12-2010

Stability of Roots of Polynomials Under Linear Combinations of Derivatives

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

Vania Mascioni

STABILITY OF ROOTS OF POLYNOMIALS UNDER LINEAR COMBINATIONS OF DERIVATIVES

BRANKO ĆURGUS AND VANIA MASCIONI

ABSTRACT. Let $T = \alpha_0 I + \alpha_1 D + \cdots + \alpha_n D^n$, where D is the differentiation operator and $\alpha_0 \neq 0$, and let f be a square-free polynomial with large minimum root separation. We prove that the roots of Tf are close to the roots of f translated by $-\alpha_1/\alpha_0$.

1. Introduction

Let n be a positive integer. Denote by $\mathcal{P}_n$ the $(n+1)$-dimensional complex vector space of all polynomials of degree at most n. By $\mathcal{P}_0$ denote the set of all constant polynomials. For a non-zero $f \in \mathcal{P}_n$ let

$$Z(f) = \{w \in \mathbb{C} : f(w) = 0\}$$

be the set of the roots of f, where $\mathbb{C}$ is the field of complex numbers. It is useful to extend the field operations from $\mathbb{C}$ onto the nonempty subsets of $\mathbb{C}$ in the following standard way: for $A, B \subset \mathbb{C}$ we set

$$A + B = \{u + v : u \in A, v \in B\} \quad \text{and} \quad A\,B = \{uv : u \in A, v \in B\}.$$

For $r > 0$ we set $\mathbb{D}(r) = \{z \in \mathbb{C} : |z| \leq r\}$. For example, the set $Z(f) + \mathbb{D}(r)$ is the union of the closed disks of radius r centered at the roots of f.

Let $\mathcal{L}(\mathcal{P}_n)$ be the set of all linear operators from $\mathcal{P}_n$ to $\mathcal{P}_n$. How does an operator $T \in \mathcal{L}(\mathcal{P}_n)$ perturb the roots of polynomials? To illustrate what we mean by this question consider two simple linear operators on $\mathcal{P}_n$. Let $\alpha, t \in \mathbb{C}$, $t \neq 0$. For $f \in \mathcal{P}_n$ we set

$$\big(S(\alpha)f\big)(z) := f(\alpha + z),$$
$$\big(H(t)f\big)(z) := f(z/t).$$

Then for all non-constant $f \in \mathcal{P}_n$ we clearly have

$$Z\big(S(\alpha)f\big) = \{-\alpha\} + Z(f),$$
$$Z\big(H(t)f\big) = \{t\}\,Z(f).$$

Hence, for these two classes of operators there is a simple relationship between the roots of the original polynomial and the roots of its image. In

Date: March 25, 2009 14:43.

2000 *Mathematics Subject Classification.* Primary: 30C15, Secondary: 26C10.

Key words and phrases. roots of polynomials, linear operators, minimum root separation.

contrast, for the differentiation operator $D : \mathcal{P}_n \to \mathcal{P}_n$ and $f \in \mathcal{P}_n$ there is no simple relation between $Z(Df)$ and $Z(f)$. The classical Gauss-Lucas theorem and its many improvements address this question; see for example the excellent monograph on the subject [9] and [6] for a recent development.

In this article we explore the relative position of $Z(Tf)$ in relation to $Z(f)$ for those invertible $T \in \mathcal{L}(\mathcal{P}_n)$ that are linear combinations of $I, D, \ldots, D^n$. To that end we first define a simple measure of the perturbation of the roots under $T \in \mathcal{L}(\mathcal{P}_n)$. For a non-constant polynomial $f \in \mathcal{P}_n$ set

$$R_T(f) := \min\bigl\{r > 0 : Z(Tf) \subset Z(f) + \mathbb{D}(r)\bigr\}.$$

Clearly, for the monomial $\phi_n(z) := z^n$ we have

$$R_T(\phi_n) = \max\bigl\{|v| : v \in Z(T\phi_n)\bigr\}.$$

In [5] we proved that the following two statements are equivalent (see Theorem 2.6 below):

(i) For all non-constant polynomials $f \in \mathcal{P}_n$ we have

$$R_T(f) \le R_T(\phi_n).$$

(ii) There exist $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that

$$\alpha_0 \neq 0 \quad \text{and} \quad T = \alpha_0 I + \alpha_1 D + \cdots + \alpha_n D^n. \tag{1.1}$$

Hence, an operator T given by (1.1) perturbs the roots of ϕ_n the most, as measured by R_T. Since ϕ_n only has one root of multiplicity n it is plausible to surmise that n distinct roots that are far apart from each other will be perturbed considerably less by T. But is this correct?

The following few lines of *Mathematica* code will help explore this question. In the code below, `T` stands for a list

$$\bigl\{\alpha_0, \alpha_1, \ldots, \alpha_n\bigr\}, \quad \alpha_0 \neq 0,$$

of the $n+1$ coefficients of the operator T in (1.1) and `W` stands for the list of the n roots of an $f \in \mathcal{P}_n$. First we calculate the roots of Tf and name them `TW`:

```
TW := z /. NSolve[
     T.(D[Times@@(z - W),{z, #}]&/@Range[0,Length[W]]) == 0,z
                ]
```

Then we plot the roots `W` as gray points and the perturbed roots `TW` as black points by the following command:

```
 Show[Graphics[{
   {AbsolutePointSize[5], GrayLevel[0.7],
                            Point[{Re[#], Im[#]}]&/@ W},
   {AbsolutePointSize[5],  Point[{Re[#], Im[#]}]&/@ TW}
                 }]]
```

Applying these two commands to sets `W` consisting of distinct points that are far apart from each other the first author observed the following surprising fact: the points of `TW` were very close to the points $\{-\alpha_1/\alpha_0\}$+`W`. That

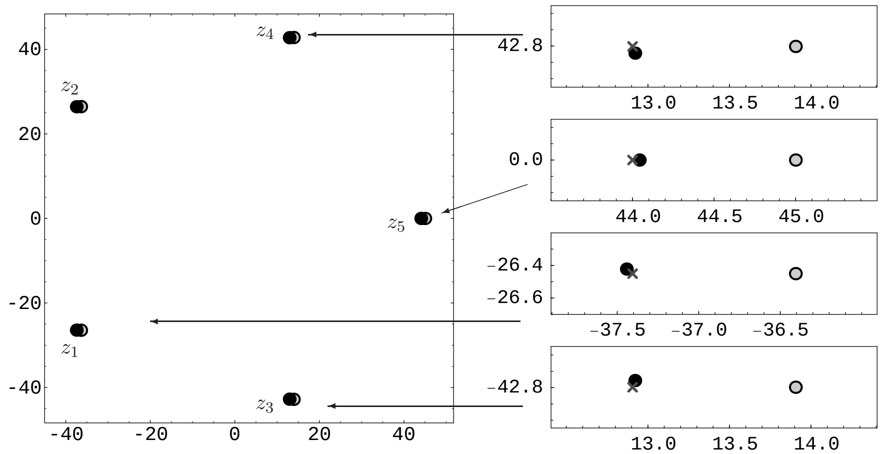

FIGURE 1. $n = 5$ and $a = 45$

FIGURE 2. A zoom-in on roots

is, the points of W had been essentially translated by $-\alpha_1/\alpha_0$. This kind of numerical experiment leads us then to believe that if the roots of $f \in \mathcal{P}_n$ are simple and far apart from each other, then the roots of Tf are close to the roots of $S(\alpha_1/\alpha_0)f$.

Next we explore this conjecture with three simple examples. Let $a > (n!)^{1/n}$, $\psi_{a,n}(z) := z^n - a^n$, and $T = I + D^n$. Here $\alpha_1 = 0$, so the roots of $T\psi_{a,n}$ and $\psi_{a,n}$ should be close for large a. Since $(T\psi_{a,n})(z) = z^n - (a^n - n!)$, if we pair the roots of $\psi_{a,n}$ and $T\psi_{a,n}$ with the same argument, then their moduli differ by $a - \big(a^n - n!\big)^{1/n}$. This quantity indeed tends to 0 as $a \to +\infty$. For an even $n > 2$ one can also consider $T = I + D^{n/2}$ and arrive at the same conclusion. For $T = I + D$ we have $(T\psi_{a,n})(z) = z^n + nz^{n-1} - a^n$ and $\alpha_1/\alpha_0 = 1$. We proceed with $n = 2$, since the roots of $T\psi_{a,2}$ are easily calculable only for this case. Then

$$Z\big(T\psi_{a,2}\big) = \Big\{-1 - \sqrt{1 + a^2}, -1 + \sqrt{1 + a^2}\Big\}.$$

To test our conjecture in this case, we consider the quantities

$$-a - \big(-1 - \sqrt{1 + a^2}\big) \quad \text{and} \quad a - \big(-1 + \sqrt{1 + a^2}\big). \tag{1.2}$$

Since both quantities converge to 1 as $a \to +\infty$ our conjecture is confirmed in this case as well. We offer Figures 1 and 2 as evidence that the same is true for $n = 5$. In the figures the gray points mark the roots of $\psi_{45,5}(z) = z^5 - 45^5$, the crosses mark the points in $\{-1\} + Z(\psi_{45,5})$, while the black dots mark the roots of $(T\psi_{45,5})(z) = z^5 + 5z^4 - 45^5$.

To formulate our conjecture as a formal statement we need to define a quantity that will play the role of a in the above simple examples and a quantity that will measure how close the roots are. To avoid the discontinuities caused by collapsed multiple roots, we will only consider polynomials

f in $\mathcal{P}_n$ with simple roots. Not surprisingly, it turns out that the appropriate generalization for a is the minimal distance between points of $Z(f)$ and points of $Z(f')$. We denote this quantity by $\tau(f)$, see Definition 2.1 below. As a measure how close the roots of two polynomials with the same degree are we will use the Fréchet distance d_F: this distance is obtained by pairing the roots of two polynomials to get the smallest possible maximal distance between the paired roots, see Definition 2.5. Then our conjecture leads us to the following proposition which is one of the main results of this article:

Proposition. *Let T be given by (1.1). Then for each $\epsilon > 0$ there exists $C_T(\epsilon) > 0$ such that for all $f \in \mathcal{P}_n$ with simple roots we have*

$$\tau(f) > C_T(\epsilon) \quad \Rightarrow \quad d_F\big(Z\big(S\big(\alpha_1/\alpha_0\big)f\big), Z(Tf)\big) < \epsilon.$$

This proposition is proved as Corollary 4.4 in Section 4. In Section 2 we collect the necessary definitions and background. Section 3 deals with operators $T \in \mathcal{P}_n$ given by (1.1) with $\alpha_1 = 0$. Here, Theorem 3.1 provides the key step towards the proof of the Proposition stated above. In Section 4 we prove more results involving Fréchet distance. We conclude with a few examples in Section 5.

A natural application of our results is towards a better estimate of the regions where the roots of polynomials are located. In Corollary 3.3 we illustrate the possible range of applications by considering the perturbations induced by operators of the form $T = I + \alpha D$, which have been traditionally of interest in the study of polynomial roots, see the books [7] and [9] for details. This result is a variation on the often quoted classical result by Takagi, see [10, part (VI)] or [9, Corollary 5.4.1 (iii)]. Our result, in some cases, gives more precise information than the classical result. This is illustrated in Example 5.3 which also gives a precise explanation of the behavior illustrated in Figures 1 and 2 above.

2. Definitions and preliminaries

Definition 2.1. Let p be a polynomial of degree n, $n \geq 2$, and assume that f has at least two distinct roots. Define

$$\operatorname{sep}(f) := \min\big\{|w - v| \ : \ w, v \in Z(f), w \neq v\big\}$$

and

$$\tau(f) := \min\big\{|w - v| \ : \ w \in Z(f), v \in Z(f')\backslash\{w\}\big\}\,.$$

We will need the following inequality which was established in [3, Theorem 4].

Theorem 2.2. *Let f be a polynomial of degree n, $n \geq 2$, and assume that f has at least two distinct roots. Then*

$$\frac{1}{n}\operatorname{sep}(f) \leq \tau(f) \leq \frac{1}{2\sin(\pi/n)}\operatorname{sep}(f).$$

Remark 2.3. The quantity $\operatorname{sep}(f)$ is known as the minimum root separation of a polynomial f. In [3] we denoted it by $\omega(f)$. Finding lower estimates for $\operatorname{sep}(f)$ is important since computing time required by an algorithm to isolate the roots of f depends inversely on $\operatorname{sep}(f)$, see for example [2]. In this sense roots of polynomials with large $\operatorname{sep}(f)$ are easy to find. Our results below indicate that such roots are also stable under the linear transformations of the form (1.1).

Remark 2.4. Most of our results below are formulated in terms of the quantity $\tau(f)$. Since we consider n fixed throughout, Theorem 2.2 yields that analogous results hold when $\tau(f)$ is replaced by $\operatorname{sep}(f)$.

Next we define the Fréchet distance (see [4, Section 3] for details and references) between multisets with the same cardinality.

Definition 2.5. Let m be a positive integer and put $\mathbb{M} = \{1, \ldots, m\}$. 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 adaptation of this definition to $d_F(A, B)$, where both A and B are multisets of size n, is then straightforward.

The following theorem is a combination of Theorems 9.1 and 11.1 from [5]. The reader should notice that $R_T(f) = d_h(Z(f), Z(Tf))$, d_h being the notation used in [5].

Theorem 2.6. *Let $T \in \mathcal{L}(\mathcal{P}_n)$. The following statements are equivalent:*

(a) *There exist $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that*

$$T = \alpha_0 I + \alpha_1 D + \cdots + \alpha_n D^n, \quad \alpha_0 \neq 0. \tag{2.1}$$

(b) *For all non-constant polynomials $f \in \mathcal{P}_n$ we have*

$$R_T(f) \leq R_T(\phi_n).$$

(c) *For all $f \in \mathcal{P}_n$ we have $\deg(Tf) = \deg(f)$ and there exists a constant C'_T such that $d_F\big(Z(f), Z(Tf)\big) \leq C'_T$ for all non-constant $f \in \mathcal{P}_n$.*

If $T \in \mathcal{L}(\mathcal{P}_n)$ is given by (2.1), then the smallest constant C'_T which satisfies (c) in Theorem 2.6 is denoted by $K_F(T)$.

3. Perturbations by a special class of operators

The following theorem is the key result which is used in the rest of the article. The main point of interest is that it treats perturbations by operators T of the type (2.1) where the D term in the expansion of T is missing. In the rest of the article we assume $n \geq 2$.

Theorem 3.1. *Let* $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ *and let* $T \in \mathcal{L}(\mathcal{P}_n)$ *be given by*

$$T = I + \alpha_1 D + \cdots + \alpha_n D^n. \tag{3.1}$$

Then $\alpha_1 = 0$ *if and only if there exists a constant* $\Gamma_T > 0$ *such that for all* $f \in \mathcal{P}_n$ *with simple roots we have*

$$\tau(f)\, R_T(f) < \Gamma_T. \tag{3.2}$$

Proof. Assume that T is given by (3.1) and that $\alpha_1 = 0$. Let $f \in \mathcal{P}_n$ be a polynomial with simple roots. We proceed with the construction of Γ_T in two steps. First, if $\tau(f) \leq 2R_T(\phi_n) + 1$, then Theorem 2.6 immediately yields

$$\tau(f) \leq 2\, R_T(\phi_n) + 1 \quad \Rightarrow \quad R_T(f)\tau(f) \leq R_T(\phi_n)\bigl(2R_T(\phi_n) + 1\bigr). \tag{3.3}$$

Next, we consider the case

$$\tau(f) > 2R_T(\phi_n) + 1. \tag{3.4}$$

Our goal is to construct an upper bound for $\tau(f)R_T(f)$ which does not depend on f. By the definition of $R_T(f)$ there exist $x_1 \in Z(f)$ and $v_1 \in Z(Tf)$ such that $R_T(f) = |x_1 - v_1|$. In addition, we can choose x_1 and v_1 so that

$$R_T(f) = |x_1 - v_1| \leq |x - v_1| \quad \text{for all} \quad x \in Z(f). \tag{3.5}$$

The polynomial g and the numbers $b_0, b_1, \dots, b_n$ defined by

$$g(z) := \bigl(S(v_1)f\bigr)(z) = b_0 + b_1 z + \frac{b_2}{2!} z + \cdots + \frac{b_n}{n!} z^n$$

will play an important role in the proof. Clearly $Z(g) = \{-v_1\} + Z(f)$. Hence, by (3.5), $w_1 = x_1 - v_1$ is a root of g with smallest modulus.

Next we explore the roots of g'. Note that $Z(g') = \{-v_1\} + Z(f')$. Let $z \in Z(f')$ be arbitrary. Since the roots of f are simple, $z \neq x_1$. By the definition of $\tau(f)$, (3.5), and (3.4) we have

$$\begin{aligned} |z - v_1| &\geq |z - x_1| - |x_1 - v_1| \\ &\geq \tau(f) - R_T(f) \\ &> 2R_T(\phi_n) + 1 - R_T(\phi_n) \\ &= R_T(\phi_n) + 1. \end{aligned}$$

Thus,

$$|u| > R_T(\phi_n) + 1 > 0 \quad \text{for all} \quad u \in Z(g'). \tag{3.6}$$

In particular,

$$b_1 = g'(0) \neq 0.$$

Now recall that $v_1 \in Z(Tf)$, $\alpha_1 = 0$, and observe that $S(v_1)$ and T commute, to deduce

$$0 = (Tf)(v_1) = \bigl(S(v_1)Tf\bigr)(0) = \bigl(Tg\bigr)(0) = b_0 + \sum_{k=2}^{n} \alpha_k b_k. \tag{3.7}$$

Since w_1 is a root of g with smallest modulus, Viète's formulas for g imply

$$|b_1||w_1| \le n|b_0|.$$

Together with (3.7) this yields

$$|w_1| \le \frac{n}{|b_1|}\sum_{k=2}^{n} |\alpha_k||b_k|. \tag{3.8}$$

Next we consider the polynomial

$$h(z) := z^{n-1}g'(1/z) = b_1 z^{n-1} + b_2 z^{n-2} + \ldots + \frac{b_{n-1}}{(n-2)!}z + \frac{b_n}{(n-1)!}.$$

We recall that $b_1 \neq 0$ and we notice that for $k = 2, \ldots, n$, the number $|b_k|/\big(|b_1|(k-1)!\big)$ is the modulus of the coefficient of z^{n-k} in the monic polynomial h/b_1. Further, let u_1 be a root of g' with minimal modulus. Then the largest modulus among the roots of h is $1/|u_1|$ and, again by Viète's formulas, we have

$$\left|\frac{b_k}{b_1(k-1)!}\right| \le \binom{n-1}{k-1}\frac{1}{|u_1|^{k-1}}, \quad k = 2, \ldots, n. \tag{3.9}$$

Since the function $x \mapsto \big((1+x)^{n-1} - 1\big)/x, x > 0$, is increasing, (3.6) implies

$$|u_1|\left(\left(1 + \frac{1}{|u_1|}\right)^{n-1} - 1\right) \le \big(R_T(\phi_n) + 1\big)\left(\left(\frac{R_T(\phi_n) + 2}{R_T(\phi_n) + 1}\right)^{n-1} - 1\right). \tag{3.10}$$

Next we use (3.8), (3.9) and (3.10) to establish an upper estimate for $|u_1 w_1|$:

$$\begin{aligned}
|u_1 w_1| &\le |u_1|\, n \sum_{k=2}^{n} \frac{|\alpha_k||b_k|}{|b_1|} \\
&\le |u_1|\, n \left(\max_{2\le k\le n} |\alpha_k|(k-1)!\right) \sum_{k=2}^{n} \left|\frac{b_k}{b_1(k-1)!}\right| \\
&\le |u_1|\, n \left(\max_{2\le k\le n} |\alpha_k|(k-1)!\right) \sum_{k=1}^{n-1} \binom{n-1}{k}\frac{1}{|u_1|^k} \\
&= |u_1|\, n \left(\max_{2\le k\le n} |\alpha_k|(k-1)!\right) \left(\left(1 + \frac{1}{|u_1|}\right)^{n-1} - 1\right) \\
&\le n \left(\max_{2\le k\le n} |\alpha_k|(k-1)!\right) \big(R_T(\phi_n) + 1\big)\left(\left(\frac{R_T(\phi_n) + 2}{R_T(\phi_n) + 1}\right)^{n-1} - 1\right).
\end{aligned}$$

For further reference we denote the last expression by Γ'_T. Hence we proved

$$|u_1 w_1| \le \Gamma'_T. \tag{3.11}$$

Recall that $u_1 \in Z(g')$ and $w_1 \in Z(g)$. Since the roots of g are simple, $w_1 \neq u_1$ and hence

$$|u_1 - w_1| \geq \tau(g) = \tau(f).$$

By Theorem 2.6 and (3.4)

$$R_T(f) \leq R_T(\phi_n) < \tau(f)/2.$$

Now, the triangle inequality and (3.5) yield

$$\begin{aligned} |u_1| &\geq |u_1 - w_1| - |w_1| \\ &\geq \tau(f) - R_T(\phi_n) \\ &> \tau(f)/2. \end{aligned}$$

Since by our choice $|w_1| = |x_1 - v_1| = R_T(f)$, (3.11) now yields

$$\tau(f) R_T(f) < 2\,|u_1|\,|w_1| \leq 2\,\Gamma'_T. \tag{3.12}$$

Thus, we have proved the implication (3.4) $\Rightarrow$ (3.12). To conclude the proof of the "only if" part of the theorem, set

$$\Gamma_T := \max\bigl\{R_T(\phi_n)\bigl(2R_T(\phi_n)+1\bigr), 2\,\Gamma'_T\bigr\}.$$

With this Γ_T the implication (3.4) $\Rightarrow$ (3.2) clearly holds. Finally, recall the implication (3.3) and the "only if" part is proved.

To prove the "if" part of the theorem, assume that T is given by (3.1) and that (3.2) holds for all $f \in \mathcal{P}_n$ with simple roots. Let $a > 0$ be arbitrary and as before $\psi_{a,2}(z) = z^2 - a^2$. Then a and $-a$ are the roots of $\psi_{a,2}$ and $\tau(\psi_{a,2}) = a$. Also,

$$(T\psi_{a,2})(z) = z^2 + 2\alpha_1 z - a^2 + 2\alpha_2.$$

The roots of $T\psi_{a,2}$ are

$$z_{a,1} = -\alpha_1 - \sqrt{\alpha_1^2 + a^2 - 2\alpha_2} \quad \text{and} \quad z_{a,2} = -\alpha_1 + \sqrt{\alpha_1^2 + a^2 - 2\alpha_2}.$$

Since clearly

$$\begin{aligned} \lim_{a\to+\infty}\Bigl|-a - \Bigl(-\alpha_1 - \sqrt{\alpha_1^2 + a^2 - 2\alpha_2}\Bigr)\Bigr| &= |\alpha_1|, \\ \lim_{a\to+\infty}\Bigl|a - \Bigl(-\alpha_1 + \sqrt{\alpha_1^2 + a^2 - 2\alpha_2}\Bigr)\Bigr| &= |\alpha_1|, \\ \lim_{a\to+\infty}\Bigl|-a - \Bigl(-\alpha_1 + \sqrt{\alpha_1^2 + a^2 - 2\alpha_2}\Bigr)\Bigr| &= +\infty, \\ \lim_{a\to+\infty}\Bigl|a - \Bigl(-\alpha_1 - \sqrt{\alpha_1^2 + a^2 - 2\alpha_2}\Bigr)\Bigr| &= +\infty, \end{aligned}$$

we conclude

$$\lim_{a\to+\infty} R_T(\psi_{a,2}) = |\alpha_1|.$$

Since by (3.2) for all $a > 0$ we have $R_T(\psi_{a,2}) \leq \Gamma_T/a$, letting $a \to +\infty$ leads to $|\alpha_1| = 0$. This completes the proof. $\square$

Corollary 3.2. *Let $\alpha_2, \dots, \alpha_n \in \mathbb{C}$. Let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by*

$$T = I + \alpha_2 D^2 + \cdots + \alpha_n D^n.$$

Let $f \in \mathcal{P}_n$ be a polynomial with simple roots such that $\tau(f) > 2R_T(\phi_n) + 1$. Then

$$Z(Tf) \subset Z(f) + \mathbb{D}\big(\Gamma'_T/\tau(f)\big),$$

with Γ'_T as defined in the sentence preceding (3.11).

Proof. The corollary is in fact a restatement of the implication (3.4)⇒(3.12) which is proved as a part of the proof of Theorem 3.1. □

The following corollary is inspired by [9, Corollary 5.4.1(iii)].

Corollary 3.3. *Let $\alpha \in \mathbb{C}$ and let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by $T = I + \alpha D$. Let $f \in \mathcal{P}_n$, $n \geq 2$, be a polynomial with simple roots such that $\tau(f) > 2|\alpha|(n-1) + 1$. Then*

$$Z(Tf) \subset \{-\alpha\} + Z(f) + \mathbb{D}\big(\gamma_\alpha/\tau(f)\big), \tag{3.13}$$

where γ_α is given by

$$2n\big(|\alpha|(n-1)+1\big)\left(\left(\frac{|\alpha|(n-1)+2}{|\alpha|(n-1)+1}\right)^{n-1} - 1\right)\max\left\{|\alpha|^k\left(1-\frac{1}{k}\right), k = 2, \dots, n\right\}.$$

Proof. Corollary 3.2 does not apply to the operator $T \in \mathcal{L}(\mathcal{P}_n)$. Therefore we consider the composition $V = S(-\alpha)T \in \mathcal{L}(\mathcal{P}_n)$ where $S(-\alpha) \in \mathcal{L}(\mathcal{P}_n)$ is defined by $\big(S(-\alpha)f\big)(z) = f(z-\alpha)$. The Taylor formula at z implies that $S(-\alpha) = \sum_{k=0}^{n}(-\alpha)^k D^k/k!$. Hence $S(-\alpha)$ and T commute. Since $D \in \mathcal{L}(\mathcal{P}_n)$ is the differentiation operator on $\mathcal{P}_n$ we have $D^{n+1} = 0$. These observations lead to the following expression for V as a linear combination of derivatives:

$$\begin{aligned}
V &= (I + \alpha D)\sum_{k=0}^{n}\frac{(-\alpha)^k}{k!}D^k \\
&= \sum_{k=0}^{n}\frac{(-\alpha)^k}{k!}D^k - \sum_{k=0}^{n}\frac{(-\alpha)^{k+1}}{k!}D^{k+1} \\
&= I + \sum_{k=1}^{n}\frac{(-\alpha)^k}{k!}(1-k)D^k \\
&= I - \frac{\alpha^2}{2}D^2 + \frac{\alpha^3}{3}D^3 - \cdots + \frac{(-1)^{n+1}\alpha^n}{(n-2)!\,n}D^n.
\end{aligned}$$

Hence, Corollary 3.2 applies to V. To show that $\Gamma'_V = \gamma_\alpha$ we first calculate $(T\phi_n)(z) = z^n + \alpha n z^{n-1}$ and deduce $Z(T\phi_n) = \{0, -\alpha n\}$. Thus, $R_T(\phi_n) = |\alpha| n$. Also,

$$Z(V\phi_n) = Z\big(S(-\alpha)T\phi_n\big) = \{\alpha\} + \{0, -\alpha n\} = \{\alpha, \alpha(1-n)\},$$

and therefore, $R_V(\phi_n) = |\alpha|(n-1)$. As we calculated the coefficients of V to be

$$\alpha_k = \frac{(-1)^{k+1}\alpha^k}{(k-2)!\,k}, \qquad k = 2,\dots,n,$$

we have $\gamma_\alpha = \Gamma'_V$.

Since we assume that $\tau(f) > 2|\alpha|(n-1)+1 = 2R_V(\phi_n)+1$, Corollary 3.2 yields

$$Z(Vf) \subset Z(f) + \mathbb{D}\big(\gamma_\alpha/\tau(f)\big).$$

To obtain the inclusion in the corollary, substitute f with $S(\alpha)f$ and notice that $\tau\big(S(\alpha)f\big) = \tau(f)$, $VS(\alpha)f = Tf$ and $Z\big(S(\alpha)f\big) = \{-\alpha\} + Z(f)$. □

Remark 3.4. In our notation the inclusion proved in [9, Corollary 5.4.1(iii)] reads

$$Z(Tf) \subset \{-\alpha\, n/2\} + Z(f) + \mathbb{D}\big(|\alpha|\, n/2\big). \tag{3.14}$$

Notice that if $n > 1$ and

$$\tau(f) > \max\left\{2|\alpha|(n-1)+1, \frac{\gamma_\alpha}{|\alpha|(n-1)}\right\},$$

then the intersection of the right-hand sides of (3.13) and (3.14) provides improved information about the location of the roots of Tf. Moreover, if

$$\tau(f) > \max\left\{2|\alpha|(n-1)+1, \frac{\gamma_\alpha}{|\alpha|}\right\},$$

then (3.13) is an improvement of (3.14), the improvement being considerable for large $\tau(f)$.

Example 3.5. In this example we give a hint of the problems that can arise if we consider polynomials with multiple roots. For $a > 0$ set

$$g_a(z) = z^2(z-a)^2.$$

Let

$$T = I + D^2.$$

Since $g'_a(z) = 2z(z-a)(2z-a)$, we have $\tau(g_a) = a/2$. On the other hand, we have

$$\begin{aligned}(Tg_a)(z) &= g_a(z) + g''_a(z)\\ &= z^4 - 2az^3 + \left(a^2+12\right)z^2 - 12az + 2a^2\\ &= \frac{a^4}{16} - a^2 + \left(12 - \frac{a^2}{2}\right)\left(z - \frac{a}{2}\right)^2 + \left(z - \frac{a}{2}\right)^4.\end{aligned}$$

Thus the roots of Tg_a are symmetric with respect to the real axis and to the vertical line $\mathrm{Re}(z) = a/2$. So, it is sufficient to calculate one root of Tg_a with a positive imaginary part and real part less than $a/2$. For $a > 3\sqrt{2}$ such a root is given by

$$z_{a,1} = \frac{1}{2}\left(a - \sqrt{a^2 - 24 - 4i\sqrt{2a^2-36}}\right),$$

and the other three roots are

$$z_{a,2} = \overline{z_{a,1}}, \quad z_{a,3} = z_{a,1} + a, \quad z_{a,4} = \overline{z_{a,1}} + a.$$

By rationalizing and then simplifying the expression for $z_{a,1}$ it is not difficult to prove that

$$\lim_{a \to +\infty} z_{a,1} = \sqrt{2}\, i.$$

Consequently,

$$\begin{aligned} \lim_{a \to +\infty} \big(z_{a,2} - (-\sqrt{2}\, i)\big) &= 0, \\ \lim_{a \to +\infty} \big(z_{a,3} - (a + \sqrt{2}\, i)\big) &= 0, \\ \lim_{a \to +\infty} \big(z_{a,4} - (a - \sqrt{2}\, i)\big) &= 0. \end{aligned}$$

Since $Z(g_a) = \{0, a\}$, the last four equalities imply

$$\lim_{a \to +\infty} R_T(g_a) = \sqrt{2}.$$

As $\tau(g_a) = a/2$, the function $a \mapsto \tau(g_a) R_T(g_a)$, $a > 0$, is unbounded. Hence, the assumption of the simplicity of the roots in Theorem 3.1 cannot be dropped.

4. Results involving the Fréchet distance

In the previous section we used $R_T(f)$ as a measure of the distance between the roots of f and the roots of Tf. In this section we work with the Fréchet distance between the roots of polynomials f and Tf. Here $Z(f)$ and $Z(Tf)$ are considered as multisets of roots. The Fréchet distance is defined only for nonconstant polynomials with equal degrees. In each of the cases below, the fact that the degrees of f and Tf are equal follows from Theorem 2.6. In particular, $S(\alpha)$ does not change the degree of a polynomial, and thus all the Fréchet distances used below are well defined. Recall that the number $K_F(T)$ is defined immediately after Theorem 2.6. The following lemma provides a connection with the results from Section 3.

Lemma 4.1. *Let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by*

$$T = \alpha_0 I + \alpha_1 D + \cdots + \alpha_n D^n, \quad \alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{C}, \quad \alpha_0 \neq 0. \tag{4.1}$$

Let $f \in \mathcal{P}_n$ be a polynomial with simple roots such that

$$R_T(f) < 1 \quad \text{and} \quad \tau(f) > \frac{1 + K_F(T)}{\sin(\pi/n)}. \tag{4.2}$$

Then

$$d_F\big(Z(f), Z(Tf)\big) = R_T(f). \tag{4.3}$$

Proof. Let $f \in \mathcal{P}_n$ be a polynomial with simple roots satisfying (4.2). Then Theorem 2.2 yields

$$\operatorname{sep}(f) > 2 + 2K_F(T) \geq 2. \tag{4.4}$$

Further, by the definition of $K_F(T)$, we have

$$d_F\big(Z(f), Z(Tf)\big) \leq K_F(T).$$

Now assume that the degree of f is m and let $Z(f) = \{z_1, \ldots, z_m\}$, where $z_1, \ldots, z_m$ are distinct complex numbers. By the definition of the Fréchet distance $d_F\big(Z(f), Z(Tf)\big)$ the roots $w_1, \ldots, w_m$ of Tf (recall that also $\deg Tf = m$) can be indexed in such a way that

$$|z_j - w_j| \leq K_F(T), \quad j = 1, \ldots, m. \tag{4.5}$$

Let j and k be distinct numbers from $\{1, \ldots, m\}$. Then, $z_j \neq z_k$ and the triangle inequality, (4.4) and (4.5) yield

$$\begin{aligned} |w_j - w_k| &= |(w_j - z_j) + (z_j - z_k) + (z_k - w_k)| \\ &\geq |z_j - z_k| - |(w_j - z_j) + (z_k - w_k)| \\ &\geq |z_j - z_k| - (|w_j - z_j| + |z_k - w_k|) \\ &\geq |z_j - z_k| - |z_j - w_j| - |z_k - w_k| \\ &\geq \operatorname{sep}(f) - 2K_F(T) \\ &> 2. \end{aligned}$$

Consequently,

$$\operatorname{sep}(Tf) > 2.$$

Therefore, the disks $\{w_j\} + \mathbb{D}(1)$ (for $j \in \{1, \ldots, m\}$) are pairwise disjoint. Since $R_T(f) < 1$ and $\operatorname{sep}(f) > 2$, in each disk $\{w_j\} + \mathbb{D}(1)$, $j \in \{1, \ldots, m\}$, there is exactly one root of f. Renumber the roots of f in such a way that

$$z_j \in \{w_j\} + \mathbb{D}(1), \quad j \in \{1, \ldots, m\}.$$

Then

$$d_F\big(Z(f), Z(Tf)\big) \leq \max\big\{|w_j - z_j| : j \in \{1, \ldots, m\}\big\} = R_T(f).$$

Since clearly $R_T(f) \leq d_F\big(Z(f), Z(Tf)\big)$, (4.3) follows. □

Theorem 4.2. *Let $n \geq 2$ and let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by*

$$T = I + \alpha_2 D^2 + \cdots + \alpha_n D^n, \quad \alpha_2, \ldots, \alpha_n \in \mathbb{C}. \tag{4.6}$$

Then for every $\varepsilon > 0$ there exists $C_T(\varepsilon) > 0$ such that for all $f \in \mathcal{P}_n$ with simple roots the following implication holds:

$$\tau(f) > C_T(\varepsilon) \quad \Rightarrow \quad d_F\big(Z(f), Z(Tf)\big) < \varepsilon.$$

Proof. Let $\Gamma'_T > 0$ be the constant introduced immediately above inequality (3.11). For $\varepsilon > 0$ set

$$C_T(\varepsilon) = \max\left\{R_T(\phi_n) + 1, \Gamma'_T, \frac{\Gamma'_T}{\varepsilon}, \frac{1 + K_F(T)}{\sin(\pi/n)}\right\}. \tag{4.7}$$

Let $f \in \mathcal{P}_n$ be a polynomial with simple roots such that

$$\tau(f) > C_T(\varepsilon).$$

Since $\tau(f) > R_T(\phi_n) + 1$, Corollary 3.2 implies

$$R_T(f) < \frac{\Gamma'_T}{\tau(f)}. \tag{4.8}$$

Since $\tau(f) > \Gamma'_T$, $R_T(f) < 1$. Hence, Lemma 4.1 yields

$$d_F\big(Z(f), Z(Tf)\big) = R_T(f).$$

Since $\tau(f) > \Gamma'_T/\varepsilon$, (4.8) implies

$$R_T(f) < \varepsilon.$$

The last two displayed relations prove the theorem. □

The following proposition is proved by combining the methods of proofs of Theorems 3.1 and 4.2. In the proposition we use Γ_T defined in Theorem 3.1.

Proposition 4.3. *If $T \in \mathcal{L}(\mathcal{P}_n)$ is given by* (4.6)*, then for all $f \in \mathcal{P}_n$ with simple roots we have*

$$\tau(f)\, d_F\big(Z(f), Z(Tf)\big) \le \max\left\{\Gamma_T, K_F(T)\Gamma_T, K_F(T)\frac{1 + K_F(T)}{\sin(\pi/n)}\right\}.$$

Proof. Set

$$\Gamma''_T := \max\left\{\Gamma_T, \frac{1 + K_F(T)}{\sin(\pi/n)}\right\}.$$

Let $f \in \mathcal{P}_n$ be a polynomial with simple roots. As in the proof of Theorem 3.1 we proceed in two steps. First, if $\tau(f) \le \Gamma''_T$, then, by the definition of $K_F(T)$,

$$\tau(f)\, d_F\big(Z(f), Z(Tf)\big) \le \Gamma''_T\, K_F(T) \le \max\left\{\Gamma_T, \Gamma''_T K_F(T)\right\}.$$

For the second step assume $\tau(f) > \Gamma''_T$. Recall that by Theorem 3.1

$$\tau(f)\, R_T(f) \le \Gamma_T \le \Gamma''_T. \tag{4.9}$$

Consequently, $R_T(f) < 1$. Since also $\tau(f) > \big(1 + K_F(T)\big)/(\sin(\pi/n))$, Lemma 4.1 implies

$$d_F\big(Z(f), Z(Tf)\big) = R_T(f).$$

Substituting this identity in (4.9), we conclude

$$\tau(f)\, d_F\big(Z(f), Z(Tf)\big) \le \Gamma_T \le \max\left\{\Gamma_T, \Gamma''_T K_F(T)\right\},$$

and the proposition is proved. □

Theorem 4.2 is of course a consequence of Proposition 4.3, but $C_T(\epsilon)$ given by (4.7) is smaller than the corresponding constant deduced from Proposition 4.3.

Now we have all the tools to prove the proposition stated in the Introduction.

Corollary 4.4. *Let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by* (4.1)*. Then for every $\varepsilon > 0$ there exists $C_T(\varepsilon) > 0$ such that for all $f \in \mathcal{P}_n$ with simple roots we have*

$$\tau(f) > C_T(\varepsilon) \quad \Rightarrow \quad d_F\big(Z(S(\alpha_1/\alpha_0)f), Z(Tf)\big) < \varepsilon.$$

Proof. To prove the corollary we observe the identity

$$d_F\big(Z(S(\alpha_1/\alpha_0)f), Z(Tf)\big) = d_F\big(Z(f), Z\big(S(-\alpha_1/\alpha_0)Tf\big)\big), \quad f \in \mathcal{P}_n \setminus \mathcal{P}_0.$$

Next we notice that Theorem 4.2 applies to the operator

$$T_1 := \frac{1}{\alpha_0} S(-\alpha_1/\alpha_0)T.$$

Then the corollary follows from Theorem 4.2. □

A simple calculation shows that for $t > 0$ we have

$$\tau\big(H(t)f\big) = t\,\tau(f),$$

and thus we note the following corollary.

Corollary 4.5. *Let $T \in \mathcal{L}(\mathcal{P}_n)$ be given by* (4.1)*. For an arbitrary $f \in \mathcal{P}_n$ with simple roots we have*

$$\lim_{t \to +\infty} d_F\Big(Z\big(S(\alpha_1/\alpha_0)H(t)f\big), Z\big(TH(t)f\big)\Big) = 0$$

Now we return to a simple motivating example from the Introduction. Finding the limits of the expressions in (1.2) is a standard calculus exercise. Since the expressions in (1.2) represent distances between roots of two quadratic polynomials, the following corollary can be seen as a generalization of this standard calculus exercise.

Corollary 4.6. *Let $a_1, \ldots, a_{n-1}$ be arbitrary, but fixed, complex numbers. For $a \in \mathbb{C}$ set*

$$g_a(z) = z^n + a^n \quad \text{and} \quad f_a(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a^n.$$

Then

$$\lim_{|a| \to +\infty} d_F\Big(Z\big(S(a_{n-1}/n)g_a\big), Z\big(f_a\big)\Big) = 0.$$

Proof. Setting

$$\alpha_0 = 1, \quad \alpha_k = \frac{(n-k)!}{n!}\, a_{n-k}, \quad k = 1, \ldots, n-1, \quad \alpha_n = 0,$$

in (4.1) we have $Tg_a = f_a$. Since clearly $\tau(g_a) = |a|$, the corollary follows from Corollary 4.4. □

Corollary 4.6 indicates that for a monic polynomial f_a with constant term a^n of large magnitude the approximate location of its roots depends only on a, n, and the coefficient a_{n-1}. The following corollary formalizes this observation.

Corollary 4.7. *Let $a_1, \dots, a_{n-1}$ be arbitrary, but fixed, complex numbers. For $a \in \mathbb{C}$ set*

$$h_a(z) = z^n + a_{n-1}z^{n-1} + a^n \quad \text{and} \quad f_a(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a^n.$$

Then

$$\lim_{|a|\to+\infty} d_F\big(Z(h_a), Z(f_a)\big) = 0.$$

Proof. Let $g_a(z) = z^n - a^n$. By Corollary 4.6

$$\lim_{|a|\to+\infty} d_F\Big(Z\big(S(a_{n-1}/n)g_a\big), Z\big(f_a\big)\Big) = 0$$

and

$$\lim_{|a|\to+\infty} d_F\Big(Z\big(S(a_{n-1}/n)g_a\big), Z\big(h_a\big)\Big) = 0.$$

Since by [4, Proposition 3.1], d_F is a metric on the space of root sets $Z(\cdot)$,

$$d_F\big(Z(h_a), Z(f_a)\big) \le d_F\Big(Z\big(S(a_{n-1}/n)g_a\big), Z\big(f_a\big)\Big) + d_F\Big(Z\big(S(a_{n-1}/n)g_a\big), Z\big(h_a\big)\Big).$$

The corollary follows from the last three displayed relations. □

5. Estimates and examples

Example 5.1. A simple way to see our results in action, in particular Theorem 3.1, is to look at an example for which we can exactly calculate $\tau(f)$ and $R_T(f)$. In the Introduction we considered $T = I + D^n$. Here $\alpha_1 = 0$, so Theorem 3.1 applies. Further, $R_T(\phi_n) = (n!)^{1/n}$ and, for $n \ge 3$,

$$\Gamma_T = 2n!\big((n!)^{1/n} + 1\big)\left(\left(\frac{(n!)^{1/n} + 2}{(n!)^{1/n} + 1}\right)^{n-1} - 1\right).$$

A rough estimate yields

$$\Gamma_T < 2\big(\sqrt{2} + 1\big)\, n! \left(\frac{3}{2}\right)^{n-1}.$$

Hence, by Theorem 3.1,

$$\tau(f)\, R_T(f) < 2\big(\sqrt{2} + 1\big)\, n! \left(\frac{3}{2}\right)^{n-1} \tag{5.1}$$

for all $f \in \mathcal{P}_n$ with simple zeros. Let $a > 0$ and $\psi_{a,n}(z) = z^n - a^n$. Then $\tau(\psi_{a,n}) = a$ and $(T\psi_{a,n})(z) = z^n - (a^n - n!)$. It is not difficult to see that

$$R_T(\psi_{a,n}) = \big|a - \big(a^n - n!\big)^{1/n}\big|,$$

where, for $0 < a < (n!)^{1/n}$, the formula $\big(a^n - n!\big)^{1/n}$ denotes the root with argument $\pi/(2n)$. Further, elementary considerations yield

$$\tau(\psi_{a,n})\, R_T(\psi_{a,n}) = a\,\big|a - \big(a^n - n!\big)^{1/n}\big| \le (n!)^{2/n}.$$

This is certainly much better than (5.1). However, (5.1) holds for *all* $f \in \mathcal{P}_n$ with simple zeros.

Remark 5.2. Example 5.1 indicates that the constant Γ_T in Theorem 3.1 might not be close to optimal, at all. We note that Theorem 3.1 together with Theorem 2.2 yields information about the maximum root separation $\operatorname{sep}(f)$:

$$\operatorname{sep}(f) R_T(f) \leq n\Gamma_T.$$

In [1], Collins provides us with an impressive amount of numerical evidence for what should be some general "ideal" lower bound for $\operatorname{sep}(f)$. Still his work also highlights how far the current theoretical tools, for example [8], are from getting close to the numerical conjectures. The same difficulty might also be at work here.

Example 5.3. Here we apply Corollary 3.3 to the polynomials $\psi_{a,5}(z) = z^5 - a^5$, $a > 0$, and the operator $T = I + D$ to provide a precise explanation for the behavior of the roots exhibited in Figures 1 and 2. In this case we have $\tau(\psi_{a,5}) = a$ and $(T\psi_{a,5})(z) = z^5 + 5z^{n-1} - a^5$. By Corollary 3.3, if $a > 9$ we have

$$Z(T\psi_{a,5}) \subset \{-1\} + Z(\psi_{a,5}) + \mathbb{D}\big(\gamma_1/a\big),$$

where $\gamma_1 = 42.944$. In Figures 1 and 2 we used $a = 45$. Thus

$$R_T(\psi_{45,5}) \leq \gamma_1/a = 0.95431\dot{1}.$$

In terms of the geometric objects in Figures 1 and 2 the last inequality predicts that the maximum distance between black dots and crosses, paired in the natural way, is < 0.954312. In fact *Mathematica* gives 0.046083 as the maximum distance. Our estimate therefore overestimates 0.046083 by means of 0.954312, which is not impressive. But it can still be considered satisfactory because the upper bound from Corollary 3.3, or more generally from Theorem 3.1, holds under the most general conditions.

Finally we notice that [9, Corollary 5.4.1(iii)] in this specific case would only give

$$Z(T\psi_{45,5}) \subset \{-5/2\} + Z(\psi_{45,5}) + \mathbb{D}\big(5/2\big),$$

which is much less precise information. This is illustrated in Figure 3, where the boundaries of the disks in the last inclusion are gray, the smaller disks from our estimate are outlined in black, and the roots of

$$T\psi_{45,5} = z^5 + 5z^4 - 45^5$$

are marked by black dots. In a magnified view in Figure 4 the gray point marks $45\exp(6\pi i/5)$, while the cross is at $45\exp(6\pi i/5) - 1$.

Acknowledgment. We are grateful to the referees for many valuable suggestions that led to a much improved paper.

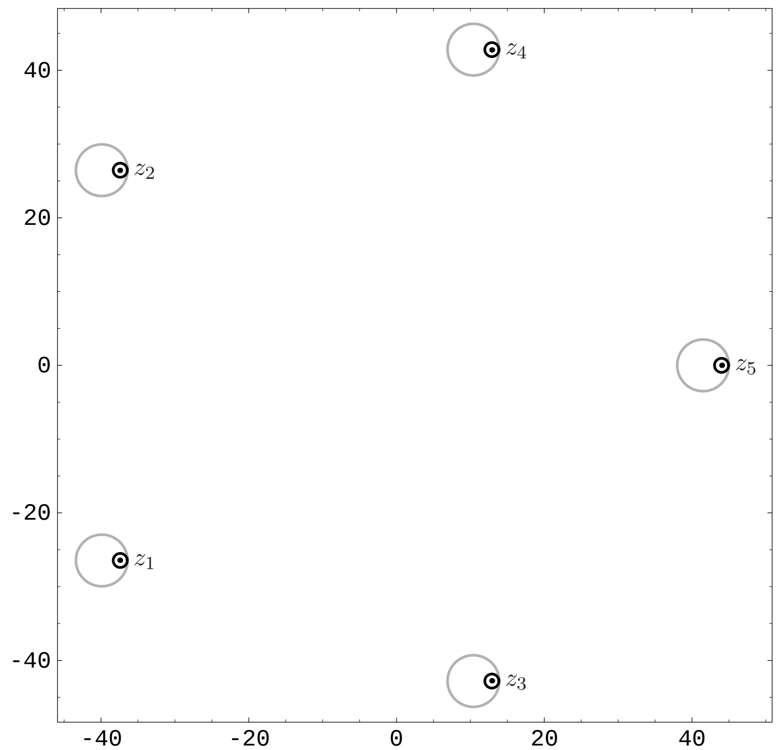

FIGURE 3. A comparison of estimates for the roots of $z^5 + 5z^4 - 45^5$

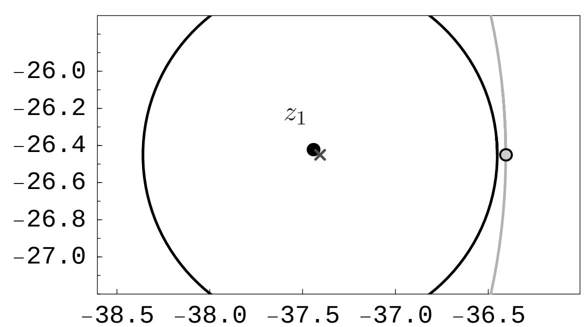

FIGURE 4. A zoom-in on z_1 with a root of $z^5 - 45^5$ marked by a gray disk

REFERENCES

[1] Collins, G.E.: Polynomial minimum root separation. J. Symbolic Comput. 32 (2001), 467–473.

[2] Collins, G. E., Horowitz, E.: The minimum root separation of a polynomial. Mathematics of Computation, 28 (1974), 589–597

[3] Ćurgus, B., Mascioni, V.: On the location of critical points of polynomials. Proc. Amer. Math. Soc. 131 (2003), 253–264.

[4] Ćurgus, B., Mascioni, V.: Roots and polynomials as homeomorphic spaces. Expo. Math. 24 (2006), 81–95.

[5] Ćurgus, B., Mascioni, V.: Perturbations of roots under linear transformations of polynomials. Constructive Approximation 25 (2007), 255–277.
[6] Malamud, S. M.: Inverse spectral problem for normal matrices and the Gauss-Lucas theorem. Trans. Amer. Math. Soc. 357 (2005), 4043–4064.
[7] Marden, M.: *Geometry of polynomials* . Second edition reprinted with corrections, American Mathematical Society, Providence, 1985.
[8] Mignotte, M.: Some useful bounds, in Computer algebra (ed. B. Buchberger, G. E. Collins and R. Loos), 259–263, Springer, 1982.
[9] Rahman, Q. I., Schmeisser, G.: *Analytic theory of polynomials*. Oxford University Press, 2002.
[10] Takagi, T.: Note on the algebraic equations. Proceedings of the Physico-Mathematical Society of Japan 3 (1921) 175–179.

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

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