A NEW PERSPECTIVE ON THE QUATERNIONIC NUMERICAL RANGE OF NORMAL MATRICES

LUÍS CARVALHO, CRISTINA DIOGO, AND SÉRGIO MENDES

Abstract. A new geometric proof of a known result characterizing the quaternionic numerical range of normal matrices is proposed. Our proof can be interpreted in probabilistic terms.

1. Introduction and Preliminaries

Let $\mathbb{F} \in \{\mathbb{H}, \mathbb{C}\}$, where $\mathbb{H}$ and $\mathbb{C}$ stand, respectively, for the quaternionic and the complex fields. Let $\mathbb{S}_{\mathbb{F}^n} = \{\boldsymbol{x} \in \mathbb{F}^n : \|\boldsymbol{x}\| = 1\}$. For a square matrix A of size $n > 1$ over $\mathbb{F}$, $A \in \mathcal{M}_n(\mathbb{F})$, the set

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

is called the numerical range of A in $\mathbb{F}$. Recall that every quaternionic normal matrix is unitarily equivalent to a complex diagonal matrix $A = diag(d_k)_{k=1}^n$, where $d_k = h_k + s_k i$ ($s_k \geq 0$) are the eigenvalues of A (see [1, p.178]). Hence, we can write $A = H + Si$, where H and S are diagonal real matrices. Since the numerical range is invariant under unitary equivalence (see [2, theorem 3.5.4]), we will consider normal matrices in diagonal form. We will also assume, without loss of generality, that $h_1 = \min\{h_k : k = 1, \ldots, n\}$ and, until theorem 2.4, that $s_k > 0$.

So, Thompson and Zhang ([1, Main Theorem, p.192]) proved that the upper bild is the convex hull of the eigenvalues and certain real numbers, constructed from pairs of non-real eigenvalues, named cone vertices. The main idea of their proof was to define an optimization problem with side conditions and use Lagrange multipliers. In 1995, two independent proofs were presented by Au-Yeung [3] and Zhang [7]. In this article, we propose a new geometric proof of the same result, bearing inspiration from probability theory.

We start by characterizing the elements in the quaternionic numerical range. It is an easy exercise to show that any element in the quaternionic

Date: March 8, 2021.

2010 *Mathematics Subject Classification.* 15B33, 47A12.

Key words and phrases. quaternions, numerical range, complex matrices.

The second author was partially supported by FCT through project UID/MAT/04459/2019 and the third author was partially supported by FCT through CMA-UBI, project PEst-OE/MAT/UI0212/2020.

numerical range can be written as

$$\boldsymbol{z}^* A \boldsymbol{z} = \boldsymbol{x}^* A \boldsymbol{x} + \boldsymbol{y}^* A^* \boldsymbol{y} + \boldsymbol{x}^* (A - A^*) \boldsymbol{y} j, \tag{1.1}$$

using the decomposition $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} j \in \mathbb{H}^n$, with $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Taking into account that A can be written as $A = H + Si$, with $H = \frac{A+A^*}{2}$ and $S = \frac{A-A^*}{2i}$, it follows that $B(A) = W_{\mathbb{H}}(A) \cap \mathbb{C}$, the bild of A, is given by (see [4]),

$$B(A) = \{\boldsymbol{x}^* A \boldsymbol{x} + \boldsymbol{y}^* A^* \boldsymbol{y} : \boldsymbol{x}^* \boldsymbol{x} + \boldsymbol{y}^* \boldsymbol{y} = 1, \text{ and } \boldsymbol{x}^* S \boldsymbol{y} = 0\}.$$

For future reference denote the above conditions by:

$$\boldsymbol{x}^* \boldsymbol{x} + \boldsymbol{y}^* \boldsymbol{y} = 1 \tag{I}$$

$$\boldsymbol{x}^* S \boldsymbol{y} = 0. \tag{II}$$

Since

$$\boldsymbol{x}^* \boldsymbol{x} + \boldsymbol{y}^* \boldsymbol{y} = \sum_{i=1}^{n} |x_i|^2 + |y_i|^2 = 1 \text{ and } \boldsymbol{x}^* A \boldsymbol{x} + \boldsymbol{y}^* A^* \boldsymbol{y} = \sum_{i=1}^{n} d_i |x_i|^2 + d_i^* |y_i|^2,$$

we conclude that $B(A) \subseteq \operatorname{conv}\{\mathcal{D}, \mathcal{D}^*\}$, with $\mathcal{D} = \{d_1, \ldots, d_n\}$. As usual, $\operatorname{conv}(\mathcal{S})$ denotes the convex combination of the elements of $\mathcal{S}$. To figure out the shape of $B(A)$ we need to know which convex combinations of $\mathcal{D}$ and $\mathcal{D}^*$ can be generated by pairs $(\boldsymbol{x}, \boldsymbol{y})$ that satisfy (I) and (II). For example, $W_{\mathbb{C}}(A) = \operatorname{conv}\{\mathcal{D}\} \subseteq B(A)$, generated with $\boldsymbol{y} = \boldsymbol{0}$, and therefore (I) and (II) are trivially satisfied. Analogously, $W_{\mathbb{C}}(A^*) \subseteq B(A)$, generated with $\boldsymbol{x} = 0$.

It is useful to work instead with the upper bild $B^+(A) = B(A) \cap \mathbb{C}^+$ since this allows us to use convexity (see [5]). From equation (1.1), each element $\omega \in B^+(A)$ is a convex combination of elements $\omega_x \in W_{\mathbb{C}}(A) \subseteq \mathbb{C}^+$ (because $s_k > 0$) and $\omega_y \in W_{\mathbb{C}}(A^*) \subseteq \mathbb{C}^-$, that is, $\omega = \alpha\omega_x + (1-\alpha)\omega_y$, with $\alpha \in (0, 1]$. Since there is a real $\omega_r = \beta\omega_x + (1-\beta)\omega_y$, $0 < \beta < \alpha$, that lies in the same segment $[\omega_x, \omega_y]$, we conclude that ω is a convex combination of ω_x and ω_r. Therefore, $B^+(A) = \operatorname{conv}\{W_{\mathbb{C}}(A), B(A) \cap \mathbb{R}\}$ and, since $B(A) \cap \mathbb{R}$ is a closed interval, $B^+(A) = \operatorname{conv}\{W_{\mathbb{C}}(A), \underline{v}, \overline{v}\}$, (see theorem 2.1) with

$$\underline{v} = \min B(A) \cap \mathbb{R}, \qquad \overline{v} = \max B(A) \cap \mathbb{R}.$$

In other words, $B^+(A) = \operatorname{conv}\{d_1, \ldots, d_n, \underline{v}, \overline{v}\}$ and, in order to determine the shape of $B^+(A)$ we just need to obtain the minimum and maximum reals (we will focus on $\underline{v}$) in the numerical range. For this matter, it is important to characterize the real elements in the bild. It follows from (1.1) and the decomposition $A = H + Si$, that a further condition on $(\boldsymbol{x}, \boldsymbol{y})$ must be satisfied:

$$\boldsymbol{x}^* S \boldsymbol{x} = \boldsymbol{y}^* S \boldsymbol{y}. \tag{III}$$

Together with (I) and (II), the three conditions are necessary and sufficient for an element of the form (1.1) to belong to the real part of the numerical range. Some of these real elements are given by $c_{i,j} = [d_i, d_j^*] \cap \mathbb{R}$. That is,

$$\begin{aligned} c_{i,j} &= \alpha_{i,j} d_i + (1 - \alpha_{i,j}) d_j^* \\ &= \alpha_{i,j} h_i + (1 - \alpha_{i,j}) h_j, \end{aligned} \tag{1.2}$$

where $\alpha_{i,j} \in (0,1)$, satisfies

$$\alpha_{i,j} s_i - (1-\alpha_{i,j}) s_j = 0. \tag{1.3}$$

We now define the relevant values of $c_{i,j}$ for future use,

$$\underline{c} = \min\{c_{i,j} : i \neq j\}, \qquad \overline{c} = \max\{c_{i,j} : i \neq j\}.$$

If we take $\boldsymbol{x} = \sqrt{\alpha_{i,j}}\boldsymbol{e_i}$ and $\boldsymbol{y} = \sqrt{1-\alpha_{i,j}}\boldsymbol{e_j}$ ($\boldsymbol{e_k}$ denotes the k-th canonical unit vector) which satisfy (I) – (III), when $i \neq j$, we see that $c_{i,j} \in B(A) \cap \mathbb{R}$. Therefore, $\underline{v} \leq \underline{c}$. Notice that, since $\omega \in \operatorname{conv}\{\mathcal{D}, \mathcal{D}^*\}$ and $d_1 = h_1 + s_1 i$ with $h_1 \leq h_j, \forall j \neq 1$, we have $\underline{c} \in [d_1, d_k^*]$, for some $k \neq 1$. Above, we proved that $c_{1,j} \in W(A)$, for any $j \neq 1$. Since any other ω must be a convex combination of elements of $\mathcal{D}$ and $\mathcal{D}^*$, the only possible real value smaller than all $c_{1,j}$ would be $c_{1,1}$. However, as we will prove, $c_{1,1}$ does not belong to the numerical range, as it is obtained from a pair $(\boldsymbol{x}, \boldsymbol{y})$ which does not satisfy conditions (I) – (III).

In order to motivate our approach we now introduce some concepts in a heuristic way. A convex combination can be seen as the expected value of a probability distribution. For instance, $c_{i,j} \in [d_i, d_j^*] \cap \mathbb{R}$ is the expected value of a probability distribution over $\{d_i, d_j^*\}$, with $\alpha_{i,j}$ the probability of d_i and $1-\alpha_{i,j}$ the probability of d_j^*. The argument of our proof is partially supported on this observation. In particular, we will interpret an element from the $B(A) \cap \mathbb{R}$ as the expected value of the probability distribution $\boldsymbol{\gamma} = (\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$ over $\mathcal{D} \cup \mathcal{D}^*$, where $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}_{\mathbb{C}^{2n}}$, with $\boldsymbol{x}_{||} = (|x_1|^2, \ldots, |x_n|^2)$ and $\boldsymbol{y}_{||} = (|y_1|^2, \ldots, |y_n|^2)$. However, we will look at this probability distribution differently. Namely, we will use the probability distribution that arises from the process of first choosing randomly a pair (d_i, d_j^*), using a probability $\boldsymbol{\theta} \in \Delta(\mathcal{D} \times \mathcal{D}^*)$, and then choosing randomly one element from that pair, d_i with probability $\alpha_{i,j}$ and d_j^* with probability $1-\alpha_{i,j}$. Here, $\Delta(\mathcal{S})$ denotes the set of probability distributions over $\mathcal{S}$. This process creates a new probability distribution $\alpha(\boldsymbol{\theta})$ over $\mathcal{D} \cup \mathcal{D}^*$, i.e. $\alpha(\boldsymbol{\theta}) \in \Delta(\mathcal{D} \cup \mathcal{D}^*)$. The choice of $\boldsymbol{\theta}$ should be coherent with the initial probability $\boldsymbol{\gamma}$, in the sense that $\alpha(\boldsymbol{\theta}) = \boldsymbol{\gamma}$.

Using the law of total probability we have that the probability $\alpha(\boldsymbol{\theta})$ for the element d_i is

$$\alpha_{d_i}(\boldsymbol{\theta}) = \operatorname{prob}(d_i) = \sum_k \operatorname{prob}\big(d_i, d_k^*\big)\operatorname{prob}\big(d_i | (d_i, d_k^*)\big) = \sum_k \theta_{i,k}\alpha_{i,k}. \tag{1.4}$$

Analogously, for the element d_j^*, we have $\alpha_{d_j^*}(\boldsymbol{\theta}) = \operatorname{prob}(d_j^*) = \sum_k \theta_{k,j}\alpha_{j,k}$. In this way, we define the function $\alpha : (\mathcal{D} \cup \mathcal{D}^*) \times \Delta(\mathcal{D} \cup \mathcal{D}^*) \to [0,1]$. We say that a probability distribution $\boldsymbol{\theta} \in \Delta(\mathcal{D} \times \mathcal{D}^*)$ is *coherent* with $\boldsymbol{\gamma} \in \Delta(\mathcal{D} \cup \mathcal{D}^*)$ if $\alpha_d(\boldsymbol{\theta}) = \boldsymbol{\gamma}(d)$, for all $d \in \mathcal{D} \cup \mathcal{D}^*$. Accordingly, we define $\Theta_{\mathcal{D}}(\boldsymbol{\gamma})$ to be the set of probability distributions $\boldsymbol{\theta} \in \Delta(\mathcal{D} \times \mathcal{D}^*)$ coherent with $\boldsymbol{\gamma} \in \Delta(\mathcal{D} \cup \mathcal{D}^*)$, that is,

$$\Theta_{\mathcal{D}}(\boldsymbol{\gamma}) = \Big\{\boldsymbol{\theta} \in \Delta(\mathcal{D} \times \mathcal{D}^*) : \alpha(\boldsymbol{\theta}) = \boldsymbol{\gamma}\Big\}. \tag{1.5}$$

We will show that the set of coherent probability distributions is non empty (see lemma 2.2). Moreover, we will find out that, if $(\boldsymbol{x}, \boldsymbol{y})$ satisfies

(I) – (III), there is a coherent $\boldsymbol{\theta}$ such that $\boldsymbol{\theta}(d_1, d_1^*) = \theta_{1,1} = 0$, meaning that this distribution gives probability zero to the pair (d_1, d_1^*), see lemma 2.3. Using (I) – (III) on (1.1), a real element in the bild can be written as $\mathfrak{h} = \boldsymbol{x}^* H \boldsymbol{x} + \boldsymbol{y}^* H \boldsymbol{y}$. Rewriting $\mathfrak{h}$ as a convex combination of $c_{i,j}$'s, $\mathfrak{h} = \sum_{i,j} \theta_{i,j} c_{i,j}$, we find that $\mathfrak{h} \geq \underline{c}$, since $\theta_{1,1} = 0$. This is the content of the main result of the paper, see theorem 2.4. We finalize by observing that the ideal case where $\underline{v} = h_1$ is not attainable and the minimum is, morally speaking, the second best case.

2. Numerical range of normal matrices

We begin with a characterization of the upper bild in terms of the complex numerical range and two real values.

Theorem 2.1. *Let $A = diag(d_k)_{k=1}^n \in \mathcal{M}_n(\mathbb{C})$, with $d_k = h_k + s_k i$ and $s_k > 0$. Then,*

$$B^+(A) = \operatorname{conv}\{W_{\mathbb{C}}(A), \underline{v}, \overline{v}\}.$$

Proof. According to equation (1.1), for any $\omega \in B^+(A)$ there is $\omega_x \in W_{\mathbb{C}}(A) = \operatorname{conv}\{d_1, \ldots, d_n\}$, $\omega_y \in W_{\mathbb{C}}(A^*)$ and $\alpha \in (0, 1]$, such that $\omega = \alpha\omega_x + (1-\alpha)\omega_y$. Then, there are $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}_{\mathbb{C}^{2n}}$, satisfying (I) and (II) such that $\omega_x = \boldsymbol{x}_{\mathbb{S}}^* A \boldsymbol{x}_{\mathbb{S}}$ and $\omega_y = \boldsymbol{y}_{\mathbb{S}}^* A^* \boldsymbol{y}_{\mathbb{S}}$, where $\boldsymbol{x}_{\mathbb{S}}$ and $\boldsymbol{y}_{\mathbb{S}}$ are unitary vectors in $\mathbb{C}^n$ with $\boldsymbol{x} = \|\boldsymbol{x}\| \boldsymbol{x}_{\mathbb{S}}$ and $\boldsymbol{y} = \|\boldsymbol{y}\| \boldsymbol{y}_{\mathbb{S}}$.

On the other hand, there is $\beta \in (0, \alpha)$ such that $\omega_r = \beta\omega_x + (1-\beta)\omega_y \in \mathbb{R}$. Thus, $\omega_r \in B^+(A) \cap \mathbb{R}$, since the vector $(\sqrt{\beta}\boldsymbol{x}_{\mathbb{S}}, \sqrt{1-\beta}\boldsymbol{y}_{\mathbb{S}})$ also satisfies (I) and (II).

It is easy to see that $\omega = \psi\omega_x + (1-\psi)\omega_r$, with $\psi = \frac{\alpha-\beta}{1-\beta}$. That is, any $\omega \in B^+(A)$ also belongs to $\operatorname{conv}\{W_{\mathbb{C}}(A), W_{\mathbb{H}}(A) \cap \mathbb{R}\}$. Since, $B(A) \cap \mathbb{R} \neq \emptyset$ (see [6, Corollary 3.3]), by convexity and compactness of the upperbild, the set $B(A) \cap \mathbb{R}$ is a closed interval, and so $\omega \in \operatorname{conv}\{W_{\mathbb{C}}(A), \underline{v}, \overline{v}\}$. We then have that $B^+(A) \subseteq \operatorname{conv}\{W_{\mathbb{C}}(A), \underline{v}, \overline{v}\}$. The converse inclusion follows trivially from the convexity of the upperbild. ∎

To characterize the reals $\underline{v}$ and $\overline{v}$, we need some technical lemmas. We start by noticing that the set $\Theta_{\mathcal{D}}(\boldsymbol{\gamma})$ of probability distributions over $\mathcal{D} \times \mathcal{D}^*$ coherent with $\boldsymbol{\gamma} = (\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$ (see (1.5)) is, under mild conditions, non empty. This result is a special case of a more general lemma that we prove in the appendix.

Lemma 2.2. *For any $(\boldsymbol{x}, \boldsymbol{y})$ satisfying (III), $\Theta_{\mathcal{D}}(\boldsymbol{x}_{||}, \boldsymbol{y}_{||}) \neq \emptyset$.*

Next result shows that there exists a coherent distribution $\boldsymbol{\theta}$ with $\theta_{1,1} = 0$, that is, the probability given to the pair (d_1, d_1^*) is, in this case, zero. This means that choosing the segment joining d_1 with d_1^* is not compatible with restrictions (I) – (III).

Lemma 2.3. *Let $(\boldsymbol{x}, \boldsymbol{y})$ satisfying (I) – (III). Then, there exists $\boldsymbol{\theta} \in \Theta_{\boldsymbol{\mathcal{D}}}(\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$ with $\theta_{1,1} = 0$.*

Proof. We start by noting that

$$s_1|x_1|^2 > \sum_{i=2}^{n} s_i|y_i|^2 \text{ and } s_1|y_1|^2 > \sum_{i=2}^{n} s_i|x_i|^2, \tag{2.1}$$

implies $\boldsymbol{x}^*S\boldsymbol{y} \neq 0$.

In fact, if (2.1) holds, by Holder's and triangle's inequalities we have

$$s_1^2|x_1|^2|y_1|^2 > \Big(\sum_{i=2}^{n} s_i|y_i|^2\Big)\Big(\sum_{i=2}^{n} s_i|x_i|^2\Big) \geq \Big(\sum_{i=2}^{n} s_i|y_i||x_i|\Big)^2 \geq \Big|\sum_{i=2}^{n} s_i y_i x_i^*\Big|^2.$$

Thus, $s_1 x_1^* y_1 \neq -\sum_{i=2}^{n} s_i y_i x_i^*$, and $\boldsymbol{x}^*S\boldsymbol{y} \neq 0$. Therefore, if (II) holds, then

$$s_1|x_1|^2 \leq \sum_{i=2}^{n} s_i|y_i|^2 \text{ or } s_1|y_1|^2 \leq \sum_{i=2}^{n} s_i|x_i|^2. \tag{2.2}$$

From (III), which can be written as $\sum_{i=1}^{n} s_i(|x_i|^2 - |y_i|^2) = 0$, the two conditions in (2.2) are equivalent. We can assume that $s_1|x_1|^2 \leq \sum_{i\geq 2} s_i|y_i|^2$, and pick $K \in \{2,\dots,n\}$ such that

$$\sum_{i=2}^{K-1} s_i|y_i|^2 \leq s_1|x_1|^2 \leq \sum_{i=2}^{K} s_i|y_i|^2.$$

If $K = 2$ the left hand side is zero. For $i = 2,\dots,K-1$, set

$$\theta_{1,i} = \frac{|y_i|^2}{1-\alpha_{1,i}} \quad \text{and} \quad \theta_{1,K} = \frac{s_1|x_1|^2 - \sum_{i\geq 2}^{K-1} s_i|y_i|^2}{s_K(1-\alpha_{1,K})}.$$

All the others $\theta_{i,j}$ are zero. From (1.3) we have $s_1\alpha_{1,i} = s_i(1-\alpha_{1,i})$. Hence

$$\begin{aligned} s_1\sum_{i=2}^{K} \theta_{1,i}\alpha_{1,i} &= \sum_{i=2}^{K} \theta_{1,i}s_i(1-\alpha_{1,i}) \\ &= \sum_{i=2}^{K-1} s_i|y_i|^2 + \Big(\theta_{1,K}s_K(1-\alpha_{1,K})\Big) = s_1|x_1|^2. \end{aligned}$$

We then easily conclude that $\boldsymbol{\theta} \in \Theta_{\mathcal{D}}(\boldsymbol{\omega}_{\boldsymbol{x}})$, with

$$\boldsymbol{\omega}_{\boldsymbol{x}} = \left(|x_1|^2, 0,\dots,0,0,|y_2|^2,\dots,|y_{K-1}|^2, \frac{s_1|x_1|^2 - \sum_{i\geq 2}^{K-1} s_i|y_i|^2}{s_K}, 0,\dots,0\right).$$

We should stress that $\theta_{1,1} = 0$. Let $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega} - \boldsymbol{\omega}_{\boldsymbol{x}} \in \mathbb{R}_+^{2n}$, with $\boldsymbol{\omega} = (\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$. Notice that $\sum_{i=1}^{n} s_i(\hat{\omega}_i - \hat{\omega}_{n+i}) = 0$, since $\sum_{i=1}^{n} s_i(\omega_i - \omega_{n+i}) = \sum_{i=1}^{n} s_i(\omega_{x,i} - \omega_{x,n+i}) = 0$. Then, lemma 2.2 implies the existence of $\hat{\boldsymbol{\theta}} \in \Theta_{\mathcal{D}}(\hat{\boldsymbol{\omega}})$. Since $\alpha_{d_1}(\hat{\boldsymbol{\theta}}) = \hat{\omega}_1 = 0$, from (1.4) it follows that $\hat{\theta}_{1,1}\alpha_{1,1} = 0$. On the other hand since $s_1 > 0$, $\alpha_{1,1} = 1/2$, and thus $\hat{\theta}_{1,1} = 0$. Then $\boldsymbol{\theta}' = \boldsymbol{\theta} + \hat{\boldsymbol{\theta}} \in \Theta_{\mathcal{D}}(\boldsymbol{\omega}_{\boldsymbol{x}} + \hat{\boldsymbol{\omega}}) = \Theta_{\mathcal{D}}(\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$, is such that $\theta'_{1,1} = 0$. ∎

We may now prove the main result of the article.

Theorem 2.4. *Let $A = H + Si \in \mathcal{M}_n(\mathbb{C})$ be a normal matrix with eigenvalues d_k of the form $d_k = h_k + s_k i$, $s_k \geq 0$, $k = 1, \dots, n$. We have:*

$$\underline{v} = \underline{c} \quad \text{and} \quad \overline{v} = \overline{c}.$$

Proof. First suppose that $S > 0$. Let $\mathfrak{h} \in W_{\mathbb{H}}(A) \cap \mathbb{R}$. There is $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}_{\mathbb{C}^{2n}}$ such that $\mathfrak{h} = \boldsymbol{x}^* H \boldsymbol{x} + \boldsymbol{y}^* H \boldsymbol{y} = \sum_{i=1}^n h_i(\gamma_i + \gamma_{n+i})$, with $\boldsymbol{\gamma} = (\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$. Note that $(\boldsymbol{x}, \boldsymbol{y})$ satisfy (I)–(III).

From lemma 2.3, we can take $\boldsymbol{\theta} \in \Theta_{\mathcal{D}}(\boldsymbol{x}_{||}, \boldsymbol{y}_{||})$, with $\theta_{1,1} = 0$. Therefore, using (1.2) and the definition of $\alpha_{d_i}(\boldsymbol{\theta})$ one easily shows that $\sum_{i,j} \theta_{i,j} c_{i,j} = \mathfrak{h}$ and $\sum_{i,j} \theta_{i,j} = 1$. Since $\min\{c_{i,j} : (i,j) \neq (1,1)\} = \underline{c}$, we have $\mathfrak{h} \geq \underline{c}$. As we proved in the introduction, $\underline{c} \in W_{\mathbb{H}}(A)$. Therefore, $\underline{v} = \underline{c}$.

We will now consider the case where $S \geq 0$. Let $A = A_1 \oplus A_2$, where $A_j = H_j + S_j i$ $(j = 1, 2)$ with $S_1 = 0$ and $S_2 > 0$. Without loss of generality, we may assume $A_1 = [a_{i,j}]_{i,j=1,\dots,k}$ and $A_2 = [a_{ij}]_{i,j=k+1,\dots,n}$. It is clear that $W_{\mathbb{H}}(A_1) \subseteq \mathbb{R}$ and that

$$W_{\mathbb{H}}(A) = \Big\{ \alpha a_1 + (1-\alpha) a_2 : a_1 \in W_{\mathbb{H}}(A_1), a_2 \in W_{\mathbb{H}}(A_2), \alpha \in [0,1] \Big\}.$$

Then, if $a = \alpha a_1 + (1-\alpha) a_2 \in W_{\mathbb{H}}(A) \cap \mathbb{R}$, necessarily $a_2 \in W_{\mathbb{H}}(A_2) \cap \mathbb{R}$. Let $\underline{v}_j = \min W_{\mathbb{H}}(A_j) \cap \mathbb{R}$, $j = 1, 2$. Since $a_j \geq \underline{v}_j$ we have $\underline{v} = \min W_{\mathbb{H}}(A) \cap \mathbb{R} \geq \min\{\underline{v}_1, \underline{v}_2\}$. On the other hand, $\underline{v}_1, \underline{v}_2 \in W_{\mathbb{H}}(A)$. Therefore $\underline{v} = \min\{\underline{v}_1, \underline{v}_2\}$. We know that $\underline{v}_1$ is the smallest entry on the diagonal of A_1. From the previous case, since $S_2 > 0$, we have $\underline{v}_2 = \min\{c_{i,j} : i \neq j, i,j \in \{k+1, \dots, n\}\}$. Taking into account (1.2) and (1.3), $c_{i,j} = h_i \geq \underline{v}_1$ when $1 \leq i \leq k$ and $k+1 \leq j \leq n$, and the conclusion that $\underline{v} = \min\{c_{i,j} : i \neq j\} = \underline{c}$ follows. The proof for the maximum goes along the same lines. ∎

Appendix A.

We will now prove a technical lemma which contains lemma 2.2 as a special case.

For a given $\boldsymbol{a} \in \mathbb{R}^N$ such that $a_1 a_2 \dots a_N \neq 0$, let $J_+ = \{i : a_i > 0\}$ and $J_- = \{i : a_i < 0\}$. For each $(i,j) \in J_+ \times J_-$ define $\alpha_{i,j} \in [0,1]$ by

$$\alpha_{i,j} a_i + (1 - \alpha_{i,j}) a_j = 0.$$

As before, $\Delta(J_+ \times J_-)$ is the set of probability distributions over $J_+ \times J_-$. For a given $\boldsymbol{\theta} \in \Delta(J_+ \times J_-)$ let

$$\alpha_i(\boldsymbol{\theta}) = \begin{cases} \sum_{j \in J_-} \theta_{i,j} \alpha_{i,j}, & \text{if } i \in J_+ \\ \sum_{j \in J_+} \theta_{j,i} \alpha_{i,j}, & \text{if } i \in J_-. \end{cases}$$

Lemma A.1. *Let $\boldsymbol{\gamma} \in \mathbb{R}_+^N$ and $\boldsymbol{a} \in \mathbb{R}^N$ be such that $a_1 a_2 \dots a_N \neq 0$. If $\sum_{i=1}^N a_i \gamma_i = 0$ then there is a $\boldsymbol{\theta} \in \Delta(J_+ \times J_-)$ such that $\alpha_i(\boldsymbol{\theta}) = \gamma_i$, for any $i = 1, \dots, N$.*

Proof. We will prove this result inductively. For $N = 2$, assume without loss of generality that $a_1 > 0$. From definition of $\alpha_{i,j}$ and the hypothesis, we conclude that $\alpha_{1,2} = \frac{\gamma_1}{\gamma_1+\gamma_2}$. Then, the result holds with $\theta_{1,2} = \gamma_1 + \gamma_2$. We will now assume the result holds for $N-1$. Pick i with the smallest $|a_i\gamma_i|$ and assume, without loss of generality, that $i = N$ and a_N is positive. Since $\sum_{k=1}^N a_k\gamma_k = 0$, there is a $j \in J_-$ such that $a_j\gamma_j + a_N\gamma_N \leq 0$. Let $\hat{\boldsymbol{\theta}} \in \Delta(J_+ \times J_-)$ be such that $\hat{\theta}_{N,j}\alpha_{N,j} = \gamma_N$, and $\hat{\theta}_{p,k} = 0$, for any $(p,k) \neq (N,j)$. Clearly, $\alpha_i(\hat{\boldsymbol{\theta}}) = \hat{\gamma}_i$ with $\hat{\boldsymbol{\gamma}}$ given by $\hat{\gamma}_N = \gamma_N$, $\hat{\gamma}_j = \hat{\theta}_{N,j}(1-\alpha_{N,j})$ and $\hat{\gamma}_k = 0$, for $k \neq j, N$. From $\sum_{i=1}^N a_i\gamma_i = 0$, we get

$$\begin{aligned}
0 &= \sum_{i\neq j,N} a_i\gamma_i + \big(a_j\gamma_j + a_N\gamma_N\big)\\
&= \sum_{i\neq j,N} a_i\gamma_i + \left(a_j\big(\gamma_j - \hat{\theta}_{N,j}(1-\alpha_{N,j})\big) + \hat{\theta}_{N,j}\big(\underbrace{\alpha_{N,j}a_n + (1-\alpha_{N,j})a_j}_{=0}\big)\right)\\
&= \sum_{i\neq j,N} a_i\gamma_i + a_j\big(\gamma_j - \hat{\theta}_{N,j}(1-\alpha_{N,j})\big) = \sum_{i=1}^{N-1} a_i\gamma_i',
\end{aligned}$$

where $\gamma_i' = \gamma_i$ for $i \neq j, N$ and $\gamma_j' = (\gamma_j - \hat{\theta}_{N,j}(1-\alpha_{N,j}))$. Notice that $\gamma_j' \geq 0$ since we know that $\alpha_{N,j} = \frac{-a_j}{a_N - a_j}$ and

$$\hat{\theta}_{N,j}(1-\alpha_{N,j}) = \gamma_N\Big(\frac{1}{\alpha_{N,j}} - 1\Big) = \frac{\gamma_N a_N}{-a_j} \leq \frac{-\gamma_j a_j}{-a_j} = \gamma_j.$$

By induction hypothesis there is a $\boldsymbol{\theta'} \in \Delta\Big((J_+ \setminus \{N\}) \times J_-\Big)$ such that $\alpha_i(\boldsymbol{\theta'}) = \gamma_i'$. Extend $\boldsymbol{\theta'} \in \Delta\Big((J_+ \setminus \{N\}) \times J_-\Big)$ to $\boldsymbol{\theta''} \in \Delta(J_+ \times J_-)$ by setting $\theta''_{N,k} = 0$, for any $k \in J_-$. Thus, $\alpha_i(\boldsymbol{\theta''}) = \gamma_i''$, where $\boldsymbol{\gamma''} \in \mathbb{R}_+^N$ extends $\boldsymbol{\gamma'} \in \mathbb{R}_+^{N-1}$ putting $\gamma_N' = 0$. Let $\boldsymbol{\theta} = \boldsymbol{\theta''} + \hat{\boldsymbol{\theta}}$ and notice that$\boldsymbol{\gamma''} + \hat{\boldsymbol{\gamma}} = \boldsymbol{\gamma}$. By linearity of $\alpha_i(.)$, we have $\alpha_i(\boldsymbol{\theta''} + \hat{\boldsymbol{\theta}}) = \boldsymbol{\gamma}_i$. ■

To see that lemma 2.2 follows from the above result, it is enough to consider $\boldsymbol{\gamma} = (\boldsymbol{x}_{||}, \boldsymbol{y}_{||}) \in \mathbb{S}_{\mathbb{R}^N}$ and $\boldsymbol{a} = (\boldsymbol{s}, -\boldsymbol{s}) \in \mathbb{R}^N$ with $N = 2n$. In particular, condition (III) can be written as $\sum_{i=1}^n s_i|x_i|^2 = \sum_{j=1}^n s_j|y_j|^2$. A simple computation shows that this is equivalent to $\sum_{i=1}^N a_i\gamma_i = 0$.

References

[1] So W, Thompson R, Zhang F. The numerical range of normal matrices with quaternion entries. Linear Multilinear Algebra. 1994;37:175–195.

[2] Rodman L. Topics in Quaternion Linear Algebra, Princeton University Press; 2014.

[3] Au-Yeung Y. A short proof of a theorem on the numerical range of a normal quaternionic matrix. Linear Multilinear Algebra. 1995;39(3):279–284.

[4] Carvalho L, Diogo C, Mendes S. Quaternionic numerical range of complex matrices, preprint.

[5] So W, Thompson R. Convexity of the upper complex plane part of the numerical range of a quaternionic matrix. Linear Multilinear Algebra. 1996;41:303-365.

[6] Carvalho L, Diogo C, Mendes S. The star-center of the quaternionic numerical range. Linear Algebra Appl. 2020;603:166–185.

[7] Zhang F. On numerical range of normal matrices of quaternions. J. Mathematical and Physical Sciences. 1995;29:235–251.

Luís Carvalho, Business Research Unit (BRU-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Av. das Forças Armadas, 1649-026, Lisbon, Portugal
Email address: luis.carvalho@iscte-iul.pt

Cristina 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
Email address: cristina.diogo@iscte-iul.pt

Sérgio 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
Email address: sergio.mendes@iscte-iul.pt