A Generalization of the Numerical Range of a Matrix

H. W. J. Lenferink* and M. N. Spijker**
Department of Mathematics and Computer Science
University of Leiden
Niels Bohrweg 1
2333 CA Leiden, The Netherlands

Submitted by Harald Wimmer

ABSTRACT

A new generalization of the classical numerical range of a matrix is introduced. Various properties and basic characterizations of this generalized numerical range are established. One of these characterizations is used in proving a resolvent inequality for a discretized version of the differential operator $d^2/dx^2$.

1. INTRODUCTION

1.1. The Classical Numerical Range and Some Known Generalizations

The classical numerical range of a given complex $n \times n$ matrix $A$ is defined as the set of complex numbers

$$\{x^*Ax : x \in \mathbb{C}^n \text{ with } x^*x = 1\}.$$

Here $x^*$ denotes the Hermitian adjoint of the column vector $x \in \mathbb{C}^n$. The classical numerical range is known to be a convex set. It allows various generalizations.

---

*This research has been supported by the Netherlands Organization for Scientific Research (NWO).

**Address correspondence to M. N. Spijker.
Let $|\cdot|$ denote an arbitrary norm on $\mathbb{C}^n$, and $|\cdot|^*$ its dual norm, given by $|y|^* = \sup\{|y^*x|: x \in \mathbb{C}^n \text{ with } |x| = 1\}$ for arbitrary $y \in \mathbb{C}^n$. The Bauer field of values [1], also called the spatial numerical range [3–5], is defined as the set

$$\mathcal{F}[A] = \{y^*Ax: x, y \in \mathbb{C}^n \text{ with } y^*x = |x| = |y|^* = 1\}. \quad (1.1)$$

If the norm on $\mathbb{C}^n$ equals the Euclidean norm $|x| = \sqrt{x^*x}$, the spatial numerical range reduces to the classical numerical range, and thus a convex set. However, in general the spatial numerical range is not convex [18, 23], which may be inconvenient.

Let $\|\cdot\|$ denote the norm on the algebra $\mathbb{C}^{n \times n}$ of all $n \times n$ matrices that is induced by the given vector norm $|\cdot|$ on $\mathbb{C}^n$, and let $\|\cdot\|^*$ denote the dual norm of $\|\cdot\|$. For any $n \times n$ matrix $A$ we thus may write $\|A\| = \sup\{|Ax|: x \in \mathbb{C}^n \text{ with } |x| = 1\}$ and $\|A\|^* = \sup\{|\text{tr}(AX)|: X \in \mathbb{C}^{n \times n} \text{ with } \|X\| = 1\}$, where $\text{tr}(\cdot)$ denotes the trace function (e.g. [15]). It is easily seen that the set

$$\{\text{tr}(XA): X \in \mathbb{C}^{n \times n} \text{ with } \|X\|^* = \text{tr}(X) = 1\} \quad (1.2)$$

coincides with the so-called algebra numerical range, which was studied extensively e.g. in [3–5]. If the norm on $\mathbb{C}^n$ satisfies $|x| = \sqrt{x^*x}$, this set again reduces to the classical numerical range. In general, the algebra numerical range equals the convex hull of the spatial numerical range, i.e. the intersection of all convex subsets of the complex plane containing $\mathcal{F}[A]$. We note that also the terms Gerschgorin domain [21], Hausdorff set, and field of values (e.g. [11]) occur in the literature to designate sets that are defined differently from (1.2) but can be seen to coincide with the algebra numerical range.

For complex $\gamma$ and $\rho \geq 0$ we introduce the disk

$$D[\gamma, \rho] = \{\zeta: \zeta \in \mathbb{C} \text{ with } |\zeta - \gamma| \leq \rho\}. \quad (1.3)$$

It follows from [4, p. 42] or [5, p. 44] that the set (1.2) is equal to

$$\bigcap_{\gamma \in \mathbb{C}} D[\gamma, \|A - \gamma I\|], \quad (1.3)$$

where $I$ denotes the $n \times n$ identity matrix. The formula (1.3) is the starting point for the generalization of the algebra numerical range to be studied in the present paper.

We refer to [15, p. 332, 333; 5; 12; 13] for generalizations in other directions.
1.2. The M-Numerical Range

We recall that a norm \( \| \cdot \| \) on the vector space \( \mathbb{C}^{n,n} \) is called a matrix norm if, for all \( n \times n \) matrices \( A, B \), we have \( \| A B \| \leq \| A \| \cdot \| B \| \) [15, p. 290].

Let \( \| \cdot \| \) be such a matrix norm, and \( M \) a given positive constant. Assume \( A \) is a given \( n \times n \) matrix. We focus on disks \( D[\gamma, \rho] \) with arbitrary \( \gamma \in \mathbb{C} \), \( \rho > 0 \) such that
\[
\|(A - \gamma I)^k\| \leq M \rho^k \quad (k = 1, 2, 3, \ldots).
\]
(1.4.a)

We define the \( M \)-numerical range of \( A \) by
\[
\tau_M[A; \| \cdot \|] = \bigcap D[\gamma, \rho],
\]
(1.4.b)
where the intersection is over all disks \( D[\gamma, \rho] \) with the property (1.4.a).

When \( M = 1 \), we see that the set (1.4.b) coincides with (1.3), so that the concept of \( M \)-numerical range generalizes the concept of algebra numerical range. The authors were led to consider the above generalization as a result of their experience that the \( M \)-numerical range is a handy and natural tool in the stability analysis of various numerical processes [16, 17].

For all positive \( M \) the set \( \tau_M[A; \| \cdot \|] \) can be seen to be compact and convex, with
\[
\text{conv} \sigma[A] \subset \tau_M[A; \| \cdot \|],
\]
(1.5.a)
where \( \sigma[A] \) denotes the spectrum of \( A \) and \( \text{conv} \) stands for the convex hull. Further, it is easy to see that
\[
\tau_N[A; \| \cdot \|] \subset \tau_M[A; \| \cdot \|] \quad \text{(for } M < N),
\]
(1.5.b)
\[
\tau_M[\xi_0 I + \xi_1 A; \| \cdot \|] = \xi_0 + \xi_1 \cdot \tau_M[A; \| \cdot \|] \quad \text{(for } \xi_0, \xi_1 \in \mathbb{C}).
\]
(1.5.c)

1.3. Scope of the Rest of This Paper

In Section 2.1 we present Theorem 2.1, which constitutes the main result of our paper. This theorem provides us with three basic characterizations of the \( M \)-numerical range, which are different from (1.4). Section 2.2 contains a series of technical lemmata needed to prove our main result. In Section 2.3 we give the actual proof of Theorem 2.1.

In Section 3.1 we discuss various interpretations and applications of Theorem 2.1. An application with \( M = 1 \) shows that the Gerschgorin domain and the Hausdorff set, mentioned above in Section 1.1, coincide with the
algebra numerical range (1.3). Applications with $M > 1$ lead to generalizations of material to be found e.g. in [9, 18]. In Section 3.2 we apply the $M$-numerical range (with $M > 1$) in proving a resolvent inequality for a discrete version of the second order derivative operator $d^2/dx^2$.

2. FORMULATION AND PROOF OF THE MAIN THEOREM

2.1. Basic Characterizations of the $M$-Numerical Range

We start with some definitions that are used in the subsequent to formulate new characterizations of the $M$-numerical range.

Let $V$ be an arbitrary convex subset of $\mathbb{C}$, and let $\zeta \in \mathbb{C}$. The distance from $\zeta$ to $V$ is defined by

$$d(\zeta, V) = \inf\{|\zeta - \xi| : \xi \in V\}.$$ 

If $\xi$ belongs to the boundary $\partial V$ of $V$ and

$$\Re\{e^{-i\theta}(\zeta - \xi)\} \leq 0 \quad \text{for all} \quad \zeta \in V,$$

where $\theta$ is a real constant, then $\theta$ is called a normal direction to $V$ at $\xi$. We recall that a matrix norm is called unital if the norm of the identity matrix equals one.

We shall deal with the following four conditions on $A$ with respect to $V \subset \mathbb{C}$.

\begin{align*}
\tau_M[A; \cdot ||.] &\subset V, \quad (2.1.a) \\
(\xi I - A) \text{ is regular and } ||(\xi I - A)^{-k}|| \leq M \cdot [d(\xi, V)]^{-k} \quad &\text{(for all } \xi \not\in V \text{ and } k = 1, 2, 3, \ldots\text{)}, \quad (2.1.b) \\
||\exp[te^{-i\theta}(A - \xi I)]|| &\leq M \quad \text{(for all } t \geq 0, \xi \in \partial V, \text{ and normal directions } \theta \text{ to } V \text{ at } \xi\text{)}, \quad (2.1.c) \\
\text{There is a unital matrix norm } ||\cdot||_1 \text{ on } \mathbb{C}^{n \times n} \text{ with } \tau_1[A; ||\cdot||_1] &\subset V \text{ and } M^{-1}||X|| \leq ||X||_1 \leq M||X|| \quad \text{(for all } X \in \mathbb{C}^{n \times n}\text{).} \quad (2.1.d)
\end{align*}

**Theorem 2.1.** Let $||\cdot||$ be a matrix norm (not necessarily unital) on $\mathbb{C}^{n \times n}$, and let $A$ be an $n \times n$ matrix. Let $M \geq ||I||$, and $V$ a given nonempty,
closed, and convex subset of \( \mathbb{C} \). Then the conditions (2.1.a)-(2.1.d) are equivalent to each other.

We note that the assumption \( M \geq \|I\| \) in the theorem cannot be omitted. This can be seen by choosing \( t = 0 \) in the first inequality of (2.1.c).

Clearly, \( \tau_M[A; \| \cdot \|] \) is the smallest nonempty, closed, and convex set \( V \subset \mathbb{C} \) with the property (2.1.a). Therefore, the above theorem reveals three new characterizations of the \( M \)-numerical range. We see that \( \tau_M[A; \| \cdot \|] \) equals the smallest nonempty, closed, convex set \( V \subset \mathbb{C} \) with the property (2.1.b), and the same holds with regard to the properties (2.1.c) and (2.1.d).

2.2. Technical Lemmata

In the following, \( \| \cdot \| \) is a matrix norm on \( \mathbb{C}^{n,n} \), \( A \) an \( n \times n \) matrix, and \( M \) a given constant with \( M \geq \|I\| \).

**Definition 2.2.** A disk \( D[\gamma, \rho] \) is called suitable if \( \gamma, \rho \) are such that (1.4.a) holds.

**Lemma 2.3.** Let the disk \( D[\gamma, \rho] \) be suitable. Then

\[
\| \exp[te^{-i\theta}(A - \xi I)] \| \leq M \quad \text{(for all } t \geq 0 \text{ and } \xi = \gamma + \rho e^{i\theta} \text{ with real } \theta). \]

**Proof.** Let \( t > 0 \) and \( \xi = \gamma + \rho e^{i\theta} \) for some real \( \theta \). We have

\[
\| \exp[te^{-i\theta}(A - \gamma I)] \| \leq \left\| \sum_{k=0}^{\infty} \frac{1}{k!} \left[ te^{-i\theta}(A - \gamma I) \right]^k \right\|
\leq \|I\| + \sum_{k=1}^{\infty} \frac{1}{k!} t^k \| (A - \gamma I) \| \leq Me^{\|I\|}. 
\]

Hence,

\[
\| \exp[te^{-i\theta}(A - \xi I)] \| = \| e^{-i\theta} \exp[te^{-i\theta}(A - \gamma I)] \| \leq M. \]

**Lemma 2.4.** Let \( \theta_1, \theta_2 \) be given with \( 0 < \theta_2 - \theta_1 < \pi \). If

\[
\| \exp(te^{-i\theta_1}A) \| \leq M, \quad \| \exp(te^{-i\theta_2}A) \| \leq M \quad \text{(for all } t \geq 0), \]

then also

\[ \| \exp( t e^{-i\theta} ) \| \leq M \quad \text{for all } t \geq 0 \text{ and all } \theta \in [\theta_1, \theta_2] \].

**Proof.** Define \( \mathcal{D} = \{ \zeta : \zeta \in \mathbb{C} \text{ and } \zeta = 0 \text{ or } -\theta_2 \leq \arg \zeta \leq -\theta_1 \} \) and \( \theta_0 = -\frac{1}{2}(\theta_1 + \theta_2) \). Let \( F \) be any complex linear functional on \( \mathbb{C}^{n,n} \) with \( \| F \| = \sup_{\|X\| = 1} |F(X)| = 1 \). We shall prove that

\[ |F(e^{\xi A})| < M \quad \text{(2.2)} \]

for all \( \zeta \in \mathcal{D} \).

Let \( \varepsilon > 0 \), and consider

\[ f_\varepsilon(\zeta) = \exp(-\varepsilon \zeta e^{-i\theta_0}) F(e^{\xi A}) \quad \text{for } \zeta \in \mathcal{D}. \]

Since (2.2) is valid for \( \zeta \in \partial \mathcal{D} \), we see that also

\[ |f_\varepsilon(\zeta)| \leq M \quad \text{for } \zeta \in \partial \mathcal{D}. \]

Let \( c > 0 \) be so large that \( |\exp(-\varepsilon \zeta e^{-i\theta_0})| \leq M^{-1} \) for all \( \zeta \in \mathcal{D} \) with \( |\zeta| \geq c \).

For any such \( \zeta \), there exist \( \zeta_1, \zeta_2 \in \partial \mathcal{D} \) with \( \zeta_1 + \zeta_2 = \zeta \). Hence

\[ |f_\varepsilon(\zeta)| \leq M^{-1} |F(\exp(\zeta_1 A) \exp(\zeta_2 A))| \leq M \quad \text{for } \zeta \in \mathcal{D} \text{ with } |\zeta| \geq c. \]

Since \( f_\varepsilon(\zeta) \) is analytic for \( -\theta_2 < \arg \zeta < -\theta_1 \), \( 0 < |\zeta| < c \), we see, by applying the principle of the maximum modulus, that also

\[ |f_\varepsilon(\zeta)| \leq M \quad \text{for } \zeta \in \mathcal{D} \text{ with } |\zeta| < c. \]

Hence,

\[ |f_\varepsilon(\zeta)| \leq M \quad \text{for all } \zeta \in \mathcal{D}. \]

By letting \( \varepsilon \downarrow 0 \) we conclude that (2.2) holds for all \( \zeta \in \mathcal{D} \).

The inequality (2.2) holds for all \( F \) with the above properties. Therefore, a corollary of the Hahn-Banach theorem (e.g. [20]) implies that \( \| e^{\xi A} \| \leq M \) (for all \( \xi \in \mathcal{D} \)). This completes the proof of the lemma. \( \blacksquare \)
Lemma 2.5. For every $\varepsilon > 0$ there exist suitable disks $D_1, D_2, \ldots, D_m$ such that if $\zeta \in D_1 \cap D_2 \cap \cdots \cap D_m$, then

$$d(\zeta, \tau_M[A; \|\cdot\|]) < \varepsilon.$$ 

Proof. Let $\Omega$ be the collection of pairs $(\gamma, \rho)$ such that the disk $D[\gamma, \rho]$ is suitable, and let $\varepsilon > 0$ be given. The disk $D[\alpha] = D[0, \|A\|]$ is suitable. Put $K_\varepsilon = \{\zeta : \zeta \in D[\alpha] \text{ with } d(\zeta, \tau_M[A; \|\cdot\|]) < \varepsilon\}$. Then $D[\alpha] \setminus K_\varepsilon$ is compact and is covered by the open sets $\mathbb{C} \setminus D[\omega]$, where $\omega$ runs through $\Omega$. Hence $D[\alpha] \setminus K_\varepsilon \subset \bigcup_{\omega \in \Psi \cup \{\alpha\}} \mathbb{C} \setminus D[\omega]$ for some finite $\Psi \subset \Omega$ and

$$\bigcap_{\omega \in \Psi \cup \{\alpha\}} D[\omega] \subset \{\zeta : \zeta \in \mathbb{C} \text{ with } d(\zeta, \tau_M[A; \|\cdot\|]) < \varepsilon\}.$$ 

The following lemma and proof are related to material in [3, p. 21].

Lemma 2.6. Let $\mathcal{R}$ be a collection of $n \times n$ matrices such that

(a) $I \in \mathcal{R}$,
(b) $\|R\| \leq M$ (for all $R \in \mathcal{R}$),
(c) For every $R_0, R_1 \in \mathcal{R}$ there exist $\lambda \in \mathbb{C}, R_2 \in \mathcal{R}$ with

$$|\lambda| \leq 1, \quad R_0 R_1 = \lambda R_2.$$ 

Then there exists a unital matrix norm $\|\cdot\|_1$ on $\mathbb{C}^{n \times n}$ with

(i) $\|R\|_1 \leq 1$ (for all $R \in \mathcal{R}$),
(ii) $M^{-1}\|X\|_1 \leq \|X\|_1 \leq M\|X\|_1$ (for all $X \in \mathbb{C}^{n \times n}$).

Proof. For arbitrary $n \times n$ matrices $X$ we define

$$\|X\|_0 = \sup\{\|RX\| : R \in \mathcal{R}\}, \quad \|X\|_1 = \sup\{\|XY\|_0 : Y \in \mathbb{C}^{n \times n} \setminus \{0\}\}.$$ 

Using properties (a), (b) it can be proved that $\|\cdot\|_0, \|\cdot\|_1$ are matrix norms with $\|I\|_1 = 1$, and

$$\|X\|_1 \leq \|X\|_0 \leq M\|X\|.$$ 

Since

$$\|X\| \leq \|X\|_0 \leq \|X\|_1 \leq \|X\| \leq M\|X\|_1,$$

property (ii) has thus been proved.
For \( R_1 \in \mathcal{R} \) we have, in view of property (c),
\[
\|R_1\|_1 = \sup\{\sup\{\|R_0R_1Y\| : R_0 \in \mathcal{R}\} : Y \in \mathbb{C}^{n,n}, \|Y\|_0 \leq 1\}
\leq \sup\{\sup\{\|HY\| : R \in \mathcal{R}\} : Y \in \mathbb{C}^{n,n}, \|Y\|_0 \leq 1\} = 1.
\]
This completes the proof.

2.3. The Proof of Theorem 2.1

Let \( \|\cdot\| \), \( A \), \( M \), and \( V \) be as in Theorem 2.1.

2.3.1. The Equivalence of (2.1.b) and (2.1.c). The condition (2.1.b) can be reformulated as

Let \( \xi \in \partial V \), let \( \theta \) be a normal direction to \( V \) at \( \xi \), and let \( B = e^{-\theta}(A - \xi I) \). Then \( \lambda I - B \) is regular and
\[
\|(\lambda I - B)^{-k}\| = M\lambda^{-k} \quad \text{for all real } \lambda > 0 \text{ and integers } k \geq 1.
\]

Let (2.1.b') be satisfied. Using the formula \( e^{tb} = \lim_{k \to \infty} (I - tk^{-1}B)^{-k} \) (e.g. [9]), we obtain (2.1.c).

Conversely, assuming (2.1.c), we easily arrive at (2.1.b') by using the formula
\[
(\lambda I - B)^{-k} = \frac{1}{(k - 1)!} \int_0^\infty t^{k-1} e^{-\lambda t} e^{tb} dt \quad \text{for } \lambda > 0
\]
(e.g. [9]).

2.3.2. (2.1.c) Implies (2.1.d). Let \( \mathcal{R} = \{\exp[te^{-i\theta}(A - \xi I)] : t \geq 0, \xi \in \partial V, \text{and } \theta \text{ a normal direction to } V \text{ at } \xi\} \). Under the assumption (2.1.c) this set \( \mathcal{R} \) satisfies conditions (a) and (b) of Lemma 2.6. We shall prove that condition (c) is fulfilled as well. Let \( R_j = \exp[t_j e^{-i\theta_j}(A - \xi_j I)] \) for some \( t_j \gg 0, \xi_j \in \partial V, \) and normal direction \( \theta_j \) to \( V \) at \( \xi_j \) (for \( j = 0,1 \)). Defining \( t_2 \gg 0, \theta_2 \in \mathbb{R} \) such that \( t_2 e^{-i\theta_2} = t_0 e^{-i\theta_0} + t_1 e^{-i\theta_1} \), we have \( R_0R_1 = \exp[t_0 e^{-i\theta_0}(\xi_2 - \xi_0) + t_1 e^{-i\theta_1}(\xi_2 - \xi_1)]\exp[t_2 e^{-i\theta_2}(A - \xi_2 I)] \) for any \( \xi_2 \in \mathbb{C} \).

It is possible to choose \( \xi_2 \) such that \( \xi_2 \in \partial V \) and \( \theta_2 \) is a normal direction to \( V \) at \( \xi_2 \). With such a \( \xi_2 \) we have, since \( \theta_j \) is a normal direction to \( V \) at \( \xi_j \) (\( j = 0,1 \)), the inequalities \( \Re[te^{-i\theta}(\xi_2 - \xi_j)] \leq 0 \) (\( i = 0,1 \)). Choose \( R_2 = \exp[t_2 e^{-i\theta_2}(A - \xi_2 I)], \lambda = \exp[t_0 e^{-i\theta_0}(\xi_2 - \xi_0) + t_1 e^{-i\theta_1}(\xi_2 - \xi_1)]. \) Then \( R_0R_1 = \lambda R_2 \) with \( |\lambda| \leq 1 \) and \( R_2 \in \mathcal{R} \). Thus, all conditions of Lemma 2.6 are
satisfied with this choice of $\mathcal{B}$, and there exists a matrix norm $\|\cdot\|_1$ as in the conclusion of that Lemma.

It remains to show that $\tau_1[A;\|\cdot\|_1] \subset V$. Let $\zeta \in \mathbb{C} \setminus V$ be given. We will show that $\zeta \not\in \tau_1[A;\|\cdot\|_1]$. There exist $\lambda > 0$, $\xi \in \partial V$, and a normal direction $\theta$ to $V$ at $\xi$ such that $\zeta = \xi + \lambda e^{i\theta}$. Since $\|\exp[te^{-i\theta}(A - \xi I)]\|_1 \leq 1$ for $t > 0$, we have

$$h^{-1}\left[\|I + h[e^{-i\theta}(A - \xi I)]\|_1 - 1\right] = O(h) \quad \text{for } h \downarrow 0.$$ 

Hence, for some $h_0 > 0$,

$$h_0^{-1}\left[\|I + h_0[e^{-i\theta}(A - \xi I)]\|_1 - 1\right] \leq \frac{\lambda}{2}.$$ 

This implies that (1.4.a) holds with the matrix norm $\|\cdot\|_1$ and with $\gamma = \xi - h_0^{-1}e^{i\theta}$, $M = 1$, $\rho = h_0^{-1} + \frac{\lambda}{2}$. However, $\zeta$ is not contained within the disk $D[\xi - h_0^{-1}e^{i\theta}, h_0^{-1} + \frac{1}{2}\lambda]$. Hence, $\zeta \not\in \tau_1[A;\|\cdot\|_1]$. This completes the proof of (2.1.d).

2.3.3. The Inclusion (2.1.a) Is Implied by (2.1.d). Let $\|\cdot\|_1$ be as in (2.1.d). For any $\gamma \in \mathbb{C}$, $\rho > 0$ with $\|(A - \gamma I)^k\|_1 \leq \rho^k$ ($k = 1, 2, 3, \ldots$), we also have $\|(A - \gamma I)^k\| \leq M\rho^k$ ($k = 1, 2, 3, \ldots$). Hence $\tau_M[A;\|\cdot\|] \subset \tau_1[A;\|\cdot\|_1]$, from which we see that (2.1.d) implies (2.1.a).

2.3.4. The Inclusion (2.1.a) Implies (2.1.b). Assume (2.1.a). Let $\varepsilon > 0$ be given. By Lemma 2.5, there exist suitable disks $D_1, \ldots, D_m$, such that, with $E = \bigcap_{i=1}^m D_i$, 

$$d(\zeta, \tau_M[A;\|\cdot\|]) < \varepsilon \quad \text{(for all } \zeta \in E).$$

By applying Lemma 2.3 to the disks $D_1, D_2, \ldots, D_m$, and applying Lemma 2.4 (with $A$ replaced by $A - \xi I$, where $\xi \in \partial E$), it can be seen that

$$\|\exp[te^{-i\theta}(A - \xi I)]\| \leq M$$

for all $t > 0$, $\xi \in \partial E$, and normal directions $\theta$ to $E$ at $\xi$. Since the conditions (2.1.b) and (2.1.c) are equivalent for any convex $V \subset \mathbb{C}$, we may choose $V = E$
so as to conclude that

$$\xi I - A \text{ is regular and } \|(\xi I - A)^{-k}\| \leq M[d(\xi, E)^{-k}]$$

(for all $\xi \not\in E$ and $k = 1, 2, 3, \ldots$).

In view of the inequalities

$$d(\xi, E) \geq d(\xi, \tau_M[A; \|\cdot\|]) - \varepsilon \geq d(\xi, V) - \varepsilon \quad \text{(for all } \xi \not\in V)$$

we thus arrive at (2.1.b), by letting $\varepsilon$ tend to zero.

3. REMARKS AND APPLICATIONS

3.1. Special Topics

3.1.1. The $M$-numerical Range with $M = 1$. In Sections 3.1.1 and 3.1.2, $\|\cdot\|$ denotes a matrix norm on $\mathbb{C}^{n \times n}$ that is induced by some vector norm $|\cdot|$ on $\mathbb{C}^n$. Further, $A = (a_{ij})$ denotes a complex $n \times n$ matrix, and $V$ a nonempty, closed, and convex subset of $\mathbb{C}$.

We shall show that Theorem 2.1 relates $\tau_1[A; \|\cdot\|]$ to various sets considered in the literature.

In [21], the Gerschgorin domain $\mathcal{G}[A]$ of $A$ was defined by

$$\mathcal{G}[A] = \bigcap_{\theta \in \mathbb{R}} \{\zeta : \zeta \in \mathbb{C} \text{ with } \Re(e^{-i\theta}\zeta) \leq \mu_\theta(A)\},$$

where $\mu_\theta(A) = \lim_{t \to 0} t^{-1}([I + te^{-i\theta}A]-1)$. For any $\theta \in \mathbb{R}$, $\xi \in \mathbb{C}$, and $A \in \mathbb{C}^{n \times n}$, we have $\|\exp[te^{-i\theta}(A - \xi I)]\| \leq 1$ (for all $t \geq 0$) iff $\mu_\theta(A) \leq \Re(e^{-i\theta}\xi)$ (e.g. [5]). Thus, for any $V$ satisfying (2.1.c) with $M = 1$, using the fact that $V$ equals the intersection of its supporting half planes $H(\xi, \theta) = \{\zeta : \zeta \in \mathbb{C} \text{ with } \Re e^{-i\theta}(\xi - \xi) \leq 0\}$ (where $\theta$ is any normal direction to $V$ at any $\xi \in \partial V$), it can be seen that $\mathcal{G}[A] \subset V$. Further, using e.g. the fact that any half plane $\{\zeta : \zeta \in \mathbb{C} \text{ with } \Re(e^{-i\theta}\zeta) \leq \mu_\theta(A)\}$ is a supporting half plane of $\mathcal{G}[A]$ (see [21, p. 34]), one can conclude that $\mathcal{G}[A]$ is the smallest $V$ satisfying (2.1.c) with $M = 1$. 

The Hausdorff set or field of values $\mathcal{H}[A]$ of $A$ was defined in [11, p. 466] by

$$\mathcal{H}[A] = \bigcap_{\theta \in \mathbb{R}} \{ \zeta : \zeta \in \mathbb{C} \text{ with } \text{Re}(e^{-i\theta}\zeta) \leq \alpha_\theta(A) \},$$

where $\alpha_\theta(A) = \inf \{ r : r \in \mathbb{R} \text{ with } \| (A - \rho e^{i\theta})^{-1} \| \leq (\rho - r)^{-1} \text{ for all real } \rho > r \}$. In [10], the convex hull of the Bauer field of values, $\text{conv}\mathcal{F}[A]$, was proved to coincide with the smallest set $V$ satisfying (2.1.b) (with $M = 1$).

Applying Theorem 2.1, one can see that

$$\mathcal{J}[A] = \mathcal{H}[A] = \text{conv}\mathcal{F}[A] = \tau_1[A; \| \cdot \|]. \quad (3.1)$$

As an illustration of (3.1), consider the maximum norm $\| \cdot \|_\infty$ on $\mathbb{C}^n$, given by

$$\| (\xi_1, \xi_2, \ldots, \xi_n)^T \|_\infty = \max \{ |\xi_1|, |\xi_2|, \ldots, |\xi_n| \}$$

$$\left[ \text{for all } (\xi_1, \xi_2, \ldots, \xi_n)^T \in \mathbb{C}^n \right].$$

Let $\| \cdot \|_\infty$ denote the corresponding induced matrix norm on $\mathbb{C}^{n \times n}$. Denote, for $1 \leq j \leq n$, by $D_j$ the Gerschgorin disk $D_j = D(\alpha_{jj}, \sum_{k \neq j} |\alpha_{jk}|)$. In [18] it was proved that $\text{conv}\mathcal{F}[A] = \text{conv}(D_1 \cup D_2 \cup \cdots \cup D_n)$, and in [21] that $\mathcal{J}[A] = \text{conv}(D_1 \cup D_2 \cup \cdots \cup D_n)$. So, by (3.1),

$$\tau_1[A; \| \cdot \|_\infty] = \text{conv}(D_1 \cup D_2 \cup \cdots \cup D_n).$$

3.1.2. Special Choices for the Set $V$. Suppose $V$ is given by $V = \{ \zeta : \zeta \in \mathbb{C} \text{ with } \text{Re } \zeta \leq 0 \}$. In this case the equivalence of the statements (2.1.b) and (2.1.c) can be viewed as the content of the Hille-Yosida-Phillips theorem (e.g. [9, p. 67]) for finite dimensional vector spaces. Further, in this case the equivalence of (2.1.c) to (2.1.d) is related to the fact that any one-parameter bounded semigroup on a Banach space is also a one-parameter contraction semigroup with respect to some norm on the Banach space that is equivalent to the original one (e.g. [5, p. 30; 6, p. 3; 9, p. 52]).

Next suppose $V = \{ \zeta : \zeta \in \mathbb{C} \text{ with } \zeta = 0 \text{ or } |\arg(-\zeta)| \leq \frac{1}{2}\pi - \alpha \}$, where $0 < \alpha \leq \frac{1}{2}\pi$. Now the equivalence of (2.1.b) and (2.1.c) is related to a basic characterization of holomorphic semigroups (e.g. [9]).

In view of the above it is clear that part of Theorem 2.1 can be viewed as an extension to general $V$ of results in semigroup theory well known for $V$ of some special types.
3.1.3. Relating the M-Numerical Range to the Spectrum. Let $\|\cdot\|$ be an arbitrary, fixed matrix norm on $\mathbb{C}^{n\times n}$, and let $A$ be an $n \times n$ matrix. We shall relate the $M$-numerical range of $A$ to its spectrum, $\sigma[A]$.

Choose any $\gamma \in \mathbb{C}$, $\rho > 0$ such that $\text{conv} \sigma[A]$ lies in the interior of the disk $D[\gamma, \rho]$. Using (e.g. [15, p. 299]), we see that there exists an $M > 0$ for which (1.4.a) is fulfilled. With this $M$ we thus have $\tau_M[A;\|\cdot\|] \subset D[\gamma, \rho]$.

Since the intersection of all $D[\gamma, \rho]$, with $\gamma, \rho$ as above, equals $\text{conv} \sigma[A]$, we obtain, in view of (1.5.a,b),

$$\lim_{M \to \infty} \tau_M[A;\|\cdot\|] = \bigcap_{M > 0} \tau_M[A;\|\cdot\|] = \text{conv} \sigma[A]. \quad (3.2)$$

Under the assumptions of Theorem 2.1, the relation (2.1.a) implies (2.1.d). Consequently, for each $M > \|A\|$ there is a unital matrix norm $\|\cdot\|_1$ with $\tau_1[A;\|\cdot\|_1] \subset \tau_M[A;\|\cdot\|]$. In view of (3.2), (1.5.a), we thus have

$$\bigcap_{\|\cdot\|_1} \tau_1[A;\|\cdot\|_1] = \text{conv} \sigma[A], \quad (3.3)$$

where the intersection is over all unital matrix norms $\|\cdot\|_1$ on $\mathbb{C}^{n\times n}$.

We note that (3.3) also follows from results in [3, p. 23] or [18, p. 361].

3.2. A Discrete Version of the Differential Operator $d^2/dx^2$

Let $u$ be a sufficiently smooth function of $x$ on $[0, 1]$ with $u(0) = u(1) = 0$. Let $n \geq 1$ be an integer and $h = (n + 1)^{-1}$. The value of $(d^2/dx^2)u$ at $x = ih$ $(i = 1, 2, \ldots, n)$ can be approximated by $h^{-2}[u((i - 1)h) - 2u(ih) + u((i + 1)h)]$. The corresponding discretized version of the operator $d^2/dx^2$ may be represented by the $n \times n$ matrix $A = (\alpha_{ij})$, where

$$\alpha_{ii} = -2h^{-2} \quad (1 \leq i \leq n), \quad (3.4.a)$$

$$\alpha_{ij} = h^{-2} \quad (|i - j| = 1), \quad (3.4.b)$$

$$\alpha_{ij} = 0 \quad (|i - j| \geq 2). \quad (3.4.c)$$
The numerical range of such a matrix is related to stability questions in the numerical solution of partial differential equations (e.g. [2, 7, 16, 17]). The following theorem provides us with a set $V$ containing the $M$-numerical range of the above $A$ with respect to the matrix norm $\| \cdot \|_\infty$ (defined in Section 3.1.1). This theorem is used in [17] to derive various stability estimates. Moreover, at the end of this section, the theorem will be used to establish an interesting resolvent inequality for the matrix $A$.

**Theorem 3.1.** Let $\alpha$ be a real number, $0 \leq \alpha < \frac{1}{2} \pi$. Then the $n \times n$ matrix $A = (a_{ij})$, given by (3.4), satisfies the condition (2.1.a) with $\| \cdot \| = \| \cdot \|_\infty$, $M$ as in (3.6), and

$$V = D[-\rho_0 h^{-2} e^{-ia}, \rho_0 h^{-2}] \cap D[-\rho_0 h^{-2} e^{ia}, \rho_0 h^{-2}],$$

where $\rho_0$ is defined by (3.5). The quantities $M$, $\rho_0$ depend on $\alpha$, but not on $n \geq 1$.

**Proof.** Let $C^\infty$ be the space of infinite sequences of complex numbers $z = (\ldots, \xi_{-1}, \xi_0, \xi_1, \ldots)$ such that $\sup_j |\xi_j| < \infty$. The function $|z|_\infty = \sup_j |\xi_j|$ ($z \in C^\infty$) is a norm on $C^\infty$. In addition to $A$, we shall use the operator $\bar{A}: C^\infty \to C^\infty$ given by

$$(\bar{A}z)_j = h^{-2}(\xi_{j+2} - 2\xi_j + \xi_{j-1}) \quad (\text{for } j \in \mathbb{Z} \text{ and } z \in C^\infty).$$

Further, we define $E: C^n \to C^n$ by $(Ez)_j = \xi_j$ ($1 \leq j \leq n$), $(Ez)_{k(n+1)} = 0$ ($k \in \mathbb{Z}$), $(Ez)_{-j} = - (Ez)_j$ ($j \in \mathbb{Z}$), $(Ez)_j = (Ez)_{j+2n+2}$ ($j \in \mathbb{Z}$) [for $z = (\xi_1, \xi_2, \ldots, \xi_n)^T \in C^n$].

Let $0 \leq \alpha \leq \frac{1}{2} \pi$. There exists $\rho_0 \in \mathbb{R}$ such that

$$\rho_0 > 0, \quad |1 - 4\rho_0^{-1} e^{ia}| = \sin \alpha. \quad (3.5)$$

Defining $f(t) = 1 + (e^{-it} - 2 + e^{it})\rho_0^{-1} e^{ia}$, we have $f(0) = 1$, $|f(\pi)| < 1$, which implies $|f(t)| < 1$ (for all real $t$ with $0 < |t| \leq \pi$). One easily sees that $f$ satisfies the condition on the characteristic polynomial of Theorem 1 in
This implies that

$$M = \sup_{k \geq 1} \left\| \left( I + p^{-1}e^{iaA} \right)^k \right\|_{\infty}$$  \hfill (3.6)

is a finite quantity, which is independent of $n$. Here $p = \rho \rho^{-2}$, and $\| \cdot \|_{\infty}$ is the operator norm induced by the norm $| \cdot |_{\infty}$ on $\mathbb{C}^n$.

Since

$$E \left[ (p^{-1}e^{iaA} + I)z \right] = (p^{-1}e^{iaA} + I) Ez \quad \text{(for all } z \in \mathbb{C}^n),$$

we have

$$\left| (p^{-1}e^{iaA} + I)^k z \right|_{\infty} \leq M |Ez|_{\infty} = M |z|_{\infty} \quad \text{(for all } z \in \mathbb{C}^n).$$

Consequently,

$$\left\| (p^{-1}e^{iaA} + I)^k \right\|_{\infty} \leq M \quad (k = 1, 2, \ldots).$$

Similarly, we have

$$\left\| (p^{-1}e^{-iaA} + I)^k \right\|_{\infty} \leq M \quad (k = 1, 2, \ldots).$$

Hence, $D_1 = D[-\rho e^{-ia}, \rho]$ and $D_2 = D[-\rho e^{ia}, \rho]$ are suitable disks (Definition 2.2) for the matrix $A$. Therefore, (2.1.a) holds with $\| \cdot \| = \| \cdot \|_{\infty}, V = D_1 \cap D_2$, and $M$ as in (3.6).

**Corollary 3.2.** Let $\alpha \in [0, \frac{1}{2} \pi)$ be given. Then there is a constant $C_\alpha$ such that for all $n \geq 1$ the corresponding $A$ of type (3.4) satisfies the resolvent inequality

$$\left\| (\zeta I - A)^{-1} \right\|_{\infty} \leq C_\alpha |\zeta|^{-1} \quad \text{for all } \zeta \in \mathbb{C} \text{ with } \zeta \neq 0, |\text{arg}(\zeta)| < \alpha + \frac{1}{2} \pi.$$

**Proof.** This corollary follows easily from Theorems 3.1 [with $\alpha$ replaced by e.g. $\alpha' = \frac{1}{2}(\alpha + \frac{1}{2} \pi)$] and 2.1 [(2.1.a) is equivalent to (2.1.b)].
We note that the above corollary has also been proved recently by M. Crouzeix [8] by some clever direct calculations without using the theory of [22]. The corollary can be viewed as a discrete version of an important resolvent inequality that is known to hold for the differential operator $d^2/dx^2$ [19, p. 217; 14]).

The authors wish to thank M. Crouzeix for stimulating discussions on the topic of this paper.

REFERENCES


*Received 14 December 1988; final manuscript accepted 19 August 1989*