

Eigenvalue estimates for the resolvent of a non-normal matrix

Oleg Szehr¹

Abstract. We investigate the relation between the spectrum of a non-normal matrix and the norm of its resolvent. We provide spectral estimates for the resolvent of matrices whose largest singular value is bounded by 1 (so-called Hilbert space contractions) and for power-bounded matrices. In the first case our estimate is optimal and we present explicit matrices that achieve equality in the bound. This result recovers and generalizes previous estimates obtained by E.B. Davies and B. Simon in the study of orthogonal polynomials on the unit circle. In case of power-bounded matrices we achieve the strongest estimate so far. Our result unifies previous approaches, where the resolvent was estimated in certain restricted regions of the complex plane. To achieve our estimates we relate the problem of bounding the norm of a function of a matrix to a Nevanlinna–Pick interpolation problem in a corresponding function space. In case of Hilbert space contractions this problem is connected to the theory of compressed shift operators to which we contribute by providing explicit matrix representations for such operators. Finally, we apply our results to study the sensitivity of the stationary states of a classical or quantum Markov chain with respect to perturbations of the transition matrix.

Mathematics Subject Classification (2010). Primary: 15A60; Secondary: 65F35, 65J05.

Keywords. Resolvent, Non-normal matrix, Markov chain.

Contents

1	Introduction	784
2	Preliminaries	785
3	Hilbert space contractions	790
4	Power-bounded operators	803
5	Stability of Markov chains	809
	References	811

¹The author acknowledges financial support by the Elite Network of Bavaria (ENB) project QCCC and the CHIST-ERA/BMBF project CQC.

1. Introduction

The contribution of this article is to provide new estimates on the norm of the resolvent of a matrix A and to prove their optimality under certain conditions. We derive bounds of the form

$$\|(\zeta - A)^{-1}\| \leq \Phi(\zeta, n, \sigma(A)), \quad (1)$$

where Φ is a function of $\zeta \in \mathbb{C}$, the dimension n and the spectrum $\sigma(A)$ of A . In the first part of the article, (cf. Section 3) we assume that the largest singular value (the spectral norm) of A is bounded by 1 i.e. $\|A\|_\infty \leq 1$. Note that this can always be achieved by a suitable normalization. Under this assumption we obtain optimal bounds for $\zeta \in \mathbb{C} - \sigma(A)$ and present explicit matrices that establish equality in (1). Thus we identify the relation between the localization of the spectrum of A and the norm of its resolvent. In the second part (cf. Section 4) we study (1) under the assumption that each power of A can be bounded with respect to *any* given norm by the same constant, $\sup_{k \geq 0} \|A^k\| \leq C$. In this case we derive the strongest estimates so far.

The problem of finding good functions Φ was studied extensively in the literature before [6, 29, 17, 5]. Our approach is based on the theory of certain (Hilbert/Banach) function spaces. We associate to a given class of matrices Γ a certain Banach algebra \mathcal{A} of functions and instead of working with matrices directly we estimate the norm of a representative function in the function algebra. A key role is played by inequalities of the type

$$\|f(A)\| \leq C \|f\|_{\mathcal{A}}, \quad (2)$$

which relate for a given $A \in \Gamma$ the norm $\|f(A)\|$ to the norm of f in \mathcal{A} . At first glance this appears to be of little use since the right hand side no longer depends on A . However, it is possible to exploit spectral properties of A to significantly strengthen the inequality in (2). Let m_A be the minimal polynomial of A . For any $f, g \in \mathcal{A}$ we have then that $\|(f + m_A g)(A)\|_{\mathcal{A}} = \|f(A)\|_{\mathcal{A}}$ and an application of (2) reveals that for all $g \in \mathcal{A}$ we have $\|f(A)\| \leq C \|f + m_A g\|_{\mathcal{A}}$. This relates the problem of bounding $\|f(A)\|$ to the problem of finding the least norm function $f + m_A g$ in \mathcal{A} , which is equivalent to a Nevanlinna–Pick interpolation problem in \mathcal{A} ; see [17]. If $\|A\|_\infty \leq 1$ the resulting interpolation problem can be solved using an operator theoretic approach pioneered by D. Sarason [22, 1]. This approach is intrinsically connected to the theory of compressed shift operators on Hardy space. We contribute to this theory by providing a framework that allows us to compute explicit matrix representations for functions of model operators. In case

that A is power-bounded we choose a rational approximation function in \mathcal{A} and bound its norm to achieve our result.

Bounds on the norm of a resolvent occur in various situations in pure and applied mathematics. For example in operator theory, when constructing a functional calculus [17],

$$f \mapsto f(X) = \frac{1}{2\pi i} \int_{\gamma} f(\zeta)(\zeta - X)^{-1} d\zeta$$

in the theory of orthogonal polynomials, when studying the location of zeros of random orthogonal polynomials on the unit circle [6]. In computational linear algebra the following are classical problems that can be approached through appropriate estimates for $\|(\zeta - A)^{-1}\|$.

- (1) To analyze the stability of solutions x of the matrix equation $Ax - \zeta x = b$ under perturbations in b and A , see [5].
- (2) To study whether an approximate eigenvalue ζ of A (in the sense that, for some vector $x \neq 0$, $\|Ax - \zeta x\| \leq \varepsilon \|x\|$) is close to an actual eigenvalue of A , see [28, 19, 5].
- (3) To estimate the distance of the spectrum of a matrix B to the spectrum of a matrix A in terms of $B - A$, see [21, 19, 2].

Our resolvent bounds are stronger than the ones used for example in [21] to obtain estimates on the spectral variation of non-normal matrices. In Section 5 we apply our estimate for power-bounded matrices to study the sensitivity of stationary states of a classical or quantum Markov chain under perturbations of the transition matrix. We recover known stability results for classical Markov chains and prove new estimates in the quantum case. A similar approach, based on the power-boundedness of the transition matrix, was previously applied in [26] to investigate spectral convergence properties of classical and quantum Markov chains.

Acknowledgements. The author is thankful to Michael M. Wolf for creating conditions that made this work possible and to Alexander Müller-Hermes for proof-reading the manuscript and for pointing out the simpler proof for Lemma 3.7. The author is equally thankful to E. B. Davies for valuable comments on a previous version of the manuscript.

2. Preliminaries

We will take a function space based approach to the problem of bounding the norm of the resolvent of a certain matrix. This section lays down the required definitions and basic results.

2.1. Notation. We denote by \mathcal{M}_n the set of $n \times n$ matrices with complex entries. For $A \in \mathcal{M}_n$ we denote by $\sigma(A)$ its spectrum and by m its minimal polynomial. We write $|m|$ for the degree of m . To the minimal polynomial m we associate the Blaschke product

$$B(z) := \prod_i \frac{z - \lambda_i}{1 - \bar{\lambda}_i z}.$$

The product is taken over all i such that (respecting multiplicities) the corresponding linear factor $z - \lambda_i$ occurs in the minimal polynomial m . Thus, the numerator of B as defined here is exactly the associated minimal polynomial.

We denote by $\|A\|$ any particular norm of A while the ∞ -norm is defined by

$$\|A\|_\infty = \sup_{\|v\|_2=1} \|Av\|_2,$$

where $\|v\|_2^2 = \sum_i |v_i|^2$ is the usual Euclidean norm. That means $\|A\|_\infty$ simply denotes the largest singular value of A . We will slightly abuse nomenclature and call matrices with

$$\|A\|_\infty \leq 1$$

Hilbert space contractions, although of course the underlying space always has finite dimension. Similarly, the class of $A \in \mathcal{M}_n$ with

$$\sup_{k \geq 0} \|A^k\| \leq C < \infty$$

will be called Banach space power-bounded operators with respect to $\|\cdot\|$ and constant C . (Note that here the norm is general.)

Let $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ denote the open unit disk in the complex plane and $\bar{\mathbb{D}}$ its closure. The space of analytic functions on \mathbb{D} is denoted by $\text{Hol}(\mathbb{D})$. The Hardy spaces considered here are

$$H_2 := \left\{ f \in \text{Hol}(\mathbb{D}) \mid \|f\|_{H_2}^2 := \sup_{0 \leq r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\phi})|^2 d\phi < \infty \right\},$$

and

$$H_\infty := \{f \in \text{Hol}(\mathbb{D}) \mid \|f\|_{H_\infty} := \sup_{z \in \mathbb{D}} |f(z)| < \infty\}.$$

The H_2 -norm can be written in terms of the Taylor coefficients of the analytic function f . We write $f(z) = \sum_{k \geq 0} \hat{f}(k)z^k$ and use Plancherel's identity to conclude that

$$\sup_{0 \leq r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\phi})|^2 d\phi = \sum_{k \geq 0} |\hat{f}(k)|^2.$$

Thus, $f \in \text{Hol}(\mathbb{D})$ is in H_2 if and only if $\sum_{k \geq 0} |\hat{f}(k)|^2 < \infty$. The Wiener algebra is defined as the subset of $\text{Hol}(\mathbb{D})$ of absolutely convergent Taylor series,

$$W := \left\{ f = \sum_{k \geq 0} \hat{f}(k)z^k \mid \|f\|_W := \sum_{k \geq 0} |\hat{f}(k)| < \infty \right\}.$$

2.2. Model spaces and operators. Let $A \in \mathcal{M}_n$ with $\sigma(A) \subset \mathbb{D}$ and let B be the Blaschke product associated to the minimal polynomial of A . We define the $|m|$ -dimensional *model space*

$$K_B := H_2 \ominus BH_2 := H_2 \cap (BH_2)^\perp,$$

where we employ the usual scalar product from the Hilbert space $L^2(\partial\mathbb{D})$,

$$\langle f \mid g \rangle := \int_{\partial\mathbb{D}} f(z)\overline{g(z)} \frac{|dz|}{2\pi}.$$

If the zeros $\{\lambda_i\}_{i=1, \dots, |m|}$ of B are distinct (that is A can be diagonalized) it is not difficult to verify that K_B is spanned by the *Cauchy kernels*

$$K_B = \text{span} \left\{ \frac{1}{1 - \bar{\lambda}_i z} \right\}_{i=1, \dots, |m|}.$$

Thus K_B is a space of rational functions f of the form

$$f(z) = \frac{p(z)}{\prod_i (1 - \bar{\lambda}_i z)},$$

where $p(z)$ is a polynomial of degree at most $|m| - 1$. If the zeros of B are not distinct the above remains valid but the Cauchy kernels have to be replaced by

$$\frac{z^{k-1}}{(1 - \bar{\lambda}_i z)^k}, \quad 1 \leq k \leq k_i,$$

where k_i denotes the multiplicity of λ_i . In our consecutive proofs, however, we omit this case and assume that A is diagonalizable. This does not result in any difficulties since upper bounds obtained in the special case extend by continuity to

bounds for non-diagonalizable matrices. The assumption that A can be diagonalized is not principal; virtually all computations in the manuscript can be carried out in the more general case. We rely on continuity only for notational convenience.

One natural orthonormal basis $\{e_k\}_{k=1,\dots,|m|}$ for K_B is the Malmquist–Walsh basis with ([16], page 117)

$$e_k(z) := \frac{(1 - |\lambda_k|^2)^{1/2}}{1 - \bar{\lambda}_k z} \prod_{i=1}^{k-1} \frac{z - \lambda_i}{1 - \bar{\lambda}_i z},$$

where, as it will remain throughout the manuscript, the empty product is defined to be 1 i.e.

$$e_1(z) = \frac{(1 - |\lambda_1|^2)^{1/2}}{1 - \bar{\lambda}_1 z}.$$

The *model operator* M_B acts on K_B as

$$\begin{aligned} M_B : K_B &\longrightarrow K_B, \\ f &\longmapsto M_B(f) := P_B(zf), \end{aligned}$$

where P_B denotes the orthogonal projection on K_B . In other words, M_B is the compression of the multiplication operation by z to the model space K_B (see [16] for a detailed discussion of model operators and spaces). As multiplication by z has operator norm 1 it is clear that M_B is a Hilbert space contraction. Moreover, it is not hard to show that the eigenvalues of M_B are exactly the zeros of the corresponding Blaschke product (see [18], page 228 and Proposition 3.5 in the article at hand).

2.3. Spectral bounds on the norm of a function of a matrix. This subsection contains a brief outline of methods to obtain spectral bounds on a function of a matrix. For a more detailed account see [17, 16, 18] and the references therein. Suppose that f is holomorphic on a domain containing all eigenvalues of A and let γ be a smooth curve in this domain that encloses the eigenvalues. The matrix $f(A)$ is defined by the Dunford–Taylor integral [9]

$$f(A) = \frac{1}{2\pi i} \int_{\gamma} f(\zeta)(\zeta - A)^{-1} d\zeta.$$

It is easily seen that if $f(z) = \sum_{k=0}^l a_k z^k$ is a polynomial then $f(A) = \sum_{k=0}^l a_k A^k$ and that the correspondence $f \mapsto f(A)$ is an algebra homomorphism from the algebra of holomorphic functions (on the given domain) to \mathcal{M}_n

i.e. $(f + g)(T) = f(T) + g(T)$ and $(fg)(T) = f(T)g(T)$ (see [9], Chapter I.6). A unital Banach algebra \mathcal{A} with elements in $\text{Hol}(\mathbb{D})$ will be called a function algebra if

- (i) \mathcal{A} contains all polynomials and $\lim_{n \rightarrow \infty} \|z^n\|_{\mathcal{A}}^{1/n} = 1$ and
- (ii) $(f \in \mathcal{A}, \lambda \in \mathbb{D}, f(\lambda) = 0)$ implies that $\frac{f}{z-\lambda} \in \mathcal{A}$.

Following the conventions of [17] we say that a set of matrices Γ obeys an \mathcal{A} functional calculus with constant C if

$$\|f(A)\| \leq C \|f\|_{\mathcal{A}},$$

holds for any $A \in \Gamma$ and $f \in \mathcal{A}$. Here $\|f\|_{\mathcal{A}}$ denotes the norm of f in \mathcal{A} . Clearly, this is only possible if all eigenvalues of A are contained in $\overline{\mathbb{D}}$. For us, two instances of such inequalities will be important. In the first example we consider Hilbert space contractions, while the second one treats power-bounded Banach space operators.

i) The family of Hilbert space contractions $\Gamma = \{A \in \mathcal{M}_n \mid \|A\|_{\infty} \leq 1\}$ is related to an H_{∞} functional calculus, since by von Neumann’s inequality [14, 5] we have for any f in the disk algebra $H_{\infty} \cap \mathcal{C}(\overline{\mathbb{D}})$ (the set of bounded holomorphic functions on \mathbb{D} that admit a continuous extension to the boundary) and $A \in \Gamma$ with $\sigma(A) \subset \mathbb{D}$

$$\|f(A)\|_{\infty} \leq \|f\|_{H_{\infty}}.$$

ii) Consider a family $\Gamma = \{A \in \mathcal{M}_n \mid \|A^k\| \leq C \ \forall k \in \mathbb{N}\}$ of Banach space operators that are power bounded by some constant $C < \infty$. This family admits a Wiener algebra functional calculus since, for any $f \in W$ and $A \in \Gamma$,

$$\begin{aligned} \|f(A)\| &= \left\| \sum_{k \geq 0} \hat{f}(k) A^k \right\| \leq \sum_{k \geq 0} |\hat{f}(k)| \|A^k\| \\ &\leq C \sum_{k \geq 0} |\hat{f}(k)| = C \|f\|_W \end{aligned}$$

holds.

At first glance, the inequalities of *i)* and *ii)* seem to be of little use when it comes to finding spectral bounds on $\|f(A)\|$ since the obtained upper bounds do not depend on A anymore. To obtain better estimates one can rely on the following insight. Instead of considering the function f directly, we add multiples of m (or any other annihilating polynomial) to this function and consider $h = f + mg$, $g \in \mathcal{A}$ instead of f . It is immediate that $\|f(X)\| = \|h(X)\|$. The following simple but crucial lemma summarizes this point.

Lemma 2.1 ([17] Lemma 3.1). *Let $m \neq 0$ be a polynomial and let Γ be a set of matrices that obey an \mathcal{A} functional calculus with constant C and that satisfy $m(A) = 0 \forall A \in \Gamma$. Then*

$$\|f(A)\| \leq C \|f\|_{\mathcal{A}/m\mathcal{A}}, \quad \text{for all } A \in \Gamma,$$

where $\|f\|_{\mathcal{A}/m\mathcal{A}} = \inf \{\|h\|_{\mathcal{A}} \mid h = f + mg, g \in \mathcal{A}\}$.

Proof. For any $g \in \mathcal{A}$ we have that $\|f(A)\| = \|(f + mg)(A)\| \leq C \|f + mg\|_{\mathcal{A}}$. □

If $\sigma(A) \subset \mathbb{D}$ (and A can be diagonalized) it follows directly from the definition of the function algebra (see also [17], Section 3.1 (iii) or [30], Section 1.2 P4) that

$$\|f\|_{\mathcal{A}/m\mathcal{A}} = \inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g(\lambda_i) = f(\lambda_i) \forall \lambda_i \in \sigma(A)\},$$

which is a Nevanlinna–Pick type interpolation problem in \mathcal{A} . If the eigenvalue λ_i carries a multiplicity $k_i > 1$ in m the above remains valid but at λ_i the first $k_i - 1$ derivatives of f and g must coincide. Since for $\sigma(A) \subset \mathbb{D}$ the Blaschke product is holomorphic on a set containing $\bar{\mathbb{D}}$ we can define $\|f\|_{\mathcal{A}/B\mathcal{A}}$ as in Lemma 2.1 and note ([30], Lemma 3.1) that as before

$$\|f\|_{\mathcal{A}/B\mathcal{A}} = \inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g(\lambda_i) = f(\lambda_i) \forall \lambda_i \in \sigma(A)\}.$$

In the special case $\mathcal{A} = H_\infty$ it is possible to evaluate $\|f\|_{H_\infty/BH_\infty}$ using Sarason’s approach to the Nevanlinna–Pick problem [22, 1] or the Commutant lifting theorem of B. Sz.-Nagy and C. Foiaş [8, 13, 1].

Lemma 2.2 ([17] Theorem 3.12, [18] Theorem 3.1.11). *For any $f \in H_\infty$ it holds that*

$$\|f\|_{H_\infty/BH_\infty} = \|f(M_B)\|_\infty.$$

3. Hilbert space contractions

Spectral bounds on the resolvent of a Hilbert space contraction were derived in [6]. The authors provide an upper bound in terms of a certain Toeplitz matrix, compute the norm of this matrix and present a sequence of matrices that approaches their upper bound. The following theorem summarizes the basic three assertions from the discussion of Hilbert spaces contractions in [6].

Theorem 3.1 ([6, 5]). (1) Let A be an $n \times n$ matrix with $\|A\|_\infty \leq 1$ and $1 \notin \sigma(A)$. Then

$$\|(\mathbb{1} - A)^{-1}\|_\infty \leq \frac{\|M_n\|_\infty}{\min_{\lambda_i \in \sigma(A)} |1 - \lambda_i|},$$

with the $n \times n$ matrix

$$M_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 2 & \dots & 2 & 1 \end{pmatrix}.$$

(2) It holds that $\|M_n\|_\infty = \cot(\frac{\pi}{4n})$.

(3) For any $a \in (0, 1)$ there are $n \times n$ matrices $A_n(a)$ with $\|A_n(a)\|_\infty \leq 1$ and $\sigma(A) = \{a\}$ such that

$$\lim_{a \rightarrow 1} (1 - a)(\mathbb{1} - A_n(a))^{-1} = M_n.$$

In this paper we recover the statements (1) and (3) using a unified approach based on the techniques developed in [17]. Our strategy is to directly compute and bound the entries of the model operator in Malmquist–Walsh basis. Our approach has the advantage that it yields spectral bounds for any $\zeta \in \mathbb{C} - \sigma(A)$ and that the optimality statement (3) is automatic. Concerning the second point of the theorem we present a technique going back to [7] in order to compute the norm of Toeplitz matrices of the form

$$M_n(\beta) := \begin{pmatrix} 1 & 0 & \dots & 0 \\ \beta & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \beta & \dots & \beta & 1 \end{pmatrix}, \quad \beta \in [0, 2]. \tag{3}$$

Theorem 3.2. Let A be an $n \times n$ matrix with $\|A\|_\infty \leq 1$ and minimal polynomial $m = \prod_{i=1}^{|m|} (z - \lambda_i)$ with $\sigma(A) \subset \mathbb{D}$. Then for any $\zeta \in \mathbb{C} - \sigma(A)$ it holds that

$$\|(\zeta - A)^{-1}\|_\infty \leq \|(\zeta - M_B)^{-1}\|_\infty$$

and

$$((\zeta - M_B)^{-1})_{ij} = \begin{cases} 0 & \text{if } i < j, \\ \frac{1}{\zeta - \lambda_i} & \text{if } i = j, \\ \frac{(1 - |\lambda_i|^2)^{1/2} (1 - |\lambda_j|^2)^{1/2}}{\zeta - \lambda_i} \prod_{\mu=j+1}^{i-1} \left(\frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right) & \text{if } i > j, \end{cases}$$

with respect to the Malmquist–Walsh basis. (The empty product is defined to be 1.)

To compare our new result Theorem 3.2 to Theorem 3.1 we note that for any $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$, the condition $|a_{ij}| \leq b_{ij} \forall i, j$ implies that $\|A\|_\infty \leq \|B\|_\infty$. Suppose for instance that $|\zeta| \leq 1$. Then we can estimate the off-diagonal components of $(\zeta - M_B)^{-1}$ by

$$\begin{aligned} & \left| \frac{(1 - |\lambda_i|^2)^{1/2} (1 - |\lambda_j|^2)^{1/2}}{1 - \bar{\lambda}_i \zeta} \frac{1}{1 - \bar{\lambda}_j \zeta} \prod_{\mu=j}^i \left(\frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right) \right| \\ & \leq \max_i \frac{1 - |\lambda_i|^2}{|1 - \bar{\lambda}_i \zeta|^2} \prod_{\mu=1}^{|m|} \left| \frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right| \\ & \leq \max_i \frac{1}{|1 - \bar{\lambda}_i \zeta|} \max_i \frac{1 - |\lambda_i|^2}{|1 - \bar{\lambda}_i \zeta|} \prod_{\mu=1}^{|m|} \left| \frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right| \\ & \leq \max_i \frac{2}{|1 - \bar{\lambda}_i \zeta|} \prod_{\mu=1}^{|m|} \left| \frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right|, \end{aligned}$$

which yields the component-wise estimate

$$|((\zeta - M_B)^{-1})_{ij}| \leq \frac{1}{\min_{\lambda_k \in \sigma(A)} |1 - \bar{\lambda}_k \zeta|} \prod_{\mu=1}^{|m|} \left| \frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right| \cdot \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j, \\ 2 & \text{if } i > j. \end{cases}$$

Corollary 3.3. *Under the assumptions of Theorem 3.2 suppose that $|\zeta| \leq 1$. It follows*

$$\|(\zeta - A)^{-1}\|_\infty \leq \frac{\|M_{|m|}\|_\infty}{\min_{\lambda_k \in \sigma(A)} |1 - \bar{\lambda}_k \zeta|} \frac{1}{|B(\zeta)|},$$

where

$$B(\zeta) = \prod_{i=1}^{|m|} \frac{\zeta - \lambda_i}{1 - \bar{\lambda}_i \zeta}$$

is the Blaschke product associated with m .

We can pass to the general case $\sigma(A) \subset \bar{\mathbb{D}}$ by continuous extension. Setting $\zeta = 1$ Corollary 3.3 is the first assertion of Theorem 3.1 with the bonus that on the right hand side the norm of an $|m| \times |m|$ matrix occurs (compare [6] Section 6 B). However, if $\max_i \frac{1-|\lambda_i|^2}{|1-\lambda_i|} = \beta$ is given we have (with the same computation as above) for $\zeta = 1$

$$|((1 - M_B)^{-1})_{ij}| \leq \frac{1}{\min_{\lambda_k \in \sigma(A)} |1 - \lambda_k|} \cdot \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j, \\ \beta & \text{if } i > j, \end{cases}$$

and we can improve the bound in Theorem 3.1 if we can compute $\|M_n(\beta)\|_\infty$ (see (3)). The following theorem generalizes the discussion of Toeplitz matrices in [6]. It establishes an indirect possibility to compute $\|M_n(\beta)\|_\infty$.

Proposition 3.4. *Let $M_n(\beta)$ with $\beta \in (0, 2]$ be the $n \times n$ Toeplitz matrix introduced in (3). Then the equation*

$$\beta \cot(n\theta) + (2 - \beta) \cot(\theta/2) = 0, \quad \theta \in \mathbb{R} \tag{4}$$

has a unique solution $\theta^* \in [\frac{2n-1}{2n}\pi, \pi)$ and

$$\|M_n(\beta)\|_\infty = \frac{1}{2} \sqrt{(\beta - 2)^2 + \frac{\beta^2}{\cot^2(\theta^*/2)}}.$$

In particular it holds that

$$\begin{aligned} \|M_n(0)\|_\infty &= 1, \\ \|M_n(1)\|_\infty &= \frac{1}{2 \sin\left(\frac{\pi}{4n+2}\right)} \end{aligned}$$

and

$$\|M_n(2)\|_\infty = \cot\left(\frac{\pi}{4n}\right).$$

It is possible to expand $\cot(n\theta)$ in Equation (4) in terms of $\cot(\theta/2)$, which yields a polynomial equation in $\cot(\theta/2)$. Since $\|M_n(\beta)\|_\infty$ only depends on $\cot(\theta/2)$ (and β) computing $\|M_n(\beta)\|_\infty$ is equivalent to finding the unique zero of the resulting polynomial in the interval $(0, \cot(\frac{2n-1}{4n}\pi)]$ as a function of β .

Finally, statement (3) of Theorem 3.1 can be recovered from Theorem 3.2 with the choice of a minimal polynomial $m = (z - a)^n$, $a \in (0, 1)$ and setting $A_n(a) = M_B$. In this case we have for $1 \leq i, j \leq n$ that

$$|((\mathbb{1} - M_B)^{-1})_{ij}| = \frac{1}{1 - a} \cdot \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j, \\ 1 + a & \text{if } i > j. \end{cases}$$

Letting $a \rightarrow 1$ proves item (3) of Theorem 3.1. In the following Subsection 3.1 we compute the entries of M_B with respect to the Malmquist–Walsh basis. This yields a simple form for matrices that achieve equality in Theorem 3.2 i.e. for A with largest $\|(\zeta - A)^{-1}\|_\infty$ for a given spectrum.

Proposition 3.5. *The components of the model operator M_B with respect to Malmquist–Walsh basis are given by*

$$(M_B)_{ij} = \begin{cases} 0 & \text{if } i < j, \\ \lambda_i & \text{if } i = j, \\ (1 - |\lambda_i|^2)^{1/2}(1 - |\lambda_j|^2)^{1/2} \prod_{\mu=j+1}^{i-1} (-\bar{\lambda}_\mu) & \text{if } i > j. \end{cases}$$

Hence, one explicit form of the matrices $A_n(a)$ in Theorem 3.1 is

$$A_n(a) := \begin{pmatrix} a & 0 & \dots & \dots & 0 \\ 1 - a^2 & a & \ddots & & \vdots \\ -a(1 - a^2) & 1 - a^2 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ (-1)^n a^{(n-2)}(1 - a^2) & \dots & -a(1 - a^2) & 1 - a^2 & a \end{pmatrix}.$$

Finally, we note that Theorem 3.2 is stronger than Theorem 3.1 in that it holds for *general* ζ and yields an *optimal* bound for *general* spectra.

The rest of this section is organized in two subsections. The first, Subsection 3.1, contains a proof of Theorem 3.2 and Proposition 3.5 while in Subsection 3.2 we prove Proposition 3.4.

3.1. A model operator approach to resolvent bounds. As mentioned before our approach is to bound a function of a matrix in terms of the norm of a representative function. A key role is played by Lemma 2.1, which however requires that $f \in \mathcal{A}$. In order to derive upper bounds for rational functions such as the resolvent we need to extend Lemma 2.1. The following is based on the techniques of [17], Lemma 3.2 for the discussion of inverses. Here, we present an extension, which is adapted to our purposes.

Lemma 3.6. *Let A be an $n \times n$ matrix with $\sigma(A) \subset \mathbb{D}$ and let ψ be a rational function with poles $(\xi_i)_{i=1,\dots,k}$ such that $\bigcup_i \{\xi_i\} \cap \sigma(A) = \emptyset$.*

(1) *If A obeys an \mathcal{A} -functional calculus with constant C then*

$$\|\psi(A)\| \leq C \inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g(\lambda_i) = \psi(\lambda_i) \ i = 1, \dots, n\}.$$

(2) *If $\|A\|_{\infty} \leq 1$ holds then $\|\psi(A)\|_{\infty} \leq \|\psi(M_B)\|_{\infty}$.*

Proof. We extend Lemma 2.1 to the situation, when ψ is rational. Define

$$\varphi := \psi \cdot \prod_{j=1}^k \left(\frac{m(\xi_j) - m}{m(\xi_j)} \right)^{k_j},$$

where k_j denotes the multiplicity of the pole at ξ_j and note that φ is polynomial and that $\psi(A) = \varphi(A)$. It follows using Lemma 2.1 that

$$\begin{aligned} \|\psi(A)\| &= \|\varphi(A)\| \\ &\leq C \|\varphi\|_{\mathcal{A}/m\mathcal{A}} = C \inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g(\lambda_i) \\ &= \varphi(\lambda_i) \ i = 1, \dots, n\} \\ &= C \inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g(\lambda_i) = \psi(\lambda_i) \ i = 1, \dots, n\}, \end{aligned}$$

which proves the first assertion. For the second one we consider the same φ as above and note that

$$\|\psi(A)\|_{\infty} = \|\varphi(A)\|_{\infty} \leq \|\varphi\|_{H_{\infty}/BH_{\infty}} = \|\varphi(M_B)\|_{\infty},$$

where we applied Lemma 2.2 in the last step. But as $m(M_B) = 0$ it follows that $\varphi(M_B) = \psi(M_B)$. □

Let us remark that Lemma 3.6 remains valid if the eigenvalue λ_i carries degeneracy k_i in m . The point here is to replace the inf on the right hand side of (1) with $\inf\{\|g\|_{\mathcal{A}} \mid g \in \mathcal{A}, g^{(k)}(\lambda_i) = \psi^{(k)}(\lambda_i), 0 \leq k < k_i\}$, where the superscript k denotes the k -th derivative.

Lemma 3.7. *Let $\{\lambda_i\}_{i=1,\dots,n} \subset \mathbb{D}$ and let $\zeta \in \mathbb{C} - \{\lambda_i\}_{i=1,\dots,n}$ and $j < i$ then*

$$\sum_{\mu=j}^i \frac{1}{\zeta - \lambda_\mu} \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_\mu)}{\prod_{v: v \neq \mu} (\lambda_\mu - \lambda_v)} = \frac{1}{(1 - \bar{\lambda}_i \zeta)(1 - \bar{\lambda}_j \zeta)} \prod_{\mu=j}^i \left(\frac{1 - \bar{\lambda}_\mu \zeta}{\zeta - \lambda_\mu} \right).$$

Proof of Lemma 3.7. We present two proofs for this lemma. The first one arises naturally in the context of H_2 spaces (see the proof of Theorem 3.2), while the second one is more direct and simple. We define

$$t(z) := \frac{z}{\zeta - z} \frac{1}{(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)}$$

and the (truncated) Blaschke product

$$B_{ji}(z) := \prod_{\mu=j}^i \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z}$$

and compute the $L_2(\partial\mathbb{D})$ scalar product. Suppose for now that $|\zeta| > 1$ then

$$\begin{aligned} \langle t \mid B_{ji} \rangle &= \int_0^{2\pi} t(z) \overline{B_{ji}(z)} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \int_0^{2\pi} t(z) \prod_{\mu} \frac{1 - \bar{\lambda}_\mu z}{z - \lambda_\mu} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{(\zeta - z)(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)} \prod_{\mu} \frac{1 - \bar{\lambda}_\mu z}{z - \lambda_\mu} dz \\ &= \sum_{\mu=j}^i \frac{1}{\zeta - \lambda_\mu} \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_\mu)}{\prod_{v: v \neq \mu} (\lambda_\mu - \lambda_v)}, \end{aligned}$$

where in the last step we applied the Residue theorem and made use of the assumption $|\zeta| > 1$. On the other hand

$$\begin{aligned} \langle B_{ji} \mid t \rangle &= \int_0^{2\pi} B_{ji}(z) \overline{t(z)} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \prod_{\mu} \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z} \frac{1}{\bar{\zeta} z - 1} \frac{z}{(z - \lambda_i)(z - \lambda_j)} dz \\ &= \frac{1}{(1 - \lambda_i \bar{\zeta})(1 - \lambda_j \bar{\zeta})} \prod_{\mu} \frac{1 - \lambda_\mu \bar{\zeta}}{\bar{\zeta} - \bar{\lambda}_\mu}. \end{aligned}$$

Clearly, $\langle t \mid B_{ji} \rangle = \overline{\langle B_{ji} \mid t \rangle}$ from which the lemma follows for $|\zeta| > 1$. In case that $|\zeta| < 1$ we compute similarly

$$\begin{aligned} \langle t \mid B_{ji} \rangle &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{(\zeta - z)(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)} \prod_{\mu} \frac{1 - \bar{\lambda}_{\mu} z}{z - \lambda_{\mu}} dz \\ &= \sum_{\mu=j}^i \frac{1}{\zeta - \lambda_{\mu}} \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_{\mu})}{\prod_{v: v \neq \mu} (\lambda_{\mu} - \lambda_v)} - \frac{1}{(1 - \bar{\lambda}_i \zeta)(1 - \bar{\lambda}_j \zeta)} \prod_{\mu} \frac{1 - \bar{\lambda}_{\mu} \zeta}{\zeta - \lambda_{\mu}} \end{aligned}$$

and

$$\langle B_{ji} \mid t \rangle = 0.$$

The case $|\zeta| = 1$ follows by continuity. For the second proof we multiply both sides of the lemma with $\prod_{\mu=j}^i (\zeta - \lambda_{\mu})$ to obtain a polynomial equation in ζ

$$\sum_{\mu=j}^i \prod_{v: v \neq \mu} (\zeta - \lambda_v) \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_{\mu})}{\prod_{v: v \neq \mu} (\lambda_{\mu} - \lambda_v)} = \prod_{\mu: \mu \neq i, \mu \neq j} (1 - \bar{\lambda}_{\mu} \zeta).$$

The polynomial on the left hand side has degree at most $i - j$ and the degree of the polynomial on the right hand side is $i - j - 1$. Two polynomials of a certain degree n are the same if and only if they coincide at $n + 1$ nodes. We choose the $i - j + 1$ values $\{\lambda_{\alpha}\}_{j \leq \alpha \leq i}$ and verify that for this choice equality indeed holds:

$$\begin{aligned} &\sum_{\mu=j}^i \prod_{v: v \neq \mu} (\zeta - \lambda_v) \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_{\mu})}{\prod_{v: v \neq \mu} (\lambda_{\mu} - \lambda_v)} \Big|_{\zeta=\lambda_{\alpha}} \\ &= \prod_{v: v \neq \alpha} (\lambda_{\alpha} - \lambda_v) \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_{\alpha})}{\prod_{v: v \neq \alpha} (\lambda_{\alpha} - \lambda_v)} \\ &= \prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_{\alpha}). \quad \square \end{aligned}$$

We are now ready to present a proof of Theorem 3.2.

Proof of Theorem 3.2. The first assertion follows directly from Lemma 3.6. To compute the matrix entries of $(\zeta - M_B)^{-1}$ with respect to Malmquist–Walsh basis we recall that

$$(\zeta - M_B)^{-1} = \varphi(M_B),$$

where $\varphi(z) := \frac{1}{\zeta - z} \frac{m(\zeta) - m(z)}{m(\zeta)}$ is a polynomial. We have that

$$\begin{aligned}
 & ((\zeta - M_B)^{-1})_{ij} \\
 &= \langle \varphi(M_B) e_j \mid e_i \rangle \\
 &= \langle P_B(\varphi e_j) \mid e_i \rangle \\
 &= \langle \varphi e_j \mid e_i \rangle \\
 &= \int_0^{2\pi} \varphi(z) e_j(z) \overline{e_i(z)} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\
 &= \frac{((1 - |\lambda_i|^2)(1 - |\lambda_j|^2))^{1/2}}{2\pi i} \int_{\partial\mathbb{D}} \varphi(z) \frac{1}{(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)} \\
 &\qquad\qquad\qquad \prod_{\mu=1}^{j-1} \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z} \prod_{v=1}^i \frac{1 - \bar{\lambda}_v z}{z - \lambda_v} dz.
 \end{aligned} \tag{5}$$

In case that $j > i$ the integrand is holomorphic on \mathbb{D} . Hence, the integral in (5) is zero. If $j = i$ we have that

$$\frac{(1 - |\lambda_i|^2)}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{\zeta - z} \frac{m(\zeta) - m(z)}{m(\zeta)} \frac{1}{(1 - \bar{\lambda}_i z)(z - \lambda_i)} dz = \frac{1}{\zeta - \lambda_i}.$$

Finally if $j < i$, then the integral in (5) equals

$$\begin{aligned}
 & \frac{((1 - |\lambda_i|^2)(1 - |\lambda_j|^2))^{1/2}}{2\pi i} \int_{\partial\mathbb{D}} \mathfrak{X} \prod_{v=j}^i \frac{1 - \bar{\lambda}_v z}{z - \lambda_v} dz \\
 &= ((1 - |\lambda_i|^2)(1 - |\lambda_j|^2))^{1/2} \sum_{\mu=j}^i \frac{1}{\zeta - \lambda_\mu} \frac{\prod_{v: v \neq i, v \neq j} (1 - \bar{\lambda}_v \lambda_\mu)}{\prod_{v: v \neq \mu} (\lambda_\mu - \lambda_v)},
 \end{aligned}$$

where

$$\mathfrak{X} = \frac{1}{\zeta - z} \frac{m(\zeta) - m(z)}{m(\zeta)} \frac{1}{(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)}.$$

An application of Lemma 3.7 concludes the proof of Theorem 3.2. □

Proposition 3.5 is verified via a direct calculation.

Proof of Proposition 3.5. We proceed as in the derivation of Theorem 3.2 and conclude

$$(M_B)_{ij} = ((1 - |\lambda_i|^2)(1 - |\lambda_j|^2))^{1/2} \int_0^{2\pi} \frac{z^2}{(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)} \prod_{\mu=1}^{j-1} \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z} \prod_{v=1}^i \frac{1 - \bar{\lambda}_v z}{z - \lambda_v} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi}.$$

If $j > i$ the Residue theorem reveals that the integral is zero. Similarly, if $i = j$ the integral is given by λ_i . Finally if $i > j$ we compute

$$\begin{aligned} & \int_0^{2\pi} \frac{z^2}{(1 - \bar{\lambda}_i z)(1 - \bar{\lambda}_j z)} \prod_{\mu=j}^i \frac{1 - \bar{\lambda}_\mu z}{z - \lambda_\mu} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \int_0^{2\pi} \frac{1}{(z - \lambda_i)(z - \lambda_j)} \prod_{\mu=j}^i \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{z(z - \lambda_i)(z - \lambda_j)} \prod_{\mu=j}^i \frac{z - \lambda_\mu}{1 - \bar{\lambda}_\mu z} dz \\ &= \prod_{\mu=j+1}^{i-1} (-\lambda_\mu), \end{aligned}$$

where the last step again uses the Residue theorem. □

3.2. Computing the norm of certain Toeplitz matrices. In this subsection we prove Proposition 3.4 with a direct computation of $\|M_n(\beta)\|_\infty$. Our approach is guided by the techniques developed in [7]. $\|M_n(1)\|_\infty$ and $\|M_n(2)\|_\infty$ are computed in [6] and [5] (Lemma 9.6.5) following a different approach.

Proof of Proposition 3.4. Instead of working with

$$M_n(\beta) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \beta & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \beta & \dots & \beta & 1 \end{pmatrix}$$

directly, we consider the matrix

$$\tilde{M}_n(\beta) := \begin{pmatrix} \beta & \dots & \beta & 1 \\ \vdots & \ddots & 1 & 0 \\ \beta & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}$$

and note that

$$\|M_n(\beta)\|_\infty = \|\tilde{M}_n(\beta)\|_\infty.$$

As $\tilde{M}_n(\beta)$ is Hermitian all its eigenvalues are real and its ∞ -norm is simply the largest in magnitude eigenvalue. The eigenvalues of $\tilde{M}_n(\beta)^2$ are the eigenvalues of $\tilde{M}_n(\beta)$ squared. Hence, we are looking for the largest λ^2 such that

$$0 = \det(\tilde{M}_n(\beta)^2 - \lambda^2 \mathbb{1}) = \det(\tilde{M}_n(\beta) - \lambda \mathbb{1})(\tilde{M}_n(\beta) + \lambda \mathbb{1}).$$

Direct computation reveals that

$$\begin{aligned} & (\tilde{M}_n(\beta) - \lambda \mathbb{1})(\tilde{M}_n(\beta) + \lambda \mathbb{1}) \\ &= \begin{pmatrix} \beta-\lambda & \beta & \beta & \dots & \beta & 1 \\ \beta & \beta-\lambda & & & 1 & 0 \\ \beta & & & & & 0 \\ \vdots & & & & & \vdots \\ \beta & 1 & & & -\lambda & 0 \\ 1 & 0 & 0 & \dots & 0 & -\lambda \end{pmatrix} \cdot \begin{pmatrix} \beta+\lambda & \beta & \beta & \dots & \beta & 1 \\ \beta & \beta+\lambda & & & 1 & 0 \\ \beta & & & & & 0 \\ \vdots & & & & & \vdots \\ \beta & 1 & & & \lambda & 0 \\ 1 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \\ &= \begin{pmatrix} (n-1)\beta^2-\lambda^2+1 & (n-2)\beta^2+\beta & (n-3)\beta^2+\beta & \dots & \beta^2+\beta & \beta \\ (n-2)\beta^2+\beta & (n-2)\beta^2-\lambda^2+1 & (n-3)\beta^2+\beta & \dots & \beta^2+\beta & \beta \\ (n-3)\beta^2+\beta & (n-3)\beta^2+\beta & & & & \beta \\ \vdots & \vdots & & & & \vdots \\ \beta^2+\beta & \beta^2+\beta & & & \beta^2-\lambda^2+1 & \beta \\ \beta & \beta & \beta & \dots & \beta & -\lambda^2+1 \end{pmatrix}. \end{aligned}$$

We rearrange the resulting determinant by subtracting successively the second column from the first, the third from the second, the n -th from the $n - 1$ -th and leave the n -th unchanged. This yields

$$\begin{aligned} & \det(\tilde{M}_n(\beta)^2 - \lambda^2 \mathbb{1}) \\ &= \det \begin{pmatrix} \beta^2-\beta-\lambda^2+1 & \beta^2 & \beta^2 & \dots & \beta^2 & \beta \\ \beta+\lambda^2-1 & \beta^2-\beta-\lambda^2+1 & \beta^2 & \dots & \beta^2 & \beta \\ 0 & \beta+\lambda^2-1 & & & & \beta \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & & & \beta^2-\beta-\lambda^2+1 & \beta \\ 0 & 0 & 0 & \dots & \beta+\lambda^2-1 & -\lambda^2+1 \end{pmatrix}. \end{aligned}$$

Similarly, we subtract the second row from the first, the third from the second, the n -th from the $n - 1$ -th and leave the n -th unchanged. We conclude

$$\begin{aligned}
 & \det(\tilde{M}_n(\beta)^2 - \lambda^2 \mathbb{1}) \\
 &= \det \begin{pmatrix} \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 & 0 & \dots & 0 & 0 \\ \beta + \lambda^2 - 1 & \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 & 0 & \dots & 0 \\ 0 & \beta + \lambda^2 - 1 & & & & 0 \\ \vdots & 0 & & & & \vdots \\ 0 & \vdots & & & \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 \\ 0 & 0 & 0 & \dots & \beta + \lambda^2 - 1 & -\lambda^2 + 1 \end{pmatrix} \\
 &= \det \begin{pmatrix} \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 & 0 & \dots & 0 & 0 \\ \beta + \lambda^2 - 1 & \beta^2 - 2\beta - 2\lambda^2 + 2 & & \dots & & 0 \\ 0 & \beta + \lambda^2 - 1 & & & & 0 \\ \vdots & 0 & & & & \vdots \\ 0 & \vdots & & & \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 \\ 0 & 0 & 0 & \dots & \beta + \lambda^2 - 1 & \beta^2 - 2\beta - 2\lambda^2 + 2 \end{pmatrix} \tag{6} \\
 &+ \det \begin{pmatrix} \beta^2 - 2\beta - 2\lambda^2 + 2 & \beta + \lambda^2 - 1 & 0 & \dots & 0 & 0 \\ \beta + \lambda^2 - 1 & \beta^2 - 2\beta - 2\lambda^2 + 2 & & \dots & & 0 \\ 0 & \beta + \lambda^2 - 1 & & & & 0 \\ \vdots & 0 & & & & \vdots \\ 0 & \vdots & & & \beta^2 - 2\beta - 2\lambda^2 + 2 & 0 \\ 0 & 0 & 0 & 0 & \dots & \beta + \lambda^2 - 1 & \lambda^2 - (\beta - 1)^2 \end{pmatrix},
 \end{aligned}$$

where the last equality is a consequence of the linearity of \det in the last column. The following is a classical formula for the determinant of an $n \times n$ tri-diagonal Toeplitz matrix [7, 20]

$$\det \begin{pmatrix} x & 1 & 0 & \dots & 0 \\ 1 & x & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & x & 1 \\ 0 & \dots & 0 & 1 & x \end{pmatrix} = \frac{\sin(n + 1)\theta}{\sin \theta}, \quad x = 2 \cos \theta. \tag{7}$$

To apply this result we exclude the trivial case $\beta = 0$ and note that we can always assume that $\lambda^2 \geq 1$ such that $\beta + \lambda^2 - 1 > 0$ and $\frac{\beta^2}{\beta + \lambda^2 - 1} \in (0, \beta]$. Hence, we can divide all columns of both determinants of (6) by $\beta + \lambda^2 - 1$. We then expand the second determinant along its last column and apply (7) to both terms resulting

from (6). We find

$$\det(\tilde{M}_n(\beta)^2 - \lambda^2 \mathbb{1}) = (\beta + \lambda^2 - 1)^n \left(\frac{\sin(n+1)\theta}{\sin \theta} + \frac{\lambda^2 - (\beta - 1)^2}{\lambda^2 + (\beta - 1)} \frac{\sin n\theta}{\sin \theta} \right) \quad (8)$$

with

$$2 \cos \theta = \frac{\beta^2 - 2\beta - 2\lambda^2 + 2}{\lambda^2 + \beta - 1} = \frac{\beta^2}{\lambda^2 + \beta - 1} - 2.$$

Solving the latter for λ^2 gives

$$\lambda^2 = \frac{1}{4}((\beta - 2)^2 + \beta^2 \tan^2(\theta/2)),$$

where $\beta \neq 0$ implies that θ is such that the tangent is well defined. This enables us to eliminate λ^2 from (8) as

$$\frac{\lambda^2 - (\beta - 1)^2}{\lambda^2 + (\beta - 1)} = \frac{1}{\beta}(-\beta + 2 - 2\beta \cos \theta + 2 \cos \theta).$$

It follows that (8) is zero if and only if

$$\begin{aligned} 0 &= \beta \frac{\sin(n+1)\theta}{\sin \theta} + (-\beta + 2 - 2\beta \cos \theta + 2 \cos \theta) \frac{\sin n\theta}{\sin \theta} \\ &= \beta \cos n\theta + (2 - \beta)(1 + \cos \theta) \frac{\sin n\theta}{\sin \theta}, \end{aligned}$$

which in turn is equivalent to

$$\cot n\theta = \frac{\beta - 2}{\beta} \cot(\theta/2). \quad (9)$$

In total, we are looking for the solution θ^* of (9) such that λ^2 is maximal i.e. $\cot^2(\theta^*/2)$ is minimal. Since for any $\theta \in [\frac{2n-1}{2n}\pi, \pi)$ we have $\frac{\beta-2}{\beta} \cot(\theta/2) \leq 0$ with $\cot(\pi/2) = 0$ and $\lim_{\theta \uparrow \pi} \cot n\theta \rightarrow -\infty$ and $\cot \frac{2n-1}{2}\pi = 0$, it follows that there is a unique solution $\theta^* \in [\frac{2n-1}{2n}\pi, \pi)$ of Equation (9). Moreover, by the same fact, $\cot(\pi/2) = 0$, this solution maximizes λ^2 as desired.

Sometimes it is possible to obtain a solution for Equation (9) in closed form. Suppose $\beta = 2$, then $\cot n\theta^* = 0$ and $\theta^* = \frac{2n-1}{2n}\pi$. It follows

$$\|M_n(2)\|_\infty^2 = \tan^2\left(\frac{2n-1}{4n}\pi\right) = \cot^2(\pi/4n)$$

as in [6]. If $\beta = 1$ we have

$$\lambda^2 = \frac{1}{4 \cos^2(\theta/2)}$$

and

$$\sin(2n + 1)\theta/2 = 0$$

such that $\theta^* = \frac{2n\pi}{2n+1}$. It follows

$$\|M_n(1)\|_\infty^2 = \frac{1}{4 \cos^2\left(\frac{n\pi}{2n+1}\right)} = \frac{1}{4 \sin^2\left(\frac{\pi}{4n+2}\right)}$$

as in [7]. The trivial fact $\|M_n(0)\|_\infty = 1$ can be recovered by continuous extension as $\beta \rightarrow 0$. □

4. Power-bounded operators

It is natural to ask if power-boundedness of A is sufficient to obtain estimates on $\|(\zeta - A)^{-1}\|$ qualitatively similar to the results of Theorem 3.1, 3.2 and Corollary 3.3. In this section we prove that this is indeed the case and present a new bound on the norm of the resolvent of a power-bounded operator.

Theorem 4.1. *Let A be an $n \times n$ matrix with minimal polynomial m of degree $|m|$ and let $\|\cdot\|$ be an arbitrary matrix norm with $\sup_{k \geq 0} \|A^k\| = C < \infty$. For any $\zeta \in \bar{\mathbb{D}} - \sigma(A)$ it holds that*

$$\|(\zeta - A)^{-1}\| \leq \frac{2|m|C}{\mathfrak{D}} \left(\frac{4e}{|B(\zeta)|^2} - 1 \right)^{1/2},$$

where

$$\mathfrak{D} = \min_{\lambda_i \in \sigma(A)} |1 - \bar{\zeta}\lambda_i|^{1/2} (2|m| - 2|m||\zeta|^2 + |\zeta|^2) \min_{\lambda_i \in \sigma(A)} |1 - \bar{\zeta}\lambda_i|^{1/2}$$

and where

$$B(\zeta) = \prod_{i=1}^{|m|} \frac{\zeta - \lambda_i}{1 - \bar{\lambda}_i \zeta}$$

is the Blaschke product associated with m . For $|\zeta| > 1$, we have the obvious estimate $\|(\zeta - A)^{-1}\| \leq \frac{C}{|\zeta| - 1}$.

Theorem 4.1 is the analogue of Corollary 3.3 for power-bounded operators. Spectral bounds on the norm of the resolvent of a power-bounded operator are well studied in the literature. Theorem 6.4 of [6] treats the same problem in the special case that A is power-bounded with respect to operator norm $\|\cdot\|_\infty$. In [17,

Theorem 3.24] the behavior of $\|(\zeta - A)^{-1}\|$ is studied for $|\zeta| < 1$ and in [29] an upper bound is derived for $|\zeta| \geq 1$. Theorem 4.1 unifies the mentioned results and yields a quantitatively better bound in each case. To compare suppose that $|\zeta| < 1$ and note that in this case

$$1 - |\zeta|^2 + \frac{1}{2|m|}|\zeta|^2 \min_{\lambda_i \in \sigma(A)} |1 - \bar{\zeta}\lambda_i| \geq (1 - |\zeta|)^2$$

and of course $\min_{\lambda_i \in \sigma(A)} |1 - \bar{\zeta}\lambda_i| \geq 1 - |\zeta|$. Hence, it follows

$$\|(\zeta - A)^{-1}\| \leq \frac{\sqrt{8e|m|}C}{(1 - |\zeta|)^{3/2}} \frac{1}{|B(\zeta)|},$$

which is qualitatively the same as Theorem 3.24 in [17] but has a better numerical prefactor. If we choose $|\zeta| = 1$ it follows $|B(\zeta)| = 1$ and therefore

$$\|(\zeta - A)^{-1}\| \leq \frac{\sqrt{16e - 4} |m| C}{\min_{\lambda_i \in \sigma(A)} |\zeta - \lambda_i|}. \tag{10}$$

This bound improves on the result in [29] (which in turn is stronger than [6, Theorem 6.4]) as the new bound only grows linearly with $|m|$ as opposed to $|m|^{3/2}$ in [29]. That for power-bounded $A \in \mathcal{M}_n$ the correct asymptotic growth order for an upper bound is $\mathcal{O}(n)$ was already suspected in [6] and [31]. The bound obtained almost reaches the optimal estimate of Theorem 3.1 for Hilbert-space contractions. In the latter case we have that $\cot(\frac{\pi}{4n})/n \leq \frac{4}{\pi}$, while the prefactor of (10) is $\sqrt{16e - 4} \approx 6.28$. However, as is clear from the derivation, Inequality (10) is not optimal. We will use Inequality (10) to study the sensitivity of a classical or quantum Markov chain to perturbations in Section 5.

To prove Theorem 4.1 we take a similar approach as to Theorem 3.2. We note that power-bounded operators admit a Wiener algebra functional calculus. Thus an application of Lemma 3.6 reveals that

$$\|(\zeta - A)^{-1}\| \leq C \inf \left\{ \|g\|_W \mid g \in W, g(\lambda_i) = \frac{1}{\zeta - \lambda_i} \right\}. \tag{11}$$

The strategy of our proof will be to consider one specific representative function g in (11) and to bound its norm. To achieve this we employ the following method. Instead of considering g directly we choose a “smoothing parameter” r and pass to a “stretched” interpolation function. Given any function $f \in H_2$ and $r \in (0, 1)$, we write

$$f_r(z) := f(rz) = \sum_{k \geq 0} \hat{f}(k) r^k z^k$$

and observe that by the Cauchy-Schwarz inequality and the Plancherel identity

$$\|f_r\|_W \leq \sqrt{\sum_{k \geq 0} |\hat{f}(k)|^2} \sqrt{\frac{1}{1-r^2}} = \|f\|_{H_2} \sqrt{\frac{1}{1-r^2}}. \tag{12}$$

This idea was used to obtain bounds on the inverse and resolvent of a power-bounded operator in [17] and to study spectral convergence bounds for bounded semigroups in [27]. We use the Blaschke products

$$B(z) = \prod_i \frac{z - \lambda_i}{1 - \bar{\lambda}_i z} \quad \text{and} \quad \tilde{B}(z) = \prod_i \frac{z - r\lambda_i}{1 - r\bar{\lambda}_i z},$$

where in the latter product the spectrum is stretched by a factor of r . (The products are taken over all prime factors of m , but to avoid cumbersome notation we do not write this explicitly.) Consider now the function g with

$$g(z) = \sum_k \left(\frac{1}{\zeta - \lambda_k} \frac{\prod_j (1 - \bar{\lambda}_j \lambda_k)}{\prod_{j \neq k} (\lambda_k - \lambda_j)} \right) \frac{B(z)}{z - \lambda_k}.$$

Note that g is analytic in the unit disc and $g(\lambda_i) = \frac{1}{\zeta - \lambda_i}$ for all $\lambda_i \in \sigma(A)$. In order to use the estimate (12) we perform the aforementioned smoothing. We define the modified function \tilde{g} by

$$\tilde{g}(z) = \sum_k \left(\frac{1}{\zeta - \lambda_k} \frac{\prod_j (1 - r^2 \bar{\lambda}_j \lambda_k)}{\prod_{j \neq k} (r\lambda_k - r\lambda_j)} \right) \frac{\tilde{B}(z)}{z - r\lambda_k}$$

and observe that \tilde{g}_r enjoys the same basic properties as g i.e. \tilde{g}_r is analytic in \mathbb{D} and $\tilde{g}_r(\lambda_i) = \frac{1}{\zeta - \lambda_i}$ for any $\lambda_i \in \sigma(A)$. Thus, by Inequality (11), we have that $\|(\zeta - A)^{-1}\| \leq C \|\tilde{g}_r\|_W$ and it follows from Inequality (12) that

$$\|\tilde{g}_r\|_W \leq \sqrt{\frac{1}{1-r^2}} \|\tilde{g}\|_{H_2}.$$

It turns out that one can directly compute $\|\tilde{g}\|_{H_2}$. The computation relies on two combinatorial observations similar to Lemma 3.7, which we shall prove before we proceed with our discussion of $\|\tilde{g}\|_{H_2}$.

Lemma 4.2. *Let $|m| \in \mathbb{N} - \{0\}$ and $\{\lambda_i\}_{i=1, \dots, |m|} \subset \mathbb{D}$. Furthermore, let $\zeta \in \mathbb{C} - \{\lambda_i\}_{i=1, \dots, |m|}$ and $r \in (0, 1)$. It follows that*

$$\sum_{i=1}^{|m|} \frac{1}{\zeta - \lambda_i} \frac{\prod_{j: j \neq i} (1 - r^2 \bar{\lambda}_j \lambda_i)}{\prod_{j: j \neq i} (r\lambda_i - r\lambda_j)} = \frac{r}{1 - r^2 \bar{\lambda}_l \zeta} \prod_{i=1}^{|m|} \frac{1 - r^2 \bar{\lambda}_i \zeta}{r\zeta - r\lambda_i}, \tag{*}$$

$$\begin{aligned} & \sum_{i=1}^{|m|} \frac{1}{\zeta - \lambda_i} \frac{1}{1 - r^2 \bar{\zeta} \lambda_i} \frac{\prod_j (1 - r^2 \bar{\lambda}_j \lambda_i)}{\prod_{j: j \neq i} (r \lambda_i - r \lambda_j)} \\ &= \frac{r}{1 - r^2 |\zeta|^2} \left(\prod_{i=1}^{|m|} \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} - \prod_{i=1}^{|m|} \frac{r \bar{\zeta} - r \bar{\lambda}_i}{1 - r^2 \lambda_i \bar{\zeta}} \right), \end{aligned} \tag{**}$$

and

$$\|\tilde{g}\|_{H_2}^2 = \frac{r^2}{1 - r^2 |\zeta|^2} \left(\prod_{i=1}^{|m|} \left| \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \right|^2 - 1 \right). \tag{***}$$

Our proof is based on the Residue theorem. (It is also possible to prove the lemma with the second technique outlined in the proof of Lemma 3.7. However, the approach via the Residue theorem is more convenient for the second assertion.)

Proof. For the first assertion set $t_1(z) := \frac{rz}{r\zeta - z} \frac{1}{1 - r\bar{\lambda}_l z}$ and suppose for now that $r|\zeta| < 1$. We have that

$$\begin{aligned} \langle t_1 \mid \tilde{B} \rangle &= \int_0^{2\pi} \frac{rz}{r\zeta - z} \frac{1}{1 - r\bar{\lambda}_l z} \prod_i \frac{1 - r\bar{\lambda}_i z}{z - r\lambda_i} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{r}{r\zeta - z} \frac{1}{1 - r\bar{\lambda}_l z} \prod_i \frac{1 - r\bar{\lambda}_i z}{z - r\lambda_i} dz \\ &= \sum_i \frac{1}{\zeta - \lambda_i} \frac{\prod_{j: j \neq l} (1 - r^2 \bar{\lambda}_j \lambda_i)}{\prod_{j: j \neq i} (r \lambda_i - r \lambda_j)} - \frac{r}{1 - r^2 \bar{\lambda}_l \zeta} \prod_i \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \end{aligned}$$

and that

$$\begin{aligned} \langle \tilde{B} \mid t_1 \rangle &= \int_0^{2\pi} \prod_i \frac{z - r\lambda_i}{1 - r\bar{\lambda}_i z} \frac{r}{r\bar{\zeta} z - 1} \frac{z}{z - r\lambda_l} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \prod_i \frac{z - r\lambda_i}{1 - r\bar{\lambda}_i z} \frac{r}{r\bar{\zeta} z - 1} \frac{1}{z - r\lambda_l} dz \\ &= 0. \end{aligned}$$

Hence, for $r|\zeta| < 1$

$$\sum_i \frac{1}{\zeta - \lambda_i} \frac{\prod_{j: j \neq l} (1 - r^2 \bar{\lambda}_j \lambda_i)}{\prod_{j: j \neq i} (r \lambda_i - r \lambda_j)} = \frac{r}{1 - r^2 \bar{\lambda}_l \zeta} \prod_i \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i}$$

as claimed. An identical computation reveals that the above remains correct if $r|\zeta| > 1$ and the case $r|\zeta| = 1$ follows by continuity. For the second assertion suppose again that $r|\zeta| < 1$ and set $t_2(z) := \frac{rz}{r\zeta - z} \frac{1}{1 - r\bar{\zeta}z}$ and compute

$$\begin{aligned} \langle t_2 \mid \tilde{B} \rangle &= \int_0^{2\pi} \frac{rz}{r\zeta - z} \frac{1}{1 - r\bar{\zeta}z} \prod_i \frac{1 - r\bar{\lambda}_i z}{z - r\lambda_i} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{r}{r\zeta - z} \frac{1}{1 - r\bar{\zeta}z} \prod_i \frac{1 - r\bar{\lambda}_i z}{z - r\lambda_i} dz \\ &= \sum_i \frac{1}{\zeta - \lambda_i} \frac{1}{1 - r^2\bar{\zeta}\lambda_i} \frac{\prod_j (1 - r^2\bar{\lambda}_j\lambda_i)}{\prod_{j:j \neq i} (r\lambda_i - r\lambda_j)} - \frac{r}{1 - r^2|\zeta|^2} \prod_i \frac{1 - r^2\bar{\lambda}_i\zeta}{r\zeta - r\lambda_i}. \end{aligned}$$

Similarly,

$$\langle \tilde{B} \mid t_2 \rangle = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \prod_i \frac{z - r\lambda_i}{1 - r\bar{\lambda}_i z} \frac{r}{r\bar{\zeta}z - 1} \frac{1}{z - r\zeta} dz = \frac{r}{r^2|\zeta|^2 - 1} \prod_i \frac{r\bar{\zeta} - r\lambda_i}{1 - r^2\bar{\lambda}_i\zeta}.$$

It follows that

$$\begin{aligned} &\sum_i \frac{1}{\zeta - \lambda_i} \frac{1}{1 - r^2\bar{\zeta}\lambda_i} \frac{\prod_j (1 - r^2\bar{\lambda}_j\lambda_i)}{\prod_{j:j \neq i} (r\lambda_i - r\lambda_j)} \\ &= \frac{r}{1 - r^2|\zeta|^2} \left(\prod_i \frac{1 - r^2\bar{\lambda}_i\zeta}{r\zeta - r\lambda_i} - \prod_i \frac{r\bar{\zeta} - r\bar{\lambda}_i}{1 - r^2\bar{\lambda}_i\bar{\zeta}} \right). \end{aligned}$$

The same computations prove the validity of this statement for $r|\zeta| > 1$. One can make sense of the formula in case that $r|\zeta| = 1$ by continuous extension. Using these observations one can compute

$$\begin{aligned} \|\tilde{g}\|_{H_2}^2 &= \int_0^{2\pi} \tilde{g}(z)\overline{\tilde{g}(z)} \Big|_{z=e^{i\phi}} \frac{d\phi}{2\pi} \\ &= \frac{1}{2\pi i} \sum_{k,l} \frac{1}{\zeta - \lambda_k} \frac{\prod_\mu (1 - r^2\bar{\lambda}_\mu\lambda_k)}{\prod_{\mu \neq k} (r\lambda_k - r\lambda_\mu)} \frac{1}{\zeta - \lambda_l} \frac{\prod_\nu (1 - r^2\bar{\lambda}_\nu\lambda_l)}{\prod_{\nu \neq l} (r\lambda_l - r\lambda_\nu)} \\ &\qquad \int_{\partial\mathbb{D}} \frac{1}{z - r\lambda_k} \frac{1}{1 - r\bar{\lambda}_l z} dz \\ &= \sum_l \left(\frac{1}{\zeta - \lambda_l} \frac{\prod_\nu (1 - r^2\bar{\lambda}_\nu\lambda_l)}{\prod_{\nu \neq l} (r\lambda_l - r\lambda_\nu)} \right. \\ &\qquad \left. \sum_k \left(\frac{1}{\zeta - \lambda_k} \frac{1}{1 - r^2\bar{\lambda}_l\lambda_k} \frac{\prod_\mu (1 - r^2\bar{\lambda}_\mu\lambda_k)}{\prod_{\mu \neq k} (r\lambda_k - r\lambda_\mu)} \right) \right) \end{aligned}$$

$$= \prod_i \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \overline{\left(\sum_l \frac{1}{\zeta - \lambda_l} \frac{r}{1 - r^2 \lambda_l \bar{\zeta}} \frac{\prod_v (1 - r^2 \bar{\lambda}_v \lambda_l)}{\prod_{v \neq l} (r \lambda_l - r \lambda_v)} \right)} \tag{13}$$

$$= \frac{r^2}{1 - r^2 |\zeta|^2} \prod_i \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \overline{\left(\prod_i \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} - \prod_i \frac{r \bar{\zeta} - r \bar{\lambda}_i}{1 - r^2 \lambda_i \bar{\zeta}} \right)} \tag{14}$$

$$= \frac{r^2}{1 - r^2 |\zeta|^2} \left(\prod_i \left| \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \right|^2 - 1 \right),$$

where we used the first assertion of the lemma for (13) and the second assertion for (14). Note that for all $\zeta \in \mathbb{C} - \sigma(A)$ and $r \in (0, 1)$ the final quantity is real and positive. \square

With this preparatory work done a proof of Theorem 4.1 is simple.

Proof of Theorem 4.1. We assume that $\sigma(A) \subset \mathbb{D}$. From Equations (11), (12) and Lemma 4.2 we have that for any $\zeta \in \mathbb{C} - \sigma(A)$

$$\begin{aligned} \|(\zeta - A)^{-1}\| &\leq C \sqrt{\frac{1}{1 - r^2}} \|\tilde{g}\|_{H_2} \\ &= C \sqrt{\frac{1}{1 - r^2}} \sqrt{\frac{r^2}{1 - r^2 |\zeta|^2} \left(\prod_{i=1}^{|m|} \left| \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \right|^2 - 1 \right)}. \end{aligned} \tag{15}$$

Clearly,

$$\prod_{i=1}^{|m|} \left| \frac{1 - r^2 \bar{\lambda}_i \zeta}{r \zeta - r \lambda_i} \right|^2 = \frac{1}{r^{2|m|}} \frac{1}{|B(\zeta)|^2} \prod_{i=1}^{|m|} \left| 1 + \bar{\lambda}_i \zeta \frac{1 - r^2}{1 - \bar{\lambda}_i \zeta} \right|^2.$$

To obtain an upper bound we assume that $\zeta \in \bar{\mathbb{D}} - \sigma(A)$ and choose $r \in (0, 1)$ such that

$$1 - r^2 = \frac{\min_i |1 - \bar{\zeta} \lambda_i|}{2|m|}.$$

It follows that

$$\prod_{i=1}^{|m|} \left| 1 + \bar{\lambda}_i \zeta \frac{1 - r^2}{1 - \bar{\lambda}_i \zeta} \right|^2 \leq \left(1 + \frac{1}{2|m|} \right)^{2|m|} \leq e$$

and that (for $|m| \geq 2$)

$$r^{2|m|} = \left(1 - \frac{\min_i |1 - \bar{\zeta} \lambda_i|}{2|m|} \right)^{|m|} \geq (1 - 1/2)^2 = 1/4.$$

We conclude that

$$\begin{aligned} & \|(\zeta - A)^{-1}\| \\ & \leq C \left(\frac{2|m|}{\min_i |1 - \bar{\zeta}\lambda_i|} \right)^{1/2} \left(\frac{1}{1 - |\zeta|^2 \left(1 - \frac{\min_i |1 - \bar{\zeta}\lambda_i|}{2|m|}\right)} \right)^{1/2} \left(\frac{4e}{|B(\zeta)|^2} - 1 \right)^{1/2}, \end{aligned}$$

which is claimed in the theorem. As always the general case $\sigma(A) \subset \bar{\mathbb{D}}$ follows by continuous extension. Finally, we note that for $|\zeta| > 1$ one can choose $r = \sqrt{\frac{1}{|\zeta|}}$ in (15) and recover the obvious estimate

$$\|(\zeta - A)^{-1}\| \leq C \frac{1}{|\zeta| - 1} \left(1 - \prod_{i=1}^{|m|} \left| \frac{|\zeta| - \bar{\lambda}_i \zeta}{\sqrt{|\zeta|\zeta} - \sqrt{|\zeta|\lambda_i}} \right|^2 \right)^{1/2} \leq C \frac{1}{|\zeta| - 1}. \quad \square$$

5. Stability of Markov chains

If T is a classical stochastic matrix or a quantum channel (a trace-preserving and completely positive map, see [15]) the sequence $\{T^n\}_{n \geq 0}$ can be regarded as a finite and homogeneous classical or quantum Markov chain with transition map T . In this section we apply Theorem 4.1 to study the stability of the stationary states of a Markov chain to perturbations in the transition map. A core observation is that the transition matrix of the Markov chain is power-bounded with respect to the 1-to-1 norm and constant 1, i.e. the Markov chain constitutes a bounded semigroup, see [27]. A similar approach based on this observation was taken in [26] to prove spectral convergence estimates for classical and quantum Markov chains. We begin by recalling the basic framework of sensitivity analysis of Markov chains. A detailed introduction, however, is beyond the scope of this article. We refer to [4] and the references therein for an overview of the existing perturbation bounds for classical Markov chains and to [27] for an introduction to the stability theory of quantum Markov chains.

Let T, \tilde{T} denote two classical stochastic matrices or two quantum channels. The inequality

$$\|\rho - \tilde{\rho}\| \leq \kappa \|T - \tilde{T}\|$$

relates the distance between two stationary states ρ and $\tilde{\rho}$ arising from T and \tilde{T} , $\rho = T(\rho)$, $\tilde{\rho} = \tilde{T}(\tilde{\rho})$, to the distance between T and \tilde{T} . Commonly T is considered to be the transition matrix of the Markov chain of interest while \tilde{T} is a

small perturbation thereof. The *condition number* κ measures the relative sensitivity of the stationary states to perturbations. If T has a unique stationary state the above inequality quantifies the stability of the asymptotic behavior of $\{T^n\}_{n \geq 0}$ with respect to perturbations in the transition matrix. Elementary linear algebra shows that if T has a unique stationary state one can choose (see [23, 24, 25]) the condition number

$$\kappa_{\text{cl}} = \sup_{\substack{\delta \in \mathbb{R}^n \\ (1, \dots, 1) \cdot \delta = 0}} \frac{\|Z(\delta)\|_1}{\|\delta\|_1}, \quad Z := (\mathbb{1} - T + T^\infty)^{-1}$$

in the classical case and similarly (see [27])

$$\kappa_{\text{qu}} = \sup_{\substack{\sigma = \sigma^\dagger \in \mathcal{M}_n \\ \text{tr}(\sigma) = 0}} \frac{\|Z(\sigma)\|_1}{\|\sigma\|_1}, \quad Z := (\mathbb{1} - T + T^\infty)^{-1}$$

in the quantum setup. Here, T^∞ denotes the projection onto the stationary state of T and $\|\cdot\|_1$ denotes the absolute entry sum in the classical and the Schatten 1-norm in the quantum case. In either case the spectral properties of T and T^∞ guarantee that the map Z exists.

If the transition matrix has a unique stationary state and a subdominant eigenvalue of this matrix is close to 1 it is clear that the chain is ill conditioned in the sense that the stationary state is sensitive to perturbations in the transition map. It is a well-studied question (see [10, 25, 12, 11, 27]) whether the reverse conclusion also holds, i.e. whether the closeness of the sub-dominant eigenvalues of T to 1 provides complete information about the sensitivity of $\{T^n\}_{n \geq 0}$. It was established that this is indeed the case by deriving spectral lower and upper bounds for certain choices of κ . In particular, as shown in [25] it holds that

$$\frac{1}{\min_{\lambda_i \in \sigma(T - T^\infty)} |1 - \lambda_i|} \leq \kappa_{\text{cl}} \leq \frac{n}{\min_{\lambda_i \in \sigma(T - T^\infty)} |1 - \lambda_i|}.$$

A similar quantum bound occurs in [27].

The techniques developed in this article yield a direct approach to spectral stability estimates in both the classical and quantum case. The core observation is that if T is a stochastic matrix or a quantum channel the map $T - T^\infty$ is power bounded with (see [26] Lemma III.1)

$$\|(T - T^\infty)^n\|_{1 \rightarrow 1} = \|T^n - (T^\infty)^n\|_{1 \rightarrow 1} \leq \|T^n\|_{1 \rightarrow 1} + \|(T^\infty)^n\|_{1 \rightarrow 1} \leq 2,$$

where $\|A\|_{1 \rightarrow 1} = \sup_v \frac{\|Av\|_1}{\|v\|_1}$. With an application of Inequality (10) we conclude that

$$\kappa_{\text{cl}} \leq \|Z\|_{1 \rightarrow 1} \leq \frac{2\sqrt{16e - 4n}}{\min_{\lambda_i \in \sigma(T - T^\infty)} |1 - \lambda_i|},$$

which is qualitatively the same as the estimate in [25] but has a worse numerical prefactor ($2\sqrt{16e-4}$ instead of 1). However, the bound in [25] uses the additional properties a classical stochastic matrix has as well as the fact that the supremum in the definition of κ_{cl} is taken over vectors with 0 column sum. Our bound proves that in this case power-boundedness alone is sufficient and the additional assumptions on T and κ_{cl} are basically superfluous. Other spectral stability estimates for classical Markov chains such as [10] are weaker than (10). In the quantum context we can use Inequality (10) to improve on the spectral stability estimates of [27].

Theorem 5.1. *Let T be a trace-preserving, positive linear map on \mathcal{M}_n and $\Lambda := \sigma(T) - \{1\}$ the set of its non-unit eigenvalues. Then*

$$\frac{1}{\min_{\lambda_i \in \Lambda} |1 - \lambda_i|} \leq \kappa_{\text{qu}} \leq \frac{2\sqrt{16e - 4n^2}}{\min_{\lambda_i \in \Lambda} |1 - \lambda_i|}.$$

The proof of the theorem is identical as in [27] up to an application of (10) instead of the theorem in [29].

References

- [1] J. Agler and J. E. McCarthy, *Pick Interpolation and Hilbert function spaces*. Graduate Studies in Mathematics, 44. American Mathematical Society, Providence, RI, 2002. [MR 1882259](#) [Zbl 1010.47001](#)
- [2] O. F. Bandtlow, Estimates for norms and resolvents and an application to the perturbation of spectra. *Math. Nachr.* **267** (2004), 3–11. [MR 2047381](#) [Zbl 1056.47004](#)
- [3] R. Bhatia, *Matrix analysis*. Graduate Texts in Mathematics, 169. Springer, New York, N.Y., 1996. [MR 1477662](#) [Zbl 0863.15001](#)
- [4] G. E. Cho and C. D. Meyer, Comparison of perturbations bounds for the stationary distributions of a Markov chain. *Linear Algebra Appl.* **335** (2001), 137–150. [MR 1850819](#) [Zbl 0983.60062](#)
- [5] E. B. Davies, *Linear operators and their spectra*. Cambridge Studies in Advanced Mathematics, 106. Cambridge University Press, Cambridge, 2007. [MR 2359869](#) [Zbl 1138.47001](#)
- [6] E. B. Davies and B. Simon, Eigenvalue estimates for non-normal matrices and the zeros of random orthogonal polynomials on the unit circle. *J. Approx. Theory* **141** (2006), 189–213. [MR 2252099](#) [Zbl 1127.47037](#)
- [7] E. Egerváry, Über gewisse Extremumprobleme der Funktionentheorie. *Math. Ann.* **99** (1928), 542–561. [MR 1512465](#) [JFM 54.0332.03](#)
- [8] F. Foaïaș and A. E. Frazho, *The commutant lifting approach to interpolation problems*. Operator Theory: Advances and Applications, 44. Birkhäuser Verlag, Basel, 1990. [MR 1120546](#) [Zbl 0718.47010](#)

- [9] T. Kato, *Perturbation theory for linear operators*. Reprint of the corr. print of the 2nd ed. 1980. Classics in Mathematics. Springer, Berlin etc., 1995. [MR 1335452](#) [Zbl 0836.47009](#)
- [10] C. D. Meyer, Sensitivity of the stationary distribution of a Markov chain. *SIAM J. Matrix Anal. Appl.* **15** (1994), 715–728. [MR 1282690](#) [Zbl 0809.65143](#)
- [11] A. Yu. Mitrophanov, Stability and exponential convergence of continuous-time Markov chains. *J. Appl. Probab.* **40** (2003), 970–979. [MR 2012680](#) [Zbl 1049.60069](#)
- [12] A. Yu. Mitrophanov, Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Probab.* **42** (2005), 1003–1014. [MR 2203818](#) [Zbl 1092.60027](#)
- [13] B. Sz.-Nagy and C. Foiaş, Commutants de certains opérateurs. *Acta Sci. Math. (Szeged)* **29** (1968), 1–17. [MR 0242011](#) [Zbl 0162.45601](#)
- [14] B. Sz.-Nagy, C. Foiaş, H. Bercovici, and L. Kerchy, *Harmonic analysis of operators on Hilbert spaces*. 2nd ed. Revised and enlarged ed. Universitext. Springer, New York, N.Y., 2010. [MR 2760647](#) [Zbl 1234.47001](#)
- [15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000. [MR 1796805](#) [Zbl1049.81015](#)
- [16] N. K. Nikol'skiĭ, *Treatise on the shift operator*. Spectral function theory. With an appendix by S. V. Hruščev and V. V. Peller. Translated from the Russian by J. Peetre. Grundlehren der Mathematischen Wissenschaften, 273. Springer, Berlin etc., 1986. [MR 0827223](#) [Zbl 0587.47036](#)
- [17] N. K. Nikol'skiĭ, Condition numbers of large matrices and analytic capacities. *Algebra i Analiz* **17** (2005), 125–180. In Russian. English. trans., *St. Petersburg Math. J.* **17** (2006), 641–682. [MR 2173939](#) [Zbl 1098.15002](#)
- [18] N. K. Nikol'skiĭ, *Operators, functions, and systems: an easy reading*. Vol. 1. Hardy, Hankel, and Toeplitz. Translated from the French by A. Hartmann. Mathematical Surveys and Monographs, 92. American Mathematical Society, Providence, R.I., 2002. Vol. 2. Model operators and systems. Translated from the French by A. Hartmann and revised by the author. Mathematical Surveys and Monographs, 93. American Mathematical Society, Providence, R.I., 2002. [MR 1864396](#) (vol. 1) [MR 1892647](#) (vol. 2) [Zbl 1007.47001](#) (vol. 1) [Zbl 1007.47002](#) (vol. 2)
- [19] Henrici P, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices. *Numer. Math.* **4** (1962), 24–40. [MR 0135706](#) [Zbl 0102.01502](#)
- [20] E. Pascal, *Die Determinanten*. Eine Darstellung ihrer Theorie und Anwendungen mit Rücksicht auf die neueren Forschungen. Berechtigte deutsche Ausgabe von H. Leitzmann. B. G. Teubner, Leipzig, 1900. [JFM 31.0154.02](#)
- [21] D. Phillips, Improving Spectral-Variation bounds with Chebyshev Polynomials. *Linear Algebra Appl.* **133** (1990), 165–173. [MR 1058112](#) [Zbl 0707.15004](#)
- [22] D. Sarason, Generalized interpolation in H^∞ . *Trans. Amer. Math. Soc.* **127** (1967), 179–203.
- [23] P. Schweitzer, Perturbation theory and finite Markov chains. *J. Appl. Probability* **5** (1968), 401–413. [MR 0234527](#) [Zbl 0196.19803](#)

- [24] E. Seneta, Perturbation of the stationary distribution measured by ergodicity coefficients. *Adv. in Appl. Probab.* **20** (1988), 228–230. [MR 0932541](#) [Zbl 0645.60070](#)
- [25] E. Seneta, Sensitivity of finite Markov chains under perturbation. *Statist. Probab. Lett.* **17** (1993), 163–168. [MR 1223539](#) [Zbl 0777.60065](#)
- [26] O. Szehr, D. Reeb, and M. M. Wolf, Spectral convergence bounds for classical and quantum Markov processes. Preprint 2013. [arXiv:1301.4827](#) [math-ph]
- [27] O. Szehr and M. M. Wolf, Perturbation bounds for quantum Markov processes and their fixed points. *J. Math. Phys.* **54** (2013), Article Id. 032203, 10 pp. [MR 3059439](#) [Zbl 1281.81050](#)
- [28] H. Wielandt, Inclusion theorems for eigenvalues. *Nat. Bur. Standards Appl. Math.* **29** (1951), 75–78. [MR 0059238](#) [Zbl 0052.25602](#)
- [29] R. Zarouf, Une amélioration d’un résultat de E. B. Davies et B. Simon. *C. R. Math. Acad. Sci. Paris* **347** (2009), 939–942. [MR 2542898](#) [Zbl 1167.15311](#)
- [30] R. Zarouf, Effective H^∞ interpolation. *Houston J. Math.* **39** (2013), 487–514. [MR 3080451](#) [Zbl 1271.30009](#)
- [31] R. Zarouf, Iterated resolvent estimates for power bounded matrices. Preprint 2011. [arXiv:1103.5019v1](#) [math.FA]

Received 2013 July, 3; revised 2014 March, 15

Oleg Szehr, Department of Mathematics, Technische Universität München, 85748 Garching, Germany

e-mail: oleg.szehr@posteo.de