SIAM J. Matrix Anal. Appl.
Vol. 21, No. 1, pp. 145–165

THE KREISS MATRIX THEOREM ON A GENERAL COMPLEX DOMAIN*

KIM-CHUAN TOH† AND LLOYD N. TREFETHEN‡

Abstract. Let A be a bounded linear operator in a Hilbert space $\mathcal{H}$ with spectrum $\Lambda(A)$. The Kreiss matrix theorem gives bounds based on the resolvent norm $\|(zI-A)^{-1}\|$ for $\|A^n\|$ if $\Lambda(A)$ is in the unit disk or for $\|e^{tA}\|$ if $\Lambda(A)$ is in the left half-plane. We generalize these results to a complex domain Ω, giving bounds for $\|F_n(A)\|$ if $\Lambda(A)\subset\Omega$, where F_n denotes the nth Faber polynomial associated with Ω. One of our bounds takes the form

$$\tilde{\mathcal{K}}(\Omega) \;\le\; 2\sup_n \|F_n(A)\|, \qquad \|F_n(A)\| \;\le\; 2\,e\,(n+1)\,\tilde{\mathcal{K}}(\Omega),$$

where $\tilde{\mathcal{K}}(\Omega)$ is the "Kreiss constant" defined by

$$\tilde{\mathcal{K}}(\Omega) = \inf\left\{ C \,:\, \|(zI-A)^{-1}\| \;\le\; C/\mathrm{dist}(z,\Omega)\ \forall\, z\notin\Omega\right\}.$$

By means of an inequality due originally to Bernstein, the second inequality can be extended to general polynomials p_n. In the case where $\mathcal{H}$ is finite-dimensional, say, $\dim(\mathcal{H}) = N$, analogous results are also established in which $\|F_n(A)\|$ is bounded in terms of N instead of n when the boundary of Ω is twice continuously differentiable.

Key words. Kreiss matrix theorem, conformal mapping, Faber polynomials, polynomials of a matrix, Krylov subspaces

AMS subject classifications. 65F35, 65F10, 15A60, 30E10

PII. S0895479897324020

1. Introduction.

Let A be a bounded linear operator in a Hilbert space $\mathcal{H}$ with spectrum $\Lambda(A)$, and let $\|\cdot\|$ denote the operator 2-norm. The Kreiss matrix theorem, originally published in 1962 [9], concerns the problem of characterizing families of bounded linear operators that are uniformly power-bounded, with spectra contained in the closed unit disk D. Let us define

$$\mathcal{K}(D) = \inf\left\{ C \,:\, \|(zI-A)^{-1}\| \;\le\; \frac{C}{\mathrm{dist}(z,D)}\ \ \forall\, z\notin D\right\},$$

the *Kreiss constant* of A with respect to D. If $\mathcal{H}$ is finite-dimensional, say, $\dim(\mathcal{H}) = N$, the current, sharp form of the theorem reads as follows [17], [20]:

$$\mathcal{K}(D) \;\le\; \sup_n \|A^n\| \;\le\; e\,N\,\mathcal{K}(D). \tag{1}$$

Since its original appearance, the Kreiss matrix theorem has been one of the fundamental results for establishing numerical stability of discrete evolution process $\{\|A^n\|\}$.

When $\mathcal{H}$ is infinite-dimensional, unfortunately, the upper bound of (1) becomes vacuous. But there is another upper bound in the following (more elementary) variant of the Kreiss matrix theorem:

$$\|A^n\| \le e\,(n+1)\,\mathcal{K}(D). \tag{2}$$

*Received by the editors July 7, 1997; accepted for publication (in revised form) by L. Reichel August 31, 1998; published electronically August 26, 1999.
http://www.siam.org/journals/simax/21-1/32402.html

†Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260 (mattohkc@math.nus.edu.sg).

‡Oxford University Computing Laboratory, Wolfson Bldg., Pashs Road, Oxford OX1 3QD, UK (lnt@comlab.ox.ac.uk).

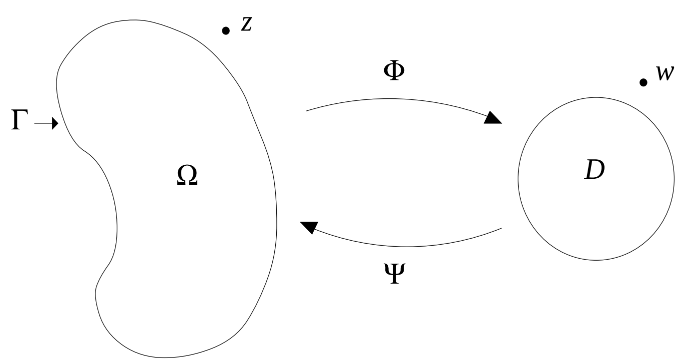

FIG. 1. *Conformal maps between the exterior of* Ω *and the exterior of the closed unit disk* D. Γ *is the boundary of* Ω.

In this paper, we generalize both (1) and (2) to a general complex domain Ω, giving bounds for $\|F_n(A)\|$ if $\Lambda(A) \subset \Omega$, where F_n is the nth Faber polynomial associated with Ω. On the unit disk D, the Faber polynomials are just the monomials z^n, and on the interval $[-1, 1]$, or on any ellipse with foci ± 1, they are twice the Chebyshev polynomials $T_n(z)$.

Let Ω be a compact set (possibly having empty interior) with boundary Γ in the complex plane $\mathbb{C}$ that contains $\Lambda(A)$, and assume that its complement Ω^c is simply connected in the extended complex plane. (This will be our assumption throughout this paper, except section 2.) By the Riemann mapping theorem [6], there exists a unique conformal map normalized by $\Phi(\infty) = \infty$ and $\Phi'(\infty) > 0$,

$$w \;=\; \Phi(z) \;=\; dz + d_0 + \sum_{k=1}^{\infty} \frac{d_k}{z^k} \quad (d > 0), \quad z \in \Omega^c,$$

from Ω^c to the exterior of the closed unit disk D (Figure 1). If $\Psi : D^c \longrightarrow \Omega^c$ is the inverse map of Φ, then Ψ has a similar expansion

$$z \;=\; \Psi(w) \;=\; cw + c_0 + \sum_{k=1}^{\infty} \frac{c_k}{w^k}, \qquad w \in D^c,$$

where $c = 1/d$. The positive constant c is known as the *logarithmic capacity* of Ω. For each nonnegative integer n, the polynomial part of $[\Phi(z)]^n$ is a polynomial of degree n. This polynomial is known as the nth Faber polynomial associated with Ω and is generally denoted by $F_n(z)$. For more details on Faber polynomials, see [2], [6], [16].

We define the Kreiss constant of A with respect to the region Ω by the following formula:

$$\mathcal{K}(\Omega) = \inf \left\{ C \,:\, \|(zI - A)^{-1}\| \;\le\; \frac{C\,|\,\Phi'(z)\,|}{|\,\Phi(z)\,| - 1} \;\forall\, z \in \Omega^c \right\}. \tag{3}$$

Our generalizations of the Kreiss matrix theorem read as follows.

THEOREM 1.1 (Kreiss matrix theorem I). *Suppose* A *is a bounded linear operator in a Hilbert space with spectrum* $\Lambda(A) \subset \Omega$ *and* $\mathcal{K}(\Omega) < \infty$. *Then* $\forall n \ge 0$,

$$\|F_n(A)\| \le e\,(n+1)\,\mathcal{K}(\Omega). \tag{4}$$

Conversely, if $\sup_{n\geq 0} \|F_n(A)\| < \infty$*, then* $\Lambda(A) \subset \Omega$*,* $\mathcal{K}(\Omega)$ *is finite, and*

$$\mathcal{K}(\Omega) \leq \sup_{n\geq 0} \|F_n(A)\|. \tag{5}$$

Theorem 1.2 (Kreiss matrix theorem II). *Suppose* A *is bounded linear operator in a Hilbert space* $\mathcal{H}$ *with* $\Lambda(A) \subset \Omega$ *and* $\mathcal{K}(\Omega) < \infty$*. If* $\mathcal{H}$ *is finite-dimensional, say,* $\dim(\mathcal{H}) = N$*, and the boundary of* Ω *is twice continuously differentiable, then* $\forall n \geq 0$*,*

$$\|F_n(A)\| \leq C_\Omega \, e \, N \, \mathcal{K}(\Omega), \tag{6}$$

where the constant C_Ω *depends only on* Ω*.*

Remark. If the boundary of Ω is not smooth, then (6) holds with the constant C_Ω replaced by a constant of the form $C_\Omega\,(1 + \log n/N)$. We do not know if this latter bound is sharp.

Note that if Ω is the unit disk, with $F_n(A)$ equal to A^n, then we get back the standard results (1) and (2) from Theorems 1.1 and 1.2, aside from the introduction of an unspecified constant in Theorem 1.2.

In applications such as Krylov subspace iterations [4], we may not know the Faber polynomials for any particular region Ω of interest. Fortunately, this is not an essential restriction. By means of an inequality due to Bernstein in 1912, Theorem 1.1 can be extended to general polynomials p_n. We have the following theorem.

Theorem 1.3 (Kreiss matrix theorem for p_n). *Suppose* A *is a bounded linear operator in a Hilbert space with spectrum* $\Lambda(A) \subset \Omega$ *and* $\mathcal{K}(\Omega) < \infty$*. Then for any polynomial* p_n *of any degree* n*,*

$$\|p_n(A)\| \leq e\,(n+1)\,\mathcal{K}(\Omega)\;\|p_n\|_\Omega. \tag{7}$$

Given the simplicity of the definition of the Kreiss constant with respect to the unit disk, one may wonder why we did not define the Kreiss constant for A with respect to Ω simply by

$$\tilde{\mathcal{K}}(\Omega) = \inf\left\{ C \,:\, \|(zI - A)^{-1}\| \;\leq\; \frac{C}{\operatorname{dist}(z,\Omega)} \;\forall\, z \in \Omega^c \right\}. \tag{8}$$

The answer is that, as we shall see in sections 4 and 5, it is the constant $\mathcal{K}(\Omega)$ that is mathematically the most natural generalization of $\mathcal{K}(D)$. On the other hand, the constant $\tilde{\mathcal{K}}(\Omega)$ has undeniable appeal because of its geometric simplicity. In particular, $\tilde{\mathcal{K}}(\Omega)$ does not involve the conformal map Φ. We shall see that although we have two different versions of the Kreiss constant, they are equivalent in the sense that they differ by at most a factor of 2 (in either direction; see Theorem 3.2). If Ω is the unit disk itself, or any other disk, then $\tilde{\mathcal{K}}(\Omega) = \mathcal{K}(\Omega)$ for any A.

Besides being more memorable and intuitive, for most practical purposes, it is the constant $\tilde{\mathcal{K}}(\Omega)$ that is easier to use. Thus it is worthwhile for the reader to bear in mind that the Kreiss matrix theorems 1.1, 1.2, and 1.3 have a more memorable version where the constant $\mathcal{K}(\Omega)$ can be replaced by $2\tilde{\mathcal{K}}(\Omega)$ or $\frac{1}{2}\tilde{\mathcal{K}}(\Omega)$, depending on whether it appears as part of an upper bound or lower bound.

The paper is organized as follows. In section 2, we discuss some properties of $\tilde{\mathcal{K}}(\Omega)$. In section 3, we show the equivalence of the Kreiss constants $\mathcal{K}(\Omega)$ and $\tilde{\mathcal{K}}(\Omega)$. In section 4, we prove Theorems 1.1 and 1.3. In section 5, we prove Theorem 1.2. In section 6, we consider the special case $\Omega = [-1, 1]$, obtaining a Kreiss matrix theorem for the Chebyshev polynomials T_n. An example is also given to show that the linear dependence on N in the upper bound of this Kreiss matrix theorem is sharp.

Throughout, we write $\mathcal{K}(\Omega)$ simply as $\mathcal{K}$ when there is no danger of confusion.

2. Some properties of the Kreiss constant $\tilde{\mathcal{K}}(\Omega)$. In this section, Ω will denote any compact subset of the complex plane. Note that $\tilde{\mathcal{K}}(\Omega)$ remains well defined without the assumption that Ω^c is simply connected in the extended complex plane. In particular, it is well defined for $\Omega = \Lambda(A)$.

The Kreiss constant $\tilde{\mathcal{K}}(\Omega)$, being defined geometrically, can be computed without the knowledge of the conformal map Φ. This is a great advantage over $\mathcal{K}(\Omega)$ since, except for special domains such as ellipses or polygons, Φ cannot be computed easily. In practice, one might want to choose Ω to be an ϵ-pseudospectrum [19] of A for some ϵ, in which case Ω is rarely one of these special domains.

Besides this computational advantage, we shall see in this section that $\tilde{\mathcal{K}}(\Omega)$ has some appealing geometrical properties. The most important of these is that it has a geometrical interpretation: It measures the largest amount by which some ϵ-pseudospectrum of A protrudes outside the region Ω, relative to ϵ. We discuss this in Proposition 2.1. Also, $\tilde{\mathcal{K}}(\Omega) \geq 1$ for all Ω and A, since in the limit $z \to \infty$, $\|(zI - A)^{-1}\|$ and $\text{dist}(z, \Omega)$ tend to $|z|^{-1}$ and $|z|$, respectively.

PROPOSITION 2.1. *For each $\epsilon \geq 0$, let $\Lambda_\epsilon(A)$ be the ϵ-pseudospectrum of A defined by*

$$\Lambda_\epsilon(A) \; = \; \{z \in \mathbb{C} \, : \, \|(zI - A)^{-1}\| \geq 1/\epsilon\}.$$

Then the following statements are equivalent:

1. $\|(zI - A)^{-1}\| \; \leq \; C/\text{dist}(z, \Omega) \; \forall \; z$ *such that* $\text{dist}(z, \Omega) > 0$;
2. $\text{dist}(z, \Omega) \; \leq \; C\,\epsilon \; \forall \, z \in \Lambda_\epsilon(A), \; \epsilon \geq 0$;
3. $\Lambda_\epsilon(A) \; \subset \; \Omega + C\,D_\epsilon \; \forall \; \epsilon \geq 0$, *where D_ϵ is the closed disk with radius ϵ.*

Therefore, an equivalent definition for $\tilde{\mathcal{K}}(\Omega)$ is

$$\tilde{\mathcal{K}}(\Omega) = \sup_{\epsilon > 0} \frac{\text{dist}(\Lambda_\epsilon(A), \Omega)}{\epsilon}, \tag{9}$$

where

$$\text{dist}(\Lambda_\epsilon(A), \Omega) \;\; = \;\; \max_{z \in \Lambda_\epsilon(A)} \; \text{dist}(z, \Omega).$$

Proof. This is an easy consequence of the definitions of $\tilde{\mathcal{K}}(\Omega)$ and $\Lambda_\epsilon(A)$. ∎

Other appealing properties of $\tilde{\mathcal{K}}(\Omega)$ include monotonicity and the fact that in the special case when A is normal and $\Omega = \Lambda(A)$ or Ω is the field of values $\mathcal{F}(A)$, $\tilde{\mathcal{K}}(\Omega)$ is equal to 1. (The field of values $\mathcal{F}(A)$ is the set $\{x^*Ax/x^*x : x \in \mathbb{C}^N, x \neq 0\}$.)

PROPOSITION 2.2. *Suppose Ω_1 and Ω_2 are subsets of $\mathbb{C}$ such that $\Lambda(A) \subset \Omega_1 \subset \Omega_2$. Then*

$$\tilde{\mathcal{K}}(\Omega_2) \leq \tilde{\mathcal{K}}(\Omega_1).$$

Proof. This result follows from the definitions of $\tilde{\mathcal{K}}(\Omega_1)$ and $\tilde{\mathcal{K}}(\Omega_2)$:

$$\begin{aligned} \tilde{\mathcal{K}}(\Omega_2) &= \sup_{z \notin \Omega_2} \; \text{dist}(z, \Omega_2) \, \|(zI - A)^{-1}\| \\ &\leq \sup_{z \notin \Omega_2} \; \text{dist}(z, \Omega_1) \, \|(zI - A)^{-1}\| \\ &\leq \sup_{z \notin \Omega_1} \; \text{dist}(z, \Omega_1) \, \|(zI - A)^{-1}\| \\ &= \tilde{\mathcal{K}}(\Omega_1). \qquad \square \end{aligned}$$

Proposition 2.3. *If A is a diagonalizable matrix with an eigenvalue decomposition $A = VDV^{-1}$, then*

$$\tilde{\mathcal{K}}(\Lambda(A)) \leq \kappa(V),$$

where $\kappa(V)$ is the condition number of V.

Proof. It is readily shown that

$$(zI - A)^{-1} = V\,(zI - D)^{-1}\,V^{-1}.$$

Taking norms on both sides gives

$$\|(zI - A)^{-1}\| \leq \frac{\kappa(V)}{\operatorname{dist}(z, \Lambda(A))}.$$

This implies that $\Lambda_\epsilon(A) \subset \Lambda(A) + \kappa(V)\,D_\epsilon$ for all $\epsilon \geq 0$, where D_ϵ is the closed disk of radius ϵ. Thus by Proposition 2.1, $\tilde{\mathcal{K}}(\Lambda(A)) \leq \kappa(V)$. ◻

Proposition 2.4. *If A is a normal matrix and $\Lambda(A) \subset \Omega$, then*

$$\tilde{\mathcal{K}}(\Omega) = 1.$$

Proof. Since $\tilde{\mathcal{K}}(\Omega) \geq 1$ for any Ω and A, we need only show $\tilde{\mathcal{K}}(\Omega) \leq 1$. This follows from Propositions 2.2 and 2.3. ◻

Proposition 2.5. *If $\mathcal{F}(A)$ is the field of values of A, then*

$$\tilde{\mathcal{K}}(\mathcal{F}(A)) = 1.$$

Proof. This is a restatement in a nonstandard language of mathematically standard material that is, for example, essentially the Hille–Yosida theorem of functional analysis [13]. ◻

3. Equivalence of the Kreiss constants $\tilde{\mathcal{K}}(\Omega)$ and $\mathcal{K}(\Omega)$. We have defined two different Kreiss constants: $\tilde{\mathcal{K}}(\Omega)$, which is defined geometrically, and $\mathcal{K}(\Omega)$, whose definition depends on the conformal maps Φ and Ψ. Thanks to results in geometric function theory stemming from the Koebe one-quarter theorem, these two constants are equivalent in the sense that they differ by at most a factor of 2.

Theorem 3.1 (R. Kühnau [11]). *Suppose that Ψ is a conformal map of the exterior of the closed unit disk D to the exterior of a compact set Ω with boundary Γ and Ω^c is simply connected, with $\Psi(\infty) = \infty$. Then for any $|\,w_0\,| > 1$,*

$$\frac{1}{2}\,(|\,w_0\,| - 1) \;\leq\; \frac{\operatorname{dist}(z_0, \Gamma)}{|\,\Psi'(w_0)\,|} \;\leq\; 2\,(|\,w_0\,| - 1), \tag{10}$$

where $z_0 = \Psi(w_0)$. If Ψ is an "interval map" (defined below), then the factor $1/2$ on the left is attained in the limit $|\,w_0\,| \to 1$, and if Ψ is a "circular arc map" (likewise), then the factor 2 on the right is attained in the limit $|\,w_0\,| \to 1$.

Proof. We establish the right-hand inequality first. Let w_0 be fixed, and consider the "circular arc map"

$$g(w) = w\frac{w - w_0}{1 - \bar{w}_0 w}, \quad w \in D^c.$$

This function maps the exterior of D onto the exterior of an arc S of the unit circle, and w_0 to 0, as shown in Figure 2. The angle subtended by S at the origin is $4 \arctan 1/(|\,w_0\,|^2 - 1)^{1/2}$.

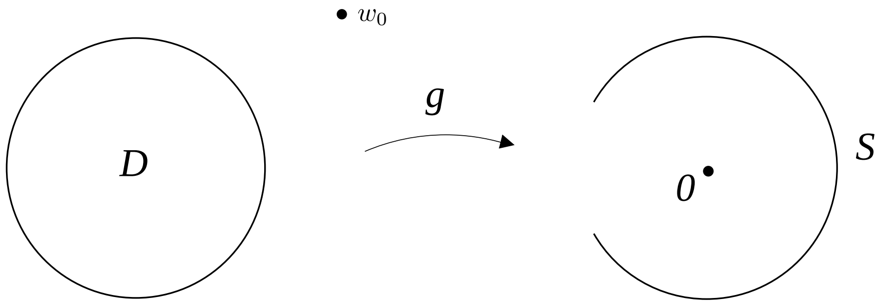

FIG. 2. *"Circular arc map," mapping the exterior of D onto the exterior of an arc S of the unit circle.*

On the contour $|w| = 1$, we have $|\Psi(w) - \Psi(w_0)| = |\Psi(w) - z_0| \geq \operatorname{dist}(z_0, \Gamma)$. Thus

$$\frac{|\Psi(w) - \Psi(w_0)|}{|g(w)|} \geq \operatorname{dist}(z_0, \Gamma) \quad \forall \ |w| = 1. \tag{11}$$

By the minimum principle applied to the analytic function $(\Psi(\cdot) - \Psi(w_0))/g(\cdot)$ defined on D^c, (11) holds $\forall\, |w| \geq 1$. In particular,

$$\lim_{w \to w_0} \left| \frac{\Psi(w) - \Psi(w_0)}{g(w)} \right| = \frac{|\Psi'(w_0)|}{|g'(w_0)|} \geq \operatorname{dist}(z_0, \Gamma). \tag{12}$$

Substituting $g'(w_0) = w_0/(1 - |w_0|^2)$ into (12) gives

$$\frac{\operatorname{dist}(z_0, \Gamma)}{|\Psi'(w_0)|} \leq \frac{|w_0|^2 - 1}{|w_0|} \leq 2 \operatorname{dist}(w_0, D). \tag{13}$$

If Ψ is a circular arc map, the left-hand inequality of (13) is actually an equality for all $|w_0| > 1$. In this case, the factor 2 is attained in the limit as $|w_0| \to 1$.

To be precise in the above proof, we actually have to consider contours which are inside D^c, say, $|w| = r$ for $1 < r < |w_0|$, and then take the limit as r tends to 1. We omit the details.

Next we show the left-hand inequality in (10). Again let w_0 be fixed, and this time, consider the "interval map"

$$h(w) = \frac{(w - w_0)(1 - \bar{w}_0 w)}{4w\,(|w_0| - 1)^2}, \quad w \in D^c.$$

This function maps the exterior of D onto the exterior of an interval I on the real line, and w_0 to 0, as shown in Figure 3. It is easily shown that

$$I = \left[\frac{1}{4}, \frac{(|w_0| + 1)^2}{4(|w_0| - 1)^2}\right] = \partial I, \tag{14}$$

$$h'(w_0) = \frac{1 + |w_0|}{4w_0\,(1 - |w_0|)}. \tag{15}$$

Now consider the conformal map $f = \Psi \circ h^{-1} \circ \kappa$, where $\kappa(\xi) = -\xi/(1 - \xi)^2$ is the Koebe function defined on the open unit disk $\mathring{D}$ (a conformal map of $\mathring{D}$ onto

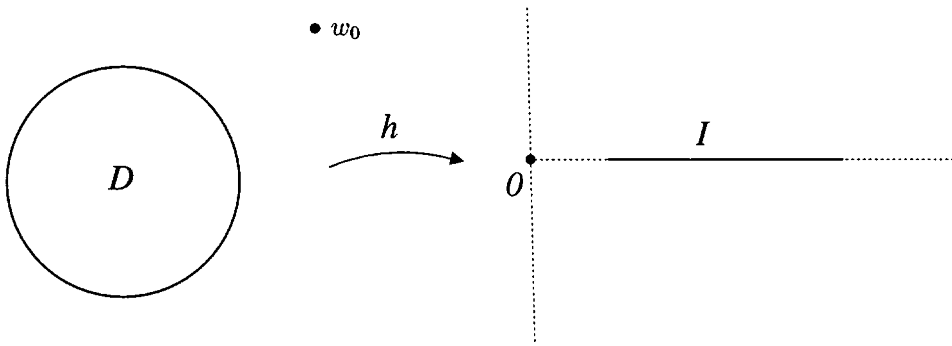

FIG. 3. *"Interval map," mapping the exterior of D onto the exterior of the interval* $I = \left[1/4,\, (|w_0|+1)^2/4(|w_0|-1)^2\right]$.

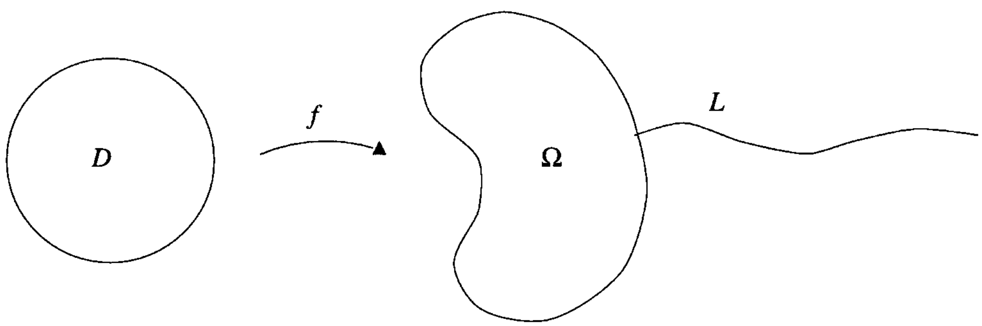

FIG. 4. *The function f maps the open unit disk $\mathring{D}$ onto the exterior of the set $\Omega \cup L$.*

$\mathbb{C} \setminus [1/4, \infty))$. The function f maps the open unit disk onto the exterior of $\Omega \cup L$, where L is a curve extending from the boundary of Ω to ∞, as shown in Figure 4.

By the Koebe one-quarter theorem [1, p. 29] applied to the function f, we have

$$\operatorname{dist}(f(0), \partial f(\mathring{D})) \geq \frac{1}{4}\,|f'(0)|\,. \tag{16}$$

Now

$$f(0) \;=\; \Psi(w_0), \quad f'(0) \;=\; \frac{\Psi'(w_0)}{h'(w_0)}, \quad \partial f(\mathring{D}) \;=\; \Gamma \cup L, \tag{17}$$

where

$$L = \Psi \circ h^{-1}\left(\frac{(|w_0|+1)^2}{4(|w_0|-1)^2}, \infty\right).$$

Substituting (17) into (16) gives

$$\operatorname{dist}(\Psi(w_0), \Gamma) \;\geq\; \operatorname{dist}(\Psi(w_0), \Gamma \cup L) \;\geq\; \frac{1}{4}\frac{|\Psi'(w_0)|}{|h'(w_0)|}. \tag{18}$$

Next, substituting (15) into (18) gives

$$\operatorname{dist}(\Psi(w_0), \Gamma) \geq \frac{|\Psi'(w_0)|}{1 + 1/\,|w_0|}\operatorname{dist}(w_0, D) \tag{19}$$

$$\geq \frac{1}{2}\ |\Psi'(w_0)|\ \operatorname{dist}(w_0, D).$$

If Ψ is an interval map, the inequality of (19) is actually an equality. In this case, the factor $1/2$ is attained in the limit as $|w_0| \to 1$. ∎

With the above theorem, it is now easy to show the equivalence of $\tilde{\mathcal{K}}(\Omega)$ and $\mathcal{K}(\Omega)$.

THEOREM 3.2. *Suppose A is a bounded linear operator in a Hilbert space with spectrum $\Lambda(A) \subset \Omega$. Then the Kreiss constants* (3) *and* (8) *are related by*

$$\frac{1}{2} \ \leq\ \frac{\tilde{\mathcal{K}}(\Omega)}{\mathcal{K}(\Omega)} \ \leq\ 2. \tag{20}$$

If Ω is an interval, then the factor $1/2$ on the left is attained in the limit as Ω becomes an infinitely long interval in the complex plane. If Ω is an arc of a circle, then the factor 2 on the right is attained in the limit as Ω approaches the full circle.

Proof. The result follows readily from Theorem 3.1 and the definitions of $\mathcal{K}(\Omega)$ and $\tilde{\mathcal{K}}(\Omega)$. ∎

4. The Kreiss matrix theorem (I). In this section, our goal is to establish Theorems 1.1 and 1.3 for a bounded linear operator A in a Hilbert space. The tools that we need are Faber series, the matrix analogue of the Cauchy integral formula, and Bernstein's lemma.

We shall begin with some basic results on Faber series. The following notation will be used in this and later sections:

- C_r: For each $r \geq 1$, C_r is the circle of radius r centered at the origin. The unit circle C_1 is also denoted by C.
- Γ_r: For each $r \geq 1$, Γ_r denotes the curve $\Psi(C_r)$. In the special case $r = 1$, Γ_1 is the boundary of Ω, and we denote this boundary also by Γ.
- Ω_r: For each $r \geq 1$, Ω_r denotes the closed region in the z-plane enclosed by $\Psi(C_r)$.
- $\|f\|_E$: For any continuous complex-valued function f defined on a compact subset E of the complex plane, $\|f\|_E$ denotes the maximum absolute value of f over E, i.e., $\|f\|_E = \max_{z\in E}\ |f(z)|$.
- P_n: The symbol P_n denotes the set of polynomials of degree at most n.

Suppose f is a function that is analytic in the interior of Ω and continuous on Ω. The *Faber series* associated with f is the formal series [2], [6, p. 44], [16]

$$f(\xi) \sim \sum_{n=0}^{\infty} a_n F_n(\xi), \quad \xi \in \Omega,$$

where

$$a_n = \frac{1}{2\pi i}\oint_C \frac{f(\Psi(w))}{w^{n+1}} dw.$$

Under suitable conditions, the Faber series actually converges to f in Ω. One sufficient condition for this is that the series converges uniformly in Ω [6, pp. 51–52]. For example, the function $(z-\xi)^{-1}$, where $z \in \Omega^c$ is fixed and $\xi \in \Omega$, can be expressed in terms of its Faber series because this series converges uniformly in Ω.

LEMMA 4.1. *For a fixed $z \in \Omega^c$, we have*

$$(z-\xi)^{-1} = \sum_{n=0}^{\infty} F_n(\xi) \frac{\Phi'(z)}{\Phi^{n+1}(z)} \quad \textit{for} \ \ \xi \in \Omega. \tag{21}$$

Proof. We omit the proof. ◻

We shall establish Theorem 1.3 first and then (4) of Theorem 1.1. Before doing so, let us state the Bernstein lemma.

LEMMA 4.2 (Bernstein). *Let* $p_n \in P_n$ *be arbitrary. Then*

$$\|p_n\|_{\Omega_r} \le r^n \, \|p_n\|_\Omega \quad \forall \quad r \ge 1.$$

Proof. See [6, p. 27]. ◻

Now we can prove Theorem 1.3.

Proof of Theorem 1.3. Under the assumption that $\Lambda(A) \subset \Omega$, the matrix $p_n(A)$ can be represented in terms of the resolvent of A via a Cauchy integral (see [5, pp. 555–577])

$$p_n(A) = \frac{1}{2\pi i} \oint_{\Gamma_r} (zI - A)^{-1} \, p_n(z) dz \tag{22}$$

for any $r > 1$. Taking norms on both sides, we obtain the inequalities

$$\begin{aligned} \|p_n(A)\| &\le \frac{1}{2\pi} \oint_{\Gamma_r} \|(zI - A)^{-1}\| \, |p_n(z)| \, |dz| \\ &\le \frac{\|p_n\|_{\Omega_r}}{2\pi} \oint_{\Gamma_r} \|(zI - A)^{-1}\| \, |dz| \, . \end{aligned} \tag{23}$$

On the contour Γ_r, the resolvent norm $\|(zI - A)^{-1}\|$ is bounded above by $\mathcal{K} \, |\Phi'(z)| \, /(r-1)$, and $\|p_n\|_{\Omega_r}$ is bounded above by $r^n \|p_n\|_\Omega$ (by the Bernstein lemma). The inequality (23) thus reduces to

$$\|p_n(A)\| \;\le\; \mathcal{K} \, \frac{\|p_n\|_\Omega}{2\pi} \, \frac{r^n}{r-1} \oint_{\Gamma_r} |\Phi'(z)| \, |dz| \;\le\; \mathcal{K} \, \|p_n\|_\Omega \, \frac{r^{n+1}}{r-1}. \tag{24}$$

In deriving (24), we have used the equality

$$\oint_{\Gamma_r} |\Phi'(z)| \, |dz| \;=\; \oint_{C_r} |dw| \;=\; 2\pi r.$$

Taking $r = 1 + 1/(n+1)$ in (24) and noting that $r^{n+1} \le e$ completes the proof. ◻

The proof of Theorem 1.3 is essentially based on ideas taken from [12], with the additional application of the Bernstein lemma. The Cauchy integral formula (22) is the standard tool that provides a link between the matrix or operator $p_n(A)$ and the resolvent $(zI - A)^{-1}$ of A, and this is the main tool that we use in this paper.

Proof of (4) *of Theorem* 1.1. The result follows immediately by applying the proof of Theorem 1.3 to the formula

$$F_n(A) = \frac{1}{2\pi i} \oint_{C_r} w^n \, \Psi'(w) \, (\Psi(w) I - A)^{-1} \, dw, \quad r > 1. \qquad ◻$$

Next we complete the proof of Theorem 1.1 by establishing the inequality (5), which states that the Kreiss constant $\mathcal{K}(\Omega)$ is bounded by the supremum of the norms $\{\|F_n(A)\|\}$. In the case of the standard Kreiss matrix theorem, a proof of such a result depends on the power series expansion of $(zI - A)^{-1}$. Here, the Faber series expansion of $(zI - A)^{-1}$ is required.

LEMMA 4.3. *Suppose* $\Lambda(A) \subset \Omega$. *Then*

$$(zI - A)^{-1} = \sum_{n=0}^{\infty} \frac{\Phi'(z)}{\Phi^{n+1}(z)} F_n(A) \quad \forall\ z \in \Omega^c. \tag{25}$$

Proof. By Theorem 16 in [5, p. 571], (25) is an easy consequence of Lemma 4.1. □

Proof of (5) *of Theorem* 1.1. Suppose $\sup_{n\geq 0} \|F_n(A)\| < \infty$. First we prove that $\Lambda(A) \subset \Omega$. Suppose there exists $\lambda \in \Lambda(A) \cap \Omega^c$. Then

$$\sup_n |F_n(\lambda)| \leq \sup_n \|F_n(A)\| < \infty.$$

But we also have (see [16])

$$\lim_{n\to\infty} |F_n(\lambda)|^{1/n} = |\Phi(\lambda)|,$$

which implies that $\sup_n |F_n(\lambda)| = \infty$ since $|\Phi(\lambda)| > 1$. Thus we have a contradiction. Hence $\Lambda(A) \subset \Omega$.

Now we prove the inequality (5). Let $z \in \Omega^c$ be fixed. Taking norms on both sides of (25) gives

$$\begin{aligned} \|(zI - A)^{-1}\| &\leq \sup_{n\geq 0} \|F_n(A)\| \sum_{k=0}^{\infty} \left| \frac{\Phi'(z)}{\Phi^{n+1}(z)} \right| \\ &= \sup_{n\geq 0} \|F_n(A)\| \frac{|\Phi'(z)|}{|\Phi(z)| - 1}, \end{aligned} \tag{26}$$

and (5) follows from (26). □

An analogue of (5) for a sequence of polynomials $\{p_n : p_n \in P_n\}$ can be derived from (5), but it is much less elegant. We give the result in the next corollary, for the sake of completeness.

COROLLARY 4.1. *Suppose* $\{p_n : p_n \in P_n,\ n \geq 0\}$ *is a sequence of polynomials,* $\Lambda(A) \subset \Omega$, *and* $\sup_{n\geq 0} \|p_n(A)\| < \infty$. *Let* $\mathbf{a}$ *denote the upper triangular array of Faber coefficients associated with* $\{p_n\}$, *i.e., for* $0 \leq k \leq n,\ n = 0, 1, 2, \ldots,$

$$a_{kn} = \frac{1}{2\pi i} \oint_C \frac{p_n(\Psi(w))}{w^{k+1}} dw.$$

If the inverse $\mathbf{b}$ *of* $\mathbf{a}$ *exists and* $\|\mathbf{b}\|_1 := \sup_{n\geq 0} \sum_{k=0}^{n} |b_{kn}| < \infty$, *then*

$$\mathcal{K}(\Omega) \leq \|\mathbf{b}\|_1 \sup_{n\geq 0} \|p_n(A)\|. \tag{27}$$

Proof. Let $\mathcal{P} = \sup\{\|p_n(A)\| : n \geq 0\}$. From the Faber series representation of p_n, we have

$$p_n(A) = \sum_{k=0}^{n} a_{kn} F_k(A), \quad n = 0, 1, 2, \ldots.$$

Thus

$$F_n(A) = \sum_{k=0}^{n} b_{kn}\, p_k(A), \quad n = 0, 1, 2, \ldots,$$

$$\|F_n(A)\| \leq \sum_{k=0}^{n} |\, b_{kn} \,|\ \|p_k(A)\| \ \leq\ \mathcal{P} \sum_{k=0}^{n} |\, b_{kn} \,| \,.$$

This implies that

$$\sup_{n\geq 0} \|F_n(A)\| \leq \mathcal{P} \sup_{n\geq 0} \sum_{k=0}^{n} |\, b_{kn} \,| \ =\ \mathcal{P}\ \|\mathbf{b}\|_1.$$

By Theorem 1.1, we have established (27). ∎

The Kreiss matrix theorem as formulated so far is scaled to the domain Ω. However, it is an easy matter to extend this to the domain Ω_r for any fixed $r \geq 1$ by scaling the conformal map Φ of Ω by r. We conclude this section with a theorem summarizing this extension.

THEOREM 4.1. *Suppose $\Lambda(A) \subset \Omega$ and $\mathcal{K}(\Omega) < \infty$. For each fixed $r \geq 1$, let $\mathcal{K}_r$ be the Kreiss constant with respect to the domain Ω_r. If p_n is a polynomial of degree n and F_n is the nth Faber polynomial associated with Ω, then the following bounds hold for any $n \geq 0$:*

$$\|p_n(A)\| \leq e\,(n+1)\, r^n\, \mathcal{K}_r\, \|p_n\|_{\Omega}; \tag{28}$$

$$\|F_n(A)\| \leq e\,(n+1)\, r^n\, \mathcal{K}_r; \tag{29}$$

$$\mathcal{K}_r \leq \sup_{n\geq 0}\ \|F_n(A)\|/r^n. \tag{30}$$

In addition,

$$\mathcal{K}_r = \inf \left\{ C \,:\, \|(zI - A)^{-1}\| \ \leq\ \frac{C\,|\,\Phi'(z)\,|}{|\,\Phi(z)\,| - r} \quad \forall\ z \notin \Omega_r \right\}, \tag{31}$$

where Φ is the conformal map associated with Ω.

Proof. First, we note that (28) follows from Theorem 1.3 with an application of the Bernstein lemma.

Next we note that the conformal map associated with the domain Ω_r is simply $\Phi(\cdot)/r$. Hence the nth Faber polynomial associated with Ω_r is $F_n(\cdot)/r^n$. With this observation, (29) and (30) follow from Theorem 1.1.

Finally, (31) follows from the definition of $\mathcal{K}_r$. ∎

5. The Kreiss matrix theorem (II). Let $N = \dim(\mathcal{H})$. In this section, we shall assume that N is finite and that the boundary Γ of Ω is a rectifiable Jordan curve or a slit in the complex plane. We assume further that Γ is of bounded total rotation. The total rotation of Γ, denoted by V, is the total variation in the argument of the tangent to the curve Γ as the curve is described once; see [6, p. 45] for details. If Ω is a convex domain, then V is 2π; if it is a polygon, then V is the total exterior turning angle. A region Ω for which Γ is of bounded total rotation need not have an interior; it might be, for example, the interval $[-1, 1]$ or a circular arc.

The Kreiss matrix theorem of the last section gives an upper bound for $\|F_n(A)\|$ that depends linearly on n. Here we establish another upper bound for $\|F_n(A)\|$ that depends linearly on N and at most logarithmically on n; if Γ is smooth, it depends only on N. The proof of these results depend on a sequence of lemmas established in Appendix A.

THEOREM 5.1. *Suppose A is a bounded linear operator in a finite-dimensional Hilbert space $\mathcal{H}$ with* $\dim(\mathcal{H}) = N$, $\Lambda(A) \subset \Omega$, *and* $\mathcal{K}(\Omega) < \infty$. *Let V be the total rotation of* Γ. *Then* $\forall n \geq 0$,

$$\|F_n(A)\| \leq \mathcal{K}\, e\, \frac{V}{2\pi}\,(4N + 1 + \alpha_n), \tag{32}$$

where

$$\alpha_n := \frac{2}{V} \int_{|w|=1+1/n} \frac{|\Psi''|}{|\Psi'|}|dw| \leq \frac{4}{\pi}\,[1 + \mathrm{arcsinh(2n+1)}]. \tag{33}$$

Proof. As in the proof of Theorem 1.1, we can express $F_n(A)$ in terms of the resolvent of A via the Cauchy integral formula,

$$F_n(A) = \frac{1}{2\pi i}\int_{C_r} w^n\,(\Psi(w)I - A)^{-1}\,\Psi'(w)\,dw \quad \forall\, r > 1.$$

Let u, v be N-vectors such that $\|u\| = \|v\| = 1$. Then

$$v^* F_n(A) u = \frac{1}{2\pi i}\int_{C_r} w^n R(\Psi(w))\,\Psi'(w)\,dw \quad \forall\ r > 1, \tag{34}$$

where $R(z) = v^*(zI - A)^{-1}u$. It can be shown that $R(z)$ is a rational function of order $(N-1, N)$ with the characteristic polynomial $\det(zI - A)$ as its denominator (hence $R(z)$ has poles only inside Ω).

Performing integration by parts in (34) gives

$$2\pi i\,(v^*\,F_n(A)\,u) = \frac{-1}{n+1}\int_{C_r} w^{n+1}\frac{d}{dw}\,[R(\Psi(w))\Psi']\,dw,$$

implying that

$$\begin{aligned} |2\pi\, v^* F_n(A) u| &\leq \frac{r^{n+1}}{n+1}\int_{C_r}\left|\frac{d}{dw}[R(\Psi(w))\Psi']\right|\,|dw| \\ &\leq \frac{r^{n+1}}{n+1}\left[\int_{C_r} |\Psi'^2 R'(\Psi)|\ |dw| + \int_{C_r} |\Psi'' R(\Psi)|\ |dw|\right]. \end{aligned} \tag{35}$$

Using the inequality of (58) in Lemma A.1, we get

$$\begin{aligned} |2\pi\, v^* F_n(A) u| &\leq \frac{r^{n+1}}{n+1}\left[\int_{C_r} |h'|\,|\Psi' R(\Psi)|\ |dw| + 2\int_{C_r} |\Psi'' R(\Psi)|\ |dw|\right] \\ &\leq \frac{r^{n+1}}{n+1}\,\|R(\Psi)\Psi'\|_{C_r}\left[\int_{C_r} |h'|\ |dw| + 2\int_{C_r}\frac{|\Psi''|}{|\Psi'|}\,|dw|\right], \end{aligned}$$

where $h(w) = \arg(w\,\Psi'(w)^2\,R'(\Psi(w)))$.

On the contour C_r, $|R(\Psi(w))\Psi'(w)| \leq \mathcal{K}/(r-1)$. Thus we have

$$|2\pi\, v^* F_n(A) u| \leq \frac{\mathcal{K}\, r^{n+1}}{(n+1)(r-1)}\left[\int_{C_r} |h'|\ |dw| + 2\int_{C_r}\frac{|\Psi''|}{|\Psi'|}\,|dw|\right].$$

Taking $r = 1 + 1/n$ and using the inequality of (59) in Lemma A.2, we get

$$|\, v^* F_n(A) u \,| \le \mathcal{K}\, e \left[\frac{V}{2\pi}\,(4N+1) \;+\; \frac{1}{\pi} \int_{|w|=1+1/n} \frac{|\,\Psi''\,|}{|\,\Psi'\,|} |dw| \right]$$

$$\le \mathcal{K}\, e\, \frac{V}{2\pi}\,(4N \,+\, 1 \,+\, \alpha_n). \tag{36}$$

Since $\|F_n(A)\|$ is the supremum of $|\, v^* F_n(A) u \,|$ over all unit vectors u and v, (32) follows from (36), and by Lemma A.3, it is easy to show that α_n satisfies the inequality of (33). This completes the proof. □

Ideally, one would like to have a bound that depends only on N in Theorem 5.1, as is the case with the standard Kreiss matrix theorem. However we have been unable to prove such a bound without making a stronger assumption on Γ. If we assume that Γ is *twice continuously differentiable*, then indeed we have a bound that depends only on N. This is the content of Theorem 1.2, whose proof follows easily from Theorem 5.1.

Proof of Theorem 1.2. From Theorem 5.1, we have

$$\|F_n(A)\| \le \mathcal{K}\, e\, \frac{V}{2\pi}\,(4N + 1 + \alpha), \tag{37}$$

where α is a constant independent of n, defined by

$$\alpha \;:=\; \sup_{n \ge 0} \alpha_n \;=\; \frac{2}{V} \int_C \frac{|\,\Psi''(w)\,|}{|\,\Psi'(w)\,|}\; |\,dw\,|\,. \tag{38}$$

(Note that α_n increases with n.) Under the assumption that Γ is twice continuously differentiable, α is finite, and therefore Theorem 1.2 is proved with

$$C_\Omega = \frac{V}{2\pi}\left(4 \,+\, \frac{1+\alpha}{N}\right). \qquad \square$$

Again, some of the ideas in the proof of Theorem 5.1 are taken from [12].

Like the first version of the Kreiss matrix theorem in the last section, analogous results of Theorem 4.1 scaled to the domain Ω_r can also be established for Theorem 5.1. This goes exactly as before, so we omit stating the results.

6. The Kreiss matrix theorem for the unit interval $[-1, 1]$. We have shown in section 5 that if the boundary of Ω is smooth, then $\|F_n(A)\|$ can be bounded in terms of N, independently of n. For a domain Ω such that the associated conformal map Ψ has the form

$$\Psi(w) = c\, w \;+\; \sum_{k=0}^{p} \frac{c_k}{w^k}, \tag{39}$$

where p is some nonnegative integer, it is also possible to bound $\|F_n(A)\|$ in terms of N, independently of n, even if $\partial\Omega$ is not smooth. In this section, we illustrate how this can be done for the special case $\Omega = [-1, 1]$. With this special case, it should be clear to the reader how an analogous result can be established for the general case.

For the unit interval $\Omega = [-1, 1]$, $\Psi(w)$ and $\Psi'(w)$ are given by

$$\Psi(w) = \frac{1}{2}\left(w + \frac{1}{w}\right), \qquad |w| > 1,$$

$$\Psi'(w) = \frac{1}{2}\left(1 - \frac{1}{w^2}\right).$$

It is well known that for $n \geq 1$, the nth Faber polynomial $F_n(z)$ associated with $[-1,1]$ is given by $F_n(z) = 2T_n(z)$, where $T_n(z) := \cos(n \arccos(z))$ is the nth Chebyshev polynomial. By means of Spijker's lemma, we shall establish a Kreiss matrix theorem for $[-1,1]$, giving a bound for $\|T_n(A)\|$ in terms of N. We state Spijker's lemma first and then prove the Kreiss matrix theorem.

LEMMA 6.1 (Spijker's lemma). *Suppose R is a rational function of order N. For any $r > 0$ such that R has no poles on C_r, there holds*

$$\int_{C_r} |R'(w)|\,|dw| \leq 2\pi N\,\|R\|_{C_r}.$$

Proof. See [17] or [20]. □

THEOREM 6.1. *Suppose A is a bounded linear operator in a finite-dimensional Hilbert space H with $\dim(\mathcal{H}) = N$, $\Lambda(A) \subset [-1,1]$, and $\mathcal{K}(\Omega) < \infty$. Then*

$$\|T_n(A)\| \leq e\,(N+1)\,\mathcal{K}(\Omega), \tag{40}$$

where $T_n(z) = \cos(n \arccos(z))$, for $n = 1, 2, \ldots$. Consequently,

$$\sup_{n \geq 0} \|T_n(A)\| \leq e\,(N+1)\,\mathcal{K}(\Omega). \tag{41}$$

Proof. For any unit N-vectors u and v, we have, as in (35),

$$\begin{aligned} |\,2\pi v^*(2T_n(A))u\,| &\leq \frac{r^{n+1}}{n+1} \int_{C_r} \left| \frac{d}{dw}[R(\Psi(w))\Psi'(w)] \right| |dw| \\ &= \frac{r^{n+1}}{n+1} \int_{C_r} |q'(w)|\,|dw|, \end{aligned}$$

where R is a rational function of order $(N-1, N)$ and $q(w) = R(\Psi(w))\Psi'(w)$. Since Ψ and Ψ' are rational functions of order $(2,1)$ and $(2,2)$, respectively, it is readily shown that q is a rational function of order $(2N+1, 2N+2)$. Applying Spijker's lemma to the rational function q and noting that $\|q\|_{C_r} \leq \mathcal{K}/(r-1)$ on the contour $C_r = \{w : |w| = r\}$, we obtain

$$|\,v^* T_n(A) u\,| \leq \frac{r^{n+1}}{(n+1)(r-1)}\,(N+1)\,\mathcal{K}. \tag{42}$$

Now (40) follows by taking $r = 1 + 1/(n+1)$ in (42). □

Next we give an example, analogous to a similar example in [12] for the standard Kreiss matrix theorem, to show that the linear dependence N in (41) is sharp.

Consider the $N \times N$ matrix

$$A = \begin{bmatrix} x_1 & \gamma & & \\ & \ddots & \ddots & \\ & & x_{N-1} & \gamma \\ & & & x_N \end{bmatrix}, \tag{43}$$

where γ is a positive real number greater than 3 and

$$x_i = \cos\left(\frac{i-1}{N-1}\pi\right), \quad i = 1, \ldots, N. \tag{44}$$

For this particular matrix, we have

$$(zI - A)^{-1} = \begin{bmatrix} \frac{1}{z-x_1} & \frac{\gamma}{\prod_{i=1}^{2}(z-x_i)} & \cdots & \frac{\gamma^{N-1}}{\prod_{i=1}^{N}(z-x_i)} \\ & \ddots & & \vdots \\ & & \frac{1}{z-x_{N-1}} & \frac{\gamma}{\prod_{i=N-1}^{N}(z-x_i)} \\ & & & \frac{1}{z-x_N} \end{bmatrix} \tag{45}$$

and

$$\begin{aligned} T_n(A) &= \frac{1}{2\pi i}\int T_n(z)(zI-A)^{-1}dz \\ &= \frac{1}{2\pi i}\begin{bmatrix} \int \frac{T_n(z)}{(z-x_1)}dz & \cdots & \int \frac{\gamma^{N-1}\, T_n(z)}{\prod_{i=1}^{N}(z-x_i)}dz \\ & \ddots & \vdots \\ & & \int \frac{T_n(z)}{(z-x_N)}dz \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} T_n[x_1] & \gamma T_n[x_1, x_2] & \cdots & \gamma^{N-1}\, T_n[x_1, \cdots, x_N] \\ & \ddots & & \vdots \\ & & T_n[x_{N-1}] & \gamma\, T_n[x_{N-1}, x_N] \\ & & & T_n[x_N] \end{bmatrix}, \tag{46}$$

where $T_n[x_i, \ldots, x_j]$ denotes the divided difference of the polynomial T_n with respect to the points $x_i, \ldots, x_j$.

To show that the linear dependence on N in (41) is sharp, we need to establish a result of the form $\sup_{n\geq 0} \|T_n(A)\| \geq c\, N\, \mathcal{K}$ for some constant c independent of N.

THEOREM 6.2. *For the matrix A of* (43)*, we have*

$$\frac{\gamma - 3}{2\pi\gamma}\,(N-1)\,\mathcal{K} \leq \sup_{n\geq 0}\, \|T_n(A)\|, \tag{47}$$

where $T_n(z) = \cos(n\cos^{-1}(z))$, $z \in [-1,1]^c$.

Proof. For readability, we divide the proof into three parts.

(a) It is clear that

$$\begin{aligned} \sup_{n\geq 0}\, \|T_n(A)\| &\geq \|T_{N-1}(A)\| \\ &\geq \gamma^{N-1}\, T_{N-1}[x_1, \ldots, x_N] \;\; = \;\; \gamma^{N-1}\, 2^{N-2}. \end{aligned} \tag{48}$$

(b) We need an upper bound for $\mathcal{K}$. But first we must obtain an upper bound for $\|(zI-A)^{-1}\|_2$, where $z \in [-1,1]^c$. For this purpose, we shall use the inequality

$$\|(zI-A)^{-1}\|_2^2 \leq \|(zI-A)^{-1}\|_1\, \|(zI-A)^{-1}\|_\infty.$$

It is easily shown that

$$\|(zI-A)^{-1}\|_1 = \frac{1}{\gamma}\sum_{k=1}^{N}\prod_{i=N-k+1}^{N}\frac{\gamma}{|z-x_i|}, \tag{49}$$

$$\|(zI-A)^{-1}\|_\infty = \frac{1}{\gamma}\sum_{k=1}^{N}\prod_{i=1}^{k}\frac{\gamma}{|z-x_i|}. \tag{50}$$

Now consider the region E defined by

$$E = \{z \in \mathbb{C} : |z-x_i| \le 3 \quad \forall\, i = 1,\dots,N\}.$$

Noting that if $z \in E^c$, then $|z-x_i| \ge 1$ for all $i = 1,\dots,N$, and with some algebraic manipulations in (49) and (50), we obtain

$$\|(zI-A)^{-1}\|_1,\ \|(zI-A)^{-1}\|_\infty \ \le \begin{cases} \dfrac{\gamma^N}{(\gamma-1)\,\mathrm{dist}(z,\Omega)} & \text{if } z \in E^c, \\[2ex] \dfrac{1}{(\gamma-3)}\ \dfrac{\gamma^N}{\prod_{i=1}^{N}\ |z-x_i|} & \text{if } z \in E \cap [-1,1]^c. \end{cases}$$

Therefore

$$\|(zI-A)^{-1}\|_2 \ \le \begin{cases} \dfrac{\gamma^N}{(\gamma-1)\,\mathrm{dist}(z,\Omega)} & \text{if } z \in E^c, \\[2ex] \dfrac{1}{(\gamma-3)}\ \dfrac{\gamma^N}{\prod_{i=1}^{N}\ |z-x_i|} & \text{if } z \in E \cap [-1,1]^c. \end{cases} \tag{51}$$

(c) Finally, recall that $\tilde{\mathcal{K}}$ is the supremum of the function $\mathrm{dist}(z,\Omega)\,\|(zI-A)^{-1}\|_2$ over the region $[-1,1]^c$. From (51), we get

$$\tilde{\mathcal{K}} \ \le\ \frac{\gamma^N}{\gamma-3}\ \max\left\{\sup_{z\in[-1,1]^c}\ \frac{\mathrm{dist}(z,\Omega)}{\prod_{i=1}^{N}\ |z-x_i|},\ 1\right\} \ =\ \pi\,\frac{\gamma^N}{\gamma-3}\,\frac{2^{N-2}}{N-1}, \tag{52}$$

where (52) follows from Lemma 6.2. Comparing (48) and (52), we get

$$\tilde{\mathcal{K}}\ (N-1)\ \frac{\gamma-3}{\pi\gamma} \le \sup_{n\ge 0}\|T_n(A)\|. \tag{53}$$

By Theorem 3.2, $\mathcal{K} \le 2\,\tilde{\mathcal{K}}$, and thus (47) follows from (53). ◻

The next lemma is a technical result used in the proof of Theorem 6.2. The reader is encouraged to skip the proof.

LEMMA 6.2. *Suppose* $\Omega = [-1,1]$ *and* x_i, $i = 1,\dots,N$, *are given as in* (44). *If* $N \ge 3$, *then* $\forall z \in \Omega^c$,

$$\frac{\mathrm{dist}(z,\Omega)}{\prod_{i=1}^{N}\ |z-x_i|} \le \pi\,\frac{2^{N-2}}{N-1}. \tag{54}$$

Proof. First we note that the zeros of $T'_{N-1}(z)$ are the extrema $x_2,\dots,x_{N-1}$ of $T_{N-1}(z)$ in the open interval $(-1,1)$. Thus

$$T'_{N-1}(z) = (N-1)\ 2^{N-2}\prod_{i=2}^{N-1}\ (z-x_i).$$

We shall make use of this equation without explicitly referring to it in this proof. Note also that $|T'_{N-1}(1)| = (N-1)^2$.

We shall prove (54) for z in two separate regions, E and F, to be defined below. Note that by symmetry, it is sufficient to consider these two regions.

(a) Let E be the region defined by

$$E = \{z = x + iy \,:\, x \geq 1, y \in \mathbb{R}\}.$$

For $z \in E$, $\mathrm{dist}(z, \Omega) = |z-1|$, and thus

$$\frac{\mathrm{dist}(z,\Omega)}{\prod_{k=1}^{N} |z - x_k|} = \frac{1}{\prod_{k=2}^{N} |z - x_k|}. \tag{55}$$

Since $1/\prod_{k=2}^{N}(z - x_k)$ is an analytic function in the interior of E and continuous in E, it is sufficient to show that for $z \in \partial E = \{1 + iy \,:\, y \in \mathbb{R}\}$, the function of (55) satisfies the bound of (54). Let $f(y)$ be the function defined by

$$f(y) = \frac{1}{\prod_{k=2}^{N} |(1 - x_k) + iy|}, \qquad y \in \mathbb{R}.$$

Elementary calculus shows that $y = 0$ is the maximizer of f and

$$f(0) \;=\; \frac{(N-1)\,2^{N-3}}{|T'_{N-1}(1)|} \;=\; \frac{2^{N-3}}{(N-1)}. \tag{56}$$

This establishes (54) for $z \in E$.

(b) Let $\theta = \pi/(N-1)$ and $I_j = [\xi_j,\ \xi_{j-1}]$, $j = 1, \ldots, N$, where $\xi_0 = 1$, $\xi_N = -1$, $\xi_j = \cos(j - \frac{1}{2})\theta$ for $j = 1, \ldots, N-1$.

Now consider the region $F = \{x + iy \,:\, -1 \leq x \leq 1,\ y > 0\}$. It is clear that $F = \cup_{j=1}^{N} F_j$, where

$$F_j \;=\; \{x + iy \,:\, x \in I_j,\ y > 0\}.$$

For $z = x + iy \in F_j$, we have

$$\frac{\mathrm{dist}(z,\Omega)}{\prod_{k=1}^{N} |z - x_k|} \;\leq\; \frac{1}{g_j(x)} \;\leq\; \frac{1}{\min_{x \in I_j}\, g_j(x)},$$

where $g_j(x)$, $x \in (x_{j+1},\ x_{j-1})$, is the function defined by

$$\begin{aligned} g_j(x) &:= \prod_{k=1,\, k\neq j}^{N} |x - x_k| \\ &= \frac{T'_{N-1}(x)}{(N-1)\,2^{N-1}}\ \frac{|1 - x^2|}{|x - x_j|}. \end{aligned}$$

It is easy to see that $g_j(x)$ has the form shown in Figure 5 in the interval (x_{j+1}, x_{j-1}) and the minimizer of g_j in the interval I_j is located at the end points of I_j. With some tedious algebra, it can be shown that

$$\min_{x \in I_j} g_j(x) = \min(g_j(\xi_{j-1}),\ g_j(\xi_j)) \;=\; g_j(\xi_{j-1})$$

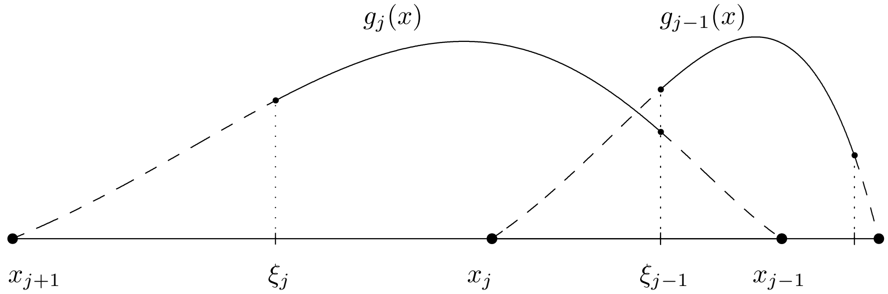

FIG. 5. *Proof of Lemma* 6.2.

$$
\begin{aligned}
&= \frac{1}{2^{N-1}} \, \frac{\sin(j-\frac{3}{2})\theta}{\sin\theta/4 \, \sin(j-\frac{5}{4})\theta} \\
&\geq \frac{1}{2^{N-1}} \, \frac{\sin\theta/2}{\sin\theta/4 \, \sin 3\theta/4} \\
&\geq \frac{1}{2^{N-1}} \, \frac{8}{3\theta} \;\; \geq \;\; \frac{(N-1)}{2^{N-2}} \, \frac{1}{\pi}. \qquad (57)
\end{aligned}
$$

By (57), we have established (54) for $z \in F_j$, $j = 1, \ldots, N$. This completes the proof. ◻

7. Discussion. We have generalized the Kreiss matrix theorem from the unit disk to a general complex domain, giving bounds for the associated Faber polynomials as well as for arbitrary polynomials. There are other generalizations of the Kreiss matrix theorem in the literature, in particular for the purpose of linear stability analysis of semidiscrete methods for time-dependent PDEs. In these problems, one is concerned with the boundedness of $\{\|\phi(\delta t\, A)^n\|\}$, where ϕ is a fixed rational function determined by the time-discretization scheme, A is determined by the space discretization scheme, and δt is the time step. The Kreiss matrix theorem has been generalized to the stability region S associated with the time-discretization scheme by a transplantation of the Kreiss matrix theorem from the unit disk to S [14]. Such a result is prototypical of the kind of generalizations of the Kreiss matrix theorem to regions in the complex plane for the purpose of linear stability analysis. For more details in this subject, see the survey article by van Dorsselaer, Kraaijevanger, and Spijker [3]. In the special case when ϕ is a fixed polynomial, our generalized Kreiss matrix theorems give the same results as those obtained from the transplantation technique, except for some technical differences.

We end with a note on the case when Ω consists of several connected components each bounded by a Jordan curve. We believe that analogues of our Kreiss matrix theorems can be established for this case. The main reason is that results such as the Bernstein lemma and the Cauchy integral formula that we have used in establishing our Kreiss matrix theorems continue to hold for this more general Ω; see [5], [21], [22].

Appendix A.

In this appendix, we present a sequence of lemmas whose results are applied in the proof of Theorem 5.1. The reader may skip these lemmas without any loss of understanding.

LEMMA A.1. *Suppose f is a analytic function defined in Ω^c. Then for each*

$r > 1$,

$$\int_{C_r} \left|\Psi'^2 f'(\Psi)\right| |dw| \le \int_{C_r} |h'|\,|\Psi' f(\Psi)|\ |dw| \ + \int_{C_r} |\Psi'' f(\Psi)|\ |dw|, \tag{58}$$

where $h(w) = \arg[w\Psi'^2(w)f'(\Psi(w))]$.

Proof. It is readily shown that

$$\int_{C_r} \left|\Psi'^2 f'(\Psi)\right| |dw| = -i\int_{C_r} \Psi'(w)\frac{df(\Psi(w))}{dw}e^{-ih(w)}dw$$
$$= -i\int_{C_r} \Psi'(w)e^{-ih(w)}\,df(\Psi(w)).$$

Integration by parts gives

$$\int_{C_r} \left|\Psi'^2 f'(\Psi)\right| |dw| = \int_{|w|=r} h'\Psi' f(\Psi)e^{-ih}dw \ + \ i\int_{|w|=r} \Psi'' f(\Psi)e^{-ih}dw.$$

From this equation, (58) follows readily. ∎

LEMMA A.2. *Suppose R is a rational function of order $(N-1, N)$ with no poles inside Ω^c. Then*

$$\int_{C_r} |h'(w)|\,|dw| \le (4N+1)V \quad \forall\ r > 1, \tag{59}$$

where $h(w) = \arg[w\Psi'^2(w)R'(\Psi(w))]$.

Proof. Since R is a rational function of order $(N-1, N)$, this implies that R' is a rational function of order $(2N-2, 2N)$. Hence $R'(\Psi)$ can be written as a product

$$R'(\Psi(w)) = \prod_{k=1}^{2N-2}(a_k\Psi(w)+b_k)\prod_{k=2N-1}^{4N-2}\frac{1}{(a_k\Psi(w)+b_k)}.$$

This implies that

$$h(w) \ = \ 2\arg(w\Psi'(w)) \ - \ \arg w \ + \ \sum_{k=1}^{2N-2}\arg\left(a_k\Psi(w)+b_k\right) \ - \sum_{k=2N-1}^{4N-2}\arg\left(a_k\Psi(w)+b_k\right).$$

Therefore the total variation $\mathrm{TV}[h]$ of h around the circle C_r satisfies

$$\mathrm{TV}[h] \le 2\,\mathrm{TV}[\arg(w\Psi'(w))] \ + \ \mathrm{TV}[\arg w] \ + \ \sum_{k=1}^{2N-2}\mathrm{TV}\left[\arg\left(a_k\Psi(w)+b_k\right)\right]$$
$$+ \sum_{k=2N-1}^{4N-2}\mathrm{TV}\left[\arg\left(a_k\Psi(w)+b_k\right)\right]$$
$$= 2\,\mathrm{TV}[\arg(w\Psi'(w))] \ + \ 2\pi \ + \ \sum_{k=1}^{4N-2}\mathrm{TV}\left[\arg\left(\Psi(w)+\gamma_k\right)\right],$$

where $\gamma_k = b_k/a_k$.

Now by Lemma 1 of [8],

$$\mathrm{TV}[\arg(\Psi(w)+\gamma)] \le \mathrm{TV}[\arg(w\Psi'(w))] \quad \forall\ \ \gamma \in \mathbb{C}.$$

Thus

$$\mathrm{TV}[h] \le 4N\ \mathrm{TV}[\arg(w\Psi'(w))] + 2\pi. \tag{60}$$

Noting that $\mathrm{TV}[\arg(w\Psi'(w))]$ is the total variation in the argument of the tangent to the curve Γ_r as the curve is described counterclockwise once, and by definition, this is the total rotation $V(\Gamma_r)$ of Γ_r. It follows from (60) that

$$\mathrm{TV}[h] \ \le\ 4N\,V(\Gamma_r) \ +\ 2\pi \ \le\ (4N\ +\ 1)\,V(\Gamma_r).$$

It is known that $V(\Gamma_r) \le V\ \forall\ r > 1$ (see [10]). Thus we have established (59). ∎

LEMMA A.3. *For each $r > 1$,*

$$\int_{C_r} \frac{|\,\Psi''(w)\,|}{|\,\Psi'(w)\,|}|dw| \le \frac{2V}{\pi}\left[1 + \operatorname{arcsinh}\left(\frac{r+1}{r-1}\right)\right]. \tag{61}$$

Proof. From Lemma 5 of [10, p. 202], we have

$$\log\Psi'(w) = \frac{1}{\pi}\int_0^{2\pi} \log\left(1 - \frac{e^{i\theta}}{w}\right) du(\theta) \quad \forall\ \ |\,w\,|> 1, \tag{62}$$

where $u(\theta)$ is a function of bounded variation such that $\int_0^{2\pi} |\,du(\theta)\,| = V$. Differentiating (62) with respect to w gives

$$\frac{\Psi''(w)}{\Psi'(w)} = \frac{1}{\pi w}\int_0^{2\pi} \frac{e^{i\theta}}{w - e^{i\theta}}\, du(\theta).$$

Thus

$$\begin{aligned} \int_{C_r} \frac{|\,\Psi''(w)\,|}{|\,\Psi'(w)\,|}|dw| &\le \frac{1}{\pi r}\int_0^{2\pi}\int_{|w|=r} \frac{|dw|}{|\,w - e^{i\theta}\,|}\ \ |\,du(\theta)\,| \\ &\le \frac{4V}{\pi(r+1)}\left[1 + \operatorname{arcsinh}\left(\frac{r+1}{r-1}\right)\right], \end{aligned} \tag{63}$$

because

$$\int_{C_r} \frac{|dw|}{|\,w - e^{i\theta}\,|} = \frac{4r}{r+1}\left[1 + \operatorname{arcsinh}\left(\frac{r+1}{r-1}\right)\right].$$

Since $r > 1$, (61) follows from (63). ∎

Acknowledgments. We are grateful to Professor R. Kühnau for providing us the proof of Theorem 3.1 and to D. Gaier for helpful comments on a draft of this paper.

REFERENCES

[1] L.V. AHLFORS, *Conformal Invariants: Topics in Geometric Function Theory*, McGraw-Hill, New York, 1973.

[2] J.H. CURTISS, *Faber polynomials and the Faber series*, Amer. Math. Monthly, 78 (1971), pp. 577–596.

[3] J.L.M. VAN DORSSELAER, J.F.B.M. KRAAIJEVANGER, AND M.N. SPIJKER, *Linear stability analysis in the numerical solution of initial value problems*, Acta Numer. 1993, pp. 199–237.

[4] 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.
[5] N. Dunford and J. Schwartz, *Linear Operators* I, Wiley–Interscience, New York, 1957.
[6] D. Gaier, *Lectures on Complex Approximation*, Birkhäuser, Boston, 1985.
[7] A. Greenbaum and L.N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comp., 15 (1994), pp. 359–368.
[8] R.R. Hall and St. Ruscheweyh, *On transforms of functions with bounded boundary rotation*, Indian J. Pure Appl. Math., 16 (1985), pp. 1317–1325.
[9] H.O. Kreiss, *Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differenzialgleichungen approximieren*, BIT, 2 (1962), pp. 153–181.
[10] T. Kövari and Ch. Pommerenke, *On Faber polynomials and Faber expansions*, Math. Zeitschr., 99 (1967), pp. 193–206.
[11] R. Kühnau, *private communication*, University of Halle, Germany, 1994.
[12] R.J. LeVeque and L.N. Trefethen, *On the resolvent condition in the Kreiss matrix theorem*, BIT, 24 (1984), pp. 584–591.
[13] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
[14] S.C. Reddy and L.N. Trefethen, *Stability of the method of lines*, Numer. Math., 62 (1992), pp. 235–267.
[15] E.B. Saff, *Polynomial and rational approximation in the complex domain*, Proc. Sympos. Appl. Math., 36 (1986), pp. 21-49.
[16] V.I. Smirnov and N.A. Lebedev, *Functions of a Complex Variable: Constructive Theory*, MIT Press, Cambridge, MA, 1968.
[17] M.N. Spijker, *On a conjecture by LeVeque and Trefethen related to the Kreiss matrix theorem*, BIT, 31 (1991), pp. 551–555.
[18] P.K. Suetin, *The basic properties of Faber polynomials*, Russian Math. Surveys, 19 (1964), pp. 121–149.
[19] L.N. Trefethen, *Spectra and Pseudospectra: The Behavior of Non-Normal Matrices and Operators*, book in preparation.
[20] E. Wegert and L.N. Trefethen, *From the Buffon needle problem to the Kreiss matrix theorem*, Amer. Math. Monthly, 101 (1994), pp. 132–139.
[21] J.L. Walsh, *Interpolation and Approximation*, Amer. Math. Soc. Colloq. Publ., New York, 1935.
[22] J.L. Walsh, *A generalization of Faber's polynomials*, Math. Ann., 136 (1958), pp. 23–33.