SIAM J. MATRIX ANAL. APPL.
Vol. 40, No. 3, pp. 1087–1101

SPECTRAL SETS: NUMERICAL RANGE AND BEYOND*

MICHEL CROUZEIX† AND ANNE GREENBAUM‡

Abstract. We extend the proof in [M. Crouzeix and C. Palencia, *SIAM J. Matrix Anal. Appl.*, 38 (2017), pp. 649–655] to show that other regions in the complex plane are K-spectral sets. In particular, we show that various annular regions are $(1+\sqrt{2})$-spectral sets and that a more general convex region with a circular hole or cutout is a $(3+2\sqrt{3})$-spectral set. We demonstrate how these results can be used to give bounds on the convergence rate of the GMRES algorithm for solving linear systems and on that of rational Krylov subspace methods for approximating $f(A)b$, where A is a square matrix, b is a given vector, and f is a function that can be uniformly approximated on such a region by rational functions with poles outside the region.

Key words. numerical range, spectral set

AMS subject classifications. 47A25, 47A30

DOI. 10.1137/18M1198417

1. Introduction.

Let us consider a closed subset $X \subset \mathbb{C}$ of the complex plane and a bounded linear operator A in a complex Hilbert space $(H, \langle , \rangle, \| \|)$. We will say that X is a K-spectral set for A if the spectrum of A is contained in X and if the inequality

$$\|f(A)\| \le K \sup_{z \in X} |f(z)| \tag{1}$$

holds for all rational functions f bounded in X. Note that $f(A)$ is naturally defined for such f since, being bounded, f has no pole in X. Let us denote by $\mathcal{A}(X)$ the set of uniform limits in X of bounded rational functions; then, by continuity, this inequality allows us to define $f(A)$ for $f \in \mathcal{A}(X)$, and inequality (1) still holds.

If X is the closure of the numerical range $W(A)$, defined by

$$W(A) := \{\langle Aq, q\rangle : q \in H,\ \|q\| = 1\}, \tag{2}$$

then it was shown in [8] that X is a $(1+\sqrt{2})$-spectral set for A. In this paper, we extend this result to show that other regions in the complex plane are K-spectral sets. In particular, we show that various annular regions are $(1+\sqrt{2})$-spectral sets and that a more general convex region with a circular hole or cutout is a $(3+2\sqrt{3})$-spectral set. We demonstrate how these results can be used to give bounds on the convergence rate of the GMRES algorithm for solving linear systems and on that of rational Krylov subspace methods for approximating $f(A)b$, where A is a square matrix, b is a given vector, and f is a function that can be uniformly approximated on such a region by rational functions with poles outside the region.

Now we consider a bounded open subset $\Omega \subset \mathbb{C}$; we assume that its boundary $\partial\Omega$ is rectifiable and has a finite number of connected components; more precisely, $\partial\Omega = \{\sigma(s) : s \in \partial\omega\}$, where $\partial\omega \subset \mathbb{R}$ is a finite union of disjoint segments $[a_j, b_j]$,

*Received by the editors July 5, 2018; accepted for publication (in revised form) by M. Embree July 12, 2019; published electronically September 12, 2019.
https://doi.org/10.1137/18M1198417

†University of Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France (michel.crouzeix@univ-rennes1.fr).

‡Department of Applied Mathematics, University of Washington, Seattle, WA 98195 (greenbau@uw.edu).

$\sigma(a_j) = \sigma(b_j)$, and s is the arc length of $\sigma(s)$ with a counterclockwise orientation. Then, if A is a bounded linear operator with spectrum $\mathrm{Sp}(A)$ contained in Ω, it follows from the Cauchy formula that (1) holds with X being the closure of Ω and $K = \frac{1}{2\pi}\int_{\partial\Omega} \|(\sigma I - A)^{-1}\|\,|d\sigma|$. But this estimate is often very pessimistic, and we are looking for a better one. For that, we start with a rational function f (bounded in Ω), and we will consider the Cauchy formulae (for $z \in \Omega$)

$$f(z) = \frac{1}{2\pi i}\int_{\partial\Omega} f(\sigma)\frac{d\sigma}{\sigma - z}, \quad f(A) = \frac{1}{2\pi i}\int_{\partial\Omega} f(\sigma)(\sigma I - A)^{-1} d\sigma.$$

We will also introduce the Cauchy transforms of the complex conjugates of f

$$g(z) := C(\overline{f}, z) := \frac{1}{2\pi i}\int_{\partial\Omega} \overline{f(\sigma)}\frac{d\sigma}{\sigma - z}, \quad g(A) := \frac{1}{2\pi i}\int_{\partial\Omega} \overline{f(\sigma)}(\sigma I - A)^{-1} d\sigma, \tag{3}$$

and finally the transforms of f by the double layer potential kernel

$$S(f, z) := \int_{\partial\omega} f(\sigma(s))\mu(\sigma(s), z)\, ds, \quad S = S(f, A) := \int_{\partial\omega} f(\sigma(s))\mu(\sigma(s), A)\, ds. \tag{4}$$

Here μ denotes the kernel[1] given by

$$\mu(\sigma(s), z) := \frac{1}{\pi}\frac{d\arg(\sigma(s) - z)}{ds} = \frac{1}{2\pi i}\Big(\frac{\sigma'(s)}{\sigma(s) - z} - \frac{\overline{\sigma'(s)}}{\overline{\sigma(s)} - \bar{z}}\Big), \tag{5}$$

$$\mu(\sigma(s), A) := \frac{1}{2\pi i}\big(\sigma'(s)(\sigma(s)I - A)^{-1} - \overline{\sigma'(s)}(\overline{\sigma(s)}I - A^*)^{-1}\big). \tag{6}$$

Note that $\sigma'(s)$ exists and $|\sigma'(s)| = 1$ for almost every $s \in \partial\omega$ since s is an arc length. These relations also assume $\sigma(s) \neq z$; thus, $\mu(\sigma(s), z)$ is defined a.e. $s \in \partial\omega$, and $\mu(\sigma(s), A)$ is defined for almost all s such that $\sigma(s)$ does not belong to the spectrum of A . Note also that $\mu(\sigma(s), z)$ is real-valued and $\mu(\sigma(s), A)$ is self-adjoint.

From these definitions, it is clear that (for $z \in \Omega$)

$$f(z) + \overline{g(z)} = S(f, z) \quad \text{and} \quad S^* = f(A)^* + g(A). \tag{7}$$

Note also that if we choose the constant function $f = 1$, then $g = 1$, $f(A) = g(A) = I$,

$$\int_{\partial\omega} \mu(\sigma(s), z)\, ds = S(1, z) = 2 \quad \text{if } z \in \Omega, \quad \text{and} \quad \int_{\partial\omega} \mu(\sigma(s), A)\, ds = S(1, A) = 2I. \tag{8}$$

We now introduce the assumption

$$\int_{\partial\omega} \mu(\sigma(s), \sigma_0)\, ds = 1 \quad \text{for almost all } \sigma_0 \in \partial\Omega, \tag{9}$$

which will be shown in section 2 to hold under a mild geometric condition on $\partial\Omega$. We will use the following lemma.

LEMMA 1. *Assume that f is a rational function satisfying $|f| \leq 1$ in Ω and that assumption* (9) *holds. Then g defined in* (3) *is holomorphic in Ω and admits a continuous extension to $\overline{\Omega}$. Furthermore, if we set*

$$c_1 := \sup\Big\{\max_{z\in\overline{\Omega}} |C(\overline{f}, z)| \;:\; f \text{ rational function, } |f| \leq 1 \text{ in } \Omega\Big\},$$

[1] Note that μ is twice the usual kernel associated to the double layer potential.

this constant satisfies

$$c_1 \leq \operatorname*{supess}_{\sigma_0\in\partial\Omega} \int_{\partial\omega} |\mu(\sigma(s),\sigma_0)|\,ds.$$

For the next theorem, we will assume that the set of uniform limits of rational functions bounded in Ω is the algebra

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

This is automatically satisfied when $\mathbb{C}\backslash\Omega$ is connected since from Mergelyan's theorem the set of polynomial functions is then dense in $\mathcal{A}(\overline{\Omega})$. In the non–simply connected case, this requires an assumption on the analytic capacity of the inner boundary curves; note that this condition is satisfied for smooth inner boundary curves [16].

THEOREM 2. *Assume that $Sp(A)\subset\Omega$ and that, for all rational functions f satisfying $|f|\leq 1$ in Ω, there exists $\gamma(f)\in\mathbb{C}$ such that $\|S(f,A)+\gamma(f)I\|\leq 2\,c_2$ with c_2 independent of f. Then $\overline{\Omega}$ is a K-spectral set for the operator A with a constant*

$$K = c_2 + \sqrt{c_2^2+c_1+\hat{\gamma}} \quad \textit{with} \quad \hat{\gamma} := \max\{|\gamma(f)|\,:\,|f|\leq 1 \textit{ in } \Omega\}.$$

One way to verify the hypotheses of Theorem 2 is to introduce

$$\lambda_{min}(\mu(\sigma,A)) = \min\{\lambda\,:\,\lambda\in \mathrm{Sp}(\mu(\sigma,A))\}$$

and use

$$\gamma = \gamma(f) := -\int_{\partial\omega} f(\sigma(s))\lambda_{min}(\mu(\sigma(s),A))\,ds.$$

Then we use the following result from [5].

LEMMA 3. *Assume that f is a rational function satisfying $|f|\leq 1$ in Ω and that $Sp(A)\subset\Omega$; then it holds that*

$$\|S(f,A)+\gamma(f)I\| \leq 2+\delta \quad \textit{with} \quad \delta = -\int_{\partial\omega}\lambda_{min}(\mu(\sigma(s),A))\,ds.$$

Thus, the hypotheses of Theorem 2 are satisfied with $c_2 = 1+\delta/2$ and $\hat{\gamma} = \int_{\partial\omega}|\lambda_{min}(\mu(\sigma(s),A))|\,ds$.

The paper is organized as follows. In section 2 we provide a proof of Lemma 1; the less obvious part is the continuity which can be found in the literature but under stronger smoothness assumptions on the boundary and in a more general context; see, for instance, [11] or Carl Neumann [13] for the original proof; we also show that assumption (9) is realized under a weak geometric condition. Section 3 contains a proof of Theorem 10 using a technique of Schwenninger [14] that was also incorporated in [5] to improve on the original result there. Section 4 gives some estimates of $\lambda_{min}(\mu(\sigma,A))$, and these are used in sections 5 and 6 to show that various annular-like regions and regions with circular cutouts are K-spectral sets and to bound the value of K. Finally, in section 7, we show how these new theoretical results provide better bounds on the convergence rate of the GMRES algorithm for solving linear systems and on that of the rational Arnoldi algorithm for approximating $f(A)b$, where A is a square matrix, b is a given vector, and f is a uniform limit of bounded rational functions on a class of regions discussed in section 5 or 6. Note that some similar arguments have been used in [2] for bounding Faber polynomials of an operator.

2. Proof of Lemma 1. Recall that f is a rational function bounded by 1 in Ω and that

$$g(z) := \frac{1}{2\pi i}\int_{\partial\Omega} \overline{f(\sigma)}\frac{d\sigma}{\sigma - z} \qquad \text{for } z \in \Omega.$$

Since f is continuous on the boundary $\partial\Omega$, it is clear that g is holomorphic in Ω. It remains to show that g has a continuous extension to $\overline{\Omega}$ and that it is bounded by c_1. For that, we will first remark that there exists a finite constant γ_f such that

$$\gamma_f = \max\{|f[z_1, z_2]| \, : z_1, z_2 \in \overline{\Omega}\}.$$

Here we use the divided difference $f[z_1, z_2] := \frac{f(z_1)-f(z_2)}{z_1-z_2}$ if $z_1 \neq z_2$, $f[z_1, z_1] := f'(z_1)$, which is a continuous function of its two variables. Now, we extend g on the boundary by setting

$$g(\sigma_0) = \int_{\partial\omega} \big(\overline{f(\sigma(s))} - \overline{f(\sigma_0)}\big)\, \mu(\sigma(s), \sigma_0)\, ds + \overline{f(\sigma_0)} \qquad \text{for } \sigma_0 \in \partial\Omega. \tag{10}$$

(a) *Proof of: the restriction of g onto $\partial\Omega$ is continuous.* Clearly, it suffices to show the continuity with respect to σ_0 of the integral part in (10). Note that $|f(\sigma(s)) - f(\sigma_0)| \leq \gamma_f|\sigma(s) - \sigma_0|$ and

$$|\mu(\sigma(s), \sigma_0)| = \frac{1}{\pi}\Big|\operatorname{Im}\frac{\sigma'(s)}{\sigma(s)-\sigma_0}\Big| \leq \frac{1}{|\pi(\sigma(s)-\sigma_0)|} \quad \text{for almost every } s$$

since, s being an arclength, $|\sigma'(s)| = 1$ a.e. $s \in \partial\omega$; the integrand being continuous with respect to σ_0 for almost every s and bounded, the continuity of g in restriction to $\partial\Omega$ follows from the dominated convergence theorem. ∎

(b) *Proof of: g is continuous in $\overline{\Omega}$.* It suffices to show that if $z_n \to \sigma_0 \in \partial\Omega$ with $z_n \in \Omega$, then $g(z_n) \to g(\sigma_0)$. For that we associate to each z_n a point $\sigma_n \in \partial\Omega$ such that

$$|z_n - \sigma_n| = \min\{|z_n - \sigma| \, : \sigma \in \partial\Omega\}.$$

Clearly, it holds that $\sigma_n \to \sigma_0$, whence, from part (a), it suffices to show that $g(z_n) - g(\sigma_n) \to 0$. Using, from (8) that $\int_{\partial\omega} \mu(\sigma(s), z)\, ds = 2$ if $z \in \Omega$, we can write

$$\begin{aligned}
g(z_n) - g(\sigma_n) &= \int_{\partial\omega} \overline{f(\sigma(s))}\; \mu(\sigma(s), z_n)\, ds - \overline{f(z_n)} - g(\sigma_n)\\
&= \int_{\partial\omega} \overline{f(\sigma(s)) - f(\sigma_n)}\; \mu(\sigma(s), z_n)\, ds + 2\overline{f(\sigma_n)} - \overline{f(z_n)} - g(\sigma_n)\\
&= \int_{\partial\omega} \overline{f(\sigma(s)) - f(\sigma_n)}\; (\mu(\sigma(s), z_n) - \mu(\sigma(s), \sigma_n))\, ds + \overline{f(\sigma_n)} - \overline{f(z_n)}.
\end{aligned}$$

It remains to observe that $f(\sigma_n) - f(z_n) \to 0$ since $\sigma_n - z_n \to 0$ and that the absolute value of the integrand is bounded by

$$\gamma_f|\sigma(s) - \sigma_n|\frac{|z_n - \sigma_n|}{\pi|\sigma(s) - \sigma_n|\,|\sigma(s) - z_n|} \leq \frac{\gamma_f}{\pi} \quad \text{for almost every } s.$$

Hence, $g(z_n) - g(\sigma_n)$ tends to zero by the dominated convergence theorem. ∎

(c) *Proof of: the bound.* From assumption (9), we can rewrite (10) as

$$g(\sigma_0) = \int_{\partial\omega} \overline{f(\sigma(s))}\, \mu(\sigma(s), \sigma_0)\, ds \qquad \text{for almost every } \sigma_0 \in \partial\Omega.$$

This implies the bound

$$|g| \leq \operatorname*{supess}_{\sigma_0 \in \partial\Omega} \int_{\partial\omega} |\mu(\sigma(s), \sigma_0)|\, ds$$

on the boundary and then in the interior by the maximum principle. ∎

We now turn to a (very weak) geometric hypothesis implying assumption (9). We will say that $\sigma_0 = \sigma(s_0)$ is a g-regular point of the boundary if the unit tangent $\sigma'(s_0)$ exists and if, for some $\delta > 0$, the following two conditions hold:

(i) $\min\{|\sigma - \sigma_0| : \sigma \in \partial\Omega,\ d_g(\sigma, \sigma_0) \geq \varepsilon\} > 0$ for all $\varepsilon > 0$;

(ii) $\sigma(s_0) + it\sigma'(s_0) \in \Omega$ for all $t \in (0, \delta)$.

Here, $d_g(\sigma, \sigma_0)$ denotes the geodesic distance from σ to σ_0 along $\partial\Omega_j$ if σ and σ_0 belong to the same connected component $\partial\Omega_j$ of the boundary, $d_g(\sigma, \sigma_0) = +\infty$ otherwise.

Remark. Recall that $\partial\omega = \cup[a_j, b_j]$. If s_0 is not an a_j or a b_j, condition (i) means that $\sigma(s) \neq \sigma(s_0)$ if $s \neq s_0$. If for some j, $s_0 = a_j$ or $s_0 = b_j$, this condition means that $\sigma(s) \neq \sigma(s_0)$ if $s \neq s_0$, $s \neq a_j$, and $s \neq b_j$. Condition (ii) means that the inward-pointing normal $\{\sigma(s_0) + it\sigma'(s_0) : t > 0\}$ is inside Ω close to $\sigma(s_0)$. These two conditions are clearly satisfied if Ω is convex and s_0 is not a corner point.

LEMMA 4. *If σ_0 is a g-regular point of the boundary $\partial\Omega$, then*

$$\int_{\partial\omega} \mu(\sigma(s), \sigma_0)\, ds = 1.$$

Proof. Without loss of generality, we may assume that $\sigma_0 = \sigma(0)$, $\sigma'(0) = 1$, $[-\varepsilon_0, \varepsilon_0] \subset \partial\omega$; thus, $d_g(\sigma(s), \sigma_0) = |s|$ if $d_g(\sigma(s), \sigma_0) \leq \varepsilon_0$. We set $z_t = it$ and $\partial\omega_\varepsilon = \partial\omega \setminus [-\varepsilon, \varepsilon]$. If $t \in (0, \delta)$, then $\sigma_0 + z_t \in \Omega$; therefore, using (8),

$$2 = \int_{\partial\omega} \mu(\sigma(s), z_t)\, ds = \int_{\partial\omega_\varepsilon} \mu(\sigma(s), z_t)\, ds + \frac{1}{\pi} \int_{-\varepsilon}^{\varepsilon} \frac{d}{ds} \arg(\sigma(s) - it)\, ds.$$

We can also assume that $s \operatorname{Re} \sigma(s) > 0$ for $0 < |s| \leq \varepsilon_0$ since $\sigma'(0) = 1$. So, we can choose continuously the determination of $\arg(\sigma(s) - it)$ in the interval $(-\frac{3\pi}{2}, \frac{\pi}{2})$. We deduce

$$\int_{-\varepsilon}^{\varepsilon} \frac{d}{ds} \arg(\sigma(s) - it)\, ds = \arg(\sigma(\varepsilon) - it) - \arg(\sigma(-\varepsilon) - it);$$

hence,

$$\lim_{t \to 0} \int_{-\varepsilon}^{\varepsilon} \frac{d}{ds} \arg(\sigma(s) - it)\, ds = \arg(\sigma(\varepsilon)) - \arg(\sigma(-\varepsilon)).$$

From the assumptions, $\alpha(\varepsilon) := \min\{|\sigma(s) - \sigma_0| : s \in \partial\omega_\varepsilon\}$ is positive; for $s \in \partial\omega_\varepsilon$ and $0 < t < \alpha(\varepsilon)/2$, it holds that $|\sigma(s) - z_t| \geq \alpha(\varepsilon)/2$, whence $|\mu(\sigma(s), z_t)| \leq \frac{1}{\pi|\sigma(s) - z_t|} \leq \frac{2}{\pi\alpha(\varepsilon)}$. This shows, using the dominated convergence theorem, that

$$\int_{\partial\omega_\varepsilon} \mu(\sigma(s), \sigma_0)\, ds = \lim_{t \to 0} \int_{\partial\omega_\varepsilon} \mu(\sigma(s), z_t)\, ds = 2 - \frac{\arg(\sigma(\varepsilon)) - \arg(\sigma(-\varepsilon))}{\pi}.$$

The lemma follows by noticing that $\lim_{\varepsilon \to 0} \arg(\sigma(\varepsilon)) = 0$ and $\lim_{\varepsilon \to 0} \arg(\sigma(-\varepsilon)) = -\pi$. ∎

Therefore, assumption (9) is realized if almost all points of $\partial\Omega$ are g-regular.

3. Proof of Theorem 2. Let A be a bounded operator satisfying $\operatorname{Sp}(A) \subset \Omega$. We set

$$K = K(A) = \sup\{\|f(A)\| : f \in \mathcal{A}(\Omega),\ |f| \leq 1 \text{ in } \Omega\}.$$

We have a first estimate $K \le \frac{1}{2\pi}\int_{\partial\Omega}\|(\sigma I{-}A)^{-1}\|\,|d\sigma|$. We have seen in (7) that $S^* = f(A)^* + g(A)$; thus,

$$(f(A)^*f(A))^2 = f(A)^*f(A)\,(S{+}\gamma(f)I)^*f(A) - f(A)^*(f(g+\overline{\gamma(f)})f)(A).$$

If we assume $|f|\le 1$ in Ω, then we can use the bounds $\|f(A)\|\le K$ and, using that g is a uniform limit of rational functions, $\|(f(g{+}\overline{\gamma(f)})f)(A)\| \le K\sup_\Omega |f(g{+}\overline{\gamma(f)})f| \le K(c_1{+}\hat\gamma)$ to get

$$\|(f(A)^*f(A))^2\| \le 2\,K^3c_2 + K^2(c_1{+}\hat\gamma),$$

whence, for the supremum, $K^4 \le 2\,c_2K^3{+}(c_1{+}\hat\gamma)K^2$, which shows that $K \le c_2{+}\sqrt{c_2^2{+}c_1{+}\hat\gamma}$.

4. Some estimates of $\boldsymbol{\lambda_{min}(\mu(\sigma, A))}$. In this section, we fix a point $\sigma_0 = \sigma(s_0)\in\partial\Omega$, where the unit tangent $\sigma_0' = \sigma'(s_0)$ exists; the half-plane $\Pi_0 = \{z\in\mathbb{C} : \operatorname{Im}\big(\sigma_0'(\overline{\sigma_0}{-}\bar z)\big)\ge 0\}$ has the same outward normal as Ω at σ_0. Note that

$$\mu(\sigma_0, A) = \frac{1}{2\pi i}\big(\sigma_0'(\sigma_0 I{-}A)^{-1} - \overline{\sigma_0'}(\overline{\sigma}_0 I{-}A^*)^{-1}\big) \tag{11}$$

depends on σ_0 and σ_0' but not on the other values of $\sigma(\cdot)$.

LEMMA 5. *Assume $W(A)\subset\Pi_0$; then $\lambda_{min}(\mu(\sigma_0, A))\ge 0$. If furthermore $\sigma_0\in\partial W(A)$, then $\lambda_{min}(\mu(\sigma_0, A)) = 0$.*

Proof. Let us consider $v\in H$, $v\ne 0$. We set $u = (\sigma_0 I{-}A)^{-1}v$ and $\alpha = \|u\|$; then $z = \langle Au, u\rangle/\alpha^2\in W(A)$, whence

$$\pi\langle\mu(\sigma_0, A)v, v\rangle = \operatorname{Im}\big(\sigma_0'\langle u, (\sigma_0 I{-}A)u\rangle\big) = \alpha^2\operatorname{Im}\big(\sigma_0'(\overline{\sigma_0}{-}\bar z)\big)\ge 0,$$

which shows that $\lambda_{min}(\mu(\sigma_0, A))\ge 0$. If furthermore $\sigma_0 = \langle Au_0, u_0\rangle\in W(A)$, $\|u_0\| = 1$, then choosing $v = (\sigma_0 I{-}A)u_0$ and thus $z = \sigma_0$, we obtain $\langle\mu(\sigma_0, A)v, v\rangle = 0$, whence $\lambda_{min}(\mu(\sigma_0, A)) = 0$. ◻

Remark. In particular, if Ω is a convex open set which contains $W(A)$, then we deduce $\mu(\sigma, A)\ge 0$ for all $\sigma\in\partial\Omega$; thus, $\|S(f, A)\|\le\big\|\int_{\partial\omega}\mu(\sigma, A)\,ds\big\| = 2$ for every f with $|f|\le 1$ in Ω. Since Ω is convex, almost all points $\sigma_0\in\partial\Omega$ are g-regular and satisfy $\mu(\sigma(s), \sigma_0) = \frac{1}{\pi}\frac{d}{ds}(\arg\sigma(s){-}\sigma_0)\ge 0$, whence $c_1 = \operatorname{supess}_{\sigma_0\in\partial\Omega}\int_{\partial\omega}\mu(\sigma(s), \sigma_0)\,ds = 1$. Therefore, we deduce from Theorem 2, used with $\gamma(f) = 0$ and $c_2 = 1$, that Ω is a K-spectral set for A with $K\le 1+\sqrt{2}$. In particular, using a decreasing sequence of convex Ω tending to $W(A)$, we recover the Crouzeix–Palencia estimate: The numerical range is a $(1{+}\sqrt{2})$-spectral set.

LEMMA 6. *Assume $|\sigma_0{-}\omega| = R$ and $\{z\in\mathbb{C} : |z{-}\omega|\le R\}\subset\Pi_0$. If $\|A{-}\omega I\|\le R$, then $\lambda_{min}(\mu(\sigma_0, A))\ge\frac{1}{2\pi R}$.*

Proof. Without loss of generality, we may assume $\omega = 0$, $\sigma_0 = R\,e^{i\theta}$, $\sigma_0' = i\,e^{i\theta}$. Then, using (11) with this value of σ_0, we may write

$$\begin{aligned}
2\pi\mu(\sigma_0, A) - \frac{1}{R}I &= e^{i\theta}(Re^{i\theta}I - A)^{-1} + e^{-i\theta}(Re^{-i\theta}I - A^*)^{-1} - \frac{1}{R}I\\
&= (Re^{i\theta}I - A)^{-1}\big[e^{i\theta}(Re^{-i\theta}I - A^*) + e^{-i\theta}(Re^{i\theta}I - A)\\
&\qquad - (1/R)(Re^{i\theta}I - A)(Re^{-i\theta}I - A^*)\big]\,(Re^{-i\theta}I - A^*)^{-1}\\
&= \frac{1}{R}(\sigma_0 I - A)^{-1}[R^2I - AA^*](\bar\sigma_0 I - A^*)^{-1}.
\end{aligned}$$

Since $\|A\|\le R$, this matrix is positive semidefinite; i.e., $\lambda_{min}(\mu(\sigma_0, A))\ge\frac{1}{2\pi R}$. ◻

Remark. In particular, if Ω is the disk $\{z \in \mathbb{C} : |z-\omega| < R\}$, where $\|A-\omega I\| \le R$, then with the notation of Lemma 3 we obtain $\delta \le -1$. It is shown in [5, section 6.1] that in this case $\|f(A)\| \le \max(1, 2+\delta)$, whence, from this lemma, we get $K = 1$. This is just the famous von Neumann inequality: Ω is a spectral set for A.

The following two lemmas will be needed for special cases in section 5. Note that if σ_0' denotes the unit tangent and if σ_0'/i is the outward normal at a boundary point σ_0 of an open set Ω, the assumption $\sigma_0-\omega = i\sigma_0'/R$ used in these lemmas means that Ω and the exterior of disk $\{z \in \mathbb{C} : |z-\omega|^{-1} \le R\}$ are tangent at σ_0 and have the same outward normal at this point.

LEMMA 7. *Assume* $\sigma_0-\omega = i\sigma_0'/R$ *with* $R > 0$. *If* $\|(A-\omega I)^{-1}\| \le R$, *then* $\lambda_{min}(\mu(\sigma_0, A)) \ge -\frac{R}{2\pi}$.

Proof. It suffices to consider the case $\omega = 0$, $\sigma_0 = r\, e^{-i\theta}$, $\sigma_0' = -i\, e^{-i\theta}$ with $r = 1/R$. Then, with $B = A^{-1}$,

$$\begin{aligned} 2\pi\mu(\sigma_0, A) + RI &= -\big(e^{-i\theta}(\sigma_0 I - A)^{-1} + e^{i\theta}(\bar\sigma_0 I - A^*)^{-1} - RI\big) \\ &= -R(\sigma_0 I - A)^{-1}(r^2 I - AA^*)(\bar\sigma_0 I - A^*)^{-1} \\ &= r(\sigma_0 I - A)^{-1} A(R^2 I - BB^*)A^*(\bar\sigma_0 I - A^*)^{-1} \ge 0 \end{aligned}$$

since $\|B\| \le R$. ∎

Define the numerical radius $w(A)$ by

$$w(A) := \sup\{|\langle Av, v\rangle| : v \in H, \|v\| = 1\}.$$

We now replace the norm of $(A - \omega I)^{-1}$ in Lemma 7 by its numerical radius.

LEMMA 8. *Assume* $\sigma_0-\omega = i\sigma_0'/R$ *with* $R > 0$. *If* $w((A-\omega I)^{-1}) \le R$, *then* $\lambda_{min}(\mu(\sigma_0, A)) \ge -\frac{R}{\pi}$.

Proof. It suffices to consider the case $\omega = 0$, $\sigma_0 = r\, e^{-i\theta}$, $\sigma_0' = -i\, e^{-i\theta}$ with $r = 1/R$. We set $B = r\, A^{-1}$; then

$$\begin{aligned} 2\pi\mu(\sigma_0, A) + 2RI &= R\big(2I - re^{-i\theta}(\sigma_0 I - A)^{-1} - re^{i\theta}(\bar\sigma_0 I - A^*)^{-1}\big) \\ &= R(\sigma_0 I - A)^{-1}(2\, AA^* - re^{-i\theta}A^* - re^{i\theta}A)(\bar\sigma_0 I - A^*)^{-1} \\ &= R(\sigma_0 I - A)^{-1} A(2\, I - e^{-i\theta}B - e^{i\theta}B^*)A^*(\bar\sigma_0 I - A^*)^{-1} \ge 0 \end{aligned}$$

since $w(B) \le 1$. ∎

5. Example 1: An annulus. We consider the annulus $\Omega = \mathcal{A}_R = \{z \in \mathbb{C} : r < |z| < R\}$ with $R > 1$, $r = 1/R$, and an invertible operator A which satisfies $\|A\| < R$ and $\|A^{-1}\| < R$. Let us denote by $\Gamma_R = \{z \in \mathbb{C} : |z| = R\}$ and $\Gamma_r = \{z \in \mathbb{C} : |z| = r\}$ the two components of the boundary. In order to lighten the notation, in the following we will write $\int_\Gamma f(s)\, ds$ in place of $\int_{\sigma(s)\in\Gamma} f(s)\, ds$. We first remark that $c_1 \le 3$; indeed,

$$\text{if } \sigma_0 \in \Gamma_R,\ \text{then } \int_{\Gamma_R} |\mu(\sigma(s), \sigma_0)|\, ds = 1 \text{ and } \int_{\Gamma_r} |\mu(\sigma(s), \sigma_0)|\, ds = \frac{4}{\pi}\arcsin(r^2) < 2;$$
$$\text{if } \sigma_0 \in \Gamma_r,\ \text{then } \int_{\Gamma_r} |\mu(\sigma(s), \sigma_0)|\, ds = 1 \text{ and } \int_{\Gamma_R} |\mu(\sigma(s), \sigma_0)|\, ds = 2.$$

These relations come from the remark that $\int_\Gamma |\mu(\sigma(s), \sigma_0)|\, ds = \frac{1}{\pi}\int_\Gamma |d\arg(\sigma(s)-\sigma_0)|$ is $\frac{1}{\pi}$ times the total variation of $\arg(\sigma(s)-\sigma_0)$ as $\sigma(s)$ runs along Γ. Note that for

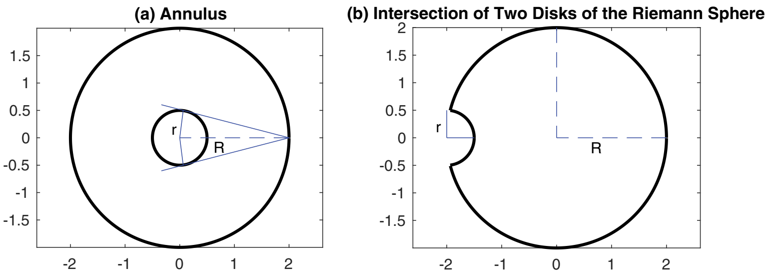

FIG. 1. (a) *Annulus with outer radius* $R > 1$ *and inner radius* $r = 1/R$. *For fixed* σ_0 *on* Γ_R, *as* $\sigma(s)$ *traverses* Γ_r, $\arg(\sigma(s) - \sigma_0)$ *goes from* 0 *to* $\arcsin(r/R)$ *back to* 0, *then to* $-\arcsin(r/R)$ *and back to* 0, *for a total variation of* $4\arcsin(r/R)$. (b) *Intersection of disk of radius* R *with exterior of disk of radius* r *centered at* $(-R, 0)$.

$\sigma_0 \in \Gamma_R$ and $\sigma(s)$ traversing Γ_r, the total variation of $\arg(\sigma(s)-\sigma_0)$ is $4\arcsin(r/R) = 4\arcsin(r^2)$ since the inner circle lies inside a cone with vertex σ_0 of angle $2\arcsin(r^2)$. This is illustrated in Figure 1(a).

In fact, we can improve on this estimate and show that $c_1 = 1$.

LEMMA 9. *For any rational function* f *bounded by* 1 *in* $\mathcal{A}_R$, *the associated function* $g(z)$ *defined by*

$$g(z) = \frac{1}{2\pi i}\int_{\partial\mathcal{A}_R} \overline{f(\sigma)}\frac{d\sigma}{\sigma - z}, \quad z \in \mathcal{A}_R, \tag{12}$$

satisfies $|g(z)| \le 1$ *in* $\mathcal{A}_R$.

Proof. Recall that g has a continuous extension to the boundary given for $\sigma_0 \in \partial\mathcal{A}_R$ by

$$g(\sigma_0) = \int_{\partial\mathcal{A}_R} \overline{f(\sigma(s))}\mu(\sigma(s),\sigma_0)\,ds \quad \text{with } \mu(\sigma(s), z) = \frac{1}{2\pi i}\left(\frac{\sigma'(s)}{\sigma(s) - z} - \frac{\overline{\sigma'(s)}}{\overline{\sigma(s)} - \bar{z}}\right),$$

where s denotes arclength on $\partial\mathcal{A}_R$. Let $f_\theta(z) = f(ze^{i\theta})$, $g_\theta(z) = g(ze^{i\theta})$, $\tilde{f}(z) = f(1/z)$, and $\tilde{g}(z) = g(1/z)$. Then it is easily verified that if we replace f by f_θ (resp., by $\tilde{f}$), the associated function g in (12) is replaced by g_θ (resp., by $\tilde{g}$). From this and the maximum principle, it suffices to show that $|f| \le 1$ in $\mathcal{A}_R$ implies $|g(r)| \le 1$. Note that

$$\mu(\sigma, r) = -\frac{R}{2\pi} \quad \text{if } \sigma = re^{-i\theta},$$

$$\mu(\sigma, r) = \frac{R}{\pi}\frac{R^2 - \cos\theta}{R^4 - 2R^2\cos\theta + 1} \quad \text{if } \sigma = Re^{i\theta}.$$

On Γ_r, we write $\sigma(s) = re^{-i\theta(s)}$, where $s = r\theta(s)$, $ds = rd\theta$, $d\sigma = -i\sigma d\theta = -i\sigma R ds$. Then

$$\begin{aligned}\overline{g(r)} &= \int_{\Gamma_R} f(\sigma(s))\mu(\sigma(s), r)\,ds - \frac{R}{2\pi}\int_{\Gamma_r} f(\sigma(s))\,ds \\ &= \int_{\Gamma_R} f(\sigma(s))\mu(\sigma(s), r)\,ds + \frac{1}{2\pi i}\int_{\Gamma_r}\frac{f(\sigma)}{\sigma}\,d\sigma.\end{aligned}$$

On Γ_R, writing $\sigma(s) = Re^{-i\theta(s)}$, where $s = R\theta(s)$, $ds = Rd\theta$, and $d\sigma = i\sigma d\theta = (i\sigma/R)ds$, and using the fact that $\int_{\partial\mathcal{A}_R} \frac{f(\sigma)}{\sigma}\, d\sigma = 0$, we obtain

$$\overline{g(r)} = \int_{\Gamma_R} f(\sigma(s))\mu(\sigma(s), r)\, ds - \frac{1}{2\pi i}\int_{\Gamma_R} \frac{f(\sigma)}{\sigma}\, d\sigma = \int_{\Gamma_R} f(\sigma(s)) \left(\mu(\sigma(s), r) - \frac{1}{2\pi R}\right) ds.$$

Finally, we remark that if $\sigma = Re^{i\theta}$, then $\mu(\sigma, r) - \frac{1}{2\pi R} = \frac{1}{2\pi R}\frac{R^4-1}{R^4-2R^2\cos\theta+1} > 0$, which implies that

$$|g(r)| \le \int_{\Gamma_R} \left(\mu(\sigma(s), r) - \frac{1}{2\pi R}\right) ds = 1. \qquad \square$$

Now, we introduce the self-adjoint operator $\nu(\sigma, A) = \mu(\sigma, A) - \frac{1}{2\pi i}\frac{\sigma'}{\sigma} I$. If $\sigma \in \Gamma_R$, we may write $\sigma = Re^{i\theta}$, $s = R\,\theta$; thus, $\nu(\sigma, A) = \mu(\sigma, A) - \frac{1}{2\pi R} I \ge 0$ follows from Lemma 5. Similarly, if $\sigma \in \Gamma_r$, we may write $\sigma = r\,e^{-i\theta}$, $s = r\,\theta$; thus, $\nu(\sigma, A) = \mu(\sigma, A) + \frac{R}{2\pi} I \ge 0$ follows from Lemma 6.

Now, we consider a rational function f bounded by 1 in Ω and note that $\int_{\partial\Omega} f(\sigma)/\sigma\, d\sigma = 0$, whence $\int_{\partial\omega} f(\sigma)(\mu(\sigma, A) - \nu(\sigma, A))\, ds = \frac{1}{2\pi i}\int_{\partial\Omega} f(\sigma)/\sigma\, d\sigma\ I = 0$. We deduce that

$$\begin{aligned}\|S(f, A)\| &= \Big\| \int_{\partial\omega} f(\sigma)\nu(\sigma(s), A)\, ds \Big\| \\ &\le \Big\| \int_{\partial\omega} \nu(\sigma(s), A)\, ds \Big\| = \Big\| \int_{\partial\omega} \mu(\sigma(s), A)\, ds \Big\| = 2,\end{aligned}$$

where we have used the fact that $S(1, A) = 2I$ from (8). We apply Theorem 2 with $c_1 = 1$, $\gamma = 0$, $c_2 = 1$ and obtain that $\mathcal{A}_R$ is a $K(R)$-spectral set for A with some optimal constant $K(R) \le 1+\sqrt{2}$.

Remark. If we assume only the weak inequalities $\|A\| \le R$ and $\|A^{-1}\| \le R$, then, for all $R' > R$, $\mathcal{A}_{R'}$ is a $(1 + \sqrt{2})$-spectral set for A; taking the limit as $R' \to R$, we obtain that $\mathcal{A}_R$ is a $(1 + \sqrt{2})$-spectral set for A.

Remark. This bound, $K(R) \le 1+\sqrt{2}$, improves the previous one given in [1]:

$$K(R) \le \min\left(2 + \frac{R+1}{\sqrt{R^2+R+1}}, \max\left(3, 2 + \sum_{n\ge1} \frac{4}{1+R^{2n}}\right)\right).$$

A first bound for this constant had been obtained by Shields [15], but this bound was unbounded for R close to 1. Note also that a lower bound is known [1]:

$$K(R) \ge 2(1-R^{-2}) \prod_{n\ge1} \left(\frac{1-R^{-8n}}{1-R^{4-8n}}\right)^2.$$

The result that $K(R) \le 1 + \sqrt{2}$ for an annulus is still true if Ω is the intersection of two disks of the Riemann sphere, as pictured in Figure 1(b).

THEOREM 10. *Let us consider $\Omega = D_1 \cap D_2$ with $D_1 = \{z \in \mathbb{C} : |z-\omega_1| < R_1\}$, $D_2 = \{z \in \mathbb{C} : |z-\omega_2| > 1/R_2\}$. If the operator A satisfies $\|A-\omega_1 I\| \le R_1$ and $\|(A-\omega_2 I)^{-1}\| \le R_2$, then Ω is a $(1 + \sqrt{2})$-spectral set for A.*

Proof. We first consider the case where $\partial D_2 \subset D_1$. Then there exist R and a Moebius function $\varphi(z) = \frac{az+b}{cz+d}$ such that φ is one to one from Ω onto $\mathcal{A}_R$, from D_1 onto $\{z \in \mathbb{C} : |z| < R\}$, and from D_2 onto $\{z \in \mathbb{C} : |z| > 1/R\}$; we set $B = \varphi(A)$.

The von Neumann inequality shows that $\|B\| \le R$ and $\|B^{-1}\| \le R$; therefore, $\mathcal{A}_R$ is a $(1+\sqrt{2})$-spectral set for B, which is clearly equivalent to Ω is a $(1+\sqrt{2})$-spectral set for A.

We now consider the case where the intersection $\partial D_1 \cap \partial D_2$ is two distinct points. Then, using $\varphi(z) = 1/(z-c)$ with $c \in \partial D_2$, $c \notin D_1$, $D_1' = \varphi(D_1)$ is some disk $\{z : |z-\gamma_1| < R_1'\}$, and $D_2' = \varphi(D_2)$ is a half-plane. From von Neumann, $B = \varphi(A)$ satisfies $\|B-\gamma_1 I\| \le R_1'$ and $W(B) \subset D_2'$; a fortiori, $W(B) \subset \Omega' := \overline{D_1' \cap D_2'}$. Since Ω' is convex, Ω' is a $(1+\sqrt{2})$-spectral set for B; thus, Ω is a $(1+\sqrt{2})$-spectral set for A.

The case where the intersection $\partial D_1 \cap \partial D_2$ is only one point follows from the case $\partial D_2 \subset D_1$ by increasing D_1 in D_1' and then letting D_1' tend to D_1. ∎

6. Example 2: Another domain with a hole or cutout. We now consider the case where $\Omega = \Omega_1 \cap \Omega_2$ is the intersection of a bounded convex domain Ω_1 with the exterior of a disk $\Omega_2 = \{z \in \mathbb{C} : |z-\omega|^{-1} < R\}$. Then, arguing as at the start of the previous section, it can be seen that $\max_{\sigma_0} \int_{\partial\Omega} |\mu(\sigma,\sigma_0)|\, ds = 3$; therefore, $c_1 \le 3$. We now assume that $\overline{W(A)} \subset \Omega_1$, $w((A-\omega I)^{-1}) < R$ and that either $\partial\Omega_2 \subset \Omega_1$ or the number of intersection points of $\partial\Omega_1$ and $\partial\Omega_2$ is finite.

Let f be a rational function bounded by 1 in Ω. We consider $\Gamma_1 = \partial\Omega_1 \cap \overline{\Omega_2}$ and $\Gamma_2 = \partial\Omega_2 \cap \overline{\Omega_1}$; then $\partial\Omega = \Gamma_1 \cup \Gamma_2$. We write $S(f,A) = S_1+S_2+S_3$ with

$$S_1 = \int_{\Gamma_1} f(\sigma(s))\mu(\sigma(s),A)\, ds, \quad S_2 = \int_{\Gamma_2} f(\sigma(s))\nu(\sigma(s),A)\, ds,$$
$$S_3 = -\frac{R}{\pi}\int_{\Gamma_2} f(\sigma(s))\, ds\, I,$$

with $\nu(\sigma,A) = \mu(\sigma,A) + \frac{R}{\pi}I$. If $\sigma \in \partial\Omega_1$, it holds that $\mu(\sigma,A) \ge 0$ since $\overline{W(A)} \subset \Omega_1$. We deduce that

$$\|S_1\| \le \left\| \int_{\Gamma_1} \mu(\sigma(s),A)\, ds \right\| \le \left\| \int_{\partial\Omega_1} \mu(\sigma(s),A)\, ds \right\| = 2,$$

where we have used the fact that $S(1,A) = 2I$ from (8). If $\sigma \in \Gamma_2$, then since $\partial\Omega_2$ is a circle, Lemma 7 shows that $\nu(\sigma,A) \ge 0$; hence,

$$\|S_2\| \le \left\| \int_{\Gamma_2} \nu(\sigma(s),A)\, ds \right\| \le \left\| \int_{\partial\Omega_2} \nu(\sigma(s),A)\, ds \right\| = \frac{R}{\pi}\int_{\partial\Omega_2} ds = 2,$$

since the assumption that $w((A-\omega I)^{-1}) < R$ ensures that the spectrum of A lies outside Ω_2 and hence that $\int_{\partial\Omega_2} \mu(\sigma(s),A)\, ds = 0$. It is clear that $\|S_3\| \le 2$.

Therefore, $c_2 \le 3$; applying Theorem 2 with $c_1 = 3$, $\gamma = 0$, $c_2 = 3$, we obtain that $\overline{\Omega}$ is a $(3+2\sqrt{3})$-spectral set for A.

In particular, we can apply this result to the annulus $\mathcal{A}_R$, but now with $c_1 = 1$, $\gamma = 0$, $c_2 = 3$: Under the assumptions $w(A) \le R$ and $w(A^{-1}) \le R$, the annulus is a $(3+\sqrt{10})$-spectral set for A. Note that this bound is larger than the one in the previous section, where R was assumed to be greater than or equal to $\max\{\|A\|, \|A^{-1}\|\}$. However, it improves, for $1 < R < 1.8837$, the previous estimates; a uniform bound was not known up to now. The previous estimates were based on the splitting

$$f(z) = f_1(z)+f_2(z) \quad \text{with} \quad f_1(z) = \sum_{n\ge 0} a_n z^n,\ f_2(z) = \sum_{n<0} a_n z^n,$$

and an estimate of $\|f_1\|_{D_1} + \|f_2\|_{D_2}$. Here $D_1 = \{z \in \mathbb{C} : |z| < R\}$, $D_2 = \{z \in \mathbb{C} : |z| > R^{-1}\}$, $\|f\|_D = \sup\{|f(z)| : z \in D\}$. From our assumptions, D_1 and D_2 are 2-spectral sets for A; therefore, $\|f(A)\| \le \|f_1(A)\| + \|f_2(A)\| \le 2(\|f_1\|_{D_1} + \|f_2\|_{D_2})$. Two estimates $\|f_1\|_{D_1} + \|f_2\|_{D_2} \le \max(3, 2+\psi(R))$ with $\psi(R) = \sum_{n\ge1} \frac{4}{R^{2n}-1}$ and $\|f_1\|_{D_1} + \|f_2\|_{D_2} \le 2 + \frac{1}{\pi}\int_0^\pi \left|\frac{R^2+e^{i\theta}}{R^2-e^{i\theta}}\right| d\theta$ follow from [7, Lemma 2.1(a) and (b)]. Therefore,

$$K(R) \le \min\left\{3+\sqrt{10},\ 4+\frac{2}{\pi}\int_0^\pi \left|\frac{R^2+e^{i\theta}}{R^2-e^{i\theta}}\right| d\theta,\ \ \max[6, 4+2\psi(R)]\right\}.$$

We summarize in the following theorem.

THEOREM 11. *Assume that $w(A) \le R$ and $w(A^{-1}) \le R$; then the annulus $\mathcal{A}_R$ is a $K(R)$-spectral set for A with*

$$\begin{aligned}
&K(R) \le 3+\sqrt{10} \simeq 6.1623 && \text{if } 1 < R \le 1.8837\ldots,\\
&K(R) \le 4 + \frac{2}{\pi}\int_0^\pi \left|\frac{R^2+e^{i\theta}}{R^2-e^{i\theta}}\right| d\theta && \text{if } 1.8837\ldots \le R \le 2.3639\ldots,\\
&K(R) \le 4+\sum_{n\ge1}\frac{8}{R^{2n}-1} && \text{if } 2.3639\ldots \le R \le 2.3912\ldots,\\
&K(R) \le 6 && \text{if } 2.3912\ldots \le R.
\end{aligned}$$

7. Some applications. The K-spectral sets derived in the previous sections can be used to give bounds on the norm of the residual in the GMRES algorithm for solving a nonsingular linear system $Ax = b$ or on the error in an approximation to $f(A)b$ generated by the rational Arnoldi algorithm.

7.1. GMRES. The GMRES algorithm generates, at each step k, an approximate solution x_k for which the 2-norm of the residual, $r_k := b - Ax_k$, is minimized over a Krylov subspace, that is,

$$\|r_k\| = \min\{\|p_k(A)r_0\| : p_k \in \mathcal{P}_k,\ p_k(0) = 1\},$$

where $\mathcal{P}_k$ is the set of polynomials of degree at most k. A bound independent of the initial residual r_0 is

$$\frac{\|r_k\|}{\|r_0\|} \le \min\{\|p_k(A)\| : p_k \in \mathcal{P}_k,\ p_k(0) = 1\}.$$

It follows from [8] that if $0 \notin W(A)$, then

$$\frac{\|r_k\|}{\|r_0\|} \le (1+\sqrt{2}) \min\left\{\max_{z\in W(A)} |p_k(z)| : p_k \in \mathcal{P}_k,\ p_k(0) = 1\right\}, \tag{13}$$

and one thus obtains a bound on the GMRES residual norm in terms of an approximation problem in the complex plane: How small can a kth-degree polynomial with value 1 at the origin be on $W(A)$?

If $W(A)$ contains the origin, however, the bound (13) is not useful since it is greater than 1, but it is always the case that $\|r_k\|/\|r_0\| \le 1$. One way to avoid this problem was devised in [6]: Note that if $B = A^{1/m}$, then for m large enough, $W(B)$ will not contain the origin, nor will the set $W(B)^m := \{z^m : z \in W(B)\}$. If $\varphi(z) = z^m$, then it follows from [8] that

$$\begin{aligned}
\|p_k(A)\| = \|p_k \circ \varphi(B)\| &\le (1+\sqrt{2}) \max\{|p_k \circ \varphi(z)| : z \in W(B)\}\\
&= (1+\sqrt{2}) \max\{|p_k(\zeta)| : \zeta \in W(B)^m\}.
\end{aligned}$$

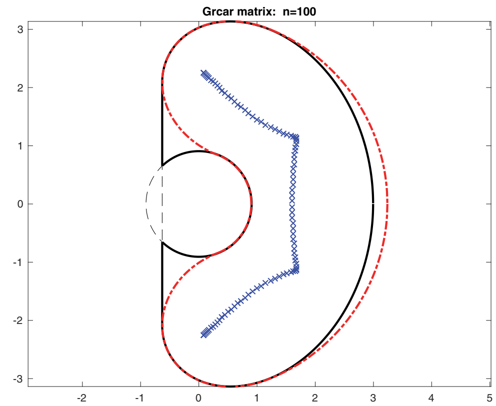

Fig. 2. *Eigenvalues (x); boundary of* $W(A)$ *and circle about* 0 *of radius* $1/R$, *where* $R = w(A^{-1})$ *(thin dashed curves); and boundary of intersection of* $W(A)$ *with exterior of disk (thick solid curve). This is a* $(3+2\sqrt{3})$*-spectral set for* A. *Also shown is the boundary of the set* $\exp(W(\log(A)))$ *(thick dash-dot curve), which was shown in* [6] *(after applying* [8]*) to be a* $(1+\sqrt{2})$*-spectral set for* A.

Unfortunately, this bound requires knowledge of $W(B)$. Also, it may happen that while $W(B)^m$ does not contain the origin, it completely surrounds the origin. Then, by the maximum principle, we still have

$$\min\left\{\max_{z\in W(B)^m} |p_k(z)| : p_k \in \mathcal{P}_k,\ p_k(0) = 1\right\} = 1.$$

A region described in section 6, consisting of the intersection of $W(A)$ and the exterior of a disk about the origin of radius $1/R$, where R is the numerical radius of A^{-1}, may provide a better bound. Note that as long as the distance from the origin to the boundary of $W(A)$ is is less than $1/R$, this set not only excludes the origin but also does not surround the origin; that is, the origin lies in the unbounded complement of this set. It follows that there is always a polynomial (of some degree) with value 1 at the origin that has maximum magnitude strictly less than 1 on the closure of this set and hence that the GMRES algorithm, applied to an infinite dimensional linear operator A satisfying $\mathrm{dist}(0, \partial W(A)) < 1/w(A^{-1})$, converges with an asymptotic rate given by $\exp(-g(0)) < 1$, where g is the Green's function of this set with pole at ∞ [9]. In Figure 2 we plot this region for the Grcar matrix[2] of order $n = 100$. As shown in section 6, this is a $(3+2\sqrt{3})$-spectral set for A. Also shown in the figure is the set $\exp(W(\log(A)))$, which was shown in [6] to be $\lim_{m\to\infty}[W(A^{1/m})]^m$ and hence (after applying the result in [8]) to be a $(1+\sqrt{2})$-spectral set for A.

7.2. Rational Arnoldi algorithm. Rational Krylov space methods (as well as standard Krylov space methods like the Arnoldi algorithm) can be used to approximate the product of a function of a matrix with a given vector: $f(A)b$. The approximation at iteration m is of the form $r_m(A)b$, where $r_m = p_{m-1}/q_{m-1}$ is a rational function with a prescribed denominator polynomial $q_{m-1} \in \mathcal{P}_{m-1}$. The rational Krylov space of order m associated with A, b, and q_{m-1} is defined as

$$\mathcal{Q}_m(A, b) := [q_{m-1}(A)]^{-1}\mathrm{span}\{b, Ab, \ldots, A^{m-1}b\}.$$

See, for example, [12] for an excellent review article.

[2] gallery('grcar',100) in MATLAB.

Let $V_m \in \mathbb{C}^{n\times m}$ be an orthonormal basis for $\mathcal{Q}_m(A,b)$. The rational Arnoldi approximation to $f(A)b$ from $\mathcal{Q}_m(A,b)$ is

$$f_m^{\mathrm{RA}} := V_m f(A_m) V_m^* b, \quad \text{where} \quad A_m := V_m^* A V_m.$$

It is shown in [12] (see also [10, 3]) that if $f(A)b$ lies in the rational Krylov subspace $\mathcal{Q}_m(A,b)$, then the rational Arnoldi approximation at step m will be exact: $f_m^{\mathrm{RA}} = f(A)b$. This is then used to show near-optimality of the rational Arnoldi approximation to $f(A)b$. Since $r_m(A)b = V_m r_m(A_m) V_m^* b$ for every rational function $r_m \in \mathcal{P}_{m-1}/q_{m-1}$, we can write

$$\begin{aligned} \|f(A)b - f_m^{\mathrm{RA}}\| &= \|f(A)b - V_m f(A_m)V_m^* b - (r_m(A)b - V_m r_m(A_m)V_m^* b)\| \\ &\le \|f(A)b - r_m(A)b\| + \|V_m(f(A_m) - r_m(A_m))V_m^* b\| \\ &\le (\|f(A) - r_m(A)\| + \|f(A_m) - r_m(A_m)\|)\|b\|. \end{aligned} \tag{14}$$

Since $W(A_m) \subset W(A)$, it follows, using the result in [8], that $W(A)$ is a $(1+\sqrt{2})$-spectral set for both A and A_m and hence that

$$\|f(A)b - f_m^{\mathrm{RA}}\| \le 2(1+\sqrt{2})\|b\| \min\left\{ \max_{z\in W(A)} |f(z) - r_m(z)| : r_m \in \mathcal{P}_{m-1}/q_{m-1} \right\}. \tag{15}$$

To use the estimate (15), the rational function r_m should have no poles in $W(A)$. But if the function f to be approximated has a pole in $W(A)$, then it would be reasonable for r_m to have one at the same point. Here the annulus of section 5, as well as the annulus or cutout region in section 6, might be useful for bounding the error in the approximation f_m^{RA}. While these regions are K-spectral sets for A, however, they might not be (with the same value of K) for A_m. The norm and numerical radius of A_m are less than or equal to those of A, but it is *not* guaranteed that the norm or numerical radius of A_m^{-1} is less than or equal to that of A^{-1}. Still, this is often the case, and assuming that it is, one can use (14) to bound the error in the rational Arnoldi approximation to $f(A)b$ in terms of the best uniform approximation to f on one of these regions.

Taking $f(z) = 1/(1-e^z)$ so that $f(A) = (I - e^A)^{-1}$, we used the RKToolkit [4] to find a rational approximation to $f(A)b$ for a random real vector b, again taking A to be the Grcar matrix of size $n = 100$. We limited the number of poles to $m-1 = 5$ and ran routine rkfit to find good pole placements for the rational Arnoldi algorithm. As expected, it returned 0 (which is inside $W(A)$) as one of the poles, and using the poles that it returned as the roots of q_{m-1}, we constructed an orthonormal basis V_m for $\mathcal{Q}_m(A,b)$ and formed the rational Arnoldi approximation $f_m^{\mathrm{RA}} = V_m f(A_m) V_m^* b$. The error $\|f(A)b - f_m^{\mathrm{RA}}\|/\|b\|$ was about $2.7e-7$. Evaluating the differences $|f(z) - \hat{r}_m(z)|$, where $\hat{r}_m$ is the rational function from routine rkfit, for z in the annulus of section 5, with outer radius $\|A\| \approx 3.2$ and inner radius $1/\|A\|$, we found the maximum difference to be about $9.4e-5$, leading to the upper bound

$$\|f(A)b - f_m^{\mathrm{RA}}\|/\|b\| \le 2(1+\sqrt{2})\ 9.4e-5.$$

The cutout region of section 6, which is the intersection of $W(A)$ with the exterior of a disk of radius $1/w(A^{-1}) \approx 0.9$, provides a better bound. The maximum value of $|f(z) - \hat{r}_m(z)|$ on this set was about $2.7e-6$, leading to the error bound

$$\|f(A)b - f_m^{\mathrm{RA}}\|/\|b\| \le 2(3+2\sqrt{3})\ 2.7e-6,$$

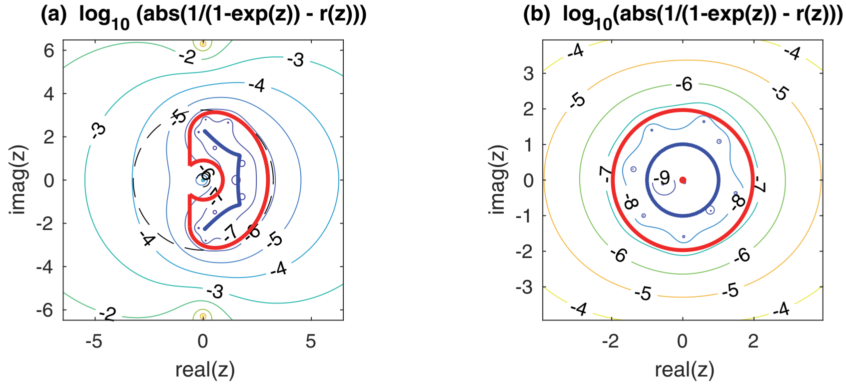

FIG. 3. *Contour plot of* $|f(z) - \hat{r}_m(z)|$ *for* (a) *the Grcar matrix and* (b) *the 'smoke' matrix. Dashed curve in* (a) *is boundary of annulus with outer radius* $\|A\|$, *inner radius* $1/\|A\|$. *Thick solid curve in each plot is boundary of intersection of* $W(A)$ *with exterior of disk about* 0 *of radius* $1/w(A^{-1})$. *Eigenvalues are marked with dots inside these regions. (Note that in* (b) *the hole in* $W(A)$ *is the tiny circle in the middle, while the dotted eigenvalues lie on the middle circle.)*

which must hold for every vector b (provided that A_m satisfies $w(A_m^{-1}) \leq w(A^{-1})$, as it did in this case, so that this region is also a $(3+2\sqrt{3})$-spectral set for A_m). (Here we assume that the same pole placement, that is, the same denominator polynomial q_{m-1}, is used for each b.) A contour plot of $|f(z) - \hat{r}_m(z)|$ is shown in Figure 3(a), along with the annulus of section 5 and the cutout region of section 6. While $\hat{r}_m$ is *not* the best uniform approximation to f on either of these regions, it is small enough to provide a reasonable upper bound for the error in the rational Arnoldi approximation.

As another example, again taking $f(z) = 1/(1-e^z)$ but now taking A to be the matrix generated in MATLAB by typing 'gallery('smoke',100)' (which is a 100 by 100 matrix with 1's on the superdiagonal, a 1 in position $(100, 1)$, and powers of roots of unity along the diagonal), we ran routine rkfit to find $m - 1 = 5$ poles to use in a rational Arnoldi approximation to $f(A)b$, and it again returned 0 (which is inside $W(A)$) as one of the poles. Using these poles in the rational Arnoldi algorithm with a random vector b led to an error $\|f(A)b - f_m^{\mathrm{RA}}\|/\|b\| \approx 2.0e-8$. While the difference $|f(z) - \hat{r}_m(z)|$ was large on the annulus of section 5, with outer radius $\max\{\|A\|, \|A^{-1}\|\} \approx 31.8$ and inner radius the reciprocal of this, it was small on the region of section 6 consisting of the intersection of $W(A)$ and the exterior of a disk about 0 of radius $1/w(A^{-1}) \approx 0.04$. Now this disk was a subset of $W(A)$, so this was a different region with a hole in it. The maximum value of $|f(z) - \hat{r}_m(z)|$ on this region was about $7.9e-8$, leading to the bound

$$\|f(A)b - f_m^{\mathrm{RA}}\|/\|b\| \leq 2(3+2\sqrt{3})\ 7.9e-8,$$

which holds for all b (again assuming that $w(A_m^{-1}) \leq w(A^{-1})$, as it was in this case). Figure 3(b) shows a contour plot of $|f(z) - \hat{r}_m(z)|$.

Acknowledgments. The authors thank the American Institute of Mathematics (AIM) for hosting a Workshop on Crouzeix's Conjecture at which many of the ideas in this paper originated. They also thank the referees for their careful reading of the manuscript and their helpful comments and suggestions. Finally, they thank Mark Embree for some excellent suggestions to improve the presentation.

REFERENCES

[1] C. Badea, B. Beckermann, and M. Crouzeix, *Intersections of several disks of the Riemann sphere as K-spectral sets*, Commun. Pure Appl. Anal., 8 (2009), pp. 37–54.

[2] B. Beckermann and M. Crouzeix, *Faber polynomials of matrices for non-convex sets*, Jaen J. Approx., 6 (2014), pp. 219–231.

[3] B. Beckermann and L. Reichel, *Error estimates and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.

[4] M. Berljafa, S. Elsworth, and S. Güttel, *Rational Krylov Toolbox for MATLAB*, http://guettel.com/rktoolbox.

[5] T. Caldwell, A. Greenbaum, and K. Li, *Some extensions of the Crouzeix-Palencia result*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 769–780.

[6] D. Choi and A. Greenbaum, *Roots of matrices in the study of GMRES convergence and Crouzeix's conjecture*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 289–301.

[7] M. Crouzeix, *The annulus as a K-spectral set*, Commun. Pure Appl. Anal., 11 (2012), pp. 2291–2303.

[8] M. Crouzeix and C. Palencia, *The numerical range is a* $(1+\sqrt{2})$*-spectral set*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 649–655.

[9] T. A. Driscoll, K. C. Toh, and L. N. Trefethen, *From potential theory to matrix iterations in six steps*, SIAM Rev., 40 (1998), pp. 547–578.

[10] V. Druskin and L. Knizhnerman, *Extended Krylov subspaces: Approximation of the matrix square root and related functions*, SIAM J. Matrix Anal. Appl., 19, pp. 755–771.

[11] G. B. Folland, *Introduction to Partial Differential Equations*, Princeton University Press, Princeton, NJ, 1995.

[12] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM-Mitt., 36 (2013), pp. 8–31.

[13] C. Neumann, *Untersuchungen über das logarithmische und Newton'sche Potential*, Math. Ann., 13 (1878), pp. 255–300.

[14] T. Ransford and F. L. Schwenninger, *Remarks on the Crouzeix-Palencia proof that the numerical range is a* $(1+\sqrt{2})$*-spectral set*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 342–345.

[15] A. L. Shields, *Weighted Shift Operators and Analytic Function Theory*, in Topics in Operator Theory Math. Surveys 13, American Mathematical Society, Providence, RI, 1974, pp. 49–128.

[16] A. G. Vitushkin, *The analytic capacity of sets in problems of approximation theory*, Russian Math. Surveys, 22 (1967), pp. 139–200.