Numerische Mathematik 6, 355—365 (1964)

The Bauer Fields of Values of a Matrix*

By

NICHOLAS NIRSCHL and HANS SCHNEIDER**

1. Introduction

Let V be a finite dimensional unitary vector space over the real or complex number field with inner product (y, x). If A is a linear transformation of V into itself the classical field of values $F[A]$ of the transformation A is the set of all complex numbers (x, Ax), with $(x, x) = 1$. Here $F[A]$ contains $P[A]$, the convex hull of the eigenvalues of A (TOEPLITZ [10]), and is a convex set (HAUSDORFF [6]).

In 1952 W. GIVENS [4] showed that $P[A]$ is the intersection of certain transforms of the field of values. More precisely, setting

$$F_H[A] = \{(x, HAx) : (x, Hx) = 1\},$$

for a positive definite hermitian H, he proved the following:

Theorem A. $P[A] = \bigcap_H F_H[A]$, where the intersection is taken over all positive definite hermitian H.

Theorem B. There exists a positive definite hermitian H such that $F_H[A] = P[A]$ if and only if all the eigenvalues of A lying on the boundary of $P[A]$ have simple elementary divisors.

Since the quadratic forms (x, x) and (x, Hx) are really the squares of vector norms, $F[A]$ and $F_H[A]$ depend not only on the transformation A but also on these norms. Recently F. L. BAUER [2], generalized the concept of field of values for norms not necessarily related to an inner product. Independently G. LUMER [9] has made a closely related generalization for operators on a general Banach space. It is the purpose of this paper to extend to Bauer fields of values the results of GIVENS.

2. Bauer Fields of Values

A vector norm $\|\cdot\|$ is said to be *weakly homogeneous* if $\|\gamma x\| = \gamma \|x\|$ is required only of scalars $\gamma \geqq 0$, and *strictly homogeneous* if $\|\gamma x\| = |\gamma|\, \|x\|$ is required of all complex scalars γ. In this paper we shall be interested only in strictly homogeneous vector norms.

If $\|\cdot\|$ is any vector norm the dual vector norm $\|\cdot\|'$ is defined by

$$\|y\|' = \sup_{x \neq 0} |(y, x)|. \tag{1}$$

Then we immediately have the inequality

$$|(y, x)| \leqq \|y\|' \, \|x\|, \quad \text{all} \quad x, y \in V. \tag{2}$$

* To ALSTON S. HOUSEHOLDER on his sixtieth birthday.

** This research was partly supported by the National Science Foundation, under grant NSF GP-2273 and by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-11-022-ORD-2059.

Since in a finite dimensional vector space the unit sphere is compact, for every non-zero $x \in V$ there exists $y \in V$ such that

$$(y, x) = \|y\|' \|x\| = 1. \tag{3}$$

Definition. If x and y satisfy (3) we say they are *dual vectors* and write $y \| x$.

We should point out that the definition of a dual norm usually used by HOUSEHOLDER [8], BAUER [2], and others would have $|(y, x)|$ replaced by $\mathrm{Re}(y, x)$ in (1). The latter definition is natural when n-dimensional complex space is considered as $2n$-dimensional real space. Also, BAUER'S definition of dual vectors would have (y, x) replaced by $\mathrm{Re}(y, x)$ in (3). He calls vectors satisfying (3) strictly dual vectors. However, since in the case of a strictly homogeneous vector norm (as he points out in [2]), dual vectors are also strictly dual, we shall not make the distinction.

Remark. As the notation $y \| x$ suggests, the duality relation generalizes the concept of parallel vectors. We give a geometric interpretation for real spaces, which is well known, but which is included because later in paragraph 4 we have an application to smooth bodies. Let K be the convex body belonging to the norm $\|\cdot\|$, K' the convex body belonging to the dual norm, and let x be on the boundary of K. If S is any support plane of K at x, let y be the vector on the boundary of K' orthogonal to S (i.e. $(y, u-x) = 0$ if $u \in S$), on the same side of the origin as S. Then $y \| x$. For proof simply note that $(y, z) \leqq (y, x)$ for all $z \in K$, and $(y, z) = 1$ for some $z \in K$. Hence $(y, x) = 1$. It is easy to see that every y dual to x is given by this construction, and of course x is also orthogonal to the support plane S' of K' at y.

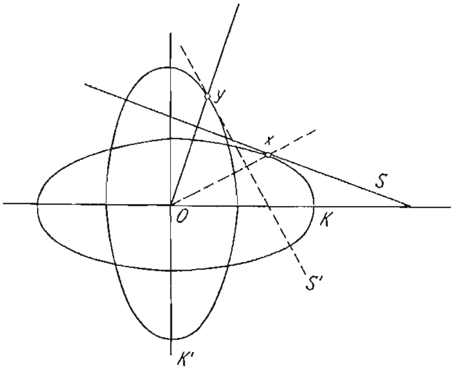

Fig. 1

Suppose $y \| x$. If we set $y_0 = \frac{1}{\|y\|'} y$ and $x_0 = \frac{1}{\|x\|} x$ we still have $y_0 \| x_0$, with y_0 and x_0 boundary vectors of K' and K. Since our principal concern will be with complex numbers of the form (y, Ax) and since $(y_0, Ax) = \frac{(y, Ax)}{\|y\|' \|x\|} = (y, Ax)$, we may restrict ourselves to dual pairs lying on the boundaries of K and K' whenever this is convenient. In this case $\|y\|' = \|x\| = 1$.

Definition. The *Bauer field of values* of a linear transformation A with respect to a norm $\|\cdot\|$ defined by

$$F[A, \|\cdot\|] = \{(y, Ax) : y \| x\}.$$

Remark. It is interesting to note that $F[A, \|\cdot\|]$ depends only on the norm $\|\cdot\|$ and the matrix A, and not on the inner product defined on V. For suppose $\pi(y, x)$ is any other unitary inner product defined on V. Then it is well known that $\pi(y, x) = (y, Hx) = (Hy, x)$ for some positive definite hermitian H. Denote by $\|\cdot\|^\pi$ the norm dual to $\|\cdot\|$ with respect to $\pi(y, x)$ and by $y \|^\pi x$ a pair of vectors dual with respect to $\pi(y, x)$. Since

$$\|y\|^\pi = \sup_{x \neq 0} \frac{|\pi(y, x)|}{\|x\|} = \sup_{x \neq 0} \frac{|(y, Hx)|}{\|x\|} = \sup_{x \neq 0} \frac{|(Hy, x)|}{\|x\|} = \|Hy\|',$$

it follows that $y \|^{\pi} x$ if and only if $H y \| x$. Let D be the set of all vector pairs $y \neq 0$, $x \neq 0$ such that $y \| x$, and let D_π be the set of all vector pairs $y \neq 0$, $x \neq 0$ such that $y \|^{\pi} x$. We have

$$\bigcup_{D_\pi} \frac{\pi(y, A x)}{\|y\|^{\pi} \|x\|} = \bigcup_{D_\pi} \frac{(H y, A x)}{\|H y\|' \|x\|} = \bigcup_{D} \frac{(y, A x)}{\|y\|' \|x\|}.$$

From this it now follows that $F[A, \|\cdot\|]$ is independent of the unitary inner product defined on V.

It will be convenient for us throughout the rest of this paper to assume that V is an n-tuple space with inner product $(y, x) = \sum_{i=1}^{n} \bar{y}_i x_i$. Thus once and for all we pick out a preferred basis $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0), \ldots$, $e_n = (0, \ldots, 0, 1)$ and a preferred inner product. Instead of linear transformations we may consider matrices.

As BAUER pointed out in [2], $F[A, \|\cdot\|]$ is not necessarily convex. We give an example when it is not. Let

$$A = \begin{bmatrix} 1+i, & -1 \\ 1 \quad, & 1-i \end{bmatrix},$$

and let us choose the norm $\|x\|_\infty = \max(|x_1|, |x_2|)$, with dual norm $\|y\|_1 = |y_1| + |y_2|$ (e.g. HOUSEHOLDER [8]). It is easy to show (cf. Theorem 6), that $F[A, \|\cdot\|_\infty]$ contains i and $-i$ but no other point on the imaginary axis.

However, $F[A, \|\cdot\|]$ does always contain the eigenvalues of A. To show this, let λ be any eigenvalue of A, and let x be any right eigenvector corresponding to it. We can choose y so that $y \| x$ and write $(y, A x) = (y, \lambda x) = \lambda (y, x) = \lambda$. Hence if we denote the convex closure of $F[A, \|\cdot\|]$ by $F^c[A, \|\cdot\|]$ we can say that $P[A] \subseteq F^c[A, \|\cdot\|]$.

3. The Givens Field of Values as a Special Case of the Bauer Field of Values

In [7] HOUSEHOLDER pointed out that if $\|\cdot\|$ is any vector norm and G is any non-singular matrix, then

$$\|x\|_G = \|G x\| \tag{4}$$

is also a vector norm called the *G-transform* of $\|\cdot\|$. He pointed out further that if $\|\cdot\|'$ is dual to $\|\cdot\|$ then

$$\|y\|'_G = \|G^{-1*} y\|' \tag{5}$$

is dual to $\|\cdot\|_G$. This follows from the definition of a dual (1) and the fact that for non-singular G the sup over all $x \neq 0$ is the same as the sup over all $G x \neq 0$. For

$$\|y\|'_G = \sup_{x \neq 0} \frac{|(y, x)|}{\|x\|_G} = \sup_{x \neq 0} \frac{|(y, x)|}{\|G x\|} = \sup_{x \neq 0} \frac{|(y, G^{-1} G x)|}{\|G x\|}$$

$$= \sup_{x \neq 0} \frac{|(G^{-1*} y, G x)|}{\|G x\|} = \sup_{x \neq 0} \frac{|(G^{-1*} y, x)|}{\|x\|} = \|G^{-1*} y\|'.$$

We shall need the following lemmas for the proof of theorem 1.

Lemma 1. The relation $y\|x$ with respect to $\|\cdot\|_G$ is equivalent to $G^{-1*}y\|Gx$ with respect to $\|\cdot\|$.

Proof. We have $y\|x$ with respect to $\|\cdot\|_G$ if and only if $(y, x) = \|y\|'_G \|x\|_G = 1$. But from (4) and (5) this is equivalent to $(G^{-1*}y, Gx) = (y, G^{-1}Gx) = \|G^{-1*}y\|' \|Gx\| = 1$, which is necessary and sufficient for $G^{-1*}y\|Gx$ with respect to $\|\cdot\|$.

Lemma 2. If $C = GAG^{-1}$ then $F[A, \|\cdot\|_G] = F[C, \|\cdot\|]$.

Proof. Let $\varrho \in F[A, \|\cdot\|_G]$. Then $\varrho = (y, Ax)$, where $y\|x$ with respect to $\|\cdot\|_G$. Let $\xi = Gx$ and $\eta = G^{-1*}y$. By lemma 1, $\eta\|\xi$ with respect to $\|\cdot\|$. Hence

$$\varrho = (y, G^{-1}GAG^{-1}Gx) = (\eta, GAG^{-1}\xi) = (\eta, C\xi) \in F[C, \|\cdot\|],$$

so that $F[A, \|\cdot\|_G] \subseteq F[C, \|\cdot\|]$. The reverse inclusion follows in the same way.

Now we recall two well known facts. First, the Euclidean vector norm $\|x\| = (x, x)^{\frac{1}{2}}$ is self dual, that is, $\|\cdot\| = \|\cdot\|'$. In this case $y\|x$ if and only if $y = \frac{x}{\|x\|^2}$. Next, a hermitian matrix H is positive definite if and only if there exists a non-singular matrix G such that $H = G^*G$.

Theorem 1. Let $\|\cdot\|$ be the Euclidean vector norm and $H = G^*G$ be any positive definite hermitian matrix. Then $F_H[A] = F[A, \|\cdot\|_G]$.

Proof. By definition $F_H[A] = \bigcup_{x \neq 0} \frac{(x, HAx)}{(x, Hx)}$. Set $\xi = Gx$. Since G is non-singular, x ranging over all non-zero vectors is the same as ξ ranging over all non-zero vectors. Hence using Lemma 2 and the fact that the Euclidean norm is self dual we can write

$$\begin{aligned} F[A, \|\cdot\|_G] &= F[GAG^{-1}, \|\cdot\|] = \bigcup_D \frac{(y, GAG^{-1}x)}{\|y\|' \|x\|} \\ &= \bigcup_{x \neq 0} \frac{(x, GAG^{-1}x)}{\|x\|^2} = \bigcup_{\xi \neq 0} \frac{(\xi, GAG^{-1}\xi)}{\|\xi\|^2} = \bigcup_{x \neq 0} \frac{(Gx, GAG^{-1}Gx)}{\|Gx\|^2} \\ &= \bigcup_{x \neq 0} \frac{(x, G^*GAx)}{(Gx, Gx)} = \bigcup_{x \neq 0} \frac{(x, HAx)}{(x, Hx)} = F_H[A]. \end{aligned}$$

It follows from this theorem that the Givens field of values is a special case of the Bauer field of values.

4. Absolute Norms and *B*-Norms

A vector norm $\|\cdot\|$ is said to be *absolute* if it depends only on the moduli of the components of the vector, that is, if $\|x\| = \|x_a\|$ for all x, where $x_a = (|x_1|, |x_2|, \ldots, |x_n|)$. It is clear that all Hölder norms

$$(6) \qquad \begin{cases} \|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p}, & 1 \leqq p < \infty \\ \|x\|_\infty = \max_i |x_i| \end{cases}$$

are absolute. It is also clear that all absolute vector norms are strictly homogeneous. Geometrically an absolute norm is one whose associated convex body is symmetric about the coordinate axes.

A vector norm $\|\cdot\|$ is said to be *monotonic* if for any two vectors $x=(x_1, x_2, \dots, x_n)$ and $z=(z_1, z_2, \dots, z_n)$, with $|x_i| \leqq |z_i|$ for all i, the inequality $\|x\| \leqq \|z\|$ holds.

These concepts are due to BAUER, STOER, and WITZGALL [1]. In that paper they showed, among other things:

(7) an absolute vector norm is monotonic and vice versa;

(8) the dual of an absolute vector norm is again absolute.

The following two lemmas exhibit important geometric properties of absolute vector norms.

Lemma 3. If $y\|x$ with respect to an absolute vector norm $\|\cdot\|$, then for every component $\bar{y}_i x_i \geqq 0$.

Proof. Suppose not. Then we have $1=(y, x)=\sum_i \bar{y}_i x_i = \sum_i \delta_i e^{i\vartheta_i}$, with $\delta_i \geqq 0$ and $\vartheta_i \neq 0$ for at least two indices i for which $\delta_i \neq 0$. It follows that $\sum_i \delta_i > 1$. Call $D=\operatorname{diag}(e^{i\vartheta_1}, e^{i\vartheta_2}, \dots, e^{i\vartheta_n})$, and $z=Dy$. By the inequality (2) we have

$$\|z\|' \, \|x\| \geqq |(z, x)| = \sum_i \delta_i > 1 = \|y\|' \, \|x\|. \tag{9}$$

But by (8), $\|z\|'=\|y\|'$, and hence the strict inequality in (9) yields a contradiction.

Lemma 4. Let $\|\cdot\|$ be any absolute vector norm and $y\|x$ with respect to $\|\cdot\|$. If for some component i there exists a vector z such that $|z_i| > |y_i|$, $|z_k| = |y_k|$ $k \neq i$, and $\|z\|'=\|y\|'$ then $x_i=0$.

Proof. Since $\|\cdot\|$ is absolute it is no restriction to assume that the arguments of the components of z are such that $\bar{z}_k x_k \geqq 0$ for all k. By lemma 3, $\bar{y}_i x_i \geqq 0$. Hence if $x_i \neq 0$ we have $\bar{z}_i x_i > \bar{y}_i x_i > 0$. But then $1=(y, x)<(z, x) \leqq \|z\|' \, \|x\| = \|y\|' \, \|x\|$, contradicting the fact that $y\|x$. Therefore $x_i=0$ and the lemma is proved.

A *coordinate* subspace V' of V is a subspace spanned by a subset of the set of n unit vectors $e_1=(1, 0, \dots, 0), \dots, e_n=(0, \dots, 0, 1)$. Clearly each coordinate subspace V' has a unique complementary coordinate subspace V''. If $x \in V$, we shall decompose $x=x' \oplus x''$, where $x' \in V'$ and $x'' \in V''$.

Definition. A vector pair y', x' is defined to be *embeddable* with respect to a vector norm $\|\cdot\|$ if y' and x' are projections into some coordinate subspace of a pair of vectors dual with respect to $\|\cdot\|$.

More precisely, y', x' is embeddable with respect to $\|\cdot\|$ if there exists a pair y'', x'' such that $y' \oplus y'' \| x' \oplus x''$ with respect to $\|\cdot\|$. It is clear that the pair y'', x'' would also be embeddable. For a given vector norm we denote by $E(\|\cdot\|)$ the set of all embeddable vector pairs. Note that for $y', x' \in E(\|\cdot\|)$, $(y', x') \geqq 0$ by lemma 3.

Definition. A *B-norm* is defined to be any absolute vector norm with the property

(B) there exists a bound M such that for all $y', x' \in E(\|\cdot\|)$ we have $\|y'\|' \, \|x'\| \leqq M(y', x')$.

Intuitively, $y\|x$ if y is "parallel" to x. Thus in the case of a B-norm, the projections of "parallel" vectors on coordinate subspaces are bounded away from orthogonality. It should be noted that the restriction to coordinate subspaces is essential here. For if $\dim V \geqq 3$, then an easy continuity argument shows that given any two non-zero vectors x, y with $x \neq y$, there exists a pair of orthogonal subspaces V', V'' with $\dim V' = 2$ such that the projections x', y' of x, y on V' along V'' are non-zero and orthogonal.

The following lemma is crucial in the proof of theorem 5 and exhibits an essential property of B-norms.

Lemma 5. Let $\|\cdot\|$ be any B-norm. If $y', x' \in E(\|\cdot\|)$ and $(y', x') = 0$, then either $y' = 0$ or $x' = 0$.

Proof. Suppose not. Then both $\|y'\|' > 0$ and $\|x'\| > 0$. Hence $0 < \|y'\|' \|x'\| \leqq M(y', x') = 0$, a contradiction.

Remark. Suppose the norm bodies K and K' are smooth, i.e., at each boundary point there is a unique support plane. (If V is complex n-space we treat V as $2n$-dimensional real space.) It follows from a remark in paragraph 1 that the support plane at x is defined by $(y, x) = 1$, where y is the unique y such that $y \| x$. A theorem (BUSEMANN [3], p. 6) asserts that y depends continuously on x. In this case we can replace (B) by an apparently weaker property:

(B′) there exists a bound M such that for all $y', x' \in E(\|\cdot\|)$ with $(y', x') > 0$, we have $\|y'\|' \|x'\| \leqq M(y', x')$.

To show that under our assumptions on K and K', (B′) implies (B), let $y', x' \in E(\|\cdot\|)$ with $(y', x') = 0$. We must show that $\|y'\|' \|x'\| = 0$. If $y' = 0$ there is nothing to prove. So suppose $y' \neq 0$. There exists i such that $x'_i = 0$ and $y'_i \neq 0$. We can find $x(t)$ depending continuously on t such that $x(0) = x$, the vector of which x' is a projection, and such that $x(t)_i \neq 0$ if $t \neq 0$. Let $y(t) \| x(t)$. For sufficiently small t, $y(t)_i \neq 0$ by continuity. Hence for $t \neq 0$, $(y(t)', x(t)') > 0$. Thus by property (B′) $\|y(t)'\|' \|x(t)'\| \leqq M(y(t)', x(t)')$. Hence by continuity at $t = 0$, $\|y'\|' \|x'\| = 0$.

We now show that all Hölder norms (6) are B-norms, and in fact have bound $M = 1$. It is well known that the norm dual to the Hölder norm $\|x\|_1$ is the norm $\|y\|_\infty = \max_i |y_i|$, and for $p > 1$ the norm dual to $\|x\|_p$ is $\|y\|_q$, where $1/p + 1/q = 1$.

Theorem 2. Let $\|\cdot\|_p$ be any Hölder norm. If $y', x' \in E(\|\cdot\|_p)$ then $\|y'\|_q \|x'\|_p = (y', x')$.

Proof. Denote by I the subset of components defining the coordinate subspace containing y' and x'. Let $y \| x$ with respect to $\|\cdot\|_p$ and suppose, as we may, $\|y\|_q = \|x\|_p = 1$.

(i) Let $p > 1$.

By the definition of duality (3) we have

$$\sum_{i=1}^{n} \bar{y}_i x_i = \Big(\sum_{i=1}^{n} |y_i|^q\Big)^{1/q} \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}.$$

By the Hölder inequality for complex numbers (see for example, HARDY, POLYA, and LITTLEWOOD [5], p. 26) this implies that $|y_i|^q$ and $|x_i|^p$ are proportional. From $\|y\|^q = \|x\|^p = 1$ it easily follows that

$$|y_i|^q = |x_i|^p. \tag{10}$$

Since $\frac{q}{p}+1=q$, and since by lemma 3, $\bar{y}_i\, x_i \geqq 0$ we can write

(11) $$\bar{y}_i\, x_i = |\bar{y}_i\, x_i| = |y_i|\,|x_i| = |y_i|\,|y_i|^{q/p} = |y_i|^q.$$

From (10) and (11) we obtain

$$(y', x') = \sum_{i\in I} \bar{y}_i\, x_i = \Big(\sum_{i\in I} |y_i|^q\Big)^{1/q} \Big(\sum_{i\in I} |x_i|^p\Big)^{1/p} = \|y'\|'\, \|x'\|,$$

and the lemma is proved for $p>1$.

(ii) Let $p=1$ and $q=\infty$.

First we note that if for some i, $|y_i|<1$, then $x_i=0$. For proof, note that we can define a vector z to be the same as y except that $|z_i|=1$. Then since $\|z\|_\infty = \|y\|_\infty = 1$ the hypothesis for lemma 4 is satisfied and $x_i=0$.

Thus for any $y', x' \in E(\|\cdot\|_1)$, for all $i\in I$ such that $x_i \neq 0$ we have $|y_i|=1$. Because of this and the fact that by lemma 3 $\bar{y}_i\, x_i \geqq 0$ we can write

$$\|y'\|_\infty \|x'\|_1 = \big(\max_{i\in I} |y_i|\big)\Big(\sum_{i\in I} |x_i|\Big) = \sum_{i\in I} |y_i|\,|x_i| = \sum_{i\in I} \bar{y}_i\, x_i = (y', x').$$

The theorem is proved.

To show that the class of B-norms is a proper sub-class of the class of absolute norms we give an example of an absolute norm which is not a B-norm. Let V be real 3-space and let $\|x\| = \max(|x_1|+|x_2|, |x_3|)$. Then the dual norm is $\|y\|' = (\max(|y_1|, |y_2|)) + |y_3|$. Choose $x=(1, 0, 1)$ and $y=(1, 1, 0)$. Then $y\|x$. If $x'=(0, x_2, x_3)$ then $(y', x')=0$ but $\|y'\|'\, \|x'\|=1$.

We point out that the convex body associated with the absolute norm in the previous example is not smooth.

5. Extension of Givens' Results to Bauer Fields of Values

Denote by $P[A]+\Gamma_\varepsilon$ the set of all points of the complex plane whose distance from some point of $P[A]$ is no greater than ε. The following theorem extends theorem A to all absolute vector norms.

Theorem 3. Let $\|\cdot\|$ be any absolute vector norm. Then $P[A] = \bigcap_G F^c[A, \|\cdot\|_G]$, the intersection being taken over all non-singular G.

Proof. If we show that for any $\varepsilon>0$ there exists a non-singular G such that $F[A, \|\cdot\|_G] \subseteq P[A]+\Gamma_\varepsilon$ the theorem will then follow.

Let Q be the non-singular matrix such that $Q A Q^{-1} = \Lambda + U$, the Jordan canonical form of A. Here Λ is a diagonal matrix whose diagonal elements are the eigenvalues λ_i of A, and U is a matrix with 1's or 0's on the first super diagonal and 0's elsewhere. Since the norm bodies K and K' are bounded, there exists a positive number b such that for all x, y satisfying $y\|x$ with respect to $\|\cdot\|$ we have $|\bar{y}_i\, x_{i+1}| \leqq b$, $i=1, 2, \ldots, n-1$. Define $D=\operatorname{diag}(d_1, d_2, \ldots, d_n)$ such that $d_i>0$ for all i and $\frac{d_i}{d_{i+1}} \leqq \frac{\varepsilon}{(n-1)\,b}$ for $i=1, 2, \ldots, n-1$. Set $G=DQ$.

By lemma 2, $F[A, \|\cdot\|_G] = F[D Q A Q^{-1} D^{-1}, \|\cdot\|] = F[D(\Lambda+U)D^{-1}, \|\cdot\|]$. For any $\varrho \in F[D(\Lambda+U)D^{-1}, \|\cdot\|]$ we can write $\varrho = (y, D(\Lambda+U)D^{-1}x) = (y, \Lambda x) + (y, DUD^{-1}x) = \sum_{i=1}^{n} \bar{y}_i\, x_i\, \lambda_i + \sum_{i=1}^{n-1} \bar{y}_i\, x_{i+1} \frac{d_i}{d_{i+1}}$. By lemma 3 and the definition of duality (3) it follows that $\sum_{i=1}^{n} \bar{y}_i x_i \lambda_i$ is a convex combination of the eigenvalues of A and hence an element of $P[A]$. And from the definition of D it follows that

$$\left|\sum_{i=1}^{n-1} \bar{y}_i x_{i+1} \frac{d_i}{d_{i+1}}\right| \leqq b \sum_{i=1}^{n-1} \frac{d_i}{d_{i+1}} \leqq \frac{(n-1)\, b\, \varepsilon}{(n-1)\, b} = \varepsilon .$$

Hence $\varrho \in P[A] + \Gamma_\varepsilon'$. This completes the proof.

Although the necessity part of theorem B holds for all strictly homogeneous vector norms, we have been able to extend the sufficiency part only to B-norms. We show these results in the following two theorems.

Theorem 4. Let $\|\cdot\|$ be any strictly homogeneous vector norm. If $F^c[A, \|\cdot\|] = P[A]$, then the elementary divisors corresponding to all eigenvalues lying on the boundary of $P[A]$ must be simple.

Proof. Let λ be an eigenvalue on the boundary of $P[A]$ to which there corresponds a non-simple elementary divisor. We shall construct $\varrho \in F^c[A, \|\cdot\|]$ such that $\varrho \notin P[A]$. It follows from the structure of the Jordan form of A that there exist two linearly independent vectors z' and z'' such that

$$A z'' = \lambda z'' \quad \text{and} \quad A z' = \lambda z' + z''. \tag{12}$$

Let ϑ be a direction normal to $P[A]$ at λ, i.e., perpendicular to a support line of $P[A]$. Set $z = \delta e^{i\vartheta} z' + z''$. As $\delta \to 0$, $z \to z''$, and since the norm is a continuous function $\|z\| \to \|z''\|$. We choose $\delta > 0$ small enough so that

$$\|z''\| \leqq 2\|z\|, \quad \text{and} \quad \delta < \frac{\|z''\|}{\|z'\|}. \tag{13}$$

It follows from (12) that

$$A z = \lambda z + \delta e^{i\vartheta} z''. \tag{14}$$

Set $x = \frac{1}{\|z\|} z$, and let $y \| x$. Then $\|y\|' = 1$ since $\|x\| = 1$. We have

$$1 = (y, x) = \frac{1}{\|z\|}(y, z) = \frac{1}{\|z\|}(y, \delta e^{i\vartheta} z' + z'')$$
$$= \frac{\delta e^{i\vartheta}}{\|z\|}(y, z') + \frac{1}{\|z\|}(y, z''), \quad \text{or} \quad \frac{1}{\|z\|}(y, z'') = 1 - \frac{\delta e^{i\vartheta}}{\|z\|}(y, z').$$

From (2), (13), and the fact that $\|y\|' = 1$ it follows that

$$\left|\frac{\delta e^{i\vartheta}(y, z')}{\|z\|}\right| = \frac{\delta}{\|z\|}|(y, z')| \leqq \delta \frac{\|z'\|}{\|z\|} \leqq \frac{2\delta\|z'\|}{\|z''\|} < \frac{2\|z''\|\,\|z'\|}{2\|z'\|\,\|z''\|} = 1 .$$

Hence we can write

$$\frac{1}{\|z\|}(y, z'') = 1 - \varepsilon e^{i\varphi}, \qquad 0 \leqq \varepsilon < 1, \qquad 0 \leqq \varphi < 2\pi. \tag{15}$$

Now $\varrho = (y, A x) \in F^c[A, \|\cdot\|]$. But by (14) and (15)

$$\varrho = \frac{1}{\|z\|}(y, A z) = \frac{1}{\|z\|}(y, \lambda z + \delta e^{i\vartheta} z'') = \lambda + \frac{\delta e^{i\vartheta}}{\|z\|}(y, z'')$$
$$= \lambda + \delta e^{i\vartheta}(1 - \varepsilon e^{i\varphi}) = \lambda + \delta e^{i\vartheta} - \delta \varepsilon e^{i(\vartheta + \varphi)}.$$

But since $\delta\varepsilon<\delta$ and since ϑ is a direction normal to $P[A]$ it follows that ϱ must lie outside $P[A]$.

Theorem 5. Let $\|\cdot\|$ be any B-norm. If A is any matrix whose eigenvalues lying on the boundary of $P[A]$ all have simple elementary divisors, then there exists a non-singular matrix G such that $F^c[A, \|\cdot\|_G]=P[A]$.

Proof. It is well known that by means of a similarity transformation any matrix can be put into a form like the Jordan canonical form, but with the elements on the super diagonal arbitrarily small. That is, there exists a non-singular matrix G such that

$$GAG^{-1}=D_0\oplus\sum_r\oplus\,(D_r+U_r). \tag{16}$$

Here D_0 is a diagonal matrix made up of the eigenvalues of A with simple elementary divisors, $D_r=\lambda_r I_r$, where λ_r is an eigenvalue of A with non-simple elementary divisors, and U_r is a matrix with non-zero elements only in the first super diagonal.

Let $\|U_r\|=\sup\limits_{x\neq 0}\dfrac{\|U_r x\|}{\|x\|}$. Call δ the minimum distance from the interior eigenvalues to the boundary of $P[A]$. Let M be the bound associated with $\|\cdot\|$ by property (B). Since G can be chosen so that the elements of the matrices U_r are arbitrarily small, and since the norm of a matrix is a continuous function of its elements, we can choose G so that $\|U_r\|\leqq\dfrac{\delta}{M}$ for all r. Set $C=GAG^{-1}$.

Now let ϱ be any element of $F[A,\|\cdot\|_G]=F[C,\|\cdot\|]$. We partition the vector space V according to the orders of the direct sum in (16). Counting multiplicities, suppose A has k eigenvalues with simple elementary divisors, and m eigenvalues with non-simple elementary divisors. We can write

$$\varrho=\left(y,\left[D_0\oplus\sum_{r=1}^{m}\oplus\,(D_r+U_r)\right]x\right)=\sum_{i=1}^{k}\bar{y}_i x_i\lambda_i+\sum_{r=1}^{m}\left(y_r\left[D_r+U_r\right]x_r\right). \tag{17}$$

The y_i and x_i of the first summation are components of y and x, and the y_r and x_r of the second summation are projections of y and x into the coordinate subspaces of V defined by the partition. Hence $y_r, x_r\in E(\|\cdot\|)$. Each of the summands of the second summation of (17) can be written

$$\left(y_r, [D_r+U_r]\,x_r\right)=(y_r, x_r)\,\lambda_r+(y_r, U_r x_r). \tag{18}$$

If for some r, $(y_r, x_r)=0$, then by lemma 5 either $y_r=0$ or $x_r=0$, and hence $(y_r, U_r x_r)=0$.

For all r such that $(y_r, x_r)\neq 0$ we can express (18) as

$$\left(y_r, [D_r+U_r]\,x_r\right)=(y_r, x_r)\,(\lambda_r+\varrho_r),\quad\text{where}\quad \varrho_r=\frac{(y_r, U_r x_r)}{(y_r, x_r)}. \tag{19}$$

But from (2), the definition of $\|U_r\|$, and property (B) we have

$$|\varrho_r|=\frac{|(y_r, U_r x_r)|}{(y_r, x_r)}\leqq\frac{\|y_r\|'\,\|U_r x_r\|}{(y_r, x_r)}\leqq\frac{\|y_r\|'\,\|x_r\|\,\|U_r\|}{(y_r, x_r)}\leqq M\|U_r\|\leqq\delta.$$

Hence $\lambda_r+\varrho_r\in P[A]$.

Since $\sum_{i=1}^{k} \bar{y}_i x_i + \sum_{r=1}^{m} (y_r, x_r) = (y, x) = 1$ it follows from (17) and (19) that ϱ is a convex combination of elements of $P[A]$. Hence $F[A, \|\cdot\|_G] \subseteq P[A]$. And since $P[A] \subseteq F^c[A, \|\cdot\|_G]$ it follows that $F^c[A, \|\cdot\|_G] = P[A]$.

6. The Gerschgorin Circles: Important Special Cases

For the dual Hölder norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ the Bauer fields of values become important special cases. For any matrix $A = (a_{ij})$ let $R_i = \sum_{\substack{j \neq i \\ j=1}}^{n} |a_{ij}|$ and $C_i = \sum_{\substack{j \neq i \\ j=1}}^{n} |a_{ji}|$. Let $H_i = \{z\colon |a_{ii} - z| \leqq R_i\}$, $H'_i = \{z\colon |a_{ii} - z| \leqq C_i\}$ and define $H = \bigcup_{i=1}^{n} H_i$ and $H' = \bigcup_{i=1}^{n} H'_i$. The circular regions H_i and H'_i are respectively the Gerschgorin row circles and column circles of the matrix A. It is well known that the convex hulls H^c and H'^c both contain $P[A]$.

Theorem 6. $F^c[A, \|\cdot\|_\infty] = H^c$ and $F^c[A, \|\cdot\|_1] = H'^c$.

Proof. We shall prove the first statement of the theorem; the proof of the second follows in an analogous manner. Let $\varrho \in F[A, \|\cdot\|_\infty]$. Then

$$\varrho = \sum_{i,j} \bar{y}_i a_{ij} x_j = \sum_i \Big(\bar{y}_i x_i a_{ii} + \bar{y}_i \sum_{j \neq i} a_{ij} x_j\Big).$$

Just as in part (ii) of the proof of theorem 2 it follows from lemma 4 that if $\bar{y}_i \neq 0$ then $|x_i| = 1$. Hence we can write

$$\varrho = \sum_i \Big[\bar{y}_i x_i \Big(a_{ii} + \sum_{j \neq i} a_{ij} \frac{x_j}{x_i}\Big)\Big] = \sum_i \bar{y}_i x_i (a_{ii} + \varrho_i),$$

where $|\varrho_i| \leqq R_i$. It follows that $a_{ii} + \varrho_i \in H_i$ and hence $\varrho \in H^c$. Since $F^c[A, \|\cdot\|_\infty]$ is the smallest convex set containing $F[A, \|\cdot\|_\infty]$ it follows that $F^c[A, \|\cdot\|_\infty] \subseteq H^c$.

Next let $\varrho \in H$. Then for some i, $\varrho = a_{ii} + r e^{i\vartheta}$, where $r \leqq R_i$. Suppose $a_{ij} = |a_{ij}| e^{i\varphi_j}$, $j \neq i$. Let x be the vector with $x_i = 1$ and $x_j = \frac{r}{R_i} e^{-i(\varphi_j - \vartheta)}$, $j \neq i$. Let y be the vector with $y_i = 1$ and $y_j = 0$, $j \neq i$. Then $y \| x$ with respect to $\|\cdot\|_\infty$ and

$$(y, A x) = a_{ii} + \sum_{j \neq i} a_{ij} x_j = a_{ii} + \frac{r}{R_i} \sum_{j \neq i} |a_{ij}| e^{i\vartheta} = a_{ii} + r e^{i\vartheta} = \varrho.$$

Hence $H \subseteq F[A, \|\cdot\|_\infty]$, which implies $H^c \subseteq F^c[A, \|\cdot\|_\infty]$. The theorem is proved.

References

[1] BAUER, F. L., J. STOER, and C. WITZGALL: Absolute and Monotonic Norms. Numer. Math. 3, 257–264 (1961).
[2] — On the Field of Values Subordinate to a Norm. Numer. Math. 4, 103–111 (1962).
[3] BUSEMANN, H.: Convex Surfaces. New York: Interscience Publishers 1958.
[4] GIVENS, W.: Fields of Values of a Matrix. Proc. Amer. Math. Soc. 3, 206–209 (1952).
[5] HARDY, G. E., J. E. LITTLEWOOD, and G. POLYA: Inequalities, 2nd ed. Cambridge: Cambridge University Press 1959.

[6] HAUSDORFF, F.: Der Wertevorrat einer Bilinearform. Math. Z. 3, 314–316 (1919).
[7] HOUSEHOLDER, A. S.: The Approximate Solution of Matrix Problems. J. Assoc. Comp. Mach. 5, 205–243 (1958).
[8] — The Theory of Matrices in Numerical Analysis. New York: Blaisdell Publishing Co. 1964.
[9] LUMER, G.: Semi-inner-product Spaces. Trans. Amer. Math. Soc. **100**, 29–43 (1961).
[10] TOEPLITZ, O.: Das Algebraische Analogon zu einem Satz von FEJER. Math. Z. **2**, 187–197 (1918).

Mathematics Department
University of California
Santa Barbara, California/USA
and
St. Norbert College
West De Pere, Wisconsin/USA

(Received April 13, 1964)

Druck der Universitätsdruckerei H. Stürtz AG., Würzburg