THE NUMERICAL RANGE IS A \((1 + \sqrt{2})\)-SPECTRAL SET

M. CROUZEIX† AND C. PALENCIA‡

Abstract. It is shown that the numerical range of a linear operator on a Hilbert space is a (complete) \((1 + \sqrt{2})\)-spectral set. The proof relies, among other things, on the behavior of the Cauchy transform of the conjugates of holomorphic functions.

Key words. spectral set, numerical range, Crouzeix’s conjecture

AMS subject classifications. 47A25, 47A30

DOI. 10.1137/17M1116672

1. Introduction. Let us consider a smooth, bounded, convex domain \(\Omega \subset \mathbb{C}\). In a seminal paper [12], Bernard and François Delyon showed that there exists a best constant \(C_{\Omega}\) such that, for all rational functions \(f\), there holds

\[
\|f(A)\| \leq C_{\Omega} \sup_{z \in \Omega} |f(z)|,
\]

whenever \(A\) is a bounded linear operator in a complex Hilbert space \((H, \langle \cdot, \cdot \rangle, \|\cdot\|)\) whose numerical range

\[
W(A) := \{ \langle Av, v \rangle : v \in H, \|v\| = 1 \}
\]
satisfies \(\overline{W(A)} \subset \Omega\) (here and in the rest of this paper \(\overline{X}\) stands for the closure of a subset \(X \subset \mathbb{C}\)).

Their work has inspired the conjecture \([4]\) \(Q := \sup_{\Omega} C_{\Omega} = 2\), and it has been shown in [5] that \(2 \leq Q \leq 11.08\). Although there is numerical support for it [10, 11], the conjecture \(Q = 2\) remains an open problem, and we refer to [2, 3, 6, 11] and the bibliography in [6] for relevant background on the issue.

The aim of this paper is to present the improvement (see Theorem 3.1)

\[
2 \leq Q \leq 1 + \sqrt{2}.
\]

Note that, due to Mergelyan’s theorem, estimate (1) is valid, not only for rational functions \(f\), but also for any \(f\) belonging to the algebra

\[
\mathcal{A}(\Omega) := \{ f : f \text{ is holomorphic in } \Omega \text{ and continuous in } \overline{\Omega} \}.
\]

Furthermore, by using a sequence of smooth, convex domains \(\Omega_{n} \supset \overline{W(A)}\) converging to \(\overline{W(A)}\), from (2) we easily get

\[
\|f(A)\| \leq (1 + \sqrt{2}) \sup_{z \in \overline{W(A)}} |f(z)|,
\]
which shows that the numerical range \( W(A) \) is a \((1 + \sqrt{2})\)-spectral set for the operator \( A \). Furthermore, since \( C_\Omega \) is uniformly bounded, (1) is still valid for all convex domains, even for unbounded ones, which allows us to extend (3) to unbounded operators under suitable classical conditions.

For the sake of simplicity, we work with complex-valued functions, but there is no difficulty in generalizing the proof we give of (2) to matrix-valued mappings \( f \), without changing the constant. Therefore, the homomorphism \( f \mapsto f(A) \), from the algebra \( \mathcal{A}(W(A)) \) into \( B(H) \), is completely bounded by \( 1 + \sqrt{2} \) [17]. In other words, the numerical range \( W(A) \) is a complete \((1 + \sqrt{2})\)-spectral set for the operator \( A \).

Let us recall that the numerical radius \( w(B) \) of a linear operator \( B \) in the Hilbert space \( H \) is the number
\[
w(B) = \sup_{z \in W(B)} |z|.
\]
Then, given (2), the interesting result [7, Theorem 3.1] implies
\[
(4) \quad w(f(A)) \leq \sqrt{2} \sup_{z \in W(A)} |f(z)|
\]
for all rational functions bounded in \( W(A) \), an estimate which also holds in a completely bounded form. In the terminology of [7], this means that \( W(A) \) is a complete \( \sqrt{2} \)-numerical radius set for the operator \( A \).

Let us point out that our approach to (2) is based on the Cauchy transform and only uses elementary tools. In particular, we do not use dilation theory, which has shown its efficiency in the case where \( \Omega \) is a disk.

This paper is organized in two sections. Section 2 is devoted to some auxiliary lemmata; one of them (Lemma 2.1) studies the behavior of the Cauchy transform \( g \) of the conjugate of \( f \in \mathcal{A}(\Omega) \) up to the boundary of \( \Omega \), an interesting issue that is addressed in the maximum norm setting by using double layer potential theory. This lemma together with a representation for the sum \( f(A) + g(A)^* \) (Lemma 2.3) are the tools for the proof of the main result (2), presented in section 3. The proof of (2) also shows that
\[
\|f(A)\| \leq 2 \|f\|_\infty
\]
if \( f \) takes values in some sector with vertex at the origin and angle \( \pi/2 \), as commented in the final remark, Remark 3.5.

2. Auxiliary lemmata. The boundary \( \partial \Omega \) of the open, bounded, convex set \( \Omega \subset \mathbb{C} \) is assumed to be smooth. In the following, the algebra \( \mathcal{A}(\Omega) \) is provided with the norm
\[
\|f\|_\infty = \max\{ |f(z)| : z \in \overline{\Omega} \}, \quad f \in \mathcal{A}(\Omega).
\]
Besides, \( \mathcal{C}(\partial \Omega) \) stands for the set of the complex continuous functions on \( \partial \Omega \), endowed with the norm
\[
\|\varphi\|_{\partial \Omega} := \max\{ |\varphi(\sigma)| : \sigma \in \partial \Omega \}, \quad \varphi \in \mathcal{C}(\partial \Omega).
\]
For \( \sigma \in \partial \Omega \), the corresponding unit outward normal vector is denoted by \( \nu = \nu(\sigma) \) and, for \( \sigma \in \partial \Omega \) and \( z \in \mathbb{C} \setminus \{\sigma\} \), we introduce the double layer potential
\[
\mu(\sigma, z) = \frac{1}{2\pi} \left( \frac{\nu}{\sigma - z} + \frac{\nu}{\sigma - \bar{z}} \right),
\]
where \( \nu = \nu(\sigma) \). It is geometrically clear that the set
\[
\Pi_\sigma := \{ z \in \mathbb{C} : \text{Re}(\nu \cdot (\sigma - z)) > 0 \} = \{ z \in \mathbb{C} \setminus \{\sigma\} : \mu(\sigma, z) > 0 \}
\]
is the open half-plane containing $\Omega$ which is tangent to $\partial \Omega$ at point $\sigma$. Therefore, since $\Omega$ is convex, the claim $z \in \Omega$ is equivalent to saying that $\mu(\sigma, z) > 0$ for all $\sigma \in \partial \Omega$. Analogously, we also note that $\mu(\sigma, \sigma_0) \geq 0$ for $\sigma, \sigma_0 \in \partial \Omega$ such that $\sigma \neq \sigma_0$.

Furthermore, we use a counterclockwise oriented arclength parametrization $\sigma(s)$ of $\partial \Omega$. Then, $\sigma(\cdot)$ is $L$-periodic, with $L$ the length of $\partial \Omega$. It is noteworthy that $\nu(\sigma(s)) = \sigma'(s)/i$ and

$$
\frac{1}{\pi} \frac{d\arg(\sigma(s) - z)}{ds} = \mu(\sigma(s), z) \quad \forall z \neq \sigma(s),
$$

where arg stands for any continuous branch of the argument function defined in some neighborhood of $\sigma(s) - z \neq 0$. In the light of this identity, it is also geometrically clear that

$$
\int_{\partial \Omega} \mu(\sigma, z) \, ds = 2 \quad \forall z \in \Omega \quad \text{and} \quad \int_{\partial \Omega \setminus \{\sigma_0\}} \mu(\sigma, \sigma_0) \, ds = 1 \quad \forall \sigma_0 \in \partial \Omega,
$$

and in particular, $\mu(\sigma(\cdot), \sigma_0)$ is a probability density for $\sigma_0 \in \partial \Omega$.

We define the Cauchy transform of a complex function $\varphi$, defined at least on the boundary of $\Omega$ and continuous on it, as

$$
C(\varphi, z) = \frac{1}{2\pi i} \int_{\partial \Omega} \varphi(\sigma) \frac{d\sigma}{\sigma - z}, \quad \text{for } z \in \Omega.
$$

Whereas $C(\varphi, \cdot)$ is holomorphic in $\Omega$, the behavior of $C(\varphi, z)$ as $z \in \Omega$ approaches a boundary point, in general, is not clear. In Lemma 2.1 below, we address this issue when $\varphi$ is the boundary value of the conjugate of $f \in A(\Omega)$, which is the situation of interest in this paper.

**Lemma 2.1.** Assume that $f \in A(\Omega)$. Then $g = C(\bar{f}, \cdot)$ admits a continuous extension (also denoted by $g$) which belongs to $A(\Omega)$ and satisfies

$$
\|g\|_\infty \leq \|f\|_\infty.
$$

Furthermore, $g(\partial \Omega) = \{ g(\sigma) : \sigma \in \partial \Omega \}$ is contained in the convex hull $\text{conv}(\bar{f}(\partial \Omega))$ of $f(\partial \Omega) = \{ \bar{f}(\sigma) : \sigma \in \partial \Omega \}$.

**Proof.** Clearly, $g$ is holomorphic in $\Omega$. Besides, conjugating in the Cauchy formula,

$$
f(z) = \frac{1}{2\pi i} \int_{\partial \Omega} \frac{\bar{f}(\sigma)}{\sigma - z} \, d\sigma = \frac{1}{2\pi} \int_{\partial \Omega} \frac{f(\sigma)\nu(\sigma)}{\sigma - z} \, ds, \quad z \in \Omega,
$$

leads to

$$
g(z) = \int_{\partial \Omega} \bar{f}(\sigma) \mu(\sigma, z) \, ds - \bar{f}(z).
$$

According to the well-known jump formula (discovered by Gauss around 1815), if $z$ tends to $\sigma_0 \in \partial \Omega$, then $g(z)$ tends to $g(\sigma_0)$ defined on the boundary by

$$
g(\sigma_0) = \int_{\partial \Omega \setminus \{\sigma_0\}} \bar{f}(\sigma) \mu(\sigma, \sigma_0) \, ds, \quad \text{for } \sigma_0 \in \partial \Omega.
$$

It is known since Neumann [15] that with this extension the function $g$ is continuous in $\Omega$; see, for instance, [9, Theorem 3.22]. (As noticed by Neumann, the existence of the nontangential limit at boundary points follows easily while the global continuity requires careful analysis.)

Since $\mu(\sigma, \sigma_0) \, ds$ is a probability measure on $\partial \Omega$, it follows from (7) that $g(\partial \Omega)$ is contained in the closed convex hull of $\bar{f}(\partial \Omega)$ which, by compactness, coincides with $\text{conv}(\bar{f}(\partial \Omega))$. Using the maximum principle, we get $\|g\|_\infty \leq \|f\|_\infty$. \( \square \)
Remark 2.2. The first part of the previous lemma is still valid if $\Omega$ is unbounded but, generally in this case, $\mu(\sigma,A)$ is no longer a probability measure, so that the convex hull part is not guaranteed.

The rest of this section concerns the Hilbert space setting. Given a complex Hilbert space $(H,\langle \cdot,\cdot \rangle)$, both the norm on $H$ and the induced norm on the algebra $B(H)$ of bounded linear operators on $H$ are denoted by $\|\cdot\|$. For a self-adjoint operator $X$ on $H$, we use the notation $X > 0$ to indicate that $\langle Xx,x \rangle > 0$ for all $x \in H$, $x \neq 0$.

For $A \in B(H)$ and $\sigma \in \partial\Omega$ in the resolvent set of $A$, we set

$$\mu(\sigma,A) := \frac{1}{2\pi} (\nu(\sigma I - A)^{-1} + \nu(\sigma I - A^*)^{-1}).$$

If the spectrum of $A$ has no point on the boundary $\partial\Omega$, it can be seen that

$$W(A) \subset \Omega \iff \mu(\sigma,A) > 0 \quad \forall \sigma \in \partial\Omega. \quad (8)$$

Under the assumption $\overline{W(A)} \subset \Omega$, it is also meaningful to define

$$C(\varphi, A) = \frac{1}{2\pi i} \int_{\partial\Omega} \varphi(\sigma)(\sigma I - A)^{-1} d\sigma \in B(H).$$

Notice that $C(\varphi,A)$ does not correspond to $\varphi(A)$, unless $\varphi$ can be extended to some function $f \in A(\Omega)$. In this restricted case it holds that

$$C(f,A) = f(A) \quad \text{and} \quad C(\overline{f},A) = g(A),$$

with $g = C(\overline{f},\cdot)$.

Lemma 2.3. For $\varphi \in C(\partial\Omega)$, let us set

$$S(\varphi,A) = C(\varphi,A) + C(\overline{\varphi},A)^* \in B(H).$$

Then we have

$$\|S(\varphi,A)\| \leq 2\|\varphi\|_{\partial\Omega}.$$  

Proof. Let $\varphi : \partial\Omega \mapsto \mathbb{C}$ be continuous and let us assume, without loss of generality, that $\|\varphi\|_{\partial\Omega} = 1$. It follows from the definition of $C(\varphi,A)$ that

$$S(\varphi,A) = \int_{\partial\Omega} \varphi(\sigma) \mu(\sigma,A) \, ds. \quad (9)$$

The remaining part of this proof is classical, but we include it for the readers convenience. We first observe that, by the Cauchy formula, there holds

$$\int_{\partial\Omega} \mu(\sigma,A) \, ds = 2I.$$

Besides, for $\sigma \in \partial\Omega$, the operator $\mu(\sigma,A)$ is self-adjoint and positive (8). Therefore, for $x, y \in H$, we have

$$\|S(\varphi,A)x,y\| = \left| \int_{\partial\Omega} \varphi(\sigma) \langle \mu(\sigma,A)x,y \rangle \, ds \right| \leq \int_{\partial\Omega} \|\mu(\sigma,A)x\| \|y\| \, ds$$

$$\leq \int_{\partial\Omega} \langle \mu(\sigma,A)x,x \rangle^{1/2} \langle \mu(\sigma,A)y,y \rangle^{1/2} \, ds$$

$$\leq \left( \int_{\partial\Omega} \langle \mu(\sigma,A)x,x \rangle \, ds \right)^{1/2} \left( \int_{\partial\Omega} \langle \mu(\sigma,A)y,y \rangle \, ds \right)^{1/2}$$

$$= \left( \int_{\partial\Omega} \mu(\sigma,A) \, ds \right)^{1/2} \left( \int_{\partial\Omega} \mu(\sigma,A) \, ds \right)^{1/2}$$

$$= 2 \|x\| \|y\|,$$
so that
\[ \|S(f, A)\| = \sup_{\|x\| = 1, \|y\| = 1} |(S(\varphi, A)x, y)| \leq 2. \]
This concludes the proof. \( \square \)

**Remark 2.4.** In the particular case of the unit disk \( \Omega = \mathbb{D} \) and for \( f \in \mathcal{A}(\Omega) \), it follows from the Cauchy formula that \( C(f, z) = \mathcal{I}(0) \) and \( C(f, A)^{\ast} = \mathcal{I}(0) \). Then, using the representation \( f(A) = S(f, A) - \mathcal{I}(0) \), a direct application of Lemma 2.3 yields the famous Berger–Stampfli estimate [1]
\[ \|f(A)\| \leq 2\|f\|_{\infty} \quad \text{if} \quad f(0) = 0. \]
Let us point out that a further result of Okubo and Ando [16] shows that this estimate remains valid even if \( f(0) \neq 0 \). Note that the proof of Berger and Stampfli, like those of Okubo and Ando, is based on dilation theory.

**Remark 2.5.** If \( \Omega \) is unbounded, there exists a largest \( \alpha \geq 0 \) such that \( \Omega \) contains a sector of angle \( 2\alpha \). Then this lemma is still valid with the improvement
\[ \|S(\varphi, A)\| \leq 2 \frac{\pi - \alpha}{\pi} \|\varphi\|_{\partial \Omega}. \]

### 3. Main result

In this section, \( H \) will denote a complex Hilbert space, \( A \) a bounded operator on \( H \), and \( \Omega \) a bounded convex domain of \( \mathbb{C} \) with smooth boundary. As we commented in the introduction, this smoothness assumption, convenient to avoid technical difficulties in the proofs, may be easily relaxed later.

**Theorem 3.1.** The uniform bound \( C_{\Omega} \leq 1 + \sqrt{2} \) holds in (1).

**Proof.** It suffices to look at the case \( C_{\Omega} > 1 \). Then, since \( C_{\Omega} \) is the best constant, given \( 0 < \varepsilon < C_{\Omega} - 1 \), there exist \( f \in \mathcal{A}(\Omega) \), with \( \|f\|_{\infty} = 1 \), a Hilbert space \( H \) and an operator \( A \in B(H) \) with \( \mathcal{W}(A) \subset \Omega \) such that \( \lambda = \|f(A)\| \geq C_{\Omega} - \varepsilon > 1 \).

We keep the notation of section 2. Since \( f \) and \( g := C(f, \cdot) \) belong to \( \mathcal{A}(\Omega) \), there holds
\[ S := S(f, A) = C(f, A) + C(f, A)^{\ast} = f(A) + g(A)^{\ast}. \]
Besides, for the multiplication \( fg \), we also have \( fg(A) = f(A)g(A) \), so that we can write
\[ \lambda^2 I - f(A)^{\ast} f(A) = \lambda^2 I - S^{\ast} f(A) + f g(A). \]
Moreover, by Lemma 2.1, we have \( \|fg\|_{\infty} \leq 1 \), and since
\[ |\lambda^2 + fg| \geq \lambda^2 - 1 > 0, \]
we deduce that the mapping \( \lambda^2 + fg \in \mathcal{A}(\Omega) \) never vanishes. Therefore, the operator \( \lambda^2 I + fg(A) \) is invertible and we can transform (10) into
\[ \lambda^2 I - f(A)^{\ast} f(A) = (I - S^{\ast} h(A))(\lambda^2 I + fg(A)), \]
where \( h = f/(\lambda^2 + fg) \in \mathcal{A}(\Omega) \). Next, we observe that the operator \( \lambda^2 I - f(A)^{\ast} f(A) \) is singular, whence its factor \( (I - S^{\ast} h(A)) \) is also singular. Therefore, \( 1 \leq \|S^{\ast} h(A)\| \) and then, in view of Lemma 2.3 and (11), we obtain
\[ 1 \leq 2\|h(A)\| \leq 2C_{\Omega}\|h\|_{\infty} \leq \frac{2C_{\Omega}}{\lambda^2 - 1}. \]
This shows that \( (C_{\Omega} - \varepsilon)^2 = \lambda^2 \leq 2C_{\Omega} + 1 \), which, by letting \( \varepsilon \to 0^{+} \), readily yields
\[ C_{\Omega}^2 \leq 2C_{\Omega} + 1, \quad \text{and thus} \quad C_{\Omega} \leq 1 + \sqrt{2}. \] \( \square \)
Let us point out that, with minor changes, the proof of the previous theorem also shows the existence of the optimal constant $C_{\Omega} < +\infty$ (see the proof of Remark 3.5 below). Moreover, as is easily checked, Theorem 3.1 remains valid in the complete version. Therefore, since

$$\frac{1}{2} \left( 1 + \sqrt{2} + \frac{1}{1 + \sqrt{2}} \right) = \sqrt{2},$$

application of Theorem 3.1 and [7, Theorem 3.1] readily leads to the following corollary.

**Corollary 3.2.** Assume that $W(A) \subset \overline{\Omega}$. Then there holds

$$w(f(A)) \leq \sqrt{2} \|f\|_{\infty} \quad \forall f \in A(\Omega).$$

**Remark 3.3.** In principle, estimating $\|f(A)\|$ from Corollary 3.2, in a direct way, would give $\|f(A)\| \leq 2w(f(A)) \leq 2\sqrt{2}\|f\|_{\infty}$. However, Corollary 3.2 holds in its complete version indeed and then [7, Theorem 3.1] shows that the statements in Theorem 3.1 and in Corollary 3.2 are equivalent.

**Remark 3.4.** In the particular case of the unit disk $\Omega = \mathbb{D}$, Drury [8] has obtained a more accurate estimate $w(f(A)) \leq 5\|f\|_{\infty}$ (see also [14]); this constant $\frac{5}{2}$ is optimal. Previously, it was known that, if furthermore $f(0) = 0$, then $w(f(A)) \leq \|f\|_{\infty}$; this result was obtained independently by Kato [13] and by Berger and Stampfli [1].

**Remark 3.5.** For $\Sigma \subset \mathbb{C}$, set

$$A(\Omega, \Sigma) = \{ f \in A(\Omega) : f(\Omega) \subset \Sigma \}.$$  

With very few changes, the proof of Theorem 3.1 also shows that

$$\|f(A)\| \leq 2\|f\|_{\infty} \quad \forall f \in A(\Omega, \Sigma),$$

whenever $\Sigma \subset \mathbb{C}$ is a sector with vertex at the origin and angle $\pi/2$. To see this, we first fix $A$ and set

$$C_{\Omega, \Sigma} = \sup \{ \|f(A)\| : f \in A(\Omega, \Sigma), \|f\|_{\infty} = 1 \} < C_{\Omega},$$

so that we must prove that $C_{\Omega, \Sigma} \leq 2$. It suffices to consider the case $1 < C_{\Omega, \Sigma}$, and we also note that there is no loss of generality in assuming that $\Sigma$ is the sector $|\arg(z)| \leq \pi/4$.

Given $\epsilon > 0$, we can select $f \in A(\Omega, \Sigma)$ such that $\|f\|_{\infty} = 1$ and

$$\lambda = \|f(A)\| \geq C_{\Omega, \Sigma} - \epsilon > 1.$$  

Following the same steps and notation as in the proof of Theorem 3.1, we first observe that, in view of Lemma 2.3, the mapping $g = C(f, \cdot)$ also takes values in $\Sigma$. Besides, at points $z \in \Omega$ where $f(z) \neq 0$, we have

$$h(z) = \frac{f(z)}{\lambda^2 + f(z)g(z)} = \frac{|f(z)|^2}{\lambda^2 f(z) + |f(z)|^2 g(z)},$$

and since $\overline{f(z)} + |f(z)|^2 g(z) \in \Sigma$, we deduce that $h(z) \in \Sigma$. We thus conclude that $h \in A(\Omega, \Sigma)$. Furthermore, it is also clear that $\text{Re}(fg) \geq 0$, so that $|\lambda^2 + fg| \geq \lambda^2$, which gives $\|h\|_{\infty} \leq 1/\lambda^2$. 

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
Finally, since \( h \in \mathcal{A}(\Omega, \Sigma) \), we have \( \|h(A)\| \leq C_{\Omega,\Sigma} \|h\|_{\infty} \leq C_{\Omega,\Sigma}/\lambda^2 \) and, recalling that \( 1 \leq 2\|h(A)\| \), we obtain

\[
1 \leq 2\|h(A)\| \leq C_{\Omega,\Sigma}/\lambda^2 \Rightarrow \lambda^2 \leq 2C_{\Omega,\Sigma},
\]

whence

\[
(C_{\Omega,\Sigma} - \epsilon)^2 \leq \lambda^2 \leq 2C_{\Omega,\Sigma}.
\]

REFERENCES


