

**Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig**

**Convexity of the Joint Numerical  
Range: Topological and Differential  
Geometric Viewpoints**

(revised version: June 2004)

by

*Eugene Gutkin, Edmond Jonckheere, and  
Michael Karow*

Preprint no.: 1

2003





# Convexity of the Joint Numerical Range: Topological and Differential Geometric Viewpoints

Eugene Gutkin

California Institute of Technology (Pasadena)

and

Max-Planck-Institut for Mathematics in the Sciences (Leipzig)

gutkin@mis.mpg.de

Edmond A. Jonckheere \*

Department of Electrical Engineering—Systems

University of Southern California

Los Angeles, CA 90089-2563

jonckhee@eudoxus.usc.edu

Michael Karow

Technical University of Berlin

karow@math.TU-Berlin.de

December 1, 2002

## Abstract

The purpose of this paper is to show that the joint numerical range of a  $m$ -tuple of  $n \times n$  hermitian matrices is convex whenever the largest eigenvalue of an associated family of hermitian matrices parameterized by the  $(m - 1)$ -dimensional sphere has constant multiplicity and, as a more technical condition, the union over the sphere of the largest eigenvalue eigenspaces does not fill the full  $n$ -dimensional complex

---

\*Partially supported by the National Science Foundation grant ECS-98-02594.

vector space. It is this global, as opposed to local, behavior of the eigenvalues that makes the problem essentially topological. For  $m \leq 3$ , it is shown that the set of hermitian matrices with simple eigenvalues is open and dense in the space of all hermitian matrices, from which it already follows that the numerical range is generically convex for  $m \leq 3$ . From there on, an additional argument shows that convexity always holds when  $m \leq 3$  and  $n \geq 3$ . Furthermore, our sufficient condition for convexity is in fact a criterion for stable convexity, in the sense that should the sufficient condition fail while convexity holds, the latter can be destroyed by an arbitrarily small perturbation of the data.

## 1 Introduction

In the beautiful paper, “Das algebraische Analogon zu einem Satze von Fejér” (Math. Zeitschrift **2** (1918), 187–197), O. Töplitz introduced and studied the *numerical range* of a complex matrix. If  $C$  is a  $n \times n$  matrix, its numerical range  $\mathcal{F}(C)$  is the set of complex numbers of the form  $z^*Cz$ , where  $z$  is a  $n$ -tuple of unit norm. Töplitz proved, among other things, that the *outer boundary* of the compactum  $\mathcal{F}(C)$  is a convex curve. He conjectured that the numerical range itself was convex, and shortly after, in another beautiful paper, F. Hausdorff proved it. (See F. Hausdorff, “Der Wertvorrat einer Bilinearform”, Math. Zeitschrift **3** (1919), 314–316.) This result, which carries the name of Töplitz-Hausdorff theorem [12, 9], launched the thriving subject of numerical range. Its vitality is due, in particular, to the many extensions of Töplitz’ original setting.

An especially natural extension is the *joint numerical range* of a collection of hermitian matrices. Let  $A = (A_1, \dots, A_m)^*$  be hermitian  $n \times n$  matrices. Their joint numerical range,  $\mathcal{F}(A)$ , is the set of vectors in  $\mathbb{R}^m$  of the form  $v = (z^*A_1z, \dots, z^*A_mz)^*$ , where  $z$  is a unit vector in the complex space of  $n$  dimensions. In view of the representation  $C = A_1 + jA_2$ , the set  $\mathcal{F}(C)$  is the joint numerical range of  $(A_1, A_2)^*$ . Already in 1918, Töplitz and Hausdorff knew that the joint numerical range is not, in general, convex. Töplitz in his paper pointed out that the convexity fails if  $A_1, \dots, A_{n^2}$  is a basis of the vector space  $\mathcal{H}(n)$  of hermitian  $n \times n$  matrices. Hausdorff observed that Töplitz’ idea and the result of his own paper combine to prove the convexity of the outer boundary of the joint numerical range of any triple of hermitian matrices.

Applications of the subject of numerical range to robust control theory [7, 6, 9, 26, 31] gave a powerful impetus to the mathematical investigation of the joint numerical range for arbitrary  $m$ -tuples of hermitian matrices. The

robust stability of a feedback system consisting of  $n$  loops and  $m$  block uncertainties involves the joint numerical range of an associated collection of  $m$  hermitian  $n \times n$  matrices [7]. There is a vast mathematical literature on the subject of convexity, or the lack thereof, of the joint numerical range. Below we will briefly survey the main points.

The discussion in Töplitz' and Hausdorff's papers implies that: a) The joint numerical range,  $\mathcal{F}$ , of a triple  $(A_1, A_2, A_3)^*$  of hermitian  $2 \times 2$  matrices is typically not convex; b) For any triple of hermitian  $n \times n$  matrices, the outer boundary of  $\mathcal{F}$  is convex. Let now  $n > 2$ . Binding [1] and Fan and Tits [8] proved the convexity of  $\mathcal{F}(A_1, A_2, A_3)$ . (See § 2.2 below where the methods of these papers are recast in our setup. In § 5.1, we will re-establish this result as a byproduct of our approach; see Theorem 5.4.) The situation becomes drastically different as we move on to the 4-tuples of hermitian matrices, and more generally, to the joint ranges  $\mathcal{F}(A_1, \dots, A_m)$ , where  $m \geq 4$ . In this case, the joint numerical range  $\mathcal{F}(A_1, A_2, A_3, \dots, A_m)$  is, typically, not convex. (See Examples 3, 4, and Proposition 2.10 in § 2.2.) In view of these counterexamples, the emphasis in the study of the joint numerical range of  $m$ -tuples,  $m > 3$ , of hermitian matrices was redirected towards: a) The study of conditions ensuring that  $\mathcal{F}(A_1, \dots, A_m)$  is convex; b) The study of the outer boundary and the convex hull of  $\mathcal{F}(A_1, \dots, A_m)$ . We refer to [27, 28] and the references cited therein for more recent trends and developments.

The major result of this paper is that the joint numerical range of  $m$  hermitian matrices is convex if the largest eigenvalue of the family of hermitian matrices,  $A(\eta) := \sum_i \eta_i A_i$ ,  $\eta \in S^{m-1}$ , parameterized by the unit sphere in dimension  $m-1$  has constant multiplicity (along with another more technical condition) and that the property of simple eigenvalue for the entire family is open and dense if  $m \leq 3$ . The *global*, as opposed to local (see [6, Sec. 5]), multiplicity behavior of the eigenvalues is clearly a differential/algebraic topological issue. Thus, from the first standpoint, we continue and considerably extend the material of the publication [19], which introduced the differential topological approach in the special case of the numerical range of a complex matrix. Alongside the differential topology, we widely use algebraic topological methods, in particular the theory of fiber bundles to describe the relationship between a constant dimensional eigenspace of  $A(\eta)$  and  $\eta$ . The topological approach to the convexity of the joint numerical range goes back to [1]. The genericity issue, which is here strengthened to openness and density, goes back to [24].

We will now discuss the contents of the paper in some detail. In § 2, we establish the setting and the basic properties of the numerical range. In particular, in § 2.1, we introduce our approach to the joint numerical range as the range of a real analytic map defined on a complex projective space. In § 2.2, we motivate our approach with simple, but essential, examples. Also, we prove a few propositions that will be crucially used in the body of the paper; see, in particular, Proposition 2.11 and Corollary 2.12.

§ 3 consists of two subsections. In § 3.1, we study the convex hull of a compactum in the euclidean space from the viewpoint of *support functions*. This material is still preliminary; see, for instance, [14]. The differentiability of a support function plays an important role in our approach.

The body of the paper starts in § 3.2. From there on, we specialize our analysis to the joint numerical range,  $\mathcal{F}(A)$ , of a  $m$ -tuple  $A$  of hermitian matrices. It is in this part that the family  $A(\eta)$  is introduced. We show that the support function of  $\mathcal{F}(A)$  is the highest eigenvalue of  $A(\eta)$ . Thus, our investigation of convexity and the related properties of the joint numerical range hinges on the study of eigenvalues of certain families of hermitian matrices. Under the crucial assumption that the family in question has a *block of eigenvalues of constant multiplicity* (see Proposition 3.10), we carry over this study to § 4. Let  $A(\eta)$  satisfy the assumption, and let  $\mu$  be the multiplicity. In § 4.2 we associate with the numerical range  $\mathcal{F}(A)$  a fiber bundle over the unit sphere of  $m - 1$  dimensions whose fiber is the unit sphere in the complex  $\mu$ -dimensional eigenspace. See Theorem 4.5. In order to use the results of § 4.2 to study the convexity of numerical ranges, we investigate in § 4.3 the issue of the multiplicity of eigenvalues of a linear family of hermitian matrices. It turns out that, for  $m < 4$ , all eigenvalues of  $A(\eta)$ ,  $\eta \in S^{m-1}$ , are simple, generically. See Proposition 4.10.

Theorem 5.1 in § 5 is the main result of the paper. It says, essentially, that if the highest eigenvalue of  $A(\eta)$ ,  $\eta \in S^{m-1}$ , has constant multiplicity, then the numerical range  $\mathcal{F}(A)$  is convex. The additional technical assumption of Theorem 5.1 is automatically satisfied unless  $m = n + 1$ , and the highest eigenvalue has multiplicity  $n/2$ . From the latter and the essentially topological fact of genericity, we recover as a particular case the known result that the numerical range of any triple of  $n \times n$ ,  $n \geq 3$ , matrices is convex. See Theorem 5.4. In § 5.2 we show that Theorem 5.1 actually yields a criterion of *stable convexity*. Namely, if  $A$  does not satisfy the constant multiplicity assumption, but  $\mathcal{F}(A)$  is nevertheless convex, then the convexity can be destroyed by an arbitrarily small perturbation of  $A$ . See Theorem 5.6.

The remaining two sections are the Appendices. There we prove two important technical theorems that we have crucially used in the body of the paper. See especially the proof of Theorem 3.7 in § 7.

## 2 Preliminaries and the setting

The notation is standard. By  $\mathbb{N}$ ,  $\mathbb{R}$  and  $\mathbb{C}$  we denote the set of natural numbers, real numbers and complex numbers respectively. By  $\mathbb{F}^{n \times m}$  we denote the space of  $n \times m$ -matrices with entries in  $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ . If  $A \in \mathbb{F}^{n \times m}$ , then  $A^T$  (resp.  $A^*$ ) denotes its (resp. conjugate) transpose, and  $A^\dagger$  stands for its generalized inverse in the sense of Moore-Penrose [30]. We denote by  $\|\cdot\|$  the Euclidean vector norm in  $\mathbb{C}^n$  unless otherwise stated. By  $S^{m-1}$  we denote the unit sphere in  $\mathbb{R}^m$ . If  $U$  is a subspace of  $\mathbb{C}^n$  or  $\mathbb{R}^n$  then  $U^\perp$  denotes its orthogonal complement with respect to the standard inner product. By  $\mathbb{C}\mathbb{P}^{n-1}$  we denote the projective space of  $\mathbb{C}^n$ . We use the notation  $[z] \in \mathbb{C}\mathbb{P}^{n-1}$  for the element defined by  $z \in \mathbb{C}^n \setminus \{0\}$ .

By  $\mathcal{H}(n)$  we denote the real vector space of hermitian  $n \times n$ -matrices,  $\dim \mathcal{H}(n) = n^2$ . If  $A \in \mathcal{H}(n)$  then  $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$  are its eigenvalues, and  $E_k(A)$ ,  $1 \leq k \leq n$ , are the corresponding eigenspaces. Note that  $E_i(A) = E_j(A)$  if  $\lambda_i(A) = \lambda_j(A)$ .

We will use the terms  $\mathcal{C}^r$ -manifold,  $\mathcal{C}^r$ -mapping, etc for any  $r \in \mathbb{N} \cup \{\infty, \omega\}$ . Let  $M, N$  be  $\mathcal{C}^r$ -manifolds, and let  $f : M \rightarrow N$  be a  $\mathcal{C}^r$ -map. Let  $x \in M$ . Then  $T_x M$  denotes the tangent space of  $M$  at  $x$ , and  $d_x f : T_x M \rightarrow T_{f(x)} N$  is the differential. We will denote by  $d_x^2 g : T_x M \times T_x M \rightarrow \mathbb{R}$  the second differential, whenever it is defined. Note that  $d_x^2 g$  is a symmetric bilinear form.

### 2.1 Basic properties of the joint numerical range

We introduce the main object of study.

**Definition 2.1** *Let  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(n)^m$  be an  $m$ -tuple of hermitian matrices. Set*

$$\mathcal{F}(A) = \mathcal{F}(A_1, \dots, A_m) = \left\{ (z^* A_1 z, \dots, z^* A_m z)^T \mid z \in \mathbb{C}^n, \|z\| = 1 \right\}.$$

*Then  $\mathcal{F}(A) \subset \mathbb{R}^n$  is the joint numerical range of matrices  $A_1, \dots, A_m$ .*

We will also say that  $\mathcal{F}(A)$  is the numerical range of  $A$ . Note that for any unitary matrix  $U \in \mathbb{C}^{n \times n}$  we have

$$\mathcal{F}(A_1, \dots, A_m) = \mathcal{F}(U^* A_1 U, \dots, U^* A_m U). \quad (1)$$

The formula

$$F_A([z]) = \left( \frac{z^* A_1 z}{\|z\|^2}, \dots, \frac{z^* A_m z}{\|z\|^2} \right)^T.$$

defines a real analytic mapping  $F_A : \mathbb{C}\mathbb{P}^{n-1} \rightarrow \mathbb{R}^m$ , and the compact, connected set  $\mathcal{F}(A)$  is the range of  $F_A$ . If  $m = 1$  then  $\mathcal{F}(A)$  is a classical object.

**Proposition 2.2** *Let  $A \in \mathcal{H}(n)$ . Then  $\mathcal{F}(A) = [\lambda_n(A), \lambda_1(A)]$ .*

We will recall the basic general properties of  $\mathcal{F}(A)$  and  $F_A$ . To this end we introduce the following notation. Let  $M = [\mu_{ik}] \in \mathbb{R}^{p \times m}$  and  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(n)^m$ . Set

$$MA = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{p1} & \cdots & \mu_{pm} \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^m \mu_{1k} A_k \\ \vdots \\ \sum_{k=1}^m \mu_{pk} A_k \end{bmatrix} \in \mathcal{H}(n)^p.$$

**Lemma 2.3** *Let  $A \in \mathcal{H}(n)^m$ . Let  $M \in \mathbb{R}^{p \times m}$  be an arbitrary matrix viewed as a mapping,  $M : \mathbb{R}^m \rightarrow \mathbb{R}^p$ . Then  $M \circ F_A = F_{MA}$ .*

Despite its simplicity, Lemma 2.3 yields important consequences.

**Corollary 2.4** *1. Let  $A \in \mathcal{H}(n)^m$  and  $M \in \mathbb{R}^{p \times m}$ . Then  $\mathcal{F}(MA) = M\mathcal{F}(A)$ .  
2. Let  $A \in \mathcal{H}(n)^m, B \in \mathcal{H}(n)^r$  be such that  $\text{span}(A_1, \dots, A_m) = \text{span}(B_1, \dots, B_r)$ . Then there exist linear maps  $\mathbb{R}^m \xrightarrow{\phi} \mathbb{R}^r \xrightarrow{\psi} \mathbb{R}^m$  such that  $\phi(\mathcal{F}(A)) = \mathcal{F}(B)$ ,  $\psi(\mathcal{F}(B)) = \mathcal{F}(A)$ , and  $(\psi \circ \phi)|_{\mathcal{F}(A)} = (\phi \circ \psi)|_{\mathcal{F}(B)} = \text{id}$ .*

**Proof:** Claim 1 is immediate from Lemma 2.3. Under the assumptions of claim 2, there exist matrices  $\phi, \psi$  such that  $B = \phi A$  and  $A = \psi B$ . They satisfy the requirements.  $\square$

**Remark 1.** Let  $A, B$  satisfy the assumptions of claim 2 above. Then  $\mathcal{F}(A)$  and  $\mathcal{F}(B)$  are affinely equivalent. In particular, one of them is convex if and only if the other one is.

**Corollary 2.5** *Let  $A \in \mathcal{H}(n)^m$ , let  $\eta \in \mathbb{R}^m$  be a nonzero vector, and let  $c \in \mathbb{R}$ . Set  $H = \{y \in \mathbb{R}^m \mid \eta^T y = c\}$ . Then: 1. We have  $\eta^T \circ F_A = F_{\eta^T A}$ ; 2. We have  $\{\eta^T y \mid y \in \mathcal{F}(A)\} = [\lambda_n(\eta^T A), \lambda_1(\eta^T A)]$ ; 3. The inclusion  $\mathcal{F}(A) \subset H$  holds if and only if  $\eta^T A = cI_n$ .*

**Proof:** The first claim is a special case of Lemma 2.3. Combining it with Proposition 2.2, we obtain the second. It implies the third.  $\square$

If  $K \subseteq \mathbb{R}^m$ , we denote by  $\text{aff}(K)$  its affine hull.



**Proposition 2.6** Let  $A \in \mathcal{H}(n)^m$ . Denote by  $V_A \subset \mathbb{R}^m$  the subspace defined by  $V_A = \{ \eta \in \mathbb{R}^m \mid \eta^T A \in \mathbb{R} I_n \}$ . Then:

1. The set  $\mathcal{F}(A)$  is a singleton if and only if  $V_A = \mathbb{R}^m$ ;
2. We have  $V_A = \{0\}$  if and only if  $I_n, A_1, \dots, A_m$  are linearly independent if and only if  $\text{aff}(\mathcal{F}(A)) = \mathbb{R}^m$ .
3. Let  $\ell = \text{codim} V_A$ . Suppose that  $\ell \neq 0, m$ . Let  $Q = (\eta_1, \dots, \eta_m)$  be an orthonormal basis of  $\mathbb{R}^m$  such that  $(\eta_{\ell+1}, \dots, \eta_m)$  is a basis of  $V_A$ . Then  $\eta_j^T A = c_j I_n$  for  $\ell + 1 \leq j \leq m$ . Set  $B_j = \eta_j^T A$  for  $1 \leq j \leq \ell$ , and let  $B = (B_1, \dots, B_\ell)^*$ . Define the affine mapping  $\alpha : \mathbb{R}^\ell \rightarrow \mathbb{R}^m$  by  $\alpha(x) = Q(x^T, c_{\ell+1}, \dots, c_m)^T$ . Then  $I_n, B_1, \dots, B_\ell$  are linearly independent,  $\mathcal{F}(A) = \alpha(\mathcal{F}(B))$  and

$$\text{aff}(\mathcal{F}(A)) = \alpha(\mathbb{R}^\ell) = \{y \in \mathbb{R}^m \mid \eta_j^T y = c_j, \ell + 1 \leq j \leq m\}.$$

4. We have  $\dim(\text{aff}(\mathcal{F}(A))) = \ell$ .

**Proof:** Claim 1 is obvious, as well as the former equivalence in claim 2, while the latter is immediate from Corollary 2.5. Since  $V_B = \{0\}$ , by claim 2,  $I_n, B_1, \dots, B_\ell$  are linearly independent. The definition of  $Q$  implies that  $Q^T A = (B_1, \dots, B_\ell, c_{\ell+1} I_n, \dots, c_m I_n)^T$ . Hence, for any unit vector  $z \in \mathbb{C}^n$  we have  $Q^T F_A([z]) = F_{Q^T A}([z]) = (z^* B_1 z, \dots, z^* B_\ell z, c_{\ell+1}, \dots, c_m)^T$ , implying  $Q^T \mathcal{F}(A) = \alpha(\mathcal{F}(B))$ . We leave the rest to the reader.  $\square$

Let  $\mathcal{K}_m$  be the metric space of nonempty compact subsets of  $\mathbb{R}^m$ , endowed with the Hausdorff metric. The formula  $A \mapsto \mathcal{F}(A)$  defines a mapping  $\mathcal{F} : \mathcal{H}(n)^m \rightarrow \mathcal{K}_m$ . Recall that a mapping  $f : X \rightarrow Y$  of metric spaces is Lipschitz if there exists  $c \geq 0$  such that for any  $x, x' \in X$  we have  $d(f(x), f(x')) \leq c d(x, x')$ . This notion depends only on the equivalence classes of the metrics. Any norm on the vector space  $\mathcal{H}(n)^m$  induces a metric on it. All these metrics are equivalent.

**Proposition 2.7** The mapping  $\mathcal{F} : \mathcal{H}(n)^m \rightarrow \mathcal{K}_m$  is Lipschitz with respect to the natural metrics.

**Proof:** Let  $B = (B_1, \dots, B_{n^2})^*$  be a basis of  $\mathcal{H}(n)$ . Then for each  $A \in \mathcal{H}(n)^m$  there is a unique matrix  $M_A \in \mathbb{R}^{m \times n^2}$  such that  $A = M_A B$ . The map  $A \mapsto M_A$  is linear. Define a norm,  $\nu$ , on  $\mathcal{H}(n)^m$  by  $\nu(A) = \|M_A\|$ . Set  $c = \max_{y \in \mathcal{F}(B)} \|y\|$ . Then for all  $A, A' \in \mathcal{H}(n)^m$  and all  $[z] \in \mathbb{C}\mathbb{P}^{n-1}$ ,

$$\|F_A([z]) - F_{A'}([z])\| = \|M_{A-A'} F_B([z])\| \leq c \nu(A - A').$$

The claim follows.  $\square$

For  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(n_1)^m$  and  $B = (B_1, \dots, B_m)^* \in \mathcal{H}(n_2)^m$  let  $A \oplus B = (A_1 \oplus B_1, \dots, A_m \oplus B_m)^* \in \mathcal{H}(n_1 + n_2)^m$ . For  $X, Y \subseteq \mathbb{R}^m$  we set

$$\text{co}(X, Y) = \{ \alpha_1 x + \alpha_2 y \mid x \in X, y \in Y, \alpha_k \geq 0, \alpha_1 + \alpha_2 = 1 \}.$$

The set  $\text{co}(X, Y)$  is not convex, in general. If  $X$  and  $Y$  are convex, then  $\text{co}(X, Y) = \text{co}(X \cup Y)$ , where  $\text{co}(\cdot)$  denotes the convex hull of a set.

**Proposition 2.8** *Let  $A \in \mathcal{H}(n_1)^m$  and  $B \in \mathcal{H}(n_2)^m$ . Then  $\mathcal{F}(A \oplus B) = \text{co}(\mathcal{F}(A), \mathcal{F}(B))$ .*

**Proof:** A unit vector  $z \in \mathbb{C}^{n_1+n_2}$  can be written in the form  $z = [\sqrt{\alpha_1} z_1^T, \sqrt{\alpha_2} z_2^T]^T$ , where  $z_k \in \mathbb{C}^{n_k}$  are unit vectors and  $\alpha_k \geq 0, \alpha_1 + \alpha_2 = 1$ . We have  $F_{A \oplus B}([z]) = \alpha_1 F_A([z_1]) + \alpha_2 F_B([z_2])$ .  $\square$

The following is an immediate consequence of Proposition 2.8.

**Corollary 2.9** *Let  $A \in \mathcal{H}(n_1)^m, B \in \mathcal{H}(n_2)^m$ . Then*

- a) *If  $\mathcal{F}(A)$  and  $\mathcal{F}(B)$  are convex, then  $\mathcal{F}(A \oplus B)$  is convex;*
- b) *If  $\mathcal{F}(A) = \mathcal{F}(B)$ , then  $\mathcal{F}(A \oplus B)$  is convex.*

## 2.2 Basic examples

In this subsection we present a few examples of the (joint) numerical ranges. They demonstrate the difficulties and the pitfalls of the subject. In what follows, the meaning of the parameters  $n, m$  comes from the notation  $\mathcal{H}(n)^m$ .

Let  $A = (A_1, A_2)^* \in \mathcal{H}(n)^2$ . Introducing  $\mathcal{A} = A_1 + j A_2$ , we identify  $\mathcal{F}(A)$  with the classical numerical range of  $\mathcal{A}$ . The celebrated Töplitz-Hausdorff theorem [12] yields, in particular, that the joint numerical range of any two hermitian matrices is convex. See [19, 13] for the differential geometry of the map  $F_A$  in this case.

Let  $m$  be arbitrary. Suppose that  $A_1, \dots, A_m \in \mathcal{H}(n)$  commute. Simultaneously diagonalizing  $A_1, \dots, A_m$  by a unitary matrix, we obtain that  $\mathcal{F}(A)$  is a convex polytope. The converse also holds: If  $\mathcal{F}(A)$  is a polytope then the matrices  $A_1, \dots, A_m$  commute [2].

**Example 1.** We will now consider a specific example with  $m = 3$ . Let

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

be the Pauli spin-matrices. They form an orthonormal basis in the space of traceless matrices in  $\mathcal{H}(2)$ . Let  $\mu \geq 1$ . For  $1 \leq k \leq 3$  set  $A_k = \sigma_k \otimes I_d$ .

Let  $A = (A_1, A_2, A_3)^* \in \mathcal{H}(2\mu)^3$ . Let  $z = (z_1, z_2) \in \mathbb{C}^{2\mu}$ , where  $\|z\|^2 = \|z_1\|^2 + \|z_2\|^2$ . Let  $S^2 \subset B^3 \subset \mathbb{R}^3$  be the unit sphere and the unit ball. Since

$$F_A([z]) = \begin{bmatrix} z^* A_1 z \\ z^* A_2 z \\ z^* A_3 z \end{bmatrix} = \begin{bmatrix} 2 \Re(z_1^* z_2) \\ 2 \Im(z_1^* z_2) \\ \|z_1\|^2 - \|z_2\|^2 \end{bmatrix}$$

we have  $\|F_A([z])\|^2 = 4|z_1^* z_2|^2 + (\|z_1\|^2 - \|z_2\|^2)^2 \leq \|z\|^4 = 1$ . We will show that  $\mathcal{F}(A) = S^2$  if  $\mu = 1$ , and  $\mathcal{F}(A) = B^3$  if  $\mu \geq 2$ . Set  $F_A([z]) = (\rho \cos \phi, \rho \sin \phi, r)^T$ . It suffices to find a solution  $z = (z_1, z_2) \in \mathbb{C}^{2\mu}$  of this equation for any  $r \in [-1, 1], \phi \in \mathbb{R}, \rho = \sqrt{1-r^2}$  if  $\mu = 1$ , and  $0 \leq \rho \leq \sqrt{1-r^2}$  if  $\mu > 1$ . For  $\mu = 1$  set  $z_1 = \sqrt{(1+r)/2}, z_2 = \sqrt{(1-r)/2}e^{j\phi}$ . For  $\mu > 1$  set  $z_1 = \sqrt{(1+r)/2}v_1, z_2 = \sqrt{(1-r)/2}e^{j\phi}v_2$ , where  $v_1, v_2 \in \mathbb{C}^\mu$  are unit vectors, such that  $v_1^* v_2 = 0$ , if  $|r| = 1$ , and  $v_1^* v_2 = \rho/\sqrt{1-r^2}$  otherwise.

**Example 2 [1, 8, 6].** Let now  $n = 2$ , and  $m$  arbitrary. Let  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(2)^m$ . For  $1 \leq k \leq m$  set  $A_k = \begin{bmatrix} a_k & \overline{w_k} \\ w_k & b_k \end{bmatrix}$ , where  $a_k, b_k \in \mathbb{R}, w_k = x_k + jy_k \in \mathbb{C}$ . Define

$$M = \begin{bmatrix} x_1 & y_1 & \frac{a_1-b_1}{2} \\ \vdots & \vdots & \vdots \\ x_m & y_m & \frac{a_m-b_m}{2} \end{bmatrix} \in \mathbb{R}^{m \times 3}, \quad p = \frac{1}{2} \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_m + b_m \end{bmatrix} \in \mathbb{R}^m. \quad (2)$$

Let  $\sigma = (\sigma_1, \sigma_2, \sigma_3)^*$  be as in Example 1. Then

$$A = \begin{bmatrix} M & p \end{bmatrix} \begin{bmatrix} \sigma \\ I_2 \end{bmatrix} = M \sigma + p I_2.$$

Let  $\alpha : \mathbb{R}^3 \rightarrow \mathbb{R}^m$  be the affine mapping given by  $\alpha(\xi) = M\xi + p$ . Then, by Corollary 2.4 and the preceding example

$$\mathcal{F}(A) = M\mathcal{F}(\sigma) + p = \alpha(\mathcal{F}(\sigma)) = \alpha(S^2). \quad (3)$$

In what follows we do not distinguish between an ellipsoid in  $\mathbb{R}^3$  (resp. ellipse in  $\mathbb{R}^2$ ) and its image under an isometry  $i : \mathbb{R}^3 \rightarrow \mathbb{R}^m$  (resp.  $i : \mathbb{R}^2 \rightarrow \mathbb{R}^m$ ). Let  $M = U \text{diag}(s_1, s_2, s_3)V$  be a singular value decomposition. Since  $U \in \mathbb{R}^{m \times 3}$  and  $V \in \mathbb{R}^{3 \times 3}$  are isometries, we obtain the following classification.

- a) If  $\text{rank } M = 3$ , then  $\mathcal{F}(A)$  is an ellipsoid with semi-axes  $s_1 \geq s_2 \geq s_3$ .
- b) If  $\text{rank } M = 2$ , then  $\mathcal{F}(A)$  is a solid ellipse with semi-axes  $s_1, s_2$ .
- c) If  $\text{rank } M = 1$ , then  $\mathcal{F}(A)$  is a segment of length  $s_1$ .
- d) If  $M = 0$ , then  $\mathcal{F}(A)$  is a point.

In particular,  $\mathcal{F}(A)$  is convex if and only if  $\text{rank } M < 3$ . If  $m \geq 3$ , then  $\text{rank } M = 3$ , generically, and  $\mathcal{F}(A)$  is nonconvex.

**Example 3.** In this example  $n = m = 4$ . For  $k = 1, 2, 3$  set  $A_k = \sigma_k \oplus \sigma_k$ , and let  $A_4 = I_2 \oplus (-I_2)$ . By Proposition 2.8 and Example 1

$$\mathcal{F}(A) = \text{co}(S^2 \times \{1\}, S^2 \times \{-1\}) \subset \mathbb{R}^4.$$

Let  $c \in \mathbb{R}$ . The hyperplanes  $H_c = \{(y, c) \mid y \in \mathbb{R}^3\}$  are parallel in  $\mathbb{R}^4$ . Then  $\mathcal{F}(A) \cap H_c = \{(y, c) \mid |c| \leq \|y\| \leq 1\}$ . In particular,  $\mathcal{F}(A)$  is not convex.

**Example 4.** Let now  $n$  be arbitrary, and  $m \geq 4$ . Set

$$A_k = \begin{cases} \sigma_k \oplus 0_{(n-2) \times (n-2)} & 1 \leq k \leq 3, \\ 0_{2 \times 2} \oplus I_{n-2} & k = 4, \\ 0_{n \times n} & 4 < k \leq m. \end{cases}$$

Proposition 2.8 and Example 1 yield

$$\mathcal{F}(A) = \text{co}\left(S^2 \times \{0\}, \left\{ \begin{bmatrix} 0_{3 \times 1} \\ 1 \end{bmatrix} \right\}\right) \times \{0_{(m-4) \times 1}\}.$$

Since

$$\text{co}\left(S^2 \times \{0\}, \left\{ \begin{bmatrix} 0_{3 \times 1} \\ 1 \end{bmatrix} \right\}\right) = \left\{ \begin{bmatrix} rx \\ 1-r \end{bmatrix} \mid x \in S^2, r \in [0, 1] \right\},$$

the set  $\mathcal{F}(A)$  is a nonconvex cone. Hence, we obtain the following.

**Proposition 2.10** *For any  $m \geq 4$ , there exist  $A \in \mathcal{H}(n)^m$  such that  $\mathcal{F}(A)$  is not convex.*

**Example 5.**

Let  $n$  be arbitrary, and let  $e_1, \dots, e_n$  be the standard basis of  $\mathbb{C}^n$ . We define  $\sigma^{(n)} = (\sigma_1^{(n)}, \dots, \sigma_{2n-1}^{(n)})^* \in \mathcal{H}(n)^{2n-1}$  by  $\sigma_k^{(n)} = e_k e_n^T + e_n e_k^T$ ,  $\sigma_{n-1+k}^{(n)} = j(e_n e_k^T - e_k e_n^T)$  for  $1 \leq k \leq n-1$ , and

$$\sigma_{2n-1}^{(n)} = \begin{bmatrix} I_{n-1} & 0 \\ 0 & -1 \end{bmatrix}.$$

If  $n = 2$  we recover the Pauli spin matrices. Let  $z = (z_1, \dots, z_n)^T \in \mathbb{C}^n$ . For  $1 \leq k \leq n-1$  we have

$$z^* \sigma_k^{(n)} z = 2\Re(\overline{z_k} z_n), \quad z^* \sigma_{n-1+k}^{(n)} z = 2\Im(\overline{z_k} z_n); \quad z^* \sigma_{2n-1}^{(n)} z = \left( \sum_{k=1}^{n-1} |z_k|^2 \right) - |z_n|^2.$$

Since  $\|z\| = 1$ , we have

$$\|F_{\sigma^{(n)}}([z])\|^2 = \sum_{j=1}^{2n-1} (z^* \sigma_j^{(n)} z)^2 = 4 \sum_{k=1}^{n-1} |\overline{z_k} z_n|^2 + \left( \sum_{k=1}^{n-1} |z_k|^2 - |z_n|^2 \right)^2 = 1.$$

Let  $y = (y_1, \dots, y_{2n-1})^T \in S^{2(n-1)}$ . Let  $y_{2n-1} = r$ , and for  $1 \leq k \leq n-1$  set

$$y_k = \varrho_k \cos(\phi_k), \quad y_{n-1+k} = \varrho_k \sin(\phi_k).$$

The parameters thus introduced are constrained only by  $r \in [-1, 1]$ ,  $\phi_k \in \mathbb{R}$ ,  $\varrho_k \geq 0$ ,  $\sum_{k=1}^{n-1} \varrho_k^2 = 1 - r^2$ . For  $1 \leq k \leq n-1$  set  $z_k = 0$  if  $r = -1$ ,  $z_k = \sqrt{1/(2(1-r))} \rho_k$  if  $r \neq -1$ , and let  $z_n = \sqrt{(1-r)/2} e^{j\phi_k}$ . Then  $F_{\sigma^{(n)}}([z]) = y$ . Since  $\mathcal{F}(\sigma^{(n)}) \subset S^{2(n-1)}$ , we obtain  $\mathcal{F}(\sigma^{(n)}) = S^{2(n-1)}$ .

Let  $w \in \mathbb{C}^{n-1}$  and  $a, b \in \mathbb{R}$ . Extending the calculations of Example 2, we obtain

$$\begin{bmatrix} (\Re w)^T & (\Im w)^T & \frac{a-b}{2} & \frac{a+b}{2} \end{bmatrix} \begin{bmatrix} \sigma^{(n)} \\ I_n \end{bmatrix} = \begin{bmatrix} aI_{n-1} & \bar{w} \\ w^T & b \end{bmatrix}. \quad (4)$$

Combining this and Examples 2 and 5, we obtain the following.

**Proposition 2.11** *Let  $n$  and  $m$  be arbitrary. For  $1 \leq k \leq m$  let  $a_k, b_k \in \mathbb{R}$ ,  $w_k = x_k + jy_k \in \mathbb{C}^{n-1}$ ,  $A_k = \begin{bmatrix} a_k I_{n-1} & \bar{w}_k \\ w_k^T & b_k \end{bmatrix}$ , and let  $A = (A_1, \dots, A_m)^*$ . Set*

$$M = \begin{bmatrix} x_1^T & y_1^T & \frac{a_1-b_1}{2} \\ \vdots & \vdots & \vdots \\ x_m^T & y_m^T & \frac{a_m-b_m}{2} \end{bmatrix} \in \mathbb{R}^{m \times (2n-1)}, \quad p = \frac{1}{2} \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_m + b_m \end{bmatrix} \in \mathbb{R}^m.$$

Define  $\alpha : \mathbb{R}^{2n-1} \rightarrow \mathbb{R}^m$  by  $\alpha(\xi) = M\xi + p$ . Then  $A = M \begin{bmatrix} \sigma^{(n)} \\ I_n \end{bmatrix} + pI_n$  and, finally,  $\mathcal{F}(A) = \alpha(S^{2n-2})$ .

Proposition 2.11 serves as a source of examples of numerical ranges. We give a simple criterion of convexity for these examples.

**Corollary 2.12** *Let  $n$  and  $m$  be arbitrary. Let  $A = (A_1, \dots, A_m)^*$  and  $M \in \mathbb{R}^{m \times (2n-1)}$  be as in Proposition 2.11. Then  $\mathcal{F}(A)$  is convex if and only if  $\text{rank } M < 2n-1$ . In particular, if  $m < 2n-1$ , then  $\mathcal{F}(A)$  is convex.*

**Proof:** In view of Proposition 2.11,  $\mathcal{F}(A)$  is a translation of the image of the unit sphere under the linear mapping  $M : \mathbb{R}^{2n-1} \rightarrow \mathbb{R}^m$ . We use the singular value decomposition of  $M$ , and extend the argument of Example 2. If  $\text{rank } M \geq 2n-1$ , then  $\mathcal{F}(A) \subset \mathbb{R}^m$  is an isometric image of a  $(2n-2)$ -dimensional ellipsoid. If  $\text{rank } M < 2n-1$ , then  $\mathcal{F}(A) \subset \mathbb{R}^m$  is an isometric image of a solid ellipsoid of dimension  $\text{rank } M$ . We leave the details to the reader.  $\square$

### 3 Convex hull of the joint numerical range

In this section, we derive results relevant to the convex hull of  $\mathcal{F}(A)$ . Besides being of interest in their own rights, these results will be used in § 5. In § 3.1, we review those propositions from differentiable convex analysis that will be used later. In § 3.2, we specialize these results to the joint numerical range.

#### 3.1 Differentiability of support functions

We work in the space  $\mathbb{R}^m$ . For  $\eta \in \mathbb{R}^m, \eta \neq 0$ , and  $c \in \mathbb{R}$  set  $H(\eta, c) = \{y \in \mathbb{R}^m \mid \eta^T y = c\}$ ,  $H^-(\eta, c) = \{y \in \mathbb{R}^m \mid \eta^T y \leq c\}$ . Let  $\mathcal{C}_m \subset \mathcal{K}_m$  be the set of convex subsets, and let  $\mathcal{S}_m$  denote the set of convex functions  $s : \mathbb{R}^m \rightarrow \mathbb{R}$  satisfying  $s(\alpha\eta) = \alpha s(\eta)$  for  $\alpha \geq 0$ . The latter are called *support functions*. In what follows, we review basic facts about compact (convex) sets in  $\mathbb{R}^m$  and their support functions. We refer the reader to [14] for details.

Let  $K \in \mathcal{K}_m$ , and set  $s(\eta) = \max_{y \in K} \eta^T y$ . Then  $s \in \mathcal{S}_m$  is the support function of  $K$ . For  $\eta \in S^{m-1}$  the halfspace (resp. hyperplane)  $H^-(\eta, s(\eta))$  (resp.  $H(\eta, s(\eta))$ ) is the *supporting halfspace* (resp. *supporting hyperplane*) of  $K$  in direction  $\eta$ . We have [14, Theorem 2.2.2]

$$\text{co}(K) = \bigcap_{\eta \in S^{m-1}} H^-(\eta, s(\eta)) = \{y \in \mathbb{R}^m \mid \eta^T y \leq s(\eta) : \eta \in \mathbb{R}^m\}.$$

For  $\eta \in S^{m-1}$  we set  $\Phi_\eta(K) = K \cap H(\eta, s(\eta))$ . By  $\partial K$  we denote the boundary of  $K$  and by  $\partial_0 K$  the outer boundary. The latter is the boundary between  $K$  and the unbounded component of  $\mathbb{R}^m \setminus K$ . Obviously,  $\Phi_\eta(K) \subseteq \partial_0 K$ . Let  $C \in \mathcal{C}_m$ . Then the convex sets  $\Phi_\eta(C)$  are called the *exposed faces* of  $C$ . We have [14, Proposition 3.1.15]

$$\partial C = \bigcup_{\eta \in S^{m-1}} \Phi_\eta(C). \quad (5)$$

The set  $C \in \mathcal{C}_m$  is said to be strictly convex if  $\Phi_\eta(C)$  is a singleton for any  $\eta \in S^{m-1}$ .

**Proposition 3.1** *Let  $K \in \mathcal{K}_m$ . Then the following claims hold.*

- i) For any  $\eta \in S^{m-1}$ , we have  $\Phi_\eta(\text{co}(K)) = \text{co}(\Phi_\eta(K))$ .*
- ii) The set  $\text{co}(K)$  is strictly convex if and only if  $\Phi_\eta(K)$  is a singleton for all  $\eta \in S^{m-1}$ .*
- iii) We have  $\partial \text{co}(K) = \bigcup_{\eta \in S^{m-1}} \text{co}(\Phi_\eta(K))$ .*
- iv) The inclusion  $\partial \text{co}(K) \subset \partial K$  holds if and only if  $\partial \text{co}(K) = \partial_0 K$  if and only if  $\Phi_\eta(K)$  is convex for any  $\eta \in S^{m-1}$ .*

**Proof:** The claims *ii*) – *iv*) follow from claim *i*) and relation (5). We prove the first claim. Let  $s$  be the support function of  $K$  and let  $y \in \Phi_\eta(\text{co}(K))$ . Then  $\eta^T y = s(\eta)$  and  $y = \sum_{k=1}^p \alpha_k y_k$ , where  $y_k \in K$ ,  $\alpha_k > 0$ , and  $\sum \alpha_k = 1$ . Therefore for all indices  $\eta^T y_j \leq s(\eta)$ . Suppose that  $\eta^T y_k < s(\eta)$  for at least one index. Then  $s(\eta) = \eta^T y = \sum \alpha_j \eta^T y_j < \sum \alpha_j s(\eta) = s(\eta)$ . Thus,  $\eta^T y_j = s(\eta)$  for all  $j$ , and hence  $y \in \text{co}(\Phi_\eta(K))$ , implying that  $\Phi_\eta(\text{co}(K)) \subseteq \text{co}(\Phi_\eta(K))$ . The opposite inclusion follows from  $\Phi_\eta(K) \subseteq \Phi_\eta(\text{co}(K))$ .  $\square$

In the rest of this subsection we study the gradient and the Hessian of support functions. We will use the notation  $\nabla f(x)$  for the gradient at  $x \in \mathbb{R}^m$ .

**Proposition 3.2** *Let  $K \in \mathcal{K}_m$ , and let  $s$  be its support function. Suppose that  $s$  is continuously differentiable on an open set,  $U \subset \mathbb{R}^m$ . Then for  $\eta \in U \cap S^{m-1}$  the set  $\Phi_\eta(K)$  is a singleton:  $\Phi_\eta(K) = \{\nabla s(\eta)\}$ .*

**Proof:** Let  $\eta \in U$ . For any  $y \in \Phi_\eta(K)$ , we have  $\eta^T y = s(\eta)$ . For any such  $y$ , set  $f(\xi) = s(\xi) - y^T \xi$ . Then the function  $f \geq 0$  is continuously differentiable on  $U$ , and  $f(\eta) = 0$ . Hence  $\nabla s(\eta) = y$ . Since the latter holds  $\forall y \in \Phi_\eta(K)$ , the claim follows.  $\square$

Now, we consider the case when the support function is at least twice differentiable.

**Proposition 3.3** *Let  $K \in \mathcal{K}_m$ , and let  $s$  be its support function. Let  $r \geq 2$ , and assume that  $s$  is a  $\mathcal{C}^r$ -function on  $\mathbb{R}^m \setminus \{0\}$ . Suppose that  $d_\eta^2 s$  is positive definite on  $\eta^\perp$  for any  $\eta$ . Then*

- i) The map  $\eta \mapsto \nabla s(\eta)$  is a  $\mathcal{C}^{r-1}$ -embedding of  $S^{m-1}$  into  $\mathbb{R}^m$ .*
- ii) The set  $\text{co}(K)$  is strictly convex. The range of the above map is  $\partial \text{co}(K)$ . We have  $\partial \text{co}(K) = \partial_0 K$ .*

The proof is based on the following lemma.

**Lemma 3.4** *Let  $s$  be the support function of  $K \in \mathcal{K}_m$ . Assume that  $s$  is a  $\mathcal{C}^2$ -function on  $\mathbb{R}^m \setminus \{0\}$ . Let  $\eta_1, \eta_2 \in S^{m-1}$  and suppose that  $\nabla s(\eta_1) \in H(\eta_2, s(\eta_2))$ . Then the differential of  $\nabla s$  at  $\eta_1$  satisfies  $(d_{\eta_1} \nabla s)(\eta_1^\perp) \subseteq \eta_2^\perp$ .*

**Proof:** Let  $\xi \in T_{\eta_1} S^{m-1} = \eta_1^\perp$  and let  $\gamma : (-\epsilon, \epsilon) \rightarrow S^{m-1}$  be a differentiable curve satisfying  $\gamma(0) = \eta_1$  and  $\gamma'(0) = \xi$ . By assumption we have  $\nabla s(\gamma(0)) \in H(\eta_2, s(\eta_2))$ , and Proposition 3.2 yields that  $\nabla s(\gamma(t)) \in K \subseteq H^-(\eta_2, s(\eta_2))$  for all  $t \in (-\epsilon, \epsilon)$ . Thus the function  $t \mapsto \eta_2^T \nabla s(\gamma(t))$  attains its maximum,  $s(\eta_2)$ , at  $t = 0$ . Thus  $0 = \frac{d}{dt}(\eta_2^T \nabla s(\gamma(t)))|_{t=0} = \eta_2^T (d_{\gamma(0)} \nabla s)(\gamma'(0)) = \eta_2^T (d_{\eta_1} \nabla s)(\xi)$ .  $\square$

**Proof of Propostion 3.3:** The second claim is a consequence of Propositions 3.1 and 3.2. We will prove the first. By Proposition 3.2 we have  $\nabla s(\eta) \in H(\eta, s(\eta))$ . Thus, by the lemma,  $(d_\eta \nabla s)(\eta^\perp) \subseteq \eta^\perp$ . Now, the positive definiteness of the quadratic form  $\eta^\perp \ni \xi \mapsto d_\eta^2 s(\xi, \xi) = \xi^T (d_\eta \nabla s)(\xi)$  yields

$$(d_\eta \nabla s)(\eta^\perp) = \eta^\perp. \quad (6)$$

Thus, the map  $\eta \mapsto \nabla s(\eta)$ ,  $\eta \in S^{m-1}$ , is an immersion. We now prove that it is injective. Let  $\tilde{\eta} \in S^{m-1}$ ,  $\tilde{\eta} \neq \eta$ , and suppose that  $\nabla s(\eta) = \nabla s(\tilde{\eta})$ . Then  $\nabla s(\eta) \in H(\tilde{\eta}, s(\tilde{\eta}))$ . Thus, by the lemma,  $(d_\eta \nabla s)(\eta^\perp) \subseteq \tilde{\eta}^\perp$ . If  $\tilde{\eta} \neq -\eta$  this contradicts equation (6). If  $\tilde{\eta} = -\eta$ , then  $K$  is contained in the hyperplane  $H(\eta, s(\eta))$ . Then, by Proposition 3.2,  $K$  is a singleton, which contradicts equation (6) again. Thus, we have shown that the map  $\nabla s(\eta)$  is an injective immersion, and since it applies to a closed manifold it is an embedding.  $\square$

### 3.2 Support function of a joint numerical range

Let  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(n)^m$ . We will apply the material of the preceding section to  $\mathcal{F}(A)$ . If  $V \subset \mathbb{C}^n$  is a subspace, we denote by  $\mathcal{F}(A; V)$  the numerical range of the restriction of  $A$  to  $V$ , i.e.  $\mathcal{F}(A; V) = \{ F_A([z]) \mid 0 \neq z \in V \}$ . The following result is basic.

**Proposition 3.5** *Let  $A \in \mathcal{H}(n)^m$ , let  $\mathcal{F}(A)$  be its numerical range, and let  $s \in \mathcal{S}_m$  be the support function of  $\mathcal{F}(A)$ . Then*

$$s(\eta) = \lambda_1(\eta^T A). \quad (7)$$

Let  $\eta \in S^{m-1}$ . Then

$$\mathcal{F}(A) \cap H(\eta, \lambda_1(\eta^T A)) = \mathcal{F}(A; E_1(\eta^T A)). \quad (8)$$

**Proof:** Let  $z \in \mathbb{C}^n$ ,  $\|z\| = 1$ , and let  $\eta \in \mathbb{R}^m$ . Then  $\eta^T F_A([z]) = z^*(\eta^T A)z \leq \lambda_1(\eta^T A)$ . Equality holds if and only if  $z \in E_1(\eta^T A)$ .  $\square$

Specializing Proposition 3.1 to the numerical range with the help of Proposition 3.5 yields the following useful claim.

**Corollary 3.6** *Let  $A \in \mathcal{H}(n)^m$ . Then the following claims hold.*

*i) Let  $T \subseteq \mathbb{R}^m$  be a proper subset. Then  $T$  is an exposed face of  $\text{co}(\mathcal{F}(A))$  if and only if there exists  $\eta \in \mathbb{R}^m$  such that  $T = \text{co}(\mathcal{F}(A; E_1(\eta^T A)))$ .*



ii) The set  $\text{co}(\mathcal{F}(A))$  is strictly convex if and only if  $\mathcal{F}(A; E_1(\eta^T A))$  is a singleton for any  $\eta \in S^{m-1}$ .

iii) We have  $\partial \text{co}(\mathcal{F}(A)) = \bigcup_{\eta \in S^{m-1}} \text{co}(\mathcal{F}(A; E_1(\eta^T A)))$ .

iv) The inclusion  $\partial \text{co}(\mathcal{F}(A)) \subset \partial \mathcal{F}(A)$  holds if and only if  $\partial \text{co}(\mathcal{F}(A)) = \partial_0 \mathcal{F}(A)$  if and only if  $\mathcal{F}(A; E_1(\eta^T A))$  is convex for any  $\eta \in S^{m-1}$ .

In order to apply Propositions 3.2 and 3.3, we need a technical, but useful, theorem. For positive integers  $k, \mu, n$  such that  $k - 1 + \mu \leq n$ , let  $\mathcal{H}_{k,\mu}(n)$  denote the set of hermitian  $n \times n$ -matrices  $A$  such that  $\lambda_j(A) = \lambda_k(A)$  if and only if  $k \leq j \leq k + \mu - 1$ . For  $A \in \mathcal{H}_{k,\mu}(n)$ , the eigenvalue  $\lambda_k(A) = \dots = \lambda_{k+\mu-1}$  has multiplicity  $\mu$ , while all other eigenvalues have arbitrary multiplicity.

**Theorem 3.7** *Let  $M$  be a  $\mathcal{C}^r$ -manifold and let  $H : M \rightarrow \mathcal{H}(n)$  be a  $\mathcal{C}^r$ -map such that  $H(M) \subset \mathcal{H}_{k,\mu}(n)$ . Then the following claims hold.*

1. *Let  $x_0 \in M$  and let  $(z_{01}, \dots, z_{0\mu})$  be an orthonormal basis of  $E_k(H(x_0))$ . Then there exists an open neighbourhood  $U_0$  of  $x_0$  and  $\mathcal{C}^r$ -maps  $z_j : U_0 \rightarrow \mathbb{C}^n$ ,  $1 \leq j \leq \mu$ , such that for any  $x \in U_0$  the set  $(z_1(x), \dots, z_\mu(x))$  is an orthonormal basis of  $E_k(H(x))$ ,  $(z_1(x_0), \dots, z_\mu(x_0)) = (z_{01}, \dots, z_{0\mu})$ , and if  $\xi \in T_{x_0}M$ , then*

$$(d_{x_0} z_j)(\xi) = (\lambda_k(H(x_0))I_n - H(x_0))^\dagger (d_{x_0} H(\xi)) z_{0j}.$$

2. *The composition  $\lambda_k \circ H$  is a  $\mathcal{C}^r$ -function on  $M$ . Let  $x \in M$ ,  $\xi \in T_x M$ , and let  $z, w \in E_k(H(x))$ . Then*

$$d_x(\lambda_k \circ H)(\xi) w^* z = w^* d_x H(\xi) z. \quad (9)$$

*In particular, if  $\|z\| = 1$  then*

$$d_x(\lambda_k \circ H)(\xi) = z^* d_x H(\xi) z.$$

3. *Let  $r \geq 2$ , and suppose that the second differentials below are defined at  $x \in M$ . Let  $\xi_1, \xi_2 \in T_x M$ . Then for any unit vector  $z \in E_k(H(x))$ , we have*

$$\begin{aligned} (d_x^2(\lambda_k \circ H))(\xi_1, \xi_2) &= 2z^*(d_x H(\xi_1) (\lambda_k(H(x))I_n - H(x))^\dagger d_x H(\xi_2))z \\ &\quad + z^*(d_x^2 H(\xi_1, \xi_2))z. \end{aligned}$$

In order not to interrupt the flow of exposition, we defer the proof of Theorem 3.7 to Appendix A. Specializing Proposition 3.2 and the above Theorem to the numerical range, and using Proposition 3.5, we obtain the following result.

**Theorem 3.8** *Let  $A \in \mathcal{H}(n)^m$ . Suppose that  $\lambda_1(\eta^T A)$  has constant multiplicity for  $\eta \in U \subset \mathbb{R}^m$ , an open set. Then the support function of  $\mathcal{F}(A)$  is real analytic on  $U$ . Let  $z \in E_1(\eta^T A)$  be any unit vector. Then, in the notation of equation (7), for any  $\eta \in U$ , we have*

$$d_\eta s(\xi) = z^*(\xi^T A)z \quad (10)$$

$$d_\eta^2 s(\xi_1, \xi_2) = 2z^*(\xi_1^T A)(s(\eta)I_n - \eta^T A)^\dagger(\xi_2^T A)z. \quad (11)$$

If  $\eta \in U \cap S^{m-1}$ , then the intersection of  $\text{co}(\mathcal{F}(A))$  with  $H(\eta, s(\eta))$  is a singleton, and

$$\mathcal{F}(A; E_1(\eta^T A)) = \{\nabla s(\eta)\} = \{(z^* A_1 z, \dots, z^* A_m z)^T\} = \{F_A([z])\}.$$

Now, we consider the situation where  $\lambda_1(\eta^T A)$ ,  $\eta \in S^{m-1}$ , has constant multiplicity  $\mu$ . Since  $\lambda_n(\eta^T A) = -\lambda_1(-\eta^T A)$ , its multiplicity is  $\mu$  as well. If  $\mu > \frac{n}{2}$ , then all matrices  $A_j$  are scalar, and  $\mathcal{F}(A)$  is a point. Hence, we assume in what follows that  $\mu \leq \frac{n}{2}$ . The proposition below is the main result of this section. We use the notation of equation (7) in its formulation.

**Theorem 3.9** *Let  $A \in \mathcal{H}(n)^m$ . Let  $\mu \leq \frac{n}{2}$  be the multiplicity of  $\lambda_1(\eta^T A)$ . Then the map  $\eta \mapsto \nabla s(\eta)$  is a real analytic embedding of  $S^{m-1}$  into  $\mathbb{R}^m$ . The range of the map is the boundary of the strictly convex set  $\text{co}(\mathcal{F}(A))$ . The latter coincides with the outer boundary of  $\mathcal{F}(A)$ .*

**Proof:** In view of Proposition 3.3 and Theorem 3.8, it suffices to show that, for any  $\eta \in S^{m-1}$  and  $\xi \in \eta^\perp \setminus \{0\}$ , there exists a unit vector  $z \in E_1(\eta^T A)$  such that

$$0 < d_\eta^2 s(\xi, \xi) = 2(\xi^T A z)^* (\lambda_1(\eta^T A)I_n - \eta^T A)^\dagger (\xi^T A z). \quad (12)$$

Since the operator  $(\lambda_1(\eta^T A)I_n - \eta^T A)^\dagger$  is positive semidefinite, the preceding inequality is equivalent to  $(\xi^T A)z \notin E_1(\eta^T A)$ . This follows from the third claim in the proposition below.  $\square$

**Proposition 3.10** *Let  $A \in \mathcal{H}(n)^m$ . Suppose that there are  $1 \leq k, \mu \leq n$ ,  $k + \mu - 1 \leq n/2$ , such that  $\eta^T A \in \mathcal{H}_{k, \mu}(n)$  for any  $\eta \in \mathbb{R}^m \setminus \{0\}$ . Let  $\eta_1, \eta_2 \in S^{m-1}$  be linearly independent vectors and let  $1 \leq j \leq n$ . Then the following claims hold.*

- i) We have  $E_k(\eta_1^T A) \cap E_j(-\eta_1^T A) \neq \{0\}$  if and only if  $E_k(\eta_1^T A) = E_j(-\eta_1^T A)$ . The latter holds if and only if  $n - k - \mu + 2 \leq j \leq n - k + 1$ .*
- ii) We have  $E_k(\eta_1^T A) \cap E_j(\eta_2^T A) = \{0\}$ .*
- iii) We have  $E_k(\eta_1^T A) \cap (\eta_2^T A)E_k(\eta_1^T A) = \{0\}$ .*

**Proof:** The first claim is immediate from the definition of  $\mathcal{H}_{k,\mu}(n)$  and the fact that  $E_j(-\eta_1^T A) = E_{n+1-j}(\eta_1^T A)$ . Let  $0 \neq z \in E_k(\eta_1^T A)$ .

Suppose that  $z$  is also an eigenvector of  $\eta_2^T A$ . Then  $z$  is an eigenvector of  $\eta^T A$  for all  $\eta \in \text{span}\{\eta_1, \eta_2\}$ . For each  $\eta \neq 0$  the eigenvalue  $\lambda_k(\eta^T A)$  belongs to an isolated group of  $\mu$  identical eigenvalues. Thus  $z \in E_k(\eta^T A)$  for all  $\eta \in \text{span}\{\eta_1, \eta_2\}$ . In particular,  $z \in E_k(-\eta_1^T A)$ . This contradicts the first claim.

Suppose now that  $(\eta_2^T A)z \in E_k(\eta_1^T A)$ . Set  $f(\eta) = \lambda_k(\eta^T A)$ . Since  $z \in E_k(\eta_1^T A)$ , the relation (9) in Theorem 3.7 implies that  $w^*(\eta_2^T A)z = d_{\eta_1} f(\eta_2)w^*z$  for all  $w \in E_k(\eta_1^T A)$ . We conclude that  $(\eta_2^T A)z = d_{\eta_1} f(\eta_2)z$ . Thus  $z$  is an eigenvector of  $\eta_2^T A$ , a contradiction to the second claim.  $\square$

## 4 The viewpoint of differential topology

### 4.1 The boundary of a joint numerical range

In this section we study the critical points and the critical values of the numerical range map  $F_A$ , and obtain information about the boundary of  $\mathcal{F}(A)$ , which will be written  $\partial\mathcal{F}(A)$ . We begin with necessary preliminaries. Let  $M, N$  be smooth manifolds without boundary, and let  $f : M \rightarrow N$  be a differentiable map. Then  $x \in M$  is a critical point if  $d_x f : T_x M \rightarrow T_x N$  is not surjective. The set of critical points will be written  $\mathcal{C}(f) \subset M$ . A point  $y \in N$  is a critical value if  $f^{-1}(y)$  contains a critical point. The following fact is basic.

**Theorem 4.1** *Let the setting be as above. Then  $f^{-1}(\partial f(M)) \subset \mathcal{C}(f)$ .*

Let  $z \in \mathbb{C}^n$  be a unit vector. The differential at  $\delta = 0$  of the map  $\delta \mapsto [z + \delta]$ ,  $\delta \in [z]^\perp$ , induces a linear isomorphism of  $[z]^\perp = \{w \in \mathbb{C}^n \mid w^*z = 0\}$  onto the tangent space  $T_{[z]}\mathbb{C}\mathbb{P}^{n-1}$ . Replacing  $z$  by  $e^{j\theta}z$  changes the isomorphism in question by the factor  $e^{-j\theta}$ . We will use these isomorphisms to identify  $T_{[z]}\mathbb{C}\mathbb{P}^{n-1}$  with  $[z]^\perp$ .

**Proposition 4.2** *Let  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(n)^m$ . Let  $z \in \mathbb{C}^n$  be a unit vector, and let  $[z] \in \mathbb{C}\mathbb{P}^{n-1}$  be the corresponding point. Identify  $T_{[z]}\mathbb{C}\mathbb{P}^{n-1}$  with  $[z]^\perp$  via the linear isomorphism determined by  $z$ . Let  $\delta \in [z]^\perp$ . Then*

$$d_{[z]}F_A(\delta) = 2(\Re(z^*A_1\delta), \dots, \Re(z^*A_m\delta))^T. \quad (13)$$

*The space  $(\text{range } d_{[z]}F_A)^\perp$  consists of  $\eta \in \mathbb{R}^m$  such that  $z$  is an eigenvector of  $\eta^T A$ .*

**Proof:** Equation (13) is immediate from the special case  $m = 1$  and the relation  $F_A([z]) = (F_{A_1}([z]), \dots, F_{A_m}([z]))^T$ . The following chain of equivalences yields the other claim:  $\eta \in (\text{range } d_{[z]}F_A)^\perp$  iff  $\forall \delta \in [z]^\perp$  we have  $\eta^T d_{[z]}F_A(\delta) = 0$  iff  $\forall \delta \in [z]^\perp$  we have  $d_{[z]}F_{\eta^T A}(\delta) = 0$  iff  $[z]$  is a critical point of  $F_{\eta^T A}$  iff  $z$  is an eigenvector of  $\eta^T A$ . The second equivalence in the chain holds by Corollary 2.5. The last equivalence was proved in [19].  $\square$

The corollary below follows directly from Theorem 4.1 and Proposition 4.2.

**Corollary 4.3** *Let  $A \in \mathcal{H}(n)^m$  and let  $F_A : \mathbb{C}\mathbb{P}^{n-1} \rightarrow \mathbb{R}^m$  be the corresponding mapping. Then the set of critical points of  $F_A$  is*

$$\mathcal{C}(F_A) = \{ [z] \in \mathbb{C}\mathbb{P}^{n-1} \mid \exists \eta \in S^{m-1} : z \text{ is an eigenvector of } \eta^T A \}.$$

*Let  $y \in \partial\mathcal{F}(A)$ . Then for each  $[z] \in F_A^{-1}(y)$  there exists  $\eta \in S^{m-1}$  such that  $z$  is an eigenvector of  $\eta^T A$ .*

The case  $m = 2$  of Corollary 4.3 is contained in [19].

## 4.2 Eigenvalues of constant multiplicity

If  $X$  is a vector space, we denote by  $\mathbb{P}(X)$  the corresponding projective space. If  $X \subset Y$  is a subspace, then  $\mathbb{P}(X) \subset \mathbb{P}(Y)$ . The following Theorem is the main result of this subsection.

**Theorem 4.4** *Let  $A \in \mathcal{H}(n)^m$ . Suppose that there are  $1 \leq k, \mu \leq n$ ,  $k + \mu - 1 \leq n/2$ , such that  $\eta^T A \in \mathcal{H}_{k,\mu}(n)$  for any  $\eta \in \mathbb{R}^m \setminus \{0\}$ . Then the following claims hold.*

- i) Let  $\eta_1, \eta_2 \in S^{m-1}$  be linearly independent vectors and let  $1 \leq j \leq n$ . Then the projective spaces  $\mathbb{P}(E_k(\eta_1^T A))$  and  $\mathbb{P}(E_j(\eta_2^T A))$  are disjoint. Furthermore,  $\mathbb{P}(E_k(\eta_1^T A)) \cap \mathbb{P}(E_j(-\eta_1^T A)) \neq \emptyset$  if and only if  $\mathbb{P}(E_k(\eta_1^T A)) = \mathbb{P}(E_j(-\eta_1^T A))$ . The latter holds if and only if  $n - k - \mu + 2 \leq j \leq n - k + 1$ .*
- ii) The disjoint union  $\mathcal{P} = \bigcup_{\eta \in S^{m-1}} \mathbb{P}(E_k(\eta^T A))$  is a closed real analytic submanifold of  $\mathbb{C}\mathbb{P}^{n-1}$ .*
- iii) Define  $p : \mathcal{P} \rightarrow S^{m-1}$  by  $p([z]) = \eta$  if  $z \in E_k(\eta^T A)$ . Then the triple  $(\mathcal{P}, p, S^{m-1})$  is a real analytic, locally trivial fiber bundle with fiber  $\mathbb{C}\mathbb{P}^{\mu-1}$ .*

Theorem 4.4 has a reformulation in terms of sphere bundles. We will use it in § 5. For convenience of the reader, we formulate the theorem below.

**Theorem 4.5** *Let  $A \in \mathcal{H}(n)^m$ , and let the assumptions be as in Theorem 4.4. Then the spheres  $S(E_k(\eta^T A)) = \{ z \in E_k(\eta^T A) \mid \|z\| = 1 \}$ ,  $\eta \in S^{m-1}$ , are*

pairwise disjoint. Their union  $\mathcal{S} = \bigcup_{\eta \in S^{m-1}} S(E_k(\eta^T A))$  is a compact submanifold of  $\mathbb{C}^n$ . The map  $q : \mathcal{S} \rightarrow S^{m-1}$ , where  $q(z) = \eta$  if  $z \in S(E_k(\eta^T A))$ , is well defined, and  $(\mathcal{S}, q, S^{m-1})$  is a real analytic, locally trivial sphere bundle.

**Proof of Theorem 4.4:** The first claim i) is immediate from Proposition 3.10. It implies that the subsets  $\mathbb{P}(E_k(\eta^T A)) \subset \mathbb{C}\mathbb{P}^{n-1}$ ,  $\eta \in S^{m-1}$ , are pairwise disjoint. Therefore the projection  $p : \mathcal{P} \rightarrow S^{m-1}$  is well defined.

Regarding claims ii) and iii), we first establish the purely topological properties of  $\mathcal{P}$  and  $p$ . Let  $[z_j], j \geq 1$ , be a sequence in  $\mathcal{P}$  converging to  $[z] \in \mathbb{C}\mathbb{P}^{n-1}$ . Multiplying the vectors  $z_j$  by suitable scalars, we can assume that  $\lim z_j = z$ . Set  $\eta_j = p([z_j]) \in S^{m-1}$ . Then

$$(\eta_j^T A)z_j = \lambda_k(\eta_j^T A)z_j. \quad (14)$$

By compactness of  $S^{m-1}$ , the sequence  $\eta_j$  has a converging subsequence,  $\eta_{j_\ell}$ . Let  $\eta$  be its limit. By equation (14) and by continuity,  $(\eta^T A)z = \lambda_k(\eta^T A)z$ . Hence,  $z \in \mathcal{P}$  and  $p([z]) = \eta = \lim p([z_{j_\ell}])$ . Thus,  $\mathcal{P} \subset \mathbb{C}\mathbb{P}^{n-1}$  is a closed subset. Now, suppose that the sequence  $\eta_j$  does not converge to  $\eta$ . Then there exists an open set  $U \ni \eta$  and a subsequence of  $\eta_j$ , contained in  $S^{m-1} \setminus U$  and converging to  $\tilde{\eta} \neq \eta$ . By equation (14),  $(\tilde{\eta}^T A)z = \lambda_k(\tilde{\eta}^T A)z$ . Thus  $z \in E_k(\eta^T A) \cap E_k(\tilde{\eta}^T A)$ , a contradiction. Hence,  $p$  is continuous.

Next, we address the submanifold properties of  $\mathcal{P}$ . For any open subset  $V \subset S^{m-1}$  set  $\mathcal{P}_V = p^{-1}(V)$ . We show that for any  $[z] \in \mathcal{P}_V$  there exists an open set  $U \subset \mathbb{C}\mathbb{P}^{n-1}$  containing  $[z]$ , such that  $U \cap \mathcal{P} = U \cap \mathcal{P}_V$ . Assume the opposite. Then there exists a sequence  $[z_j] \subset \mathcal{P} \setminus \mathcal{P}_V$  converging to  $[z]$  and such that  $p([z_j]) \notin V$  for all  $j$ . But, by continuity of  $p$ ,  $\lim p([z_j]) = p([z]) \in V$ , a contradiction. Thus, in order to prove that  $\mathcal{P}$  is a submanifold [21, 2.7], it suffices to show that for each  $\eta \in S^{m-1}$  there exists an open neighbourhood  $V \subset S^{m-1}$  of  $\eta$  such that  $\mathcal{P}_V$  is a submanifold. However, by the construction below there is an open neighborhood  $V(\eta)$  of  $\eta$  and a real analytic embedding  $\Psi_\eta : V(\eta) \times \mathbb{C}\mathbb{P}^{\mu-1} \rightarrow \mathbb{C}\mathbb{P}^{n-1}$  such that  $\mathcal{P}_{V(\eta)} = \Psi_\eta(V(\eta) \times \mathbb{C}\mathbb{P}^{\mu-1})$ . Thus  $\mathcal{P}_{V(\eta)}$  is a real analytic submanifold of  $\mathbb{C}\mathbb{P}^{n-1}$ .

We are now going to construct the embedding  $\Psi_\eta$ . According to Theorem 3.7, there is an open neighborhood  $V \subset S^{m-1}$  of  $\eta \in S^{m-1}$  and real analytic functions  $z_1, \dots, z_\mu : V \rightarrow \mathbb{C}^n$  such that  $z_1(\tilde{\eta}), \dots, z_\mu(\tilde{\eta})$  is an orthonormal basis of  $E_k(\tilde{\eta}^T A)$  for all  $\tilde{\eta} \in V$ . Moreover, the differentials of the  $z_j$  satisfy  $d_\eta z_j(\xi) = G_\eta(z_j(\eta), \xi)$ , where

$$G_\eta : E_k(\eta^T A) \times \eta^\perp \rightarrow E_k(\eta^T A)^\perp, \quad G_\eta(z, \xi) := (\lambda_k(\eta^T A)I_n - \eta^T A)^\dagger (\xi^T A)z.$$

Note that for each  $z \in E_k(\eta^T A) \setminus \{0\}$  the linear map  $G_\eta(z, \cdot) : \eta^\perp \rightarrow E_k(\eta^T A)^\perp$  is injective since  $G_\eta(z, \xi) = 0$  with  $\xi \in \eta^\perp \setminus \{0\}$  implies that  $(\xi^T A)z \in$

$E_k(\eta^T A)$ . The latter fails by Proposition 3.10 iii) because  $(\xi^T A)z \in E_k(\eta^T A)$  and  $(\xi^T A)z \in (\xi^T A)E_k(\eta^T A)$ . Let  $Z(\tilde{\eta}) = [z_1(\tilde{\eta}), \dots, z_\mu(\tilde{\eta})] \in \mathbb{C}^{n \times \mu}$ . Then  $Z(\tilde{\eta})$  is unitary and  $\text{range}(Z(\tilde{\eta})) = E_k(\tilde{\eta}^T A)$ . Now define the real analytic map

$$\Psi : V \times \mathbb{C}\mathbb{P}^{\mu-1} \rightarrow \mathbb{C}\mathbb{P}^{n-1}, \quad \Psi(\tilde{\eta}, [w]) := [Z(\tilde{\eta})w].$$

Then  $\Psi(\{\tilde{\eta}\} \times \mathbb{C}\mathbb{P}^{\mu-1}) = \mathbb{P}(E_k(\tilde{\eta}^T A))$  for every  $\tilde{\eta} \in V$ . Thus  $\Psi$  is injective. Its inverse  $\Psi^{-1} : \mathcal{P}_V \rightarrow V \times \mathbb{C}\mathbb{P}^{\mu-1}$  has the form

$$\Psi^{-1}([z]) = (p([z]), [Z(p([z]))^* z]), \quad \|z\| = 1.$$

Thus  $\Psi^{-1}$  is continuous, and  $\Psi$  is a homeomorphism onto its image.

We now show that an appropriate restriction of  $\Psi$  is an immersion. A direct computation of the differential of  $\Psi$  at  $(\eta, [w])$  yields

$$d_{(\eta, [w])}\Psi(\xi, \delta) = \underbrace{G_\eta(Z(\eta)w, \xi)}_{\in E_k(\eta^T A)^\perp} + \underbrace{Z(\eta)\delta}_{\in E_k(\eta^T A)},$$

where  $\xi \in \eta^\perp$ ,  $\delta \in T_{[w]}\mathbb{C}\mathbb{P}^{\mu-1} \cong w^\perp$ . The relation  $d_{(\eta, [w])}\Psi(\xi, \delta) = 0$  implies  $(\xi, \delta) = 0$ . Thus,  $d_{(\eta, [w])}\Psi$  is injective. Therefore, for every  $[w] \in \mathbb{C}\mathbb{P}^{\mu-1}$ , there exist open neighborhoods  $V_w \subset V \subset S^{m-1}$  and  $U_w \subset \mathbb{C}\mathbb{P}^{\mu-1}$  of  $\eta$  and  $[w]$ , respectively, such that  $\Psi|_{V_w \times U_w}$  is an immersion. By compactness, there are  $w_1, \dots, w_r \in \mathbb{C}\mathbb{P}^{\mu-1}$  such that  $\bigcup_{1 \leq j \leq r} U_{w_j} = \mathbb{C}\mathbb{P}^{\mu-1}$ . Set  $V(\eta) := \bigcap_{1 \leq j \leq r} V_{w_j}$ . Then  $V(\eta) \times \mathbb{C}\mathbb{P}^{\mu-1} \subseteq \bigcup_{1 \leq j \leq r} (V_{w_j} \times U_{w_j})$ , and hence the map  $\Psi_\eta := \Psi|_{V(\eta) \times \mathbb{C}\mathbb{P}^{\mu-1}}$  is a real analytic immersion. But,  $\Psi_\eta$  is also a homeomorphism onto its image,  $\mathcal{P}_{V(\eta)}$ . Thus,  $\Psi_\eta$  is a real analytic embedding.

The maps  $\Psi_\eta$ ,  $\eta \in S^{m-1}$ , are local parametrizations of  $\mathcal{P}$ . Their inverses  $\Psi_\eta^{-1} : \mathcal{P}_{V(\eta)} \rightarrow V(\eta) \times \mathbb{C}\mathbb{P}^{\mu-1}$ ,  $\eta \in S^{m-1}$ , are bundle charts which endow the triple  $(\mathcal{P}, p, S^{m-1})$  with the structure of a real analytic projective fiber bundle [21]. To see this first note that for each  $\tilde{\eta} \in V(\eta)$  the restriction  $\Psi_\eta^{-1}|_{E_k(\tilde{\eta}^T A)} : \mathbb{P}(E_k(\tilde{\eta}^T A)) \rightarrow \{\tilde{\eta}\} \times \mathbb{C}\mathbb{P}^{\mu-1}$  is the projectivization of a linear isomorphism. Consider now two parametrizations

$$\Psi_{\eta_i} : V(\eta_i) \times \mathbb{C}\mathbb{P}^{\mu-1} \rightarrow \mathcal{P}_{V(\eta_i)}, \quad \Psi_{\eta_i}(\tilde{\eta}, [w]) := [Z_i(\tilde{\eta})w], \quad i = 1, 2,$$

where  $Z_i(\tilde{\eta})$  is unitary. Suppose the sets  $V(\eta_1)$  and  $V(\eta_2)$  overlap. Then the change of charts satisfies  $\Psi_{\eta_2}^{-1} \circ \Psi_{\eta_1}(\tilde{\eta}, [w]) = (\tilde{\eta}, [T(\tilde{\eta})w])$ , where  $T(\tilde{\eta}) := Z_2(\tilde{\eta})^* Z_1(\tilde{\eta})$  is unitary, so that the transition function depends analytically on  $\tilde{\eta}$ . (See [21, Remark 2.5.7, Section 6.4].)  $\square$

The special case of Theorem 4.4, when  $\mu = 1$  is especially useful.

**Corollary 4.6** *Let  $A \in \mathcal{H}(n)^m$ . Suppose that for an index  $k \neq \frac{n+1}{2}$  and all  $\eta \in S^{m-1}$  the eigenvalue  $\lambda_k(\eta^T A)$  is simple. Let  $\phi(\eta) = E_k(\eta^T A) \in \mathbb{C}\mathbb{P}^{n-1}$ ,  $\psi(\eta) = E_{n+1-k}(\eta^T A)$ . Then  $\phi, \psi : S^{m-1} \rightarrow \mathbb{C}\mathbb{P}^{n-1}$  are real analytic embeddings. They satisfy the identity  $\phi(-\eta) = \psi(\eta)$ . If  $j \neq k, n+1-k$  then  $\phi(S^{m-1}) \cap \mathbb{P}(E_j(\eta^T A)) = \emptyset$  for all  $\eta \in S^{m-1}$ .*

We conclude this section with a few remarks and examples. The following example illustrates the fact that the condition  $k \neq \frac{n+1}{2}$  in Corollary 4.6 is necessary.

**Example 6.** Let  $\sigma_k$  be the Pauli spin-matrices, and set

$$A = \left( \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \sigma_2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \sigma_3 & 0 \\ 0 & 0 \end{bmatrix} \right) \in \mathcal{H}(3)^3.$$

Then  $\lambda_1(\eta^T A) = 1, \lambda_2(\eta^T A) = 0, \lambda_3(\eta^T A) = -1$  for all  $\eta \in S^2$ . Thus, all eigenvalues are simple. However, for all  $\eta \in S^2$  we have  $\phi(\eta) = E_2(\eta^T A) = \mathbb{C}(0, 0, 1)^T$ .

Let  $A \in \mathcal{H}(n)^m$ . Suppose that  $A$  has a block of eigenvalues of constant multiplicity,  $\mu$ . The following proposition imposes some restrictions on the parameters.

**Proposition 4.7** *Let the notation be as in Theorem 4.4, and let  $A \in \mathcal{H}(n)^m$  satisfy the assumptions of the Theorem. Suppose that  $(k, m, \mu) \neq (1, n+1, \frac{n}{2})$ . Then*

$$m \leq 2(n - \mu). \tag{15}$$

**Proof:** We use the notation of Theorem 4.4. Denote by  $\dim$  the real dimension. Then

$$2(n-1) = \dim \mathbb{C}\mathbb{P}^{n-1} \geq \dim \mathcal{P} = \dim S^{m-1} + \dim \mathbb{C}\mathbb{P}^{\mu-1} = m-1 + 2(\mu-1).$$

Suppose that  $\dim \mathbb{C}\mathbb{P}^{n-1} = \dim \mathcal{P}$ . Then  $\mathcal{P} \subset \mathbb{C}\mathbb{P}^{n-1}$  is open. Since  $\mathcal{P}$  is closed, and  $\mathbb{C}\mathbb{P}^{n-1}$  is connected,  $\mathcal{P} = \mathbb{C}\mathbb{P}^{n-1}$ . By Theorem 4.4, this is possible only if  $k = 1, \mu = n/2$ . But then  $m = n+1$ .  $\square$

Proposition 4.7 is proved in [10] by a different method. The following example shows that if the eigenvalue in Proposition 4.7 is simple, then the bound in equation (15) is sharp.

**Example 7.** (Compare with [10], page 395.) Set

$$X_n = \left\{ \begin{bmatrix} 0_{n-1} & x \\ x^* & 0 \end{bmatrix} \mid x \in \mathbb{C}^{n-1} \right\} \subset \mathcal{H}(n).$$

For all  $x, y \in \mathbb{C}^{n-1}$  with  $x^*y = 0$  we have

$$\begin{bmatrix} 0_{n-1} & x \\ x^* & 0 \end{bmatrix} \begin{bmatrix} x \\ \pm \|x\| \end{bmatrix} = \pm \|x\| \begin{bmatrix} x \\ \pm \|x\| \end{bmatrix}, \quad \begin{bmatrix} 0_{n-1} & x \\ x^* & 0 \end{bmatrix} \begin{bmatrix} y \\ 0 \end{bmatrix} = 0.$$

Hence, if  $A = (A_1, \dots, A_{2(n-1)})^*$  is a basis of  $X_n$  over  $\mathbb{R}$ , the largest and the smallest eigenvalues of any  $\eta^T A$  are simple.

Examples of  $A \in \mathcal{H}(n)^m$  with a block of eigenvalues of arbitrarily high constant multiplicity can be constructed via tensor products. The proposition below does this for an important special case.

**Proposition 4.8** *Let  $B \in \mathcal{H}(n)^m$ . Suppose that  $\lambda_1(\eta^T B)$  is simple for all  $\eta \in S^{m-1}$ . Let  $r > \mu \geq 1$  be arbitrary, and let  $C \in \mathcal{H}(r)$  be a positive semi-definite matrix such that  $\lambda_1(C) = \lambda_\mu(C) > \lambda_{\mu+1}(C)$ . Set  $A_j = B_j \otimes C$  for  $1 \leq j \leq m$ , and let  $A = (A_1, \dots, A_m)^* \in \mathcal{H}(rn)^m$ . Then  $\lambda_1(\eta^T A)$  has multiplicity  $\mu$  for all  $\eta \in S^{m-1}$ .*

**Proof:** The eigenvalues of  $\eta^T A$  are  $\lambda_i(\eta^T B)\lambda_j(C)$ . □

For completeness we mention the method given in [11] to construct linear families of hermitian matrices with eigenvalues of constant multiplicities. Let  $m = \rho(n, \mathbb{C})$ , where  $\rho(n, \mathbb{C})$  is defined as in [11]. Then there exists a  $m$ -tuple  $U = (U_1, \dots, U_m)^*$  of unitary  $n \times n$ -matrices such that  $\eta^T U$  is unitary for all  $\eta \in S^{m-1}$ . Let  $A_0 \in \mathcal{H}(n)$  and set  $A = (A_1, \dots, A_m)^*$  where  $A_k = \begin{bmatrix} 0 & A_0 U_k^* \\ U_k A_0 & 0 \end{bmatrix}$ . Then all matrices  $\eta^T A$ ,  $\eta \in S^{m-1}$ , have the same eigenvalues: let  $x \in \mathbb{C}^n$  be an eigenvector of  $A_0$  such that  $A_0 x = \lambda_0 x$ . Then we have for all  $\eta \in S^{m-1}$ ,  $(\eta^T A) [x^T, \pm((\eta^T U)x)^T]^T = (\pm \lambda_0) [x^T, \pm((\eta^T U)x)^T]^T$ .

### 4.3 Genericity of simple eigenvalues

This subsection deals with the likelihood of having eigenvalue crossing in a linear  $m$ -parameters family of hermitian matrices, a problem initiated by von Neumann and Wigner [24]. They correctly pointed out that, for  $m \leq 3$ , eigenvalue crossing does not “in general” occur; here, we further prove the openness and density of the noncrossing property in Proposition 4.10. The dimension formula of Theorem 4.12 for a specific crossing pattern is available in [24]; here, we further investigate the topological properties of the set of matrices exhibiting the crossing pattern.

Let  $\mathcal{H}_0(n, m) \subset \mathcal{H}(n)^m$  be the set of  $A \in \mathcal{H}(n)^m$  such that for any  $\eta \in \mathbb{R}^m \setminus \{0\}$  all eigenvalues of  $\eta^T A$  are simple. The following statement is an immediate



consequence of Corollary 4.3 and Corollary 4.6. We denote a disjoint union by  $\uplus$ .

**Corollary 4.9** *Let  $A \in \mathcal{H}_0(n, m)$ . For  $1 \leq k \leq n$  and  $\eta \in \mathbb{R}^m \setminus \{0\}$  set  $\phi_k(\eta) = E_k(\eta^T A)$ . For  $k \neq (n+1)/2$  the maps  $\phi_k : S^{m-1} \rightarrow \mathbb{C}\mathbb{P}^{n-1}$  are real analytic embeddings. We have  $\mathcal{C}(F_A) = \uplus_{k=1}^{\lfloor (n+1)/2 \rfloor} \phi_k(S^{m-1})$ .*

Corollary 4.9 is subordinate to the simplicity condition of all eigenvalues, which as we show here below is generically satisfied when  $m \leq 3$ .

**Proposition 4.10** *Let  $m \leq 3$ . Then  $\mathcal{H}_0(n, m)$  is open and dense in  $\mathcal{H}(n)^m$ .*

Our proof of Proposition 4.10 relies on the Lemma and the Theorem below. Let  $M$  be a differentiable manifold. A subset,  $N \subset M$ , has measure 0 if for any coordinate chart  $U \subset M$  the Lebesgue measure of  $U \cap N$  is 0. If  $f : M_1 \rightarrow M_2$  is a smooth map of differentiable manifolds, and  $\dim M_1 < \dim M_2$  then  $f(M_1) \subset M_2$  is a set of measure 0. See, e. g., [3, 23]. A submanifold  $M \subset \mathbb{R}^n$  is  $\mathbb{R}^*$ -homogeneous if  $\mathbb{R}^*M = M$ , where  $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ .

**Lemma 4.11** *Let  $X$  be a finite dimensional real vector space, and let  $M \subset X$  be a  $\mathbb{R}^*$ -homogeneous submanifold. Let  $N_j$  be the set of  $j$ -tuples  $(x_1, \dots, x_j) \in X^j$  such that  $\text{span}\{x_1, \dots, x_j\} \cap M \neq \emptyset$ . If  $1 \leq j \leq \text{codim } M$ , then  $N_j \subset X^j$  has measure zero.*

**Proof:** The projectivization of  $M$  is a proper submanifold (and hence a subset of measure 0) of the projective space  $\mathbb{P}(X)$ . Thus the claim holds for  $j = 1$ . Now, let  $j$  be arbitrary. For  $\xi = (x_1, \dots, x_j) \in X^j$ , let  $X_\xi^c$  be a subspace of  $X$  such that  $X = X_\xi^c \oplus \text{span}(\xi)$  and let  $P_\xi^c : X \rightarrow X_\xi^c$  be the linear projection onto  $X_\xi^c$  along  $\text{span}(\xi)$ . Furthermore, let  $Q_j(\xi) = P_\xi^c(M) + \text{span}(\xi) \subseteq X$ . If  $j < \text{codim } M$ , then  $\dim M < \dim X_\xi^c$ , and hence  $P_\xi^c(M)$  is a subset of measure 0 of  $X_\xi^c$ . Then, by Fubini's Theorem,  $Q_j(\xi)$  is a subset of measure 0 of  $X$ . Let  $x, \tilde{x} \in X$  and suppose that  $\tilde{x} \notin \text{span}(\xi)$ . Then  $\mathbb{R}\tilde{x} \subset \text{span}(\xi, x)$  iff  $P_\xi^c(\mathbb{R}\tilde{x}) = P_\xi^c(\mathbb{R}x)$  iff  $x \in P_\xi^c(\mathbb{R}\tilde{x}) + \text{span}(\xi)$ . Using these equivalences with  $\tilde{x} \in M$ , it is straightforward to verify that the sets  $N_j = \{\xi \in X^j \mid M \cap \text{span}(\xi) \neq \emptyset\}$  satisfy  $N_{j+1} = (N_j \times X) \cup \{(\xi, x) \mid \xi \in X^j \setminus N_j, x \in Q_j(\xi)\}$ . Thus, we obtain the following statement: If  $N_j$  has measure 0 and  $j < \text{codim } M$ , then  $N_{j+1}$  as measure 0. But  $N_1$  has measure 0.  $\square$

For  $n_1, \dots, n_r \in \mathbb{N}$  such that  $\sum_{j=1}^r n_j = n$ , let  $\mathcal{H}(n; n_1, \dots, n_r)$  denote the set of  $A \in \mathcal{H}(n)$  such that  $\lambda_1(A) = \lambda_{n_1}(A) > \lambda_{n_1+1}(A)$ ,  $\lambda_{n_1+1}(A) = \lambda_{n_1+n_2}(A) > \lambda_{n_1+n_2+1}(A)$ , etc.

**Theorem 4.12** Any  $\mathcal{H}(n; n_1, \dots, n_r)$  is a real analytic  $\mathbb{R}^*$ -homogenous submanifold of  $\mathcal{H}(n)$ . Furthermore,

$$\text{codim } \mathcal{H}(n; n_1, \dots, n_r) = \left( \sum_{j=1}^r n_j^2 \right) - r.$$

We defer the proof of this Theorem to Appendix B. Note that, if  $n_j > 1$  for at least one  $j$ , then  $\text{codim } \mathcal{H}(n; n_1, \dots, n_r) \geq 3$ . Equality holds if there is an index  $j_0$  such that  $n_{j_0} = 2$  and  $n_j = 1$  for all  $j \neq j_0$ . The union of the  $\mathcal{H}(n; n_1, \dots, n_r)$ 's over all sequences except  $1, \dots, 1$  is the real algebraic variety,  $\mathcal{V}_n$ , of hermitian  $n \times n$  matrices with multiple eigenvalues. The following corollary is immediate from Theorem 4.12.

**Corollary 4.13** The variety  $\mathcal{V}_n$  has codimension 3 in  $\mathcal{H}(n)$ .

**Proof of Proposition 4.10:** The set  $\mathcal{H}_0(n, m)$  is open for arbitrary values of  $n$  and  $m$ . Thus, it suffices to show that for  $m \leq 3$  and  $n \geq 2$  the set  $\mathcal{C}_0(n, m) := \mathcal{H}(n)^m \setminus \mathcal{H}_0(n, m)$  has an empty interior. By definition,  $\mathcal{C}_0(n, m)$  is the set of  $A \in \mathcal{H}(n)^m$  such that  $\text{span}\{A_1, \dots, A_m\} \cap \mathcal{V}_n \neq \emptyset$ . By Corollary 4.13 and Lemma 4.11,  $\mathcal{C}_0(n, 1)$ ,  $\mathcal{C}_0(n, 2)$ , and  $\mathcal{C}_0(n, 3)$  are sets of measure zero in  $\mathcal{H}(n)^m$  for  $m = 1, 2, 3$ , respectively.  $\square$

## 5 Convexity of numerical range

### 5.1 The highest eigenvalue and the convexity

The following theorem is the main result of this paper.

**Theorem 5.1** Let  $A \in \mathcal{H}(n)^m$  be such that  $\lambda_1(\eta^T A)$ ,  $\eta \in S^{m-1}$ , has constant multiplicity. Suppose, in addition, that  $\bigcup_{\eta \in S^{m-1}} E_1(\eta^T A) \neq \mathbb{C}^n$ . Then  $\mathcal{F}(A)$  is convex.

**Remark 2.** If the dimensional parameters of  $A \in \mathcal{H}(n)^m$  satisfy the inequality (15), then, by the proof of Proposition 4.7,  $\dim \mathcal{P} < \dim \mathbb{C}\mathbb{P}^{n-1}$  and the extra assumption in the above theorem is fulfilled. Thus, by Proposition 4.7, the additional assumption is redundant, unless  $m = n+1$  and the multiplicity of the highest eigenvalue is  $n/2$ .

We will need the propositions below for the proof of Theorem 5.1.

**Proposition 5.2** Let  $A \in \mathcal{H}(n)^m$ . Suppose that the largest eigenvalue of  $\eta^T A$ ,  $\eta \in S^{m-1}$ , has constant multiplicity. Then  $\partial_0 \mathcal{F}(A) \subset \mathbb{R}^m$  is a real analytic submanifold diffeomorphic to  $S^{m-1}$ . Let  $s$  be as in Proposition 3.5, let

$\mathcal{S}$  and  $q$  be as in Theorem 4.5, and define  $\pi := (\nabla s) \circ q : \mathcal{S} \rightarrow \partial_0 \mathcal{F}(A)$ . Then  $(\mathcal{S}, \pi, \partial_0 \mathcal{F}(A))$  is a real analytic, locally trivial sphere bundle.

**Proof:** The claims follow from Theorems 3.8, 3.9, and 4.5. □

We will need the following basic fact. For convenience of the reader, we will sketch a proof.

**Proposition 5.3** *Let  $C \in \mathcal{C}_m$ , and suppose that  $\text{int}(C) \neq \emptyset$ . Let  $x_0 \in \text{int}(C)$ . For  $x \in \mathbb{R}^m, x \neq x_0$ , let  $R(x)$  denote the ray from  $x_0$  containing  $x$ . Then  $R(x) \cap \partial C$  consists of a unique point,  $r(x)$ . The map  $r : \mathbb{R}^m \setminus \{x_0\} \rightarrow \partial C$  is a continuous retraction.*

**Proof:** We assume, without loss of generality, that  $x_0 = 0$ . Set  $f(x) = \inf\{t > 0 \mid x/t \in C\}$ . Then  $r(x) = x/f(x)$ . The function  $f > 0$  is convex, hence continuous. See, e. g., [16, Theorem 2.1.23]. □

**Proof of Theorem 5.1:** Let  $\mathcal{S}$  and  $\pi : \mathcal{S} \rightarrow \partial \text{co}(\mathcal{F}(A))$  be as in Proposition 5.2. Assume that the claim fails. Then there exists a point  $y_0 \in \text{int}(\text{co}(\mathcal{F}(A))) \setminus \mathcal{F}(A)$ . For  $y \in \mathbb{R}^m \setminus \{y_0\}$ , let  $r(y)$  be the point of intersection of the ray  $R(y) = \{y_0 + t(y - y_0) \mid t \geq 0\}$  with  $\partial \text{co}(\mathcal{F}(A))$ . By Proposition 5.3, the map  $r : \mathbb{R}^m \setminus \{y_0\} \rightarrow \partial \text{co}(\mathcal{F}(A))$  is a continuous retraction. By the additional assumption of Theorem 5.1, there exists a unit vector  $z_0 \in \mathbb{C}^n \setminus \mathcal{S}$ . Then  $(1 - t)z + tz_0 \neq 0$  and  $F_A([(1 - t)z + tz_0]) \neq y_0$ , for all  $t \in [0, 1]$  and any  $z \in \mathcal{S}$ . Thus, we have constructed the homotopy:

$$h : [0, 1] \times \mathcal{S} \rightarrow \partial \text{co}(\mathcal{F}(A)), \quad h(t, z) = r \circ F_A([(1 - t)z + tz_0]).$$

Since  $h(0, \cdot) = \pi$  and  $h(1, \cdot)$  is a constant map, it follows that  $\pi$  is homotopically trivial. Since  $\pi : \mathcal{S} \rightarrow \partial \text{co}(\mathcal{F}(A))$  is a sphere bundle over a sphere, the latter is impossible. <sup>1</sup> □

**Remark 3.** The examples of § 2.2 show that we cannot suppress the additional assumption in Theorem 5.1. Set  $A = (\sigma_1, \sigma_2, \sigma_3)^*$ . Then the highest eigenvalue of  $\eta^T A$  is simple for all  $\eta$ , but  $\bigcup_{\eta \in S^2} E_1(\eta^T A) = \mathbb{C}^2$ . Thus, the additional assumption of the Theorem does not hold. And, indeed,  $\mathcal{F}(A) = S^2$ .

---

<sup>1</sup>We sketch a proof: The homotopy  $h : [0, 1] \times \mathcal{S} \rightarrow \partial \text{co}(\mathcal{F}(A))$  has a continuous lift  $g : [0, 1] \times \mathcal{S} \rightarrow \mathcal{S}$  such that  $h = \pi \circ g$  and  $g(0, \cdot) = \text{id}_{\mathcal{S}}$  [33, Theorem 7.13]. The map  $g(1, \cdot)$  has degree 1 and is therefore surjective [15]. The latter implies that  $h(1, \cdot)$  is surjective and hence non constant.

As an application of the preceding results, we will provide a new proof of the following known theorem [1, 8]:

**Theorem 5.4** *Let  $A \in \mathcal{H}(n)^m$ , where  $1 \leq m \leq 3$  and where  $n$  is arbitrary unless  $m = 3$ , in which case  $n \geq 3$ . Then  $\mathcal{F}(A)$  is convex.*

We need some preliminaries. For any  $n, m$  let  $\mathcal{C}(n, m)$  be the set of  $A \in \mathcal{H}(n)^m$  such that  $\mathcal{F}(A)$  is convex.

**Proposition 5.5** *The set  $\mathcal{C}(n, m)$  is closed in  $\mathcal{H}(n)^m$ .*

**Proof:** By Proposition 2.7, the mapping  $A \mapsto \mathcal{F}(A)$  from  $\mathcal{H}(n)^m$  to  $\mathcal{K}_m$  is continuous. But  $\mathcal{C}(n, m)$  is the preimage of the closed set  $\mathcal{C}_n \subset \mathcal{K}_m$ .  $\square$

**Proof of Theorem 5.4:** The case  $m = 1$  is trivial. For  $m = 2$ , this is the celebrated Töplitz- Hausdorff theorem. See, e. g., [12]. We will prove the claim for  $2 \leq m \leq 3$ . The inequality  $m \leq 2(n - 1)$  makes the extra assumption of Theorem 5.1 redundant. Thus,  $\mathcal{H}_0(n, m) \subset \mathcal{C}(n, m)$ . Since  $m \leq 3$ , Proposition 4.10 applies. Hence,  $\mathcal{C}(n, m)$  is dense in  $\mathcal{H}(n)^m$ . But, by Proposition 5.5,  $\mathcal{C}(n, m)$  is closed.  $\square$

## 5.2 Stable convexity

Let  $\|\cdot\|_S$  and  $\|\cdot\|_F$  be the spectral norm and the Frobenius norm on the space of  $n \times n$  matrices, respectively. Let  $d_S, d_F$  denote the translation invariant distance functions on  $\mathcal{H}(n)^m$  defined by  $d_\bullet(A, 0) = \max_{\eta \in S^{m-1}} \|\eta^T A\|_\bullet$ , where  $\bullet = S, F$ . The following theorem is the main result of this Section.

**Theorem 5.6** *Let  $m \geq 4$ . Then*

$$\text{int } \mathcal{C}(n, m) = \left\{ A \in \mathcal{H}(n)^m \mid \lambda_1(\eta^T A) \text{ is simple for all } \eta \in S^{m-1} \right\}.$$

*Let  $A \in \mathcal{C}(n, m)$ . Then*

$$\sqrt{2}d_F(A, \partial\mathcal{C}(n, m)) = 2d_S(A, \partial\mathcal{C}(n, m)) = \min_{\eta \in S^{m-1}} (\lambda_1(\eta^T A) - \lambda_2(\eta^T A)).$$

We will need a few auxilliary results.

**Lemma 5.7** *Let  $m \geq 4$ . Let  $A_1 \in \mathcal{H}(n)$  be such that  $\dim E_1(A_1) = 2$ . Set  $\mathcal{N}(A_1) = \left\{ (A_2, \dots, A_m)^* \in \mathcal{H}(n)^{m-1} \mid \mathcal{F}(A_1, A_2, \dots, A_m) \text{ is not convex} \right\}$ . Then  $\mathcal{N}(A_1)$  is open and dense in  $\mathcal{H}(n)^{m-1}$ .*

**Proof:** Set  $A = (A_1, A_2, \dots, A_m)^*$ . Let  $(z_1, z_2)$  be an orthonormal basis of  $E_1(A_1)$ . For  $1 \leq j \leq m$ , let  $B_j$  be the Gram matrix of  $A_j$  with respect to  $(z_1, z_2)$ . Then  $B_1 = \lambda_1(A_1)I_2$ . Set  $B = (B_2, \dots, B_m)^*$ . By Example 2,  $B = M\sigma + pI_2$ ,  $p \in \mathbb{R}^{m-1}$ , where

$$M = \begin{bmatrix} \Re(z_2^* A_2 z_1) & \Im(z_2^* A_2 z_1) & \frac{1}{2}(z_1^* A_2 z_1 - z_2^* A_2 z_2) \\ \vdots & \vdots & \vdots \\ \Re(z_2^* A_m z_1) & \Im(z_2^* A_m z_1) & \frac{1}{2}(z_1^* A_m z_1 - z_2^* A_m z_2) \end{bmatrix} \in \mathbb{R}^{(m-1) \times 3}.$$

By Proposition 3.5 and Example 2,  $\mathcal{F}(A) \cap H(e_1, \lambda_1(A_1)) = \mathcal{F}(A, E_1(A_1)) = \mathcal{F}(B_1, \dots, B_m) = \{\lambda_1(A_1)\} \times (M(S^2) + p)$ . If this intersection is not convex (that is,  $\text{rank}(M) = 3$  by Example 2), then  $\mathcal{F}(A)$  is not convex. It follows that  $\mathcal{N}(A_1)$  contains those  $(A_2, \dots, A_m)^* \in \mathcal{H}(n)^{m-1}$  such that  $\text{rank } M = 3$ . Hence, the complement of  $\mathcal{N}(A_1)$  is contained in a closed subvariety of  $\mathcal{H}(n)^{m-1}$ .  $\square$

**Proposition 5.8** *Let  $m \geq 4$ . Let  $A \in \mathcal{H}(n)^m$  be such that  $\dim E_1(\eta^T A) > 1$  for some  $\eta \in S^{m-1}$ . Then either  $\mathcal{F}(A)$  is nonconvex, or  $A \in \partial\mathcal{C}(n, m)$ .*

**Proof:** We can assume that  $A_1$  is diagonal and  $\dim E_1(A_1) > 1$ . Suppose that  $A \in \mathcal{C}(n, m)$ . Let  $A'_1$  be a diagonal matrix such that  $\dim E_1(A'_1) = 2$ . By Lemma 5.7, arbitrarily close to  $(A'_1, A_2, \dots, A_m)^*$ , there are  $A' = (A'_1, A'_2, \dots, A'_m)^* \in \mathcal{H}(n)^m \setminus \mathcal{C}(n, m)$ . Since  $A'_1$  can be chosen arbitrarily close to  $A_1$ , we obtain a sequence  $A^{(k)} \in \mathcal{H}(n)^m \setminus \mathcal{C}(n, m)$  converging to  $A$ .  $\square$

Let  $\mathcal{M} \subset \mathcal{H}(n)$  be the set of  $M$  such that  $\lambda_1(M)$  is a multiple eigenvalue.

**Lemma 5.9** *Let  $A_0 \in \mathcal{H}(n)$ . Then*

$$\sqrt{2}d_F(A_0, \mathcal{M}) = 2d_S(A_0, \mathcal{M}) = \lambda_1(A_0) - \lambda_2(A_0). \quad (16)$$

**Proof:** Let  $A, B \in \mathcal{H}(n)$  be arbitrary. Then  $\|A - B\|_S \geq |\lambda_k(A) - \lambda_k(B)|$  for any  $1 \leq k \leq n$ , and  $\|A - B\|_F^2 \geq \sum_{k=1}^n (\lambda_k(B) - \lambda_k(A))^2$ . See, e. g., [30, Cor. 4.10, Cor. 4.13]. Applying this to a pair  $A_0, M$ , where  $M \in \mathcal{M}$ , we obtain  $2d_S(A_0, M) \geq |\lambda_1(A_0) - \lambda_2(A_0)|$ ,  $2d_F(A_0, M)^2 \geq (\lambda_1(A_0) - \lambda_2(A_0))^2$ . This yields lower bounds on the distances in equation (16).

The distance functions are invariant under the conjugation by unitary matrices. The set  $\mathcal{M}$  is also invariant. Hence, we can assume that  $A_0$  is diagonal. Letting  $M$  vary over the set of diagonal matrices in  $\mathcal{M}$ , we attain

the bounds. □

**Proof of Theorem 5.6:** Let  $\mathcal{S}(n, m) \subset \mathcal{H}(n)^m$  be the open subset of  $A$ 's such that  $\lambda_1(\eta^T A)$  is simple for all  $\eta \in S^{m-1}$ . By Theorem 5.1,  $\mathcal{S}(n, m) \subset \mathcal{C}(n, m)$ . By Proposition 5.8,  $\mathcal{C}(n, m) \setminus \mathcal{S}(n, m) \subset \partial\mathcal{C}(n, m)$ . This proves the first claim. The second is immediate from Lemma 5.9. □

## 6 Conclusion

The main point of this paper is that, in view of the no crossing criterion for the largest eigenvalue of a family of matrices parameterized by a sphere, convexity of the joint numerical range is essentially a topological issue. This “noncrossing” issue is in fact very general and appears in a variety of other problems—e.g., system balancing [32], quantum mechanics [24], etc.

## 7 Appendix A

We will derive Theorem 3.7 from the following result.

**Theorem 7.1** *Let  $A_0 \in \mathcal{H}_{k,\mu}(n)$ , and let  $\lambda_0 = \lambda_k(A_0)$ . Let  $(v_{01}, \dots, v_{0\mu})$  be an orthonormal basis of  $E_k(A_0)$ . Then there exists an open neighbourhood  $U \subset \mathcal{H}(n)$  of  $A_0$  and real analytic functions  $v_j : U \rightarrow \mathbb{C}^n$ ,  $1 \leq j \leq \mu$ , such that for any  $A \in U$  the vectors  $v_1(A), \dots, v_\mu(A)$  form an orthonormal basis of  $\sum_{j=k}^{k+\mu-1} E_j(A)$  satisfying  $(v_1(A_0), \dots, v_\mu(A_0)) = (v_{01}, \dots, v_{0\mu})$ . The differential of  $v_j$  at  $A_0$  satisfies*

$$d_{A_0} v_j(\Delta) = (\lambda_0 I_n - A_0)^\dagger \Delta v_{0j}, \quad \Delta \in \mathcal{H}(n). \quad (17)$$

**Proof:** Let  $V_0 := [v_{01}, \dots, v_{0\mu}] \in \mathbb{C}^{n \times \mu}$  and

$$h : \mathbb{C}^{n \times \mu} \rightarrow \mathcal{H}(\mu), \quad h(V) := j(V_0^* V - V^* V_0).$$

Note that  $h$  is onto since  $h(-\frac{1}{2}V_0 X) = X$  for all  $X \in \mathcal{H}(\mu)$ . Therefore

$$\dim_{\mathbb{R}} \ker h = 2n\mu - \mu^2. \quad (18)$$

Let

$$f : \mathcal{H}(n) \times (\ker h \times \mathbb{C}^{\mu \times \mu}) \rightarrow \mathbb{C}^{n \times \mu} \times \mathcal{H}(\mu), \quad f(A, (V, L)) := \begin{bmatrix} AV - VL \\ V^* V - I_\mu \end{bmatrix}.$$

We are going to prove the theorem by applying the implicit function theorem to the equation  $f(A, (V, L)) = 0$ . By our assumptions on  $V_0$ , we already have  $f(A_0, (V_0, \lambda_0 I_\mu)) = 0$ . The differential of  $f$  at  $(A, (V, L))$  is

$$d_{(A, (V, L))}f(\Delta_1, \Delta_2, \Delta_3) = \begin{bmatrix} \Delta_1 V + A\Delta_2 - \Delta_2 L - V\Delta_3 \\ \Delta_2^* V + V^* \Delta_2 \end{bmatrix},$$

where  $(\Delta_1, \Delta_2, \Delta_3) \in \mathcal{H}(n) \times (\ker h \times \mathbb{C}^{\mu \times \mu})$ . In particular

$$d_{(A_0, (V_0, \lambda_0 I_\mu))}f(0, \Delta_2, \Delta_3) = \begin{bmatrix} (A_0 - \lambda_0 I_n)\Delta_2 - V_0\Delta_3 \\ \Delta_2^* V_0 + V_0^* \Delta_2 \end{bmatrix}.$$

The columns of  $V_0$  form an orthonormal basis of  $\ker(A_0 - \lambda_0 I_n)$ . Thus  $I_n - V_0 V_0^*$  is the orthogonal projector onto  $(\ker(A_0 - \lambda_0 I_n))^\perp$ . Hence

$$(A_0 - \lambda_0 I_n)^\dagger (A_0 - \lambda_0 I_n) = (A_0 - \lambda_0 I_n)(A_0 - \lambda_0 I_n)^\dagger = I_n - V_0 V_0^*.$$

Using this fact, it is easily verified that

$$d_{(A_0, (V_0, \lambda_0 I_\mu))}f\left(0, \frac{1}{2}V_0 Y + (A_0 - \lambda_0 I_n)^\dagger X, -V_0^* X\right) = \begin{bmatrix} X \\ Y \end{bmatrix}$$

for all  $X \in \mathbb{C}^{n \times \mu}$ ,  $Y \in \mathcal{H}(\mu)$ . Thus the map

$$d_{(A_0, (V_0, \lambda_0 I_\mu))}f(0, \cdot, \cdot) : \ker h \times \mathbb{C}^{\mu \times \mu} \rightarrow \mathbb{C}^{n \times \mu} \times \mathcal{H}(\mu)$$

is onto. However, from (18), it follows that  $\ker h \times \mathbb{C}^{\mu \times \mu}$  and  $\mathbb{C}^{n \times \mu} \times \mathcal{H}(\mu)$  have the same real dimension. Thus  $d_{(A_0, (V_0, \lambda_0 I_\mu))}f(0, \cdot, \cdot)$  is bijective and

$$\left(d_{(A_0, (V_0, \lambda_0 I_\mu))}f(0, \cdot, \cdot)^{-1}\right)\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}V_0 Y + (A_0 - \lambda_0 I_n)^\dagger X \\ -V_0^* X \end{bmatrix}.$$

Now, the implicit function theorem for real analytic functions [4, Theorem 10.2.4] yields existence of a neighbourhood  $U' \subset \mathcal{H}(n)$  of  $A_0$  and a real analytic map  $(V, L) : U' \rightarrow \mathbb{C}^{n \times \mu} \times \mathcal{H}(\mu)$  such that

$$(V(A_0), L(A_0)) = (V_0, \lambda_0 I_\mu) \tag{19}$$

and

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = f(A, (V(A), L(A))) = \begin{bmatrix} AV(A) - V(A)L(A) \\ V(A)^* V(A) - I_\mu \end{bmatrix} \tag{20}$$

for all  $A \in U'$ . The differential of the map  $(V, L)$  at  $A_0$  is given by

$$\begin{aligned} \begin{bmatrix} d_{A_0} V(\Delta) \\ d_{A_0} L(\Delta) \end{bmatrix} &= -\left(d_{(A_0, (V_0, \lambda_0 I_\mu))}f(0, \cdot, \cdot)\right)^{-1} d_{(A_0, (V_0, \lambda_0 I_\mu))}f(\Delta, 0, 0) \\ &= \begin{bmatrix} -(A_0 - \lambda_0 I_n)^\dagger \Delta V_0 \\ V_0^* \Delta V_0 \end{bmatrix}, \quad \Delta \in \mathcal{H}(n). \end{aligned} \tag{21}$$

Let  $(x_1(A), \dots, x_\mu(A))$  be an orthonormal basis of eigenvectors of  $L(A)$  such that

$$L(A)x_j(A) = \lambda_j(L(A))x_j(A), \quad j = 1, \dots, \mu.$$

Then it follows from (20) that  $(V(A)x_1(A), \dots, V(A)x_\mu(A))$  is an orthonormal system of eigenvectors of  $A$  corresponding to the same eigenvalues. Thus

$$\{\lambda_1(L(A)), \dots, \lambda_\mu(L(A))\} \subseteq \{\lambda_1(A), \dots, \lambda_n(A)\}. \quad (22)$$

By (19) we have

$$\lambda_1(L(A_0)) = \dots = \lambda_\mu(L(A_0)) = \lambda_0 = \lambda_k(A_0) = \dots = \lambda_{k+\mu-1}(A_0). \quad (23)$$

The eigenvalue functions  $A \mapsto \lambda_j(A)$  are continuous. Hence (22), (23) and the fact that  $\lambda_0 \neq \lambda_\ell(A_0)$  for  $\ell \notin \{k, \dots, k + \mu - 1\}$  imply that

$$\lambda_j(L(A)) = \lambda_{k+j-1}(A), \quad j = 1, \dots, \mu, \quad (24)$$

$$\lambda_j(L(A)) \neq \lambda_\ell(A), \quad \ell \notin \{k, \dots, k + \mu - 1\} \quad (25)$$

for all  $A$  in a neighbourhood  $U \subset U' \subset \mathcal{H}(n)$  of  $A_0$ . Therefore

$$\text{range } V(A) = \text{span}\{V(A)x_1(A), \dots, V(A)x_\mu(A)\} = \sum_{j=k}^{k+\mu-1} E_j(A). \quad (26)$$

Finally, let  $v_j(A)$  denote the  $j^{\text{th}}$  column of  $V(A)$ . Then by (19), (21) and (26), the functions  $v_j : U \rightarrow \mathbb{C}^n$  have the properties required in the theorem.  $\square$

**Proof of Theorem 3.7:** Using charts, the proof of Theorem 3.7 can be reduced to the case where the manifold  $M$  is an open subset of a real Banach space  $X$  and  $H : M \rightarrow \mathcal{H}_{k,\mu}(n)$  is a  $\mathcal{C}^r$ -map,  $r \in \mathbb{N}^* \cup \{\infty, \omega\}$ .

Let  $x_0 \in M$  and let  $(z_{01}, \dots, z_{0\mu})$  be an orthonormal basis of the eigenspace  $E_k(H(x_0))$ . As in Theorem 7.1, choose an open neighbourhood  $U \subset \mathcal{H}(n)$  of  $H(x_0)$  and functions  $v_j : U \rightarrow \mathbb{C}^n$ ,  $j = 1, \dots, \mu$ , such that

$$(v_1(H(x_0)), \dots, v_\mu(H(x_0))) = (z_{01}, \dots, z_{0\mu}).$$

Set  $z_j = v_j \circ H|_{U_0}$ , where  $U_0 = H^{-1}(U \cap \mathcal{H}_{k,\mu}(n))$ . Then  $(z_1(x), \dots, z_\mu(x))$  is an orthonormal basis of  $E_k(H(x))$  for all  $x \in U_0$ . Applying the chain rule to (17), we obtain

$$d_{x_0} z_j(\xi) = (\lambda_k(H(x_0))I_n - H(x_0))^\dagger d_{x_0} H(\xi) z_{0j}, \quad \xi \in T_{x_0} M = X. \quad (27)$$



Now set  $w(x) = \sum_{j=1}^{\mu} \alpha_j z_j(x)$  and  $z(x) = \sum_{j=1}^{\mu} \beta_j z_j(x)$ , where  $\alpha_j, \beta_j \in \mathbb{C}$ . Then, for all  $x \in U_0$ ,

$$(\lambda_k \circ H)(x) w(x)^* z(x) = w(x)^* H(x) z(x). \quad (28)$$

Moreover, the function  $x \mapsto w(x)^* z(x) = \sum_{j=1}^{\mu} \bar{\alpha}_j \beta_j$  is constant. Differentiating the relation (28), we obtain

$$\begin{aligned} d_x(\lambda_k \circ H)(\xi) w(x)^* z(x) &= w(x)^* d_x H(\xi) z(x) \\ &\quad + d_x w(\xi)^* H(x) z(x) + w(x)^* H(x) d_x z(\xi) \\ &= w(x)^* d_x H(\xi) z(x) \\ &\quad + \lambda_k(H(x)) \underbrace{(d_x w(\xi)^* z(x) + w(x)^* d_x z(\xi))}_{=d_x(w(\cdot)^* z(\cdot))(\xi)=0} \\ &= w(x)^* d_x H(\xi) z(x). \end{aligned}$$

In the special case  $z(x) = w(x)$ ,  $\|z(x)\| = 1$ , we have

$$d_x(\lambda_k \circ H)(\xi) = z(x)^* d_x H(\xi) z(x).$$

Suppose now that  $H$  is twice differentiable. Set  $g(x) = z(x)^* d_x H(\xi_1) z(x)$  for a fixed  $\xi_1 \in T_x M$ . Differentiating  $g$ , we obtain the second derivative of  $\lambda_k \circ H$  as

$$d_x^2(\lambda_k \circ H)(\xi_1, \xi_2) = d_x g(\xi_2) = z(x)^* d_x^2 H(\xi_1, \xi_2) z(x) + 2z(x)^* d_x H(\xi_1) d_x z(\xi_2).$$

From (27), it follows that  $d_x z(\xi_2) = (\lambda_k(H(x))I_n - H(x))^\dagger d_x H(\xi_2) z(x)$ . Combining the latter two equations, the final result follows.  $\square$

The above theorem says that, if  $M(x) \in \mathcal{H}_{k,\mu}(n)$  is a smooth family, then the eigenvectors *associated with the constant multiplicity eigenvalue*  $\lambda_k(M(x))$  are still smooth, whereas, in general, the eigenvectors associated with a smooth family cannot even be guaranteed to be continuous (see [20, Remark II.6.9]).

## 8 Appendix B

**Proof of Theorem 4.12:** Let  $n, n_1, \dots, n_r \in \mathbb{N}$  be such that  $\sum_{k=1}^r n_k = n$ . Set  $\mathcal{D} = \{ \text{diag}(\mu_1 I_{n_1}, \dots, \mu_r I_{n_r}) \mid \mu_k \in \mathbb{R}, \mu_1 > \dots > \mu_r \} \subset \mathcal{H}(n)$  and  $\mathcal{D}' = \{ \text{diag}(A_1, \dots, A_r) \mid A_k \in \mathbb{C}^{n_k \times n_k} \} \subset \mathbb{C}^{n \times n}$ . It is easily seen that  $\mathcal{D}'$  is the centralizer of each element of  $\mathcal{D}$  in  $\mathbb{C}^{n \times n}$ , i.e., for all  $D \in \mathcal{D}$  and all

$M \in \mathbb{C}^{n \times n}$ ,  $MD = DM$  iff  $M \in \mathcal{D}'$ . By  $\mathcal{U}(n)$  we denote the set of all unitary  $n \times n$  matrices. It is a compact connected real Lie group of real dimension  $\dim_{\mathbb{R}} \mathcal{U}(n) = n^2$ . Its tangent spaces are

$$T_U \mathcal{U}(n) = \{ jAU \mid A \in \mathcal{H}(n) \}, \quad U \in \mathcal{U}(n).$$

$\mathcal{D}$  is a submanifold of  $\mathcal{H}(n)$  of dimension  $\dim_{\mathbb{R}} \mathcal{D} = r$ . Its tangent spaces are

$$T_D \mathcal{D} = \{ \text{diag}(\delta_1 I_{n_1}, \dots, \delta_r I_{n_r}) \mid \delta_k \in \mathbb{R} \}, \quad D \in \mathcal{D}.$$

In order to show that  $\mathcal{H}(n; n_1, \dots, n_r)$  is a submanifold of  $\mathcal{H}(n)$ , we consider the map

$$\psi : \mathcal{U}(n) \times \mathcal{D} \rightarrow \mathcal{H}(n), \quad \psi(U, D) := UDU^*.$$

Obviously,  $\mathcal{H}(n; n_1, \dots, n_r) = \psi(\mathcal{U}(n) \times \mathcal{D})$ . We will show that the differential of  $\psi$  has constant rank  $\rho$ , where  $\rho := n^2 + r - \sum_{k=1}^r n_k^2$ . To this end, we need the following easily verified lemma.

**Lemma 8.1** *For  $U \in \mathcal{U}(n)$  and  $A \in \mathcal{H}(n)$  set  $f_U(A) = UAU^*$ . Then  $f_U : \mathcal{H}(n) \rightarrow \mathcal{H}(n)$  is a linear isomorphism of  $\mathcal{H}(n)$ , and we have for all  $D \in \mathcal{D}$ ,*

$$\{ A \in \mathcal{H}(n) \mid A\psi(U, D) = \psi(U, D)A \} = f_U(\mathcal{D}' \cap \mathcal{H}(n)).$$

Moreover,  $\dim f_U(\mathcal{D}' \cap \mathcal{H}(n)) = \dim(\mathcal{D}' \cap \mathcal{H}(n)) = \sum_{k=1}^r n_k^2$ .

The differential of the real analytic map  $\psi$  at  $(U, D) \in \mathcal{U}(n) \times \mathcal{D}$  in the direction  $(\Delta_1, \Delta_2) = (jAU, \text{diag}(\delta_1 I_{n_1}, \dots, \delta_r I_{n_r})) \in T_U \mathcal{U}(n) \times T_D \mathcal{D}$  is

$$\begin{aligned} d_{(U, \mathcal{D})} \psi(\Delta_1, \Delta_2) &= \Delta_1 D U^* + U D \Delta_1^* + U \Delta_2 U^* \\ &= j(A\psi(U, D) - \psi(U, D)A) + U \Delta_2 U^* \\ &= U(j(f_U^*(A)D - Df_U^*(A)) + \Delta_2)U^*. \end{aligned}$$

Since the diagonal elements of  $f_U^*(A)D - Df_U^*(A)$  are zero, the lemma above yields that the kernel of the differential  $d_{(U, \mathcal{D})} \psi : T_U \mathcal{U}(n) \times T_D \mathcal{D} \rightarrow \mathcal{H}(n)$  is  $\ker d_{(U, \mathcal{D})} \psi = \{ (jAU, 0) \in T_U \mathcal{U}(n) \times T_D \mathcal{D} \mid A \in f_U(\mathcal{D}' \cap \mathcal{H}(n)) \}$ . Thus,  $\text{rank } d_{(U, \mathcal{D})} \psi = \dim(\mathcal{U}(n) \times \mathcal{D}) - \dim \ker d_{(U, \mathcal{D})} \psi = \rho$ . We will need the following fact.

**Proposition 8.2** *Let  $(U_0, D_0) \in \mathcal{U}(n) \times \mathcal{D}$  and let  $V \subset \mathcal{U}(n) \times \mathcal{D}$  be an open neighbourhood of  $(U_0, D_0)$ . Then there is an open neighbourhood  $W \subset \mathcal{H}(n)$  of  $\psi(U_0, D_0)$  such that  $\psi(V) \cap W = \mathcal{H}(n; n_1, \dots, n_r) \cap W$ .*

**Proof:** Suppose the claim fails. Then there are sequences  $D_k \in \mathcal{D}, U_k \in \mathcal{U}(n)$  such that (a)  $\lim_{k \rightarrow \infty} \psi(U_k, D_k) = \psi(U_0, D_0)$ , and (b)  $\psi(U_k, D_k) \notin \psi(V)$  for all  $k$ . Since the eigenvalues are continuous functions it follows from (a) that  $\lim_{k \rightarrow \infty} D_k = D_0$ . Since  $\mathcal{U}(n)$  is compact, we can assume that  $\lim_{k \rightarrow \infty} U_k = \tilde{U}$  for some  $\tilde{U} \in \mathcal{U}(n)$ . Consider now the sequence  $\tilde{U}_k = U_k \tilde{U}^* U_0$ . We have  $\lim_{k \rightarrow \infty} \tilde{U}_k = U_0$ . From the relation  $\psi(\tilde{U}, D_0) = \lim_{k \rightarrow \infty} \psi(U_k, D_k) = \psi(U_0, D_0)$  it follows that  $\tilde{U}^* U_0 \in \mathcal{D}'$ . The latter implies that  $\psi(U_k, D_k) = \psi(\tilde{U}_k, D_k)$  for all  $k$ . Thus, by (b),  $(\tilde{U}_k, D_k) \notin V$  for all  $k$ , a contradiction.  $\square$

We are now in the position to show that  $\mathcal{H}(n; n_1, \dots, n_r)$  is a submanifold of  $\mathcal{H}(n)$  of dimension  $\rho$ . Let  $q = \dim(\mathcal{U}(n) \times \mathcal{D})$  and  $(U_0, D_0) \in \mathcal{U}(n) \times \mathcal{D}$ . Recall that  $\dim \mathcal{H}(n) = n^2$ . We have seen that the differential of  $\psi$  has constant rank  $\rho$ . By the Rank Theorem [4, Theorem 10.3.1],[21, Theorem 2.5.3] there are neighborhoods  $V$  of  $(U_0, D_0)$  and  $W$  of  $\psi(U_0, D_0)$  and analytic diffeomorphisms  $\phi_1 : V \rightarrow \phi_1(V) \subset \mathbb{R}^q$ ,  $\phi_2 : W \rightarrow \phi_2(W) \subset \mathbb{R}^{n^2}$  such that for all  $(x_1, \dots, x_q) \in \phi_1(V)$ ,

$$\phi_2 \circ \psi \circ \phi_1^{-1}(x_1, \dots, x_q) = (y_1, \dots, y_\rho, 0, \dots, 0).$$

Thus  $\phi_2(W \cap \psi(V)) = \mathbb{R}^\rho \times \{0\}$ . By Proposition 8.2 we may assume that  $W \cap \psi(V) = W \cap \mathcal{H}(n; n_1, \dots, n_r)$ . Hence  $\phi_2$  is a chart for  $\mathcal{H}(n; n_1, \dots, n_r)$  about  $\psi(U_0, D_0)$ .  $\square$

## References

- [1] Binding, Paul: Hermitian forms and the fibration of spheres. Proc. Am. Math. Soc. 94, 581-584 (1985).
- [2] Binding, Paul and Li, Chi-Kwong: Joint ranges of Hermitian matrices and simultaneous diagonalization. Linear Algebra Appl. 151, 157-167 (1991)
- [3] Bröcker T. and Jänich, K.: Introduction to differential topology. Cambridge University Press. VII (1982).
- [4] Dieudonné, Jean: Foundations of modern analysis. New York, London: Academic Press (1960).
- [5] Dodson, C.T.J.; Parker, Phillip E.: A user's guide to algebraic topology. Dordrecht: Kluwer Academic Publishers (1997).

- [6] Doyle, J.: Analysis of feedback systems with structured uncertainties. IEE Proceedings, vol. 129, no. 6, pp. 242-250, 1982.
- [7] Fan, Michael K.H. and Tits, Andre L.: m-form numerical range and the computation of the structured singular value . IEEE Trans. Autom. Control AC-33, No.3, 284-289 (1988).
- [8] Fan, M.K.H.; Tits, A.L.: On the generalized numerical range. Linear & Mult. Algebra 21, 313-320 (1987).
- [9] Feintuch, A.; Markus, A.: The Töplitz-Hausdorff theorem and robust stability theory. The Mathematical Intelligencer 21, No. 3, 33-36 (1999).
- [10] Friedland, S.; Loewy, R.: Subspaces of symmetric matrices containing matrices with a multiple first eigenvalue. Pac. J. Math. 62, 389-399 (1976).
- [11] Friedland, S.; Robbin, J.W.; Sylvester, J.H.: On the crossing rule. Commun. Pure Appl. Math. 37, 19-37 (1984).
- [12] Halmos, P.: A Hilbert space problem book. New York: Springer-Verlag (1982).
- [13] Hillman, J.A.; Jefferies, B.R.F.; Ricker, W.J.; Straub, B .: Differential properties of the numerical range map of pairs of matrices. Linear Algebra Appl. 267, 317-334 (1997).
- [14] Hiriart-Urruty, J.-B. and Lemarechal, C.: Fundamentals of convex analysis. Springer-Verlag, Berlin 2001.
- [15] Hirsch, M.W.: Differential Topology. New York: Springer-Verlag (1976)
- [16] Hörmander, Lars: Notions of convexity. Progress in Mathematics (Boston, Mass.) 127. Basel: Birkhäuser (1994).
- [17] Hu, S.-T.: Homotopy theory. New York: Academic Press (1959).
- [18] Johnson, Charles R.: Numerical determination of the field of values of a general complex matrix. SIAM J. Numer. Anal. 15, 595-602 (1978).
- [19] Jonckheere, Edmond A.; Ahmad, Farooq; Gutkin, Eugene: Differential topology of numerical range. Linear Algebra Appl. 279, No.1-3, 227-254 (1998).
- [20] Kato, T.: *Perturbation Theory for Linear Operators*. Berlin, Heidelberg, New York: Springer Classics in Mathematics (1995).

- [21] Krantz, S. G. and Parks, H. R.: A Primer of Real Analytic Functions (Second Edition). Boston, Basel, Berlin: Birkhäuser Advanced Texts (2002).
- [22] Li, Chi-Kwong; Poon, Yiu-Tung: Convexity of the joint numerical range. SIAM J. Matrix Anal. Appl. 21, No.2, 668-678 (2000).
- [23] Matsushima, Y.: *Differentiable manifolds*. New York: Marcel Dekker (1972).
- [24] von Neumann, J.; Wigner, E.: Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. Phys. Zschr. 30, 1929, pp. 467-470
- [25] Nomizu, K.: Characteristic roots and vectors of a differentiable family of symmetric matrices. Linear & Mult. Algebra 1, 159-162 (1973).
- [26] Packard, A.; Doyle, J.: The complex structured singular value. Automatica 29, No.1, 71-109 (1993).
- [27] Poon, Yiu Tung: Generalized numerical ranges, joint positive definiteness and multiple eigenvalues. Proc. Am. Math. Soc. 125, No.6, 1625-1634 (1997)
- [28] Poon, Yiu-Tung. On the convex hull of the multiform numerical range. Linear & Mult. Algebra 37, No.1-3, 221-223 (1994)
- [29] Rockafellar, R. T.: Convex Analysis. Princeton Landmarks in Mathematics (1996).
- [30] Stewart, G. W.; Sun, J.: Matrix perturbation theory. Computer Science and Scientific Computing. Boston: Academic Press (1990).
- [31] Safonov, M.G.: Stability and robustness of multivariable feedback systems. The MIT Press Series in Signal Processing, Optimization, and Control, 3. Cambridge, Massachusetts - London: The MIT Press. XI, (1980).
- [32] Shokoohi, S., Silverman, L. M., Van Dooren, P.: Linear time-variable systems: balancing and model reduction. IEEE Trans. Aut. Contr., vol. AC-28, No. 8, pp. 810-822, 1983.
- [33] Whitehead, G.W.: Elements of Homotopy Theory. Berlin: Springer-Verlag (1978)