MATRICIAL AND VECTORIAL NORMS

by

Gurdeep Singh Kahlon

A Thesis Submitted in Partial Fulfilment of The Requirements for the Degree of Master of Arts in the Department of Mathematics

We accept this thesis as conforming to the required standard

The University of British Columbia, July, 1972
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at this University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Mathematics
The University of British Columbia
Vancouver 8, Canada

Date 5/4/72
Supervisor:  Dr. R. Westwick

Abstract

Matricial norms, minimal matricial norms, vectorial norms and vectorial norms subordinate to matricial norms, which are respectively generalizations of matrix norms, minimal matrix norms, vector norms and vectorial norms subordinate to matrix norms, are defined and their various applications and properties are discussed.
ACKNOWLEDGEMENTS

I sincerely thank Dr. R. Westwick for his comments and suggestions, and Dr. R. Ree for reading this thesis.

I also thank the Mathematics Department, University of British Columbia, for giving me financial assistance.
Dedicated

to

my parents
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Introduction</th>
<th>(vi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Matricial norms</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Zeroes of polynomials</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Vectorial norms</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
</tr>
</tbody>
</table>
Introduction

The concept of matricial norm was mentioned by Wielandt as a suggestion and later studied by Bauer and Robert, and recently by Deutsch [1,2]. Certain types of mappings which satisfy matricial norm axioms were also studied by Ostrowski [12] and Robert [14].

The advantage of introducing matricial norm is that it gives us a better upper bound for the spectral radius of the companion matrix of \( f(z)=a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n, a_i \in \mathbb{C} \). In fact, with the application of matricial norm, upper bounds have been found for the spectral radius which are better than the bounds due to Cauchy [11], Wilf [17], Kojima [11] and Fujivara [11].

Vectorial norm was introduced by Kautorovitch [10] and later studied by Robert [13,14], Stoer, and recently by Deutsch [3,4]. Other norms such as matricial norm subordinate to vectorial norms and matrix norm subordinate to vector norm are discussed in Chapter 3. Various properties such as regularity and duality of different norms are also discussed in Chapter 3.

In this thesis many of the propositions and results left unproved in the recent papers where they appear are proved here and examples are worked out in detail. Many examples are given
and to make this paper more complete, examples, definitions, and results have been added wherever necessary. In fact, every effort has been made to make things complete and explicit.
Chapter I

Matricial norms

In this chapter the definition of a matricial norm introduced by Deutsch [1] is given. Various properties of minimal matricial norms and generalized matricial norms are discussed, and a proof for the existence of a minimal matricial norm is included. Theorem 2 generalizes a result of Householder [8]. Since zeroes of \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + a_n z^n \), \( a_i \in \mathbb{C} \) are eigenvalues of its companion matrix so different bounds for the spectral radius are obtained. Enough examples are given to see that these bounds for the spectral radius are better than the bounds due to Kojima [11], Cauchy [11], Fujwara [11] and Wilf [17].

Let \( M_n \) denote the set of all \( n \times n \) matrices over the complex field \( \mathbb{C} \). We know that \( M_n \) is a non-commutative complex algebra with identity. Also, let \( \mathbb{R}^+ \) be the set of all non-negative real numbers and \( M_n^+ \) be the set of all \( n \times n \) matrices with entries from \( \mathbb{R}^+ \). Then it is easy to see that \( M_n^+ \) is a partially ordered set elementwise i.e., For \( A = (a_{ij}) \in M_n^+ \), \( B = (b_{ij}) \in M_n^+ \), define \( (a_{ij}) \leq (b_{ij}) \) if \( a_{ij} \leq b_{ij} \).

A matrix norm is a mapping from \( M_n \to \mathbb{R}^+ \) such that for all \( A, B \in M_n \) and \( \alpha \in \mathbb{C} \), we have
i) \( A \neq 0 \Rightarrow \|A\| \neq 0 \)

ii) \( \|\alpha A\| = |\alpha| \cdot \|A\| \)

iii) \( \|A + B\| \leq \|A\| + \|B\| \)

iv) \( \|AB\| \leq \|A\| \cdot \|B\| \)

A mapping which satisfies i), ii), iii) but not iv) is called a generalized matrix norm. Our object in this chapter is to generalize matrix norm to the matricial norm and see the applications of the latter. Define the matrix norms \( \psi_{\text{row}}, \psi_{\text{col}}, \) and \( \psi_{\text{euc.}} \) on \( M_n \) by

\[
\psi_{\text{row}}(A) = \max_{i} \sum_{j=1}^{n} |a_{ij}|
\]

\[
\psi_{\text{col}}(A) = \max_{j} \sum_{i=1}^{n} |a_{ij}|
\]

\[
\psi_{\text{euc.}}(A) = (\sum_{i,j=1}^{n} |a_{ij}|^2)^{1/2}
\]

That these are actually matrix norms is easy to verify. We call these norms, respectively, row norm, column norm, and euclidian norm.

**Definition:** A matricial norm is a mapping \( \mu : M_n \to M_k^+ \) which satisfies the following axioms

(M·1) \( A \neq 0 \Rightarrow \mu(A) \neq 0 \)

(M·2) \( \mu(\alpha A) = |\alpha| \mu(A) \)

(M·3) \( \mu(A + B) \leq \mu(A) \mu(B) \quad \forall A, B \in M_n \)

(M·4) \( \mu(\alpha B) \leq \mu(A) \mu(B) \quad \alpha \in C. \)
k is called the order of the matricial norm $\mu$.

Remark: If $k=1$ Then $\mu$ is matrix norm on $M_n$.

Remark: From (M·3) and (M·4) one gets by induction

$$
\mu(A_1 + A_2 + \ldots + A_n) \leq \mu(A_1) + \mu(A_2) + \ldots + \mu(A_n)
$$

$$
\mu(A_1 A_2 \ldots A_n) \leq \mu(A_1) \mu(A_2) \ldots \mu(A_n)
$$

In particular $\mu(nA) \leq n\mu(A)$ and $\mu(A^n) \leq \mu(A)^n$ $\forall A \in M_n$.

A mapping $\mu : M_n \rightarrow M_{k^+}$ which satisfies (M·1), (M·2), and (M·3) is called a generalized matricial norm on $M_n$. Certain mappings which satisfy (M·2) and (M·3) are called pseudo-norms on $M_n$.

Example 1. Define $\mu : M_n \rightarrow M_k^+$ such that

$$
\mu(A) = (|a_{ij}|), A = (a_{ij}) \in M_n. \text{ Then } \mu \text{ is a matricial norm of order } n \text{ on } M_n.
$$

(M·1) and (M·2) are obvious. To see (M·3) and (M·4) let $A = (a_{ij}), B = (b_{ij}), i, j = 1, 2, \ldots, n$. Then $A+B = (a_{ij} + b_{ij})$

$$
\Rightarrow \mu(A+B) = (|a_{ij} + b_{ij}|) \leq (|a_{ij}| + |b_{ij}|) = (|a_{ij}|) + (|b_{ij}|) = \mu(A) + \mu(B) \text{ and } AB = (C_{ij}) \text{ where } C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \text{ Therefore } \mu(AB) = (|C_{ij}|) \leq \sum_{n=1}^{\infty} |a_{ik}| |b_{kj}| = \mu(A) \mu(B).
$$

Example 2. Given a matrix norm $\psi$ on $M_n$ and a $k \times k$ matrix $A$ such that $\lambda^2 \geq A > 0$, define $\mu : M_n \rightarrow M_k^+$ such that $\mu(A) = \psi(A)A$ $\forall A \in M_n$. Then $\mu$ is a matricial norm on $M_n$ of order $k$. 
Again (M.1) and (M.2) are obvious. To check (M.3) and (M.4) we have

\[ u(A+B) = \psi(A+B) \Delta \leq (\psi(A) + \psi(B)) \Delta = \psi(A) + \psi(B) \]

Therefore \( u(AB) \leq u(A) \mu(B) \) \( \forall A, B \in M_n \).

**Proposition 1.** Let \( \mu : M_n \rightarrow M^+ \) be a matricial norm. Then \( \mu^t : M_n \rightarrow M^k \) such that \( \mu^t(A) = (\mu(A^t))^t, A \in M_n \) is also a matricial norm on \( M_n \). (\( A^t \) denotes transpose of matrix A).

**Proof.** To check (M.1) and (M.2) we note that \( A \neq 0 \Rightarrow A^t \neq 0 \) \( \Rightarrow \mu(A^t) \neq 0 \) since \( \mu \) is a matricial norm. Therefore \( \mu^t(A) \neq 0 \), and \( \mu^t(aA) = (\mu((aA)^t))^t = (\mu(aA^t))^t = |a|\mu(A^t)^t = |a|\mu^t(A) \). For (M.3) we have \( \mu^t(A+B) = (\mu((A+B)^t))^t = (\mu(A^t+B^t))^t \leq (\mu(A^t) + \mu(B^t))^t = (\mu(A))^t + (\mu(B))^t = \mu^t(A) + \mu^t(B) \).

To see (M.4), \( \mu^t(AB) = (\mu(AB)^t)^t = (\mu(B^tA^t))^t \leq (\mu(A^t)\mu(B^t))^t = (\mu(A))^t\mu(B)^t = \mu^t(A)\mu^t(B) \) and this completes the proof.

**Proposition 2.** Let \( \mu \) be a matricial norm on \( M_n \) and let \( G \) be an invertible \( n \times n \) matrix. Then the function \( \mu_G : M_n \rightarrow M^+_k \) such that \( \mu_G(A) = \mu(GAG^{-1}) \) \( \forall A \in M_n \) is a matricial norm.

**Remark:** \( \mu_G \) is called \( G \)-transform of \( \mu \).

**Proof:** Let \( A \neq 0 \). Then \( GAG^{-1} \neq 0 \) and therefore \( \mu(GAG^{-1}) \neq 0 \) since \( \mu \) is a matricial norm. Thus \( \mu_G(A) = \mu(GAG^{-1}) \neq 0 \).

Also \( \mu_G(aA) = \mu((GaA)^{-1}) = |a|\mu(GA^{-1}) = |a|\mu_G(A) \). Further, for \( A, B \in M_n \) we have

\[ \mu_G(A+B) = \mu(G(A+B)G^{-1}) = \mu(GA^{-1} + GBG^{-1}) \leq \mu(GA^{-1}) + \mu(GB^{-1}) = \mu_G(A) + \mu_G(B) \] and \( \mu_G(AB) = \mu(GABG^{-1}) = \mu(GA^{-1}GB^{-1}) \leq \mu(GA^{-1}) \mu(GB^{-1}) = \mu_G(A)\mu_G(B) \) which proves
Remark: Proposition 2 shows that similar matrices have the same matricial norm.

The next proposition shows that given a direct-sum decomposition of $C^n = X_1 \oplus \cdots \oplus X_k$ of $C^n$ (vector space of $n$-tuples over $C$) and a matrix norm on $M_n$, we can generate a matricial norm of order $k$ on $M_n$.

**Proposition 3:** Let $E_1, E_2, \ldots, E_k$ be the projections associated with the direct-sum decomposition $C^n = X_1 \oplus \cdots \oplus X_k$ and let $\psi$ be a matrix norm on $M_n$. Then the mapping $\mu : M_n \rightarrow M^*_k$ such that

$$\mu(A) = (\psi(E_i A E_j))_{i,j=1,2,\ldots,k} \quad A \in M_n$$

is a matricial norm of order $k$ on $M_n$.

Remark: The function $\mu$ generated in this way is called the matricial norm induced by the direct-sum decomposition $C^n = X_1 \oplus \cdots \oplus X_k$ and the matrix norm $\psi$.

**Proof:** To prove (M.1), suppose $\mu(A) = 0$. Then $\psi(E_i A E_j) = 0$ for all $i,j = 1,2,\ldots,k$ and therefore $E_i A E_j = 0$. Now since $A = \sum_{i,j} E_i A E_j$ we have $A = 0$ and so $\mu$ satisfies (M.1). It is easy to see (M.2).

To check (M.3) we have for all $A,B \in M_n$

$$\mu(A+B) = (\psi(E_i (A+B) E_j))_{i,j=1,2,\ldots,k} = \psi(E_i A E_j + E_i B E_j)_{i,j=1,2,\ldots,k} \leq (\psi(E_i A E_j) + \psi(E_i B E_j))_{i,j=1,2,\ldots,k} = (\psi(E_i A E_j) + \psi(E_i B E_j))_{i,j=1,2,\ldots,k} = \psi(A) + \psi(B).$$
Since $E_1, E_2, \ldots, E_k$ are associated projections,
\[ E_i^2 = E_i, E_i E_j = E_j E_i = 0 \text{ and } \sum_{i=1}^{k} E_i = 1. \]
Let $\mu_{ij}(AB)$ denote the absolute value of $(i,j)$ element of matrix $AB$. Then
\[ \mu_{ij}(AB) = \mu_{ij}(A E_i B) \leq \sum_{a=1}^{k} \mu_{ij}(A E_a B) \ldots (1). \]

Now consider the element $\mu_{ij}(A E_a B) = \psi(E_i A E_a B_j) = \psi(E_i A E_a B_j) 
\leq \psi(E_i A E_a) \psi(E_a B_j) = \mu_{ia}(A) \mu_{aj}(B).$

Hence from (1) we have
\[ \mu_{ij}(AB) \leq \sum_{a=1}^{k} \mu_{ia}(A) \mu_{aj}(B) \text{ and so } \mu(AB) \leq \mu(A) \mu(B) \]
and this completes the proof.

**Proposition 4:** Let $\mu : M_n \rightarrow M_k^+$ be a generalized matricial norm
and let $\{A_m\}$ be a sequence of complex $n \times n$ matrices. Then
\[ \lim_{m \to \infty} A_m = 0 \iff \lim_{m \to \infty} \mu(A_m) = 0. \]

**Proof:** Let $A_m = \sum_{i,j} \alpha_{ij} E_{ij}$ where $\{E_{ij}\}$ is a standard bases for $M_n$. Since $\mu_{ij}$ is a pseudo-norm on $M_n$, therefore $\mu_{ij}(\sum_{i,j} \alpha_{ij} E_{ij}) 
\leq \sum_{i,j} |\alpha_{ij}| \mu_{ij}(E_{ij})$. Now if $A_m \to 0$ then $\mu_{ij}(A_m) \to 0$ and so
\[ \mu(A_m) \to 0. \]

Conversely suppose $\mu(A_m) \to 0$. If $E_i = \begin{bmatrix} \alpha_{ij} & 0 \\ 0 & 0 \end{bmatrix}$, then
\[ \mu(E_i A E_{ij}) = \mu \begin{bmatrix} 0 & 0 \\ \alpha_{ij} & 0 \end{bmatrix} = \mu(\alpha_{ij} E_{ij}) = |\alpha_{ij}| \mu(E_{ij}) \text{ and } \mu(E_i A E_{ij}) = \mu(E_i) \mu(A) \mu(E_{ij}). \]
Therefore $\mu(E_i)\mu(A)\mu(E_j) = |\alpha_{ij}|\mu(E_{ij})$. Hence $\mu(A^n) \to 0$ and $A_m \to 0$ follows.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of the matrix $A$. Define the spectral radius of matrix $A$ as the maximum of the absolute value of the eigenvalues of $A$, i.e.,

$$r(A) = \text{spectral radius of } A = \max_i |\lambda_i|.$$ 

**Theorem 1** Let $\mu: M_n \to M_k^+$ be a matricial norm. Then $r(A) \leq r(\mu(A))$ for all $A \in M_n$.

**Proof:** Note that if $r(A) \geq 1$ then $A$ does not tend to the zero matrix as $n$ increases while if $r(A) < 1$ then $A^n \to 0$ as $n \to \infty$.

Hence if $r(\mu(A)) < 1$ then $\mu(A)^n \to 0$ and so $\mu(A^n) \to 0$. Therefore by proposition 4, $A^n \to 0 \Rightarrow r(A^n) < 1$. Thus $r(\mu(A)) < 1 \Rightarrow r(A) < 1$. On the other hand if $r(\mu(A^m)) < 1$ then also we have $r(A) < 1$.

Let $\epsilon > 0$. Write $B = A/ r(\mu(A)) + \epsilon$. Then $\mu(B) = \mu\{A(\mu(A)) + \epsilon\}^{-1} = (\mu(A)) + \epsilon^{-1}\mu(A)$.

Therefore $r(\mu(B)) = (r(\mu(A)) + \epsilon)^{-1}r(\mu(A)) = \frac{r(\mu(A))}{r(\mu(A)) + \epsilon}$.
< 1 since ε > 0. Hence by the above argument we have \( r(B) < 1 \).

So \( \frac{r(A)}{r(u(A))} + \epsilon < 1 \Rightarrow r(A) < r(u(A)) + \epsilon \). Since \( \epsilon \) is arbitrary so \( r(A) \leq r(u(A)) \).

The next proposition gives us a better bound for the spectral radius than given by Theorem 1. We have seen that for all \( A \in M_n \), \( r(A) \leq r(u(A)) \). Therefore \( r(A)^p = r(A^p) \leq r(u(A^p)) \) and we have \( r(A) \leq \frac{1}{p} r(u(A^p)) \).

On the other hand, \( \frac{1}{p} r(u(A^p)) \leq r(u(A)) \Rightarrow r(u(A)) \leq r(u(A^p)) \Rightarrow r(A) \leq \frac{1}{p} r(u(A^p)) \leq r(u(A)) \). Hence \( r(A) \leq \frac{1}{p} r(u(A)) \). Thus we have proved

**Proposition 5.** If \( u: M_n \to M_k^+ \) is a matricial norm, then \( r(A) \leq \frac{1}{p} r(u(A)) \) for all \( A \in M_n^+ \) and \( p \) any positive integer.

**Example 3**

Let

\[
B = \begin{bmatrix}
2 & 0 & -2 \\
-1 & -1 & 0 \\
0 & 2 & 2
\end{bmatrix}
\]

The eigenvalues of matrix \( B \) are \( \lambda = 0, 0, 3 \) and therefore \( r(B) = 3 \).

Now consider the matricial norm \( u: M_3 \to M_2^+ \) defined by

\[
u(A) = \begin{cases}
|a_{11}| & |a_{12}| + |a_{13}| \\
\max(|a_{21}|, |a_{31}|) & \max(|a_{22}| + |a_{23}|, |a_{32}| + |a_{33}|)
\end{cases}
\]

for \( A = (a_{ij}) \in M_3^+ \).
Therefore \( u(B) = \begin{pmatrix} 2 & 2 \\ 1 & 4 \end{pmatrix} \) and \( r(u(B)) = 4.73 \).

Again \( B^2 = \begin{pmatrix} 4 & -4 & -8 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{pmatrix} \), \( u(B^2) = \begin{pmatrix} 4 & 12 \\ 2 & 6 \end{pmatrix} \), \( r(u(B^2)) = 10 \)

\( \begin{pmatrix} 12 & -12 & -24 \\ -3 & 3 & 6 \\ -6 & 6 & 12 \end{pmatrix} \), \( u(B^3) = \begin{pmatrix} 12 & 36 \\ 6 & 18 \end{pmatrix} \), \( r(u(B^3)) = 30 \)

Therefore \( \frac{1}{2} [r(u(B^2))]^2 = 3.16 \) and \( \frac{1}{3} [r(u(B^3))]^3 = 3.11 \).

Similarly \( \frac{1}{4} [r(u(B^4))]^4 = 3.08 \). This agrees with the conclusions of Proposition 5. We will see in the next Proposition that as \( n \to +\infty \), \( \frac{1}{n} [r(u(A^n))]^n \to r(A) \). It is also interesting to verify that \( \psi_{\text{row}}(B) = 4 \), \( \frac{1}{2} [\psi_{\text{row}}(B^2)]^2 = 4 \), \( \frac{1}{3} [\psi_{\text{row}}(B^3)]^3 = 3.63 \)

and \( \frac{1}{4} [\psi_{\text{row}}(B^4)]^4 = 3.46 \). Thus we see that in the above cases \( r(u(B)) \) is the weakest upperbound for \( r(B) \). The following Lemma is needed for Proposition 6 and Theorem 2. Its proof can be found in [5, P.113].

**Lemma:** Spectral radius of a matrix cannot exceed the value of any of its norms.
Proposition 6: Let $\nu: M_n \to M_k^+$ be a matricial norm. Then for all $A \in M_n$, $r(A) = \lim_{m \to \infty} \frac{1}{m} [r(\nu(A^m))]^\frac{1}{m}$.

Proof: Let $\psi: M_k^+ \to R^+$ be the euclidean matrix norm on $M_k$.

Define $\mu_0: M_n \to R^+$ such that

$$\mu_0(A) = \psi(\nu(A)), \quad \forall A \in M_n.$$ 

Then it is routine to verify that $\mu_0$ is a matrix norm on $M_n$. Now by the above Lemma $r(\nu(A^m)) \leq \mu_0(A^m)$ from which we have

$$[r(\nu(A^m))]^\frac{1}{m} \leq [\mu_0(A^m)]^\frac{1}{m} \leq [r(A^m)]^\frac{1}{m}.$$ 

Therefore by Proposition 5, $r(A) \leq [r(\nu(A^m))]^\frac{1}{m} \leq [\mu_0(A^m)]^\frac{1}{m}$.

Now in general if $|| \cdot ||$ is matrix norm on $M_n$, then [15,P.312]

$$r(A) = \lim_{n \to \infty} ||A^n||^\frac{1}{n}.$$ 

Therefore $r(A) \leq [r(\nu(A^m))]^\frac{1}{m} \leq r(A)$ as $m \to \infty \implies r(A) = \lim_{m \to \infty} [r(\nu(A^m))]^\frac{1}{m}$.

The next Theorem generalizes a well-known theorem of matrix norms [8,P.46].

Theorem 2. Let $\mathcal{M}_{n,k}$ denote the set of all matricial norms on $M_n$ of orders $k$ where $1 \leq k \leq n$. Then

$$r(A) = \inf_{\mu \in \mathcal{M}_{n,k}} \{r(\mu(A))\}, \quad \forall A \in M_n.$$ 

Proof: By Theorem 1 $r(A) \leq r(\nu(A))$, $\forall \nu \in \mathcal{M}_{n,k}$.

Now $C$, the field of complex numbers, is algebraically closed. Therefore all the eigenvalues of matrix $A$ lie in $C$ and so there exists an invertible matrix $B$ such that $BAB^{-1}$ is in the Jordan Form. Let $\epsilon > 0$ and $T=BAB^{-1}$. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $A$ (not necessarily distinct). Consider the diagonal matrix
\[
D = \begin{bmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
0 & e & 0 & 0 & \cdots & 0 \\
0 & 0 & e^2 & \cdots & 0 \\
& & & & & \\
& & & & & \\
0 & 0 & 0 & 0 & \cdots & e^{n-1}
\end{bmatrix}
\]

\[\det D = e, e^2, e^3, \ldots, e^{n-1}\] and

\[
D^{-f} = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & \frac{1}{e} & 0 & \cdots & 0 \\
0 & 0 & \frac{1}{e^2} & \cdots & 0 \\
& & & & & \\
& & & & & \\
0 & 0 & 0 & \cdots & \frac{1}{e^{n-1}}
\end{bmatrix}
\]

\[
DTD^{-1} = \begin{bmatrix}
\lambda_1 & \gamma_1 & 0 & 0 & \cdots & 0 \\
0 & \lambda_2 & \gamma_2 & 0 & \cdots & 0 \\
& & & & & \\
& & & & & \\
0 & 0 & 0 & \cdots & \gamma_{n-1} \\
& & & & & \\
& & & & & \\
0 & 0 & 0 & \cdots & \lambda_n
\end{bmatrix}
\]

Where \(\gamma_i = \frac{1}{e_i}, i = 1, 2, \ldots, n-1\)

\[
\gamma_i > 0.
\]

Let \(e_i\) be the vectors in \(C^n\) such that \(e_i = 1\) in the \(i\)-th place.

Let \(X_1, X_2, \ldots, X_{k-1}\) be the subspaces of \(C^n\) spanned by the vectors \(e_1, e_2, \ldots, e_{k-1}\) and \(X_k\) spanned by \(e_k, e_{k+1}, \ldots, e_n\). Then obviously \(C^n = X_1 \oplus X_2 \oplus \ldots \oplus X_k\). Let \(\mu_0 : M_n \to M_k^+\) be the matricial norm induced by this direct-sum decomposition and the
column norm on $M_n$ (Prop. 3).

Therefore $\nu_0(DTD^{-1}) = (\psi(e_i DTD^{-1} e_j))_{i,j=1,2,\ldots,k}$

$$\begin{pmatrix}
|\lambda_1| & \gamma_1 & 0 & 0 & \cdots & 0 & 0 \\
0 & |\lambda_2| & \gamma_2 & 0 & \cdots & 0 & 0 \\
0 & 0 & |\lambda_3| & \cdots & 0 & 0 & 0 \\
0 & 0 & 0 & \cdots & |\lambda_{k-1}| & \gamma_{k-1} & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & \alpha
\end{pmatrix}$$

where $\alpha = \max \{ |\lambda_k|, |\lambda_i| + \gamma_{i-1} \}$ for $i=k+1, \ldots, n$.

By taking the column norm of the above matrix and applying the above Lemma we have

$$r(\nu_0(DTD^{-1})) \leq \max_{i=1,2,\ldots,k-2} |\lambda_{i+1} + \gamma_i| + |\alpha + \gamma_{k-1}|$$

$$\leq \max_{i=1,2,\ldots,n-1} |\gamma_i| + \max_{i=1,2,\ldots,n-1} |\lambda_i|$$

$$\leq \frac{2}{\epsilon} + \max_{i=1,2,\ldots,n} |\lambda_i| = \frac{2}{\epsilon} + r(A) \quad \text{(1)}$$

Now define the mapping $\mu : M_n \to M_k^+$ such that $\mu(Q) = \nu_0(DBQB^{-1}D^{-1})$, where $Q \in M_n$.

Since $\mu$ is a DB-transform of $\nu_0$ [Proposition 2], $\mu$ is a matricial norm on $M_n$. Set $Q=A$. Therefore
\[ y(A) = p_0(DBAB^{-1}D^{-1}) = p_0(DTD^{-1}) \Rightarrow r(y(A)) = r(p_0(DTD^{-1})) < \frac{2}{\varepsilon} + r(A) \] by (1).

Therefore \[ r(A) \leq r(y(A)) < \frac{2}{\varepsilon} + r(A) \Rightarrow r(A) = \inf r(y(A)). \forall \mu \in \mathbb{W}_{n,k}

Let \( \mu: M_n \rightarrow M^+_k \) be a generalized matricial norm. Define

\[ K_{ij}(\mu) = \{ A \in M_n | \mu_{ij}(A) = 0 \} \]

\[ W_{ij}(\mu) = \{ K_{\alpha \beta}(\mu) | (\alpha, \beta) \neq (i, j) \} \]

\[ W(\mu) = \sum_{ij=1}^{k} W_{ij}(\mu) \]

Claim \( K_{ij}(\mu), W_{ij}(\mu) \) and \( W(\mu) \) are subspaces of \( M_n \). Let \( A \) and \( B \in K_{ij}(\mu) \). Then \( \mu_{ij}(A) = 0 = \mu_{ij}(B) \), and since \( \mu_{ij} \) is a pseudo-norm on \( M_n \), we have

\[ \mu_{ij}(A+B) \leq \mu_{ij}(A) + \mu_{ij}(B) = 0 \]

and \( \mu_{ij}(\alpha A) = |\alpha| \mu_{ij}(A) = 0 \). Therefore \( K_{ij}(\mu) \) is of \( M_n \).

Now since sum and intersection of subspaces is a subspace, therefore \( W_{ij}(\mu) \) and \( W(\mu) \) are subspaces of \( M_n \). These subspaces were introduced by Robert [14] in his investigation of vectorial norms.

**Proposition 7:** Let \( \mu: M_n \rightarrow M^+_k \) be a matricial norm such that

\[ W_{ij}(\mu) \neq 0, j = 1, 2 \ldots k. \] Then
i) $A \in W(\mu)$ and $A = \sum_{j=1}^{k} A_{jj}, A_{jj} \in W_{jj}(\mu) \Rightarrow \mu_{jj}(A) = \mu_{jj}(A_{jj})$.

ii) There exists an $n \times n$ matrix $A$ such that $\mu(A) = I_k$.

iii) $\mu(I_n) \geq I_k$.

**Proof:**

i) $\mu_{jj}$ is a pseudo-norm on $M_n$. Therefore $\mu_{jj}(A) = \mu_{jj}(A_{11} + A_{22} + \ldots + A_{kk}) = \mu_{jj}(A_{jj})$ since in general $\mu_{jj}(b_{jj}) = 0 \Rightarrow \mu_{jj}(a_{jj} + b_{jj}) \leq \mu_{jj}(a_{jj}) + \mu_{jj}(b_{jj}) = \mu_{jj}(a_{jj})$.

On the other hand $\mu_{jj}(a_{jj}) = \mu_{jj}(a_{jj} + b_{jj} - b_{jj}) \leq \mu_{jj}(a_{jj} + b_{jj}) + \mu_{jj}(b_{jj}) = \mu_{jj}(a_{jj}) + \mu_{jj}(b_{jj}) = \mu_{jj}(a_{jj})$.

ii) Since $W_{jj}(\mu)$ is a subspace of $M_n$ and $\mu_{jj}$ is a norm on $M_n$, we have $\mu_{jj}|_{W_{jj}(\mu)}$ is a norm. Let $A_{jj} \in W_{jj}(\mu)$ such that $A_{jj} \neq 0$. Let $\Delta = \sum_{j=1}^{k} \frac{1}{\mu_{jj}(A_{jj})} A_{jj}$.

Then $\mu_{jj}(\Delta) = 1 \Rightarrow \mu(\Delta) = I_k$.

iii) $\mu(I_n) = \mu(I_n) I_k = \mu(I_n) \cdot \mu(\Delta) \geq \mu(I_n \Delta) = \mu(\Delta) = I_k \Rightarrow \mu(I_n) \geq I_k$.

**Remark:** In Proposition 7, the assumption that $W_{jj}(\mu) \neq 0$ cannot be dropped. For example let $\psi$ be some matrix norm on $M_n$. Then the mappings $\mu, \nu: M_n \to M_k^+$ defined by

$$
\mu(A) = \begin{cases} 
\psi(A) & 0 \\
0 & 2\psi(A)
\end{cases}, \quad \nu(A) = \begin{cases} 
\psi(A) & 0 \\
0 & 0
\end{cases}
$$

for $A \in M_n$,
are matricial norms of order 2. But there exists no matrix \( \Delta \in M_n \) such that \( \nu(\Delta) = I_2 \) and we do not have \( \nu(I_n) \geq I_2 \).

We introduce the following minimality notion on matrix norms. A matrix norm \( \psi \) is said to be minimal for a complex \( n \times n \) matrix \( A \) if the spectral radius of \( A \) is equal to the matrix norm of \( A \) i.e., \( r(A) = \psi(A) \). It is obvious that there exists matrices for which no matrix norm is minimal. For example, nilpotent matrices (because \( A \neq 0 \Rightarrow \psi(A) \neq 0 \), and if \( A \) is nilpotent then \( r(A^m) = 0 \) and so \( \psi(A) \neq r(A) \)). Our object is to extend the notion of minimal matrix norm to that of minimal matricial norm. We will see that for any \( n \times n \) complex matrix \( A \) there exists a minimal matricial norm for \( A \).

**Definition:** A matricial norm \( \nu : M_n \rightarrow M_k^+ \) is said to be minimal for the matrix \( A \in M_n \), if \( r(A) = r(\nu(A)) \).

**Example 4:** Let

\[
A = \begin{pmatrix}
1 & 2 & 0 & 0 \\
2 & 4 & 0 & 0 \\
0 & 1 & 1 & 2 \\
0 & 0 & 2 & 4
\end{pmatrix}.
\]

The eigenvalues of \( A \) are given by the equation \( x^4 - 10x^3 + 25x^2 = 0 \) \( \Rightarrow x = 0, 0, 5, 5 \Rightarrow r(A) = 5 \). Define the matricial norm \( \nu : M_4 \rightarrow M_2^+ \) such that for all \( B = (b_{ij}) \in M_4 \),
\[
\mu(B) = \left( \begin{array}{c}
(b_{11}^2 + b_{12}^2 + b_{21}^2 + b_{22}^2)^{\frac{1}{2}} & (b_{13}^2 + b_{14}^2 + b_{23}^2 + b_{24}^2)^{\frac{1}{2}} \\
(b_{31}^2 + b_{32}^2 + b_{41}^2 + b_{42}^2)^{\frac{1}{2}} & (b_{33}^2 + b_{34}^2 + b_{43}^2 + b_{44}^2)^{\frac{1}{2}}
\end{array} \right).
\]

Therefore \( \mu(A) = \begin{bmatrix} 5 & 0 \\ 1 & 5 \end{bmatrix} \Rightarrow r(\mu(A)) = 5 = r(A) \) which shows that \( \mu \) is minimal for \( A \).

**Proposition 8:** Let \( \mu : M_n \to M_k^+ \) be a matricial norm. Then \( \mu \) is minimal for \( A \) \( \Rightarrow \mu \) is minimal for \( A^m \), \( m=1,2,3 \ldots \).

**Proof:** Since \( \mu \) is minimal for \( A \), therefore \( r(A) = r(\mu(A)) \).

Also by Theorem 1, \( \forall A \in M_n \), \( r(A) \leq r(\mu(A)) \Rightarrow r(A^m) \leq r(\mu(A^m)) \)

\[
\leq r(\mu(A)^m) = [r(\mu(A))]^m = r(A^m) = r(A)^m.
\]

Therefore \( r(A^m) = r(\mu(A^m)) \), and hence the Proposition.

**Proposition 9:** Let \( \mu : M_n \to M_k^+ \) be matricial norm. Suppose \( \mu(A^m) = \mu(A)^m \) \( \forall \ m = 1,2,3 \ldots \). Then \( \mu \) is minimal for \( A \).

**Proof:** Since \( \mu(A)^m = \mu(A^m) \) we have \( r(\mu(A^m)) = r(\mu(A)^m) \).

Therefore \( r(\mu(A)) = [r(\mu(A^m))]^{\frac{1}{m}} \) for \( m=1,2,3 \ldots \). Now by Proposition 6 we have \( r(A) = [r(\mu(A^m))]^{\frac{1}{m}} = r(\mu(A)) \), and therefore \( \mu \) is minimal for \( A \).

**Remark:** The condition in Proposition 9 is necessary but not sufficient. For example, it is shown in example 5 that \( \mu \) is minimal for \( A \). However,

\[
\mu(A) = \begin{bmatrix} 5 & 0 \\ 1 & 5 \end{bmatrix}, \quad \mu(A)^2 = \begin{bmatrix} 25 & 0 \\ 10 & 25 \end{bmatrix}
\]
\[
A^2 = \begin{pmatrix}
5 & 10 & 0 & 0 \\
10 & 20 & 0 & 0 \\
2 & 5 & 5 & 10 \\
0 & 2 & 10 & 20 \\
\end{pmatrix}
\Rightarrow \mu(A^2) = \begin{pmatrix}
25 & 0 \\
\sqrt{33} & 25 \\
\end{pmatrix},
\]

and so \(\mu(A^2) \neq \mu(A)^2\). We show below that if \(\mu(A^m)\) is irreducible, then the condition is sufficient.

**Definition**: Square matrix \(A\) is said to be reducible, if there exists a permutation matrix \(P\) such that

\[
P^T A P = \begin{pmatrix}
A_{11} & 0 \\
0 & A_{22} \\
\end{pmatrix}
\]

where diagonal matrices are square matrices and \(P^T\) denotes the transpose of \(P\). Matrices which are not reducible are called irreducible.

**Proposition 10**: Let \(\mu : M_\mathbb{R}^n \rightarrow M_\mathbb{R}^k\) be a minimal matricial norm for a matrix \(A\). Suppose \(\mu(A^m)\) is irreducible where \(m\) is some positive integer. Then \(\mu(A^m) = \mu(A)^m\).

**Proof**: Suppose \(\mu(A^m) \neq \mu(A)^m\). Then \(\mu(A^m) < \mu(A)^m\). Therefore [16, P.30], \(r(\mu(A^m)) < r(\mu(A)^m) = [r(\mu(A))]^m = r(A)^m\) since \(\mu\) is minimal. Also \(r(A^m) \leq r(\mu(A^m))\). Thus \(r(\mu(A^m)) < r(A^m) \leq r(\mu(A^m))\) which is not possible of course. Hence we must have \(\mu(A^m) = \mu(A)^m\).

**Remark**: In the above Proposition the condition \(\mu(A^m)\) is irreducible cannot be dropped. In the above remark, \(\mu(A^2)\) is reducible, \(\mu\) is minimal for \(A\) and we do not have \(\mu(A^2) \neq \mu(A)^2\).

**Proposition 11**: Let \(\mu\) be a minimal matricial norm. Suppose \(G\) is an invertible \(n \times n\) complex matrix. Then the \(G\)-transform of \(\mu\) is minimal for \(G^{-1}AG\).
Proof: Since similar matrices have the same characteristic polynomials, therefore \( r(A) = r(G^{-1}AG) \). Define \( \mu_G : M_n \rightarrow M_k^+ \) such that \( \mu_G(A) = \mu(GAG^{-1}) \) for all \( A \in M_n \).

By Proposition 2 \( \mu_G \) is a matricial norm. Now the \( G \)-transform of \( \mu \) for \( G^{-1}AG \) is \( \mu_G(G^{-1}AG) = \mu(GG^{-1}AGG^{-1}) = \mu(A) \) and therefore \( r(\mu_G(G^{-1}AG)) = r(\mu(A)) = r(A) = r(G^{-1}AG) \). Hence the result.

The next proposition guarantees the existence of a matricial norm of order \( n \) which is minimal for \( A \).

**Proposition 12:** Given \( A \in M_n \), there exists a matricial norm \( \mu : M_n \rightarrow M_n^+ \) of order \( n \) which is minimal for \( A \).

**Proof:** Given \( A \in M_n \) we can find an invertible matrix \( G \) such that \( GAG^{-1} \) is in Jordan form. Define the matricial norm \( \mu : M_n \rightarrow M_n^+ \) such that \( \mu(B) = (|b_{ij}|) \) for all \( B \in M_n \).

Now if \( B \) is a triangular matrix, then, since the eigenvalues of triangular matrix lie on the main diagonal, \( \mu \) is minimal for \( B \). In particular \( \mu \) is minimal for \( GAG^{-1} \). So by the above Proposition, \( \mu \) is minimal for \( A \).

**Definition:** Let \( k \) be the least positive integer such that there exists a matricial norm \( \mu : M_n \rightarrow M_k^+ \) which is minimal for \( A \). Then \( A \) is said to be of class \( k \).

The existence of such a positive integer is guaranteed by Proposition 12.
Remark: The matrix of example 5 is of Class 2.

Remark: Matrices of Class 1 are also called of Class M [8, P. 46].

Example 5: Diagonalizable matrices i.e., matrices similar to diagonal matrices are of Class 1.

Let $A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1k} \\ A_{21} & A_{22} & \ldots & A_{2k} \\ & & \ldots & \ldots \\ A_{k1} & A_{k2} & \ldots & A_{kk} \end{pmatrix}$ be any partition of the matrix $A \in M_n$. Then it is easy to check directly that the mapping $\phi : M_n \rightarrow M_k^+$ such that $\phi(A) = (\psi(A_{ij}))_{ij=1,2,\ldots,k}$ for all $A \in M_n$, where $\psi$ is matrix norm is a matricial norm on $M_n$.

The next theorem is a well-known theorem, and its proof can be found in most books on linear algebra.

Theorem 3: Let $V$ be a finite dimensional vector space over $\mathbb{C}$. Let $T$ be the linear transformation on $V$. Suppose that $f(z) = a_0 + a_1z + \ldots + a_{n-1}z^{n-1} + z^n$ is the minimal polynomial of $T$ over $\mathbb{C}$. Then there exists bases of $V$ in which the matrix of $T$ is
and that the roots of \( f(z) \) are precisely the eigenvalues of the matrix \( F \).

**Note:** For convenience, we have taken \( f(z) \) to be monic.

The matrix \( F \) is called the companion matrix of \( f(z) \).

Let \( e(f) \) be the largest of the absolute values of the zeroes of \( f(z) \). Consider the matrix

\[
D = \begin{pmatrix}
\alpha_0 & 0 & 0 & \ldots & 0 & 0 \\
0 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & \alpha_{n-2} & 0 \\
0 & 0 & \ldots & 0 & 1
\end{pmatrix}
\]

\[
\text{Det } D = \alpha_0 \alpha_1 \ldots \alpha_{n-2}
\]
\[ D^{-1} = \begin{bmatrix} \frac{1}{a_0} & 0 & 0 & \ldots & 0 & 0 \\ 0 & \frac{1}{a_1} & 0 & \ldots & 0 & 0 \\ & & & \ddots & & \\ & & & & \frac{1}{a_{n-2}} & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \]

Therefore \( D^{-1}FD = \begin{bmatrix} 0 & 0 & \ldots & -a_0/a_0 \\ \frac{a_0}{a_1} & 0 & \ldots & -a_1/a_1 \\ 0 & \frac{a_1}{a_2} & \ldots & -a_2/a_2 \\ & & \ddots & \ddots \\ & & & \frac{a_{n-3}}{a_{n-2}} & 0 \end{bmatrix} \).

Partition this matrix as follows

\[ D^{-1}FD = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ \frac{a_0}{a_1} & 0 & 0 & \ldots & 0 & 0 \\ 0 & \frac{a_1}{a_2} & 0 & \ldots & 0 & 0 \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & \frac{a_{n-3}}{a_{n-2}} & 0 \\ 0 & 0 & 0 & \ldots & \alpha_{n-2} & 0 \\ 0 & 0 & 0 & \ldots & 0 & \alpha_{n-1} \end{bmatrix} \]

Now since \( F \) and \( D^{-1}FD \) are similar matrices, therefore they have
the same spectral radius. Thus,

\[ e(f) = r(F) = r(D^{-1}FD) \] .... (2)

For the partition given by (1), define the matricial norm 
\( \phi: M_n \to M_2^+ \) such that

\[ \phi(D^{-1}FD) = \begin{bmatrix} \beta & \gamma \\ \alpha_{n-2} & |\alpha_{n-1}| \end{bmatrix} \] .... (3)

where \( \beta = \max\{ \frac{\alpha_0}{\alpha_1}, \frac{\alpha_1}{\alpha_2}, \ldots, \frac{\alpha_{n-3}}{\alpha_{n-2}} \} \) .... (4)

\[ \gamma = \max\{ \frac{|a_0|}{a_0}, \frac{|a_1|}{a_1}, \ldots, \frac{|a_{n-2}|}{a_{n-2}} \} \] .... (5)

Now by Theorem 1 \( r(A) \leq r(\mu(A)) \Rightarrow r(D^{-1}FD) \leq r(\phi(D^{-1}FD)) \), and therefore

\[ e(f) \leq r(\phi(D^{-1}FD)) \] .... (6)

Let \( \sigma \) be any matrix norm on \( M_2 \). Since the spectral radius cannot exceed any of the matrix norms, therefore by (16),

\[ e(f) \leq \sigma(\phi(D^{-1}FD)) \] .... (7)

The upperbound given by (7) cannot be better than (6). Let us evaluate (6) and (7). Eigenvalues of the matrix given by (3) are the zeroes of 

\[ x^2 - (\beta + |a_{n-1}|)x + \beta |a_{n-1}| - \alpha_{n-2} \gamma = 0, \]

namely

\[ x = \frac{\beta + |a_{n-1}| \pm \sqrt{(\beta + |a_{n-1}|)^2 - 4(\beta |a_{n-1}| - \alpha_{n-2} \gamma)}}{2} \]

\[ = \frac{1}{2} \left[ (\beta + |a_{n-1}| \pm \sqrt{(\beta - |a_{n-1}|)^2 + 4\alpha_{n-2} \gamma}) \right] \]

Taking positive value to get the maximum, we have by (6)

\[ e(f) \leq \frac{1}{2} \left[ \beta + |a_{n-1}| + \sqrt{(|a_{n-1}| - \beta)^2 + 4\alpha_{n-2} \gamma} \right]. \]
To find (7), we take respectively row norm, column norm and euclidean norm of (3), we have

\[
e(f) \leq \max \{ \beta + \gamma, \quad a_{n-2} + |a_{n-1}| \}
\]

\[
e(f) \leq \max \{ \beta + a_{n-2}, \quad \gamma + |a_{n-1}| \}
\]

\[
e(f) \leq (\beta^2 + \gamma^2 + |a_{n-1}|^2 + a_{n-2}^2)^{\frac{1}{2}}
\]

Putting all the pieces together, we obtain

Theorem 4: Let \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n, a_i \in \mathbb{C} \) and let \( \alpha_0, \alpha_1, \ldots, \alpha_{n-2} \) be positive real numbers. Then

\[
e(f) \leq \frac{1}{\sqrt{2}} [\beta + |a_{n-1}| + \sqrt{(|a_{n-1}| - \beta)^2 + 4\alpha_{n-2} \gamma}]
\]  

.... (8)

\[
e(f) \leq \max \{ \beta + \gamma, a_{n-2} + |a_{n-1}| \}
\]  

.... (9)

\[
e(f) \leq \max \{ \beta + a_{n-2}, \gamma + |a_{n-1}| \}
\]  

.... (10)

\[
e(f) \leq (\beta^2 + \gamma^2 + |a_{n-1}|^2 + a_{n-2}^2)^{\frac{1}{2}}
\]  

.... (11)

where \( \beta \) and \( \gamma \) are given by (4) and (5).

Since (9), (10) and (11) are obtained by applying different norms, therefore these inequalities cannot give better bounds than that given by (8). Substitute for \( \beta \) and \( \gamma \) in (9), we have

\[
e(f) \leq \max \{ \frac{|a_0|}{\alpha_0}, \quad \frac{|a_1| + a_0}{\alpha_1}, \quad \ldots \quad \frac{|a_{n-2} + a_{n-3}|}{\alpha_{n-2}}, \quad \alpha_{n-2} + |a_{n-1}| \}
\]

which is Wilf's inequality [17]. Another proof will be given in the next chapter using Frobenius theory.
Taking \( \alpha_0 = \alpha_1 = \ldots \alpha_{n-2} = 1 \) i.e., when the diagonal matrix \( D \) is the identity matrix, we obtain from Theorem 4, Cor. 1 Let \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n, a_i \in \mathbb{C} \). Set \( \alpha_i = 1 \) \( \forall i = 0,1 \ldots n-2 \). Then we have

\[
\eta(f) \leq \frac{1}{2} \left[ 1 + |a_{n-1}| + \sqrt{(|a_{n-1}| - 1)^2 + 4M} \right] \quad \ldots \quad (12)
\]

\[
\eta(f) \leq \max\{|a_0|, 1 + |a_1|, \ldots, 1 + |a_{n-1}|\} \quad \ldots \quad (13)
\]

\[
\eta(f) \leq \max\{2, |a_0| + |a_{n-1}|, \ldots, |a_{n-2}| + |a_{n-1}| \} \ldots \quad (14)
\]

\[
\eta(f) \leq (2 + M^2 + |a_{n-1}|^2)\frac{1}{2} \quad \ldots \quad (15)
\]

where \( M = \max\{|a_0|, |a_1|, \ldots, |a_{n-2}|\} \).

(12) is Rehman's inequality [9] and (13) is Cauchy's inequality [11]. Another proof of these inequalities can be found in the next chapter. The following example shows that (14) can give a better upper bound than Cauchy's (13).

**Example 6**: Consider \( f(z) = -2 - 2z - 0.5z^2 + z^3 \)

Cauchy's (13) gives \( \eta(f) \leq 3 \).

(14) gives \( \eta(f) \leq 2.5 \).

It is noteworthy to see that all inequalities from (8) to (15) contain coefficients of \( z^{n-1} \) and if we set \( a_{n-1} = 0 \) the inequalities are simplified. In particular, if \( |a_{n-1}| = 1 \) in Rehman's inequality, we have \( \eta(f) \leq 1 + \sqrt{M} \).

Assume that all the coefficients in \( f(z) \) are non-zero and real. Set \( \alpha_i = \alpha_{i+1} \) \( \forall i = 0,1, \ldots, n-2 \), we have from theorem 4.
Cor. 2: Let \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \) for \( a_i \neq 0 \) \( \forall i \).

Set \( a_i = a_{i+1} \) for \( i = 0, 1, \ldots, n-2 \). We obtain

\[
e(f) \leq \frac{1}{2} \left[ \beta' + |a_{n-1}| + \sqrt{(|a_{n-1}| - \beta')^2 + 4\gamma' |a_{n-1}|} \right] 
\]

\[\ldots \ldots (16)\]

\[
e(f) \leq \max\{\beta' + \gamma', |a_{n-1}| + a_{n-1}\} 
\]

\[\ldots \ldots (17)\]

\[
e(f) \leq \max\{\beta' + a_{n-1}, \gamma' + |a_{n-1}|\} 
\]

\[\ldots \ldots (18)\]

\[
e(f) \leq (\beta'^2 + \gamma'^2 + 2|a_{n-1}|^2) \frac{1}{2} 
\]

\[\ldots \ldots (19)\]

where \( \beta' = \max\left\{ \frac{|a_1|}{|a_2|}, \frac{|a_2|}{|a_3|}, \ldots, \frac{|a_{n-2}|}{|a_{n-1}|} \right\} \)

\( \gamma' = \max\left\{ \frac{|a_0|}{|a_1|}, \frac{|a_1|}{|a_2|}, \ldots, \frac{|a_{n-2}|}{|a_{n-1}|} \right\} \)

Observe that \( \beta' < \gamma' \) but if \( \beta' = \gamma' \) then (16) and (18) coincide provided \( a_{n-1} > 0 \), for then in each case \( e(f) \leq \beta' + a_{n-1} \).

Substitute for \( \beta' \) and \( \gamma' \) in (17), and set \( a_i = |a_{i+1}| \) \( i = 0, 1, \ldots, n-2 \). We obtain the following Kojima's inequality [11]

\[
e(f) \leq \max\{ 2|a_{n-1}|, \frac{|a_0|}{|a_1|}, \frac{2|a_1|}{|a_2|}, \ldots, \frac{2|a_{n-2}|}{|a_{n-1}|} \} 
\]

For another proof of this inequality see next chapter. Inequalities (16) to (19) can still be modified if \( a_i \) are real and some of the co-efficients are zero. In that case, set
The following example shows that (18) can give a better upper bound than Kojima's inequality.

**Example 7:** \( f(z) = -1 + 2z + 3z^2 - 4z^3 + z^4 \)

(18) gives \( e(f) \leq 4 + \max \left\{ \frac{1}{2}, \frac{2}{3}, \frac{3}{4} \right\} = 4.75 \)

Kojima's inequality gives \( e(f) \leq 8 \).

Let \( t \neq 1 \) be any positive number and set \( a_i = t^n - i.1 \)

\( i = 0, 1, \ldots n-2 \) in Theorem 4, we obtain

**Cor. 3** Let \( f(z) = a_0 + a_1 z + \ldots a_{n-1}^n z^n + z \), \( a_i \in \mathbb{C} \). Let \( t \) be a positive number different from 1. Set \( a_i = t^n - i.1 \)

\( i = 0, 1, \ldots n-2 \). We obtain

\[
e(f) \leq \frac{1}{t} \left[ t + |a_{n-1}| + t \left( |a_{n-1}| - t \right)^2 + 4 \delta t \right]
\]

\[
e(f) \leq \max \left\{ \frac{|a_0|}{t^{n-1}}, t + \frac{|a_1|}{t^n}, \ldots t + \frac{|a_1|}{t^n}, t + |a_{n-1}| \right\}
\]

\[
e(f) \leq \max \{ 2t, |a_{n-1}| + \frac{|a_0|}{t^{n-1}}, \ldots |a_{n-1}| + \frac{|a_{n-2}|}{t} \}
\]

\[
e(f) \leq (2t^2 + \delta^2 + |a_{n-1}|^2)^{\frac{1}{2}}
\]

where \( \delta = \max \left\{ \frac{|a_0|}{t^{n-1}}, \frac{|a_1|}{t^n}, \ldots, \frac{|a_{n-2}|}{t} \right\} \).

(21) is Wilf's inequality [17]. We will give another proof in Chapter 2. The next example shows that (22) can give a
better bound than Wilf's (21).

Example 8: Let \( f(z) = -24 - 38z - 13z^2 + 2z^3 + z^4 \)

(22) gives \( e(f) \leq \max \{ 6, 2 + \frac{24}{7}, 2 + \frac{38}{9}, 2 + \frac{13}{3} \} = 6.34 \)

(21) gives \( e(f) \leq \max \{ \frac{24}{7}, 7.22, 7.33, 5 \} = 7.33 \)

In fact it can be easily verified that Wilf's bound (21) is
the weakest of the above bounds.

Let \( |a_{n-1}| \neq 0 \). If \( t \) is so chosen that \( t = |a_{n-1}| \)
then we have from Cor. 3

\[
e(f) \leq |a_{n-1}| + \sqrt{\delta'} \] .... (24)

\[
e(f) \leq \max \left\{ \frac{|a_0|}{|a_{n-1}|} n-1, \frac{|a_{n-1}| + |a_1|}{|a_{n-1}|} n-2, \ldots 2|a_{n-1}| \right\} .... (25)

\[
e(f) \leq |a_{n-1}| + \max \left\{ \frac{|a_{n-1}| + |a_{n-2}|}{|a_{n-1}|} \right\} \] .... (26)

\[
e(f) \leq (3|a_{n-1}|^2 + \frac{\delta'^2}{|a_{n-1}|^2})^2 \] .... (27)

where \( \delta' = \max \left\{ \frac{|a_0|}{|a_{n-1}|} n-2, \frac{|a_{n-1}| + |a_{n-2}|}{|a_{n-1}|} n-3, \ldots |a_{n-2}| \right\} \).

Let \( N = \max \left\{ \frac{1}{n}, \frac{1}{n-1}, \ldots \frac{1}{2} \right\} \) .... (28)

Then \( N^{n-i} \geq |a_i| \) \( \forall i = 0, 1 \ldots n-2 \) with equality for at least
one \( i \) as \( N \) is maximum. If \( t \) is so chosen that \( t=N \), then from
Cor. 3 we have
\[ e(f) \leq \frac{1}{2} \left( N + |a_{n-1}| + \sqrt{(|a_{n-1}| - N)^2 + 4N^2} \right) \quad \ldots \quad (29) \]

\[ e(f) \leq \max \left\{ \frac{|a_0|}{N^{n-1}}, \frac{|a_1|}{N^{n-2}}, \ldots, N + |a_{n-1}| \right\} \]

\[ = \max N \left\{ \frac{|a_0|}{N^n}, \frac{|a_1|}{N^{n-2}}, \ldots, \frac{|a_{n-1}|}{N^{n-1}} \right\} \]

\[ \leq N \max \left\{ 1, 1 + \frac{1}{2}, \ldots, 1 + \frac{1}{n} \right\} \text{ as } |a_i| \leq 1 \]

\[ = 2N \quad \ldots \quad (30) \]

\[ e(f) \leq N + \max \left\{ N, |a_{n-1}| \right\} \quad \ldots \quad (31) \]

\[ e(f) \leq (3N^2 + |a_{n-1}|^2)^{\frac{1}{2}} \quad \ldots \quad (32) \]

where \( N \) is given by (28) and \( \delta = N \).

From (31) we have

\[ e(f) \leq \max \left\{ |a_{n-i}|^{\frac{1}{i}}, \max \left\{ |a_{n-i}|^{\frac{1}{i}}, |a_{n-1}| \right\} \right\}_{i=2}^{n} \]

\[ = \max \left\{ |a_{n-i}|^{\frac{1}{i}}, \max \left\{ |a_{n-i}|^{\frac{1}{i}} \right\} \right\}_{i=2}^{n} \]

\[ \leq 2 \max \left\{ |a_{n-i}|^{\frac{1}{i}} \right\} \quad \text{for } i = 1, 2, \ldots, n \]

which is Fujwara's inequality [11]. Another proof can be found in the next chapter.

The following example shows that (29) can give a better bound than Fujwara's inequality. But before we give an example, we note that if the maximum in (28) is \( |a_{n-1}| \) then all the four inequalities coincide and in each case we have

\[ e(f) \leq 2N = 2|a_{n-1}|. \]
Example 9: \[ f(z) = -2 - 2z - 0.5z^2 + z^3 \]

\[ N = \sqrt{2} \]

(29) gives \( \varepsilon(f) \leq 2.36 \)

Fujwara's inequality gives \( \varepsilon(f) \leq 3.46 \)

Even (32) gives a better bound than Fujwara.
Chapter II

Zeroes of Polynomials

The main object in this chapter is to deal with the location of the zeroes of the polynomial $f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n$, $a_i \in \mathbb{C}$ and in particular with the trinomial equation $1 - z + \lambda z^n = 0, \lambda \neq 0$. Proofs of different inequalities promised earlier are also given.

The following theorem is a classical result due to Cauchy [11].

**Theorem 1:** All the zeroes of the polynomial $f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n$, $a_i \in \mathbb{C}$ lie in the circle $|z| < 1 + M$ where $M = \max |a_i|, 0 \leq i \leq n-1$.

**Proof:** We have $|z|^n = |f(z) - (a_0 + a_1 z + \ldots + a_{n-1} z^{n-1})|$

$$\leq |f(z)| + |a_0 + a_1 z + \ldots + a_{n-1} z^{n-1}|$$

$$\leq |f(z)| + |a_0| + |a_1||z| + \ldots + |a_{n-1}||z^{n-1}|$$

$$
\Rightarrow f(z) \geq |z|^n - \sum_{i=0}^{n-1} |a_i||z|^i = |z|^n (1 - \sum_{i=0}^{n-1} \frac{|a_i|}{|z|^{n-i}})
$$

$$
\Rightarrow f(z) \geq |z|^n (1 - \sum_{i=1}^{\infty} M|z|^{-i}) \text{ since } M = \max |a_i|
$$

$$
\Rightarrow |f(z)| > \frac{|z|^n}{|z|-1} (1 - \sum_{i=1}^{\infty} \frac{M}{|z|^i}) \text{ if } |z| > 1
$$

$$\Rightarrow |f(z)| > \frac{|z|^n}{|z|-1} (1 - 1 - M)
$$

Now if $|z| \geq 1 + M$ then $|f(z)| > 0$. Therefore no root of $f(z)$ lies in the circle $|z| \geq 1 + M$, and so all roots lie in $|z| < 1 + M$. 
The next theorem is due to Rehman [9]. It gives a better (smaller) circle than given in Theorem 1.

**Theorem 2**: All zeroes of the polynomial \( f(z) = a_0, a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \) lie in the circle \( |z| \leq \frac{1}{2} \left[ 1 + |a_{n-1}| + \sqrt{(|a_{n-1}| - 1)^2 + 4M} \right] \)

where \( M = \max\{|a_i| \mid 0 \leq i \leq n-1\} \).

**Proof**: Suppose \( |z| > \frac{1}{2} \left[ 1 + |a_{n-1}| + \sqrt{(|a_{n-1}| - 1)^2 + 4M} \right] \).

Then \( (2|z| - 1 - |a_{n-1}|)^2 > (|a_{n-1}| - 1)^2 + 4M \),

and \( (|z| - 1) \left( |z| - |a_{n-1}| \right) - M > 0 \implies |z| - |a_{n-1}| - \frac{M}{|z| - 1} > 0 \).

Multiply by \( |z|^{n-1} \) we have \( |z|^{n-1} - |a_{n-1}| |z|^{n-1} - \frac{M |z|^{n-1}}{|z| - 1} > 0 \).

Consider \( \frac{M |z|^{n-1}}{|z| - 1} = \frac{M |z|^{n-1}}{|z| \left( 1 - \frac{1}{|z|} \right)} = M |z|^{n-2} \left( 1 + \frac{1}{|z|} + \frac{1}{|z|^2} + \ldots \right) \)

\( > M |z|^{n-2} \left( 1 + \frac{1}{|z|} + \ldots + \frac{1}{|z|^{n-2}} \right) > M (1 + |z| + \ldots |z|^{n-2}) \)

\( \geq |a_0| + |a_1| |z| + \ldots |a_{n-2}| |z|^{n-2} \)

\( \geq |a_0 + a_1 z + \ldots + a_{n-2} z^{n-2}|. \)

Also \( |z|^n - |a_{n-1}| |z|^{n-1} \leq |z^n + a_{n-1} z^{n-1}|. \) Hence from (1) we have

\( |z^n + a_{n-1} z^{n-1}| - |a_0 + a_1 z + \ldots a_{n-2} z^{n-2}| > 0 \implies |f(z)| > 0 \)
Thus no root of \( f(z) \) lies in the circle \(|z| > \frac{1}{2} \left[ 1 + \left| a_{n-1} \right| + \sqrt{(\left| a_{n-1} \right| - 1)^2 + 4M} \right] \). Therefore all zeroes lie in

\[ |z| \leq \frac{1}{2} \left[ 1 + \left| a_{n-1} \right| + \sqrt{(\left| a_{n-1} \right| - 1)^2 + 4M} \right]. \]

If we replace \( z \) by \( \frac{1}{z} \) in \( f(z) \), we have

\[ z^n f\left(\frac{1}{z}\right) = 1 + a_{n-1} z + \ldots + a_1 z^{n-1} + z^n, \quad |a_0| = 1. \]

Let \( N = \max\{\left| a_i \right| \} \). By applying Theorem 2 to \( z^n f\left(\frac{1}{z}\right) \), we get all zeroes of \( 1 + a_{n-1} z + \ldots + a_1 z^{n-1} + z^n \) to lie in the circle

\[ \frac{1}{|z|} \leq \frac{1}{2} \left[ 1 + \left| a_1 \right| + \sqrt{(1 - \left| a_1 \right|)^2 + 4N} \right]. \]

Therefore \(|z| \geq \frac{2}{1 + \left| a_1 \right| + \sqrt{(1 - \left| a_1 \right|)^2 + 4N}} \) we have proved

\[ \text{Cor. 1} \quad \text{Let } N = \max\{\left| a_i \right| \}. \text{ Then the polynomial } g(z) = 1 + a_{n-1} z + \ldots + a_1 z^{n-1} + z^n \text{ has no root in the circle} \]

\[ |z| < \frac{2}{1 + \left| a_1 \right| + \sqrt{(1 - \left| a_1 \right|)^2 + 4N}}. \]

If we apply theorem 2 to the polynomial \((z - a_{n-1}) f(z)\), we have

\[ \text{Cor. 2} \quad \text{All the zeroes of } f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \text{ lie in the circle } |z| \leq \frac{1}{2} \left[ 1 + \sqrt{1 + 4N'} \right] \text{ where } N' = \max\{\left| a_i - 1 - a_{n-1} a_i \right|, 0 < i < n-1 \}, \]

\[ a_{-1} = 0. \]

The following theorem has been proved in different ways
by Zedek and Rubinstien separately. The following proof is simpler than the both.

Theorem 3: Let \( f(z) = a_0 + a_1z + \ldots + a_{m-1}z^{m-1} + z^m, \)
\( g(z) = b_0 + b_1z + \ldots + b_{n-1}z^{n-1} + z^n. \) Suppose their zeroes lie in the circles \(|z-c_1| \leq r_1\) and \(|z-c_2| \leq r_2\) respectively.

Let \( m > n \geq 1 \) and \( \epsilon \) be the unique positive root of the equation
\[ h(z) = z^m - |\lambda| (z+|c_2-c_1|+r_1+r_2)^n. \]
Then the \( m \) zeroes of \( f(z) + \lambda g(z) \) lie in the circle \(|z-c_1| \leq r_1+\epsilon_1\).

**Proof:** Let \( \xi_1, \xi_2, \ldots \xi_m \) and \( \eta_1, \eta_2, \ldots \eta_n \) be the zeroes of \( f(z) \) and \( g(z) \) respectively. So \(|\xi_i - c_1| \leq r_1\) and \(|\eta_i - c_2| \leq r_2\).

Suppose \(|z-c_1| = \epsilon + r_1\). Then
\[ |z-i| = |z-c_1 + c_1 - c_2 - c_2 - \eta_i| \leq |z-c_1| + |c_2-c_1| + |\eta_i - c_2| \]
\[ = \epsilon + r_1 + |c_2-c_1| + |\eta_i - c_2| \leq r_1 + r_2 + \epsilon + |c_2-c_1|. \]

Thus \(|\lambda g(z)| = |\lambda| \prod_{i=1}^{n} |z-\eta_i| \leq |\lambda|(r_1 + r_2 + \epsilon + |c_2-c_1|)^n \) .... (1)

Also \(|f(z)| = \prod_{i=1}^{m} |z-\xi_i| = \prod_i (|z-c_1 + c_1 - \xi_i| \geq \prod_i (|z-c_1| - |\xi_i - c_1|) \]
\[ \geq \prod_i (r_1 + \epsilon - r_1) = \epsilon^m \] .... (2)

Therefore if \( \epsilon > \epsilon_1 \) then since \( \epsilon_1 \) (but not \( \epsilon \) ) is the root of \( h(z) \) we have \( \epsilon^m > |\lambda|^n (\epsilon + |c_2-c_1| + r_1 + r_2)^n. \)

Thus from (1) and (2) we have \(|\lambda g(z)| < \epsilon^m < |f(z)|.\)

Thus when \(|z-c_1| = \epsilon + r_1\) and \( \epsilon > \epsilon_1 \) then \( \lambda g(z) + f(z) \) cannot vanish for any value of \( z \) in \(|z-c_1| > r_1 + \epsilon_1. \) Therefore all the zeroes must lie in the circle \(|z-c_1| \leq r_1 + \epsilon_1. \)
Theorem 4: If \( a_n \geq a_{n-1} \geq \ldots \geq a_1 \geq a_0 \). Then \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + a_n z^n \) has all the zeroes in the circle \( |z| \leq \frac{(a_n - a_0 + |a_0|)}{|a_n|} \).

Proof: Consider the polynomial
\[
(1-z)f(z) = -a_n z^{n+1} + z^n (a_n - a_{n-1}) + \ldots (a_1 - a_0) z + a_0 = g(z) - a_n z^{n+1}
\]
where \( g(z) = (a_n - a_{n-1}) z^{n+1} + \ldots (a_1 - a_0) z + a_0 \).

Now on the unit circle \( |z|=1 \), we have \( |g(z)| \leq |a_n - a_{n-1}| + \ldots + |a_1 - a_0| + |a_0| \) and therefore \( |g(z)| \leq a_n - a_{n-1} + a_{n-1} - a_{n-2} + \ldots + a_1 - a_0 + |a_0| \). Again on the unit circle \( |z|=1 \), we have
\[
|z^n g(1)| \leq a_n - a_0 + |a_0|.
\]
Since \( z^n g(1) \) is polynomial it is analytic in \( |z| \leq 1 \). Thus
\[
|z^n g(1)| \leq a_n - a_0 + |a_0| \frac{1}{|z|^n} \text{ for } |z| \leq 1 \Rightarrow |g(z)| \leq (a_n - a_0 + |a_0|) |z|^n.
\]
for \( |z| \geq 1 \). Hence \( |(1-z)f(z)| = |g(z) - a_n z^{n+1}| \geq |a_n z^{n+1}| - |g(z)| \geq |a_n| |z|^{n+1} - |z^n (a_n - a_0 + |a_0|)| = |a_n| |z|^n (|z| - a_n - a_0 + |a_0|) > 0 \text{ if } |z| \geq a_n - a_0 + |a_0|.
\]
Thus for \( |z| > a_n - a_0 + |a_0| \), \( f(z) \) does not vanish. Hence all the zeroes of \( f(z) \) lie in the circle \( |z| \leq \frac{a_n - a_0 + |a_0|}{|a_n|} \).
Remark: Suppose $a_0 \geq 0$. Then $|z| \leq \frac{a_n - a_0 + a_0}{a_n} = 1$. Therefore the above theorem reduces to the following Kakeya and Enestrom theorem.

Theorem 4': If $a_n \geq a_{n-1} \geq \ldots \geq a_1 \geq a_0 \geq 0$, then all the zeroes of $f(z) = a_n + a_{n-1}z + \ldots + a_1z + a_0$ lie in the unit circle.

We now consider a special case, the trinomial equation $1 - z + \lambda z^n = 0, \lambda \neq 0$. We consider first the case $n=2$, i.e., $1 - z + \lambda z^2 = 0$. Setting $z = 1 + \xi$, we have

$$\lambda \xi^2 + (2\lambda - 1) \xi + \lambda = 0.$$ 

Let $\xi_1$ and $\xi_2$ be its roots. Then $\xi_1 \xi_2 = 1 \Rightarrow \xi_2 = \frac{1}{\xi_1}$ and $\xi_1$, $\xi_2$ are the roots and their product is unity. Hence these are inverse points with respect to the unit circle. Being inverse points both cannot lie in $|\xi| \leq 1$ i.e., $|z-1| \leq 1$. Thus the equation must have a zero in $|z-1| > 1$.

Now consider the case $n \geq 3$. Suppose, if possible, all the zeroes of $1 - z + \lambda z^n = 0$ lie in $|z-1| \leq 1$. Then by Gaus-Lucas theorem all the zeroes of the derived equation $\lambda_n z^{n-1} - 1 = 0$ must also lie in $|z-1| \leq 1$. But this is impossible for $n \geq 3$ (by the above argument). Hence the trinomial equation must have a root in $|z-1| > 1$. Still it remains to see that it has a root in $|z-1| \leq 1$.

Let $\xi = z - 1$. Therefore we are to show that $\lambda(\xi + 1)^n - \xi = 0$ has a root in $|\xi| \geq 1$. Change $\xi$ to $\frac{\xi}{1 - \xi}$. So we are to show that $\lambda(\xi + 1)^n - \xi^{n-1} = 0$ has a root in $|\xi| \geq 1$ or equivalently
\( \lambda z^n - (z-1)^{n-1} = 0 \) has a root in \( |z-1| \geq 1 \). To show this suppose 
\( \lambda z^n - (z-1)^{n-1} = 0 \) has all the roots in \( |z-1| < 1 \). Then again by Gauss-Lucas theorem the successive derived equations have all roots in \( |z-1| < 1 \). In particular \( 1 - z + \mu z^2 = 0 \) has all the roots in \( |z-1| < 1 \) whatever the coefficient of \( z^2 \) may be. But we have seen above for \( n=2 \) that the quadratic equation has a root in \( |z-1| \geq 1 \). Hence we have proved

**Theorem 5**: The trinomial equation \( 1 - z + \lambda z^n = 0 \) has a root in both the regions \( |z-1| \geq 1 \) and \( |z-1| \leq 1 \).

In fact, it has been proved that every general trinomial equation \( a_0 + a_1 z + a_n z^n = 0 \), \( a_1 a_n \neq 0 \), \( n \geq 2 \) has at least one zero in the circle \( |z| \leq 2 \left| \frac{a_0}{a_1} \right| \), and that every quadrinomial equation \( a_0 + a_1 z + a_m z^m + a_n z^n = 0 \), \( a_1 a_m a_n \neq 0 \), \( 2 \leq m < n \) has at least one zero in the circle \( |z| \leq \frac{17}{2} \left| \frac{a_0}{a_1} \right| \). For detailed study of this type see Marden [11].

Let us see the application of Perron-Frobenius theorem to find the upper bounds for the zeroes of the polynomial \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \). We need the following Lemma A and Lemma B. Their proofs can be found in Marden [11] and Gantmacher [7] respectively.

**Lemma 1** All the zeroes of the polynomial \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \), \( a_0 \neq 0 \) lie in the circle \( |z| \leq r \) where \( r \) is the real positive root of the equation \( g(z) = -z^n + |a_{n-1}| z^{n-1} + \ldots + |a_1| z + |a_0| \).
Lemma 2 Let \( A \) and \( B \) be square matrices of order \( n \) where \( A \) is irreducible and \( B^+ \leq A \). Then \( |\gamma| \leq r \) where \( \gamma \) is any eigenvalue of \( B \) and \( r \) is the largest eigenvalue of the matrix \( A \).

The companion matrix of \( f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \) is

\[
B = \begin{pmatrix}
-a_{n-1} & -a_{n-2} & \cdots & -a_0 \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 1 \\
0 & 0 & \cdots & 0 \\
\end{pmatrix}
\]

\[
B^+ = \begin{pmatrix}
|a_{n-1}| & |a_{n-2}| & \cdots & |a_0| \\
1 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & 1 \\
\end{pmatrix}
\]

By Lemma 1 the eigenvalues of \( B \) (zeroes of \( f(z) \)) lie in the circle \( |z| \leq r \) where \( r \) is the zero of \( g(z) \). Now \( a_0 \neq 0 \Rightarrow f(z) \) is irreducible \( \Rightarrow B^+ \) is irreducible. Also \( B \preceq B^+ \). Therefore by Lemma 2, the eigenvalues of matrix \( B \) are dominated by the largest eigenvalue of \( B^+ \) i.e., the real positive root of the equation \( h(z) = -|a_0|-|a_1|z \ldots -|a_{n-1}|z^{n-1}+z^n \). Now apply Perron-Frobenius theorem to the matrix \( B^+ \) we have
Theorem 6: The polynomial \( h(z) \) has a zero \( r \) which is real, simple, and exceeds in modulus all other eigenvalues of \( h(z) \).

If precisely \( p \) zeroes of \( h(z) \) have modulus \( r \), then each of these satisfies the polynomial

\[
z^p - r^p = 0 \quad \text{.... (1)}
\]

and then the set of all zeroes of \( h(z) \) is carried into itself by a rotation of the complex plane through an angle \( \frac{2\pi}{p} \).

The number \( r \) is given by

\[
r = \min \left\{ \max \left\{ \frac{[B^+X]_i}{X_i} \right\} : X_i \geq 0, 1 \leq i \leq n, X_i \neq 0 \} \right\} \quad \text{.... (2)}
\]

where \( X \) is an \( n \)-tuple of non-negative entries.

Now replace \( B^+ \) by its transpose, namely

\[
\begin{pmatrix}
|a_{n-1}| & 1 & 0 & \ldots & 0 \\
|a_{n-2}| & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
|a_1| & 0 & \ldots & 0 & 1 \\
|a_0| & 0 & \ldots & 0 & 0
\end{pmatrix}
\]

we have

Cor. The number \( r = \min \left\{ \max \left\{ \frac{|a_{n-i}|X_i}{X_i^2} + \frac{X_i + 1}{X_i} \right\} : X_i \geq 0, 1 \leq i \leq n, X_n + 1 = 0 \} \right\} \)

is the real positive dominated root of \( h(z) \). If \( z \) is a zero of \( f(z) \) then since all the zeroes of \( f(z) \) lie in \( |z| \leq r \), we have
\[ |z| \leq \max_{1 \leq i \leq n} \left( \frac{|a_{n-i}|}{X_i} X_1 + \frac{X_{i+1}}{X_i} \right), \quad X_{n+1} = 0 \quad \ldots \quad (3) \]

**Applications**

(a) Take \( X_i = 1, \ i = 1, 2, \ldots, n \) in (3), we have

\[ |z| \leq \max \{ 1 + |a_{n-1}|, 1 + |a_{n-2}|, \ldots, 1 + |a_1|, |a_0| \} \quad \ldots \quad (4) \]

which is Wilf's inequality proved on Page 23.

(b) Take \( X_i = |a_{n-i+1}| X_1, \ i = 1, 2, \ldots, n \) in (3), we get

\[ |z| \leq \max \{ \frac{|a_0|}{a_1}, 2 \frac{|a_1|}{a_2}, \ldots, 2 \frac{|a_{n-2}|}{a_{n-1}} \} \quad \ldots \quad (5) \]


(c) Take \( X_i = t^i, \ t > 0, \ i = 1, 2, \ldots, n \)

\[ |z| \leq \max \{ t + |a_{n-1}|, t + |a_{n-2}|, \ldots, t + |a_1|, |a_0| \} \quad \ldots \quad (6) \]

which is Wilf's inequality (21), proved on Page 26. (4) is a special case of (6).

From (6) we have \( |z| \leq t + \max \{ \frac{|a_{n-i}|}{t^i-1} \quad i = 1, 2, \ldots, n \} \).

Choose \( t \) such that two terms are equal. Therefore

\[ |z| \leq t + \max \{ \frac{1}{t^{n-i-1}} \} \leq 2 \max \{ \frac{1}{t^{n-i}}, i = 1, 2, \ldots, n \} \]

Chapter III
Vectorial Norms

The present chapter is devoted to the study of vectorial norms, dual of vectorial norms, regular vectorial norms, vectorial pseudo-norms, matrix norms subordinate to vector norms and matricial norms subordinate to vectorial norms. Recently these norms have been studied by Deutsch [3,4]. The main object is to discuss the different properties such as regularity and duality. Theorem 3 establishes the relation between two matricial norms, one generated from a vectorial norm and the second generated from a matrix norm.

Let $C^n$ denote the vector space of $n$-tuples over the complex numbers and $R^+_k$ be $k$-tuples of non-negative real numbers partially ordered component-wise.

**Definition:** A vectorial norm of order $k$ is a mapping $v : C^n \rightarrow R^+_k$ such that

(V.1) $v(\alpha x) = |\alpha|v(x)$ $\forall \alpha \in C$, $x \in C^n$.
(V.2) $v(x+y) \leq v(x) + v(y)$ $\forall x,y \in C^n$.
(V.3) $x \neq 0 \Rightarrow v(x) \neq 0$.

**Remark:** If $k=1$ then $v$ is a vector norm on $C^n$, so the notion of a vectorial norm is a generalization of vector norm.

**Remark:** A generalized matricial norm of order $k$ on $M_n$ is a vectorial norm of order $k^2$ on $C^{n^2}$.

A mapping $v : C^n \rightarrow R^+_k$ which satisfies (V.1) and (V.2) is
called a vectorial pseudo-norm of order $k$ on $C^n$. Let $v_1(x), \ldots, v_k(x)$ be the components of $v(x)$.

**Proposition 1:** $v$ is vectorial pseudo-norm iff the mapping $x \mapsto v_i(x)$ ($x \in C^n$) is a pseudo-norm on $C^n$ for each $i=1,2,\ldots,k$.

**Proof:** Let $v_i(x)$ be a pseudo-norm for each $i$. Define the mapping $v: C^n \to R^k_+$ such that

$$v(x) = (v_1(x), v_2(x), \ldots, v_k(x)) \quad \forall x \in C^n.$$ 

Then $v(x+y) \leq (v_1(x) + v_i(y), \ldots, v_k(x) + v_k(y))$

$$= (v_1(x), \ldots, v_k(x)) + (v_1(y), \ldots, v_k(y)) = v(x) + v(y).$$

$v(ax) = (v_1(ax), \ldots, v_k(ax)) = (|a|v_1(x), \ldots, |a|v_k(x)) = |a|v(x).$

Conversely, if $v$ is a vectorial pseudo-norm, then $v_i(x+y) \leq v_i(x) + v_i(y)$ and $v_i(ax) = |a|v_i(x)$, so $v_i$ is a vectorial pseudo-norm also.

Let $v: C^n \to R^k_+$ be vectorial norm on $C^n$ of order $k$. Define

$$K_j(v) = \{ x \in C^n | v_j(x) = 0 \}$$

$$K(v) = \bigcap_{i=1}^k K_i(v), \quad W_j(v) = \bigcap_{i \neq j} K_i(v)$$

$$W(v) = \sum_{i=1}^k W_i(v)$$

**Proposition 2:**

i) $K_j(v), K(v), W_j(v), W(v)$ are subspaces of $C^n$.

ii) $v$ is vectorial norm iff $K(v) = \{ 0 \}$.

**Proof:** i) since sum and intersection of subspaces is a subspace, it is enough to show that $K_j(v)$ is a subspace of $C^n$.

Let $x, y \in K_j(v)$. Then since $v_j$ is a pseudo-norm, we have $v_j(x+y) \leq v_j(x)$ and $v_j(y) = 0$ and $v_j(ax) = |a|v_j(x) = 0 \Rightarrow$
\[ ax + \beta y \in K_j(v) \Rightarrow K_j(v) \text{ is subspace of } C^n. \]

ii) Let \( v \) be vectorial norm. Then \( v(x) = 0 \Rightarrow v_i(x) = 0 \) \( \forall i = 1, 2, \ldots k \Rightarrow K_i(v) = 0 \Rightarrow K(v) = 0. \) On the other hand if \( v(x) \neq 0 \) then \( v_i(x) \neq 0 \) for at least one \( i. \) Therefore \( K(v) \neq 0. \)

**Definition:** A vectorial norm \( v \) is said to be a regular vectorial norm if \( C^n = W_1(v) \oplus \cdots \oplus W_k(v) \) and a vectorial pseudo-norm \( v \) is said to be a regular vectorial pseudo-norm if \( C^n = W_1(v) + \cdots + W_k(v). \)

The following example shows that not every regular vectorial pseudo-norm is a regular vectorial norm.

**Example 1:** Consider the mapping \( v : C^3 \rightarrow R^2 \) such that

\[ v(\alpha, \beta, \gamma) = (|\alpha|, |\gamma|) \quad \forall \alpha, \beta, \gamma \in C. \]

Then \( v \) is a vectorial pseudo-norm. \( v((\alpha, \beta, \gamma) + (x, y, z)) = v(\alpha + x, \beta + y, \gamma + z) = (|\alpha + x|, |\gamma + z|) \leq (|\alpha|, |\gamma|) + (|x|, |z|) = v(\alpha, \beta, \gamma) + v(x, y, z) \) and \( v(\lambda x, \lambda y, \lambda z) = (|\lambda x|, |\lambda z|) = |\lambda| v(x, y, z). \)

Now \( K_1(v) = \{ (\alpha, \beta, \gamma) | v_1(\alpha, \beta, \gamma) = 0 \} = \{ (0, \beta, \gamma) | \beta, \gamma \in C \} = W_2(v) \), and \( K_2(v) = \{ (\alpha, \beta, \gamma) | v_2(\alpha, \beta, \gamma) = 0 \} = \{ (\alpha, \beta, 0) | \alpha, \beta \in C \} = W_1(v). \)

Now \( C^3 = W_1(v) + W_2(v) \) and so \( v \) is a regular vectorial pseudo-norm. However \( v \) is not a regular vectorial norm since \( C^3 \neq W_1(v) \oplus W_2(v). \)

**Definition:** Two vectorial pseudo-norms \( u,v \) are said to be equivalent if \( W_j(v) = W_j(u) \quad \forall j = 1, 2, \ldots k. \)
Proposition 3: Let \( v \) be a vectorial pseudo-norm on \( \mathbb{C}^n \) and \( G \) an \( n \times n \) complex matrix. Then the mapping \( v_G : \mathbb{C}^n \rightarrow \mathbb{R}_+^k \) defined by
\[
v_G(x) = v(Gx) \quad \forall x \in \mathbb{C}^n
\]
is a vectorial pseudo-norm on \( \mathbb{C}^n \).

Remark: \( v_G \) is called the \( G \)-transform of \( v \). Note the similarity of \( v_G \) with \( v \).

Proof: \( v_G(ax) = v(Gax) = |a|v(Gx) = |a|v_G(x) \quad \forall x \in \mathbb{C}^n, a \in \mathbb{C} \). \( v_G(x+y) = v(Gx+Gy) \leq v(Gx)+v(Gy) = v_G(x) + v_G(y) \quad \forall x, y \in \mathbb{C}^n \).

Proposition 4: Let \( v : \mathbb{C}^n \rightarrow \mathbb{R}_+^k \) be a vectorial norm and \( G \) an invertible \( n \times n \) complex matrix. Then

1) \( v_G \) is vectorial norm.
2) \( K_j(v_G) = G^{-1}K_j(v) \quad \forall j = 1, 2, \ldots, k. \)
3) \( W_j(v_G) = G^{-1}W_j(v) \quad \forall j = 1, 2, \ldots, k. \)
4) \( W(v_G) = G^{-1}W(v) \).
5) \( v \) is a regular vectorial norm \( \Rightarrow v_G \) is a regular vectorial norm.

Proof: (1) \( x \neq 0 \Rightarrow Gx \neq 0 \Rightarrow v(Gx) \neq 0 \Rightarrow v_G(x) \neq 0 \). Therefore by Proposition 3, \( v_G \) is a vectorial norm.

(2) \( K_j(v_G) = \{x \in \mathbb{C}^n | v_{Gj}(x) = 0\} = \{x \in \mathbb{C}^n | v_j(Gx) = 0\} = \{x \in \mathbb{C}^n | Gx \in K_j(v)\} = G^{-1}K_j(v). \)

(3) \( W_j(v_G) = \bigcap_{i \neq j} K_i(v_G) = \bigcap_{i \neq j} G^{-1}K_i(v) = G^{-1} \bigcap_{i \neq j} K_i(v) = G^{-1}W_j(v). \)

(4) \( W(v_G) = \sum_{i=1}^{k} W_i(v_G) = \sum_{i=1}^{k} G^{-1}W_i(v) = G^{-1} \sum_{i=1}^{k} W_i(v) = G^{-1}W(v). \)
(5) \( v \) is a regular vectorial norm \( \Rightarrow W(v) = \sum_{i=1}^{k} \theta W_{i}(v) \Rightarrow \)
\[ G^{-1}W(v) = \sum_{i=1}^{k} \theta G^{-1}W_{i}(v) \Rightarrow \sum_{i=1}^{k} \theta W_{i}(v) = W(v) \] (using
(3) and (4) \( ) \Rightarrow v_{G} \) is a regular vectorial norm.

Proposition 5: Let \( p \) and \( q \) be vectorial pseudo-norms on \( C^{n} \).
Assume \( p \leq q \) (i.e., \( p(x) \leq q(x) \ \forall x \in C^{n} \)). Then

(i) \( p \) is a vectorial norm \( \Rightarrow q \) is a vectorial norm.

(ii) \( q \) is a regular vectorial pseudo-norm \( \Rightarrow p \) is a regular
vectorial pseudo-norm.

Proof: (i) Let \( x \in K_{j}(q) \Rightarrow q_{j}(x) = 0 \Rightarrow p_{j}(x) = 0 \) since
\( p \leq q \), therefore \( x \in K_{j}(p) \Rightarrow K_{j}(q) \subseteq K_{j}(p) \Rightarrow \bigcap_{j} K_{j}(q) \subseteq \bigcap_{j} K_{j}(p) \). Therefore by Proposition 2(ii) we
have \( K(q) = 0 \Rightarrow q \) is a vectorial norm.

(ii) \( x \in W_{j}(q) = \bigcap_{i \neq j} K_{i}(q) \Rightarrow q_{i}(x) = 0 \)
\( \forall i \neq j \Rightarrow p_{i}(x) = 0 \) since \( p \leq q \Rightarrow x \in \bigcap_{i \neq j} K_{i}(p) \Rightarrow x \in W_{j}(p) \)
\( \Rightarrow W_{j}(q) \subseteq W_{j}(p) \ \forall j = 1, 2, \ldots k. \) Since \( q \) is regular we have
\( p \) is regular.

Proposition 6: Let \( v: C^{n} \rightarrow R_{+}^{k} \) be a regular vectorial norm. Then
\[ K_{j}(v) = \sum_{i \neq j} \theta W_{i}(v) \ \forall i = 1, 2, \ldots k. \]

Proof: \( x \in K_{j}(v) \Rightarrow v_{j}(x) = 0, \ \forall x \in C^{n}. \) Since \( v \) is a
regular vectorial norm, therefore \( C^{n} = W(v) = W_{1}(v) \theta \ldots \theta W_{k}(v). \) Then \( x = x_{1} + x_{2} + \ldots x_{k}, \ x_{i} \in W_{i}(v) \Rightarrow v_{j}(x) = \)
\[ v_j(x_1 + x_2 + \ldots + x_k) = v_j(x_j) \quad \text{[since if } v(y) = 0 \text{ then } v(x+y) \leq v(x), \text{ on the other hand } v(x) = v(x+y-y) \leq v(x+y) \Rightarrow v(x+y) = v(x)] \]

Hence \( 0 = v_j(x) = v_j(x_j) \Rightarrow x_j = 0 \). Now \( x = x_1 + x_2 + \ldots \)

\[ x_{j-1} + x_j + x_{j+1} + \ldots + x_k = x_1 + \ldots + x_{j-1} + x_{j+1} + \ldots + x_k \in \sum_{i \neq j} W_i(v), \]

Since \( v \) is regular therefore \( x \in \sum_{i \neq j} \Theta W_i(v) \). Thus \( K_j(v) \subseteq \sum_{i \neq j} \Theta W_i(v) \). On the other hand we always have \( \sum_{i \neq j} \Theta W_i(v) \subseteq K_j(v) \)

\[ \Rightarrow K_j(v) = \sum_{i \neq j} \Theta W_i(v). \]

Remark: The converse of the above Proposition is false.

Consider the vectorial norm \( v : C^4 \to R^3 \) such that

\[ v(a_1,a_2,a_3,a_4) = (|a_1| + |a_2| + |a_3| + |a_4|) \]

It is routine to see that \( v \) so defined is a vectorial norm.

Now \( K_1(v) = \{ x \in C^n | v_1(x) = 0 \} \Rightarrow a_1 = a_4 = 0 \)

\[ \Rightarrow K_1(v) = \{(0,a_2,a_3,0) | a_2, a_3 \in C \}. \]

Similarly \( K_2(v) = \{(a_1,0,a_3,0) | a_1, a_3 \in C \} \)

\[ K_3(v) = \{(a_1,a_2,0,0) | a_1, a_2 \in C \}. \]

Now \( W_1(v) = K_2(v) \cap K_3(v) = (a_1,0,0,0) \)

\[ W_2(v) = K_1(v) \cap K_3(v) = (0,a_2,0,0) \]

\[ W_3(v) = K_1(v) \cap K_2(v) = (0,0,a_3,0). \]

Obviously \( K_1(v) = W_2(v) \Theta W_3(v), K_2(v) = W_1(v) \Theta W_3(v), K_3(v) = W_2(v) \Theta W_1(v). \) But \( v \) is not regular since \( C^4 \neq W_1(v) \Theta W_2(v) \Theta W_3(v). \)
Theorem 1: Let $p$ be a vectorial norm on $\mathbb{C}^n$. The following are equivalent:

1. $p$ is a regular vectorial norm.
2. There exists a norm $v$ on $\mathbb{C}^n$ and a direct-sum decomposition $\mathbb{C}^n = X_1 \oplus \ldots \oplus X_k$ with associated projections $E_1, E_2, \ldots, E_k$ such that $p(x) = (v(E_1 x), v(E_2 x), \ldots, v(E_k x))$ for all $x \in \mathbb{C}^n$.
3. If $p(x) = u + a$, $u, a \in \mathbb{R}_+$, then there exist vectors $y$ and $z \in \mathbb{C}^n$ such that $x = y + z$, $p(y) = u$ and $p(z) = a$.
4. $\mathbb{C}^n = K_1(p) \oplus \ldots \oplus K_k(p)$.

Proof: (1) $\Rightarrow$ (2). Suppose $p$ is a regular vectorial norm. Let $X_i = W_i(p)$. Therefore $\mathbb{C}^n = W_1(p) \oplus \ldots \oplus W_k(p)$. Let $E_1, \ldots, E_k$ be the projections associated with this direct-sum decomposition. Then $p_i(x) = p_i(E_1 x + E_2 x + \ldots + E_k x) = p_i(E_i x)$. Now consider the mapping $v$ on $\mathbb{C}^n$ defined by

$$x \mapsto v(x) = \max_i p_i(x).$$

Then it is easy to see that this mapping defines a norm on $\mathbb{C}^n$.

Now $v(E_j x) = \max_h \{p_h(E_j x)\} = \max_{h=1,2,\ldots,k} \{p_h(E_h E_j x)\} = p_j(E_j x) = p_j(x)$ for all $j = 1, 2, \ldots, k$. Hence $p(x) = (v(E_1 x), \ldots, v(E_k x))$.

(2) $\Rightarrow$ (3) Let $u = (u_1, u_2, \ldots, u_k)$, $a = (a_1, \ldots, a_k)$. Then since $p(x) = u + a$ we have $p_i(x) = u_i + a_i \Rightarrow (by (2)) v(E_i x) = u_i + a_i$. If $v(E_i x) = 0$ then $u = 0 = a$ so we are done.
If \( v(E_i x) \neq 0 \), then let \( v(E_i x) = w_i > 0 \). Set \( y_i = \frac{u_i}{w_i} E_i x \) and \( z_i = \frac{a_i}{w_i} E_i x \). Then \( y_i + z_i = \frac{1}{w_i} (u_i + a_i) E_i x \). Now since \( u_i + a_i = p_i(x) = v(E_i x) = w_i \) therefore \( y_i + z_i = E_i x \)

\[ \psi i = 1, 2, \ldots k \Rightarrow \sum_{i=1}^{k} y_i + \sum_{i=1}^{k} z_i = \sum_{i=1}^{k} E_i x \Rightarrow y + z = x. \]

Now \( y = \sum_{i=1}^{k} \frac{u_i}{w_i} E_i x \Rightarrow p_i(y) = \frac{p_i(u_i)p_i(E_i x)}{w_i} = \frac{p_i(u_i) v(E_i x)}{v(E_i x)} = p_i(u_i) \quad \psi i = 1, 2, \ldots k \Rightarrow p(y) = u. \)

Similarly \( p(z) = a. \)

(3) \( \Rightarrow \) (4) Let \( x \in C^n \) and \( u = (p_1(x), \ldots, p_{j-1}(x), 0, p_{j+1}(x), \ldots, p_k(x)) \) and \( a = p(x) - u = (0, 0, \ldots, p_j(x), 0 \ldots 0) \).

Now \( p(x) = u + a \Rightarrow \) by hypothesis that there exists \( y, z \in C^n \) such that \( x = y + z, p(y) = u, p(z) = a. \) Now \( u = p(y) \Rightarrow p_j(y) = u_j = 0 \Rightarrow y \in K_j(p). \) Also \( a = p(z) \Rightarrow p_i(z) = a_i = 0 \)

\( \psi i \neq j \Rightarrow z \in W_j(p). \) Thus \( x = y + z \in K_j(p) \) \( + W_j(p) \Rightarrow C^n \subseteq K_j(p) + W_j(p) \). But we always have \( K_j(p) + W_j(p) \subseteq C^n. \)

Therefore \( C^n = K_j(p) + W_j(p). \) Since \( W_j(p) = \bigcap_{i \neq j} K_i(p) \) therefore the sum is direct.

(4) \( \Rightarrow \) (1). \( \sum_{i \neq j} W_i(p) \subseteq K_j(p) \Rightarrow \sum_{i=1}^{k} W_i(p) \subseteq K_j(p) \) \( \cap W_j(p) \)

\[ C^n = \sum_{i=1}^{k} W_i(p) \subseteq C^n. \] On the other hand \( x \in C^n \Rightarrow x \in K_j(p) \)
Theorem 1 shows that all these mappings are exactly regular vectorial norms.

Proposition 7: Let \( p \) and \( q \) be regular vectorial norms of order \( k \) such that \( p \leq q \). Then

1. \( K_j(p) = K_j(q) \) \( \forall j = 1,2, \ldots k \)
2. \( p \) and \( q \) are equivalent, i.e., \( W_j(p) = W_j(q) \) \( \forall j = 1,2, \ldots k \).

Proof: Since \( p \) and \( q \) are regular vectorial norms therefore by Theorem 1 we have

\[
C^n = K_j(p) \oplus W_j(p) = K_j(q) \oplus W_j(q).
\]

It was proved in proposition 5 that \( K_j(q) \subseteq K_j(p) \) and \( W_j(q) \subseteq W_j(p) \). Since \( K_j(q) \oplus W_j(q) \) and \( K_j(p) \oplus W_j(p) \) have the same dimension, \( K_j(p) = K_j(q) \) and \( W_j(p) = W_j(q) \).

We now define the notion of the dual of a vectorial norm. Let \( p \) be vectorial norm on \( C^n \). Consider the mappings

\[ q_j: C^n \to R \text{ such that } \]

\[ q_j(y) = \sup_{x \neq 0} \frac{|y^*x|}{p_j(x)}, j = 1,2, \ldots k \] (\( y^* \) is adjoint of \( y \))

Further, define the mapping \( p^d: C^n \to R^k_+ \) such that

\[ p^d(y) = (q_1(y), \ldots, q_k(y)) \quad \forall y \in C^n. \]
Now we claim \( p^d \) so defined is a vectorial pseudo-norm on \( \mathbb{C}^n \).

\[
p^d_j(ay) = q_j(ay) = \sup_{x \neq 0} \frac{|(ay)^* x|}{p_j(x)} = \sup_{x \neq 0} \frac{|ay^* x|}{p_j(x)}
\]

\[
= |a| \sup_{x \neq 0} \frac{|y^* x|}{p_j(x)} = |a| p^d_j(y) \Rightarrow p^d(ay) = |a| p^d(y).
\]

Also \( p^d(y+z) = (q_1(y+z), \ldots, q_k(y+z)) \forall y, z \in \mathbb{C}^n \).

Now \( q_j(y+z) = \sup_{x \neq 0} \frac{|(y+z)^* x|}{p_j(x)} = \sup_{x \neq 0} \frac{|y^* x + z^* x|}{p_j(x)} \leq \sup_{x \neq 0} \frac{|y^* x|}{p_j(x)} + \sup_{x \neq 0} \frac{|z^* x|}{p_j(x)} = p^d_j(y) + p^d_j(z). \) Thus

\[
p^d(y+z) \leq p^d(y) + p^d(z) \Rightarrow p^d \text{ is a vectorial pseudo-norm on } \mathbb{C}^n.
\]

**Definition**: The mapping \( p^d \) defined above is called the dual of \( p \).

**Proposition 8**: Let \( p \) be a vectorial norm on \( \mathbb{C}^n \). Then

1. \( K_j(p^d) = (W_j(p)) \quad \forall j = 1, 2, \ldots, k \)
2. \( W_j(p^d) \supseteq (K_j(p)) \)
3. \( K(p^d) = (W(p)) \)
4. \( p^d \) is regular vectorial pseudo-norm.
5. \( p^d \) is vectorial norm iff \( p \) is regular vectorial norm.
Proof: (1) $K_j(p^d) = \{ y \in \mathbb{C}^n | p_j^d(y) = 0 \} = \{ y \in \mathbb{C}^n | q_j(y) = 0 \}$

= $\{ y \in \mathbb{C}^n | y^*x = 0 \} = (W_j(p))$.

(2) $W_j(p^d) = \bigcap_{i \neq j} K_i(p^d) = \bigcap_{i \neq j} (W_i(p)) = (\sum_{i \neq j} W_i(p))$.

Now since $\sum_{i \neq j} W_i(p) \subseteq K_j(p)$, $(\sum_{i \neq j} W_i(p)) \subseteq (K_j(p))$ therefore

$W_j(p^d) \subseteq (K_j(p))$.

(3) $K(p^d) = \bigcap_i K_i(p^d) = \bigcap_i (W_i(p)) = (\sum_i W_i(p)) = (W(p))$.

(4) We have shown already that $p^d$ is vectorial pseudo-norm. We show that $W(p^d) = \mathbb{C}^n$. Now $W(p^d) = \sum_i W_i(p^d) \supseteq \sum_i (K_i(p^d)) = (\bigcap_i K_i(p)) = \bigcap_i \{ x \in \mathbb{C}^n | p_i(x) = 0 \} = \mathbb{C}^n \Rightarrow W(p^d) \supseteq \mathbb{C}^n$.

But we always have $W(p^d) \subseteq \mathbb{C}^n$. Therefore $W(p^d) = \mathbb{C}^n$.

(5) Using Proposition (2) (ii), we have $p^d$ is vectorial norm iff $K(p^d) = 0$ iff $W(p) \downarrow = 0$ iff $W(p) = \mathbb{C}^n$ iff $p$ is regular vectorial norm. Proposition 9: Let $p$ and $q$ be vectorial norms on $\mathbb{C}^n$ such that $p \leq q$. Then $p^d \geq q^d$.

Proof: We have $p_j^d(y) = q_j(y) = \sup_{x \neq 0} \frac{|y^*x|}{p_j(x)} \forall y \in \mathbb{C}^n$. Since $p(x) \leq q(x)$ therefore $\frac{1}{p_j(x)} \geq \frac{1}{q_j(x)}$. Thus $p_j^d(y) \geq \sup_{x \neq 0} \frac{|y^*x|}{q_j(x)} = q_j^d(y) \Rightarrow p^d \geq q^d$. 
Proposition 10: Let $p$ be a regular vectorial norm on $\mathbb{C}^n$ such that $W_i(p) \perp W_j(p)$ $\forall i \neq j$. Then

a) $W_j(p^d) = (K_j(p))$

b) $K_j(p^d) = K_j(p)$

c) $W_j(p^d) = W_j(p)$ i.e., $p$ and $p^d$ are equivalent.

Proof: (a) Using Proposition 8 (ii) and Proposition 6, we have $W_j(p^d) \supseteq (K_j(p)) \supseteq \bigcap_{i \neq j} W_i(p)) = \bigcap_{i \neq j} K_i(p^d) = W_j(p^d) \Rightarrow W_j(p^d) = (K_j(p))$.

(b) By Proposition 6, we have $K_j(p) = \bigoplus_{i \neq j} W_i(p)$ and therefore by hypothesis we obtain $K_j(p) \perp W_j(p)$ and since by Theorem 1(iv) $\mathbb{C}^n = K_j(p) \oplus W_j(p)$ therefore $K_j(p) = W_j(p)$, $W_j(p) = K_j(p)$.

Now using Proposition 8(i) we get $K_j(p^d) = W_j(p) = K_j(p)$.

c) Using (a) we have $W_j(p^d) = (K_j(p)) = W_j(p)$.

Proposition 11: Let $p$ be a vectorial norm and $G$ an invertible $n \times n \text{ complex matrix}$. Then

(1) the dual of the $G$-transform of $p$ is the $(G^{-1})^*$ - transform of the dual of $p$.

(2) If $G$ is a unitary matrix, then the dual of the $G$-transform of $p$ is equal to the $G$-transform of the dual of $p$. 

Proof: (1) The dual of the G-transform of the jth component of p is given by \((p^d_G)_j(y) = \sup_{x \neq 0} \frac{|y^*x|}{(P_G)_j(x)} = \sup_{x \neq 0} \frac{|y^*x|}{p_j(Gx)}\) 

\[= \sup_{z \neq 0} \frac{|y^*G^{-1}z|}{p_j(z)} \] (where \(Gx = z\)),

\[= (p^d_G)_j(y) \] (since \((G^{-1}y)^* = y^*G^{-1})\),

\[= (p^d_G^{-1})_j(y). \] Therefore \(p^d_G = p^d_G^{-1}\) is the \(G^{-1}\) - transform of \(p^d\).

(2) Since \(G\) is a unitary matrix, therefore \(GG^* = 1 \Rightarrow G^* = G^{-1} \Rightarrow (G^{-1})^* = G\). Thus by (1) the dual of the G-transform of \(p\) is equal to the G-transform of the dual of \(p\).

Define the norm of a matrix \(A\) as the maximum of the norms \([5,113]\) of the vectors \(Ax\) where the vector \(x\) runs over the set of all vectors whose norm is 1 i.e., \(||A|| = \max \frac{||Ax||}{||x|| = 1}\). 

Let \(A \neq 0\). Then a vector \(x\) can be found with \(||x|| = 1\) such that \(Ax \neq 0 \Rightarrow ||Ax|| \neq 0\). Thus \(||A|| = \max \frac{||Ax||}{||x|| = 1}\).

Also \(||\alpha A|| = \max \frac{||\alpha Ax||}{||x|| = 1} = |\alpha| \max \frac{||Ax||}{||x|| = 1} = |\alpha|||A||.\)

Further \(||A+B|| = \max \frac{||(A+B)x||}{||x|| = 1} = \max \frac{||Ax+Bx||}{||x|| = 1} \leq \max \frac{||Ax|| + \max \frac{||Bx||}{||x|| = 1}}{||x|| = 1} = ||A|| + ||B||\),

and \(||AB|| = \max \frac{||ABx||}{||x|| = 1} \leq \max \frac{||AxBx||}{||x|| = 1}||x||=1 \neq\)
\[ \max \frac{||Ax||}{||x||} \cdot \max \frac{||Bx||}{||x||} = \frac{||A||}{||B||}. \]

**Definition**: The norm of a matrix constructed in this way is called the matrix norm sub-ordinate to the vector norm and is written as \( \text{lub}_v \) (least upper bound of vector norm \( v \)).

Next we introduce the idea of a matricial norm sub-ordinate to a vectorial norm which is a generalization of matrix norm subordinate to a vector norm. Let \( p \) be a regular vectorial norm on \( \mathbb{C}^n \). Define the mapping \( \text{lub}_p : M_n \to M^n_+ \) such that

\[ \text{lub}_p(A) = (m_{ij}(A))_{i,j = 1,2, \ldots k} \quad \forall A \in M_n \]

where \( m_{ij}(A) = \sup_{x \neq 0} \frac{p_i(Ax)}{x \in W_j(p)} \).

Then \( \text{lub}_p(A) \) is a regular matricial norm on \( M_n \).

Let \( A \neq 0 \). Then since \( x \neq 0 \) we have \( Ax \neq 0 \Rightarrow m_{ij}(A) \neq 0 \)

\( \forall i,j = 1,2, \ldots k \Rightarrow \text{lub}_p(A) \neq 0 \) which satisfies (M·1).

Let \( a \in \mathbb{C} \). Therefore \( m_{ij}(aA) = |a|m_{ij}(A) \) for all \( i,j = 1,2, \ldots k \).

Hence \( \text{lub}_p(aA) = |a| \text{lub}_p(A) \) which verifies (M·2).

To see (M·3) and (M·4) we have \( \forall A,B, \in M_n \)

\[ \text{lub}_p(A+B) = m_{ij}(A+B)_{i,j = 1,2, \ldots k} \]
\[
\begin{align*}
&= \sup_{x \neq 0, x \in W_j(p)} \frac{p_i((A+B)x)}{p_j(x)} \leq \sup_{x \neq 0, x \in W_j(p)} \frac{p_i(Ax)}{p_j(x)} + \sup_{x \neq 0, x \in W_j(p)} \frac{p_i(Bx)}{p_j(x)} \\
&= \sup_{x \neq 0, x \in W_j(p)} \frac{p_i(Ax)}{p_j(x)} + \sup_{x \neq 0, x \in W_j(p)} \frac{p_i(Bx)}{p_j(x)} = \operatorname{lub}_p(A) + \operatorname{lub}_p(B).
\end{align*}
\]

Where \( \operatorname{lub}_p(AB) = -(m_{ij}(AB)) \) for \( i, j = 1, 2, \ldots, k \) is given by:

\[
\sup_{x \neq 0, x \in W_j(p)} \frac{p_i((AB)x)}{p_j(x)}.
\]

Since \( p \) is a regular vectorial norm therefore \( \operatorname{lub}_p(A) \) is a regular matricial norm.

**Definition:** The mapping \( \operatorname{lub}_p \) (least upper bound of \( p \)) defined above is called the matricial norm subordinate to the vectorial norm \( p \).

**Example 2:** Consider the vectorial norm \( p: C^3 \rightarrow R^2 \) defined by \( p(x_1, x_2, x_3) = (|x_1|, |x_2| + |x_3|) \) for all \((x_1, x_2, x_3) \in C^3\).

It is easy to see that \( K_2(p) = W_1(p), K_1(p) = W_2(p) \) and \( C^3 = W_1(p) \odot W_2(p) \). Therefore \( p \) is a regular vectorial norm.

Define the mapping \( \operatorname{lub}_p: M_3 \rightarrow M_2^+ \) such that

\[
\operatorname{lub}_p(A) = \begin{bmatrix} m_{11}(A) & m_{12}(A) \\ m_{21}(A) & m_{22}(A) \end{bmatrix}
\]

where \( m_{ij} = \sup_{x \neq 0, x \in W_j(p)} \frac{p_i(Ax)}{p_j(x)} \)
Now \( p_1(x) = |x_1| \) and \( p_1(Ax) = |a_{11}| |x_1| \). Therefore
\[
m_{11}(A) = |a_{11}|.
\]
Similarly \( m_{12}(A) = |a_{12}| + |a_{13}| \), \( m_{21} = |a_{21}| + |a_{31}| \), \( m_{22}(A) = |a_{22}| + |a_{23}| + |a_{32}| + |a_{33}| \).

Hence \( \text{lub}_p(A) = \begin{pmatrix} |a_{11}| & |a_{12}| + |a_{13}| \\ |a_{21}| + |a_{31}| & |a_{22}| + |a_{23}| + |a_{32}| + |a_{33}| \end{pmatrix} \).

**Theorem 2:** Let \( p: \mathbb{C}^n \to R_+^k \) be a regular vectorial norm on \( \mathbb{C}^n \).

Let \( G \) be invertible \( n \times n \) complex matrix. Then the matricial norm subordinate to the \( G \)-transform of \( p \) is equal to the \( G \)-transform of \( \text{lub}_p \). In other words \( \text{lub}_{p_G} = (\text{lub}_p)_G \).

**Proof:** Set \( p_G = q \). Since \( p \) is a regular vectorial norm therefore by Proposition 4 (v) \( q \) is a regular vectorial norm and so
\[
\mathbb{C}^n = W_1(p) \oplus W_2(p) \oplus \ldots \oplus W_k(p)
\]
\[
\mathbb{C}^n = W_1(q) \oplus W_2(q) \oplus \ldots \oplus W_k(q).
\]
Let \( E_1, E_2, \ldots, E_k \) and \( F_1, F_2, \ldots, F_k \) be the associated projections of the above decompositions. Again by Proposition 4(iii) \( W_j(q) = G^{-1}W_j(p) \). Now \( \forall x \in \mathbb{C}^n \), we have
\[
F_j(x) \in W_j(q) = G^{-1}W_j(p).
\]
Also \( E_j \) \( G \times \in W_j(p) \) therefore \( G^{-1} E_j, Gx \in G^{-1} W_j(p) \) and since the above sum is direct, therefore
\[
G^{-1} E_j G = F_j \quad \forall j = 1, 2, \ldots, k.
\]
Now \((\text{lub}_{qA})_{ij} = (i, j)\) element of matrix \( \text{lub}_{qA} \)
\[
= \sup_{x \in W_j(q)} q_i(Ax) = \sup_{x \in G^{-1}W_j(p)} q_i(F_jAx)
\]
(since \( W_j(q) = G^{-1}W_j(p) \) and \( q_i(y) = q_i(F_iy) \)),
\[
= \sup_{x \in G^{-1}W_j(p)} \frac{q_i(F_iAx)}{q_j(x)} \quad (\text{since} \ \forall x \in W_j(q), F_j(x) = x),
\]
\[
= \sup_{x \in G^{-1}W_j(p)} \frac{q_i(G^{-1}E_jG^{-1}E_jGx)}{q_j(G^{-1}E_jGx)} \quad (\text{since} \ F_j = G^{-1} E_j G),
\]
\[
= \sup_{y \in W_j(p)} \frac{p_i(E_iG^{-1}E_jy)}{p_j(E_jy)} \quad (\text{where} \ y = Gx, p_i(Gx) = q_i(x)
\text{i.e.,} \ q_i(G^{-1}x) = p_i(x)),
\]
\[
= \sup_{y \in W_j(p)} \frac{p_i(E_iG^{-1}y)}{p_j(y)} \quad (\text{since} \ \forall y \in W_j(p), E_j(y) = y),
\]
\[
= \sup_{y \in W_j(p)} \frac{p_i(G^{-1}y)}{p_j(y)} \quad (\text{since} \ p_i(E_iy) = p_i(y)),
\]
Thus $(\text{lub}_p A) = (\text{lub}_p GAG^{-1}) = (\text{lub}_p G)$ and the proof is complete.

Example 3: Consider $p$ the vectorial norm, and $\text{lub}_p$ the matricial norm subordinate to $p$, given by the example 2.

$$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Take $G = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 3 \\ 0 & 0 & 3 \end{pmatrix}$.

Now the $G$-transform of $p$ is given by

$$p_G(x_1, x_2, x_3) = (3|x_1|, 2x_1 + x_2, x_3)$$

Now $W_1(p_G) = G^{-1}W_1(p) = (x_1, 0, 0)$ and $W_2(p_G) = G^{-1}W_2(p) = (0, x_2, x_3)$. Since $(x_1, x_2, x_3) = (x_1, 0, 0) \oplus (0, x_2, x_3)$ therefore $p_G$ is a regular vectorial norm.

Now by the above theorem $\text{lub}_{p_G}(A) = \text{lub}_{p_G}(GAG^{-1}) = \begin{pmatrix} m_{11}(GAG^{-1}) & m_{12}(GAG^{-1}) \\ m_{21}(GAG^{-1}) & m_{22}(GAG^{-1}) \end{pmatrix}$ .... (1)
\[ GAG^{-1} = \frac{1}{3} \begin{bmatrix} 3a_{11} - 6a_{12} & 9a_{12} & 9a_{13} \\ 2a_{11} + a_{21} - 4a_{12} - 2a_{22} & 6a_{12} + 3a_{22} \\ a_{31} - 2a_{32} & 3a_{32} & 3a_{33} \end{bmatrix} \]

Now \( m_1(GAG^{-1}) = \sup_{x \neq 0, x \in W_1(p)} \frac{p_1(GAG^{-1}x)}{p_1(x)} \frac{1}{3} = \frac{1}{3} \mid 3a_{11} - 6a_{12} \mid = \mid a_{11} - 2a_{12} \mid \) since \( x \in W_1(p) \).

Similarly \( m_2(GAG^{-1}) = \frac{1}{3} \{ \mid 2a_{11} + a_{21} - 4a_{12} - 2a_{22} \mid + \mid a_{31} - 2a_{32} \mid \} \)

\[ m_1(GAG^{-1}) = 3 \max \{ \mid a_{12} \mid, \mid a_{13} \mid \} \]

\[ m_2(GAG^{-1}) = \max \{ \mid 2a_{12} + a_{22} \mid + \mid a_{32} \mid, \mid 2a_{13} + a_{23} \mid + \mid a_{33} \mid \} \]

Substituting these values in (1) we get a matricial norm subordinate to \( p_G \).

We have seen that starting with a vector norm \( v \) on \( \mathbb{C}^n \) and the given direct-sum decomposition \( \mathbb{C}^n = X_1 \oplus \ldots \oplus X_k \),
we can now generate two matricial norms of order \( k \) on \( M_n \) as follows:

(1) Given the vectorial norm \( p \) on \( \mathbb{C}^n \) induced by \( v \) and the given direct-sum decomposition, we form the matricial norm \( \text{lub}_p \) subordinate to \( p \).

(2) Given the matrix norm \( \text{lub}_v \) subordinate to \( v \), we form the matricial norm \( \mu \) induced by \( \text{lub}_v \) and the given decomposition of \( \mathbb{C}^n \).

The following theorem gives the relationship between these two matricial norms.
Theorem 3: Let $E_1, E_2, \ldots, E_k$ be the projections associated with the decomposition $C^n = X_1 \oplus \ldots \oplus X_k$ of $C^n$. Let $v: C^n \rightarrow \mathbb{R}_+$ be a vector norm on $C^n$ and $p: C^n \rightarrow \mathbb{R}_+^k$ be the vectorial norm induced by $\{X_1, X_2, \ldots, X_k\}$ and $v$. Further, let $\mu: M_n \rightarrow \mathbb{R}$ be the matricial norm induced by $\{X_1, X_2, \ldots, X_k\}$ and matrix norm $\text{lub}_v$. Then

(i) $\mu(AB) \leq (\text{lub}_p A)\mu(B)$ $\quad \forall A, B \in M_n$

(ii) $(\text{lub}_p A) \leq \mu(A) \leq (\text{lub}_p A)\mu(I_n)$ $\quad \forall A \in M_n$

(iii) If $\text{lub}_v E_j = 1$ $\quad \forall j = 1, 2, \ldots, k$ then $\mu = \text{lub}_p$ i.e.,

$$\text{lub}_p (A) = \mu(A).$$

Proof: "(1) We have $(\mu(AB))_{ij} = (\mu(AE_iB + \ldots + AE_kB))_{ij}$

$$\leq (\mu(AE_1B))_{ij} + \ldots + (\mu(AE_kB))_{ij} \quad \cdots (1)$$

Now by definition $(\mu(A))_{ij} = \text{lub}_v (E_iAE_j) = \sup_{x \neq 0} v(E_iAE_j x) x \in C^n \ \Rightarrow (\mu(A)_{ij} \quad \cdots (2)$$

Since $E_j$ is identity map for $X_j$ therefore $E_j(x) = x \quad \forall x \in X_j$.

Also since $p$ is the induced vectorial norm therefore $p_i(x) = v(E_i x)$ and so $p_i(Ax) = v(E_iAx) = v(E_iAE_j x)$.

Hence $(\text{lub}_p (A))_{ij} = \sup_{x \neq 0} \frac{p_i(Ax)}{p_j(x)} = \sup_{x \neq 0} \frac{v(E_iAE_j x)}{v(x)} \quad \cdots (3)$
Let $0 \neq x \in X_m$ where $1 \leq m \leq k$. Therefore from (2) and (3) we have

$$\left( \text{lub}_p A \right)_{im} (\mu(B))_{mj} \geq \frac{v(E_iAE_m x)}{v(x)} \cdot \frac{v(E_mE_jz)}{v(z)} \quad \ldots \ldots \quad (4)$$

Now choose $z \in C^n$ such that $E_mE_jz \neq 0$ and set $x = E_mE_jz$.

Then from (4)

$$\frac{(\text{lub}_p A)_{im} (\mu(B))_{mj}}{v(z)} \geq \frac{v(E_iAE_m x)}{v(z)} = \frac{v(E_iAE_mE_jz)}{v(z)} \quad \ldots \ldots \quad (5)$$

Now replace $A$ by $AE_mE_jB$ in (2). We have from (5) since $z \in C^n$

$$(\text{lub}_p A)_{im} (\mu(B))_{mj} \geq (\mu(AE_mB))_{ij}.$$ 

Hence from (1) we have

$$v(AB)_{ij} \leq (\text{lub}_p A)_{im} (\mu(B))_{mj} \ldots (\text{lub}_p A)_{ik} (\mu(B))_{kj} \Rightarrow$$

$$v(AB) \leq (\text{lub}_p A) \mu(B) \quad \forall A, B \in M_n.$$ 

(ii) From (3) and (2) we have

$$(\text{lub}_p A)_{ij} \leq (\mu(A))_{ij} \Rightarrow (\text{lub}_p A) \leq (\mu(A)).$$

Now in (i) let $B = I_n$. Therefore $\mu(A) \leq (\text{lub}_p A) \mu(I_n).$

Hence $\mu(A) \leq (\text{lub}_p A) \mu(I_n) \quad \forall A \in M_n.$

(iii) Since

$$\mu(I_n) = \begin{bmatrix}
\text{lub}_v E_1 & 0 & 0 & \ldots & 0 \\
0 & \text{lub}_v E_2 & 0 & \ldots & 0 \\
& & & \cdots & \\
& & & & \\
0 & 0 & \ldots & \ldots & \text{lub}_v E_k
\end{bmatrix}.$$
and we are given that \( \text{lub}_v E_j = 1 \) \( \forall j = 1, 2, \ldots, k \) therefore 
\( \nu(I_n) = I_k \). Hence from (ii) we have 
\[
(\text{lub}_p A) \leq \nu(A) \leq (\text{lub}_p A) \Rightarrow \text{lub}_p (A) = \nu(A) \quad \forall A \in M_n.
\]

The following example shows that the inequalities (i) and (ii) can be strict.

Example 4: Let \( C^3 = X_1 \oplus X_2 \) where 
\[
X_1 = \{ (a_1, -2a_2, 0) | a \in C \},
X_2 = \{ (0, \beta, \gamma) | \beta, \gamma \in C \}.
\]
Consider the vector norm \( v : C^3 \to \mathbb{R}_+ \) defined by 
\[
v(a_1, a_2, a_3) = |a_1| + |a_2| + |a_3|.
\]

The vectorial norm \( p \) induced by \( \{X_1, X_2\} \) and \( v \) is (Theorem 1) 
\[
p(x) = (v(E_1 x), v(E_2 x)) \quad \forall x \in C^3.
\]
Now \( E_1(x) = (a, -2a, 0) \) therefore \( v(E_1 x) = 3|a_1| \). Similarly 
\[
v(E_2 x) = |2a_1 + a_2| + |a_3|.
\]
Thus \( p(x) = (3|a_1|, |2a_1 + a_2| + |a_3|) \) where \( p : C^3 \to \mathbb{R}_+^2 \).

Now since this vectorial norm \( p \) is the same as \( p_G \) of example 3 therefore the matricial norm \( \text{lub}_p \) subordinate to \( p \) is given by
\[
\text{lub}_p (A) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}
\]
where 
\[
m_{11} = |a_{11} - 2a_{12}|,
m_{12} = 3 \max \{ |a_{12}|, |a_{13}| \},
m_{21} = \frac{1}{3} \left( |2a_{11} + a_{21} - 4a_{12} - 2a_{22}| + |a_{31} - 2a_{32}| \right),
m_{22} = \max \{ |2a_{12} + a_{22}| + |a_{32}|, |2a_{13} + a_{23}| + |a_{33}| \}.
\]
Now it is proved [5, P.109] that \( \text{lub}_v(A) = \max_j \sum_{i=1}^{3} |a_{ij}| \)

where \( A = (a_{ij}) \in M_3 \) is the matrix norm \( \text{lub}_v \) subordinate to \( v \).

Thus the matricial norm \( \mu : M_3 \to M_2 \) induced by \( \{X_1, X_2\} \) and \( \text{lub}_v \) is given by

\[
\mu(A) = (\text{lub}_v(E_j A E_j))_{ij=1,2} \quad \forall A \in M_3.
\]

Let us evaluate the matrix \( \mu(A) \). Now \( E_1 x = (a_1, -2a_1, 0) \)

therefore \( AE_1 = \begin{pmatrix} a_{11}a_1 - 2a_{12}a_1 \\ a_{21}a_1 - 2a_{22}a_1 \\ a_{31}a_1 - 2a_{32}a_1 \end{pmatrix} \)

\( (a_{11} - 2a_{12}, -2(a_{11} - 2a_{12}), 0) \).

Thus \( \text{lub}_v E_1 AE_1 = 3|a_{11} - 2a_{12}| \geq m_{11} \).

Similarly since \( E_2 AE_1 = (0, 2(a_{11} - 2a_{12}) + a_{21} - 2a_{22}, a_{31} - 2a_{32}) \),

therefore \( \text{lub}_v E_2 AE_1 = |2a_{11} + a_{21} - 2a_{22} - 4a_{12}| + |a_{31} - 2a_{32}| \geq m_{21} \)

Similarly \( \text{lub}_v E_1 AE_2 = 3 \max(2|a_{12}|, |a_{13}|) \geq m_{12} \)

\[
\text{lub}_v E_2 AE_2 = \max(2|2a_{12} + a_{22}| + 2|a_{32}|, |2a_{13} + a_{23}| + |a_{33}|) \geq m_{22}.
\]

Obviously \( \mu(A) > \text{lub}_p(A) \).

Remark: From (iii) of Theorem 3 we see that when \( \text{lub}_v E_j = 1 \)

\( \forall j=1,2, \ldots k \) then \( \mu = \text{lub}_p \), that is the matricial norm \( \mu \)
generated by the matrix norm \( \text{lub}_v \) subordinate to vector norm

(Proposition 3, Chapter 1) is equal to the matricial norm \( \text{lub}_p \)
generated by the vectorial norm \( p \), where \( p \) itself is generated
from vector norm \( v \) (Theorem 1). In fact it is easy to see that the following diagram is commutative

\[
\begin{array}{c}
\text{\( v \)} \\
\downarrow \\
\text{\( \text{lub}_v \)}
\end{array}
\quad \begin{array}{c}
\text{\( p \)} \\
\downarrow \\
\text{\( \mu = \text{lub}_p \)}
\end{array}
\]

**Example 5:** Let \( C^n = X_1 \oplus X_2 \oplus \ldots \oplus X_k \) be an orthogonal direct-sum decomposition of \( C^n \) and let \( E_1, E_2, \ldots, E_k \) be the associated projections. Further let \( \| \| \) be the euclidean norm on \( C^n \) i.e., \( \| x \|^2 = \sum_{i=1}^{n} |x_i|^2 \forall x \in C^n \). Then the vectorial norm on \( C^n \) induced by \( \{X_1, X_2, \ldots, X_k\} \) and the euclidean norm \( \| \| \) is given by (Theorem 1) \( p: C^n \to \mathbb{R}_+^k \) such that

\[
p(x) = (\| E_1 x \|, \ldots, \| E_k x \|) \forall x \in C^n.
\]

Now it is proved [5,P.109] that the matrix norm \( \| \| \) subordinate to the euclidean norm \( \| \| \) is given by

\[
\| \| A \| \| = \sqrt{r(A^*A)} \forall A \in M_n \] where \( A^* \) is the adjoint of matrix \( A \) and \( r(B) \) is the spectral radius of matrix \( B \). Since \( \{X_1, X_2, \ldots, X_k\} \) is an orthogonal decomposition of
C^n therefore $E_j^* = E_j \quad \forall j = 1, 2, \ldots, k$, and so $E_j^* E_j = E_j E_j^* = E_j^* = E_j \quad \forall j = 1, 2, \ldots, k$. Hence the matricial norm subordinate to the vectorial norm $\| \cdot \|$ is (by definition)

$$\text{lub}_p A = (\| |E_i A E_j| |)_{i,j = 1, 2, \ldots, k} = \left( \sqrt{r((E_i A E_j)^* E_i A E_j)} \right)_{i,j = 1, 2, \ldots, k} = \left( \sqrt{r(E_i^* A E_i E_j)} \right)_{i,j = 1, 2, \ldots, k} \quad \forall A \in M_n.$$


