

NUMERICAL RANGES ENCIRCLED BY ANALYTIC CURVES

BRIAN LINS

(Communicated by J. Ball)

Abstract. Let D be a bounded convex domain in \mathbb{C} with a regular analytic boundary. Suppose that the numerical range $W(A)$ of a bounded linear operator A is contained in \overline{D} . If $W(A)$ intersects the boundary ∂D at infinitely many points while the essential numerical range $W_{\text{ess}}(A)$ does not intersect ∂D , then $W(A) = \overline{D}$. This generalizes some infinite dimensional analogues of a result of Anderson [2, 4].

1. Introduction

Let \mathcal{H} be a complex Hilbert space and let $\mathcal{B}(\mathcal{H})$ denote the set of bounded linear operators on \mathcal{H} . Let $S_{\mathcal{H}}$ denote the unit sphere in \mathcal{H} . For an operator $A \in \mathcal{B}(\mathcal{H})$, the numerical range of A , denoted $W(A)$, is the range of the map $f_A : S_{\mathcal{H}} \rightarrow \mathbb{C}$ defined by

$$f_A(x) := \langle Ax, x \rangle.$$

According to the Toeplitz-Hausdorff theorem, the numerical range is a convex set. It is also well known that $\overline{W(A)}$ contains the spectrum of A , $\sigma(A)$. In 1970 Anderson proved, but did not publish, the following striking theorem.

THEOREM 1.1. *If the numerical range of a matrix $A \in \mathbb{C}^{n \times n}$ is contained in the closed unit disk $\overline{\mathbb{D}}$ and $W(A)$ intersects $\partial \mathbb{D}$ at more than n points, then $W(A) = \overline{\mathbb{D}}$.*

At its heart, this theorem is an algebraic result since both the boundary of the numerical range and the unit circle are algebraic curves [6]. As recounted in [4], Anderson's original proof combined Bézout's theorem with Kippenhahn's algebraic description of the boundary curves of the numerical range.

The following infinite dimensional analogue of Anderson's theorem was proved by Gau and Wu [4].

THEOREM 1.2. *If $A \in \mathcal{B}(\mathcal{H})$ is a compact operator with $W(A) \subseteq \overline{\mathbb{D}}$ and $\overline{W(A)}$ intersects $\partial \mathbb{D}$ at infinitely many points, then $W(A) = \overline{\mathbb{D}}$.*

Mathematics subject classification (2010): Primary 47A12.

Keywords and phrases: Numerical range, essential numerical range.

The boundary of $W(A)$ need not be an algebraic curve when \mathcal{H} is infinite dimensional. Instead, the proof of Theorem 1.2 relies on an observation, first made by Narcowich [10], that the boundary of the numerical range of a compact operator consists of a countable number of regular analytic curves. Recall that a curve is *analytic* if it is the range of an analytic function γ from an interval $I \subseteq \mathbb{R}$ into \mathbb{C} . It is *regular* if $\gamma'(t) \neq 0$ for all $t \in I$.

Recently, Theorem 1.2 has been generalized to compact plus normal operators [2]. For non-compact operators, the boundary of the numerical range need not be piecewise analytic. In fact, if $N \in \mathcal{B}(\mathcal{H})$ is normal, then $\overline{W(N)}$ is the convex hull of $\sigma(N)$ [1]. Therefore one can choose N such that $\overline{W(N)}$ is any compact convex subset of \mathbb{C} . To avoid pathological cases, Birbonshi et al. place conditions on the essential spectrum in [2, Theorems 3 and 4]. The main contribution of this paper is to generalize the results of [2, 4] to arbitrary bounded operators by utilizing the essential numerical range. Note that [9] independently applies the essential numerical range to extend Theorem 1.2, although their results are limited to certain special classes of operators. Because these results are analytic rather than algebraic in nature, we also observe that any bounded convex domain with a regular analytic boundary can take the place of the unit disk.

The *essential numerical range* of A is

$$W_{\text{ess}}(A) := \bigcap \overline{W(A + K)},$$

where the intersection is taken over all compact operators $K \in \mathcal{B}(\mathcal{H})$. The essential numerical range was introduced in [14]. Several equivalent characterizations of $W_{\text{ess}}(A)$ were given in [3], one of which is the following.

$$W_{\text{ess}}(A) = \{ \lambda \in \mathbb{C} : \exists x_k \in S_{\mathcal{H}} \text{ with } x_k \xrightarrow{w} 0 \text{ and } \langle Ax_k, x_k \rangle \rightarrow \lambda \}.$$

From the definition, it is clear that $W_{\text{ess}}(A)$ is a nonempty, compact, convex subset of \mathbb{C} , unless \mathcal{H} is finite dimensional in which case $W_{\text{ess}}(A) = \emptyset$. If A is a compact operator on an infinite dimensional Hilbert space, then $W_{\text{ess}}(A) = \{0\}$. The essential numerical range $W_{\text{ess}}(A)$ contains the *essential spectrum* $\sigma_{\text{ess}}(A)$, which is the set of all $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is not Fredholm [14] (see also [3] for a discussion of alternative notions of essential spectra). When A is self-adjoint, $\sigma(A) \setminus \sigma_{\text{ess}}(A) = \sigma_{\text{disc}}(A)$ where $\sigma_{\text{disc}}(A)$ is the *discrete spectrum* of A consisting of all isolated eigenvalues of A with finite multiplicity [11, Section VII.3].

Our main result is the following theorem which generalizes [2, Theorem 3] (see Remark 2.4).

THEOREM 1.3. *Let $A \in \mathcal{B}(\mathcal{H})$ and suppose that $W(A) \subseteq \overline{D}$ where D is a convex set. Let Γ be a compact regular analytic curve contained in ∂D such that none of the tangent lines to Γ intersect $W_{\text{ess}}(A)$ and Γ is not a line segment. If $\overline{W(A)}$ intersects Γ at infinitely many points, then $\Gamma \subset W(A)$.*

An immediate corollary is the following infinite dimensional analogue of Anderson’s theorem that generalizes both Theorem 1.2 and [2, Theorem 4].

COROLLARY 1.4. *Let D be a bounded open convex subset of \mathbb{C} with a regular analytic boundary curve. Suppose that $W(A) \subseteq \overline{D}$ and $W_{\text{ess}}(A) \subset D$ for some $A \in \mathcal{B}(\mathcal{H})$. If $\overline{W(A)}$ intersects ∂D at infinitely many points, then $W(A) = \overline{D}$.*

2. Proof of main result

Before proving Theorem 1.3, we collect some known facts related to numerical ranges. For any operator $B \in \mathcal{B}(\mathcal{H})$, recall that the *real part of B* is the self-adjoint operator $\text{Re}(B) = \frac{1}{2}(B + B^*)$. The following result is well known. See, for example, [10, Theorem 3.1] and its proof.

LEMMA 2.1. *Let $A \in \mathcal{B}(\mathcal{H})$ and let $\mu(\theta)$ denote the maximum of the spectrum of $\text{Re}(e^{-i\theta}A)$ for all $\theta \in \mathbb{R}$. Then,*

$$\overline{W(A)} = \bigcap_{0 \leq \theta < 2\pi} \{z \in \mathbb{C} : \text{Re}(e^{-i\theta}z) \leq \mu(\theta)\}.$$

Furthermore, any $z \in \partial W(A)$ lies on a support line L_θ where

$$L_\theta := \{z \in \mathbb{C} : \text{Re}(e^{-i\theta}z) = \mu(\theta)\}$$

for some $\theta \in [0, 2\pi)$.

The following lemma uses the same notation as Lemma 2.1. We don't claim that this is a new result, but it doesn't appear to be widely known, so we include a proof (see also [8, Lemma 4.1]).

LEMMA 2.2. *There is a point $z \in W_{\text{ess}}(A) \cap L_\theta$ if and only if $\mu(\theta)$ is in the essential spectrum of $\text{Re}(e^{-i\theta}A)$.*

Proof. Since $W(e^{-i\theta}A) = e^{-i\theta}W(A)$ and $W_{\text{ess}}(e^{-i\theta}A) = e^{-i\theta}W_{\text{ess}}(A)$, we may assume without loss of generality that $\theta = 0$. Suppose that $z \in W_{\text{ess}}(A) \cap L_0$. There is a sequence $x_k \in S_{\mathcal{H}}$ such that x_k converges weakly to zero, and $\langle Ax_k, x_k \rangle \rightarrow z$. By Lemma 2.1, $\text{Re} \langle Ax_k, x_k \rangle = \langle \text{Re}(A)x_k, x_k \rangle$ converges to $\mu(0)$. Thus $\mu(0) \in W_{\text{ess}}(\text{Re}(A))$. For a self-adjoint operator, the essential numerical range is the convex hull of the essential spectrum [14, Corollary 1], therefore $\mu(0)$ is in the essential spectrum of $\text{Re}(A)$.

Conversely, if $\mu(0) \in \sigma_{\text{ess}}(\text{Re}(A))$, then $\mu(0) \in W_{\text{ess}}(\text{Re}(A))$ so there is a sequence $x_k \in S_{\mathcal{H}}$ such that $x_k \xrightarrow{w} 0$ and $\langle \text{Re}(A)x_k, x_k \rangle \rightarrow \mu(0)$. By passing to a subsequence, we may assume that $\langle Ax_k, x_k \rangle$ converges to some $z \in L_0$ and since the sequence x_k converges weakly to zero, $z \in W_{\text{ess}}(A)$. \square

The final crucial ingredient of our proof is the perturbation theory for eigenvalues of an analytic family of self-adjoint operators. First introduced by Rellich [12] and later refined by Kato [5] and Sz.-Nagy [16], this theory was used by Narcowich [10] to show that the boundary of the numerical range of a compact operator is a countable union of regular analytic arcs. In fact, although it is not explicitly stated, the results in

[10] imply that for $A \in \mathcal{B}(\mathcal{H})$, any closed arc of $\partial W(A)$ whose support lines do not intersect $W_{\text{ess}}(A)$ is a finite union of regular analytic curves that are contained in $W(A)$ [8, Proposition 4.2]. For a readable introduction to analytic perturbation theory and a proof of the following theorem, see [13, Section 136].

THEOREM 2.3. *Suppose that $A(t)$ is a self-adjoint operator valued function of a real parameter t given by*

$$A(t) = A^{(0)} + A^{(1)}t + A^{(2)}t^2 + \dots$$

where the coefficients $A^{(k)} \in \mathcal{B}(\mathcal{H})$ are self-adjoint and $\sum_{k \in \mathbb{N}} t^k \|A^{(k)}\|$ converges for all t in a neighborhood of 0. If $A(0)$ has an isolated eigenvalue $\lambda^{(0)}$ of finite multiplicity m , then there exists $\varepsilon > 0$ such that when $-\varepsilon < t < \varepsilon$, the spectrum of $A(t)$ in a neighborhood of $\lambda^{(0)}$ consists of m (not necessarily distinct) real values $\lambda_j(t)$ given by

$$\lambda_j(t) = \lambda^{(0)} + \lambda_j^{(1)}t + \lambda_j^{(2)}t^2 + \dots$$

Furthermore there are m corresponding orthogonal $\varphi_j(t) \in S_{\mathcal{H}}$ given by

$$\varphi_j(t) = \varphi_j^{(0)} + \varphi_j^{(1)}t + \varphi_j^{(2)}t^2 + \dots$$

such that $A(t)\varphi_j(t) = \lambda_j(t)\varphi_j(t)$ for all j and $-\varepsilon < t < \varepsilon$.

Proof of Theorem 1.3. If $\overline{W(A)}$ intersects Γ at infinitely many points, then there is a sequence of distinct points $z_k \in \overline{W(A)} \cap \Gamma$ that converge to some z . Since both $\overline{W(A)}$ and Γ are closed sets, $z \in \overline{W(A)} \cap \Gamma$. The tangent line to Γ at z is a support line for the convex set \overline{D} , and since $W(A) \subseteq \overline{D}$, it must also be a support line for $\overline{W(A)}$. We may assume without loss of generality that this line is L_0 , as defined in Lemma 2.1. By assumption, L_0 does not intersect $W_{\text{ess}}(A)$. Therefore the maximum element of the spectrum of $\text{Re}(A)$, which is $\text{Re}(z)$, is an isolated eigenvalue with finite multiplicity by Lemma 2.2.

We can now apply Theorem 2.3 to the analytic family of self-adjoint operators $\text{Re}(e^{-i\theta}A)$. There is a finite collection of real analytic functions $\lambda_j(\theta)$ defined in an ε -neighborhood of $\theta = 0$ such that $\lambda_j(0) = \text{Re}(z)$ for all j , and each $\lambda_j(\theta)$ is an eigenvalue of $\text{Re}(e^{-i\theta}A)$ for all $-\varepsilon < \theta < \varepsilon$. All other elements of the spectrum of $\text{Re}(e^{-i\theta}A)$ are strictly less than the minimum $\lambda_j(\theta)$ when $\theta \in (-\varepsilon, \varepsilon)$. To each $\lambda_j(\theta)$, there is a corresponding eigenvector $\varphi_j(\theta) \in S_{\mathcal{H}}$, and these eigenvectors are also analytic functions of θ in the interval $(-\varepsilon, \varepsilon)$. Each pair (λ_j, φ_j) has an associated analytic curve $\zeta_j(\theta) := f_A(\varphi_j(\theta))$ contained in $W(A)$. Each of the curves $\zeta_j(\theta)$ is either regular or its range is a single point [10, Lemma 3.1].

The boundary of $W(A)$ in the vicinity of z consists of either a single analytic curve corresponding to one of the $\zeta_j(\theta)$, or to two analytic curves which can either be line segments with one endpoint at z or curves that correspond to some of the $\zeta_j(\theta)$ [10, Theorem 5.1]. Since the sequence z_k approaches z along the boundary of $W(A)$, and the z_k are not contained in a line segment, we conclude that there are infinitely many z_k

contained in an arc of $\partial W(A)$ that is parametrized by one of the functions $\zeta_j(\theta)$. This means that the curve given by $\zeta_j(\theta)$ intersects Γ at infinitely many points.

Both $\zeta_j(\theta)$ and Γ are regular analytic curves, and both have a vertical tangent line at z , so we may parameterize both curves analytically in a neighborhood of z using their imaginary coordinate as the common parameter. Then, since the two analytic curves intersect infinitely many times and the intersection points accumulate at z , we conclude that the two curves are identical in a neighborhood of z (possibly one-sided if z is an endpoint of Γ). Every point of Γ in that neighborhood also lies on the curve $\zeta_j(\theta)$ for some θ , and is therefore contained in $W(A)$. We can repeat this argument at the endpoint(s) of the neighborhood where $\zeta_j(\theta)$ and Γ coincide to analytically continue the functions λ_j , φ_j , and ζ_j until the range of $\zeta_j(\theta)$ contains all of Γ . Therefore $\Gamma \subset W(A)$. \square

REMARK 2.4. If $A = N + K$ where $N \in \mathcal{B}(\mathcal{H})$ is normal and K is compact, then $W_{\text{ess}}(A) = W_{\text{ess}}(N)$. The essential numerical range of a normal operator is the convex hull of its essential spectrum, and this essential spectrum does not change when N is replaced by $N + K$ [3]. Therefore the essential numerical range of A is the convex hull of the essential spectrum of A . So Theorem 1.3 implies [2, Theorems 3 and 4].

3. Examples

The following examples illustrate the how the conditions of Theorem 1.3 and Corollary 1.4 can fail.

EXAMPLE 3.1. It is easy to construct a normal operator $A \in \mathcal{B}(\mathcal{H})$ such that $W(A)$ is contained in the closed unit circle $\overline{\mathbb{D}}$ and $W(A)$ intersects $\partial\mathbb{D}$ at infinitely many points without $W(A)$ equaling $\overline{\mathbb{D}}$. For example, take a diagonal operator on $\ell_2(\mathbb{N})$ with diagonal entries $e^{i/k}$, $k \in \mathbb{N}$. For such an operator, the accumulation point of $W(A) \cap \partial\mathbb{D}$ is contained in the essential numerical range $W_{\text{ess}}(A)$, and therefore the conditions of of Theorem 1.3 and Corollary 1.4 are not satisfied.

EXAMPLE 3.2. A weighted shift operator $S: \ell_2(\mathbb{N}) \rightarrow \ell_2(\mathbb{N})$ is defined by $Se_k := s_k e_{k+1}$ where $\{e_k\}$ is the elementary basis of $\ell_2(\mathbb{N})$ and s_k is a bounded sequence of weights. It is possible to choose the s_k so that the numerical range $W(S)$ is the open unit disk \mathbb{D} [15, Proposition 6]. In that case, $\overline{W(S)}$ intersects every point on $\partial\mathbb{D}$ but Corollary 1.4 does not apply since every point in $\partial\mathbb{D}$ is an element of the essential numerical range $W_{\text{ess}}(S)$ by [7, Theorem 1].

The requirement that Γ not be a line segment in Theorem 1.3 is essential, as the following two examples show.

EXAMPLE 3.3. Suppose that D is the convex hull of $\{0, 2, 2i\}$ and A is the 3-by-3 diagonal matrix with diagonal entries 0, 1, and i . Clearly, $W(A) \subset D$. The line segment $[0, 2]$ is a regular analytic curve contained in ∂D that intersects $W(A)$ at infinitely many points but is not contained in $W(A)$.

EXAMPLE 3.4. Consider the compact diagonal operator A on $\ell_2(\mathbb{N})$ with diagonal entries, 1 and i/k for $k \in \mathbb{N}$. Then $\overline{W(A)}$ is the convex hull of $\{0, 1, i\}$, but the line segment $[0, 1)$ is not in $W(A)$.

Acknowledgement. The author would like to thank Dr. Ilya Spitkovsky for his encouragement and many helpful suggestions.

REFERENCES

- [1] S. K. BERBERIAN, *The numerical range of a normal operator*, Duke Math. J., 31 (3): 479–483, 1964.
- [2] R. BIRBONSHI, I. M. SPITKOVSKY, AND P. D. SRIVASTAVA, *A note on Anderson's theorem in the infinite-dimensional setting*, J. Math. Anal. Appl., 461 (1): 349–353, 2018.
- [3] P. A. FILLMORE, J. G. STAMPFLI, AND J. P. WILLIAMS, *On the essential numerical range, the essential spectrum, and a problem of Halmos*, Acta Sci. Math. (Szeged), 33: 179–192, 1972.
- [4] H. GAU AND P. Y. WU, *Anderson's theorem for compact operators*, Proc. Amer. Math. Soc., 134 (11): 3159–3162, 2006.
- [5] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1995, reprint of the 1980 edition.
- [6] R. KIPPENHAHN, *Über den Wertevorrat einer Matrix*, Math. Nachr., 6: 193–228, 1951.
- [7] J. S. LANCASTER, *The boundary of the numerical range*, Proc. Amer. Math. Soc., 49: 393–398, 1975.
- [8] B. LINS AND I. M. SPITKOVSKY, *Inverse continuity of the numerical range map for Hilbert space operators*, Oper. Matrices, 14 (6): 77–90, 2020.
- [9] M. NAIMI AND M. BENHARRAT, *Anderson's theorem for some class of operators*, Khayyam Journal of Mathematics, 6 (2): 236–242, 2020.
- [10] F. J. NARCOWICH, *Analytic properties of the boundary of the numerical range*, Indiana Univ. Math. J., 29 (1): 67–77, 1980.
- [11] M. REED AND B. SIMON, *Methods of modern mathematical physics, I*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980, Functional analysis.
- [12] F. RELICH, *Störungstheorie der spektralzerlegung*, Mathematische Annalen, 113 (1): 600–619, Dec 1937.
- [13] F. RIESZ AND B. S. NAGY, *Functional Analysis*, Dover Books on Mathematics, Dover Publications, 2012.
- [14] J. G. STAMPFLI AND J. P. WILLIAMS, *Growth conditions and the numerical range in a Banach algebra*, Tôhoku Math. J. (2), 20: 417–424, 1968.
- [15] Q. F. STOUT, *The numerical range of a weighted shift*, Proc. Amer. Math. Soc., 88 (3): 495–502, 1983.
- [16] B. SZ.-NAGY, *Perturbations des transformations autoadjointes dans l'espace de Hilbert*, Comment. Math. Helv., 19: 347–366, 1946.

(Received March 11, 2020)

Brian Lins
Hampden-Sydney College
e-mail: blins@hsc.edu