4-2007

Perturbations of Roots under Linear Transformations of Polynomials

Branko Ćurgus
Western Washington University, branko.curgus@wwu.edu

Vania Mascioni
Ball State University

PERTURBATIONS OF ROOTS UNDER LINEAR TRANSFORMATIONS OF POLYNOMIALS

BRANKO ĆURGUS AND VANIA MASCIONI

ABSTRACT. Let $\mathcal{P}_n$ be the complex vector space of all polynomials of degree at most n. We give several characterizations of the linear operators $T : \mathcal{P}_n \to \mathcal{P}_n$ for which there exists a constant $C > 0$ such that for all nonconstant $f \in \mathcal{P}_n$ there exist a root u of f and a root v of Tf with $|u - v| \leq C$. We prove that such perturbations leave the degree unchanged and, for a suitable pairing of the roots of f and Tf, the roots are never displaced by more than a uniform constant independent on f. We show that such "good" operators T are exactly the invertible elements of the commutative algebra generated by the differentiation operator. We provide upper bounds in terms of T for the relevant constants.

1. INTRODUCTION

Let n be a positive integer, and denote by $\mathcal{P}_n$ the $(n+1)$-dimensional complex vector space of all polynomials of degree at most n. Let T be a linear operator from $\mathcal{P}_n$ to $\mathcal{P}_n$. In [1] we proved that for each non-constant polynomial $f \in \mathcal{P}_n$ the polynomials f and Tf have at least one common root if and only if T is a non-zero constant multiple of the identity on $\mathcal{P}_n$. In other words, if T is not a multiple of the identity, then there exists a polynomial $f \in \mathcal{P}_n$ such that f and Tf do not share any roots. A natural question to ask is: How far apart are the roots of Tf from the roots of f?

This requires that we introduce a measure of distance between finite subsets of the complex plane $\mathbb{C}$. In Section 8 we introduce four such distances, among which are two common ones: d_H, the Hausdorff distance and d_F, the Fréchet distance. The main result of this article is the characterization of the set of those linear operators $T : \mathcal{P}_n \to \mathcal{P}_n$ for which there exists a constant $C > 0$ such that for all $f \in \mathcal{P}_n$, the distance between the roots of polynomials f and Tf is at most C. Here the distance can be any of the four distances that we introduce, which implies that this set of "good" operators will turn out not to depend on the distance used.

A simple example of a "bad" operator is the operator $R : \mathcal{P}_n \to \mathcal{P}_n$ which changes the sign of the independent variable, defined by

$$(Rf)(z) := f(-z), \quad z \in \mathbb{C}, \quad f \in \mathcal{P}_n.$$

2000 *Mathematics Subject Classification.* Primary: 30C15, Secondary: 26C10.
Key words and phrases. roots of polynomials, linear operators.

If the roots of f are large positive numbers, then the roots of Rf are negative numbers with large moduli, making any of the distances that we consider as large as we want.

To illustrate our result let α be a complex number and consider the linear operator $S(\alpha) : \mathcal{P}_n \to \mathcal{P}_n$ corresponding to the additive shift of the independent variable. It is defined by

$$\big(S(\alpha)f\big)(z) := f(z+\alpha), \quad z \in \mathbb{C}, \quad f \in \mathcal{P}_n. \tag{1.1}$$

It will be quite clear that (with respect to any of the four distances) the distance between the roots of f and the roots of $S(\alpha)f$ will be at most $|\alpha|$ (see Proposition 9.4). The Taylor formula at z implies that the operator $S(\alpha)$ can be expressed as

$$S(\alpha) = I + \frac{\alpha}{1!}\,D + \frac{\alpha^2}{2!}\,D^2 + \cdots + \frac{\alpha^n}{n!}\,D^n,$$

where $D : \mathcal{P}_n \to \mathcal{P}_n$ is the operator of differentiation with respect to the complex variable. This example hints at the main result of this article stated in Theorems 7.3 and 11.1. We paraphrase it below.

Let $T \in \mathcal{L}(\mathcal{P}_n)$, $T \neq 0$, and let $Z(f)$ denote the set of the roots of a non-constant $f \in \mathcal{P}_n$. The following statements are equivalent.

(i) There exists a constant $C > 0$, which depends on the distance d, such that $d\big(Z(f), Z(Tf)\big) \leq C$ for each non-constant $f \in \mathcal{P}_n$.

(ii) There exist $a_0, a_1, \ldots, a_n \in \mathbb{C}$, $a_0 \neq 0$, such that

$$T = a_0\,I + a_1\,D + a_2\,D^2 + \cdots + a_n\,D^n. \tag{1.2}$$

In (i) the symbol d can be replaced with any of the distances d_m, d_h, d_H, d_F from Section 8. Thus (i) really stands for four equivalent statements. Moreover, for T described in (ii) and for each of the distances, in Theorem 10.4 we give an estimate for the maximum possible distance between the roots of f and the roots of Tf in terms of T.

Surprisingly, we found only one article, [14], which considers the relationship between (i) and (ii) as stated above. In [14] an entirely different method was used to prove that (ii) implies (i) with the distance $d = d_F$. The converse was not considered in [14]. Also, no specific estimate for C is given there, which is in part due to the use of "soft" theorems from complex function theory.

It is not surprising, though, that the location of the roots of Tf in relation to the roots of f for T as in (ii) has been extensively researched, see [12, Sections 5.3 and 5.4]. In fact, the implication (ii)⇒(i), with the distance $d = d_h$, is a consequence of Grace's theorem, [12, Theorem 5.3.1]. For completeness we include the details in Sections 2 and 5 below. Furthermore, [12, Corollary 5.4.1] is fundamental for the proof that (ii) implies (i) with the distance $d = d_F$.

The article is organized in twelve short sections; the first section being this introduction. In Section 2 we recall Grace's theorem and one of its consequences. This consequence is of interest to us since in Section 5 we

restate it in terms of operators on $\mathcal{P}_n$. To this end, in Section 3 we study the algebra of operators given by (1.2) and a connection between this algebra and $\mathcal{P}_n$ is explored in Section 4. Since the proofs in Sections 3 and 4 are short and interesting we have not omitted them. In Section 5 we present a version of Grace's theorem for linear operators on $\mathcal{P}_n$. This theorem and a result from Section 6 are the main tools in Section 7, in which the first version of our main result is formulated as Theorem 7.3. This is where we prove the equivalence of (i) and (ii) stated above. In addition, Theorem 7.3, among several equivalent statements, contains the converse of Grace's theorem for linear operators. In Theorem 7.3 we do not use the concepts of distances from Section 8. We wanted to keep the first part of the article, Sections 2 through 7, independent of these concepts.

However, the second part depends heavily on the four distances from Section 8. Why these four distances? The distance d_m is the simplest ("the two closest points distance"), d_h and d_F are implicitly already present in theorems about roots of polynomials, and the Hausdorff distance d_H is probably the simplest distance which is a metric. In Sections 9 and 10, for each distance, we give exact calculations and estimates for the maximum possible distance between the roots of f and the roots of Tf in terms of T. Finally, in Section 11 we present the main theorem, Theorem 11.1. We conclude with several examples in Section 12.

We now introduce the basic notation. By $\deg(f)$ we denote the degree of a polynomial $f \in \mathcal{P}_n$. For a non-zero polynomial $f \in \mathcal{P}_n$, $Z(f) \subset \mathbb{C}$ will denote the multiset of all the roots of f, that is, each root of f appears in $Z(f)$ as many times as its multiplicity as a root of f. Thus $Z(f)$ has exactly $\deg(f)$ elements, and these are not necessarily distinct. The distinction between sets and multisets is essential only when we consider the Fréchet distance d_F. In all other cases $Z(f)$ can be considered simply as the set of roots of f. For completeness we set $Z(0) = \mathbb{C}$.

By $\mathcal{L}(\mathcal{P}_n)$ we denote the set of all linear operators from $\mathcal{P}_n$ to $\mathcal{P}_n$. We shall simply refer to elements in $\mathcal{L}(\mathcal{P}_n)$ as operators. Whenever we need a basis for $\mathcal{P}_n$ we shall use the basis $\{\phi_0, \phi_1, \ldots, \phi_n\}$ (in this listed order) where $\phi_k(z) := z^k/k!$, $k = 0, 1, \ldots, n$.

By $\mathbb{D}(w, r)$ we denote the closed disk in $\mathbb{C}$ centered at $w \in \mathbb{C}$ with radius $r > 0$. The letter z always stands for a complex number. For $A, B \subset \mathbb{C}$ we define $A + B := \{u + v : u \in A, v \in B\}$ and $-A := \{-u : u \in A\}$. By $\operatorname{conv}(A)$ we denote the convex hull of A. Thus, for $f \in \mathcal{P}_n$, $\operatorname{conv}\big(Z(f)\big)$ is the convex hull of the roots of f.

Finally, we thought it would be in reader's interest to try to have as many references as possible pointing to a single source, and the recent monograph by Rahman and Schmeisser [12] is perfectly adapted to the task. Another standard reference in this field is [7].

2. Grace's Theorem

We begin with the definition of the $*$-product of polynomials which appears in [13, page 375] (see also pages 148 and 178 in [12]).

Definition 2.1. Let $f, g \in \mathcal{P}_n$ be polynomials such that $\deg f = \deg g = m > 0$. Let $a_k = f^{(k)}(0)$ and $b_k = g^{(k)}(0)$ for $k = 0, \dots, n$, be the coordinates of f and g with respect to the basis $\{\phi_0, \phi_1, \dots, \phi_n\}$ of $\mathcal{P}_n$. Set

$$(f * g)(z) := \sum_{k=0}^{m} b_k f^{(m-k)}(z) = \sum_{k=0}^{m} a_k g^{(m-k)}(z) = \sum_{k=0}^{m} \left(\sum_{j=0}^{m-k} a_{k+j}\, b_{m-j} \right) \frac{z^k}{k!}.$$

The reader can easily verify (or see [13]) that the three sums that appear in the definition are equal for all $z \in \mathbb{C}$.

Definition 2.1 requires that the polynomials which are being $*$-multiplied have the same degree. The definition depends on the common degree and the $*$-product is a polynomial of the same degree. It is clear that $*$ is commutative.

The next definition is equivalent to the standard one, see [12, Definition 3.3.1]. As before, $(Rf)(z) = f(-z)$.

Definition 2.2. Two polynomials f and g with equal positive degrees are *apolar* if $\big(f * (Rg)\big)(0) = 0$.

The symmetry of the apolarity relation follows from the straightforward equality $g * (Rf) = (-1)^m f * (Rg)$.

The most important result about apolar polynomials is Grace's theorem, see [12, Theorem 3.4.1].

Theorem 2.3 (Grace)**.** *Let f and g be apolar polynomials. If Ω is a circular domain and $Z(g) \subset \Omega$, then $0 \in Z(f) - \Omega$.*

With $S(\alpha)$ as defined in (1.1) we clearly have $Z\big(S(\alpha)p\big) = \{-\alpha\} + Z(p)$ for any polynomial p. Let now f and g be polynomials with the same positive degree. It is easy to verify that $\big(S(\alpha)f\big) * (Rg) = S(\alpha)\big(f * (Rg)\big)$. Combining the last two equalities we conclude that $S(\alpha)f$ and g are apolar if and only if $\alpha \in Z(f * (Rg))$. Applying Grace's theorem to $S(\alpha)f$ and g yields that $Z\big(f * (Rg)\big) \subset Z(f) - \Omega$ whenever Ω is a circular domain and $Z(g) \subset \Omega$. Since $Z(Rg) = -Z(g)$ this leads to the following theorem, see [12, Theorem 5.3.1].

Theorem 2.4. *Let f and g be polynomials with the same positive degree. If Ω is a circular domain and $Z(g) \subset \Omega$, then $Z(f * g) \subset Z(f) + \Omega$.*

For a fixed $g \in \mathcal{P}_n$ with degree n, the mapping $f \mapsto f * g$, $f \in \mathcal{P}_n \backslash \mathcal{P}_{n-1}$, is a restriction of a linear combination of derivatives. We shall explore the relationship between the $*$-product and operators on $\mathcal{P}_n$ further. For that purpose we first study linear combinations of derivatives.

3. The commutative algebra $\mathcal{D}(\mathcal{P}_n)$

The following definition introduces our main object of study.

Definition 3.1. Let $D : \mathcal{P}_n \to \mathcal{P}_n$ be the operator of differentiation on $\mathcal{P}_n$. By $\mathcal{D}(\mathcal{P}_n)$ we denote the linear span in $\mathcal{L}(\mathcal{P}_n)$ of the operators $I, D, \dots, D^n$. If $\alpha_0, \dots, \alpha_n \in \mathbb{C}$ and

$$T = a_0\, I + a_1\, D + \cdots + a_n\, D^n \in \mathcal{D}(\mathcal{P}_n), \tag{3.1}$$

then we write $T = T(a_0, \dots, a_n)$.

To get familiar with the operators in $\mathcal{D}(\mathcal{P}_n)$ we first obtain their matrix representation with respect to the basis of $\mathcal{P}_n$ defined as

$$\big\{\phi_0, \phi_1, \dots, \phi_n\big\}, \qquad \phi_k(z) := z^k/k!, \quad k = 0, \dots, n. \tag{3.2}$$

For each $m \in \{0, \dots, n\}$ we clearly have

$$D^k\phi_m = \phi_{m-k},\ 0 \le k \le m, \quad \text{and} \quad D^k\phi_m = 0,\ m < k \le n. \tag{3.3}$$

Consequently, for T given by (3.1), we have

$$T\phi_m = a_m\phi_0 + a_{m-1}\phi_1 + \cdots + a_1\phi_{m-1} + a_0\phi_m, \quad m = 0, \dots, n, \tag{3.4}$$

and therefore $a_m = \big(T\phi_m\big)(0)$, $m = 0, \dots, n$. Equalities (3.4) imply that the matrix of T with respect to the basis in (3.2) of $\mathcal{P}_n$ is the following upper triangular Toeplitz matrix

$$\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 0 & a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ 0 & 0 & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_0 & a_1 \\ 0 & 0 & 0 & \cdots & 0 & a_0 \end{bmatrix}.$$

Additional basic information about $\mathcal{D}(\mathcal{P}_n)$ is provided in the next three statements.

Proposition 3.2. *Let $T \in \mathcal{L}(\mathcal{P}_n)$. Then $T \in \mathcal{D}(\mathcal{P}_n)$ if and only if T commutes with D.*

Proof. All elements of $\mathcal{D}(\mathcal{P}_n)$ clearly commute with D. To prove the converse, set $\mathcal{C}_n = \big\{T \in \mathcal{L}(\mathcal{P}_n) : TD = DT\big\}$. Clearly $\mathcal{C}_n$ is a subspace of $\mathcal{L}(\mathcal{P}_n)$ and $\mathcal{D}(\mathcal{P}_n) \subset \mathcal{C}_n$. Let $T \in \mathcal{C}_n$. By (3.3), $T\phi_k = TD^{n-k}\phi_n = D^{n-k}T\phi_n$, for $k = 0, \dots, n$. Hence $T \in \mathcal{C}_n$ is uniquely determined by $T\phi_n \in \mathcal{P}_n$. Consequently, the evaluation operator $T \mapsto T\phi_n$, $T \in \mathcal{C}_n$, is an injection. Therefore $\dim \mathcal{C}_n \le \dim \mathcal{P}_n = n + 1$. Since $I, D, \dots, D^n$, are linearly independent elements of $\mathcal{C}_n$, it follows that $\dim\big(\mathcal{C}_n\big) = n + 1$. Consequently $\mathcal{C}_n = \mathcal{D}(\mathcal{P}_n)$. □

Corollary 3.3. *$\mathcal{D}(\mathcal{P}_n)$ is a maximal commutative subalgebra of $\mathcal{L}(\mathcal{P}_n)$.*

The proposition below can be proved in different ways. See, for example, the last paragraph in Section 4. We include this proof since its method is also used in Section 10.

Proposition 3.4. *Let $T \in \mathcal{D}(\mathcal{P}_n)$. The operator T is invertible if and only if $T\phi_0 \neq 0$. If T is invertible, then $T^{-1} \in \mathcal{D}(\mathcal{P}_n)$.*

Proof. The "only if" part of the first statement is obvious. To prove the "if" part assume that $T\phi_0 \neq 0$. First note that since $D \in \mathcal{D}(\mathcal{P}_n)$ is nilpotent, each operator $I - \gamma\, D \in \mathcal{D}(\mathcal{P}_n)$ is invertible and

$$(I - \gamma D)^{-1} = I + \gamma D + \cdots + \gamma^n D^n \in \mathcal{D}(\mathcal{P}_n). \tag{3.5}$$

Now let $T \in \mathcal{D}(\mathcal{P}_n)$ be given by (3.1) and assume $T\phi_0 \neq 0$. Then, by (3.4), $a_0 = T\phi_0 \neq 0$. Following [12, Section 5.4, p. 151], let $\gamma_1, \ldots, \gamma_n$ be the roots of

$$a_0\, z^n + a_1\, z^{n-1} + \cdots + a_{n-1}\, z + a_n$$

counted according to their multiplicities. Then clearly,

$$a_0 + \cdots + a_n\, z^n = a_0 z^n \prod_{j=1}^{n} \big(z^{-1} - \gamma_j\big) = a_0 \prod_{j=1}^{n} (1 - \gamma_j z)\,,$$

and therefore,

$$T = T(a_0, \ldots, a_n) = a_0 \prod_{j=1}^{n} \big(I - \gamma_j\, D\big). \tag{3.6}$$

As a product of invertible operators, T is invertible. Since the inverse of each of its invertible factors is in $\mathcal{D}(\mathcal{P}_n)$, Corollary 3.3 implies $T^{-1} \in \mathcal{D}(\mathcal{P}_n)$. □

4. The algebra $\mathcal{D}(\mathcal{P}_n)$ and the vector space $\mathcal{P}_n$

In this section we explore the relationship between $\mathcal{D}(\mathcal{P}_n)$ and $\mathcal{P}_n$.

Definition 4.1. Define $\varpi : \mathcal{D}(\mathcal{P}_n) \to \mathcal{P}_n$ by

$$\varpi(T) := T\phi_n, \qquad T \in \mathcal{D}(\mathcal{P}_n).$$

If $T = T(a_0, \ldots, a_n)$, then, by (3.4) with $m = n$,

$$\varpi(T) = a_0\, \phi_n + a_1\, \phi_{n-1} + \cdots + a_{n-1}\, \phi_1 + a_n\, \phi_0. \tag{4.1}$$

Results from Section 3 and (4.1) yield the following proposition.

Proposition 4.2. *The operator ϖ is a linear bijection. The image under ϖ of the set of all invertible operators in $\mathcal{D}(\mathcal{P}_n)$ is the set of all polynomials of degree n.*

Since $\mathcal{D}(\mathcal{P}_n)$ is a commutative algebra, it is natural to use ϖ to equip $\mathcal{P}_n$ with an algebra structure. We do that next.

Definition 4.3. The $\star$-product is defined on $\mathcal{P}_n$ by

$$f \star g := \varpi\big(\varpi^{-1}(f)\varpi^{-1}(g)\big), \qquad f, g \in \mathcal{P}_n. \tag{4.2}$$

The properties of $\mathcal{D}(\mathcal{P}_n)$ and ϖ yield the following corollary.

Corollary 4.4. (a) *The vector space $\mathcal{P}_n$ equipped with the $\star$-product is a commutative algebra with unit ϕ_n.*
(b) *The operator $\varpi : \mathcal{D}(\mathcal{P}_n) \to \mathcal{P}_n$ is an algebra isomorphism.*
(c) *The $\star$-invertible polynomials are exactly the polynomials of degree n.*
(d) *For each $m \in \{0, 1, \dots, n-1\}$, the set*

$$\Big\{ f \in \mathcal{P}_n : \deg f = n,\ f^{(k)}(0) = 0,\ k = 0, \dots, n-m-1 \Big\} \tag{4.3}$$

is a $\star$-subgroup of the $\star$-group $\mathcal{P}_n \backslash \mathcal{P}_{n-1}$.

Now we are ready to establish the connection with the operators on $\mathcal{P}_n$.

Proposition 4.5. *Let $T \in \mathcal{D}(\mathcal{P}_n)$. Then,*

$$\begin{aligned} Tf &= f \star \big(\varpi(T)\big) = \big(\varpi(T)\big) \star f, \qquad f \in \mathcal{P}_n, \\ (Tf) \star g &= f \star (Tg) = T(f \star g), \qquad f, g \in \mathcal{P}_n. \end{aligned}$$

Proof. Let $T \in \mathcal{D}(\mathcal{P}_n)$ and $f \in \mathcal{P}_n$. If $f = \varpi(V) = V\phi_n$, then $\varpi^{-1}(f) = V$, and therefore $\big(\varpi^{-1}(f)\big)(\phi_n) = f$. Using successively the commutativity of $\star$, the fact that ϖ is an algebra isomorphism, the definition of ϖ, and the last equality, we calculate

$$f \star \big(\varpi(T)\big) = \varpi(T) \star f = \varpi\big(T\varpi^{-1}(f)\big) = T\big((\varpi^{-1}(f))(\phi_n)\big) = Tf.$$

Now the second claim follows from the associativity of the $\star$-product. □

The definition in (4.2) is convenient since it emphasizes the connection between $\mathcal{D}(\mathcal{P}_n)$ and $\mathcal{P}_n$. However, the formula for the $\star$-product in terms of the coordinates with respect to $\{\phi_0, \dots, \phi_n\}$ is also useful. Let $f, g \in \mathcal{P}_n$ and $a_k = f^{(k)}(0)$ and $b_k = g^{(k)}(0)$ for $k = 0, \dots, n$. Now we first use (4.1) to express $\varpi^{-1}(f)$ and $\varpi^{-1}(g)$ in terms of the coordinates of f and g, then we calculate the composition $\varpi^{-1}(f)\varpi^{-1}(g)$, and again, we use (4.1) to get

$$(f \star g)(z) = \sum_{k=0}^{n} \left(\sum_{j=0}^{n-k} a_{k+j}\, b_{n-j} \right) \frac{z^k}{k!}. \tag{4.4}$$

If $\deg f = \deg g = n$, a comparison of (4.4) and Definition 2.1 with $m = n$, yields that $f * g = f \star g$. More generally, the following proposition holds.

Proposition 4.6. *Let $f, g \in \mathcal{P}_n$, $\deg f = m$ and $\deg g = n$. Then*

$$f * (D^{n-m} g) = f \star g. \tag{4.5}$$

For each $m \in \{1, \dots, n\}$ the set $\mathcal{P}_m \backslash \mathcal{P}_{m-1}$ with the $*$-product is a commutative group. The mapping D^{n-m} restricted to the set (4.3) *is an isomorphism between the commutative group* (4.3) *equipped with the $\star$-product and $\mathcal{P}_m \backslash \mathcal{P}_{m-1}$ with the $*$-product.*

Proof. Set $a_k = f^{(k)}(0)$ and $b_k = g^{(k)}(0)$ for $k = 0, \dots, n$. Now regroup the terms in (4.4) and use $a_{m+1} = \cdots = a_n = 0$ to get a proof of (4.5):

$$f \star g = \sum_{j=0}^{n} a_j g^{(n-j)} = \sum_{j=0}^{m} a_j \big(D^{n-m} g\big)^{(m-j)} = f * D^{n-m} g.$$

The mapping D^{n-m} restricted to the set (4.3) is clearly a bijection between that set and $\mathcal{P}_m \setminus \mathcal{P}_{m-1}$. Let now h and g be polynomials in the set (4.3). Using (4.5) and Proposition 4.5 we calculate

$$(D^{n-m} h) * (D^{n-m} g) = (D^{n-m} h) \star g = D^{n-m}(h \star g).$$

This proves the last claim of the proposition. □

The fact that $\mathcal{P}_m \setminus \mathcal{P}_{m-1}$ with the $*$-product is a commutative group was proved in [13]. Moreover, in [13] the reader can find a nice formula for the inverses.

5. Grace's theorem for linear operators

The next theorem is a restatement of Theorem 2.4 in terms of operators on $\mathcal{P}_n$. The role of g in Theorem 2.4 is now played by an invertible operator $T \in \mathcal{D}(\mathcal{P}_n)$. The union of the sets $Z(T\phi_k)$, $k = 1, \dots, n$, plays the role of $Z(g)$.

Theorem 5.1. *Let $T \in \mathcal{L}(\mathcal{P}_n)$. If $T \in \mathcal{D}(\mathcal{P}_n)$ and T is invertible, then $Z(Tf) \subset Z(f) + \Omega$ for all $f \in \mathcal{P}_n \setminus \{0\}$ and for all circular domains Ω such that $Z(T\phi_k) \subset \Omega$, $k = 1, \dots, n$.*

Proof. Let T be an invertible operator in $\mathcal{D}(\mathcal{P}_n)$. For a constant non-zero f the theorem is obvious. Let f be a non-constant polynomial in $\mathcal{P}_n$. Set $m = \deg f$. Then, by (3.4) and Proposition 3.4, $\deg(Tf) = m$ and $\deg\big(\varpi(T)\big) = \deg\big(T\phi_n\big) = n$. Propositions 4.5 and 4.6 yield

$$Tf = f \star \big(T\phi_n\big) = f * \big(D^{n-m} T\phi_n\big). \tag{5.1}$$

Let Ω be a circular domain such that $Z(T\phi_k) \subset \Omega$ for all $k = 1, \dots, n$. Since by Proposition 3.2 and (3.3), $D^{n-m} T\phi_n = T D^{n-m} \phi_n = T\phi_m$, the theorem follows from (5.1) and Theorem 2.4. □

Corollary 5.2. *Let T be an invertible operator in $\mathcal{D}(\mathcal{P}_n)$ and let Ω be a convex circular domain such that $Z(T\phi_n) \subset \Omega$. Then $Z(Tf) \subset Z(f) + \Omega$ for all $f \in \mathcal{P}_n \setminus \{0\}$.*

Proof. By the Gauss-Lucas theorem, see [12, Theorem 2.1.1], $Z\big(D^{n-k} T\phi_n\big) \subset \operatorname{conv}\big(Z(T\phi_n)\big)$, $k = 1, \dots, n$. Hence, for a convex Ω, $Z(T\phi_n) \subset \Omega$ implies $Z(T\phi_k) \subset \Omega$, $k = 1, \dots, n$, and Theorem 5.1 applies. □

Theorem 5.1 is a motivation for the following definition.

Definition 5.3. An operator $T \in \mathcal{L}(\mathcal{P}_n)$ will be called *Grace operator* if there exists a finite set $A \subset \mathbb{C}$ such that $Z(Tf) \subset Z(f) + \Omega$ for all $f \in \mathcal{P}_n \setminus \{0\}$ and for all circular domains Ω such that $A \subset \Omega$.

The question whether each Grace operator is an invertible operator in $\mathcal{D}(\mathcal{P}_n)$ will be answered by Theorem 7.3.

6. The first step towards the main result

Lemma 6.1. *Let $T \in \mathcal{L}(\mathcal{P}_n)$, $T \neq 0$. Assume that there exists a constant $C > 0$ such that for each non-constant polynomial $f \in \mathcal{P}_n$ there exist $u \in Z(f)$ and $v \in Z(Tf)$ such that $|u - v| \leq C$. Then the matrix of T with respect to the basis $\{\phi_0, \dots, \phi_n\}$ of $\mathcal{P}_n$ is upper triangular and the main diagonal entries are all equal to the same non-zero constant $T\phi_0$.*

Remark 6.2. Note that the hypothesis of Lemma 6.1 implies that $Z(Tf) \neq \emptyset$ for all non-constant $f \in \mathcal{P}_n$. Further, the conclusion of the lemma implies that T maps constant polynomials into constants.

Proof of Lemma 6.1. Let $m \in \{1, \dots, n\}$ and $t > 0$ be arbitrary. Consider the polynomials $\phi_m - (t^m/m!)\phi_0$ and $T\big(\phi_m - (t^m/m!)\phi_0\big)$. By hypothesis these two polynomials have roots which are at most C apart. Thus, for each $t > 0$ there exists an m-th root of unity $\theta(t)$ such that the polynomial $T\phi_m - (t^m/m!)T\phi_0$ has a root $w(t)$ in the disc $\mathbb{D}(t\,\theta(t), C)$. If we assume that $T\phi_0 = 0$, then the last statement would imply $T\phi_m = 0$. Since $m \in \{1, \dots, n\}$ is arbitrary, this would yield $T = 0$. But $T \neq 0$; hence, $T\phi_0 \neq 0$ holds.

Set $v(t) = w(t) - t\,\theta(t)$, $t > 0$. Since $v(t) \in \mathbb{D}(0, C)$ and $\theta(t) \in \mathbb{D}(0, 1)$, for each $k = 0, \dots, n$,

$$(T\phi_k)\big(t\,\theta(t) + v(t)\big) = O\big(t^{\deg(T\phi_k)}\big), \quad t \to +\infty, \tag{6.1}$$

and $\deg(T\phi_k)$ is the smallest power of t for which (6.1) holds. The special case of (6.1), with $k = 0$, implies, for each $m = 0, \dots, n$,

$$(t^m/m!)(T\phi_0)\big(t\,\theta(t) + v(t)\big) = O\big(t^{m+\deg(T\phi_0)}\big), \quad t \to +\infty, \tag{6.2}$$

and $m + \deg(T\phi_0)$ is the smallest power of t for which (6.2) holds. Recall that by the definition of $\theta(t)$ and $v(t)$ we have

$$(T\phi_m)\big(t\,\theta(t) + v(t)\big) = (t^m/m!)(T\phi_0)\big(t\,\theta(t) + v(t)\big), \quad t > 0. \tag{6.3}$$

This, (6.1) and (6.2) imply

$$\deg(T\phi_m) = m + \deg(T\phi_0), \quad m = 0, 1, \dots, n.$$

Since $\deg(T\phi_n) \leq n$, the last equality with $m = n$ implies $\deg(T\phi_0) = 0$; that is, $T\phi_0$ is constant. Consequently,

$$\deg(T\phi_m) = m, \quad m = 0, 1, \dots, n. \tag{6.4}$$

Hence, the matrix of T with respect to the basis $\{\phi_0, \dots, \phi_n\}$ of $\mathcal{P}_n$ is upper triangular.

The main diagonal entries of this matrix are equal to the limits

$$\lim_{z \to \infty} \frac{(T\phi_m)(z)}{\phi_m(z)}, \quad m = 0, 1, \dots, n,$$

existence of which is a consequence of (6.4). Since $v(t) \in \mathbb{D}(0,C)$ and $\theta(t) \in \mathbb{D}(0,1)$, (6.3) implies that these limits all equal to the constant $T\phi_0$. The lemma is proved. □

The following theorem is the first step towards a complete answer to the question posed in the Introduction.

Theorem 6.3. *Let $T \in \mathcal{L}(\mathcal{P}_n)$, $T \neq 0$. Assume that there exists a constant $C > 0$ such that for each non-constant $f \in \mathcal{P}_n$ there exist $u \in Z(f)$ and $v \in Z(Tf)$ such that $|u - v| \leq C$. Then*

$$T = a_0\, I + a_1\, D + \cdots + a_n\, D^n, \tag{6.5}$$

where

$$a_0 = (T\phi_0)(0) \neq 0 \quad \text{and} \quad a_k = (T\phi_k)(0), \quad k = 1, \ldots, n. \tag{6.6}$$

Proof. Let T be as in the hypothesis. Define the coefficients a_k, $k = 0, \ldots, n$, by (6.6). Next we shall prove (6.5).

From Lemma 6.1 we know that the matrix of T with respect to the basis $\{\phi_0, \ldots, \phi_n\}$ of $\mathcal{P}_n$ is upper triangular and the main diagonal entries are all equal to the same non-zero number $a_0 := (T\phi_0)(0)$ (and thus the statement $a_0 \neq 0$ in (6.6) is justified). This implies that T is invertible and

$$T\phi_0 = \big((T\phi_0)(0)\big)\phi_0 = a_0\phi_0. \tag{6.7}$$

(Remember that $\phi_0(z) := 1$ for all $z \in \mathbb{C}$).

Since $\{\phi_0, \ldots, \phi_n\}$ is a basis for $\mathcal{P}_n$, (6.5) is equivalent to

$$T\phi_m = a_0\, \phi_m + a_1\, \phi_{m-1} + \cdots + a_{m-1}\, \phi_1 + a_m\, \phi_0, \quad m = 0, \ldots, n. \tag{6.8}$$

We prove (6.8) by complete induction with respect to m. By (6.7) equality (6.8) holds for $m = 0$. Let $k \in \{1, \ldots, n\}$ and assume that (6.8) is true for $m = 0, \ldots, k-1$. We need to prove that (6.8) holds for $m = k$. By Lemma 6.1, $T\phi_k$ is a polynomial of degree k:

$$\big(T\phi_k\big)(z) = \sum_{j=0}^{k} b_{k-j}\, \phi_j(z)$$

where $b_j \in \mathbb{C}$, $j = 0, \ldots, k$. The rest of the proof is devoted to calculating the coefficients b_j, $j = 0, \ldots, k$.

Let $w \in \mathbb{C}$ be arbitrary and consider the polynomial $p_k(z) := \phi_k(z - w)$. Using the binomial expansion of $(z-w)^k$ and the induction hypothesis, for all $z \in \mathbb{C}$, we get the identity

$$\begin{aligned}\big(Tp_k\big)(z) &= \sum_{j=0}^{k} \frac{1}{j!}(-w)^j \big(T\phi_{k-j}\big)(z) \\ &= \big(T\phi_k\big)(z) + \sum_{j=1}^{k} \frac{1}{j!}(-w)^j \sum_{l=0}^{k-j} a_l \phi_{k-j-l}(z)\end{aligned}$$

$$
\begin{aligned}
&= b_k + \sum_{l=0}^{k-1} b_l\, \phi_{k-l}(z) + \sum_{j=1}^{k} \frac{1}{j!}(-w)^j \sum_{l=0}^{k-j} \frac{a_l}{(k-j-l)!} z^{k-j-l} \\
&= b_k + \sum_{l=0}^{k-1} \frac{b_l}{(k-l)!} z^{k-l} \\
&\qquad + \sum_{l=0}^{k-1} \frac{a_l}{(k-l)!} \left(\sum_{j=1}^{k-l} \binom{k-l}{j} \Big(-\frac{w}{z}\Big)^j z^{k-l} \right) \\
&= b_k + \sum_{l=0}^{k-1} \left[\frac{b_l}{(k-l)!} + \frac{a_l}{(k-l)!} \sum_{j=1}^{k-l} \binom{k-l}{j} \Big(-\frac{w}{z}\Big)^j \right] z^{k-l}.
\end{aligned}
$$

By hypothesis, for each $w \in \mathbb{C}$ there exists $u(w) \in Z(Tp_k)$ such that $|w - u(w)| \le C$. Put $v(w) = u(w) - w$ and note that $|v(w)| \le C$ for all $w \in \mathbb{C}$. The substitution $z = u(w) = w + v(w)$ in the last long displayed identity yields

$$
b_k + \sum_{l=0}^{k-1} \frac{1}{(k-l)!} \left[b_l + a_l \sum_{j=1}^{k-l} \binom{k-l}{j} \Big(\frac{-w}{w+v(w)}\Big)^j \right] \big(w + v(w)\big)^{k-l} = 0,
$$

which simplifies to

$$
b_k + \sum_{l=0}^{k-1} \frac{1}{(k-l)!} \left[b_l + a_l \left(\Big(\frac{v(w)}{w+v(w)}\Big)^{k-l} - 1 \right) \right] \big(w + v(w)\big)^{k-l} = 0, \tag{6.9}
$$

using

$$
\sum_{j=1}^{k-l} \binom{k-l}{j} \left(-\frac{w}{w+v(w)}\right)^j = \left(1 - \frac{w}{w+v(w)}\right)^{k-l} - 1.
$$

Regrouping terms in (6.9) yields

$$
b_k + \sum_{l=0}^{k-1} \frac{b_l - a_l}{(k-l)!} \big(w + v(w)\big)^{k-l} + \sum_{l=0}^{k-1} \frac{a_l}{(k-l)!} v(w)^{k-l} = 0. \tag{6.10}
$$

Since $|v(w)| \le C$ for all $w \in \mathbb{C}$ the last sum in (6.10) is a bounded function of w. Therefore (6.10) implies

$$
\sum_{l=0}^{k-1} \frac{b_l - a_l}{(k-l)!} \big(w + v(w)\big)^{k-l} = O(1), \qquad |w| \to +\infty.
$$

Again, since $|v(w)| \le C$ for all $w \in \mathbb{C}$, the last displayed relation yields

$$
b_l - a_l = 0, \quad l = 0, 1, \ldots, k-1.
$$

Since clearly $b_k = (T\phi_k)(0) = a_k$, we have proved that (6.8) holds for $m = k$. By induction, (6.8) holds for all $m = 0, 1, \ldots, n$, and the theorem is proved. □

7. The first version of the main theorem

Let f be a non-constant polynomial. The number

$$\varrho[f] := \max\big\{|u| : u \in Z(f)\big\}$$

is called the *root radius* of f.

Proposition 7.1. *Let T be an in invertible operator in $\mathcal{D}(\mathcal{P}_n)$. Then for every non-constant $f \in \mathcal{P}_n$ and for each $v \in Z(Tf)$ there exists $u \in Z(f)$ such that $|v-u| \le \varrho[T\phi_n]$.*

Proof. Since the proposition is trivial for a non-zero constant multiple of the identity operator, we assume that $T \neq a_0 I$. Then, by (3.4), $\varrho[T\phi_n] > 0$. Let $f \in \mathcal{P}_n$ be a non-constant polynomial. Since $\mathbb{D}\big(0, \varrho[T\phi_n]\big)$ is a convex circular domain and $Z(T\phi_n) \subset \mathbb{D}\big(0, \varrho[T\phi_n]\big)$, Corollary 5.2 yields

$$Z(Tf) \subset Z(f) + \mathbb{D}\big(0, \varrho[T\phi_n]\big).$$

Thus, for each non-constant $f \in \mathcal{P}_n$, for every $v \in Z(Tf)$ there exists $u \in Z(f)$ such that $|u-v| \le \varrho[T\phi_n]$. □

Remark 7.2. The conclusion of Proposition 7.1 can also be expressed as

$$Z(Tf) \subset \bigcup_{u \in Z(f)} \mathbb{D}\big(u, \varrho[T\phi_n]\big).$$

The following theorem is the main result of this section.

Theorem 7.3. *Let $T \in \mathcal{L}(\mathcal{P}_n)$, $T \neq 0$. The following statements are equivalent.*

(a) *T is an invertible operator in $\mathcal{D}(\mathcal{P}_n)$.*
(b) *T is a Grace operator.*
(c) *There exists a constant $C_1 > 0$ such that for each non-constant $f \in \mathcal{P}_n$ and for each $w \in Z(Tf)$ there exists $v \in Z(f)$ such that $|w-v| \le C_1$.*
(d) *There exists a constant $C_2 > 0$ such that for each non-constant $f \in \mathcal{P}_n$ there exist $u \in Z(f)$ and $v \in Z(Tf)$ such that $|v-u| \le C_2$.*
(e) *T is invertible and it commutes with the differentiation operator D.*

Proof. With $A = \bigcup\big\{Z(T\phi_k) : k = 1, \dots, n\big\}$, the implication (a)⇒(b) follows from Theorem 5.1.

The following short proof of (b)⇒(c) is similar to the proof of Proposition 7.1. Assume (b) and let $A \subset \mathbb{C}$ be a finite set from Definition 5.3. Let $C_1 > 0$ be such that $A \subset \mathbb{D}(0, C_1)$. Let $f \in \mathcal{P}_n \setminus \mathcal{P}_0$. Since $\mathbb{D}(0, C_1)$ is a circular domain, by Definition 5.3 we have $Z(Tf) \subset Z(f) + \mathbb{D}(0, C_1)$. Hence, for each $v \in Z(Tf)$ there exists $u \in Z(f)$ such that $v - u \in \mathbb{D}(0, C_1)$. This proves (c).

The implication (c)⇒(d) is obvious and (d)⇒(a) was proved in Theorem 6.3. Since (a)⇔(e) is an immediate consequence of Proposition 3.2, the theorem is proved. □

8. Distances

Recall the following standard definition:

Definition 8.1. A function $d : X \times X \to [0, +\infty)$ is a *metric* on a nonempty set X, if for all $x, y, z \in X$ we have

(a) $d(x, x) = 0$;
(b) $d(x, y) = 0$ implies that $x = y$;
(c) $d(x, y) = d(y, x)$;
(d) $d(x, z) \le d(x, y) + d(y, z)$.

The problem of measuring the distance between two finite sets of points has been considered in several seemingly unrelated areas of research. For a recent account and references see [3].

The best known metric on the family of finite nonempty subsets of $\mathbb{C}$ is the *Hausdorff metric* defined as

$$d_H(A, B) := \max\Big\{\max_{x\in A}\min_{y\in B}|x - y|,\ \max_{x\in B}\min_{y\in A}|x - y|\Big\},$$

where A and B are nonempty finite subsets of $\mathbb{C}$. That the function d_H is really a metric on the family of finite nonempty subsets of $\mathbb{C}$ is a simple exercise.

To connect the Hausdorff metric to Theorem 7.3 we introduce two related functions. We informally call these functions *distances* since they indicate the location of points of one set in relation to the other. For two nonempty finite subsets A and B of $\mathbb{C}$ define

$$d_m(A, B) := \min\big\{|x - y| \ :\ x \in A,\ y \in B\big\},$$
$$d_h(A, B) := \max\big\{d_m\big(A, \{y\}\big) \ :\ y \in B\big\}.$$

Now, the Hausdorff metric can be expressed as

$$d_H(A, B) = \max\big\{d_h(A, B), d_h(B, A)\big\}.$$

Clearly neither of the functions d_m and d_h is a metric. The function d_m satisfies only (a) and (c) and d_h satisfies only (a) and (d) in Definition 8.1. The function d_h is sometimes called asymmetric Hausdorff distance. Both distances d_m and d_h appear implicitly in Theorem 7.3.

The definitions of d_m and d_h can be extended to include the empty set and the entire complex plane, which correspond to $Z(f)$ for $f \in \mathcal{P}_0\backslash\{0\}$ and $Z(0)$. For both $d = d_m$ and $d = d_h$, we set $d(\emptyset, \emptyset) = 0$ and $d(A, \emptyset) = d(\emptyset, A) = +\infty$ whenever $A \neq \emptyset$. For either $A = \mathbb{C}$ or $B = \mathbb{C}$ the original definitions make sense, giving the value 0. It is in this extended sense that d_m, d_h and d_H will be used in the rest of the article.

Another well known metric is the Fréchet metric (see [5] where a similar definition was first introduced, or [4, Chapter 6] where an analogous metric is defined for curves). Let m be a positive integer and put $\mathbb{M} = \{1, \dots, m\}$. By Π_m we denote the set of all permutations of $\mathbb{M}$. For two functions

$u, v : \mathbb{M} \to \mathbb{C}$ we define

$$d_F(u, v) := \min_{\sigma \in \Pi_m} \max_{k \in \mathbb{M}} \big|u(k) - v(\sigma(k))\big|.$$

The function d_F is not a metric on $\mathbb{C}^m$ since it does not satisfy (b) in Definition 8.1. But d_F is a metric on the factor set $\mathbb{C}^m/\!\sim$, where $u \sim v \Leftrightarrow d_F(u, v) = 0$. The elements of the factor set $\mathbb{C}^m/\!\sim$ can be identified with "unordered" m-tuples, that is with multisets of m complex numbers in which a same element can appear more than once. This concept fits well with the sets of roots of polynomials where roots can occur with multiplicities. In this context the Fréchet metric is defined for two multisets of m complex numbers $U = \{u_1, \dots, u_m\}$ and $V = \{v_1, \dots, v_m\}$ by

$$d_F(U, V) := \min_{\sigma \in \Pi_m} \max_{k \in \mathbb{M}} \big|u_k - v_{\sigma(k)}\big|.$$

In some ways this distance is a natural distance when perturbation of the roots of polynomials are studied, see [10, Theorem, p. 276], [6] and [2]. For a simple proof that d_F is a metric see [2].

We are interested in the question of what happens (in a quantitative sense) to the roots of polynomials under linear operators on $\mathcal{P}_n$. The following numbers give a "one number summary" answer to this question for $T \in \mathcal{L}(\mathcal{P}_n)$:

$$K_m(T) := \sup\big\{d_m\big(Z(f), Z(Tf)\big) : f \in \mathcal{P}_n\big\},$$
$$K_h(T) := \sup\big\{d_h\big(Z(f), Z(Tf)\big) : f \in \mathcal{P}_n\big\},$$
$$K_H(T) := \sup\big\{d_H\big(Z(f), Z(Tf)\big) : f \in \mathcal{P}_n\big\}.$$

The power of these definitions is in the fact that the statements from theorems in Sections 6 and 7 can now be formulated in a more compact way. For example the hypothesis of Theorem 6.3 is: $K_m(T) < +\infty$ and the conclusion of Proposition 7.1 is: $K_h(T) \leq \varrho[T\phi_n]$.

Since the Fréchet metric is defined only for multisets with the same number of elements, we can define $K_F(T)$ only for $T \in \mathcal{L}(\mathcal{P}_n)$ with the property $\deg(Tf) = \deg(f)$ for every $f \in \mathcal{P}_n$. For such T we define

$$K_F(T) := \sup\big\{d_F\big(Z(f), Z(Tf)\big) : f \in \mathcal{P}_n \setminus \mathcal{P}_0\big\}.$$

Proposition 8.2. *Let $T, T_1 \in \mathcal{L}(\mathcal{P}_n)$ and let α be a non-zero complex number. Then*

(a) $K_m(T) \leq K_h(T) \leq K_H(T)$.
(b) $K_h(T_1 T) \leq K_h(T) + K_h(T_1)$.
(c) $K_H(T_1 T) \leq K_H(T) + K_H(T_1)$.
(d) *For invertible T,* $K_H(T) = \max\big\{K_h(T), K_h(T^{-1})\big\} = K_H(T^{-1})$.

If $\deg(f) = \deg(Tf) = \deg(T_1 f)$ for all $f \in \mathcal{P}_n$, then

(e) $K_F(T_1 T) \leq K_F(T) + K_F(T_1)$.
(f) $K_H(T) \leq K_F(T)$.
(g) *T is invertible and* $K_F(T) = K_F(T^{-1})$.

Proof. The proofs follow directly from the definitions, properties of the underlying distances and properties of the supremum. A proof of (b) follows:

$$\begin{aligned} K_h(T_1T) &= \sup\{d_h\big(Z(f), Z(T_1Tf)\big) : f \in \mathcal{P}_n\} \\ &\le \sup\{d_h\big(Z(f), Z(Tf)\big) : f \in \mathcal{P}_n\} \\ &\qquad + \sup\{d_h\big(Z(Tf), Z(T_1Tf)\big) : f \in \mathcal{P}_n\} \\ &\le K_h(T) + K_h(T_1). \end{aligned}$$

To prove (d) assume that T is invertible. Then $T^{-1}\mathcal{P}_n = \mathcal{P}_n$ and therefore

$$\begin{aligned} K_h(T^{-1}) &= \sup\{d_h\big(Z(f), Z(T^{-1}f)\big) : f \in \mathcal{P}_n\} \\ &= \sup\{d_h\big(Z(Tg), Z(g)\big) : g \in \mathcal{P}_n\}. \end{aligned}$$

Now the first equality in (d) follows from the definition of $K_H(T)$. The second equality follows from the first when T is substituted by T^{-1}.

The remaining statements are proved similarly. □

9. Exact Calculations

We start with general results for K_h and K_H.

Theorem 9.1. *Let T be an invertible operator in $\mathcal{D}(\mathcal{P}_n)$. Then*

$$K_h(T) = \varrho[T\phi_n] = d_h\big(Z(\phi_n), Z(T\phi_n)\big).$$

Proof. By Proposition 7.1,

$$d_h\big(Z(f), Z(Tf)\big) \le \varrho[T\phi_n], \qquad f \in \mathcal{P}_n \backslash \mathcal{P}_0.$$

Since T maps constants onto constants, it follows that $K_h(T) \le \varrho[T\phi_n]$. Clearly, $d_h\big(Z(\phi_n), Z(T\phi_n)\big) = \varrho[T\phi_n]$, and therefore, $K_h(T) \ge \varrho[T\phi_n]$. □

Proposition 8.2(d) now yields:

Corollary 9.2. *Let T be an invertible operator in $\mathcal{D}(\mathcal{P}_n)$. Then*

$$\begin{aligned} K_H(T) &= \max\{\varrho[T\phi_n], \varrho[T^{-1}\phi_n]\} \\ &= \max\{d_H\big(Z(\phi_n), Z(T\phi_n)\big), d_H\big(Z(T^{-1}\phi_n), Z(\phi_n)\big)\}. \end{aligned}$$

Remark 9.3. Theorem 9.1 conveys that the worst possible perturbation of roots measured by the distances d_h occurs at the polynomial ϕ_n. That is, for all $f \in \mathcal{P}_n$,

$$d_h\big(Z(f), Z(Tf)\big) \le d_h\big(Z(\phi_n), Z(T\phi_n)\big).$$

For the distance d_H, by Corollary 9.2, the worst possible perturbation of roots occurs either at ϕ_n or at $T^{-1}\phi_n$. That is, for all $f \in \mathcal{P}_n$,

$$d_H\big(Z(f), Z(Tf)\big) \le \max\{d_H\big(Z(\phi_n), Z(T\phi_n)\big), d_H\big(Z(T^{-1}\phi_n), Z(\phi_n)\big)\}.$$

It would be interesting to know whether the last two inequalities must be strict when f is not a multiple of ϕ_n.

Exact calculations are possible only for simple operators in $\mathcal{D}(\mathcal{P}_n)$. We study two such classes. Let $\alpha, \gamma \in \mathbb{C}$. As before, $S(\alpha) \in \mathcal{D}(\mathcal{P}_n)$ is the operator that shifts the independent variable by α defined in (1.1). Further, we define

$$H_k(\gamma) := I - \gamma D^k\,, \quad k = 1, \dots, n.$$

Proposition 9.4. *Let $\alpha \in \mathbb{C}$ and consider $S(\alpha) \in \mathcal{D}(\mathcal{P}_n)$. Then*

$$K_m\big(S(\alpha)\big) = K_h\big(S(\alpha)\big) = K_H\big(S(\alpha)\big) = K_F\big(S(\alpha)\big) = |\alpha|.$$

Proof. For $f \in \mathcal{P}_n$ we have $Z\big(S(\alpha)f\big) = \{-\alpha\} + Z(f)$. Therefore,

$$d_m\big(Z(f), Z(S(\alpha)f)\big) \leq |\alpha| \quad \text{and} \quad d_m\big(Z(\phi_n), Z(S(\alpha)\phi_n)\big) = |\alpha|.$$

Hence $K_m\big(S(\alpha)\big) = |\alpha|$. The same argument can be used for K_F.

Since $\varpi\big(S(\alpha)\big) = S(\alpha)\phi_n$, by Theorem 9.1 we have

$$K_h\big(S(\alpha)\big) = \varrho\big[S(\alpha)\phi_n\big] = |\alpha|.$$

By Proposition 8.2 (d),

$$K_H\big(S(\alpha)\big) = \max\big\{K_h\big(S(\alpha)\big), K_h\big(S(-\alpha)\big)\big\} = |\alpha|. \qquad \square$$

Proposition 9.5. *Let $\gamma \in \mathbb{C}$ and consider $H_k(\gamma) \in \mathcal{D}(\mathcal{P}_n)$ for $k = 1, \dots, n$. Then*

$$K_h\big(H_k(\gamma)\big) = \sqrt[k]{\frac{|\gamma|\, n!}{(n-k)!}}.$$

Proof. By the definition of ϖ, $\varpi\big(H_k(\gamma)\big) = H_k(\gamma)\phi_n$ and

$$\big(H_k(\gamma)\phi_n\big)(z) = \phi_n(z) - \gamma\, \phi_{n-k}(z) = \frac{z^{n-k}}{n!}\left(z^k - \frac{\gamma\, n!}{(n-k)!}\right).$$

Consequently, $\varrho\big[H_k(\gamma)\phi_n\big] = \sqrt[k]{|\gamma|\, n!/(n-k)!}$, and the proposition follows from Theorem 9.1. $\square$

10. Estimates

Let $a_0, \dots, a_n \in \mathbb{C}$, $a_0 \neq 0$, and let $T = T(a_0, \dots, a_n)$ be the corresponding invertible operator in $\mathcal{D}(\mathcal{P}_n)$. In this section we give estimates for the quantities $K_h(T)$, $K_H(T)$ and $K_F(T)$ in terms of the coefficients $a_0, \dots, a_n$.

Proposition 10.1. *Let $\gamma \in \mathbb{C}$. Then*

$$(n!)^{1/n}|\gamma| \leq K_h\big(H_1(\gamma)^{-1}\big) \leq n|\gamma|. \tag{10.1}$$

Proof. To prove the first inequality in (10.1) consider the polynomial

$$g = \gamma^{n-1}\, \phi_1 + \gamma^{n-2}\, \phi_2 + \cdots + \gamma\, \phi_{n-1} + \phi_n.$$

A straightforward computation gives $H_1(\gamma)g = g - \gamma\, g' = \phi_n - \gamma^n$. Hence, g has a root at $z = 0$, while $H_1(\gamma)g$ has all of its roots on the circle of radius

$(n!)^{1/n}|\gamma|$. If we put $f = H_1(\gamma)g$, then $g = H_1(\gamma)^{-1}f$ and the previous observation about the location of the roots implies that

$$d_h\big(Z(f), Z\big(H_1(\gamma)^{-1}f\big)\big) \geq (n!)^{1/n}|\gamma|.$$

This implies the first inequality in (10.1).

To prove the second inequality notice that by (3.5),

$$\big(H_1(\gamma)^{-1}\phi_n\big)(z) = \sum_{k=0}^{n} \gamma^{n-k}\,\phi_k(z) = \gamma^n \sum_{k=0}^{n} \phi_k\big(z/\gamma\big).$$

The roots of $\sum_{k=0}^{n}\phi_k$ have been researched extensively; see for example [11] and the references therein. Here we only need that the root radius of the polynomial $\sum_{k=0}^{n}\phi_k$ is smaller than or equal to n. Consequently, $\varrho\big[H_1(\gamma)^{-1}\phi_n\big] \leq n\,|\gamma|$ and the second inequality in (10.1) follows from Theorem 9.1. □

Corollary 10.2. *Let $\gamma \in \mathbb{C}$. Then*

$$K_H\big(H_1(\gamma)\big) = K_H\big(H_1(\gamma)^{-1}\big) = n|\gamma|.$$

Proof. The corollary follows from Propositions 8.2 (d), 9.5 and 10.1. □

Proposition 10.3. *Let $\gamma \in \mathbb{C}$. Then*

$$K_F\big(H_1(\gamma)\big) \leq n^2|\gamma|.$$

Proof. Let f be a non-constant polynomial in $\mathcal{P}_n$ and $\gamma \in \mathbb{C}$. Consider the set

$$\Omega = \bigcup_{w\in Z(f)} \mathbb{D}\Big(w + \frac{n\,\gamma}{2}, \frac{n\,|\gamma|}{2}\Big).$$

By [12, Corollary 5.4.1(iii)] we have $Z\big(H_1(\gamma)f\big) \subset \Omega$, and in each connected component $\Omega_1, \Omega_2, \ldots, \Omega_k$ of Ω the polynomials f and $H_1(\gamma)f$ have the same number of zeros, counted according to their multiplicities. Therefore

$$d_F\big(Z(f), Z(H_1(\gamma)f)\big) \leq \max\big\{\operatorname{diam}(\Omega_j) : j = 1, \ldots, k\big\}.$$

Since for each $j = 1, \ldots, k$,

$$\operatorname{diam}(\Omega_j) \leq 2n\frac{n\,|\gamma|}{2} = n^2\,|\gamma|,$$

we conclude that $K_F(H_1(\gamma)) \leq n^2\,|\gamma|$. □

Theorem 10.4. *Let $a_0, \ldots, a_n \in \mathbb{C}$, $a_0 \neq 0$. Let $T = T(a_0, \ldots, a_n)$ be the corresponding invertible operator in $\mathcal{D}(\mathcal{P}_n)$. Let $\gamma_1, \ldots, \gamma_n$ be the roots of*

$$a_0\,z^n + a_1\,z^{n-1} + \cdots + a_{n-1}\,z + a_n$$

counted according to their multiplicities. Then

$$K_h(T) \leq n\big(|\gamma_1| + \cdots + |\gamma_n|\big)\,, \tag{10.2}$$

$$K_H(T) \leq n\big(|\gamma_1| + \cdots + |\gamma_n|\big)\,, \tag{10.3}$$

$$K_F(T) \leq n^2\big(|\gamma_1| + \cdots + |\gamma_n|\big). \tag{10.4}$$

Proof. Lemma 6.1 implies that $\deg(Tf) = \deg(f)$ for all $f \in \mathcal{P}_n$. Thus $K_F(T)$ is defined and $T\big(\mathcal{P}_n \setminus \mathcal{P}_0\big) = \mathcal{P}_n \setminus \mathcal{P}_0$. By repeated application of statements (b),(c), and (e) of Proposition 8.2 to (3.6) we get

$$K(T) \leq K\big(H_1(\gamma_1)\big) + \cdots + K\big(H_1(\gamma_n)\big),$$

where K can be either K_h, K_H, or K_F. Now (10.2) follows from Proposition 9.5, (10.3) follows from Corollary 10.2, and (10.4) follows from Proposition 10.3. □

Remark 10.5. The problem of estimating the sum of the absolute values of all the roots of a given polynomial was studied by Berwald, see [8, Theorem 2.3].

Proposition 10.6. *Let T be an invertible operator in $\mathcal{D}(\mathcal{P}_n)$. Then*

$$K_H(T) \leq K_F(T) \leq (e\, n^3 \ln n)\, K_H(T). \tag{10.5}$$

Proof. The first inequality was proved in Proposition 8.2. Let us now fix $T \in \mathcal{D}(\mathcal{P}_n)$ and assume that

$$T = I + \frac{\alpha_1}{1!}\, D + \cdots + \frac{\alpha_{n-1}}{(n-1)!}\, D^{n-1} + \frac{\alpha_n}{n!}\, D^n.$$

By the known estimate [12, (8.1.12)], all the roots γ_j of the polynomial

$$z^n + \alpha_1\, z^{n-1} + \cdots + \frac{\alpha_{n-1}}{(n-1)!}\, z + \frac{\alpha_n}{n!}$$

are smaller than $\sum_j \big(|\alpha_j|/j!\big)^{1/j}$. By (10.4) we then have

$$K_F(T) \leq n^3 \sum_{j=1}^{n} \left(\frac{|\alpha_j|}{j!}\right)^{1/j}. \tag{10.6}$$

On the other hand, by Theorem 9.1 we have

$$K_h(T) = d_h\big(Z(\phi_n), Z(T(\phi_n))\big) = \varrho[T\phi_n]. \tag{10.7}$$

A lower estimate for $\varrho[T\phi_n]$, where

$$\big(T\phi_n\big)(z) = \frac{1}{n!}\left(\alpha_n + \binom{n}{1}\alpha_{n-1} z + \ldots + \binom{n}{n-1}\alpha_1 z^{n-1} + z^n\right),$$

is given by another classical inequality (see [12, (8.1.1)]):

$$\max_{1\leq j\leq n} |\alpha_j|^{1/j} \leq \varrho[T\phi_n]. \tag{10.8}$$

Combining (10.6), (10.8) and (10.7) and letting $\mu = \sum_{j=1}^{n} (j!)^{-1/j}$ gives

$$K_F(T) \leq n^3 \mu \max_{1\leq j\leq n} |\alpha_j|^{1/j} \leq n^3 \mu\, K_h(T) \leq n^3 \mu\, K_H(T).$$

To estimate μ, note that for $j > 1$ we have

$$\left(\frac{1}{j!}\right)^{1/j} < \frac{e}{j} < \frac{e}{j - \frac{1}{2}} < e \ln \frac{j}{j-1},$$

where the first inequality easily follows from Stirling's approximation [9, p. 183], while the last one is a special case of [9, 3.6.17]. Since $1+\frac{1}{\sqrt{2}} < e \ln 2$, adding up gives

$$\mu = \sum_{j=1}^{n} \left(\frac{1}{j!}\right)^{1/j} < e \ln n$$

and thus, finally, $K_F(T) \le (e\, n^3 \ln n) K_H(T)$ as we needed to prove. □

The factor $e\, n^3 \ln n$ in the second inequality of (10.5) is plausibly far from being best possible, but we have not pursued this line of enquiry.

11. The main theorem

We can now finally give a comprehensive and compact answer to the question posed in the Introduction.

Theorem 11.1. *Let $T \in \mathcal{L}(\mathcal{P}_n)$, $T \neq 0$. The following statements are equivalent.*

(a) *$T \in \mathcal{D}(\mathcal{P}_n)$ and T is invertible.*
(b) *T is a Grace operator.*
(c) *$K_h(T) < +\infty$.*
(d) *$K_m(T) < +\infty$.*
(e) *T is invertible and $TD = DT$.*
(f) *$K_H(T) < +\infty$.*
(g) *$\deg(Tf) = \deg(f)$ for all $f \in \mathcal{P}_n$ and $K_F(T) < +\infty$.*

Proof. The statements (a) through (e) are equivalent by Theorem 7.3.

The implication (a)⇒(g) follows from Proposition 3.4, equalities (3.4) and Theorem 10.4. Proposition 8.2 (f) yields (g)⇒(f) and (f)⇒(c) follows from Proposition 8.2 (a). The theorem is proved. □

As $\mathcal{D}(\mathcal{P}_n)$ is a commutative subalgebra of $\mathcal{L}(\mathcal{P}_n)$ we have the following corollary.

Corollary 11.2. *If $T, T_1 \in \mathcal{L}(\mathcal{P}_n)\backslash\{0\}$ satisfy any of the equivalent conditions* (a)-(g) *in Theorem* 11.1*, then T and T_1 commute.*

Remark 11.3. By Proposition 3.4 the statement (a) in Theorem 11.1 is equivalent to T^{-1} is invertible and $T^{-1} \in \mathcal{D}(\mathcal{P}_n)$. Therefore, the statements about the operator T in Theorem 11.1 are equivalent to the corresponding statements about the operator T^{-1}.

12. Examples

In the next two examples we consider the impact of the operators from $\mathcal{D}(\mathcal{P}_n)$ on the roots of polynomials in $\mathcal{P}_n$ for $n = 1, 2$, where the computations can be carried out in all detail. While the situation for $n = 1$ is as trivial as expected, the case $n = 2$ gives a good insight into why direct calculations for larger matrices are bound to be unwieldy.

Example 12.1. In $\mathcal{P}_1$ we have

$$T(a_0,a_1)(\phi_1 - w\phi_0) = a_0\phi_1 - (a_0 w - a_1)\phi_0.$$

Provided that $a_0 \neq 0$, the polynomial $T(a_0,a_1)(\phi_1 - w\phi_0)$ has a root at $w - a_1/a_0$. Thus, the operator $T(a_0,a_1)$ shifts all the roots by exactly a_1/a_0. This also follows from $T(a_0,a_1) = a_0 S(a_1/a_0)$; see (1.1).

Example 12.2. Let w_1, w_2 be arbitrary complex numbers. For the operator $T(a_0,a_1,a_2) \in \mathcal{D}(\mathcal{P}_2)$ with $a_0 \neq 0$ we have

$$\begin{aligned}\Big(T(a_0,a_1,a_2)\big((\phi_1 - w_1\phi_0)(\phi_1 - w_2\phi_0)\big)\Big)(z)& \\ = a_0\left(z^2 + 2\left(\frac{a_1}{a_0} - \frac{w_1+w_2}{2}\right)z + w_1 w_2 - \frac{a_1}{a_0}(w_1+w_2) + \frac{a_2}{a_0}\right)&\end{aligned}$$

and so the roots of $T(a_0,a_1,a_2)\big((\phi_1 - w_1\phi_0)(\phi_1 - w_2\phi_0)\big)$ are

$$\frac{w_1+w_2}{2} - \frac{a_1}{a_0} \pm \sqrt{\left(\frac{w_1-w_2}{2}\right)^2 + \left(\frac{a_1}{a_0}\right)^2 - \frac{a_2}{a_0}}\,. \tag{12.1}$$

In this case we see that, as predicted by our main result, *both* of the new roots are uniformly close to the original w_1, w_2. Also, we see that if all the constants are real and $a_1^2 \geq a_0 a_2$ then an invertible operator in $\mathcal{D}(\mathcal{P}_2)$ "sends real roots into real roots". To check the statement on the uniform displacement of the roots, do as follows:

Call the roots in (12.1) z_1 (with $+$) and z_2 (with $-$). Since we are dealing with complex numbers, let us think of the square root as having one definite value (then the $\pm$ takes care of the second root). Also, to simplify the algebra, define

$$\delta_1 := \frac{a_1}{a_0} \qquad \text{and} \qquad \delta_2 := \left(\frac{a_1}{a_0}\right)^2 - \frac{a_2}{a_0}.$$

So, for instance, we have

$$|z_1 - w_1| \leq |\delta_1| + \left|\sqrt{\left(\frac{w_1-w_2}{2}\right)^2 + \delta_2} - \frac{w_1-w_2}{2}\right|$$

and similarly we can estimate $z_2 - w_1$. Now, we have

$$\left(\sqrt{\left(\frac{w_1-w_2}{2}\right)^2 + \delta_2} - \frac{w_1-w_2}{2}\right)\left(\sqrt{\left(\frac{w_1-w_2}{2}\right)^2 + \delta_2} + \frac{w_1-w_2}{2}\right) = \delta_2$$

and so we see that at least one of the factors on the left hand side has modulus less than or equal to $\sqrt{|\delta_2|}$. By what noted above, this means that

$$\min\big\{|z_1 - w_1|, |z_2 - w_1|\big\} \leq |\delta_1| + \sqrt{|\delta_2|}$$

which (together with the similar estimates derived with w_2 instead of w_1) translates into the statement that the roots of

$$T(a_0, a_1, a_2)\big((\phi_1 - w_1\phi_0)(\phi_1 - w_2\phi_0)\big)$$

are not farther away from w_1 and w_2 than the uniform quantity

$$\left|\frac{a_1}{a_0}\right| + \sqrt{\left|\left(\frac{a_1}{a_0}\right)^2 - \frac{a_2}{a_0}\right|}.$$

Example 12.3. This example shows that $K_H(T) < K_F(T)$. It was found with the help of Mathematica. We consider $\mathcal{P}_3$ and the operator T and its inverse given by

$$T = I + \frac{2}{3}D + \frac{2}{9}D^2 - \frac{4}{27}D^3, \quad T^{-1} = I - \frac{2}{3}D + \frac{2}{9}D^2 + \frac{4}{27}D^3.$$

Set $f(z) = (z-1)^2(z+1)$ and $g(z) = (z+1)^2(z-1)$. Then $Tf = g$ and $d_F(Z(f), Z(g)) = 2$. Mathematica calculated that

$$\varrho\big[T\phi_3\big] = \varrho\big[T^{-1}\phi_3\big] = \frac{2}{3}\big(1 + \sqrt[3]{2}\big) \approx 1.506614. \tag{12.2}$$

Therefore

$$K_H(T) = \frac{2}{3}\big(1 + \sqrt[3]{2}\big) < 2 \leq K_F(T)$$

Since both $d_F\big(Z(\phi_3), Z(T\phi_3)\big)$ and $d_F\big(Z(\phi_3), Z(T^{-1}\phi_3)\big)$ are equal to the number in (12.2), this example also shows $K_F(T)$ is not related to ϕ_3 in the sense of Corollary 9.2. We leave it as an open problem to explore the existence and uniqueness of a polynomial f such that $K_F(T) = d_F\big(Z(f), Z(Tf)\big)$.

Example 12.4. Several easy examples of "bad behavior" from an operator T which does not satisfy condition (a) in Theorem 7.3 can be obtained as follows.

Define T by $T(f) := f + f(0)\phi_n$ for $f \in \mathcal{P}_n$. Then T is clearly linear, invertible and and does not commute with D. Thus, by Proposition 3.2, $T \notin \mathcal{D}(\mathcal{P}_n)$. If $w \in \mathbb{C} \setminus \{0, 1\}$ is arbitrary, then $Z\big(T(\phi_n - w\phi_0)\big)$ consists of all the n-th roots of $n!w/(1-w)$, while all the roots of $\phi_n - w\phi_0$ have modulus $\sqrt[n]{|w|}$, thus ensuring that the distance (with respect to each of the distances introduced in Section 8) between the two multisets of roots can be as large as desired (just let w approach 1).

The operator D belongs to $\mathcal{D}(\mathcal{P}_n)$ but it is not invertible. With the same polynomials as above we have $Z\big(D(\phi_n - w\phi_0)\big) = \{0, \ldots, 0\}$ ($n-1$ zeros) again showing that the distance between $Z\big(D(\phi_n - w\phi_0)\big)$ and $Z(\phi_n - w\phi_0)$ (with respect to each of the distances, except d_F which is not defined) can be as large as we wish.

The operator defined by $T(f) = f + (a_{n-1}/n)\phi_0$ for $f(z) = a_0 + a_1 z + \cdots + a_n z^n$, is clearly linear, invertible and its matrix with respect to the basis $\{\phi_0, \ldots, \phi_n\}$ of $\mathcal{P}_n$ is upper triangular. This operator does not belong $\mathcal{D}(\mathcal{P}_n)$

since it does not commute with D. Polynomials that show that none of (d), (c), (f) and (g) of Theorem 11.1 hold are $(\phi_1 - w\phi_0)^n$, $w \in \mathbb{C}$. Namely, the roots of $T\big((\phi_1 - w\phi_0)^n\big)$ are on the circle centered at w with radius $\sqrt[n]{|w|}$, and $Z\big((\phi_1 - w\phi_0)^n\big) = \{w, \dots, w\}$ (n times).

Acknowledgment The authors thank the referees for pointing out several related references, the article [13] in particular. This resulted in numerous improvements.

References

[1] Ćurgus, B., Mascioni, V.: Root preserving transformations of polynomials. To appear in Math. Mag.
[2] Ćurgus, B., Mascioni, V.: Roots and polynomials as homeomorphic spaces. Expo. Math. 24 (2006), 81–95.
[3] Eiter, T., Mannila H.: Distance measures for point sets and their applications. Acta Informatica 34 (1997), 109–133.
[4] Ewing, G. M.: *Calculus of variations with applications*. Corrected reprint of the 1969 original. Dover Publications, 1985.
[5] Fréchet, M.: Sur quelques points du calcul fonctionnel. Rend. Circ. Mat. Palermo 22 (1906), 1–74.
[6] Krause, G. M.: Bounds for the variation of matrix eigenvalues and polynomial roots. Linear Algebra Appl. 208/209 (1994), 73–82.
[7] Marden, M.: *Geometry of polynomials* . Second edition reprinted with corrections, American Mathematical Society, 1985.
[8] Milovanović, G. V., Rassias, T. M.: Inequalities for polynomial zeros. Survey on classical inequalities, 165–202, Math. Appl., 517, Kluwer Acad. Publ., 2000.
[9] Mitrinović, D. S.: *Analytic Inequalities*, Springer-Verlag 1970.
[10] Ostrowski, A. M.: *Solution of equations in Euclidean and Banach spaces*. Third edition of *Solution of equations and systems of equations*. Academic Press, 1973.
[11] Pritsker, I. E., Varga, R. S.: The Szegö curve, zero distribution and weighted approximation. Trans. Amer. Math. Soc. 349 (1997), 4085–4105.
[12] Rahman, Q. I., Schmeisser, G.: *Analytic theory of polynomials*, Oxford University Press, 2002.
[13] Specht, W.: Die Lage der Nullstellen eines Polynoms III, Math. Nachr. 16 (1957), 369–389.
[14] Tulovsky, V.: On perturbations of roots of polynomials. J. Analyse Math. 54 (1990), 77–89.

Department of Mathematics, Western Washington University,
Bellingham, WA 98225, USA
E-mail address: curgus@cc.wwu.edu

Department of Mathematical Sciences, Ball State University,
Muncie, IN 47306-0490, USA
E-mail address: vdm@cs.bsu.edu