Numerische Mathematik 6, 302—314 (1964)

# On the Characterization of Least Upper Bound Norms in Matrix Space

By
JOSEF STOER

## 1.

It is well known that every vector norm $\|x\|$ in $C^n$ gives rise to a matrix norm $\operatorname{lub}(A)$ in the space $C^{n^2}$ of all square matrices of order $n$

$$\operatorname{lub}(A) := \sup_{x \neq 0} \frac{\|A\,x\|}{\|x\|},$$

which is *consistent* with the underlying vector norm:

$$\|A\,x\| \leqq \operatorname{lub}(A) \cdot \|x\|$$

and is a *multiplicative* matrix norm: The inequality

$$\operatorname{lub}(A\,B) \leqq \operatorname{lub}(A) \cdot \operatorname{lub}(B)$$

holds for all matrices $A$ and $B$.

Any vector norm $\|x\|$ uniquely defines a convex body

$$B := \{x \mid \|x\| \leqq 1\} \subseteq C^n,$$

which is compact and contains the origin as an interior point. Conversely, for any compact convex neighbourhood $B$ of the origin in $C^n$, there exists a vector norm $\|x\|$

$$\|x\| := \inf\{\omega \geqq 0 \mid x \in \omega\,B\}, \qquad \text{where} \qquad \omega\,B := \{\omega\,x \mid x \in B\},$$

with $B = \{x \mid \|x\| \leqq 1\}$. Likewise, every matrix norm $\nu(A)$ in $C^{n^2}$ is associated with a compact convex neighbourhood

$$H := \{A \mid \nu(A) \leqq 1\}$$

of the origin in $C^{n^2}$, and vice versa. But clearly, not every convex compact neighbourhood of the origin in $C^{n^2}$ belongs to a lub norm. The principal problem to be considered here is the geometric characterization of those convex bodies in $C^{n^2}$ which are associated with lub norms. The question arises because of the fact that many of the properties of vector norms fall quite easily and naturally out of a consideration of the associated convex bodies, and it is hoped that these results will throw some light on the rather more difficult questions related with multiplicative matrix norms and lub norms.

This paper incorporates several unpublished results of Dr. A. S. HOUSEHOLDER, Prof. H. SCHNEIDER and Prof. F. L. BAUER. The author admits gratefully that this paper has been initiated by their investigations and he wishes to thank them for communicating to him their results and for many discussions on this subject.

## 2.

In the sequel, we shall assume that all vector norms $\|x\|$ considered are

*strictly homogeneous*,

i.e. $\|\alpha x\| = |\alpha| \cdot \|x\|$ holds for all complex $\alpha$. Since in this case $\mathrm{lub}(A)$ is also strictly homogeneous, it is natural to require hereafter that all matrix norms $\nu(A)$ encountered are also strictly homogeneous.

Now, a first result can be stated about matrix norms and multiplicative matrix norms:

*For any matrix norm $\nu(A)$, there exists a scalar $\sigma > 0$ such that*

$$\|A\| := \sigma \cdot \nu(A)$$

*is a multiplicative matrix norm.*

This theorem does not seem to be in the literature. A special case has been stated by Gastinel [5].

For the proof, we remark only that in view of the compactness of $\{A \mid \nu(A) = 1\}$,

$$\varkappa = \max\{\nu(AB) \mid \nu(A) = \nu(B) = 1\}$$

exists and is finite. Let $\sigma \geqq \varkappa$. Evidently every such $\sigma$ is effective. Moreover, $\sigma = \varkappa$ is effective and optimal.

Given any vector norm $\|x\|$, a matrix norm $\nu(A)$ is said to be consistent with it, if

$$\|Ax\| \leqq \nu(A) \cdot \|x\|.$$

If $\nu(A)$ is any multiplicative matrix norm, then for any fixed vector $a \neq 0$,

$$\|x\| := \nu(x a^H)$$

defines a vector norm, and the matrix norm $\nu(A)$ is consistent with it because of the multiplicativity $\nu(AB) \leqq \nu(A) \cdot \nu(B)$. Thus every multiplicative matrix norm is consistent with some vector norm. Evidently, $\mathrm{lub}(A)$ is by definition the smallest matrix norm which is consistent with $\|x\|$. Hence, $\nu(A) \geqq \mathrm{lub}(A)$ holds for every matrix norm $\nu(A)$ which is consistent with the vector norm $\|x\|$ generating $\mathrm{lub}(A)$. Geometrically this means that if $H$ is the convex body associated with the lub norm, then it contains the convex body associated with any other matrix norm consistent with the given vector norm. Thus convex bodies associated with lub norms are in this sense maximal.

## 3.

A deeper insight into the nature of lub norms is gained by using the powerful tool of duality. Associated with any vector norm $\|x\|$ in $n$-space, which may be identified with the space of all $n$-dimensional column vectors, is the dual norm

$$\|y^H\|^D := \sup_{x \neq 0} \frac{\mathrm{Re}\, y^H x}{\|x\|} = \sup_{x \neq 0} \frac{\mathrm{Re}\,\mathrm{tr}(y^H x)}{\|x\|} = \sup_{x \neq 0} \frac{\mathrm{Re}\,\mathrm{tr}(x y^H)}{\|x\|} \tag{1}$$

defined in the linear space of all $n$-dimensional row vectors $y^H$. From this the Hölder inequality

$$\mathrm{Re}\, y^H x \leqq \|y^H\|^D \cdot \|x\|$$

follows. Since it is assumed that $\|x\|$ is strictly homogeneous, this is equivalent to

$$(2) \qquad |y^H x| \leqq \|y^H\|^D \cdot \|x\| \quad \text{and} \quad \|y^H\|^D = \sup_{x \neq 0} \frac{|y^H x|}{\|x\|} .$$

Two vectors $x \neq 0$, $y^H \neq 0$ yielding equality in (2) are said to be mutually dual, which is denoted by $x \| y^H$.

The convex bodies $B$ and $B^D$ belonging to the norms $x$ and $\|y^H\|^D$, respectively, are related by polarity:

$$B^D = \{y^H \mid \operatorname{Re} y^H x \leqq 1 \text{ for all } x \in B\}.$$

The concept of the polar

$$K^D := \{y^H \mid \operatorname{Re} y^H x \leqq 1 \text{ for all } x \in K\}$$

is meaningful for arbitrary sets $K$ in $n$-space. For later use, we note the well known relation (see for instance EGGLESTON [4])

$$(3) \qquad K^{DD} = \{x \mid \operatorname{Re} y^H x \leqq 1 \text{ for all } y^H \in K^D\} = \overline{\mathfrak{H}(K \cup \{0\})},$$

i.e. $K^{DD}$ is the closure of the convex hull of $K$ and 0. Since for each vector norm $\|x\|$ the set $B = \{x \mid \|x\| \leqq 1\}$ is a closed convex neighbourhood of the origin, the relation

$$B^{DD} = \overline{\mathfrak{H}(K \cup \{0\})} = B$$

shows that

$$(4) \qquad \|x\|^{DD} = \sup_{y^H \neq 0} \frac{\operatorname{Re} y^H x}{\|y^H\|^D} = \max_{y^H \neq 0} \frac{\operatorname{Re} y^H x}{\|y^H\|^D} = \|x\|$$

holds for any vector norm $\|x\|$.

Clearly, the concepts of a dual norm, and of polarity of sets have their counterparts in the linear space $C^{n^2}$ of all square matrices of order $n$: If $\nu(A)$ is a matrix norm, its dual norm $\nu^D(A)$ may be defined by

$$\nu^D(A) = \sup_{B \neq 0} \frac{\operatorname{Re} \operatorname{tr}(AB)}{\nu(B)} .$$

If $K$ is a set in this matrix space, its polar $K^D$ is

$$K^D := \{A \mid \operatorname{Re} \operatorname{tr}(A B) \leqq 1 \text{ for all } B \in K\}.$$

These definitions are the natural extensions of the corresponding definitions in $n$-space.

In the sequel, the class of matrices of rank 1 is very important. Their lub norm is easily calculated, since the norm is strictly homogeneous:

*If $A = x y^H$ is any matrix of rank 1, then*

$$\operatorname{lub}(A) = \|y^H\|^D \cdot \|x\|.$$

In fact, the definition of $\operatorname{lub}(x y^H)$ and (2) yield immediately

$$\operatorname{lub}(x y^H) = \sup_{u \neq 0} \frac{\|x y^H u\|}{\|u\|} = \sup_{u \neq 0} \frac{|y^H u| \cdot \|x\|}{\|u\|} = \|y^H\|^D \cdot \|x\|.$$

Denote by

$$P := \{x y^H \mid \operatorname{lub}(x y^H) \leqq 1\}$$

the set of all matrices $A = x y^H$ with $\operatorname{lub}(A) \leqq 1$.

Then, it is easy to show

$$\operatorname{lub}(A) = \sup_{B\in P} \operatorname{Re}\operatorname{tr}(A\,B)\,. \tag{5}$$

Indeed, we obtain by (4)

$$\begin{aligned}\sup_{B\in P} \operatorname{Re}\operatorname{tr}(A\,B) &= \sup_{\substack{x\neq 0\\ y^H\neq 0}} \frac{\operatorname{Re}\operatorname{tr}(A\,x\,y^H)}{\|x\|\,\|y^H\|^D} = \sup_{\substack{x\neq 0\\ y^H\neq 0}} \frac{\operatorname{Re} y^H A\,x}{\|y^H\|^D\,\|x\|}\\ &= \sup_{x\neq 0}\left\{\frac{1}{\|x\|}\sup_{y^H\neq 0}\frac{\operatorname{Re} y^H A\,x}{\|y^H\|^D}\right\} = \sup_{x\neq 0}\frac{\|A\,x\|}{\|x\|}\\ &= \operatorname{lub}(A)\,.\end{aligned}$$

But, this implies immediately

$$\begin{aligned}P^D &= \{A\,|\operatorname{Re}\operatorname{tr}(A\,B)\leqq 1 \text{ for all } B\in P\}\\ &= \{A\,|\operatorname{lub}(A)\leqq 1\},\end{aligned}$$

and therefore the inclusion

$$P\subseteq P^D$$

holds.

The main result can now be stated:

**Theorem 1.** *Every lub norm has the following properties:*

a) $\{A\,|\operatorname{lub}(A)\leqq 1\} = P^D$.

b) $\{A\,|\operatorname{lub}^D(A)\leqq 1\} = P^{DD} = \mathfrak{H}\{A = x\,y^H\,|\operatorname{lub}^D(A)\leqq 1\}$.

c) $\operatorname{lub}(A)\leqq \operatorname{lub}^D(A)$ *for all square matrices* $A$ *of order* $n$.

d) $\operatorname{lub}(A) = \operatorname{lub}^D(A)$ *for all matrices* $A = x\,y^H$ *of rank* 1 or 0.

e) $\operatorname{lub}^D(A) = \inf\left\{\sum_i \lambda_i \,\middle|\, A = \sum_i \lambda_i B_i,\ \lambda_i\geqq 0,\ B_i\in P\right\}$.

f) $\operatorname{lub}^D(A)$ *is a multiplicative matrix norm, which is consistent with the vector norm* $\|x\|$ *generating* $\operatorname{lub}(A)$.

*Proof.* Property a) has already been established. Since a) is true, and $P\subseteq P^D$, the definition of the dual norm $\operatorname{lub}^D(A)$ of $\operatorname{lub}(A)$ implies at once

$$\begin{aligned}\operatorname{lub}^D(A) := \sup_{B\neq 0}\frac{\operatorname{Re}\operatorname{tr}(A\,B)}{\operatorname{lub}(B)} &= \sup_{B\in P^D}\operatorname{Re}\operatorname{tr}(A\,B)\\ &\geqq \sup_{B\in P}\operatorname{Re}\operatorname{tr}(A\,B) = \operatorname{lub}(A)\,,\end{aligned}$$

which establishes c), and by (3)

$$P^{DD} = \{A\,|\operatorname{lub}^D(A)\leqq 1\} = \overline{\mathfrak{H}\,(P\cup\{0\})}\,.$$

Obviously, $0\in P$ and $P$ is a compact set, which proves

$$P^{DD} = \overline{\mathfrak{H}\,(P\cup\{0\})} = \mathfrak{H}\,(P)\,, \tag{6}$$

since the convex hull of a compact set is also compact (see EGGLESTON [4]).

d) is a consequence of c), since for matrices $A = x\,y^H$

$$\begin{aligned}\operatorname{lub}^D(A) &= \sup_{B\neq 0}\frac{\operatorname{Re}\operatorname{tr}(A\,B)}{\operatorname{lub}(B)} = \sup_{B\neq 0}\frac{\operatorname{Re} y^H B\,x}{\operatorname{lub}(B)} \leqq \sup_{B\neq 0}\frac{\|y^H\|^D\cdot\operatorname{lub}(B)\cdot\|x\|}{\operatorname{lub}(B)}\\ &= \operatorname{lub}(A)\,.\end{aligned}$$

Thus, $\operatorname{lub}(x y^H) = \operatorname{lub}^D(x y^H)$, and therefore

$$P^{DD} = \{A \mid \operatorname{lub}^D(A) \leqq 1\} = \mathfrak{H}(P) = \mathfrak{H}(\{A = x y^H \mid \operatorname{lub}^D(A) \leqq 1\}),$$

which proves b).

e) is implied by (6): If $\operatorname{lub}^D(A) = 1$, then $A \in \mathfrak{H}(P)$, and $A$ is a convex combination of matrices $B_i \in P$:

$$A = \sum_i \lambda_i B_i, \quad \text{where} \quad \operatorname{lub}(B_i) = 1, \quad \lambda_i \geqq 0, \quad \sum_i \lambda_i = 1.$$

On the other hand, if

$$A = \sum_i \mu_i B_i, \quad B_i \in P, \quad \mu_i \geqq 0$$

is another decomposition of $A$, then by the norm properties of $\operatorname{lub}^D$:

$$\operatorname{lub}^D(A) \leqq \sum_i \mu_i \operatorname{lub}^D(B_i) \leqq \sum_i \mu_i,$$

because of d). This establishes e).

f) The consistency of $\operatorname{lub}^D(A)$ follows directly from c). It remains to be shown that $\operatorname{lub}^D(A)$ is multiplicative. But this is implied by c):

$$\begin{aligned}\operatorname{lub}^D(AB) &= \sup_{C \neq 0} \frac{\operatorname{Re}\operatorname{tr}(ABC)}{\operatorname{lub}(C)} \leqq \sup_{C \neq 0} \frac{\operatorname{lub}^D(A)\operatorname{lub}(BC)}{\operatorname{lub}(C)} \\ &\leqq \sup_{C \neq 0} \frac{\operatorname{lub}^D(A)\operatorname{lub}(B)\operatorname{lub}(C)}{\operatorname{lub}(C)} = \operatorname{lub}^D(A)\operatorname{lub}(B) \\ &\leqq \operatorname{lub}^D(A)\operatorname{lub}^D(B),\end{aligned}$$

which completes the proof of Theorem 1.

Theorem 1 leads to

**Lemma 1.** *For every matrix norm $\nu(A)$ that is consistent with the vector norm $\|x\|$ and satisfies*

$$\nu(A) \leqq \operatorname{lub}(A) \quad \textit{for all } A = x y^H,$$

*the inequalities*

$$\operatorname{lub}(A) \leqq \nu(A) \leqq \operatorname{lub}^D(A)$$

$$\operatorname{lub}(A) \leqq \nu^D(A) \leqq \operatorname{lub}^D(A)$$

*hold for all matrices $A$.*

Indeed, the consistency of $\nu(A)$ with $x$ means that $\operatorname{lub}(A) \leqq \nu(A)$ for all $A$, and therefore, by the hypotheses of the Lemma, $\operatorname{lub}(A) = \nu(A)$ for all $A = x y^H$. But now Theorem 1, b), d) implies $\nu(A) \leqq \operatorname{lub}^D(A)$ for all $A$ proving the first inequality

$$\operatorname{lub}(A) \leqq \nu(A) \leqq \operatorname{lub}^D(A).$$

The second one follows by forming the dual of the previous inequality:

$$\operatorname{lub}(A) \leqq \nu^D(A) \leqq \operatorname{lub}^D(A).$$

A further interesting consequence of Theorem 1 is the

**Theorem of STRANG** [*6*]. *If* $\operatorname{lub}_1(A)$ *and* $\operatorname{lub}_2(A)$ *are two lub norms and* $\operatorname{lub}_1(A) \leqq \operatorname{lub}_2(A)$ *for all $A$, then* $\operatorname{lub}_1(A) = \operatorname{lub}_2(A)$.

In fact, $\operatorname{lub}_1(A) \leqq \operatorname{lub}_2(A)$ implies $\operatorname{lub}_1^D(A) \geqq \operatorname{lub}_2^D(A)$ for all $A$, and therefore by Theorem 1, c),

$$\operatorname{lub}_1(A) \leqq \operatorname{lub}_2(A) \leqq \operatorname{lub}_2^D(A) \leqq \operatorname{lub}_1^D(A).$$

Since $\operatorname{lub}_1(A) = \operatorname{lub}_1^D(A)$ for all $A = x y^H$, the last relation, and Theorem 1, b) lead at once to

$$\operatorname{lub}_2^D(A) = \operatorname{lub}_1^D(A) \quad \text{for all } A,$$

which proves $\operatorname{lub}_1(A) = \operatorname{lub}_2(A)$ for all $A$.

Strang's Theorem can be used in order to find a first characterization of lub norms. For this purpose, we call a multiplicative matrix norm $\nu(A)$

*minimal multiplicative*

if it is minimal among all multiplicative matrix norms — that is, if there is any multiplicative matrix norm $\mu(A)$ satisfying $\mu(A) \leqq \nu(A)$ for all $A$, then $\mu(A) = \nu(A)$ for all $A$. Then we can state, following an idea by Gastinel [5]:

**Theorem 2.** *Every lub norm is minimal multiplicative and vice versa.*

*Proof.* We note first that if $\mu(A)$ is a multiplicative matrix norm, then we can find a lub norm $\operatorname{lub}_\mu(A)$ with

$$\operatorname{lub}_\mu(A) \leqq \mu(A) \quad \text{for all } A.$$

Indeed, define the vector norm

$$\|x\|_\mu := \mu(x a^H)$$

for some fixed vector $a \neq 0$. Then, by a familiar argument, the multiplicativity of $\mu(A)$ yields

$$\operatorname{lub}_\mu(A) = \sup_{x \neq 0} \frac{\|A x\|_\mu}{\|x\|_\mu} \leqq \mu(A).$$

In order to show that every lub norm is minimal multiplicative, assume that $\mu(A)$ is a multiplicative matrix norm such that

$$\mu(A) \leqq \operatorname{lub}(A) \quad \text{for all } A.$$

Then, by definition of $\operatorname{lub}_\mu(A)$,

$$\operatorname{lub}_\mu(A) \leqq \mu(A) \leqq \operatorname{lub}(A) \quad \text{for all } A,$$

and the Theorem of Strang implies

$$\operatorname{lub}_\mu(A) = \mu(A) = \operatorname{lub}(A) \quad \text{for all } A.$$

Conversely, let $\nu(A)$ be a minimal multiplicative matrix norm. Then $\operatorname{lub}_\nu(A) \leqq \nu(A)$ for all $A$. Since $\operatorname{lub}_\nu(A)$ is a multiplicative matrix norm and $\nu$ is minimal, we conclude

$$\operatorname{lub}_\nu(A) = \nu(A) \quad \text{for all } A,$$

which completes the proof of Theorem 2.

It is easy to show that the properties b) and c) of Theorem 1 are sufficient to characterize the lub norms among all multiplicative matrix norms:

**Theorem 3.** *If $\nu(A)$ is a multiplicative matrix norm satisfying*

a) $\nu(A) \leqq \nu^D(A)$

*and*

b) $\{A \mid \nu^D(A) \leqq 1\} = \mathfrak{H}(\{A = x y^H \mid \nu^D(A) \leqq 1\})$,

*then* $\nu(A) = \operatorname{lub}(A)$, *where*

$$\operatorname{lub}(A) = \sup_{x \neq 0} \frac{\|A x\|}{\|x\|}$$

*is the lub norm subordinate to the vector norm* $\|x\| := \nu(x a^H)$ *for some fixed* $a \neq 0$, *and* $\|x\|$ *is, up to a factor, independent of* $a$.

*Proof.* Let for some vector $a \neq 0$, $\|x\| = \nu(x a^H)$. Then by the multiplicativity of $\nu(A)$, we obtain

$$\operatorname{lub}(A) \leqq \nu(A) \quad \text{for all } A.$$

Therefore, $\operatorname{lub}^D(A) \geqq \nu^D(A)$, which shows

$$\operatorname{lub}(A) \leqq \nu(A) \leqq \nu^D(A) \leqq \operatorname{lub}^D(A).$$

Theorem 1, d) yields

$$\operatorname{lub}(A) = \nu(A) = \nu^D(A) = \operatorname{lub}^D(A) \quad \text{for all} \quad A = x y^H,$$

and by the hypothesis b) of Theorem 3, and Theorem 1, b)

$$\operatorname{lub}^D(A) = \nu^D(A) \quad \text{for all } A,$$

proving $\operatorname{lub}(A) = \nu(A)$ for all $A$ and for arbitrary $a \neq 0$.

A similar result can be proved for matrix norms:

**Theorem 4.** *If $\nu(A)$ is a matrix norm which has the properties*

a) $\nu(A) = \nu^D(A)$ *for all* $A = x y^H$,

b) $\{A \mid \nu^D(A) \leqq 1\} = \mathfrak{H}(\{A = x y^H \mid \nu^D(A) \leqq 1\})$,

c) *there is a vector* $a \neq 0$ *such that*

$$\nu(x y^H u a^H) \leqq \nu(x y^H)\, \nu(u a^H)$$

*for all vectors* $x, y, u$ *("weak multiplicativity"),*

*then* $\nu(A) = \operatorname{lub}(A)$, *where* $\operatorname{lub}(A)$ *is generated by the vector norm* $\|x\| := \nu(x a^H)$.

To begin the proof, we obtain from c)

$$\operatorname{lub}(A) \leqq \nu(A) \quad \text{for all} \quad A = x y^H,$$

and therefore

$$\operatorname{lub}^D(A) \leqq \nu^D(A) \quad \text{for all} \quad A = x y^H,$$

since $\nu^D(x y^H) = \nu(x y^H)$ and $\operatorname{lub}(x y^H) = \operatorname{lub}^D(x y^H)$. Thus, b) and Theorem 1, b), d) imply

$$\operatorname{lub}^D(A) \leqq \nu^D(A) \quad \text{for all } A.$$

Together with Theorem 1, this proves

$$\nu(A) \leqq \operatorname{lub}(A) \leqq \operatorname{lub}^D(A) \leqq \nu^D(A) \quad \text{for all } A.$$

Hence, by a),

$$\operatorname{lub}^D(A) = \nu^D(A) \quad \text{for all} \quad A = x y^H,$$

and therefore, by b) and Theorem 1, b),

$$\operatorname{lub}^D(A)=\nu^D(A) \quad \text{for all } A,$$

which shows $\operatorname{lub}(A)=\nu(A)$, and Theorem 4 is proved.

Theorem 1 shows that if the matrix norm $\nu(A)$ satisfies

(7) $$\operatorname{lub}(A)\leqq\nu(A)\leqq\operatorname{lub}^D(A)$$

then

$$\nu(AB)=\operatorname{lub}(AB)\leqq\operatorname{lub}(A)\operatorname{lub}(B)=\nu(A)\cdot\nu(B)$$

for all matrices $A=x y^H$, $B=u v^H$ of rank 1; i.e. $\nu(A)$ is weakly multiplicative (see Theorem 4). Thus the question arises whether relation (7) implies that $\nu(A)$ is a multiplicative matrix norm. The following counterexample shows that this is not true:

Take as a vector norm

$$\|x\|:=\max_i |x_i|.$$

The corresponding $\operatorname{lub}(A)$ and $\operatorname{lub}^D(A)$ are easily determined:

$$\operatorname{lub}(A)=\max_i\sum_k |a_{ik}|, \qquad \operatorname{lub}^D(A)=\sum_k\max_i |a_{ik}|.$$

For the matrices

$$A:=\begin{pmatrix}1&1\\2&0\end{pmatrix}, \quad A^2=\begin{pmatrix}3&1\\2&2\end{pmatrix}$$

we have, therefore,

$$\operatorname{lub}(A)=2, \qquad \operatorname{lub}^D(A)=3$$
$$\operatorname{lub}(A^2)=4, \qquad \operatorname{lub}^D(A^2)=5.$$

Define the matrix norm $\nu(C)$ by setting

$$\{C\,|\,\nu(C)\leqq 1\}:=\mathfrak{H}\left(\{C\,|\operatorname{lub}^D(C)\leqq 1\}\cup\left\{C=e^{i\varphi}\cdot\frac{A}{2}\,\middle|\,\varphi \text{ real}\right\}\right).$$

Clearly, because of $\operatorname{lub}\left(e^{i\varphi}\cdot\frac{A}{2}\right)=1$ and $\operatorname{lub}^D\left(e^{i\varphi}\cdot\frac{A}{2}\right)=\frac{3}{2}>1$, this construction of $\nu(C)$ guarantees that $\nu(C)$ is a strictly homogeneous matrix norm satisfying

$$\operatorname{lub}(C)\leqq\nu(C)\leqq\operatorname{lub}^D(C) \quad \text{for all } C$$

and simultaneously $\nu(A)=2$. Moreover, the relation $\nu(A^2)=5$, which will be proved at once, shows that

$$\nu(A^2)=5>\nu(A)^2=4$$

and, therefore, $\nu(C)$ is not multiplicative. In order to show $\nu(A^2)=\operatorname{lub}^D(A^2)=5$, note that the inequality ($I=$ identity matrix)

$$\operatorname{Re}\operatorname{tr}(IC)\leqq\operatorname{Re}\operatorname{tr}\left(I\frac{A^2}{5}\right)=1 \quad \text{for all } C \text{ with } \operatorname{lub}^D(C)\leqq 1$$

exhibits the hyperplane

$$E:=\{C\,|\operatorname{Re}\operatorname{tr}(IC)=\operatorname{Re}\operatorname{tr}(C)=1\}$$

in matrix space as a supporting plane of the convex body

$$B:=\{C\,|\operatorname{lub}^D(C)\leqq 1\}$$

through the boundary point $A^2/5$ of $B$. Moreover, since

$$\operatorname{Re}\operatorname{tr}\left(I \cdot e^{\varphi i} \cdot \frac{A}{2}\right) \leqq \operatorname{tr}\left(\frac{A}{2}\right) = \frac{1}{2} < 1$$

for all real $\varphi$, $E$ is also a supporting plane of the convex body $B' := \{C \mid \nu(C) \leqq 1\}$. This proves that $A^2/5$ lies also on the boundary of $B'$, and therefore $\nu(A^2) = 5$, which was to be shown.

The maximum norm $\|x\| = \max_i |x_i|$ is also interesting for another reason. In theorem 1 we have proved

$$\operatorname{lub}(A) \leqq \operatorname{lub}^D(A) \quad \text{for all } A,$$

but

$$\operatorname{lub}(A) = \operatorname{lub}^D(A) \quad \text{for all } A = x\,y^H.$$

This suggests the question, whether equality

$$\operatorname{lub}(A) = \operatorname{lub}^D(A)$$

holds only for matrices $A = x\,y^H$ of rank 1 or 0. The matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ shows that this suggestion is not true for the maximum norm. However, the next theorem demonstrates that this irregularity cannot occur if the vector norm $\|x\|$ is smooth enough:

**Theorem 5.** *If the vector norm $\|x\|$ or its dual $\|y^H\|^D$ is differentiable for all points with exception of the origin, then*

$$\operatorname{lub}(A) = \operatorname{lub}^D(A)$$

*holds if and only if $A$ is a matrix of rank 1 or 0.*

*Proof.* Because of Theorem 1 we need only show the "only if" part of Theorem 5. We assume without loss of generality that

$$\operatorname{lub}(A) = \operatorname{lub}^D(A) = 1,$$

and that $\|x\|$ is a differentiable vector norm in $C^n$. The case of a differentiable $\|y^H\|^D$ can be treated analogously.

Because of $\operatorname{lub}^D(A) = 1$, we get by Theorem 1, e)

$$A = \sum_i \lambda_i\, x_i\, y_i^H, \quad \text{where} \quad \|x_i\| \cdot \|y_i^H\|^D = 1, \quad \lambda_i \geqq 0, \quad \sum_i \lambda_i = 1.$$

Now, by Theorem 1, a),

$$\operatorname{lub}(A) = \sup_{B \in P} \operatorname{Re}\operatorname{tr}(A\,B)$$

and there exists a matrix $B_0 = x_0\, y_0^H \in P$ with $\operatorname{lub}(B_0) \leqq 1$ such that

$$\begin{aligned}
1 = \operatorname{lub}(A) &= \operatorname{Re}\operatorname{tr}(A\,B_0) = \operatorname{Re} y_0^H A\, x_0 \\
&= \sum_i \lambda_i\, |y_0^H\, x_i| \cdot |y_i^H\, x_0| \\
&\leqq \sum_i \lambda_i \|y_0^H\|^D \|x_i\| \cdot \|y_i^H\|^D \|x_0\| \\
&= \sum_i \lambda_i \operatorname{lub}(x_0\, y_0^H) \operatorname{lub}(x_i\, y_i^H) \\
&\leqq \sum_i \lambda_i = 1.
\end{aligned}$$

But this can only be true if

$$|y_0^H x_i| = \|y_0^H\|^D \|x_i\|, \quad \text{or} \quad x_i \| y_0^H$$
$$|y_i^H x_0| = \|y_i^H\|^D \|x_0\|, \quad \text{or} \quad x_0 \| y_i^H$$

for all $i=1, 2, \ldots$, that is, the hyperplanes

$$E_i := \left\{ x \,\middle|\, \operatorname{Re} \frac{y_i^H}{\|y_i^H\|^D} x = \|x_0\| \right\}$$

are supporting planes of the convex set

$$B := \{x \mid \|x\| \leqq \|x_0\|\}$$

passing through the boundary point $x_0$ of $B$. It is well known that the differentiability of $\|x\|$ implies at once that there is exactly one such supporting plane through $x_0$. Therefore,

$$y_i^H = \mu_i \bar{y}^H$$

must be valid for some constants $\mu_i > 0$ and a vector $\bar{y}^H \neq 0$. This proves

$$A = \sum_i \lambda_i x_i y_i^H = \left(\sum_i \lambda_i \mu_i x_i\right) \cdot \bar{y}^H,$$

and $A$ is a matrix of rank 1 or 0.

## 4.

As an interesting application of these ideas we shall exhibit a result of Bauer [1] concerning the function

$$m(A) := \sup\{\operatorname{Re}\operatorname{tr}(A\,B) \mid B = x\,y^H \,\&\, \operatorname{lub}(B) = \operatorname{Re} y^H x = \operatorname{Re}\operatorname{tr}(B) = 1\}$$

as a special case of a general theorem. This function was considered by Bauer in order to obtain localisation theorems for the eigenvalues of the matrix $A$. He proved the following relation between $m(A)$ and $\operatorname{lub}(A)$:

**Theorem.** *The equation*

$$m(A) = \lim_{\tau\to\infty} \left(\operatorname{lub}(A + \tau I) - \tau\right), \qquad I = \textit{identity matrix},$$

*holds for every square matrix $A$ of order $n$.*

*Proof.* We note that for every matrix $B = x y^H$ the Hölder inequality for lub and $\operatorname{lub}^D$ yields

$$\begin{aligned}\operatorname{Re} y^H x &= \operatorname{Re}\operatorname{tr}(B) = \operatorname{Re}\operatorname{tr}(I\,B) \leqq \operatorname{lub}^D(B)\operatorname{lub}(I)\\ &= \operatorname{lub}^D(B) = \operatorname{lub}(B),\end{aligned}$$

by Theorem 1, d). Therefore, the plane

$$E := \{A \mid \operatorname{Re}\operatorname{tr}(I\,A) = 1\} \tag{8}$$

is a supporting plane of the convex body

$$P^{DD} = \{A \mid \operatorname{lub}^D(A) \leqq 1\} = \mathfrak{H}(P) = \mathfrak{H}\left(\{B = x\,y^H \mid \operatorname{lub}^D(B) = \operatorname{lub}(B) \leqq 1\}\right).$$

Hence, $m(A)$ can be equivalently defined by

$$(9)\qquad m(A) := \sup\{\operatorname{Re}\operatorname{tr}(A\,B) \mid \operatorname{lub}^D(B) = \operatorname{Re}\operatorname{tr}(B) = 1\}$$
$$= \sup\{\operatorname{Re}\operatorname{tr}(A\,B) \mid B \in E \cap P^{DD}\}.$$

As noted in [1], we remark further that the relation maintained in the Theorem is equivalent to

$$m(A) = \lim_{\sigma \to 0} \frac{\operatorname{lub}(I + \sigma A) - \operatorname{lub}(I)}{\sigma}$$
$$=: \operatorname{lub}'(I; A),$$

where $\operatorname{lub}'(I; A)$ is the directional derivative, or "Richtungsderivierte" (see BONNESEN, FENCHEL [3]) of the convex function lub at the point $I$ in the direction of the matrix $A$. It is well known (see [3, 4]) that $\operatorname{lub}'(I; A)$ exists always.

Now, a result of BONNESEN and FENCHEL states that if

$$f(u^H) := \max_{x \in K} \operatorname{Re} u^H x$$

is defined as the *support function* of the compact convex set $K$ then the directional derivative

$$f'(u_0^H; u^H) := \lim_{\sigma \to 0} \frac{f(u_0^H + \sigma u^H) - f(u_0^H)}{\sigma}$$

is the support function of the convex set $K_1 := E_{u_0} \cap K$, where $E_{u_0} := \{x \mid \operatorname{Re} u_0^H x = f(u_0^H)\}$ is a supporting plane of $K$, i.e.

$$f'(u_0^H; u^H) = \max\{\operatorname{Re} u^H x \mid x \in E_{u_0} \cap K\}.$$

When applied to (8) and (9), this relation yields at once

$$m(A) = \operatorname{lub}'(I; A),$$

since $\operatorname{lub}(A) = \sup\{\operatorname{Re}\operatorname{tr}(A\,B) \mid \operatorname{lub}^D(B) \leqq 1\}$ is the support function of the convex set $P^{DD} = \{B \mid \operatorname{lub}^D(B) \leqq 1\}$.

## 5.

Finally, let us consider $\operatorname{lub}(A)$ and $\operatorname{lub}^D(A)$ subordinate to the euclidean norm $\|x\| := (x^H x)^{\frac{1}{2}}$. It is well known that given a matrix $A$ there exists a unitary matrix $U_0$ and a hermitian matrix $H_0$ such that

$$(10)\qquad A = H_0 U_0 \quad \text{and} \quad \operatorname{tr}(H_0) = \sum_{i=1}^{n} \mu_i,$$

the so called polar decomposition, where $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_n \geqq 0$ are the singular values of $A$. Moreover, there exist unitary matrices $U_1$ and $V_1$ such that

$$(11)\qquad U_1^H A\, V_1 = D := \operatorname{diag}(\mu_1, \ldots, \mu_n).$$

Further, it is well known that $\operatorname{lub}(A) = \mu_1$.

Now, if

$$A = \sum_i \lambda_i\, x_i\, y_i^H, \qquad x_i^H x_i \leqq 1, \qquad y_i^H y_i \leqq 1, \qquad \lambda_i \geqq 0,$$

and if $U$ and $V$ are any unitary matrices, then

$$U^H A V = \sum_i \lambda_i (U^H x_i)(y_i^H V), \qquad (U^H x_i)^H (U^H x_i) \leqq 1, \qquad (V^H y_i)^H (V^H y_i) \leqq 1,$$

so that, by Theorem 1, e)

$$\operatorname{lub}^D(A) = \operatorname{lub}^D(U^H A V),$$

hence the norm $\operatorname{lub}^D$ is unitarily invariant. Thus, by (11), there is no restriction in considering matrices of the form $D$ only.

But the euclidean norm $\|x\|$ is an absolute norm (i.e. $\|x\|$ depends only on the moduli $|x_i|$ of the components $x_i$ of $x$), and therefore, by a result of [2],

$$\operatorname{lub}^D\bigl(\operatorname{diag}(d_1, \ldots, d_n)\bigr) = \sum_{i=1}^{n} |d_i|$$

holds for every diagonal matrix. This implies

$$\operatorname{lub}^D(A) = \operatorname{lub}^D(D) = \sum_{i=1}^{n} \mu_i.$$

Thus the norm $\operatorname{lub}^D(A)$ is the sum of the singular values.

Further, by (10), we have

$$\operatorname{lub}^D(A) = \sup\{\operatorname{Re}\operatorname{tr}(A\,U) \mid U^H U = I\}.$$

This implies that the convex body associated with $\operatorname{lub}(A)$ is the convex hull of all unitary matrices:

$$\{A \mid \operatorname{lub}(A) \leqq 1\} = \mathfrak{H}\{U \mid U^H U = 1\}.$$

As a consequence, the following remarkable formula for $\operatorname{lub}(A)$ is obtained (compare Theorem 1, e)):

$$\operatorname{lub}(A) = \mu_1 = \inf\Bigl\{\sum_i \lambda_i \Bigm| A = \sum_i \lambda_i U_i, \quad \text{where} \quad U_i^H U_i = I \quad \text{and} \quad \lambda_i \geqq 0\Bigr\}.$$

Such a minimal decomposition of a matrix $A$ into a weighted sum of unitary matrices $U_i$ can be obtained in the following way. Assume first $A = D = \operatorname{diag}(\mu_1, \ldots, \mu_n)$, where $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_n \geqq 0$. Then, the decomposition of $D$ is

$$D = \sum_{i=1}^{n} \varkappa_i E_i, \qquad \text{where} \qquad E_i := \operatorname{diag}(\sigma_1^{(i)}, \ldots, \sigma_n^{(i)})$$

and

$$\sigma_k^{(i)} := \begin{cases} 1 & \text{if} \quad k \leqq i \\ -1 & \text{if} \quad k > i, \end{cases}$$

$$\varkappa_i := \frac{\mu_i - \mu_{i+1}}{2} \quad \text{for} \quad i < n,$$

$$\varkappa_n := \frac{\mu_1 + \mu_n}{2}.$$

Clearly $\varkappa_i \geqq 0$ holds and $\sum_{i=1}^{n} \varkappa_i = \mu_1 = \operatorname{lub}(D)$. Now, by (11), a decomposition of an arbitrary matrix $A$ is

$$A = U_1 D V_1^H = \sum_{i=1}^{n} \varkappa_i U_1 E_i V_1^H = \sum_{i=1}^{n} \varkappa_i U_1 V_1^H (I - 2 \cdot V_1 P_i V_1^H),$$

where $P_i$ is the projector defined by

$$P_i := \frac{I - E_i}{2}.$$

## References

[1] BAUER, F. L.: On the Field of Values subordinate to a Norm. Num. Math. **4**, 103—113 (1962).
[2] — J. STOER, and C. WITZGALL: Absolute and Monotonic Norms. Num. Math. **3**, 257—264 (1961).
[3] BONNESEN, T., u. W. FENCHEL: Theorie der konvexen Körper. Berlin: Springer 1934.
[4] EGGLESTON, H. G.: Convexity. Cambridge: Cambridge University Press 1958.
[5] GASTINEL, N.: Thèses présentées à la faculté des sciences de l'université de Grenoble (1960): Matrices du second degré et normes générales en analyse numérique linéaire.
[6] SCHNEIDER, H., and W. G. STRANG: Comparison Theorems for Supremum Norms. Num. Math. **4**, 15—20 (1962).

Mathematisches Institut der TH München
8 München 2, Arcisstraße 21

*(Received April 27, 1964)*