Numer. Math. 26, 99—105 (1976)

A Symmetric Numerical Range for Matrices

B. David Saunders and Hans Schneider*

Received October 2, 1975

Summary. For each norm ν on $\mathbb{C}^n$, we define a numerical range Z_ν, which is symmetric in the sense that $Z_\nu = Z_{\nu^D}$, where ν^D is the dual norm.

We prove that, for $a \in \mathbb{C}^{nn}$, $Z_\nu(a)$ contains the classical field of values $V(a)$. In the special case that ν is the l_1-norm, $Z_\nu(a)$ is contained in a set $G(a)$ of Gershgorin type defined by C. R. Johnson.

When a is in the complex linear span of both the Hermitians and the ν-Hermitians, then $Z_\nu(a)$, $V(a)$ and the convex hull of the usual ν-numerical range $V_\nu(a)$ all coincide. We prove some results concerning points of $V(a)$ which are extreme points of $Z_\nu(a)$.

1. Introduction

In this note we introduce a numerical range Z_ν for matrices, where ν is a norm on $\mathbb{C}^n$. We call this numerical range symmetric since $Z_\nu = Z_{\nu^D}$, where ν^D defined by $\nu^D(y) = \sup\{|y^* x|/\nu(x) : 0 \neq x \in \mathbb{C}^n\}$ is the dual norm of ν. For $a \in \mathbb{C}^{nn}$, we compare $Z_\nu(a)$ with numerical ranges already in the literature:

(1.1) The classical field of values (Hausdorff [8], Toeplitz [13]) is

$$V(a) = \left\{\frac{x^* a x}{x^* x} : 0 \neq x \in \mathbb{C}^n\right\}.$$

(1.2) The generalization of $V(a)$ introduced by Bauer [1] is

$$V_\nu(a) = \{y^* a x : (x, y) \in \Pi_\nu\},$$

where ν is a norm on $\mathbb{C}^n$ and

$$\Pi_\nu = \{(x, y) \in \mathbb{C}^n \times \mathbb{C}^n : \nu(x) = \nu^D(y) = y^* x = 1\}.$$

For variants of Π_ν see Lumer [10], Bonsall-Duncan [3], Bauer [2] and Deutsch-Zenger [5].

(1.3) The convex hull of the "mixed" Gershgorin circles introduced by C. R. Johnson [9] is

$$G(a) = \operatorname{conv} \bigcup_{j=1}^{n} G_j(a),$$

* Part of this research was done while the authors were at the Mathematische Institut, Technische Universität, München, West Germany. The first author presented these results at the Seminar on Matrix Theory (Positivity and Norms) held in Munich in December 1974. The second author also acknowledges support from the National Science Foundation under grant GP 37978X

where conv stands for convex hull and for $j=1, \ldots, n$

$$G_j(a)=\{\xi\in\mathbb{C}: |\xi-a_{jj}| \leqq \tfrac{1}{2}\sum_{\substack{i=1\\ i\neq j}}^{n} (|a_{ij}|+|a_{ji}|)\}.$$

Johnson proved that $V(a) \subseteq G(a)$, for all $a\in\mathbb{C}^{nn}$. We show that $V(a) \subseteq Z_\nu(a)$, for all norms ν. If ν is the l_1-norm (or the l_∞-norm) we show that $Z_\nu(a) \subseteq G(a)$. Further if a is Hermitian then, for all norms ν, $Z_\nu(a) \subseteq \mathbb{R}$. We show by an example that there exists $a\in\mathbb{C}^{nn}$ and a norm ν such that $Z_\nu(a) \subset V_\nu(a) \cap V_{\nu D}(a)$. Thus Z_ν is of numerical interest. In the special case that ν is the l_1-norm, our bounds for $V(a)$ are sharper though less easy to compute in general than Johnson's.

In [16], Zenger gave an axiomatic treatment of eigenvalue inclusion sets. This approach is appropriate here, since Z_ν, G, and conv V_ν share basic properties. Thus we begin Section 2 by listing axioms for convex numerical ranges, and then we prove the results outlined above. Section 3 is motivated by the result that for each norm ν there is a p-transform ν_p such that the ν_p-Hermitians are Hermitian, see Deutsch-Schneider [6, Proposition (4.1)]. Under the hypothesis that h and k are both ν-Hermitian and Hermitian, we show in Section 3 that $Z_\nu(h+i\,k)=\operatorname{conv} V_\nu(h+i\,k)=V(h+i\,k)$. Our theorem then leads us to investigate in Section 4 extreme points α of $Z_\nu(a)$ which also belong to $V(a)$. If α is such a point, and Π_ν is the set defined in (1.1), then we show that there exist $(x,y)\in\Pi_\nu$ such that

$$\alpha=x^*a\,x/x^*x=y^*a\,y/y^*y=\tfrac{1}{2}(y^*a\,x+x^*a\,y).$$

This results has an application to eigenvectors of norm-Hermitians.

The chief tool in these investigations is a result due to Cain-Saunders-Schneider [4, Theorem 6] from which it follows that for each $z\in\mathbb{C}^n$ there is a positive multiple of z which is of the form $x+y$, where $(x,y)\in\Pi_\nu$.

2. The Symmetric Numerical Range

(2.1) Definition. A set valued mapping W: $\mathbb{C}^{nn}\to\mathscr{P}(\mathbb{C})$ is called a (homogeneous, unital, compact) *convex numerical range* if

(i) $W(a)$ contains the spectrum of a, for all $a\in\mathbb{C}^{nn}$,

(ii) $W(a+b) \subseteq W(a)+W(b)$, for all $a, b\in\mathbb{C}^{nn}$,

(iii) $W(\lambda a)=\lambda W(a)$, for $\lambda\in\mathbb{C}$, $a\in\mathbb{C}^{nn}$,

(iv) $W(1)=\{1\}$,

(v) $W(a)$ is compact for $a\in\mathbb{C}^{nn}$,

(vi) $W(a)$ is convex for $a\in\mathbb{C}^{nn}$.

Remark. Conditions (ii), (iii) and (iv) of (2.1) imply that

$$W(0)=\{0\},$$

and

$$W(a+\lambda)=W(a)+\lambda, \quad \text{for } \lambda\in\mathbb{C},\ a\in\mathbb{C}^{nn}.$$

Examples of convex numerical ranges are V, G, and conv V_ν, as defined in (1.1)–(1.3).

(2.2) Definition. Let ν be a norm. For $a\in\mathbb{C}^{nn}$, let

$$Z_\nu(a)=\operatorname{conv}\{\tfrac{1}{2}(y^* a x+x^* a y):(x,y)\in\Pi_\nu\}.$$

It is easy to see that Z_ν satisfies (ii), (iii), and (iv) of Definition (2.1). The compactness of $Z_\nu(a)$ is a consequence of the compactness of Π_ν. Clearly $Z_\nu(a)$ is convex. That Z_ν satisfies (i) is a consequence of Theorem (2.4) below.

By taking $\delta=0$ in Theorem 6 of [4], we immediately obtain the following lemma.

(2.3) *Decomposition Lemma.* Let ν be a norm on $\mathbb{C}^n$, and let $z\in\mathbb{C}^n$. Then there exist unique $(x,y)\in\Pi_\nu$ and $t>0$ such that $tz=x+y$. □

We shall call (x,y) the *Π_ν-decomposition* of the direction z.

(2.4) Theorem. Let ν be a norm. Then, for all $a\in\mathbb{C}^{nn}$,

$$Z_\nu(a)\supseteq V(a).$$

Proof. Suppose first that Π^+ is the closed right-half plane and that $0\in V(a)\subseteq\Pi^+$. Let $z^*az=0$ where $z\in\mathbb{C}^n$ and $z^*z=1$, and suppose that (x,y) is the Π_ν-decomposition of the direction z. Thus, for some $t>0$, $tz=x+y$ and so

$$0=t^2 z^* a z=x^* a x+y^* a y+x^* a y+y^* a x.$$

Since $\operatorname{Re}(x^* a x+y^* a y)\geqq 0$ it follows that α defined by $\alpha=\frac{1}{2}(x^* ay+y^* a x)$ satisfied $\operatorname{Re}\alpha\leqq 0$. Since $\alpha\in Z_\nu(a)$, we have $Z_\nu(a)\cap\Pi^-\neq\phi$, where Π^- is the closed left-half plane. Let S be any supporting half-space of $Z_\nu(a)$ in the plane. There exists θ, $0\leqq\theta<2\pi$, such that $e^{i\theta}S$ is a translate of Π^+. If β is a point of $V(e^{i\theta}a)$ with minimal real part, then, for $b=e^{i\theta}a-\beta$, we have $0\in V(b)\subseteq\Pi^+$. The previous argument shows that $Z_\nu(b)$ contains a point of Π^- and since $e^{i\theta}S-\beta$ is a supporting half-space of $Z_\nu(b)$, we have $\Pi^+\subseteq e^{i\theta}S-\beta$. Thus $V(b)\subseteq e^{i\theta}S-\beta$. It follows that $V(a)\subseteq S$. Since $Z_\nu(a)$ is the intersection of its supporting half-spaces, the theorem follows. □

(2.5) Corollary. For any norm ν, Z_ν is a convex numerical range. □

Since $Z_\nu=Z_{\nu D}$ we call Z_ν the *symmetric numerical range*. If ν is the Euclidean norm, then $Z_\nu=Z_{\nu D}=V_\nu=V_{\nu D}=V$. Also, if $h\in\mathbb{C}^{nn}$ is Hermitian, $y^* hx=x^* hy$, for $(x,y)\in\Pi_\nu$, whence $Z_\nu(h)\subseteq\mathbb{R}$.

(2.6) Theorem. Let ν be the l_1 or l_∞-norm on $\mathbb{C}^n$ and let G be defined as in (1.3). Then, for all $a\in\mathbb{C}^{nn}$,

$$Z_\nu(a)\subseteq G(a).$$

Proof. It is enough to prove the result when ν is the l_1-norm, since then ν^D is the l_∞-norm. So let $(x,y)\in\Pi_\nu$. Suppose

$$x=(r_1 e^{i\theta_1},\ldots,r_n e^{i\theta_n})^t,\qquad y=(s_1 e^{i\phi_1},\ldots,s_n e^{i\phi_n})^t,$$

where $r_i\geqq 0$, $s_i\geqq 0$, $i=1,\ldots,n$, and $0\leqq\theta_i,\ \phi_i<2\pi$. Then $\sum_{i=1}^n r_i=1$, and $0\leqq s_i\leqq 1$, $i=1,\ldots,n$. Let $N=\{1,\ldots n\}$ and $E=\{i\in N: r_i>0\}$. Then

$$\begin{aligned}\tfrac{1}{2}(y^* ax+x^* ay)&=\tfrac{1}{2}\sum_{i,j\in N}(s_i a_{ij} r_j e^{i(\theta_j-\phi_i)}+r_j a_{ji} s_i e^{-i(\theta_i-\phi_i)})\\&=\sum_{j\in E} r_j\xi_j,\end{aligned}$$

where, for $j\in E$,

$$\xi_j=\tfrac{1}{2}\sum_{i\in N}(s_i a_{ij} e^{i(\theta_j-\phi_i)}+a_{ji} s_i e^{-i(\theta_i-\phi_i)})$$
$$=a_{jj}+\tfrac{1}{2}\sum_{i\in N\backslash\{j\}}(s_i a_{ij} e^{i(\theta_j-\phi_i)}+a_{ji} s_i e^{-i(\theta_j-\phi_i)}),$$

since, for $j\in E$, we have $s_j=1$ and $\theta_j=\phi_j$.

Hence

$$|\xi_j-a_{jj}|\leqq\tfrac{1}{2}\sum_{i\in N\backslash\{j\}}(|a_{ij}|+|a_{ji}|).$$

Since $\sum_{j\in E} r_j=1$, it follows that $\tfrac{1}{2}(y^* a x+x^* a y)\in G(a)$. □

(2.7) *Example.* Let

$$a=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}$$

and let ν be the l_1 norm. Then

$$V(a)\subset Z_\nu(a)\subset V_\nu(a)\cap V_{\nu D}(a)\subset G(a)$$

where all the containments are strict.

For, $V(a)=[0,5]$, $Z_\nu(a)=[-1,6]$, $G(a)=\operatorname{conv}(G_1(a), G_2(a))$, where $G_1(a)$ and $G_2(a)$ are the circles of radius 2 with center 1 and 4 respectively. By Nirschl-Schneider [11], $V_\nu(a)\supseteq G_i(a)$, $i=1,2$ and $V_\nu(a)$ is non-convex since the segment $[1+2i, 4+2i]$ intersects $V_\nu(a)$ in the points $\{1+2i, 4+2i\}$ only. We also have $V_\nu(a)=V_{\nu D}(a)$. The last remark is a consequence of the next result whose proof is easy.

(2.8) Lemma. Let ν be a norm on $\mathbb{C}^n$ such that $\nu(x)=\nu(\bar{x})$, for all $x\in\mathbb{C}^n$. If $a\in\mathbb{C}^{nn}$ is symmetric (viz. $a=a^t$) then $V_\nu(a)=V_{\nu D}(a)$. □

3. Comparison Between Numerical Ranges

(3.1) Definition. Let ν be a norm on $\mathbb{C}^n$. Then

$$H(\nu)=\{h\in\mathbb{C}^{nn}: V_\nu(h)\subseteq\mathbb{R}\},$$

and

$$J(\nu)=\{h+i\,k: h, k\in H(\nu)\}.$$

If ν is the l_2-norm, then $H(\nu)$ consists of the set of Hermitian matrices and we shall denote this set by H. For general ν, Vidav [14] (cf. Bonsall-Duncan [3, p. 51]) proved that if $h\in H(\nu)$ then $V_\nu(h)=\operatorname{conv}(\operatorname{spec} h)$. This equality motivates the more general assumption of our next lemma.

(3.2) Lemma. Let W_1, W_2 be convex numerical ranges. Let $h, k\in\mathbb{C}^{nn}$ and suppose that for all $\alpha,\beta\in\mathbb{R}$

$$W_1(\alpha h+\beta k)=W_2(\alpha h+\beta k)\subseteq\mathbb{R}.$$

Then for all $\eta,\zeta\in\mathbb{C}$,

$$W_1(\eta h+\zeta k)=W_2(\eta h+\zeta k).$$

Proof. Let $a=\eta\, h+\zeta\, k$ where $\eta,\zeta\in\mathbb{C}$ and put $\eta=\alpha+i\gamma$, $\zeta=\beta+i\delta$, where $\alpha,\beta,\gamma,\delta\in\mathbb{R}$. For $k=1,2$, $W_k(a)\subseteq W_k(\alpha h+\beta k)+i\,W_k(\gamma h+\delta k)$, whence Re $W_k(a)\subseteq$

$W_k(\alpha h+\beta k)$. But also $W_k(\alpha h+\beta k) \subseteq W_k(a) - i W_k(\gamma h+\delta k)$, whence $\mathrm{Re}\, W_k(a) = W_k(\alpha h+\beta k)$. It follows that $\mathrm{Re}\, W_1(a) = \mathrm{Re}\, W_2(a)$.

But the same argument shows that for each θ, $0 \leqq \theta < 2\Pi$, $\mathrm{Re}\, W_1(e^{i\theta} a) = \mathrm{Re}\, W_2(e^{i\theta} a)$. Since $W_1(a)$ and $W_2(a)$ are convex subsets of the plane, we deduce that $W_1(a) = W_2(a)$. □

If ν is a norm on $\mathbb{C}^n$, and $p \in \mathbb{C}^{nn}$ is non-singular, we define the norm ν_p by $\nu_p(x) = \nu(p x)$, for all $x \in \mathbb{C}^n$. If p is the Loewner-John matrix for ν as defined in Deutsch-Schneider [6], then $H(\nu_p) \subseteq H$, see [6, Proposition 4.1]. This motivates the hypothesis of our next theorem. We first state a simple lemma.

(3.3) Lemma. Let ν be a norm on $\mathbb{C}^n$ and let $h, k \in H(\nu) \cap H$. If $(x,y) \in \Pi_\nu$ then $y^*(h+ik)\, x = x^*(h+ik)\, y$.

Proof. $y^*(h+ik)\, x = y^* h x + i y^* k x = x^* h^* y + i x^* k^* y = x^* h y + i x^* k y = x^*(h+ik)\, y$. □

(3.4) Theorem. Let ν be a norm on $\mathbb{C}^n$ and suppose $h, k \in H(\nu) \cap H$. Then

$$Z_\nu(h+ik) = \mathrm{conv}\, V_\nu(h+ik) = V(h+ik).$$

Proof. Let $a = h+ik$. By Lemma (3.3), $y^* a x = \frac{1}{2}(y^* a x + x^* a y)$, for all $(x, y) \in \Pi_\nu$. Hence $Z_\nu(a) = \mathrm{conv}\, V_\nu(a)$. To prove the second equality, observe that

$$V_\nu(\alpha h+\beta k) = \mathrm{conv}\, \mathrm{spec}(\alpha h+\beta k) = V(\alpha h+\beta k),$$

by Vidav's Lemma [14]. Hence by Lemma (3.2), $\mathrm{conv}\, V_\nu(h+ik) = V(h+ik)$. □

Remark. A considerably better result holds: $Z_\nu(h+ik) = V_\nu(h+ik)$, under the hypothesis of Theorem (3.4). For, using Theorem (3.4), B. D. Saunders [12] shows that for $h, k \in H(\nu)$, $V_\nu(h+ik)$ is itself convex.

Since $V_{\nu_p}(p^{-1} a p) = V_\nu(a)$ (cf. [11, Lemma 2] it follows by [6, Proposition 4.1] that:

(3.5) Corollary. Let ν be a norm on $\mathbb{C}^n$ and let p be the Loewner-John matrix for ν. If $a \in J(\nu)$, then

$$Z_{\nu_p}(p^{-1} a p) = \mathrm{conv}\, V_\nu(a) = V(p^{-1} a p).$$

Proof. By Theorem (3.4),

$$Z_{\nu_p}(p^{-1} a p) = V(p^{-1} a p) = \mathrm{conv}\, V_{\nu_p}(p^{-1} a p) = \mathrm{conv}\, V_\nu(a). \quad \square$$

4. Extreme Points

We next prove a theorem concerning extreme points of $Z_\nu(a)$ which also belong to $V(a)$.

(4.1) Theorem. Let ν be a norm, let $a \in \mathbb{C}^{nn}$, and let z be such that $z^* z = 1$ and $z^* a z$ is an extreme point of $Z_\nu(a)$. If (x, y) is the Π_ν-decomposition of the direction z, then

$$z^* a z = \frac{x^* a x}{x^* x} = \frac{y^* a y}{y^* y} = \frac{1}{2}(y^* a x + x^* a y).$$

Proof. Let $\alpha = z^* a z$ and put $b = a - \alpha$. For some $t > 0$, we have $tz = x + y$ and so

$$0 = t^2 z^* b z = x^* x \alpha_1 + y^* y \alpha_2 + 2\alpha_3,$$

where

$$\alpha_1 = x^* b x / x^* x, \qquad \alpha_2 = y^* b y / y^* y,$$

and

$$\alpha_3 = \tfrac{1}{2}(y^* b x + x^* b y).$$

Then, for $i = 1, 2, 3$, we have by Theorem (2.4) that $\alpha_i \in Z_\nu(b)$. Let $s = x^* x + y^* y + 2$. If $r_1 = s^{-1} x^* x$, $r_2 = s^{-1} y^* y$, $r_3 = 2s^{-1}$, then $0 = \sum_{i=1}^{3} r_i \alpha_i$ and $\sum_{i=1}^{3} r_i = 1$, $r_i > 0$ for $i = 1, 2, 3$. Since 0 is an extreme point of $Z_\nu(b)$ it follows that $\alpha_1 = \alpha_2 = \alpha_3 = 0$. The result follows. □

(4.2) Corollary. Let ν be a norm on $\mathbb{C}^n$, and suppose that $h, k \in H(\nu) \cap H$. Let $a = h + ik$. If α is an extreme point of $\operatorname{conv} V_\nu(a)$ then there exist $(x, y) \in \Pi_\nu$ such that

$$\alpha = y^* a x = \frac{x^* a x}{x^* x} = \frac{y^* a y}{y^* y}.$$

Proof. By Theorem (3.4), α is an extreme point of $Z_\nu(a)$ and $\alpha \in V(a)$. Let $z \in \mathbb{C}^n$ such that $z^* z = 1$ and $z^* a z = \alpha$. Since $y^* a x = x^* a y$, by Lemma (3.3), the result follows from Theorem (4.1). □

(4.3) Corollary. Let ν be a norm on $\mathbb{C}^n$ and let $h \in H(\nu) \cap H$. Let $\lambda = \min(\operatorname{spec} h)$ (or $\lambda = \max(\operatorname{spec} h)$) and suppose $hz = \lambda z$, where $0 \neq z \in \mathbb{C}^n$. If (x, y) is the Π_ν-decomposition of the direction z, then $hx = \lambda x$ and $hy = \lambda y$.

Proof. By Corollary (4.2), $x^* h x / x^* x = y^* h y / y^* y = \lambda$, it is easy to show that if $x^* h x / x^* x = \lambda$ then $h x = \lambda x$ ([7, p. 142]). Similarly we may prove that $h y = \lambda y$. □

Corollary (4.3) is related to a result proved under more general hypotheses concerning the existence of pairs $(x, y) \in \Pi_\nu$ which are pairs of eigenvectors of a matrix, see Zenger [15, 1.8].

References

1. Bauer, F. L.: On the field of values subordinate to a norm. Numer. Math. **4**, 103–113 (1962)
2. Bauer, F. L.: Fields of value and Gershgorin discs. Numer. Math. **12**, 91–95 (1968)
3. Bonsall, F. F., Duncan, J.: Numerical ranges of operators on normed spaces and of elements of normed algebras. Lond. Math. Soc. Lecture Note Series, 2, Cambridge: U. Press 1971
4. Cain, B. E., Saunders, B. D., Schneider, H.: On the geometry of dual pairs. Studies App. Math. (to appear)
5. Deutsch, E., Zenger, C.: On Bauer's generalized Gershgorin discs. Numer. Math. **24**, 63–70 (1975)
6. Deutsch, E., Schneider, H.: Bounded groups and norm-hermitian matrices. Linear Algebra and Appl. **9**, 9–27 (1975)
7. Franklin, J. N.: Matrix Theory. Prentice-Hall 1968
8. Hausdorff, F.: Der Wertvorrat einer Bilinearform. Math. Zeitschrift **3**, 314–316 (1919)

9. Johnson, C. R.: A Gershgorin inclusion set for the field of values of a finite matrix. Proc. Amer. Math. Soc. **41**, 57–60 (1973)
10. Lumer, G.: Semi-inner product spaces. Trans. Amer. Math. Soc. **10**, 29–43 (1961)
11. Nirschl, N., Schneider, H.: The Bauer field of values of a matrix. Numer. Math. **6**, 355–365 (1964)
12. Saunders, B. D.: A condition for the convexity of the norm-numerical range of a matrix. Linear Algebra and Appl. (to appear)
13. Toeplitz, O.: Das algebraische Analogon zu einem Satze von Fejer. Math. Zeitschrift **2**, 187–197 (1918)
14. Vidav, I.: Eine metrische Kennzeichnung der selbst-adjungierten Operatoren. Math. Zeit. **66**, 185–193 (1956)
15. Zenger, C.: On convexity properties of the Bauer field of values of a matrix. Numer. Math. **12**, 96–105 (1968)
16. Zenger, C.: Minimal subadditive inclusion domains for the eigenvalues of matrices. Linear Algebra and Appl. (to appear)

B. David Saunders
Mathematical Sciences Department
Rensselaer Polytechnic Inst.
Troy, NY 12181
U.S.A.

Hans Schneider
Mathematics Department
University of Wisconsin
Madison, WI 53706
U.S.A.