CONSTRUCTIVE MATHEMATICAL ANALYSIS

Volume V
Issue III

CMA
CONSTRUCTIVE MATHEMATICAL ANALYSIS

ISSN 2651-2939

https://dergipark.org.tr/en/pub/cma

VOLUME V ISSUE III
ISSN 2651-2939

September 2022
https://dergipark.org.tr/en/pub/cma

CONSTRUCTIVE MATHEMATICAL ANALYSIS

Contents

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 3, pp. 119-133
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

The disconnectedness of certain sets defined after uni-variate polynomials

VLADIMIR PETROV KOSTOV*

ABSTRACT. We consider the set of monic real uni-variate polynomials of a given degree d with non-vanishing coefficients, with given signs of the coefficients and with given quantities pos of their positive and neg of their negative roots (all roots are distinct). For $d \geq 6$ and for signs of the coefficients $(+,-,+,+,\dots,+,+,-,+)$, we prove that the set of such polynomials having two positive, $d-4$ negative and two complex conjugate roots, is not connected. For $pos + neg \leq 3$ and for any d, we give the exhaustive answer to the question for which signs of the coefficients there exist polynomials with such values of pos and neg.

Keywords: Real polynomial in one variable, hyperbolic polynomial, Descartes' rule of signs, discriminant set.

2020 Mathematics Subject Classification: 26C10, 30C15.

1. INTRODUCTION

We consider questions about the general family of monic uni-variate real degree d polynomials: $Q_d := x^d + \sum_{j=0}^{d-1} a_j x^j$. In the space $\mathbb{R}^d$ of the coefficients a_j, one defines the *discriminant set* Δ_d as the set of their values for which the polynomial Q_d has a multiple real root. More precisely, if Δ_d^1 is the set of values of the coefficients for which Q_d has a multiple root (real or complex), then this is the set of the zeros of the determinant of the Sylvester matrix of the polynomials Q_d and Q_d'. One has to set $\Delta_d := \Delta_d^1 \setminus \Delta_d^2$, where Δ_d^2 is the set of values of the coefficients a_j for which there is a multiple complex conjugate pair of roots of Q_d and no multiple real root. It is true that $\dim(\Delta_d) = \dim(\Delta_d^1) = d-1$ and $\dim(\Delta_d^2) = d-2$.

The set

$$R_{1,d} := \mathbb{R}^d \setminus \Delta_d$$

consists of $[d/2]+1$ open components of dimension d ($[\cdot]$ stands for the integer part of $\cdot$). The polynomials Q_d from a given component have one and the same number μ of real roots (which are all distinct); the number ν of complex conjugate pairs can range from 0 to $[d/2]$, because $\mu + 2\nu = d$. Given two polynomials with one and the same number ν, one can continuously deform the roots of the first polynomial into the roots of the second one by keeping the real roots distinct throughout the deformation. This proves that to any possible number ν corresponds exactly one component of the set $R_{1,d}$.

In the same way one can consider the components of the set

$$R_{2,d} := \mathbb{R}^d \setminus (\Delta_d \cup \{a_0 = 0\}) .$$

Received: 29.04.2022; Accepted: 02.08.2022; Published Online: 12.08.2022
*Corresponding author: Vladimir Petrov Kostov; vladimir.kostov@unice.fr
DOI: 10.33205/cma.1111247

The polynomials from one and the same open component (also of dimension d) have one and the same numbers pos of positive and neg of negative roots (and no vanishing roots). When deforming the roots of one polynomial into the roots of another one, one has to keep the same numbers pos and neg throughout the deformation. To each pair (pos, neg) corresponds exactly one component of the set $R_{2,d}$. As $pos + neg = \mu$, $0 \leq pos, neg \leq \mu$ and $\mu + 2\nu = d$, there are

$$(d+1) + (d-1) + (d-3) + \cdots = ([d/2]+1)([(d+1)/2]+1)$$

components of the set $R_{2,d}$.

A more complicated task is to study the components of the set

$$R_{3,d} := \mathbb{R}^d \setminus (\Delta_d \cup \{a_0 = 0\} \cup \{a_1 = 0\} \cup \cdots \cup \{a_{d-1} = 0\})$$

of monic uni-variate polynomials with no multiple real roots and no zero coefficients.

Definition 1. A *sign pattern* of length $d+1$ is a sequence of $d+1$ symbols $+$ and/or $-$ beginning with a $+$. We say that a polynomial Q_d with no vanishing coefficients defines the sign pattern $\sigma_0 := (+, \beta_{d-1}, \beta_{d-2}, \ldots, \beta_0)$, $\beta_j = +$ or $-$, (notation: $\sigma(Q_d) = \sigma_0$), if $\mathrm{sign}(a_j) = \beta_j$, $j = 0, \ldots, d-1$.

One can ask the question to which couple (sign pattern, pair (pos, neg)) (we call them *couples* for short) corresponds at least one component of the set $R_{3,d}$. The polynomials from a given component of $R_{3,d}$ have one and the same couple. All components are of dimension d.

When considering the set $R_{3,d}$, it is self-understood that the couples have to be defined in accordance with Descartes' rule of signs. This rule states that a real uni-variate polynomial Q_d has not more positive roots counted with multiplicity than the number c of sign changes in the sequence of its coefficients; the difference $c - pos$ is even, see [4, 9, 10, 11, 15, 16, 19, 20, 28] or [30]. Hence the sign of the constant term is $(-1)^{pos}$. When the polynomial has no zero coefficients, Descartes' rule of signs applied to $Q_d(-x)$ implies that Q_d has not more negative roots counted with multiplicity than the number p of sign preservations in that sequence (hence $c + p = d+1$), and the difference $p - neg$ is also even.

Definition 2. A pair (pos, neg) satisfying these conditions w.r.t. a given sign pattern σ_0 is called *compatible* with σ_0 (and vice versa), and the couple $(\sigma_0, (pos, neg))$ is also called *compatible*. For a monic polynomial Q_d with no vanishing coefficients, with pos positive simple and neg negative simple roots and no other real roots, we say that Q_d *realizes* the couple $(\sigma(Q_d), (pos, neg))$.

Yet this compatibility is just a necessary condition which turns out not to be sufficient. That is, there exist cases when to certain compatible couples correspond no components of $R_{3,d}$. So we formulate the first problem which we consider in the present paper:

Problem 1. *For a given degree d, for which compatible couples do there exist monic polynomials realizing these couples? In other words, to which of the compatible couples there corresponds at least one component of the set $R_{3,d}$?*

Some results in relationship with Problem 1 are formulated in the next section. The problem seems to have been stated for the first time in [2]. The first example when to a compatible couple corresponds no component of the set $R_{3,d}$ (this is an example with $d = 4$), and the exhaustive answer to the problem for $d = 4$, are to be found in [18]. For $d = 5$ and $d = 6$, the result is given in [1]. For $d = 7$ and partially for $d = 8$ (resp. completely for $d = 8$), the answer is formulated and proved in [12] and [13] (resp. in [22]). Different aspects concerning Descartes' rule of signs are treated in papers [23, 5, 6, 7, 8] and [14].

Of particular importance is the class of *hyperbolic polynomials*, i. e. real polynomials whose roots are all real. The *hyperbolicity domain* Π_d is the set of values of the coefficients a_j for which

the polynomial Q_d is hyperbolic. For properties of hyperbolic polynomials and the domain Π_d see [3, 17, 21, 29] and [24].

In what follows we are also interested in another problem:

Problem 2. *For a given degree d, to which compatible couples correspond two or more components of the set $R_{3,d}$?*

To formulate our first result connected with Problem 2, we introduce the following notation:

Notation 1. For $d \geq 4$, we consider $\mathbb{R}^d$ as the set $\{(a_{d-1}, a_{d-2}, \ldots, a_0) | a_j \in \mathbb{R}\}$ of d-tuples of coefficients (excluding the leading one) of polynomials Q_d. We denote by $\sigma_\bullet$ the sign pattern $(+,-,+,+,\ldots,+,+,-,+)$ and by $\Pi_d^*(\sigma_\bullet)$ (resp. by $A(\sigma_\bullet,(2,d-4))$) the subset of $\mathbb{R}^d$ of polynomials with signs of the coefficients (all non-zero) as defined by $\sigma_\bullet$ and having four positive and $d-4$ negative distinct real roots (resp. two positive and $d-4$ negative distinct roots and one complex conjugate pair). Hence the polynomials of the set $\Pi_d^*(\sigma_\bullet)$ are hyperbolic while the ones of the set $A(\sigma_\bullet,(2,d-4))$ are not.

The following theorem is proved in Section 3.

Theorem 1. *(1) For $d \geq 6$, the set $A(\sigma_\bullet,(2,d-4))$ is non-empty and consists of more than one component of the set $R_{3,d}$. Hence the set $A(\sigma_\bullet,(2,d-4))$ is not connected.*

(2) For $d=4$ and 5, the respective sets $A(\sigma_\bullet,(2,0))$ and $A(\sigma_\bullet,(2,1))$ are connected.

Remarks 1. (1) One can mention cases in which the components of the set $R_{3,d}$ are contractible and to each compatible couple corresponds exactly one component of the set $R_{3,d}$ (see [26]). Namely, such are the cases of hyperbolic polynomials and of polynomials having exactly one or no real roots at all.

(2) In the case of polynomials having exactly two real distinct roots (hence $pos+neg=2$) to each compatible couple corresponds either one or no component of $R_{3,d}$, and all components are contractible. See more details in the next section or in [26]. Whether in the case of exactly three real roots to each compatible couple corresponds at most one component of the set $R_{3,d}$ is an open question.

(3) For $d=4$ and $d=5$, pictures of the set Δ_d^1 (from which one can deduce the form of the set $A(\sigma_\bullet,(2,d-4))$) can be found in [27] and [8] respectively.

2. Comments and further results

Given a sign pattern $\hat{\sigma}$ with c sign changes and p sign preservations (hence $c+p=d$), Descartes' rule of signs implies that any hyperbolic polynomial with sign pattern $\hat{\sigma}$ has exactly c positive and exactly p negative roots counted with multiplicity. We define the *canonical order of moduli* corresponding to $\hat{\sigma}$. The sign pattern $\hat{\sigma}$ is read from the right and to each sign change (resp. sign preservation) one puts in correspondence the letter P (resp. the letter N).

For example, for $\hat{\sigma}=\sigma_\dagger:=(+,-,-,-,+,+)$ (resp. for $\hat{\sigma}=\sigma_\bullet$) this gives the string $NPNNP$ (resp. $PPNN\cdots NNPP$, $d-4$ times N). After this one inserts the symbol $<$ between any two consecutive letters which in the cases of $\sigma_\dagger$ and $\sigma_\bullet$ gives

$$N<P<N<N<P \quad \text{and} \quad P<P<N<N<\cdots<N<N<P<P$$

respectively. If one denotes by α_j and β_j the moduli of the positive and negative roots, then one replaces the letters P and N by these moduli which in the case of $\sigma_\dagger$ defines the canonical order

$$\beta_1<\alpha_1<\beta_2<\beta_3<\alpha_2$$

whereas the canonical order corresponding to $\sigma_\bullet$ is given by (3.1).

It is true that for any sign pattern σ_0 of length $d+1$, there exists a degree d monic hyperbolic polynomial T with $\sigma(T)=\sigma_0$ whose roots define the respective canonical order of moduli, see Proposition 1 in [25].

Our next step is to consider the cases when the polynomial Q_d has not more than three real roots, i. e. $pos+neg\le 3$ (and hence in the case of equality the possible values of the pair (pos,neg) are $(3,0)$, $(2,1)$, $(1,2)$ and $(0,3)$). For the cases $pos=neg=0$ and $pos+neg=1$, see part (1) of Remarks 1. For $pos+neg=2$ (hence d is even), we remind some of the results of [26].

Definition 3. For $pos+neg=2$, we define *Case 1)* (resp. *Case 2)*) by the conditions the constant term to be positive, all coefficients of monomials of odd degree to be positive (resp. negative), the pair (pos,neg) to equal $(2,0)$ (resp. $(0,2)$) and the coefficient of at least one monomial of even degree to be negative.

Theorem 2. (see [26]). *For d even and $pos+neg=2$,*

(1) *A given compatible couple is realizable if and only if it does not correspond to Case 1) or 2).*
(2) *If the constant term is positive (hence $(pos,neg)=(2,0)$ or $(0,2)$) and one is not in Case 1) or 2), a given compatible couple is realizable by polynomials having any ratio different from 1 between the moduli of the two real roots.*
(3) *If the constant term is negative (hence $(pos,neg)=(1,1)$) and there are two monomials of odd degree with coefficients of opposite signs, then such a compatible couple is realizable by polynomials with any ratio of the moduli α and β of its positive and negative root respectively.*
(4) *If the constant term is negative and all coefficients of monomials of odd degree are positive (resp. negative), then such a compatible couple is realizable by polynomials with any ratio $\alpha/\beta<1$ (resp. $\alpha/\beta>1$) and not realizable by polynomials with $\alpha/\beta\ge 1$ (resp. $\alpha/\beta\le 1$).*

To formulate the new results about the situation with $pos+neg=3$, we introduce the following notion:

Definition 4. For a given degree d, the $\mathbb{Z}_2\times\mathbb{Z}_2$*-action* on the set of compatible couples is defined by two commuting involutions. The first of them maps a polynomial Q_d into $(-1)^dQ_d(-x)$ (this changes the pair (pos,neg) into (neg,pos), it changes the signs of the coefficients of x^{d-1}, x^{d-3}, ... and preserves the signs of the other coefficients). The second involution maps Q_d into $x^dQ_d(1/x)/Q_d(0)$ (the pair (pos,neg) is preserved and the sign pattern, eventually multiplied by -1, is read from the right; the roots of $x^dQ_d(1/x)/Q_d(0)$ are the reciprocals of the roots of Q_d). An *orbit* of the $\mathbb{Z}_2\times\mathbb{Z}_2$-action consists of 2 or 4 compatible couples which are simultaneously realizable or not. This allows to formulate the results only for one of the 2 or 4 couples of a given orbit.

Theorem 3. *Suppose that the pair $(2,1)$ is compatible with the sign pattern $\sigma_\triangle$ (hence the constant term is positive). Then*

(1) *The couple $\mathcal{C}:=(\sigma_\triangle,(2,1))$ is realizable.*
Denote by $-\beta<0$, $\alpha_1>0$ and $\alpha_2>0$ the three real roots of a polynomial realizing the couple $\mathcal{C}$.
(2) *If there are monomials x^{2m} and x^{2n-1} with negative coefficients (one can have $2m<2n-1$ or $2n-1<2m$), then for any of the five possibilities*

$$\beta<\alpha_1<\alpha_2\,,\quad \beta=\alpha_1<\alpha_2\,,\quad \alpha_1<\beta<\alpha_2\,,\quad \alpha_1<\alpha_2=\beta\quad\text{and}\quad \alpha_1<\alpha_2<\beta\,,$$

there exist polynomials realizing the couple $\mathcal{C}$.

(3) *If all odd monomials have positive coefficients, then only the possibility* $\beta < \alpha_1 < \alpha_2$ *is realizable.*
(4) *If all even monomials have positive coefficients, then only the possibility* $\alpha_1 < \alpha_2 < \beta$ *is realizable.*

The theorem is proved in Section 4. The compatibility of the sign pattern with the pair $(2,1)$ implies that in part (3) (resp. in part (4)) of the theorem there is at least one even (resp. odd) monomial whose coefficient is negative.

Notation 2. For d odd, we denote by $D(a,b,c)$ the sign pattern consisting of $2a$ pluses followed by b pairs "$-,+$" followed by $2c$ minuses, where $1 \leq a$, $1 \leq b$, $1 \leq c$ and $2a+2b+2c = d+1$.

Theorem 4. *Suppose that the pair* $(3,0)$ *is compatible with the sign pattern* $\sigma_\diamond$ *which is not of the form* $D(a,b,c)$. *Then the couple* $(\sigma_\diamond, (3,0))$ *is realizable.*

The theorem is proved in Section 5.

Theorem 5. *For* $j = 1, 2, \ldots, b$, *the couple* $(D(a,b,c), (2j+1,0))$ *is not realizable.*

The theorem is proved in Section 6. Its proof resembles the proof of part (i) of Theorem 4 in [27] which treats a particular case of Theorem 5. However the proof of Lemma 1 (used in the proof of Theorem 5) is more complicated than the proof of its analog which is Lemma 6 of [27]. This renders indispensable giving the whole proof of Theorem 5.

Remark 1. *For the sign pattern* $D(a,b,c)$, *compatible are the following pairs* (pos, neg):
1) *the ones mentioned in Theorem 5;*
2) *the pair* $(1,0)$;
3) *the pairs* $(2j+1, 2r)$, $r = 1, 2, \ldots, a+c-1$, $j = 0, 1, \ldots, b$.

Realizability of the couples $(D(a,b,c), (pos, neg))$ *with* (pos, neg) *as in 2) and 3) can be proved by analogy with the proof of parts (ii) and (iii) of Theorem 4 in* [27].

3. Proof of Theorem 1

Part (1). A) For $d \geq 6$, the set $\Pi_d^*(\sigma_\bullet)$ is non-empty, see Proposition 1 in [25]. Fix a polynomial $Q^* \in \Pi_d^*(\sigma_\bullet)$. By Proposition 1 of [25], one can choose Q^* such that the moduli of its positive and negative roots (denoted by $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$ and $\beta_1 < \beta_2 < \cdots < \beta_{d-5} < \beta_{d-4}$ respectively) satisfy the string of inequalities

$$\alpha_1 < \alpha_2 < \beta_1 < \beta_2 < \cdots < \beta_{d-5} < \beta_{d-4} < \alpha_3 < \alpha_4 \,. \tag{3.1}$$

So the negative roots of Q^* are $-\beta_{d-4} < -\beta_{d-5} < \cdots < -\beta_1 < 0$. Starting with Q^*, we construct two polynomials Q^1 and Q^2 of the set $A(\sigma_\bullet, (2, d-4))$ (so this set is non-empty) about which we show that they belong to different components of $R_{3,d}$. This implies the theorem.

B) We consider the one-parameter family of polynomials

$$\tilde{Q}_t := Q^* + tx^2(x+\beta_1)(x+\beta_2)\cdots(x+\beta_{d-4}) \,, \quad t \geq 0 \,.$$

For any $t \geq 0$, one has $\sigma(\tilde{Q}_t) = \sigma_\bullet$. As t increases, the roots $-\beta_1, -\beta_2, \ldots, -\beta_{d-4}$ of $\tilde{Q}_t$ do not move. The roots α_1 and α_3 move to the right while α_2 and α_4 move to the left. For some $t_0 > 0$, either α_1 coalesces with α_2 or α_3 coalesces with α_4 or both these things take place. Indeed, the values of $\tilde{Q}_t$ for each fixed $x \geq \alpha_1$ increase at least as fast as $t\alpha_1^2 \prod_{i=1}^{d-4}(\alpha_1 + \beta_i)$.

If for $t = t_0$, α_1 and α_2 coalesce and α_3 and α_4 remain positive and distinct, then one can fix $t_1 > t_0$ sufficiently close to t_0 for which the roots α_1 and α_2 have given birth to a complex

conjugate pair while α_3 and α_4 are still positive and distinct. We set $Q^1 := \tilde{Q}_{t_1}$. Hence the polynomial Q^1 has $d-2$ real roots

$$\begin{aligned} &-\beta_{d-4} < -\beta_{d-5} < \cdots < -\beta_1 < 0 < \alpha_3 < \alpha_4 \quad \text{such that} \\ &0 < \beta_1 < \beta_2 < \cdots < \beta_{d-4} < \alpha_3 < \alpha_4 \end{aligned} \tag{3.2}$$

and a complex conjugate pair. After this we set $Q^2_* := x^d Q^1(1/x)$. The sequence of coefficients of Q^1, when read from the right, is the string of coefficients of Q^2_*. After this we set $Q^2 := Q^2_*/Q^1(0)$, so Q^2 is monic. The sign pattern $\sigma_\bullet$ is center-symmetric, therefore $\sigma(Q^2) = \sigma_\bullet = \sigma(Q^1)$. The roots of the polynomial Q^2 are the reciprocals of the roots of Q^1. The real roots of Q^2 satisfy the conditions

$$\begin{aligned} &-\beta_{d-4} < -\beta_{d-5} < \cdots < -\beta_1 < 0 < \alpha_3 < \alpha_4 \quad \text{and} \\ &0 < \alpha_1 < \alpha_2 < \beta_1 < \beta_2 < \cdots < \beta_{d-4}; \end{aligned} \tag{3.3}$$

the polynomial Q^2 has also a complex conjugate pair.

If for $t = t_0$, α_3 and α_4 coalesce while α_1 and α_2 remain positive and distinct, then for some $t_1 > t_0$ sufficiently close to t_0 we obtain the polynomial Q^2 with exactly two positive and $d-4$ negative roots which satisfy conditions (3.3). After this we set $Q^1_* = x^d Q^2(1/x)$ and $Q^1 := Q^1_*/Q^2(0)$. The real roots of Q^1 satisfy conditions (3.2).

Finally, if for $t = t_0$, one has $\alpha_1 = \alpha_2 = a > 0$ and $\alpha_3 = \alpha_4 = b > a$, then one constructs the polynomials

$$Q^\pm := \tilde{Q}_{t_0} \pm \varepsilon(x - (a+b)/2)\,, \quad \varepsilon > 0\,.$$

For ε small enough,

1) the coefficients of $Q^\pm$ are non-zero and $\sigma(Q^\pm) = \sigma_\bullet$;
2) each of the polynomials $Q^\pm$ has $d-4$ distinct negative roots close to $-\beta_i$;
3) Q^+ has two distinct positive roots close to a and a complex conjugate pair close to b;
4) and vice versa for Q^-.

We set $Q^1 := Q^-$ and $Q^2 := Q^+$.

C) Suppose that the two polynomials Q^1 and Q^2 belong to one and the same component of the set $R_{3,6}$. Then it is possible to connect them by a continuous path (homotopy) within this component: Q^s, $s \in [1,2]$. Along the path the two positive, the $d-4$ negative and the two complex conjugate roots of Q^s depend continuously on s while remaining distinct throughout the homotopy. We denote the negative roots by $-\tilde{\beta}_j$, $j = 1, \ldots, d-4$, and the two positive roots by $\tilde{\gamma}_j$, $j = 1, 2$, where

$$\text{for} \quad s = 1\,, \quad \text{one has} \quad \tilde{\beta}_j = \beta_j\,, \quad \tilde{\gamma}_j = \alpha_{2+j}\,;$$

$$\text{for} \quad s = 2\,, \quad \text{one has} \quad \tilde{\beta}_j = \beta_j\,, \quad \tilde{\gamma}_j = \alpha_j\,.$$

Hence there exists $s = s_0 \in (1,2)$ such that for $s = s_0$, $\tilde{\beta}_{d-4} = \tilde{\gamma}_2$. This means that the polynomial Q^{s_0} has exactly $d-2$ real roots such that

$$-\tilde{\beta}_{d-4} < \cdots < -\tilde{\beta}_1 < 0 < \tilde{\gamma}_1 < \tilde{\gamma}_2 \quad , \quad \tilde{\beta}_{d-4} = \tilde{\gamma}_2\,.$$

Using a linear change $x \mapsto hx$, $h > 0$, we achieve the condition $\tilde{\beta}_{d-4} = \tilde{\gamma}_2 = 1$.

D) Suppose that d is even. The fact that ± 1 are roots of Q_d implies the two conditions:

$$a_d + a_{d-2} + a_{d-4} + \cdots + a_2 + a_0 = 0 \quad \text{and} \quad a_{d-1} + a_{d-3} + \cdots + a_3 + a_1 = 0\,.$$

The first of them is possible only if all even coefficients are 0, because in the corresponding positions the sign pattern $\sigma_\bullet$ contains $(+)$-signs. However $a_d = 1$. This contradiction means that the homotopy Q^s does not exist, so Q^1 and Q^2 belong to different components of the set $R_{3,d}$ and the set $A(\sigma_\bullet, (2, d-4))$ is not connected. One can observe that this reasoning is not valid for $d = 2$ or $d = 4$, because in these cases there are no negative roots at all.

E) Suppose that $d \geq 7$ is odd. Set $\delta := \tilde{\beta}_{d-5} > 0$ and $Q_d^{s_0} = (x + \delta)U(x)$, where $U = x^{d-1} + \sum_{j=0}^{d-2} u_j x^j$. The polynomial U has an even number of positive roots, so $u_0 > 0$. The conditions

$$0 > a_{d-1} = \delta + u_{d-2} \quad \text{and} \quad \delta > 0$$

imply $u_{d-2} < 0$ whereas from

$$0 < a_{d-2} = \delta u_{d-2} + u_{d-3}\,, \quad \delta > 0 \quad \text{and} \quad u_{d-2} < 0$$

one deduces that $u_{d-3} > 0$. In the same way one has

$$0 > a_1 = \delta u_1 + u_0\,, \quad \delta > 0\,, \quad u_0 > 0\,, \quad \text{so} \quad u_1 < 0 \quad \text{and}$$

$$0 < a_2 = \delta u_2 + u_1\,, \quad \delta > 0\,, \quad u_1 < 0\,, \quad \text{so} \quad u_2 > 0\,.$$

The first three and the last three of the coefficients of the polynomial $U(-x)$ are positive. By Descartes' rule of signs it has not more than $d-5$ positive roots, and it has exactly $d-5$ positive roots only if it has $d-5$ sign changes. On the other hand, one knows that $U(-x)$ has exactly $d-5$ positive roots $-\tilde{\beta}_j$, $j = 1, 2, \ldots, d-6, d-4$. Hence $U(x)$ has $d-5$ sign preservations, therefore $u_k > 0$ for $2 \leq k \leq d-3$.

Thus $\sigma(U) = \sigma_\bullet$ (but here the sign pattern $\sigma_\bullet$ is meant to be of length d, not $d+1$). Suppose that the homotopy Q^s exists. Along this homotopy the root $-\tilde{\beta}_{d-5}$ is a continuous negative-valued function. As division of Q^s by $x + \delta$ gives the polynomials U, there exists a homotopy between the polynomial U corresponding to Q^1 and the one corresponding to Q^2. We denote them by U^1 and U^2. They are of even degree $d - 1 \geq 6$, each of them has exactly two positive roots $\tilde{\gamma}_1 < \tilde{\gamma}_2$, exactly $d-5$ negative roots and one complex conjugate pair. For the moduli of the real roots one has

$$\tilde{\beta}_j < \tilde{\gamma}_1 \quad \text{for} \quad U^1 \quad \text{and} \quad \tilde{\gamma}_2 < \tilde{\beta}_j \quad \text{for} \quad U^2\,, \quad j = 1,\ 2, \ldots\,,\ d-6,\ d-4$$

(see (3.2) and (3.3)). This, however, is impossible, see D). □

Part (2). F) For $d = 4$, for each polynomial $Q \in A(\sigma_\bullet, (2, 0))$, there exists a unique quantity $g > 0$ such that for $g' \in [0, g)$, one has $Q + g' \in A(\sigma_\bullet, (2, 0))$ and for $g' = g$, the polynomial $Q + g'$ has a multiple positive root.

On the other hand, for each polynomial $Q \in A(\sigma_\bullet, (2, 0))$, there exists a unique quantity $h > 0$ such that for $h' \in [0, h)$, one has $Q - h' \in A(\sigma_\bullet, (2, 0))$ and for $h' = h$, Q has either a zero root or a multiple positive root. The quantities g and h are continuous functions of the coefficients of Q.

Denote by $A^*(\sigma_\bullet)$ the set of monic polynomials whose coefficients have signs as defined by the sign pattern $\sigma_\bullet$ and which have a multiple positive root and a complex conjugate pair. Hence the set $A(\sigma_\bullet, (2, 0))$ is homeomorphic to the direct product of the set $A^*(\sigma_\bullet)$ and an open interval. Therefore if $A^*(\sigma_\bullet)$ is connected, then such is $A(\sigma_\bullet, (2, 0))$ as well.

Denote by $A_0^*(\sigma_\bullet)$ the subset of $A^*(\sigma_\bullet)$ for which the multiple root of Q is at 1. Each polynomial $Q \in A^*(\sigma_\bullet)$ can be transformed into a polynomial of $A_0^*(\sigma_\bullet)$ by a linear change of the variable x followed by a multiplication with a non-zero constant. Hence $A^*(\sigma_\bullet)$ is homeomorphic to $A_0^*(\sigma_\bullet) \times (0, \infty)$.

Any polynomial $Q \in A_0^*(\sigma_\bullet)$ is of the form

$$(x-1)^2(x^2+Ax+B) = x^4 + (A-2)x^3 + (B-2A+1)x^2 + (A-2B)x + B\,,$$

where $A^2 - 4B \leq 0$. The set $A_0^*(\sigma_\bullet)$ is defined by the conditions

$$A < 2\,, \quad B - 2A + 1 > 0\,, \quad A - 2B < 0 \quad \text{and} \quad B \geq A^2/4\,.$$

This is the set of points in the plane (A, B) which are to the left of the vertical line $A = 2$ and above or on the graph of the function (of the argument $A \in (-\infty, 2)$) $\max(2A-1, A/2, A^2/4)$; strictly above for $A \in [0, 2)$ and above or on the graph for $A < 0$. This is a contractible set.

G) For $d = 5$, we denote by $A^\dagger(\sigma_\bullet)$ the set of monic polynomials the signs of whose coefficients are defined by the sign pattern $\sigma_\bullet$ and which have a simple negative root, a double positive root and a complex conjugate pair. Denote by $A_0^\dagger(\sigma_\bullet)$ its subset for which the double root is at 1. By complete analogy with part F) of the proof we show that connectedness of $A_0^\dagger(\sigma_\bullet)$ implies the one of $A(\sigma_\bullet, (2,1))$.

Any polynomial $Q \in A_0^\dagger(\sigma_\bullet)$ is of the form

$$(x-1)^2(x+A)(x^2+Bx+C) = x^5 + \sum_{j=0}^{4} f_j x^j\,, \quad \text{where}$$

$$f_4 = A + B - 2\,, \qquad f_3 = AB - 2A - 2B + C + 1\,,$$

$$f_2 = -2AB + AC + A + B - 2C\,, \quad f_1 = AB - 2AC + C$$

$$\text{and} \qquad f_0 = AC\,,$$

with $A > 0$ and $B^2 - 4C < 0$. For any $\rho > 0$ and $r > 0$, the polynomial $Q_{\rho,r} := Q + \rho(x-1)^2 + rx^3(x-1)^2$ defines the sign pattern $\sigma_\bullet$ and belongs to the set $A_0^\dagger(\sigma_\bullet)$. Indeed, it is non-negative for $x \geq 0$, with equality only for $x = 1$; its second derivative at $x = 1$ is positive, so $x = 1$ is a double root; the sign pattern $\sigma_\bullet$ and Descartes' rule of sign imply that $Q_{\rho,r}$ has not more than one negative root, so it has exactly one such root. Hence one can choose ρ and r such that $f_2 = f_3$. The set $A_0^\dagger(\sigma_\bullet)$ is connected if and only if its subset defined by the condition $f_2 = f_3$ is connected.

H) The condition $f_2 = f_3$ allows to express A as a function of B and C:

$$A = T_0/D\,, \quad \text{where} \quad T_0 = 3B - 3C - 1 \quad \text{and} \quad D = 3B - C - 3\,.$$

For the coefficients f_i with $A = T_0/D$ one finds

$$f_4 = T_4/D\,, \qquad T_4 = 3B^2 - BC - 6B - C + 5\,,$$

$$f_3 = f_2 = T_3/D\,, \quad T_3 = -3B^2 + 2BC - C^2 + 2B + 2C - 1\,,$$

$$f_1 = T_1/D\,, \qquad T_1 = 3B^2 - 6BC + 5C^2 - B - C\,.$$

In Fig. 1 and 2 we represent the following sets:

– $\mathcal{L}_0 : T_0 = 0$ (in solid line) and $\mathcal{L} : D = 0$ (in dashed line) are straight lines;

– $\mathcal{E}_3 : T_3 = 0$ (in dashed line) and $\mathcal{E}_1 : T_1 = 0$ (in dotted line) are ellipses;
– $\mathcal{H} : T_4 = 0$ (in solid line) is a hyperbola;
– $\mathcal{P} : C = B^2/4$ is a parabola (in dash-dotted line).

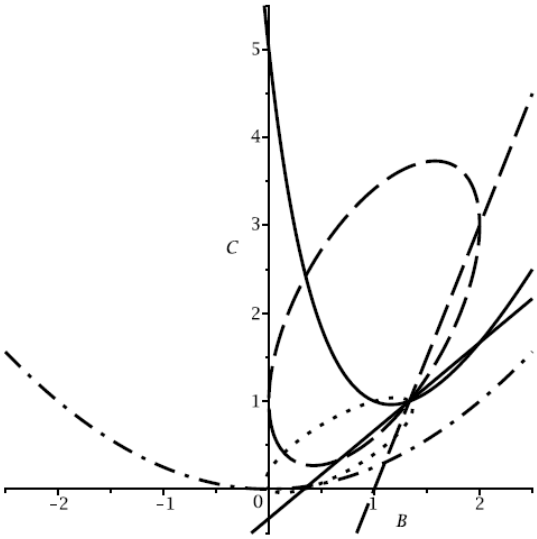

FIGURE 1. The set $A_0^{\dagger}(\sigma_\bullet)$ subdued to the condition $f_2 = f_3$ (global view).

Remark 2. As $C > 0$, only the branch of $\mathcal{H}$ belonging to the upper half-plane is represented in Fig. 1 and 2. The asymptotes of $\mathcal{H}$ are the lines $B = -1$ and $C = 3B - 6$. We denote by $\mathrm{Int}(\mathcal{E}_i)$ and $\mathrm{Out}(\mathcal{E}_i)$ the intersections with the half-plane $C > 0$ of the interior and the exterior of the ellipse $\mathcal{E}_i$. By $\mathrm{Int}(\mathcal{H})$ we denote the part of the upper half-plane which is above and by $\mathrm{Out}(\mathcal{H})$ the part which is below the branch of $\mathcal{H}$ with $C > 0$. Notice that

$$\mathrm{Int}(\mathcal{E}_3) : T_3 > 0,\ C > 0, \quad \mathrm{Int}(\mathcal{E}_1) : T_1 < 0,\ C > 0, \quad \mathrm{Int}(\mathcal{H}) : T_4 < 0,\ C > 0,$$

$$\mathrm{Out}(\mathcal{E}_3) : T_3 < 0,\ C > 0, \quad \mathrm{Out}(\mathcal{E}_1) : T_1 > 0,\ C > 0, \quad \mathrm{Out}(\mathcal{H}) : T_4 > 0,\ C > 0.$$

The ellipse $\mathcal{E}_1$ intersects the C-axis at $(0, 0)$ and $(0, 1/5)$ while $\mathcal{E}_3$ is tangent to the C-axis at $(0, 1)$. The leftmost point of the ellipse $\mathcal{E}_1$ is at

$$((8 - \sqrt{70})/12 = -0.030\ldots, (10 - \sqrt{70})/20 = 0.081\ldots) .$$

The point $(4/3, 1)$ is a common point for $\mathcal{L}$, $\mathcal{L}_0$, $\mathcal{H}$, $\mathcal{E}_1$ and $\mathcal{E}_3$.

The intersecting lines $\mathcal{L}_0$ and $\mathcal{L}$ define two pairs of opposite sectors. The ones of opening $> \pi/2$ are denoted by $\mathcal{S}_u : T_0 < 0, D < 0$ (upper) and $\mathcal{S}_\ell : T_0 > 0, D > 0$ (lower). One has $A > 0$ exactly when the point (B, C) belongs to one of these two sectors.

I) The signs of the coefficients f_i and of the quantities $A > 0$ and $C > 0$ imply that one must have one of the two systems of conditions:

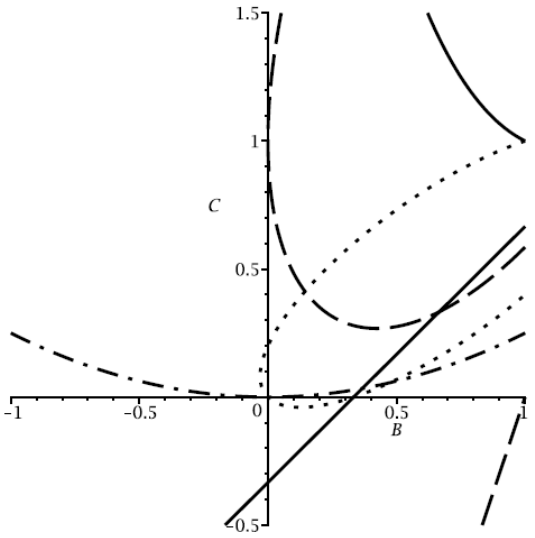

FIGURE 2. The set $A_0^\dagger(\sigma_\bullet)$ subdued to the condition $f_2 = f_3$ (local view).

$$\text{(i)} \ : \ (B,C) \in \mathcal{S}_\ell \cap \mathrm{Int}(\mathcal{E}_1) \cap \mathrm{Int}(\mathcal{E}_3) \cap \mathrm{Int}(\mathcal{H}) \ , \quad \text{i.e.}$$

$$T_0 > 0 \ , \quad D > 0 \ , \quad T_1 < 0 \ , \quad T_3 > 0 \quad \text{and} \quad T_4 < 0 \quad \text{or}$$

$$\text{(ii)} \ : \ (B,C) \in \mathcal{S}_u \cap \mathrm{Out}(\mathcal{E}_1) \cap \mathrm{Out}(\mathcal{E}_3) \cap \mathrm{Out}(\mathcal{H}) \ , \quad \text{i.e.}$$

$$T_0 < 0 \ , \quad D < 0 \ , \quad T_1 > 0 \ , \quad T_3 < 0 \quad \text{and} \quad T_4 > 0 \ .$$

The possibility (i) is to be excluded. Indeed, one has

$$\mathcal{E}_3 \cap \mathcal{L}_0 = \{(2/3, 1/3) \ , \quad (4/3, 1)\} \quad \text{and} \quad \mathcal{E}_3 \cap \mathcal{L} = \{(4/3, 1) \ , \quad (2, 3)\} \ ,$$

see Fig. 1 and 2, so $\mathrm{Int}(\mathcal{E}_3)$ intersects with $\mathcal{S}_u$, but not with $\mathcal{S}_\ell$.

J) We describe the set obtained in case (ii). For $B \leq -1$, this is the part of the upper plane which is above the parabola $\mathcal{P}$. For $-1 < B < (8-\sqrt{70})/12$, this is its part between the parabola $\mathcal{P}$ from below and the hyperbola $\mathcal{H}$ from above, see Fig. 1. For each $(8-\sqrt{70})/12 \leq B < 0$, this is the union of two intervals whose endpoints belong to $\mathcal{H}$ and $\mathcal{E}_1$ for the upper and to $\mathcal{E}_1$ and $\mathcal{P}$ for the lower interval. For $B \geq 0$, this is the union of two curvilinear triangles, each with one rectilinear side which is part of the C-axis. The above triangle has vertices at $(0, 1)$, $(0, 5)$ and $(0.34\ldots, 2.42\ldots)$. The latter point, together with $(4/3, 1)$, is the intersection $\mathcal{H} \cap \mathcal{E}_3$.

The lower triangle has vertices at $(0, 1/5)$, $(0, 1)$ and $(0.14\ldots, 0.41\ldots)$. The latter point, together with $(4/3, 1)$, is the intersection $\mathcal{E}_1 \cap \mathcal{E}_3$.

To see that there is no other point of the set defined in case (ii) with $B > 0$, one has to observe the order on $\mathcal{P}$ of the intersection points of

$$\mathcal{P} \cap \mathcal{L}_0 = \{(0.36\ldots, \ 0.03\ldots), (3.63\ldots, \ 3.29\ldots)\}$$

and

$$\mathcal{P} \cap \mathcal{E}_1 = \{(0,\ 0), (0.47\ldots,\ 0.22\ldots)\}.$$

The connectedness of the set obtained in case (ii) follows from its description. □

4. Proof of Theorem 3

Part (1). The last component of $\sigma_\triangle$ is a $+$. Suppose that there is a minus sign in $\sigma_\triangle$ corresponding to x^{2m}, $1 \le m \le [d/2]$. The polynomial $-x^{2m}+1$ has exactly two real roots, namely ± 1, and they are simple. For $\varepsilon > 0$ small enough, the polynomial $P_0 := \varepsilon x^d - x^{2m} + 1$ has exactly three real roots two of which are close to ± 1 and the third is > 1. (One can notice that by Descartes' rule of signs it has not more than two positive and not more than one negative root.)

Fix a degree d polynomial P_1 with $\sigma(P_1) = \sigma_\triangle$. Then for $0 < \eta \ll \varepsilon$, the polynomial $P_0 + \eta P_1$ has signs of the coefficients as defined by $\sigma_\triangle$ and has exactly one negative and two positive simple roots and $(d-3)/2$ complex conjugate pairs counted with multiplicity. Thus $P_0 + \eta P_1$ realizes the couple $(\sigma_\triangle, (2,1))$.

Suppose now that there are $(+)$-signs in $\sigma_\triangle$ corresponding to all monomials of even degrees. Then there is a monomial x^{2m+1}, $1 \le 2m+1 < d$, whose sign is negative. The polynomial $P_2 := x^d - x^{2m+1}$ has simple roots at ± 1 and a $(2m+1)$-fold root at 0. For $\varepsilon > 0$ small enough, the polynomial $P_2 + \varepsilon$ has exactly three real roots (two positive and one negative) all of which are simple. Then with P_1 and η as above, the polynomial $P_2 + \varepsilon + \eta P_1$ realizes the couple $(\sigma_\triangle, (2,1))$.

Part (2). We construct a polynomial of the form $V := x^d - Ax^{2m} - Bx^{2n-1} + C$, $A > 0$, $B > 0$, $C > 0$, such that $V(1) = V'(1) = V(-1) = 0$:

$$\begin{aligned} &1 - A - B + C = 0\,, \quad -1 - A + B + C = 0\,, \qquad d - 2mA - (2n-1)B = 0 \\ &\text{hence} \qquad\qquad A = C = (d - 2n + 1)/2m\,, \quad B = 1\,. \end{aligned} \tag{4.4}$$

By Descartes' rule of signs, V has no other real roots. After this one decreases C: $C \mapsto C - t$, $t \ge 0$. For $t = 0$, the root -1 moves with a finite speed to the right while the double root at 1 splits into two real roots moving for $t = 0$ with infinite speeds to the left and right respectively. Hence for $t > 0$ close to 0, one has $\alpha_1 < \beta < \alpha_2$. The linear system (4.4) with unknown variables A, B and C has non-zero determinant. Hence for $\varepsilon > 0$ small enough, one can obtain polynomials V satisfying the conditions

$$V(1) = V'(1) = V(-1 \pm \varepsilon) = 0 \quad (\text{resp.} \quad V(1) = V'(1) \pm \varepsilon = V(-1) = 0)$$

which after decreasing C yield polynomials satisfying the inequalities $\beta < \alpha_1 < \alpha_2$ or $\alpha_1 < \alpha_2 < \beta$ (resp. the conditions $\beta = \alpha_1 < \alpha_2$ or $\alpha_1 < \alpha_2 = \beta$). It remains to construct the polynomial $V + \eta P_1$, where $0 < \eta \ll \varepsilon$ and $\sigma(P_1) = \sigma_\triangle$.

Part (3). There exists a monomial x^{2m} with negative coefficient. Then for $\varepsilon > 0$ small enough, the polynomial $W := x^{2m-1}(x-1)(x-2) + \varepsilon$ has exactly one negative and two positive roots whose moduli satisfy the condition $\beta < \alpha_1 < \alpha_2$. Its four non-zero coefficients have the signs as defined by $\sigma_\triangle$. After this one constructs the polynomial $W + \eta P_1$ with η and P_1 as above.

The inequality $\beta \ge \alpha_1$ is impossible. Indeed, represent a polynomial W realizing the couple $\mathcal{C}$ in the form $W = W_o + W_e$, where W_o is the odd and W_e is the even part of W. Then for $x \in (-\beta, 0)$, one has $W_e(x) = W_e(-x)$ and $W_o(x) < W_o(-x)$. As $W(x) > 0$ for $x \in (-\beta, 0)$, one cannot have $W(\alpha_1) = 0$. This is a contradiction.

Part (4). Changing the polynomial $Y(x)$ with $\sigma(Y) = \sigma_\triangle$ which realizes the couple $\mathcal{C}$ to $Y_1 := x^d Y(1/x)/Y(0)$ (we set $\sigma^R_\triangle := \sigma(Y_1)$), one obtains a polynomial realizing the couple $(\sigma^R_\triangle, (2,1))$, where all odd monomials have positive signs, see Definition 4. The roots of Y_1 are the reciprocals of the roots of Y, so one deduces part (4) from part (3).

5. Proof of Theorem 4

The last sign of $\sigma_\diamond$ is a $-$. Suppose that there are two monomials x^{2m} and x^{2p}, $m > p > 0$, whose signs defined by $\sigma_\diamond$ are $-$ and $+$ respectively. Consider the polynomial $P_3 := -x^{2m} + Ax^{2p} - B$, $A > 0$, $B > 0$. By Descartes' rule of signs it has at most two positive and at most two negative roots. We define A and B such that P_3 has double roots at 1 and (-1):

$$-1 + A - B = 0\,, \quad -2m + 2pA = 0 \qquad \text{hence}$$

$$A = m/p > 0\,, \qquad B = (m-p)/p > 0\,.$$

Then for $\varepsilon > 0$ small enough, the polynomial $P_3 + \varepsilon x^d$ has exactly three real roots, all simple and positive. Suppose that P_4 is a degree d polynomial such that $\sigma(P_4) = \sigma_\diamond$. Then for $0 < \eta \ll \varepsilon$, the polynomial $P_3 + \varepsilon x^d + \eta P_4$ has sign pattern $\sigma_\diamond$ and has exactly three real roots, all simple and positive.

Suppose that there are no monomials x^{2m} and x^{2p} as above. Then the signs of the first a even monomials are positive and the ones of the last $(d+1-2a)/2$ of them are negative, $0 \le a \le (d-1)/2$. Suppose that there are monomials $x^{2\nu}$, $x^{2\mu-1}$ and $x^{2\theta}$, $2\nu > 2\mu - 1 > 2\theta$, whose signs defined by $\sigma_\diamond$ are $-$, $+$ and $-$ respectively. By Descartes' rule of signs a polynomial of the form $P_5 := -x^{2\nu} + Cx^{2\mu-1} - Dx^{2\theta}$, $C > 0$, $D > 0$, has at most two positive roots and no negative roots; clearly it has a (2θ)-fold root at 0. One can choose C and D such that the positive roots are at 1 and 2:

$$-1 + C - D = 0\,, \qquad -2^{2\nu} + 2^{2\mu-1}C - 2^{2\theta}D = 0 \quad \text{hence}$$

$$D = (2^{2\nu} - 2^{2\mu-1})/(2^{2\mu-1} - 2^{2\theta}) > 0\,, \quad C = D + 1 > 0\,.$$

For $\varepsilon > 0$ small enough, the polynomial $P_5 + \varepsilon x^d$ has three positive simple roots and no other real roots, and the polynomial $P_6 := P_5 + \varepsilon x^d + \eta P_4$ with η and P_4 as above has three positive simple roots, no other real roots and $\sigma(P_6) = \sigma_\diamond$.

So now we suppose that there are no monomials x^{2m} and x^{2p}, and no monomials $x^{2\nu}$, $x^{2\mu-1}$ and $x^{2\theta}$ as above. Suppose that there are monomials x^{2u-1} and x^{2v-1}, $d > 2u-1 > 2v-1 > 0$, such that their signs are $-$ and $+$ respectively. One can construct a polynomial $P_7 := x^d - Ex^{2u-1} + Fx^{2v-1}$, $E > 0$, $F > 0$, having double roots at ± 1, a $(2v-1)$-fold root at 0 and no other real roots:

$$1 - E + F = 0\,, \qquad d - (2u-1)E + (2v-1)F = 0 \quad \text{hence}$$

$$F = (d - 2u + 1)/2(u - v) > 0\,, \quad E = F + 1 > 0\,.$$

The absence of other real roots is guaranteed by Descartes' rule of signs. Hence for $0 < \eta \ll \varepsilon \ll 1$, the polynomial $P_7 - \varepsilon + \eta P_4$ has sign pattern $\sigma_\diamond$, three simple positive roots and no other real roots (recall that $P_4(0) < 0$).

Suppose that there are no couples or triples of monomials x^{2m}, x^{2p} or $x^{2\nu}$, $x^{2\mu-1}$, $x^{2\theta}$ or x^{2u-1}, x^{2v-1}. Then the signs of the first $h_o \ge 1$ odd monomials (including x^d) are positive and the signs of the remaining $(d+1-2h_o)/2$ odd monomials are negative. The signs of the

first $h_e \geq 0$ even monomials are positive and the signs of the other $(d+1-2h_e)/2$ ones are negative. The absence of triples $x^{2\nu}$, $x^{2\mu-1}$, $x^{2\theta}$ implies $h_o \leq h_e + 1$. The cases $h_o = h_e + 1$ and $h_o = h_e$ are impossible, because there is only one sign change in the sign pattern. Therefore $1 \leq h_o \leq h_e - 1$. This means that the sign pattern is $D(a,b,c)$ with $a = h_o$, $b = h_e - h_o$ and $c = (d+1-2a-2b)/2$.

6. Proof of Theorem 5

Suppose that a polynomial $P := \sum_{j=0}^{d} a_j x^j$ realizes the couple $(D(a,b,c),(3,0))$. Denote by

$$P_o := \sum_{\nu=0}^{(d-1)/2} a_{2\nu+1} x^{2\nu+1} \quad \text{and} \quad P_e := \sum_{\nu=0}^{(d-1)/2} a_{2\nu} x^{2\nu}$$

its odd and even parts respectively. In each of the sequences $\{a_{2\nu+1}\}_{\nu=0}^{(d-1)/2}$ and $\{a_{2\nu}\}_{\nu=0}^{(d-1)/2}$ there is exactly one sign change. Descartes' rule of signs implies that the polynomial P_o has exactly three real roots, namely $-x_o$, 0 and x_o, $x_o > 0$, while the polynomial P_e has exactly two real roots $\pm x_e$, $x_e > 0$; all these five roots are simple.

Remarks 2. (1) The polynomial P_e is positive and increasing on (x_e, ∞) and negative on $[0, x_e)$. The polynomial P_o is positive and increasing on (x_o, ∞) and negative on $(0, x_o)$.

(2) One has $x_o \neq x_e$, otherwise $P(-x_o) = 0$, i.e. P has a negative root which is a contradiction.

(3) One can assume that all positive roots of P are distinct. Indeed, if this is not the case, then one can perturb P to make all its positive roots distinct without changing the signs of its coefficients as follows. If P has an ℓ-fold root $\lambda > 0$ ($\ell > 1$), i.e. $P = (x-\lambda)^\ell P^0$, $P^0(\lambda) \neq 0$, then for $\varepsilon > 0$ small enough, the polynomial $(x-\lambda)^{\ell-1}(x-\lambda-\varepsilon)P^0$ has the same sign pattern and its ℓ-fold root has split into an $(\ell-1)$-fold and a simple real roots. It remains to iterate this construction sufficiently many times.

Notation 3. We denote by $0 < \xi_1 < \xi_2 < \xi_3$ the smallest three of the positive roots of P and by ζ a positive number different from x_o and x_e.

It is clear that $P(\zeta) > 0$ for $\zeta \in (\xi_1, \xi_2)$ and $P(\zeta) < 0$ for $\zeta \in (\xi_2, \xi_3)$. For $\zeta \in (\xi_1, \xi_2)$, it is impossible to have $P_e(\zeta) \leq 0$ and $P_o(\zeta) \leq 0$ (with at most one equality, see part (2) of Remarks 2). It is also impossible to have $P_e(\zeta) \geq 0$ and $P_o(\zeta) \geq 0$. Indeed, this would imply that $x_e \leq \zeta < \xi_2$ and $x_o \leq \zeta < \xi_2$ which means that for $x \in (\xi_2, \xi_3)$, one has $P_e(x) \geq 0$ and $P_o(x) \geq 0$, i.e. $P(x) > 0$. This is a contradiction.

Two possible situations are left:

a) $P_e(\zeta) > 0$, $P_o(\zeta) < 0$;

b) $P_e(\zeta) < 0$, $P_o(\zeta) > 0$

(we skip the cases of equalities, because they were already taken into account).

Situation a) cannot take place, because this would mean that

$$P(-\zeta) = P_e(\zeta) - P_o(\zeta) > 0\,,$$

and since $P(0) < 0$ and $P(x) \to -\infty$ for $x \to -\infty$, in each of the intervals $(-\infty, -\zeta)$ and $(-\zeta, 0)$ the polynomial P would have at least one root – a contradiction.

So suppose that we are in situation b), so $x_o < \zeta < x_e$. Without loss of generality one can assume that $\xi_1 = 1$; this can be achieved by a rescaling $x \mapsto \xi_1 x$. Hence $P_o(1) = \beta > 0$ and $P_e(1) = -\beta$. Considering the polynomial P/β instead of P, one can assume that $\beta = 1$. One deduces from Lemma 1 which follows that there are no real roots of P larger than 1 (one can use the Taylor series of P at 1); this contradiction completes the proof.

Lemma 1. *Under the above assumptions, $P^{(m)}(1) > 0$, for any $m = 1, 2, \dots, d$.*

Proof of Lemma 1. In the proof we allow zero values of the coefficients as well. This is because we need to deal with compact sets on which minimization arguments are to be applied.

Suppose that the sum $\delta_1 := a_1 + a_3 + \dots + a_{2b+2c-1}$ is fixed (recall that these are all the negative coefficients of P_o). Then for any $m = 1, 2, \dots, d$, it is true that $P_o^{(m)}(1)$ is minimal for

$$a_{2b+2c-1} = \delta_1 \ , \quad a_1 = a_3 = \dots = a_{2b+2c-3} = 0 \ .$$

Indeed, when computing the values of the derivatives at $x = 1$, monomials of larger degree in x are multiplied by larger factors (equal to these degrees). We apply here $(d-3)/2$ times the fact that for $A + B$ fixed, the inequalities $A \geq 0$, $B \geq 0$ and $\lambda > \mu > 0$ imply that the sum $\lambda A + \mu B$ is maximal when $B = 0$.

Similarly, if the sum $\delta_2 := a_{2b+2c+1} + a_{2b+2c+3} + \dots + a_d$ of all positive coefficients of P_o is fixed, then $P_o^{(m)}(1)$ is minimal for $a_{2b+2c+1} = \delta_2$, $a_{2b+2c+3} = \dots = a_d = 0$.

For the polynomial P_e, we obtain in the same way that if the sums

$$\delta_3 := a_0 + a_2 + \dots + a_{2c-2} \quad \text{and} \quad \delta_4 := a_{2c} + \dots + a_{d-1}$$

are fixed, then $P_e^{(m)}(1)$ is minimal for $a_{2c-2} = \delta_3$, $a_0 = a_2 = \dots = a_{2c-4} = 0$, $a_{2c} = \delta_4$, $a_{2c+2} = \dots = a_{d-1} = 0$. Thus the polynomials P_o and P_e are of the form

$$P_o = Ex^{2b+2c+1} - Fx^{2b+2c-1} \quad , \quad P_e = Gx^{2c} - Hx^{2c-2} \ ,$$

with $E := a_{2b+2c+1} \geq 0$, $-F := a_{2b+2c-1} \leq 0$, $G := a_{2c} \geq 0$ and $-H := a_{2c-2} \leq 0$. Recall that

$$P(1) = 0 \ , \quad P_o(1) = 1 \quad \text{and} \quad P_e(1) = -1 \ , \quad \text{i. e.} \quad E - F = 1 \quad \text{and} \quad G - H = -1 \ .$$

The values of the derivatives at $x = 1$ are of the form

$$P^{(m)}(1) = u_m E - v_m F + w_m G - t_m H \ , \quad u_m > v_m > w_m > t_m \ ,$$

with $u_m, \ v_m, w_m, \ t_m \in \mathbb{N}$. Hence

$$\begin{aligned} P^{(m)}(1) &= (u_m - v_m)E + v_m(E - F) + (w_m - t_m)G + t_m(G - H) \\ &= (u_m - v_m)E + (w_m - t_m)G + (v_m - t_m) > 0 \ . \end{aligned}$$

□

References

[1] A. Albouy, Y. Fu: *Some remarks about Descartes' rule of signs*, Elem. Math., **69** (2014), 186-194.

[2] B. Anderson, J. Jackson and M. Sitharam: *Descartes' rule of signs revisited*, Am. Math. Mon., **105** (1998), 447-451.

[3] V. I. Arnold: *Hyperbolic polynomials and Vandermonde mappings*, Funct. Anal. Appl., **20** (1986), 52-53.

[4] F. Cajori: *A history of the arithmetical methods of approximation to the roots of numerical equations of one unknown quantity*, Colorado College Publication: Science Series, (1910) 171-215.

[5] H. Cheriha, Y. Gati and V. P. Kostov: *A nonrealization theorem in the context of Descartes' rule of signs*, Annual of Sofia University "St. Kliment Ohridski", Faculty of Mathematics and Informatics, **106** (2019), 25-51.

[6] H. Cheriha, Y. Gati and V. P. Kostov: *Descartes' rule of signs, Rolle's theorem and sequences of compatible pairs*, Studia Scientiarum Mathematicarum Hungarica, **57** (2) (2020), 165-186.

[7] H. Cheriha, Y. Gati and V. P. Kostov: *On Descartes' rule for polynomials with two variations of sign*, Lithuanian Math. J., **60** (2020), 456-469.

[8] H. Cheriha, Y. Gati and V. P. Kostov: *Degree 5 polynomials and Descartes' rule of signs*, Acta Universitatis Matthiae Belii, series Mathematics, **28** (2020), 32-51.

[9] D. R. Curtiss: *Recent extensions of Descartes' rule of signs*, Ann. of Math., **19** (4) (1918), 251-278.

[10] J. -P. de Gua de Malves: *Démonstrations de la Règle de Descartes, Pour connoître le nombre des Racines positives & négatives dans les Équations qui n'ont point de Racines imaginaires*, Memoires de Mathématique et de Physique tirés des registres de l'Académie Royale des Sciences (1741), 72-96.
[11] R. Descartes: *The Geometry of René Descartes: with a facsimile of the first edition*, translated by D. E. Smith and M. L. Latham, Dover Publications, New York (1954).
[12] J. Forsgård, V. P. Kostov and B. Shapiro: *Could René Descartes have known this?*, Exp. Math., **24** (4) (2015), 438-448.
[13] J. Forsgård, V. P. Kostov and B. Shapiro: *Corrigendum: "Could René Descartes have known this?"*, Exp. Math., **28** (2) (2019), 255-256.
[14] J. Forsgård, D. Novikov and B. Shapiro: *A tropical analog of Descartes' rule of signs*, Int. Math. Res. Not., **12** (2017), 3726-3750.
[15] J. Fourier: *Sur l'usage du théorème de Descartes dans la recherche des limites des racines. Bulletin des sciences par la Société philomatique de Paris* (1820), 156-165, 181-187; œuvres 2, 291-309, Gauthier-Villars (1890).
[16] C. F. Gauss: *Beweis eines algebraischen Lehrsatzes,* J. Reine Angew. Math., **3** (1-4) (1828); Werke 3, 67-70, Göttingen (1866).
[17] A. B. Givental: *Moments of random variables and the equivariant Morse lemma (Russian)*, Uspekhi Mat. Nauk, **42** (1987), 221-222.
[18] D. J. Grabiner: *Descartes' Rule of Signs: Another Construction*, Am. Math. Mon., **106** (1999), 854-856.
[19] J. L. W. Jensen: *Recherches sur la théorie des équations*, Acta Math., **36** (1913), 181-195.
[20] V. Jullien: Descartes La "Geometrie" de 1637.
[21] V. P. Kostov: *On the geometric properties of Vandermonde's mapping and on the problem of moments*. Proceedings of the Royal Society of Edinburgh, 112A, (1989), 203-211.
[22] V. P. Kostov: *On realizability of sign patterns by real polynomials*, Czechoslovak Math. J., **68** (3) (2018), 143, 853–874.
[23] V. P. Kostov: *Polynomials, sign patterns and Descartes' rule of signs*, Math. Bohem., **144** (1) (2019), 39-67.
[24] V. P. Kostov: T*opics on hyperbolic polynomials in one variable*. Panoramas et Synthèses 33, vi + 141 p. SMF (2011).
[25] V. P. Kostov: *Hyperbolic polynomials and canonical sign patterns*, Serdica Math. J., **46** (2) (2020), 135-150.
[26] V. P. Kostov: *Univariate polynomials and the contractibility of certain sets*, Annual of Sofia University "St. Kliment Ohridski", Faculty of Mathematics and Informatics, **107** (2020), 75-99.
[27] V. P. Kostov, B. Z. Shapiro: *Polynomials, sign patterns and Descartes' rule*, Acta Universitatis Matthiae Belii, series Mathematics, **27** (2019), 1-11.
[28] E. Laguerre: *Sur la théorie des équations numériques*, Journal de Mathématiques pures et appliquées, s. 3, 9, 1883, 99-146; œuvres 1, Paris, 1898, Chelsea, New-York, 3–47 (1972).
[29] I. Méguerditchian: *Thesis - Géométrie du discriminant réel et des polynômes hyperboliques*, thesis defended in 1991 at the University Rennes 1.
[30] B. E. Meserve: Fundamental Concepts of Algebra, Dover Publications, New York (1982).

VLADIMIR PETROV KOSTOV
UNIVERSITÉ CÔTE D'AZUR
DEPARTMENT OF MATHEMATICS
CNRS, LJAD, FRANCE
ORCID: 0000-0001-5836-2678
E-mail address: vladimir.kostov@unice.fr

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 3, pp. 134-140
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

On the Poisson equation in exterior domains

WERNER VARNHORN*

ABSTRACT. We construct a solution of the Poisson equation in exterior domains $\Omega \subset \mathbb{R}^n,\ n \geq 2$, in homogeneous Lebesgue spaces $L^{2,q}(\Omega), ; 1 < q < \infty$, with methods of potential theory and integral equations. We investigate the corresponding null spaces and prove that its dimensions are equal to $n+1$ independent of q.

Keywords: Poisson equation, potential theory, homogeneous Lebesgue spaces.

2020 Mathematics Subject Classification: 31B10, 31B30, 35C15, 35J05.

1. INTRODUCTION

Let $G \subset \mathbb{R}^n$ $(n \geq 2)$ be an exterior domain with a smooth boundary ∂G of class C^2. We consider Poisson's equation concerning some scalar function u:

$$-\Delta u = f \text{ in } G,\ u_{|\partial G} = \Phi. \tag{1.1}$$

Here f is given in G and Φ is the boundary value prescribed on ∂G. As usual, Δ denotes the Laplacian in $\mathbb{R}^n$.

It is well-known that in unbounded domains the treatment of partial differential equations causes special difficulties, and that the usual Sobolev spaces $W^{m,q}(G)$ are not adequate in this case: Even for the Laplacian in $\mathbb{R}^n$ we find [6] that the operator $\Delta : W^{m,q}(\mathbb{R}^n) \to W^{m-2,q}(\mathbb{R}^n)$ is not a Fredholm operator in general, as it is in the case of bounded domains [16]. Thus in exterior domains, the equations (1.1) have mostly been studied in connection with weight functions: Either (1.1) has been solved in weighted Sobolev spaces directly [7, 12, 14] or it has first been multiplied by some weights and then been solved in standard Sobolev spaces [17].

It is the aim of the present note to prove the solvability of (1.1) in homogeneous spaces $L^{2,q}(G)$ $(1 < q < \infty)$ of the following type [5, 11]: Let $L^q(G)$ be the space of functions defined almost everywhere in G such that the norm

$$\|f\|_{q,G} = \left(\int_G |f(x)|^q \, \mathrm{d}x \right)^{1/q}$$

is finite. Then $L^{2,q}(G)$ is the space of all functions being locally in $L^q(\overline{G})$ and having all second order distributional derivatives in $L^q(G)$. We show that for f given in $L^q(G)$ and some boundary value $\Phi \in W^{2-1/q,q}(\partial G)$ (see the notations below) there exists always a solution $u \in L^{2,q}(G)$. Concerning the uniqueness of this solution we prove that the space of all $u \in L^{2,q}(G)$ satisfying (1.1) with $f = 0$ and $\Phi = 0$ has the dimension $n+1$, independent of q. This result also holds for the case $n = 2$.

Received: 14.06.2022; Accepted: 02.08.2022; Published Online: 12.08.2022
*Corresponding author: Werner Varnhorn; varnhorn@mathematik.uni-kassel.de
DOI: 10.33205/cma.1143800

Throughout this paper $G \subset \mathbb{R}^n$ $(n \geq 2)$ is an exterior domain, i.e. a domain whose complement is compact. Let $\overline{G}$ denote its closure in $\mathbb{R}^n$ and ∂G its boundary, which we assume to be of class C^2 [1, p. 67].

In the following, all function spaces contain real valued functions. Let $D \subset \mathbb{R}^n$ be any domain with a compact boundary ∂D of class C^2, or let $D = \mathbb{R}^n$. Besides the spaces $L^q(D)$ we need the well-known function spaces $C^\infty(D)$, $C_0^\infty(D)$, and the space $C_0^\infty(\overline{D})$, containing the restrictions $f_{|\overline{D}}$ of functions $f \in C_0^\infty(\mathbb{R}^n)$.

We call a function u locally in $L^q(\overline{D})$ $(1 < q < \infty)$ and write $u \in L^q_{\mathrm{loc}}(\overline{D})$ if $u \in L^q(D \cap B)$ for every ball $B \subset \mathbb{R}^n$. Note that this space does not coincide with the usual space $L^q_{\mathrm{loc}}(D)$ in general (except for $D = \mathbb{R}^n$).

By $W^{m,q}(D)$ $(m = 0, 1, 2; W^{0,q}(D) = L^q(D))$ we mean the usual Sobolev space of functions u such that $D^\alpha u \in L^q(D)$ for all multiindices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n = \{0, 1, \dots\}^n$ with $\alpha_1 + \cdots + \alpha_n \leq m$ [1]. Here we use

$$D^\alpha u = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} u, \; D_i = \partial/\partial x_i \; (i = 1, \dots, n; \; x = (x_1, \dots, x_n) \in \mathbb{R}^n).$$

The spaces $W^{m,q}_{\mathrm{loc}}(D)$ and $W^{m,q}_{\mathrm{loc}}(\overline{D})$ are defined analogously.

We need the fractional order space $W^{2-1/q,q}(\partial D)$, which contains the trace $u_{|\partial D}$ of all $u \in W^{2,q}_{\mathrm{loc}}(\mathbb{R}^n)$ [1, p. 216]. The norm in $W^{2-1/q,q}(\partial D)$ is denoted by $\| \cdot \|_{2-1/q,q,\partial D}$.

The term $\nabla u = (D_j u)_{j=1,\dots,n}$ represents the gradient of u and $\nabla^2 u = (D_i D_j u)_{i,j=1,\dots,n}$ means the system of all second order derivatives of u. For these terms we define the seminorms

$$\|\nabla u\|_{q,D} = \left(\sum_{k=1}^n \|D_k u\|_{q,D}^q \right)^{1/q}, \; \|\nabla^2 u\|_{q,D} = \left(\sum_{j,k=1}^n \|D_j D_k u\|_{q,D}^q \right)^{1/q},$$

and introduce for $m = 1, 2$ and $1 < q < \infty$ the homogeneous spaces

$$L^{m,q}(D) = \left\{ u \in L^q_{\mathrm{loc}}(\overline{D}) \mid \|\nabla^m u\|_{q,D} < \infty \right\}. \tag{1.2}$$

Finally, concerning the norms and seminorms, we sometimes omit the domain of definition if it is obvious and use $\| \cdot \|_q$ or $\| \cdot \|_{2-1/q,q}$ instead of $\| \cdot \|_{q,G}$ or $\| \cdot \|_{2-1/q,q,\partial G'}$ for example.

2. Potential theory

Besides the Poisson equation (1.1) we also consider the special case of Laplace' equation with Dirichlet boundary condition

$$-\Delta u = 0 \text{ in } G, \quad u_{|\partial G} = \Phi. \tag{2.3}$$

These equations have mostly been studied with methods of potential theory (see for example [8, 15]). We collect some well-known facts in this section.

Let E_n $(n \geq 2)$ in the following denote the fundamental solution of the Laplacian such that $-\Delta E_n(x) = \delta(x)$ where δ is Dirac's distribution in $\mathbb{R}^n$. It is well-known that

$$E_2(x) = -\frac{\ln|x|}{\omega_2} \; (n = 2), \quad E_n(x) = \frac{|x|^{2-n}}{(n-2)\omega_n} \; (n \geq 3), \tag{2.4}$$

where ω_n is the area of the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$ $(n \geq 2)$.

Proposition 2.1. *Let $G \subset \mathbb{R}^n$ $(n \geq 2)$ be an exterior domain with boundary ∂G of class C^2, and let $\Phi \in W^{2-1/q,q}(\partial G)$ be given $(1 < q < \infty)$. Then there exists a unique function $u \in L^{2,q}(G)$ satisfying*

(2.3) *in G, if we require the following decay conditions as $|x| \to \infty$:*

$$u(x) - a\ln|x| = 0(1)\ (n=2), \quad u(x) = 0(|x|^{2-n})\ (n \geq 3), \tag{2.5}$$
$$\nabla^m u(x) = O(|x|^{2-n-m})\ (n \geq 2;\ m = 1, 2).$$

Here $a \in \mathbb{R}$ is a fixed prescribed constant.

Proof. To prove uniqueness let $u = u^1 - u^2$ be the difference of two solutions u^1 and u^2 with the required decay properties above. Define the bounded domain $G_r = G \cap B_r(0)$ where $B_r(0) \subset \mathbb{R}^n$ denotes an open ball with center at zero and radius r such that $\partial G \subset B_r(0)$. From the local regularity theory we find $D_j u \in L^2_{\text{loc}}(\overline{G})$ $(j = 1, \ldots, n)$. Thus in G_r we may apply Greens first identity, obtaining

$$\int_{G_r} |\nabla u|^2 \, \mathrm{d}x = \int_{\partial B_r} (\partial_N u) u \, \mathrm{d}o, \tag{2.6}$$

because the boundary integral over ∂G vanishes. Here N denotes the outward (with respect to G_r) unit normal vector on the boundary $\partial B_r = \partial B_r(0)$ and $\partial_N u$ is the normal derivative of u. Now do to the decay properties of u, the right hand side in (2.6) tends to zero as $r \to \infty$. This is obvious if $n \geq 3$. For $n = 2$, using the expansion theorem for harmonic functions at infinity [15, p. 523], we find $u(x) = 0(1)$ and $\nabla u(x) = 0(|x|^{-2})$ as $|x| \to \infty$, which implies the assertion above, too. It follows $\nabla u = o$ in G, hence $u = 0$ in G because u vanishes on the boundary ∂G. This proves the uniqueness.

To show the existence of a solution with the required properties we use the boundary integral equations method: Let us define the simple layer potential

$$(E^n \Theta)(x) = \int_{\partial G} E_n(x - y) \Theta(y) \, \mathrm{d}o_y, \ (x \notin \partial G),$$

the double layer potential

$$(D^n \Theta)(x) = -\int_{\partial G} \partial_{N(y)} E_n(x - y) \Theta(y) \, \mathrm{d}o_y \ (x \notin \partial G),$$

and the normal derivative of the simple layer potential

$$(H^n \Theta)(x) = -\int_{\partial G} \partial_{N(x)} E_n(x - y) \Theta(y) \, \mathrm{d}o_y \ (x \notin \partial G).$$

Here and in the following, $N = N(z)$ is the outward (with respect to the bounded domain $G_b = \mathbb{R}^n / \overline{G}$) unit normal vector in $z \in \partial G$, and $\Theta \in W^{2-1/q,q}(\partial G)$ is the unknown source density. Then we have the continuity relation

$$(E^n \Theta)^e = (E^n \Theta)^i = E^n \Theta \quad \text{on } \partial G \tag{2.7}$$

and the jump relations

$$D^n \Theta - (D^n \Theta)^e = (D^n \Theta)^i - D^n \Theta = 1/2 \Theta \quad \text{on } \partial G, \tag{2.8}$$

$$H^n \Theta - (H^n \Theta)^e = (H^n \Theta)^i - H^n \Theta = -1/2 \Theta \quad \text{on } \partial G. \tag{2.9}$$

The index e stands for the limit from outside, and the index i for the limit from inside. Now let us first assume $n \geq 3$. Following [3, 10] (here for the case of Helmholtz' equation), for the solution of (2.3) we choose in G the ansatz

$$u = D^n \Theta - \alpha E^n(\Theta) \quad (0 < \alpha \in \mathbb{R}).$$

Then by means of (2.7), (2.8) we obtain the second kind Fredholm boundary integral equation

$$\Phi = -1/2 \Theta + D^n \Theta - \alpha E^n \Theta \quad \text{on } \partial G \tag{2.10}$$

for the unknown source density $\Theta \in W^{2-1/q,q}(\partial G)$. To see that (2.10) is uniquely solvable for all boundary values $\Phi \in W^{2-1/q,q}(\partial G)$, let $0 \neq \Psi$ be a solution of the homogeneous adjoint integral equation

$$0 = -1/2\Psi + H^n\Psi - \alpha E^n\Psi \quad \text{on } \partial G. \tag{2.11}$$

By (2.7) and (2.9), this implies $\alpha(E^n\Psi)^i = (H^n\Psi)^i = -(\partial_N E^n\Psi)^i$, and Green's first identity yields $\int_{G_b} |\nabla(E^n\Psi)|^2 \,\mathrm{d}x = \int_{\partial G}(E^n\Psi)^i(\partial_N E^n\Psi)^i \,\mathrm{d}o = -\alpha \int_{\partial G} |E^n\Psi|^2 \,\mathrm{d}o$, hence $E^n\Psi = 0$ in $\overline{G}_b$. This implies $(E^n\Psi)^e = 0$ using (2.7), and the uniqueness statement above yields $E^n\Psi = 0$ in G, too. Thus $E^n\Psi = 0$ in the whole $\mathbb{R}^n$, and we obtain $\Psi = 0$ by (2.9), as asserted. This proves the existence in the case $n \geq 3$.

Now let $n = 2$. As in [9] (for the case of Stokes' equations) we use in G the ansatz

$$u = -a\omega_2 E^2 1 + D^2\Theta - \alpha E^2\Theta^* - \beta b_\Theta \quad (0 < \alpha \in \mathbb{R},\ 0 \neq \beta \in \mathbb{R}).$$

Here $a \in \mathbb{R}$ is the prescribed constant from (2.5), $E^2 1$ is the simple layer potential with constant density $\Psi = 1$,

$$b_\Theta = \int_{\partial G} \Theta(y) \,\mathrm{d}o_y$$

is some constant, and the source density Θ^* is defined by

$$\Theta^*(x) = \Theta(x) - b_\Theta/(\mathrm{meas}(\partial G)), \tag{2.12}$$

which implies $b_{\Theta^*} = \int_{\partial G} \Psi^*(y) \,\mathrm{d}o_y = 0$. Note that the decay properties (2.5) are fulfilled in this case. Here again, (2.7) and (2.8) lead to the second kind Fredholm boundary integral equation

$$\Phi + a\omega_2 E^2 1 = -1/2\Theta + D^2\Theta - \alpha E^2\Theta^* - \beta b_\Theta \quad \text{on } \partial G. \tag{2.13}$$

To see that (2.13) has a unique solution $\Theta \in W^{2-1/q,q}(\partial G)$ for all boundary values $\Phi \in W^{2-1/q,q}(\partial G)$ and all $a \in \mathbb{R}$, let $0 \neq \Psi$ solve the homogeneous adjoint integral equation

$$0 = -1/2\Psi + H^2\Psi - \alpha E^2\Psi^* - \beta b_\Psi \quad \text{on } \partial G.$$

Because for any constant $c \in \mathbb{R}$ we have $-1/2c + D^2 c = 0$ [15, p. 511] and $E^2 c^* = 0$ (see (2.12) for the definition of c^*), we find

$$0 = \langle c, -1/2\Psi + H^2\Psi - \alpha E^2\Psi^* - \beta b_\Psi \rangle = -\beta \langle c, b_\Psi \rangle,$$

where here $\langle \psi, \varphi \rangle = \int_{\partial G} \psi(y)\varphi(y) \,\mathrm{d}o$ denotes the corresponding duality. It follows $b_\Psi = 0$ and $\Psi^* = \Psi$, hence Ψ is a solution of

$$0 = -1/2\Psi + H^2\Psi - \alpha E^2\Psi \quad \text{on } \partial G,$$

too. Now the same arguments as for (2.11) in the case $n \geq 3$ yield the assertion and the proposition is proved. □

3. The Poisson equation

The first theorem ensures the solvability of Poisson's equation (1.1) in the space $L^{2,q}(G)$, defined by (1.2).

Theorem 3.1. *Let $G \subset \mathbb{R}^n$ $(n \geq 2)$ be an exterior domain with boundary ∂G of class C^2, and let $1 < q < \infty$. Then for every $f \in L^q(G)$ and $\Phi \in W^{2-1/q,q}(\partial G)$ there exists some $u \in L^{2,q}(G)$ satisfying the Poisson equation (1.1) in G.*

Proof. Setting $f = 0$ in $\mathbb{R}^n/G$ we obtain some function $\tilde{f} \in L^q(\mathbb{R}^n)$ with $\tilde{f}_{|G} = f$ in G. Let $\tilde{f}_i \in C_0^\infty(\mathbb{R}^n)$ denote a sequence such that $\tilde{f}_i \to \tilde{f}$ in $L^q(\mathbb{R}^n)$ as $i \to \infty$. Consider now for fixed i the equation $-\Delta \tilde{u}_i = \tilde{f}_i$ in $\mathbb{R}^n$. We can solve it by convolution with E_n (see (2.4)), obtaining $x \in \mathbb{R}^n$ the representation

$$\tilde{u}_i(x) = (E_n * \tilde{f}_i)(x) = \int_{\mathbb{R}^n} E_n(x-y)\tilde{f}_i(y)\,\mathrm{d}y.$$

Moreover, by the theorem of Calderon-Zygmund [4], for the second order derivatives we obtain the estimate $\|\nabla^2 \tilde{u}_i\|_q \le c\|\tilde{f}_i\|_q$ with some constant c independent of $i \in \mathbb{N}$, which implies $\|\nabla^2(\tilde{u}_i - \tilde{u}_k)\|_q \to 0$ as $i, k \to \infty$.

Next consider a sequence of open balls $(B_j)_j$ with $B_j \subset B_{j+1}$ $\bigcup_{j=1}^\infty B_j = \mathbb{R}^n$. Let us define the space

$$\mathbb{P} = \{P : x \to P(x) = a + b \cdot x \mid b, x \in \mathbb{R}^n, a \in \mathbb{R}\} \tag{3.14}$$

of linear functions $P : \mathbb{R}^n \to \mathbb{R}$. Then by the generalized Poincaré inequality (compare [11, p. 22] or [13, p. 112]) we obtain for every $v \in L^{2,q}(\mathbb{R}^n)$ the estimate

$$\|v\|_{L^q(B_j)/\mathbb{P}} := \inf_{P \in \mathbb{P}} \|v + P\|_{L^q(B_j)} \le c_j \|\nabla^2 v\|_{L^q(B_j)n^2} \tag{3.15}$$

with some constants $c_j > 0$. Because $\tilde{u}_i \in L^{2,q}(\mathbb{R}^n)$ we conclude that $(\tilde{u}_i)_i$ is a Cauchy sequence with respect to the norm $\|\cdot\|_{L^q(B_1)/\mathbb{P}}$ on the left hand side of (3.15) for fixed $j = 1$. This implies the existence of linear functions $P_i \in \mathbb{P}$ such that $(\tilde{u}_i + P_i)_i$ is a Cauchy sequence in $L^q(B_1)$. Repeating this argument now for $j = 2$, there exist linear functions $Q_i \in \mathbb{P}$ such that $\tilde{u}_i + Q_i$ is a Cauchy sequence in $L^q(B_2)$, hence in $L^q(B_1)$, because $B_1 \subset B_2$. Thus the difference $(P_i - Q_i)_i$ is a Cauchy sequence in $L^q(B_1)$, and using the representation

$$P_i(x) = \alpha_i + B_i \cdot x,\ Q_i(x) = \gamma_i + \delta_i \cdot x,$$

we obtain that $(\alpha_i - \gamma_i)_i$ and $(\beta_i - \delta_i)_i$ are Cauchy sequences in $\mathbb{R}$ and in $\mathbb{R}^n$, respectively. From this we find that $(P_i - Q_i)_i$ is a Cauchy sequence in $L^q(B_2)$, and thus also $(\tilde{u}_i + P_i)_i = (\tilde{u}_i + Q_i)_i + (P_i - Q_i)_i$. Repeating this procedure it follows that $(\tilde{u}_i + P_i)_i$ is a Cauchy sequence in $L^q(B_j)$ for all $j = 1, 2, \ldots$. Thus we can find some $\tilde{u} \in L^{2,q}(\mathbb{R}^n)$ such that $(\tilde{u}_i + P_i) \to \tilde{u}$ in $L^q_{\mathrm{loc}}(\mathbb{R}^n)$ and $\|\nabla^2(\tilde{u} - \tilde{u}_i)\|_{q,\mathbb{R}^n} \to 0$ as $i \to \infty$. Moreover, $\tilde{u}$ satisfies $-\Delta\tilde{u} = \tilde{f}$ in $\mathbb{R}^n$ and the estimate $\|\nabla^2 \tilde{u}\|_q \le c\|\tilde{f}\|_q$. Since $\tilde{u} \in W^{2,q}_{\mathrm{loc}}(\mathbb{R}^n)$ we conclude from the usual trace theorem [1, p. 217] that $\tilde{u}_{|\partial G} \in W^{2-1/q,q}(\partial G)$. Following Proposition 2.1 there is a function $w \in L^{2,q}(G)$ satisfying the equations

$$-\Delta w = 0 \text{ in } G, \quad w_{|\partial G} = \tilde{u}_{|\partial G} - \Phi,$$

where $\Phi \in W^{2-1/q,q}(\partial G)$ is the prescribed boundary value. Now setting $u = \tilde{u}_{|G} - w$ we obtain the desired solution and the theorem is proved. □

Because functions $u \in L^{2,q}(G)$ have no suitable decay properties at infinity, in general we cannot expect uniqueness for the solution of (1.1) constructed in Theorem 3.1. Thus we consider in G the homogeneous equations and defined the nullspace of (1.1) by

$$N_q(G) = \{u \in L^{2,q}(G) \mid -\Delta u = 0 \text{ in } G,\ u_{|\partial G} = 0\}. \tag{3.16}$$

Theorem 3.2. *Let $G \subset \mathbb{R}^n$ $(n \ge 2)$ be an exterior domain with boundary ∂G of class C^2, and let $1 < q < \infty$. Then for the dimension $\dim N_q(G)$ of the nullspace defined in (3.16) we have $\dim N_q(G) = n + 1$ independent of q.*

Proof. Consider the space $\mathbb{P}$ of linear functions defined in (3.14). Because for every $P \in \mathbb{P}$ we have $P(x) = a + b \cdot x$ with some $a \in \mathbb{R}$ and some vector $b \in \mathbb{R}^n$ we find $\dim \mathbb{P} = n + 1$. Let u^P denote the uniquely determined solution of the equation

$$-\Delta u = 0,\ u_{|\partial G} = -P_{|\partial G}$$

with $P \in \mathbb{P}$, according to Lemma 2.1. Here in the case $n = 2$ we require

$$u(x) - a \ln|x| = 0(1) \quad \text{as } |x| \to \infty, \tag{3.17}$$

where the constant a is choosen from $P(x) = a + b \cdot x$. Setting

$$Mq(G) = \{u^P + P_{|\overline{G}} \mid P \in \mathbb{P}\}$$

we obtain $M_q(G) \subset N_q(G)$, obviously. Furthermore, we have $\dim M_q(G) = \dim \mathbb{P} = n + 1$, which can be shown as follows: Let $p(x) = a + b \cdot x$ and let $u^P + P_{|\overline{G}} = 0$ in $\overline{G}$. Then from the decay properties of u^P and ∇u^P established in Lemma 2.1 we find $a = 0$ and $b = 0$, hence $P = 0$. Here in the case $n = 2$ we obtain $a = 0$ due to the special choice of the number a in (3.17). Together with the uniqueness statement in Lemma 2.1 this means that, if B is a basis of $\mathbb{P}$, then

$$B_q(G) = \{u^P + P_{|\overline{G}} \mid P \in B\}$$

is a basis of $M_q(G)$. Thus it remains to show

$$N_q(G) \subset M_q(G). \tag{3.18}$$

To do so, let us first determine the null space

$$N_q(\mathbb{R}^n) = \{u \mid u \in L^{2,q}(\mathbb{R}^n) \text{ with } -\Delta u = 0 \quad \text{in } \mathbb{R}^n\}.$$

From $\Delta u = 0$, hence $\Delta \nabla^2 u = 0$ with $D^2_{jk} u \in L^q(\mathbb{R}^n)$ $(j, k = 1, \ldots, n)$ we obtain $\nabla^2 u = 0$ in $\mathbb{R}^n$, which implies $u = P$ for some $P \in \mathbb{P}$. Thus we have shown that

$$N_q(\mathbb{R}^n) = \mathbb{P}. \tag{3.19}$$

Now let $u \in N_q(G)$. We extend u on the whole space obtaining a function $\tilde{u} \in L^{2,q}(\mathbb{R}^n)$ with $\tilde{u}_{|G} = u$ [1, p. 83]. Moreover, this function satisfies on $\mathbb{R}^n$ the identity $-\Delta \tilde{u} = \tilde{f} \in L^q(\mathbb{R}^n)$, where the function $\tilde{f}$ has a compact support in the bounded domain $\mathbb{R}^n \setminus \overline{G}$. Consider the equations

$$-\Delta w = \tilde{f} \quad \text{in } \mathbb{R}^n. \tag{3.20}$$

Again, it can be solved by convolution with the fundamental solution E_n of the Laplacian: We obtain $w = E_n * \tilde{f}$ in $\mathbb{R}^n$ and the Calderon-Zygmund theorem implies $D^2_{jk} w \in L^r(\mathbb{R}^n)$ for all $1 < r \leq q$ $(j, k = 1, \ldots, n)$. Here we used $\tilde{f} \in L^r(\mathbb{R}^n)^n$ for all $1 < r \leq q$ due to its compact support. Now using a well-known estimate of Hardy-Littlewood-Sobolev-type [2, p. 242] we find $w \in L^s(\mathbb{R}^n)$ for some $s \geq q$, hence $w \in L^s_{\text{loc}}(\mathbb{R}^n) \subset L^q_{\text{loc}}(\mathbb{R}^n)$. Thus we have constructed some solution w of (3.20) such that $w \in L^{2,q}(\mathbb{R}^n)$. Setting $W = \tilde{u} - w$ we obtain $W \in N_q(\mathbb{R}^n)$, and (3.19) leads to $\tilde{u} = w + P$ for some $P \in \mathbb{P}$. Because $\tilde{u}_{|\partial G} = 0$ and since $\tilde{u}_{|G} = u$ we find $u \in M_q(G)$, which proves (3.18) and thus the theorem. □

REFERENCES

[1] R. A. Adams: *Sobolev Spaces*, New York-San Francisco-London: Academic Press (1975).
[2] L. Bers, F. John and M. Schechter: *Partial differential equations*, Providence R.I.: Wiley (1979).
[3] H. Brakhage, P. Werner: *Über das Dirichlet'sche Außenraumproblem für die Helmholtz'sche Schwingungschleichung*, Arch. Math., **16** (1965), 325–329.
[4] A. P. Calderon, A. Zygmund: *On singular integrals*. Amer. J. Math., **78** (1956) 289–309.
[5] J. Deny, J.-L. Lions: *Les espaces du type de Beppo Levi*, Ann. Inst. Fourier, **5** (1953–1954) 305–370.
[6] D. Fortunato: *On the index of elliptic partial differential operators in* $\mathbb{R}^n$, Ann. Mat. Pura Appl., **119** (1979), 317–331.
[7] J. Giroire: *Etude de quelques problemes aux limites exterieurs et resolution par equations integrales. Thèse de doctorat d'etat es sciences mathematiques*. Paris 6: Universite Pierre et Marie Curie (1987).
[8] N. M. Günter: *Die Potentialtheorie und ihre Anwendungen auf Grundaufgaben der mathematischen Physik*, Leipzig: Verlagsgessellschaft (1957).
[9] G. C. Hsiao, R. Kress: *On an integral equation for the two-dimensional exterior Stokes problem*. NAM-Bericht 33. Universität Göttingen (1983).
[10] R. Leis: *Zur Eindeutigkeit der Randwertaufgaben der Helmholtz'schen Schwingungsgleichung*, Math. Z., **85** (1964) 141–153.
[11] V. G. Maz'ja: *Sobolev spaces*, Berlin Heidelberg New York Tokyo, Springer Verlag (1985).
[12] R. McOwen: *The behaviour of the Laplacian on weighted Sobolev spaces*, Comm. Pure Appl. Math., **32** (1979), 783–795.
[13] J. Nečas: *Les mèthodes directes en thèorie des èquations elliptiques*, Prague, Academia (1967).
[14] J. Saranen, K. J. Witsch: *Exterior boundary value problems for elliptic equations*. Ann. Acad. Sci. Fennicae Series A. I. Mathematica, **8** (1983) 3–42.
[15] W. I. Smirnow: *Lehrgang der höheren Mathematik 4*, Berlin: Deutscher Verlag der Wissenschaften (1979).
[16] C. G. Simader: *On Dirichlet's boundary value problem*, Berlin Heidelberg New York, Springer Lecture Notes 268 (1972).
[17] W. Varnhorn: *The Poisson equation with weights in exterior domains of* $\mathbb{R}^n$, Applic. Anal., **43** (1992), 135–145.

WERNER VARNHORN
KASSEL UNIVERSITY
INSTITUTE OF MATHEMATICS, FACULTY OF MATHEMATICS AND NATURAL SCIENCES
34127 KASSEL, GERMANY
ORCID: 0000-0001-9486-1319
E-mail address: varnhorn@mathematik.uni-kassel.de

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 3, pp. 141-153
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Improvements of some Berezin radius inequalities

MEHMET GÜRDAL* AND MOHAMMAD W. ALOMARI

ABSTRACT. The Berezin transform $\widetilde{A}$ and the Berezin radius of an operator A on the reproducing kernel Hilbert space over some set Q with normalized reproducing kernel $k_\eta := \frac{K_\eta}{\|K_\eta\|}$ are defined, respectively, by $\widetilde{A}(\eta) = \langle Ak_\eta, k_\eta\rangle$, $\eta \in Q$ and $\mathrm{ber}(A) := \sup_{\eta\in Q}\left|\widetilde{A}(\eta)\right|$. A simple comparison of these properties produces the inequalities $\frac{1}{4}\|A^*A + AA^*\| \leq \mathrm{ber}^2(A) \leq \frac{1}{2}\|A^*A + AA^*\|$. In this research, we investigate other inequalities that are related to them. In particular, for $A \in \mathcal{L}(\mathcal{H}(Q))$ we prove that

$$\mathrm{ber}^2(A) \leq \frac{1}{2}\|A^*A + AA^*\|_{\mathrm{ber}} - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|}(\eta)\right) - \left(\widetilde{|A^*|}(\eta)\right)\right)^2.$$

Keywords: Mixed Schwarz inequality, Berezin radius, Furuta inequality.

2020 Mathematics Subject Classification: 47A30, 47A63.

1. INTRODUCTION

Many mathematicians and physicists are interested in the Berezin symbol of an operator defined with the help of a reproducing kernel Hilbert space. Several mathematicians have done extensive study on the Berezin radius inequality, which is presented in (1.1) (see [23]). Indeed, researchers are eager to obtain refinements and additions to this inequality given by (1.1) ([20], [32]). We show various inequalities for Berezin transformations of operators on the reproducing kernel Hilbert space $\mathcal{H}(Q)$ over some set Q in this study. By using Berezin transforms, we study several sharp inequalities involving powers of Berezin radius of some operators.

Remember that a reproducing kernel Hilbert space (abbreviated RKHS) is the Hilbert space $\mathcal{H} = \mathcal{H}(Q)$ of complex-valued functions on some set Q in which:

(a) the evaluation functionals

$$\varphi_\eta(f) = f(\eta), \eta \in Q,$$

are continuous on $\mathcal{H}$;

(b) for every $\eta \in Q$, there exists a function $f_\eta \in \mathcal{H}$ such that $f_\eta(\eta) \neq 0$.

Then, via the classical Riesz representation theorem, we know that if $\mathcal{H}$ is a RKHS on Q, there is a unique element $K_\eta \in \mathcal{H}$ such that $h(\eta) = \langle h, K_\eta\rangle_{\mathcal{H}}$ for every $\eta \in Q$ and all $h \in \mathcal{H}$. The reproducing kernel at η denoted by the element K_η. In addition, we shall refer to the normalized reproducing kernel at η as $k_\eta := \frac{K_\eta}{\|K_\eta\|}$. Let $\mathcal{L}(\mathcal{H})$ be the Banach algebra of all

Received: 28.04.2022; Accepted: 03.08.2022; Published Online: 12.08.2022
*Corresponding author: Mehmet Gürdal; gurdalmehmet@sdu.edu.tr
DOI: 10.33205/cma.1110550

bounded linear operators on a complex Hilbert space $\mathcal{H}$ including the identity operator $1_{\mathcal{H}}$ in $\mathcal{L}(\mathcal{H})$. The Berezin transform (symbol) of A is the complex-valued function on Q defined by

$$\widetilde{A}(\eta) := \langle Ak_\eta, k_\eta \rangle$$

for an operator $A \in \mathcal{L}(\mathcal{H})$.

The Berezin transform $\widetilde{A}$ is obviously a bounded function on Q and $\sup_{\eta\in Q} \left|\widetilde{A}(\eta)\right|$, which is known as the Berezin radius (number) of operator A, i.e.,

$$\operatorname{ber}(A) := \sup_{\eta\in Q} \left|\widetilde{A}(\eta)\right|.$$

The Berezin transform $\widetilde{A}$ is a bounded real-analytic function on Ω for any bounded operator A on $\mathcal{H}$. Properties of the operator A are often reflected in properties of the Berezin transform $\widetilde{A}$. F. Berezin proposed the Berezin transform in [7], and it has proven to be a valuable tool in operator theory, since many fundamental features of significant operators are stored in their Berezin transforms. The Berezin set and number, also known as $\operatorname{Ber}(A)$ and $\operatorname{ber}(A)$, were allegedly first publicly proposed by Karaev in [22].

The range of the Berezin transform $\widetilde{A}$, which is stated to be the Berezin set of operator A, is also obvious from the definition of the Berezin transform, i.e.,

$$\operatorname{Ber}(A) := \operatorname{Range}\left(\widetilde{A}\right) = \left\{\widetilde{A}(\eta) : \eta \in Q\right\}.$$

Recall that the numerical radius of operator A is defined by

$$w(A) := \sup_{\|x\|=1} |\langle Ax, x\rangle|.$$

It is well-known that

$$\operatorname{ber}(A) \le w(A) \le \|A\| \tag{1.1}$$

for any $X \in \mathcal{L}(\mathcal{H})$. See [1, 2, 9, 8, 17, 19, 24, 25, 30, 35] for further details.

Berezin set and Berezin radius of operators are new numerical properties of RKHS operators presented by Karaev in [21]. See [5, 12, 23] for the fundamental features and information about these new categories.

In 2021, Huban et al. [20] improved the inequality (1.1) by proving that

$$\operatorname{ber}(A) \le \frac{1}{2}\left(\|A\|_{\operatorname{ber}} + \left\|A^2\right\|_{\operatorname{ber}}^{1/2}\right) \tag{1.2}$$

for any $A \in \mathcal{L}(\mathcal{H})$.

It has been shown in [20] that if $A \in \mathcal{B}(\mathcal{H})$, then

$$\frac{1}{4}\|A^*A + AA^*\| \le \operatorname{ber}^2(A) \le \frac{1}{2}\|A^*A + AA^*\|. \tag{1.3}$$

Inspired by the numerical radius inequalities in [3], this study proves an extension of the inequality (1.3). In particular, for $A \in \mathcal{L}(\mathcal{H}(Q))$ we prove that

$$\operatorname{ber}^2(A) \le \frac{1}{2}\|A^*A + AA^*\|_{\operatorname{ber}} - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|}(\eta)\right) - \left(\widetilde{|A^*|}(\eta)\right)\right)^2.$$

Other general-related outcomes have also been established.

2. Auxiliary Theorems

The following chain of corollaries is required to attain our aim.

According to the basic Schwarz inequality for positive operators, if $A \in \mathcal{L}(\mathcal{H})$ is a positive operators, then

$$|\langle Ax_1, x_2\rangle|^2 \leq \langle Ax_1, x_1\rangle \langle Ax_2, x_2\rangle \tag{2.4}$$

for any $x_1, x_2 \in \mathcal{H}$.

Reid [28] demonstrated an inequality in 1951 that may be regarded a version of the Schwarz inequality. In fact, he proved that for all $x_1 \in \mathcal{H}$

$$|\langle ABx_1, x_2\rangle| \leq \|B\| \langle Ax_1, x_1\rangle \tag{2.5}$$

for any operators $A \in \mathcal{L}(\mathcal{H})$ where A is positive and AB is selfadjoint.

Kato [27] established a companion inequality of (2.4), the mixed Schwarz inequality, in 1952, which claims

$$|\langle Ax_1, x_2\rangle|^2 \leq \left\langle |A|^{2\alpha} x_1, x_1\right\rangle \left\langle |A^*|^{2(1-\alpha)} x_2, x_2\right\rangle, \ 0 \leq \alpha \leq 1 \tag{2.6}$$

for every operators $A \in \mathcal{L}(\mathcal{H})$ and any vectors $x_1, x_2 \in \mathcal{H}$, where $|A| = (A^*A)^{1/2}$.

In 1988, Kittaneh [24] proved a very interesting extension combining both the Halmos-Reid inequality (2.5) and the mixed Schwarz inequality (2.6). His result reads that

$$|\langle ABx_1, x_2\rangle| \leq r(B) \|f(|A|) x_1\| \|g(|A^*|) x_2\| \tag{2.7}$$

for any vectors $x_1, x_2 \in \mathcal{H}$, where $A, B \in \mathcal{L}(\mathcal{H})$ such that $|A|B = B^*|A|$ and f, g are nonnegative continuous functions defined on $[0, \infty)$ satisfying that $f(t)g(t) = t$ $(t \geq 0)$. Clearly, when we select $f(t) = t^\alpha$ and $g(t) = t^{1-\alpha}$ with $B = 1_{\mathcal{H}}$, we are referring to the inequality (2.6). Furthermore, several alterations that are chosen $\alpha = \frac{1}{2}$ pertain to the Halmos version of the Reid inequality.

Furuta [11] demonstrated another extension of Kato's inequality (2.6) in 1994, as follows:

$$\left|\left\langle A|A|^{\alpha+\beta-1} x_1, x_2\right\rangle\right|^2 \leq \left\langle |A|^{2\alpha} x_1, x_1\right\rangle \left\langle |A|^{2\beta} x_2, x_2\right\rangle \tag{2.8}$$

for any $x_1, x_2 \in \mathcal{H}$ and $\alpha, \beta \in [0, 1]$ with $\alpha + \beta \geq 1$.

The inequality (2.8) was generalized for any $\alpha, \beta \geq 0$ with $\alpha + \beta \geq 1$ by Dragomir in [10]. Indeed, as Dragomir pointed out, Furuta adopted the condition $\alpha, \beta \in [0, 1]$ to match with the Heinz-Kato inequality, which reads:

$$|\langle Ax_1, x_2\rangle| \leq \|T^\alpha x_1\| \left\|S^{1-\alpha} x_2\right\|$$

for any $x_1, x_2 \in \mathcal{H}$ and $\alpha \in [0, 1]$ where T and S are positive operators such that $\|Ax_1\| \leq \|Tx_1\|$and $\|A^*x_2\| \leq \|Sx_2\|$ for any $x_1, x_2 \in \mathcal{H}$.

Lemma 2.1. *If $B \in \mathcal{L}(\mathcal{H})$, $B \geq 0$ and $x_1 \in H$ is any unit vector, then there's*

$$\langle Bx, x\rangle^r \leq (\geq) \langle B^r x, x\rangle, \ r \geq 1 \ (0 \leq r \leq 1). \tag{2.9}$$

Kittaneh and Manasrah [26] discovered this conclusion, which is a refinement of the scalar Young inequality.

Lemma 2.2. *If $a, b \geq 0$, and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, then we have*

$$ab + \min\left\{\frac{1}{p}, \frac{1}{q}\right\} \left(a^{p/2} - b^{q/2}\right)^2 \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{2.10}$$

Sheikhhosseini et al. [31] recently found the following generalization of (2.10).

Lemma 2.3. *If $a,b>0$, and $p,q>1$ such that $\frac{1}{p}+\frac{1}{q}=1$, then for $m=1,2,3,...,$*

$$\left(a^{1/p}b^{1/q}\right)^m+r_0^m\left(a^{m/2}-b^{m/2}\right)^2\leq\left(\frac{a^r}{p}+\frac{b^r}{q}\right)^{m/r},\ r\geq 1, \tag{2.11}$$

where $r_0=\min\left\{\frac{1}{p},\frac{1}{q}\right\}$. In particular, if $p=q=2$, then

$$\left(a^{1/2}b^{1/2}\right)^m+2^{-m}\left(a^{m/2}-b^{m/2}\right)^2\leq 2^{-m/r}\left(a^r+b^r\right)^{m/r}.$$

For $m=1$,

$$\left(a^{1/2}b^{1/2}\right)+2^{-1}\left(a^{1/2}-b^{1/2}\right)^2\leq 2^{-1/r}\left(a^r+b^r\right)^{1/r}.$$

3. Main result

We are now prepared to provide this section's primary results. The section's next finding is a revised Berezin radius inequality.

Theorem 3.1. *If $A\in\mathcal{L}(\mathcal{H}(Q))$ and $\alpha,\beta\geq 0$ such that $\alpha+\beta\geq 1$, then we get*

$$\begin{aligned}\operatorname{ber}^m\left(A\left|A\right|^{\alpha+\beta-1}\right)&\leq\frac{1}{2^{m/r}}\left\|\left|A\right|^{2r\alpha}+\left|A^*\right|^{2r\beta}\right\|_{\operatorname{ber}}^{m/r}\\&\quad-\frac{1}{2^m}\inf_{\eta\in Q}\left(\left(\widetilde{\left|A\right|^{2\alpha}}(\eta)\right)^{m/2}-\left(\widetilde{\left|A^*\right|^{2\beta}}(\eta)\right)^{m/2}\right)^2\end{aligned} \tag{3.12}$$

for all $r\geq 1$.

Proof. For all $m\geq 1$, on choosing $x_1=k_\eta$ and $x_2=k_\tau$ in the inequality (2.8), we get

$$\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta,k_\tau\right\rangle\right|^m\leq\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{\frac{m}{2}}\left\langle\left|A^*\right|^{2\beta}k_\tau,k_\tau\right\rangle^{\frac{m}{2}}.$$

By the inequalities (2.9) and (2.11), for $\eta\in Q$ with $\eta=\tau$ we have

$$\begin{aligned}\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta,k_\eta\right\rangle\right|^m&\leq\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{\frac{m}{2}}\left\langle\left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^{m/2}.\\&\leq\left(\frac{\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^r+\left\langle\left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^r}{2}\right)^{m/r}\\&\quad-\frac{1}{2^m}\left(\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{m/2}-\left\langle\left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^{m/r}\right)^2\\&\leq\left(\frac{\left\langle\left|A\right|^{2r\alpha}k_\eta,k_\eta\right\rangle+\left\langle\left|A^*\right|^{2r\beta}k_\eta,k_\eta\right\rangle}{2}\right)^{m/r}\\&\quad-\frac{1}{2^m}\left(\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{m/2}-\left\langle\left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^{m/2}\right)^2,\end{aligned}$$

and

$$\begin{aligned}\sup_{\eta\in Q}\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta,k_\eta\right\rangle\right|^m&\leq\frac{1}{2^{m/r}}\sup_{\eta\in Q}\left(\left\langle\left|A\right|^{2r\alpha}k_\eta,k_\eta\right\rangle+\left\langle\left|A^*\right|^{2r\beta}k_\eta,k_\eta\right\rangle\right)^{m/r}\\&\quad-\frac{1}{2^m}\inf_{\eta\in Q}\left(\left\langle\left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{m/2}-\left\langle\left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^{m/2}\right)^2\end{aligned}$$

which is equivalent to

$$\operatorname{ber}^m\left(A|A|^{\alpha+\beta-1}\right) \leq \frac{1}{2^{m/r}}\left\||A|^{2r\alpha}+|A^*|^{2r\beta}\right\|_{\mathrm{ber}}^{m/r} - \frac{1}{2^m}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{m/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{m/2}\right)^2,$$

and completes the theorem's proof. □

We get the following result by putting $m=2$ in (3.12).

Corollary 3.1. *If $A\in\mathcal{L}(\mathcal{H}(Q))$ and $\alpha,\beta\geq 0$ such that $\alpha+\beta\geq 1$, then we have*

$$\operatorname{ber}^2\left(A|A|^{\alpha+\beta-1}\right) \leq \frac{1}{2^{2/r}}\left\||A|^{2r\alpha}+|A^*|^{2r\beta}\right\|_{\mathrm{ber}}^{2/r} - \frac{1}{4}\inf_{\eta\in Q}\left(\widetilde{|A|^{2\alpha}}(\eta)-\widetilde{|A^*|^{2\beta}}(\eta)\right)^2 \tag{3.13}$$

for all $r\geq 1$.

By choosing $r=1$ in (3.13), we get

$$\operatorname{ber}^2\left(A|A|^{\alpha+\beta-1}\right) \leq \frac{1}{4}\left\||A|^{2\alpha}+|A^*|^{2\beta}\right\|_{\mathrm{ber}}^{2} - \frac{1}{4}\inf_{\eta\in Q}\left(\widetilde{|A|^{2\alpha}}(\eta)-\widetilde{|A^*|^{2\beta}}(\eta)\right)^2 \tag{3.14}$$

for all $\alpha,\beta\geq 0$ such that $\alpha+\beta\geq 1$.

Also for $\alpha=\beta=\frac{1}{2}$ in (3.14), we get

$$\operatorname{ber}^2(A)\leq\frac{1}{4}\||A|+|A^*|\|_{\mathrm{ber}}^2-\frac{1}{4}\inf_{\eta\in Q}\left(\widetilde{|A|}(\eta)-\widetilde{|A^*|}(\eta)\right)^2.$$

In particular, take $\alpha=\beta=1$, we have

$$\operatorname{ber}^2(A|A|)\leq\frac{1}{4}\left\||A|^2+|A^*|^2\right\|_{\mathrm{ber}}^2-\frac{1}{4}\inf_{\eta\in Q}\left(\widetilde{|A|^2}(\eta)-\widetilde{|A^*|^2}(\eta)\right)^2$$

and

$$\operatorname{ber}^2(A|A|)\leq\frac{1}{4}\|A^*A+AA^*\|_{\mathrm{ber}}^2-\frac{1}{4}\inf_{\eta\in Q}\left([A^*\widetilde{A-A}A^*](\eta)\right)^2.$$

A generalization of the above findings may be expressed as follows:

Theorem 3.2. *If $A\in\mathcal{L}(\mathcal{H}(Q))$ and $\alpha,\beta\geq 0$ such that $\alpha+\beta\geq 1$, then we have*

$$\operatorname{ber}^{2s}\left(A|A|^{\alpha+\beta-1}\right) \leq \frac{1}{2^{2/r}}\left\||A|^{2rs\alpha}+|A^*|^{2rs\beta}\right\|_{\mathrm{ber}}^{2/r} - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2sr\alpha}}(\eta)\right)-\left(\widetilde{|A^*|^{2sr\beta}}(\eta)\right)\right) \tag{3.15}$$

for all $r,s\geq 1$.

Proof. Setting $x_1 = x_2 = k_\eta$ in (2.8) and then using Lemma 2.3 with $p = q = 2$ and $m = 2$, we get

$$\begin{aligned}
\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta, k_\eta\right\rangle\right|^{2s} &\leq \left\langle \left|A\right|^{2\alpha}k_\eta, k_\eta\right\rangle^{s}\left\langle \left|A^*\right|^{2\beta}k_\eta, k_\eta\right\rangle^{s} \ (t^s \text{ increasing}) \\
&\leq \left\langle \left|A\right|^{2s\alpha}k_\eta, k_\eta\right\rangle \left\langle \left|A^*\right|^{2s\beta}k_\eta, k_\eta\right\rangle \text{ (by convexity of } t^s) \\
&\leq \frac{1}{2^{2/r}}\left(\left\langle \left|A\right|^{2s\alpha}k_\eta, k_\eta\right\rangle^{r} + \left\langle \left|A^*\right|^{2s\beta}k_\eta, k_\eta\right\rangle^{r}\right)^{2/r} \\
&\quad \text{(by the inequality (2.11))} \\
&\quad - \frac{1}{4}\left[\left\langle \left|A\right|^{2sr\alpha}k_\eta, k_\eta\right\rangle - \left\langle \left|A^*\right|^{2rs\beta}k_\eta, k_\eta\right\rangle\right] \\
&\leq \frac{1}{2^{2/r}}\left(\left\langle \left|A\right|^{2rs\alpha}k_\eta, k_\eta\right\rangle + \left\langle \left|A^*\right|^{2rs\beta}k_\eta, k_\eta\right\rangle\right)^{2/r} \\
&\quad \text{(by the inequality (2.9))} \\
&\quad - \frac{1}{4}\left[\left\langle \left|A\right|^{2sr\alpha}k_\eta, k_\eta\right\rangle - \left\langle \left|A^*\right|^{2rs\beta}k_\eta, k_\eta\right\rangle\right].
\end{aligned}$$

Equivalenty, we may write

$$\begin{aligned}
\left|\widetilde{A\left|A\right|^{\alpha+\beta-1}}(\eta)\right|^{2s} &\leq \frac{1}{2^{2/r}}\left(\widetilde{\left|A\right|^{2rs\alpha}}(\eta) + \widetilde{\left|A^*\right|^{2rs\beta}}(\eta)\right)^{2/r} \\
&\quad - \frac{1}{4}\inf_{\eta\in Q}\left[\left(\widetilde{\left|A\right|^{2sr\alpha}}(\eta)\right) - \left(\widetilde{\left|A^*\right|^{2sr\beta}}(\eta)\right)\right].
\end{aligned}$$

By taking the supremum over $\eta \in Q$, we obtain

$$\begin{aligned}
\mathrm{ber}^{2s}\left(A\left|A\right|^{\alpha+\beta-1}\right) &\leq \frac{1}{2^{2/r}}\left\|\left|A\right|^{2rs\alpha} + \left|A^*\right|^{2rs\beta}\right\|_{\mathrm{ber}}^{2/r} \\
&\quad - \frac{1}{4}\inf_{\eta\in Q}\left[\left(\widetilde{\left|A\right|^{2sr\alpha}}(\eta)\right) - \left(\widetilde{\left|A^*\right|^{2sr\beta}}(\eta)\right)\right],
\end{aligned}$$

which completes the proof. □

We get the following result by setting $r = 1$ in (3.15).

Corollary 3.2. *If $A \in \mathcal{L}(\mathcal{H}(Q))$ and $\alpha, \beta \geq 0$ such that $\alpha + \beta \geq 1$, then we have*

$$\begin{aligned}
\mathrm{ber}^{2s}\left(A\left|A\right|^{\alpha+\beta-1}\right) &\leq \frac{1}{4}\left\|\left|A\right|^{2s\alpha} + \left|A^*\right|^{2s\beta}\right\|_{\mathrm{ber}}^{2} \\
&\quad - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{\left|A\right|^{2s\alpha}}(\eta)\right) - \left(\widetilde{\left|A^*\right|^{2s\beta}}(\eta)\right)\right) \qquad (3.16)
\end{aligned}$$

for all $s \geq 1$.

In (3.16), let $\alpha = \beta = \frac{1}{2}$ we get

$$\mathrm{ber}^{2s}(A) \leq \frac{1}{4}\left\||A|^s + |A^*|^s\right\|_{\mathrm{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^s}(\eta)\right) - \left(\widetilde{|A^*|^s}(\eta)\right)\right)$$

for every $s \geq 1$. We have, in particular, for $s = 1$

$$\mathrm{ber}^{2}(A) \leq \frac{1}{4}\left\||A| + |A^*|\right\|_{\mathrm{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|}(\eta)\right) - \left(\widetilde{|A^*|}(\eta)\right)\right).$$

By choosing $\alpha = \beta = \frac{1}{s}$, $(s \geq 1)$, in (3.16) we get

$$\operatorname{ber}^{2s}\left(A|A|^{\frac{2}{s}-1}\right) \leq \frac{1}{4}\left\||A|^2 + |A^*|^2\right\|_{\operatorname{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right) - \left(\widetilde{|A^*|^2}(\eta)\right)\right). \tag{3.17}$$

Also for $s = 1$ in (3.17), we get

$$\operatorname{ber}^2(A|A|) \leq \frac{1}{4}\left\||A|^2 + |A^*|^2\right\|_{\operatorname{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right) - \left(\widetilde{|A^*|^2}(\eta)\right)\right), \tag{3.18}$$

and

$$\operatorname{ber}^2(A|A|) \leq \frac{1}{4}\|A^*A + AA^*\|_{\operatorname{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right) - \left(\widetilde{|A^*|^2}(\eta)\right)\right).$$

Remark 3.1. *By choosing $\alpha = \beta = \frac{1}{2}$, $s = 1$, $r = 2$ in (3.16), we have*

$$\operatorname{ber}^2(A) \leq \frac{1}{2}\||A| + |A^*|\|_{\operatorname{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right) - \left(\widetilde{|A^*|^2}(\eta)\right)\right)$$

or

$$\operatorname{ber}^2(A) \leq \frac{1}{2}\|A^*A + AA^*\|_{\operatorname{ber}}^2 - \frac{1}{4}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right) - \left(\widetilde{|A^*|^2}(\eta)\right)\right). \tag{3.19}$$

This improves the upper bound of the inequality (1.2).

Theorem 3.3. *If $A \in \mathcal{L}(\mathcal{H}(Q))$ and $\alpha, \beta \geq 0$ such that $\alpha + \beta \geq 1$, then we have*
(i)

$$\operatorname{ber}^{2s}\left(A|A|^{\alpha+\beta-1}\right) \leq \left\|\frac{1}{p}|A|^{2sp\alpha} + \frac{1}{q}|A^*|^{2sq\beta}\right\|_{\operatorname{ber}} - r_0\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2s\alpha}}(\eta)\right)^{p/2} - \left(\widetilde{|A^*|^{2s\beta}}(\eta)\right)^{q/2}\right)^2 \tag{3.20}$$

for every $s \geq 1$ and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, where $r_0 := \min\left\{\frac{1}{p}, \frac{1}{q}\right\}$.
(ii)

$$\operatorname{ber}^{2s}\left(A|A|^{\alpha+\beta-1}\right) \leq \frac{1}{2}\left\||A|^{4s\alpha} + |A^*|^{4s\beta}\right\|_{\operatorname{ber}} - \frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2s\alpha}}(\eta)\right) - \left(\widetilde{|A^*|^{2s\beta}}(\eta)\right)\right)^2. \tag{3.21}$$

Proof. Now, as in (2.8) but with $x_1 = x_2 = k_\eta$, we have by convexity of t^s

$$\begin{aligned}
\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta,k_\eta\right\rangle\right|^{2s} &\leq \left\langle \left|A\right|^{2\alpha}k_\eta,k_\eta\right\rangle^{s}\left\langle \left|A^*\right|^{2\beta}k_\eta,k_\eta\right\rangle^{s}\\
&\text{(by the inequaltiy (2.8))}\\
&\leq \left\langle \left|A\right|^{2s\alpha}k_\eta,k_\eta\right\rangle\ \left\langle \left|T^*\right|^{2s\beta}k_\eta,k_\eta\right\rangle\\
&\leq \frac{1}{p}\left\langle \left|A\right|^{2s\alpha}k_\eta,k_\eta\right\rangle^{p}+\frac{1}{q}\left\langle \left|A^*\right|^{2s\beta}k_\eta,k_\eta\right\rangle^{q}\\
&\text{(by the inequality (2.10))}\\
&-r_0\left(\left\langle \left|A\right|^{2s\alpha}k_\eta,k_\eta\right\rangle^{\frac{p}{2}}-\left\langle \left|A^*\right|^{2s\beta}k_\eta,k_\eta\right\rangle^{\frac{q}{2}}\right)^2\\
&\leq \frac{1}{p}\left\langle \left|A\right|^{2sp\alpha}k_\eta,k_\eta\right\rangle+\frac{1}{q}\left\langle \left|A^*\right|^{2sq\beta}k_\eta,k_\eta\right\rangle\\
&\text{(by the inequality (2.9))}\\
&-r_0\left(\left\langle \left|A\right|^{2s\alpha}k_\eta,k_\eta\right\rangle^{\frac{p}{2}}-\left\langle \left|A^*\right|^{2s\beta}k_\eta,k_\eta\right\rangle^{\frac{q}{2}}\right)^2
\end{aligned}$$

for $s \geq 1$. Thus,

$$\begin{aligned}
\left|\left\langle A\left|A\right|^{\alpha+\beta-1}k_\eta,k_\eta\right\rangle\right|^{2s} &\leq \frac{1}{p}\left\langle \left|A\right|^{2sp\alpha}k_\eta,k_\eta\right\rangle+\frac{1}{q}\left\langle \left|A^*\right|^{2sq\beta}k_\eta,k_\eta\right\rangle\\
&-r_0\left(\left\langle \left|A\right|^{2s\alpha}k_\eta,k_\eta\right\rangle^{\frac{p}{2}}-\left\langle \left|A^*\right|^{2s\beta}k_\eta,k_\eta\right\rangle^{\frac{q}{2}}\right)^2,
\end{aligned}$$

and by taking supremum over $\eta \in Q$, we then obtain the first inequality

$$\operatorname{ber}^{2s}\left(A\left|A\right|^{\alpha+\beta-1}\right) \leq \frac{1}{2}\left\|\left|A\right|^{4s\alpha}+\left|A^*\right|^{4s\beta}\right\|_{\text{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{\left|A\right|^{2s\alpha}}(\eta)\right)-\left(\widetilde{\left|A^*\right|^{2s\beta}}(\eta)\right)\right)^2$$

as required. Taking $p = q = 2$, we get the particular case (3.21). □

Several intriguing particular situations might be drawn from this (3.12).
When we put $\alpha = \beta = \frac{1}{2}$ in (3.13), we get

$$\operatorname{ber}^{2s}(A) \leq \frac{1}{2}\left\|\left|A\right|^{2s}+\left|A^*\right|^{2s}\right\|_{\text{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\widetilde{\left|A\right|^{s}}(\eta)-\widetilde{\left|A^*\right|^{s}}(\eta)\right)^2$$

for every $s \geq 1$. We have, in particular, for $s = 1$

$$\operatorname{ber}^{2}(A) \leq \frac{1}{2}\left\|\left|A\right|^{2}+\left|A^*\right|^{2}\right\|_{\text{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\widetilde{\left|A\right|}(\eta)-\widetilde{\left|A^*\right|}(\eta)\right)^2,$$

which can be written as

$$\operatorname{ber}^{2}(A) \leq \frac{1}{2}\left\|A^*A+AA^*\right\|_{\text{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\widetilde{\left|A\right|}(\eta)-\widetilde{\left|A^*\right|}(\eta)\right)^2. \tag{3.22}$$

and this refines the upper bound of the refinement of the inequality (1.3). Clearly, (3.22) is better than (3.19) which in turn bettern that (1.2).

Remark 3.2. *(i) When we set $\alpha=\beta=1$ in (3.20), we get*

$$\operatorname{ber}^{2s}(A|A|)\leq\left\|\frac{1}{p}|A|^{2sp}+\frac{1}{q}|A^*|^{2sq}\right\|_{\operatorname{ber}}-r_0\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2s}}(\eta)\right)^{p/2}-\left(\widetilde{|A^*|^{2s}}(\eta)\right)^{q/2}\right)^2$$

for every $s\geq 1$ and $p,q>1$ such that $\frac{1}{p}+\frac{1}{q}=1$, where $r_0:=\min\left\{\frac{1}{p},\frac{1}{q}\right\}$.

(ii) Choose $s=1$ and $p=q=2$ in the above inequality, we get

$$\operatorname{ber}^2(A|A|)\leq\left\|\frac{1}{p}|A|^4+\frac{1}{q}|A^*|^4\right\|_{\operatorname{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right)-\left(\widetilde{|A^*|^2}(\eta)\right)\right)^2.$$

The Berezin radius inequality of Hilbert space operators of a certain kind for commutators may be proven as follows:

Theorem 3.4. *If $A,B\in\mathcal{L}(\mathcal{H}(Q))$ and $\alpha,\beta,\gamma,\delta\geq 0$ such that $\alpha+\beta\geq 1$ and $\gamma+\delta\geq 1$, then we have*

$$\begin{aligned}(3.23)\qquad &\operatorname{ber}\left(A|A|^{\alpha+\beta-1}+B|B|^{\gamma+\delta-1}\right)\\ \leq&\frac{1}{2^{1/r}}\left\||A|^{2r\alpha}+|A^*|^{2r\beta}\right\|_{\operatorname{ber}}^{1/r}+\frac{1}{2^{1/r}}\left\||B|^{2r\gamma}+|B^*|^{2r\delta}\right\|_{\operatorname{ber}}^{1/r}\\ -&\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{1/2}\right)^2\\ -&\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^{2\gamma}}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^{2\delta}}(\eta)\right)^{1/2}\right)^2\end{aligned}$$

for all $r\geq 1$.

Proof. Using the triangle inequality, we get

$$\begin{aligned}&\left|\left\langle\left(A|A|^{\alpha+\beta-1}+B|B|^{\gamma+\delta-1}\right)k_\eta,k_\eta\right\rangle\right|\\ \leq&\left|\left\langle A|A|^{\alpha+\beta-1}k_\eta,k_\eta\right\rangle\right|+\left|\left\langle B|B|^{\gamma+\delta-1}k_\eta,k_\eta\right\rangle\right|\\ \leq&\left\langle|A|^{2\alpha}k_\eta,k_\eta\right\rangle^{\frac{1}{2}}\left\langle|A^*|^{2\beta}k_\eta,k_\eta\right\rangle^{\frac{1}{2}}\left\langle|B|^{2\gamma}k_\eta,k_\eta\right\rangle^{\frac{1}{2}}\left\langle|B^*|^{2\delta}k_\eta,k_\eta\right\rangle^{1/2}\\ &\text{(by the inequality (2.8))}\\ \leq&\frac{1}{2^{1/r}}\left(\left\langle|A|^{2\alpha}k_\eta,k_\eta\right\rangle^r+\left\langle|A^*|^{2\beta}k_\eta,k_\eta\right\rangle^r\right)^{1/r}\\ &\text{(by the inequality (2.11))}\\ -&\frac{1}{2}\left(\left\langle|A|^{2\alpha}k_\eta,k_\eta\right\rangle^{1/2}-\left\langle|A^*|^{2\beta}k_\eta,k_\eta\right\rangle^{1/2}\right)^2\\ +&2^{\frac{-1}{r}}\left(\left\langle|B|^{2\gamma}k_\eta,k_\eta\right\rangle^r+\left\langle|B^*|^{2\delta}k_\eta,k_\eta\right\rangle^r\right)^{1/r}\\ -&\frac{1}{2}\left(\left\langle|B|^{2\gamma}k_\eta,k_\eta\right\rangle^{1/2}-\left\langle|B^*|^{2\delta}k_\eta,k_\eta\right\rangle^{1/2}\right)^2\end{aligned}$$

$$
\begin{aligned}
&\leq 2^{\frac{-1}{r}}\left(\left\langle |A|^{2r\alpha}k_\eta,k_\eta\right\rangle+\left\langle |A^*|^{2r\beta}k_\eta,k_\eta\right\rangle\right)^{1/r}\\
&\quad\text{(by the inequality (2.9))}\\
&-\frac{1}{2}\left(\left\langle |A|^{2\alpha}k_\eta,k_\eta\right\rangle^{1/2}-\left\langle |A^*|^{2\beta}k_\eta,k_\eta\right\rangle^{1/2}\right)^2\\
&+\frac{1}{2^{1/r}}\left(\left\langle |B|^{2r\gamma}k_\eta,k_\eta\right\rangle+\left\langle |B^*|^{2r\delta}k_\eta,k_\eta\right\rangle\right)^{1/r}\\
&-\frac{1}{2}\left(\left\langle |B|^{2\gamma}k_\eta,k_\eta\right\rangle^{1/2}-\left\langle |B^*|^{2\delta}k_\eta,k_\eta\right\rangle^{1/2}\right)^2,
\end{aligned}
$$

and so

$$
\begin{aligned}
\left|\left(A\,|A|^{\widetilde{\alpha+\beta-1}}+B\,|B|^{\gamma+\delta-1}\right)(\eta)\right| &\leq \frac{1}{2^{1/r}}\left(\widetilde{|A|^{2r\alpha}}(\eta)+\widetilde{|A^*|^{2r\beta}}(\eta)\right)^{1/r}\\
&+\frac{1}{2^{1/r}}\left(\widetilde{|B|^{2r\gamma}}(\eta)+\widetilde{|B^*|^{2r\delta}}(\eta)\right)^{1/r}\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{1/2}\right)^2\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^{2\gamma}}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^{2\delta}}(\eta)\right)^{1/2}\right)^2.
\end{aligned}
$$

By taking supremum over $\eta\in Q$ above inequality, we have

$$
\begin{aligned}
\sup_{\eta\in Q}\left|\left(A\,|A|^{\widetilde{\alpha+\beta-1}}+B\,|B|^{\gamma+\delta-1}\right)(\eta)\right| &\leq \frac{1}{2^{1/r}}\sup_{\eta\in Q}\left(\widetilde{|A|^{2r\alpha}}(\eta)+\widetilde{|A^*|^{2r\beta}}(\eta)\right)^{1/r}\\
&+\frac{1}{2^{1/r}}\sup_{\eta\in Q}\left(\widetilde{|B|^{2r\gamma}}(\eta)+\widetilde{|B^*|^{2r\delta}}(\eta)\right)^{1/r}\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{1/2}\right)^2\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^{2\gamma}}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^{2\delta}}(\eta)\right)^{1/2}\right)^2
\end{aligned}
$$

which clearly implies that

$$
\begin{aligned}
\operatorname{ber}\left(A\,|A|^{\alpha+\beta-1}+B\,|B|^{\gamma+\delta-1}\right) &\leq \frac{1}{2^{1/r}}\left\||A|^{2r\alpha}+|A^*|^{2r\beta}\right\|_{\operatorname{ber}}^{1/r}\\
&+\frac{1}{2^{1/r}}\left\||B|^{2r\gamma}+|B^*|^{2r\delta}\right\|_{\operatorname{ber}}^{1/r}\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{1/2}\right)^2\\
&-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^{2\gamma}}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^{2\delta}}(\eta)\right)^{1/2}\right)^2.
\end{aligned}
$$

Then the desired result has been obtained. □

Using $r=1$ in the proof of Theorem 3.4, we achieve the desired result.

Corollary 3.3. *If $A,B\in\mathcal{L}(\mathcal{H}(Q))$ and $\alpha,\beta,\gamma,\delta\geq 0$ such that $\alpha+\beta\geq 1$ and $\gamma+\delta\geq 1$, then we have*

$$\begin{aligned}(3.24)\qquad &\operatorname{ber}\left(A|A|^{\alpha+\beta-1}+B|B|^{\gamma+\delta-1}\right)\\ \leq&\frac{1}{2}\left\||A|^{2\alpha}+|A^*|^{2\beta}+|B|^{2\gamma}+|B^*|^{2\delta}\right\|_{\operatorname{ber}}\\ -&\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^{2\alpha}}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^{2\beta}}(\eta)\right)^{1/2}\right)^2\\ -&\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^{2\gamma}}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^{2\delta}}(\eta)\right)^{1/2}\right)^2.\end{aligned}$$

Remark 3.3. *(i) Setting $\alpha=\beta=\gamma=\delta=\frac{1}{2}$ in (3.24), we get*

$$\begin{aligned}\operatorname{ber}(A+B)\leq&\frac{1}{2}\||A|+|A^*|+|B|+|B^*|\|_{\operatorname{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|}(\eta)\right)^{1/2}\right)^2\\ &-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|}(\eta)\right)^{1/2}\right)^2.\end{aligned}$$

(ii) In particular, take $B=A$, we get

$$\operatorname{ber}(A)\leq\frac{1}{2}\||A|+|A^*|\|_{\operatorname{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|}(\eta)\right)^{1/2}\right)^2.$$

(iii) Setting $\alpha=\beta=\gamma=\delta=1$ in (3.24), we get

$$\begin{aligned}\operatorname{ber}(A|A|+B|B|)\leq&\frac{1}{2}\left\||A|^2+|A^*|^2+|B|^2+|B^*|^2\right\|_{\operatorname{ber}}\\ &-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^2}(\eta)\right)^{1/2}\right)^2\\ &-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|B|^2}(\eta)\right)^{1/2}-\left(\widetilde{|B^*|^2}(\eta)\right)^{1/2}\right)^2.\end{aligned}$$

(iv) In particular, take $B=A$, we get

$$\begin{aligned}\operatorname{ber}(A|A|)\leq&\frac{1}{2}\left\||A|^2+|A^*|^2\right\|_{\operatorname{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^2}(\eta)\right)^{1/2}\right)^2\\ =&\frac{1}{2}\|A^*A+AA^*\|_{\operatorname{ber}}-\frac{1}{2}\inf_{\eta\in Q}\left(\left(\widetilde{|A|^2}(\eta)\right)^{1/2}-\left(\widetilde{|A^*|^2}(\eta)\right)^{1/2}\right)^2.\end{aligned}$$

For more recent research on Berezin radius inequalities for operators and other relevant results, we recommend [4, 6, 13, 14, 15, 16, 18, 29, 33, 34].

References

[1] M. W. Alomari: *On the generalized mixed Schwarz inequality*, Proc. Inst. Math. Mech., **46** (1) (2020), 3–15.

[2] M. W. Alomari: *Refinements of some numerical radius inequalities for Hilbert space operators*, Linear Multilinear Algebra, **69** (7) (2021), 1208–1223.

[3] M. W. Alomari: *Improvements of some numerical radius inequalities*, Azerb. J. Math., **12** (1) (2022), 124–137.

[4] M. Bakherad: *Some Berezin number inequalities for operators matrices*, Czechoslovak Math. J., **68** (143) (2018), 997–1009.
[5] M. Bakherad, M. T. Garayev: *Berezin number inequalities for operators*, Concr. Oper., **6** (1) (2019), 33–43.
[6] M. Bakherad, M. Hajmohamadi, R. Lashkaripour and S. Sahoo: *Some extensions of Berezin number inequalities on operators*, Rocky Mountain J. Math., **51** (6) (2021), 1941–1951.
[7] F. A. Berezin: *Covariant and contravariant symbols for operators*, Math. USSR-Izvestiya, **6** (1972), 1117–1151.
[8] S. S. Dragomir: *Inequalities for the numerical radius of linear operators in Hilbert spaces*, SpringerBriefs in Mathematics (2013).
[9] S. S. Dragomir: *Some inequalities for the norm and the numerical radius of linear operators in Hilbert spaces*, Tamkang J. Math., **39** (2008), 1–7.
[10] S. S. Dragomir: *Some Inequalities generalizing Kato's and Furuta's results*, Filomat, **28** (1) (2014), 179–195.
[11] T. Furuta: *An extension of the Heinz-Kato theorem*, Proc. Amer. Math. Soc., **120** (3) (1994), 785–787.
[12] M. T. Garayev, M. W. Alomari: *Inequalities for the Berezin number of operators and related questions*, Complex Anal. Oper. Theory, **15**, 30 (2021).
[13] M. Garayev, F. Bouzeffour, M. Gürdal and C. M. Yangöz: *Refinements of Kantorovich type, Schwarz and Berezin number inequalities*, Extracta Math., **35** (2020), 1–20.
[14] M. T. Garayev, M. Gürdal and A. Okudan: *Hardy-Hilbert's inequality and a power inequality for Berezin numbers for operators*, Math. Inequal. Appl., **19** (2016), 883–891.
[15] M. T. Garayev, M. Gürdal and S. Saltan: *Hardy type inequaltiy for reproducing kernel Hilbert space operators and related problems*, Positivity, **21** (6) (2017), 1615–1623.
[16] M. T. Garayev, H. Guedri, M. Gürdal and G. M. Alsahli: *On some problems for operators on the reproducing kernel Hilbert space*, Linear Multilinear Algebra, **69** (11) (2021), 2059–2077.
[17] K. E. Gustafson, D. K. M. Rao: *Numerical Range*, Springer-Verlag, New York (1997).
[18] M. Hajmohamadi, R. Lashkaripour and M. Bakherad: *Improvements of Berezin number inequalities*, Linear Multilinear Algebra, **68** (6) (2020), 1218–1229.
[19] P. R. Halmos: *A Hilbert space problem book*, Van Nostrand Company, Inc., Princeton (1967).
[20] M. B. Huban, H. Başaran and M. Gürdal: *New upper bounds related to the Berezin number inequalities*, J. Inequal. Spec. Funct., **12** (3) (2021), 1–12.
[21] M. T. Karaev: *Berezin set and Berezin number of operators and their applications*, The 8th Workshop on Numerical Ranges and Numerical Radii WONRA -06, Bremen (Germany) (2006), p.14.
[22] M. T. Karaev: *Berezin symbol and invertibility of operators on the functional Hilbert spaces*, J. Funct. Anal., **238** (2006), 181–192.
[23] M. T. Karaev: *Reproducing kernels and Berezin symbols techniques in various questions of operator theory*, Complex Anal. Oper. Theory, **7** (2013), 983–1018.
[24] F. Kittaneh: *A numerical radius inequality and an estimate for the numerical radius of the Frobenius companion matrix*, Studia Math., **158** (2003), 11–17.
[25] F. Kittaneh: *Numerical radius inequalities for Hilbert space operators*, Studia Math., **168** (1) (2005), 73–80
[26] F. Kittaneh, Y. Manasrah: *Improved Young and Heinz inequalities for matrices*, J. Math. Anal. Appl., **361** (1) (2010), 262–269.
[27] T. Kato: *Notes on some inequalities for linear operators*, Math. Ann., **125** (1952), 208–212.
[28] W. Reid: *Symmetrizable completely continuous linear transformations in Hilbert space*, Duke Math., **18** (1951), 41–56.
[29] S. Sahoo, M. Bakherad: *Some extended Berezin number inequalities*, Filomat, **35** (6) (2021), 2043–2053.
[30] M. Sattari, M. S. Moslehian and T. Yamazaki: *Some generalized numerical radius inequalities for Hilbert space operators*, Linear Algebra Appl., **470** (2015), 216–227.
[31] A. Sheikhhosseini, M. S. Moslehian and K. Shebrawi: *Inequalities for generalized Euclidean operator radius via Young's inequality*, J. Math. Anal. Appl., **445** (2) (2017), 1516–1529.
[32] R. Tapdigoglu: *New Berezin symbol inequalities for operators on the reproducing kernel Hilbert space*, Oper. Matrices, **15** (3) (2021), 1445–1460.
[33] U. Yamancı, M. Gürdal and M. T. Garayev: *Berezin number inequality for convex function in reproducing kernel Hilbert space*, Filomat, **31** (2017), 5711–5717.
[34] U. Yamancı, R. Tunç and M. Gürdal: *Berezin numbers, Grüss type inequalities and their applications*, Bull. Malays. Math. Sci. Soc., **43** (2020), 2287–2296.
[35] T. Yamazaki: *On upper and lower bounds of the numerical radius and an equality condition*, Studia Math., **178** (2007), 83–89.

MEHMET GÜRDAL
SULEYMAN DEMIREL UNIVERSITY
DEPARTMENT OF MATHEMATICS
32260, ISPARTA, TURKEY
ORCID: 0000-0003-0866-1869
E-mail address: gurdalmehmet@sdu.edu.tr

MOHAMMAD WAJEEH ALOMARI
IRBID NATIONAL UNIVERSITY
DEPARTMENT OF MATHEMATICS
2600 IRBID 21110, JORDAN
ORCID: 0000-0002-6696-9119
E-mail address: mwomath@gmail.com

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 3, pp. 154-167
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Rational generalized Stieltjes functions

IVAN KOVALYOV*

ABSTRACT. The rational meromorphic functions on $\mathbb{C}\backslash\mathbb{R}$ are studied. We consider the some classes of one, as the generalized Nevanlinna $\mathbf{N}_\kappa$ and generalized Stieltjes $\mathbf{N}_\kappa^k$ classes. By Euclidean algorithm, we can find indices κ and k, i.e. determine which class the function belongs to $\mathbf{N}_\kappa^k$.

Keywords: Rational function, generalized Nevanlinna function, generalized Stieltjes function.

2010 Mathematics Subject Classification: 30A05, 46C20, 47A57, 47A56.

1. INTRODUCTION

Recall a generalized Nevalinna class $\mathbf{N}_\kappa$ and a generalized Stieltjes class $\mathbf{N}_\kappa^k$.

Definition 1.1. *A function f meromorphic on $\mathbb{C}\backslash\mathbb{R}$ with the set of holomorphy $\mathfrak{h}_f$ is said to be in the generalized Nevanlinna class $\mathbf{N}_\kappa$ ($\kappa \in \mathbb{N}$), if for every set $z_i \in \mathbb{C}_+ \cap \mathfrak{h}_f$ ($j = 1, \dots, n$) the form*

$$\sum_{i,j=1}^{n} \frac{f(z_i) - \overline{f(z_j)}}{z_i - \overline{z}_j} \xi_i \overline{\xi}_j$$

has at most κ and for some choice of z_i ($i = 1, \dots, n$) it has exactly κ negative squares. For $f \in \mathbf{N}_\kappa$, let us write $\kappa_-(f) = \kappa$. In particular, if $\kappa = 0$ then the class $\mathbf{N}_0$ coincides with the class $\mathbf{N}$ of Nevanlinna functions. A function $f \in \mathbf{N}_\kappa$ is said to belong to the class $\mathbf{N}_\kappa^+$ (see [8, 9]) if $zf \in \mathbf{N}_k$ and to the class $\mathbf{N}_\kappa^k$ ($k \in \mathbb{N}$) if $zf \in \mathbf{N}_\kappa^k$ (see [3], [4]). In particular, if $k = 0$, then $\mathbf{N}_\kappa^0 := \mathbf{N}_\kappa^+$. The function $f \in \mathbf{N}_\kappa^{-k}$, if $f \in \mathbf{N}_\kappa$ and $\frac{1}{z} f \in \mathbf{N}_k$ (see [5]).

Recall some properties of the generalized Nevanlinna functions and generalized Stieltjes functions.

Proposition 1.1. *([8]) Let $f \in \mathbf{N}_\kappa$, $f_1 \in \mathbf{N}_{\kappa_1}$, $f_2 \in \mathbf{N}_{\kappa_2}$. Then*

(1) $-f^{-1} \in \mathbf{N}_\kappa$.
(2) $f_1 + f_2 \in \mathbf{N}_{\kappa'}$, where $\kappa' \le \kappa_1 + \kappa_2$.
(3) If, in addition, $f_1(iy) = o(y)$ as $y \to \infty$ and f_2 is a polynomial, then

$$f_1 + f_2 \in \mathbf{N}_{\kappa_1 + \kappa_2}. \tag{1.1}$$

(4) Every real polynomial $P(t) = p_\nu t^\nu + p_{\nu-1} t^{\nu-1} + \ldots + p_1 t + p_0$ of degree ν belongs to a class $\mathbf{N}_\kappa$, where the index $\kappa = \kappa_-(P)$ can be evaluated by (see [8, Lemma 3.5])

$$\kappa_-(P) = \begin{cases} \left[\frac{\nu+1}{2}\right], & \text{if } p_\nu < 0; \text{ and } \nu \text{ is odd}; \\ \left[\frac{\nu}{2}\right], & \text{otherwise}. \end{cases} \tag{1.2}$$

Received: 13.05.2022; Accepted: 03.08.2022; Published Online: 15.08.2022
*Corresponding author: Ivan Kovalyov; i.m.kovalyov@gmail.com
DOI: 10.33205/cma.1116322

Proposition 1.2. *([2]) Let $f \in \mathbf{N}_\kappa^k$. Then the following equivalences hold:*

(1) $f \in \mathbf{N}_\kappa^k \iff -\frac{1}{f} \in \mathbf{N}_\kappa^{-k}$;

(2) $f \in \mathbf{N}_\kappa^k \iff zf(z) \in \mathbf{N}_k^{-\kappa}$.

Lemma 1.1 ([7, Lemma 3.2])**.** *Let $P(z)$ be a polynomial of the degree ν and let $\alpha \in \mathbb{R}$. Then:*

(1) if $zP(z) \in \mathbf{N}_\kappa$, then

$$(z-\alpha)P(z) \in \mathbf{N}_\kappa; \tag{1.3}$$

(2) if $P(z) \in \mathbf{N}_\kappa$, then

$$\frac{(z-\alpha)}{z}P(z) \in \mathbf{N}_{\kappa'}, \quad \text{where} \quad \kappa' = \kappa + \kappa_-\left(-\frac{\alpha P(0)}{z}\right); \tag{1.4}$$

(3) if $((z-\alpha)P(z) - g(z)) \in \mathbf{N}_{\kappa'}^k$, then

$$(-\alpha P(0) - g(z)) \in \mathbf{N}_{\kappa-\kappa_1}^{(k-k_1)} \quad \text{and} \quad (\alpha P(0) + g(z))^{-1} \in \mathbf{N}_{\kappa-\kappa_1}^{-(k-k_1)}, \tag{1.5}$$

where $\kappa_1 = \kappa_-(zP(z))$ and $k_1 = \kappa_-(P(z))$.

The indefinite Hamburger moment in the generalized Nevanlinna class $\mathbf{N}_\kappa$ was studied in [10]. The indefinite Stieltjes moment problem in the generalized Stieltjes class $\mathbf{N}_\kappa^k$ was studied in [11], [1], [2], [6] and [7]. One is based on the Schur algorithm, i.e. the description of the solutions are found in terms of the continued fractions. In the present paper, the rational generalized Stieltjes functions are investigated. The goal is to determine class $\mathbf{N}_\kappa^k$, such that the some rational generalized Stieltjes function f belongs to one (i.e. find the indices κ and k).

2. Finding the index

2.1. **Euclidean algorithm.** Let us recall an Euclidean algorithm. Let P_0 and Q_0 be the polynomials, such that $\deg(P_0) = n_0$ and $\deg(Q_0) = m_0$, where $n_0, m_0 \in \mathbb{Z}_+$ and let $m_0 \le n_0$. By Euclidean algorithm, we obtain

$$\begin{aligned}
P_0(z) &= Q_0(z)a_0(z) + r_1(z),\\
Q_0(z) &= r_1(z)a_1(z) + r_2(z),\\
r_1(z) &= r_2(z)a_2(z) + r_3(z),\\
&\vdots\\
r_{n-2}(z) &= r_{n-1}(z)a_{n-1}(z) + r_n(z),\\
r_{n-1}(z) &= r_n(z)a_n(z),
\end{aligned} \tag{2.6}$$

where r_i are polynomials. Consequently, the ratio $\dfrac{P_0(z)}{Q_0(z)}$ can be represented as a continued fraction

$$\frac{P_0(z)}{Q_0(z)} = a_0(z) + \cfrac{1}{a_1(z) + \cfrac{1}{a_2(z) + \cdots + \cfrac{1}{a_n(z)}}}. \tag{2.7}$$

2.2. **Rational generalized Nevanlinna function and its index κ.**

Theorem 2.1. *Let P_0 and Q_0 be the polynomials, such that $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 < n_0$. Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$. Then f belongs to the class $\mathbf{N}_\kappa$ and the index κ is calculated by*

$$\kappa = \sum_{j=0}^{n} \kappa_-((-1)^{j+1}a_j(z)). \tag{2.8}$$

Proof. Assume, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ is meromorphic function on $\mathbb{C}\backslash\mathbb{R}$, where the P_0 and Q_0 are the polynomials of the power $\deg(P_0) = n_0$ and $\deg(Q_0) = m_0$, respectively. By Definition 1.1, $f \in \mathbf{N}_\kappa$.

Calculating index κ. Due to (2.7), we can rewrite f as follows

$$f(z) = \frac{1}{\frac{P_0(z)}{Q_0(z)}} = -\cfrac{1}{-a_0(z) - \cfrac{1}{a_1(z) - \cfrac{1}{-a_2(z) - \cdots - \cfrac{1}{(-1)^{n+1}a_n(z)}}}}. \tag{2.9}$$

By Proposition 1.1 (see (1.2))

$$\kappa_j = \kappa_-((-1)^{j+1}a_j(z)), \quad j = \overline{0,n},$$

i.e. $(-1)^{j+1}a_j(z) \in \mathbf{N}_{\kappa_j}$. Moreover, by Proposition 1.1 (see items (1) and (3)), we obtain

$$\begin{aligned} &-\frac{1}{(-1)^{j+1}a_j(z)} \in \mathbf{N}_{\kappa_j} \quad \text{for all } j = \overline{0,n}, \\ &(-1)^n a_{n-1}(z) - \frac{1}{(-1)^{n+1}a_n(z)} \in \mathbf{N}_{\kappa_n+\kappa_{n-1}}. \end{aligned} \tag{2.10}$$

Let us construct a recursive sequence as

$$\begin{aligned} f_n(z) :=& (-1)^n a_{n-1}(z) - \frac{1}{(-1)^{n+1}a_n(z)}, \\ f_{n-1}(z) :=& (-1)^{n-1}a_{n-2}(z) - \frac{1}{f_n(z)}, \\ &\vdots \\ f_{n-2}(z) :=& (-1)^{n-2}a_{n-3}(z) - \frac{1}{f_{n-1}(z)}, \\ f_1(z) :=& -a_0(z) - \frac{1}{f_2(z)}. \end{aligned} \tag{2.11}$$

Hence (see Proposition 1.1)

$$f_n \in \mathbf{N}_{\kappa_n+\kappa_{n-1}},\ f_{n-1} \in \mathbf{N}_{\kappa_n+\kappa_{n-1}+\kappa_{n-2}},\ \dots,\ f_1 \in \mathbf{N}_{\kappa_n+\kappa_{n-1}+\ldots+\kappa_0}. \tag{2.12}$$

By the recursive sequence, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ can be rewritten as

$$f(z) = -\frac{1}{f_1(z)}. \tag{2.13}$$

Therefore $f \in \mathbf{N}_\kappa$, where the index $\kappa = \sum_{j=0}^{n} \kappa_-((-1)^{j+1}a_j(z))$. This completes the proof. □

Corollary 2.1. *Let P_0 and Q_0 be the polynomials, such that $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 \le n_0$. Let $f(z) = \dfrac{P_0(z)}{Q_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$. Then f belongs to the class $\mathbf{N}_\kappa$ and the index κ is calculated by*

$$\kappa = \sum_{j=0}^{n} \kappa_-((-1)^j a_j(z)). \tag{2.14}$$

Proof. Let the rational function $f(z) = \dfrac{P_0(z)}{Q_0(z)}$ is meromorphic function on $\mathbb{C}\backslash\mathbb{R}$, where the numerator P_0 and denominator Q_0 are the polynomials of the power $\deg(P_0) = n_0$ and $\deg(Q_0) = m_0$, respectively. Hence, f belongs to the generalized Nevanlinna class $\mathbf{N}_\kappa$ (see Definition 1.1).

Let us find the index κ. By the representation (2.7), we obtain

$$f(z) = \frac{P_0(z)}{Q_0(z)} = a_0(z) - \cfrac{1}{-a_1(z) - \cfrac{1}{a_2(z) - \cdots - \cfrac{1}{(-1)^n a_n(z)}}}. \tag{2.15}$$

By Theorem 2.1 (see (2.10)-(2.13)), $f \in \mathbf{N}_\kappa$ and the index κ is calculated by (2.14). This completes the proof. □

3. Rational generalized Stieltjes function and its indices κ, k

First of all, we study the simple case of the rational functions, which belong to the generalized Stieltjes classes $\mathbf{N}_\kappa^{\pm k}$ and find the formulas for the indices κ and k.

3.1. Rational function of the generalized Stieltjes class $\mathbf{N}_\kappa^k$.

Theorem 3.2. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 < n_0$. Let f admit the representation (2.9) and let $a_{2i}(z)$ vanish at zero for all $i = \overline{0, [n/2]}$ (i.e. $a_{2i}(0) = 0$). Then f belongs to the class $\mathbf{N}_\kappa^k$, where the index κ is calculated by (2.8) and index k is found by*

$$k = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_-\left(-\frac{a_{2j}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_-(z a_{2j+1}(z)), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_-\left(-\frac{a_{2j}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]} \kappa_-(z a_{2j+1}(z)), & \text{if } n \text{ is odd.} \end{cases} \tag{3.16}$$

Proof. By Definition 1.1, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ meromorphic on $\mathbb{C}\backslash\mathbb{R}$ belongs to the generalized Stiektjes class $\mathbf{N}_\kappa^k$ (i.e. $f \in \mathbf{N}_\kappa$ and $zf \in \mathbf{N}_k$) and by Theorem 2.1, the index κ is calculated by (2.8).

Let us find an index k. Assume f admits the representation (2.9) and $a_{2i}(0) = 0$ for all $i = \overline{0, [n/2]}$. Hence, we get the two cases, where n is even or odd.

First of all we consider the even case (i.e. $n = 2m$, $m \in \mathbb{Z}_+$), we obtain

$$zf(z) = \frac{Q_0(z)}{P_0(z)} = \frac{1}{\frac{P_0(z)}{zQ_0(z)}} = -\frac{1|}{|-\frac{a_0(z)}{z}} - \frac{1|}{|za_1(z)} - \cdots - \frac{1|}{|za_{2m-1}(z)} - \frac{1|}{|-\frac{a_{2m}(z)}{z}}. \tag{3.17}$$

The terms $-\dfrac{a_{2i}(z)}{z}$ are polynomials, i.e. $a_{2i}(0) = 0$ for all $i = \overline{0, [n/2]}$. By Theorem 2.1 , we get

$$k = \sum_{j=0}^{[n/2]} \kappa_- \left(-\frac{a_{2j}(z)}{z}\right) + \sum_{j=0}^{[n/2]-1} \kappa_-(za_{2j+1}(z)).$$

The next step, let n is odd (i.e. $n = 2m+1$, $m \in \mathbb{Z}_+$). Consequently

$$zf(z) = -\frac{1|}{|-\frac{a_0(z)}{z}} - \frac{1|}{|za_1(z)} - \cdots - \frac{1|}{|-\frac{a_{2m}(z)}{z}} - \frac{1|}{|za_{2m+1}(z)}. \tag{3.18}$$

Similarly, $-\dfrac{a_{2i}(z)}{z}$ are the polynomials and the index k is

$$\sum_{j=0}^{[n/2]} \kappa_- \left(-\frac{a_{2j}(z)}{z}\right) + \sum_{j=0}^{[n/2]} \kappa_-(za_{2j+1}(z)).$$

This completes the proof. □

Corollary 3.2. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $n_0 \le m_0$. Then f belongs to the class $\mathbf{N}_\kappa^k$ and admits the representation (2.15).*

Moreover, the index κ is calculated by (2.14). In addition, if the all polynomials $a_{2i+1}(z)$ vanish at zero in the representation (2.15), then the index k is found by

$$k = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(za_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_-(-\frac{a_{2j+1}(z)}{z}), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(za_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]} \kappa_-(-\frac{a_{2j+1}(z)}{z}), & \text{if } n \text{ is odd.} \end{cases} \tag{3.19}$$

Proof. By Definition 1.1, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ belongs to the generalized Stieltjes class $\mathbf{N}_\kappa^k$ and by Corollary 2.1, f admits the representation (2.15) and the index κ is calculated by (2.14). By Theorem 3.2, the index k can be found by (3.19). This completes the proof. □

Corollary 3.3. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 + 1 < n_0$. Then the rational function $zf(z)$ admits the following representation*

$$zf(z) = -\frac{1|}{|-\widetilde{a}_0(z)} - \frac{1|}{|\widetilde{a}_1(z)} - \cdots - \frac{1|}{|(-1)^{n+1}\widetilde{a}_n(z)} \tag{3.20}$$

and f belongs to the class $\mathbf{N}_\kappa^k$.

Furthermore, in addition, if $\widetilde{a}_{2i+1}$ vanish at zero for all $i = \overline{0,[n/2]}$, then the indices κ and k can be found by

$$k = \sum_{j=0}^{n} \kappa_-((-1)^{j+1}\widetilde{a}_j(z)), \tag{3.21}$$

$$\kappa = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-z\widetilde{a}_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_- \left(\frac{\widetilde{a}_{2j+1}(z)}{z}\right), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-z\widetilde{a}_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]} \kappa_- \left(\frac{\widetilde{a}_{2j+1}(z)}{z}\right), & \text{if } n \text{ is odd.} \end{cases} \tag{3.22}$$

Proof. By Euclidean algorithm, the rational function $zf(z) = \frac{zQ_0(z)}{P_0(z)}$ admits the representation (3.20). By Theorem 3.2, the rational function f belongs to the generalized Stieltjes class $\mathbf{N}_\kappa^k$, where the indices k and κ are found by (3.21) and (3.22), respectively. This completes the proof. □

Corollary 3.4. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $n_0 \le m_0 + 1$. Then the rational function $zf(z)$ admits the following representation*

$$zf(z) = \hat{a}_0(z) - \frac{1|}{|-\hat{a}_1(z)} - \frac{1|}{|\hat{a}_2(z)} - \cdots - \frac{1|}{|(-1)^n\hat{a}_n(z)} \tag{3.23}$$

and f belongs to the class $\mathbf{N}_\kappa^k$.

Furthermore, if $\hat{a}_{2i}$ vanish at zero for all $i = \overline{1,[n/2]}$, then the indices κ and k can be found by

$$k = \sum_{j=0}^{n} \kappa_-((-1)^j\hat{a}_j(z)), \tag{3.24}$$

$$\kappa = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-z\hat{a}_{2j+1}(z)\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_- \left(\frac{\hat{a}_{2j}(z)}{z}\right), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-z\hat{a}_{2j+1}(z)\right) + \sum\limits_{j=0}^{[n/2]} \kappa_- \left(\frac{\hat{a}_{2j}(z)}{z}\right), & \text{if } n \text{ is odd.} \end{cases} \tag{3.25}$$

3.2. Rational function of the generalized Stieltjes class $\mathbf{N}_\kappa^{-k}$.

Theorem 3.3. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 \le n_0$. Let f admits the representation (2.9) and let the all odd polynomials $a_{2i+1}(z)$ vanish at zero (i.e. $a_{2i+1}(0) = 0$). Then f belongs to the class $\mathbf{N}_\kappa^{-k}$, where the index κ is calculated by (2.8) and index k is found by*

$$k = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-za_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_- \left(\frac{a_{2j+1}(z)}{z}\right), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-za_{2j}(z)\right) + \sum\limits_{j=0}^{[n/2]} \kappa_- \left(\frac{a_{2j+1}(z)}{z}\right), & \text{if } n \text{ is odd.} \end{cases} \tag{3.26}$$

Proof. By Definition 1.1, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ meromorphic on $\mathbb{C}\backslash\mathbb{R}$ belongs to the generalized Stieltjes class $\mathbf{N}_\kappa^{-k}$ (i.e. $f \in \mathbf{N}_\kappa$ and $\frac{f}{z} \in \mathbf{N}_k$) and by Theorem 2.1, the index κ is calculated by (2.8).

Suppose f admits representation (2.9) and the all odd polynomials $a_{2i}(0)$ vanish at zero (i.e. $a_{2i+1}(0)=0$).

If n is even (i.e. $n=2m$, $m\in\mathbb{Z}_+$), then

$$\begin{aligned}\frac{f(z)}{z}&=\frac{Q_0(z)}{zP_0(z)}\\&=\frac{1}{\frac{zP_0(z)}{Q_0(z)}}\\&=-\frac{1|}{\Big|-za_0(z)}-\frac{1|}{\Big|\frac{a_1(z)}{z}}-\frac{1|}{\Big|-za_2(z)}-\cdots-\frac{1|}{\Big|\frac{a_{2m-1}(z)}{z}}-\frac{1|}{\Big|-za_{2m}(z)}.\end{aligned}\tag{3.27}$$

Due to the all odd polynomials $a_{2i+1}(0)=0$, $\frac{a_{2j+1}}{z}$ are polynomials and by Theorem 2.1, we obtain

$$k=\sum_{j=0}^{[n/2]}\kappa_-\left(-za_{2j}(z)\right)+\sum_{j=0}^{[n/2]-1}\kappa_-\left(\frac{a_{2j+1}(z)}{z}\right).$$

If n is odd (i.e. $n=2m+1$, $m\in\mathbb{Z}_+$), then

$$\frac{f(z)}{z}=-\frac{1|}{\Big|-za_0(z)}-\frac{1|}{\Big|\frac{a_1(z)}{z}}-\cdots-\frac{1|}{\Big|\frac{a_{2m-1}(z)}{z}}-\frac{1|}{\Big|-za_{2m}(z)}-\frac{1|}{\Big|\frac{a_{2m+1}(z)}{z}}.$$

Obviously, $a_{2i+1}(0)=0$, $\frac{a_{2j+1}}{z}$ are polynomials and by Theorem 2.1, we find index k as follow

$$k=\sum_{j=0}^{[n/2]}\kappa_-\left(-za_{2j}(z)\right)+\sum_{j=0}^{[n/2]}\kappa_-\left(\frac{a_{2j+1}(z)}{z}\right).$$

This completes the proof. □

Corollary 3.5. *Let the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $n_0\le m_0$. Then f belongs to the class $\mathbf{N}_\kappa^{-k}$ and admits the representation (2.15).*

Moreover, the index κ is calculated by (2.14). In addition, if the all polynomials $a_{2i}(z)$ vanish at zero in the representation (2.15), then the index k is culculated by

$$k=\begin{cases}\sum\limits_{j=0}^{[n/2]-1}\kappa_-\left(-za_{2j+1}(z)\right)+\sum\limits_{j=0}^{[n/2]}\kappa_-\left(\frac{a_{2j}(z)}{z}\right), & \text{if } n \text{ is even};\\ \sum\limits_{j=0}^{[n/2]}\kappa_-\left(-za_{2j+1}(z)\right)+\sum\limits_{j=0}^{[n/2]}\kappa_-\left(\frac{a_{2j}(z)}{z}\right), & \text{if } n \text{ is odd}.\end{cases}\tag{3.28}$$

Proof. By Definition 1.1, the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ belongs to $\mathbf{N}_\kappa^{-k}$ and by Corollary 2.1, f admits representation (2.15) and the index κ can be calculated by (2.14). By Theorem 3.3, the index k can be found by (3.19). This completes the proof. □

Corollary 3.6. *Let the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ be meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $m_0<n_0+1$. Then the rational function $zf(z)$ admits the following representation*

$$\frac{f(z)}{z}=\frac{Q_0(z)}{zP_0(z)}=-\frac{1|}{|-\widetilde{a}_0(z)}-\frac{1|}{|\widetilde{a}_1(z)}-\cdots-\frac{1|}{|(-1)^{n+1}\widetilde{a}_n(z)}\tag{3.29}$$

and f belongs to the class $\mathbf{N}_\kappa^{-k}$.

Furthermore, if $\widetilde{a}_{2i}$ vanish at zero for all $i = \overline{0,[n/2]}$, then the indices κ and k can be found by

$$k = \sum_{j=0}^{n} \kappa_-((-1)^{j+1}\widetilde{a}_j(z)), \tag{3.30}$$

$$\kappa = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-\frac{\widetilde{a}_{2j}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_-(z\widetilde{a}_{2j+1}(z)), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-\frac{\widetilde{a}_{2j}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]} \kappa_-(z\widetilde{a}_{2j+1}(z)), & \text{if } n \text{ is odd.} \end{cases} \tag{3.31}$$

Proof. By Euclidean algorithm, the rational function $\frac{f(z)}{z} = \frac{Q_0(z)}{zP_0(z)}$ admits the representation (3.29). By Theorem 3.3, the rational function f belongs to the generalized Stieltjes class $\mathbf{N}_\kappa^{-k}$, the indices k and κ are found by (3.30) and (3.31), respectively. This completes the proof. □

Corollary 3.7. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $n_0 + 1 \le m_0$. Then the rational function $zf(z)$ admits the following representation*

$$\frac{f(z)}{z} = \hat{a}_0(z) - \frac{1|}{|-\hat{a}_1(z)} - \frac{1|}{|\hat{a}_2(z)} - \cdots - \frac{1|}{|(-1)^n\hat{a}_n(z)} \tag{3.32}$$

and f belongs to the class $\mathbf{N}_\kappa^{-k}$.

Furthermore, if $\hat{a}_{2i+1}$ vanish at zero, then the indices κ and k can be found by

$$k = \sum_{j=0}^{n} \kappa_-((-1)^j\hat{a}_j(z)), \tag{3.33}$$

$$\kappa = \begin{cases} \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-\frac{\hat{a}_{2j+1}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]-1} \kappa_-(z\hat{a}_{2j}(z)), & \text{if } n \text{ is even;} \\ \sum\limits_{j=0}^{[n/2]} \kappa_- \left(-\frac{\hat{a}_{2j+1}(z)}{z}\right) + \sum\limits_{j=0}^{[n/2]} \kappa_-(z\hat{a}_{2j}(z)), & \text{if } n \text{ is odd.} \end{cases} \tag{3.34}$$

4. General cases

4.1. General case in the class $\mathbf{N}_\kappa^k$.

Proposition 4.3. *Let the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ be the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $m_0 < n_0$, let f admits representation (2.9). Then f belongs to the class $\mathbf{N}_\kappa^k$, such that*

$$\kappa = \sum_{j=0}^{n} \kappa_j \text{ and } k \le \sum_{i=0}^{n} k_i + \sum_{i=0}^{[n/2]} k_i^0, \tag{4.35}$$

where the indices κ_i, k_i and k_i^0 can be found by

$$\begin{aligned} &\kappa_i = \kappa_-((-1)^{i+1}a_i(z)), \quad k_{2i} = \kappa_-\left(-\frac{a_{2i}(z)-a_{2i}(0)}{z}\right), \\ &k_{2i+1} = \kappa_-(za_{2i+1}(z)), \quad k_i^0 = \begin{cases} 1, & \text{if } a_{2i}(0) < 0; \\ 0, & \text{if } a_{2i}(0) > 0. \end{cases} \end{aligned} \tag{4.36}$$

Proof. (i) The first case. Let $n = 2m+1$, $m \in \mathbb{Z}_+$, then the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ can be rewritten by formula (2.9) as follows

$$f(z) = -\cfrac{1}{\Big|-a_0(z)-\cfrac{1}{a_1(z)}} - \cfrac{1}{\Big|-a_2(z)-\cfrac{1}{a_3(z)}} - \cdots - \cfrac{1}{\Big|-a_{2m}(z)-\cfrac{1}{a_{2m+1}(z)}}. \tag{4.37}$$

Setting

$$f_m(z) := -\cfrac{1}{-a_{2m}(z)-\cfrac{1}{a_{2m+1}(z)}},$$

then zf_m takes the following form

$$\begin{aligned} zf_m(z) &= -\cfrac{z}{-a_{2m}(z)-\cfrac{1}{a_{2m+1}(z)}} \\ &= -\cfrac{1}{-\cfrac{a_{2m}(z)-a_{2m}(0)}{z} - \cfrac{a_{2m}(0)}{z} - \cfrac{1}{za_{2m+1}(z)}}. \end{aligned}$$

By Proposition 1.1 and Proposition 1.2, $f_m \in \mathbf{N}_{\widetilde{\kappa}_m}^{\widetilde{k}_m}$, where

$$\widetilde{\kappa}_m = \kappa_-(-a_{2m}) + \kappa_-(a_{2m+1}) \text{ and } \widetilde{k}_m \le k_{2m} + k_{2m+1} + k_m^0,$$

where

$$\begin{aligned} &k_{2m} := \kappa_-\left(-\frac{a_{2m}(z)-a_{2m}(0)}{z}\right), \ k_{2m+1} = \kappa_-(za_{2m+1}), \\ &k_m^0 := \kappa_-\left(-\frac{a_{2m}(0)}{z}\right) = \begin{cases} 1, & \text{if } a_{2m}(0) < 0; \\ 0, & \text{if } a_{2m}(0) > 0. \end{cases} \end{aligned} \tag{4.38}$$

The next step. Let us define the function f_{m-1} by

$$f_{m-1}(z) = -\cfrac{1}{-a_{2m-2}(z)-\cfrac{1}{a_{2m-1}(z)+f_m(z)}}.$$

Consequently, zf_{m-1} takes the following form

$$zf_{m-1}(z) = -\cfrac{1}{-\cfrac{a_{2m-2}(z)-a_{2m-2}(0)}{z} - \cfrac{a_{2m-2}(0)}{z} - \cfrac{1}{za_{2m-11}(z)+zf_m(z)}}.$$

Hence $f_{m-1} \in \mathbf{N}_{\widetilde{\kappa}_{m-1}}^{\widetilde{k}_{m-1}}$ (see Propositions 1.1 and 1.2), where the indices $\widetilde{\kappa}_{m-1}$ and $\widetilde{k}_{m-1}$ are

$$\widetilde{\kappa}_{m-1} = \widetilde{\kappa}_m + \kappa_-(-a_{2m-2}) + \kappa_-(a_{2m-1}) \text{ and } \widetilde{k}_{m-1} \le k_{2m-2} + k_{2m-1} + k_{m-1}^0 + \widetilde{k}_m,$$

where

$$k_{2m} := \kappa_-\left(-\frac{a_{2m}(z)-a_{2m}(0)}{z}\right),\ k_{2m+1}=\kappa_-(za_{2m+1}),$$
$$k_m^0 := \kappa_-\left(-\frac{a_{2m}(0)}{z}\right)=\begin{cases}1, & \text{if } a_{2m}(0)<0;\\ 0, & \text{if } a_{2m}(0)>0.\end{cases} \tag{4.39}$$

Step-by-step, we obtain that $f\in\mathbf{N}_\kappa^k$ and (4.35)–(4.36) hold.

(ii) The second case. Let $n=2m+2$, $m\in\mathbb{Z}_+\cup\{-1\}$, then the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ can be rewritten by

$$f(z)=-\cfrac{1}{\left|-a_0(z)-\cfrac{1}{a_1(z)}\right.}-\cdots-\cfrac{1}{\left|-a_{2m}(z)-\cfrac{1}{a_{2m+1}(z)}\right.}-\cfrac{1}{\left|-a_{2m+2}(z)\right.}. \tag{4.40}$$

Let us set the function f_{m+1} by

$$f_{m+1}(z)=-\frac{1}{-a_{2m+2}(z)}.$$

Hence, the function zf_{m+1} takes the form

$$zf_{m+1}(z)=-\frac{z}{-a_{2m+2(z)}}=-\frac{1}{-\dfrac{a_{2m+2}(z)-a_{2m+2}(0)}{z}-\dfrac{a_{2m+2}(0)}{z}}.$$

By Proposition 1.1 and Proposition 1.2, $f_{m+1}\in\mathbf{N}_{\widetilde{\kappa}_{m+1}}^{\widetilde{k}_{m+1}}$, where the indices $\widetilde{\kappa}_{m+1}$ and $\widetilde{k}_{m+1}$ are defined by

$$\widetilde{\kappa}_{m+1}=\kappa_-(-a_{2m+2}),$$
$$\widetilde{k}_{m+1}\le\kappa_-\left(-\frac{a_{2m+2}(z)-a_{2m+2}(0)}{z}\right)+\kappa_-\left(-\frac{a_{2m+2}(0)}{z}\right).$$

By the first case (i), we obtain $f\in\mathbf{N}_\kappa^k$, where the indices κ and k satisfy the formulas (4.35)–(4.36). This completes the proof. □

Corollary 4.8. *Let the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ be the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $n_0\le m_0$. Then f admits representation*

$$f(z)=a_{-1}(z)-\cfrac{1|}{|-a_0(z)}-\cfrac{1|}{|a_1(z)}-\cdots-\cfrac{1|}{|(-1)^{n+1}a_n(z)}. \tag{4.41}$$

Furthermore, f belongs to the class $\mathbf{N}_\kappa^k$, such that

$$\kappa=\sum_{j=-1}^{n}\kappa_j \text{ and } k\le\sum_{i=-1}^{n}k_i+\sum_{i=-1}^{[n/2]}k_i^0, \tag{4.42}$$

where the indices κ_i, k_i and k_i^0 can be found by

$$k_{2i+1}=\kappa_-(za_{2i+1}(z)),\quad \kappa_i=\kappa_-((-1)^{i+1}a_i(z)),$$
$$k_i^0=\begin{cases}1, & \text{if } a_{2i}(0)<0;\\ 0, & \text{if } a_{2i}(0)>0.\end{cases}\qquad k_{2i}=\kappa_-\left(-\frac{a_{2i}(z)-a_{2i}(0)}{z}\right). \tag{4.43}$$

Proof. Assume the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ be the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $n_0\le m_0$. By Euclidean algorithm, the function f admits representation (4.41).

By Proposition 1.1, $a_{-1}\in\mathbf{N}_{\kappa_{-1}}^{k_{-1}}$, where indices κ_{-1} and k_{-1} are defined by (4.43).

By Proposition 4.3, $(f-a_{-1})\in\mathbf{N}_{\widetilde{\kappa}}^{\widetilde{k}}$, where the indices $\widetilde{\kappa}$ and $\widetilde{k}$ are defined by formulas (4.35)–(4.36). Therefore, the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ belongs to the class $\mathbf{N}_\kappa^k$ and the formulas (4.42)–(4.43) hold. This completes the proof. □

Theorem 4.4. *Let $\tau\in\mathbf{N}_{\kappa^*}^{k^*}$ and let $f(z)=\dfrac{Q_0(z)}{P_0(z)}+\tau(z)$, where the P_0 and Q_0 are polynomials, such that $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $m_0<n_0$. Then $f\in\mathbf{N}_\kappa^k$, where*

$$\kappa\le\kappa^*+\sum_{j=0}^{n}\kappa_j \text{ and } k\le k^*+\sum_{i=0}^{n}k_i+\sum_{i=0}^{[n/2]}k_i^0, \tag{4.44}$$

where the indices κ_i, k_i and k_i^0 can be found by (4.43).

Proof. This proof is based on Proposition 4.3 and Proposition 1.1. □

4.2. **General case in the class $\mathbf{N}_\kappa^{-k}$.**

Proposition 4.4. *Let the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ be the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0)=n_0$, $\deg(Q_0)=m_0$ and $m_0<n_0$ and let f admits representation (2.9). Then f belongs to the class $\mathbf{N}_\kappa^{-k}$, such that*

$$\kappa=\sum_{j=0}^{n}\kappa_j \text{ and } k\le\sum_{i=0}^{n}k_i+\sum_{i=0}^{[n/2]}k_i^0, \tag{4.45}$$

where the indices κ_i, k_i and k_i^0 can be found by

$$\begin{aligned}&\kappa_i=\kappa_-((-1)^{i+1}a_i(z)),\quad k_{2i+1}=\kappa_-\left(\frac{a_{2i+1}(z)-a_{2i+1}(0)}{z}\right),\\ &k_{2i}=\kappa_-(-za_{2i}(z)),\quad k_i^0=\begin{cases}1, & \text{if } a_{2i+1}(0)>0;\\ 0, & \text{if } a_{2i+1}(0)<0.\end{cases}\end{aligned} \tag{4.46}$$

Proof. By Euclidean algorithm, the rational function $f(z)=\dfrac{Q_0(z)}{P_0(z)}$ admits the representation (2.9) and by Theorem (2.1), $f\in\mathbf{N}_\kappa$, where the index κ are calculated by

$$\kappa=\sum_{j=0}^{n}\kappa_j=\sum_{j=0}^{n}\kappa_-((-1)^{j+1}a_j(z)).$$

By Defenition (1.1), the function f is the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, then $f\in\mathbf{N}_\kappa^{-k}$. Find index k.

(i) The first case. Let $n=2m+1$ in (2.9), then

$$\frac{f(z)}{z}=-\cfrac{1}{-za_0(z)-\cfrac{1}{\cfrac{a_1(z)}{z}}}-\cdots-\cfrac{1}{-za_{2m}(z)-\cfrac{1}{\cfrac{a_{2m+1}(z)}{z}}}.$$

Setting

$$\phi_m(z) = -\cfrac{1}{-za_{2m}(z) - \cfrac{1}{\cfrac{a_{2m+1}(z)}{z}}}$$
$$= -\cfrac{1}{-za_{2m}(z) - \cfrac{1}{\cfrac{a_{2m+1}(z) - a_{2m+1}(0)}{z} + \cfrac{a_{2m+1}(z)}{z}}},$$

by Proposition 1.1, $\phi_m \in \mathbf{N}_{\widetilde{k}_m}$, where the index $\widetilde{k}_m$ is defined by

$$\widetilde{k}_m \leq k_{2m} + k_{2m+1} + k_m^0,$$

where the indices k_{2m}, k_{2m+1} and k_m^0 can be calculated by

$$k_{2m} = \kappa_-(-za_{2m}(z)), \quad k_{2m+1} = \kappa_- \left(\frac{a_{2m+1}(z) - a_{2m+1}(0)}{z} \right),$$
$$k_m^0 = \begin{cases} 1, \text{ if } a_{2m+1}(0) > 0; \\ 0, \text{ if } a_{2m+1}(0) < 0. \end{cases}$$

So, let ϕ_{m-1} is defined by

$$\phi_{m-1}(z) = -\cfrac{1}{-za_{2m-2}(z) - \cfrac{1}{\cfrac{a_{2m-11}(z)}{z} + \phi_m(z)}}$$
$$= -\cfrac{1}{-za_{2m-2}(z) - \cfrac{1}{\cfrac{a_{2m-1}(z) - a_{2m-1}(0)}{z} + \cfrac{a_{2m-1}(z)}{z} + \phi_m(z)}}.$$

Due to Proposition 1.1, $\phi_{m-1} \in \mathbf{N}_{\widetilde{k}_{m-1}}$, where the index $\widetilde{k}_{m-1}$ is

$$\widetilde{k}_m \leq k_{2m-2} + k_{2m-11}k_{2m} + k_{2m+1} + k_{m-1}^0 + k_m^0,$$

where the indices k_{2m-2}, k_{2m-1} and k_m^0 are defined by

$$k_{2m-2} = \kappa_-(-za_{2m-2}(z)), \quad k_{2m-1} = \kappa_- \left(\frac{a_{2m-1}(z) - a_{2m-1}(0)}{z} \right),$$
$$k_m^0 = \begin{cases} 1, \text{ if } a_{2m-1}(0) > 0; \\ 0, \text{ if } a_{2m-1}(0) < 0. \end{cases}$$

By induction, we obtain the sequence $\phi_m, \phi_{m-1}, ..., \phi_1$, where $\phi_1(z) = \frac{f(z)}{z}$ and $\phi_1 \in \mathbf{N}_k$ and k is defined by (4.45)–(4.46). Therefore, the function $f \in \mathbf{N}_\kappa^{-k}$, where the indices κ and k are generated by (4.45)–(4.46).

(ii) The second case. Let $n = 2m + 2$ in (2.9), then

$$\frac{f(z)}{z} = -\cfrac{1}{\Big| -za_0(z) - \cfrac{1}{\cfrac{a_1(z)}{z}}} - \cdots - \cfrac{1}{\Big| -za_{2m}(z) - \cfrac{1}{\cfrac{a_{2m+1}(z)}{z}}} - \cfrac{1}{\Big| -za_{2m+2}(z)}.$$

Let us set

$$\phi_{m+1}(z) = -\frac{1}{-za_{2m+2}(z)}.$$

By Proposition 1.1, $\phi_{m+1} \in \mathbf{N}_{k_{2m+2}}$, where $k_{2m+2} = \kappa_-(-za_{2m+2}(z))$. The next step, we apply the first case (i) and obtain $f \in \mathbf{N}_\kappa^{-k}$, where the indices κ and k satisfy (4.45)–(4.46). This completes the proof. □

Corollary 4.9. *Let the rational function* $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ *be the meromorphic on* $\mathbb{C}\backslash\mathbb{R}$, *where* $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ *and* $n_0 \leq m_0$. *Then* f *admits representation*

$$f(z) = a_{-1}(z) - \frac{1|}{|-a_0(z)} - \frac{1|}{|a_1(z)} - \cdots - \frac{1|}{|(-1)^{n+1}a_n(z)}. \tag{4.47}$$

Furthermore, f *belongs to the class* $\mathbf{N}_\kappa^{-k}$, *such that*

$$\kappa = \sum_{j=-1}^{n} \kappa_j \text{ and } k \leq \sum_{i=-1}^{n} k_i + \sum_{i=-1}^{[n/2]} k_i^0, \tag{4.48}$$

where the indices κ_i, k_i *and* k_i^0 *can be found by*

$$\begin{gathered} k_{2i} = \kappa_-(-za_{2i}(z)),\ \kappa_i = \kappa_-((-1)^{i+1}a_i(z)),\ k_i^0 = \begin{cases} 1, a_{2i+1}(0) > 0; \\ 0, a_{2i+1}(0) < 0, \end{cases} \\ k_{-1} = \kappa_{-1}\left(\frac{a_{-1}(z) - a_{-1}(0)}{z}\right), \quad k_{2i+1} = \kappa_-\left(\frac{a_{2i+1}(z) - a_{2i+1}(0)}{z}\right). \end{gathered} \tag{4.49}$$

Proof. Suppose the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ is the meromorphic on $\mathbb{C}\backslash\mathbb{R}$, where $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ and $n_0 \leq m_0$. By Euclidean algorithm, the function f admits representation (4.41).

We can rewrite the ratio $\frac{a_{-1}(z)}{z}$ as

$$\frac{a_{-1}(z)}{z} = \frac{a_{-1}(z) - a_{-1}(0)}{z} + \frac{a_{-1}(0)}{z}.$$

By Proposition 1.1 and Proposition 1.2, $a_{-1} \in \mathbf{N}_{\kappa_{-1}}^{-\widetilde{k}_{-1}}$, where $\widetilde{k}_{-1} \leq k_{-1} + k_{-1}^0$ and κ_{-1}, k_{-1}, k_{-1}^0 are defined by (4.49).

By Proposition 4.3, $(f - a_{-1}) \in \mathbf{N}_{\widetilde{\kappa}}^{\widetilde{k}}$, where the indices $\widetilde{\kappa}$ and $\widetilde{k}$ are defined by formulas (4.45)–(4.46). Therefore, the rational function $f(z) = \dfrac{Q_0(z)}{P_0(z)}$ belongs to the class $\mathbf{N}_\kappa^{-k}$ and the formulas (4.48)–(4.49) hold. This completes the proof. □

Theorem 4.5. *Let* $\tau \in \mathbf{N}_{\kappa^*}^{-k^*}$ *and let* $f(z) = \dfrac{Q_0(z)}{P_0(z)} + \tau(z)$, *where the* P_0 *and* Q_0 *are polynomials, such that* $\deg(P_0) = n_0$, $\deg(Q_0) = m_0$ *and* $m_0 < n_0$. *Then* $f \in \mathbf{N}_\kappa^{-k}$, *where*

$$\kappa \leq \kappa^* + \sum_{j=0}^{n} \kappa_j \text{ and } k \leq k^* + \sum_{i=0}^{n} k_i + \sum_{i=0}^{[n/2]} k_i^0, \tag{4.50}$$

where the indices κ_i, k_i *and* k_i^0 *can be found by* (4.46)

Proof. This proof is based on Proposition 4.4 and Proposition 1.1. □

REFERENCES

[1] V. Derkach: *On indefinite moment problem and resolvent matrices of Hermitian operators in Krein spaces*, Math. Nachr., **184** (1997), 135–166.

[2] V. Derkach, I. Kovalyov: *The Schur algorithm for indefinite Stieltjes moment problem*, Math. Nachr., (2017). DOI: 10.1002/mana.201600189.

[3] V. Derkach: *Generalized resolvents of a class of Hermitian operators in a Krein space*, Dokl. Akad. Nauk SSSR, **317** (4) (1991), 807–812.

[4] V. Derkach: *On Weyl function and generalized resolvents of a Hermitian operator in a Krein space*, Integral Equations Operator Theory, **23** (1995), 387–415 .

[5] V. A. Derkach, M. M. Malamud: *Generalized resolvents and the boundary value problems for Hermitian operators with gaps*, J. Funct.Anal., **95** (1) (1991) 1–95.

[6] I. Kovalyov: *A truncated indefinite Stieltjes moment problem*, J. Math. Sci., **224** (2017), 509–529.

[7] I Kovalyov: *Regularization of the indefinite Stieltjes moment problem*, Linear Algebra Appl., **594** (2020), 1–28.

[8] M. G. Krein, H. Langer: *Über einige Fortsetzungsprobleme, die eng mit der Theorie Hermitscher Operatoren in Raume* Π_κ *zusammenhängen, I. Einige Fuktionenklassen und ihre Dahrstellungen*, Math. Nachr., **77** (1977), 187–236.

[9] M. G. Krein, H. Langer: *Über einige Fortsetzungsprobleme, die eng mit der Theorie Hermitscher Operatoren in Raume* Π_κ *zusammenhängen, II,* J. of Funct. Analysis, **30** (1978), 390–447.

[10] M. G. Krein, H. Langer: *On some extension problem which are closely connected with the theory of Hermitian operators in a space* Π_κ *III. Indefinite analogues of the Hamburger and Stieltjes moment problems, Part I*, Beiträge zur Anal., **14** (1979), 25–40.

[11] M. G. Krein, H. Langer: *On some extension problem which are closely connected with the theory of Hermitian operators in a space* Π_κ *III. Indefinite analogues of the Hamburger and Stieltjes moment problems, Part II*, Beiträge zur Anal., **15** (1981), 27–45.

IVAN KOVALYOV
UNIVERSITÄT OSNABRÜCK
DEPARTMENT OF MATHEMATICS
SNEUER GRABEN/SCHLOSS, 49074 OSNABRÜCK, GERMANY
ORCID: 0000-0001-8464-3377
E-mail address: i.m.kovalyov@gmail.com

CONSTRUCTIVE MATHEMATICAL ANALYSIS
5 (2022), No. 3, pp. 168-182
http://dergipark.org.tr/en/pub/cma
ISSN 2651 - 2939

Research Article

Lower estimates on the condition number of a Toeplitz sinc matrix and related questions

LUDWIG KOHAUPT* AND YAN WU

ABSTRACT. As one new result, for a symmetric Toeplitz sinc $n \times n$-matrix $A(t)$ depending on a parameter t, lower estimates (tending to infinity as t vanishes) on the pertinent condition number are derived. A further important finding is that prior to improving the obtained lower estimates it seems to be more important to determine the lower bound on the parameter t such that the smallest eigenvalue $\mu_n(t)$ of $A(t)$ can be reliably computed since this is a precondition for determining a reliable value for the condition number of the Toeplitz sinc matrix. The style of the paper is expository in order to address a large readership.

Keywords: Condition number, eigenvalues and eigenvectors, inverse power method, power method, Toeplitz sinc matrix.

2020 Mathematics Subject Classification: 15B05, 65F15, 65F22.

1. INTRODUCTION

This paper is organized as follows. In Section 2, a symmetric Toeplitz sinc $n \times n$-matrix $A(t) = A_n(t)$ is defined and the problem with its pertinent condition number $\kappa_2(t)$ is described. The entries of this $n \times n$-matrix are made up of $s(0) := 1$ and $s(jt) := \sin(j\pi t)/(j\pi t),\ j = 1, \ldots, n-1$ and are investigated for $0 < t < 1$. Such a matrix appears frequently in the study of minimum phase filter designs [10] and numerical integration/differentiation of bandlimited systems [11]. As properties of the matrices $A(t)$, we found that the limit $\lim_{t\to 0} A(t) = A$ exists and also that, for the eigenvalues $\mu_j(t),\ j = 1, \ldots, n$ of $A(t)$, the limits $\lim_{t\to 0} \mu_j(t) = \mu_j = \mu_j(A),\ j = 1, \ldots, n$ exist and, further, that the values of the entries of A and $\mu_j,\ j = 1, \ldots, n$ can be given explicitly. In Section 3, two-sided estimates on $\mu_j(t),\ j = 1, \ldots, n$ are derived. The eigenvalues are arranged according to $\mu_1(t) \geq \cdots \geq \mu_n(t)$ and $\mu_1 \geq \cdots \geq \mu_n$. In Section 4, two upper bounds on the smallest eigenvalue $\mu_n(t)$ are obtained. Thereby, in Section 5, three lower estimates on the condition number $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ can be derived. These lower bounds are new and tend to infinity as t tends to zero. For comparison reasons, in Section 6, a lower bound on $\mu_1(t)$ and an upper bound on $\mu_n(t)$ are stated from a paper of D. Hertz delivering an upper bound on the condition number. Section 7 contains numerical verifications of the obtained estimates on $\kappa_2(t)$ for some examples. In Section 8, linearly independent eigenvectors of the matrix $A = \lim_{t\to 0} A(t)$ are derived that form a basis of $\mathbb{R}^n$. Then, in Section 9, appropriate computational methods for the determination of $\mu_n(t)$ and $\mu_1(t)$ are presented, and in Section 10, these computational methods are applied to a series of matrices $A(t) = A_n(t)$. Finally, Section 11 contains the conclusions followed by the References.

Received: 10.07.2022; Accepted: 08.08.2022; Published Online: 15.08.2022
*Corresponding author: Ludwig Kohaupt; lkohaupt4@web.de
DOI: 10.33205/cma.1142905

2. Some Properties of a Toeplitz Sinc Matrix $A(t)$

A Toeplitz sinc matrix is defined as

$$A(t) = A_n(t) = \begin{bmatrix} s(0) & s(t) & s(2\,t) & \cdots & s((n-2)\,t) & s((n-1)\,t) \\ s(t) & s(0) & s(t) & \cdots & s((n-3)\,t) & s((n-2)\,t) \\ s(2\,t) & s(t) & s(0) & \cdots & & s((n-3)\,t) \\ & & & \vdots & & \\ s((n-1)\,t) & s((n-2)\,t) & & \cdots & s(t) & s(0) \end{bmatrix}, \tag{2.1}$$

where $0 < t < 1$ and $s(0) = 1$ as well as $s(t) = \operatorname{sinc}(t) = \sin(\pi t)/(\pi t)$. From [9, Theorem 2.2], it follows that this matrix is positive definite by setting there $t_1 = 0$, $t_i = (i-1)\,t$, $i = 2, \ldots, n$ and taking into account that s(-t)=s(t) for $0 < t < 1$. As t gets smaller, the condition number of this matrix deteriorates quickly. The question that one might ask therefore is: Can one find a way to estimate the largest and smallest eigenvalues of this matrix? This would help us to monitor the condition number of the above Toeplitz sinc matrix and is the starting point of our investigation.

The following theorem presents some properties of the matrices $A(t)$.

Theorem 2.1. *Let the matrix $A(t) = A_n(t)$ in (2.1) be given. Further, let the eigenvalues $\mu_j(t) = \mu_j(A(t)) = \mu_j(A_n(t))$ be arranged according to*

$$\mu_1(t) \geq \mu_2(t) \geq \cdots \geq \mu_n(t). \tag{2.2}$$

Then, the limits

$$A := \lim_{t \to 0} A(t) \tag{2.3}$$

as well as

$$\mu_j = \mu_j(A) := \lim_{t \to 0} \mu_j(t) = \lim_{t \to 0} \mu_j(A(t)), \quad j = 1, \ldots, n \tag{2.4}$$

exist and

$$A = A_n = \lim_{t \to 0} A(t) = \lim_{t \to 0} A_n(t) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \\ & & & & & \\ & & & \vdots & & \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}. \tag{2.5}$$

Further, the limits in (2.4) are eigenvalues of A, and with appropriate enumeration of the eigenvalues $\mu_j := \mu_j(A)$, $j = 1, \ldots, n$, one has

$$\lim_{t \to 0} \mu_1(A(t)) = \lim_{t \to 0} \mu_1(A_n(t)) = \mu_1(A) = \mu_1 = n, \tag{2.6}$$

$$\lim_{t \to 0} \mu_j(A(t)) = \lim_{t \to 0} \mu_j(A_n(t)) = \mu_j(A) = \mu_j = 0, \; j = 2, \ldots, n. \tag{2.7}$$

Proof. (2.5): The Toeplitz matrix $A(t) = A_n(t) \in \mathbb{R}^{n \times n}$ according to (2.1) reads

$$A(t) = A_n(t) = \begin{bmatrix} s(0) & s(t) & s(2\,t) & \cdots & s((n-2)\,t) & s((n-1)\,t) \\ s(t) & s(0) & s(t) & \cdots & s((n-3)\,t) & s((n-2)\,t) \\ s(2\,t) & s(t) & s(0) & \cdots & & s((n-3)\,t) \\ & & & \vdots & & \\ s((n-1)\,t) & s((n-2)\,t) & & \cdots & s(t) & s(0) \end{bmatrix}$$

for $0 < t < 1$ and $s(0) := \lim_{t\to 0} s(t) := \lim_{t\to 0} \operatorname{sinc}(t) := \lim_{t\to 0} \sin(\pi\, t)/(\pi\, t) = 1$. From this, apparently (2.5) follows.

(2.6) and (2.7): This is seen as follows. Matrix $A = A_n$ in (2.5) is a rank-one symmetric matrix. Hence, there is only one non-zero eigenvalue, namely $\mu_1(A)$, and therefore $\mu_1(A)$ must equal $tr(A) = n$. Since the limits $\lim_{t\to 0} \mu_j(t)$ exist and are equal to μ_j for $j = 1, \ldots, n$ if one chooses an appropriate enumeration, the assertion follows.

(2.3): This follows immediately from (2.5).

(2.4): This follows immediately from (2.6) and (2.7).

So, on the whole, Theorem 2.1 is proven. □

For later use, we arrange the eigenvalues $\mu_j = \mu_j(A)$ according to

$$\mu_1 \geq \cdots \geq \mu_n. \tag{2.8}$$

Remark 2.1. *Another elementary proof of Theorem 2.1 will be given at the end of Section 3.*

Remark 2.2. *Theorem 2.1 also follows from* [4, Theorem 17, p. 263] *that is a much more general result.*

3. Two-Sided Estimates on the Eigenvalues $\mu_j(t)$, $j = 1, \ldots, n$ of $A(t)$

(i) Upper Estimate on $\mu_1(t)$ According to [5, Section 5.4, Formula (7), p. 89], we have

$$\mu_1(t) = |\mu_1(t)| \leq \|A(t)\|_\infty = \max_{j=1,\cdots,n} \sum_{k=1}^{n} |a_{jk}(t)|\,.$$

Now,

$$s((n-1)\,t) < \cdots < s(2\,t) < s(t) < s(0) = 1$$

yielding the upper estimate

$$0 < \mu_1(t) \leq \sum_{k=0}^{n-1} 1 = n. \tag{3.9}$$

(ii) Lower Estimate on $\mu_1(t)$

We use [5, Section 5.4, Formula (28), p. 94]. Thereby, employing (2.2) and (2.8),

$$|\mu_1 - \mu_1(t)| \leq \|A - A(t)\|_\infty = \max_{j=1,\cdots,n} \sum_{k=1}^{n} |a_{jk} - a_{jk}(t)|$$

with

(3.10)

$$A - A(t) = \begin{bmatrix} 0 & 1-s(t) & 1-s(2\,t) & \cdots & 1-s((n-2)\,t) & 1-s((n-1)\,t) \\ 1-s(t) & 0 & 1-s(t) & \cdots & 1-s((n-3)\,t) & 1-s((n-2)\,t) \\ 1-s(2\,t) & 1-s(t) & 0 & \cdots & & 1-s((n-3)\,t) \\ & & & \vdots & & \\ 1-s((n-1)\,t) & 1-s((n-2)\,t) & & \cdots & 1-s(t) & 0 \end{bmatrix}.$$

Because of

$$1 > s(t) > s(2\,t) > \cdots > s((n-1)\,t),$$

we obtain

$$-s(t) < -s(2\,t) < \cdots < -s((n-1)\,t)$$

and thus

$$1 - s(t) < 1 - s(2\,t) < \cdots < 1 - s((n-1)\,t). \tag{3.11}$$

Therefore,

$$\begin{aligned} \mu_1(t) &= \mu_1(t) - \mu_1 + \mu_1 \geq \mu_1 - |\mu_1 - \mu_1(t)| \\ &\geq \mu_1 - \|A - A(t)\|_\infty \geq \mu_1 - n\,[1 - s((n-1)\,t)] \\ &= n - n\,[1 - s((n-1)\,t)] = n\,s((n-1)\,t) \end{aligned}$$

so that we obtain the lower estimate

$$\mu_1(t) \geq n\,s((n-1)\,t)\,. \tag{3.12}$$

(iii) Two-Sided Estimate on $\mu_1(t)$

On the whole, we have the two-sided estimate

$$n\,s((n-1)\,t) \leq \mu_1(t) \leq n\,. \tag{3.13}$$

(iv) Two-Sided Estimates on $\mu_j(t),\ j = 2, \ldots, n$

Since $\mu_j = 0,\ j = 2, \ldots, n$, one has

$$0 < \mu_j(t) = |-\mu_j(t)| = |\mu_j - \mu_j(t)| \leq \|A - A(t)\|_\infty \leq \max_{j=1,\ldots,n} \sum_{k=1}^{n} |a_{jk} - a_{jk}(t)|.$$

Along with (3.10) and (3.11), we herewith conclude that

$$0 < \mu_j(t) < (n-1)\,[1 - s((n-1)\,t)],\ j = 2, \ldots, n\,. \tag{3.14}$$

(v) Elementary Proof of $\lim_{t\to 0} \mu_1(t) = n$

Taking the limit as $t \to 0$ in the two-sided estimate (3.13), we get

$$\lim_{t\to 0} \mu_1(t) = n \tag{3.15}$$

since

$$\lim_{t\to 0} s((n-1)\,t) = 1\,.$$

(vi) Elementary Proof of $\lim_{t\to 0} \mu_j(t) = 0,\ j = 2, \ldots, n$

Taking the limit as $t \to 0$ in the two-sided estimate (3.14), we get

$$\lim_{t\to 0} \mu_j(t) = 0,\ j = 2, \ldots, n \tag{3.16}$$

since

$$\lim_{t\to 0} s((n-1)\,t) = 1\,.$$

(vii) Elementary Proof of $\lim_{t\to 0} \mu_1(t) = \mu_1 = n$

One has the chain of implications

$$det(A(t) - \mu_1(t)\,I) = 0$$

$\Rightarrow$

$$\lim_{t\to 0} det(A(t) - \mu_1(t)\,I) = 0$$

$\Rightarrow$

$$det(\lim_{t\to 0} A(t) - \lim_{t\to 0} \mu_1(t)\,I) = 0$$

$\Rightarrow$

$$det(A - \lim_{t\to 0} \mu_1(t)\,I) = 0.$$

Thus, $\lim_{t\to 0} \mu_1(t)$ is an eigenvalue of A that is denoted by μ_1. Therefore,

$$det(A - \mu_1\,I) = 0.$$

Together with

$$\mu_1 = n,$$

one obtains

$$\lim_{t\to 0} \mu_1(t) = n = \mu_1\,. \tag{3.17}$$

(viii) Elementary Proof of $\lim_{t\to 0} \mu_j(t) = \mu_j = 0,\ j = 2,\ldots,n$

The Proof is similar to that in (vii).

Remark 3.3. *The points (v) - (viii) deliver an elementary proof of Theorem 2.1. This is because they show that the limits* $\lim_{t\to 0} \mu_j(t),\ j = 1,\ldots,n$ *exist and are eigenvalues of A. In particular, the elementary proof is independent of* [4] *the application of which is in a way like using a sledge-hammer to crack a nut.*

4. Two Upper Estimates on Smallest Eigenvalue $\mu_n(t)$

(i) First Upper Estimate

As is known,

$$0 < |A(t)| := det(A(t)) = \mu_1(t)\,\mu_2(t)\,\cdots\,\mu_n(t) < 1$$

at least for sufficiently small t in $0 < t < 1$ since from Section 3 we know that $\mu_j = \mu_j(A) = 0,\ j = 2,\ldots,n$ so that in particular $\mu_n(t) \to 0$ as $t \to 0$. This entails

$$0 < [\mu_n(t)]^n \le |A(t)| < 1$$

for $0 < t \le t_1$ with sufficiently small t_1 or the first upper estimate

$$0 < \mu_n(t) \le |A(t)|^{\frac{1}{n}} < 1 \tag{4.18}$$

for $0 < t \le t_1$ with sufficiently small t_1.

(ii) Second Upper Estimate

The derivation of the second upper estimate is based on [3, Corollary 8.1,4, p.411] that, in turn, is proven in [8, pp. 103-104] using the Courant-Fischer Minimax Theorem. The cited corollary is called Theorem 4.2 here and, in our notation, reads as follows:

Theorem 4.2. *If* A_r *denotes the leading r-by-r principal submatrix of an n-by-n symmetric matrix A, then for* $r = 1 : n-1$ *the following* interlacing property *holds:*

$$\mu_{r+1}(A_{r+1}) \le \mu_r(A_r) \le \mu_r(A_{r+1}) \le \cdots \le \mu_2(A_{r+1}) \le \mu_1(A_r) \le \mu_1(A_{r+1})\,.$$

For $r = n-1$, Theorem 4.2 delivers

$$\mu_n(A_n) \le \mu_{n-1}(A_{n-1}),$$

where $A_n = A$. Now, we apply the last estimate to the Toeplitz sinc matrix (for short: Tsinc matrix) $A(t) = A_n(t) \in \mathbb{R}^{n\times n}$ and remark that the leading $(n-1)\times(n-1)$ submatrix of this matrix is the Tsinc matrix $A_{n-1}(t)$. This entails the chain of inequalities

$$\mu_n(A_n(t)) \le \mu_{n-1}(A_{n-1}(t)),$$
$$\mu_{n-1}(A_{n-1}(t)) \le \mu_{n-2}(A_{n-2}(t)),$$
$$\cdots$$
$$\mu_3(A_3(t)) \le \mu_2(A_2(t)),$$

where

$$\mu_2(A_2(t)) = 1 - s(t)$$

and

$$\mu_1(A_2(t)) = 1 + s(t)$$

which follows from

$$|A_2(t) - \mu(t)I| = \left| \begin{array}{c|c} 1-\mu(t) & s(t) \\ \hline s(t) & 1-\mu(t) \end{array} \right| = 0\,.$$

This yields the second upper estimate

$$0 < \mu_n(t) := \mu_n(A_n(t)) \le 1 - s(t), \quad 0 < t < 1\,. \tag{4.19}$$

5. Three Lower Estimates on Condition Number $\kappa_2(t) := \mu_1(t)/\mu_n(t)$

(i) First Lower Estimate on $\kappa_2(t)$

From the first upper estimate on $\mu_n(t)$, we obtain

$$\frac{1}{\mu_n(t)} \ge \frac{1}{|A(t)|^{\frac{1}{n}}} > 1$$

for $0 < t \le t_1$ with sufficiently small t_1. This yields the first lower estimate

$$\kappa_2(t) = \mu_1(t)/\mu_n(t) \ge \frac{n\, s((n-1)\, t)}{|A(t)|^{\frac{1}{n}}} := e_1(t) \tag{5.20}$$

for $0 < t \le t_1$ with sufficiently small t_1.

(ii) Second Lower Estimate on $\kappa_2(t)$

From the second upper estimate on $\mu_n(t)$, we obtain

$$\frac{1}{\mu_n(t)} \ge \frac{1}{1 - s(t)}$$

for $0 < t \le t_1$ with sufficiently small t_1. This yields the second lower estimate

$$\kappa_2(t) = \mu_1(t)/\mu_n(t) \ge \frac{n\, s((n-1)\, t)}{1 - s(t)} := e_2(t) \tag{5.21}$$

for all t in $0 < t < 1$.

(iii) Third Lower Estimate on $\kappa_2(t)$

Combining the preceding results, one gets the third lower estimate

$$\kappa_2(t) = \mu_1(t)/\mu_n(t) \ge \max\{e_1(t), e_2(t)\} := e_3(t) \tag{5.22}$$

for $0 < t \le t_1$ with sufficiently small t_1.

6. Bounds Stated by D. Hertz and Application

In this section, we apply the bounds on the extreme eigenvalues of Toeplitz matrices stated in [1] to our symmetric Toeplitz matrix $A(t)$ defined in (2.1). As application, one obtains upper bounds on $\kappa_2(t) = \mu_1(t)/\mu_n(t)$.

Let

$$a(t) = [a_1(t),\, a_2(t), \ldots,\, a_n(t)], \qquad 0 < t < 1 \tag{6.23}$$

be the first row of $A(t)$, and define

$$\tilde{a}(t) := [a_1(t),\, |a_2(t)|, \ldots,\, |a_n(t)|], \qquad 0 < t < 1. \tag{6.24}$$

From (2.1), we have

$$\tilde{a}(t) = a(t), \qquad 0 < t < 1. \tag{6.25}$$

Further, define

$$\overline{\lambda}_k = -\underline{\lambda}_k = 2\cos\left(\frac{\pi}{\text{floor}[(n-1)/(k-1)]+2}\right), \quad k = 2,\ldots,n. \tag{6.26}$$

As in Section 2, we assume that the eigenvalues $\mu_k(t)$, $k = 1,\ldots,n$ are arranged according to (2.2). Then, one has, in our notation, the following theorem.

Theorem 6.3. *The maximal eigenvalue $\mu_1(t)$ of the symmetric Toeplitz matrix $A(t)$ in (2.1) is bounded from above by the inner product*

$$\mu_1(t) \le (a(t), \overline{w}), \qquad 0 < t < 1, \tag{6.27}$$

where $a(t)$ is as in (6.23) and the vector $\overline{w}$ is defined by

$$\overline{w} = [1, \overline{\lambda}_2, \ldots, \overline{\lambda}_n] \tag{6.28}$$

and $\overline{\lambda}_k$ is as in (6.26).

Proof. The theorem is a direct consequence of [1, Theorem 1]. □

Remark 6.4. *Theorem 6.3 can hold only if $(a(t), \overline{w}) > 0$, of course.*

Further, one has the following theorem.

Theorem 6.4. *The minimal eigenvalue $\mu_n(t)$ of the symmetric Toeplitz matrix $A(t)$ in (2.1) is bounded from below by the inner product*

$$\mu_n(t) \ge (a(t), \underline{w}), \qquad 0 < t < 1, \tag{6.29}$$

where $a(t)$ is as in (6.23) and the vector $\underline{w}$ is defined by

$$\underline{w} = [1, \underline{\lambda}_2, \ldots, \underline{\lambda}_n]. \tag{6.30}$$

Note that using (6.26), we obtain

$$\underline{w} = [1, -\overline{\lambda}_2, \ldots, -\overline{\lambda}_n]. \tag{6.31}$$

Proof. The theorem is a direct consequence of [1, Theorem 2]. □

Remark 6.5. *Theorem 6.4 can hold only if $(a(t), \underline{w}) > 0$, of course.*

Remark 6.6. *From Theorems 6.3 and 6.4, we get the upper estimates*

$$\kappa_2(t) = \mu_1(t)/\mu_n(t) \le (a(t), \overline{w})/(a(t), \underline{w}), \quad 0 < t < 1 \tag{6.32}$$

provided that $(a(t), \overline{w}) > 0$ and $(a(t), \underline{w}) > 0$ for $0 < t < 1$.

7. Numerical Verification of the Estimates on $\kappa_2(t) := \mu_1(t)/\mu_n(t)$ for Some Examples

In this section, we present estimates on $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ for fixed $t = 0.1$ and $n = 2,\ldots,6$.

For this, corresponding Matlab computations were carried out. The expressions $e_1(t)$, $e_2(t)$, $e_3(t)$ are estimates from below (tending to ∞ as $n \to \infty$ and $t \to 0$) on $\kappa_2(t)$, expression $e_4(t)$ is defined as condition number $\kappa_2(t) = \mu_1(t)/\mu_n(t)$, whereas expression $e_5(t)$ is an estimate from above on $\kappa_2(t)$ provided that $(a(t), \overline{w}) > 0$ and $(a(t), \underline{w}) > 0$. Its derivation follows from two theorems stated by D. Hertz. The pertinent upper estimate should at least be positive since $\kappa_2(t)$ is so. But, it turns out to be negative since $(a(t), \underline{w}) < 0$ for $n \ge 3$. Consequently, $e_5(t)$ cannot deliver an upper bound on $\kappa_2(t)$.

In the following estimate $e_1(t)$, the determinant $|A(t)| = \det(A(t))$ enters. This is computed in two ways, namely first with Matlab routine *det* and second, for comparison reasons, as a product

of the eigenvalues of $A(t)$. From numerical considerations, it is clear that the determination of $|A(t)|$ can be achieved through elementary operations by casting matrix $A(t)$ into triangular form without changing the determinant so that the product of the diagonal elements gives the determinant. We think that this technique is behind the Matlab routine *det*. The second way via the product of the eigenvalues that is computationally much more costly is used only for comparison reasons. This is because if one computes the determinant via the product of eigenvalues $\mu_j(t)$, $j = 1, \dots, n$, then one has immediately $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ and needs no estimates.

Now, the details of the computations follow.

For $n = 2$, we obtain

$$A(t) = \begin{bmatrix} 1.000000000000000 & 0.983631643083466 \\ 0.983631643083466 & 1.000000000000000 \end{bmatrix},$$

$$\mu(t) = \begin{bmatrix} 0.016368356916534 \\ 1.983631643083466 \end{bmatrix},$$

$$d(n,t) := \mu_1(t)\,\mu_2(t) = 0.032468790724921,$$

$$|A(t)| = \det(A(t)) = det(A_n(t)) = 0.032468790724921,$$

and

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \end{bmatrix} := \begin{bmatrix} n\,s((n-1)t)/|A(t)|^{\frac{1}{n}} \\ n\,s((n-1)t)/(1-s(t)) \\ \max\{e_1(t), e_2(t)\} \\ \mu_1(t)/\mu_n(t) \\ (a(t), \overline{w})/(a(t), \underline{w}) \end{bmatrix} = \begin{bmatrix} 10.917656600748638 & \\ 1.201869739399289 & \times 10^2 \\ 1.201869739399289 & \times 10^2 \\ 1.211869739399293 & \times 10^2 \\ 1.211869739399306 & \times 10^2 \end{bmatrix}.$$

For $n = 3$, we obtain

$$A(t) = \begin{bmatrix} 1.000000000000000 & 0.983631643083466 & 0.935489283788639 \\ 0.983631643083466 & 1.000000000000000 & 0.983631643083466 \\ 0.935489283788639 & 0.983631643083466 & 1.000000000000000 \end{bmatrix},$$

$$\mu(t) = \begin{bmatrix} 0.000145422566712 \\ 0.064510716211361 \\ 2.935343861221926 \end{bmatrix},$$

$$d(n,t) := \prod_{j=1}^{n} \mu_1(t) \dots \mu_j(t) = 0.032468790724921,$$

$$|A(t)| = \det(A(t)) = det(A_n(t)) = 0.032468790724921,$$

and

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \end{bmatrix} := \begin{bmatrix} n\,s((n-1)t)/|A(t)|^{\frac{1}{n}} \\ n\,s((n-1)t)/(1-s(t)) \\ \max\{e_1(t), e_2(t)\} \\ n\,s((n-1)t)/(1-s(t)) \\ \mu_1(t)/\mu_n(t) \\ (a(t), \overline{w})/(a(t), \underline{w}) \end{bmatrix} = \begin{bmatrix} 92.936401783180301 & \\ 1.714569071090478 & \times 10^2 \\ 1.714569071090478 & \times 10^2 \\ 2.018492677986877 & \times 10^2 \\ -2.507665165149629 & \end{bmatrix}.$$

For $n = 4$, we obtain

$$A(t) = \begin{bmatrix} 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 \\ 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 \\ 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 \\ 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 \end{bmatrix},$$

$$\mu(t) = \begin{bmatrix} 0.000001113119258 \\ 0.000870415161304 \\ 0.157973552463136 \\ 3.841154919256300 \end{bmatrix},$$

$$d(n,t) := \prod_{j=1}^{n} \mu_1(t) \dots \mu_j(t) = 5.879147433554857 \times 10^{-1},$$

$$|A(t)| = \det(A(t)) = det(A_n(t)) = 5.879147434765446 \times 10^{-1},$$

and

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \end{bmatrix} := \begin{bmatrix} n\, s((n-1)t)/|A(t)|^{\frac{1}{n}} \\ n\, s((n-1)t)/(1-s(t)) \\ \max\{e_1(t), e_2(t)\} \\ \mu_1(t)/\mu_n(t) \\ (a(t), \overline{w})/(a(t), \underline{w}) \end{bmatrix} = \begin{bmatrix} 6.972971989701032 & \times 10^2 \\ 2.097690551864878 & \times 10^2 \\ 6.972971989701032 & \times 10^2 \\ 3.450802681898942 & \times 10^6 \\ -1.838422415548006 & \end{bmatrix}.$$

For $n = 5$, we obtain

$$A(t) = \begin{bmatrix} 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 & 0.756826728640657 \\ 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 \\ 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 \\ 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 \\ 0.756826728640657 & 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 \end{bmatrix},$$

$$\mu(t) = \begin{bmatrix} 0.000000008008103 \\ 0.000008896854399 \\ 0.003035674827786 \\ 0.307675090716305 \\ 4.689280329593409 \end{bmatrix},$$

$$d(n,t) := \prod_{j=1}^{n} \mu_1(t) \dots \mu_j(t) = 3.120469248845038 \times 10^{-16},$$

$$|A(t)| = \det(A(t)) = det(A_n(t)) = 3.120469141684447 \times 10^{-16},$$

and

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \end{bmatrix} := \begin{bmatrix} n\, s((n-1)t)/|A(t)|^{\frac{1}{n}} \\ n\, s((n-1)t)/(1-s(t)) \\ \max\{e_1(t), e_2(t)\} \\ \mu_1(t)/\mu_n(t) \\ (a(t), \overline{w})/(a(t), \underline{w}) \end{bmatrix} = \begin{bmatrix} 4.776639739123309 & \times 10^3 \\ 2.311859194236440 & \times 10^2 \\ 4.776639739123309 & \times 10^3 \\ 5.855669475721616 & \times 10^8 \\ -1.549163591209945 & \end{bmatrix}.$$

For $n = 6$, we obtain

$$A(t) = \begin{bmatrix} 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 & 0.756826728640657 & 0.636619772367581 \\ 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 & 0.756826728640657 \\ 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 & 0.858393691334140 \\ 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 & 0.935489283788639 \\ 0.756826728640657 & 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 & 0.983631643083466 \\ 0.636619772367581 & 0.756826728640657 & 0.858393691334140 & 0.935489283788639 & 0.983631643083466 & 1.000000000000000 \end{bmatrix},$$

$$\mu(t) = \begin{bmatrix} 0.000000000055683 \\ 0.000000080040530 \\ 0.000039987658742 \\ 0.008055094169584 \\ 0.521314905500388 \\ 5.470589932575071 \end{bmatrix},$$

$$d(n,t) := \prod_{j=1}^{n} \mu_1(t) \dots \mu_j(t) = 4.094114597934044 \times 10^{-24},$$

$$|A(t)| = \det(A(t)) = det(A_n(t)) = 4.094097171079637 \times 10^{-24},$$

and

$$\begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \end{bmatrix} := \begin{bmatrix} n\, s((n-1)t)/|A(t)|^{\frac{1}{n}} \\ n\, s((n-1)t)/(1-s(t)) \\ \max\{e_1(t), e_2(t)\} \\ \mu_1(t)/\mu_n(t) \\ (a(t), \overline{w})/(a(t), \underline{w}) \end{bmatrix} = \begin{bmatrix} 3.019986597952274 & \times & 10^4 \\ 2.333599306077634 & \times & 10^2 \\ 3.019986597952274 & \times & 10^4 \\ 9.824603336342802 & \times & 10^{10} \\ -1.460059391749488 & & \end{bmatrix}.$$

Discussion of the Computational Results on the Estimates on $\kappa_2(t) := \mu_1(t)/\mu_n(t)$ for the Examples

The computational results underpin the theoretical findings. In particular, they show that the lower estimates $e_1(t)$, $e_2(t)$, $e_3(t)$ on $e_4(t) = \kappa_2(t) = \mu_1(t)/\mu_n(t)$ tend to ∞ as $t \to 0$, as it must be. Further, apparently expression $e_3(t)$ is the best lower bound out of the lower bounds $e_j(t)$, $j = 1, 2, 3$. But, with growing dimension n, it underestimates the condition number $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ significantly. In order to find out more on the reason for this, in the next sections, it will be investigated for what values of t and to how many decimal places the eigenvalues $\mu_1(t)$, $\mu_n(t)$, and the condition number $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ can be determined. We hope that the pertinent results will deliver upper bounds on n and lower bounds on t that form estimates for the applicability of the best estimate $e_3(t)$ on $e_4(t) = \kappa_2(t)$.

The estimates stated by D. Hertz for $n \geq 3$ are not applicable since $(a(t), \overline{w}) < 0$ for $n \geq 3$.

8. The Eigenvectors of $A = \lim_{t\to 0} A(t)$

For symmetric matrices $A(t)$ and A, when $A(t) \to A$ $(t \to 0)$, one uses, as a rule, the eigenvectors of $A(t)$ associated with an eigenvalue $\mu(t)$ of $A(t)$ as an approximation of an eigenvector of A provided the eigenvectors of $A(t)$ can be determined much easier than those of A. Here, it is almost the other way around. The reason for this is that the eigenvalues of matrix A can be determined very simply, and various linearly independent associated eigenvectors can likewise be determined very easily.

This will be shown in the present section.

In the next section, these eigenvectors of A will be used as initial vectors for the power method resp. the inverse power method to compute $\mu_1(t)$ resp. $\mu_n(t)$.

Now, the computational details follow. $n = 3$:

(i) Determination of the eigenvector w_1 associated with $\mu_1 = 3$

One has

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

so that

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore, the eigenvector w_1 pertinent to $\mu_1 = 3$ is equal to

$$w_1 = e := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The associated normed eigenvector reads

$$w_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^3.$$

The generalization to the case $A \in \mathbb{R}^{n \times n}$ clearly is

$$w_1 = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n.$$

(ii) Determination of the eigenvector w_2 and w_3 associated with $\mu_2 = 0$ and $\mu_3 = 0$

From

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

we obtain

$$v_1 + v_2 + v_3 = 0$$

or

$$v_3 = -v_1 - v_2.$$

$v_1 = 1$, $v_2 = 1$:

With these values,

$$v = \begin{bmatrix} v_1 \\ v_2 \\ -v_1 - v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

$v_1 = 1$, $v_2 = -1$:

With these values,

$$v = \begin{bmatrix} v_1 \\ v_2 \\ -v_1 - v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

The normed eigenvectors are thus

$$w_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad w_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \quad , \quad w_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Apparently,

$$(w_j, w_k) = \delta_{j,k}, \; j, k = 1, 2, 3\,.$$

There are other eigenvectors, for example,

$$w_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \; w_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \; w_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

$\underline{n=5:}$ Let

$$w_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \; w_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \; w_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \; w_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \; w_5 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then, $w_1 \in \mathbb{R}^n = \mathbb{R}^5$ is a normed eigenvector corresponding to the largest eigenvalue $\mu_1 = n = 5$ of $A \in \mathbb{R}^{n\times n} = \mathbb{R}^{5\times 5}$, whereas w_j, $j = 2, \ldots, n = 5$ are linearly independent normed eigenvectors corresponding to the eigenvalues $\mu_j = 0$, $j = 2, \ldots, n = 5$ that are linearly independent, but not pairwise orthogonal. However, one has

$$(w_1, w_j) = 0,\ j = 2, \ldots, 5\,.$$

Let

$$e_{\pm 1} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; e_{\pm 2} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \; e_{\pm 3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \; e_{\pm 4} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \; e_{\pm 5} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^n = \mathbb{R}^5$$

and

$$e = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^n = \mathbb{R}^5.$$

Then, the components of $e_{\pm j}$, $j = 2, 3, 4, 5$ are cyclic permutations of $e_{\pm 1}$:

$$e_{\pm 2} = P(23451)(e_{\pm 1}),\ e_{\pm 3} = P(34512)(e_{\pm 1}),\ e_{\pm 4} = P(45123)(e_{\pm 1}),\ e_{\pm 5} = P(51234)(e_{\pm 1})$$

so that

$$w_1 = \frac{1}{\sqrt{5}}\, e,\ w_2 = \frac{1}{\sqrt{2}}\, e_{\pm 2},\ w_3 = \frac{1}{\sqrt{2}}\, e_{\pm 3},\ w_4 = \frac{1}{\sqrt{2}}\, e_{\pm 4},\ w_5 = \frac{1}{\sqrt{2}}\, e_{\pm 5}.$$

A set of pairwise orthogonal eigenvectors can be obtained when we apply Schmidt's orthogonalization method to these linear independent eigenvectors w_j, $j = 1, \ldots, 5$.

The generalization from $n = 5$ to arbitrary $n \in \mathbb{N}$ of eigenvectors w_j, $j = 1, \ldots, n$ as above can be done in a straightforward way.

9. Appropriate Computational Methods for the Determination of $\mu_n(t)$ and $\mu_1(t)$

Since $\mu_n(t) \to 0$ $(t \to 0)$ and $\mu_1(t) \to n$ $(t \to 0)$, it is clear that $\kappa_2(t) \to \infty$ $(t \to 0)$ which posed the problem to determine lower estimates on $\kappa_2(t)$. A related important question is how $\mu_n(t)$ and $\mu_1(t)$ can be computed such that the outcome is reliable.

For the determination of the largest eigenvalue $\mu_1(t)$ of $A(t) \in \mathbb{R}^{n\times n}$, the *power method* is appropriate as described, for example, in [5, Section 10.1.1] and for compact symmetric operators in [7, Section 7]. As initial vector $x_0 \in \mathbb{R}^n$, one can use every non-zero real n-vector. However, the eigenvector $w_1 = (1/\sqrt{n})\,[1, \ldots, 1]^T \in \mathbb{R}^n$ corresponding to $\mu_1 = \mu_1(A)$ seems to be especially advantageous as initial vector x_0.

For the determination of the smallest eigenvalue $\mu_n(t)$ of $A(t)$, the *inverse iteration* can be used as described in [5, Section 10.1.3]. This is a modification of the power method where the

n	t	$\mu_n(t)$	$\mu_1(t)$
5	0.1	0.800810×10^{-10}	4.68928
10	0.3	0.586776×10^{-12}	3.33055
15	0.6	0.669565×10^{-8}	1.66666

TABLE 1. Computational Results for $\mu_n(t)$ and $\mu_1(t)$

power method is applied to the inverse of a non-singular matrix. For short, we call this method *inverse power method*.

Based on these methods, pertinent Matlab programs were developed. For comparative reasons, also Matlab routine eig.m is applied that computes not only the largest and smallest eigenvalues of a square matrix, but all eigenvalues which is computationally disadvantages, of course.

10. Application of the Computational Methods to a Series of Matrices $A(t)$

First, with the inverse power method mentioned in Section 9, for $n = 5$, 10, 15 and $t = 0.1$, we tried to determine the smallest eigenvalues $\mu_n(t)$. For $n = 5$ and $t = 0.1$, this was possible. For $n = 10$ and $t = 0.1$, the developed Matlab program issued the **error code NaN** meaning **Not a Number**. This error code is typically put out by Matlab, for instance, when a division by zero is tried. For short, the determination of $\mu_n(t)$ by the inverse power method was not possible for $n = 10$ and $t = 0.1$. It was neither possible for $n = 10$ and $t = 0.2$. However, the determination of $\mu_n(t)$ was possible for $n = 10$ and $t = 0.3$. Similarly for $n = 15$ and $t = 0.1, \ldots, 0.5$, $\mu_n(t)$ could not be determined by the inverse power method. However, $\mu_n(t)$ could be successfully determined for $n = 15$ and $t = 0.6$.

Further, for all those pairs (n, t) the smallest eigenvalues $\mu_n(t)$ could be computed successfully for, also the pertinent largest eigenvalues $\mu_1(t)$ could be determined by the power method. In Table 1, the computational results are compiled. For comparison reasons, we applied also the Matlab routine eig.m.

For $n = 5$, $t = 0.1$, we obtained the following vector of not-arranged eigenvalues

$$\mu(t) = [-0.1224,\ -0.6286,\ -0.5261,\ 0.3238,\ 0.4326]^T.$$

Since $\mu_j(t) < 0,\ j = 1, 2, 3$, program eig.m delivers a false result without issuing a warning or error code.

For $n = 10$, $t = 0.3$, we obtained the following vector of not-arranged eigenvalues

$$\mu(t) = [0.0075,\ -0.1683,\ -0.4915,\ -0.2763,\ 0.2189,\ 0.2698, „; 0.3038,\ -0.2039,\ -0.3721,\ 0.0683]^T.$$

Since $\mu_j(t) < 0,\ j = 2, 3, 4, 8, 9$, program eig.m delivers a false result without issuing a warning or error code.

For $n = 15$, $t = 0.6$, we obtained the following vector of not-arranged eigenvalues

$$\mu(t) = [0.0037,\ 0.0758,\ -0.3355,\ 0.2519,\ 0.2716,\ 0.1537,\ 0.0616,\ -0.0000,\ 0.0219,\ -0.0291,\ 0.1980,\ -0.4128,\ 0.2169,\ 0.0254,\ -0.0006]^T.$$

Since $\mu_j(t) < 0,\ j = 3, 8, 10, 12, 15$, program eig.m delivers a false result without issuing a warning or error code.

As we see, the computation of $\mu_j(t),\ j = 1, \ldots, n$ by the Matlab routine eig.m is not only costly since it computes all eigenvalues, but it also delivers false results without any error warning.

n	$t = \underline{t}$	$\mu_n(t)$	$\mu_1(t)$
5	0.004256	0.01148×10^{-17}	4.999
10	0.3300	0.3514×10^{-11}	3.028
15	0.5350	0.1698×10^{-10}	1.868

TABLE 2. Determination of minimal $t = \underline{t}$ in $0 < t < 1$ such that $\mu_n(t)$ can be reliably computed

Consequences of the Computational Results

The computational results of this section shows that the critical point in the determination of the condition number $\kappa_2(t) = \mu_1(t)/\mu_n(t)$ is the smallest eigenvalue $\mu_n(t)$ of $A(t) = A_n(t)$. As a consequence, instead of trying to derive better closed-form lower estimates on $\kappa_2(t)$ than those we have already obtained, the efforts should be laid on the reliable computation of the smallest eigenvalue $\mu_n(t)$ as a function of n and t.

For $n = 5$, 10, 15, we have determined the minimal $t = \underline{t}$ *up to four significant places* such that $\mu_n(t)$ can be computed by the inverse power method. For these values of t, we then determined also $\mu_1(t)$. The results are assembled in Table 2.

11. Conclusions

Starting point of this paper was the aim to derive lower estimates on the condition number $\kappa_2(t)$ of the symmetric Toeplitz sinc matrix $A(t) = A_n(t)$. This is of interest since $\kappa_2(t) \to \infty$. The aim was achieved, but numerical calculations showed that the derived lower estimates significantly underestimate the condition number with growing n and vanishing t. Thus, this finding shifted the effort to the problem of effectively and reliably determining the smallest eigenvalue $\mu_n(t)$ of the symmetric Toeplitz sinc matrix $A(t) = A_n(t)$. It turned out that the inverse power method is most appropriate to do this. The pertinent computational experiments showed, for instance, that for $n = 5$ and $t = 0.1$, $\mu_n(t)$ can be determined by this method. But, for $n = 10$ and $t = 0.1$ and $t = 0.2$, this was not possible. However, for $n = 10$ and $t = 0.3$, the inverse power method was successful in determining $\mu_n(t)$. For $n = 15$ and $t = 0.1, \ldots, 0.5$, again $\mu_n(t)$ could not be determined, but for $n = 15$ and $t = 0.6$, this was possible. These results were somehow surprising since, for example, $t = 0.6$ is not near zero so that the problems begin (depending on n) with much larger values of t than we thought. The reason for the numerical problems are, of course, that the computations are done with a restricted number of digital places of the used machine numbers as opposed to the computation with real numbers that have an unlimited number of places. Comparative computations with the Matlab routine eig.m delivered false results for all the mentioned pairs (n,t) since some of the eigenvalues were negative, which cannot be correct because $A(t)$ is positive definite. The most important implication of all these results is that, for calculations with machine numbers (i.e., on computers), priority should be given to the determination of the lower bound $\underline{t} := \inf t$ of the parameter t such that $\mu_n(t)$ can be reliably computed for $0 < \underline{t} = \inf t \le t < 1$. This was done for $n = 5$, 10, 15 with a precision of four significant places by applying the inverse power method. So, one can also say that, for calculations on computers, the expression $\lim\limits_{\substack{t \to 0 \\ 0<t<1}} \kappa_2(t)$ has to be replaced by $\lim\limits_{\substack{t \to \underline{t} \\ 0<\underline{t}<t<1,\ t \in \mathbb{M}}} \kappa_2(t)$, where $\mathbb{M}$ is the set of machine numbers of the used computer, and further that the problems begin already with around $n = 15$ in the sense that with $n = 15$, the minimal value $t = \underline{t}$ reads $\underline{t} = 0.5350 > 0.5$ indicating that problems must be expected when using machine numbers, i.e., when using a computer in calculations involving the Tsinc matrix for

$n \geq 15$ such as the solution of a system of linear equations. The calculations were carried out in *single precision*. Corresponding computations in double precision might deliver better results, but were not done because we think that this would not give new insight in the problem. For information on the effects of finite precision arithmetic on numerical algorithms, the reader is referred to [2, Chapters 1 and 2] or [6, Sections 13 and 14]. We mention that in the English translation of the First Edition [5], Sections 13 and 14 are not yet contained.

Acknowledgements
The authors would like to thank one of the referees for his/her valuable hints to improve the paper.

References

[1] D. Hertz: *Simple Bounds on the Extreme Eigenvalues of Toeplitz Matrices*, IEEE Transactions on Information Theory, **38** (1) (1992), 175–176.
[2] N. J. Higham: *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia (1996).
[3] G. H. Goloub, Ch. F. van Loan: *Matrix Computations*, The Johns Hopkins University Press, Baltimore and London (1989).
[4] F. Stummel: *Diskrete Approximation linear Operatoren*. II (Discrete Approximation of Linear Operators. II). Math. Z., **120** (1971), 231–264.
[5] F. Stummel, K. Hainer: *Introduction to Numerical Analysis (English Translation by E.R. Dawson of the First Edition of the German Original of 1971)*, Scottish Academic Press, Edinburgh (1980).
[6] F. Stummel, K. Hainer: *Praktische Mathematik (Introduction to Numerical Analysis)*, Second Edition, B.G. Teubner, Stuttgart (1982).
[7] F. Stummel, L. Kohaupt: *Eigenwertaufgaben in Hilbertschen Räumne. Mit Aufgaben und vollständigen Lösungen, (Eigenvalue Problems in Hilbert spaces. With Exercises and Complete Solutions)*, Logos Verlag, Berlin (2021).
[8] J. H. Wilkinson: *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford (1965).
[9] Y. Wu: *On the positiveness of a functional symmetric matrix used in digital filter design*, Journal of Circuits, Systems, and Computers, **13** (5) (2004), 1105–1110.
[10] Y. Wu, D. H. Mugler: *A robust DSP integrator for acceleration signals*, IEEE Transactions on Biomedical Engineering, Vol. **51** (2) (2004), 385–389.
[11] Y. Wu, N. Sepehri: *Interpolation of bandlimited signals from uniform or non-uniform integral samples*, Electronic Letters, **47** (1) (2011), 6th Jan.

Ludwig Kohaupt
Berlin University of Technology (BHT)
Department of Mathematics
Luxemburger Str. 10, 13353 Berlin, Germany
ORCID: 0000-0003-4364-9144
E-mail address: lkohaupt4@web.de

Yan Wu
Georgia Southern University
Department of Mathematical Sciences
65 Georgia Ave, Statesboro, GA, 30460, USA
ORCID: 0000-0002-7202-8980
E-mail address: yan@georgiasouthern.edu