Tree Breadth of the Continued Fractions Root Finding Method

Citation for published version:
Kalorkoti, K 2020, Tree Breadth of the Continued Fractions Root Finding Method. in *New Trends in Algebras and Combinatorics: Proceedings of the 3rd International Congress in Algebras and Combinatorics (ICAC2017).* World Scientific, pp. 203-227, The Third International Congress in Algebras and Combinatorics, Hong Kong, Hong Kong, 25/08/17. https://doi.org/10.1142/9789811215476_0014

Digital Object Identifier (DOI):
10.1142/9789811215476_0014

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
New Trends in Algebras and Combinatorics

Tree Breadth of the Continued Fractions Root Finding Method

K. Kalorkoti[1]

Abstract: The continued fractions algorithm for isolating the real roots of polynomials with certainty is one of the most efficient known and widely used. It can be viewed as exploring a tree created by a sequence of simple transformations. In this paper we produce new upper bounds for the breadth of the tree that are significantly smaller than the degree of the input polynomial. We also consider the expected breadth under a reasonable distribution and derive a bound, subject to a plausible assumption, that grows logarithmically with the degree and coefficient size.

Keywords: Polynomial roots, continued fractions, tree breadth.

AMS Subject Classification: 12Y05

1 Introduction

The continued fractions algorithm for isolating the (positive) real roots of polynomials is one of the most efficient known and widely used, excluding algorithms that are subject to numerical instability. Its ultimate source is a result of Vicent [21] which guarantees the termination of a sequence of transformations that enable us to explore the real roots in $(0, 1)$ and $(1, \infty)$. Alesina and Galuzzi [4] give a modern proof of the result as well as historical information. Rather than reiterate the extensive bibliography we will refer the reader to a few papers that cover it in great detail (in particular Krandick and Mehlhorn [14], Tsigaridas and Emiris [20]).

Krandick and Mehlhorn [14] analyse a version that uses homothetic transformations given by Collins and Akritas [6], obtaining new bounds on the number of recursive subdivisions. They also prove that the breadth of the recursion tree is bounded by the degree of the input square free polynomial.

By constrast, Tsigaridas and Emiris [20] give complexity and implementation results for the method without homothetic transformations (see below for further details on this method). Their analysis relies on a conjecture regarding the continued fractions expansions of non-quadratic algebraic irrationals (see p.161 of [20]). They also discuss an earlier analysis by Akritas [2, 3].

The original continued fractions method (without homothetic transformations) can be seen as exploring a tree of transformations with vertices labeled by polynomials and edges by one of two transformations. In this paper we provide bounds for the tree breadth of individual polynomials as well as for the average tree breadth (subject to an assumption) under a reasonable distribution.

For individual polynomials, Lemma 2.4 provides a bound in terms of the number of certain types of roots, which immediately implies that the degree is an upper bound. Theorem 3.2 provides a bound in terms of the degree and the size of coefficients which is good for cases of sequences where the coefficients have sufficiently controlled growth, the bound is even better if the number of non-zero coefficients is bounded.

For the average case we consider drawing uniformly at random square free polynomials of degree n with integer coefficients in $[-B, B]$ where $B \geq 1$, we also consider primitive square free polynomials. Theorem 4.1 provides an upper bound that is logarithmic in n and B but this is subject to an assumption on the distribution of roots in the unit disc (discussed at the start of §4.1, see also §4.2).

2 Definitions

Throughout we consider non-zero polynomials in z with coefficients from $\mathbb{R}$ (in practice the coefficients are from $\mathbb{Q}$ or, equivalently for root finding, from $\mathbb{Z}$ and we will assume this in some places). If

$$f = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

then we define $\mathrm{vc}(f)$, the *variation of coefficients*, to be the number of sign changes in the sequence $a_n, a_{n-1}, \ldots, a_0$ of coefficients (ignoring as usual any occurrences of 0). The continued fractions

[1]School of Informatics, 10 Crichton Street, Edinburgh EH8 9AB, UK (kk@inf.ed.ac.uk)

method of finding the non-negative real roots of a square free f is based on the transformations

$$f^T = f(z+1) = a_n(z+1)^n + a_{n-1}(z+1)^{n-1} + \cdots + a_0,$$

and

$$f^I = f(1/(z+1))(z+1)^n = a_n + a_{n-1}(z+1) + \cdots + a_0(z+1)^n,$$

where we have assumed that n is the degree of f (for the sake of completeness we can set $0^T = 0^I = 0$). For simplicity later on, if $f^T(0) = 0$ then we replace $f^T(z)$ by $f^T(z)/z$ and similarly for $f^I(z)$. This ensures that 0 is not a root of f^T of of f^I and so it has non-zero constant coefficient enabling us to express various bounds using this (the alternative is to use the trailing coefficient). For the same reasons we will assume that $a_0 \neq 0$.

In practice transformations of the first type are replaced by $z \mapsto z + c$ where c is a good integer lower bound on the smallest positive root of f, see [2, 3]. Without this optimisation the algorithm is provably exponential, see [6]. For the purposes of this paper it makes no difference to the overall situation so we will stay with the definition as given (our analysis holds unchanged no matter which version is used). Vincent's Theorem [21] shows that if f is square free then for all sufficiently long sequences of transformations involving T and I we produce a polynomial with variation of coefficients either 0 or 1. This is false if the polynomial is not square free, for example if $f = (2z^2 - 1)^2$ then $f^{IIT} = f^I = z^4 + 4z^3 + 2z^2 - 4z + 1$. In connection with non-square free polynomials see the result cited by [20] as Theorem 5 of that paper. A polynomial f is called *terminal* if and only if $\mathrm{vc}(f) \leq 1$.

In order to avoid confusion we note here that we will employ two forms of notation for transforms. The exponent form is compact and makes for ease of readability, but it must be borne in mind that it conforms to the algebraic notation of writing the argument on the left with order of application being as shown. Thus f^{IT} means that we apply I first and then T. In some situations it is more convenient to use the standard function notation with the argument on the right so that the order of application is the reverse of the written one. Thus if we wish to represent z^{IT} in standard notation it becomes $T(I(z))$, which can be abbreviated to $TI(z)$; this will be significant when considering sequences of transformations as single Moebius transforms.

Given a polynomial f we associate with it a binary tree, denoted by $\mathrm{tree}(f)$, of the possible sequences of transformations of f as follows. If f is terminal the tree is empty otherwise the root is labeled with f. At any vertex v labeled with the polynomial g if g^T not terminal then there is a right child with the edge labeled T and the vertex labeled with g^T. Similarly if g^I is not terminal there is a left child with the edge labeled I and the vertex labeled with g^I.

We define the depth of a vertex to be the number of edges from the unique path to it starting at the root. Thus the root has depth 0 and its children, if any, have depth 1. The breath at depth d of the tree is the number of vertices of depth d. The breadth of a tree is 0 if it is empty otherwise it is the maximum breadth over all depths (if this is unbounded we take the breadth to be ∞, this does not happen in our case). We denote the breadth of $\mathrm{tree}(f)$ by $\mathrm{br}(f)$.

We define $\mathbf{R}^+(f)$ to be the set of strictly positive real roots of f. We will say that α is complex to mean that $\alpha \in \mathbb{C} - \mathbb{R}$. Let $\mathbf{C}(f)$ denote the set of complex roots α of f such that $\mathrm{Re}(\alpha) > 0$ and $\mathrm{Re}(\alpha) - \lfloor \mathrm{Re}(\alpha) \rfloor > |\alpha - \lfloor \mathrm{Re}(\alpha) \rfloor|^2$. We also define $\mathbf{c}(f)$ to be 0 if all complex roots α of f that do not belong to $\mathbf{C}(f)$ have $\mathrm{Re}(\alpha) \leq 0$ and to be 1 otherwise. Note that $\mathrm{vc}((z-\alpha)(z-\overline{\alpha})) = 0$ if and only if $\mathrm{Re}(\alpha) \leq 0$ since $(z-\alpha)(z-\overline{\alpha}) = z^2 - 2\,\mathrm{Re}(\alpha)z + |\alpha|^2$. It follows that if $|\,\mathbf{R}^+(f)| + \mathbf{c}(f) = 0$ then $\mathrm{vc}(f) = 0$.

Lemma 2.1 *For all polynomials f*

1. $\mathbf{C}(f)$ *consists of all complex roots α of f with $Re(\alpha) > 0$ that are in the open disc* $|z - \lfloor Re(z) \rfloor - 1/2| < 1/2$.
2. *If $\alpha \notin \mathbf{C}(f)$ is complex then $Re(\alpha) \leq |\alpha|^2$.*

PROOF. Let α be a root of f and set $m = \lfloor \mathrm{Re}(\alpha) \rfloor$ throughout. For the first part, we have α with $\mathrm{Re}(\alpha) > 0$ and α is in the given disc if and only if

$$(\alpha - m - 1/2)(\overline{\alpha} - m - 1/2) < 1/4 \iff (\alpha - m)(\overline{\alpha} - m) - (\alpha - m)/2 - (\overline{\alpha} - m)/2 + 1/4 < 1/4$$
$$\iff |\alpha - m|^2 - (\mathrm{Re}(\alpha) - m) < 0$$

which completes the proof.

For the second part, if $\mathrm{Re}(\alpha) \leq 0$ the claim is trivial. So assume that $\mathrm{Re}(\alpha) > 0$. Since $\alpha \notin \mathbf{C}(f)$ we have

$$\begin{aligned}\mathrm{Re}(\alpha) - m &\leq |\alpha - m|^2 \\ &= (\alpha - m)(\overline{\alpha} - m) \\ &= |\alpha|^2 - 2m\,\mathrm{Re}(\alpha) + m^2\end{aligned}$$

Thus $\mathrm{Re}(\alpha) \leq |\alpha|^2 - m(2\,\mathrm{Re}(\alpha) - m - 1)$. If $m = 0$ the claim follows immediately. Otherwise if $m > 0$ then $2\,\mathrm{Re}(\alpha) - m - 1 \geq 2m - m - 1 = m - 1 \geq 0$ and hence $\mathrm{Re}(\alpha) \leq |\alpha|^2$. □

We note that the preceding lemma remains true for the set $\widehat{\mathbf{C}}(f)$ that is defined is the same way as $\mathbf{C}(f)$ but without the condition that its members must be complex. This will be useful in §4.1.

Define $\mathbf{C}_m(f)$ to be the set of points of $\mathbf{C}(f)$ satisfying $|z - m - 1/2| < 1/2$ for $m = 0, 1, 2, \ldots$. Note that $\mathbf{C}_0(f)$ is the same as the set C of Definition 19 in [14]. It follows from Lemma 2.1 that $\mathbf{C}(f) = \cup_{m=0}^{\infty} \mathbf{C}_m(f)$. We define $\widehat{\mathbf{C}}_m(f)$ similarly. Clearly the sets $\mathbf{C}_i(f)$ are disjoint as are the sets $\widehat{\mathbf{C}}_i(f)$. The next Lemma follows from Lemma 2.6, the proof given here is more direct.

LEMMA 2.2 *$\mathbf{C}(f^T) = \cup_{m=1}^{\infty} T^{-1}(\mathbf{C}_m(f))$ and $\mathbf{C}(f^I) \subseteq I^{-1}(\mathbf{C}_0(f))$. In particular $\mathbf{C}_m(f^T) = T^{-1}(\mathbf{C}_{m+1}(f))$. Furthermore a root α of f cannot both be mapped by T a root in $\mathbf{C}(f^T)$ and mapped by I to a root in $\mathbf{C}(f^I)$.*

PROOF. We have for a complex β that

$$\begin{aligned}\beta \in \mathbf{C}_m(f^T) &\Leftrightarrow f^T(\beta) = 0 \;\&\, \mathrm{Re}(\beta) > 0 \;\&\; |\beta - m - 1/2| < 1/2 \\ &\Leftrightarrow f(\alpha) = 0 \;\&\; \beta = \alpha - 1 \;\&\, \mathrm{Re}(\alpha) > 1 \;\&\; |\alpha - 1 - m - 1/2| < 1/2 \\ &\Leftrightarrow \alpha \in \mathbf{C}_{m+1}(f) \;\&\; \beta = T^{-1}(\alpha).\end{aligned}$$

This proves the claims regarding f^T. For the claim regarding f^I, suppose $\beta \in \mathbf{C}(f^I)$. Then $\beta = 1/\alpha - 1 = \overline{\alpha}/|\alpha|^2 - 1$ where $f(\alpha) = 0$. If $\alpha \notin \mathbf{C}_0(f)$ then one of three possibilities holds: (i) $\mathrm{Re}(\alpha) \leq 0$, (ii) $0 < \mathrm{Re}(\alpha) < 1$ and $\mathrm{Re}(\alpha) \leq |\alpha|^2$ or (iii) $\mathrm{Re}(\alpha) \geq 1$. If any of these conditions hold then $\mathrm{Re}(\beta) \leq 0$ which contradicts the assumption that $\beta \in \mathbf{C}(f^I)$.

The final claim follows from the description of $\mathbf{C}(f^T)$ and $\mathbf{C}(f^I)$ given by the first part. □

Note that we cannot strengthen the second containment to an equality, e.g., if $f = 8z^2 - 4z + 1$ then $\mathbf{C}_0(f) = \{\, 1/4 + i/4, 1/4 - i/4 \,\}$. However $f^I = z^2 - 2z + 5$ and $\mathbf{C}(f^I) = \emptyset$ since the roots are $1 \pm 2i$. Indeed it follows from Lemma 2.6 that $\mathbf{C}_m(f^I)$ consists of all $I^{-1}(\alpha)$ where α is a root of f that is in the open disc $|z - (2m+3)/2(m+1)(m+2)| < 1/2(m+1)(m+2)$. Just as above, the preceding lemma holds for $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{C}}_m$.

LEMMA 2.3 *For all polynomials f we have*

1. *If $\alpha \in \mathbf{R}^+(f)$ but $T^{-1}(\alpha) \notin \mathbf{R}^+(f^T)$ and $I^{-1}(\alpha) \notin \mathbf{R}^+(f^I)$ then $\alpha = 1$. Moreover $T(\mathbf{R}^+(f^T)) \cap I(\mathbf{R}^+(f^I)) = \emptyset$*
2. *$|\,\mathbf{R}^+(f)| \geq |\,\mathbf{R}^+(f^T)| + |\,\mathbf{R}^+(f^I)|$.*
3. *$|\mathbf{C}(f)| \geq |\mathbf{C}(f^T)| + |\mathbf{C}(f^I)|$.*
4. *$|\mathbf{C}_0(f)| \geq |\mathbf{C}(f^I)| + \mathbf{c}(f^I)$ and $\mathbf{c}(f) \geq \mathbf{c}(f^T)$.*

PROOF. For the first claim note that $T^{-1}(1) = I^{-1}(1) = 0 \notin T^{-1}(\mathbf{R}^+(f^T)) \cup I^{-1}(\mathbf{R}^+(f^I))$. If $\alpha > 1$ then $T^{-1}(\alpha) = \alpha - 1 > 0$ while $I^{-1}(\alpha) = 1/\alpha - 1 < 0$ and so $T^{-1}(\alpha) \in \mathbf{R}^+(f^T)$ but $I^{-1}(\alpha) \notin \mathbf{R}^+(f^I)$. If $\alpha < 1$ the same argument shows that $T^{-1}(\alpha) \notin \mathbf{R}^+(f^T)$ but $I^{-1}(\alpha) \in \mathbf{R}^+(f^I)$.

The second claim follows from the first.

For the third claim, suppose a complex root $\alpha \notin \mathbf{C}(f)$ then under I it is mapped to $1/\alpha - 1 = \overline{\alpha}/|\alpha|^2 - 1$. From Lemma 2.1 we have $\mathrm{Re}(\alpha) \leq |\alpha|^2$. It follows that $\mathrm{Re}(1/\alpha - 1) \leq 0$ and so $1/\alpha - 1 \notin \mathbf{C}(f^T) \cup \mathbf{C}(f^I)$. Under T the root is mapped to $\alpha - 1$ and again this cannot be in $\mathbf{C}(f^T) \cup \mathbf{C}(f^I)$, e.g., by the second part of the preceding Lemma. It follows from Lemma 2.2 that if $\alpha \in \mathbf{C}(f)$ then we cannot have both $1/\alpha - 1 \in \mathbf{C}(f^I)$ and $\alpha - 1 \in \mathbf{C}(f^T)$. The claim now follows.

We now deal with the first part of the fourth claim. By Lemma 2.2 we have $\mathbf{C}(f^I) \subseteq I^{-1}(\mathbf{C}_0(f))$ and so if $\mathbf{c}(f^I) = 0$ the claim follows immediately. Suppose now that $\mathbf{c}(f^I) = 1$ so there is a complex root β of f^I with $\mathrm{Re}(\beta) > 0$ and $\beta \notin \mathbf{C}(f^I)$. It follows that $\beta = 1/\alpha - 1$ for some complex root α of f. Thus $\mathrm{Re}(\beta) = \mathrm{Re}(\alpha)/|\alpha|^2 - 1$ and so $\mathrm{Re}(\alpha) > |\alpha|^2$. It follows that $0 < \mathrm{Re}(\alpha) < 1$ and hence $\alpha \in \mathbf{C}_0(f)$. Since $\mathbf{C}(f^I) \subseteq I^{-1}(\mathbf{C}_0(f))$ and $\beta \notin \mathbf{C}_0(f^I)$ the containment is strict. The claim now follows. The second part of fourth claim follows immediately from Lemma 2.2. □

LEMMA 2.4 *$br(f) \leq |\mathbf{R}^+(f)| + |\mathbf{C}(f)| + \mathbf{c}(f)$ and hence $br(f) \leq \deg(f)$.*

PROOF. If $\mathrm{tree}(f)$ is empty the claim is trivial. We now assume that $\mathrm{tree}(f)$ is not empty and use induction on the depth d. If $d = 0$ then $|\mathbf{R}^+(f)| + |\mathbf{C}(f)| + \mathbf{c}(f) \geq 1$ since f is not terminal and so $\mathrm{br}(f) = 1 \leq |\mathbf{R}^+(f)| + |\mathbf{C}(f)| + \mathbf{c}(f)$. Assume now that $d > 0$. By induction $\mathrm{br}(f^T) \leq |\mathbf{R}^+(f^T)| + |\mathbf{C}(f^T)| + \mathbf{c}(f^T)$ and $\mathrm{br}(f^I) \leq |\mathbf{R}^+(f^I)| + |\mathbf{C}(f^I)| + \mathbf{c}(f^I) \leq |\mathbf{R}^+(f^I)| + |\mathbf{C}_0(f)|$ by the fourth part of Lemma 2.3. Now, using Lemma 2.2 in the third line below and Lemma 2.3 in the last line,

$$\begin{aligned}
\mathrm{br}(f) &\leq \mathrm{br}(f^T) + \mathrm{br}(f^I) \\
&\leq |\mathbf{R}^+(f^T)| + |\mathbf{C}(f^T)| + \mathbf{c}(f^T) + |\mathbf{R}^+(f^I)| + |\mathbf{C}_0(f)| \\
&\leq |\mathbf{R}^+(f^T)| + |\mathbf{R}^+(f^I)| + \sum_{m=1}^{\infty} |\mathbf{C}_m(f)| + |\mathbf{C}_0(f)| + \mathbf{c}(f) \\
&\leq |\mathbf{R}^+(f)| + |\mathbf{C}(f)| + \mathbf{c}(f),
\end{aligned}$$

which establishes the main inequality. The consequence is immediate since f has at least $|\mathbf{R}^+(f)| + |\mathbf{C}(f)| + \mathbf{c}(f)$ distinct roots. □

Krandick and Mehlhorn [14] prove that the breadth is bounded by the degree of the input square free polynomial for the variant method that also employs homothetic transformations (see their Theorem 29).

2.1 Effect of Möbius Transforms

For the reader's convenience we collect together some simple results on the effect of Möbius transforms on certain discs.

LEMMA 2.5 *Assume that $(ck - cr - a)(ck + cr - a) > 0$. Then the set of values satisfying $|(az + b)/(cz + d) - k| < r$ is given by the open disc*

$$\left| z - \frac{(ad + bc)k - ab - cd(k^2 - r^2)}{(ck - cr - a)(ck + cr - a)} \right| < \frac{|ad - bc|r}{(ck - cr - a)(ck + cr - a)}$$

In particular if $k > r \geq 0$ then under the transform $z \mapsto 1/z$ the disc $|z - k| < r$ goes to $|z - k/(k^2 - r^2)| < r/(k^2 - r^2)$.

PROOF. The given disc is the same as $|(az+b)-k(cz+d)|<r|cz+d|$. Setting $z=u+iv$ and squaring both sides, the disc is given by $((a-kc)u+b-kd)^2+(a-kc)^2v^2-r^2(cu+d)^2-r^2c^2v^2<0$, since $(ck-cr-a)(ck+cr-a)>0$ the derived inequality is equivalent to $(u-C)^2+v^2-R^2<0$ where

$$C=\frac{(ad+bc)k-ab-cd(k^2-r^2)}{(ck-cr-a)(ck+cr-a)} \text{ and } R=\frac{|ad-bc|r}{(ck-cr-a)(ck+cr-a)},$$

which is the claimed disc.

The rest follows by noting that $a=0$, $b=1$, $c=1$, $d=0$ so that the condition $(ck-cr-a)(ck+cr-a)>0$ reduces to $(k-c)(k+c)>0$ and since $k>c\geq 0$ it is satisfied. Substituting the values of a,b,c,d into the general derived disc we obtain claimed disc. □

We note that if M consists of a sequence of T^{-1} and I^{-1} transformations then $ad-bc=(-1)^s$ where s is the number of occurrences of I^{-1}, cf. Theorem 8 of Collins and Krandick [5]. This can be shown by a straightforward induction. Thus, in this situation, the disc in the preceding lemma can be written as

$$\left|z-\frac{(ad+bc)k-ab-cd(k^2-r^2)}{(ck-cr-a)(ck+cr-a)}\right|<\frac{r}{(ck-cr-a)(ck+cr-a)}$$

LEMMA 2.6 *Let $M(z)$ be a Möbius transform composed of T and I and set $M^{-1}(z)=(az+b)/(cz+d)$. Then $C_m(f^M)$ consists of all $M^{-1}(\alpha)$ where α is a complex root of f that is in the open disk*

$$\left|z-\frac{(ad+bc)(m+1/2)-ab-cdm(m+1)}{(cm-a)(c(m+1)-a)}\right|<\frac{1}{2(cm-a)(c(m+1)-a)}$$

PROOF. The roots β of f^M are precisely all $\beta=M^{-1}(\alpha)$ where α is a root of f. Now a root β belongs to $\mathbf{C}_m(f^M)$ if and only $|\beta-(m+1/2)|<1/2$, i.e., $|M^{-1}(\alpha)-(m+1/2)|<1/2$. Since M is a composition of T and I we have $M(z)=(Az+B)/(Cz+D)$ with $A,B,C,D\geq 0$ and not both C,D are 0 (similarly for A,B). We have $M^{-1}(z)=(Dz-B)/(-Cz+A)$. Thus the inequality $(ck-cr-a)(ck+cr-a)>0$ of Lemma 2.5 becomes $(-C(m+1/2)+C/2-D)(-C(m+1/2)-C/2-D)>0$ which is equivalent to $(Cm+D)(C(m+1)+D)>0$. Since $C,D\geq 0$ and at least one is non-zero the inequality holds. The result now follows from Lemma 2.5. □

2.2 Bounds on the number of roots

This section summarises some well known results for the reader's convenience. The Mahler measure of a polynomial $f=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_0$ with roots $\alpha_1,\ldots,\alpha_n$ is $M(f)=|a_n|\prod_{j=1}^n\max\{1,|\alpha_j|\}$. As is well known, Jensen's formula (see, e.g., Ahlfors [1]) yields the bound $M(f)\leq\sum_{j=0}^n|a_j|$. Suppose $\rho>1$ and set $\epsilon(f,\rho)=|\{\alpha_j\mid 1\leq j\leq n\ \&\ |\alpha_j|>\rho\}|$. Then

$$\epsilon(f,\rho)\leq\frac{1}{\log\rho}\Big(\log\sum_{j=0}^n|a_j|-\log|a_n|\Big). \tag{1}$$

This follows from the simple observation that $\rho^{\epsilon(f,\rho)}\leq|a_n|^{-1}M(f)\leq|a_n|^{-1}\sum_{j=0}^n|a_j|$. Now suppose that $\rho<1$ and set $\mu(f,\rho)=|\{\alpha_j\mid 1\leq j\leq n\ \&\ |\alpha_j|<\rho\}|$. It follows from (1) and the transformation $z\mapsto 1/z$ that

$$\mu(f,\rho)\leq\frac{1}{\log(1/\rho)}\Big(\log\sum_{j=0}^n|a_j|-\log|a_0|\Big). \tag{2}$$

Hughes and Nikeghbali [10] give the bounds

$$\epsilon(f,1/(1-\rho))\leq\frac{1}{\rho}\big(\log\sum_{i=0}^n|a_i|-\log|a_n|\big),$$

$$\mu(f,1-\rho)\leq\frac{1}{\rho}\big(\log\sum_{i=0}^n|a_i|-\log|a_0|\big),$$

These are slightly weaker than the ones above since $\log(1/(1-\rho)) = \rho + \rho^2/2 + \rho^3/3 + \cdots$.

It will be helpful to set

$$L(f) = \frac{1}{\sqrt{|a_0||a_n|}} \sum_{j=0}^{n} |a_j|.$$

LEMMA 2.7 *Let $f = a_n z^n + \cdots + a_0$ where $a_n a_0 \neq 0$ and assume that $\rho > 1$. Then*

$$\epsilon(f, \rho) + \mu(f, 1/\rho) \leq \frac{2}{\log \rho} L(f).$$

PROOF. Straightforward application of (1) and (2). □

LEMMA 2.8 *Let S be a set of univariate polynomials of degree n with non-zero constant term and coefficients bounded in absolute value from above by B. Consider drawing polynomials at random from S (using any probability distribution). Then*

$$\mathbb{E}[\epsilon(f, \rho)] \leq \frac{\log(n+1)B}{\log \rho},$$

when $\rho > 1$ and

$$\mathbb{E}[\mu(f, \rho)] \leq \frac{\log(n+1)B}{\log(1/\rho)},$$

when $\rho < 1$.

PROOF. The inequalities are an immediate consequence of (1) and (2). □

LEMMA 2.9 *Let $f = a_n z^n + \cdots + a_0$ where $f \in \mathbb{Z}[z]$ and $a_n a_0 \neq 0$, then the number of positive integer roots of f (counted with multiplicity) is at most $1 + \log|a_0|/\log 2$.*

PROOF. Suppose $m_1, \ldots, m_r$ are integer roots of f, necessarily non-zero. A simple argument based on Gauss's Lemma for the content and primitive part of polynomials shows that $m_1 \cdots m_r \mid a_0$. Consider now the positive integer roots of f other than 1 and assume there are s of them. Since each root is an integer which is at least 2, it follows that $s \leq \log|a_0|/\log 2$. Taking the possibility that 1 is a root into account now yields the result. □

3 Bound on the tree breadth for a polynomial

We will use a famous result of Erdős and Turàn [8] in the form given by Rahman and Schmeisser [17], Theorem 11.6.4: denote by $n_f[\theta, \phi)$ the number of zeros in the sector $\{z \mid \theta \leq \arg z < \phi\}$, where $0 < \phi - \theta \leq 2\pi$. Then

$$\left| n_f[\theta, \phi) - \frac{\phi - \theta}{2\pi} n \right| \leq C\sqrt{n \log L(f)},$$

where $C = \sqrt{2\pi/G} < 2.62$ and $G = \sum_{m=0}^{\infty} (-1)^{m-1}(2m+1)^{-2}$ is Catalan's constant. The original paper [8] had $C = 16$, the improved constant is due to Ganelius [9].

We are interested in the number of roots $N_f(\phi)$ within a wedge defined by the angles $-\phi$, ϕ. Note that in this we include those roots α with $\arg(\alpha) = \pm\phi$. It follows that

$$N_f(\phi) < \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\epsilon\phi}{\pi} n \tag{3}$$

for all $\epsilon > 1$.

Suppose now that f has no more than k non-zero coefficients. Proposition 11.2.4 of [17] states: for $0 < \phi - \theta < 2\pi$, denote by $n_f(\alpha, \beta)$ the number of zeros of f in the sector $\{z \mid \theta < \arg z < \phi\}$ then

$$\left| n_f(\theta, \phi) - \frac{\phi - \theta}{2\pi} n \right| \leq k.$$

It follows that

$$N_f(\phi) < k + \frac{\epsilon\phi}{\pi} n \tag{4}$$

for all $\epsilon > 1$.

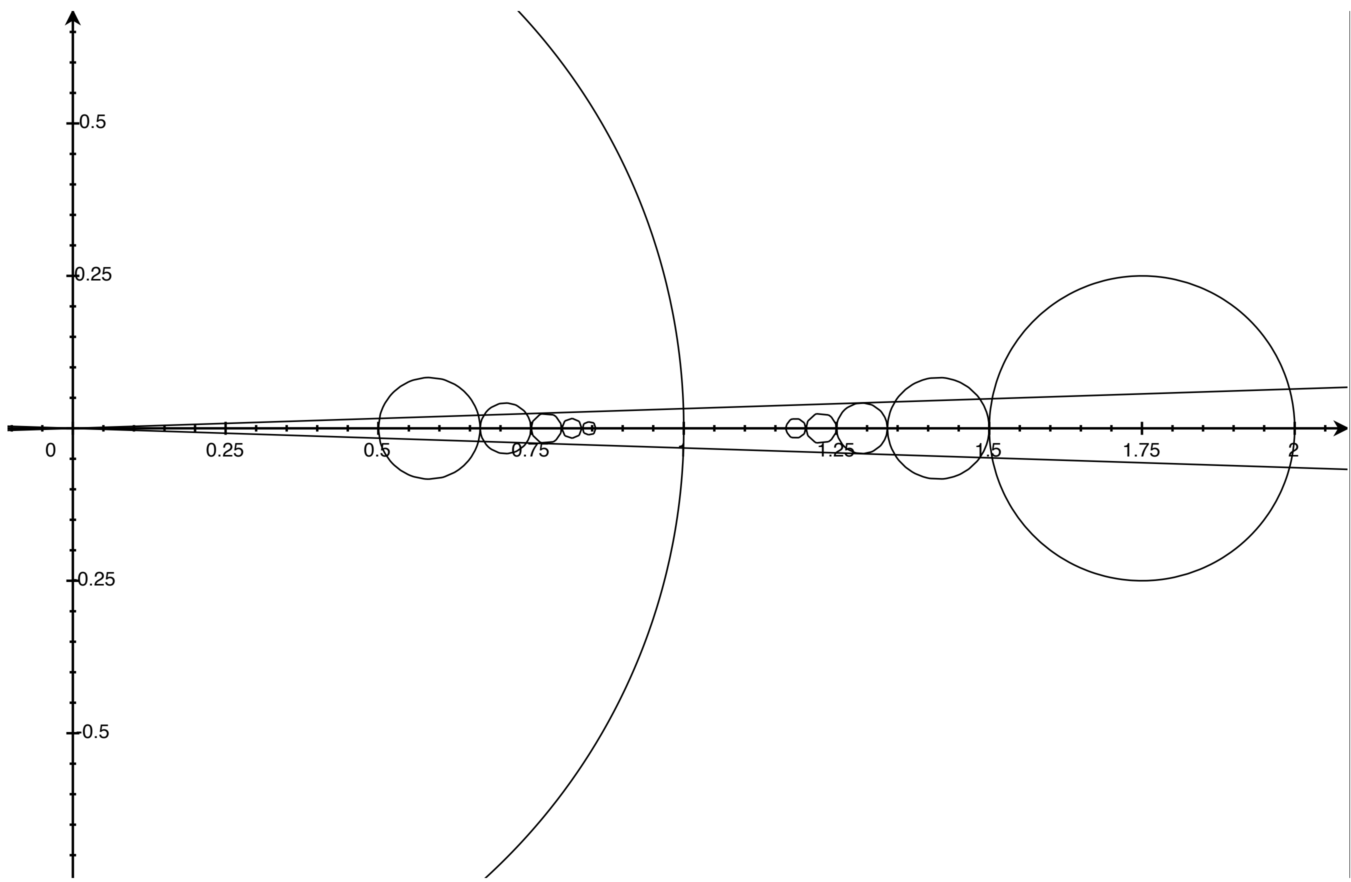

Figure 1: The discs corresponding to $\mathbf{C}_m(f^M)$ for $M(z) = II(z) = z^{II}$ (left of unit circle) and $M(z) = IT(z) = z^{TI}$ (right of unit circle) and $m = 0, 1, 2, 3, 4$. Also shown is the wedge enclosing the discs corresponding to $\mathbf{C}_m(f^M)$ for $m \geq 2$

THEOREM 3.1 *Let f be any polynomial of degree n with real coefficients and non-zero constant term. For $m \geq 0$ define $\phi = \arctan\left(1/2(m+2)\sqrt{(m+1)(m+3)}\right)$. Then*

1. *the breadth of tree(f) satisfies*

$$br(f) \leq \frac{2}{\log\left(\frac{m+2}{m+1}\right)} \log L(f) + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\phi}{\pi} n + 4.$$

Suppose a satisfies $0 < a < 4$ and $b \geq 0$. Choose m_0 such that $a(m+b)^4 < 4(m+1)(m+2)^2(m+3)$ for all $m \geq m_0$. If $\sqrt{a}(m_0+b)^2 \geq 2/\pi$, then

$$br(f) < \frac{1}{\sqrt[4]{a}\log^2(2)\sqrt{\phi}} \log L(f) + \frac{\phi}{\pi} n + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} - \frac{1}{\log 2}\left(\frac{b}{\log 2} - 2\right) \log L(f) + 4.$$

2. *Suppose that the number of non-zero coefficients of f is no more than k. Then*

$$br(f) \leq \frac{2}{\log\left(\frac{m+2}{m+1}\right)} \log L(f) + \frac{\epsilon\phi}{\pi} n + k + 4.$$

Let a, b, ϕ and m_0 be as above. If $\sqrt{a}(m_0+b)^2 \geq 2/\pi$, then

$$br(f) < \frac{1}{\sqrt[4]{a}\log^2(2)\sqrt{\phi}} \log L(f) + \frac{\epsilon\phi}{\pi} n + k - \frac{1}{\log 2}\left(\frac{b}{\log 2} - 2\right) \log L(f) + 4.$$

PROOF. We look first at item 1. Consider tree(f). If it is not complete at depth 2 (i.e., one of f^I, f^T, f^{II}, f^{IT}, f^{TI}, f^{TT} is terminal) we may extend it artificially by adding the appropriate vertices and paths. It follows from Lemma 2.4 that br(f) is bounded from above by the maximum of 2 and

$$\begin{aligned}
\mathrm{br}(f^{II}) + \mathrm{br}(f^{IT}) + \mathrm{br}(f^{TI}) + \mathrm{br}(f^{TT}) &\leq |\mathbf{C}(f^{II})| + |\mathbf{C}(f^{IT})| + |\mathbf{C}(f^{TI})| + |\mathbf{C}(f^{TT})| + \\
&\quad \mathbf{R}^+(f^{II}) + \mathbf{R}^+(f^{IT}) + \mathbf{R}^+(f^{TI}) + \mathbf{R}^+(f^{TT}) + \\
&\quad \mathbf{c}(f^{II}) + \mathbf{c}(f^{IT}) + \mathbf{c}(f^{TI}) + \mathbf{c}(f^{TT}) \\
&\leq |\mathbf{C}(f^{II})| + |\mathbf{C}(f^{IT})| + |\mathbf{C}(f^{TI})| + |\mathbf{C}(f^{TT})| + \mathbf{R}^+(f) + 4
\end{aligned}$$

where the final line is justified by the second part of Lemma 2.3. Thus the preceding expression acts as an upper bound for br(f) in any case.

Using Lemmas 2.2 and 2.6 we have the following cases.

1. $\mathbf{C}_m(f^{II})$ is a subset of the set of all $I^{-1}I^{-1}(\alpha)$ where α ranges over all roots α of f in $\mathbf{C}(f)$ that lie in the open disc

$$\left| z - \frac{2m^2+8m+7}{2(m+2)(m+3)} \right| < \frac{1}{2(m+2)(m+3)}.$$

2. $\mathbf{C}_m(f^{IT})$ is a subset of the set of all $I^{-1}T^{-1}(\alpha)$ where α ranges over all roots α of f in $\mathbf{C}(f)$ that lie in the open disc

$$\left| z - \frac{2m+5}{2(m+2)(m+3)} \right| < \frac{1}{2(m+2)(m+3)}.$$

3. $\mathbf{C}_m(f^{TI})$ is a subset of the set consists of all $T^{-1}I^{-1}(\alpha)$ where α ranges over all roots α of f in $\mathbf{C}(f)$ that lie in the open disc

$$\left| z - \frac{2m^2+8m+7}{2(m+1)(m+2)} \right| < \frac{1}{2(m+1)(m+2)}.$$

4. $\mathbf{C}(f^{TT})$ consists of all $T^{-1}T^{-1}(\alpha)$ where α ranges over all roots α of f in $\cup_{m=2}^{\infty}\mathbf{C}_m(f)$ that lie in any of the open discs

$$\left| z - \frac{2m+5}{2} \right| < \frac{1}{2}.$$

Note that, by Lemma 2.5, cases 1 and 3 are dual to each other via the transformation $z \mapsto 1/z$ and similarly for 2 and 4; however this observation does not lead to any advantage over the proof below. We can deal with the second and fourth cases easily. For the second case the disc for m is contained in a disc $D(0, 1/(m+2)$ and hence $\mathbf{C}(f^{IT})$ is contained in the disc $D(0, 1/2)$. For the fourth case, all the roots α in question satisfy $|\alpha| > 2$. Thus, by Lemma 2.7, $|\mathbf{C}(f^{IT})| + |\mathbf{C}(f^{TT})| \le (2/\log 2)\log L(f)$.

For the other two cases we consider a wedge that encloses $C_m(f^{II})$. Suppose the lines $y = \pm\tan(\phi)x$, where $\phi > 0$, are tangent to the boundary of $\mathbf{C}_m(f^{II})$ with centre c and radius r, see Figure 1. It follows easily that

$$\tan^2(\phi) = \frac{r^2}{c^2 - r^2} = \frac{1}{4(m+1)(m+2)^2(m+3)}. \tag{5}$$

Clearly such a wedge will enclose $\mathbf{C}_s(f^{II})$ for all $s \ge m$. Furthermore it does the same for $\mathbf{C}_s(f^{TI})$, since the ratio of the centre to the radius of the disc for $\mathbf{C}_m(f^{TI})$ is the same as that for $\mathbf{C}_m(f^{II})$ and $\tan^2(\alpha) = 1/(c^2/r^2 - 1)$.

The boundary of the disc corresponding to $\mathbf{C}_m(f^{II})$ crosses the x-axis at $x = (m+1)/(m+2)$, nearest to the origin, while the boundary for the disc corresponding to $\mathbf{C}_m(f^{TI})$ crosses the x-axis at $x = (m+2)/(m+1)$, furthest from the origin. Thus $|\cup_{i=0}^{m-1}\mathbf{C}_m(f^{II})|$ is bounded from above by the number of roots of f in the interior of the disc with radius $\sigma = (m+1)/(m+2)$. Similarly $|\cup_{i=0}^{m-1}\mathbf{C}_m(f^{TI})|$ is bounded from above by the number of roots of f in the exterior of the disc with radius $1/\sigma$. By Lemma 2.7 the number of such roots is bounded from above by $2\log L(f)/\log((m+2)/(m+1))$.

Since the wedge includes $\mathbf{R}^+(f)$ it follows from (3) that for all $\epsilon > 1$.

$$|\mathbf{R}^+(f)| + |\mathbf{C}(f^{II})| + |\mathbf{C}(f^{IT})| + |\mathbf{C}(f^{TI})| + |\mathbf{C}(f^{TT})| < \frac{2}{\log\left(\frac{m+2}{m+1}\right)}\log L(f) + \sqrt{\frac{2\pi}{G}}\sqrt{n\log L(f)} + \frac{\epsilon\phi}{\pi}n.$$

The first claim of item 1 now follows from (5) and the observation regarding $\mathrm{br}(f)$ at the start of this proof by taking ϵ arbitrarily close to 1.

For the second claim, It is easily seen that $\tan(\phi) > \phi$ for $0 < \phi < \pi/2$ and so, by (5), we have $4(m+1)(m+2)^2(m+3) < 1/\phi^2$. For all large enough m, say $m \ge m_0 \ge 0$, we have $a(m+b)^4 \le 4(m+1)(m+2)^2(m+3)$. Thus it suffices to have $m < 1/\sqrt[4]{a}\sqrt{\phi} - b$ in order to ensure the condition $4(m+1)(m+2)^2(m+3) < 1/\phi^2$. So for a given $\phi > 0$ the corresponding value of m satisfies $m < 1/\sqrt[4]{a}\sqrt{\phi} - b$. In order to have $m \ge m_0$ we need $\phi < 1/\sqrt{a}(m_0+b)^2 \le \pi/2$.

We claim that $\log((m+2)/(m+1)) \ge 2\log^2 2/(m + 2\log 2)$. To see this consider the function $h(x) = \log((x+2)/(x+1)) - 2\log^2 2/(x + 2\log 2)$ defined over the non-negative reals. The denominator of dh/dx is positive and the numerator is $(c_1 x + c_0)x$ where $c_1 = 2\log^2 2 - 1 = -0.390\ldots$ and $c_0 = 6\log^2 2 - 4\log 2 = 0.110\ldots$. Hence h is increasing for $0 \le x \le -c_0/c_1 \approx 2.817$ and decreasing for $x \ge -c_0/c_1$. Now $h(0) = 0$ and then becomes positive while $\lim_{x\to\infty} h(x) = 0$. Hence $h(x) \ge 0$ for all $x \ge 0$ and the claim is established. By the first part of item 1 we have

$$\begin{aligned}
\mathrm{br}(f) &\le 2\left(\frac{1}{\log 2} + \frac{m}{2\log^2 2}\right)\log L(f) + \sqrt{\frac{2\pi}{G}}\sqrt{n\log L(f)} + \frac{\phi}{\pi}n + 4 \\
&< \frac{2}{\log 2}\log L(f) + \frac{1}{\log^2 2}\left(\frac{1}{\sqrt[4]{a}\sqrt{\phi}} - b\right)\log L(f) + \sqrt{\frac{2\pi}{G}}\sqrt{n\log L(f)} + \frac{\phi}{\pi}n + 4 \\
&= \frac{1}{\sqrt[4]{a}\log^2(2)\sqrt{\phi}}\log L(f) + \frac{\phi}{\pi}n + \sqrt{\frac{2\pi}{G}}\sqrt{n\log L(f)} - \frac{1}{\log 2}\left(\frac{b}{\log 2} - 2\right)\log L(f) + 4
\end{aligned}$$

The claims of item 2 follow by using (4) instead of (3) in the derivation above. □

THEOREM 3.2 *Let f be any polynomial of degree n with real coefficients and non-zero constant term. Suppose a satisfies $0 < a < 4$ and $b \geq 0$. Choose m_0 such that $a(m+b)^4 \leq 4(m+1)(m+2)^2(m+3)$ and set $c = (b/\log 2 - 2)/\log 2$. If $\sqrt{a}(m_0+b)^2 \geq 2/\pi$ and $\log L(f) \leq 2n\log^2(2)/\pi\sqrt{a}(m_0+b)^3$ then*

$$\begin{aligned} br(f) &< \frac{3}{\sqrt[3]{4\pi\sqrt{a}\log^4 2}}\sqrt[3]{n\log^2 L(f)} + \sqrt{\frac{2\pi}{G}}\sqrt{n\log L(f)} - c\log L(f) + 4 \\ &< \frac{2.11}{\sqrt[6]{a}}\sqrt[3]{n\log^2 L(f)} + 2.62\sqrt{n\log L(f)} - c\log L(f) + 4 \end{aligned}$$

Suppose that the number of non-zero coefficients of f is no more than k. Then

$$\begin{aligned} br(f) &< \frac{3}{\sqrt[3]{4\pi\sqrt{a}\log^4 2}}\sqrt[3]{n\log^2 L(f)} - c\log L(f) + k + 4 \\ &< \frac{2.11}{\sqrt[6]{a}}\sqrt[3]{n\log^2 L(f)} - c\log L(f) + k + 4 \end{aligned}$$

PROOF. The function $g(\phi) = u/\sqrt{\phi} + v\phi$ has its minimum at $\phi = \big(u/2v\big)^{2/3}$. As seen in the proof of Theorem 3.1 the r.h.s. needs to be bounded from above by $1/\sqrt{a}(m_0+b)^2$ and this holds provided that $u < 2v/\sqrt[4]{a^3}(m_0+b)^3$. Thus the minimum value of $g(\phi)$ is $(2u^2v)^{1/3} + (u^2v/4)^{1/3} = 3(u^2v/4)^{1/3}$. Setting $u = \log L(f)/\sqrt[4]{a}\log^2 2$ and $v = n/\pi$ the inequality is $\log L(f) < 2n\log^2 2/\pi\sqrt{a}(m_0+b)^3$ and the value of $g(\phi)$ is $3(n\log^2 L(f)/4\pi\sqrt{a}\log^4 2)^{1/3}$. Substituting into the second inequality of item 1 of Theorem 3.1 yields the desired bound. The second bound follows by substituting into the second inequality of item 2 of Theorem 3.1. □

The assumption that $\log L(f) \leq 2n\log^2(2)/\pi\sqrt{a}(m_0+b)^3$ ensures that the bounds of the preceding lemmas are strictly positive; a simple argument shows that under the assumption $c\log L(f) < 3\sqrt[3]{n\log^2 L(f)}/\sqrt[3]{4\pi\sqrt{a}\log^4 2}$. If the condition on $L(f)$ is satisfied then, since $\sqrt{a}(m_0+b)^2 \geq 2/\pi$, we have $\log L(f) \leq n\log^2(2)/(m_0+b)$. Thus $L(f) \leq 2^{n\log(2)/(m_0+b)}$. If f has coefficients bounded in absolute value by B (the situation of the next section) than $\log L(f) \leq \log(n+1)B$ and the condition on $L(f)$ is satisfied if $B \leq 2^{n\log(2)/(m_0+b)}/(n+1)$, i.e., for all polynomials of sufficiently high degree. As a simple illustration we may take $a = 9/8$, $b = 3$ and $m_0 = 1$. The condition $\sqrt{a}(m_0+b)^2 \geq 2/\pi$ is satisfied while the condition on $L(f)$ is satisfied if $\log L(f) \leq 0.05n$. The second bound of the preceding lemma becomes

$$L(f) < 2.07\sqrt[3]{n\log^2 L(f)} - 3.35\log L(f) + k + 4.$$

4 BOUND ON THE EXPECTED BREADTH OF THE TRANSFORMATION TREE

Let $B > 0$ be an integer and let $P(n,B)$ be the set of non-zero polynomials from $\mathbb{Z}[z]$ of degree n with non-zero constant term and coefficients from $[-B,B]$. It will be convenient to set $M = 2B+1$ in various places below. We set

$$\begin{aligned} P^+(n,B) &= \{f \in P(n,B) \mid f \text{ is square free and has positive leading coefficient}\}, \\ P^\circ(n,B) &= \{f \in P^+(n,B) \mid f \text{ is primitive}\}, \\ \mathrm{sq}(n,B) &= |P^+(n,B)|, \\ \mathrm{psq}(n,B) &= |P^\circ(n,B)|. \end{aligned}$$

This situation is typical of experiments to measure statistics for algorithms on polynomials, we generate them by choosing the coefficients uniformly at random from the chosen range $[-B,B]$

and accept the polynomial if it satisfies any further conditions (for efficiency we draw the leading coefficient from $[1, B]$). The requirement in $P^+(n, B)$ that f has positive leading coefficient and $f(0) \neq 0$ is clearly reasonable in the context of root finding. Insisting that such a polynomial is primitive is also well motivated, but note that the procedure of first generating coefficients does not guarantee uniform sampling, for example all 8 members of $P^+(1, 2)$ are square free and all but $2z \pm 2$ are primitive. We can correct this as follows. For a polynomial f define $|f|$ to be the maximum absolute value of its coefficients. A generated polynomial f is accepted with probability $1/c$ where $c = \lfloor B/|\operatorname{pp}(f)|\rfloor$ and $\operatorname{pp}(f)$ is the primitive part of f. The reason for this is that a primitive polynomial f of degree n and non-zero constant term accounts for $\lfloor B/|f|\rfloor$ members of $P^+(n, B)$.

Lemma 4.1 *Let p be any prime that divides $M = 2B + 1$. Then*

1. $\operatorname{sq}(n, B) \geq (1 - 1/p)^2 B(2B + 1)^n$.
2. $\operatorname{psq}(n, B) \geq (1 - 1/p)^2 (2B + 1)^n$.

PROOF. It suffices to show that $|P(n, B)| \geq (1 - 1/p)^2(M - 1)M^n$ since clearly $\operatorname{sq}(n, B) = |P(n, B)|/2$.

Define $\phi : P(n, B) \to \mathbb{F}_p[z]$ by $f \mapsto f \bmod p$, where $\mathbb{F}_p$ is the field of integers modulo p. If $g \in \mathbb{F}_p[z]$ is square free and has degree n then the same holds for every $f \in \phi^{-1}(g)$. For if $f = h^2 q$ where h is not a constant then $p \nmid \operatorname{lc}(h)$ since $p \nmid \operatorname{lc}(f)$ and so $\phi(h)$ is not a constant. Hence $\phi(f) = \phi(h)^2\phi(q)$ is not square free, which is a contradiction.

Clearly the image of ϕ contains all polynomials in $\mathbb{F}_p[z]$ of degree n. Furthermore $|\phi^{-1}(g)| = (M/p)^{n+1}$ for all $g \in \mathbb{F}_p[z]$ with $n = \deg g$ and $g(0) \neq 0$. For if $g(0) \neq 0$ then $f(0) \neq 0$, for all $f \in \phi^{-1}(g)$. The cardinality claim follows from the fact that if r is a residue modulo p then there are M/p numbers in $[-B, B]$, equivalently in $[0, 2B]$, with residue r; namely $r + qp$ for $q = 0, 1, \ldots M/p - 1$. Similarly $|\phi^{-1}(g)| = (M/p)^n(M/p - 1)$ for all $g \in \mathbb{F}_p[z]$ with $n = \deg g$ and $g(0) = 0$. For if $f \in \phi^{-1}(g)$ then $f(0) \neq 0$ provided the constant term is non-zero and this is so for $(M/p)^n(M/p - 1)$ members of $\phi^{-1}(g)$.

As is well known, the number of monic square free polynomials in $\mathbb{F}_p[z]$ of degree $n \geq 2$ is $p^n - p^{n-1}$, e.g., see Mignotte [15]. The number of these that have 0 as a root is the number of square free monic polynomials of degree $n-1$. Hence, for $n \geq 3$, the number of monic square free polynomials that do not have 0 as a root is $(p^n - p^{n-1}) - (p^{n-1} - p^{n-2}) = (1 - 1/p)^2 p^n$. For $n = 2$ the number that have 0 as a root is the number of those of form $z^2 + az$ with $a \neq 0$, i.e., there are $p - 1$ of them. Hence the number of monic square free polynomials of degree 2 that do not have 0 as a root is $(p^2 - p) - (p - 1) = (1 - 1/p)^2 p^2$. It follows that the formula $(1 - 1/p)^2 p^n$ applies to the case $n = 2$ as well. Hence the number of square free such polynomials is $(p - 1)(1 - 1/p)^2 p^n$. These correspond to $(p - 1)(1 - 1/p)^2 p^n (M/p)^{n+1} = (1 - 1/p)^3 M^{n+1}$ members of $P(n, B)$. Similarly the number of square free polynomials in $\mathbb{F}_p[z]$ of degree $n \geq 2$ with 0 as a root is $(p - 1)(p^{n-1} - p^{n-2})$. These correspond to $(p - 1)(p^{n-1} - p^{n-2})(M/p)^n(M/p - 1) = (1 - 1/p)^2 M^n (M/p - 1)$ members of $P(n, B)$. Thus the number polynomials in $P(n, B)$ with $n \geq 2$ is at least

$$(1 - 1/p)^3 M^{n+1} + (1 - 1/p)^2 M^n (M/p - 1) = (1 - 1/p)^2 (M - 1)M^n.$$

For $n = 1$ we have $|P(1, B)| = (M - 1)^2$. Since $p \leq M$ we have $(1 - 1/p)^2(M - 1)M < (M - 1)^2$. For $n = 0$ we have $|P(0, B)| = M - 1 > (1 - 1/p)^2(M - 1)$.

For the second claim suppose that $f \in P^\circ(n, B)$. This corresponds to all $cf \in P^+(n, B)$ for all c with $1 \leq c \leq \lfloor B/|f|\rfloor$, i.e., to $\lfloor B/|f|\rfloor \leq B$ distinct members of $P^+(n, B)$. Since different f give rise to different sets of members of $P^+(n, B)$ it follows that $|P^\circ(n, B)| \geq |P^+(n, B)|/B$ and the claim follows from part 1. □

We note that the assumption that p divides M is for the sake of simplicity, if $p \leq M$ is any prime then the preceding argument shows that $\operatorname{sq}(n, B) \geq (1 - 1/p)^2 p^n \lfloor M/p\rfloor^n (p\lfloor M/p\rfloor - 1)$.

An alternative approach to a (weaker) lower bound for $\operatorname{sq}(n, B)$ is to observe that $f \in P(n, B)$ is square free provided that the leading coefficient or the trailing coefficient is a square free integer.

Using $Q(B)$ to denote the number of square free integers between 1 and B we then have that

$$|P(n,B)| \geq 2Q(B)M^n + 4Q(B)(B - Q(B))M^{n-1}$$

Now $6B/\pi^2 \leq Q(B) < 6B/\pi^2 + \sqrt{B}$, see Moser and McLeod [16], so that

$$\begin{aligned} |P(n,B)| &> \frac{12}{\pi^2}BM^n + \frac{24}{\pi^2}B^2\left(1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}}\right)M^{n-1} \\ &= \frac{6}{\pi^2}(M-1)M^n + \frac{6}{\pi^2}\left(1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}}\right)(M-1)^2M^{n-1} \end{aligned}$$

Since $\mathrm{sq}(n,B) = |P(n,B)|/2$ we have

$$\mathrm{sq}(n,B) \geq \frac{6}{\pi^2}B(2B+1)^n + \frac{12}{\pi^2}\left(1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}}\right)B^2(2B+1)^{n-1}.$$

4.1 Expected breadth

We define two open discs parametrised by m. S_m is the disc given by

$$\left|z - \frac{2m^2+8m+7}{2(m+2)(m+3)}\right| < \frac{1}{2(m+2)(m+3)}.$$

This is the disc in Case 1 of Theorem 3.1. L_m is the disc given by

$$|z| < \frac{m+2}{m+3}.$$

Note that L_m is the smallest disc centred at the origin that encloses S_m.

Root Distribution Assumption: Let N_m denote the expected number of roots in L_m of members of $P^+(n,B)$ drawn uniformly at random. There is a constant m_0 such that for all $m \geq m_0$ the expected number of roots in S_m is at most $N_m(s_m/l_m)^2$ where s_m is the radius of S_m and l_m is the radius of L_m. The same applies to the situation in which members of $P^\circ(n,B)$ are drawn uniformly at random.

It follows from the second bound in Lemma 2.8 that $N_m \leq \log((n+1)B)/\log((m+3)/(m+2))$.

Shepp and Vanderbei [18] study the expected number of complex roots in a region for the case of coefficients drawn from a normal distribution while Ibragimov and Zeitouni [11] prove a more general result. Unfortunately it does not seem possible to deduce the assumption stated above form these results. An intuitive justification is that roots cluster around the unit disc and are distributed fairly uniformly by angle, see the experimental evidence presented in §4.4.

Lemma 4.2 *Assume that members of $P^+(n,B)$ are drawn uniformly at random and that the root distribution assumption holds. Then*

$$\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|] < (4.44 + m_0)\log(n+1)B.$$

The same inequality holds for $P^\circ(f)$.

Proof. As in Theorem 3.1, we consider the four cases of $\widehat{\mathbf{C}}_m(f^{II})$, $\widehat{\mathbf{C}}_m(f^{IT})$, $\widehat{\mathbf{C}}_m(f^{TI})$ and $\widehat{\mathbf{C}}_m(f^{TT})$, but with $\widehat{\mathbf{C}}$ instead of $\mathbf{C}$, using the same enumeration as there. It is shown there that $|\widehat{\mathbf{C}}(f^{IT})| + |\widehat{\mathbf{C}}(f^{TT})| \leq (2/\log 2)\log L(f)$; the inequality there is observed for $\mathbf{C}$ but it is clearly valid for $\widehat{\mathbf{C}}$.

For the first case, the disc for m is contained in the disc L_m of the root distribution assumption. The discs for $m < m_0$ are all contained in the disc centred at the origin with radius $(m_0+1)/(m_0+2)$. Thus the total number of roots of any polynomial is no more than $\log(n+1)B/\log((m_0+2)/(m_0+1))$. The transformation $f(z) \mapsto (-1)^{\deg f}f(-z)$ preserves $P^+(n,B)$ as well as $P^\circ(n,B)$

and shows that the number of roots with strictly positive real part is no more than half the total number of roots. Thus $|\cup_{0\le i<m_0}|\widehat{\mathbf{C}}_m(f^{II})|\le \log(n+1)B/2\log((m_0+2)/(m_0+1))$. Assuming that $m_0>0$,

$$\begin{aligned}\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] &= \mathbb{E}\left[\sum_{m=0}^{m_0-1}|\widehat{\mathbf{C}}_m(f)|\right]+\mathbb{E}\left[\sum_{m=m_0}^{\infty}|\widehat{\mathbf{C}}_m(f)|\right]\\ &\le \frac{\log(n+1)B}{2\log\left(\frac{m_0+2}{m_0+1}\right)}+\log(n+1)B\sum_{m=m_0}^{\infty}\frac{1}{4(m+2)^2(m+3)^2}\left(\frac{m+3}{m+2}\right)^2\frac{1}{\log\left(\frac{m+3}{m+2}\right)}.\end{aligned}$$

A simple argument shows that $\log(1+x)>2x/(x+2)$ for all $x>0$, hence $1/\log(1+x)<1/x+1/2$. Now

$$\begin{aligned}\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] &< \log(n+1)B\left(\frac{2m_0+3}{4}+\sum_{m=m_0}^{\infty}\frac{m+5/2}{4(m+2)^4}\right)\\ &\le \log(n+1)B\left(\frac{2m_0+3}{4}+\frac{m_0+5/2}{4(m_0+2)^4}+\int_{m_0}^{\infty}\frac{m+5/2}{4(m+2)^4}dm\right)\\ &= \log(n+1)B\left(\frac{2m_0+3}{4}+\frac{m_0+5/2}{4(m_0+2)^4}+\left[\frac{-3m-7}{24(m+2)^3}\right]_{m_0}^{\infty}\right)\\ &= \left(\frac{m_0+2}{2}-\frac{1}{4}+\frac{1}{8(m_0+2)^2}+\frac{7}{24(m_0+2)^3}+\frac{1}{8(m_0+2)^4}\right)\log(n+1)B\end{aligned}$$

As pointed out in the proof of Theorem 3.1, the first and third cases transform to each other by the transform $f(z)\mapsto(-z)^{\deg(f)}f(1/z)$. Hence $|\widehat{\mathbf{C}}(f^{II})|\le|\widehat{\mathbf{C}}(f^{TI})|$ and $|\widehat{\mathbf{C}}(f^{TI})|\le|\widehat{\mathbf{C}}(f^{II})|$, hence $|\widehat{\mathbf{C}}(f^{TI})|=|\widehat{\mathbf{C}}(f^{II})|$. Thus bound above applies to $\mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|]$ and so

$$\begin{aligned}\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] &+ \mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|]\\ &\le 2\left(\frac{1}{\log 2}-\frac{1}{4}+\frac{m_0+2}{2}+\frac{1}{8(m_0+2)^2}+\frac{7}{24(m_0+2)^3}+\frac{1}{8(m_0+2)^4}\right)\log(n+1)B\\ &\le 2\left(\frac{1}{\log 2}-\frac{1}{4}+\frac{m_0+2}{2}+\frac{1}{8\cdot 3^2}+\frac{7}{24\cdot 3^3}+\frac{1}{8\cdot 3^4}\right)\log(n+1)B\\ &= \left(\frac{2}{\log 2}+\frac{503}{324}+m_0\right)\log(n+1)B\\ &< (4.44+m_0)\log(n+1)B\end{aligned}$$

If $m_0=0$ the argument above becomes simpler and leads to

$$\begin{aligned}\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] &+ \mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|]\\ &\le 2\left(\frac{1}{\log 2}+\frac{29}{384}\right)\log(n+1)B<3.037\log(n+1)B.\end{aligned}$$

□

THEOREM 4.1 *Assume that members f of $P^+(n,B)$ or of $P^\circ(f)$ are drawn uniformly at random and that the root distribution assumption holds. Then*

$$\mathbb{E}[br(f)]<(4.44+m_0)\log(n+1)B+1.45\log B+5.$$

PROOF. As in Theorem 3.1 we have

$$\mathbb{E}[\mathrm{br}(f)]\le\mathbb{E}[|\,\mathrm{br}(f^{II})|]+\mathbb{E}[|\,\mathrm{br}(f^{IT})|]+\mathbb{E}[|\,\mathrm{br}(f^{TI})|]+\mathbb{E}[|\,\mathrm{br}(f^{TT})|].$$

It follows from the first part of Lemma 2.3 that $\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|]+\mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|]$ accounts for all the real roots except for the integer roots of f. By Lemmas 2.4 and 2.9 we now have

$$\mathbb{E}[\mathrm{br}(f)] \leq \mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|] + 1 + \log B/\log 2 + 4.$$

The result follows by Lemma 4.2. $\square$

4.2 Alternative hypothesis

The experimental evidence of §4.4 suggests that the Root Distribution Assumption is true, with one exception, for all $m \geq 4$ provided we replace the upper bound estimate $N_m(s_m/l_m)^2$ by $N_m(s_m/l_m)^{1+\epsilon}$ for some $\epsilon > 0$. Under this assumption the first bound in Lemma 4.2 becomes

$$\begin{aligned}
\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] &\leq \log(n+1)B\left(\sum_{m=0}^{\infty} \frac{m+5/2}{2^{1+\epsilon}(m+2)^{2(1+\epsilon)}}\right) \\
&\leq \log(n+1)B\left(\sum_{m=0}^{4} \frac{m+5/2}{2(m+2)^2} + \int_{m=4}^{\infty} \frac{m+5/2}{2^{1+\epsilon}(m+2)^{2(1+\epsilon)}}dm\right) \\
&= \log(n+1)B\left(\frac{12209}{14400} + \left[\frac{-((1+2\epsilon)m+2+5\epsilon)}{\epsilon(1+2\epsilon)2^{2+\epsilon}(m+2)^{1+2\epsilon}}\right]_4^{\infty}\right) \\
&< \log(n+1)B\left(\frac{12209}{14400} + \frac{6+13\epsilon}{\epsilon(1+2\epsilon)2^{2+\epsilon}6^{1+2\epsilon}}\right)
\end{aligned}$$

A simple argument shows that $(6+13\epsilon)/\epsilon(1+2\epsilon)2^{2+\epsilon}6^{1+2\epsilon} \leq 1/2\epsilon\log 2$ and so

$$\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|] < \log(n+1)B\left(\frac{12209}{14400} + \frac{1}{2\epsilon\log 2}\right).$$

Thus

$$\begin{aligned}
\mathbb{E}[|\widehat{\mathbf{C}}(f^{II})|]+&\mathbb{E}[|\widehat{\mathbf{C}}(f^{IT})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TI})|] + \mathbb{E}[|\widehat{\mathbf{C}}(f^{TT})|] \\
&< 2\left(\frac{1}{\log 2} + \frac{12209}{14400} + \frac{1}{2\epsilon\log 2}\right)\log(n+1)B.
\end{aligned}$$

The bound of Theorem 4.1 now becomes

$$\begin{aligned}
\mathbb{E}[\mathrm{br}(f)] &< 2\left(\frac{1}{\log 2} + \frac{12209}{14400} + \frac{1}{2\epsilon\log 2}\right)\log(n+1)B + 1.45\log B + 5 \\
&< \left(4.59 + \frac{1}{\epsilon\log 2}\right)\log(n+1)B + 1.45\log B + 5.
\end{aligned}$$

4.3 Dependence of expected breadth on the real roots

Lemma 2.4 shows that $\mathbb{E}[\mathrm{br}(f)] \leq \mathbb{E}[\mathbf{R}^+(f)] + \mathbb{E}[\mathbf{C}(f)] + 1$. In this section we show that $\mathbb{E}[\mathbf{R}^+(f)]$ is in line with the bound of Theorem 4.1 but without any assumptions.

Consider the situation in which we draw the coefficients of non-zero polynomials of degree at *most* n uniformly at random from $[-B, B]$, where $B \geq 1$. Thus the probability of drawing any given polynomial is $1/(M^{n+1} - 1)$ since the zero polynomial is excluded. Let $N(n, B)$ be the expected number of real roots. It is shown by Ibragimov and Maslova [12] that if we draw the coefficients independently from the same distribution that has mean zero and which belongs to the domain of attraction of the normal law then the expected number of real roots is equal to

$$\frac{2}{\pi}\log n + o(\log n).$$

This generalises the corresponding result of Kac [13] for the standard normal distribution. The situation of [12] covers our case, as observed by Cucker and Roy [7]. Since the transformation

$z \mapsto (-1)^{\deg(f)} z$ swaps positive and negative roots while keeping the class of polynomials invariant it follows that the number of expected positive real roots, $N^+(n, B)$, is equal to

$$\frac{1}{\pi} \log n + o(\log n).$$

Consider now drawing the members of $P^+(n, B)$, respectively $P^\circ(n, B)$, uniformly at random and denote the expected number of real roots by $N^+(n, B)$, respectively $N^\circ(n, B)$.

LEMMA 4.3 *Let p be a prime that divides $M = 2B + 1$. Then*

$$N^+(n, B) < \frac{(2 + 1/B)}{\pi(1 - 1/p)^2} \log n + o(\log n) < 2.13 \log n + o(\log n),$$

and

$$N^\circ(n, B) < \frac{(2B + 1)}{\pi(1 - 1/p)^2} (\log n + o(\log n)) < 0.72(2B + 1) (\log n + o(\log n)).$$

PROOF. Using the first part of Lemma 4.1 in the fourth line below, we have

$$\begin{aligned} N^+(n, B) &= \sum_{f \in P^+(n,B)} \frac{\mathbf{R}^+(f)}{\mathrm{sq}(n, B)} \\ &= \frac{M^{n+1} - 1}{\mathrm{sq}(n, B)} \sum_{f \in P^+(n,B)} \frac{\mathbf{R}^+(f)}{M^{n+1} - 1} \\ &\leq \frac{M^{n+1} - 1}{\mathrm{sq}(n, B)} \sum_{f \in P(n,B)} \frac{\mathbf{R}^+(f)}{M^{n+1} - 1} \\ &< \frac{M^{n+1}}{(1 - 1/p)^2 B M^n} N^+(n, B) \\ &= \frac{(2 + 1/B)}{\pi(1 - 1/p)^2} \log n + o(\log n). \end{aligned}$$

The numerical constant follows from the fact that $M \geq 3$ is odd and so $p \geq 3$. The second inequality follows similarly by using the second part of Lemma 4.1. □

4.4 Experimental evidence

The results presented here are based on data produced by Zhao Zheng, a student at the School of Informatics in Edinburgh working under the supervision of the author. The root finding package used from [19] is based on Laguerre's method (org.apache.commons.math3.analysis.solvers.LaguerreSolver).

In the tables we present the expected number of roots found in the discs S_m and L_m of the Root Distribution Assumption (denoted by $\mathbb{E}[S_m]$ and $\mathbb{E}[L_m]$ reespectively) in the first two columns. The third column gives the estimated upper bound on $\mathbb{E}[S_m]$ represented by $u_m = \mathbb{E}[L_m](s_m/l_m)^2$. The fourth column gives $e_m = (\mathbb{E}[S_m]/\mathbb{E}[L_m])/\log(s_m/l_m)$ from which we may derive an approximate value of ϵ in the modified conjecture of §4.2. The polynomials generated were both square free and primitive. A comparison was made with purely square free polynomials (for two cases), these showed no significant difference with the corresponding cases of square free and primitive polynomials, we present only the data for the square free primitive case in Tables 1–9.

In each experiment polynomials with the appropriate degree and coefficient bound were generated uniformly at random then accepted with appropriate probability if they were square free and primitive. Each experiment was run for a maximum time, this is the reason that the sample sizes vary somewhat. Once the roots were found for each polynomial the polynomial was evaluated at each root and this sample was rejected if any value was not close to 0 (within 10^{-8}). As can be seen from the tables the two hypotheses hold after an initial period and stay this way. One

n	B	$\mathbb{E}[\lvert\widehat{\mathbf{C}}\rvert]/\log(n+1)B$
10	1	1.047 349
	10	0.642 732
	100	0.435 704
20	1	1.334 795
	10	0.833 969
	100	0.586 072
30	1	1.547 431
	10	1.016 662
	100	0.740 089

Figure 2: Ratio of estimate for $\mathbb{E}[\lvert\widehat{\mathbf{C}}\rvert]$ and $\log(n+1)B$.

exception is the case for $n = 30$ and $B = 100$ (Table 9) where the result for $m = 1024$ is somewhat anomalous. It is not clear if this is a genuine trend or a software issue, unfortunately further investigation was not possible due to time constraints but it will be investigated in the future.

The bounds obtained in §§4.1, 4.2 are dominated by $c\log(n+1)B$ for some constant c. By Lemmas 2.3 and 2.4 $\mathbb{E}[\mathrm{br}(f)]$ is dominated by $\mathbb{E}[\lvert\widehat{\mathbf{C}}\rvert]$. Figure 2 shows the ratio of $\mathbb{E}[\lvert\widehat{\mathbf{C}}\rvert]$ to $\log(n+1)B$. In all cases this is well below the theoretical constant, however there is an increase as n increases for any given B.

Acknowledgement. The author is grateful to an anonymous referee for helpful comments.

References

[1] L. Ahlfors, *Complex Analysis*, McGraw-Hill Book Company, second ed., 1966.

[2] A. G. Akritas, *An implementation of Vincent's theorem*, Numerische Mathematik, 36 (1980), pp. 53–62.

[3] ——, *Elements of Computer Algebra with Applications*, J. Wiley & Sons, 1989.

[4] A. Alesina and M. Galuzzi, *A new proof of Vincent's theorem*, L'Enseignement Matheématique, 44 (1998).

[5] G. Collins and W. Krandick, *On the computing time of the continued fractions method*, Journal of Symbolic Computation, 47 (2012), pp. 1372–1412.

[6] G. E. Collins and A. G. Akritas, *Polynomial real root isolation using Descarte's rule of signs*, in Proceedings of the 1976 Symposium on Symbolic and Algebraic Computation, R. D. Jenks, ed., ACM Press, 1976, pp. 272–275.

[7] F. Cucker and M.-F. Roy, *A theorem on random polynomials and some consequences in average complexity*, J. Symbolic Computation, 10 (1990), pp. 405–409.

[8] P. Erdős and P. Turán, *On the distribution of roots of polynomials*, Ann. Math., 51 (1950), pp. 105–119.

[9] T. Ganelius, *Sequences of analytic functions and their zeros*, Ark. Math., 3 (1954), pp. 1–50.

[10] C. P. Hughes and A. Nikeghbali, *The zeros of random polynomials cluster uniformly near the unit circle*, Composition Math., 144 (2008), pp. 734–746.

[11] I. Ibragimov and O. Zeitouni, *On roots of random polynomials*, Trans. of the American Math. Soc., 349 (1997), pp. 2427–2441.

[12] I. A. Ibragimov and N. B. Maslova, *The mean number of real zeros of random polynomials. I. Coefficients with zero mean*, Theor. Probability Appl., 16 (1971), pp. 228–248.

[13] M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc., 49 (1943), pp. 314–320 and 938.

[14] W. Krandick and K. Mehlhorn, *New bounds for the Descartes method*, J. Symbolic Computation, 41 (2006).

[15] M. Mignotte, *Mathematics for Computer Algebra*, Springer, 1992.

[16] L. Moser and R. A. McLeod, *The error term for square free integers*, Canadian Mathematics Bulletin, 9 (1966), pp. 303–306.

[17] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, no. 26 in London Mathematical Society Monographs, New Series, Clarendon Press, Oxford, 2002.

[18] L. Shepp and R. Vanderbei, *The complex zeros of random polynomials*, Trans. of the American Math. Soc., 347 (1995), pp. 4365–4384.

[19] The Apache Commons Mathematics Library, *http://commons.apache.org/proper/commons-math/*.

[20] E. P. Tsigaridas and I. Z. Emiris, *On the complexity of real root isolation using continued fractions*, Theoretical Computer Science, 392 (2008), pp. 158–173.

[21] A. J. H. Vincent, *Sur la résolution des équations numériques*, Journal de mathématiques pures et appliquées, 1 (1836), pp. 341–372.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	0.700 392	0.056 823	0.002 162	0.868 987
2	1.183 154	0.038 298	0.001 155	0.989 846
4	1.953 484	0.023 188	3.768 294E−4	1.036 732
8	2.887 957	0.010 172	7.219 892E−5	1.066 128
16	3.715 613	0.002 733	8.848 721E−6	1.114 440
32	4.209 781	0.000 000	7.875 604E−7	—
64	4.465 587	0.000 000	5.883 596E−8	—
128	4.595 769	0.000 000	4.022 766E−9	—
256	4.652 807	0.000 000	2.625 284E−10	—
512	4.681 456	0.000 000	1.676 750E−11	—
1024	4.695 251	0.000 000	1.059 277E−12	—

Table 1: Results for $n = 10$, $B = 1$ based on 162803 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 2.511\,434$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	1.213 440	0.054 137	0.003 745	1.075 881
2	1.649 975	0.039 199	0.001 611	1.079 093
4	2.343 683	0.021 803	4.520 993E−4	1.093 707
8	3.193 095	0.009 437	7.982 736E−5	1.099 237
16	3.920 581	0.003 261	9.336 850E−6	1.095 493
32	4.412 287	8.946 619E−4	8.254 449E−7	1.097 806
64	4.691 315	2.783 392E−4	6.181 002E−8	1.072 740
128	4.838 881	4.638 987E−5	4.235 567E−9	1.108 062
256	4.909 407	0.000 000	2.770 067E−10	—
512	4.943 935	0.000 000	1.770 762E−11	—
1024	4.960 595	0.000 000	1.119 140E−12	—

Table 2: Results for $n = 10$, $B = 10$ based on 150895 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 3.021\,147$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	1.258 061	0.054 619	0.003 883	1.085 311
2	1.689 504	0.038 297	0.001 650	1.092 643
4	2.387 732	0.021 852	4.605 965E−4	1.097 535
8	3.229 419	0.009 366	8.073 548E−5	1.102 794
16	3.955 142	0.003 108	9.419 158E−6	1.104 266
32	4.444 238	8.447 456E−4	8.314 222E−7	1.106 149
64	4.722 547	2.034 789E−4	6.222 152E−8	1.108 002
128	4.871 740	5.549 424E−5	4.264 329E−9	1.091 527
256	4.949 247	6.166 026E − 6	2.792 546E−10	1.152 268
512	4.988 075	0.000 000	1.786 572E−11	—
1024	5.007 467	0.000 000	1.129 714E−12	—

Table 3: Results for $n = 10$, $B = 100$ based on 162179 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 3.051\,264$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	0.709 483	0.058 185	0.002 190	0.865 255
2	1.238 127	0.040 825	0.001 209	0.984 513
4	2.231 143	0.026 331	4.303 903E−4	1.038 078
8	3.910 605	0.013 800	9.776 512E−5	1.065 765
16	5.989 258	0.005 232	1.426 340E−6	1.087 910
32	7.700 692	0.001 387	1.440 635E−7	1.113 115
64	8.732 849	7.429 467E−5	1.150 589E−7	1.286 814
128	9.267 578	0.000 000	8.112 092E−9	—
256	9.538 061	0.000 000	5.381 722E−10	—
512	9.673 401	0.000 000	3.464 709E−11	—
1024	9.741 851	0.000 000	2.197 820E−12	—

Table 4: Results for $n = 20$, $B = 1$ based on 161519 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 4.063\,813$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	1.245 698	0.055 668	0.003 845	1.075 312
2	1.742 496	0.040 453	0.001 702	1.085 750
4	2.710 376	0.025 895	5.228 349E−4	1.087 480
8	4.341 311	0.014 121	1.085 328E−4	1.081 147
16	6.326 965	0.005 408	1.506 765E−5	1.091 261
32	7.928 642	0.001 556	1.483 280E−6	1.102 010
64	8.922 128	4.694 172E−4	1.175 527E−7	1.085 985
128	9.465 688	8.359 484E−5	8.285 501E−9	1.115 934
256	9.747 794	2.572 149E−5	5.500 061E−10	1.088 665
512	9.888 832	0.000 000	3.541 869E−11	—
1024	9.958 228	0.000 000	2.246 636E−12	—

Table 5: Results for $n = 20$, $B = 10$ based on 155512 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 4.459\,321$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	0.057 730	0.057 730	0.004 015	1.077 719
2	1.797 625	0.041 200	0.001 755	1.089 463
4	2.763 467	0.026 917	5.330 762E−4	1.082 962
8	4.399 762	0.013 158	1.099 941E−4	1.097 000
16	6.368 575	0.005 174	1.516 674E−5	1.099 099
32	7.962 050	0.001 518	1.489 530E−6	1.105 755
64	8.948 142	4.457 163E−4	1.178 955E−7	1.092 017
128	9.487 465	8.397 553E−5	8.304 563E−9	1.115 719
256	9.771 858	1.291 931E−5	5.513 639E−10	1.147 234
512	9.915 928	0.000 000	3.551 574E−11	—
1024	9.987 991	0.000 000	2.253 350E−12	—

Table 6: Results for $n = 20$, $B = 100$ based on 154807 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 4.483\,273$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	0.712 845	0.062 848	0.002 200	0.840 221
2	1.236 715	0.043 455	0.001 208	0.966 166
4	2.237 816	0.027 267	4.316 775E−4	1.030 611
8	4.165 344	0.015 311	1.041 336E−4	1.058 073
16	7.188 359	0.006 984	1.711 905E−5	1.071 477
32	10.283 021	0.002 031	1.923 734E−6	1.101 200
64	12.413 197	4.651 498E−4	1.635 490E−7	1.123 390
128	13.588 310	2.620 562E−5	1.189 411E−8	1.261 841
256	14.189 264	0.000 000	8.006 101E−10	—
512	14.489 239	0.000 000	5.189 591E−11	—
1024	14.638 729	0.000 000	3.302 585E−12	—

Table 7: Results for $n = 30$, $B = 1$ based on 152639 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 5.313\,858$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	1.278 439	0.064 402	0.003 946	1.033 863
2	1.779 597	0.044 520	0.001 738	1.064 189
4	2.778 407	0.027 307	5.359 581E−4	1.080 859
8	4.693 874	0.015 662	1.173 468E−4	1.076 335
16	7.673 410	0.006 568	1.827 420E−5	1.091 035
32	10.658 205	0.002 218	1.993 923E−6	1.094 462
64	12.713 808	5.888 049E−4	1.675 097E−7	1.100 044
128	13.879 733	1.766 415E−4	1.214 920E−8	1.080 896
256	14.492 444	2.616 910E−5	8.177 166E−10	1.120 814
512	14.799 885	0.000 000	5.300 854E−11	—
1024	14.951 463	0.000 000	3.373 139E−12	—

Table 8: Results for $n = 30$, $B = 10$ based on 152852 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 5.832\,158$.

m	$\mathbb{E}[S_m]$	$\mathbb{E}[L_m]$	u_m	e_m
1	1.349 239	0.084 909	0.004 164	0.956 872
2	1.859 514	0.054 657	0.001 816	1.017 675
4	2.849 841	0.027 416	5.497 379E−4	1.085 871
8	4.753 587	0.015 555	1.188 397E−4	1.080 022
16	7.703 725	0.007 004	1.834 640E−5	1.081 732
32	10.677 306	0.002 077	1.997 496E−6	1.103 179
64	12.735 517	5.296 352E−4	1.677 957E−7	1.111 906
128	13.895 209	1.533 154E−4	1.216 275E−8	1.094 584
256	14.506 910	0.000 000	8.185 328E−10	—
512	14.822 266	0.000 000	5.308 870E−11	—
1024	14.980 327	6.968 884E−6	3.379 651E−12	1.001 429

Table 9: Results for $n = 30$, $B = 100$ based on 143495 square free and primitive polynomials generated at random. $\mathbb{E}[|\widehat{\mathbf{C}}|] = 5.949\,692$.