



**HAL**  
open science

## Analysis And Matrices Revisited

Antoine Mhanna

► **To cite this version:**

| Antoine Mhanna. Analysis And Matrices Revisited. 2015. hal-01286126

**HAL Id: hal-01286126**

**<https://inria.hal.science/hal-01286126>**

Submitted on 10 Mar 2016

**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.

*Analysis And Matrices  
Revisited*

✦ Antoine Mhanna ✦



© 2016

Antoine Mhanna

ALL RIGHTS RESERVED

# ACKNOWLEDGMENTS

I must thank the Lebanese University and staff for letting me publish the work from my thesis with slight modifications as a book; I also mostly thank my parents because without them non of this could be achieved.

# TABLE OF CONTENTS

Cover .....	i
Acknowledgments .....	iii
List of Figures.....	vi
Introduction.....	vii
<b>Chapter</b>	
<b>CHAPTER 1 Basics in Matrix Analysis</b>	<b>1</b>
1.1. Basics .....	1
1.2. General Results .....	4

1.3.	Polar Decomposition.....	7
1.4.	Jensen Inequality and Applications .....	10
1.5.	Numerical Range.....	16
<b>CHAPTER 2</b>	<b>Some Characterizations</b>	<b>23</b>
2.1.	Eigenvalues.....	23
2.2.	Rayleigh Quotient .....	25
2.3.	Modulus of a Matrix .....	27
2.4.	Minimax principle .....	30
<b>CHAPTER 3</b>	<b>Matrices and Norms</b>	<b>36</b>
3.1.	Definitions.....	36
3.2.	Numerical Radius.....	38
3.3.	Ky Fan Norms.....	43
	References.....	48
<b>Appendices</b>		
<b>CHAPTER A</b>	<b>The Appendix</b>	<b>50</b>

# LIST OF FIGURES

## Figure

1.1. $W(A)$ of a normal (unitary) matrix $A$ .....	18
1.2. $W(A)$ of the $2 \times 2$ matrix $A$ .....	22
3.1. The line $f$ and it's projection $L$ (in red) separating $O$ and $W(A)$ in $\mathbb{C}$ . .....	43

# Introduction

The work I present in this book is a deep recapitulation of ideas (with proofs) on analysis of matrices. The terminology used in this work matches the definitions given in ([9], [4]).

The first chapter will cover the basics in matrix analysis, we state some well known results, the Spectral Decomposition, Polar Decomposition and The Singular Value Decomposition Theorems and give a short discussion on the numerical range of a square matrix with examples to get familiar with.

In chapter two and three we discuss mainly the variational properties of Hermitian matrices, their eigenvalues and survey norm properties on the space of matrices.



## CHAPTER 1

# Basics in Matrix Analysis

### 1.1. Basics

Let  $E$  and  $F$  be two vector spaces over  $\mathbb{C}$  or  $\mathbb{R}$  of dimension  $n$  and  $m$  respectively. A linear transformation from  $E$  to  $F$  is identified to an  $m \times n$  matrix, the set of these matrices is denoted by  $\mathbb{M}_{m,n}(\mathbb{C})$ . In this chapter the norm  $\|\cdot\|$  on the Hilbert space  $\mathbb{C}^n$  is the one associated to the complex scalar product (see A.4).

**Definition 1.1.** *The singular values of a matrix  $A$  are the nonnegative square roots of the eigenvalues of  $A^*A$ , where  $A^* = (\overline{A})^T$  is the adjoint of  $A$ .*

**Proposition 1.2.** *The eigenvalues of  $A^*A$  are all nonnegative.*

*Proof.*

$$x^*(A^*A)x = \langle Ax, Ax \rangle = (Ax)^*(Ax) = \|Ax\|^2 \geq 0. \quad (1.1)$$

The proposition follows by taking  $x$  to be an eigenvector of  $A^*A$ .  $\square$

If we denote by  $\sigma_i$  a singular value and by  $\lambda_i$  an eigenvalue.

$$\sigma_i(A) = \sqrt{\lambda_i(A^*A)}.$$

We begin by stating without proof the Singular Value Decomposition Theorem:

**Theorem 1.3 (S.V.D.).** [9] *Let  $A \in \mathbb{M}_{m,n}$  with nonzero singular values  $\sigma_1, \dots, \sigma_r$ . Then there exist an  $m \times n$  unitary  $W$  and an  $n \times n$  unitary  $V$  such that*

$$A = W \overbrace{\begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}}^{m \times n} V = W \Sigma V, \quad (1.2)$$

where  $D_r = \text{diag}(\sigma_1, \dots, \sigma_r)$ .

**Definition 1.4.** *A square matrix  $A$  is said to be congruent (resp unitarily congruent, thus similar) to  $B$  if there exist an invertible matrix  $U$  (resp. a unitary matrix  $U$ ) such that*

$$A = U^* B U$$

As one of the consequences of (1.2), is that for any square matrix  $A$ ,  $A^*A$  and  $AA^*$  are unitarily congruent (Remark 1.16).

**Definition 1.5.** • *An element  $U \in \mathbb{M}_n(\mathbb{C})$  is called unitary if  $U^*U = Id$ , ( in this case  $U^* = U^{-1}$ ).*

• *An element  $U \in \mathbb{M}_n(\mathbb{C})$  is called isometry if  $\|U(x)\| = \|x\|$  for all  $x \in \mathbb{C}^n$ .*

• *An element  $N \in \mathbb{M}_n(\mathbb{C})$  is called normal if  $NN^* = N^*N$ .*

- An element  $U \in \mathbb{M}_n(\mathbb{C})$  is called Hermitian or self-adjoint if  $U = U^*$ .

**Definition 1.6.** Let  $A, B \in \mathbb{M}_n(\mathbb{C})$  be two self-adjoint operators, we write  $A \leq B$  if  $\forall u \in \mathbb{C}^n$  we have  $\langle Au, u \rangle \leq \langle Bu, u \rangle$  and we say  $A$  is positive semi-definite (resp positive definite) if  $A \geq 0$  (resp  $A > 0$ ).

We easily conclude if  $A \in \mathbb{M}_n(\mathbb{C})$  is positive semi-definite ( respectively positive definite) if and only if :

$$\langle Au, u \rangle \geq 0 \quad (\text{resp } \langle Au, u \rangle > 0). \quad (1.3)$$

And by taking  $u$  to be an eigenvector of the matrix  $A$  1.3 shows that all eigenvalues of  $A$  are nonnegative (respectively strictly positive).

**Lemma 1.7.** Let  $\mathcal{H}$  be a vector space over  $\mathbb{C}$  and  $b$  a sesquilinear Hermitian form over  $\mathcal{H} \times \mathcal{H}$ . Then  $b$  is identically null over  $\mathcal{H} \times \mathcal{H}$  if and only if it is null on the diagonal of  $\mathcal{H} \times \mathcal{H}$ .

*Proof.* This lemma is a consequence of the polarization formula :

$$b(u, v) = \frac{1}{4} [b(u + v, u + v) - b(u - v, u - v) + ib(u + iv, u + iv) - ib(u - iv, u - iv)]. \quad (1.4)$$

□

**Proposition 1.8.** Let  $A \in \mathbb{M}_n(\mathbb{C})$ , we have the two equivalences:

i)  $A$  is self-adjoint.

ii)  $\forall u \in \mathbb{C}^n$ ,  $\langle Au, u \rangle \in \mathbb{R}$ .

*Proof.*  $A$  is self adjoint so  $A = A^*$ . However

$$\langle Au, u \rangle = \langle u, A^*u \rangle = \langle u, Au \rangle = \overline{\langle Au, u \rangle},$$

for any  $u \in \mathbb{C}^n$ . Hence  $\langle Au, u \rangle \in \mathbb{R}$  for all  $u \in \mathbb{C}^n$ .

Assume that *ii*) is verified, hence

$$\langle (A - A^*)u, u \rangle = 0 \text{ for all } u \in \mathbb{C}^n.$$

To conclude the proposition apply Lemma 1.7 to the form

$$b(u, v) = \left\langle \frac{A - A^*}{2i}u, v \right\rangle \tag{1.5}$$

Thus  $A - A^* = 0$ ,  $A = A^*$  and  $A$  is Hermitian. □

The binary relation on the vector space of self-adjoint square matrices; that is,  $A \geq B$  if and only if  $(A - B)$  is positive semi-definite (P.S.D) and  $A > B$  if and only if  $(A - B)$  is positive definite (P.D.), is a partial order and many results can be investigated in the sense of matrix partial orders, this and more related Theorems and applications can be found in [5].

## 1.2. General Results

The main goal of this section is to prove the well known Schur Decomposition Theorem.

**Lemma 1.9.** *Let  $Q$  be a subspace of a finite-dimensional space  $H$ , and let  $A$  be an operator over  $H$ . If  $Q$  is invariant under  $A$ , then there exist a certain scalar  $\lambda$  and at least one vector  $v_i$  in  $Q$  such that:  $Av_i = \lambda v_i$ .*

*Proof.* Let  $(s_1, \dots, s_n)$  be a basis of  $Q$ , and assume the converse; that is, there is no  $v \in Q$  such that  $Av$  is a multiple of  $v$ , we will construct an

infinite sequence of linearly independent vectors in  $Q$  which is of course a contradiction. Take for example a vector  $z_1$  belonging to  $Q$ , then we have:

$$\begin{cases} Az_1 = a_1 z_2 \\ Az_2 = a_2 z_3 \\ Az_3 = a_3 z_4 \\ \vdots \\ \vdots \end{cases} \quad \begin{cases} \vdots \\ \vdots \\ Az_{i-1} = a_i z_i \\ \vdots \\ \vdots \end{cases}$$

Where  $(z_1, \dots, z_i, \dots)$  are some vectors in  $Q$  different from the zero tuple and  $(a_1, \dots, a_i, \dots)$  some non zero complex numbers. Since  $A : Q \rightarrow Q$  and  $Q$  is finite dimensional space, we deduce by applying this process that there is at least one  $j$  and one  $k$  such that  $z_j = m z_k$ , for some scalar  $m$ , without loss of generality we assume that  $k = j - s$  this way the vector  $z_{j-1}$  is an eigenvector of  $A^s$ ; that is,  $A^s z_{j-1} = f z_{j-1}$  for some scalar  $f \neq 0$  but an eigenvector of  $A$  is an eigenvector of any power of  $A$  thus we deduce that there exist a vector  $w = z_{j-1}$  in  $Q$  such that  $Aw = f^{\frac{1}{s}} w$ .  $\square$

**Theorem 1.10.** [9] *Let  $A$  and  $B$  be square matrices of the same size. If  $AB = BA$ , then there exists a unitary matrix  $U$  such that  $U^*AU$  and  $U^*BU$  are both upper-triangular.*

*Proof.* By a classic induction on  $n$ . If  $n = 1$ , we have nothing to show. Suppose that the assertion is true for  $n - 1$ .  $A$  has at least one eigenvector in  $\mathbb{C}^n$  for,

$$Ax = \lambda x, x \neq 0, \quad \text{if and only if} \quad \det(\lambda I - A) = 0, \quad (1.6)$$

which has a solution in  $\mathbb{C}$ .

For each eigenvalue  $\mu$  of  $B$ , consider the eigenspace of  $B$

$$V_\mu = \{v \in \mathbb{C}^n : Bv = \mu v\},$$

and for every  $v \in V_\mu$

$$BAv = ABv = \mu(Av).$$

Thus if  $v \in V_\mu$ ,  $Av \in V_\mu$  so  $V_\mu$  is invariant under  $A$ . As a linear transformation on a finite dimensional subspace we conclude from Lemma 1.9 that  $A$  has an eigenvalue say  $\lambda$  and a corresponding eigenvector  $v_1$  in  $V_\mu$ . Assuming that  $v_1$  is a unit vector, extending  $v_1$  to a unitary matrix  $U_1$ , we get

$$U_1^*AU_1 = \begin{pmatrix} \lambda & \alpha \\ 0 & C \end{pmatrix} \quad \text{and} \quad U_1^*BU_1 = \begin{pmatrix} \mu & \beta \\ 0 & D \end{pmatrix},$$

where  $C, D \in \mathbb{M}_{n-1}$ , and  $\alpha, \beta$  are some row vectors.

It follows from  $AB = BA$  that  $CD = DC$ . The induction hypothesis guarantees the existence of a unitary matrix  $U_2 \in \mathbb{M}_{n-1}$ , such that  $U_2^*CU_2$  and  $U_2^*DU_2$  are both upper-triangular. Let

$$U = U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}.$$

Then  $U$  is unitary and  $U^*AU$  and  $U^*BU$  are both upper-triangular.  $\square$

**Corollary 1.11 (Schur Decomposition).** *Let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be eigenvalues of  $A \in \mathbb{M}_n(\mathbb{C})$ . Then there exists a unitary matrix  $U \in \mathbb{M}_n(\mathbb{C})$  such that  $U^*AU$  is an upper-triangular matrix.*

**Remark 1.12.** *If  $A^* = A$ ; that is,  $A$  is self-adjoint, one can see that the upper-triangular matrix is in fact a real diagonal matrix.*

**Theorem 1.13.** [9] Let  $A \in \mathbb{M}_n(\mathbb{C})$  a positive semi-definite operator. There exist a unique positive semi-definite operator  $B$  such that  $B^2 = A$  we denote it by  $B = \sqrt{A}$

We obtain the result easily by diagonalizing  $A$ , if we had  $A = U^* \text{diag}(\lambda_1, \dots, \lambda_n) U$  then  $\sqrt{A} = U^* \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) U$ . For uniqueness (see [9]).

### 1.3. Polar Decomposition

We begin by this important Lemma.

**Lemma 1.14.** If  $A$  and  $B$  are linear operators on  $\mathbb{C}^n$  and if  $A^*A \leq B^*B$  then  $\text{Ker}(B) \subset \text{Ker}(A)$ .

*Proof.* Let  $u \in \text{Ker}(B)$ , then  $Bu = 0$ . However  $A^*A \leq B^*B$  thus

$$\langle A^*Au, u \rangle - \langle B^*Bu, u \rangle \leq 0.$$

Hence  $\|Au\|^2 \leq \|Bu\|^2$ . And this completes the proof.  $\square$

Note that if we replaced the large inequality in lemma 1.14 by an equality then  $\text{Ker}(B) \equiv \text{Ker}(A)$ . We denote by  $\sqrt{A^*A}$  by  $|A|$  and it is called Modulus of  $A$ . Similarly we have:

$$\forall u \in \mathbb{C}^n \quad \| |A|u \|^2 = \langle A^*Au, u \rangle = \|Au\|^2.$$

In particular  $\text{Ker}A = \text{Ker}(|A|)$ , which implies that:  $n - \text{rank}(A) = n - \text{rank}(|A|)$  and if we denote by  $m_0$  the algebraic multiplicity of 0 as eigenvalue, we have  $\text{rank}(A) = \text{rank}(|A|) = \text{rank}(A^*A) = \text{rank}(AA^*)$  since the

rank of an  $n \times n$  Hermitian (diagonalisable) matrix equals  $n - m_0$  and 0 is an eigenvalue of  $|A|$  if and only if it is an eigenvalue of  $A^*A$ , if and only if it is an eigenvalue of  $AA^*$ . For  $A \in \mathbb{M}_n(\mathbb{C})$ , we shall see that the Spectral Decomposition Theorem and The S.V.D Theorem are consequences of the Polar Decomposition Theorem.

**Corollary 1.15 (Polar Decomposition).** *Let  $A \in \mathbb{M}_n(\mathbb{C})$  and  $\sqrt{A^*A} = |A|$ .*

*There exist a unitary matrix  $U \in \mathbb{M}_n(\mathbb{C})$  verifying:  $A = U|A|$ . This decomposition is also unique, if  $A = VB$  with  $B \geq 0$ ,  $V$  a unitary matrix, then  $B = |A|$  and  $V = U$ .*

*Proof.* Giving that  $|A|$  is self-adjoint, we have

$$|A| = V^*\Sigma V \tag{1.7}$$

for some unitary matrix  $V$  and in terms of singular value decomposition of  $A$ ,  $A = W\Sigma V$ , let  $U = WV$  then  $A = U|A|$ . For the uniqueness, if  $A = VB$  then  $A^* = BV^*$  and  $A^*A = BV^*VB$ , but  $V^*V = Id$  thus  $B = |A|$ . Of course if  $|A|$  is invertible we deduce that  $V = U$ , otherwise by noticing that the set of unitary matrices is a closed set, we get by a continuity argument that  $U = V$ . □

**Remark 1.16.** *Since  $A = U|A|$  for some unitary  $U$ , then  $A^* = |A|U^*$ ,  $A^*A = |A|U^*U|A| = |A||A|$ , and  $AA^* = U|A||A|U^*$  so if  $A$  is a square matrix then  $A^*A$  is unitarily congruent to  $AA^*$ . In addition we get as a*



consequence of the previous corollary:

$$W^*|A^*|W = W^*U|A|U^*W = V|A|V^* = \Sigma$$

$$A = U|A|U^*U = WV|A|V^*W^*U = W\Sigma W^*U = |A^*|U$$

**Theorem 1.17** (Spectral Decomposition Theorem). *Let  $A$  be an  $n$ -square complex matrix with eigenvalues  $\lambda_1, \dots, \lambda_n$ . Then  $A$  is normal if and only if  $A$  is unitarily diagonalizable; that is, there exists a unitary matrix  $U$  such that*

$$U^*AU = \text{diag}(\lambda_1, \dots, \lambda_n)$$

*In particular,  $A$  is Hermitian if and only if the  $\lambda_i$  are all real and is positive semi-definite if and only if the  $\lambda_i$  are all nonnegative.*

*Proof.* The matrix  $A$  is normal if and only if  $|A| = |A^*|$ , in particular if  $A = U\text{diag}(\lambda_1, \dots, \lambda_n)U^*$  for some unitary  $U$  then  $A$  is normal, let us show the converse, here the matrix notations are those of Corollary 1.15, we have  $U|A| = |A|U$ , From Theorem 1.10 it is possible to diagonalize  $U$  with a unitary matrix  $S$ , to see this note that  $|A|$  is Hermitian so  $S|A|S^*$  is the diagonal matrix  $\Sigma$ , in the other hand  $SUS^*$  is an upper triangular matrix  $O$ , but since  $U$  is unitary *i.e.*  $U^*U = Id$ ,  $O$  is necessarily a unitary diagonal matrix. Putting  $Q = S^*$ , with  $\Upsilon = O.\Sigma$  and  $(O.\Sigma)(O.\Sigma)^* = \Upsilon.\Upsilon^* = \Sigma^2$ , we rewrite the polar decomposition as

$$A = (QOQ^*)(Q\Sigma Q^*)$$

that way  $A$  has a spectral decomposition

$$A = Q\Upsilon Q^*$$

□

## 1.4. Jensen Inequality and Applications

A real function  $f(t)$  defined on a vector space  $E$  is said to be convex if for all  $x, y \in E$  and all  $\alpha, \beta$  positive verifying  $\alpha + \beta = 1$  we have

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y). \quad (1.8)$$

A function is strictly convex if the above strict inequality holds whenever  $x \neq y$ . A general equivalent, yet important, form of (1.8) is known as Jensen's inequality.

**Theorem 1.18 (Jensen's inequality).** *Let  $f$  be a convex function*

*$(x_1, \dots, x_n) \in E, (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^+$  such that  $\sum_{i=1}^n \lambda_i = 1$ , then*

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Inequality is reversed if  $f$  is concave (i.e  $-f$  is convex).

*Proof.* By induction on  $n$ , the cases  $n = 1$  and  $2$  are trivial. Suppose it is true for  $n - 1$ , one of the sums  $\lambda_i + \lambda_j$ , ( $i \neq j$ ) is different from zero, for example  $(\lambda_1 + \lambda_2) \neq 0$ , and we may write by the induction hypothesis :

$$f\left((\lambda_1 + \lambda_2)\frac{\lambda_1 x_1 + \lambda_2 x_2}{\lambda_1 + \lambda_2} + \dots + \lambda_n x_n\right) \leq f\left(\frac{\lambda_1 x_1 + \lambda_2 x_2}{\lambda_1 + \lambda_2}\right) (\lambda_1 + \lambda_2) + \sum_{i=3}^n \lambda_i x_i$$

Applying hypothesis for  $n = 2$  to the first term of the right hand side yields to the inequality.  $\square$

**Theorem 1.19.** *Let  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  be a strictly convex function with  $f(0) \leq 0$ . If  $x_1, \dots, x_n$  are nonnegative numbers and at least two  $x_i$  are nonzero, then :*

$$\sum_{i=1}^n f(x_i) < f\left(\sum_{i=1}^n x_i\right)$$

*Proof.* Also by induction, let us prove it first for  $n = 2$  and the generalization follows easily. Let  $x_1, x_2 > 0$ . Write  $x_1 = \frac{x_1}{x_1+x_2}(x_1 + x_2) + \frac{x_2}{x_1+x_2} \cdot 0$ , similar for  $x_2$  we have from the convexity of  $f$ :

$$(1) \quad f(x_1) < \frac{x_1}{x_1+x_2}f(x_1 + x_2) + \frac{x_2}{x_1+x_2}f(0).$$

$$(2) \quad f(x_2) < \frac{x_2}{x_1+x_2}f(x_1 + x_2) + \frac{x_1}{x_1+x_2}f(0)$$

By adding 1 and 2 we get  $f(x_1) + f(x_2) < f(x_1 + x_2) + f(0)$  with  $f(0) \leq 0$  and thus  $f(x_1) + f(x_2) < f(x_1 + x_2)$ .  $\square$

Some consequences of Theorem 1.18 are: Arithmetic geometric mean inequality, Hadamard inequality, Minkowski Determinant inequality, Hölder and Minkowski inequalities.

**Proposition 1.20** (Arithmetic Geometric mean inequality). *If  $(x_1, \dots, x_n)$  is an  $n$ -tuple of nonnegative real numbers, then:*

$$\frac{1}{n} \sum_{i=1}^n x_i \geq \sqrt[n]{\prod_{i=1}^n x_i}.$$

*Proof.* It is well known that the function  $\ln(x)$  is strictly concave on

$\mathbb{R}_+^*$ . Therefore

$$\ln\left(\sum_{i=1}^n \lambda_i x_i\right) \geq \sum_{i=1}^n \lambda_i \ln(x_i)$$

where  $\lambda_i = \frac{1}{n}$ . By taking the exponential of each side of the inequality we obtain the result.  $\square$

**Proposition 1.21** (Hadamard Inequality). *Let  $H = (h_{ij}) \in \mathbb{M}_n$  be a positive semi-definite matrix, then*

$$\det H \leq \prod_{j=1}^n h_{jj}.$$

*Proof.* If  $\det H = 0$  there is nothing to prove, because  $h_{jj}$  are nonnegative. Otherwise  $H$  is positive definite and one has  $h_{jj} > 0$  let  $D = \text{diag}\left(\frac{1}{\sqrt{h_{11}}}, \dots, \frac{1}{\sqrt{h_{nn}}}\right)$  and  $H' = DHD$ . Then we have:

$$\frac{\det H}{\prod_j h_{jj}} = \det DHD = \det H'. \quad (1.9)$$

The diagonal entries of  $H'$  are all equal 1 and its eigenvalues  $\mu_1, \dots, \mu_n$  are strictly positive and their sum is  $n$ . Require to proof that  $\det H' \leq 1$ .

But

$$\frac{1}{n} \ln(\det H') = \frac{1}{n} \sum_j \ln \mu_j \leq \ln \left( \frac{1}{n} \sum_j \mu_j \right) = 0$$

again taking exponential at each side completes the proof. Since the concavity is strict (  $\ln$  is strictly concave) the equality holds only if  $\mu_1 = \dots = \mu_n = 1$  but then  $H'$  is similar thus equal to  $I_n$ . In this case from (1.9)  $H$  is diagonal.  $\square$

**Proposition 1.22.** [1] *If  $A$  is an  $n \times n$  P.D. matrix, then :*

$$(\det A)^{\frac{1}{n}} = \min \left\{ \frac{\text{tr}(AB)}{n} : B \geq 0 \text{ and } \det B = 1 \right\}. \quad (1.10)$$

*If  $A$  is P.S.D., then the same relation holds with min replaced by inf.*

*Proof.* It suffices to prove the statement about P.D. matrices the P.S.D case follow by continuity argument because the function in braces is a continuous function over the set of matrices and the set of P.D. matrices is dense in the set of P.S.D. matrices. Using the spectral decomposition theorem and the cyclicity of the trace, we may assume that  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ , using the arithmetic geometric mean inequality we have

$$\frac{\text{tr}(AB)}{n} = \frac{1}{n} \sum_{i=1}^n \lambda_i b_{ii} \geq \left( \prod_{i=1}^n \lambda_i \right)^{\frac{1}{n}} \left( \prod_{i=1}^n b_{ii} \right)^{\frac{1}{n}} \geq (\det A)^{\frac{1}{n}} (\det B)^{\frac{1}{n}} \quad (1.11)$$

for any positive matrix  $B$  such that  $\det B = 1$ . We will have an equality in (1.11) if  $B = (\det A)^{\frac{1}{n}} A^{-1}$ .  $\square$

**Corollary 1.23** (The Minkowski Determinant Inequality). *If  $A, B$  are  $n \times n$  positive matrices then :*

$$(\det(A + B))^{\frac{1}{n}} \geq (\det A)^{\frac{1}{n}} + (\det B)^{\frac{1}{n}}$$

*Proof.* For any  $D \geq 0$  with  $\det D = 1$ , we have  $\frac{\text{tr}(A + B)D}{n} = \frac{\text{tr}(AD)}{n} + \frac{\text{tr}(BD)}{n}$ , let us denote by  $\mathbb{P}_n$  the set of positive semi-definite matrices  $D$

of order  $n$  whose determinant equals one, thus

$$(\det(A + B))^{\frac{1}{n}} = \min \left\{ \frac{\operatorname{tr}(AD)}{n} + \frac{\operatorname{tr}(BD)}{n} : D \in \mathbb{P}_n \right\} \quad (1.12)$$

$$\geq \min \left\{ \frac{\operatorname{tr}(AD)}{n} : D \in \mathbb{P}_n \right\} + \min \left\{ \frac{\operatorname{tr}(BD)}{n} : D \in \mathbb{P}_n \right\} \quad (1.13)$$

$$= (\det A)^{\frac{1}{n}} + (\det B)^{\frac{1}{n}} \quad (1.14)$$

and that completes the proof, we deduce also that  $(\det(A+B)) \geq (\det A) + (\det B)$ .  $\square$

Another consequence is the following :

**Corollary 1.24.** *If  $A = B + iC \geq 0$  ( P.S.D.) and  $B$  and  $C$  are real matrices, then*

$$\det A \leq \det B.$$

*Proof.*  $C$  must have all diagonal entries equals zero because the diagonal entries of  $A$ ,  $B$  and  $C$  are real. Now taking the adjoint of  $A$ :

$$A^* = A = B^* - iC^* = B^T - iC^T = B + iC$$

Therefore

$$C = -C^T \text{ and } B = B^T.$$

Now put the Hermitian matrix  $B$  into diagonal form by a unitary  $U$ , so  $U^*AU = U^*BU - iU^*CU \geq 0 \implies U^*CU$  has also 0 as diagonal entries for the same reason as  $C$ , this implies that all eigenvalues of  $B$  are nonnegative therefore  $B \geq 0$ . Since the diagonal entries of  $U^*BU - iU^*CU$  are the

eigenvalues of  $B$ , from the Hadamard inequality we obtain  $\det U^*AU = \det(U^*BU - iU^*CU) \leq \det B$ .  $\square$

**Proposition 1.25.** *For any two Hermitian matrices  $A$  and  $B$  we have this inequality:*

$$\operatorname{tr}(AB)^2 \leq \operatorname{tr}(A^2.B^2).$$

*Proof.* Since  $\operatorname{tr}(AB) = \operatorname{tr}(U^*ABU) = \operatorname{tr}(U^*AUU^*BU)$  for any unitary matrix  $U$  we can consider that  $B$  is a diagonal matrix. Consequently we have

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \overline{a_{1,2}} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1,n}} & \overline{a_{2,n}} & \cdots & a_{n,n} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1} & 0 & \cdots & 0 \\ 0 & b_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{n,n} \end{pmatrix}$$

A direct computation shows that

$$\operatorname{tr}(AB)^2 - \operatorname{tr}(A^2.B^2) = \frac{\sum_{i,j=1}^n |a_{i,j}|^2 (2b_{i,i}b_{j,j} - b_{i,i}^2 - b_{j,j}^2)}{2}.$$

By the arithmetic geometric mean inequality we have for all  $1 \leq i \leq n$ ,  $1 \leq j \leq n$

$$2b_{i,i}b_{j,j} - b_{i,i}^2 - b_{j,j}^2 \leq 0$$

which implies that  $\operatorname{tr}(AB)^2 - \operatorname{tr}(A^2.B^2) \leq 0$ . Equality holds if and only if  $AB = BA$ .  $\square$

## 1.5. Numerical Range

A square matrix  $A$  has many characterizations, one of them is his numerical range.

**Definition 1.26.** *The numerical range also known as field of values of  $A \in \mathbb{M}_n$  is*

$$W(A) \equiv \{x^*Ax; x \in \mathbb{C}, x^*x = 1\} = R_A(x)$$

Thus  $W(A)$  is a function from  $\mathbb{M}_n$  into subsets of the complex plane .

***Properties:***

(1) For all  $A \in \mathbb{M}_n$  and  $\alpha \in \mathbb{C}$ ,  $W(A + \alpha I) = W(A) + \alpha$ .

(2)  $W(\alpha A) = \alpha W(A)$ .

(3)  $\rho(A) \subset W(A)$ , where  $\rho(A)$  stands for the set of eigenvalues of  $A$ .

(4)  $W(A)$  is invariant under unitary congruences, that is:  $W(U^*AU) = W(A)$ .

(5)  $F(A) = \{x^*Ax, x \in \mathbb{C}^n\}$  is invariant under unitary congruences; that is,  $F(A) = F(U^*AU)$ .

*Proof.*

(1) We have  $W(A + \alpha I) = \{x^*(A + \alpha I)x; x^*x = 1\} = \{x^*Ax + \alpha x^*x; x^*x = 1\} = W(A) + \alpha$ .

(2) We have  $\langle \alpha Ax, x \rangle = \alpha \langle Ax, x \rangle$ . for all  $\alpha \in \mathbb{C}$ ,  $x \in \mathbb{C}^n$ .



(3) Let  $(x_1, \dots, x_n)$  be the set of normalized eigenvectors corresponding to  $(\lambda_1, \dots, \lambda_n)$  respectively, we have  $\langle Ax_i, x_i \rangle = \lambda_i \|x_i\|^2 = \lambda_i$ .

(4) First  $\|Ux\| = \|x\|$  whenever  $U$  is unitary, thus  $U$  leaves invariant the surface of the Euclidean unit ball, and we have for  $x$  such that  $\|x\| = 1$ ,  $W(U^*AU) = \langle U^*AUx, x \rangle = \langle AUx, Ux \rangle = W(A)$ .

(5)  $U$  is bijective and  $F(U^*AU) = \langle U^*AUx, x \rangle = \langle AUx, Ux \rangle = F(A)$ .

□

As  $W(A)$  is the range of a compact set by a continuous function,  $W(A)$  is compact. The well known result is:

**Theorem 1.27 (Toeplitz-Hausdorff).** [9] *The numerical range of a square matrix is a convex compact subset of the complex plane .*

We denote by  $Co(\rho(A))$  the convex hull of the set of eigenvalues of  $A \in \mathbb{M}_n(\mathbb{C})$

**Proposition 1.28.** *If  $A \in \mathbb{M}_n$  is normal, then*

$$W(A) = Co(\rho(A)).$$

*Proof.* If  $A$  is normal, then  $A = U^*\Lambda U$ , where  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  and  $U$  is unitary. By the unitary similarity property,  $W(A) = W(\Lambda)$ , since

$$x^* \Lambda x = \sum_{i=1}^n \bar{x}_i x_i \lambda_i = \sum_{i=1}^n |x_i|^2 \lambda_i$$

and  $\|x\| = 1$  implies that  $\sum_{i=1}^n |x_i|^2 = 1$ ,  $W(A)$  is the set of all convex combination of the eigenvalues of  $A$ . □

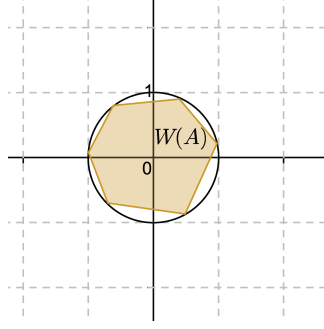


Figure 1.1.:  $W(A)$  of a normal (unitary) matrix  $A$

Since  $W(A)$  is convex; one can see that  $\frac{1}{n}\text{tr}(A) = \frac{1}{n}\lambda_1 + \dots + \frac{1}{n}\lambda_n \in W(A)$  as convex combination of the eigenvalues (elements belonging to  $W(A)$ ). Let  $|\cdot|$  denote the modulus of a vector in  $\mathbb{C}$ .

**Lemma 1.29.** *Let  $A = (a_{ij})$  be an  $n \times n$  complex matrix with numerical radius defined as*

$$s(A) := \max_{\|x\|=1} |\langle Ax, x \rangle|.$$

*If  $A$  has nonnegative entries then  $s(A) = \max_{\|x\|=1} \{\langle Ax, x \rangle, x \in \mathbb{R}^n\}$*

*Proof.* It is clear that there exist a certain unit vector  $x_0 = (\xi_1, \dots, \xi_n)^T$  such that  $s(A) = |\langle Ax, x \rangle|$ . Since  $A$  is positive entriwise and the vector  $y := (|\xi_1|, \dots, |\xi_n|)^T$  in  $\mathbb{R}^n$  has norm 1, by the Cauchy Schwartz inequality we have

$$s(A) \leq |\langle Ax_0, x_0 \rangle| = \left| \sum_{i,j} \langle a_{ij} \xi_j, \xi_i \rangle \right| \leq \sum_{i,j} |a_{ij}| |\xi_j| |\xi_i| \quad (1.15)$$

$$= |\langle Ay, y \rangle| \leq s(A). \quad (1.16)$$

and that completes the proof . □

**Proposition 1.30.** Let  $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Then  $W(N)$  is the closed disc  $D := \{x \in \mathbb{C} : |x| \leq \frac{1}{2}\}$ , and  $F(N)$  is the whole complex plan.

*Proof.* The direct proof goes as follows: since  $x \in \mathbb{C}^2 \simeq \mathbb{R}^4$  and  $\|x\| = 1$ , we can choose to represent  $x$  as if  $x \in \mathbb{S}^3$  see Theorem A.1, but Lemma 1.29 make it much easier,

$$W(N) = (\bar{x}_1 \quad \bar{x}_2) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \langle x_2, x_1 \rangle$$

By the Cauchy-Schwartz inequality

$$2|\langle x_2, x_1 \rangle| \leq 2|x_1||x_2|$$

but  $|x_1|^2 + |x_2|^2 = 1$  thus we can write

$$2|\langle x_2, x_1 \rangle| \leq 2|\sin \theta \cos \theta| := |\sin 2\theta| \leq 1 \quad (1.17)$$

If  $x_2 = \sin \theta$  and  $x_1 = \cos \theta$  we get an equality in 1.17, for any  $\alpha \in [0, 2\pi]$  let  $x_1 = \frac{\sqrt{2}}{2}e^{i\alpha}$  and  $x_2 = \frac{\sqrt{2}}{2}$ , consequently  $|x_1|^2 + |x_2|^2 = 1$  and  $\langle x_2, x_1 \rangle = \frac{1}{2}e^{i\alpha}$ , thus the circle of radius  $\frac{1}{2}$  is in the range of  $W(A)$ , and we can see from the convexity of the numerical range that it is the entire closed disc of radius  $\frac{1}{2}$ , to complete the proof notice that any vector  $x$  of norm  $\kappa$  can be written as  $\kappa \frac{x}{\|x\|}$ . □

**Example 1.** For  $x \in \mathbb{C}^n, x_j = a_j + ib_j, j = 1, \dots, n$  and  $\|x\| = 1$

- $W \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  is the set  $\{|a_1 + ib_1|\} \equiv [0, 1]$
- $W \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$  is the set  $\{2(\bar{x}_1 \cdot x_2) : \|x\| = 1\} \equiv \{x \in \mathbb{C} : |x| \leq 1\}$

**Proposition 1.31** (Rotation). *Let  $A \in \mathbb{M}_2(\mathbb{R})$  then there exist a real rotation matrix  $U$  such that  $U^T A U$  have equal diagonal entries. In addition if  $A$  has real eigenvalues, there exist a real rotation matrix  $U$  such that  $U^T A U$  have the diagonal entries equal to  $\frac{1}{2} \text{tr}(A)$ .*

*Proof.* Let  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_2(\mathbb{R})$ . A direct computation shows that:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} l & * \\ * & l \end{pmatrix}$$

where  $l$  is any real number and  $(*)$  stands for unspecified entries, if and only if

$$\begin{aligned} a \cos^2 \theta - (c + b) \cos \theta \sin \theta + d \sin^2 \theta &= d \cos^2 \theta + (c + b) \cos \theta \sin \theta \\ &\quad + a \sin^2 \theta \end{aligned} \tag{1.18}$$

$$(a - d)(\cos^2 \theta - \sin^2 \theta) = 2(c + b) \sin \theta \cos \theta \tag{1.19}$$

$$(a - d)(\cos(2\theta)) = (c + b)(\sin(2\theta)) \tag{1.20}$$

$$\frac{a - d}{c + b} = \tan(2\theta) \tag{1.21}$$

since  $\tan(2\theta)$  can be any real number for some  $\theta$ , we get the result. As for the second assertion let us take  $U_2$  to be of the form  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  for some real number  $\theta$ . By considering  $A' = A - \frac{1}{2} \text{tr}(A) I$  we may assume that  $\text{tr}(A) = 0$ , by the Schur Decomposition Theorem,  $A$  is unitarily congruent to an upper triangular matrix  $D$  i.e.  $\exists U_1$  unitary such that  $U_1^* A U_1 = D$  but the eigenvalues of  $A$  are the real numbers  $\pm \lambda$  and  $A$  have real entries thus we can assume that  $U_1^T = U_1^*$ . It suffices to find  $\theta$  of another orthogonal

matrix  $U_2$  such that  $U_2^T \begin{pmatrix} \lambda & g \\ 0 & -\lambda \end{pmatrix} U$  has zero diagonal entries, here  $g$  is some real number, after computation we get

$$\lambda \cos^2 \theta + g \cos \theta \sin \theta - \lambda \sin^2 \theta = 0 \quad (1.22)$$

$$\lambda(\cos^2 \theta - \sin^2 \theta) = -g \cos \theta \sin \theta \quad (1.23)$$

$$\frac{2\lambda}{-g} = \tan(2\theta) \quad (1.24)$$

and this has a solution also because the range of the function  $\theta \rightarrow \tan \theta$  is all the real line. The result follows by taking  $U = U_1 U_2$ .  $\square$

Noting the non triviality of finding  $W(A)$  it is known that (see [12]) for any  $2 \times 2$  matrix  $W(A)$  is an elliptical disc whose parameters are entirely defined by  $A$ .

**Lemma 1.32.** *Let  $A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ , Then  $W(A)$  is the closed elliptical disc with foci at  $(0, 0)$  and  $(1, 0)$  minor axis of length 1, major axis of length  $\sqrt{2}$ .*

*Proof.* Recall from Proposition 1.30 that  $W(N)$  is the disc of center 0 and radius  $\frac{1}{2}$ .  $G = A - \frac{1}{2}I_2$  is a real matrix with real eigenvalues and a trace equal 0, thus from Proposition 1.31 it is unitarily congruent to a matrix  $S$  whose diagonal entries are 0 *i.e.*  $S = aN + bN^T$  for some real numbers  $a$  and  $b$ .  $G^T$  is upper triangular and has the same numerical range as  $G$ , applying the second part of Proposition 1.31 we obtain by a direct computation that  $a = \frac{\sqrt{2} + 1}{2}$  and  $b = \frac{\sqrt{2} - 1}{2}$ , by translation property and invariance under unitary similarity,  $W(A)$  is the range of the

transformation  $h$ , where

$$h : \{x \in \mathbb{C}^2, \|x\| = 1\} \rightarrow \frac{1}{2} + ax_1 + b\bar{x}_1$$

consequently  $|h(x)| \leq \frac{1}{2} + a|x_1| + b|\bar{x}_1|$  so the upper bound (right hand side) is maximal if  $x_2 = 0$  but in this case we can consider  $(x_1) := (|a_1|, |b_1|) \in \mathbb{S}_1$  which implies that  $x_1 = \cos \theta + i \sin \theta$  and

$$h(x) = \frac{1}{2} + a \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + b \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix} = W(A),$$

with  $a = \frac{\sqrt{2} + 1}{2}$  and  $b = \frac{\sqrt{2} - 1}{2}$ , upon replacing in  $h(x)$  we have the parametric equations of the boundary of the convex set  $W(A)$  in the complex plane (see Figure 1.2). □

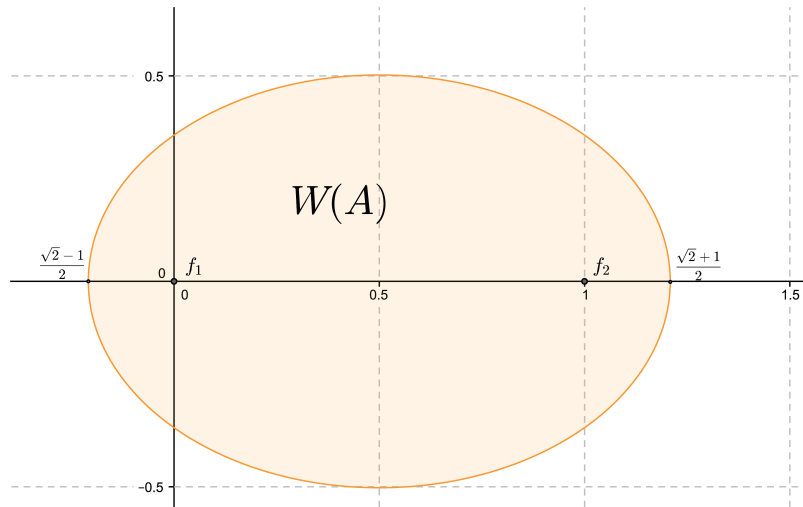


Figure 1.2.:  $W(A)$  of the  $2 \times 2$  matrix  $A$

## CHAPTER 2

# Some Characterizations

### 2.1. Eigenvalues

In this section we will state an important Theorem that help us find the eigenvalues of certain matrices.

**Theorem 2.1.** *Let  $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$  be any square matrix written by blocks of same size, if  $AC = CA$  then  $\det(M) = \det(AD - CB)$ .*

*Proof.* Suppose first that  $A$  is invertible, let us write  $M$  as

$$M = \begin{pmatrix} Z & 0 \\ V & I \end{pmatrix} \begin{pmatrix} I & E \\ 0 & F \end{pmatrix}, \quad (2.1)$$

upon calculation we find that:  $Z = A, V = C, E = A^{-1}B, F = D - CA^{-1}B$

taking the determinant on each side of (2.1) we get:

$$\det(M) = \det(A(D - CA^{-1}B)) = \det(AD - CB).$$

the result follows by a continuity argument since the Determinant function is a continuous function. □

If  $X$  is any  $m \times n$  matrix, we say that  $X \in \mathbb{M}_{m,n}$ , the following is an application of Theorem 2.1:

**Lemma 2.2.** *Let  $X \in \mathbb{M}_{m,n}$  then the eigenvalues of  $H = \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}$  are the real numbers  $(-\sigma_1 \cdots, -\sigma_n, \sigma_n, \cdots, \sigma_1)$  where  $\sigma_n \leq \cdots \leq \sigma_1$  are the singular values of  $X$ .*

*Proof.* For any matrix  $X$ ,  $X^*$  will commute with  $A$  if  $A$  has all of its components zero, the eigenvalues of  $H$  are the roots of

$$\begin{aligned} \det(H - \lambda I_{m+n}) &= \det(\lambda^2 I_n - X^* X) \\ &= \det(\lambda I_n - (X^* X)^{\frac{1}{2}})(\lambda I_n + (X^* X)^{\frac{1}{2}}) = 0 \end{aligned}$$

from this, we see that  $(-\sigma_1 \cdots, -\sigma_n, \sigma_n, \cdots, \sigma_1)$  are the eigenvalues of  $H$ . □

Let  $I_k$  denotes the  $k \times k$  identity matrix, for any matrix  $R$  written by blocks such that  $R = \begin{pmatrix} A & X \\ Y & B \end{pmatrix}$  with  $A \in \mathbb{M}_m$  and  $B \in \mathbb{M}_n$  we have

$PRP^T = \begin{pmatrix} B & Y \\ X & A \end{pmatrix}$  where  $P$  is the orthogonal matrix  $\begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix}$ . Let  $H = \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}$  and suppose that  $X$  is not a square matrix; that is,  $X \in \mathbb{M}_{m,n}$

and  $m \neq n$ , since the eigenvalues of  $\begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}$  are the same as those of  $\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}$ , we will assume that  $m > n$ .

**Lemma 2.3.** *Let  $X \in \mathbb{M}_{m,n}$  with  $m > n$ , then the eigenvalues of  $H = \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}$  are the real numbers  $(-\sigma_1 \cdots, -\sigma_n, 0_{m-n}, \sigma_n, \cdots, \sigma_1)$  where  $\sigma_n \leq \cdots \leq \sigma_1$  are the singular values of  $X$ , and  $0_{m-n}$  is the  $1 \times (m - n)$  vector which has  $m - n$  zero components.*



*Proof.* We have  $H = \begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}$ , and  $X \in \mathbb{M}_{m,n}$ , in order to apply Theorem 2.1 we add to  $H$  an  $(m-n) \times (m-n)$  zero block  $O$  to obtain  $H'$  whose eigenvalues are the same of those of  $H$  union those of the zero block  $O$ , so  $H' = \begin{pmatrix} 0 & X & 0 \\ X^* & 0 & \vdots \\ 0 & \dots & O \end{pmatrix} = \begin{pmatrix} 0 & Y \\ Y^* & 0 \end{pmatrix}$  where  $Y$  is a square matrix of order  $m$ . Applying the previous result the eigenvalues of  $H'$  are  $(-\sigma_1 \cdots, -\sigma_n, 0_{2(m-n)}, \sigma_n, \cdots, \sigma_1)$  and we deduce that the eigenvalues of  $H$  are precisely  $(-\sigma_1 \cdots, -\sigma_n, 0_{m-n}, \sigma_n, \cdots, \sigma_1)$ .  $\square$

## 2.2. Rayleigh Quotient

Let  $A$  be a Hermitian matrix. By the Spectral Decomposition Theorem the eigenvalues of  $A$  are all real, we order them (here and after) in decreasing order, that is  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . The numerical range of a Hermitian matrix  $A$  is called the Rayleigh Quotient  $R_A(v)$  which is the application (Rayleigh quotient):

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

In fact this application is well defined because  $v^* v = \|v\|^2$  is real and  $v^* A v$  is real due to the fact that:  $\overline{v^* A v} = (v^* A v)^* = v^* A^* (v^*)^* = v^* A v$ .

**Proposition 2.4.** *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  $A v_i = \lambda_i v_i$ . 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:

$$(3) \text{ Let } V_k \text{ be the subspace of } \mathbb{C}^n \text{ spanned by } \{v_1, v_2, \dots, v_k\},$$

$$\text{and } W_k = \text{span} \{v_{k+1}, \dots, v_n\} = V_k^\perp$$

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

**Remark 2.5.** Note the difference between (1) and (2) : in (2) the two eigenvalues are uniquely determined by  $B$  and  $C$  respectively. Obviously the dimension of  $W_{k-1}$  is  $n - k + 1$ , and  $\dim(V_k)$  is equal to  $k$ .

*Proof.* Recall that for a Hermitian matrix  $A$  there is an orthonormal basis in which  $A$  is a diagonal matrix :  $\Delta = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  so taking any vector  $v$  in  $\mathbb{C}^n$  and writing  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  as its coordinates in that base:

$$v^* A v = (\alpha_1, \alpha_2, \dots, \alpha_n) \Delta \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \sum_{i=1}^n \lambda_i |\alpha_i|^2. \quad (2.2)$$

We get this formula:

$$R_A(v) = \frac{\sum_{i=1}^n \lambda_i |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2}.$$

As a matter of fact this implies  $R_A(\text{span}\{v_i, \dots, v_j\}) = [\lambda_i, \lambda_j]$  and that completes the proof.  $\square$

### 2.3. Modulus of a Matrix

The main result of this section is that for any square matrices  $A$  and  $B$  of the same size, there exist unitary matrices  $U$  and  $V$  such that

$$|A + B| \leq U^*|A|U + V^*|B|V.$$

Notice that for any square (Hermitian) matrix  $A$ ,  $A \leq |A|$ . In fact there exist  $A, B \in \mathbb{M}_n$  such that  $|A+B| \leq |A|+|B|$  is not true, (see Example 3). Unless otherwise we will denote here and after  $\lambda^\downarrow(A)$  the vector in  $\mathbb{R}^n$  of eigenvalues of  $A$  in decreasing order, that is  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , and  $\lambda^\uparrow(A)$  eigenvalues of  $A$  arranged in increasing order, and  $\lambda(A)$  short means the 'eigenvalue' vector but in any index ordering. The same will be annotated for any vector in  $\mathbb{R}^n$ . For Hermitian matrices we will see in the next section that

$$\text{If } A \geq B \text{ then } \lambda_i(A) \geq \lambda_i(B), \forall i = 1, \dots, n. \quad (2.3)$$

However the inequalities  $\lambda_i(A) \geq \lambda_i(B)$ , for all  $i$  cannot ensure  $A \geq B$  as the following example shows:

**Example 2.**

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

Then  $\lambda_1(A) = \lambda_1(B)$  and  $\lambda_2(A) \geq \lambda_2(B)$ . But  $A - B$  is not P.S.D since it has a negative eigenvalue.

**Example 3.** Take  $A = \begin{pmatrix} 0 & 2 \\ \frac{1}{2} & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$  then  $|A|+|B| = \begin{pmatrix} 2.5 & 0 \\ 0 & 2 \end{pmatrix}$

$$\lambda_1(|A + B|) \approx 2.85 > 2.5 = \lambda_1(|A| + |B|)$$

and

$$\lambda_2(|A + B|) \approx 0.35 < 2 = \lambda_2(|A| + |B|).$$

By (2.3) we can't have  $|A + B| \leq |A| + |B|$ .

**Theorem 2.6.** Let  $A, B \in \mathbb{M}_n(\mathbb{C})$  be Hermitian matrices. If  $\lambda_i(A) \geq \lambda_i(B)$  for all  $i = 1, \dots, n$ , then there exist a unitary matrix  $U$  such that  $U^*AU \geq B$ .

*Proof.* Let  $P$  and  $Q$  be unitary matrices such that:

$$A = P^* \text{diag}(\lambda_1(A), \dots, \lambda_n(A))P \quad (2.4)$$

$$B = Q^* \text{diag}(\lambda_1(B), \dots, \lambda_n(B))Q \quad (2.5)$$

The condition  $\lambda_i(A) \geq \lambda_i(B)$   $i = 1, \dots, n$  implies that

$$\text{diag}(\lambda_1(A), \dots, \lambda_n(A)) \geq \text{diag}(\lambda_1(B), \dots, \lambda_n(B)) \quad (2.6)$$

Therefore

$$Q^* \text{diag}(\lambda_1(A), \dots, \lambda_n(A))Q \geq Q^* \text{diag}(\lambda_1(B), \dots, \lambda_n(B))Q = B \quad (2.7)$$

Hence

$$Q^*PAP^*Q \geq B.$$

Thus  $U^*AU \geq B$  where  $U = P^*Q$  is unitary. □

**Theorem 2.7.** *Let  $A$  be a complex square matrix. Then*

$$\lambda_i \left( \frac{A^* + A}{2} \right) \leq \lambda_i(|A|) \quad i = 1, 2, \dots, n$$

*Proof.* Take  $v_1, v_2, \dots, v_n$  and  $w_1, w_2, \dots, w_n$  to be orthonormal sets of eigenvectors of  $\frac{A^*+A}{2}$  and  $A^*A$ , respectively. For each  $1 \leq k \leq n$ , let  $S_1 = \text{span}\{v_1, v_2, \dots, v_k\}$ .  $S_2 = \text{span}\{w_k, \dots, w_n\}$  since  $\dim(S_1) + \dim(S_2) > n$ , we have some unit vector  $x \in S_1 \cap S_2$ . Of course by Proposition 2.4

$$x^* \left( \frac{A^* + A}{2} \right) x \geq \lambda_k \left( \frac{A^* + A}{2} \right)$$

and

$$x^*(A^*A)x \leq \lambda_k(A^*A) := (\lambda_k(|A|))^2.$$

We have consequently:

$$\lambda_k \left( \frac{A^* + A}{2} \right) \leq x^* \left( \frac{A^* + A}{2} \right) x = \frac{x^*Ax + \overline{x^*Ax}}{2} = \text{Re}(x^*Ax) \quad (2.8)$$

$$\leq |x^*Ax| \stackrel{\text{C-S}}{\leq} \sqrt{x^*A^*Ax} \leq \lambda_k(|A|) \quad (2.9)$$

□

**Remark 2.8.** *Consequently, from Theorem 2.6 we know that there exist a unitary matrix  $U$  such that  $\frac{A^* + A}{2} \leq U^*|A|U$ .*

**Theorem 2.9 (Thompson).** *For any square complex matrices  $A$  and  $B$  of same size, there exist unitary matrices  $U$  and  $V$  such that*

$$|A + B| \leq U^*|A|U + V^*|B|V.$$

*Proof.* By the Polar Decomposition we may write  $A + B = W|A + B|$  where  $W$  is a unitary matrix.

$$\Leftrightarrow |A + B| = \overbrace{W^*(A + B)}^{\text{Hermitian}} = \frac{1}{2}(W^*(A + B) + (A + B)^*W) \quad (2.10)$$

$$= \frac{1}{2}(A^*W + W^*A) + \frac{1}{2}(B^*W + W^*B) \quad (2.11)$$

$$\leq U^*|W^*A|U + V^*|W^*B|V \quad \text{By Remark 2.8} \quad (2.12)$$

$$= U^*|A|U + V^*|B|V \quad (2.13)$$

□

## 2.4. Minimax principle

In this section let  $\lambda^\downarrow(A) = (\lambda_1, \dots, \lambda_n)$  denote the eigenvalues and  $(v_1, \dots, v_n)$  the corresponding eigenvectors of an  $n \times n$  Hermitian matrix  $A$ . Let  $\sigma^\downarrow(A) = (\sigma_1, \dots, \sigma_n)$  be the singular value vector of  $A$ . Here is another yet equivalent form of Proposition 2.4.

**Proposition 2.10.** *Let  $A$  be an  $n \times n$  Hermitian matrix then*

$$\lambda_1(A) = \max\{R_A(v)\} \text{ so } \lambda_1 = R_A(v_1) \text{ for some } v_1 \in \mathbb{C}^n.$$

⋮

$$\lambda_k(A) = \max\{R_A(v), v \perp v_1, v_2, \dots, v_{k-1}\} \text{ and}$$

$$\lambda_k = R_A(v_k) \text{ for some } v_k \perp v_1, v_2, \dots, v_{k-1}.$$

*Proof.* The characterization of  $\lambda_1$  is already proved, since the unit ball is a compact set of  $\mathbb{C}^n$  and every closed set in a compact is compact, the

continuous function  $x^*Ax$  will attain its maximum on all described sets, we need to show the proposition for  $\lambda_2$  and the rest follows easily; take any vector  $w$  such that  $w \perp v_1$  i.e  $w$  orthogonal to  $v_1$ , then  $w = \sum_{i=2}^n \alpha_i v_i$  but from Proposition 2.4  $R_A(\text{span}\{v_n, \dots, v_2\}) = [\lambda_n, \lambda_2]$ , thus  $R_A(w) \leq R_A(v_2)$  since  $v_2 \perp v_1$ , this completes the proof.  $\square$

**Corollary 2.11.** *Let  $A$  be an  $m \times n$  matrix, and let  $\sigma_1 \geq \dots \geq \sigma_n$  denote the singular values of  $A$ , we have*

$$\sigma_1(A) = \max\{\|Ax\|, \|x\| = 1\} \text{ so } \sigma_1 = \|Ax_1\| \text{ for some } x_1 \in \mathbb{C}^n.$$

$\vdots$

$$\sigma_k(A) = \max\{\|Ax\|, x \perp x_1, x_2, \dots, x_{k-1}, \|x\| = 1\} \text{ and}$$

$$\sigma_k = \|Ax_k\| \text{ for some } x_k \perp x_1, x_2, \dots, x_{k-1}.$$

*Proof.* Since  $\sigma_i^2(A)$  are the eigenvalues of  $A^*A$ , and  $\|Ax\|^2 = \langle Ax, Ax \rangle = \langle A^*Ax, x \rangle$  for every  $x \in \mathbb{C}^n$ , we get the desired representation from Proposition 2.10.  $\square$

**Theorem 2.12** (Courant-Fisher I). *Let  $A$  be an  $n \times n$  Hermitian matrix and let  $\lambda_1 \geq \dots \geq \lambda_n$  its eigenvalues then*

$$\lambda_k = \max_{\dim(S_k)=k} \min_{x \in S_k} \frac{x^*Ax}{x^*x} \tag{2.14}$$

$$= \max_{\dim(S)=n-k} \min_{x \in S^\perp} \frac{x^*Ax}{x^*x} \tag{2.15}$$

*Proof.* The two representations are equivalent by noticing that  $\dim S^\perp = k$  if and only if  $\dim S = n - k$ . Let  $S_k$  be a  $k$  dimensional subspace of  $\mathbb{C}^n$  and  $V_{n-k+1}$  denote the span of the eigenvectors associated with  $n - k + 1$

smallest eigenvalues; *i.e.*  $V_{n-k+1} = \text{span}\{v_k, \dots, v_n\}$ . Then, their intersection  $V_{n-k+1} \cap S_k$  must be not trivial because  $\dim(V_{n-k+1}) + \dim(S_k) > n$ . Choose a vector  $v \in V_{n-k+1} \cap S_k$ ,  $v$  is a linear combination of the eigenvectors  $v_k, \dots, v_n$ . The minimum of the Rayleigh quotient over  $S_k$  can be bounded by

$$\min_{x \in S_k} \frac{x^* A x}{x^* x} \leq \frac{v^* A v}{v^* v} \leq \lambda_k.$$

This bound is attained by  $S_k = \text{span}\{v_1, \dots, v_k\}$  so:

$$\lambda_k = \max_{\dim(S_k)=k} \min_{x \in S_k} \frac{x^* A x}{x^* x}.$$

□

**Theorem 2.13** (Courant-Fisher II). *Let  $A$  be an  $n \times n$  Hermitian matrix and  $\lambda_1 \geq \dots \geq \lambda_n$  its eigenvalues then*

$$\lambda_k = \min_{\dim(S)=n-k+1} \max_{x \in S} \frac{x^* A x}{x^* x} \tag{2.16}$$

$$= \min_{\dim(S)=k-1} \max_{x \in S^\perp} \frac{x^* A x}{x^* x} \tag{2.17}$$

*Proof.* Let us show the first representation as the second follows immediately, let  $S$  be an  $n - k + 1$  dimensional subspace of  $\mathbb{C}^n$  and  $V_k$  denote the span of the eigenvectors associated with  $k$  biggest eigenvalues; *i.e.*  $V_k = \text{span}\{v_1, \dots, v_k\}$ . Then, their intersection  $V_k \cap S$  must be not trivial because  $\dim(V_k) + \dim(S) > n$ . Choose a vector  $v \in V_k \cap S$ ,  $v$  is a linear combination of the eigenvectors  $v_1, \dots, v_k$ . The maximum of the Rayleigh



quotient over  $S$  can be bounded by

$$\lambda_k \leq \frac{v^* Av}{v^* v} \leq \max_{x \in S} \frac{x^* Ax}{x^* x}.$$

This bound is attained by  $S = \text{span} \{v_k, \dots, v_n\}$  so:

$$\lambda_k = \min_{\dim(S)=n-k+1} \max_{x \in S} \frac{x^* Ax}{x^* x}.$$

□

**Remark 2.14.** If  $B$  is any given Hermitian matrix such that its eigenvalues are ordered as  $\nu_1 \leq \dots \leq \nu_n$ , by considering  $-(\lambda(A)^\downarrow)$  and preserving the indexing we get an increasing ordering of the eigenvalue vector  $\lambda(A)$ , and for example for any  $\lambda(A)^\downarrow$  and any Hermitian matrix  $A$ :

$$-\lambda_k = - \max_{\dim(S_k)=k} \min_{x \in S_k} \frac{x^* Ax}{x^* x} \iff -\lambda_k = \min_{\dim(S_k)=k} \max_{x \in S_k} - \frac{x^* Ax}{x^* x}$$

So without loss of generality we have:  $\nu_k = \min_{\dim(S_k)=k} \max_{x \in S_k} \frac{x^* Bx}{x^* x}$ .

**Theorem 2.15.** Let  $A \in \mathbb{M}_{m,n}(\mathbb{C})$  If  $A = W\Sigma V$  is the singular value decomposition,  $W = \begin{pmatrix} t_1 & \dots & t_m \\ \vdots & \dots & \vdots \end{pmatrix}$  and  $V^* = \begin{pmatrix} s_1 & \dots & s_n \\ \vdots & \dots & \vdots \end{pmatrix}$  we have:

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

for all  $k \leq \min(m, n)$ .

*Proof.* Let  $x$  and  $y$  both be unit vectors,  $x \in \mathbb{C}^m$  and  $y \in \mathbb{C}^n$ ,  $W$  (respectively  $V$ ) is an  $m \times m$  (respectively  $n \times n$ ) unitary matrix, we first expand the quadratic form in terms of the S.V.D of  $A$ :

$$|x^* Ay| = |x^* W \Sigma V y| = \left| \sum_{i=1}^n \sigma_i (x^* t_i) (s_i^* y) \right| \leq \sigma_1.$$

where the last inequality holds because  $x$  and  $y$  are both unit vectors. This bound is attained by  $x = t_1$  and  $y = s_1$  so  $\sigma_1$  is the optimal value of the optimization problem:  $\max_{x,y} \frac{|x^*Ay|}{\|x\|\|y\|}$ . To characterize the smaller singular values, we constrain  $x$  and  $y$  to lie in  $\text{span}\{t_1, \dots, t_{k-1}\}^\perp$  and  $\text{span}\{s_1, \dots, s_{k-1}\}^\perp$  respectively. Then we can ignore the  $k-1$  largest singular values and

$$|x^*Ay| = \left| \sum_{i=k}^n \sigma_i (x^*t_i)(s_i^*y) \right| \leq \sigma_k.$$

this bound is attained by  $x = t_k$  and  $y = s_k$  so  $\sigma_k$  is the optimal value of the optimization problem:

$$\max_{\substack{x \in \text{span}\{t_1, \dots, t_{k-1}\}^\perp \\ y \in \text{span}\{s_1, \dots, s_{k-1}\}^\perp}} \frac{|x^*Ay|}{\|x\|\|y\|}.$$

□

**Corollary 2.16.** *Let  $A, B \in \mathbb{M}_n$  be Hermitian matrices. If  $A \geq B$  then  $\lambda_i(A) \geq \lambda_i(B)$   $i = 1, \dots, n$ .*

*Proof.*  $A \geq B$  if and only if  $x^*Ax \geq x^*Bx, \forall x \in \mathbb{C}^n$ , applying the Courant-Fisher minimax principle to both sides of the inequality completes the proof. □

**Proposition 2.17.** *Let  $A \in \mathbb{M}_n$  be Hermitian and  $B$  positive definite, then all the eigenvalues of  $AB^{-1}$  are real and*

$$\lambda_1(AB^{-1}) = \max_{x \neq 0} \frac{x^*Ax}{x^*Bx} \tag{2.18}$$

$$\lambda_n(AB^{-1}) = \min_{x \neq 0} \frac{x^*Ax}{x^*Bx} \tag{2.19}$$

*Proof.*  $\lambda_i(AB^{-1}) = \lambda_i(AB^{-\frac{1}{2}}B^{-\frac{1}{2}}) = \lambda_i(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})$  where  $B^{-\frac{1}{2}}$  denotes the operator  $(B^{-1})^{\frac{1}{2}}$ . Since  $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$  is a Hermitian matrix  $\lambda_i(AB^{-1})$  are real. Let  $y = B^{\frac{1}{2}}x$  and replacing  $x$  by  $y$  since  $B^{\frac{1}{2}}$  is an invertible matrix we get:

$$\lambda_1(AB^{-1}) = \max_{x \neq 0} \frac{x^* B^{-\frac{1}{2}} A B^{-\frac{1}{2}} x}{x^* x} \quad (2.20)$$

$$\lambda_n(AB^{-1}) = \min_{x \neq 0} \frac{x^* B^{-\frac{1}{2}} A B^{-\frac{1}{2}} x}{x^* x} \quad (2.21)$$

Equivalently:

$$\lambda_1(AB^{-1}) = \max_{y \neq 0} \frac{y^* A y}{y^* B y} \quad (2.22)$$

$$\lambda_n(AB^{-1}) = \min_{y \neq 0} \frac{y^* A y}{y^* B y} \quad (2.23)$$

□

**Corollary 2.18.** *Let  $A$  and  $B$  be two  $n \times n$  Hermitian matrices,  $B > 0$ .*

*Let  $\nu_1 \geq \dots \geq \nu_n$  be the solutions of  $\det(A - \nu B) = 0$ . Then:*

$$\nu_1 = \max_{x \neq 0} \frac{x^* A x}{x^* B x} \quad (2.24)$$

$$\nu_n = \min_{x \neq 0} \frac{x^* A x}{x^* B x} \quad (2.25)$$

*Proof.* Since we have the following equivalence

$$\det(A - \nu B) = 0 \iff \det(AB^{-1} - \nu I) = 0$$

the solutions of  $\det(A - \nu B) = 0$  are the eigenvalues of  $AB^{-1}$ . Applying the above proposition completes the proof. □

## CHAPTER 3

# Matrices and Norms

### 3.1. Definitions

#### Definition 3.1.

- *The spectral radius of a matrix  $A$  is*

$$\vartheta(A) = \max_{j \leq n} |\lambda_j|$$

- *The numerical radius of  $A$  is*

$$s(A) = \sup_{w \in W(A)} |w|$$

**Definition 3.2.** *A norm  $\|\cdot\|$  on a vector space is a function that satisfies*

*:*

- (1)  $\|x\| \geq 0$  with equality if and only if  $x = 0$ ,
- (2)  $\|c \cdot x\| \leq |c| \cdot \|x\|$  for  $c \in \mathbb{C}$ ,

$$(3) \quad \|x + y\| \leq \|x\| + \|y\|.$$

A norm on a vector space is called a vector norm if it satisfies (1)-(3)

A norm on Algebra of matrices is a matrix norm if it satisfies (1)-(3)

and if  $\|AB\| \leq \|A\| \cdot \|B\|, \forall A \in \mathbb{M}_n(\mathbb{C})$ .

**Proposition 3.3** (Spectral norm).

$$\sigma_1(A) = \sigma_{max}(A) = \sqrt{\vartheta(A^*A)} = \|A\|_s = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\| \leq 1} \|Ax\| \quad (3.1)$$

$$= \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=\|y\|=1} |\langle Ax, y \rangle|. \quad (3.2)$$

*Proof.* Using Remark 3.5, the only nontrivial equalities are the third one and the last one. Since  $A^*A$  is Hermitian thus normal:

$$\|A\|^2 = \sup_{v \neq 0} \frac{\langle Av, Av \rangle}{v^*v} = \sup_{v \neq 0} \frac{v^*A^*Av}{v^*v} = \sup_{v \neq 0} R_{A^*A}(v) = s(A^*A) = \vartheta(A^*A).$$

The last equality was proven in Theorem 2.15, here's another proof when  $A$  is a square matrix set:

$$M_A = \sup_{\|x\|=\|y\|=1} |\langle Ax, y \rangle|. \quad (3.3)$$

Suppose  $A$  is invertible, without loss of generality we can take  $y$  in (3.3) to be  $\frac{Ax}{\|Ax\|}$ , in this case  $\|A\| = M_A$ , but invertible matrices are dense in  $\mathbb{M}_n(\mathbb{C})$ , and by C-S inequality,  $M_A$  is a continuous function over  $\mathbb{M}_n(\mathbb{C})$ , thus the two functions are equal for all  $A \in \mathbb{M}_n(\mathbb{C})$ .  $\square$

**Remark 3.4.** If  $A$  is a normal matrix, using the spectral decomposition theorem and the unitary similarity invariance of the numerical range we

see that  $\vartheta(A^*A) = s(A^*A) = (\vartheta(A))^2$ , since  $W(A)$  is the convex hull of  $\rho(A)$ ,  $s(A) = \|A\|$ , thus

$$s(A) = \|A\| = \vartheta(A).$$

**Remark 3.5.** *A proof of*

$$N_A := \sup_{\|x\|=1} |\langle Ax, x \rangle| = \sup_{\|x\|\leq 1} |\langle Ax, x \rangle| \quad (3.4)$$

and

$$E_A := \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|\leq 1} \|Ax\| \quad (3.5)$$

is either directly by the C-S inequality or by noticing that  $N_A$  and  $E_A$  are homogeneous functions over  $\mathbb{R}$  (see A.5) and any vector  $v$  of norm  $\kappa$  less than 1 can be written as  $v = \kappa \frac{x}{\|x\|}$  putting this in (3.4) and (3.5) respectively gives the result.

### 3.2. Numerical Radius

For a set  $S \in \mathbb{C}$ , we denote  $Re(S)$  the set  $\{Re(s); s \in S\}$  i.e. the projection of  $S$  onto the real axis, and by  $Im(S)$  the set  $\{Im(s); s \in S\}$ , i.e. the projection of  $S$  onto the imaginary axis.

**Proposition 3.6.** *Every operator  $A \in \mathbb{M}_n(\mathbb{C})$  can be decomposed as a sum  $A = R(A) + iI(A)$  where  $R(A)$  and  $I(A)$  are self adjoint operators.*

*Proof.* Let  $R(A) = \frac{A + A^*}{2}$  and  $I(A) = \frac{A - A^*}{2i}$ ,  $A = R(A) + iI(A)$  and the proposition follows. □

**Proposition 3.7.** *For all  $A \in \mathbb{M}_n(\mathbb{C})$ :*

$$(1) \quad W(R(A)) := \operatorname{Re}(W(A)).$$

$$(2) \quad W(I(A)) := \operatorname{Im}(W(A)).$$

*Proof.* We calculate:

$$(1) \quad x^* R(A)x = x^* \frac{1}{2}(A + A^*)x = \frac{1}{2}(x^* Ax + x^* A^* x) = \frac{1}{2}(x^* Ax + \overline{x^* Ax}) \\ = \operatorname{Re}(x^* Ax).$$

$$(2) \quad x^* I(A)x = x^* \frac{1}{2i}(A - A^*)x = \frac{1}{2i}(x^* Ax - x^* A^* x) = \frac{1}{2i}(x^* Ax - \overline{x^* Ax}) \\ = \operatorname{Im}(x^* Ax). \quad \square$$

**Definition 3.8.** An operator  $A \in \mathbb{M}_n(\mathbb{C})$  is called accretive if its numerical range is contained in the closed right half-plane.

Equivalently an operator is accretive if and only if  $R(A)$  is positive semi-definite.

**Lemma 3.9.** Let  $A \in \mathbb{M}_n(\mathbb{C})$ , then:  $s(A) \leq \|A\|_s$  and

$$\|A\|_s \leq 2 \cdot s(A). \quad (3.6)$$

*Proof.* The first inequality is a consequence of the C-S inequality. The spectral norm on  $\mathbb{M}_n(\mathbb{C})$  is equal to  $P_A := \sup_{\|x\|=\|y\|=1} |\langle Ax, y \rangle|$ . To prove the inequality we can suppose that

$$\sup_{\|x\|=\|y\|=1} |\langle Ax, y \rangle| = P_A = |G_A(x_0, y_0)| = |\langle Ax_0, y_0 \rangle|.$$

for some  $x_0, y_0 \in \mathcal{H}$  and that  $G_A$  is real, because if not it suffices to multiply

$G_A$  by a complex number of norm 1 to have  $G_A \in \mathbb{R}$  where

$$G_A(x, y) = \underbrace{\langle R(A)x, y \rangle}_{Z(x,y)} + i \underbrace{\langle I(A)x, y \rangle}_C.$$

Since  $G_A$  is a real number,  $C$  will be equal 0 on the diagonal, by (1.4),  $C = 0$ ,  $Z(x_0, y_0)$  is a Hermitian form and from (1.4) we get:

$$4|G_A(x_0, y_0)| = 4|Z(x_0, y_0)| \tag{3.7}$$

$$= \left| [Z(x_0 + y_0, x_0 + y_0) - Z(x_0 - y_0, x_0 - y_0) + iZ(x_0 + iy_0, x_0 + iy_0) - iZ(x_0 - iy_0, x_0 - iy_0)] \right| \tag{3.8}$$

$$\leq s(R(A)) (\|x_0 + y_0\|^2 + \|x_0 - y_0\|^2 + \|x_0 + iy_0\|^2 + \|x_0 - iy_0\|^2) \tag{3.9}$$

$$\leq 2.s(A) (\|x_0\|^2 + \|y_0\|^2 + \|iy_0\|^2 + \|x_0\|^2). \tag{3.10}$$

(see A.8 for the last inequality), which completes the proof.  $\square$

**Proposition 3.10.** *Let  $A \in \mathbb{M}_n(\mathbb{C})$ . The numerical radius  $s(A)$  is a vector norm on  $\mathbb{M}_n(\mathbb{C})$ .*

*Proof.* We need to show that  $s(A)$  satisfies all three conditions:

- Clearly  $s(\lambda.A) = |\lambda|s(A)$ .
- From  $W(A + B) \subset W(A) + W(B)$ , we conclude that  $s(A + B) \leq s(A) + s(B)$ .
- It remains to show that  $s(A) = 0$  if and only if  $A = 0$ , this is a consequence of Lemma 3.9, since  $A = 0$  if and only if  $\|A\|_s = 0$



□

**Remark 3.11.** Since the numerical radius is a vector norm on  $\mathbb{M}_n$ , with  $s(A) \leq \|A\|_s$ , if  $A \in \mathbb{M}_n$ , a finite dimensional space, we know that all norms are equivalent thus there exist a constant  $k$  such that  $\|A\|_s \leq k.s(A)$ , from (3.6)  $k$  is equal 2, taking any two matrices  $A$  and  $B$ , we have  $s(A.B) \leq 2s(A).2s(B)$ , multiplying both sides by 4 we see that the  $4s(A)$  is also a matrix norm.

**Remark 3.12.** We know that  $s(R(A)) \leq s(A)$ , and we can have  $s(A) = s(R(A))$  by taking  $A$  to have nonnegative entries (see [11]), thus if  $A = \begin{pmatrix} 0 & 2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$  from Proposition 1.30  $\|A\| = 2s(A)$ , and 2 is the sharpest constant in (3.6), if  $A \in \mathbb{M}_n$ .

A simple rotation argument yields to:

**Proposition 3.13.** For  $A \in \mathbb{M}_n(\mathbb{C})$ ,  $|\langle Ax, x \rangle| = \sup_{\theta \in \mathbb{R}} \operatorname{Re} \{e^{i\theta} \langle Ax, x \rangle\}$ .

Next we state an important Theorem: The Separating Hyperplane Theorem<sup>1</sup> known as Hahn-Banach convex separation Theorem from Functional Analysis.

**Theorem 3.14** (Convex Separation I). [6] Let  $A$  and  $B$  be two not empty convex disjoint sets of an  $\mathbb{R}$ - normed vector space  $E$ . If  $A$  is an open set then there exist a continuous linear form  $f$  over  $E$  and a real scalar  $\alpha$  such that  $f(x) < \alpha$  for  $x \in A$  and  $f(x) \geq \alpha$  if  $x \in B$ .

**Theorem 3.15** (Convex Separation II). [6] Let  $A$  and  $B$  be two disjoint

---

<sup>1</sup>An Hyperplane is a set  $\{x/f(x) = \alpha$  for some real  $\alpha\}$

non empty convex sets of  $E$ . We suppose that  $A$  is closed and  $B$  is compact, then there exist a continuous linear form  $f \neq 0$  and  $c \in \mathbb{R}$  such that  $f(A) < c$  and  $f(B) > c$ .

What will be useful for us is when  $E$  is the complex plane ( $\equiv \mathbb{R}^2$ ) the linear form is as a matter of fact a line  $\in \mathbb{R}^3$  because  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  but it can be projected to our plane (where the numerical range  $W(A)$  lies). Despite the clear idea in our case ( $\mathbb{R}^2$ ), a formal proof uses a much powerful result from Analysis that is the two precedent Theorems. Of course  $f(0) = 0$ , the convex sets will soon be clarified.

**Theorem 3.16.** *Let  $A \in \mathbb{M}_n(\mathbb{C})$  be given. There exists a real number  $\theta$  such that the Hermitian matrix  $R(e^{i\theta}) = \frac{1}{2}[e^{i\theta}A + e^{-i\theta}A^*]$  is positive definite (respectively positive semi-definite) if and only if  $0 \notin W(A)$  (respectively  $0 \notin \text{rlin}(W(A))$ ) see A.6).*

*Proof.* If  $R(e^{i\theta}A)$  is positive definite for some  $\theta \in \mathbb{R}$ ,  $W(e^{i\theta}A)$  is in the right half plane  $\setminus \{0\}$ , and if  $R(e^{i\theta}A)$  is P.S.D. its numerical range would be a closed interval of the form  $[0, a]$  for some real number  $a$  and  $0$  can't be in the relative interior of  $W(e^{i\theta}A)$ . For the converse suppose  $0 \notin W(A)$ , by Theorem 3.15 there is a line  $L$  in the plane such that each of the non intersecting convex sets  $\{0\}$  (compact) and the closed set  $W(A)$  lies entirely within exactly one of the two open-half planes determined by  $L$ . The coordinate axes may now be rotated so that the line  $L$  is carried into a vertical line in the right half-plane. In the other hand suppose  $0 \notin$

$\text{rlin}(W(A))$  since the relative interior of  $W(A)$  is convex (as an intersection of convex sets (see A.6)) but not an open set in  $\mathbb{R}^3$  we will take without loss of generality the smallest open convex set denoted by  $D$  that contains  $\text{rlin}(W(A))$ , we are in position of applying Theorem 3.14 where the open convex set is  $D$  and  $\{0\}$  is the convex set, similarly the proof follows after rotating the axes and by a continuity argument, that is  $W(e^{i\theta}A) \subset \{Re(z) \geq 0; z \in \mathbb{C}\}$ , so  $R(e^{i\theta}A)$  is positive semi-definite.  $\square$

Here is an illustration of Theorem 3.16 and for construction sake we have chosen  $W(A)$  to be a polygon-convex as one can ask if every convex in  $\mathbb{C}$  is the numerical range of a certain operator  $A$ , speaking of  $A$ , it can be considered a  $5 \times 5$  normal matrix (by Proposition 1.28 ).

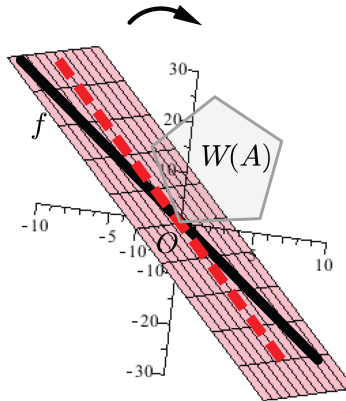


Figure 3.1.: The line  $f$  and its projection  $L$  (in red) separating  $O$  and  $W(A)$  in  $\mathbb{C}$ .

### 3.3. Ky Fan Norms

Let  $\mathbb{M}_{m,n}$  denote the space of  $m \times n$  complex matrices.

**Definition 3.17.** A norm  $\|\cdot\|$  on  $\mathbb{M}_{m,n}$  is unitarily invariant (or symmet-

ric) if  $\|A\| = \|UAV\|$  for all unitary  $U \in \mathbb{M}_m$  and  $V \in \mathbb{M}_n$ , and for all  $A \in \mathbb{M}_{m,n}$ .

If  $\|\cdot\|$  is a unitarily invariant norm, then from the singular value decomposition  $\|A\| = \|W\Sigma(A)V\| = \|\Sigma(A)\|$ , thus  $\|\cdot\|$  is a function of the singular values of  $A$  only. Conversely since  $UAV$  and  $A$  have the same singular values, a norm on  $\mathbb{M}_{m,n}$  which is a function of the singular values, is a unitarily invariant norm.

**Proposition 3.18.** *If  $\|\cdot\|$  is a given unitarily invariant norm on  $\mathbb{M}_{m,n}$ ,  $\mathcal{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$  is the augmented matrix of  $A$  by zero blocks to fill out an  $m \times n$  matrix, where  $1 \leq r \leq m$ ,  $1 \leq s \leq n$ , and  $A \in \mathbb{M}_{r,s}$ , then  $\|A\|_\star \equiv \|\mathcal{A}\|$  is well defined and unitarily invariant.*

*Proof.* The only not trivial point is the unitary invariance of the norm  $\|\cdot\|_\star$ , this is a consequence of the fact that for all  $U \in \mathbb{M}_r$  and  $V \in \mathbb{M}_s$  unitaries

$$\|UAV\|_\star = \|W\mathcal{A}P\| = \|\mathcal{A}\| = \|A\|_\star$$

where,  $W = \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix}$  and  $P = \begin{pmatrix} V & 0 \\ 0 & I \end{pmatrix}$  are two unitary matrices.  $\square$

Let  $M$  be a  $n \times m$  matrix and let  $\sigma^\downarrow(M) = (\sigma_1(M), \dots, \sigma_{\text{rank}(M)}(M))$  denote the non-zero singular values of  $M$  in decreasing order.

**Definition 3.19.** *For  $k \in \mathbb{N}$  and  $1 \leq p < \infty$ , the  $(k, p)$  Ky Fan norm is:*

$$\|M\|_{(k,p)} := \left( \sum_{k'=1}^k \sigma_{k'}(M)^p \right)^{\frac{1}{p}}.$$

*If  $p = 1$ , the Ky Fan  $k$ -norms  $\|\cdot\|_k, k = 1, \dots, n$  are:  $\sum_{i=1}^k \sigma_i$ , where*

$\sigma(M) = 0$  for  $k > \text{rank}(M)$ . The  $(\text{rank}(M), p)$  Ky Fan norm denoted  $\|\cdot\|_{(p)}$  is called the Schatten  $p$ -norm.

**Theorem 3.20.** [4] For all  $k \in \mathbb{N}$  and  $1 \leq p < \infty$ ,  $\|\cdot\|_{(k,p)}$  is a unitarily invariant vector norm on  $\mathbb{M}_{n,m}$ . If  $n = m$ , it is also a matrix norm.

For  $k = 1$  the Ky Fan norm becomes the *spectral norm*. For  $p = 2$ ,  $k = \text{rank}(M)$  we obtain the *Frobenius norm*, and for  $p = 1$ ,  $k = \text{rank}(M)$  the *Trace norm*. The *Frobenius norm* can also be defined as the one associated to the usual complex scalar product on  $\mathbb{M}_{m,n}(\mathbb{C})$ : that is for all  $A$  and  $B \in \mathbb{M}_{m,n}$ ;  $(A, B) \rightarrow \langle A, B \rangle = \text{tr}(B^*A)$ . If  $M = (m_{ij})$ ,  $\|M\|_{(2)} = \sqrt{\text{tr}(M^*M)}$ , and it is easy to verify that  $\|M\|_{(2)} = \sqrt{\sum_{i,j} (m_{ij})^2}$ .

**Remark 3.21.** The *Frobenius norm*  $:= \|\cdot\|_{(2)}$  is the one associated to a scalar product, thus it satisfies the C-S inequality, if  $B \neq \alpha A$ ,  $\alpha \geq 0$ , then  $\|A + B\|_{(2)} < \|A\|_{(2)} + \|B\|_{(2)}$ .

The dual norm of a norm  $N := \|\cdot\|_n$  (for an explanation on why we call it the dual norm see Riesz representation theorem in [6]) denoted by  $\|\cdot\|_n^D$  is defined as  $A \in \mathbb{M}_{m,n} \rightarrow \sup_{\|X\|_n \leq 1} |\langle A, X \rangle|$  and the dual-dual norm  $\|\cdot\|_n^{DD}$  is by definition:  $X \in \mathbb{M}_{n,m} \rightarrow \sup_{\|A\|_n^D \leq 1} |\langle A, X \rangle|$ , when  $N$  is the *spectral norm*  $\|\cdot\|_s^D$  is the *trace norm*. One can see that for every norm  $\|\cdot\|_n$ ,  $\|X\|_n \geq \|X\|_n^{DD}$  the other inequality can also be verified as shown in Theorem (5.5.14) in [3] where  $X \in E := \mathbb{C}^n$  but it will still be true for any  $\mathbb{C}$ -vector space of finite dimension by isomorphism. Worth mentioning

that when we have  $E$  an  $\mathbb{R}$ - vector space, also  $\|X\| = \|X\|^{D^D}$  (see [6]).

**Lemma 3.22.** *Let  $\|\cdot\|$  be a given norm on  $\mathbb{M}_{m,n}$  and let  $\|\cdot\|^D$  denote its dual norm with respect to the Frobenius inner product. Then  $\|\cdot\|$  is unitarily invariant if and only if  $\|\cdot\|^D$  is unitarily invariant.*

*Proof.* First notice that  $|\operatorname{tr}(AC^*)| = |\operatorname{tr}(C^*A)|$ , thus using the cyclicity of the trace we have:

$$\|UAV\|^D = \max \{ |\operatorname{tr}(UAV)C^*| : C \in \mathbb{M}_{m,n}, \|C\| = 1 \} \quad (3.11)$$

$$= \max \{ |\operatorname{tr}(A[U^*CV^*]^*)| : C \in \mathbb{M}_{m,n}, \|C\| = 1 \} \quad (3.12)$$

$$= \max \{ |\operatorname{tr}(AE^*)| : C \in \mathbb{M}_{m,n}, \|UEV\| = 1 \} \quad (3.13)$$

$$= \max \{ |\operatorname{tr}(AE^*)| : C \in \mathbb{M}_{m,n}, \|E\| = 1 \} \quad (3.14)$$

$$= \|A\|^D \quad (3.15)$$

The converse follows from the preceding argument and the fact that:

$$\|A\|^{D^D} = \|A\|. \quad \square$$

**Definition 3.23.** *We write  $\mathbb{R}_+^{n\downarrow} = \{x \in \mathbb{R}^n, x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$ . For  $X \in \mathbb{M}_{m,n}$  and  $\alpha \in \mathbb{R}_+^{q\downarrow}$ ,  $q = \min\{m, n\}$  we write:  $\|X\|_\alpha = \alpha_1\sigma_1 \dots + \alpha_q\sigma_q$ .*

**Theorem 3.24.** *Let  $m, n$  be given positive integers and let  $q = \min\{m, n\}$ .*

(a) *For each given nonzero  $\alpha \in \mathbb{R}_+^{q\downarrow}$ ,  $\|A\|_\alpha \equiv \alpha_1\sigma_1(A) + \dots + \alpha_q\sigma_q(A)$  is a unitarily invariant norm on  $\mathbb{M}_{m,n}$ .*

(b) *For each given unitarily invariant norm  $\|\cdot\|$  on  $\mathbb{M}_{m,n}$ :*

$$\|A\| = \max \{ \|A\|_\alpha : \alpha \in N_{\|\cdot\|} \}$$

for all  $A \in \mathbb{M}_{m,n}$ , where  $N_{\|\cdot\|}$  is the following compact set  $\{\sigma(X) : X \in \mathbb{M}_{m,n}, \|X\|^D = 1\}$ .

*Proof.* We will show the first part, interested reader can see Theorem 3.5.5 in [4] for a proof of (b). We have:

$$\|A\|_\alpha = \sum_{i=1}^q \alpha_i \sigma_i = \sum_{i=1}^{q-1} (\alpha_i - \alpha_{i+1}) \|A\|_i + \alpha_q \|A\|_q \quad (3.16)$$

which is a nonnegative linear combination of Ky Fan  $k$ -norms  $\|\cdot\|_k$  and the result follows.  $\square$

**Corollary 3.25.** *Let  $A, B \in \mathbb{M}_{m,n}$  be given and let  $q = \min\{m, n\}$ . The following are equivalent:*

- (i)  $\|A\| \leq \|B\|$  for every unitariouy invariant norm  $\|\cdot\|$  on  $\mathbb{M}_{m,n}$ .
- (ii)  $\|A\|_k \leq \|B\|_k$  for  $k = 1, \dots, q$ .
- (iii)  $\|A\|_\alpha \leq \|B\|_\alpha$  for all  $\alpha \in \mathbb{R}_+^{q \downarrow}$ .

*Proof.* Since each Ky Fan  $k$ -norm  $\|\cdot\|_k$  is unitarily invariant, (i) implies (ii). (3.16) shows that (ii) implies (iii). Finally the (b) part of Theorem 3.24 shows that (iii) implies (i).  $\square$

The same proof is valid if the large inequalities in Corollary 3.25 are replaced by strict inequalities.

## REFERENCES

- [1] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
- [2] Emmanuel Fricain, *Analyse Fonctionnelle et Théorie des Opérateurs, Cours et Exercices*, 2009-2010.
- [3] R. A. Horn and C. R Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
- [4] R. A. Horn and C. R Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.
- [5] S. K. Mitra, P. Bhimasankaram, S. B. Malik, *Matrix Partial Orders, Shorted Operators And Applications*, Series in Algebra Vol. 10. 2013
- [6] D. Robert, *Analyse Fonctionnelle*, Nantes-2005.
- [7] M. Schechter, *Principles of Functional Analysis*. New York-London Academic Press , 1971.
- [8] C. Zălinescu, *Convex Analysis In General Vector Spaces*, (2002).
- [9] F.Zhang, *Matrix Theory Basic Results and Techniques*, 2<sup>nd</sup> Edition (Universitext) Springer, 2013.
- [10] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, Part I: Lectures 4-5, SIAM, null edition, 1997.
- [11] M.Godberg, et al *Numerical Radius of Positive Matrices*, Linear Algebra and its Applications 12, pp. 209–214, 1975.
- [12] Chi-K. Li, (Communicated by Palle E. T. Jorgensen), *A simple Proof*



*Of The Numerical Range Theorem*, Proceedings Of The American  
Mathematical Society, Volume 124, Number 7, July 1996.

## APPENDIX A

# The Appendix

**Theorem A.1** (Spherical Coordinate System in  $\mathbb{R}^n$ ).

$$\left\{ \begin{array}{l} x_1 = r \cos \theta_1 \\ x_2 = r \sin \theta_1 \cos \theta_2 \\ \vdots \\ x_{n-1} = r \sin \theta_1 \cdots \sin \theta_{n-2} \cos \theta_{n-1} \\ x_n = r \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{n-1} \end{array} \right.$$

Consequently if  $r = \|x\| = 1$ ,  $x \in \mathbb{S}^3$  we have this parametrization:

$$\left\{ \begin{array}{l} x_1 = \cos \theta_1 \\ x_2 = \sin \theta_1 \cos \theta_2 \\ x_3 = \sin \theta_1 \sin \theta_2 \cos \theta_3 \\ x_4 = \sin \theta_1 \sin \theta_2 \sin \theta_3 \end{array} \right.$$

**Theorem A.2.** Let  $(X, \mathcal{M}, \mu)$  be a measurable space where  $\mu$  is a positive

measure,  $f$  and  $g: X \rightarrow \mathbb{C}$  measurable functions,  $1 < p < \infty$  and  $p'$  it's conjugate exponent we have:

$$(1) \int_X |f \cdot g| d\mu \leq \left( \int_X |f|^p \right)^{\frac{1}{p}} \left( \int_X |g|^{p'} \right)^{\frac{1}{p'}} \quad (\text{Hölder Inequality}),$$

$$(2) \left( \int_X |f + g|^p d\mu \right)^{\frac{1}{p}} \leq \left( \int_X |f|^p \right)^{\frac{1}{p}} + \left( \int_X |g|^p \right)^{\frac{1}{p}} \quad (\text{Minkowski Inequality})$$

Taking  $\mu$  to be the counting measure we have the celebrated results:

**Corollary A.3.**

$$(1) \sum_{i=1}^n |x_i \cdot y_i| \leq \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n |y_i|^{p'} \right)^{\frac{1}{p'}},$$

$$(2) \left( \sum_{i=1}^n |x_i + y_i|^p \right)^{\frac{1}{p}} \leq \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left( \sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}}.$$

**Definition A.4.** Let  $V$  be a complex vector space. A map  $b: V \times V \rightarrow \mathbb{C}$  is a sesquilinear form if for all  $x, y, z, t \in V$  and all  $a, c \in \mathbb{C}$  we have:

- $b(x + y, z + t) = b(x, z) + b(x, t) + b(y, z) + b(y, t)$ ,
- $b(ax, cy) = \bar{c} a \cdot b(x, y)$ .

If a sesquilinear form satisfies  $b(u, v) = \overline{b(v, u)}$ ,  $\forall u$  and  $v$  we call it a Hermitian form.

A complex scalar product denoted by  $\langle x, y \rangle$  is a Hermitian form that satisfies  $\langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0$  if and only if  $x = 0$ . The application  $|\cdot| : v \rightarrow |v| = \sqrt{\langle v, v \rangle}$  is a norm on  $V$  referred as the norm associated to

the scalar product.

**Definition A.5.** Let  $V$  and  $W$  be two vector spaces over a field  $F$ . A function  $f : V \rightarrow W$  is said to be homogeneous of degree  $k$  over  $F$ , if

$$f(\alpha.v) = \alpha^k.f(v).$$

for all nonzero  $\alpha \in F$  and  $v \in V$ .

**Definition A.6.** [8]

(1) Let  $E$  be a topological space. The affine hull of a set  $S \in E$  is the smallest affine set containing  $S$  denoted by  $Af(S)$ . We have

$$Af(S) = \left\{ \sum_{i=1}^k \alpha_i x_i / k > 0, x_i \in S, \alpha_i \in \mathbb{R}, \sum_{i=1}^k \alpha_i = 1 \right\}.$$

If  $E := \mathbb{R}^n$  an affine set can be defined as the translation of a vector subspace.

(2) The relative interior of a set  $S$  in  $E$  denoted by  $rlin(S)$  is defined as its interior (with respect to the topology equipped) within the affine hull of  $S$ .

$$\{x \in S : \exists \mathcal{V}_x, Af(S) \cap \mathcal{V}_x \subseteq S\}.$$

If  $E$  is a metric space, the open neighborhood  $\mathcal{V}_x$  is a certain ball of a small enough radius, centered on  $x$ .

**Remark A.7.** Since the scalar multiplication and the sum operation are continuous operations on any topological vector space, for any nonempty

convex set  $C$ ,  $\text{rlin}(C)$  can be defined as :

$$\{x \in C : \forall y \in C, \exists \lambda > 1 : \lambda x + (1 - \lambda)y \in C\}.$$

**Theorem A.8.** For all vectors  $x$  and  $y$  in an inner product  $\mathbb{C}$  (or  $\mathbb{R}$ )-Vector space  $E$ , we have:

- $|\langle x, y \rangle|^2 \leq |\langle x, x \rangle| |\langle y, y \rangle|$  Cauchy-Schwartz Inequality (C-S). The two sides are equal if and only if  $\exists \alpha \in \mathbb{C}$  such that  $x = \alpha y$ .
- $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$  Parallelogram Identity.

