CS 75

THEORY OF NORMS

BY

F. L. BAUER

TECHNICAL REPORT NO. CS 75
AUGUST 31, 1967

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

THEORY OF NORMS*

BY

F. L. Bauer

These notes are based on lectures given during the winter of 1967 as CS 233, Computer Science Department, Stanford University, The author expresses his warmly felt thanks to Mr. Stanley Eisenstat, who aided him in preparing these notes.

*Supported in part by a grant from the National Science Foundation.

Table of Contents

§1. Vector Spaces

We summarize some basic concepts from linear algebra. A

VECTOR SPACE $\mathcal{V}$ OVER A FIELD K (1.1)

(denoted $\mathcal{V}_K$) is an abelian group $(\mathcal{V},+)$ --the zero element being denoted by ϕ or simply 0 --with K as a multiplier field; i.e., with a mapping $K \times \mathcal{V} \to \mathcal{V}$ (SCALAR MULTIPLICATION) satisfying

$$\forall \alpha,\beta \in K;\ x,y \in \mathcal{V}: \quad \begin{aligned} \alpha(x+y) &= \alpha\ x + \alpha\ y \\ (\alpha + \beta)x &= a\ x + \beta\ y \\ (\alpha\ \beta)x &= \alpha(\beta\ x) \\ 1x &= x \end{aligned}$$

$x \in \mathcal{V}$ is called a

VECTOR. (1.2)

$\alpha \in K$ is called a

SCALAR. (1.3)

Examples:

(i) Let (If,+) be the additive part of a field K . Then $\mathcal{V}_K$ is a vector space over K with multiplication in K as a scalar multiplication.

(ii) Let $\mathcal{V} = K^n$ be the n-fold direct product of K; i.e., the set of all ordered n-tuples of elements of K . We may write them columnwise.

$$\mathcal{V} = K^n = \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} : \alpha_i \in K,\ i = 1, 2, \ldots . n \right\}$$

Define + over $\mathcal{V}$ componentwise in the sense of K:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}$$

Then $(\mathcal{V},+)$ is an **abelian** group with the zero element $\phi = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Let scalar multiplication be defined by components in the sense of K:

$$\alpha \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \vdots \\ \alpha\alpha_n \end{pmatrix} .$$

Then $\mathcal{V}_K$ is a vector space over K . It will be denoted K^n . Its elements in the representation given above are called <u>COLUMN VECTORS</u> .

(iii) Let C[0,1] be the set of all real-valued functions defined and continuous on the closed interval [0,1] . For $f_1, f_2 \in C[0,1]$, define $f = f_1 + f_2$ by

$$f(S) = f_1(\xi) - f_2(\xi) \qquad \forall\ \xi \in [0,1]$$

Then C[0,1] is an **abelian** group with zero element ϕ : $\phi\ (\xi) \equiv 0$. For $g \in C[0,1]$, $\alpha \in R$, define $f = \alpha\ g$ by

$$f(\xi) = a\ g(\xi) \qquad \forall\ \xi \in [0,1]$$

Then C[0,1] is a vector space over R .

[Note: The sum and multiples of continuous functions are continuous].

Example (ii) and Example (iii) are special cases of vector spaces obtained from a field K by forming ordered sets of elements, ordered according to some index set (the set (1, 2, ... n) of natural numbers and the set [0,1] of real numbers respectively). In Example (iii) moreover, an additional property is postulated (continuity) which is hereditary under the componentwise operations.

A subset $\mathcal{V}_1$ of a vector space $\mathcal{V}_K$ is called a

<u>SUBSPACE</u> <u>OF</u> $\mathcal{V}_K$ (1.4)

if it is a vector space over K; i.e., if

$$x,y \in \mathcal{V}_1 > x + y \in \mathcal{V}_1$$

$$\forall\ \alpha \in K:\ x \in \mathcal{V}_1 > \alpha x \in \mathcal{V}_1$$

or equivalently,

$$\forall\ \alpha,\beta \in K:\ x,y \in \mathcal{V}_1 > \circ\, x + \beta\, \circ \in \mathcal{V}_1\ .$$

$\alpha x + \beta y$ is called a

<u>LINEAR COMBINATION</u> (1.5)

of x and y .

A subset M of a vector space $\mathcal{V}_K$ is called a

<u>K-BASIS</u> <u>OF</u> $\mathcal{V}_K$ or simply a <u>BASIS</u> <u>OF</u> $\mathcal{V}_K$ (1.6)

if any $x \in \mathcal{V}_K$ is uniquely determined by some (finite!) linear combination ($x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m$) of elements of M ($x_i \in M$) . For Example (ii), the axis vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

are a basis for K^n.

If $\mathcal{V}_K$ has a

<u>FINITE BASIS,</u> (1.7)

i.e., a basis formed by a finite number n of elements, then every basis has n elements and these elements are K-l<u>inearly independent:</u>

$$\alpha_i \in K, \ \alpha_1 x_1 + \ldots + \alpha_n x_n = 0 \succ \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0 .$$

n is called the

<u>DIMENSION OF</u> $\mathcal{V}_K$ (1.8)

(denoted $\dim(\mathcal{V}_K)$) and $\mathcal{V}_K$ is isomorphic to K^n.

In Example (ii), the dimension of K^n is n. We call K^n an

<u>n-DIMENSIONAL COORDINATE SPACE.</u> (1.9)

In particular, we shall consider R^n and C^n, where R is the real field and C the complex field.

§2. Normed Vector Spaces

A set M is

$$\text{ORDERED BY A RELATION } \rho \text{ or p-ORDERED} \tag{2.1}$$

if there is a relation ρ over M X M, an ORDERING of M, with the properties:

TRANSITIVITY: $x \rho y \wedge y \rho z \succ x \rho z \ \forall\, x,y,z \in M$ (2.2)

REFLEXIVITY: $x \rho x \ \forall\, x \in M$. (2.3)

ANTISYMMETRY: $x \rho y \wedge y \rho x \succ x = y \ \forall\, x,y \in M$ (2.4)

An abelian group G = (M,+) is a

$$\text{p-ORDERED GROUP} \tag{2.5}$$

if the ordering ρ of M is compatible with the group composition; i.e., if

$$a \rho b \succ a + x \ \rho \ b+x \quad \forall\, a,b,x \in M \ . \tag{2.6}$$

An element of G is

$$\text{NON-NEGATIVE} \tag{2.7}$$

if Opx. An ordering ρ of M is

$$\text{LINEAR} \tag{2.8}$$

if it has the property

$$x \rho y \vee ypx \quad \forall\, x,y \in M \ .$$

If the ordering of an ordered group G is linear, then G is a

<u>LINEARLY ORDERED GROUP</u> . (2.10)

Examples:

(i) The family of all subsets of a given set is ordered, the ordering being set inclusion ⊂ : $X \subset Y :\Leftrightarrow p \in X \Rightarrow p \in Y$. It is not linearly ordered.

(ii) The set of natural numbers has a linear ordering, usually denoted by $\leq$.

(iii) The additive parts of the ring of integers Z, the rational field P, and the real field R are linearly ordered abelian groups for the ordering usually denoted by $\leq$.

(iv) Let K be a field, the additive part of which has a linear ordering < (e.g., P or R) . Then

$$x \rho y :\Leftrightarrow x_i \leq y_i, \quad i = 1, 2, \ldots, n \qquad (2.11)$$

defines an ordering ρ in K^n (<u>COMPONENTWISE ORDERING</u>); The additive part of K^n is a ρ-ordered abelian group. For $n > 1$, however, the ordering ρ is not linear. Nevertheless, we shall use the conventional sign $\leq$ to denote this ordering; i.e.,

$$x \leq y :\Leftrightarrow x_i \leq y_i, \quad i = 1, 2, \ldots, n \; ; \; x,y \in K^n \qquad (2.12)$$

In accordance with standard practice, we shall use < to denote strict inequality; i.e.,

$$x < y :\Leftrightarrow x_i < y_i, \quad i = 1, 2, \ldots, n \; ; \; x,y \in K^n \qquad (2.13)$$

Note that $x \leq y$ and $x \neq y$ together is weaker than $x < y$.

Furthermore, we shall denote by $|x|$ the vector whose components are the absolute values of the components of x :

$$|x| := \begin{pmatrix} |x_1| \\ |x_2| \\ \vdots \\ |x_n| \end{pmatrix}, \quad x \in K^n \quad . \tag{2.14}$$

A functional over a vector space $\mathcal{V}$ with values from a p-ordered abelian group G, i.e., a mapping $v : \mathcal{V} \to G$, is a

NORM (2.15)

if it is

SUBADDITIVE: $\nu(x+y) \;\rho\; \nu(x) + \nu(y) \;\; \forall x,y \in \mathcal{V}$ (2.16)

NON-NEGATIVE: $0 \;\rho\; \nu(x) \;\; \forall\, x \in \mathcal{V}$ (2.17)

DEFINITE: $x = \phi \;\Leftrightarrow\; \nu(x) = \dot{0}$. (2.18)

Examples:

(i) Let K be the primitive field of characteristic 2 with elements 0 and 1 . Define a function v over K^n with values in the $\leq$ - ordered abelian group of integers by:

If $x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ has k components which are 1,

then $v(x) = k$. (2.19)

Then $v(x)$ is a norm, the "Hamming norm" of coding theory.

(ii) In $V = R^n$ or C^n, the

TSCHEBYSHEFF NORM. or MAXIMUM NORM, (2.20)

with values in R is defined by

$$v(x) := \max_{1 \le i \le n} |x_i| . \qquad (2.21)$$

(iii) In $V = C[0,1]$, the Tschebysheff norm with values in R is defined by

$$v(f) := \max\{f(\xi): 0 \le \xi \le 1\} .$$

(iv) In $V = R^n$ or C^n, a norm with values in the vector space $G = R^n$, ordered componentwise (2.11), is defined by

$$v(x) := |x| . \qquad (2.22)$$

We will refer to this norm as the

MODULUS NORM ("BETRAGSNORM") (2.23)

of R^n or C^n. For n = 1, it reduces to the simple absolute value which is a norm over the vector spaces R and C .

In a normed vector space with a real norm, a (unsymmetric) distance is induced by

$$d(x,y) := v(x-y) . \qquad (2.24)$$

It has the properties:

TRIANGLE INEQUALITY: $d(x,z) \le d(x,y) + d(y,z)$ (2.25)

NON-NEGATIVITY: $0 \le d(x,y)$ (2.26)

DEFINITENESS: $d(x, y) = 0 \Leftrightarrow x = y$ (2.27)

Proof: $d(x,z) = \nu(x-z) = \nu((x-y) + (y-z)) \leq \nu(x-y) + \nu(y-z)$
$= d(x,y) + d(y,z)$

In particular $x \neq y \Rightarrow d(x,y) > 0$ .If the norm is

SYMMETRIC: $\nu(-x) = \nu(x)$ (2.28)

then the distance is

SYMMETRIC: $d(x, y) = d(y,x)$, (2.29)

and, by means of the distance induced by the norm, the vector space V becomes a topological space, the topology being based upon ϵ-neighborhoods

$$U_\epsilon(x) = \{y: d(x,y) < \epsilon\}$$

Moreover, the distance is

TRANSLATION-INVARIANT: $d(x+a, y+a) = d(x,y)$. (2.30)

Conversely, a distance which is translation invariant induces a norm by means of

$$\nu(x) = d(x, \phi) . \quad (2.31)$$

Examples:

(i) The usual distance in Euclidean geometry which is invariant under translation, furnishes the most important and best-known example of a norm. The

EUCLIDEAN NORM, (2.32)

given by the distance from the origin, is the natural norm of the vector space of Euclidean geometry. In an isomorphic coordinate space of dimension n it is given by

$$\nu(x) = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} \quad x \in R^n$$

$$\nu(x) = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \quad x \in C^n \,, \tag{2.33}$$

(ii) In Manhattan, the distance a car has to travel from one place to another is the sum of the distances along the streets and the avenues. In R^n or C^n, the <u>MANHATTAN DISTANCE</u> is

$$d(x,y) := |x_1-y_1| + |x_2-y_2| + \cdots + |x_n-y_n| \tag{2.34}$$

The norm in $V = R^n$ or C^n induced by this translation invariant distance, the

<u>MANHATTAN NORM,</u> or <u>SUM NORM,</u> (2.35)

is defined by

$$\nu(x) = \sum_{i=1}^{n} |x_i| \,. \tag{2.36}$$

A mapping $\varphi: V_R \times V_R \to R$ of a vector space V_R over the real field R is a

<u>SCALAR PRODUCT</u> (2.37)

if it is

<u>SYMMETRIC:</u> $\varphi(x,y) = \varphi(y,x)$ (2.38)

<u>BILINEAR:</u> $\varphi(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1\varphi(x_1,y) + \alpha_2\varphi(x_2,y)$ (2.39)

<u>DEFINITE:</u> $x \neq \phi \Rightarrow \varphi(x,x) > 0$ (definite on the diagonal) (2.40)

A scalar product defines a norm, the

SCALAR PRODUCT NORM, (2.41)

$$\nu(x) = [\varphi(x,x)]^{\frac{1}{2}} . \tag{2.42}$$

The scalar product norm of a linear combination $\alpha x + \beta y$ can be expanded using (2.42) and (2.39) as

$$\nu^2(\alpha x+\beta y) = \alpha^2\nu^2(x) + 2\alpha\beta\varphi(x,y) + \beta^2\nu^2(y) \geq 0 . \tag{2.43}$$

For a = $\nu(y)$, $\beta = -\varphi(x,y)/\nu(y)$ we obtain

$$\nu^2(x)\nu^2(y) - \varphi^2(x,y) \geq 0 ,$$

whence

SCHWARZ-BUNJAKOWSKI INEQUALITY: $|\varphi(x,y)| \leq \nu(x)\nu(y)$ (2.44)

The cosine of the ANGLE a, $0 \leq \alpha \leq \pi$, between x and y may therefore be defined by

$$\cos\alpha = \frac{\varphi(x,y)}{\nu(x)\nu(y)}, \quad x \neq \phi \wedge y \neq \phi \tag{2.45}$$

since $\left|\frac{\varphi(x,y)}{\nu(x)\nu(y)}\right| \leq 1$.

A scalar product norm has the additional property (see(2.43))

PARALLELOGRAM EQUALITY: $\nu^2(x+y) + \nu^2(x-y) = 2\nu^2(x) + 2\nu^2(y)$ (2.46)

The scalar product is reproduced from the norm by

$$\begin{aligned}\varphi(x, y) &= \tfrac{1}{2}[\nu^2(x+y) - \nu^2(x) - \nu^2(y)] \\ &= \tfrac{1}{2}[\nu^2(x) + \nu^2(y) - \nu^2(x-y)] \\ &= \tfrac{1}{4}[\nu^2(x+y) - \nu^2(x-y)] .\end{aligned} \tag{2.47}$$

Moreover, any norm ν for which the parallelogram equality holds defines by (2.47) a function which is a scalar product (Exercise 1) and therefore is a scalar product norm, $\nu(x) = [\varphi(x,x)]^{\frac{1}{2}}$.

A normed vector space with a scalar product norm and hence a scalar product is a

$$\underline{\text{HILBERT SPACE}} . \tag{2.48}$$

In the vector space R^n, any scalar product $\varphi(x,y)$, being a symmetric, bilinear, definite functional, can be written as a symmetric, bilinear, definite form in the components of x and y, i.e.,

$$\varphi(x,y) = x^T A y \tag{2.49}$$

where x^T is the transposed vector x and A is a symmetric, positive definite matrix of order n . Consequently, any scalar product norm ν can be written

$$\nu(x) = (x^T A x)^{\frac{1}{2}} \tag{2.50}$$

The Euclidean norm is a special case with A = I .

Exercise 1. Let g(x) be a real functional over V_R such that

$$g(x+y) + g(x-y) = 2g(x) + 2g(y) .$$

Show that

$$g(x+y+z) - g(x+y) - g(y+z) - g(z+x) + g(x) + g(y) + g(z) = 0 .$$

Let furthermore $\psi(x+y) := \frac{1}{2}[g(x+y) - g(x) - g(y)]$.
Show that

$$\psi(x+y,z) = \psi(x,z) + \psi(y,z) .$$

§3. Homogeneous Norms

Very often, the range of a norm is not only an ordered abelian group, but a field (such as the real field in some of the examples in $2) or a vector space (as in Example (iv) of $2) with an ordered additive part such that non-negativity is preserved under suitable multiplications. By way of definition, a field K_0 is a

$$\underline{\text{LINEARLY } \rho_0 \text{ - ORDERED FIELD}} \tag{3.1}$$

if the additive part of K_0 is a linearly ρ_0 - ordered group and the ordering ρ_0 is compatible with non-negative multipliers:

$$\text{V-a}, \beta, \gamma \in K_0 : \quad 0 \rho_0 \alpha \wedge \beta \rho_0 \gamma \succ \alpha \beta \rho_0 \alpha \gamma \tag{3.2}$$

In particular,

$$\forall \alpha, \beta \in K_0 : \quad 0 \rho_0 \alpha \wedge 0 \rho_0 \beta \succ 0 \rho_0 \alpha \beta . \tag{3.3}$$

Since we have a linear ordering and a ρ_0 $0 \succ 0$ ρ_0 (-a)

$$\forall \alpha \in K_0 : \quad 0 \rho_0 \text{ [illegible] } \vee 0 \rho_0 \text{ [illegible] } .$$

Since $(-\alpha)^2 = \alpha^2$, squares are non-negative and, in particular,

$$1 = (1)^2 \succ 0 \rho_0 1 .$$

As a consequence, the characteristic of a linearly ordered field cannot be finite:

$$0 \rho_0 1 \succ 0 \rho_0 n \cdot 1 \ (= 1 + \ldots + 1) \ \forall n \geq 1 .$$

Moreover,

$$0 \rho_0 \alpha \wedge \alpha \neq 0 \succ 0 \rho_0 \alpha^{-1}; \tag{3.4}$$

otherwise,

$$0 \;\rho_0\; \alpha \wedge \alpha^{-1} \;\rho_0\; 0 \succ 1 \;\rho_0\; 0,$$

a contradiction. Furthermore, from $\alpha^{-1} - \beta^{-1} = \alpha^{-1}\beta^{-1}(\beta - \alpha)$,

$$0 \;\rho_0\; \alpha \wedge \alpha \neq 0 \wedge \alpha \;\rho_0\; \beta \succ \beta^{-1} \;\rho_0\; \alpha^{-1} . \tag{3.5}$$

The rational field P, the real field R, and the field of all real algebraic numbers are linearly ordered fields with the conventional $<$ - ordering.

Similarly, a vector space G over a linearly ρ_0 - ordered field K_0 is a

<u>ρ - ORDERED VECTOR SPACE</u> (3.6)

if the additive part of G is a ρ - ordered abelian group and multiplication by non-negative scalars is compatible with the ordering ρ :

$$\forall\, \alpha \in K_0,\, x,y \in G:\;\; 0 \;\rho_0\; \alpha \wedge x \;\rho\; y \succ \alpha x \;\rho\; \alpha y \,. \tag{3.7}$$

In particular,

$$\forall\, \alpha \in K_0,\, x \in G:\;\; 0 \;\rho_0\; \alpha \wedge \phi \;\rho\; x \succ \phi \;\rho\; \alpha x \,. \tag{3.8}$$

Examples:

(i) R^n is a ρ - ordered vector space over the linearly $<$ - ordered field R, ρ being the $\leq$ - ordering of (2.12)

(ii) C[0,1] is a ρ - ordered vector space over the linearly $\leq$ - ordered field R, ρ being defined by:

$$f \;\rho\; G :\Leftrightarrow f(\xi) \leq G(E),\;\; \forall\, \xi \in [0,1] \,.$$

If V_K is a normed vector space and the range of the norm ν is a ρ - ordered vector space G over a linearly ρ_0 - ordered field K_0, K_0 a subfield of K, then it makes sense to define

<u>HOMOGENEITY:</u> $\forall \alpha \in K_0$, [illegible] $\in V$: $0 \; \rho_0 \; \alpha \succ$ [illegible] $x)$ [illegible] α [illegible] . (3. 9)

For homogeneous norms, (2.17) and (2.18) can be replaced by:

<u>POSITIVE DEFINITE</u>: $\forall x \in V: x \neq 0 \succ 0 \; \rho \; v(x) \wedge v(x) \neq 0$ (3.10)

Proof:

From homogeneity with $\alpha = 0$,

$\nu(\phi)$ [illegible] $\nu(0$ [illegible] $\nu(x) = 0$; i.e., $x = \phi \succ \nu(x) = 0$.

From positive definiteness,

$$\nu(x) = 0 \succ x = \phi$$

giving (2.18). This and positive definiteness give (2.17).

In R^n and C^n, the Tschebyscheff norm, the Euclidean norm and other scalar product norms, the Manhattan norm, and the modulus norm are all homogeneous. We shall assume homogeneity in succeeding paragraphs and shall speak simply of norms if G is a field (mainly the real field) and of <u>VECTORIAL NORMS</u> if G is a vector space of dimension greater than 1 over some field (again mainly the real field).

§4. Linear Mappings.

A

LINEAR MAPPING, (4.1)

i.e., a mapping φ of a vector space $\mathcal{V}_K$ into a vector space $\mathcal{V}'_K$ is called a VECTOR SPACE HOMOMORPHISM if

$$\forall\ \alpha, \beta \in K,\ x,y \in \mathcal{V}_K:\ \varphi(\alpha\ x + \beta\ y) = \alpha\ \varphi(x) + \beta\ \varphi(y)\ ; \qquad (4.2)$$

i.e., if φ is compatible with linear combinations. In particular, $\varphi(\phi) = \phi$.

The image $\varphi(\mathcal{V}_K) \subset \mathcal{V}'_K$ is itself a vector space, a subspace of $\mathcal{V}'_K$. φ induces a SURJECTIVE (onto) linear mapping of $\mathcal{V}_K$ onto $\varphi(\mathcal{V}_K)$. However, since we frequently consider homomorphisms of a vector space $\mathcal{V}_K$ into itself (ENDOMORPHISMS), it would be impractical to restrict our attention to surjective mappings only.

Let φ be a linear mapping of $\mathcal{V}_K$ into $\mathcal{V}'_K$. The set

$$\text{Ker}\ \varphi := \{x \in \mathcal{V}_K:\ \varphi(x) = \phi\}$$

is a subspace of $\mathcal{V}_K$ the KERNEL of cp . [Note that $\phi \in$ Ker φ; $\varphi(x) = \phi \wedge \varphi(y) =$ 4 $\succ \varphi(\alpha\ x + \beta\ y) = \alpha\ \varphi(x) + \beta\varphi(y) = \phi$] .

φ is INJECTIVE if

$$\varphi(x)\ _{\text{\tiny{■}}}\ \varphi(y) \succ x\ _{\text{\tiny{■}}}\ _{\text{\tiny{□}}}\ .$$

Equivalently, φ is injective $\asymp$ Ker $\varphi = \{\phi\}$. [Note that

$$\varphi(x) = \varphi(y) \asymp \varphi(x\text{-}y) = \varphi(x) - \varphi(y) = \phi\ ;$$
$$(\text{cp}(z) = \phi \succ z\text{-}= \phi) \asymp \{x \in \mathcal{V}_K : \varphi(x) = \phi\} = \{\phi\}\ ].$$

φ is ONE-TO-ONE if it is both surjective (every element of $\mathcal{V}'_K$ has at least one preimage) and injective (every element of $\mathcal{V}'_K$ has at most one preimage). Such a mapping φ is called a REGULAR mapping or ISOMORPHISM. If Ker $\varphi = \{\phi\}$, then the induced linear mapping $\hat{\varphi}: \mathcal{V}_K \rightarrow \varphi(\mathcal{V}_K)$ is an isomorphism. The set of all linear mappings of a vector space $\mathcal{V}_K$ into a vector space $\mathcal{V}'_K$, denoted

$$\text{Hom}\ (\mathcal{V}_K, \mathcal{V}'_K) \qquad (4.3)$$

is itself a vector space over K with addition and scalar multiplication defined by

$$\varphi = \varphi_1 + \varphi_2 \quad :\Leftrightarrow \quad \varphi(x) = \varphi_1(x) + \varphi_2(x)\ , \quad \forall\ x \in \mathcal{V}_K \qquad (4.4)$$

$$\varphi = \alpha\,\varphi_1 \quad :\Leftrightarrow \quad \varphi(x) = \alpha\ \varphi_1(x)\ , \qquad \forall\ x \in \mathcal{V}_K \qquad (4.5)$$

The zero element of Horn $(\mathcal{V}_K, \mathcal{V}'_K)$ is the zero mapping $0 : 0(x) \equiv \phi$.
If $\mathcal{V}'_K$ is just the field. K itself, the mapping φ is called a

$$\text{LINEAR (K - VALUED) FUNCTIONAL OF } \mathcal{V}_K \qquad (4.6)$$

and we write

$$\mathcal{V}^D_K = \text{Hom}\ (\mathcal{V}_K, K) \qquad (4.7)$$

The zero element of $\mathcal{V}^D_K$ is the ZERO FUNCTIONAL $\phi^D : \phi^D(x) \equiv 0$.

Example:

The dual vector space of the coordinate space R^n is the set of all linear functionals

$$\varphi(x) = \ell_1 x_1 + \ell_2 x_2 + \dots + \ell_n x_n = \ell^T x, \qquad (4.8)$$

where $\ell^T = (\ell_1, \ell_2, \dots, \ell_n)$ is called a ROW VECTOR. In this representation, $(R^n)^D$ is again a coordinate space of dimension n over R .

§5. Subadditive Functionals Generated by a Set of Linear Functionals

Linear mappings are trivially seen to be subadditive and homogeneous but not definite. We shall use supremum constructions which preserve subadditivity and homogeneity to generate functionals that are non-negative and even definite. We first turn our attention to the case where K is the real field R and G_{K_o} coincides with K, i.e., the real field. Thus, we discuss real-valued functionals and norms of a vector space $\mathcal{V}$ over R .

In the linearly ordered real field R, the supremum of a set of elements is defined for bounded, nonempty sets. To remove these restrictions, we form the EXTENDED REAL FIELD[1] $R^* = \{R, +\infty, -\infty\}$ and define

$$\sup R = +\infty ; \quad \sup \emptyset = -\infty \qquad (5.1)$$
$$\inf R = -\infty ; \quad \inf \emptyset = +\infty$$

where $\emptyset$ denotes the empty set. The $<$ - ordering of R^* is that of R, supplemented by

$$\forall \alpha \in R: \ -\infty \leq \alpha \wedge \alpha \leq +\infty .$$

Theorem: (5.2)

Let $S = \mathcal{V}^D = \text{Horn}(\mathcal{V}, R)$ be a set of linear real-valued functionals of a vector space U over R . Then

$$\gamma_S(x) := \sup\{\varphi(x): \varphi \in S\} \qquad (5.3)$$

is a subadditive, homogeneous functional (sometimes called a GAUGE FUNCTION) over $\mathcal{V}$ with values from extended real field R^* .

[1] Note that R^* is not a field: $(+\infty) + (-\infty)$ is not defined.

Proof:

$$\begin{aligned}\gamma_S(x+y) &= \sup\{\varphi(x+y) : \varphi \in S\} \\ &= \sup\{\varphi(x) + \varphi(y) : \varphi \in S\} \\ &\leq \sup\{\varphi(x) : \varphi \in S\} + \sup\{\varphi(y) : \varphi \in S\} \\ &= \gamma_S(x) + \gamma_S(y)\end{aligned}$$

$\therefore \gamma_S(x)$ is subadditive.

$$\begin{aligned}\gamma_S(\alpha x) &= \sup\{\varphi(\alpha x): \varphi \in S\} \\ &= \sup\{\alpha\ \varphi(x) : \varphi \in S\} \\ &= \alpha \sup\{\varphi(x): \varphi \in S\} \qquad \text{for } \alpha \geq 0 \\ &= \alpha\ \gamma_S(x)\end{aligned}$$

$\therefore \gamma_S(x)$ is homogeneous. Q.E.D.

To be a norm, $\gamma_S(x)$ must also be non-negative, definite, and real-valued (i.e., bounded). A sufficient condition for the first property is given by:

Theorem: $\Phi^D \in S \Rightarrow 0 \leq \gamma_S(x), \ \forall\, x \in \mathcal{V}$. (5.4)

Proof: $0 = \Phi^D(x) \leq \sup\{\varphi(x): \varphi \in S\} = \gamma_S(x)$.

A mapping over $\mathcal{V}$ with values from the extended real field R* is a

SEMINORM (5.5)

if it is subadditive, homogeneous, and non-negative. Obviously,

$$\forall\ \varphi \in S,\ x \in \mathcal{V}: \varphi(x) \leq \gamma_S(x) \ . \tag{5.6}$$

Moreover, some linear combinations of elements of S are bounded by γ_S. Let $\varphi_1, \varphi_2, \ldots \varphi_n \in \mathcal{V}^D$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in R$. Then $\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_n\varphi_n \in \mathcal{V}^D$ is a

CONVEX COMBINATION OF $\varphi_1, \varphi_2, \ldots, \varphi_n$ (5.7) if $0 \leq \alpha_i$ and $\alpha_1 + \alpha_2 + \ldots + a_n = 1$.

Theorem: Let φ be a convex combination of $\varphi_1, \varphi_2, \ldots, \varphi_n \epsilon S$. (5.8)

Then

$$\forall\ x \ \epsilon\ V: \varphi(x) \leq \gamma_S(x) . \tag{5.9}$$

Proof:

$$\alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) + \ldots + \alpha_n\varphi_n(x)$$
$$\leq \alpha_1\gamma_S(x) + \alpha_2\gamma_S(x) + \ldots + \alpha_n\gamma_S(x)$$
$$= (\alpha_1 + \alpha_2 + \ldots + \alpha_n)\gamma_S(x)$$
$$= \gamma_S(x) .$$

Theorem: (5.10)

If ϕ^D can be represented as a convex combination of elements of S, then $\gamma_S(x)$ is non-negative and therefore a seminorm.

The converse is not true, e.g., Example (i)(d) below.

Examples:

(i) The following subadditive, homogeneous functionals γ_S over R^2 are depicted by their contour maps in Figure 1.

(a) $S_1 = \{(0,1), (1,0)\}$ $\quad\gamma_{S_1}(x) = \max(x_1, x_2)$

(b) $S_2 = \{(1,1), (1,2), (2,1), (2,2)\}$

$$\gamma_{S_2}(x) = \max\{x_1+x_2,\ x_1+2x_2,\ 2x_1+x_2,\ 2x_1+2x_2\}$$

(c) $S_3 = \{(\ell_1,\ell_2)\ell_1 \geq 0,\ \ell_2 \geq 0,\ \ell_1^2 + \ell_2^2 = 1)$

$$\gamma_{S_3}(x) = \begin{cases} (x_1^2 + x_2^2)^{\frac{1}{2}} & x_1 \geq 0,\ x_2 \geq 0 \\ x_1 & x_1 > 0,\ x_2 < 0 \\ x_2 & x_1 \leq 0,\ x_2 \geq 0 \\ \max(x_1, x_2) & x_1 \leq 0, x_2 \leq 0 \end{cases}$$

FIGURE 1

(i)

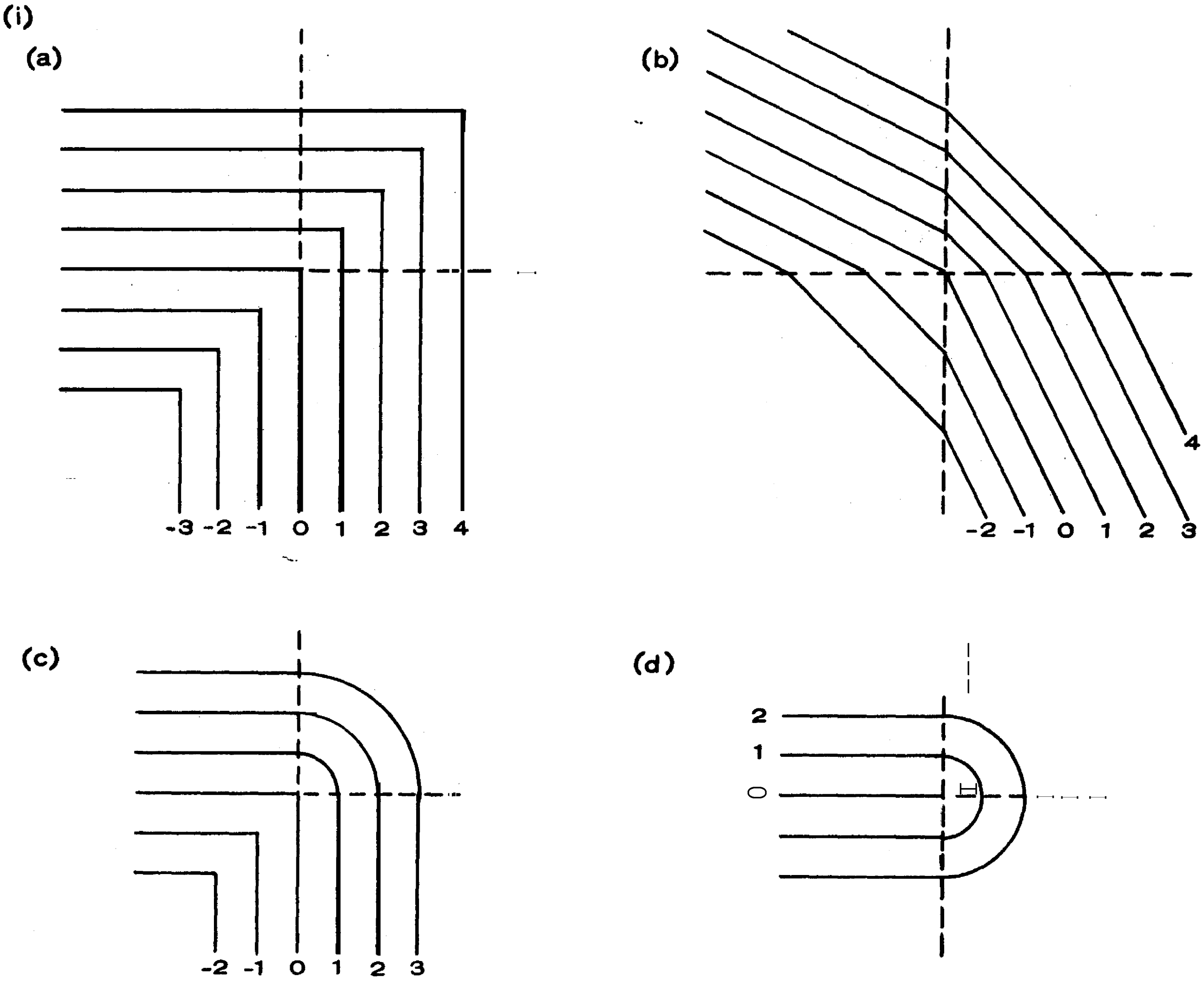
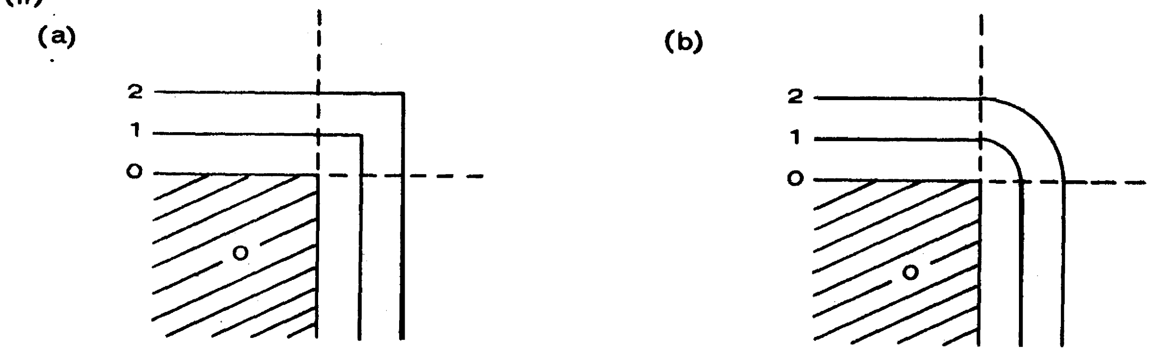


FIGURE 2

(ii)



(d) $S_4 = \{(\ell_1,\ell_2): \ell_1 > 0, \ell_1^2 + \ell_2^2 = 1\}$

$$\gamma_{S_4}(x) = \begin{cases} (x_1^2 + x_2^2)^{\frac{1}{2}} & x_1 > 0 \\ |x_2| & x_1 \leq 0 \end{cases}$$

(ii) The following seminorms γ_S over R^2 are generated by sets S which contain ϕ^D or a subset, a convex combination of which is ϕ^D (See Figure 2):

(a) $S_5 = \{(1,0), (0,1), (0,0)\}$ $\quad \gamma_{S_5}(x) = \max(x_1, x_2, 0)$

(b) $S_6 = \{(\ell_1,\ell_2): \ell_1 \geq 0, \ell_2 > 0, \ell_1^2 + \ell_2^2 = 1\} \cup (0,0)\}$

$$\gamma_{S_6}(x) = \max(\gamma_{S_3}(x), 0)$$

(iii) The following **functionals** γ_S over R^2 have $+\infty$ among their values. All except (a) are seminorms. (See Figure 3):

(a) $S_7 = \{(\ell_1,0): \ell_1 \geq 1\}$ $\quad \gamma_{S_7}(x) = \begin{cases} +\infty & x_1 > 0 \\ x_1 & x_1 \leq 0 \end{cases}$

(b) $S_8 = \{(\ell_1,0): \ell_1 \geq 0\}$ $\quad \gamma_{S_8}(x) = \begin{cases} +\infty & x_1 > 0 \\ 0 & x_1 \leq 0 \end{cases}$

(c) $S_9 \equiv \{(\ell_1,\ell_2): \ell_1 \geq 0, \ell_2 \geq 0\}$ $\quad \gamma_{S_9}(x) = \begin{cases} 0 & x_1 \leq 0, x_2 \leq 0 \\ +\infty & \text{otherwise} \end{cases}$

(d) $S_{10} = \{(\ell_1,0): \ell_1 \epsilon R\}$ $\quad \gamma_{S_{10}}(x) = \begin{cases} +\infty & x_1 \neq 0 \\ 0 & x_1 = 0 \end{cases}$

(e) $S_{11} = \{(\ell_1,\ell_2): |\ell_2| \leq 1\}$ $\quad \gamma_{S_{11}}(x) = \begin{cases} |x_2| & x_1 = 0 \\ +\infty & \text{otherwise} \end{cases}$

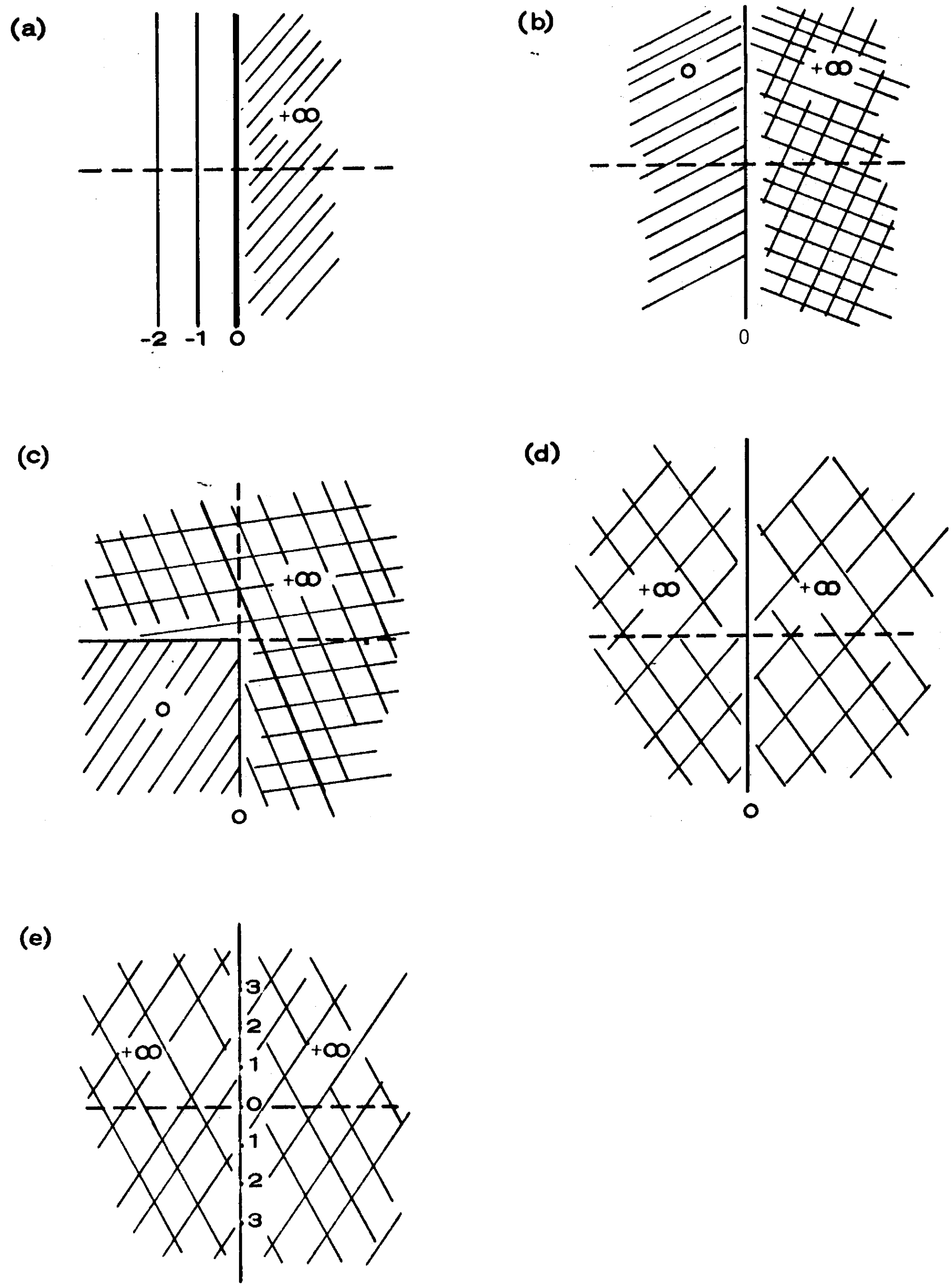
Note that in all the examples, the set

$$K_\rho := \{x : \gamma_S(x) \leq \rho\} \tag{5.11}$$

is an intersection of a family of half-planes

FIGURE 3

(iii)



$$H_{\varphi,\rho} := \{x : \varphi(x) \le \rho\} .$$

In fact

<u>Theorem</u>: $K_\rho = \bigcap_{\varphi \in S} H_{\varphi,\rho}$ (5.12)

<u>Proof</u>:

$$\begin{aligned} x \in K_\rho &\Rightarrow \gamma_S(x) \le \rho \\ &\Rightarrow \varphi(x) \le \rho, \quad \forall \varphi \in S \qquad \text{by (5.6)} \\ &\Rightarrow x \in H_{\varphi,\rho}, \quad \forall \varphi \in S \\ &\Rightarrow x \in \bigcap_{\varphi \in S} H_{\varphi,\rho} \end{aligned}$$

$$\begin{aligned} x \in \bigcap_{\varphi \in S} H_{\varphi,\rho} &\Rightarrow x \in H_{\varphi,\rho}, \quad \forall \varphi \in S \\ &\Rightarrow \varphi(x) \le \rho, \quad \forall \varphi \in S \\ &\Rightarrow \rho \text{ upper bound for } \{\varphi(x) : \varphi \in S\} \\ &\Rightarrow \rho \le \gamma_S(x) = \mathrm{lub}\{\varphi(x) : \varphi \in S\} \\ &\Rightarrow x \in K_\rho \qquad \text{Q.E.D.} \end{aligned}$$

For $\rho < 0$, K_ρ may be empty. In particular, K_0 is a

<u>CONE,</u> (5.13)

i.e., a subset of $\mathcal{V}$ such that

$$x \in K_0 \wedge \alpha \in R \wedge \alpha \ge 0 \Rightarrow \alpha x \in K_0 .$$

K_0 certainly contains Φ and may degenerate to $\{\Phi\}$:

<u>Theorem</u>: A seminorm is definite $\Leftrightarrow K_0 = \{\Phi\}$. (5.14)

If $K_0 = \{\Phi\}$, then

$$\bigcap_{\varphi \in S} H_{\varphi,0} = \{\Phi\} . \tag{5.15}$$

i.e., S "surrounds the origin of $\mathcal{V}^D$."

A seminorm $\gamma_S(x)$ is a norm if it is definite and also

$$\underline{\text{BOUNDED}}: \forall x \in \mathcal{V}: \gamma_S(x) < +\infty . \qquad (5.16)$$

A sufficient condition for boundedness is:

<u>Theorem:</u> If S is a finite set, then γ_S is bounded. (5.17)

In $\mathcal{V} = R^n, \gamma_S$ is bounded if S is componentwise bounded. If γ_S is both definite and bounded, we write the norm defined by S as

$$\nu_S(x) . \qquad (5.18)$$

In particular, we can now derive the Tschebyscheff, Euclidean, and Manhattan norms in R^n from their generation sets. Let $e_1 := (0,0, \ldots, 0,1,0, \ldots 0)^T \in R^n$ where the 1 is in the i^{th} place. The Tschebyscheff norm is defined by

$$S = \bigcup_i \{e_i^T, -e_i^T\} ; \nu_S(x) = \max_i |x_i| . \qquad (5.19)$$

The Euclidean norm is defined by

$$S = \{(\ell_1, \ldots, \ell_n) : \ell_1^2 + \ell_2^2 + \cdots + \ell_n^2 = 1\} ; \quad \nu_S(x) = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$$
(5.20)

<u>Proof:</u>

$$\nu_S(x) = \sup\{\ell^T x : \ell^T \ell = 1\}$$

$$< (x^T x)^{\frac{1}{2}} \quad \text{since by (2.49)} \quad |\ell^T x| \leq (\ell^T \ell)^{\frac{1}{2}} (x^T x)^{\frac{1}{2}}$$

For $\ell^T = x^T/(x^T x)^{\frac{1}{2}}, \ell^T x = (x^T x)^{\frac{1}{2}}$ and $\ell^T \ell = 1$.

$$\therefore \nu_S(x) = (x^T x)^{\frac{1}{2}} = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$$

The Manhattan norm is defined by

$$S = \{(\pm 1, \pm 1, \ldots, \pm 1)\} \; ; \qquad \nu_S(x) = \sum_{i=1}^{n} |x_i| \tag{5.21}$$

<u>Proof</u>:

$$\forall \; \varphi \in S: \; \varphi(x) = \sum_{i=1}^{n} |x_i|$$

$$\text{For } \ell_i = \begin{cases} 1 & x_i \geq 0 \\ -1 & x_i \leq 0 \end{cases}, \quad \varphi(x) = \ell^T x = \sum_{i=1}^{n} |x_i| \; .$$

We can now discuss real-valued norms of a vector space V over the complex field C . For the supremum construction, we can no longer use linear functionals of V over C since they are complex-valued. However, the real part of these functionals is still additive and homogeneous:

<u>Theorem</u>: Let $S \subset V^D = \mathrm{Hom}(V, C)$ be a set of complex-valued functionals on V_C . Then (5.22)

$$\gamma_S(x) = \sup\{\mathrm{Re}(\varphi(x)) : \varphi \in S\} \tag{5.23}$$

is a subadditive, homogeneous functional on V with values from the extended real field R* .

The theory develops further as in the real case. For the Tschebyscheff norm, the Euclidean norm, and the Manhattan norm, the generating sets are respectively,

$$S = \bigcup_i \{\omega e_i^T : \; |\omega| = 1\} \tag{5.24}$$

$$S = \{(\ell_1, \ldots, \ell_n) : \; |\ell_1|^2 + |\ell_2|^2 + \; . \; . \; . + |\ell_n|^2 = 1\} \tag{5.25}$$

$$S = \{(\omega_1, \omega_2, \; . \; . \; . \; . \; \omega_n) : \; |\omega_1| = 1, \; i = 1, 2, \; . \; . \; . \; . \; n) \; . \tag{5.26}$$

Before going into a similar study of the case of vectorial norms, we shall elaborate on the generation of norms somewhat further in order to investigate fields of values and eigenvalue exclusion theorems.

§6. Replete Generating Sets. Application: Fields of Values and Eigenvalue Exclusion Theorems.

We shall call the set $S \subset \mathcal{V}^D$ that generates $\gamma_S(x)$ (or $\nu_S(x)$)

$$\underline{\text{REPLETE}} \tag{6.1}$$

if

$$\forall\, x \in \mathcal{V} \;\; \exists\, \varphi \in S : \;\; \varphi(x) = \gamma_S(x) \,. \tag{6.2}$$

If S is replete, then

$$\gamma_S(x) = \max_{\varphi \in S} \varphi(x); \tag{6.3}$$

i.e., the supremum is actually attained.

Not every set is replete; in §5, Example (i)(d),

$$x = \binom{0}{1} \Rightarrow \varphi(x) < \gamma_S(x), \; \forall\, \varphi \in S \,.$$

Whether a set $S \subset \mathcal{V}^D$ can be extended to a replete set S' such that $\gamma_S(x) \equiv \gamma_{S'}(x)$ and whether

$$S' = \{\varphi \in \mathcal{V}^D : \varphi(x) \le \gamma_S(x), \; \forall\, x \in \mathcal{V}\}$$

is replete are subtle topological problems for which no general answers exist. For finite dimensional spaces, however, the <u>SUPPORT THEOREM</u> (Bonnesen - Fenchel) guarantees that every set S can be so extended. Henceforth we shall consider only replete sets in generating norms in R^n and C^n.

On the other hand, a replete set S need not consist of all linear functionals φ satisfying $\varphi(x) \le \gamma_S(x)$. The set

$$S = \bigcup_i \{e_i^T, -e_i^T\}$$

of (5.19) generates the Tschebyscheff norm in R^n as does

$$S' = \{\varphi \in \mathcal{V}^D : \varphi(x) \leq \gamma_S(x), \quad \forall\, x \in \mathcal{V}\}$$
$$= \{(\ell_1, \ell_2, \ldots, \ell_n) : \quad |\ell_1| + |\ell_2| + \ldots + |\ell_n| \leq 1\} .$$

Both sets are replete, but the additional elements in S' are convex combinations of the elements of S and are in a sense superfluous. In finite dimensional spaces, the set $\tilde{S}$ of EXTREME POINTS of any replete extension S',

$$\tilde{S} = \{\varphi \in S' : \varphi_1, \varphi_2 \in S' \wedge \varphi = \alpha\, \varphi_1 + (1 - \alpha)\varphi_2 \wedge 0 \leq \alpha \leq 1$$
$$\succ \varphi = \varphi_1 \vee \varphi = \varphi_2\},$$

contains no superfluous elements. We shall see in the following application to fields of values and eigenvalue exclusion theorems that it is important to choose the generating set S to be replete yet as small as possible.

Let A be a linear mapping of a vector space $\mathcal{V}$ over the complex field C into itself. The set of all such endomorphisms of $\mathcal{V}$, Hom($\mathcal{V}$, $\mathcal{V}$) ((4.3)), is itself a vector space over C and even a ring, multiplication being composition of mappings. If

$$Ax = Ax \tag{6.4}$$

where $x \neq 0$, then $\lambda \in C$ is called an

EIGENVALUE OF A (6.5)

and $x \in \mathcal{V}$ the corresponding

EIGENVECTOR OF A . (6.6)

We define the

FIELD OF VALUES OF A (6.7)

with respect to the replete set S which generates the norm $\nu_S(x)$ to be the set of complex numbers

$$G_S[A] := \{\varphi(Ax) : \varphi \in S,\ x \in \mathcal{V},\ \nu_S(x) = \varphi(x) = 1\} . \qquad (6.8)$$

$G_S[A]$ has the property of COVARIANCE UNDER TRANSLATION:

$$G_S[A + \sigma I] = G_S[A] + \sigma := \{\sigma + \alpha : \alpha \in G_S[A]\} . \qquad (6.9)$$

Proof:

$$\forall\ \varphi \in S, x \in \mathcal{V} \text{ such that } \nu_S(x) = \varphi(x) = 1:$$

$$\varphi((A + \sigma I)x) = \varphi(Ax) + q(x) = \varphi(Ax) + \sigma .$$

The field of values of A with respect to such a set S defines an EXCLUSION DOMAIN for the eigenvalues of A:

Exclusion Theorem: No eigenvalue of A lies outside $G_S[A]$; (6.10)
i.e., if $\lambda \in C$ is an eigenvalue of A, then $\lambda \in G_S[A]$.

Proof:

Let $x \neq 0$ be an eigenvector of A corresponding to the eigenvalue λ . Since S generates a norm, $\nu_S(x) > 0$ and $x' = x/\nu_S(x)$ is again an eigenvector with $\nu_S(x') = 1$. By repleteness, there exists $\varphi \in S$ such that $\varphi(x') = \nu_S(x') = 1$. $\varphi(Ax') = \varphi(\lambda x') = \lambda\varphi(x') = \lambda . 1 = \lambda$ whence $\lambda \in G_S[A]$.

Q.E.D.

In $\mathcal{V} = C^n$, A can be represented by an n X n complex matrix, an element of the matrix ring $C^{n \times n}$, and

$$G_S[A] := \{\ell^H A x \ :\ \ell^H \epsilon S,\ x \epsilon V,\ \nu_S(x) = \ell^H x = 1\}. \qquad (6.11)$$

Using the generating set S of the Tschebyscheff norm ((5.24)), we obtain the

<u>GERSCHGORIN FIELD OF VALUES,</u> (6.12)

a union of circular domains centered at the diagonal elements of A :

$$G_S[A] := \bigcup_{i=1}^{n} C_i[A], \qquad (6.13)$$

where

$$C_i[A] = (z\ :\ |z - a_{ii}| \leq \sum_{\mu \neq i} |a_{i\mu}|\} . \qquad (6.14)$$

<u>Proof:</u>

$$G_S[A] = \{\ell^H A x\ :\ \ell^H \epsilon S = \bigcup_i \{\omega e_i^T\ :\ |\omega| = 1\},\ x \epsilon V,\ \nu_S(x) = \ell^H x = 1\}$$

$$= \bigcup_{i=1}^{n} \{\omega e_i^T A x\ :\ |\omega| = 1,\ x \epsilon V, \nu_S(x) = \max |x_i| = 1,\ \omega e_i^T x = \omega x_i = 1\}$$

$$= \bigcup_{i=1}^{n} \{\ \omega e_i^T A x\ :\ |\omega| = 1,\ x_i = \frac{1}{\omega}\ \text{ and }\ |x_\mu| \leq 1 \text{ for } \mu \neq i\}$$

$$= \bigcup_{i=1}^{n} \{e_i^T A x\ :\ x_i = 1 \text{ and } |x_\mu| \leq 1 \text{ for } \mu \neq i\}$$

$$= \bigcup_{i=1}^{n} \{\ a_{ii} + \sum_{\mu \neq i} a_{i\mu} x_\mu\ :\ |x_\mu| \leq 1 \text{ for } \mu \neq i\}$$

$$= \bigcup_{i=1}^{n} \{\ a_{ii} + \eta \sum_{\mu \neq i} |a_{i\mu}|\ :\ 0 \leq \eta \leq 1\}$$

$$= \bigcup_{i=1}^{n} \{z : |z - a_{ii}| \leq \sum_{\mu \neq i} |a_{i\mu}|\}$$

$$= \bigcup_{i=1}^{n} C_i[A] .$$

Q.E.D

<u>Corollary:</u> Let x be an eigenvector of A with a <u>DOMINANT i-TH COMPONENT</u>: (6.15)

$$|x_\mu| \leq |x_i| , \mu = 1, 2, ... n .$$

Then the eigenvalue corresponding to x lies in $C_i[A]$.[1]

We also note that $C_i[A]$ reduces to a single point, $C_i[A] = \{a_{ii}\}$, if and only if e_i^T is a <u>LEFT EIGENVECTOR</u> of A : $e_i^T A = \lambda e_i^T$; i.e., the i-th row of A is just $a_{ii} e_i^T$. Consequently, the Gerschgorin field of values reduces to n points if and only if A is a diagonal matrix, these n points being the eigenvalues of A . The following examples show, however, that one or several of the disks $C_i[A]$ may be arbitrarily small without containing an eigenvalue:

Examples:

(i) $A = \begin{pmatrix} 1 & 1 & \epsilon^{-1} \\ 1 & 2 & 3\epsilon^{-1} \\ \epsilon & 3\epsilon & 6 \end{pmatrix}$ with eigenvalues $1, 4 \pm \sqrt{15}$ (6.16)

$C_3[A] = \{z : |z - 6| < 4\epsilon\}$ does not contain any eigenvalues of A for ϵ sufficiently small.

(ii) $A = \begin{pmatrix} 1 & \epsilon & 0 \\ \epsilon^{-1} & 2 & \epsilon^{-1} \\ 0 & \epsilon & 3 \end{pmatrix}$ with eigenvalues $2, 2 + \sqrt{3}$ (6.17)

$C_1[A] = \{z : |z - 1| \leq \epsilon\}$ and $C_3[A] = \{z : |z - 3| \leq \epsilon\}$ do not contain any eigenvalues of A for ϵ sufficiently small.

[1] The classical elementary proof of Gerschgorin's Theorem goes along this line. In practice, however, information of this kind is rarely available.

Obviously, the set S generating the Tschebyscheff norm is distinguished with respect to diagonal matrices in so far as all of its elements are left eigenvectors of a diagonal matrix. Other generating sets may contain more elements than there could be eigenvectors of a non-derogatory matrix. In general, the field of values will reduce to a finite number of points only if the matrix is a multiple of the identity matrix. This is true in particular for the set (5.25) generating the Euclidean norm. The corresponding field of values, the

$$\underline{\text{TOEPLITZ FIELD OF VALUES}}, \tag{6.18}$$

is given by

$$G_S[A] = \{x^H Ax \; : \; x^H x = 1\} \tag{6.19}$$

since by the Schwarz inequality ((2.44)),

$$\ell^H \ell = 1 \wedge \nu_S(x) = x^H x = 1 \wedge \ell^H x = 1 \Rightarrow \ell^H = x^H .$$

A classic result by Toeplitz asserts that this field of values is convex ((9.20)). If A is NORMAL (unitarily similar to a diagonal matrix), then the Toeplitz field of values of A is the convex hull (the set of all convex combinations) of the eigenvalues of A :

$$\begin{aligned} G_S[A] &= \{x^H Ax \; : \; x^H x = 1\} \\ &= \{x^H U \Lambda U^H x \; : \; x^H U U^H x = x^H x = 1\} \\ &= \{y^H \operatorname{diag}(\lambda_i) y \; : \; y^H y = 1\} \\ &= \{ \sum_{i=1}^{n} |y_i|^2 \lambda_i \; : \; \sum_{i=1}^{n} |y_i|^2 = 1\} \\ &= \{ \sum_{i=1}^{n} p_i \lambda_i \; : \; 0 \le p_i \le 1, \sum_{i=1}^{n} p_i = 1\} . \end{aligned}$$

In this case, $x^H Ax$ such that $x^H x = 1$ is called a **RAYLEIGH QUOTIENT** and we write

$$G_S[A] = H[\lambda_1, \lambda_2, \ldots \lambda_n] . \qquad (6.20)$$

For a vector space V over the real field R, we may again define $G_S[A]$ by (6.8), now giving a set of real numbers which contains all real eigenvalues of A . For the generating set (5.19), we obtain the restriction of the Gerschgorin field of values to the real axis, thus nothing new for real A . For the generating set (5.20), we obtain the restriction of the Toeplitz field of values to the real axis. If A is real and symmetric, then A is normal with real eigenvalues and we again obtain the convex hull of the two extreme eigenvalues which consists of all Rayleigh quotients.

In the real case, the set (5.21) generating the Manhattan norm gives the field of values

$$G_S[A] = \bigcup_{\ell^T \in S} T_\ell[A] , \quad S = \{(\pm 1, \pm 1, \ldots \pm 1)\} \qquad (6.21)$$

where

$$\begin{aligned}
T_\ell[A] &= \{\ell^T Ax : \nu_S(x) = |x_1| + |x_2| + \ldots + |x_n| = 1, \ell^T x = 1\} \\
&= \{\ell^T Ax : x = (\ell_1 p_1, \ell_2 p_2, \ldots \ell_n p_n) \text{ where } 0 \le p_i \le 1 \\
&\qquad \text{and } \Sigma p_i = 1) \\
&= \{\ell^T AD_\ell p : 0 \le p_i \le 1 \text{ and } \Sigma p_i = 1) \text{ where} \\
&\qquad D_\ell = \text{diag}(\ell_1, \ldots, \ell_n) \\
&= \{\Sigma p_i (\ell^T AD_\ell)_i : 0 \le p_i \le 1 \text{ and } \Sigma p_i = 1) \\
&= H[(\ell^T AD_\ell)_1, \ldots, (\ell^T AD_\ell)_n] \\
T_\ell[A] &= H[\mathbf{e}^T D_\ell AD_\ell e_1, \ldots \mathbf{e}^T D_\ell ADe_n] \text{ where } \mathbf{e}^T = (1, 1, \ldots 1) \\
&\qquad \text{and } \mathbf{e}^T D_\ell = \ell^T . \qquad (6.22)
\end{aligned}$$

As in (6.15), we can gain some additional information as to the location of eigenvalues by looking at the eigenvectors:

Theorem: Let x be a real eigenvector of a real matrix A (6.23)
and $\ell^T = (+1, +1, \ldots, +1)$ be the SIGN DISTRIBUTION of x :

$$\ell_i = \begin{cases} +1, & x_i \geq 0 \\ -1, & x_i < 0 . \end{cases}$$

Then the eigenvalue corresponding to x lies in $T_\ell[A]$.

In contrast to the previous situation, some information concerning the sign distribution of real eigenvectors of a matrix is often available as is the case with so-called oscillation matrices We shall later see that matrices with non-negative elements have at least one eigenvector which has in suitable form non-negative components. The corresponding eigenvalue (the Perron root) certainly lies in $T_{(1, 1, \ldots, 1)}$.

Example:

$$A = \begin{pmatrix} 9 & 3.6 & 2 \\ 22.5 & 18 & 15 \\ 40.5 & 48.6 & 54 \end{pmatrix}$$

$$T_{(1, 1, \ldots, 1)}[A] = H[72, 70.2, 71] = [70.2, 72]$$

contains the eigenvalue $36 + 9\sqrt{15} = 70.8568$

Note that this theorem gives good results only if the column sums of A (or rather of $D_\ell A D_\ell$) are not very different. Thus matrices are distinguished which are non-negative apart from a sign transformation and whose column sums are nearly equal.

§7. Norm Transformations and Invariance Groups

Let S be a set of linear functionals and let

$$SB := \{\varphi B: \varphi \in S\} \qquad (7.1)$$

denote the transformed set under the linear mapping $B \in Hom(V, V)$ where φB is defined by

$$\varphi B(x) := \varphi(Bx) \; . \qquad (7.2)$$

SB is again a set of linear functionals and generates the functional $\gamma_{SB}(x)$:

<u>Theorem:</u> $\gamma_{SB}(x) = \gamma_S(Bx)$. (7.3)

Moreover,

<u>Theorem:</u> γ_{SB} is a norm if and only if γ_S is a norm and B is a regular mapping (isomorphism). (7.4)

<u>Proof</u>:

$$\text{B not regular} \succ Bx = 0 \text{ for some } x \neq 0$$
$$\succ \gamma_{SB}(x) = \gamma_S(Bx) = 0 \text{ for some } x \neq 0$$
$$\succ \gamma_{SB} \text{ is not a norm.}$$

$$\text{B regular} \wedge \gamma_S \text{ not a norm} \succ \gamma_S(x) = 0 \text{ for some } x \neq 0$$
$$\succ \gamma_{SB}(y) = \gamma_S(By) = 0 \text{ for some } y = B^{-1}x \neq 0$$
$$\succ \gamma_{SB} \text{ is not a norm.}$$

$$\text{B regular} \wedge \gamma_S \text{ a norm} \succ (\gamma_{SB}(x) = \gamma_S(Bx) = 0 \succ Bx = 0 \succ x = 0)$$
$$\succ \gamma_{SB} \text{ is a norm.}$$

Q.E.D.

We shall call $\nu^B(x) := \nu(Bx)$ (in particular, $(\nu_S)^B = \nu_{SB}$) a

LINEARLY TRANSFORMED NORM or a LINEAR TRANSFORMATION OF ν. (7.5)

Let K_ρ denote the set $\{x: \nu_S(x) \leq \rho\}$ = The corresponding set for ν_{SB} is

$$\{x: \nu_{SB}(x) \leq \rho\} = \{x: \nu_S(Bx) \leq \rho\} = \{B^{-1}y: \nu_S(y) \leq \rho\} = B^{-1}K_\tau .$$

Thus,

If S is replaced by SB, then K_ρ is replaced by $B^{-1}K_\rho$. (7.6)

If B leaves the norm ν invariant (in particular, SB = S), then the linear transformation B is called a

NORM INVARIANCE TRANSFORMATION. (7.7)

The set of all such transformations is clearly a group, the

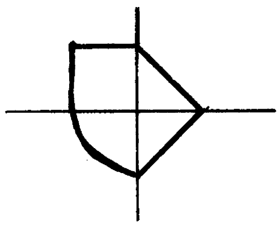
INVARIANCE GROUP (7.8)

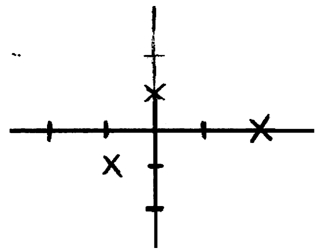
of ν (or S).

The invariance group of the Tschebyscheff and Manhattan norms in R^n is the hyperoctahedral group of permutations and sign-changes of the n objects $e_1^T, \ldots, e_n^T$. The invariance group of the Euclidean norm in R^n is somewhat larger; it is the orthogonal group, the group of all orthogonal transformations in R^n .

In C^n, the group of permutations and phase changes is the invariance group of the Tschebyscheff and Manhattan norms and the group of unitary transformations that of the-Euclidean norm.

There are norms whose invariance group consists of the identity alone; e.g., the norms in R^2 generated by the sets:





The application determines whether small or large invariance groups are desirable. In most cases, however, norms have at least some invariance properties.

A set S and a norm ν are

$$\underline{\text{SYMMETRIC}} \tag{7.9}$$

if -1 is an invariance transformation:

$$-S = S \quad \text{and} \quad \nu(x) = \nu(-x) \tag{7.10}$$

(see (2.28)). A set S and a norm ν are

$$\underline{\text{STRICTLY HOMOGENEOUS}} \tag{7.11}$$

if the field K of the vector space $\mathcal{V}_K$ is the complex field or a subfield thereof and $\{\omega I: \omega \in K, |\omega| = 1\}$ is a subgroup of the invariance group:

$$\omega \in K \wedge |\omega| = 1 \Rightarrow \omega S = S \wedge \nu(\omega x) = \nu(x) . \tag{7.12}$$

As a consequence, for a strictly homogeneous norm ν:

$$\forall \sigma \in K: \nu(\sigma x) = |\sigma| \, \nu(x) \tag{7.13}$$

since any $\sigma \in K$ can be decomposed as $\sigma = \omega|\sigma|$ with $|\sigma| \geq 0$ and $|\omega| = 1$. The Euclidean, Tschebyscheff, and Manhattan norms in C^n and R^n are strictly homogeneous.

If K is the real field R, the concepts of symmetry and strict homogeneity coincide. In C^n, the norms

$$\nu(x) = \max\{|\text{Re } x_i|, |\text{Im } x_i|\} \tag{7.14}$$

$$\nu(x) = \Sigma[|\text{Re } x_i| + |\text{Im } x_i|]$$

are symmetric but not strictly homogeneous.

Finally, we may investigate how the field of values $G_S[A]$ is changed by a regular linear transformation of the generating set S .

<u>Theorem</u>: $G_{SB}[A] = G_S[BAB^{-1}]$. (7.15)

<u>Proof</u>:

$$G_{SB}[A] = \{\ell^H Ax: \ell^H \in SB, \nu_{SB}(x) = \ell^H x = 1\}$$

$$= \{\ell^H B^{-1}(BAB^{-1})Bx: \ell^H B^{-1} \in S, \nu_S(Bx) = \ell^H B^{-1} Bx = 1\}$$

$$= \{\tilde{\ell}^H (BAB^{-1})\tilde{x}: \tilde{\ell}^H \in S, \nu_S(\tilde{x}) = \tilde{\ell}^H \tilde{x} = 1\} .$$

Q.E.D.

If A is normal (unitarily diagonalizable), then there exists a linear transformation B (dependent on A!) such that the field of values of A with respect to SB is just the field of values of the diagonal matrix BAB^{-1} with respect to **S** . For the Gerschgorin field of values, we thus obtain the set of all eigenvalues; for the Toeplitz field of values, the convex hull of this set.

The Gerschgorin field of values is frequently used to locate the eigenvalues of a normal matrix if an approximate eigenvector system is available. The success of this procedure is based upon the following theorem also due to Gerschgorin:

Theorem: If the union of k Gerschgorin disks is disjoint from the remaining disks, then this union contains exactly k eigenvalues--multiplicities being counted as the multiplicities of the zeroes in the characteristic equation. (7.16)

The proof, usually using a continuity argument, seems to be outside of norm-theoretic considerations, In particular, if one Gerschgorin disk is ISOLATED from all the others, then it contains exactly one eigenvalue. We can now obtain some information about the eigenvector corresponding to this eigenvalue:

Lemma: If' $C_i[A] \cap C_k[A] = \emptyset$, then there is no eigenvector whose i-th and k-th components are dominant. (7.17)

Proof:

If the i-th and the k-th components of the eigenvector x are dominant, then $\lambda \epsilon C_i[A]$ and $\lambda \epsilon C_k[A]$ whence $\lambda \epsilon C_i[A] \cap C_k[A] \neq \emptyset$, a contradiction.

As a consequence,

Theorem: If the Gerschgorin disk $C_i[A]$ is isolated, then if contains exactly one eigenvalue λ with a corresponding eigenvector x whose i-th component is STRICTLY DOMINANT: (7.18)

$$\mu \neq i \succ |x_\mu| < |x_i| .$$

Proof:

From the Lemma, if the component x_i is dominant, then it is strictly dominant. If x_i is not dominant, then $\exists \mu \neq i$ such that x_μ is dominant and therefore $\lambda \epsilon C_\mu[A]$. But $\lambda \epsilon C_i[A]$ whence $\lambda \epsilon C_i[A] \cap C_\mu[A] \neq 0$, a contradiction.

Q.E.D.

Among linear transformations of the generating set **S**, diagonal transformations or correspondingly SIMILARITY SCALING

$$A \to DAD^{-1}; \quad a_{ik} \to \frac{d_i}{d_k} a_{ik} \tag{7.19}$$

with $D = \text{diag}(d_1, d_2, \ldots, d_n)$ and $d_\mu \neq 0$ are of special practical interest. For the Gerschgorin field of values, they leave the centers of the Gerschgorin disks fixed and change only the radii. Assume that $C_i[A]$ is isolated and let x be an eigenvector corresponding to $\lambda \epsilon C_i[A]$. Since x_i is dominant, $|x_i| > 0$ and we may set $q_\mu = |\frac{x_\mu}{x_i}| < 1, \quad \mu \neq i$. For

$$d_i = 1 \text{ and } q_\mu \leq d_\mu^{-1} < 1, \ \mu \neq i, \tag{7.20}$$

the diagonal transformation $D = \text{diag}(d_1, d_2, \ldots . d_n)$ will decrease the radius of C_i unless it is already zero:

$$\sum_{\mu \neq i} |\frac{d_i}{d_\mu} a_{i\mu}| = \sum_{\mu \neq i} d_\mu^{-1} |a_{i\mu}| < \sum_{\mu \neq i} |a_{i\mu}|$$

provided $\sum_{\mu \neq i} |a_{i\mu}| \neq 0$. But eventually, isolation of the disk C_i will be lost, at the latest when $d_\mu^{-1} = q_\mu$ for some μ since then the transformed eigenvector Dx will have dominant i-th and μ-th components and therefore $C_i[DAD^{-1}] \cap C_\mu[DAD^{-1}] \neq \emptyset$. Varga has recently discussed this problem in detail.

Diagonal scaling is of particular importance in connection with the field of values obtained from the generating set for the Manhattan norm in R^n since diagonal scaling with positive elements leaves the sign distribution of an eigenvector invariant. Thus if Theorem (6.23) can be used to prove that the eigenvalue λ corresponding to the eigenvector x lies in $T_\ell[A]$, then it can be used to prove that λ lies in $T_\ell[DAD^{-1}]$, provided the scaling is positive. However, scaling can shrink the set T_ℓ enormously:

Let $p^T = (p_1, p_2, \ldots, p_n)$ $(p_\mu \neq 0)$ be a row vector with the sign distribution ℓ^T; i.e., $p_\mu = \ell_\mu |p_\mu|$. Let $D = \text{diag}(|p_1|, |p_2|, \ldots, |p_n|)$ so that $e^T D_\ell D = \ell^T D = p^T$. Then

$$T_\ell[DAD^{-1}] = H[e^T D_\ell DAD^{-1} D_\ell e_1, \ldots, e^T D_\ell DAD^{-1} D_\ell e_n]$$
$$= H[p^T A e_1 / p^T e_1, \ldots, p^T A e_n / p^T e_n]$$
$$T_\ell[DAD^{-1}] = H[p'_1/p_1, \ldots, p'_n/p_n] \tag{7.21}$$

where

$$(p'_1, p'_2, \ldots, p'_n) = p^T A. \tag{7.22}$$

Thus, we may reformulate (6.23):

<u>Theorem:</u> Any eigenvalue λ corresponding to an eigenvector x with sign distribution ℓ^T is contained in (7.23)

$$T_p[A] := H[p'_1/p_1, p'_2/p_2, \ldots, p'_n/p_n]$$

where $p^T = (p_1, p_2, \ldots, p_n)$ is any row vector with sign pattern ℓ^T and nonzero components and $(p'_1, p'_2, \ldots, p'_n) = p^T A$.

Note that the n quotients will coincide if and only if p^T is a left eigenvector of A with the prescribed sign pattern; the better p^T approximates such a left eigenvector, the smaller $T_p[A]$ will be. Such a left eigenvector does not exist if A has two right eigenvectors with the sign pattern ℓ^T corresponding to different eigenvalues.

Example:

$$A = \begin{pmatrix} 1 & 2 \\ -4 & 7 \end{pmatrix}; \qquad \begin{matrix} \lambda = 3 \text{ and } x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \lambda = 5 \text{ and } x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{matrix}$$

For $\ell^T = (1,1)$ and $p_1 > 0$, $p_2 > 0$:
$T_p[A]$ contains both eigenvalues but cannot shrink to a point since no $(p_1, p_2)^T$ is a left eigenvector.

For $\ell^T = (1,-1)$ and $p_1 > 0$, $p_2 < 0$:

$T_p[A] = H[1 - 4\frac{p_2}{p_1}, 7 + 2\frac{p_1}{p_2}]$ shrinks to $\{3\}$ for $p_2/p_1 = -\frac{1}{2}$ and to $\{5\}$ for $p_2/p_1 = -1$ but does not contain an eigenvalue even if p_2/p_1 approaches these values since there is no right eigenvector with this sign distribution.

For non-negative matrices, $T_p[A]$ with $p_\mu > 0$ contains an eigenvalue (the Perron root) with a corresponding non-negative right eigenvector. For positive matrices, there is only one such eigenvector and therefore only one eigenvalue in $T_p[A]$.

Example:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix} ; \quad \lambda = 1,\ 4 + \sqrt{15} = 7.87298,\ 4 - \sqrt{15} = 0.12702$$

$p^T = (1, 2.4, 4.4)$	$T_p[A] = H[7.8, 7.91, 7.86]$
$p^T = (3.9, -5.6, 2.2)$	$T_p[A] = H[0.128, 0.125, 0.136]$
$p^T = (2, 1, -1)$	$T_p[A] = H[1, 1, 1]$
$p^T = (-4, 1, 1)$	$T_p[A] = H[0.5, 1, 5]$ no eigenvector with this sign distribution.

§8. Suprema and Infima in Ordered Vector Spaces.

We shall now return to the general case of vectorial norms, norms in a vector space V_K with values from a p-ordered vector space G_{K_o} with K_o a subfield of K. In order to-generate these norms by a supremum construction, we first investigate suprema (and infima) in a p-ordered vector space G_{K_o} over a ρ_o-ordered field K_o. G_{K_o} is characterized by its POSITIVITY (NON-NEGATIVITY) CONE:

Theorem: The set G^+ of all non-negative elements of G_{K_o} (8.1)

$$G^+ := \{x \in G_{K_o} : \phi \ \rho \ x\} \tag{8.2}$$

is a CONE:--.

$$\forall\ a \in K_o,\ x \in G^+ : 0 \ \rho_o \ \alpha \succ \alpha \ x \in G^+ \tag{8.3}$$

which is

CONVEX: $x \in G^+ \wedge y \in G^+ \succ x+y \in G^+$ (8.4)

POINTED AT ϕ: $\phi \in G^+;\ x \in G^+ \wedge (-x) \in G^+ \succ x = \phi$. (8.5)

Proof:

That G^+ is a cone follows from the compatibility of multiplication by non-negative scalars with the ordering ρ((3.8)).

$$x \in G^+ \wedge y \in G^+ \succ \phi \ \rho \ x \wedge \phi \ \rho \ y$$

$\succ \phi \ \rho \ x \wedge x \ \rho \ x+y$ by compatability of ρ with addition ((2.6)

$\succ \phi \ \rho \ x+y$ by transitivity ((2.2))

$\succ x+y \in G^+$

∴ G^+ is convex.

ρ reflexive ((2.3)) $\Rightarrow \Phi \, \rho \, \Phi$
$\Rightarrow \Phi \in G^+$

$x \in G^+ \wedge (-x) \in G^+ \Rightarrow \Phi \, \rho \, x \wedge \Phi \, \rho \, (-x)$
$\Rightarrow \Phi \, \rho \, x \wedge x \, \rho \, \Phi$
$\Rightarrow x = t$ by antisymmetry ((2.4)) .

$\therefore G^+$ is pointed at Φ.

Q.E.D.

In fact, the come G^+ completely characterizes the ordering ρ :

<u>Theorem:</u> Let $G^+ \subset G_{K_0}$ be a convex cone pointed at Φ . (8.4)

Then the relation ρ defined by

$$x \, \rho \, y :\Leftrightarrow y-x \in G+ \qquad (8.7)$$

is an ordering which is compatible with vector addition and multiplication by non-negative scalars.

<u>Proof:</u>

G^+ pointed at $\Phi \Rightarrow x-x = \Phi \in G^+$
$\Rightarrow x \, \rho \, x$ (reflexivity).

$x \, \rho \, y \wedge y \, \rho \, x \Rightarrow (y-x) \in G^+ \wedge -(y-x) = x-y \in G^+$
$\Rightarrow x = y$ (antisymmetry).

$x \, \rho \, y \wedge y \, \rho \, z \Rightarrow (y-x) \in G^+ \wedge (z-y) \in G^+$
$\Rightarrow (z-x) = (z-y) + (y-z) \in G^+$
$\Rightarrow x \, \rho \, z$ (transitivity).

$x \, \rho \, y \Rightarrow (y-x) \in G^+$
$\Rightarrow (y+a) - (x+a) \in G^+$
$\Rightarrow x+a \, \rho \, y+a$ (compatibility with vector addition).

$x \, \rho \, y \wedge 0 \, \rho_0 \, \alpha \Rightarrow (y-x) \in G^+ \wedge 0 \, \rho_0 \, \alpha$
$\Rightarrow \alpha(y-x) \in G^+$
$\Rightarrow \alpha x \, \rho \, \alpha y$ (compatibility with scalar multiplication).

Q.E.D.

An element x of a p-ordered set $\mathfrak{M}$ is

$$\underline{\rho\text{-MAXIMAL}}\ (\underline{\rho\text{-MINIMAL}})\text{ or simply }\underline{\text{MAXIMAL}}\ (\underline{\text{MINIMAL}}) \qquad (8.8)$$

if it has no upper (lower) bound other than itself:

$$x \rho y \succ y = x \qquad (y \rho x \succ y = x) . \qquad (8.9)$$

Elements which are upper (lower) bounds for all elements of a subset h of $\mathfrak{M}$ are called UPPER (LOWER) BOUNDS of h :

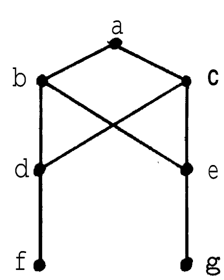
$$\mathfrak{Ub}_{\rho}(\mathfrak{h}) := \{x \in \mathfrak{M} : y \rho x, \forall y \in \mathfrak{h}\} \qquad (8.10)$$

$$\mathfrak{Lb}_{\rho}(\mathfrak{h}) := (x \in \mathfrak{M} : x \rho y, \forall y \in \mathfrak{h}\} . \qquad (8.11)$$

Either set may, of course, be empty.

Example:

For the ordering given by the Hasse diagram



the set of all upper bounds of h = {d,e} is {a,b,c} and the set of all lower bounds is empty.

Usually one is only interested in the best upper and lower bounds, best in the sense that they cannot be replaced by other bounds. Thus we define the set of <u>MINIMAL UPPER BOUNDS</u> of $h \subset \mathfrak{M}$

$$\mathrm{Sup}_\rho(h) := \{y \in \mathcal{U}b_\rho(h) : y \text{ minimal in } \mathcal{U}b_\rho(h)\} \tag{8.12}$$

and the set of <u>MAXIMAL LOWER BOUNDS</u> of h

$$\mathrm{Inf}_\rho(h) := \{y \in \mathcal{L}b_\rho(h) : y \text{ maximal in } \mathcal{L}b_\rho(h)\} . \tag{8.13}$$

In the preceding example, $\mathrm{Sup}((d,e]) = \{b,c\}$ and $\mathrm{Inf}(\{d,e\}) = \emptyset$.

In particular, we are interested in the case where all upper (lower) bounds can be replaced by one least (greatest) bound. In this case we define the

<u>LEAST UPPER BOUND</u> or <u>SUPREMUM OF h</u> (8.14)

$$a = \sup_\rho(h) :\Leftrightarrow a \in \mathcal{U}b_\rho(h) \wedge a \,\rho\, x, \forall x \in \mathcal{U}b_\rho(h) . \tag{8.15}$$

and the

<u>GREATEST LOWER BOUND</u> or <u>INFIMUM OF h</u> (8.16)

$$b = \mathrm{info}(h) :\Leftrightarrow b \in \mathcal{L}b_\rho(h) \wedge x \,\rho\, b, \forall x \in \mathcal{L}b_\rho(h) . \tag{8.17}$$

Obviously, the supremum and infimum, if they exist, are uniquely determined. Moreover,

<u>Theorem:</u> (8.18)

$$a = \sup_\rho(h) \text{ exists} \Leftrightarrow \exists a : \mathcal{U}b_\rho(h) = \{x : a \,\rho\, x\}$$

$$b = \mathrm{info}(h) \text{ exists} \Leftrightarrow \exists b : \mathcal{L}b_\rho(h) = \{x : x \,\rho\, b\} .$$

In the p-ordered vector space G_{K_o}, the set of all upper bounds of an element c is the

TRANSLATED CONE (8.19)

$$\{y : c \;\rho\; y\} = \{ y : y-c \in G^{+}\} = \{c+z : z \in G^{+}\} := c+G^{+} \quad (8.20)$$

and the set of all lower bounds is likewise $c-G^{+}$. Therefore, the set of all upper (lower) bounds of $h \subset G_{K_o}$ is an intersection of translated positivity (negativity) cones:

$$Ub_{\rho}(h) = \bigcap_{c\in h} (c+G^{+}) \; [Lb_{\rho}(h) = \bigcap_{c \in h} (c-G^{+})] \,. \quad (8.21)$$

As an immediate consequence of Theorem (8.18),

Theorem: (8.22)

$$a = \sup_{\rho}(h) \text{ exists } \Leftrightarrow \exists a\in G_{K_o} : \bigcap_{c\in h} (c+G^{+}) = a+G^{+} \quad (8.23)$$

$$b = \inf_{\rho}(h) \text{ exists } \Leftrightarrow \exists b\in G_{K_o} : \bigcap_{c\in h} (c-G^{+}) = b-G^{+} \quad (8.24)$$

Moreover, if $\sup_{\rho}(h)$ exists, then $\inf_{\rho}(-h)$ and $\sup_{\rho}(h+a)$ exist and

INVOLUTION:

$$\inf_{\rho}(-h) = -\sup_{\rho}(h) \quad (8.25)$$
$$\sup_{\rho}(-h) = -\inf_{\rho}(h)$$

TRANSLATION-COVARIANCE:

$$\sup_{\rho}(h+a) = \sup_{\rho}(h) + a \quad (8.26)$$
$$\inf_{\rho}(h+a) = \inf_{\rho}(h) + a \,.$$

Theorem (8.22) shows that a rather heavy restriction is imposed on the ordering of the vector space (to be precise, on the defining positivity cone) if the supremum of even two elements should exist. In $(R^3,\leq)$, it is intuitively clear that circular and ellipsoidal cones fail to meet this restriction (an intersection of such cones is not necessarily a cone); in fact, suprema and infima will only exist in general in $(R^3,\leq)$ if the positivity cone is triangular.

If we require of an ordered set $\mathfrak{M}$ that the supremum and infimum of any two elements (and therefore of any finite number of elements) exist, then $\mathfrak{M}$ is a

$$\underline{\text{LATTICE}} \tag{8.27}$$

and we write

$$a \sqcup b := \sup\{a,b\} \qquad (\text{"a cup b"}) \tag{8.28}$$

$$a \sqcap b := \inf\{a,b\} \qquad (\text{"a cap b"}) \ . \tag{8.29}$$

A p-ordered vector space G_{K_o} is a

<u>LATTICE-ORDERED VECTOR SPACE</u> or simply a <u>VECTOR LATTICE</u> (8.30)

if it is a lattice with respect to the ordering ρ. By Theorem (8.22) this is equivalent to

$$\forall\ a,\ b \in G_{K_o}\ \exists\ c,\ d \in G_{K_o} : \begin{array}{l} (a+G^+) \cap (b+G^+) = c+G^+ \\ (a-G^+) \cap (b-G^+) = d-G^+ \end{array} \tag{8.31}$$

Moreover,

<u>Theorem</u>: G_{K_o} is a vector lattice if and only if (8.32)

$$\forall\ a \in G_{K_o} : a^+ := \sup\{a,\phi\} \ (\text{"positive part"}) \text{ exists.}$$

<u>Proof</u>:

Involution and translation-covariance can be expressed in lattice notation by

<u>INVOLUTION</u>:

$$\begin{array}{l} (-a) \sqcap (-b) = -(a \sqcup b) \\ (-a) \sqcup (-b) = -(a \sqcap b) \end{array} \tag{8.33}$$

<u>TRANSLATION-COVARIANCE</u>:

$$\begin{array}{l} (a+c) \sqcup (b+c) = (a \sqcup b)+c \\ (a+c) \sqcap (b+c) = (a \sqcap b)+c \end{array} \tag{8.34}$$

Therefore,

$$a \sqcup b = a + [(a \sqcup b) + (-a)]$$
$$= a + [\phi \sqcup (b-a)]$$
$$= a + \sup\{\phi, b-a)$$
$$a \sqcup b = a + (b-a)^{+} = b + (a-b)^{+} \qquad (8.35)$$

Similarly,

$$a \sqcap b = a + [(a \sqcap b) - a]$$
$$= a - [(-a) \sqcup (-b) + a]$$
$$= a - [\phi \sqcup (a-b)]$$
$$a \sqcap b = a - (a-b)^{+} = b - (b-a)^{+} \qquad (8.36)$$

∴ The supremum and infimum of two elements can be expressed in terms of the positive part of their difference and conversely.

Q.E.D.

As a consequence,

<u>DEDEKIND'S PROPERTY</u>: $a \sqcup b + a \sqcap b = a + b$. (8.37)

The following result characterizes the vector lattice G_{K_o} in terms of its positivity cone:

<u>Theorem</u>: In a vector lattice G_{K_o}, every element is a difference of two non-negative elements: $a = a^{+} - a^{-}$ where $a^{-} := (-a)^{+}$. That is, $G_{K_o} = G^{+} - G^{+}$. (8.38)

<u>Proof</u>:

Taking b to be ϕ in Dedekind's property,

$$a = a + \phi = a \sqcup \phi + a \sqcap \phi = a \sqcup \phi - (-a) \sqcup \phi = a^{+} - a^{-} .$$

Q.E.D.

Moreover, a^+ and $a-$ are <u>DISJOINT</u>:

$$\begin{aligned} a^+ \sqcap a^- &= (a \sqcup \Phi) \sqcap (-a \sqcup \Phi) \\ &= [a+(-a \sqcup \Phi)] \sqcap (-a \sqcup \Phi) \\ &= a \sqcap \Phi + (-a \sqcup \Phi) \\ &= a \sqcap \Phi - a \sqcap \Phi \\ a^+ \sqcap a^- &= \mathbf{4} \end{aligned}$$

Other properties of vector lattice operations are

<u>IDEMPOTENCE</u>:	$a \sqcup a = a$	$a \sqcap a = a$	(8.39)
<u>COMMUTATIVITY</u>:	$a \sqcup b = b \sqcup a$	$a \sqcap b = b \sqcap a$	(8.40)
<u>ABSORPTIVITY</u>:	$a \sqcap (a \sqcup b) = a$	$a \sqcup (a \sqcap b) = a$	(8.41)
<u>ASSOCIATIVITY</u>:	$a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c$	$a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c$.	(8.42)

Moreover, a vector lattice is

<u>DISTRIBUTIVE</u>:

$$a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c)$$
$$a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c) . \qquad (8.43)$$

<u>Proof</u>:

The proofs of idempotence, commutativity, absorptivity, and associativity are straight-forward applications of the definitions of supremum and infimum. The proof of distributivity is more difficult:

$$a \;\rho\; (a \sqcup b) \wedge a \;\rho\; (a \sqcup c) \succ a \;\rho\; (a \sqcup b) \sqcap (a \sqcup c)$$

$$(b \sqcap c) \;\rho\; b \wedge (b \sqcap c) \;\rho\; c \succ (b \sqcap c) \;\rho\; (a \sqcup b \wedge (b \sqcap c) \;\rho\; (a \sqcup c)$$

$$\succ (b \sqcap c) \;\rho\; (a \sqcup b) \sqcap (a \sqcup c)$$

$$\therefore a \sqcup (b \sqcap c) \;\rho (a \sqcup b) \sqcap (a \sqcup c)$$

$$(a \sqcup b) \sqcap (a \sqcup c) = [a+(b-a)^+] \sqcap [a+(c-a)^+]$$
$$= a+\left[(b-a)^+ \sqcap (c-a)^+\right]$$
$$a \sqcup (b \sqcap c) = a+[(b \sqcap c) - a]^+$$
$$= a+[(b-a) \sqcap (c-a)]^+$$

Hence to prove that $(a \sqcup b) \sqcap (a \sqcup c) \,\rho\, a \sqcup (b \sqcap c)$, it sufficies to prove that $f^+ \sqcap g^+ \,\rho\, (f \sqcap g)^+$.

$$f^+ \sqcap g^+ = (f \sqcup \phi) \sqcap (g \sqcup \phi) + (f \sqcap \phi) \sqcap (g \sqcap \phi) - (f \sqcap g \sqcap \phi)$$
$$= [(f \sqcup \phi) + (f \sqcap \phi)] \sqcap [(f \sqcup \phi) + (g \sqcap \phi)]$$
$$\sqcap [(g \sqcup \phi) + (f \sqcap \phi)] \sqcap [(g \sqcup \phi) + (g \sqcap \phi)]$$
$$- (f \sqcap g \sqcap \phi)$$
$$\rho\ f \sqcap g - (f \sqcap g) \sqcap \phi$$
$$= (f \sqcap g) \sqcup \phi$$
$$= (f \sqcap g)^+$$
$$\therefore a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c).$$

The proof of the second distributive law is analogous.

Q.E.D.

Related to this is the cancellation law

$$x \sqcup y_1 = x \sqcup y_2 \wedge x \sqcap y_1 = x \sqcap y_2 \succ y_1 = y_2 .$$

which follows immediately from Dedekind's property (indeed, the assumptions give $x+y_1 = x+y_2$) . Another useful result is

$$a \,\rho\, b \succ \begin{cases} a \sqcup c \,\rho\, b \sqcup c \\ a \sqcap c \,\rho\, b \sqcap c \\ a^+ \,\rho\, b^+ \end{cases} \qquad (8.44)$$

Proof:

$$a \sqcup c \;\rho\; (a \sqcup c) \sqcup b = (a \sqcup b) \sqcup c = b \sqcup c \text{ since } a \sqcup b = b .$$

The remainder of the proof is analogous.

Examples:

(i) The real field is a well-known though trivial example of a vector lattice.

(ii) R^n is a vector lattice under the componentwise ordering $<$ of (2.11). The positivity cone is the set of all vectors with non-negative components, the "full first orthant." It is intuitively clear that the intersection of two translated orthants is again a translated orthant. Indeed,

$$a^+ = \sup\{a, \phi\} = \begin{pmatrix} \max(a_1, 0) \\ \vdots \\ \max(a_n, 0) \end{pmatrix}$$

and every finite or infinite set of elements has a supremum.

(iii) R^n is not a vector lattice under the ordering ρ defined by

$$x \;\rho\; y :\Leftrightarrow (\forall\, i : x_i < y_i) \vee (\forall\, i : x_i = y_i) .$$

The positivity cone is the set of all vectors with positive components together with the origin ϕ, the 'strict first orthant.' However, the intersection of two translated cones is in general a translated cone minus the point of that cone.

(iv) R^n is a vector lattice under the ordering ρ defined by

$$\begin{aligned} x \;\rho\; y :\Leftrightarrow\; & (x_1 < y_1) \vee (x_1 = y_1 \wedge x_2 < y_2) \\ & \vee (x_1 = y_1 \wedge x_2 = y_2 \wedge x_3 < y_3) \ldots \\ & \vee (x_1 = y_1 \wedge x_2 = y_2 \wedge \ldots \wedge x_n = y_n) \end{aligned}$$

("lexicographic" or "telephone book" ordering).

However, it has the property that there exist elements a and b such that a is "incomparably smaller" than b $(a \ll b)$:

$$na \, \rho \, b, \; \forall n > 1 \; .$$

For example, $\binom{0}{1} \ll \binom{1}{0}$ in (R^2, ρ) .

(v) C[0,1] is a vector lattice with the ordering defined by

$$f \, \rho \, g \; :\Leftrightarrow \forall x \in [0,1]: \; f(x) \leq g(x) \; .$$

The positivity cone is the set of all non-negative continuous functions on [0,1] .

Examples (ii) and (iv) are prototypes for all finite dimensional vector lattices over the real field. Mannos (1942) has shown that any n-dimensional vector lattice G_R is isomorphic to R^n with an ordering built up by direct union

$$(g, h) \, \rho \, (g', h') \; :\Leftrightarrow (g \, \rho_g \, g') \wedge (h \, \rho_h \, h')$$

and lexicographic union

$$(g, h) \, \rho \, (g', h') \; :\Leftrightarrow (g \, \rho_g \, g' \wedge g \neq g') \vee (g = g' \wedge h \, \rho_h \, h')$$

of the orderings of subspaces. If we require our ordered vector space to satisfy

$$(\forall \, \alpha \in K_o \; : \alpha \, a \, \rho \, b) \succ a = \phi$$

then lexicographic union is excluded in the construction of ρ and G_R is isomorphic to $(R^n, <)$ of Example (ii), with some one-to-one affine mapping of the full first orthant as its positivity cone.

More generally, we shall call any ordered vector space G_{K_o} which has no incomparably small elements

ARCHIMEDEAN: $(\forall \alpha \in K_0 : \alpha a \rho b) \succ a = \phi$. (8.45)

An even stronger property is

STRONGLY ARCHIMEDEAN ("integrally closed"): (8.46)

$$(\forall \alpha \in K_0 : 0 \rho_0 \alpha \succ \alpha a \rho b) \succ a \rho \phi .$$

Indeed, every strongly Archimedean ordered vector space is Archimedean.

Proof:

Assume that $\forall \alpha \epsilon K_0 : \alpha a \rho b$. If $0 \rho_0 \alpha$, then $\alpha a \rho b$. If $0 \rho_0 (-\alpha)$, then $a(-a) = (-\alpha)a \rho b$. From the strong Archimedean property, $a \rho \phi$ and $-a \rho \phi$ whence $a = \phi$.

Q.E.D.

The converse is not true in general. However,

Theorem: If G_{K_0} is a vector lattice, then G_{K_0} is strongly Archimedean if and only if G_{K_0} is Archimedean. (8.47)

Proof:

Assume that G_{K_0} is Archimedean. If $0 \rho_0 \alpha$ and $\alpha a \rho b$, then $\alpha a^+ = (\alpha a)^+ \rho b^+$. If $\alpha \rho_0 0$ and $a a \rho b$, then $\alpha a^+ \rho \phi \rho b^+$. Thus $\forall \alpha \epsilon K_0$: $\alpha a^+ \rho b^+$. From the Archimedean property, $a+ = \phi$ whence $a \rho \phi$.

Q.E.D.

To continue the discussion for the finite dimensional case, every finite dimensional Archimedean vector lattice over the real field is isomorphic to $(R^n, \leq)$, the ordering being generated by the full first orthant. The only cones which make R^n an Archimedean vector lattice are deformed full orthants or SIMPLICIAL CONES. Such a cone is the set of all convex combinations of n linearly independent vectors and non-negative multiples thereof.

Another difficulty with vector lattices is that there are always sets of elements for which no supremum exists:

Theorem: The set of all multiples of a nonzero element x (8.48)

$$\mathcal{F} := (\alpha x : \alpha \in K_o)$$

has no supremum.

Proof:

If the vector lattice is Archimedean, then not even upper bounds exist. In general, however, if $s = \sup(\mathcal{F})$ exists, then $\mathcal{F} = x + \mathcal{F}$ and

$$s = \sup(\mathcal{F}) = \sup(x + \mathcal{F}) = x + \sup(d) = x + s$$

whence $x = \phi$, a contradiction.

Q.E.D.

Thus we can only ask for the existence of the supremum and infimum of a set of elements if that set is BOUNDED, that is, has a lower bound and an upper bound. Therefore we define a vector lattice to be

COMPLETE (8.49)

if every non-empty bounded set has a supremum (and by involution, an infimum). As in the case of the real numbers, we can remove this restriction by enlarging the vector lattice G_{K_o} to the EXTENDED VECTOR LATTICE $G^*_{K_o}$ with two additional elements, $-\infty$ and $+\infty$:

$$\forall x \in G_{K_o}: \quad -\infty \,\rho\, x \,\rho\, +\infty \tag{8.50}$$

$$\inf \emptyset := +\infty \; ; \quad \sup \emptyset := -\infty \, . \tag{8.51}$$

Then every set $h \subset G^*_{K_o}$ has a supremum and infimum:

If h not bounded from below: $\inf(h) = -\infty$

If h not bounded from above: $\sup(h) = +$ co .

Of course $G^*_{K_o}$ is not a vector space since $+\infty + (-\infty)$ is not defined.

Completeness will only be needed to assure the existence of suprema and infima of infinite sets. As a consequence of completeness,

Theorem: A complete vector lattice is strongly Archimedean. (8.52)

Proof:

Assume-that $\forall\ \alpha \in K_o: 0\ \rho_o\ \alpha \succ \alpha\, a\ \rho\ b$. Then $h = \{\alpha\, a:\ 0\ \rho_o\ \alpha\}$ is bounded above whence $c = \sup(h)$ exists. But

$$\begin{aligned} c + a &= \sup\{(\alpha + 1)a:\ 0\ \rho_o\ \alpha\} \\ &= \sup\{\beta\, a:\ 1\ \rho_o\ \beta\} \\ &\rho\ \sup\{\beta\, a:\ 0\ \rho_o\ \beta\} \\ &= c \end{aligned}$$

whence $a.\ \rho\ \phi$.

Q.E.D.

The *vector* lattice $(R^n, <)$ of Example (ii) is complete. Therefore it is the only n-dimensional complete vector lattice over the real field up to isomorphism.

Exercise: Prove that $(a+b)^+\ \rho\ a^+ + b^+$.

§9. Subadditive Mappings Generated by a Set of Linear Mappings.

As was the case with scalar norms (\$5), we can now generate norms in $\mathcal{V}_K$ with values in a complete vector lattice G_{K_o} by supremum constructions over sets of linear mappings:-

Theorem: Let $S \subset \mathrm{Hom}(\mathcal{V}_K, G_K)$ be a set of linear mappings of a vector space $\mathcal{V}_K$ into the vector space G_K and let Re: $G_K \to G_{K_o}$ be an additive, &-homogeneous mapping of G_K into the complete vector lattice G_{K_o} where K_o is a subfield of K . Then (9.1)

$$\gamma_{S,\rho}(x) :\!= \sup_\rho\{\mathrm{Re}\ \varphi(x) : \varphi \in S\}$$

is a subadditive, homogeneous mapping of $\mathcal{V}_K$ into $G^*_{K_o}$, the extended vector lattice.

Remark: If S is finite, then completeness is not necessary since G_{K_o} is a vector lattice.

Lemma: Provided that the suprema exist, (9.2)

$$\mathfrak{h}_1 \subset \mathfrak{h}_2 > \sup_\rho(\mathfrak{h}_1)\ \rho\ \sup_\rho(\mathfrak{h}_2) .$$

Proof:

Since $\sup_\rho(\mathfrak{h}_2)$ is an upper bound for all elements of $\mathfrak{h}_2$, it is an upper bound for all elements of the subset $\mathfrak{h}_1$ and therefore is an upper bound of $\mathfrak{h}_1$. But $\sup_\rho(\mathfrak{h}_1)$ is the least upper bound of $\mathfrak{h}_1$.

Q.E.D.

Lemma: Let $\mathfrak{h}_1 + \mathfrak{h}_2 := \{a+b : a \in \mathfrak{h}_1, b \in \mathfrak{h}_2\}$. Then provided that the suprema exist, (9.3)

$$\sup_\rho(\mathfrak{h}_1 + \mathfrak{h}_2) = \sup_\rho(\mathfrak{h}_1) + \sup_\rho(\mathfrak{h}_2) .$$

Proof:

$$\begin{aligned}\sup_\rho(h_1 + h_2) &= \sup_\rho\{\sup_\rho(a + h_2) : a \in h_1\} \\ &= \sup_\rho\{a + \sup_\rho(h_2) : a \in h_1\} \\ &\;\; \sup_\rho(h_1) + \sup_\rho(h_2) .\end{aligned}$$

Q.E.D.

Proof of Theorem:

Re $\varphi(x)$ is an additive, &-homogeneous mapping of V_K into G_{K_0} .

$$\begin{aligned}\gamma_{S,\rho}(x+y) &= \sup_\rho\{\mathrm{Re}\,\varphi(x) + \mathrm{Re}\,\varphi(y) : \varphi \in S\} \\ &\rho \sup_\rho\{\mathrm{Re}\,\varphi_1(x) + \mathrm{Re}\,\varphi_2(y) : \varphi_1, \varphi_2 \in S\} \\ &= \sup_\rho\{\mathrm{Re}\,\varphi_1(x) : \varphi_1 \in S\} + \sup_\rho\{\mathrm{Re}\,\varphi_2(x) : \varphi_2 \in S\} \\ &= \gamma_{S,\rho}(x) + \gamma_{S,\rho}(y)\end{aligned}$$

$$\begin{aligned}\gamma_{S,\rho}(\alpha x) &= \sup_\rho\{\mathrm{Re}\,\varphi(\alpha x) : \varphi \in S\} \\ &= \sup_\rho\{\alpha\, \mathrm{Re}\,\varphi(x) : \varphi \in S\} \qquad \forall\, \alpha \in K_0 \\ &= \alpha \cdot \sup_\rho\{\mathrm{Re}\,\varphi(x) : \varphi \in S\} \qquad \forall\, \alpha \in K_0 : 0\, \rho_0\, \alpha \\ &= \alpha\, \gamma_{S,\rho}(x)\end{aligned}$$

$\therefore \gamma_{S,\rho}$ is a subadditive, &-homogeneous mapping of V_K into $G^*_{K_0}$.

Q.E.D.

Examples:

(i) Let $V_K = G_{K_0}$. Then Re: $G_K \to G_{K_0}$ is the identity mapping. For $S = \{I,0\}$ where I is the identity and 0 the zero mapping of V_K into itself,

$$\gamma_{S,\rho}(x) = \sup_\rho\{x, \phi\} = x^+$$

is subadditive and &-homogeneous:

$$(a+b)^+ \rho\ a^+ + b^+ \tag{9.4}$$

$$0 \rho_0 \alpha > (\alpha\ a)^+ = \alpha\ a^+ . \tag{9.5}$$

(ii) Let $\mathcal{V}_K = G_{K_0}$ and let $S = (I, -1)$. Then

$$\gamma_{S,\rho}(x) = x \sqcup (-x)$$
$$= (2x \sqcup \phi) - x$$
$$= 2x^+ - (x^+ - x^-)$$
$$= x^+ + x^-$$

$$\gamma_{S,\rho}(x) = |x| := x^+ + x^- . \tag{9.6}$$

Since $\gamma_{S,\rho}$ is subadditive and &-homogeneous,

$$|a+b| \rho |a| + |b| \tag{9.7}$$

$$|\alpha\ a| = \alpha|a| \quad \forall\ \alpha \in K_0 : \quad 0 \rho_0 \alpha . \tag{9.8}$$

Moreover,

$$|a-b| = (a-b)^+ + (a-b)^-$$
$$= [(a-b)^+ + b] + [(b-a)^+ - b]$$
$$|a-b| = a \sqcup b - a \sqcap b \tag{9.9}$$

From Dedekind's property ((8.36))

$$a+b = a \sqcup b + a \sqcap b,$$

we now obtain

$$a \sqcup b = \tfrac{1}{2}[a+b + |a-b|] \tag{9.10}$$

$$a \sqcap b = \tfrac{1}{2}[a+b - |a-b|] \;. \qquad (9.11)$$

In addition to being subadditive and homogeneous, $|x|$ is non-negative as a consequence of (9.6). Indeed, it is even positive definite:

$$\begin{aligned} |a-b| = 0 &\Rightarrow (a \sqcup b) - (a \sqcap b) = 0 \\ &\Rightarrow a \sqcup b = a \sqcap b \\ &\Rightarrow a = b \\ &\Rightarrow a-b = \phi \end{aligned}$$

Thus,

$$|x| \text{ is a norm.} \qquad (9.12)$$

(iii) Let $\mathcal{V}_K = R^n$ and $G_{K_o} = (R^n, \le)$, the vector lattice generated by the full first orthant. Then the norm $|x|$ of Example (ii) is just the modulus norm (Betragsnorm) in R^n ((2.23)).

Most of the results of §5 carry over to the case of vectorial norms generated by sets of linear mappings. In particular, Theorems (5.8) and (5.10) now read:

<u>Theorem:</u> If A is a convex combination of elements of S. (9.13)
then

$$\operatorname{Re} A(x) \le \gamma_{S,\rho}(x), \qquad \forall x \in \mathcal{V}_K \;. \qquad (9.14)$$

<u>Proof:</u>

Let $A = \alpha_1\varphi_1 + \alpha_2\varphi_2 + \dots + \alpha_n\varphi_n$ where $\alpha_i \in K_o$, $\varphi_i \in S$, $0 \le \alpha_i$, and $\sum_i \alpha_i = 1$. Then

$$\begin{aligned} \operatorname{Re} A(x) &= \alpha_1 \operatorname{Re} \varphi_1(x) + \alpha_2 \operatorname{Re} \varphi_2(x) + \dots + \alpha_n \operatorname{Re} \varphi_n(x) \\ &\le [\alpha_1 + \alpha_2 + \dots + \alpha_n] \sup_\rho\{\operatorname{Re} \varphi(x) : \varphi \in S\} \\ &= \gamma_{S,\rho}(x) \;. \end{aligned}$$

Q.E.D.

Theorem: If the zero mapping 0 can be represented as a convex combination of elements of S, then $\gamma_{S,\rho}(\ \cdot\ )$ is non-negative and therefore a seminorm. (9.15)

Example:

(iv) Let $V_K = C^n$ and $G_{K_0} = (R^n, <)$ as in Example (iii) and let $S = \{\alpha\ I: |\alpha| = 1)$. Then

$$|x| := \gamma_S(x) = \sup\{\text{Re}\ (\alpha\ x)\ :\ |\alpha| = 1)$$

is subadditive and strictly homogeneous. By the preceding theorem, it is non-negative $[0 = \&(I) + \frac{1}{2}(-I)]$; positive definiteness follows from its explicit representation. Thus $|x|$ is--a norm, the modulus norm in C^n .

As in §5, we may introduce the sets

$$K_p := (x \in V_K : \gamma_{S,\rho}(x)\ \rho\ p\}, \qquad p \in G_{K_0} . \tag{9.16}$$

$K_\phi = (x: \gamma_{S,\rho}(x)\ \rho\ \phi\}$ is again a cone and

Theorem: A seminorm $\gamma_{S,\rho}(x)$ is definite $\Leftrightarrow K_\phi = \{\phi\}$. (9.17)

We can still represent K_p as an intersection of domains

$$\mathcal{L}_{\varphi,p} := (x \in V_K : \text{Re}\ \varphi(x)\ \rho\ p\}: \tag{9.18}$$

Theorem: $K_p = \bigcap_{\varphi \in S} \mathcal{L}_{\varphi,p}$. (9.19)

However, $\mathcal{L}_{\varphi,p}$ is no longer a half-plane:

For $\varphi = I$, $\mathcal{L}_{\varphi,p} = \{x:\ x\ \rho\ p) = p - G+$.

For $\varphi = -1$, $\mathcal{L}_{\varphi,p} = \{x:\ -x\ \rho\ p) = p + G^+$.

A set G is

<u>CONVEX</u> (9.20)

if

$$V\ a,\ b \in G: \quad \mu a + (1-\mu)b \in G,\ 0 \rho_o \mu \rho_o 1\ . \tag{9.21}$$

<u>Theorem</u>: $\mathcal{L}_{\varphi,p}$ is convex. (9.22)

<u>Proof</u>:

Let $a, b \in \mathcal{L}_{\varphi,p}$ and $0 \rho_o \mu \rho_o 1$. Then $\text{Re}\ \varphi(a)\ \rho\ p \wedge \text{Re}\ \varphi(b)\ \rho\ p$

and

$$\begin{aligned} \text{Re}\ \varphi(\mu\ a + (1-\mu)b) &= \mu\ \text{Re}\ \varphi(a) + (1-\mu)\text{Re}\ \varphi(b) \\ &\rho\ \mu\ p + (1-\mu)\ p \\ &= p \end{aligned}$$

since $0 \rho_o \mu$ and $0 \rho_o 1-\mu$. Therefore $\mu\ a + (1-\mu)b \in \mathcal{L}_{\varphi,p}$.

Q.E.D.

In general,

$$\mathcal{L}_{\varphi,p} = \{x : \text{Re}\ \varphi(x) \in p-G^+\} = (\text{Re}\ \varphi)^{-1}(p-G^+),$$

the preimage of the translated cone p-G+ . Letting $\hat{\psi}$ denote the one-to-one mapping of $\mathcal{V}_K/\text{Ker}(\text{Re}\ \varphi)$ into $\text{Re}\ \varphi(\mathcal{V}_K)$ induced by Re φ,

$$\mathcal{L}_{\varphi,p}/\text{Ker}(\text{Re}\ \varphi) = \hat{\psi}^{-1}(\text{Re}\ \varphi(\mathcal{V}_K) \cap (p-G+))\ .$$

If G_{K_o} is a finite-dimensional, Archimedean vector lattice, we might expect the domains $\mathcal{L}_{\varphi,p}$ to be intersections of half-spaces rather than half-spaces themselves. Indeed, if $\mathcal{V}_K$ is a vector space over the real or complex field and $G_{K_o} = (R^m, \leq)$, then

$$\mathcal{L}_{\varphi,p} = \{x: \ \operatorname{Re} \varphi(x) \le p\}$$

$$= \{x: \ e_i^T \operatorname{Re} \varphi(x) \le e_i^T p, \ i = 1, 2, \ldots, m)$$

$$= \bigcap_{i=1}^{m} \{x: \ \operatorname{Re} e_i^T \varphi(x) \le e_i^T p)$$

$$\mathcal{L}_{\varphi,p} = \bigcap_{i=1}^{m} H_{\operatorname{Re} e_i^T \varphi, \ e_i^T p} \tag{9.23}$$

and therefore

$$K_p = \bigcap_{\varphi \epsilon \mathbf{S}} \mathcal{L}_{\varphi,p} = \bigcap_{i=1}^{m} \bigcap_{\varphi \epsilon \mathbf{S}} H_{\operatorname{Re} e_i^T\varphi, \ e_i^T p} \ . \tag{9.24}$$

This result is obviously a consequence of the fact that $(R^m, \le)$ is a direct union o-f the linearly ordered real field, and we shall now likewise investigate this effect on the mapping $\gamma_{\mathbf{S}}(x)$.

Let $\mathcal{V}_K$ be a vector space over the real or complex field and let $G_{K_o} = (R^m, \le)$. Then it is easily seen that each component of a subadditive, homogeneous mapping of $\mathcal{V}_K$ into G_{K_o} is itself subadditive and homogeneous. If the mapping is a norm generated by a set **S**, then each component is a bounded seminorm or even a norm; moreover

<u>Theorem:</u>

$$\gamma_{\mathbf{S}}(x) = \begin{pmatrix} \gamma_{\mathbf{S}_1}(x) \\ \vdots \\ \gamma_{\mathbf{S}_m}(x) \end{pmatrix} \tag{9.25}$$

where

$$S_i = \{e_i^T\varphi: \varphi \in S\} \subset \mathcal{V}_K^D \ . \tag{9.26}$$

<u>Proof:</u>

$$\gamma_{\mathbf{S}}(x) = \sup\{\operatorname{Re} \varphi(x) : \varphi \in S)$$

$$= \begin{pmatrix} e_1^T \sup \operatorname{Re} \varphi(x) \\ \vdots \\ e_m^T \sup \operatorname{Re} \varphi(x) \end{pmatrix}$$

$$= \begin{pmatrix} \sup \operatorname{Re} e_1^T \varphi(x) \\ \vdots \\ \sup \operatorname{Re} e_m^T \varphi(x) \end{pmatrix} \qquad \text{since } e_i^T \sup(h) = \sup(e_i^T h) \text{ in } (R^m, \leq)$$

$$= \begin{pmatrix} \gamma_{\mathbf{S}_1}(x) \\ \vdots \\ \gamma_{\mathbf{S}_m}(x) \end{pmatrix}$$

Q.E.D.

In the literature, only a special case of this result has been studied: the case where

The i-th component of the norm $\nu(x)$ is a norm on the subspace $\mathcal{V}_K^{(i)} = P_i \mathcal{V}_K$. (9.27)

Examples:

(i) Let $\mathcal{V}_K = R^2$ and $G_{K_o} = (R^2, <)$. Then the sets

$$\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\} \text{ and } \{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\}$$

give rise to the same sets $\mathbf{S}_1 = \{\pm(1,0)\}$ and $\mathbf{S}_2 = \{\pm(0,1)\}$ and therefore generate the same norm. Note, however, that the second set does not generate the first set by convex combination.

(ii) Let $\mathcal{V}_K = R^n$ or C^n and $G_{K_o} = (R^n, <)$ with $\nu(x)$ the modulus norm in $\mathcal{V}_K$. Then each component of $\nu(x)$ is a norm on the subspace formed by all scalar multiples of a coordinate axis.

(iii) Let $\mathcal{V}_K = R^3$ and $G_{K_o} = (R^2, <)$ with the norm

$$\nu\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} \max(|x_1|, x_2) \\ \max(-x_2, |x_3|) \end{pmatrix} .$$

V(x) is generated by

$$S = \{\pm\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1 & 0\\ 0 & -1 & 0\end{pmatrix}\}$$

which gives rise to the sets

$$S_1 = \{(1\ 0\ 0),\ (-1\ 0\ 0\ ),\ (0\ 1\ 0)\}$$

$$S_2 = \{(0\ -1\ 0),\ (0\ 0\ 1),\ (0\ 0\ -1))$$

In this example, neither $y_{S_1}(x)$ nor $y_{S_2}(x)$ is a norm though both are bounded seminorms.

Theorem: Let y be a symmetric seminorm on $\mathcal{V}_K$. Then there exists a subspace $\mathcal{U}_K \subset \mathcal{V}_K$ such that y restricted to $\mathcal{U}_K$ is definite. (9.28)

Proof:

Since y is non-negative, the cone $K_\phi = (x\colon Y(x)\ \rho\ \phi\}$ is the domain where $\gamma(x)$ vanishes. By symmetry $(\gamma(-x) = y(x))$, K_ϕ contains with every element x its negative $-x$. Therefore K_ϕ is a subspace of $\mathcal{V}_K$. Let $\mathcal{U}_K = \mathcal{V}_K/K_\phi$ and let P be the projection of $\mathcal{V}_K$ onto $\mathcal{U}_K$. If $x \in \mathcal{U}_K$ and $y(x) = \phi$, then $x \in K_\phi$ whence $Px = 0$ and $x = 0$ since $Px = x\ \forall\ x \in \mathcal{U}_K$. Thus y is definite on $\mathcal{U}_K$.

In conclusion, we note that the concept of linearly transformed norms carries over unchanged from §7 to vectorial norms, and that relation (7.3) is valid for the transformed generating set.

Exercise: Prove that $a \sqcap b\ \rho\ x\ \rho\ a \sqcup b \succ |x|\ \rho\ |a| \sqcup |b|$.

§10. Additional Remarks.

We have seen how norms can be generated from sets of linear mappings. The question may arise whether all norms are so generated. The following theorem is suggestive:

Theorem: Let y be a subadditive mapping of $\mathcal{V}_K$ into G_{K_o} . (10.1)

Define

$$S = \{\varphi : \varphi \text{ is additive; } \forall x \in \mathcal{V}_K : \varphi(x) \rho \gamma(x)\} . \quad (10.2)$$

Then $\gamma_S(x) \rho \gamma(x)$.

Proof:

$$\gamma_S(x) = \sup\{\varphi(x) : \varphi \in s)$$
$$= \sup\{\varphi(x) : \varphi \text{ is additive; } \forall \xi \in \mathcal{V}_K : \varphi(\xi) \rho \gamma(\xi)\}$$
$$\rho \gamma(x) .$$

Whether such a set S generates y(x), that is, whether the supremum is indeed y(x) for all x, depends on the topological properties of the space $\mathcal{V}_K$. In finite dimensional spaces over R and C, the support theorem guarantees that $\gamma_S(x)$ = y(x) .

A further remark concerns the basic triangular inequality (2.16):

$$\nu(x + y) \rho \nu(x) + \nu(y) .$$

Replacing x by x + y and y by -y, we obtain

$$\nu(x) - \nu(-y) \rho \nu(x + y), \quad (10.3)$$

or, combining the two inequalities,

$$4 - Y) \rho \nu(x + y) - \nu(x) \rho \nu(y) . \quad (10.4)$$

If Vis symmetric $(\nu(x) = \nu(-x))$, then

$$|\nu(x + y) - \nu(x)| \,\rho\, \nu(y) \quad . \tag{10. 5}$$

Replacing y by **-y** in (10.3), we obtain

$$\nu(x) - \nu(y) \,\rho\, \nu(x - y) \tag{10. 6}$$

Again, if ν is symmetric, then

$$|\nu(x) - \nu(y)| \,\rho\, \nu(x - y) \quad . \tag{10. 7}$$

§11. Mappings of Normed Vector Spaces

Let $\mathcal{V}_K$ and $\mathcal{V}'_K$ be normed vector spaces with norms $\nu : \mathcal{V}_K \to (G_{K_o}, \rho)$ and $\nu' : \mathcal{V}'_K \to (G'_{K_o}, \rho')$. Let A be a linear mapping of $\mathcal{V}_K$ into $\mathcal{V}'_K$. Then a linear mapping B: $G_{K_o} \to G'_{K_o}$ is an

UPPER BOUND FOR A or LIPSCHITZ BOUND (11.1)

if

$$\nu'(Ax)\ \rho'\ B\nu(x),\ \forall\ x \in \mathcal{V}_K\ . \qquad (11.2)$$

The situation is illustrated by the following diagram:

$$\begin{array}{ccc} \mathcal{V}_K & \xrightarrow{\nu} & (G_{K_o}, \rho) \\ A \downarrow & & \downarrow B \\ \mathcal{V}'_K & \xrightarrow{\nu'} & (G'_{K_o}, \rho') \end{array} \qquad (11.3)$$

A mapping $\lceil\,\rceil_{\nu',\nu} : \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) \to \mathrm{Hom}(G_{K_o}, G'_{K_o})$ is an

UPPER BOUND MAPPING (11.4)

if $B = \lceil A \rceil_{\nu',\nu}$ is an upper bound for A for all $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$.

Examples:

(i) Let $\mathcal{V}_K = \mathcal{V}'_K = R^n$ and $G_{K_o} = G'_{K_o} = (R^n, \leq)$ with $\nu(x) = \nu'(x) = |x|$, the modulus norm. Then $A \in \mathrm{Hom}(R^n, R^n)$ is an $n \times n$ matrix (a_{ij}) and an upper bound mapping is given by

$$|\overline{A}| = |A| := (|a_{ij}|)\ .$$

(ii) Let $\mathcal{V}_K = \mathcal{V}'_K = R^n$ and $G_{K_o} = G'_{K_o} = (R, <)$ with $\nu(x) = \nu'(x) = (\sum_i x_i^2)^{\frac{1}{2}}$, the Euclidean norm. Then

$$|\overline{A}| := \max_i |a_{ii}| + \Big(\sum_{i \neq k} a_{ik}^2 \Big)^{\frac{1}{2}}$$

is an upper bound mapping. Another upper bound mapping is the Frobenius norm ((16.17))

$$|\overline{A}| = \|A\|_F := \Big(\sum_i \sum_k a_{ik}^2\Big)^{\frac{1}{2}} .$$

Similarly, a linear mapping $C : G_{K_o} \to G'_{K_o}$ is a

LOWER BOUND FOR A (11.5)

if
$$Cv(x) \; \rho' v'(Ax), \; \forall \; x \in \mathcal{V}_K . \qquad (11.6)$$

A mapping $\lfloor\;\rfloor_{v',v} : \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) \to \mathrm{Hom}(G_{K_o}, G'_{K_o})$ is a

LOWER BOUND MAPPING (11.7)

if $C = \lfloor A \rfloor_{v',v}$ is a lower bound for A for all $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$.

Example:

Let $\mathcal{V}_K = \mathcal{V}'_K = R^2$ and $G_{Ko} = G'_{K_o} = (R^2, \leq)$ with $v(x) = v'(x) = |x|$, the modulus norm. Then

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \text{ has the lower bound } C = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \text{ since}$$

$$cv(x) = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} |x_1| \\ |x_2| \end{pmatrix} \leq \begin{pmatrix} |3x_1 + x_2| \\ |x_2 + 3x_1| \end{pmatrix} = |Ax| = v'(Ax) .$$

$C \equiv 0$ is also a lower bound for A and $\lfloor\;\rfloor_{v',v} \equiv 0$ is a lower bound mapping.

§12. Least Upper and Greater Lower Bounds I.

We shall first investigate the case of scalar norms; that is, $G_{K_o} = G'_{K_o} = (R, \leq)$. For a given mapping $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$, the set of all upper bounds for A has a least element, the <u>LEAST UPPER BOUND</u>:

$$\mathrm{lub}_{\nu',\nu}(A) := \inf\{\beta : \nu'(Ax) < \beta\,\nu(x), \forall\, x \in \mathcal{V}_K\}. \qquad (12.1)$$

Since $\nu'(A\phi) < \beta\,\nu(\phi)$ for all β,

$$\begin{aligned}\mathrm{lub}_{\nu',\nu}(A) &= \inf\{\beta : \nu'(Ax) \leq \beta\,\nu(x)\ \wedge\ x \neq \phi\}\\ &= \inf\{\beta : \nu'(A\tfrac{x}{\nu(x)}) \leq \beta \wedge \nu(x) \neq 0\}\\ &= \inf\{\beta : \nu'(Ax) \leq \beta \wedge \nu(x) = 1\ )\end{aligned}$$

$$\mathrm{lub}_{\nu',\nu}(A) = \sup\{\nu'(Ax) : \nu(x) = 1\} \qquad (12.2)$$

$$\mathrm{lub}_{\nu',\nu}(A) = \sup\{\frac{\nu'(Ax)}{\nu(x)} : x \neq \phi\} \qquad (12.3)$$

$\mathrm{lub}_{\nu',\nu}$ is, of course, an upper bound mapping and

$$\mathrm{lub}_{\nu',\nu}(A) \leq |\overline{A}|_{\nu',\nu},\ \forall\ A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) \qquad (12.4)$$

for all upper bound mappings $|\overline{\ \ }|_{\nu',\nu}$. Moreover,

<u>Theorem</u>: The mapping $\mathrm{lub}_{\nu',\nu} : \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) \to R$ is subadditive, homogeneous, and positive definite. (12.5)

<u>Proof</u>:

$$\begin{aligned}&\mathrm{lub}_{\nu',\nu}(A_1 + A_2)\\ &= \sup\{\nu'((A_1 + A_2)x) : \nu(x) = 1)\\ &< \sup\{\nu'(A_1x) + \nu'(A_2x) : \nu(x) = 1)\\ &< \sup\{\nu'(A_1x_1) + \nu'(A_2x_2) : \nu(x_1) = \nu(x_2) = 1)\\ &= \sup\{\nu'(A_1x_1) : \nu(x_1) = 1) + \sup\{\nu'(A_2x_2) : \nu(x_2) = 1)\\ &= \mathrm{lub}_{\nu',\nu}(A_1) + \mathrm{lub}_{\nu',\nu}(A_2).\end{aligned}$$

$\therefore \text{lub}_{\nu',\nu}$ is subadditive; homogeneity and nonnegativity are likewise inherited from ν'.

$$\text{lub}_{\nu',\nu}(A) = 0 \Rightarrow \nu'(Ax) \leq 0, \forall\ x \in \mathcal{V}_K$$

$$\Rightarrow \nu'(Ax) = 0, \forall\ x \in \mathcal{V}_K \qquad \text{since } \nu' \text{ is nonnegative}$$

$$\Rightarrow Ax = 0, \quad \forall\ x \in \mathcal{V}_K \qquad \text{since } \nu' \text{ is a norm}$$

$$\Rightarrow A \equiv 0 .$$

$\therefore \text{lub}_{\nu',\nu}$ is definite.

Q.E.D.

Note that $\text{lub}_{\nu',\nu}$ may not be bounded and therefore may not be a norm.

Example:

Let $\mathcal{V}_K = C^1[0,1]$, the space of once continuously differentiable functions on $[0,1]$, and let $\mathcal{V}'_K = C[0,1]$, the space of continuous functions on $[0,1]$. Take

$$\nu(f) = \nu'(f) = \max\{|f(x)| : 0 \leq x \leq 1\} .$$

Let $A = \frac{d}{dx}$. Then $\nexists \beta$:

$$\max|f'(x)| = \nu(Ax) \leq \beta\ \nu(x) = \beta . \max|f(x)|$$

for all $f \in C^1[0,1]$. Therefore, $\text{lub}_{\nu',\nu}(A) = +\infty$.

Any mapping $A \in \text{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ for which $\text{lub}_{\nu',\nu}(A) < +\infty$ is said to be bounded. That the set of all such mappings is a subspace of $\text{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ follows trivially from the subadditivity and homogeneity of $\text{lub}_{\nu',\nu}$. Thus:

Theorem: $\text{lub}_{\nu',\nu}$ is a norm on the subspace of all bounded mappings. (12. 6)

In this case, $\text{lub}_{\nu',\nu}$ is called the

LEAST UPPER BOUND NORM (12. 7)

subordinate to the norms ν' and ν.

In finite dimensional spaces, every mapping is bounded, independent of the norms ν' and ν used. This can be shown by a compactness argument or by Ostrowski's theorem that all norms over a finite dimensional space are topologically equivalent; that is, given norms ν_1 and ν_2 over $\mathcal{V}_K$, there exists a constant τ such that

$$\forall\, x \in \mathcal{V}_K \;\; \nu_1(x) < \tau \cdot \nu_2(x) \, .$$

Thus, the proof is reduced to the case where ν and ν' are the maximum norm over $\mathcal{V}_K$ and $\mathcal{V}_K'$ and follows by using the product topology.

For a given mapping $A \in \text{Hom}(\mathcal{V}_K,\mathcal{V}_K')$, the set of all lower bounds for A has a greatest element, the <u>GREATEST LOWER BOUND</u>:

$$\text{glb}_{\nu',\nu}(A) := \sup\{\gamma\colon \gamma\nu(x) \leq \nu'(Ax),\ \forall\ x \in \mathcal{V}_K\} \, . \qquad (12.\ 8)$$

$$= \inf\{\nu'(Ax)\colon \ \nu(x) = 1\} \qquad (12.\ 9)$$

$$= \inf\{\frac{\nu'(Ax)}{\nu(x)}\colon \ x \neq \Phi\} \, . \qquad (12.10)$$

$\text{glb}_{\nu',\nu}$ is, of course, a lower bound mapping and

$$\underline{A}|_{\nu',\nu} \leq \text{glb}_{\nu',\nu}(A), \quad \forall\ A \in \text{Hom}(\mathcal{V}_K,\mathcal{V}_K') \qquad (12.11)$$

for all lower bound mappings $\lfloor\;\rfloor_{\nu',\nu}$. Thus, for all $x \in \mathcal{V}_K$ with $x \neq \Phi$, we may bound the <u>MAPPING DISTORTION</u> $\frac{\nu'(Ax)}{\nu(x)}$:

$$\lfloor A\rfloor_{\nu',\nu} \leq \text{glb}_{\nu',\nu}(A) \leq \frac{\nu'(Ax)}{\nu(x)} \leq \text{lub}_{\nu',\nu}(A) \leq |\overline{A}|_{\nu',\nu} \, .$$

The mapping $\text{glb}_{\nu',\nu}$ is homogeneous and nonnegative but neither **subadditive** nor definite.' Indeed, it is not even superadditive $(\text{glb}_{\nu',\nu}(A_1) + \text{glb}_{\nu',\nu}(A_2) \leq \text{glb}_{\nu',\nu}(A_1+A_2))$ as one might expect.

Theorem: If A is not injective $(\text{Ker } A \neq \{\phi\})$, then (12.12)
$\text{glb}_{\nu',\nu}(A) = 0$.

Proof:

$$\text{Ker } A = \{x \in \mathcal{V}_K : Ax = \phi\} \neq \emptyset \Rightarrow \exists x \neq \phi : Ax = \phi$$
$$\Rightarrow \exists x \neq \phi : \frac{\nu'(Ax)}{\nu(x)} = 0$$
$$\Rightarrow \text{glb}_{\nu',\nu}(A) = 0 \text{ .}$$

If A is injective (but not necessarily surjective), then a left inverse A^{-1} exists on $A\mathcal{V}_K \subset \mathcal{V}'_K$ and

$$\text{glb}_{\nu',\nu}(A) = \inf\{\frac{\nu'(Ax)}{\nu(x)} : x\in\mathcal{V}_K \wedge x \neq \phi\}$$
$$= \inf\{\frac{\nu'(Ax)}{\nu(A^{-1}Ax)} : x\in\mathcal{V}_K \wedge x \neq \phi\}$$
$$= \inf\{\frac{\nu'(y)}{\nu(A^{-1}y)} : y\in A\mathcal{V}_K \wedge y \neq \phi\}$$
$$= 1/\sup\{\frac{\nu(A^{-1}y)}{\nu'(y)} : y\in A\mathcal{V}_K \wedge y \neq \phi\}$$
$$\geq 1/\sup\{\frac{\nu(A^{L}y)}{\nu'(y)} : y\in\mathcal{V}'_K \wedge y \neq \phi\}$$
$$= 1/\text{lub}_{\nu,\nu'}(A^L)$$

where $A^L \in \text{Hom}(\mathcal{V}'_K, \mathcal{V}_K)$ is any mapping which coincides with A^{-1} on $A\mathcal{V}_K$, an extended LEFT INVERSE of A:

$$A^L Ax = x, \quad \forall\, x \in \mathcal{V}_K \text{ .}$$

Theorem: If A is injective and A^L is any left inverse (12.13)
of A, then $\text{glb}_{\nu',\nu}(A) > 1/\text{lub}_{\nu,\nu'}(A^L)$.

Moreover, if A^L is a bounded mapping, then $\mathrm{lub}_{\nu,\nu'}(A^L) < +\infty$ and $\mathrm{glb}_{\nu',\nu}(A) \geq 1/\mathrm{lub}_{\nu,\nu'}(A^L) > 0$.

In finite dimensional spaces, all linear mappings are bounded and any injective mapping has at least one left inverse. Therefore A is injective if and only if $\mathrm{glb}_{\nu',\nu}(A) > 0$.

<u>Theorem</u>: If A is regular (injective and surjective), then (12.14)

$$\mathrm{glb}_{\nu',\nu}(A) = 1/\mathrm{lub}_{\nu,\nu'}(A^{-1}) .$$

<u>Proof</u>:

If A is regular, then $A\mathcal{V}_K = \mathcal{V}'_K$ and A^{-1} is uniquely determined. Thus, we may sharpen the proof of Theorem (12.12)

$$\begin{aligned}\mathrm{glb}_{\nu',\nu}(A) &= \inf\{\frac{\nu'(y)}{\nu(A^{-1}y)} : y \epsilon \mathcal{V}'_K \wedge y \neq \phi\} \\ &= 1/\sup\{\frac{\nu(A^{-1}y)}{\nu'(y)} : y \epsilon \mathcal{V}'_K \wedge y \neq \phi\} \\ &= 1/\mathrm{lub}_{\nu,\nu'}(A^{-1}) .\end{aligned}$$

Q.E.D.

If $\mathcal{V}_K = \mathcal{V}'_K$, then the mapping $A : \mathcal{V}_K \to \mathcal{V}'_K$ is an endomorphism and may be injective yet not surjective. In finite dimensional spaces, a dimension argument shows that this situation cannot occur and we obtain a nonsingularity criterion:

<u>Theorem</u>: If A is an endomorphism of a finite dimensional vector space, then (12.15)

$$\mathrm{glb}_{\nu',\nu}(A) = \begin{cases} 1/\mathrm{lub}_{\nu,\nu'}(A^{-1}) & \text{if A nonsingular} \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathcal{V}_K = \mathcal{V}'_K$ and $\nu=\nu'$, then the greatest lower bound and the least upper bound of the identity endomorphism are given by:

$$\mathrm{glb}_{\nu',\nu}(I) = \mathrm{lub}_{\nu',\nu}(I) = 1 .$$

Although $\text{glb}_{\nu',\nu}$ is not subadditive,

Theorem: $\text{glb}_{\nu',\nu}(A_1 + A_2) < \text{glb}_{\nu',\nu}(A_1) + \text{lub}_{\nu',\nu}(A_2)$ (12.16)

Proof:

For any $\varepsilon > 0$, there exists $\xi \in \mathscr{V}_K$ with $\nu(\xi) = 1$ such that $\text{glb}_{\nu',\nu}(A_1) = \nu'(A_1\xi) - \varepsilon$. Therefore,

$$\begin{aligned}\text{glb}_{\nu',\nu}(A_1 + A_2) &\leq \nu'((A_1 + A_2)\xi) \\ &< \nu'(A_1\xi) + \nu'(A_2\xi) \\ &= \text{glb}_{\nu',\nu}(A_1) + \varepsilon + \nu'(A_2\xi) \\ &< \text{glb}_{\nu',\nu}(A_1) + \text{lub}_{\nu',\nu}(A_2) + \varepsilon\end{aligned}$$

Q.E.D.

From (12.5) and (12.16) we obtain a result analagous to that derived for vector norms (see (10.4)):

$$\text{lub}_{\nu',\nu}(A) - \text{lub}_{\nu',\nu}(-B) < \text{lub}_{\nu',\nu}(A+B) \leq \text{lub}_{\nu',\nu}(A) + \text{lub}_{\nu',\nu}(B) \qquad (12.17)$$

$$\text{glb}_{\nu',\nu}(A) - \text{lub}_{\nu',\nu}(-B) < \text{glb}_{\nu',\nu}(A+B) \leq \text{glb}_{\nu',\nu}(A) + \text{lub}_{\nu',\nu}(B) \, . \qquad (12.18)$$

If ν and/or ν' is symmetric, then these relations simplify to

$$|\text{lub}_{\nu',\nu}(A+B) - \text{lub}_{\nu',\nu}(A)| \leq \text{lub}_{\nu',\nu}(B) \qquad (12.19)$$

$$|\text{glb}_{\nu',\nu}(A+B) - \text{glb}_{\nu',\nu}(A)| \leq \text{lub}_{\nu',\nu}(B) \qquad (12.20)$$

as a consequence of the following

Theorem: if ν and/or ν' is symmetric, then $\text{lub}_{\nu',\nu}$ is symmetric. (12.21)

Proof:

$$\nu(x) = \nu(-x) > \text{lub}_{\nu',\nu}(-A) = \sup\{\nu'(-Ax) : \nu(x) = 1)$$
$$= \sup\{\nu'(A(-x)) : \nu(-x) = 1)$$
$$= \sup(\nu'(Ay) : \nu(y) = 1)$$
$$= \text{lub}_{\nu',\nu}(A) .$$

$$\nu'(-x) = \nu'(x) > \text{lub}_{\nu',\nu}(-A) = \sup\{\nu'(-Ax) : \nu(x) = 1)$$
$$= \sup\{\nu'(Ax) : \nu(x) = 1\}$$
$$= \text{lub}_{\nu',\nu}(A)$$

Q.E.D.

Relations (12.19) and (12.20) may also be expressed as

$$|\text{lub}_{\nu',\nu}(A) - \text{lub}_{\nu',\nu}(B)| \leq \text{lub}_{\nu',\nu}(A-B)$$
$$|\text{glb}_{\nu',\nu}(A) - \text{glb}_{\nu',\nu}(B)| \leq \text{lub}_{\nu',\nu}(A-B)$$

from which it is easily seen that $\text{lub}_{\nu',\nu}$ and $\text{glb}_{\nu',\nu}$ are continuous mappings with respect to the topology generated by $\text{lub}_{\nu',\nu}$.

The effect of norm transformations on $\text{lub}_{\nu',\nu}$ and $\text{glb}_{\nu',\nu}$ is given by:

Theorem: Let ν'_Q and ν_R be the transformed norms corresponding to the nonsingular linear transformations Q and R: (12.22)

$$\nu'_Q(x) = \nu'(Qx) \quad \text{and} \quad \nu_R(x) = \nu(Rx) .$$

Then

$$\text{lub}_{\nu'_Q,\nu_R}(A) = \text{lub}_{\nu',\nu}(QAR^{-1}) \tag{12.23}$$

$$\text{glb}_{\nu'_Q,\nu_R}(A) = \text{glb}_{\nu',\nu}(QAR^{-1}) \tag{12.24}$$

Proof:

$$\text{lub}_{\nu'_Q,\nu_R}(A) = \sup\{\frac{\nu'_Q(Ax)}{\nu_R(x)} : x \neq \phi\}$$
$$= \sup\{\frac{\nu'(QAR^{-1}Rx)}{\nu(Rx)} : Rx \neq \phi\}$$

$$= \sup\{\frac{\nu'((QAR^{-1})y)}{\nu(y)} : y \neq \phi\}$$

$$= \text{lub}_{\nu',\nu}(QAR^{-1}) \ .$$

A similar argument shows $\text{glb}_{\nu'_Q,\nu_R}(A) = \text{glb}_{\nu',\nu}(QAR^{-1})$.

Q.E.D.

Thus, if ν' is invariant under the group of norm transformations $\mathcal{Q}$ and ν is invariant under the group $\mathcal{R}$, then $\text{lub}_{\nu',\nu}$ are invariant under the product group $\mathcal{Q} \times \mathcal{R} : A \to \mathcal{Q}A\mathcal{R}$.

Let A be an endomorphism of a normed vector space $\mathcal{V}_C$ over the complex field C and let $\nu' = \nu$ be strictly homogeneous. Then:

Theorem: If λ is an eigenvalue of A, then (12.25)

$$\text{glb}_{\nu',\nu}(A) \leq |\lambda| \leq \text{lub}_{\nu',\nu}(A) \ .$$

Proof:

Let λ be an eigenvalue of A and x the corresponding eigenvector. Then $Ax = \lambda x$ and

$$\text{glb}_{\nu,\nu}(A) \leq \frac{\nu(Ax)}{\nu(x)} = \frac{\nu(\lambda x)}{\nu(x)} = |\lambda| \frac{\nu(x)}{\nu(x)} = |\lambda| \leq \text{lub}_{\nu,\nu}(A)$$

Q.E.D.

The domain defined in (12.25) is an annulus in the complex plane. For real, nonnegative eigenvalues, the assumption of strict homogeneity may be dropped. In this case, if $\lambda = e^{i\theta}\tau$, $\tau \geq 0$, is an eigenvalue of A, then τ is a real, nonnegative eigenvalue of $e^{-i\theta}A$ and

$$\text{glb}_{\nu',\nu}(e^{-i\theta}A) \leq \tau \leq \text{lub}_{\nu',\nu}(e^{-i\theta}A) \ .$$

The domain is still an annulus but the bounding curves no longer need be concentric circles.

If the norm ν' is generated by the set $S' \subset V_K'^D (\nu' = \nu_{S'})$, then

$$\nu'(Ax) = \sup\{\text{Re } \varphi(Ax) : \varphi \in S'\}$$

and

$$\begin{aligned} \text{lub}_{\nu',\nu}(A) &= \sup\{\nu'(Ax) : \nu(x) = 1\} \\ &= \sup\{\sup\{\text{Re } \varphi(Ax) : \varphi \in S'\} : \nu(x) = 1\} \\ &= \sup\{\text{Re } \varphi(Ax) : \varphi \in S', \nu(x) = 1\} \\ &= \sup \text{Re } \{\varphi(Ax) : \varphi \in S', \nu(x) = 1\}, \end{aligned}$$

the <u>BILINEAR CHARACTERIZATION OF THE LEAST UPPER BOUND</u>. (12.26)

This leads us to introduce the

<u>BILINEAR FIELD OF VALUES OF A</u> (12.27)

subordinate to $\nu_{S'}$ and ν :

$$P_{S',\nu}[A] := \{\varphi(Ax) : \varphi \in a', \ \nu(x) = 1\} . \qquad (12.28)$$

Since

$$\text{lub}_{\nu',\nu}(A) = \sup \text{Re } P_{S',\nu}[A], \qquad (12.29)$$

$\text{lub}_{\nu',\nu}(A)$ characterizes the position of a parallel to the imaginary axis supporting $P_{S',\nu}[A]$ from the right. In the special case of $\nu' = \nu = \nu_{S'}$, comparing (12.28) with the more restrictive (6.8) gives:

$$G_S[A] = \{\varphi(Ax) \in P_{S,\nu}[A] : \varphi(x) = \nu(x)\} \subset P_{S,\nu_S}[A] . \qquad (12.30)$$

Thus by Theorem (6.10),

<u>Exclusion Theorem</u>: No eigenvalue of A lies outside $P_{S,\nu_S}[A]$; that is, if $\lambda \in C$ is an eigenvalue of A, then $\lambda \in P_{S,\nu_S}[A]$. (12.31)

$P_{S,\nu_S}[A]$ is not only larger than $G_S[A]$ in general but is also not covariant under translation. Like $G_S[A]$, however, it is INVARIANT UNDER SCALAR MULTIPLICATION:

$$\forall\ \tau \in K : P_{S,\nu_S}[\tau A] = \tau\, P_{S,\nu_S}[A], \tag{12.32}$$

and, for nonsingular B,

$$P_{SB,\nu_{SB}}[A] = P_{S,\nu_S}[BAB^{-1}] ,$$

More generally, if B and B' are nonsingular, then

$$P_{S'B',\nu_{SB}}[A] = P_{S',\nu_S}[B'AB^{-1}] . \tag{12.33}$$

If either ν'_S or ν is strictly homogeneous, $P_{S',\nu}[A]$ will have rotational symmetry about the origin: If $q \in P_{S',\nu}[A]$ for some choice of φ and x, then $\omega q \in P_{S',\nu}[A]$ for all ω with $|\omega| = 1$ (consider the element of $P_{S',\nu}[A]$ generated by either $\omega\varphi$ and x or φ and ωx). Thus, (12.31) is a generalization of (12.25) without the restriction of strictly homogeneous norms.

§13. Dual Norms

The order in which the suprema are taken in (12.26) may be reversed:

$$\mathrm{lub}_{\nu',\nu}(A) = \sup_x \left\{ \sup_{\varphi \in S'} \{\mathrm{Re}\ \varphi(Ax)\} : \nu(x) = 1 \right\}$$

$$= \sup_{\varphi \in S'} \sup_x \{\mathrm{Re}\ \varphi A(x) : \nu(x) = 1\}$$

If we define the mapping ν^D or $\mathcal{V}_K^D = \mathrm{Hom}(\mathcal{V}_K, K)$ by

$$\nu^D(\psi) := \sup\{\mathrm{Re}\ \psi(x) : \nu(x) = 1\} \qquad (13.1)$$

$$= \sup \left\{\frac{\mathrm{Re}\ \psi(x)}{\nu(x)} : x \neq \phi\right\} \qquad (13.2)$$

$$= \inf\{\beta : \mathrm{Re}\ \psi(x) \leq \beta \cdot \nu(x), \forall x \in \mathcal{V}_K\}, \qquad (13.3)$$

then

$$\mathrm{lub}_{\nu',\nu}(A) = \sup\{\nu^D(\varphi A) : \varphi \in S'\}, \qquad (13.4)$$

the <u>DUAL CHARACTERIZATION OF THE LEAST UPPER BOUND</u>.

The supremum of (13.1) much resembles the supremum that led to the least upper bound. In fact, since Re is additive and homogeneous, a proof analogous to that of Theorem (12.5) shows that ν^D is subadditive and homogeneous. However, since Re is neither non-negative nor definite, another argument is needed to show that ν^D is positive definite:

<u>Proof</u>:

Assume that $\nu^D(\psi) \leq 0$. Then $\forall x \in \mathcal{V}_K$ with $x \neq \phi$: $\frac{\mathrm{Re}\ \psi(x)}{\nu(x)} \leq 0$. But for such x, $\nu(x) > 0$, whence $\forall x \in \mathcal{V}_K$: $\mathrm{Re}\ \psi(x) \leq 0$. In particular, $\forall \theta \in [0,2\pi), x \in \mathcal{V}_K$: $\mathrm{Re}\ e^{i\theta}\psi(x) = \mathrm{Re}\ \psi(e^{i\theta}x) \leq 0$. Therefore $\psi(x) \equiv 0$ and $\psi = \phi^D$; that is, $\psi \neq \phi^D \succ \nu^D(\psi) > 0$.

Q.E.D.

If $\nu^D(\psi)$ is bounded, then ν^D is a norm; therefore,

Theorem: $\nu^D(\psi)$ is a norm on the subspace of bounded linear functionals of $\mathcal{V}_K$, the (13.5)

DUAL NORM TO THE NORM ν. (13.6)

In the finite dimensional case, every linear mapping is bounded and ν^D is a norm on $\mathcal{V}_K^D$ for every norm ν .

Examples:

(i) Let $\mathcal{V}_K = R^n$ (or C^n) with ν the Tschebyscheff norm. Then $\forall\psi = (\psi_1, \ldots, \psi_n) \in \mathcal{V}_K^D$,

$$\nu^D(\psi) = \sum_{i=1}^{n} |\psi_i|, \tag{13.7}$$

the Manhattan norm on the dual space.

(ii) Let $\mathcal{V}_K = R^n$ (or C^n) with ν the Manhattan norm. Then $\forall\psi = (\psi_1, \ldots, \psi_n) \in \mathcal{V}_K^D$,

$$\nu^D(\psi) = \max_{1 < i < n} |\psi_i|, \tag{13.8}$$

the Tschebyscheff norm in the dual space.

(iii) Let $\mathcal{V}_K$ be a Hilbert space with the scalar product norm $\nu(x) = (\varphi(x,x))^{\frac{1}{2}}$. Then by the Riesz Representation Theorem, $\forall\psi \in \mathcal{V}_K^D \ \exists y \in \mathcal{V}_K: \psi(x) = \varphi(x,y)$ and

$$\nu^D(\psi) = \nu(y) . \tag{13.9}$$

(iv) Let $\mathcal{V}_K = R^n$ and let A be Hermitian and positive definite so that $\nu(x) = (x^T A x)^{\frac{1}{2}}$ is a norm. Then $\forall y^T \in \mathcal{V}_K^D = R_n^D$,

$$\nu^D(y^T) = (y^T A^{-1} y)^{\frac{1}{2}} . \tag{13.10}$$

In particular, if $A = I$, then v is the Euclidean norm and

$$\nu^D(y^T) = \left(\sum_{i=1}^{n} |y_i|^2\right)^{\frac{1}{2}},$$

the Euclidean norm in the dual space.

If the set $S \subset \mathcal{V}_K^D$ generates the norm ν_S, then the unit ball of the dual norm

$$K_1^D = \{\psi \in \mathcal{V}_K^D: \ \nu_S^D(\psi) \leq 1\}$$

is closely related to S. In fact, since from (13.3)

$$\nu_S^D(\psi) \leq 1 \iff \forall x \in \mathcal{V}_K: \ \text{Re}\ \psi(x) \leq \nu_S(x),$$

it follows that

$$\begin{aligned} K_1^D &= \{\psi \in \mathcal{V}_K^D: \ \nu_S^D(\psi) \leq 1\} \\ &= \{\psi \in \mathcal{V}_K^D: \ \text{Re}\ \psi(x) \leq \nu_S(x), \ \forall x \in \mathcal{V}_K\} \\ &= \bigcap_{x \in \mathcal{V}_K} \{\psi \in \mathcal{V}_K^D: \text{Re}\ \psi(x) \leq \nu_S(x)\} \\ &= \bigcap_{x \in \mathcal{V}_K} H_{x,\nu_S(x)}^D \end{aligned} \tag{13.11}$$

where

$$H_{x,\alpha}^D := \{\psi \in \mathcal{V}_K^D: \text{Re}\ \psi(x) < \alpha\} \tag{13.12}$$

is a half-space in $\mathcal{V}_K^D$. It is clear from (9.1) that $\forall x \in \mathcal{V}_K$: $S \subset H_{x,\nu_S(x)}^D$. On the other hand, if $S \subset H_{x,\alpha}$, then $\forall \varphi \in S$: $\text{Re}\ \varphi(x) \leq \alpha$ whence $\nu_S(x) \leq \alpha$ and $H_{x,\nu_S(x)} \subseteq H_{x,\alpha}$. Therefore,

<u>Theorem</u>: K_1^D is the intersection of all half-spaces in $\mathcal{V}_K^D$ containing S, the (13.13)

BIPOLAR HULL OF $\mathbf{S}$. (13.14)

As a consequence $\mathbf{S} \subset K_1^D$ and, since the half-spaces $H_{x,\alpha}$ are convex,

$$H[\mathbf{S}] \subset K_1^D, \tag{13.15}$$

where $H[\mathbf{S}]$ denotes the CONVEX HULL OF $\mathbf{S}$, the intersection of all convex sets containing $\mathbf{S}$. In finite dimensional spaces,

$$K_1^D = \mathscr{F}H[\mathbf{S}] \tag{13.16}$$

where $\mathscr{F}$ denotes the topological closure operation.

Example:

Let $V_K = R^n$ and let

$$\mathbf{S} = \{(\ell_1, \ldots, \ell_n): \ \ell_1^2 + \ell_2^2 + \ldots + \ell_n^2 < 1\},$$

the open unit sphere. Then

$$K_1^D = \{(\ell_1, \ldots, \ell_n): \ \ell_1^2 + \ell_2^2 + \ldots + \ell_n^2 \leq 1\},$$

the closed unit sphere.

At this point it is interesting to note that the work of the preceding paragraphs could have been done using K_1^D rather than $\mathbf{S}$. However, the eigenvalue inclusion theorems gave better results for simple (minimal) sets $\mathbf{S}$. Also, it is nicer to generate norms without resorting to limit processes and this can only be done for finite sets $\mathbf{S}$.

If the norm ν' is generated by a finite set $\mathbf{S}'$, then (13.4) reduces $\text{lub}_{\nu',\nu}(A)$ to a maximum over a finite set of dual norms:

$$\text{lub}_{\nu',\nu}(A) = \max_{\varphi \in \mathbf{S}'} \{\nu^D(\varphi A)\}. \tag{13.17}$$

Examples:

(i) Let $\mathcal{V}_K = \mathcal{V}'_K = C^n$ and let ν' be the Tschebyscheff norm with generating set

$$S' = \bigcup_i \{\omega e_i^T : |\omega| = 1\} .$$

Then

$$\mathrm{lub}_{\nu',\nu}(A) = \max_{i,\, |\omega| = 1} \nu^D(\omega e_i^T A) . \qquad (13.18)$$

If ν is strictly homogeneous, then

$$\mathrm{lub}_{\nu',\nu}(A) = \max_i \nu^D(e_i^T A), \qquad (13.19)$$

the maximum of the dual norms of the rows of A . If ν is the Tschebyscheff norm, then ν^D is the Manhattan norm and

$$\mathrm{lub}_{\nu',\nu}(A) = \max_i \sum_k |a_{ik}|, \qquad (13.20)$$

the ROW SUM NORM. If ν is the Manhattan norm, then ν^D is the Tschebyscheff norm and

$$\mathrm{lub}_{\nu',\nu}(A) = \max_{i,k} |a_{i,k}|, \qquad (13.21)$$

the <u>MATRIX TSCHEBYSCHEFF NORM</u>.

(ii) Let $\mathcal{V}_K = \mathcal{V}'_K = C^n$ and let ν' be the Manhattan norm with generating set

$$S' = \{(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n})\} .$$

Then

$$\mathrm{lub}_{\nu',\nu}(A) = \max_{\ell^T \in S'} \nu^D(\ell^T A) . \qquad (13.22)$$

If ν is the Manhattan norm, then ν^D is the Tschebyscheff norm and

$$\mathrm{lub}_{\nu',\nu}(A) = \max_{\ell^T \in S'} \max_k |\ell^T A e_k|$$
$$= \max_k \max_{\ell^T \in S'} |\ell^T A e_k|$$
$$\mathrm{lub}_{\nu',\nu}(A) = \max_k \sum_i |a_{ik}|, \qquad (13.23)$$

the <u>COLUMN SUM NORM</u>. If ν is the Tschebyscheff norm, then ν^D is the Manhattan norm and

$$\mathrm{lub}_{\nu',\nu}(A) = \max_{\ell^T \in S'} \sum_k |\ell^T A e_k|$$
$$= \max_{D_\ell} \sum_k |e^T D_\ell A e_k| \qquad \text{(see (6.22))}$$
$$\mathrm{lub}_{\nu',\nu}(A) = \max_{D_\ell} \sum_k \Big| \sum_i (D_\ell A)_{ik} \Big|, \qquad (13.24)$$

the maximum of the sum of the absolute values of the column sums of A under left-sided phase transformations. One might have expected from the duality between the Manhattan and Tschebyscheff norms that

$$\mathrm{lub}_{\nu',\nu}(A) = \sum_{i,k} |a_{ik}|$$

in this case. Indeed,

$$\mathrm{lub}_{\nu',\nu}(A) \le \sum_{i,k} |a_{ik}|; \qquad (13.25)$$

however, there is equality if and only if A is non-negative up to a two-sided phase pattern transformation. Thus in general,

$$\sum_{i,k} |a_{ik}|$$

is merely an upper bound for A compatible with the Manhattan norm ν' and the Tschebyscheff norm ν .

(iii) Let $\mathcal{V}_K = \mathcal{V}'_K = C^n$ and let ν' and ν be the Euclidean norm. We introduce the

BELTRAMI-JORDAN DECOMPOSITION OF A: (13.26)

$$A = U \Sigma V^H \tag{13.27}$$

where U and V are unitary and $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \geq 0$. Since the Euclidean norm is invariant under unitary transformations,

$$\begin{aligned} \mathrm{lub}_{\nu',\nu}(A) &= \mathrm{lub}_{\nu',\nu}(U \Sigma V^H) \\ &= \mathrm{lub}_{\nu',\nu}(\Sigma) \\ &= \sup \frac{\left(\sum_i |\sigma_i x_i|^2\right)^{\frac{1}{2}}}{\left(\sum_i |x_i|^2\right)^{\frac{1}{2}}} \\ &\leq \max_i \sigma_i \ . \end{aligned}$$

This bound is achieved for $x_\mu = \begin{cases} 1 & \sigma_\mu = \max \sigma_i \\ 0 & \sigma_\mu < \max \sigma_i \end{cases}$ so that

$$\mathrm{lub}_{\nu',\nu}(A) = \max_i \sigma_i , \tag{13.28}$$

the EUCLIDEAN BOUND NORM ("SPECTRAL" NORM). The non-negative scalars σ_i are the

SINGULAR VALUES OF A . (13.29)

Since $A^H A = U \Sigma^2 U^H$ is Hermitian positive definite, the σ_i are just the non-negative square roots of the eigenvalues of $A^H A$. If A is Hermitian or normal, then the σ_i are the absolute values of the eigenvalues of A .

§14. Least Upper and Greater Lower Bounds II.

We shall now consider least upper and greatest lower bounds in the general case where G_{K_0} and G'_{K_0} are not both one-dimensional. But first we must state what we mean by such bounds.

Let $\mathcal{B}[A] \subset \mathrm{Hom}(G_{K_0}, G'_{K_0})$ denote the set of all upper bounds for A ((11.1)) and let $C[A] \subset \mathrm{Hom}(G_{K_0}, G'_{K_0})$ be the set of all lower bounds for A ((11.5)). To compare bounds within these sets we must introduce an ordering $\hat{\rho}$ in $\mathrm{Hom}(G_{K_0}, G'_{K_0})$. Thus we define

$$\forall \mathscr{S}_1, \mathscr{S}_2 \in \mathrm{Hom}(G_{K_0}, G'_{K_0}): \mathscr{S}_1 \hat{\rho} \mathscr{S}_2 \;:\gtrless\; \forall x \in G_{K_0}: \; \phi \rho x \succ \mathscr{S}_1 x \; \rho' \; \mathscr{S}_2 x \; . \tag{14.1}$$

Theorem: $\hat{\rho}$ is transitive, reflexive, and antisymmetric. (14.2)

Proof:

Transitivity ((2.2)) and reflexivity ((2.3)) are inherited directly from $\hat{\rho}$ but the proof of antisymmetry ((2.4)) is more difficult:

$$\begin{aligned} \mathscr{S}_1 \hat{\rho} \mathscr{S}_2 \wedge \mathscr{S}_2 \hat{\rho} \mathscr{S}_1 &\succ \forall x \in G^+: \; \mathscr{S}_1 x \; \rho' \; \mathscr{S}_2 x \wedge \mathscr{S}_2 x \; \rho' \; \mathscr{S}_1 x \\ &\succ \forall x \in G^+ : \mathscr{S}_1 x = \mathscr{S}_2 x \\ &\succ \forall x \in G_{K_0} : \mathscr{S}_1 x = \mathscr{S}_2 x \qquad \text{since} \quad G_{K_0} = G^+ - G^+ \\ &\succ \mathscr{S}_1 = \mathscr{S}_2 \end{aligned}$$

Q.E.D.

Thus $\hat{\rho}$ is an ordering of $\mathrm{Hom}(G_{K_0}, G'_{K_0})$, the

ORDERING INDUCED BY ρ AND ρ'. (14.3)

Moreover, it is

COMPATIBLE WITH ρ AND ρ': (14.4)

$$x_1 \rho x_2 \wedge \mathcal{B}_1 \hat{\rho} \mathcal{B}_2 > \mathcal{B}_1 x_1 \rho' \mathcal{B}_2 x_2 . \qquad (14.5)$$

Example:

Let $G_{K_0} = (R^k, \leq)$ and $G'_{K_0} = (R^\ell, \leq)$. Then the ordering $\hat{\rho}$ induced in $\mathrm{Hom}(R^k, R^\ell) = R^{k\times\ell}$ is given by

$$\mathcal{B} \hat{\rho} H \;:\Leftrightarrow\; \forall i,j:\; (\mathcal{B})_{ij} \leq (H)_{ij}, \qquad (14.6)$$

an elementwise ordering which we shall again denote by $\leq$.

As might be expected, we can obtain weaker bounds from known bounds by means of the ordering $\hat{\rho}$:

Theorem: (14.7)

$$B_1 \in \mathcal{B}[A] \wedge B_1 \hat{\rho} B_2 > B_2 \in \mathcal{B}[A]$$

$$C_1 \in C[A] \wedge C_2 \hat{\rho} C_1 > C_2 \in C[A].$$

Proof:

Assume that $B_1 \in \mathcal{B}[A]$ and $B_1 \hat{\rho} B_2$. Then $\nu'(Ax) \rho' B_1\nu(x)\rho' B_2\nu(x)$ since $0 \rho \nu(x)$. Therefore $B_2 \in \mathcal{B}[A]$.

Q.E.D.

Theorem: $\mathcal{B}[A]$ and $C[A]$ are convex. (14.8)

Proof:

Let $B_1, B_2 \in \mathcal{B}[A]$ and assume $0 \rho_0 \mu \rho_0 1$. Then

$$\nu'(Ax)\rho' B_1\nu(x) \wedge \nu'(Ax)\rho' B_2\nu(x)$$
$$> \mu\nu'(Ax) \rho' \mu B_1\nu(x) \wedge (1-\mu)\nu'(Ax) \rho'(1-\mu)B_2\nu(x)$$
$$> \nu'(Ax) \rho' [\mu B_1 + (1-\mu)B_2]\nu(x)$$
$$> \mu B_1 + (1-\mu)B_2 \in \mathcal{B}[A] .$$

Q.E.D.

The zero element 0 of $\mathrm{Hom}(G_{K_0}, G'_{K_0})$ is always a lower bound:

$$0 \in C[A] . \tag{14. 9}$$

However, this does not imply that $\forall \mathcal{B} \in \mathcal{B}[A]: 0 \hat{\rho} \mathcal{B}$. Indeed

$$H \in C[A] \wedge \mathcal{B} \in \mathcal{B}[A] \succ \forall x \in \mathcal{V}_K: H\nu(x) \rho' \mathcal{B}\nu(x),$$

whereas

$$H \hat{\rho} \mathcal{B} \approx \forall z \in G^+_{K_0}: Hz \ \rho' \ \mathcal{B}z \ .$$

Still,

Theorem: If v-(x) is SURJECTIVE ON $G^+_{K_0}$, that is, if ν is a mapping of $\mathcal{V}_K$ onto $G^+_{K_0}$, then- (14.10)

$$H \in C[A] \wedge \mathcal{B} \in \mathcal{B}[A] \succ H \hat{\rho} \mathcal{B} \ .$$

For now $\{V(x): x \in \mathcal{V}_K\} = \{z: z \in G^+_{K_0}\}$.

In the case where $G'_{K_0} = (K_0, \rho_0)$, the induced ordering $\hat{\rho}$ is an ordering of $G^D_{K_0} = \mathrm{Hom}(G_{K_0}, K_0)$, the

DUAL ORDERING (14.11)

(For $G_{K_0} = (K_0, \rho_0)$ as well, $\hat{\rho}$ reduces to the ordering ρ_0 of Ko).
The DUAL CONE of the positivity cone G^+ is then given by

$$\begin{aligned} G^D_+ &:= \{\psi \in G^D_{K_0} : 0^D \hat{\rho} \psi\} \\ &= \{\psi \in G^D_{K_0} : 0 \ \rho_0 \psi(x), \ \forall x \in G^+\} \\ G^D_+ &= \bigcap_{x \in G^+} \{\psi \in G^D_{K_0} : 0 \ \rho_0 \ \psi(x)\} \end{aligned} \tag{14.12}$$

an intersection of half-spaces. The question of whether $\mathcal{B}[A] \subset G^D_{K_0}$ has a least element and C[A] a greatest element was answered in §8:

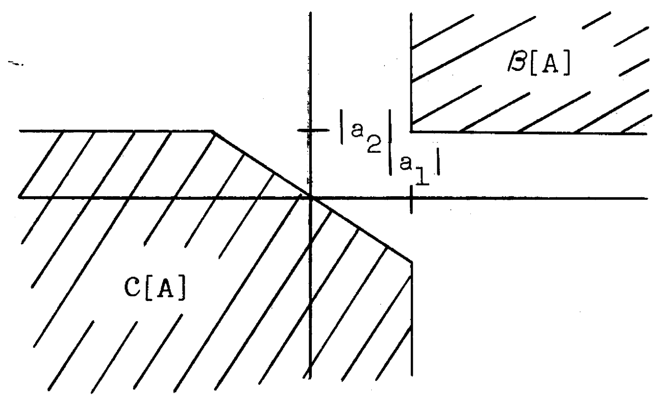
$\mathcal{B}[A]$ has a least element $\beta_0 \approx \mathcal{B}[A] = \beta_0 + G^D_+$

$C[A]$ has a greatest element $\gamma_0 \approx C[A] = \gamma_0 - G^D_+$.

Examples:

(i) Let $\mathcal{V}_K = R^2$ and $G_{K_0} = (R^2, \leq)$ with $\nu\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} |x_1| \\ |x_2| \end{pmatrix}$, (14.13) the modulus norm. Let $\mathcal{V}'_K = R$ and $G'_{K_0} = (R, \leq)$ with $\nu'(x) = |x|$. Then $\forall A = (a_1, a_2) \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$,

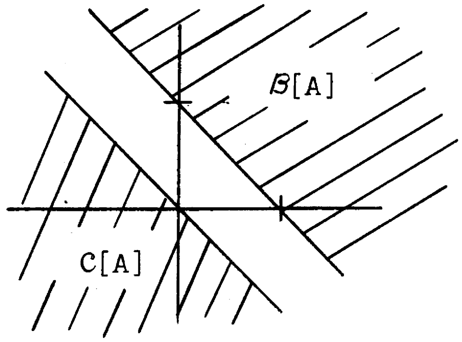
$$\mathcal{B}[A] = \{(\beta_1,\beta_2): |a_1| \leq \beta_1, |a_2| \leq \beta_2\}$$
$$C[A] = \{(\gamma_1,\gamma_2): \gamma_1 \leq |a_1|, \gamma_2 \leq |a_2|, \gamma_1|a_1| + \gamma_2|a_2| \leq 0\}$$



Since $\mathcal{B}[A]$ is a translated positivity cone, $\mathcal{B}[A]$ has a least element $\beta = (|a_1|, |a_2|)$. However, $C[A]$ has no greatest element.

(ii) Let $\mathcal{V}_K = R^2$ and $G_{K_0} = (R^2, \leq)$ with $\nu\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sqrt{x_1^2 + x_2^2} \\ \sqrt{x_1^2 + x_2^2} \end{pmatrix}$. (14.14) Let $\mathcal{V}'_K = R$ and $G'_{K_0} = (R, \leq)$ with $\nu'(x) = |x|$. Then $\forall A = (a_1, a_2) \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$,

$$\mathcal{B}[A] = \left\{(\beta_1,\beta_2): \beta_1 + \beta_2 \geq \sqrt{a_1^2 + a_2^2}\right\}$$
$$C[A] = \{(\gamma_1,\gamma_2): \gamma_1 + \gamma_2 \leq 0\} .$$



$\mathcal{B}[A]$ has no least element and $C[A]$ has no greatest element. Moreover, since the values of $v(x)$ lie on a single ray in G_{K_o}, ν is not surjective on $G^+_{K_o}$ and Theorem (14.10) does not apply. Indeed, there do exist some lower and upper bounds which are incomparable.

(iii) Let $\mathcal{V}_K = R^2$ and $G_{K_o} = (R^2,\leq)$ with (14.15)

$$\nu\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}|x_1| + |x_2|\\ \max(|x_1|,|x_2|)\end{pmatrix} .$$

Let $\mathcal{V}'_K = R$ and $G'_{K_o} = (R,\leq)$ with $V'(x) = |x|$. Then $\forall A = (a_1,a_2) \in \mathrm{Hom}\,(\mathcal{V}_K, \mathcal{V}'_K)$,

$$\mathcal{B}[A] = \{ (\beta_1,\beta_2) : \ 2\beta_1 + \beta_2 \geq |a_1| + |a_2|,\ \beta_1 + \beta_2 \geq \max(|a_1|, |a_2|)\}$$

an intersection of half-spaces not yet a translated positivity cone. Note that ν is not surjective on $G^+_{K_o}$.

Let $G'_{K_o} = (K_o,\rho_o)$ and assume that ν' is generated by a set $S' \subset \mathcal{V}'^D_K$ of linear functionals. Then in a purely formal manner, we may extend the concept of the least upper bound of $A \in \mathrm{Hom}(\mathcal{V}_K,\mathcal{V}'_K)$ with respect to the norms ν' and ν:

$$\begin{aligned}
\text{lub}_{\nu',\nu}(A) &:= \inf\{\beta\epsilon G^D_{K_o}: \ \nu'(Ax) \ \rho_o \ \beta\nu(x), \ \forall x\epsilon \mathcal{V}_K\} \\
&= \inf\{\beta\epsilon G^D_{K_o}: \ \sup_{\varphi\epsilon \mathbf{S}'} \ \text{Re} \ \varphi A(x) \ \rho_o \ \beta\nu(x), \ \forall x\epsilon \mathcal{V}_K\} \\
&= \sup_{\varphi\epsilon \mathbf{S}'} \ \inf\{\beta\epsilon G^D_{K_o}: \ \text{Re} \ \varphi A(x) \rho_o \ \beta\nu(x), \ \forall x\epsilon \mathcal{V}_K\} \\
&= \sup_{\psi\epsilon \mathbf{S}'A} \ \inf\{\beta\epsilon G^D_{K_o}: \ \text{Re} \ \psi(x) \ \rho_o \ \beta\nu(x), \ \forall x\epsilon \mathcal{V}_K\}
\end{aligned}$$

where the interchange of infimum and supremum is again purely formal. $\mathbf{S}'A \subset \mathcal{V}^D_K$ is a set of linear functionals on $\mathcal{V}_K$. Thus, provided all the necessary infima and suprema exist, we have reduced the study of upper bounds for homomorphisms $A \in \text{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ to the study of upper bounds for linear functionals $\varphi\epsilon \mathcal{V}^D_K$. It is well to remark at this point that although G_{K_o} and $G'_{K_o} = (K_o, \rho_o)$ are assumed to be vector lattices, $\text{Hom}(G_{K_o}, G'_{K_o}) = G^D_{K_o}$ is not necessarily a vector lattice. Thus, the indicated infima may not exist.

Let $\mathcal{B}[\varphi]$ denote the set of all upper bounds for $\varphi\epsilon \mathcal{V}^D_K$:

$$\begin{aligned}
\mathcal{B}[\varphi] &:= \{\beta\epsilon G^D_{K_o}: \ \text{Re} \ \varphi(x) \ \rho_o \ \beta\nu(x), \ \forall x\epsilon \mathcal{V}_K\} \\
&= \bigcap_{x\,\epsilon \mathcal{V}_K} \{\beta\epsilon G^D_{K_o}: \text{Re} \ \varphi(x) \ \rho_o \ \beta\nu(x)\}
\end{aligned} \qquad (14.16)$$

$$\mathcal{B}[\varphi] = \bigcap_{x\epsilon \mathcal{V}_K} H_x[\varphi] \qquad (14.17)$$

where

$$H_x[\varphi] = \{\beta\epsilon G^D_{K_o}: \text{Re} \ \varphi(x) \ \rho_o \ \beta\nu(x)\} \ .$$

Thus $\mathcal{B}[\varphi]$ is an intersection of half-spaces. By Theorem (8.22), $\mathcal{B}[\varphi]$ has a least element β_o if and only if $\mathcal{B}[\varphi]$ is the translated cone $\beta_o + G^D_+$. In this case we denote the least element β_o by $\nu^D(\psi)$:

Theorem: If $\nu^D(\varphi)$ exists for all $\varphi \epsilon \mathcal{V}_K^D$, then ν^D is a norm on the subspace of bounded linear functionals on $\mathcal{V}_{K'}$ the (14.18)

DUAL NORM TO THE NORM $\underline{v}$. (14.[illegible])

Proof:

Let $\varphi_1, \varphi_2 \epsilon \mathcal{V}_K^D$. Then

$$\mathrm{Re}(\varphi_1+\varphi_2)(x) = \mathrm{Re}\ \varphi_1(x) + \mathrm{Re}\ \varphi_2(x)\ \rho_o\ (\nu^D(\varphi_1) + \nu^D(\varphi_2))\nu(x) .$$

Therefore

$$\nu^D(\varphi_1) + \nu^D(\varphi_2) \in \mathcal{B}[\varphi_1+\varphi_2]$$

and

$$\nu^D(\varphi_1+\varphi_2)\ \hat{\rho}\ \nu^D(\varphi_1) + \nu^D(\varphi_2) \qquad \text{(subadditivity)}$$

since $\nu^D(\varphi_1+\varphi_2)$ is the least element of $\mathcal{B}[\varphi_1+\varphi_2]$. Homogeneity follows from the homogeneity of the mapping Re: $\forall \alpha \epsilon K_o$ with $0\ \rho_o\ \alpha$,

$$\begin{aligned}
\nu^D(\alpha\ \varphi) &= \inf\{\beta \epsilon G_{K_o}^D :\ \mathrm{Re}\ \alpha\ \varphi(x)\ \rho_o\ By(x), \quad \forall x \epsilon \mathcal{V}_K\} \\
&= \inf\{\beta \epsilon G_{K_o}^D :\ \alpha\ \mathrm{Re}\ \varphi(x)\ \rho_o\ \beta\nu(x), \quad \forall x \epsilon \mathcal{V}_K\} \\
&= \alpha \cdot \inf\{\alpha^{-1}\beta \epsilon G_{K_o}^D :\ \mathrm{Re}\ \varphi(x)\ \rho_o\ \alpha^{-1}\beta\nu(x), \quad \forall x \epsilon \mathcal{V}_K\} \\
&= \alpha \cdot \inf\{\beta \epsilon G_{K_o}^D :\ \mathrm{Re}\ \varphi(x)\ \rho_o\ \beta\nu(x), \quad \forall x \epsilon \mathcal{V}_k\} \\
&= \alpha\ \nu^D(\varphi) .
\end{aligned}$$

Assume that $\nu^D(\varphi)\ \hat{\rho}\ 0^D$. Then $\forall x \epsilon \mathcal{V}_K$: $\mathrm{Re}\ \varphi(x) \rho_o\ 0$. In particular, for each $x \epsilon \mathcal{V}_{K'}$

$$\forall\ \alpha \in K:\ \mathrm{Re}\ \alpha\ \varphi(x) = \mathrm{Re}\ \varphi(\alpha\ x) \rho_o\ 0 .$$

From this we conclude that $\varphi(x) \equiv 0$; that is, $\varphi \neq 0^D > \nu^D(\varphi) > 0$ (positive definiteness).

Q.E.D.

Moreover, as in (13.4),

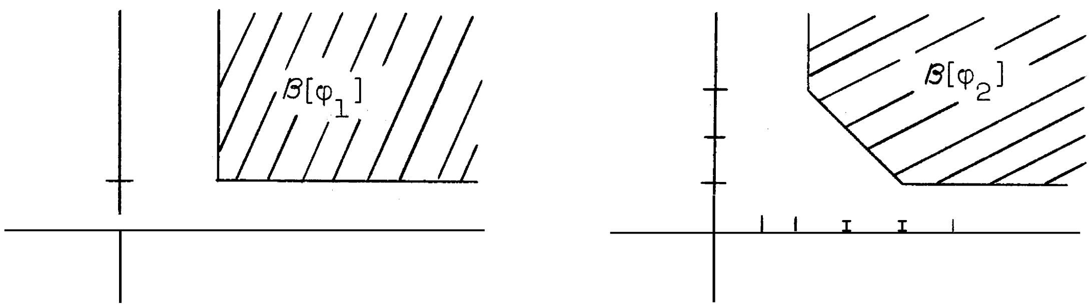
$$\mathrm{lub}_{\nu',\nu}(A) = \sup_{\varphi \in \mathbf{S}'} \nu^D(\varphi A) \quad . \tag{14.20}$$

Examples:

(i) Let $\mathcal{V}_K = R^n$ and $G_{K_o} = (R^n, \leq)$ with v the modulus norm. Let $\varphi = (a_1, a_2, \ldots, a_n) \in \mathcal{V}_K^D$. Then

$$\nu^D(\varphi) = |\varphi| := (|a_1|, |a_2|, \ldots, |a_n|) \quad .$$

(ii) Let $\mathcal{V}_K = R^3$ and $G_{K_o} = (R^2, \leq)$ with $\nu\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} |x_1| + |x_3| \\ |x_2| + |x_3| \end{pmatrix}$. Then for $\varphi_1 = (2,1,0) \in \mathcal{V}_K^D$ and $\varphi_2 = (2,1,5) \in \mathcal{V}_K^D$:



v is surjective; however, for some φ, $\mathcal{B}[\varphi]$ is not a translated cone.

As the preceding example indicates, surjectivity of the norm v is necessary but not sufficient to guarantee the existence of a least upper bound for the linear functional φ . However, in the case of a finite dimensional norm, we can prescribe a sufficient condition .

Henceforth we shall assume that ν is a finite-dimensional norm, $\nu\colon \mathcal{V}_K \to (R^k, \leq)$. By Theorems (9.25) and (9.28), each component ν_i of v is a seminorm and, if v is symmetric, even a norm on some subspace $\mathcal{U}_i$ of $\mathcal{V}_K$. Proceeding along this line, we define the norm v to be

<u>REGULAR</u> (14.21)

if $\mathcal{V}_K$ is a direct sum of subspaces

$$\mathcal{V}_K = \mathcal{U}_1 \oplus \mathcal{U}_2 \oplus \ . \ . \ . \oplus \mathcal{U}_k \tag{14.22}$$

and each component ν_i of ν is a norm on the subspace $\mathcal{U}_i$. Let

$$\mathcal{V}_K^D = \mathcal{U}_1^D \oplus \mathcal{U}_2^D \oplus \ldots \oplus \mathcal{U}_k^D \tag{14.23}$$

be the decomposition of $\mathcal{V}_K^D$ as a direct sum of subspaces $\mathcal{U}_i^D$ corresponding to (14.22). Then

<u>Theorem:</u> If ν is regular, then ν^D exists and (14.24)

$$\nu^D(\psi) = (\nu_1^D(\psi_1), \ldots) \ \nu_k^D(\psi_k)) \tag{14.25}$$

where

$$\psi = (\psi_1 | \psi_2 | \ \ldots \ |\psi_k) \in \mathcal{V}_K^D \text{ and } \psi_i \in \mathcal{U}_i^D \ .$$

<u>Proof:</u>

$$\text{Re } \psi(x) = \sum_i \text{Re } \psi_i(x_i) \leq \sum \nu_i^D(\psi_i)\nu_i(x_i)$$

Therefore $\nu^D(\psi)$ is an upper bound. If $(\beta_1, \ldots, \beta_k) \in \mathcal{B}[\psi]$, then

$$\forall x_i \in \mathcal{U}_i: \ \ \psi_i(x_i) \leq \beta_i \ \nu_i(x_i)$$

whence $\nu_i^D(\psi_i) \leq \beta_i$. Therefore $\nu^D(\psi)$ is the least upper bound.

Q.E.D.

Thus we have given a sufficient condition for the existence of ν^D . Note that a regular norm is surjective on $G_{K_0}^+$ but that the converse is not necessarily true.

Example:

Let $\mathcal{V}_K = R^n$ and $G_{K_o} = (R^n, \leq)$ with ν the modulus norm. Then an immediate consequence of Theorem (14.24) is that for $\varphi = (\varphi_1, \ldots, \varphi_n) \in \mathcal{V}^D_{K'}$

$$\nu^D(\varphi) = |\varphi| = (|\varphi_1|, \ldots, |\varphi_n|),$$

the modulus norm in the dual space.

Having found a sufficient condition for the existence of least upper bounds for linear functionals $\varphi \in V^D_K$, we now return to the study of least upper bounds for homomorphisms $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$. Let ν be a regular norm and let

$$\mathrm{Hom}\,(\mathcal{V}_K, \mathcal{V}'_K) = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \ldots \oplus \mathcal{L}_k \tag{14.26}$$

be the decomposition of $\mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ as a direct sum of subspaces

$$\mathcal{L}_i = \mathrm{Hom}(\mathcal{U}_i, \mathcal{V}'_K) \tag{14.27}$$

corresponding to (14.22). Then analogous to Theorem (14.24),

<u>Theorem:</u> $\mathrm{lub}_{\nu',\nu}$ exists and (14.28)

$$\mathrm{lub}_{\nu',\nu}(A) = (\mathrm{lub}_{\nu',\nu_1}(A_1), \mathrm{lub}_{\nu',\nu_2}(A_2), \ldots \mathrm{lub}_{\nu',\nu_k}(A_k))$$

where

$$A = (A_1 | A_2 | \ldots | A_k) \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$$

and

$$A_i \in \mathcal{L}_i .$$

We may now drop the assumption that ν' is a scalar norm and require instead that ν' be a (finite-dimensional) regular norm. Let

$$\mathcal{V}'_K = \mathcal{U}'_1 \oplus \mathcal{U}'_2 \oplus \ldots \oplus \mathcal{U}'_\ell \tag{14.29}$$

be the direct sum decomposition of $\mathcal{V}'_K$ such that each component ν'_i of ν' is a norm on the subspace $\mathcal{U}'_i$. Let

$$\mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) = \bigoplus_{i,j} \mathcal{L}_{ij} \tag{14.30}$$

where

$$\mathcal{L}_{ij} = \mathrm{Hom}(\mathcal{U}_i, \mathcal{U}'_j)$$

is the decomposition of $\mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ corresponding to (14.22) and (14.29). Then analogous to Theorem (14.28),

Theorem: $\mathrm{lub}_{\nu',\nu}$ exists and (14 30)

$$\mathrm{lub}_{\nu',\nu}(A) = \begin{pmatrix} \mathrm{lub}_{\nu'_1,\nu_1}(A_{11}) & \mathrm{lub}_{\nu'_1,\nu_2}(A_{12}) & \cdots & \mathrm{lub}_{\nu'_1,\nu_k}(A_{1k}) \\ \mathrm{lub}_{\nu'_2,\nu_1}(A_{21}) & \mathrm{lub}_{\nu'_2,\nu_2}(A_{22}) & \cdots & \mathrm{lub}_{\nu'_2,\nu_k}(A_{2k}) \\ \vdots & \vdots & & \vdots \\ \mathrm{lub}_{\nu'_\ell,\nu_1}(A_{\ell 1}) & \mathrm{lub}_{\nu'_\ell,\nu_2}(A_{\ell 2}) & \cdots & \mathrm{lub}_{\nu'_\ell,\nu_k}(A_{\ell k}) \end{pmatrix} \tag{14.32}$$

where

$$A = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1k} \\ \hline A_{21} & A_{22} & \cdots & A_{2k} \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{\ell 1} & A_{\ell 2} & \cdots & A_{\ell k} \end{array}\right) \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) \quad \text{and} \quad A_{ij} \in \mathcal{L}_{ij}\,. \tag{14.33}$$

Example:

Let $\mathcal{V}_K = R^n$ and $\mathcal{V}'_K = R^m$ with ν and ν' the modulus norms in R^n and R^m respectively., Then for $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K) = R^{n\times m}$:

$$\mathrm{lub}_{\nu',\nu}(A) = \begin{pmatrix} |a_{11}| & |a_{12}| & \cdots & |a_{1n}| \\ |a_{21}| & |a_{22}| & \cdots & |a_{2n}| \\ \vdots & \vdots & & \vdots \\ |a_{m1}| & |a_{m2}| & \cdots & |a_{mn}| \end{pmatrix} . \qquad (14.34)$$

As an immediate consequence of the norm properties of $\mathrm{lub}_{\nu_i',\nu_j}$:

<u>Theorem:</u> $\mathrm{lub}_{\nu',\nu}$ is a norm on the subspace of bounded homomorphisms $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}_K')$. (14.35)

We now shift our attention to lower bounds for homomorphisms. We shall find that, except for the case of scalar norms considered in $12, the set @[A] of lower bounds for A has in general no greatest element. This will follow &mediately from the existence of a maximal element which is not comparable with all other lower bounds.

As before, we assume initially that ν is a regular norm and that ν' is a scalar (real-valued) norm.

<u>Theorem:</u> Let (14.36)

$$\gamma_\mu^{(i)} = \begin{cases} \mathrm{glb}_{\nu',\nu_i}(A_i) & \mu = i \\ -\mathrm{lub}_{\nu',\nu_\mu}(-A_\mu) & \mu \neq i \end{cases}$$

for $A = (A_1 | A_2 | \ldots | A_k) \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}_K')$. Then $\gamma^{(i)} = (\gamma_1^{(i)}, \gamma_2^{(i)}, \ldots, \gamma_k^{(i)}) \in$ @[A] for $i = 1, 2, \ldots, n$.

<u>Proof:</u>

$$\nu'(Ax) = \nu'(A_i x_i + \sum_{\mu \neq i} A_\mu x_\mu)$$

$$\geq \nu'(A_i x_i) - \sum_{\mu \neq i} \nu'(-A_\mu x_\mu)$$

$$\geq \mathrm{glb}_{\nu',\nu_i}(A_i) \cdot \nu_i(x_i) - \sum_{\mu \neq i} \mathrm{lub}_{\nu',\nu_\mu}(-A_\mu) \cdot \nu_\mu(x_\mu)$$

$$= \mathrm{glb}_{\nu',\nu_i}(A_i) \cdot \nu_i(x_i) + \sum_{\mu \neq i} [-\mathrm{lub}_{\nu',\nu_\mu}(-A_\mu)] \cdot \nu_\mu(x_\mu)$$

$$= \gamma^{(i)} \cdot \nu(x)$$

Q.E.D.

If A_i is singular, then $\text{glb}_{\nu',\nu_i}(A_i) = 0$ and the bound $\gamma^{(i)}$ can be replaced by the bound 0 by virtue of (14.9). On the other hand, if A_i is nonsingular, then $\text{glb}_{\nu',\nu_i}(A_i) > 0$ and, following ROBERT,

<u>Theorem:</u> Let (14.37)

$$\tilde{\gamma}_\mu^{(i)} = \begin{cases} \text{glb}_{\nu',\nu_i}(A_i) & \mu = i \\ -\text{glb}_{\nu',\nu_i}(A_i) \cdot \text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) & \mu \neq i \end{cases}$$

for $A = (A_1|A_2| \ldots |A_k) \in \text{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ with A_i nonsingular. Then $\tilde{\gamma}^{(i)} = (\tilde{\gamma}_1^{(i)}, \tilde{\gamma}_2^{(i)}, \ldots \tilde{\gamma}_k^{(i)}) \in \mathcal{G}[A]$.

<u>Proof:</u>

$$\begin{aligned}
\nu'(Ax) &= \nu'(A_i(x_i + \sum_{\mu\neq i} A_i^{-1}A_\mu x_\mu)) \\
&\geq \text{glb}_{\nu',\nu_i}(A_i) \cdot \nu_i(x_i + \sum_{\mu\neq i} A_i^{-1}A_\mu x_\mu) \\
&\geq \text{glb}_{\nu',\nu_i}(A_i) \cdot [\nu_i(x_i) - \sum_{\mu\neq i} \nu_i(-A_i^{-1}A_\mu x_\mu)] \\
&\geq \text{glb}_{\nu',\nu_i}(A_i) \cdot [\nu_i(x_i) - \sum_{\mu\neq i} \text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) \cdot \nu_\mu(x_\mu)] \\
&= \text{glb}_{\nu',\nu_i}(A_i)\nu_i(x_i) + \sum_{\mu\neq i} [-\text{glb}_{\nu',\nu_i}(A_i)\, \text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu)]\nu_\mu(x_\mu) \\
&= \tilde{\gamma}^{(i)} \cdot \nu(x)
\end{aligned}$$

Q.E.D.

Using the inequality

$$\text{glb}_{\nu',\nu_i}(A_i) \cdot \text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) \leq \text{lub}_{\nu',\nu_\mu}(-A_\mu)$$

to be derived in §16, we find that the bound $\gamma^{(i)}$ of (14.36) can be replaced by the bound $\tilde{\gamma}^{(i)}$ of (14.37) provided A_i is nonsingular. Thus the bounds of (14.36) are not necessarily maximal.

Example:

Let $\mathcal{V}_K = R^3$ and $G_{K_0} = (R^2, \leq)$ with

$$\nu\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}\nu_1\begin{pmatrix}x_1\\x_2\end{pmatrix}\\ \nu_2(x_3)\end{pmatrix} = \begin{pmatrix}\left(x_1^2 + x_2^2\right)^{\frac{1}{2}}\\ |x_3|\end{pmatrix}.$$

Let $\mathcal{V}'_K = R^2$ and $G'_{K_0} = (R, \leq)$ with $\nu'\begin{pmatrix}y_1\\y_2\end{pmatrix} = \left(y_1^2 + y_2^2\right)^{\frac{1}{2}}$.

Let

$$A = (A_1|A_2) = \left(\begin{array}{cc|c}2 & 2 & 2\\2 & 5 & 3\end{array}\right).$$

Then

$$\text{lub}_{\nu',\nu_1}(A_1) = 6$$

$$\text{glb}_{\nu',\nu_1}(A_1) = 1$$

$$\text{lub}_{\nu',\nu_2}(A_2) = \sqrt{13}$$

$$\text{glb}_{\nu',\nu_2}(A_2) = \sqrt{13}$$

$$\text{lub}_{\nu_1,\nu_2}(-A_1^{-1}A_2) = \frac{\sqrt{5}}{3}$$

and

$$\gamma^{(1)} = (1, -\sqrt{13})$$

$$\gamma^{(2)} = (-6, \sqrt{13})$$

$$\tilde{\gamma}^{(1)} = (1, -\sqrt{5}/3).$$

$\tilde{\gamma}^{(2)}$ does not exist since A_2 is singular.

As previously noted, $\tilde{\gamma}^{(1)}$ is a better bound than $\gamma^{(1)}$, in this case a far better bound. From the basic inequality

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{V}_K:$$

$$\gamma \cdot \nu(x) = \gamma_1\sqrt{x_1^2 + x_2^2} + \gamma_2|x_3|$$

$$\leq \nu'(Ax) = \sqrt{(2x_1+2x_2+2x_3)^2 + (2x_1+5x_2+3x_3)^2}$$

with $x = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix},$ we may infer that every lower bound $\gamma = (\gamma_1, \gamma_2) \in C[A]$ satisfies

$$\gamma_1 \leq 1; \quad \gamma_2 \leq 2\sqrt{5}; \quad \sqrt{5}\,\gamma_1 + 3\gamma_2 \leq 0 \; .$$

Since $\tilde{\gamma}^{(1)}$ fulfills the first and third conditions sharply, $\tilde{\gamma}^{(1)}$ is a maximal lower bound.

Analogous to (14.17), it is immediate from the definition that

Theorem: $C[A]$ is an intersection of half-spaces: (14.38)

$$C[A] = \bigcap_{x \in \mathcal{V}_K} \tilde{H}_x[A] \tag{14.39}$$

where

$$\tilde{H}_x[A] = \{\gamma \in G^D_{K_o} : \gamma\nu(x) \leq \nu'(Ax)\} \; .$$

Since $C[A] \subset \tilde{H}_x[A]$ for each x, we can obtain restrictions on the set of lower bounds by choosing suitable x:

Theorem: $\forall \gamma \in C[A]: \gamma_i \leq \mathrm{glb}_{\nu', \nu_i}(A_i).$ (14.40)

Proof:

$$\forall x \in u_i: \quad \gamma_i \nu_i(x_i) = \gamma\nu(x) \leq \nu'(Ax) = \nu'(A_i x_i)$$

Thus γ_i is a lower bound for A_i and $\gamma_i \leq \mathrm{glb}_{\nu', \nu_i}(A_i)$

Q.E.D.

Theorem: If $\gamma \epsilon C[A]$ and $\gamma_i > 0$, then (14.41)

$$\frac{\gamma_\mu}{\gamma_i} \leq -\text{lub}_{\nu_i, \nu_\mu}(-A_i^{-1}A_\mu \qquad \mu \neq i . \qquad (14.42)$$

Moreover, if A_i is nonsingular, then the bound of (14.37) is maximal.

Proof:

Let $x_\mu \in \mathcal{U}_\mu$ $(\mu \neq i)$ and let $x_i = -A_i^{-1}A_\mu x_\mu \in \mathcal{U}_i$. Then for $x = x_i + x_\mu$,

$$\gamma\nu(x) = \gamma_i\nu_i(x_i) + \gamma_\mu\nu_\mu(x_\mu)$$

$$\leq \nu'(Ax) = \nu'(A_ix_i + A_\mu x_\mu) = \nu'(4) = 0$$

$$-\frac{\gamma_\mu}{\gamma_i} \geq \frac{\nu_i(x_i)}{\nu_\mu(x_\mu)} = \frac{\nu_i(-A_i^{-1}A_\mu x_\mu)}{\nu_\mu(x_\mu)} \qquad \forall x_\mu \in \mathcal{U}_\mu$$

Thus $-\frac{\gamma_\mu}{\gamma_i}$ is an upper bound for $-A_i^{-1}A_\mu$ whence

$$-\frac{\gamma_\mu}{\gamma_i} \geq \text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu)$$

$$\frac{\gamma_\mu}{\gamma_i} \leq -\text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) .$$

Assume that A_i is nonsingular and that $\exists \gamma \epsilon C[A]$ such that $\tilde{\gamma}^{(i)} \leq \gamma$:

$$\tilde{\gamma}_i^{(i)} = \text{glb}_{\nu,\nu_i}(A_i) \leq \gamma_i$$

$$\tilde{\gamma}_\mu^{(i)} = -\text{glb}_{\nu',\nu_i}(A_i)\text{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) \leq \gamma_\mu \qquad (\mu \neq i) .$$

By Theorem (14.40), $\gamma_i \leq \text{glb}_{\nu',\nu_i}(A_i)$ and therefore $\gamma_i = \text{glb}_{\nu',\nu_i}(A_i) > 0$. By the result just proved,'

$$\frac{y_\mu}{\gamma_i} \leq -\mathrm{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu)$$

$$\gamma_\mu \leq -\gamma_i \ \mathrm{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) \qquad (\mu \neq i)$$

$$= -\mathrm{glb}_{\nu',\nu_i}(A_i)\mathrm{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu) \ .$$

Thus $y_\mu = -\mathrm{glb}_{\nu',\nu_i}(A_i)\mathrm{lub}_{\nu_i,\nu_\mu}(-A_i^{-1}A_\mu)$ and $\gamma = \tilde{\gamma}^{(i)}$, whence Robert's bound is maximal.

Q.E.D.

The full importance of this result will become evident in $15. For now, we note that although 0 and Robert's bound are both lower bounds, they are usually not comparable. Thus there does not generally exist a greatest lower bound for the homomorphism $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$.

Since regular norms are surjective, Theorem (14.10) gives

$$\gamma \in C[A] \succ \gamma \leq \mathrm{lub}_{\nu',\nu}(A) \ . \qquad (14.43)$$

As before, we may now drop the assumption that ν' is a scalar norm and require only that ν and ν' are regular norms. We can now construct lower bounds from the "row-wise" bounds previously discussed.

Let $A \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}'_K)$ and write

$$A = \begin{pmatrix} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_\ell \end{pmatrix}$$

where

$$A_j \in \mathrm{Hom}(\mathcal{V}_K, \mathcal{U}'_j) \ .$$

Then if γ_j is a lower bound for $A_j (1 \leq j \leq \ell)$

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_\ell \end{pmatrix}$$

is a lower bound for A .

If $k = \ell$, then the row bounds may be taken to be the lower bounds of (14.36) in such a manner that there is exactly one element of the form $\text{glb}_{\nu_i', \nu_j}(A_{ij})$ in each row and each column. With suitable reordering, these elements will appear on the diagonal and, if the dimensions of the subspaces in the direct sum decompositions of $\mathcal{V}_K$ and $\mathcal{V}_K'$ coincide, we obtain the lower bound of FIEDLER:

$$H_{Fiedl} = \begin{pmatrix} \text{glb}_{\nu_1', \nu_1}(A_{11}) & -\text{lub}_{\nu_1', \nu_2}(-A_{12}) & \cdots & \text{lub}_{\nu_1', \nu_k}(-A_{1k}) \\ -\text{lub}_{\nu_2', \nu_1}(-A_{21}) & \text{glb}_{\nu_2', \nu_2}(A_{22}) & \cdots & -\text{lub}_{\nu_2', \nu_k}(-A_{2k}) \\ \vdots & \vdots & & \vdots \\ -\text{lub}_{\nu_k', \nu_1}(-A_{k1}) & -\text{lub}_{\nu_k', \nu_2}(-A_{k2}) & \cdots & \text{glb}_{\nu_k', \nu_k}(A_{kk}) \end{pmatrix} \quad (14.44)$$

$$= D_{Fiedl} - U_{Fiedl}$$

where

$$D_{Fiedl} = \text{diag}(\text{glb}_{\nu_1', \nu_1}(A_{11}), \ldots, \text{glb}_{\nu_k, \nu_k}(A_{kk}))$$

$$(U_{Fiedl})_{ij} = \begin{cases} \text{lub}_{\nu_i', \nu_j}(-A_{ij}) & i \neq j \\ 0 & i = j \end{cases}$$

If the A_{ii} are all nonsingular, then Fiedler's bound can be replaced by the bound of Robert

$$H_{Rob} = D_{Fiedl}(I - {}'_{Rob}) \tag{14.45}$$

where

$$(U_{Rob})_{ij} = \begin{cases} lub_{\nu_i,\nu_j}(-A_{ii}^{-1}A_{ij}) & i \neq j \\ 0 & i = j \end{cases} \tag{14.46}$$

Of course, if all subspaces are one-dimensional, then these bounds coincide as do bounds (14.36) and (14.37).

Example:

Let $\mathcal{V}_K = \mathcal{V}'_K = R^n$ and let $\nu = \nu'$ be the modulus norm. Then $VA = (a_{ij}) \in Hom(\mathcal{V}_K,\mathcal{V}'_K) = R^{n\times n}$:

$$H_{Rob} = H_{Fiedl} = \begin{pmatrix} |a_{11}| & -|a_{12}| & \dots & -|a_{1n}| \\ -|a_{21}| & |a_{22}| & \dots & -|a_{2n}| \\ \vdots & \vdots & & \vdots \\ -|a_{n1}| & -|a_{2n}| & \dots & |a_{nn}| \end{pmatrix}. \tag{14.47}$$

Finally, we consider upper and lower bounds for a sum $A_1 + A_2$ of two homomorphisms A_1 $A_2 \in Hom(\mathcal{V}_K,\mathcal{V}'_K)$:

<u>Theorem:</u> Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be upper bounds for A_1 and A_2 respectively. Then $\mathcal{B}_1 + \mathcal{B}_2$ is an upper bound for $A_1 + A_2$. (14.48)

<u>Proof:</u>

$$\nu'((A_1+A_2)x) \rho' \nu'(A_1x) + \nu'(A_2x)$$

$$\rho' \mathcal{B}_1\nu(x) + \mathcal{B}_2\nu(x)$$

$$= (\mathcal{B}_1+\mathcal{B}_2)\nu(x)$$

Q.E.D.

<u>Theorem:</u> Let H_1 be a lower bound for A_1 and let $\mathcal{B}_2$ be an upper bound for $-A_2$. Then $H_1 - \mathcal{B}_2$ is a lower bound for $A_1 + A_2$. (14.49)

<u>Proof:</u>

$$(H_1 - \mathcal{B}_2)\nu(x) = H_1\nu(x) - \mathcal{B}_2\nu(x)$$

$$\rho' \ \nu'(A_1x) - \nu'(-A_2x)$$

$$\rho' \ \nu'((A_1+A_2)x)$$

Q.E.D.

§15. Best Lower Bounds in the Sense of Robert; Applications: Nonsingularity Criteria and Eigenvalue Exclusion Theorems.

Lower bounds for an endomorphism A,. of $\mathcal{V}_K$ may be used to establish nonsingularity of the mapping A and to find upper bounds for A^{-1}. For the case of scalar norms ν and ν',

$$A \text{ singular} \Rightarrow \operatorname{glb}_{\nu',\nu}(A) = 0 .$$

Equivalently,

$$\operatorname{glb}_{\nu',\nu}(A) > 0 \Rightarrow A \text{ nonsingular} .$$

Moreover, since any left inverse for A is the unique two-sided inverse, Theorem (12.15) gives

$$\operatorname{lub}_{\nu',\nu}(A^{-1}) = 1/\operatorname{glb}_{\nu',\nu}(A),$$

an upper bound for A^{-1}.

In the general case, the situation is similar yet in a weaker sense. We shall assume that $\mathcal{V}_K = \mathcal{V}'_K$, that the norms ν and ν' are regular, and that the finite-dimensional vector lattices $(G_{K\circ},\rho)$ and $(G'_{K\circ},\rho')$ have the same dimension. Thus $A \in \operatorname{Hom}(\mathcal{V}_K,\mathcal{V}'_K)$ is an endomorphism of $\mathcal{V}_K$ and all bounds for A, subordinate to ν' and ν, are square matrices. Furthermore, both mappings and bound mappings have the property that M is nonsingular $\Leftrightarrow \exists$ two-sided inverse M^{-1}, where M is either a mapping or a bound mapping.

A lower bound $H \in C[A]$ is said to have a

$$\underline{\text{SEMIPOSITIVE INVERSE } H^{-1}} \qquad (15.1)$$

if H is nonsingular and $0 \check{\rho} H^{-1}$, where $\check{\rho}$ is the ordering induced in

$Hom(G'_{K_o}, G_{K_o})$ by ρ' and ρ . For such a bound, $H^{-1} \neq 0$ and

$$\forall x \in \mathcal{V}_K: \quad \nu(x) \ \rho \ H^{-1}\nu'(Ax) \ . \tag{15.2}$$

Theorem: If $H \in C[A]$ has a semipositive inverse, then A is nonsingular. (15.3)

Proof:

$$x \neq \phi \ > \ \nu(x) \neq 0$$

$$> H^{-1}\nu'(Ax) \neq 0 \qquad \text{by (15.2)}$$

$$> \nu'(Ax) \neq 0 \qquad \text{since } H^{-1} \text{ is nonsingular}$$

$$> Ax \neq \phi$$

Thus A is nonsingular.

Q.E.D.

Example:

Let $G_{K_o} = G'_{K_o} = (R^n, \leq)$. Then $\check{\rho}$ coincides with $\hat{\rho}$ and is just the elementwise ordering of matrices. Thus H is semipositive ⇔ it is componentwise non-negative.

Theorem: If $H \in C[A]$ has a semipositive inverse, then H^{-1} is an upper bound for A^{-1} . (15.4)

Proof:

By Theorem (15.3), A is nonsingular and letting $x = A^{-1}y$ in (15.2):

$$\forall y \in \mathcal{V}'_K: \quad \nu(A^{-1}y) \ \rho \ H^{-1}\nu'(y) \ . \tag{15.5}$$

Q.E.D.

Although H^{-1} is not necessarily the least upper bound for A^{-1} and may be a quite weak upper bound, it often does have the advantage of being more easily calculated. Moreover, it is of some importance in connection with matrix problems which have a natural decomposition into blocks; for example, finite difference approximations in multi-dimensional problems.

Example:

Let $\mathcal{V}_K = R^3$ and let $v = v'$ be the modulus norm. For

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix},$$

The lower bound of (14.47) is just

$$H = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

H has a semipositive inverse

$$H^{-1} = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

which is an upper bound for

$$A^{-1} = \begin{pmatrix} 0.4 & -0.1 & -0.1 \\ -0.1 & 0.4 & -0.1 \\ -0.1 & -0.1 & 0.4 \end{pmatrix}.$$

The least upper bound for A^{-1} is

$$\text{lub}_{\nu,\nu'}(A-1) = \begin{pmatrix} 0.4 & 0.1 & 0.1 \\ 0.1 & 0.4 & 0.1 \\ 0.1 & 0.1 & 0.4 \end{pmatrix} .$$

A mapping $H \in \text{Hom}(G_{K_0}, G'_{K_0})$ is

<u>MONOTONIC</u> (15. 6)

if

$$\forall u, v \in G_{K_0}: \quad Hu\ \rho'\ Hv \succ u\ \rho\ v \ . \tag{15. 7}$$

<u>Theorem:</u> H is monotonic $\approx$ H has a semipositive inverse. (15. 8)

<u>Proof:</u>

Assume that H has a semipositive inverse and that $Hu\ \rho'\ Hv$. Then

$$0\ \check{\rho}\ H^{-1} \quad A\ 0\ \rho' \quad H(u-v) \succ \quad 0 \quad \rho \quad (u-v)$$

whence $u\ \rho\ v$. Therefore H is monotonic. Assume that H is monotonic. Then

$$\begin{aligned} H \text{ singular} &\succ \exists w \neq 0: Hw = 0 \\ &\succ \begin{cases} 0\ \rho\ w & (15 \bullet 7) \text{ with } u=0,\ v=w \\ w\ \rho\ 0 & (15 \bullet 7) \text{ with } u=w,\ v=0 \end{cases} \\ &\succ w = 0, \end{aligned}$$

· a contradiction. Thus H is nonsingular. Moreover,

$$\begin{aligned} 0\ \rho' w \succ \quad H(0) \quad \rho' \ H(H^{-1}w) \\ \succ 0 \rho \ H-1-w \ . \end{aligned}$$

Thus $0\ \check{\rho}\ H^{-1}$ and H has a semipositive inverse.

Q.E.D.

Apart from the question of whether monotonic lower bounds exist, we may want to compare the inverses of monotonic lower bounds under the ordering $\check{\rho}$. Let

$$C^{-1}[A] := \{H^{-1}: \ H\epsilon C[A] \text{ has a semipositive inverse}\}, \qquad (15.9)$$

the set of inverses of the monotonic lower bounds of A . Then

$$C^{-1}[A] \subset \mathcal{B}[A^{-1}] \qquad (15.10)$$

and we may seek minimal or even least elements in $C^{-1}[A]$.

<u>Theorem:</u> If H_1 and H_2 are monotonic, then (15.11)

$$H_1 \hat{\rho} H_2 \succ H_2^{-1} \check{\rho} H_1^{-1} .$$

<u>Proof:</u>

$$\begin{aligned} O\rho' w &\succ 0 \ \rho \ H_2^{-1} w \\ &\succ H_1 H_2^{-1} w \ \rho' \ w \qquad \text{since } H_1 \hat{\rho} H_2 \\ &\succ H_2^{-1} w \ \rho \ H_1^{-1} w \end{aligned}$$

Therefore $H_2^{-1} \check{\rho} H_1^{-1}$.

Q.E.D.

Thus inverses of monotonic bounds which can be replaced in the sense of $\hat{\rho}$ can be replaced in the sense of $\check{\rho}$. The converse is not true as the following example indicates:

Example:

Let

$$H_1 = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \ ; \ H_2 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} \\ 0 & \frac{1}{2} \end{pmatrix} .$$

Then

$$H_1^{-1} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}; \quad H_2^{-1} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

and H_1 and H_2 are monotone. However, although $H_2^{-1} \check{\rho} H_1^{-1}$ it is not true that $H_1 \hat{\rho} H_2$. In fact, H_1 and H_2 are incomparable.

We could instead consider (15.5) from the point of view of obtaining an error estimate by means of residuals. In this case we would desire H to be such that the set

$$\mathcal{D}[H,v] := \{u \in G^+ : Hu \ \rho' \ v\} \tag{15.12}$$

is as small as possible for fixed $v = \nu'(x)$. For the size of

$$\mathcal{D}[H,\nu'(x)] = \{\nu(A^{-1}x) : H\nu(A^{-1}x) \ \rho' \ \nu'(x)\}$$

reflects the size of the set

$$\{A^{-1}z : \ \nu'(z) = \nu'(x)\},$$

the set of possible errors.

<u>Theorem:</u> $$H_1 \ \hat{\rho} \ H_2 \Rightarrow \forall v \in G^+ : \mathcal{D}[H_2,v] \subset \mathcal{D}[H_1,v] \ . \tag{15.13}$$

<u>Proof:</u>

$$\begin{aligned} u \in \mathcal{D}[H_2,v] &\Rightarrow u \in G^+ \wedge H_2 u \ \rho' \ v \\ &\Rightarrow u \in G^+ \wedge H_1 u \ \rho' \ H_2 u \ \rho' \ v \qquad \text{since } H_1 \ \hat{\rho} \ H_2 \\ &\Rightarrow u \in \mathcal{D}[H_1,v] \ . \end{aligned}$$

Q.E.D.

Note that the preceding result does not require monotonicity, but the class of monotonic bounds is again distinguished with respect to $\mathcal{D}[H,v]$:

Theorem: If H is monotonic, then $\forall v \epsilon G'^{+}$: $\mathscr{D}[H,v]$ is nonempty and bounded. (15.14)

Proof:

Since H is monotonic,

$$Hu \ \rho' \ v \succ u \ \rho \ H^{-1}v$$

and

$$\mathscr{D}[H,v] \subset \{u: \quad 0 \ \rho \ u \ \rho \ H^{-1}v\}$$

Thus $\mathscr{D}[H,v]$ is bounded from above and below. Since

$$v \epsilon G'^{+} \succ H^{-1}v \epsilon G^{+},$$

$H^{-1}v \epsilon \mathscr{D}[H,v]$ and the set is nonempty.

Q.E.D.

Thus a monotonic lower bound gives a bounded set of norms of errors and hence a bounded set of errors.

An equivalent characterization of the boundedness of $\mathscr{D}[H,v] \ \forall v \epsilon G'^{+}$ with respect to an Archimedean ordering is given by

Theorem: $\mathscr{D}[H,v]$ is bounded $\forall v \epsilon G'^{+} \approx$ (15.15)

$$(u \epsilon G^{+} \wedge Hu \ \rho' \ 0 \succ u = 0) \ .$$

Proof:

Assume that $\forall v \epsilon G'^{+}$: $\mathscr{D}[H,v]$ is bounded. Let $u \epsilon G^{+}$ with $Hu \ \rho' \ 0$. Then

$$(\forall v \epsilon G'^{+}) (\forall \ \alpha \ \epsilon \ K_0 : \quad 0 \ \rho_0 \ \alpha): \quad \alpha \ u \ \epsilon \ \mathscr{D}[H,v] \ .$$

But the set $\{\alpha u\}$ is unbounded for $u \neq 0$. Thus $u = 0$. Assume that $\exists v \epsilon G'^{+}$: $\mathscr{D}[H,v]$ is not bounded. Then since

$$\{u \epsilon G^+ : 0 \ \rho' \ Hu \ \rho' \ v\}$$

is bounded, its complement with respect to $\mathfrak{D}[H,v]$

$$\{u \epsilon G^+ : \ Hu \ \rho' \ 0)$$

is unbounded and contains a nonzero element; that is, $\exists u \epsilon G^+ : \ Hu \ \rho' \ 0$ and $u \neq 0$.

Q.E.D.

The full importance of Robert's bound (14.45) is indicated in the following two Theorems:

Theorem: The set of all monotonic lower bounds with positive diagonal and non-positive off-diagonal elements is non-empty if and only if it contains H_{Rob} . (15.16)

Proof:

Let H be a monotonic lower bound with positive diagonal and non-positive off-diagonal elements. Then $H = D(I - U)$ where D is diagonal with positive elements and U is off-diagonal with non-negative elements. By Theorem (14.41), $U \geq U_{Rob}$. Since H is monotonic, H^{-1} is semipositive and

$$0 \leq D \wedge 0 \leq H^{-1} = (I - U)^{-1} D^{-1} \succ 0 \leq (I - U)^{-1} .$$

Moreover, since $0 \leq U$, $\forall k \geq 0$: $0 \leq (I - U)^{-1} U^k$. From

$$(I - U)^{-1} U^k = (I - U)^{-1} [I - (I - U^k)]$$
$$= (I - U)^{-1} - (I + u + U^2 + . . . + U^{k-1})$$

we obtain

$$0 \leq I + u + U^2 + \ldots + U^{k-1} < (I - U)^{-1}$$

and the infinite sum being elementwise bounded and nondecreasing,

$$I + U + U^2 + U^3 + \ldots$$

converges. Since $0 \leq U_{Rob} \leq U$, the infinite sum

$$I + U_{Rob} + U^2_{Rob} + {}^3_{Rob} + \ldots$$

also converges. The limit is just $(I - U_{Rob})^{-1}$ which is element-wise non-negative:

$$0 \leq (I - U_{Rob})^{-1}$$

whence

$$0 \leq (I - U_{Rob})^{-1} D^{-1}_{Fiedl} = H^{-1}_{Rob}$$

and H_{Rob} is monotonic. The remainder of the theorem is immediately evident since H_{Rob} is a lower bound with positive diagonal and non-positive off-diagonal elements.

Q.E.D.

Theorem: Among all monotonic lower bounds H which have (15.17) positive diagonal and non-positive off-diagonal elements, the bound H_{Rob} is least in the sense of inclusion of the domains $\mathcal{D}[H,v]$:

$$\forall v \in G'^{+}: \mathcal{D}[H_{Rob}, v] \subset \mathcal{D}[H,v].$$

Proof:

Assume that the set of monotonic lower bounds with the prescribed sign pattern is non-empty. Then by Theorem (15.16)' H_{Rob} is monotonic and an element of this set. Thus

$$u \in \mathcal{D}[H_{Rob}, v] \Rightarrow H_{Rob} u \leq v$$

$$\Rightarrow (I - U_{Rob}) u \leq D^{-1}_{Fiedl} v .$$

Let $H = D(I - U)$ be any other such bound. By Theorems (14.40) and (14.41)'

$$D \leq D_{Fiedl}; \quad U \geq U_{Rob} .$$

Thus

$$(I - U)u \leq (I - U_{Rob})u \leq D^{-1}_{Fiedl}v \leq D^{-1}v$$

or $Hu \leq v$. Therefore $u \in \mathfrak{D}[H,v]$.

Q.E.D.

The monotonic lower bounds which are the subject of Theorems (15.16) and (15.17) are matrices with positive diagonal and non-positive off-diagonal elements which have non-negative inverses. These M-matrices (OSTROWSKI) have been studied in detail by OSTROWSKI, FAN, KOTELJANSKI, and FIEDLER and FTAK. GASTINEL has proposed studying the class of matrices H for which $\mathfrak{D}[H,v]$ is bounded $\forall v \in G'^{+}$, and SCHNEIDER has discussed a related class of matrices. It would be interesting to know how H_{Rob} is characterized within this class which is wider than (Theorem (15.14)) the class of M-matrices.

We shall now apply our results on bound mappings to formulate several nonsingularity criteria and eigenvalue exclusion theorems. Thus we assume that $\mathcal{V}_K = \mathcal{V}'_K$ and $G_{K_0} = G'_{K_0}$ and that the norms $\nu = \nu'$ are regular.

The dual characterization of $lub_{\nu',\nu}(A)$ ((14.20)) for scalar norms ν'

$$lub_{\nu',\nu}(A) = \sup_{\hat{\rho}} \{\nu^D(\varphi A): \varphi \in S'\}$$

is also valid for regular norms ν' . However, a bilinear characterization and the corresponding bilinear field of values do not seem to exist. Moreover, the field of values defined in §6 does not seem to allow a useful generalization.

There is a natural connection between nonsingularity criteria for a matrix A and exclusion theorems for the eigenvalues of A:

<u>Theorem:</u> Let $\wp[A]$ be a statement about an endomorphism $A \in Hom(V_K, V'_K)$. Then the nonsingularity criterion (15.18)

$$\wp[A] \succ A \text{ nonsingular} \tag{15.19}$$

and the eigenvalue exclusion theorem

$$\wp[A - \lambda I] \succ \lambda \text{ is not an eigenvalue of } A \tag{15.20}$$

are equivalent.

<u>Proof</u>:

Assume that

$$\wp[A] \succ A \text{ nonsingular}$$

and let λ be an eigenvalue of A . Then $A - \lambda I$ is singular:

$$\lambda \text{ is an eigenvalue of } A \succ \neg\, \wp[A - \lambda I] \tag{15.21}$$

an equivalent formulation of (15.20). Assume that

$$\wp[A - \lambda I] \succ \lambda \text{ is not an eigenvalue of } A \text{ .}$$

Then

$$\wp[A] \succ \lambda = 0 \text{ is not an eigenvalue of } A$$
$$\succ A \text{ is nonsingular.}$$

Q.E.D.

Example:

Let

$$\wp[A] := \forall i : |a_{ii}| > {}_{j\,i} |a_{ij}| \text{ . .}$$

Then Gerschgorin's Theorem may be stated as

$$\lambda \epsilon \{z \epsilon C: \exists i: |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|\} \qquad (15.22)$$

$$= \{z \epsilon C: \neg \wp[A - zI]\}$$

or

$$\lambda \text{ is an eigenvalue of } A \succ \neg \wp[A - zI] .$$

The equivalent nonsingularity criterion

$$\forall i : |a_{ii}| > \sum_{j \neq i} |a_{ij}| \succ A \text{ nonsingular} \qquad (15.23)$$

was discovered by LEVY in the nineteenth century.

There is a direct proof of the preceding result ((15.23)) as a special case (take. B to be the diagonal of A and consider the lub subordinate to the Tschebyscheff norm) of the following nonsingularity criterion:

Theorem: If B is nonsingular and $lub(I - B^{-1}A) < 1$, then A is nonsingular. (15.24)

Proof:

$$1 > lub(I - B^{-1}A)$$

$$\geq lub(I) - glb(B^{-1}A)$$

$$= 1 - glb(B^{-1}A)$$

$$0 < glb(B^{-1}A)$$

Thus $B^{-1}A$ is nonsingular as is $A = B \cdot B^{-1}A$.

Q.E.D.

The corresponding exclusion theorem is

$$\mathrm{lub}(I - B^{-1}(A - \lambda I)) \geq 1 \Rightarrow \lambda \text{ is not an eigenvalue of } A . \quad (15.25)$$

B is usually chosen to be $C - \lambda I$ provided $C - \lambda I$ is nonsingular. In this case, (15.25) becomes

$$\mathrm{lub}((C - \lambda I)^{-1}(C - A)) \geq 1 \Rightarrow \lambda \text{ is not an eigenvalue of } A . \quad (15.26)$$

Equivalently, the set

$$\{z \epsilon C: \ (C - \lambda I) \text{ is singular or } \mathrm{lub}((C - \lambda I)^{-1}(C - A)) < 1\} \quad (15.27)$$

contains all the eigenvalues of A . The preceding results are all comparison theorems and their usefulness depends on the choice of B or C .

The nonsingularity criterion of Theorem (15.13) leads immediately to

<u>Theorem</u>: Let $H \in C[A - xI]$ be a monotonic lower bound for $A - \lambda I$. Then λ is not an eigenvalue of A . (15.28)

This result can be applied to $H_{Fiedl}[A - \lambda I]$ and $H_{Rob}[A - \lambda I]$. In the first case, it can be slightly modified.

<u>Lemma</u>: If H_1 and H_2 have positive diagonal and non-positive off-diagonal elements and $H_1 \leq H_2$, then (15.29)

$$H_1 \text{ monotonic} \Rightarrow H_2 \text{ monotonic} .$$

<u>Proof</u>: Compare the Neumann series for H_1 and H_2 .

<u>Theorem</u>: (Fiedler - Ptak) If λ is an eigenvalue of A, then (15.30)

$$\lambda \in \bigcup_i \{z: \ \mathrm{glb}_{\nu_i', \nu_i}(A_{ii} - zI) < c_i\}$$

provided $c_1, c_2, \ldots, c_k$ are such that

$$\mathrm{diag}(c_1, c_2, \ldots, c_k) - U_{\mathrm{Fiedl}}[A]$$

is monotonic.

Proof:

Assume that $c_1, c_2, \ldots, c_k$ are such that

$$H_c = \mathrm{diag}(c_1, c_2, \ldots, c_k) - U_{\mathrm{Fiedl}}[A]$$

is monotonic. Then by the preceding Lemma, if

$$\mathrm{glb}_{\nu_i', \nu_i}(A_{ii}) \geq c_i$$

Then $H_{\mathrm{Fiedl}}[A - \lambda I]$ is monotonic and λ is not an eigenvalue of A.

Q.E.D.

In the case of the modulus norm, we get a sharpened form of Gerschgorin's Theorem due to KOTELJANSKII and FAN:

$$\lambda \in \bigcup_i \{z: |a_{ii} - z| < c_i\} . \qquad (15.31)$$

§16. Submultiplicative Functionals on Half-categories and Semigroups; Normed Categories and Rings

In preceding sections, we have considered mappings between a pair of vector spaces $\mathcal{V}_K$ and $\mathcal{V}'_K$ and, in some special cases, mappings of a vector space $\mathcal{V}_K$ into itself. If instead we have a family of vector spaces $\mathcal{V}_K^{(i)}$ where i ranges over some finite or infinite index set I, then we may consider homomorphisms between any two members of this family. If the vector spaces are normed, then these norms induce upper and lower bounds for the homomorphisms.

Let A: $\mathcal{V}_K^{(i)} \to \mathcal{V}_K^{(j)}$ and B: $\mathcal{V}_K^{(\ell)} \to \mathcal{V}_K^{(m)}$ be homomorphisms. Then, provided the range $U_K^{(j)}$ of A coincides with the domain $U_K^{(\ell)}$ of B, the product BA can be naturally defined as A composed with B . Since composition of mappings is an associative operation, if A, B, C are homomorphisms and A(BC) exists, then (AB) C also exists and (AB)c = A(BC) . We may abstract this algebraic structure of homomorphisms and define a

<u>HALF-CATEGORY</u> (16.1)

as a set $\mathfrak{M}$ together with an associative partial composition:

$$\forall A, B, C \in \mathfrak{M}: A(BC) \text{ exists} \Rightarrow (AB)C \text{ exists and } (AB)C = A(BC) . \tag{16.2}$$

Example:

The set of all finite matrices is a half-category with the usual definition of matrix multiplication (if the number of columns of A is not equal to the number of rows of B, then the product AB is not defined). It is the half-category of homomorphisms corresponding to the family of vector spaces $\{R^n\}_{n=1}^{\infty}$.

A half-category in which composition is defined for every pair of elements is a

SEMIGROUP. (16.3)

A functional N on a half-category $\mathfrak{M}$ with values from a (j-ordered half-category $\mathfrak{D}$ is

SUBMULTIPLICATIVE (16.4)

if for all A, B $\in \mathfrak{M}$ such that AB exists, N(A) N(B) also exists and

$$N(AB) \mathrel{\dot{\rho}} N(A)\, N(B) \qquad (16.5)$$

Example:

On the half-category of finite matrices, the mapping

$$A \to |A| \quad \text{where } A = (a_{ij}) \quad \text{and } |A| = (|a_{ij}|)$$

is a submultiplicative functional.

Theorem: Let $\{\mathcal{V}_K^{(i)}\}_{i\in I}$ be a family of normed vector spaces (16.6) with regular norms $\nu^{(i)}: \mathcal{V}_K^{(\cdot)} \to (G_{K_0}^{(i)}, \rho^{(i)})$. For $A \in \mathrm{Hom}(\mathcal{V}_K^{(i)}, \mathcal{V}_K^{(j)})$, define

$$\mathrm{lub}(A) := \mathrm{lub}_{\nu^{(j)},\nu^{(i)}}(A) .$$

($\mathrm{lub}_{\nu^{(j)},\nu^{(i)}}(A)$ exists since $\nu^{(\cdot)}$ and $\nu^{(j)}$ are regular norms.)

Then lub is a sub-multiplicative functional on the half-category of homomorphisms

$$\mathfrak{M} = \{A: \exists i,j\in I: A \in \mathrm{Hom}(\mathcal{V}_K^{(i)}, \mathcal{V}_K^{(j)})\}$$

with values from the i-ordered half-category

$$\mathfrak{P} = \{\mathcal{B}: \exists i,j\in I: \mathcal{B} \in \mathrm{Hom}(G_{K_0}^{(i)}, G_{K_0}^{(j)})\},$$

the ordering $\dot{\rho}$ being the ordering induced on $\mathrm{Hom}(G_{K_0}^{(i)}, G_{K_0}^{(j)})$ by $\rho^{(i)}$ and $\rho^{(j)}$.

Proof:

If A, B $\epsilon\ \mathcal{M}$ and AB exists, then the domain of A coincides with the range of B:

$$B: \ \mathcal{V}_K^{(i)} \to \mathcal{V}_K^{(j)} \ ; \quad A: \ \mathcal{V}_K^{(j)} \to \mathcal{V}_K^{(\ell)} .$$

From the definition of $\mathrm{lub}_{\nu', \nu}$,

$$\nu^{(\ell)}(ABx) \ \rho^{(\ell)} \ \mathrm{lub}_{\nu^{(\ell)}, \nu^{(j)}}(A) \ . \ \nu^{(j)}(Bx)$$

$$\rho^{(\ell)} \ \mathrm{lub}_{\nu^{(\ell)}, \nu^{(j)}}(A) \cdot \mathrm{lub}_{\nu^{(j)}, \nu^{(i)}}(B) \cdot \nu^{(i)}(x)$$

Therefore $\mathrm{lub}_{\nu^{(\ell)}, \nu^{(j)}}(A) \ . \ \mathrm{lub}_{\nu^{(j)}, \nu^{(i)}}(B)$ is an upper bound for AB and

$$\mathrm{lub}_{\nu^{(\ell)}, \nu^{(i)}}(AB) \ \dot{\rho} \ \mathrm{lub}_{\nu^{(\ell)}, \nu^{(j)}}(A) \ . \ \mathrm{lub}_{\nu^{(j)}, \nu^{(i)}}(B) \qquad (16.7)$$

where $\dot{\rho}$ is the ordering induced by $\rho^{(\ell)}$ and $\rho^{(i)}$. Equivalently,

$$\mathrm{lub}(AB) \ \dot{\rho} \ \mathrm{lub}(A) \ \mathrm{lub}(B) \ . \qquad (16.7)'$$

Q.E.D.

The inequality (16.7) may be weakened. If $\sqcap_{\nu^{(\ell)}, \nu^{(j)}}$ and $\sqcap_{\nu^{(j)}, \nu^{(i)}}$ are upper bound mappings ((11.4))' then

$$\mathrm{lub}_{\nu^{(\ell)}, \nu^{(i)}}(AB) \ \dot{\rho} \ \lceil A \rceil_{\nu^{(\ell)}, \nu^{(j)}} \ . \ \lceil B \rceil_{\nu^{(j)}, \nu^{(i)}} \qquad (16.8)$$

$$\mathrm{lub}(AB) \ \dot{\rho} \ \lceil A \rceil \lceil B \rceil \ . \qquad (16.8)'$$

The proof of Theorem (16.6) depended upon the existence of a least upper bound for AB with respect to $\nu^{(\ell)}$ and $\nu^{(i)}$. Since the greatest lower bound for a mapping does not exist in general in the case of non-scalar norms, we cannot expect a corresponding result for glb. However, if we restrict ourselves to scalar norms, then

Theorem: Let $\{\mathcal{V}_K^{(i)}\}_{i\in I}$ be a family of normed vector spaces with scalar norms $\nu^{(i)} : \mathcal{V}_K^{(i)} \to (R,\leq)$. For $A \in \mathrm{Hom}(\mathcal{V}_K^{(i)},\mathcal{V}_K^{(j)})$, define (16.9)

$$\mathrm{glb}(A) := \mathrm{glb}_{\nu^{(j)},\nu^{(i)}}(A) .$$

Then glb is a SUPEXMULTIPLICATIVE functional on the half-category of homomorphisms

$$\{A: \exists i,j\in I: A \in \mathrm{Hom}(\mathcal{V}_K^{(i)},\mathcal{V}_K^{(j)})\}$$

with values from the $\leq$ - ordered half-category R; that is

$$\mathrm{glb}(AB) \geq \mathrm{glb}(A)\,\mathrm{glb}(B) . \tag{16.10}$$

As before, inequality (16.10) may be weakened:

$$\mathrm{glb}(AB) \geq \lfloor A \rfloor \lfloor B \rfloor \tag{16.11}$$

where $\lfloor\;\rfloor$ is any lower bound mapping ((11.7)).

Theorem: (16.12)

$$\mathrm{lub}(AB) \geq \mathrm{glb}(A)\,\mathrm{lub}(B)$$

$$\mathrm{glb}(AB) \leq \mathrm{lub}(A)\,\mathrm{glb}(B) . \tag{16.13}$$

Proof:

$\forall \epsilon > 0\ \exists y \neq \Phi$:

$$\mathrm{lub}(B) - \epsilon \leq \frac{\nu^{(j)}(By)}{\nu^{(i)}(y)} .$$

Letting x = By,

$$\text{glb}(A) \cdot [\text{lub}(B) - \epsilon] \leq \frac{\nu^{(\ell)}(Ax)}{\nu^{(j)}(x)} \cdot [\text{lub}(B) - \epsilon]$$

$$\leq \frac{\nu^{(\ell)}(ABy)}{\nu^{(j)}(By)} \cdot \frac{\nu^{(j)}(By)}{\nu^{(i)}(y)}$$

$$\leq \frac{\nu^{(\ell)}(ABy)}{\nu^{(i)}(y)}$$

$$\leq \text{lub}(AB)$$

Similarly, $\forall \epsilon > 0 \exists y \neq \phi$:

$$\text{glb}(B) - \epsilon \geq \frac{\nu^{(j)}(By)}{\nu^{(i)}(y)} \quad .$$

Letting x = By,

$$\text{lub}(A) \cdot [\text{glb}(B) - \epsilon] \geq \frac{\nu^{(\ell)}(Ax)}{\nu^{(j)}(x)} \cdot \frac{\nu^{(j)}(By)}{\nu^{(i)}(y)}$$

$$\geq \frac{\nu^{(\ell)}(ABy)}{\nu^{(i)}(y)}$$

$$\geq \text{glb}(AB) \; .$$

Q.E.D.

Note that in the proof of the preceding theorem, it would not have been possible to vary x first since there might not exist a y such that By = x . However,

<u>Theorem:</u> If B is surjective, then (16.14)

$$\text{lub}(AB) \geq \text{lub}(A)\,\text{glb}(B)$$

$$\text{glb}(AB) < \text{glb}(A)\,\text{lub}(B) \; .$$

In the case where A and B are both endomorphisms, the last pair of inequalities always hold: for if B is nonsingular, then B is surjective; and if B is singular, then AB is also singular and glb(B) = glb(AB) = 0 .

We may define a second partial composition, addition, between the elements of a half-category of homomorphisms. The sum of two homomorphisms A and B is just

$$(A+B)(x) := A(x) + B(x),$$

provided the domains and ranges of A and B coincide. The resulting algebraic structure is called a CATEGORY. A subadditive, submultiplicative, definite functional on a category is a MULTIPLICATIVE NORM on that category. Restating several earlier results ((12.5) and (16.6)):

Theorem: lub is a multiplicative norm on the category of homomorphisms corresponding to a family of vector spaces. (16.15)

Examples:

(i) The set of all finite matrices is a category with the usual definitions of addition and multiplication. The mapping

$$A \to |A| \quad \text{where } A = (a_{ij}) \quad \text{and } |A| = (|a_{ij}|) \tag{16.16}$$

is a lub subordinate to the modulus norm and therefore a multiplicative norm, the MODULUS NORM on the category of finite matrices.

(ii) On the category of finite matrices' the mapping

$$\|A\|_F = \left(\sum_{\mu=1}^{m} \sum_{\nu=1}^{n} |a_{\mu\nu}|^2 \right)^{\frac{1}{2}}, \quad A \in \mathrm{Hom}(R^n, R^m) \tag{16.17}$$

is a multiplicative norm, the FROBENIUS NORM. This norm is not a lub-NORM subordinate to two vector norms and indicates that not all multiplicative norms are so generated.

(iii) If the family of vector spaces $\{V_K^{(i)}\}_{i\in I}$ consists of only one vector space U_K, then the corresponding set $\mathfrak{M}$ of homomorphisms is a set of endomorphisms. Thus addition and multiplication are defined for any pair of elements of $\mathfrak{M}$. Considered as a half-category, $\mathfrak{M}$ is a semigroup. Considered as a category, $\mathfrak{M}$ is a ring since multiplication distributes over addition .

Henceforth we shall restrict our attention to scalar (real-valued) multiplicative norms $\|\ldots\|$ over a ring $\mathfrak{R}$ with identity. We define

$$\lceil A \rceil := \|A\| \tag{16.18}$$

$$\lfloor A \rfloor := \begin{cases} 0, & \text{if } A \text{ has a right zero divisor} \\ \inf\{\|A^L\|^{-1} : A^L \text{ is a left inverse of } A], & \text{otherwise} \end{cases} \tag{16.19}$$

where we have assumed in addition that $\mathfrak{R}$ satisfies

If A has no right zero divisors, then A has at least one left inverse. (16.20)

To further simplify matters, we also assume that

a (unique) two-sided inverse A^{-1} for A exists $\Leftrightarrow$ A has no right zero divisors. (16.21)

In this case, (16.19) becomes

$$\lfloor A \rfloor = \begin{cases} \|A^{-1}\|^{-1}, & \text{if } A^{-1} \text{ exists} \\ 0\ , & \text{if } A \text{ has a right zero divisor.} \end{cases} \tag{16.19)"}$$

Theorem:

$$\lceil A \rceil - \lceil B \rceil \leq \lceil A+B \rceil \leq \lceil A \rceil + \lceil B \rceil \tag{16.22}$$

$$\lfloor A \rfloor - \lceil B \rceil \leq \lfloor A+B \rfloor \leq \lfloor A \rfloor + \lceil B \rceil \tag{16.23}$$

$$\lceil A \rceil \cdot \lfloor B \rfloor \leq \lceil AB \rceil \leq \lceil A \rceil \cdot \lceil B \rceil \tag{16.24}$$

$$\lfloor A \rfloor \cdot \lfloor B \rfloor \leq \lfloor AB \rfloor \leq \lfloor A \rfloor \cdot \lceil B \rceil \tag{16.25}$$

Proof:

(i) (16.22) follows from the subadditivity of $\|\ldots\|$ and a systematic change of variables (see (10.3)).

(ii) Assume that A^{-1} and $(A+B)^{-1}$ exist. Then

$$A^{-1} = (A+B)^{-1}(I + BA^{-1})$$
$$= (A+B)^{-1} + (A+B)^{-1}BA^{-1}$$
$$\|A^{-1}\| \leq \|(A+B)^{-1}\| \cdot (1 + \|B\| \, \|A^{-1}\|)$$
$$\|(A+B)^{-1}\|^{-1} \leq \|A^{-1}\|^{-1} + \|B\|$$
$$\lfloor A+B \rfloor \leq \lfloor A \rfloor + \lceil B \rceil \ .$$

Assume that $(A+B)^{-1}$ exists and A^{-1} does not exist. Then A has a right zero divisor $(\exists X \neq 0: \ AX=0)$ and

$$X = (A+B)^{-1}(A+B)X$$
$$= (A+B)^{-1}BX$$
$$\|X\| \leq \|(A+B)^{-1}\| \|B\| \, \|X\|$$
$$\|(A+B)^{-1}\|^{-1} \leq \|B\| \quad \text{since} \quad X \neq 0 \Rightarrow \|X\| \neq 0$$
$$\lfloor A+B \rfloor \leq \lceil B \rceil$$
$$\lfloor A+B \rfloor \leq \lfloor A \rfloor + \lceil B \rceil \ \text{ since } \ \lfloor A \rfloor = 0 \ .$$

Assume that $(A+B)^{-1}$ does not exist. Then $\lfloor A+B \rfloor = 0$ and the preceding inequality is again vaild. The left hand side of (16.23) is obtained by a systematic change of variables (see(10.3)).

(iii) The right hand side of (16.24) follows from the submultiplicativity of $\|\ldots\|$. Assume that B^{-1} exists. Then

$$A = AB \cdot B^{-1}$$

$$\|A\| \leq \|AB\| \, . \, \|B^{-1}\|$$

$$\|A\| \, . \, \|B^{-1}\|^{-1} \leq \|AB\|$$

$$\lceil A \rceil \, . \, \lfloor B \rfloor \leq \lceil AB \rceil \, .$$

Assume that B^{-1} does not exist. Then $\lfloor B \rfloor = 0$ and the preceding inequality is still satisfied.

(iv) Assume that $(AB)^{-1}$ exists. Then A^{-1} also exists and

$$A^{-1} = B(AB)^{-1}$$

$$\|A^{-1}\| \leq \|B\| \, \|(AB)^{-1}\|$$

$$\|(AB)^{-1}\|^{-1} \leq \|A^{-1}\|^{-1} \, . \, \|B\|$$

$$\lfloor AB \rfloor \leq \lfloor A \rfloor \, \lceil B \rceil$$

Assume that $(AB)^{-1}$ does not exist. Then $\lfloor AB \rfloor = 0$ and the preceding inequality again holds. Assume that A^{-1} and B^{-1} exist. Then $(AB)^{-1} = B^{-1}A^{-1}$ exists and

$$\|(AB)^{-1}\| \leq \|B^{-1}\| \, . \, \|A^{-1}\|$$

$$\|A^{-1}\|^{-1} \, . \, \|B^{-1}\|^{-1} \leq \|(AB)^{-1}\|^{-1}$$

$$\lfloor A \rfloor \cdot \lfloor B \rfloor \leq \lfloor AB \rfloor \, .$$

Assume that A^{-1} does not exist. Then $(AB)^{-1}$ does not exist and $\lfloor A \rfloor = \lfloor AB \rfloor = 0$, whence the preceding inequality is trivially satisfied. A similar situation occurs when B^{-1} does not exist.

Q.E.D.

The following inequalities are proved in an analogous manner:

$$\lfloor A \rfloor \cdot \lceil B \rceil \leq \lceil AB \rceil$$
$$\lfloor AB \rfloor \leq \lceil A \rceil \, \lfloor B \rfloor \, . \qquad (16.26)$$

Example:

Inequalities (16.22) - (16.26) are valid for any real-valued multiplicative norm on a finite matrix ring.

The question naturally arises as to whether the functionals $\lceil\ \rceil$ and $\lfloor\ \rfloor$ as defined by (16.18) and (16.19) are upper and lower bound mappings respectively in the sense of (11.4) and (11.7); that is, whether there exists a normed vector space $\mathcal{V}_K$ such that for every $A \in \mathfrak{R}$, A is an endomorphism of U_K and $\lceil A \rceil$ and $\lfloor A \rfloor$ are an upper and lower bound for the mapping distortion. For the case where $\mathfrak{R}$ is a ring over some field K (i.e., an algebra), the answer is given by

Theorem: Let $\mathfrak{R}$ be a ring over the field K and let $\mathcal{V}_K$ be (16.27) $\mathfrak{R}$ (taken additively) or some proper left ideal of $\mathfrak{R}$. Define anormon $\mathcal{V}_K$ by $v(X) = \|X\|$. Then

$$\mathfrak{R} \subseteq \mathrm{Hom}(\mathcal{V}_K, \mathcal{V}_K)$$

and

$$\lfloor A \rfloor \leq \frac{\nu(AX)}{\nu(X)} \leq \lceil A \rceil \qquad (X \neq 0) .$$

Proof:

The first assertion is an immediate consequence of the fact that multiplication on the left by an element of $\mathfrak{R}$ is an endomorphism of any left ideal of $\mathfrak{R}$ ($\mathfrak{R}$ is itself a left ideal of $\mathfrak{R}$) . From

$$\nu(AX) = \|AX\| \leq \|A\| \, \|X\| = \lceil A \rceil \, \nu(X)$$

we obtain

$$\frac{\nu(Ax)}{\nu(x)} \leq \lceil A \rceil$$

provided $\nu(X) \neq 0$ $(X \neq 0)$. Assume that A^{-1} exists. Then

$$\nu(X) = \|A^{-1}AX\|$$
$$\leq \|A^{-1}\| \cdot \|AX\|$$
$$\|A^{-1}\|^{-1} \leq \|AX\|/\nu(X)$$
$$\lfloor A \rfloor \leq \frac{\nu(AX)}{\nu(X)} \quad (X \neq 0) .$$

Assume that A^{-1} does not exist. Then $\lfloor A \rfloor = 0$ and the preceding inequality is trivially satisfied.

Q.E.D.

Henceforth we shall assume that $\mathfrak{R}$ is a ring over the field K . For any non-zero ideal $\mathfrak{J}$ of $\mathfrak{R}$ we may define $\forall A \in \mathfrak{R}$:

$$\text{lub}_{\mathfrak{J}}(A) := \sup\left\{\frac{\|AX\|}{\|X\|}: \quad X \in \mathfrak{J}, \quad X \neq 0\right\} \tag{16.28}$$

$$\text{glb}_{\mathfrak{J}}(A) := \inf\left\{\frac{\|AX\|}{\|X\|}: \quad X \in \mathfrak{J}, \quad X \neq 0\right\} . \tag{16.29}$$

Clearly,

<u>Theorem:</u> (16.30)

$$\lfloor A \rfloor \leq \text{glb}_{\mathfrak{J}}(A) \leq \text{lub}_{\mathfrak{J}}(A) \leq \lceil A \rceil .$$

Moreover, as a consequence of Lemma (9.2),

<u>Theorem</u>: If $\mathfrak{J}_1 \subset \mathfrak{J}_2$ are left ideals of $\mathfrak{R}$, then (16.31)

$$\text{lub}_{\mathfrak{J}_1}(A) \leq \text{lub}_{\mathfrak{J}_2}(A)$$
$$\text{glb}_{\mathfrak{J}_2}(A) \leq \text{glb}_{\mathfrak{J}_1}(A) .$$

Finally,

<u>Theorem:</u> $\text{lub}_{\mathfrak{R}}$ is a multiplicative norm on $\mathfrak{R}$. (16.32)

Proof:

Subadditivity and submultiplicativity are inherited from $\|\ldots\|$; the proof is analogous to that of Theorems (12.5) and (16.6) and may be carried out for $\text{lub}_{\mathcal{J}}$ for any non-zero ideal $\mathcal{J}$. Definiteness is also inherited from $\|\ldots\|$:

$$\text{lub}_{\mathfrak{R}}(A) = 0 \Rightarrow \frac{\|AX\|}{\|X\|} = 0 \quad \forall X \in \mathfrak{R}, \quad X \neq 0$$

$$\Rightarrow \|AX\| = 0 \quad \forall X \in \mathfrak{R}$$

$$\Rightarrow AX = 0 \quad \forall X \in \mathfrak{R}$$

$$\Rightarrow A = O .$$

The last step is not in general valid for a proper nonzero ideal $\mathcal{J}$ so that although $\text{lub}_{\mathcal{J}}$ is subadditive and submultiplicative, it is not usually a multiplicative norm.

Q.E.D.

Theorem: (16.33)

$$\text{glb}_{\mathfrak{R}}(A) = \begin{cases} 1/\text{lub}_{\mathfrak{R}}(A^{-1}), & \text{if } A^{-1} \text{ exists.} \\ 0 \quad , & \text{otherwise.} \end{cases}$$

Proof:

If A^{-1} exists, then the proof is the same as that of Theorem (12.14). Otherwise, A has a right zero divisor and the infimum is zero.

Q.E.D.

As a consequence of the preceding theorems, $\text{lub}_{\mathfrak{R}}$ and $\text{glb}_{\mathfrak{R}}$ satisfy inequalities (16.22) - (16.26).

Appendix: Historical and Bibliographical Notes

The concept of norms came up around the turn of the century in algebra for a sum of squares (rather than the square root of it). In vector spaces and functional spaces, O. Hölder used it for the first time in a wider sense--to include the Euclidean and the Tschebyscheff norm (which had been used somewhat implicitly by Tschebyscheff.) These norms, among others, have the property that $|x_i| \leq \|x\|$, and the Tschebyscheff norm is the best among them in the sense that these inequalities are sharp for at least one i . Later, abstract properties of norms have been used to define them, mainly in connection with metric topologies (Lindenbaum, Banach in the twenties). See

S. Banach. Théorie der Operations linéaries. Warszawa 1932.

For a modern treatment, in particular of the topological side, see

Kelly, Namioka.

A topology that has a special connection with norms, the weak topology, has been introduced by Tychonoff. Norms were introduced into numerical analysis by Faadeva and, more systematically, by Householder. In Banach spaces, partial ordering has been studied by

L. Kantorovitch. in Mat. Sbornik N. S. 44 121 - 168 (1937).

Further details, in particular about positivity cones, can be found in

M. G. Krein and M. A. Rutman. Linear operators leaving invariant a cone in a banach space (1948). English translation in Amer. Math. Soc. Transl. Series I, 10.

From a more algebraic side, partially ordered groups and vector spaces have been studied by Freudenthal (1936), inspired by Riesz. Stone, Birkhoff and Lorenzen have developed the theory further. See chapters XIV, xv of

G. Birkhoff, Lattice Theory. Revised ed. Providence 1961 (a new edition is in preparation).

and

H. Gericke. Theorie der Verb&de. Mannheim 1963.

Questions of imbedding in direct products of linearly ordered lattice groups have been studied by Mannos and Lorenzen. Rudimentary steps were already taken by Dedekind in 1897,

A. Dedekind. Werke, Vol. 2, 103 -148.

Vectorial norms seemingly were first considered by Kantorovitch ("spaces normal with the elements of a semi-ordered space"). See

L. Kantorovitch. The method of successive approximations for functional equations. Acta Math. 71, 62 - 97 (1939)

The first published results on bounds are due to Fiedler and Ptak (1960):

M. Fiedler and V. Ptak. Generalized norms of matrices and the location of the spectrum. Czech. Math. J. 1-2, 558 - 571 (1962).

More work on bounds has been done by Ostrowski (1960 Madison Report No. 138) and by Robert (to appear in Num. Math.). M-matrices, which show up in this connection, were studied by Fan, Kotelyanskii and in particular by Fiedler and Ptak in 1960:

M. Fiedler and V. Ptak. On matrices with non-positive off-diagonal elements and positive principal minors. Czech. Math. J. 12, 302 - 400 (1962).

A wide class of matrices was introduced by Hans Schneider in 1964:

H. Schneider. Positive operators and an inertia theorem, Num. Math. 7, 11 - 17 (1965).

Multiplicative norms have been studied by Gastinel, Focke and Stoer. Stoer has in particular characterized matrix norms which are also lub norms:

J. Stoer. On the characterization of least upper bound norms in matrix space. Numer. Math. 6, 302 - 314 (1964).

Further concepts in a comprehensive theory of norms would include:

Condition numbers: based on norms. These have been introduced by Householder and Bauer. Their relation to certain matrix transformations has been investigated in

F. L. Bauer. Optimally scaled matrices. Numer. Math. 5, 73 - 87 (1963).

Fields of values: Connected with the support tangents to a field of values is a functional which turns out to be a directional derivation of the lub . See

F. L. Bauer. On the field of values subordinate to a norm. Numer. Math. 4, 103 - 113 (1962).

and

N. Nirschl and H. Schneider. The Bauer fields of values of a matrix. Numer. Math. 6, 355 - 365 (1964).

Composite norms: A variety of multiplicative norms, defined by some composition, have been studied by

A. M. Ostrowski. Über Normen von Matrizen. Math. Z. 63, 2 - 18 (1955).

More recently, Maitre (to appear in Numer. Math.) has obtained more results in this direction.

Unitarily invariant norms: Multiplicative norms which are invariant under two-sided unitary transformations were studied by J. von Neumann.

Absolute norms: Absolute norms are norms which depend only on the absolute values of the coordinates. See

F. L. Bauer, J. Stoer, C. Witzgall. Absolute and monotonic norms. Numer. Math. 3 257 - 264 (1961).

and more recently

D. Gries. Über einige Klassen von Normen. Thesis. Munich, 1966.

who also discusses fields of values.

Most of the norms used in practice are absolute, and this has an important consequence: the lub of a diagonal matrix is equal to the maximum of the absolute values of the diagonal elements. For Hölder norms, which are absolute, Stoer has given an abstract characterization:

J. Stoer. A characterization of Hölder norms. J. Soc. Indust. Appl. Math. 12, 634 - 648 (1964).

In the theory of partially ordered vector spaces, some recent developments due to Birkhoff, Hopf and Ostrowski have led to an interesting submultiplicative functional or non-negative mappings which is homogeneous of degree zero. Connected with this is a bound for the oscillation of a vector. This and other concepts playing a role in this connection deserve great attention. See

F. L. Bauer. An elementary proof of the Hopf inequality for positive operators. Numer. Math. 7, 331 - 337 (1965).

Related to this theory is the generalization of the Perron-Frobenius theorem to a large class of positivity cones (Krein-Rutman). See

H. Schneider. Positive operators and an inertia theorem. Numer. Math. 7, 11 - 17 (1965).