_2^* a_{23} y_3$$

$$= y_1^*(a_{12} y_2 + a_{13} y_3) + r_2 |a_{23}|^2 |y_3|^2 + r_2 \alpha_1 y_3^* a_{13}^* a_{12} a_{23} y_3. \qquad (5.6)$$

Notice that the first two terms of (5.6) are real. Since $W(A)$ is convex, by [AY1, theorem 3], $M = \boldsymbol{y}^* A \boldsymbol{y} \in W(A) \cap \mathbb{R}$. Thus, the term $y_3^* a_{13}^* a_{12} a_{23} y_3$ is also real. This only happens if $a_{13}^* a_{12} a_{23} \in \mathbb{R}$ (since $y_3^* a_{13}^* a_{12} a_{23} y_3 = |y_3|^2 \frac{y_3^*}{|y_3|} a_{13}^* a_{12} a_{23} \frac{y_3}{|y_3|}$ and $\frac{y_3^*}{|y_3|} a_{13}^* a_{12} a_{23} \frac{y_3}{|y_3|} \sim a_{13}^* a_{12} a_{23}$) or $y_3 = 0$. The case $y_3 = 0$ can be ruled out because then $(y_1^* a_{13} + y_2^* a_{23}) y_3 = 0$ and this contradicts (5.3).

If $1 - \alpha_1 \alpha_2 |a_{12}|^2 = 0$ and since $y_3 \neq 0$, from (5.5) $\alpha_1 a_{12}^* a_{13} + a_{23} = 0$. It follows that $a_{13}^* a_{12} a_{23} \in \mathbb{R}$. $\qquad \square$

We finish with a simple example.

Example 5.4. Let $A = \begin{pmatrix} 0 & i & j \\ 0 & 0 & k \\ 0 & 0 & 0 \end{pmatrix}$. Since $(-j)ik = -1$, $W(A)$ is convex and noncircular.

References

[AY1] Au-Yeung, Yik Hoi. On the convexity of the numerical range in quaternionic Hilbert spaces. *Linear and Multilinear Algebra* **16** (1984), no. 1–4, 93–100. MR0768998, Zbl 0563.15016, doi: 10.1080/03081088408817611. 1386, 1389, 1398, 1400, 1403

[AY2] Au-Yeung, Yik Hoi. A short proof of a theorem on the numerical range of a normal quaternionic matrix. *Linear and Multilinear Algebra* **39** (1995), no. 3, 279–284. MR1365448, Zbl 0838.15017, doi: 10.1080/03081089508818402. 1386, 1389

[Be] Berge, Claude. Topological spaces. Including a treatment of multi-valued functions, vector spaces and convexity. *Dover Publications, Inc., Mineola, NY*, 1997. xiv +270 pp. ISBN: 0-486-69653-7. MR1464690, Zbl 0114.38602. 1394

[CDM] Carvalho, Luís; Diogo, Cristina; Mendes, Sérgio. A bridge between quaternionic and complex numerical ranges. *Linear Algebra Appl.* **581** (2019), 496–504. MR3990066, Zbl 07106862, doi: 10.1016/j.laa.2019.07.022. 1393, 1402

[CT] Chien, Mao-Ting; Tam, Bit-Shun. Circularity of the numerical range. *Linear Algebra Appl.* **201** (1994), 113–133. MR1274882, Zbl 0809.15012, doi: 10.1016/0024-3795(94)90110-4. 1387, 1393

[Di] Diestel, Reinhard. Graph theory. Fourth edition. Graduate Texts in Mathematics, 173. *Springer, Heidelberg*, 2010. xviii+437 pp. ISBN: 978-3-642-14278-9. MR2744811, Zbl 1204.05001. 1388

[Ed] Edwards, Charles H., Jr. Advanced calculus of several variables. *Dover Publications, Inc., New York*, 1994. xii+457 pp. ISBN: 0-486-68336-2. MR1319337, Zbl 0308.26002.

[GR] Gustafson, Karl E.; Rao, Duggirala K. M. Numerical range. The field of values of linear operators and matrices. Universitext. *Springer-Verlag, New York*, 1997. xiv+189 pp. ISBN: 0-387-94835-X. MR1417493, Zbl 0874.47003, doi: 10.1007/978-1-4613-8498-4.

[Ki] Kippenhahn, Rudolf. On the numerical range of a matrix. Translated from the German by Paul F. Zachlin and Michiel E. Hochstenbach. *Linear and*

Multilinear Algebra **56** (2008), no. 1–2, 185–225. MR2378310, Zbl 1137.47003, doi: 10.1080/03081080701553768. 1385, 1386

[MM] MATACHE, VALENTIN; MATACHE, MIHAELA T. When is the numerical range of a nilpotent matrix circular? *Appl. Math. Comput.* **216** (2010), no. 1, 269–275. MR2596155, Zbl 1191.15018, doi: 10.1016/j.amc.2010.01.052. 1386, 1389

[R] RODMAN, LEIBA. Topics in quaternion linear algebra. Princeton Series in Applied Mathematics. *Princeton University Press, Princeton, NJ*, 2014. xii+363 pp. ISBN: 978-0-691-16185-3. MR3241695, Zbl 1304.15004, doi: 10.23943/princeton/9780691161853.001.0001. 1385, 1387, 1388, 1394, 1400

[ST] SO, WASIN; THOMPSON, ROBERT C. Convexity of the upper complex plane part of the numerical range of a quaternionic matrix. *Linear and Multilinear Algebra* **41** (1996), no. 4, 303–365. MR1481908, Zbl 0876.15022, doi: 10.1080/03081089608818484. 1386, 1388

[STZ] SO, WASIN; THOMPSON, ROBERT C.; ZHANG, FUZHEN. The numerical range of normal matrices with quaternion entries. *Linear and Multilinear Algebra* **37** (1994), no. 1–3, 175–195. MR1313766, Zbl 0814.15026, doi: 10.1080/03081089408818320. 1386

[Zh] ZHANG, FUZHEN. Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251** (1997), 21–57. MR1421264, Zbl 0873.15008, doi: 10.1016/0024-3795(95)00543-9. 1389

(L. Carvalho) ISCTE - LISBON UNIVERSITY INSTITUTE, AV. DAS FORÇAS ARMADAS, 1649-026 LISBON, PORTUGAL
luis.carvalho@iscte-iul.pt

(C. Diogo) ISCTE - LISBON UNIVERSITY INSTITUTE, AV. DAS FORÇAS ARMADAS, 1649-026 LISBON, PORTUGAL, AND CENTER FOR MATHEMATICAL ANALYSIS, GEOMETRY, AND DYNAMICAL SYSTEMS, MATHEMATICS DEPARTMENT, INSTITUTO SUPERIOR TÉCNICO, UNIVERSIDADE DE LISBOA, AV. ROVISCO PAIS, 1049-001 LISBOA, PORTUGAL
cristina.diogo@iscte-iul.pt

(S. Mendes) ISCTE - LISBON UNIVERSITY INSTITUTE, AV. DAS FORÇAS ARMADAS, 1649-026 LISBON, PORTUGAL, AND CENTRO DE MATEMÁTICA E APLICAÇÕES, UNIVERSIDADE DA BEIRA INTERIOR, RUA MARQUÊS D'ÁVILA E BOLAMA, 6201-001 COVILHÃ, PORTUGAL
sergio.mendes@iscte-iul.pt