



**HAL**  
open science

# VARIATIONAL PRINCIPLES OF FUNCTIONS IN MATRIX ANALYSIS

Antoine Mhanna

► **To cite this version:**

Antoine Mhanna. VARIATIONAL PRINCIPLES OF FUNCTIONS IN MATRIX ANALYSIS. 2015.  
hal-01186456v2

**HAL Id: hal-01186456**

**<https://inria.hal.science/hal-01186456v2>**

Preprint submitted on 21 Nov 2015 (v2), last revised 16 Apr 2016 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# VARIATIONAL PRINCIPLES OF FUNCTIONS IN MATRIX ANALYSIS

ANTOINE MHANNA<sup>1\*</sup>

ABSTRACT. The main result I present is first extending by a very basic lemma the Ky-Fan principle to a larger set of functions satisfying certain conditions, second generalizing the property for functions defined over product of finite dimensional Hilbert spaces.

## 1. INTRODUCTION AND PRELIMINARIES

The numerical range of a Hermitian matrix  $A$  is the image of the Rayleigh Quotient  $R_A$  which is the application :

$$\begin{aligned} R_A : \mathbb{C}^n \setminus \{0\} &\rightarrow \mathbb{R} \\ v &\rightarrow \frac{v^* A v}{v^* v}. \end{aligned}$$

Let  $\lambda_1 \geq \dots \geq \lambda_n$  denote the eigenvalues of an  $n \times n$  Hermitian matrix  $A$ , and let  $\sigma_1 \geq \dots \geq \sigma_n$  be the singular values of  $A$ . The set of  $m \times n$  complex matrices is denoted by  $\mathbb{M}_{m,n}(\mathbb{C})$ .

**Proposition 1.1.** [1] *Let  $A$  be an  $n \times n$  Hermitian matrix, and  $(v_1, v_2, \dots, v_n)$  an orthonormal basis of eigenvectors corresponding to the set of eigenvalues  $\lambda_i$ ,  $i = 1, \dots, n$ , that is  $Av_i = \lambda_i v_i$ . Then:*

- (1)  $\lambda_k = R_A(v_k)$  and the range of  $R_A$  is  $[\lambda_n, \lambda_1]$ .
- (2)  $\lambda_1 = \sup_v (R_A(v)) = R_A(B)$ ,  $\lambda_n = \inf_v (R_A(v)) = R_A(C)$

where  $B$  is the subspace of eigenvectors corresponding to  $\lambda_1$  and  $C$  is the subspace of eigenvectors corresponding to  $\lambda_n$ .

More specifically:

- (3) Let  $V_k$  be the subspace of  $\mathbb{C}^n$  spanned by  $\{v_1, v_2, \dots, v_k\}$ , and  $W_k = \text{span} \{v_{k+1}, \dots, v_n\} = V_k^\perp$

$$\lambda_k = \inf_{v \in V_k} (R_A(v)) \quad \text{and} \quad \lambda_k = \sup_{v \in W_{k-1}} (R_A(v))$$

Another yet equivalent form of Proposition 1.1 is:

---

*Date:* 21 November 2015

\* Corresponding author.

2010 *Mathematics Subject Classification.* Primary 39B82; Secondary 44B20, 46C05.

*Key words and phrases.* Hermitian matrices, variational characterization, Ky-Fan principle.

**Proposition 1.2.** *Let  $A$  be an  $n \times n$  Hermitian matrix then*

$$\begin{aligned} \lambda_1(A) &= \max\{R_A(v)\} \text{ so } \lambda_1 = R_A(v_1) \text{ for some } v_1 \in \mathbb{C}^n. \\ &\vdots \\ \lambda_k(A) &= \max\{R_A(v), v \perp v_1, v_2, \dots, v_{k-1}\} \text{ and} \\ \lambda_k &= R_A(v_k) \text{ for some } v_k \perp v_1, v_2, \dots, v_{k-1}. \end{aligned}$$

## 2. MAIN RESULTS

**2.1. Variational characterizations.** If  $f$  is a given function that is defined over a product of vector spaces and a certain domain  $D$ , we mean hereafter by  $f(\text{span}(\mathbf{u}_1), \dots, \text{span}(\mathbf{u}_n))$  where  $\mathbf{u}_i$  is any set of vectors, the value of  $f(x_1, \dots, x_n)$  for any of the  $x_i$ 's taken to be in  $\text{span}(\mathbf{u}_i)$  and in the domain  $D$ .

**Lemma 2.1.** *Let  $\mathcal{H}$  be a  $\mathbb{K}$ -Hilbert space ( $\mathbb{K} \equiv \mathbb{R}$  or  $\mathbb{C}$ ) of dimension  $n$  equipped with the norm denoted by  $\|\cdot\|_s$  and let  $f$  be a continuous function from  $E$  to  $\mathbb{R}$ . If  $r > 0$  is any fixed real number, Set:*

$$\begin{aligned} h_1 &:= \max\{f(x), \|x\|_s = r\} \text{ so } h_1 = f(v_1) \text{ for some } v_1 \in E. \\ h_2 &:= \max\{f(x), \|x\|_s = r, x \perp v_1\} \text{ so } h_2 = f(v_2) \text{ for some } v_2 \in E. \\ &\vdots \\ h_n &:= \max\{f(x), \|x\|_s = r, x \perp v_1, v_2, \dots, v_{n-1}\} \\ &\text{ so } h_n = f(v_n) \text{ for some } v_n \perp v_1, \dots, v_{n-1}. \end{aligned}$$

and set:

$$\begin{aligned} q_1 &:= \min\{f(x), \|x\|_s = r\} \text{ so } q_1 = f(w_1) \text{ for some } w_1 \in E. \\ q_2 &:= \min\{f(x), \|x\|_s = r, x \perp w_1\} \text{ so } q_2 = f(w_2) \text{ for some } w_2 \in E. \\ &\vdots \\ q_n &:= \min\{f(x), \|x\|_s = r, x \perp w_1, w_2, \dots, w_{n-1}\} \\ &\text{ so } q_n = f(w_n) \text{ for some } w_n \perp w_1, w_2, \dots, w_{n-1}. \end{aligned}$$

$\mathcal{A} \equiv (v_1, \dots, v_n)$  and  $\mathcal{B} \equiv (w_1, \dots, w_n)$  are two orthonormal basis of  $E$ . Let  $k \leq n$ , and  $m < z$  since we study  $f$  on the ball of radius  $r$  denoted  $B_r$  we will suppose that our function  $f$  is only defined from  $B_r$  to  $\mathbb{R}$ .

- If  $f(\text{span}(v_z, \dots, v_m)) \leq rf(v_m)$  and  $f(\sum_{j=i}^n \alpha_j v_j) \leq \sum_{j=i}^n |\alpha_j|^2 f(v_j)$  (max con-

dition) for all  $i = 1, \dots, n$  then  $\sum_{i=1}^k rh_i = \max_{B_k} \sum_{i=1}^k f(x_i)$ , if  $f$  is positive

valued we have also  $\prod_{i=1}^k rh_i = \max_{B_k} \prod_{i=1}^k f(x_i)$ .

- If  $f(\text{span}(w_z, \dots, w_m)) \geq rf(w_z)$  and  $f(\sum_{j=i}^n \alpha_j w_j) \geq \sum_{j=i}^n |\alpha_j|^2 f(w_j)$  (min condition) for all  $i = 1, \dots, n$  then  $\sum_{i=1}^k r q_i = \min_{B_k} \sum_{i=1}^k f(x_i)$ , if  $f$  is positive valued we have also  $\prod_{i=1}^k r q_i = \min_{B_k} \prod_{i=1}^k f(x_i)$ .

where  $B_k = (x_1, \dots, x_k)$  denotes an orthogonal basis of dimension  $k$  such that  $\|x_i\|_s = r$  for  $1 \leq i \leq k$  and  $\alpha_i \in \mathbb{C}$ ,  $\beta_i \in \mathbb{C}$  for all  $i = 1, \dots, k$ .

*Proof.* It suffices to prove the maximum version as the other has a similar proof, we will take  $r$  to be 1, as the same proof holds if  $r \neq 1$ . Let us suppose first that  $E$  is of dimension  $n = 2$ , take two orthonormal vectors  $x_1$  and  $x_2$  and since  $(v_1, v_2)$  is a basis, necessarily we have

$$\begin{cases} x_1 = \alpha v_1 + \beta v_2 \\ x_2 = i v_1 + j v_2 \\ |\alpha|^2 + |\beta|^2 = 1 \quad (\|x_1\|_s = 1) \\ |i|^2 + |j|^2 = 1 \quad (\|x_2\|_s = 1) \\ \alpha \bar{i} = -\beta \bar{j} \quad (x_1 \perp x_2) \\ |\alpha|^2 |i|^2 = |\beta|^2 |j|^2 \end{cases} \implies \begin{cases} |\alpha|^2 |i|^2 + |\beta|^2 |j|^2 = |i|^2 \\ |\beta|^2 (|i|^2 + |j|^2) = |i|^2 \\ |\beta|^2 = |i|^2 \\ |\alpha|^2 = |j|^2. \end{cases}$$

The max condition let us write

$$\begin{aligned} f(x_1) &\leq |\alpha|^2 f(v_1) + |\beta|^2 f(v_2) \\ f(x_2) &\leq |i|^2 f(v_1) + |j|^2 f(v_2) = |\beta|^2 f(v_1) + |\alpha|^2 f(v_2), \end{aligned}$$

adding both inequalities yields to  $f(x_1) + f(x_2) \leq f(v_1) + f(v_2)$  and we have equality if  $v_1 = x_1$  and  $x_2 = v_2$ . For any  $n$ ,  $n > 2$  we can write:

$$\begin{cases} x_1 = \alpha_{1,1} v_1 + \alpha_{2,1} v_2 + \dots + \alpha_{n,1} v_n \\ x_2 = \alpha_{1,2} v_1 + \alpha_{2,2} v_2 + \dots + \alpha_{n,2} v_n \\ \vdots \\ x_k = \alpha_{1,k} v_1 + \alpha_{2,k} v_2 + \dots + \alpha_{n,k} v_n \\ |\alpha_{1,1}|^2 + \dots + |\alpha_{n,1}|^2 = 1 \quad (\|x_1\|_s = 1) \\ |\alpha_{1,2}|^2 + \dots + |\alpha_{n,2}|^2 = 1 \quad (\|x_2\|_s = 1) \\ \vdots \\ |\alpha_{1,k}|^2 + \dots + |\alpha_{n,k}|^2 = 1 \quad (\|x_k\|_s = 1). \end{cases}$$

Here we assume that  $1 < k \leq n$ ,  $|\alpha_{j,i}|^2$  can be different from zero for any  $j$  and same  $i$ , the matrix whose lines are these  $k$  vectors, completed by  $n - k$  unitary vectors orthogonal to the  $k$  vectors and mutually orthogonal is an  $n \times n$  unitary matrix, we will impose a supplementary condition over the added  $n - k$  vectors as follows; without loss of generality

$$\left\{ \begin{array}{l} x_1 = \alpha_{1,1}v_1 + \alpha_{2,1}v_2 + \cdots + \alpha_{n,1}v_n \\ x_2 = \alpha_{1,2}v_1 + \alpha_{2,2}v_2 + \cdots + \alpha_{n,2}v_n \\ \vdots \\ x_k = \alpha_{1,k}v_1 + \cdots + \alpha_{n,k}v_n \\ x_{k+1} = \alpha_{2,k+1}v_2 + \cdots + \alpha_{n,k+1}v_n \\ \vdots \\ x_n = \alpha_{n-k+1,n}v_{n-k+1} + \cdots + \alpha_{n,n}v_n \end{array} \right.$$

It is easily seen that such basis always exists, a direct consequence of the max condition is this inequality

$$f(x_1) + \cdots + f(x_n) \leq f(v_1) + \cdots + f(v_n)$$

since  $f(x_j) \leq f(v_{j-(k-1)})$  for  $n \geq j > k$  we obtain:

$$f(x_1) + \cdots + f(x_k) \leq f(v_1) + f(v_{n-(k-2)}) + \cdots + f(v_n) \leq f(v_1) + f(v_2) + \cdots + f(v_k).$$

Consequently we have discussed all possible cases to complete the proof.  $\square$

Notice that  $f(v_n) \leq \cdots \leq f(v_1)$  and  $f(w_n) \geq \cdots \geq f(w_1)$ . Here is an example showing that the conditions in Lemma 2.1 are necessary.

**Example 2.2.** Let

$$\begin{aligned} f : \quad \mathbb{R}^2 &\quad \rightarrow \quad \mathbb{R} \\ u = (x, y) &\quad \rightarrow \quad \ln(|x + \epsilon|). \end{aligned}$$

where  $\epsilon$  is a strictly positive number to be fixed, here the norm  $\|\cdot\|_s$  is the eucliden norm, we verify then that  $h_1 = \ln(|1 + \epsilon|) = \max_{\|u\|_s=1} \ln(|x + \epsilon|) = f(v_1)$ , by taking

$$v_1 = (1, 0), v_2 = (0, \pm 1), x_1 = \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \text{ and } x_2 = \left( \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) \text{ we have}$$

$v_1 \perp v_2, x_1 \perp x_2$  but

$$f(v_1) + f(v_2) = \ln(\epsilon(1 + \epsilon)) = h_1 + h_2 < f(x_1) + f(x_2) = \ln(\sqrt{2} + 0.5 + \epsilon^2)$$

whenever  $\epsilon < 0.5 + \sqrt{2}$ .

Lemma 2.1 and the previous results will entail important variational representations concerning the sum and product of eigenvalues and singular values:

**Corollary 2.3** (Ky-Fan principle). *Let  $H$  be an  $n \times n$  Hermitian matrix such that  $\lambda_1 \geq \cdots \geq \lambda_n$  are the eigenvalues of  $H$  in decreasing order. For any  $1 \leq k \leq n$*

$$\sum_{i=1}^k \lambda_i(H) = \max_{U^*U=I_k} \text{tr}(U^*HU) \quad (2.1)$$

$$\sum_{i=1}^k \lambda_{n-i+1}(H) = \min_{U^*U=I_k} \text{tr}(U^*HU) \quad (2.2)$$

And if  $H$  is P.S.D that is  $\lambda_n$  is nonnegative we have

$$\prod_{i=1}^k \lambda_i(H) = \max_{U^*U=I_k} \det(U^*HU) \quad (2.3)$$

$$\prod_{i=1}^k \lambda_{n-i+1}(H) = \min_{U^*U=I_k} \det(U^*HU). \quad (2.4)$$

*Proof.* It suffices to notice that  $\max_{U^*U=I_k} \text{tr}(U^*HU) = \max_{x_i^*x_j=\delta_{ij}} \sum_{i=1}^k x_i^*Hx_i$  applying

Proposition 1.2 and Lemma 2.1 to the function  $f(x) = x^*Hx$  completes the proof, since  $f(\text{span}(v_z, \dots, v_m)) \leq f(v_m)$  when  $m < z$  (see Proposition 1.1) and for all

$i$  we have  $f(\sum_{j=1}^i \alpha_j v_j) = \sum_{j=1}^i |\alpha_j|^2 f(v_j)$ . For the product assertion if  $H$  is P.S.D

we have  $\max_{U^*U=I_k} \det(U^*HU) = \max_{x_i^*x_j=\delta_{ij}} (x_i^*Hx_j)$  but since  $U^*HU$  is Hermitian we can suppose that it is a diagonal matrix we apply again Proposition 1.2 and Lemma 2.1 to the positive valued function  $f(x) := x^*Hx$  to complete the proof. The minimum versions are similarly proven.  $\square$

**2.2. Generalization to product spaces.** In the previous section we have taken  $\mathcal{H}$  to be any Hilbert space, hereafter we consider  $f$  a continuous function defined over  $\mathcal{H}_{1,j_1} \times \dots \times \mathcal{H}_{n,j_n}$  for any  $n$ , where  $\mathcal{H}_{l,j_l}$  denotes a given Hilbert vector-space of dimension  $j$ , with scalar product denoted by  $\langle \cdot, \cdot \rangle_{\mathcal{H}_{l,j_l}}$  and  $\|\cdot\|_{\mathcal{H}_{l,j_l}}$  the associated norm. We will restrict also our real valued functions  $f$  to be defined only over the domain  $B_{r_1} \times \dots \times B_{r_n}$  where for any  $l$ ,  $r_l$  is a fixed real number,  $B_{r_l}$  stands for the sphere of radius  $r$  in  $\mathcal{H}_{l,j_l}$ . For  $k \leq m := \min(j_1, \dots, j_n)$ , similarly we set

$$\mathfrak{G}_1 := \max\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\} \text{ so } \mathfrak{G}_1 = f(z_{1,1}, \dots, z_{n,1})$$

$$\mathfrak{G}_2 := \max\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l, x_1 \perp z_{1,1}, \dots, x_n \perp z_{n,1}\} \\ \text{ so } \mathfrak{G}_2 = f(z_{1,2}, \dots, z_{n,2})$$

$\vdots$

$$\mathfrak{G}_k := \max\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l, x_1 \perp z_{1,1}, z_{1,2} \dots, z_{1,k-1}; \\ \dots; x_n \perp z_{n,1}, z_{n,2} \dots, z_{n,k-1}\} \text{ so } \mathfrak{G}_k = f(z_{1,k}, \dots, z_{n,k})$$

and:

$$\mathfrak{D}_1 := \min\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\} \text{ so } \mathfrak{D}_1 = f(c_{1,1}, \dots, c_{n,1})$$

$$\mathfrak{D}_2 := \min\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l, x_1 \perp c_{1,1}, \dots, x_n \perp c_{n,1}\} \\ \text{ so } \mathfrak{D}_2 = f(c_{1,2}, \dots, c_{n,2})$$

$\vdots$

$$\mathfrak{D}_k := \min\{f(x_1, \dots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l, x_1 \perp c_{1,1}, c_{1,2} \dots, c_{1,k-1}; \\ \dots; x_n \perp c_{n,1}, c_{n,2} \dots, c_{n,k-1}\} \text{ so } \mathfrak{D}_k = f(c_{1,k}, \dots, c_{n,k}).$$

For a certain  $l$  and  $k$  fixed;  $\mathfrak{Z}_{l,k} \equiv (z_{l,1}, \dots, z_{l,k})$  and  $\mathfrak{D}_{l,k} \equiv (c_{l,1}, \dots, c_{l,k})$  are two orthogonal basis of  $\mathcal{H}_{l,j_l}$  each one of dimension  $k$  and such that  $\|c_{l,s}\|_{\mathcal{H}_{l,j_l}} = \|z_{l,q}\|_{\mathcal{H}_{l,j_l}} = r_l$  for all  $q$  and  $s$ . A direct generalization of Lemma 2.1 is the following:

**Theorem 2.4.** For  $k \leq \min(j_1, \dots, j_n)$ ,  $y_{i,1} \leq y_{i,2}$  for all  $i \leq n$ ,  $\nu = \min_i y_{i,1}$ ,  $\mu = \max_i y_{i,2}$ ,  $m := \min_l j_l$  and let  $\mathfrak{X}_{l,k} = (x_{l,1}, \dots, x_{l,k})$  denote an orthogonal basis of  $\mathcal{H}_{l,j_l}$  of dimension  $k$  such that  $\|x_{l,g}\|_{\mathcal{H}_{l,j_l}} = r_l$  for all  $l$  and  $g$ . Let  $f$  be any continuous function from  $B_{r_1} \times \dots \times B_{r_n}$  into  $\mathbb{R}$ .

- If  $f(\text{span}(z_{1,y_{1,1}}, \dots, z_{1,y_{1,2}}), \dots, \text{span}(z_{n,y_{n,1}}, \dots, z_{n,y_{n,2}})) \leq f(z_{1,\nu}, \dots, z_{n,\nu})$   
and  $\sum_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}, g=1}^m f(x_{1,g}, \dots, x_{n,g}) \leq \sum_{i=1}^m f(z_{1,i}, \dots, z_{n,i})$  (max condition).  
then

$$\sum_{i=1}^k \mathfrak{G}_i = \max_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}} \sum_{g=1}^k f(x_{1,g}, \dots, x_{n,g}).$$

In addition if  $f$  is positive valued we have

$$\prod_{i=1}^k \mathfrak{G}_i = \max_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}} \prod_{g=1}^k f(x_{1,g}, \dots, x_{n,g})$$

- If  $f(\text{span}(c_{1,y_{1,1}}, \dots, c_{1,y_{1,2}}), \dots, \text{span}(c_{n,y_{n,1}}, \dots, c_{n,y_{n,2}})) \geq f(c_{1,\mu}, \dots, c_{n,\mu})$   
and  $\sum_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}, g=1}^m f(x_{1,g}, \dots, x_{n,g}) \geq \sum_{i=1}^m f(c_{1,i}, \dots, c_{n,i})$  (min condition).  
then

$$\sum_{i=1}^k \mathfrak{D}_i = \min_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}} \sum_{g=1}^k f(x_{1,g}, \dots, x_{n,g}).$$

In addition if  $f$  is positive valued we have

$$\prod_{i=1}^k \mathfrak{D}_i = \min_{\mathfrak{X}_{1,k}, \dots, \mathfrak{X}_{n,k}} \prod_{g=1}^k f(x_{1,g}, \dots, x_{n,g})$$

**Proposition 2.5.** Let  $k$  be a fixed integer and let  $\alpha_{i,j}$  be any complex numbers

with  $i \leq k$ ,  $j = 1, 2$  such that  $\sum_{i=1}^k |\alpha_{i,1}|^2 \leq 1$  and  $\sum_{i=1}^k |\alpha_{i,2}|^2 \leq 1$  then

$$\left| \sum_{i=1}^k \alpha_{i,1} \alpha_{i,2} \right| \leq 1$$

*Proof.* By the Cauchy-Schwartz inequality, it suffices to prove the proposition when all the complex numbers are nonnegative numbers, we proceed by induction when  $k = 2$ , we have  $\alpha_{1,1}^2 + \alpha_{2,1}^2 \leq 1$  and  $\alpha_{1,2}^2 + \alpha_{2,2}^2 \leq 1$ , we will prove that the quantity  $M := \sqrt{1 - \alpha_{2,1}^2} \sqrt{1 - \alpha_{2,2}^2} + \alpha_{2,1} \alpha_{2,2} \leq 1$ , and consequently we will have

$\alpha_{1,1}\alpha_{1,2} + \alpha_{2,2}\alpha_{2,1} \leq 1$ , but if  $w := \alpha_{2,2} - \alpha_{2,1}$  we get:

$$M \leq 1 \iff (1 - \alpha_{2,1}^2)(1 - \alpha_{2,2}^2) \leq (1 - \alpha_{2,1}\alpha_{2,2})^2 \quad (2.5)$$

$$\iff (1 - \alpha_{2,1})(1 + \alpha_{2,2})(1 - \alpha_{2,2})(1 + \alpha_{2,1}) \leq (1 - \alpha_{2,1}\alpha_{2,2})^2 \quad (2.6)$$

$$\iff (1 - \alpha_{2,1}\alpha_{2,2} - w)(1 - \alpha_{2,1}\alpha_{2,2} + w) \leq (1 - \alpha_{2,1}\alpha_{2,2})^2 \quad (2.7)$$

$$\iff (1 - \alpha_{2,1}\alpha_{2,2})^2 - w^2 \leq (1 - \alpha_{2,1}\alpha_{2,2})^2. \quad (2.8)$$

Suppose the result true for  $k = n$  let us prove it for  $k = n + 1$ , we will take

$$\sum_{i=1}^{n+1} |\alpha_{i,1}|^2 \leq 1 \text{ and } \sum_{i=1}^{n+1} |\alpha_{i,2}|^2 \leq 1 \text{ and set without loss of generality } s_1^2 = \alpha_{n,1}^2 +$$

$$\alpha_{n+1,1}^2 \text{ and } s_2^2 = \alpha_{n,2}^2 + \alpha_{n+1,2}^2. \text{ By the induction hypothesis } \sum_{i=1}^{n-1} \alpha_{i,1}\alpha_{i,2} + s_1s_2 \leq 1$$

but then it is easy to verify that

$$\sum_{i=1}^{n+1} \alpha_{i,1}\alpha_{i,2} \leq \sum_{i=1}^{n-1} \alpha_{i,1}\alpha_{i,2} + s_1s_2 \leq 1$$

thus  $\left| \sum_{i=1}^k \alpha_{i,1}\alpha_{i,2} \right| \leq \sum_{i=1}^k |\alpha_{i,1}\alpha_{i,2}| \leq 1$  which is the desired result.  $\square$

**Lemma 2.6.** *Let  $k, h$  be two fixed integers and let  $\alpha_{i,j}$  be any complex numbers with  $i \leq k, j \leq h$  such that  $\sum_{i=1}^k |\alpha_{i,j}|^2 \leq 1$  for all  $j$ , then*

$$\left| \sum_{i=1}^k \alpha_{i,1} \cdots \alpha_{i,h} \right| \leq 1$$

*Proof.* The proof will follow from Proposition 2.5 by noticing that

$$\left| \sum_{i=1}^k \alpha_{i,1} \cdots \alpha_{i,h} \right| \leq \sum_{i=1}^k |\alpha_{i,1} \cdots \alpha_{i,h}| \leq \sum_{i=1}^k |\alpha_{i,1}\alpha_{i,2}| \leq 1$$

$\square$

As noticed in Theorem 2.4 the order at which we write the elements of the orthogonal basis is important, we associate to every basis  $(x_1, \dots, x_k)$  the group of permutation  $\mathfrak{S}_k$ .

**Definition 2.7.** Given a certain basis  $A := (x_1, \dots, x_k)$  of dimension  $k$ , the no fix point permutation basis of  $A$  is  $A' := (x_{\mathfrak{d}(1)}, \dots, x_{\mathfrak{d}(k)})$  where  $\mathfrak{d}$  is a permutation of the  $k$  elements with no fixed point.

Keeping the terminology we used in this section we have:



**Lemma 2.8.** *Let  $n$  be a fixed integer and let  $f_i$  be nonnegative real numbers for  $i \leq n$ . Let  $f$  be the function*

$$f : B_{1_1} \times \cdots \times B_{1_n} \rightarrow \mathbb{R}$$

$$(x_1, \cdots, x_n) \rightarrow \left| \sum_{i=1}^m f_i \langle x_1, s_{1,i} \rangle_{\mathcal{H}_{1,j_1}} \langle x_2, s_{2,i} \rangle_{\mathcal{H}_{2,j_2}} \cdots \langle x_n, s_{n,i} \rangle_{\mathcal{H}_{n,j_n}} \right|,$$

where  $(s_{1,1}, \cdots, s_{1,k})$  is an orthonormal basis of dimension  $k$  in  $\mathcal{H}_{1,j_1}$ , then  $\mathfrak{G}_i = f_i$ ,  $\mathfrak{D}_i = 0$  for all  $i$  and  $f$  verifies the max and min conditions of Theorem 2.4.

*Proof.* By writing any unit vector  $x_l \in \mathcal{H}_{l,j_l}$  in terms of the  $S_{l,j_l} := (s_{l,1}, \cdots, s_{l,j_l})$  basis and using Lemma 2.6 we get the required result. Note in this case that the basis  $\mathfrak{Z}_{l,k} \equiv (z_{l,1}, \cdots, z_{l,k})$  can be taken to be the arbitrarily chosen  $S_{l,k} = (s_{l,1}, \cdots, s_{l,k})$  and  $\mathfrak{D}_{l,k} \equiv (c_{l,1}, \cdots, c_{l,k})$  are also the  $S_{l,k}$  except one say  $S_{n,k}$  which will be replaced by  $S'_{n,k}$  (the no fix permutation basis).  $\square$

**Example 2.9.** Let  $A \in \mathbb{M}_{m,n}(\mathbb{C})$  fixed,  $q := \min(m, n)$ , the norm  $\|\cdot\|_s$  is the one associated to the usual scalar product on  $\mathbb{C}^k$  for any  $k$ , if  $A = W\Sigma V$  is the singular value decomposition,  $W = \begin{pmatrix} t_1 & \cdots & t_m \\ \vdots & \cdots & \vdots \end{pmatrix}$ ,  $V^* = \begin{pmatrix} s_1 & \cdots & s_n \\ \vdots & \cdots & \vdots \end{pmatrix}$  and  $\Sigma$  is the

$n \times m$  matrix  $\left( \begin{array}{ccc|c} \sigma_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \sigma_q & 0 \\ \hline 0 & & & 0 \end{array} \right)$ , then a direct computation shows that

$$x^* A y = x^* W \Sigma V y = \sum_{i=1}^n \sigma_i (x^* t_i) (s_i^* y).$$

Let us define the continuous function  $f_A$  as:

$$f_A : \{(x, y) \in \mathbb{C}^m \times \mathbb{C}^n; \|x\|_s = \|y\|_s = 1\} \rightarrow \mathbb{R}$$

$$(x, y) \rightarrow |x^* A y| = \left| \sum_{i=1}^q \sigma_i (x^* t_i) (s_i^* y) \right|.$$

By Lemma 2.8 for each  $k = 1, \cdots, q$  we have

$$\mathfrak{G}_1 = \sigma_1(A) = \max \frac{|x^* A y|}{\|x\| \|y\|} \quad \text{and} \quad \mathfrak{G}_k = \sigma_k(A) = \max_{\substack{x \in \text{span}\{t_1, \dots, t_{k-1}\}^\perp \\ y \in \text{span}\{s_1, \dots, s_{k-1}\}^\perp}} \frac{|x^* A y|}{\|x\| \|y\|},$$

and :

$$\begin{aligned} \bullet \sum_{i=1}^k \mathfrak{G}_i &= \max \{ |tr X^* A Y| : X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, X^* X = I = Y^* Y \} \\ &= \max_{B_k, C_k} \sum_{i=1}^k f_A(x_i, y_i) = \sum_{i=1}^k \sigma_i(A), \\ \bullet \prod_{i=1}^k \mathfrak{G}_i &= \max_{B_k, C_k} \prod_{i=1}^k f_A(x_i, y_i) = \prod_{i=1}^k \sigma_i(A), \end{aligned}$$

$$\begin{aligned} \bullet \sum_{i=1}^k \mathfrak{D}_i &= \min \{ |tr X^* A Y| : X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, X^* X = I = Y^* Y \} \\ &= \min_{B_k, C_k} \sum_{i=1}^k f_A(x_i, y_i) = 0, \end{aligned}$$

where  $x_i$  is a column of  $X$ ,  $y_i$  a column of  $Y$ ,  $B_k = (x_1, \dots, x_k)$  and  $C_k = (y_1, \dots, y_k)$  are any two orthonormal basis of  $\mathbb{C}^k$ .

#### REFERENCES

1. F.Zhang, *Matrix Theory Basic Results and Techniques, 2<sup>nd</sup> Edition* (Universitext) Springer, 2013.
2. R. A. Horn and C. R Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

<sup>1</sup> DEPARTMENT OF MATHEMATICS, LEBANESE UNIVERSITY HADATH, BEIRUT, LEBANON  
 E-mail address: [tmhanat@yahoo.com](mailto:tmhanat@yahoo.com)