An Efficient and Exact Subdivision Algorithm for Isolating Complex Roots of a Polynomial and its Complexity Analysis

Michael Sagraloff

Max Planck Institute for Informatics, Saarbrücken

Chee K. Yap

Courant Institute, NYU, New York

Abstract

We introduce an exact subdivision algorithm CEVAL for isolating complex roots of a square-free polynomial. The subdivision predicates are based on evaluating the original polynomial or its derivatives, and hence is easy to implement. It can be seen as a generalization of a previous real root isolation algorithm called EVAL. Under suitable conditions, the algorithm is applicable for general analytic functions.

We provide a complexity analysis of our algorithm on the benchmark problem of isolating all complex roots of a square-free polynomial with Gaussian integer coefficients. The analysis is based on a novel technique called δ-clusters. This analysis shows, somewhat surprisingly, that the simple EVAL algorithm matches (up to logarithmic factors) the bit complexity bounds of current practical exact algorithms such as those based on Descartes, Continued Fraction or Sturm methods. Furthermore, the more general CEVAL also achieves the same complexity.

Key words: root isolation, subdivision methods, complex roots, Bolzano method

1 Introduction

Root finding might be called the *Fundamental Problem of Algebra*, after the Fundamental Theorem of Algebra [39,41,44]. The literature on root finding is extremely rich, with a large classical literature. The work of Schönhage [39]

Email addresses: msagralo@mpi-inf.mpg.de (Michael Sagraloff), yap@cs.nyu.edu (Chee K. Yap).

marks the beginning of complexity-theoretic approaches to the Fundamental Problem. Pan [33] provides a history of root-finding from the complexity view point; see McNamee [22] for a general bibliography. The root finding problem can be studied as two distinct problems: root isolation and root refinement. In the complexity literature, the main focus is on what we call the **benchmark problem**, that is, isolating all the complex roots of a polynomial f of degree n with integer coefficients of at most L bits. Let $T(n, L)$ denote the (worst case) bit complexity of this problem. There are three variations on this benchmark problem:

- We can ask for only the real roots. Special techniques apply in this important case [7,16]. E.g., Sturm [20,37,9], Descartes [5,38,27,12,10], and continued fraction methods [1,40].
- We can seek the arithmetic complexity of this problem, that is, we seek to optimize the number $T_A(n, L)$ of arithmetic operations.
- We can add another parameter $p > 0$, and instead of isolation, we may seek to approximate each of the roots to p relative or absolute bits.

Schönhage achieved a bound of $T(n, L) = \widetilde{O}(n^3 L)$ for the benchmark isolation problem where $\widetilde{O}$ indicates the omission of logarithmic factors. This bound has essentially remained intact. Pan and others [33,29] have given theoretical improvements in the sense of achieving $T_A(n, L) = \widetilde{O}(n^2 L)$ and $T(n, L) = T_A(n, L) \cdot \widetilde{O}(n)$, thus achieving record bounds in both bit complexity and arithmetic complexity. Theoretical algorithms designed to achieve record bounds for the benchmark problem have so far not been used in practice. Moreover, the benchmark problem is inappropriate for some applications. For instance, we may only be interested in the first positive root (as in ray shooting in computer graphics), or the roots in some limited neighborhood. In the numerical literature, there are many algorithms that are widely used and effective in practice but lack a guarantee on the global behavior (cf. [33] for discussion). Some "global methods" such as the Weierstrass or Durant-Kerner method that simultaneously approximates all roots seem ideal for the benchmark problem and work well in practice, but their convergence and/or complexity analysis are open. Thus, the benchmark complexity, despite its theoretical usefulness, has limitation as sole criterion in evaluating the usefulness of root isolation algorithms.

There are two sources of literature on "practical" root isolation algorithms: (1) One is the exact computation literature, providing algorithms used in various algebraic applications and computer algebra systems. Such exact algorithms have a well-developed complexity analysis and there is considerable computational experience especially in the context of cylindrical algebraic decomposition. The favored root isolation algorithms here, applied to the benchmark problem, tend to lag behind the theoretical algorithms by a factor of nL. Nevertheless, current experimental data justify their use [38,16]. (2) The other is the numerical literature, mentioned above. Although numerical algorithms traditionally lack any exactness guarantees, they have many advantages that practitioners intuitively understand: compared to algebraic methods, they are easier to implement and their complexity is more adaptive. Hence, there is a growing interest to constructing numerical algorithms that are exact and efficient. *This paper is a contribution along this line.*

1.1 The Subdivision Approach

Among the practical exact root isolation algorithms, the subdivision paradigm is perhaps the most widely used. This paradigm is a generalization of binary search in which we begin with a domain (say a box $B_0 \subseteq \mathbb{C}$) and recursively subdivide the boxes to search for roots. Unlike the theoretical algorithms or global methods above, subdivision algorithms have a strong advantage of being "local" as they can restrict computational effort only to the given initial box B_0, in order to find roots near B_0. If there are few or no roots in B_0, such methods can terminate quickly. The "subdivision" terminology derives from the use of such algorithms in meshing curves and surfaces [21]; root isolation is just meshing in 1-D. The principle action of subdivision algorithms is the **subdivision phase** that operates on a queue Q containing subboxes of B_0. Initially, $Q = \{B_0\}$. In each iteration, a box B is removed from Q and tested with an **exclusion predicate** C_{out} and an **inclusion predicate** C_{in}. If $C_{out}(B)$ holds, B is discarded; if $C_{in}(B)$ holds, then B is output. Otherwise, we subdivide B into four children boxes and put them back into Q. For root isolation, $C_{out}(B)$ guarantees that B has no zeros, and $C_{in}(B)$ guarantees that there is a unique zero in B. For real roots, we would use intervals instead of boxes. The general structure of many subdivision algorithms is fairly simple; in [21], the "generic subdivision algorithm" is viewed as a sequence of four phases: boundary, subdivision (described above), refinement, and construction. For our root isolation problem, we can omit the boundary and refinement phases. The construction phase amounts to ensuring that the output boxes are pairwise disjoint. Since finding a root is metaphorically like "finding a needle in a hay stack", an efficient exclusion predicate C_{out} is crucial to the success of such algorithms. *Here, numerical forms of C_{out} such as those used in this paper are relatively cheap and have advantages over algebraic ones.*

1.2 Three Principles for Subdivision

We compare three general principles used in subdivision algorithms for real root isolation: theory of Sturm sequences, Descartes' rule of sign, and the Bolzano principle. These principles are used in the exclusion and inclusion predicates of the corresponding algorithms. Continued Fraction Solvers can be viewed as extended Descartes methods since they use Descartes' Rule of Sign as their main predicates, but combine it with an exclusion predicate based on a root bound. Although the Continued Fraction method has not been proven to achieve better worse case complexity [40,23,13], a significant speed up can be observed in practice [13,1,15]. This paper concentrates on the Bolzano principle, also known as the Bolzano theorem. It is simple and intuitive: *if a continuous real function $f(x)$ satisfies $f(a)f(b) < 0$ then there is a point c between a and b such that $f(c) = 0$. Furthermore, if f is differentiable and f' does not vanish on (a, b) then this root is unique in (a, b).* In recent years, algorithms based on the first two principles have been called (respectively) Sturm method [37,20,9] and the Descartes method [1] [5,12,19,6]. By analogy, we may call algorithms based on the Bolzano principle the **Bolzano method** [25,4,3]. Note that the Bolzano principle is an analytic one, while Sturm and Descartes are more algebraic. The complexity analysis of Bolzano methods seems to be new, prompted in part by interest in exact numerical methods in meshing algebraic surfaces [35,21,2]. Perhaps it is no surprise that Bolzano methods could outperform the more sophisticated algebraic methods in practice. *Somewhat surprisingly, the results of this paper indicate that Bolzano methods could also match the theoretical complexity of algebraic methods as well.*

There are two basic complexity measures for subdivision algorithms: the subdivision tree size $S(n, L)$ and the bit complexity $P(n, L)$ of the subdivision predicates. Thus, $T(n, L) \leq S(n, L)P(n, L)$. But the analysis in this paper shows that $T(n, L)$ may be smaller than $S(n, L)P(n, L)$ by a factor of n. Tree size in the Sturm method is optimal in a very strong sense: for any polynomial $f(x)$ and for any interval I_0, the Sturm subdivision tree is minimum in an absolute, not asymptotic, sense. For the benchmark problem where $f(x)$ has degree n and L-bit integer coefficients, this tree size was shown to be $O(n(L + \log n))$ by Davenport [8] in 1985. This is optimal if $L \geq \log n$ [12]. Modern algorithmic treatment of the Descartes method began with Collins and Akritas [5]. The tree size in the Descartes method was only recently proven to be $O(n(L + \log n))$ [12]. In this paper, we will prove that the tree size in the Bolzano method is $\widetilde{O}(n(L + \log n))$ for real roots. Furthermore, for our extension of the Bolzano method for complex roots the corresponding tree size is $\widetilde{O}(n^2(L + \log n))$. *Despite this larger tree size, we prove that both real and complex Bolzano have $\widetilde{O}(n^4L^2)$ bit complexity, matching Descartes and Sturm.*

Johnson [16] has shown empirically that the Descartes method is more efficient than Sturm. Rouillier and Zimmermann [38] implemented a highly efficient exact real root isolation algorithm based on the Descartes method. Since their theoretical complexity bounds are indistinguishable, any practical advantage of Descartes over Sturm must be derived from the fact that the predicates in the Descartes method are cheaper. We believe that the Bolzano method

[1] Note that we avoid the possessive “Descartes method” as Descartes did not envision such algorithms.

has a similar advantage over Descartes. This has not yet been demonstrated. Nevertheless, it has an advantage of a different kind: *The Bolzano method is applicable to a much wider class of functions — most common analytic functions are amenable. Furthermore, this paper shows that the Bolzano method can be extended to the domain of complex numbers.*

1.3 Contributions of this paper

1. Our root isolation algorithm is a contribution to a growing list of exact algorithms based on numerical (as opposed to algebraic) techniques and simple subdivision. Numerical subdivision methods are widely used in practice, being easy to implement and having adaptive complexity. In comparison to existing exact practical methods for real root isolation (Descartes, Sturm, Continued Fraction) it extends to most common analytic functions and also to the domain of complex numbers.

2. This paper represents one of the first complexity analysis of exact numerical subdivision methods based on the Bolzano principle. It uses a novel technique of δ-clusters, from which we expect other application as well. Surprisingly, our analysis shows that the simple Bolzano principle already yields an algorithm EVAL whose worst-case bit-complexity matches those of more sophisticated methods like Sturm or Descartes. Also unexpected is that the complex analogue CEVAL achieve the same bit complexity as EVAL (despite the fact that in terms of tree size, that of CEVAL is quadratic in the real size).

1.4 Overview of Paper

Section 2 reviews related work. The algorithm is presented in Section 3. Therein we also summarize the results of our complexity analysis accomplished in Section 5 by the use of the new concept of $\delta-$clusters. Section 4 develops basic tools for proving the correctness of the algorithm. Section 6 addresses issues in implementing our algorithm exactly. We conclude in Section 7.

2 Prior Work

The main distinction among the various subdivision algorithms is the choice[2] of tests or predicates. One approach is based on doing root isolation on the boundary of the boxes. Pinkert [34] and Wilf [43] (see also [44]) use Sturm-like sequences, while Collins and Krandick [18] considered Descartes method. Such approaches are related to topological degree methods [28], which go back to Brouwer (1924).

[2] We shall use the terms “predicate” and “test” interchangeably.

2.1 Weyl's Approach

We briefly review Pan's work [33,30–32] as it is closest to our approach. Pan regards his work as a refinement of Weyl's Exclusion Algorithm (1924); this algorithm was also the basis of work by Henrici and Gargantini (1969) and Renegar (1987) (see [33]). The predicates are based on estimating the distance λ_B from the midpoint $m(B)$ of a box B to the nearest zero of the input polynomial $f(z)$. To estimate this distance, first shift $m(B)$ to the origin by the Taylor shift $f_B(z) := f(z + m(B))$. Then consider

$$g_B(z) := z^n f_B(1/z) = \sum_{i=0}^{n} a_i z^i$$

and find an estimate on the largest absolute value of the roots $\xi_1, \dots, \xi_n$ of g_B. As the roots of g_B are the reciprocals of the roots of f_B this gives us an estimate on λ_B. One such estimate from van der Sluis (1970) is

$$\frac{T}{n} \leq \max_j |\xi_j - m(B)| < 2T$$

where $T = \max_{i\geq 1} |a_{n-i}/a_n|^{1/i}$. This gives an (relative) error factor of $2n$ between the upper and lower estimates. A more sophisticated estimate from Turan (1968), using $O(n \ln n)$ arithmetic operations, yields a constant error factor (say 5). We need to improve these error factors to $1+\epsilon$ for a small $\epsilon > 0$. To do this, apply the above proximity test to the polynomial g_N, obtained by the Graeffe iteration

$$g_0(z) := g_B(z)/a_n, \text{ computing } g_{i+1}(z) := (-1)^n g_i(\sqrt{z})g_i(-\sqrt{z})$$

for $i = 0, \dots, N - 1$. Then the zeros of $g_k(z)$ are the 2^k-th powers of the zeros of $g_B(z)$. The proximity error reduces to $5^{1/N}$ (or $(2n)^{1/N}$ for the Sluis estimate) which is smaller than $1 + \epsilon$ if N is chosen to be sufficiently large. Pan provided the following complexity analysis: let us count the number of proximity tests in depth h of the subdivision tree. There are $\leq 4nh$ tests since each zero accounts for 4 squares in each step, assuming that the relative error is less than 1.4. Since each test takes $O(n \ln n)$ arithmetic operations, so the total is $O(n^2 h \ln n)$ arithmetic operations. If 2^{-h} is less than the root separation bound, then $h = O(n(L+\ln n))$. So the number of overall arithmetic operations is $O(n^3 \ln n(L+\ln n))$. However, Pan shows that exclusion test can be combined with Newton-like accelerations to finally achieve the record bound of $O(n^2 \ln n \ln(hn))$. Concerning his method, Pan noted that "*there remains many open problems on the numerical implementation of Weyl's algorithm and its modification*" [33, p. 216]; in particular, "*proximity tests should be modified substantially to take into account numerical problems ... and controlling the precision growth*" [33, p. 193]. In contrast, the details of implementing the subdvision algorithm in the present paper will be fully fleshed out.

2.2 The EVAL Algorithm

Our current work has roots in two prior lines of work: on one hand, it is related to our work on subdivision methods based on Sturm sequences and Descartes' Rule of Sign [19,11,10,24,9,12]. On the other hand, it arose from the surface meshing algorithms of Plantinga-Vegter [35]. Indeed, EVAL is the 1-dimensional analog of mesh generation in higher dimensions [3]. EVAL is an exact computation form of a machine floating point algorithm from Mitchell [25] who used it in ray tracing. He attributes ideas to Moore [26]. The key tool in the PV algorithm and its extensions [21,2] is the use of interval functions, evaluated on axes-parallel boxes.

Because we view our complex root algorithm as generalization of EVAL, let us briefly recall the latter algorithm. Suppose f has only simple roots in an interval $I_0 = [a, b]$ and we want to isolate the roots of f in I_0. Assume that we have interval formulations of f and its derivative f', denoted $\square f, \square f'$. Following [3], we call $\square f$ a **box function** for f if $\square f$ is an inclusion function (i.e., $f(I) \subseteq \square f(I)$) and convergent (i.e., $\lim_{i\to\infty} \square f(I_i) = f(\lim_i I_i)$ where each I_{i+1} is properly contained in I_i). EVAL depends on two predicates which we call C_0 and C_1 on real intervals I:

$$\begin{cases} C_0(I) : 0 \notin \square f(I) \\ C_1(I) : 0 \notin \square f'(I). \end{cases} \tag{1}$$

Clearly, if $C_0(I)$ holds then f has no zeros in I. If $C_1(I)$ holds then f has at most one zero in I. Moreover, the interior of I has exactly one zero iff the following condition holds:

$$f(a)f(b) < 0, \qquad \text{where } I = [a, b]. \tag{2}$$

We then say that I **passes the sign-change test**. Thus C_0 is an exclusion predicate. Similarly, C_1 in combination with the sign-change test (2) provides an inclusion predicate. The algorithm uses a queue Q (a simple list suffice) for processing the intervals:

```
EVAL($I_0$):
       $Q \leftarrow \{I_0\}$.
       While $Q$ is non-empty
              Remove $I$ from $Q$.
1.            If $C_0(I)$ holds, discard $I$.
2.            Else if $C_1(I)$ holds,
3.                   If $I$ passes the sign-change test (2), output $I$.
4.                   Else, discard $I$.
5.            Else
6.                   If $f(m) = 0$, output $[m, m]$ where $m = m(I)$ is the midpoint.
7.                   Split $I$ into $I', I''$ at $m$, and put both intervals into $Q$.
```

Termination and correctness are easy to see. The output intervals either have the exact form $[m, m]$ or are regarded as open intervals (a, b). This algorithm is easy to implement exactly if we assume that all intervals are represented by dyadic numbers, and $\square f, \square f'$ are computable functions on dyadic intervals, and the sign of f on dyadic numbers are computable. Obviously, this algorithm is an analytic one – we can use it to find simple roots of most common analytic functions f.

In this paper, the predicates (1) are assumed to be implemented as

$$\begin{cases} C_0(I) : |f(m)| > \sum_{k\geq 1} \frac{|f^{(k)}(m)|}{k!} \left(\frac{w(I)}{2}\right)^k, \\ C_1(I) : |f'(m)| > \sum_{k\geq 2} \frac{|f^{(k)}(m)|}{(k-1)!} \left(\frac{w(I)}{2}\right)^{k-1}. \end{cases} \tag{3}$$

where $m = (a+b)/2$ and $w(I) = (b-a)$. This is closely related to the centered form used in [3,36].

3 A New Complex Root Algorithm and its Complexity

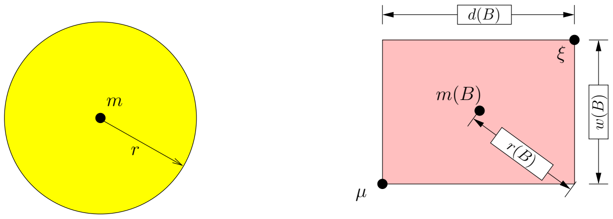
In this section, we will state our main results in three parts. We will (1) describe our main algorithm called CEVAL, (2) prove its correctness, and (3) state bounds on its complexity. Details of the correctness proof and complexity bounds are deferred to Sections 4 and 5 respectively.

Throughout this paper, we fix a square-free polynomial $f \in \mathbb{C}[z]$. We also write it as $f(z) = f(x+\mathbf{i}y) = u(x, y)+\mathbf{i}v(x, y)$ where $\mathbf{i} = \sqrt{-1}$, $x = \texttt{Re}(z)$ and $y = \texttt{Im}(z)$ are real and imaginary parts of z, and $u, v : \mathbb{R}^2 \to \mathbb{R}$. If $z' = x' + \mathbf{i}y'$ then we write $\langle z, z'\rangle = xx' + yy'$. Absolute value is denoted $|z| := \sqrt{\langle z, z\rangle}$. Sometimes, instead of viewing $u(x, y)$ as a function of two real variables, we view u as a real-valued complex function, writing $u(z) = u(x + \mathbf{i}y)$ for $u(x, y)$. Similar remarks hold for $v(x, y)$. Let $S^1 = [0, 2\pi)$ denote the set of angles in radians. Then $\arg(\xi) \in S^1$ denotes the argument of a complex number $\xi \in \mathbb{C}$. If $(x, y) \in \mathbb{R}^2$, we also write $\arg(x, y)$ for $\arg(x + \mathbf{i}y)$.

3.1 Complex Geometry

We use two basic shapes in our algorithms: disks and boxes. These are illustrated in Figure 1.

Let $\xi, \mu \in \mathbb{C}$ and $r > 0$. Let $D_r(m)$ denote the disk of radius $r > 0$ centered at $m \in \mathbb{C}$. We write "$\xi \leq \mu$" if $\texttt{Re}(\xi) \leq \texttt{Re}(\mu)$ and $\texttt{Im}(\xi) \leq \texttt{Im}(\mu)$. A subset $B \subseteq \mathbb{C}$ is called a **box** if $B = \{z \in \mathbb{C} : \xi \leq z < \mu\}$ for some $\xi \leq \mu$. Also, let $B(\xi, \mu)$ denote the smallest box that contains ξ, μ. The **midpoint** of $B(\xi, \mu)$



Disk $D_r(m)$ Box $B = B(\mu, \xi)$

Fig. 1. Two geometric shapes in the complex plane: disk and box

is $m(B) := (\xi + \mu)/2$. If $\xi < \mu$, then the **width**, **diameter**, and **radius** of $B(\xi, \mu)$ are (respectively) given by:

$$\begin{aligned} w(B) &:= \min\left\{\mathtt{Re}(\mu) - \mathtt{Re}(\xi), \mathtt{Im}(\mu) - \mathtt{Im}(\xi)\right\} \\ d(B) &:= \max\left\{\mathtt{Re}(\mu) - \mathtt{Re}(\xi), \mathtt{Im}(\mu) - \mathtt{Im}(\xi)\right\}, \\ r(B) &:= \frac{1}{2}\sqrt{w(B)^2 + d(B)^2}. \end{aligned}$$

We can split a box B into four equally dimensioned subboxes, called the **children** of B. The boundary of a region $R \subseteq \mathbb{C}$ is denoted ∂R (R is usually a disk or a box). A box B or disk D is said to be **isolating** if it contains exactly one zero of $f(z)$. Our goal is to find isolating disks for each of the complex zeros of $f(z)$ in a given box $B_0 \subseteq \mathbb{C}$.

3.2 Complex Analogues of the C_0 and C_1 Predicates

The EVAL algorithm in 2.2 is based on the interval predicates, C_0 and C_1 in (1). We now provide the complex analogues of these predicates; disks will now play the role of intervals. For $m \in \mathbb{C}$ and $K, r > 0$, define the test T_K^f:

$$T_K^f(m,r): \quad |f(m)| > K \sum_{k \geq 1} \left| \frac{f^{(k)}(m)}{k!} \right| r^k. \tag{4}$$

Since f is fixed in this paper, we simply write $T_K(m,r)$ for $T_K^f(m,r)$. Also, when f' is used in place of f, we may write $T'_K(m,r)$ for $T_K^{f'}(m,r)$. Moreover, for any disk D, we may also write $T_K(D)$ for $T_K(m(D), r(D))$, etc. Our first lemma shows that these tests (for suitable K) provide the analogues of the C_0 and C_1 predicates in (1):

Lemma 1. Consider any disk D:

(i) If $T_1(D)$ holds then D contains no zeros of f.

(ii) If $T'_{\sqrt{2}}(D)$ holds, then D has at most one zero of f.

Thus, the test $T_1(D)$ serves as an exclusion predicate for the disk D. Part(i) is obvious while Part(ii) is in Lemma 8 in Section 4.1.

3.3 The Eight Point Test

To extend the test $T'_{\sqrt{2}}(D)$ in Lemma 1 into an inclusion predicate, we need the analogue of the sign-change test (2). We now try to detect points where the curves $u = 0$ and $v = 0$ cross the boundary of D. Such points can be identified with angles as follows.

Let ϕ, ϕ', θ be angles. We say that there is a u**-crossing** of $D_r(m)$ **at** ϕ if $u(m + re^{\mathbf{i}\phi}) = 0$. We also need crossings not to be too close together: say ϕ and ϕ' are θ**-separated** if $\theta \le |\phi - \phi'| \le 2\pi - \theta$. Note that this notion is non-vacuous only if $\theta \le \pi$, and non-trivial only if $\theta \ge 0$. The notion of v**-crossing** is similarly defined.

In order to detect such crossings, we device a finitistic test based on a set of canonical points on the boundary of $D_{4r}(m)$. The **main compass points** of $D_{4r}(m)$ are the 8 points $m + 4r \cdot e^{\mathbf{i}j\pi/4}$ for $j = 0, 1, \ldots, 7$. We give them standard labels: the four **cardinal points** are N, S, E, W and four **ordinal points** are NE, SE, SW, NW, as illustrated in Figure 2.

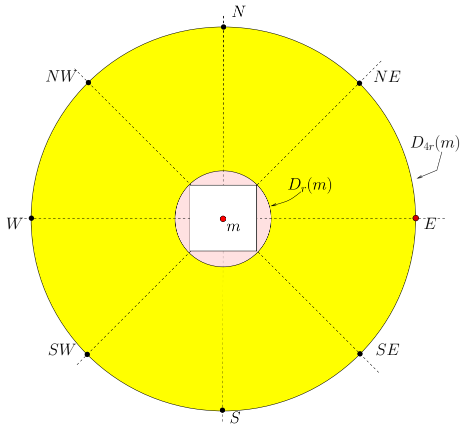


Fig. 2. 8 compass points on D_{4r}.

The boundary $\partial D_{4r}(m)$ is subdivided by the main compass points into 8 arcs

$$A_j := \left\{ m + 4re^{\mathbf{i}\theta} : j\pi/4 \le \theta < (j+1)\pi/4 \right\}.$$

For instance, the endpoints of A_0 are E and NE. We extend the idea of crossings to arcs: say there is an **arc-wise** u**-crossing** of $D_{4r}(m)$ **at** A_j if

$$\begin{cases} u(m + 4re^{\mathbf{i}j\pi/4}) \cdot u(m + 4re^{\mathbf{i}(j+1)\pi/4}) < 0, \text{ or} \\ u(m + 4re^{\mathbf{i}j\pi/4}) = 0. \end{cases} \tag{5}$$

If there is an arc-wise u-crossing at A_j then there is an u-crossing at some $\phi \in [j\pi/4, (j+1)\pi/4)$. In a similar way we also define arc-wise v-crossings.

The 8**-point test** applied to the disk $D_{4r}(m)$ amounts to the following two conditions:

- There are exactly two arcwise u-crossings at A_j and A_k and exactly two arcwise v-crossings at $A_{j'}$ and $A_{k'}$. Note that these arcs are on the boundary of $D_{4r}(m)$, not $D_r(m)$.

- These pairs of crossings are **interleaving** in the following sense: either $j < j' < k < k'$ or $j' < j < k' < k$.

If any of these conditions does not hold, we say the disk **fails** the 8-point test.

THEOREM 2 (Success of 8-Point Test). Suppose $T'_6(m, 4r)$ holds and the 8-point test is applied to $D_{4r}(m)$.

(i) If $D_{4r}(m)$ fails the 8-point test, then $D_r(m)$ is not isolating.

(ii) If $D_{4r}(m)$ passes the 8-point test, then $D_{4r}(m)$ is an isolating disk.

We view Theorem 2 as providing a "weak" inclusion predicate for the disk $D_r(m)$ because, in case the predicate holds, we do not guarantee an isolated root in $D_r(m)$, but only in $D_{4r}(m)$. Most of Section 3.6 is devoted to the proof of this theorem.

3.4 The Root Isolation Algorithm CEVAL

We present the complex analogue of EVAL, called CEVAL:

```
CEVAL($B_0, f$):
        Input: Box $B_0$, and polynomial $f(z)$ with only simple roots.
        Output: List $\mathcal{L}$ of pairwise disjoint isolating disks with centers in $B_0$.
        ____________________________________________________________
        $Q \leftarrow \{B_0\}$. $\mathcal{L} \leftarrow \emptyset$.
        While $Q$ is non-empty
                Remove $B$ from $Q$. Let $m = m(B)$ and $r = r(B)$.
1.              If $T_1(m, r)$ holds, discard $B$.
2.              Else if $T'_6(m, 4r)$ and $T'_{\sqrt{2}}(m, 8r)$ hold:
2.1                     If $D_{4r}(m)$ fails the 8-point test, discard $B$.
2.2                     Else if $D_{4r}(m)$ intersects any disk $D'$ in $\mathcal{L}$,
                                replace $D'$ by the smaller of $D_{4r}(m)$ and $D'$.
2.3                     Else insert $D_{4r}(m)$ into $\mathcal{L}$.
3.              Else
                        Split $B$ into four children and insert them into $Q$.
```

3.5 Remarks on CEVAL

Note that CEVAL is described within an algebraic RAM model of computation. To implement CEVAL exactly, we need to attend to several details. We make some preliminary remarks here, deferring other details to Section 6.

(i) This description of CEVAL is close to what one can directly implement using an arbitrary precision floating point package. Irrational operations (e.g., in the definition of $r(B)$ can easily be replaced by a dyadic approximation. E.g., Instead of $T_K(m, r(B))$, you can use the predicate $T_K(m, \delta(B))$ where $\delta(B) := \frac{3}{4}w(B) > r(B)$ is a dyadic value. Furthermore, you may replace $K = \sqrt{2}$ by $K = 3/2$ and the ordinal compass points $SE, NE, NW, SW = m + w(B)\left(\pm\frac{1}{2\sqrt{2}} \pm \frac{1}{2\sqrt{2}}\mathbf{i}\right)$ by $SE, NE, NW, SW = m + w(B)\left(\pm\frac{20}{29} \pm \frac{21}{29}\mathbf{i}\right)$, respectively. The last replacement is justified by Theorem 14.

(ii) Note that in Step 2, we not only require $T'_6(m, 4r)$, but also $T'_{\sqrt{2}}(m, 8r)$. This is to ensure that whenever two discs in $\mathcal{L}$ overlap, we can discard either one of them according to Lemma 1 (in Step 2.2, we discard the larger one).

(iii) In Step 2.2, there is an implicit search of the list $\mathcal{L}$ for disks that intersects $D_{4r}(m)$. For simplicity, we may assume a simple linear search. Since the size of $\mathcal{L}$ is non-decreasing, each search time is at most proportional to the output size of $\mathcal{L}$, which is at most n.

(iv) The reader may also note that although an isolating disk D in $\mathcal{L}$ is centered in B_0, there is no guarantee that the isolated root $z_0 \in D$ actually belongs to B_0. If we like, we could refine this algorithm with an additional parameter $\varepsilon > 0$ and guarantee that $z_0 \in B_0 \oplus D_\varepsilon(0)$ (where $\oplus$ denotes Minkowski sum). This refinement does not seem necessary in practice.

3.6 Correctness Statement

This is comprised of three claims:

THEOREM 3. *(Correctness)*

(a) *The algorithm halts.*

(b) *Throughout the algorithm, $\mathcal{L}$ is a list of pairwise disjoint isolating disks. Each disk is centered at some point of B_0.*

(c) *At termination, each zero of $f(z)$ in B_0 is isolated by some disk in $\mathcal{L}$.*

Claim (a) will be proven in a much stronger form when we give explicit complexity bounds later. However, it is instructive to see that, in general, halting is guaranteed if $f(z)$ has only simple roots in B_0. If the algorithm does not halt, then there is an infinite sequence $(B_0, B_1, B_2, \ldots)$ of boxes where B_{i+1} is a child of B_i, and each B_i fails the T_1, T'_6, and $T'_{\sqrt{2}}$ predicates. Thus the sequence converges to a point $z^* = \cap_i B_i$. By the convergence of these disk predicates, this implies that $f(z^*) = f'(z^*) = 0$, contradicting our assumption that f has simple roots in B_0.

To see (b), observe that a disk $D_{4r}(m)$ is only inserted into $\mathcal{L}$ in Steps 2.2 or 2.3. This happens after it passes the T'_6-test, the 8-point test, and the $T'_{\sqrt{2}}$-test on $D_{8r}(m)$. Lemma 1(ii) guarantees that such disks are isolating.

To see claim (c), we must show that no discarded box $B \subseteq B_0$ contains a root of f. Observe that boxes $B \subseteq B_0$ are discarded in one of three steps of the algorithm: Steps 1, 2.1, or 2.2. Step 1 is justified by Lemma 1(i) and Step 2.1 is justified by Lemma 1(i). Finally, Step 2.2 is justified by remark (ii) in 3.5.

It is important to note that a disk in $\mathcal{L}$ might isolate a root of f that lies outside B_0.

3.7 Complexity Results

We now summarize the results of the complexity analysis of our algorithm; the actual proofs are found in Section 5. For this purpose, we consider the benchmark problem of isolating all the roots of a square-free polynomial of degree n with coefficients that are L-bit Gaussian integers. In fact, there is little complexity difference between Gaussian integers and ordinary rational integers. The initial start box may be assumed to be $B_0 = B(-2^L(1+\mathbf{i}), 2^L(1+\mathbf{i}))$. According to Cauchy's bound [44] B_0 contains all the roots of f.

As noted, the efficiency of subdivision methods crucially depends on the choice of the exclusion predicate. By simple modification, you can reformulate our algorithm by using any box function. You may apply a simple method such Horner's scheme applied to the initial polynomial f at each step. Although this is relatively cheap, this approach may suffer from strong overestimation for many boxes, in particular for those where the higher derivatives are small in relation to those at the origin. This may result in a huge subdivision tree, rendering the overall algorithm useless. Our chosen predicates T_K^f are based on the Taylor expansions at the centers of boxes, and thus they profit from the local information on the values of the higher order derivatives. We remark that this is common to other efficient methods for real root isolation – there are implicit Taylor expansions in the Descartes method, for instance.

3.7.1 Cluster Analysis and Tree Size

Let us denote the subdivision trees, induced by CEVAL and EVAL by T^{CE} and T^{EV} respectively. Before stating our results that bounds the sizes of T^{CE} and T^{EV}, it is instructive to first give a crude estimate.

We start with a reformulation of the predicate T_K^f:

$$T_K^f(m,r) : \sum_{k\geq 1} \left| \frac{f^{(k)}(m)}{f(m)} \right| \frac{r^k}{k!} < \frac{1}{K}$$

It is easy to see (Section 5.2) that $\Sigma_k(m) := (\sum_i \frac{1}{|m-z_i|})^k$ constitutes an upper bound on $\lambda_k := \frac{|f^{(k)}(m)|}{|f(m)|}$ for all $k \geq 1$, where $z_1, \ldots, z_n$ denote the complex roots of f. Therefore, if $\Sigma_1(m) < \nu$ for a $\nu > 0$, then $\sum_{k\geq 1} \left| \frac{f^{(k)}(m)}{f(m)} \right| \frac{r^k}{k!} < e^{\nu r} - 1$. We easily verify that $T_K^f(m,r)$ succeeds if (see also Lemma 21)

$$\Sigma_1(m) = \sum_{i=1}^{n} \frac{1}{|m - z_i|} < \frac{1}{r} \ln\left(1 + \frac{1}{K}\right).$$

Now let us consider an arbitrary box B during the subdivision. If its midpoint $m(B)$ fulfills $|m(B) - z_i| > 2n \cdot r(B)$ for all $i = 1, \ldots, n$ then $T_1(m(B), r(B))$ succeeds according to the above consideration, thus B is discarded. For each root z_i, there exist at most $O(n^2)$ disjoint boxes B of the same size such that $|m(B) - z_i| \leq 2n \cdot r(B)$. Thus, in total, at most $O(n^3)$ boxes are retained. From this straightforward observation we immediately derive the upper bound $O(n^3)$ on the width of T^{CE}. This consideration is based on a pretty rough estimation of Σ_1 which assumes that, from a given point m, the distances to all roots z_i have roughly the same minimal value. In Section 5.1 we consider so called δ-clusters of roots which are related to the size δ of boxes at a certain subdivision level. We show that outside some "smaller" neighborhood of the roots of f the sum $\Sigma_1(m)$ is sufficiently small to guarantee the success of our exclusion predicate T_1 (see also Theorem 19):

THEOREM 4. Let $z_1, \ldots, z_n$ be points in the complex space and $\delta > 0$ an arbitrary real value. Then there exist disjoint, axes-parallel, open boxes $B_1, \ldots, B_k \subset \mathbb{C}$, $k \leq n^2$, with the following properties:

(i) The union $\mathcal{B} := \bigcup_{i=1,\ldots,k} B_i$ of all boxes covers all points $z_1, \ldots, z_n$.

(ii) $\mathcal{B}$ covers an area of less than or equal to $4n^2\delta^2$.

(iii) For each point $p \notin \mathcal{B}$ we have $\sum_{i=1}^{n} \frac{1}{|p-z_i|} \leq \frac{2(1+\ln\lceil n/2 \rceil)}{\delta}$.

From this result it follows directly that the width of T^{CE} is $O((n \ln n)^2)$ (Theorem 22). In case of the EVAL algorithm it turns out that the width of T^{EV} can be bounded by $O(n \ln n)$ (Theorem 23). A more refined argument in the proof of Theorem 22 even shows that, at a certain subdivision level h, the width of the tree is adapted to the number k_h of roots which are not isolated yet. To be more precisely, the width of T^{CE} (or T^{EV}) is upper bounded by $O((k_h \ln k_h)^2)$ (or $O(k_h \ln k_h)$).

We next apply the generalized Davenport-Mahler bound [9,10] to get a bound on k_h. This leads to the following result on the tree size (cf. Theorem 24):

Theorem 5. *(Tree Size)* For a square-free polynomial f of degree n with Gaussian integer coefficients of at most L bits, $\mathcal{T}^{CE}$ has $\widetilde{O}(n^2L)$ nodes. Similarly, $\mathcal{T}^{EV}$ has $\widetilde{O}(nL)$ nodes.

3.7.2 Bit Complexity

To analyze the bit complexity of our algorithm we have to consider the computational costs at a node of depth h. These costs are dominated by the computation of the Taylor expansion $f(z + m(B))$ at the midpoint $m(B)$ of the corresponding box B. We refer to Section 6 where we show that, assuming asymptotically fast Taylor shift, this can be achieved by $\widetilde{O}(n(L + nh))$ bit operations.

Readers familiar with the bit complexity analysis of the Descartes algorithms will notice that, up to a constant factor, this bound matches the computational costs at a node of depth h there. Our result on the bit complexity of the algorithms CEVAL and EVAL shows that the larger tree size of T^{CE}, in comparison to that of T^{EV}, does not effect the overall computational costs:

Theorem 6. *(Bit Complexity)* For a square-free polynomial f of degree n with integer coefficients of at most L bits, the algorithms CEVAL and EVAL isolate the complex (real) roots of f with a number Δ^{CE} (Δ^{EV}) of bit operations bounded by $\widetilde{O}(n^4L^2)$.

4 Proof of Correctness

This section proves the lemmas stated in Sections 3.3 and 3.4 for the correctness of CEVAL.

4.1 Basic Tools

We recall some basic facts of complex analysis. The Cauchy-Riemann equations for $f(z) = u(z) + \mathbf{i}v(z)$ say that

$$u_x = v_y, \qquad u_y = -v_x.$$

where u_x, u_y, v_x, v_y denote the partial differentiations of u, v with respect to x, y respectively. The gradient of u is given by $\nabla u = (u_x, u_y)$. Thus, $\nabla v =$

$(v_x, v_y) = (-u_y, u_x)$. Furthermore, we have complex differentiation of f satisfying

$$f'(z) := \frac{\partial f}{\partial z}(z) = u_x(z) + \mathbf{i}v_x(z) = v_y(z) - \mathbf{i}u_y(z) = u_x(z) - \mathbf{i}u_y(z). \tag{6}$$

Thus,

$$\arg f'(z) = -\arg \nabla u(z) = \arg \nabla v(z) - \frac{\pi}{2}. \tag{7}$$

Let $S^1 = [0, 2\pi)$ denote the set of angles in radians, with the usual addition modulo 2π. If $\alpha, \beta \in S^1$, let $[\alpha \pm \beta]$ denote the angular interval $\{\alpha + \theta : |\theta| \le \beta\}$. To exclude the endpoints in this interval, we write $(\alpha \pm \beta)$ for $\{\alpha + \theta : |\theta| < \beta\}$. We also write "$\xi \parallel \mu$" (parallel) if $\arg(\xi)$ is $\arg(\mu)$ or $\pi + \arg(\mu)$, and write "$\xi \perp \mu$" (perpendicular) if $\arg(\xi)$ is $\arg(\mu) + \pi/2$ or $\arg(\mu) - \pi/2$.

The lemmas in the rest of this section will be stated in terms of two constants, $K > 1$ and $L > 1$. We use these constants to define the predicate $T'_K(m, r)$ and the disk $D_{Lr}(m)$. Eventually, we will choose certain combinations of these constants, namely $(K, L) \in \{(4, 4), (3/2, 8), (1, 1)\}$.

Lemma 7. If $T'_K(m, r)$ holds then for all $\xi \in D_r(m)$, we have

$$\arg f'(\xi) \in (\arg f'(m) \pm \arcsin(1/K)).$$

Equivalently,

$$\arg \nabla u(\xi) \in (\arg \nabla u(m) \pm \arcsin(1/K)).$$

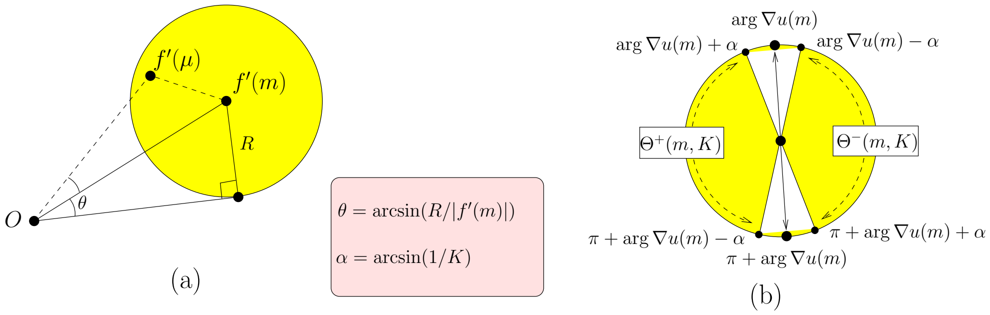

$f'(\mu)$
$f'(m)$
R
O
θ
(a)
$\theta = \arcsin(R/|f'(m)|)$
$\alpha = \arcsin(1/K)$
$\arg \nabla u(m)$
$\arg \nabla u(m) + \alpha$
$\arg \nabla u(m) - \alpha$
$\Theta^+(m, K)$
$\Theta^-(m, K)$
$\pi + \arg \nabla u(m) - \alpha$
$\pi + \arg \nabla u(m) + \alpha$
$\pi + \arg \nabla u(m)$
(b)

Fig. 3. (a) Bounding $|\arg f'(\mu) - \arg f'(m)|$ for $\mu \in D_r(m)$. (b) Forbidden range $\Theta(m, K)$ and its complement.

PROOF. Let $R = \sum_{k \ge 2} \left| \frac{f^{(k)}(m)}{(k-1)!} r^{k-1} \right|$. If $\mu \in D_r(m)$, we see that

$$f'(\mu) = f'(m) + \sum_{k \ge 2} \frac{f^{(k)}(m)}{(k-1)!} (\mu - m)^{k-1}$$

and so $|f'(\mu) - f'(m)| \le R$. Hence, from Figure 3(a), we see that

$$|\arg f'(\mu) - \arg f'(m)| \le \arcsin\left(\frac{R}{|f'(m)|}\right) < \arcsin(1/K)$$

since $T'_K(m,r)$ holds implies $|f'(m)| > KR$. The equivalent form in terms of ∇u follows from the fact that $\arg f'(\mu) = -\arg \nabla u(\mu)$.

It follows from this lemma that if $T'_K(m,r)$ holds and $\mu, \xi \in D_r(m)$ then

$$|\arg f'(\mu) - \arg f'(\xi)| < 2 \arcsin(1/K).$$

Thus, the argument of $f'(z)$ (for $z \in D_r(m)$) cannot vary by more than $2\arcsin(1/K)$.

The next property is, of course, a generalization of Lemma 1(ii).

LEMMA 8. If $K \geq \sqrt{2}$ and $T'_K(m,r)$ holds, then the disk $D_r(m)$ has at most one zero of f.

PROOF. Say a, b are two zeros of f in $D_r(m)$. As $a = b$ implies $f'(a) = 0$, which is not possible as $T'_1(m,r)$ holds, we can assume $a \neq b$. Then $f(a) = f(b) = 0$ and so $u(a) = v(a) = u(b) = v(b) = 0$. But $u(a) = u(b) = 0$ implies, by the Mean Value Theorem, that there exists $\mu \in [a,b]$ such that

$$\nabla u(\mu) \perp (b-a).$$

Similarly, $v(a) = v(b) = 0$ implies there exists $\xi \in [a,b]$ such that

$$\nabla v(\xi) \perp (b-a).$$

But $\nabla v(\xi) = (v_x(\xi), v_y(\xi)) = (-u_y(\xi), u_x(\xi))$. It follows that

$$\nabla u(\xi) \parallel (b-a).$$

Therefore $\nabla u(\mu)$ and $\nabla u(\xi)$ must be perpendicular.

On the other hand, Lemma 7 says that if $\mu, \xi \in D_r(m)$, then $\arg \nabla u(\mu)$ and $\arg \nabla u(\xi)$ differ by less than $2\arcsin(1/K)$. Since $K \geq \sqrt{2}$, they differ by less than $2\arcsin(1/\sqrt{2}) = \pi/2$. This contradicts the perpendicularity between $\arg \nabla u(\mu)$ and $\arg \nabla u(\xi)$.

4.2 Crossings of $u = 0$ *on Disk Boundary*

We next prove several lemmas that show that u-crossings of a disk $D_r(m)$ are quite restricted under the following assumption.

$$\text{The predicate } T'_K(m,r) \text{ holds for some } K \geq \sqrt{2}. \tag{8}$$

A similar argument shows corresponding results for v-crossings. To focus on the behavior of the function $u(z) = u(x,y)$ on the boundary of $D_r(m)$, it is useful to consider u there as a function of the angle ϕ

$$u_{m,r}(\phi) := u(m + re^{\mathbf{i}\phi}). \tag{9}$$

From our earlier definition, there is a u-crossing of $D_r(m)$ at ϕ iff $u_{m,r}(\phi) = 0$.

We introduce the notation

$$\Theta(m, K) := (\arg \nabla u(m) \pm \arcsin(1/K)) \cup (\pi + \arg \nabla u(m) \pm \arcsin(1/K)) \subseteq S^1$$

for the double cone of angles. In Figure 3(b), this double cone is indicated by two white sectors. We call $\Theta(m, K)$ the **forbidden range**. The complement of the forbidden range is composed of two angular ranges (see Figure 3(b)),

$$\begin{cases} \Theta^+(m, K) := [\arg \nabla u(m) + \frac{\pi}{2} \pm \arccos(1/K)] \\ \Theta^-(m, K) := [\arg \nabla u(m) - \frac{\pi}{2} \pm \arccos(1/K)]. \end{cases} \tag{10}$$

The "forbidden" terminology is partly motivated by the next lemma. We show that the derivative $u'_{m,r}(\phi) := \frac{du_{m,r}}{d\phi}(\phi)$ of $u_{m,r}(\phi)$ does not vanish if ϕ lies outside the forbidden range.

Lemma 9. Assume (8).

(i) If $u'_{m,r}(\phi) = 0$ then $\phi \in \Theta(m, K)$.

(ii) There is at most one u-crossing of $D_{m,r}$ in $\Theta^+(m, K)$, and at most one u-crossing of $D_{m,r}$ in $\Theta^-(m, K)$.

PROOF. (i) Let $\mu = m + re^{\mathbf{i}\phi}$. Note that

$$u'_{m,r}(\phi) := \frac{du_{m,r}}{d\phi}(\phi) = u_x(m + re^{\mathbf{i}\phi})(-r \sin \phi) + u_y(m + re^{\mathbf{i}\phi})(r \cos \phi).$$

Since $e^{\mathbf{i}(\phi+\pi/2)} = -\sin \phi + \mathbf{i} \cos \phi$, we conclude that

$$\begin{cases} u'_{m,r}(\phi) = 0 \Leftrightarrow \nabla u(\mu) \perp e^{\mathbf{i}(\phi+\pi/2)} \\ \qquad\qquad\quad\;\, \Leftrightarrow \nabla u(\mu) \parallel e^{\mathbf{i}(\phi)}. \end{cases} \tag{11}$$

Thus $u'_{m,r}(\phi) = 0$ implies $\arg \nabla u(\mu)$ is equal to ϕ or to $\pi + \phi$. Since (8) implies $\arg \nabla u(\mu) \in \Theta(m, K)$, we conclude that $\phi \in \Theta(m, K)$.

(ii) This is an immediate application of part (i) using Rolle's Theorem.

The preceding lemma implies that there are at most two u-crossings outside the forbidden range. And in case that there are two such crossings, they must lie on opposite sides of the circle separated by the forbidden range. The next lemma is also useful for limiting u-crossings to at most two without consideration of the forbidden range.

Lemma 10. Assume (8) and suppose there are three u-crossings of $D_r(m)$ at $\phi_1, \phi_2, \phi_3 \in S^1$. Let $a_i = m + re^{\mathbf{i}\phi_i}$. Then each side of the triangle $\Delta a_1 a_2 a_3$ is shorter than $4r/K$.

PROOF. We consider a line segment $[a_i, a_j]$. As $u(a_i) = u(a_j) = 0$, the Mean Value Theorem implies that there exists a point ξ_{ij} on $[a_i, a_j]$ where $\nabla u(\xi_{ij}) \perp (a_j - a_i)$. Now if an interior angle of the triangle $\Delta a_1 a_2 a_3$ is in between $(2\arcsin(1/K), \pi - 2\arcsin(1/K))$, at least for two of the gradients $\nabla u(\xi_{ij})$, their arguments would differ by more than $2\arcsin(1/K)$. But this would contradict Lemma 7. W.l.o.g. let us consider the angle α_3 at the point a_3. Then from the extended sine theorem we get $|a_2 - a_1|/\sin(\alpha_3) = 2r$, thus we must have $|a_2 - a_1| = 2r\sin(\alpha_3) < 2r\sin(2\arcsin(1/K)) < 4r/K$. Similarly, $|a_3 - a_1|, |a_3 - a_2| < 4r/K$, too.

A consequence of Lemma 7 is to confine the curve $u = 0$ within a certain double cone region:

Lemma 11. Assume (8) and suppose $u(\xi) = 0$ for some $\xi \in D_r(m)$. Then the curve $u = 0$ inside $D_r(m)$ is confined within the double cone $C(\xi, m, r, K)$ consisting of all $z \in D_r(m)$ that fulfill

$$|\arg(z - \xi) - \arg(\nabla u(m))| \in \left(\frac{\pi}{2} \pm \arcsin(1/K)\right). \qquad (12)$$

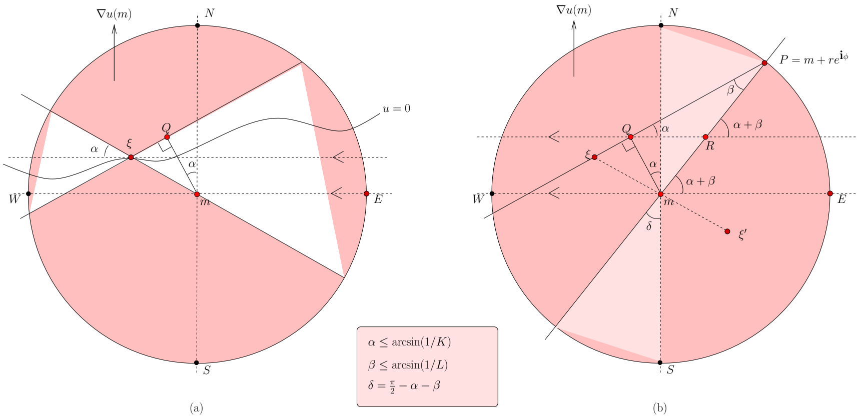


Fig. 4. (a) The double cone $C(\xi, m, r, K)$ is shaded white; (b) The separation δ between $\arg \nabla u(m)$ and u-crossing ϕ.

PROOF. In Figure 4(a), the angle $\arg \nabla u(m)$ is viewed as pointing northward, and the double cone $C(\xi, m, r, K)$ is shaded white. If $u(z) = 0$ for any $z \in D_r(m)$, then by the Mean Value Theorem, there is a point μ on the line segment $[\xi, z]$ such that

$$\nabla u(\mu) \perp (z - \xi). \tag{13}$$

By the previous lemma,

$$|\arg \nabla u(\mu) - \arg \nabla u(m)| \leq \arcsin(1/K). \tag{14}$$

But (13) and (14) is equivalent to $z \in C(\xi, m, r, K)$, as is evident from Figure 4(a).

The next two lemmas show that if the curve u passes relatively close to the center of $D_r(m)$ (say, within distance r/L for some $L > 1$) then the u-crossings are separated from $\arg \nabla u(m)$ and from the v-crossing. First we show that u-crossings are separated from $\arg \nabla u(m)$, as illustrated in Figure 4(b).

Lemma 12. Assume (8) and ξ is a root of $f(z)$ with $|\xi - m| \leq r/L$ for some $L > 1$. Then for any u-crossing ϕ of $D_r(m)$, it obtains that ϕ and $\arg \nabla u(m)$ are δ-separated where

$$\delta \geq \delta(K, L) := \frac{\pi}{2} - \arcsin(1/K) - \arcsin(1/L).$$

Similarly, ϕ and $\pi + \arg \nabla u(m)$ are δ-separated. If $\delta(K, L) > \arcsin(1/K)$ then u has exactly two u-crossing, one in $\Theta^+(m, K)$ and the other in $\Theta^-(m, K)$.

PROOF. Refer to Figure 4(b) where again we assume $\nabla u(m)$ is pointing northward. Thus, $(N - m) \parallel \nabla u(m)$ where N is the north pole of $D_r(m)$ (see Figure 4). If ϕ lies in the third or fourth quadrants, then clearly ϕ and $\arg \nabla u(m)$ are δ-separated. Otherwise, by symmetry, we may assume ϕ lies in the first quadrant. Let P be the point $m + re^{\mathbf{i}\phi}$. So by assumption, $u(P) = 0$. Consider the angle $\delta := \angle(PmN)$ (see Figure 4(b)). Thus we must prove that $\delta \geq \frac{\pi}{2} - \arcsin(1/L) - \arcsin(1/K)$.

Consider the line $\overline{Pm}$: the point ξ is either above or below the line. It is not hard to see that the minimum value of $\alpha(PmN)$ is attained only if ξ lies above $\overline{Pm}$, as seen in Figure 4(b). For instance, the point ξ' in Figure 4(b) lies below $\overline{Pm}$, but it can be replaced by $\xi := 2m - \xi$ which lies symmetrically opposite relative to center m.

Let Q be the point on the line $\overline{P\xi}$ that is closest to m. Let R be the point on the line $\overline{Pm}$ so that $(Q - R) \perp \nabla u(m)$. If we define $\alpha := \angle(RQP)$ and $\beta := \angle(RPQ)$ then it easily seen that $\delta = \frac{\pi}{2} - \alpha - \beta$. From Lemma 11, we conclude that $\alpha \leq \arcsin(1/K)$ and from the assumption that $|\xi - m| \leq r/L$, we see from examining the triangle $\Delta(P\xi m)$ that $\beta \leq \arcsin(1/L)$. These two inequalities imply

$$\delta \geq (\pi/2) - \arcsin(1/K) - \arcsin(1/L).$$

By a symmetrical argument, we also conclude that $\pi + \nabla u(m)$ and ϕ must be separated by an angle of at least $((\pi/2) - \arcsin(1/K) - \arcsin(1/L))$.

It remains to prove the claim about the number of crossings for $\delta(K, L) > \arcsin(1/K)$. As $D_r(m)$ contains a root of f the image of $\partial D_r(m)$ under the function f is a curve in $\mathbb{C}$ that circles the origin at least once, thus we must have at least two $u-$crossing on $\partial D_r(m)$. We have already shown that all $u-$crossings are separated from $\nabla u(m)$ and $\pi + \nabla u(m)$ by an angle of at least $\delta(K, L)$. Hence, from our definition of the forbidden range it follows that all $u-$crossings are located outside the forbidden range, thus the claim about exactly two u-crossings follows from Lemma 9.

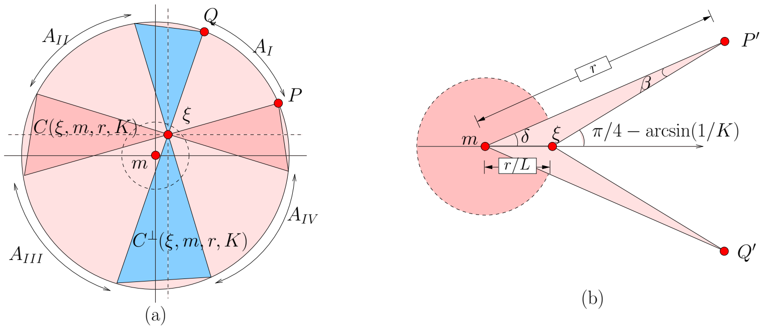
The next lemma is similar to the preceding one, except that we now show separation between u-crossings and v-crossings:

Lemma 13. Assume (8) and ξ is a root of $f(z)$ with $|\xi - m| \leq r/L$. If we further assume that

$$\frac{\pi}{2} - 2\arcsin 1/K - \arcsin 1/L > 0,$$

then there are exactly two u-crossing ϕ_1, ϕ_2 and two v-crossings ψ_1, ψ_2 on $D_r(m)$, and they are interleaving. Each $u-$crossing ϕ is separated from each $v-$crossing ψ by at least

$$\delta := \frac{\pi}{2} - 2\arcsin(1/K) - 2\arcsin\left(\frac{\sin(\pi/4 - \arcsin(1/K))}{L}\right).$$



(a)

(b)

Fig. 5. Angle separation between a u-crossing and a v-crossing.

PROOF. Wlog, assume $\arg \nabla u(m) = \pi/2$ (i.e., $\nabla u(m)$ points northward). From Lemma 12 we already know that there exists exactly two u-crossings

ϕ_1, ϕ_2 and two v-crossings ψ_1, ψ_2. Let ϕ be a u-crossing, then we have $u(P) = 0$ where $P = m + re^{\mathbf{i}\phi}$, and likewise $v(Q) = 0$ where $Q = m + re^{\mathbf{i}\psi}$, ψ a $v-$crossing. By Lemma 11, P lies in the cone $C(\xi, m, r, K)$, and similarly, Q lies in the cone $C^{\perp}(\xi, m, r, K)$, defined as the cone $C(\xi, m, r, K)$ rotated by 90° about the point ξ. See Figure 5(a). From the assumption that $\frac{\pi}{2} > 2 \arcsin 1/K$, it follows that the two cones only share the point ξ. It follows that the $u-$ and $v-$ crossings are interleaving. The complement of

$$\big(\partial C(\xi, m, r, K) \cup \partial C^{\perp}(\xi, m, r, K)\big) \cap \partial D_r(m)$$

is comprised of four arcs $A_I, A_{II}, A_{III}, A_{IV}$. Because $\nabla u(m)$ points northward, arc A_i may be associated with ith quadrant ($i \in \{I, II, III, IV\}$) in a natural way. The angle subtended by arc A_i at m is proportional to the arc length of A_i. It is not hard to see that the minimum angle $\angle(PmQ)$ is attained under the following conditions:
(a) P and Q are endpoints of one of these arcs.
(b) $|m - \xi| = r/L$.
(c) Measure of angle $\angle(P\xi Q)$ is $(\pi/2) - 2\arcsin(1/K)$.

Consider the somewhat more general situation where A' is any arc of $\partial D_r(m)$ with endpoints P' and Q' satisfying the analogues of conditions (a),(b), and (c). What is the minimum measure of $\angle(P'mQ')$? This measure is minimized when the line $\overline{m\xi}$ bisects the angle $\angle(P'\xi Q')$. Thus the exterior angle at $\angle(P'\xi m)$ has measure that is half of $\angle(P'\xi Q')$, i.e., $(\pi/4) - \arcsin(1/K)$. This optimal configuration is illustrated in Figure 5(b).

If β is the measure of $\angle(mP'\xi)$, then the sign formula for $\Delta(P'm\xi)$ shows that

$$\sin(\beta) = \frac{\sin(\pi/4 - \arcsin(1/K))}{L}$$

Let $\delta' := (\pi/4) - \arcsin(1/K) - \beta$. The lemma follows from the fact that $\arg(P' - m)$ and $\arg(Q' - m)$ are $2\delta'$-separated.

4.3 Application to the 8-Point Test

We are now ready to apply the preceding lemmas, using them to prove Theorem 2. Recall that this theorem is concerned with the success and non-success of the 8-Point Test for $D_{4r}(m)$. We now fix the constants $K = 6$ and $L = 4$.

We also want to slightly generalize the 8-Point Test by allowing some flexibility in choosing the 8 main compass points. Given $0 \le \theta_0 < \theta_1 < \cdots < \theta_7 < 2\pi$, we can define the 8 main compass points on $D_{4r}(m)$ to be $P_i := m + 4re^{\mathbf{i}\theta_i}$ ($i = 0, \ldots, 7$). Say these compass points are **δ-approximate** if each pair

(θ_i, θ_j) is δ_{ij}-separated where

$$\delta_{ij} \in [45^\circ \pm \delta].$$

We are interested in an **Approximate 8-Point Test** based on such a set of δ-separated compass points. The following is the slightly generalized version of Theorem 2:

Theorem 14 (Success of Approximate 8-Point Test). Assume the 8-Point Test is based on a set of 2.5°-separated compass points.

(i) If $T_6'(m, 4r)$ holds and $D_r(m)$ is isolating, then $D_{4r}(m)$ passes the approximate 8-Point Test.

(ii) If $T_6'(m, 4r)$ holds and $D_{4r}(m)$ passes the approximate 8-Point Test, then $D_{4r}(m)$ is isolating.

PROOF. We first prove the Part (ii). If an approximate 8-Point Test succeeds for $D_{4r}(m)$, then we must show that $D_{4r}(m)$ contains a root of f. By the assumption $T_6'(m, 4r)$, we know that $D_{4r}(m)$ has at most one root. The success of the test implies that there are two arcwise u-crossings and two arcwise v-crossings, and these interleave. Thus, there are two u-crossings ϕ^+, ϕ^- that are 42.5°-separated. A calculation shows that the distance $|P - Q|$ between $P = m + 4re^{\mathbf{i}\phi^+}$ and $Q = m + 4re^{\mathbf{i}\phi^-}$ is at least $4r\sqrt{2 - 2\cos 42.5^\circ} \approx 2.9r$

If there are any other u-crossings, then Lemma 10 implies the distance $|P-Q|$ is at most $4(4r)/6 \approx 2.67r$, which is a contradiction. Therefore, the u-curve has exactly one connected component within $D_{4r}(m)$. Similarly for the v-curve. Since the u-crossings and v-crossings are interleaving, they must intersect within $D_{4r}(m)$. This intersection is the root we seek.

We now prove Part (i), so let us assume that $D_r(m)$ contains a root ξ of f.

1. From Lemma 13 we know that there exists exactly two $u-$crossings ϕ^+, ϕ^- and two $v-$crossings ψ^+, ψ^- which are interleaving.

2. Recall that the main compass points of $D_{4r}(m)$ divides $\partial D_{4r}(m)$ into eight arcs. For any angle ϕ, let $A(\phi)$ denote the arc that contains ϕ. We claim that there is an arc-wise crossings at $A(\phi^*)$ where ϕ^* is either ϕ^+ or ϕ^-. Since ϕ^* is at least $\pi/2 - \arcsin(1/4) - \arcsin(1/6) \approx 65^\circ$-separated (see Lemma 12) from $\arg \nabla u(m)$, and $\arcsin(1/6) \approx 9.6^\circ$, we conclude that the two endpoints ϕ_1, ϕ_2 of $A(\phi^*)$ lie outside $\Theta(m, 4)$. This proves that $u_{m,4r}(\phi_1)u_{m,4r}(\phi_2) < 0$. Moreover, $A(\phi^+)$ and $A(\phi^-)$) are distinct because they are separated by the forbidden range.

3. By the same argument we see exactly two arc-wise intersections at two distinct arcs $A(\psi^+)$ and $A(\psi^-)$ for v. As we already know that the u- and v-

crossings are interleaving it remains to show that $A(\psi^*)$ and $A(\phi^*)$ are distinct for $\phi^* \in \{\phi^+, \phi^-\}$ and $\psi^* \in \{\psi^+, \psi^-\}$. If $A(\phi^*) = A(\psi^*)$, then ϕ^* and ψ^* are separated by at most 50°. But this contradicts our result in Lemma 13 which says that these crossings are separated by at least

$$\frac{\pi}{2} - 2\arcsin(1/6) - 2\arcsin\left(\frac{\sin(\pi/4 - \arcsin(1/6))}{4}\right) \approx 54.15^\circ.$$

This concludes our proof of Theorem 14.

5 Complexity Analysis

This section justifies the lemmas stated in Section 3.7 on the complexity of CEVAL and EVAL. In particular, we introduce our cluster analysis technique.

5.1 The Clustering Approach

For the complexity analysis we need non-trivial bounds on the quotients $\lambda_k := \frac{|f^{(k)}(m)|}{|f(m)|}$ where m is the midpoint of a box B in $\mathbb{C}$, as these values determine the success of our chosen predicates. It is easy to see (see Section 5.2) that $\Sigma_k(m) := (\sum_i \frac{1}{|m-z_i|})^k$ constitutes an upper bound on $\lambda_k(m)$ where $z_1, \ldots, z_n$ denote the complex roots of f. Thus, before we turn to the complexity analysis we formulate a number of useful results to estimate the sum $\Sigma_1(m)$, in particular, we derive non-trivial upper bounds when m is located outside some "small" neighborhoods of the roots z_i.

Let $\delta > 0$, and suppose $R \subseteq \mathbb{R}$ is a non-empty multiset of real numbers. Multiset means that elements of R may be duplicated, and its size is denoted $|R|$, with multiplicity counted. Then its **center of gravity** is

$$\mathrm{cg}(R) := \left(\sum_{x \in R} x\right) / |R|,$$

and **δ-interval** is

$$I_\delta(R) := (\mathrm{cg}(R) \pm |R|\delta).$$

Thus the width of the $I_\delta(R)$ is $2|R|\delta$.

A **ranking** of R is a one-one onto function $r : R \to \{1, 2, \ldots, |R|\}$. We call R a **semi δ-cluster** if there is a ranking r of R such that for all $x \in R$,

$$(\mathrm{cg}(R) + |R|\delta) - x \geq r(x)\delta. \tag{15}$$

We call R a **δ-cluster** if both R and $-R = \{-x : x \in R\}$ are δ-clusters.

Rewriting (15) as

$$x - \mathrm{cg}(R) \leq (|R| - r(x))\delta$$

we see that the right-hand side is non-negative, and the inequality is automatic when $x \leq \mathrm{cg}(R)$. We are mainly interested in clusters, but it is easier to prove properties for semi-clusters and to extend them to clusters by symmetry.

Consider the following examples:

$$\begin{aligned} R_n &= \{x_1, \ldots, x_n\}\,, \text{where } x_1 = x_i \text{ for all } i;\\ R_1 &= \{-3, 1, 2\}\,;\\ R_2 &= \{x_1, x_2\}\,;\\ R_3 &= \{-x, 0, x\}\,. \end{aligned}$$

R_n is a δ-cluster for any $\delta > 0$. R_1 is a semi 1-cluster with $\mathrm{cg}(R_1) = 0$, but it is not a 1-cluster. R_2 is a δ-cluster iff $|x_0 - x_1| \leq 2\delta$. R_3 is a δ-cluster iff $|x| \leq 2\delta$.

5.1.1 Properties of Clusters

The following is immediate:

Lemma 15. If R is a δ-cluster, then R is contained in $I_\delta(R)$. In fact, a stronger containment is true:

$$R \subseteq [\mathrm{cg}(R) \pm (|R| - 1)\delta].$$

This lemma motivates a useful definition: a collection $\mathcal{P} = \{R_1, \ldots, R_k\}$ is called a **δ-partition** (of the set $R = \bigcup_{i=1}^k R_i$) if each R_i is a δ-cluster and the intervals $I_\delta(R_i)$ are pairwise disjoint. Let $I_\delta(\mathcal{P}) := \bigcup_{i=1}^k I_\delta(R_i)$. Clearly, a δ-partition of R induces an ordinary partition of R.

Another useful property is this:

Lemma 16. If R is a δ-cluster and $p \notin I_\delta(R)$ then

$$\sum_{x \in R} \frac{1}{|p - x|} \leq \frac{1 + \ln |R|}{\delta}.$$

If $\mathcal{P} = \bigcup_{i=1,\ldots,k} R_i$ is a δ-partition of a multiset R, and $p \notin I_\delta(\mathcal{P})$ then

$$\sum_{x \in R} \frac{1}{|p - x|} \leq \frac{2(1 + \ln \lceil |R|/2 \rceil)}{\delta}.$$

PROOF. As $p \notin I_\delta(R)$, then we may, wlog, assume that $p > x$ for all $x \in R$. We only consider the first case as the case $p < x$ can be treated completely similar. If r is the ranking function that witnesses R as a semi δ-cluster then we have

$$\sum_{x\in R}\frac{1}{|p-x|}\leq\sum_{i=1}^{|R|}\frac{1}{|p-r^{-1}(i)|}\leq\sum_{i=1}^{|R|}\frac{1}{i\delta}\leq\frac{1+\ln|R|}{\delta}.$$

For the proof of the second claim we assume, wlog, that the clusters are ordered in way such that $x<y$ for all $i<j$ and $x\in R_i$, $y\in R_j$. Let $\mathcal{R}_0:=\bigcup_{i=1,\dots,k_0}R_i$ be the union of all points $x\in\mathcal{R}$ with $x<p$ and $\mathcal{R}_1:=\bigcup_{i=k_0+1,\dots,k}R_i$. Notice that p separates clusters as it is not contained in any $I_\delta(R_i)$. For $i\leq k_0$ and $x\in R_i$ we define the ranking function $r:\mathcal{R}_0\to\{1,\dots,|\mathcal{R}_0|\}$ by $r(x):=\sum_{j=i+1}^{k_0}|R_i|+r_i(x)$ where r_i denotes the ranking function that witnesses R_i as a semi δ-cluster. It follows that $|p-x|\geq r(x)\delta\geq l\delta$ if x is the l-th element of $\mathcal{R}_0$ left to p. Hence, we get

$$\sum_{x\in\mathcal{R}_0}\frac{1}{|p-x|}\leq\sum_{l=1}^{|\mathcal{R}_0|}\frac{1}{|p-r^{-1}(l)|}\leq\sum_{l=1}^{|\mathcal{R}_0|}\frac{1}{l\delta}\leq\frac{1+\ln|\mathcal{R}_0|}{\delta}.$$

In an analogous manner we also show $\sum_{x\in\mathcal{R}_1}|p-x|^{-1}\leq(1+\ln|\mathcal{R}_1|)/\delta$, and thus

$$\sum_{x\in\mathcal{R}}\frac{1}{|p-x|}\leq\frac{2+\ln|\mathcal{R}_0|+\ln|\mathcal{R}_1|}{\delta}\leq\frac{2(1+\ln\lceil|R|/2\rceil)}{\delta}.$$

Lemma 17. Let R, R' be semi δ-clusters of sizes n and n', respectively. If $|\mathrm{cg}(R)-\mathrm{cg}(R')|\leq(n+n')\delta$, then

(i) $\max\{\mathrm{cg}(R)+n\delta,\mathrm{cg}(R')+n'\delta\}\leq\mathrm{cg}(R\cup R')+(n+n')\delta$

(ii) $R\cup R'$ is a semi δ-cluster.

The union of δ-clusters R, R' is again a δ-cluster if $I_\delta(R)\cap I_\delta(R')\neq\emptyset$.

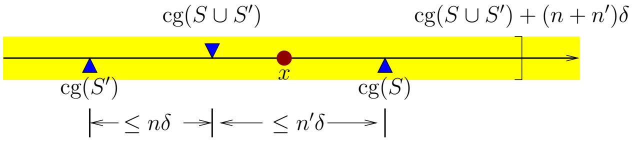


Fig. 6. The union of two δ-clusters R, R'

PROOF. Wlog, let $\mathrm{cg}(R')\leq\mathrm{cg}(R\cup R')\leq\mathrm{cg}(R)$, as in Figure 6.
(i) Clearly, $\mathrm{cg}(R')+n'\delta\leq\mathrm{cg}(R\cup S')+(n+n')\delta$. Furthermore, we have

$$\begin{aligned}(n+n')\mathrm{cg}(R\cup R')&=n\mathrm{cg}(R)+n'\mathrm{cg}(R')\\&\geq n\mathrm{cg}(R)+n'(\mathrm{cg}(R)-(n+n')\delta)\\&=(n+n')(\mathrm{cg}(R)-n'\delta)\end{aligned}$$

and thus $\mathrm{cg}(R \cup R') \geq \mathrm{cg}(R) - n'\delta$, which shows the second part of (i).
(ii) Let $r : R \to \{1, \ldots, n\}$ and $r' : R' \to \{1, \ldots, n'\}$ be the ranking functions that witness R and R' as the semi δ-clusters, respectively. We choose a new ranking function $\overline{r} : R \cup R' \to \{1, \ldots, n + n'\}$ where

$$\overline{r}(x) = \begin{cases} r(x) & \text{if } \ x \in R, \\ n + r'(x) & \text{if } \ x \in R'. \end{cases}$$

If $x \in R$, then we have

$$\mathrm{cg}(R \cup R') + (n + n')\delta - x \geq \mathrm{cg}(R) + n\delta - x \geq r(x)\delta = \overline{r}(x)\delta$$

as desired. If $x \in R'$, then we also have

$$\mathrm{cg}(R \cup R') + (n + n')\delta - x \geq (\mathrm{cg}(R') + n'\delta - x) + n\delta \geq r(x)\delta + n\delta = \overline{r}(x)\delta.$$

From the definition of $I_\delta(R)$ and $I_\delta(R')$ it is immediate that $|\mathrm{cg}(R) - \mathrm{cg}(R')| \leq (|R| + |R'|)\delta$ if $I_\delta(R) \cap I_\delta(R') \neq \emptyset$. Hence $R \cup R'$ is a δ-cluster according to (ii).

Lemma 18. Let R be a multiset that contains n points $x_1, \ldots, x_n \in \mathbb{R}$ and $\delta > 0$ an arbitrary real value. Then there exists a δ-partition $\mathcal{P}$ of R and for each $p \notin I_\delta(\mathcal{P})$ it holds that

$$\sum_{i=1}^{n} \frac{1}{|p - x_i|} \leq \frac{2(1 + \ln \lceil n/2 \rceil)}{\delta}.$$

PROOF. Let $\mathcal{P} = \{R_1, \ldots, R_k\}$ be a partition of R where each R_i is a δ-cluster. We will keep transforming $\mathcal{P}$ until it becomes a δ-partition. We start with $\mathcal{P} = \{\{x_1\}, \ldots, \{x_n\}\}$. In each step we consider clusters $R, R' \subset \mathcal{P}$ with $I_\delta(R) \cap I_\delta(R') \neq \emptyset$. Their union $R \cup R'$ is again a δ-cluster due to Lemma 17. We remove R and R' from $\mathcal{P}$ and insert $R \cup R'$. When all the intervals $I_\delta(R)$ for $R \in \mathcal{P}$ are pairwise disjoint, we have the desired δ-partition. The statement about the bound on the sum $\sum_{j=1}^{k} \frac{1}{|p - x_i|}$ follows directly from Lemma 16.

5.1.2 Complex Clusters

We now extend the concept of δ-clusters to a multiset $R = \{z_1, \ldots, z_n\}$ of complex numbers. Let $\mathtt{Re}[R]$ and $\mathtt{Im}[R]$ denote the multiset of the real and imaginary part of elements in R. We note that in our application, R is the set of roots of a square-free polynomial and hence R is just an ordinary set. Nevertheless, $\mathtt{Re}[R]$ and $\mathtt{Im}[R]$ will multisets in general.

According to Lemma 18 there exists a δ-partition $\left\{R_1, \ldots, R_{k_{\mathtt{Re}}}\right\}$ of $\mathtt{Re}[R]$. Similarly, let $\left\{\widetilde{R}_1, \ldots, \widetilde{R}_{k_{\mathtt{Im}}}\right\}$ denote a δ-partition of $\mathtt{Im}[R]$. Each interval

$I_\delta(R_i)$ $(I_\delta(\widetilde{R}_j))$ defines a vertical (horizontal) stripe (see Figure 7 on page 32) in the complex plane, containing all points $z \in \mathbb{C}$ with $\mathtt{Re}(z) \in I_\delta(R_i)$ $(\mathtt{Im}(z) \in I_\delta(\widetilde{R}_j))$. Their overlapping consists of $k := k_{\mathtt{Re}} \cdot k_{\mathtt{Im}}$ disjoint boxes which we denote by $B_1, \ldots, B_k$. For any point $p \notin \bigcup_{i=1}^k B_i$, either $\mathtt{Re}(p) \notin \bigcup_{i=1}^{k_{\mathtt{Re}}} I_\delta(R_i)$ or $\mathtt{Im}(p) \notin \bigcup_{i=1}^{k_{\mathtt{Im}}} I_\delta(\widetilde{R}_i)$, hence from Lemma 18 we get $\sum_{i=1}^n \frac{1}{|p-z_i|} \leq \frac{2(1+\ln\lceil n/2\rceil)}{\delta}$. Furthermore, let $\epsilon \geq 0$ be an arbitrary positive value and B_i^ϵ the box that is obtained by enlarging B_i by ϵ in each direction. If $\mathcal{B} := \bigcup_{i=1,\ldots,k} B_i$, then the total area covered by the union $\mathcal{B}^\epsilon := \bigcup_{B\in\mathcal{B}} B^\epsilon$ of all these enlarged boxes is upper bounded by

$$\sum_{i,j}(w(I_\delta(R_i)) + 2\epsilon)(w(I_\delta(\widetilde{R}_j)) + 2\epsilon) = \sum_i (w(I_\delta(R_i)) + 2\epsilon) \cdot \sum_j (w(I_\delta(\widetilde{R}_j)) + 2\epsilon)$$
$$\leq (2n\delta + 2n\epsilon)^2 = 4n^2(\delta+\epsilon)^2.$$

where the sum is taken over all $i = 1, \ldots, k_{\mathtt{Re}} \leq n$, $j = 1, \ldots, k_{\mathtt{Im}} \leq n$. We fix this result.

Theorem 19. Let R be a multiset consisting of n points $z_1, \ldots, z_n$ in the complex space and $\epsilon \geq 0$, $\delta > 0$ arbitrary real values. Then there exist disjoint axes-parallel boxes $B_1, \ldots, B_k \subset \mathbb{C}$, $k \leq n^2$, with the following properties:

(i) The union $\mathcal{B} := \bigcup_{i=1,\ldots,k} B_i$ of all boxes cover R.

(ii) $\mathcal{B}^\epsilon = \bigcup_{i=1,\ldots,k} B_i^\epsilon$ covers an area of less than or equal to $4n^2(\delta+\epsilon)^2$.

(iii) For each point $p \notin \mathcal{B}$ we have $\sum_{i=1}^n \frac{1}{|p-z_i|} \leq \frac{2(1+\ln\lceil n/2\rceil)}{\delta}$.

We conclude this section with another useful lemma. Again we consider a multiset R, consisting of n complex points $z_1, \ldots, z_n$. We are interested in a partition of R into multisets that consist of nearby points, only. Let $\sigma(z_i) := \min_{j\neq i} |z_i - z_j|$ denote the distance of z_i to its nearest point in R. Furthermore, for an arbitrary $\delta > 0$, we consider the multiset R_δ that contains exactly those z_i with $\sigma(z_i) \leq \delta$.

Lemma 20. There exists a partition of R_δ into disjoint multisets $R_1, \ldots, R_k$ such that $|R_{i_0}| \geq 2$ for each $i_0 \in \{1, \ldots, k\}$ and $|z_i - z_j| \leq |R_\delta|\delta$ for all $z_i, z_j \in R_{i_0}$.

PROOF. Wlog we can assume that R_δ consists of the points $z_1, \ldots, z_l$ with an $l \leq n$. We start with z_1 and define $R_1 := \{z_1\}$. We further put all points z_i in R_1 that satisfy $|z_i - z_1| \leq \delta$. Then we proceed with each point in R_1 in the same way. If no further point can be added to R_1 we consider the set $R_\delta \backslash R_1$ of the remaining points and treat it in exactly the same manner. Finally, we end up with a partition $R_1, \ldots, R_k$ of R such that for any two points in any R_{i_0}, their distance is less than or equal to $(|R_{i_0}| - 1)\delta \leq |R_\delta|\delta$. Furthermore, each of the multisets R_i must contain at least two points as $\sigma(z_i) \leq \delta$ for all $i = 1, \dots, l$.

5.2 Analysis of the Subdivision Tree

We show that our algorithm CEVAL, despite its simple predicates, is also efficient in a theoretical sense. More precisely, we consider the benchmark problem of isolating all complex roots of a degree n polynomial with L bit integer coefficients. In parallel, also the complexity analysis for its real counterpart EVAL is given. We show that both algorithms have complexity bounds that match (in $\widetilde{O}$ sense) those of known exact and practical algorithms for real root isolation.

5.2.1 Notation

In the following considerations let $f \in \mathbb{Z}[z]$ be a square-free polynomial of degree $n \in \mathbb{N}$ whose coefficients have at most L bits. The complex roots of f and its derivative f' are denoted by $z_1, \dots, z_n$ and $z'_1, \dots, z'_{n-1}$, respectively. We further define $\sigma(z_i) := \min_{j \neq i} |z_i - z_j|$ as the distance of z_i to its nearest root and call $\sigma(z_i)$ the *separation* of z_i. W.lo.g. we assume that the roots are ordered with respect to their separations, that is, z_1 has the smallest and z_n the largest separation. For a given positive value δ let $k(\delta)$ be the largest index k such that $\sigma(z_k) \leq 56n^2\delta$. This apparently strange definition is justified by the results in the Theorems 22 and 23. We further assume that we start with an initial square box B_0 (interval), centered at the origin and size $s_0 := w(B_0) = d(B_0) = 2^{L+2}$. By Cauchy's bound [44,10], B_0 contains all roots of f (real roots in case of EVAL). $\mathcal{T}^{CE}$ and $\mathcal{T}^{EV}$ denote the subdivision trees induced by CEVAL and EVAL, respectively. At a certain depth $h \in \mathbb{N}$ of the subdivision tree all boxes (intervals) B have the same size $s_h := w(B) = d(B) = 2^{L+2-h}$. We denote by $\delta_h := 3s_h/4 = 3 \cdot 2^{L-h}$ which is an upper bound on the radius of each of these boxes (intervals).

5.2.2 Width of $\mathcal{T}^{CE}$ and $\mathcal{T}^{EV}$

For a given point m and radius r the success of our exclusion predicate

$$T_K^f(m,r): \ |f(m)| - K \sum_{k \geq 1} \left| \frac{f^{(k)}(m)}{k!} r^k \right| > 0 \Longleftrightarrow \sum_{k \geq 1} \left| \frac{f^{(k)}(m)}{f(m)} \right| \frac{r^k}{k!} < 1/K$$

mainly depends on the quotients $\left|\frac{f^{(k)}(m)}{f(m)}\right|$. Each of them can be rewritten as

$$\left|\frac{f^{(k)}(m)}{f(m)}\right| = \left|\sum\nolimits'_{i_1,\dots,i_k} \frac{1}{(m-z_{i_1})\dots(m-z_{i_k})}\right| \leq \left(\sum\nolimits_{i=1}^{n} \left|\frac{1}{m-z_i}\right|\right)^k$$

where the prime means that the i_j's ($j = 1, \dots, k$) are chosen to be distinct. In the following we investigate in a good estimation of the sum $\Sigma_1(m) := \sum_{i=1}^n \frac{1}{|m-z_i|}$. We start with a simple observation:

Lemma 21. Let $\delta > 0$ be an arbitrary positive real value and $|m - z_i| \geq n\delta$ for all i, then $T_K^f(m, r)$ is true for all $K < (e^{\frac{r}{\delta}} - 1)^{-1}$. In particular,

- $T_1(m, r)$ succeeds if $|m - z_i| > 2nr$ for all $i = 1, \dots, n$
- $T'_6(m, 4r)$, $T'_{3/2}(m, 8r)$ succeed if $|m - z'_i| > 28(n-1)r$ for all $i = 1, \dots, n-1$.

PROOF. From $|m - z_i| > n\delta$ we get $\Sigma_1(m) < 1/\delta$, thus

$$\sum_{k\geq1}\left|\frac{f^{(k)}(m)}{f(m)}\right|\frac{r^k}{k!} < \sum_{k\geq1}\frac{1}{k!}\left(\frac{r}{\delta}\right)^k < e^{\frac{r}{\delta}} - 1 \leq 1/K.$$

In particular, for $\delta := 2r$ ($\delta := 28r$ applied to f') as $(e^{1/2} - 1)^{-1} > 1$ ($(e^{1/7} - 1)^{-1} > 6$, $(e^{2/7} - 1)^{-1} > 3/2$).

Now let us consider a box of a certain depth h in the subdivision tree. If the midpoint $m(B)$ of such a box B fulfills $|m(B) - z_i| > 2n\delta_h$ for all $i = 1, \dots, n$ then $T_1(m(B), \delta_h)$ succeeds according to the previous lemma, thus B is omitted. For each root z_i, there exist at most $O(n^2)$ boxes with $|m_B - z_i| \leq 2n\delta_h$. Thus, in total, at most $O(n^3)$ boxes are retained. This straightforward observation is based on a pretty rough estimation of Σ_1 which assumes that, from a given point m, the distances to all roots z_i have roughly the same minimal value. In the following, we will use our results from Section 5.1 to show that this preliminary bound can be substantially improved.

Theorem 22. (Width of $\mathcal{T}^{CE}$) For each $h \in \mathbb{N}$, there exist disjoint square open axes-parallel boxes $B_1, \dots, B_k \subset \mathbb{C}$, $k \leq k(\delta_h)$, such that

(i) The total area of all boxes is at most

$$(8k(\delta_h)(1 + \ln\lceil|k(\delta_h)/2|\rceil)\delta_h)^2$$

(ii) For all $m \in \mathbb{C}\setminus\bigcup_{i=l}^k B_l$ either $T_1(m, \delta_h)$ or both, $T'_6(m, 4\delta_h)$ and $T'_{3/2}(m, 8\delta_h)$ succeed.

(iii) The width w_h of $\mathcal{T}^{CE}$ at the depth $h \in \mathbb{N}$ is bounded by

$$w_h < (12k(\delta_{h-1})(2 + \ln(k(\delta_{h-1}) + 1)))^2 = O(k(\delta_{h-1})^2 \ln k(\delta_{h-1})^2)$$

This result shows that the width of $\mathcal{T}^{CE}$ is bounded by $O(n^2(\ln n)^2)$. We next formulate a version of the above theorem for the real case. It shows that the EVAL algorithm induces a subdivision tree whose width is $O(n \ln n)$.

Theorem 23. (Width of $\mathcal{T}^{EV}$) There exist disjoint open intervals $I_1, \ldots, I_k$, $k \le k(\delta_h)$, on the real axes such that

(i) The total length of all intervals is smaller than or equal to

$$8k(\delta_h)(1 + \ln \lceil |k(\delta_h)/2| \rceil)\delta_h$$

(ii) For all $m \in \mathbb{R} \setminus \bigcup_{l=1}^{k} I_l$ either $T_1(m, \delta_h)$ or $T_1'(m, \delta_h)$ succeeds.

(iii) The width w_h of $\mathcal{T}^{EV}$ at the depth $h \in \mathbb{N}$ is bounded by

$$w_h < 3k(\delta_{h-1})(5 + \ln(k(\delta_{h-1}) + 1)) = O(k(\delta_{h-1}) \ln k(\delta_{h-1})) = O(n \ln n).$$

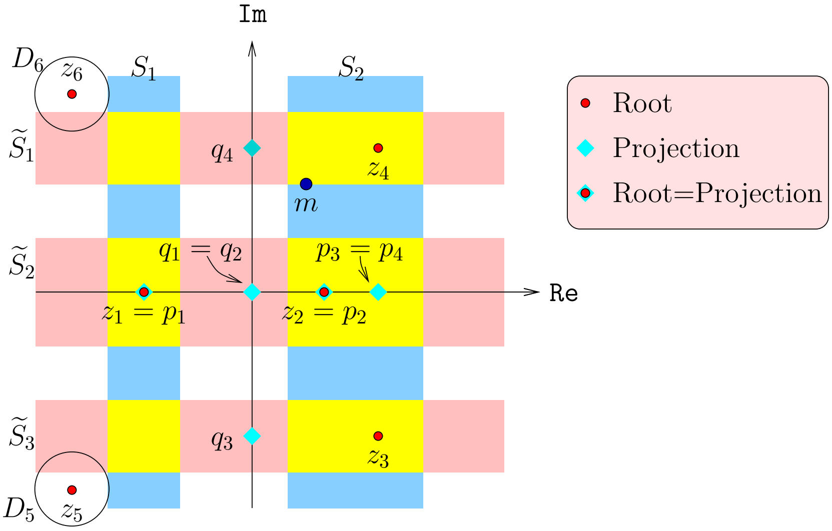


Fig. 7. The roots $z_1, \ldots, z_4$ define a multiset R with $\texttt{Re}[R] = \{p_1, \ldots, p_4\}$ and $\texttt{Im}[R] = \{q_1, \ldots, q_4\}$ the projections of R onto the real and imaginary axes. The corresponding δ-partitions define horizontal (pink) and vertical (blue) stripes $\widetilde{S}_i$ and S_j which intersect in disjoint boxes (yellow). The boxes which contains $z_1, \ldots, z_4$ are denoted by $B_1, \ldots, B_4$. Let m be a point on the boundary of one of the boxes B_i which is not contained in $D_5 \cup D_6$. Then its distance to z_5 and z_6 is larger than $28n\delta_h$, thus $\sum_{i=1}^{6} \frac{1}{|z_i - m|} < \frac{1}{2\delta_h} + \frac{1}{28\delta_h} = \frac{2}{3\delta_h}$.

We proceed with the proof of Theorem 22. Consider the set $R = \left\{z_1, \ldots, z_{k(\delta_h)}\right\}$, and

$$\delta^* := 4(1 + \ln \lceil k(\delta_h)/2 \rceil)\delta_h.$$

We apply Theorem 19 to R, using δ^* instead of δ and δ_h instead of ϵ: so there exist disjoint open axes-parallel boxes $B_1, \ldots, B_{\widetilde{k}}$, $\widetilde{k} \leq k(\delta_h)^2$, such that their union $\widetilde{\mathcal{B}} := \bigcup_{i=1,\ldots,\widetilde{k}} B_i$ has the following properties:
(a) $\widetilde{\mathcal{B}}$ contains all roots $z_1, \ldots, z_{k(\delta_h)}$.
(b) $\widetilde{\mathcal{B}}^{\delta_h}$ covers an area of at most $4k(\delta_h)^2(\delta_h + \delta^*)^2$. Here, $\widetilde{\mathcal{B}}^{\delta_h}$ denotes the union of all boxes in $\mathcal{B}$ where we enlarge each B_i by δ_h in each direction as in 5.1.2.
(c) For each point $m \notin \widetilde{\mathcal{B}}$ we have $\sum_{i=1}^{k(\delta_h)} \frac{1}{|m-z_i|} \leq \frac{1}{2\delta_h}$.

In the following we only consider those boxes which contain at least one of the roots in R. Wlog we can assume that these are the boxes $B_1, \ldots, B_k$, where $k \leq k(\delta_h)$. Obviously the properties (a) and (b) are also fulfilled for $\mathcal{B} := \bigcup_{i=1,\ldots,k} B_i$. Let $\partial\mathcal{B} := \bigcup_{i=1,\ldots,k} \partial B_i$ be the union of the boundaries of all boxes in $\mathcal{B}$, then for each $m \in \partial\mathcal{B}$ the property in (c) holds, as well.

For the remaining roots $z_{k(\delta_h)+1}, \ldots, z_n$, we consider disks $D_i := D_{28n\delta_h}(z_i)$, $i = k(\delta_h) + 1, \ldots, n$ of radius $28n\delta_h$, centered at z_i. We denote the union of all these discs by $\mathcal{D} := \bigcup_{i=k(\delta_h)+1}^{n} D_i$. Note that $\mathcal{D}$ is not necessarily disjoint from $\mathcal{B}$.

We now prove Part (ii) of Theorem 22. Let $m \in \mathbb{C}$ be an arbitrary point not contained in $\mathcal{B}$. We must show that either $T_1(m, \delta_h)$ holds, or $T'_6(m, 4\delta_h)$ and $T'_{3/2}(m, 8\delta_h)$ hold. We distinguish two cases:

- $m \in \mathcal{D}$: Wlog, we can assume that $m \in D_n$. By definition of $k(\delta_h)$, we have $\sigma(z_n) > 56n^2\delta_h$. From [10,44] we know that the distance from z_n to any root $z'_1, \ldots, z'_{n-1}$ of f' is larger than $\sigma(z_n)/n \geq 56n\delta_h$. Thus, the distance from m to any z'_i is larger than $28n\delta_h$. According to Lemma 21 the predicates $T'_6(m, 4\delta_h)$ as well as $T'_{3/2}(m, 8\delta_h)$ succeed, thus any box with center m and radius less than or equal to δ_h is terminal.
- $m \notin \mathcal{B} \cup \mathcal{D}$: On $\mathbb{C} \backslash (\mathcal{B} \cup \mathcal{D})$ each quotient $\frac{f^{(k)}}{f}$, $k = 1, \ldots, n$, defines a holomorphic function. For each of these functions we have $\lim_{z\to\infty} \frac{f^{(k)}}{f}(z) = 0$. Thus, according to the maximum principle, their maxima are either taken on the boundary of $\mathcal{B}$ or on the boundary $\partial\mathcal{D}$ of $\mathcal{D}$. Thus, in order to bound $\left|\frac{f^{(k)}}{f}(m)\right|$, we can restrict to these cases. If $m \in \partial D_i$ for one of the discs D_i then m is at least $28n\delta_h$ away from *all* roots of f and thus $\left|\frac{f^{(k)}}{f}(m)\right| \leq \left(\sum_{i=1}^{n} \frac{1}{28n\delta_h}\right)^k = \left(\frac{1}{28\delta_h}\right)^k$. It remains to discuss the case where m is on the boundary of one of the boxes. Then (c) holds and, in addition, $|m - z_i| \geq 28n\delta_h$ for all $i = k(\delta_h) + 1, \ldots, n$. It follows that

$$\left|\frac{f^{(k)}}{f}(m)\right| \le \left(\sum\nolimits_{i=1}^{k(\delta_h)} \frac{1}{|z_i - m|} + \sum\nolimits_{i=k(\delta_h)+1}^{n} \frac{1}{|z_i - m|}\right)^k$$
$$\le \left(\frac{1}{2\delta_h} + (n - k(\delta_h)) \cdot \frac{1}{28n\delta_h}\right)^k < \left(\frac{2}{3\delta_h}\right)^k.$$

Hence, in both situations we have $\left|\frac{f^{(k)}}{f}(m)\right| < \left(\frac{2}{3\delta_h}\right)^k$ and thus

$$\sum_{k=1}^{n} \left|\frac{f^{(k)}(m)}{f(m)}\right| \frac{\delta_h^k}{k!} < e^{2/3} - 1 < 1.$$

Hence, $T_1(m, \delta_h)$ succeeds and any box with center m and radius smaller than δ_h is terminal.

It remains to show (iii) about the number of boxes in Theorem 22. If the midpoint $m(B)$ of a box B of depth h is contained in $\mathcal{B}$ then B is completely contained in $\mathcal{B}^{\delta_h}$. $\mathcal{B}^{\delta_h}$ covers an area of at most $4k(\delta_h)^2(\delta_h + \delta^*) = (2k(\delta_h)\delta_h(5 + 4\ln\lceil k(\delta_h/2)\rceil))^2$. As all boxes B at depth h are pairwise disjoint and cover an area of at least $((4/3)\delta_h)^2$ it follows that at most

$$(\frac{3}{2}k(\delta_h)(5 + 4\ln\lceil k(\delta_h/2)\rceil))^2 < (6k(\delta_h)(2 + \ln(k(\delta_h) + 1)))^2$$

boxes are retained. As each non-terminal node has four children the width w_h of $\mathcal{T}^{CE}$ at height h is bounded by

$$(12k(\delta_{h-1})(2 + \ln(k(\delta_{h-1}) + 1)))^2 = O(n^2(\ln n)^2).$$

The proof of Theorem 23 is a direct consequence of our above considerations. Consider the intersection of $\mathcal{B}$ with the real axes. The overlapping consists of at most $k(\delta_h)$ intervals $I_1, \ldots, I_{\bar{k}}$ and the total length of their union $\mathcal{I} := \bigcup_{l=1,\ldots,\bar{k}} I_l$ is bounded by $2k(\delta_h)\delta^* = 8k(\delta_h)(1 + \ln\lceil k(\delta_h/2)\rceil)\delta_h$. We have already shown that for all points m outside these intervals either $T_1(m, \delta_h)$ or $T_6'(m, 4\delta_h)$ succeeds. Then trivially, either $T_1(m, \delta_h)$ or $T_1'(m, \delta_h)$ succeeds, as well. Hence, an interval I of length at most $2\delta_h$ with midpoint $m \notin \mathcal{I}$ is terminal. If an interval I with midpoint $m(I) \in \mathcal{I}$ has length at most $2\delta_h$, then it is completely contained in $\mathcal{I}^{\delta_h} := \bigcup_{l=1,\ldots,\bar{k}} I_l^{\delta_h}$, where $I_l^{\delta_h}$ is obtained by enlarging I_l by δ_h in both sides. Thus $\mathcal{I}^{\delta_h}$ has total length less than or equal to

$$2k(\delta_h)(\delta_h + \delta^*) < 2k(\delta_h)\delta_h(5 - \ln 2 + \ln(k(\delta_h) + 1)).$$

At depth h all intervals have width $\frac{4}{3}\delta_h$, thus at most $\frac{3}{2}k(\delta_h)(5 - \ln 2 + \ln(k(\delta_h) + 1))$ intervals are not terminal. As each non-terminal node in $\mathcal{T}^{EV}$ has two children the width of $\mathcal{T}^{EV}$ at depth h is bounded by $3k(\delta_{h-1})(5 + \ln(k(\delta_{h-1}) + 1))$.

5.2.3 Size of $\mathcal{T}^{CE}$ and $\mathcal{T}^{EV}$

The preceding analysis gives the width of the trees $\mathcal{T}^{CE}$ and $\mathcal{T}^{EV}$. We now bound their sizes. In particular, our result shows that the subdivision tree induced by the EVAL algorithm is, at least in terms of $\widetilde{O}$-complexity, as good as that of well-known methods for real root isolation using Descartes' Rule of Sign or Sturm sequences.

Theorem 24. For a square-free polynomial f of degree n with coefficients of at most L bits, the size of $\mathcal{T}^{CE}$ is $O((n \ln n)^2(L + \ln n)) = \widetilde{O}(n^2L)$. For $\mathcal{T}^{EV}$, the size is $O(n(L + \ln n)(\ln L + \ln n)) = \widetilde{O}(nL)$.

PROOF. We first investigate in a bound on $k(\delta_h)$. As in the proof of Theorem 22, consider the set R consisting of those roots $z_1, \ldots, z_{k(\delta_h)}$ with separation $\sigma(z_i) \leq 56n^2\delta_h$. Then according to Lemma 20, there exists a partition of R into disjoints sets $R_1, \ldots, R_k$ such that $|R_{i_0}| \geq 2$ for each $i_0 = 1, \ldots, n$ and $|z_i - z_j| \leq 56n^2\delta_h|R| \leq 56n^3\delta_h$ for all pairs $z_i, z_j \in R_{i_0}$. We consider a directed graph $\mathcal{G}_i$ on R_i which connects consecutive points of R_i in ascending order of their absolute values. We define $\mathcal{G} := (R, E)$ as the union of all $\mathcal{G}_i$. Then $\mathcal{G}$ is a directed graph on R with the following properties:

(1) each edge $(\alpha, \beta) \in E$ satisfies $|\alpha| \leq |\beta|$,
(2) $\mathcal{G}$ is acyclic, and
(3) the in-degree of any node is at most 1.

Hence, we can apply the generalized Davenport-Mahler bound [9,10] on $\mathcal{G}$:

$$\prod_{(\alpha,\beta)\in E} |\alpha - \beta| \geq \frac{1}{((n+1)^{1/2}2^L)^{n-1}} \cdot \left(\frac{\sqrt{3}}{n}\right)^{\#E} \cdot \left(\frac{1}{n}\right)^{n/2}$$

As each set R_i contains at least 2 roots, we must have $\#E \geq k(\delta_h)/2$. Furthermore, for each edge $(\alpha, \beta) \in E$ we have $|\alpha - \beta| \leq 56n^3\delta_h = 168n^32^{L-h}$. It follows that

$$\begin{aligned}\left(168n^32^{L-h}\right)^{\frac{k(\delta_h)}{2}} &\geq \frac{1}{((n+1)^{1/2}2^L)^{n-1}} \cdot \left(\frac{\sqrt{3}}{n}\right)^{k(\delta_h)} \cdot \left(\frac{1}{n}\right)^{n/2} \\ &> \frac{1}{(n+1)^n 2^{nL}} \cdot \left(\frac{3}{n^2}\right)^{k(\delta_h)/2}\end{aligned}$$

and thus

$$k(\delta_h) \cdot (5 + 5\ln n + (L - h)\ln 2) > -2n(L \ln 2 + \ln(n+1))$$

where we used the inequality $\ln 56 < 5$. Thus, for $h > L + 8(1 + \ln n) >$

$L + \frac{5}{\ln 2} + \frac{5 \ln n}{\ln 2}$, we get

$$k(\delta_h) < \frac{2n(L \ln 2 + \ln(n+1))}{(h-L)\ln 2 - 5 - 5\ln n} < \frac{3n(L + \ln(n+1))}{h - L}. \tag{16}$$

Since $h > L + 8(1 + \ln n)$, we may define $h' := h - h_0 \in \mathbb{N}$ where $h_0 := \lceil L + 8(1 + \ln n) \rceil$. Then (16) transforms into

$$k(\delta_h) < \frac{3n(L + \ln(n+1))}{h' + \lceil 8(1 + \ln n) \rceil} < \frac{3n(L + \ln(n+1))}{h'}. \tag{17}$$

For all $h \leq 2h_0$ we use the simple inequality $k(\delta_h) \leq n$ whereas for $h > 2h_0$ we use the bound on $k(\delta)$ in (17). From (17), we can bound on the height $h_{\max}$ of T^{CE} as follows. Observe by Theorem 22(iii) that when $k(\delta_h) = 0$ then the width is 0. So we may assume $k(\delta_h) \geq 1$ in (17). Therefore $h' \leq 3n(L + \ln(n+1))$. Therefore

$$h_{\max} \leq h' + h_0 \leq 3n(L + \ln(n+1)) + L + 8(1 + \ln n) = O(n(L + \ln n)).$$

Now we are able to compute the size of T^{CE}:

$$\begin{aligned}
\left|T^{CE}\right| &\leq \sum_{h=1}^{h_{\max}} (12k(\delta_{h-1})(2 + \ln(k(\delta_{h-1}) + 1)))^2 \qquad \text{(by Theorem 22)} \\
&\leq 144 \sum_{h=1}^{2h_0} (n(\ln(n+1) + 2))^2 + 144 \sum_{h'=h_0+1}^{h_{\max}-h_0} 9n^2 \left(\frac{L + \ln(n+1)}{h'}\right)^2 \cdot (2 + \ln(n+1))^2 \\
&= O(n^2(\ln n)^3 + L(n \ln n)^2) + O(n^2(L + \ln n)^2) \cdot (2 + \ln(n+1))^2 \cdot \sum_{h'=h_0+1}^{h_{\max}-h_0} \left(\frac{1}{h'}\right)^2 \\
&= O(n^2(\ln n)^3 + L(n \ln n)^2) + O(n^2(L + \ln n)^2) \cdot (2 + \ln(n+1))^2 \cdot \frac{1}{L + \ln n} \\
&= O((n \ln n)^2(L + \ln n)) = \widetilde{O}(n^2 L).
\end{aligned}$$

For the size of T^{EV} we obtain

$$\begin{aligned}
\left|T^{EV}\right| &\leq \sum_{h=1}^{h_{\max}} 3k(\delta_{h-1})(5 + \ln(k(\delta_{h-1}) + 1)) \\
&\quad \text{(by Theorem 23)} \\
&\leq 3 \sum_{h=1}^{h_0} n(\ln(n+1) + 5) + 9(5 + \ln(n+1)) \cdot \sum_{h'=1}^{h_{\max}-h_0} n \frac{L + \ln(n+1)}{h'} \\
&= O(n \ln n(L + \ln n)) + O(n(L + \ln n) \ln h_{\max} \ln n) \\
&= O(n(L + \ln n)(\ln L + \ln n) = \widetilde{O}(nL).
\end{aligned}$$

We will see that the larger tree size of T^{CE} does not lead to an asymptotically larger bit complexity when compared to T^{EV}. More precisely, both algorithms use $\widetilde{O}(n^4L^2)$ bit operations to isolate the roots of f (either real or complex).

Theorem 25. For a square-free polynomial f of degree n with integer coefficients of at most L bits, CEVAL and EVAL isolate the complex (real) roots of f with a number Δ^{CE} (Δ^{EV}) of bit operations bounded by $\widetilde{O}(n^4L^2)$.

PROOF. We refer to Section 6 where we show that, for each node v of T^{CE} (T^{EV}) of depth h, the number λ_v of bit operations is bounded by $\widetilde{O}(nL+n^2h)$. For all $h \leq 2h_0 = 2\lceil L+8(1+\ln n)\rceil = \widetilde{O}(L)$ this simplifies to $\lambda_v = \widetilde{O}(n^2L)$. Now our claim about the bit complexity derives from a simple computation (cf. proof of Theorem 24):

$$\begin{aligned}\Delta^{CE} &\leq \sum_{h=1}^{2h_0}(n(\ln(n+1)+2))^2\widetilde{O}(n^2L)\\ &+ \sum_{h'=h_0+1}^{h_{\max}-h_0} n^2\left(\frac{L+\ln(n+1)}{h'}\right)^2\widetilde{O}(nL+n^2(h'+h_0))\\ &= \widetilde{O}(n^4L^2)+\sum_{h'=h_0+1}^{h_{\max}-h_0} n^4\left(\frac{L+\ln(n+1)}{h'}\right)^2\widetilde{O}(h')\\ &= \widetilde{O}(n^4L^2)(1+\sum_{h'=h_0+1}^{h_{\max}-h_0}\frac{1}{h'}) = \widetilde{O}(n^4L^2). \qquad (18)\end{aligned}$$

In the above inequality (18), we use $\widetilde{O}(nL+n^2(h'+h_0)) = \widetilde{O}(n^2h')$ because the second summation is only summed over $h' \geq h_0 > L$.
For the EVAL algorithm, the computation turns out to be a little simpler, although the final bound is the same:

$$\begin{aligned}\Delta^{EV} &\leq \sum_{h=1}^{h_0} n(\ln(n+1)+5)\widetilde{O}(nL+n^2L)\\ &+ \sum_{h'=1}^{h_{\max}-h_0} n\frac{L+\ln(n+1)}{h'}\widetilde{O}(nL+n^2(h'+h_0))\\ &= \widetilde{O}(n^3L^2)+\widetilde{O}(n^3L^2)\sum_{h'=1}^{h_{\max}-h_0}\frac{1}{h'}+\widetilde{O}(n^3L)\sum_{h'=1}^{h_{\max}-h_0}\frac{1}{h'}\widetilde{O}(h') = \widetilde{O}(n^4L^2)\end{aligned}$$

6 Exactness and Other Implementation Issues

Our CEVAL algorithm is meant to be practical and suitable for exact implementation. In this section, we address the exactness question and also some techniques to improve the practical efficiency of CEVAL.

The basis for all our numerical computation is the set of BigFloats or dyadic numbers, $\mathbb{F} = \{m2^n : m, n \in \mathbb{Z}\} = \mathbb{Z}[\frac{1}{2}]$. The ring operations $(+, -, \times)$ are exact in $\mathbb{F}$, as is division by 2. But general division will be approximated. See [45] for discussion of the use of $\mathbb{F}$ for general real computation. In this paper, we use the obvious extension to complex dyadic numbers $\mathbb{F}[\mathbf{i}]$. All input numbers will be assumed to be dyadic; in particular, the polynomial f has coefficients in $\mathbb{F}[\mathbf{i}]$, and the initial box $B_0 = Box(\mu, \xi)$ where $\mu, \xi \in \mathbb{F}[\mathbf{i}]$. Subsequent subdivision boxes remain dyadic.

Note that m is dyadic, but the exact radius r of the box is not. But we can replace r by any dyadic upper bound: for square boxes of width w, we may use the dyadic value $3/4w$ for r.

Next we consider the 8 compass points: the cardinal points (N, S, E, W) are dyadic, but the ordinal points (NE, SE, SW, NW) are not. In fact, dyadic points are generally impossible, and we must settle for some choice of rational points. The proof on the exactness of our algorithm (cf. Theorem 14 in Appendix 4.1) shows that it is sufficient to choose a set of 8 angles $\{\theta_i : i = 0, \ldots, 7\}$ that are pairwise separated by angles in the range $[45° \pm \delta]$ such that each θ_i is **Pythagorean**, i.e., $\sin(\theta_i)$ and $\cos(\theta_i)$ are rational values. It is well known that such angles are obtained from Pythagorean triples $(x, y, z) \in \mathbb{N}^3$ where $x^2 + y^2 = z^2$, and it is also easy to generate such triples.

The amount of deviation δ depends on the choice of some constants K and L — we have not tried to optimize this choice. In Theorem 14, we show that if $(K, L) = (6, 4)$ then we can choose $\delta = 2.5°$. For our purposes, we only need to approximate the ordinal points. A useful Pythagorean triple for this purpose is $(x, y, z) = (20, 21, 29)$ Note that $\arcsin(20/29) \approx 43.60°$.

In the 8-Point Test, it is not necessary to compute $u(x, y)$ and $v(x, y)$ separately. Instead, we simply evaluate the function $f(z)$ at the compass points P. Note that P will be rational, not dyadic. The signs $u(P) := \mathtt{Re}(f(P))$ and $v(P) := \mathtt{Im}(f(P))$ can usually be obtained exactly by interval arithmetic; even exact sign can be obtained by sufficiently high accuracy approximation and using zero bounds.

Our main predicates are based on the explicit representation of $f_B(z) := f(z + m(B))$, the Taylor expansion of f at the center $m(B)$ of a box B. The direct computation of f_B from f at each node of the recursion is costly. It is better to compute f_B incrementally as follows: let B' be (wlog) the upper right child of B. We may also assume that our initial box is a square of width 2^{-k} ($k \in \mathbb{Z}$), centered at a dyadic point. Recursively, B is also a square of this form. So

$$f(m(B')+z) = f(m(B)+2^{-k-2}(1+\mathbf{i})+z) = f(m(B)+2^{-k-2}(2^{k+2}z+1+\mathbf{i})).$$

We now compute $f_{B'}$ from $f_0(z) := f_B(z)$ in three steps: First scale the function f_0 by the substitution $z \mapsto 2^{-k-2}z$ to obtain $f_1(z) := f_0(2^{-k-2}z)$. Next apply a Taylor shift by 1, and a shift by $\mathbf{i}$, to get $f_2(z) := f_1(z+1+\mathbf{i})$. Finally, we scale again with $z \mapsto 2^{k+2}z$ to yield our goal, $f_{B'}(z) = f_2(2^{k+2}z)$.

Assume the standard encoding of binary floating point numbers, each scaling $z \mapsto 2^k z$ amounts to adding k to the exponents of the polynomial coefficients. Thus the computational cost is dominated by the Taylor shifts. A Taylor shift by $\mathbf{i}$ can be realized as a Taylor shift by 1 combined with two scalings by $\mathbf{i}$, an immediate consequence of the identity $f(z+i) = f(i(-iz+1))$. Using classical Taylor shift [17], a shift by 1 requires $\widetilde{O}(n^2(n+\tau_B))$ bit operations, where the coefficients of f_B are represented by 2^{τ_B}-bit dyadic numbers. However, using asymptotically fast Taylor shift [14,10], this number reduces to $\widetilde{O}(n(n+\tau_B))$. Also, the bit complexity of the coefficients grow by $O(n)$ bits in every node. As we start with a polynomial f with integer coefficients of at most L bits, we get $\tau_B = O(L+nh)$ for each box B at depth h in the subdivision tree. For the predicate evaluations and the point evaluations in the 8-Point Test we have to compute the value of a polynomial, whose coefficients are $O(L+nh)$-bit dyadic numbers, at a point of bit complexity $O(1)$. Therefore $O(L+nh)$ bit operations are sufficient. It follows that the overall number of bit operations at a node of depth h is bounded by $\widetilde{O}(n(n+\tau_B)) = \widetilde{O}(n(L+nh))$.

Our description of the test T_K does not exploit any additional information that might be obtained while doing the test. For instance, in practice, the test $T_K(m,r)$ proceeds by computing the remainder term

$$R(m,r) := \sum_{k\geq 1} \left|\frac{f^{(k)}(m)}{k!}\right| r^k$$

and $|f(m)|$ separately. Then the test $T_K(m,r)$ amounts to checking if $|f(m)| > K \cdot R(m,r)$. Even in the case of a failure, the values $|f(m)|$ and $R(m,r)$ can be used to determine subsequent actions. In the case of the exclusion test $T_1(m,r)$, a success currently excludes the box contained in $D_r(m)$; but in fact, the computed value may justify the exclusion of a much larger region. Furthermore, from the Taylor expansion at m we can also derive approximate information (via root bounds) about the largest r that fulfills $|f(m)| > R(m,r)$ which allows us to exclude disc regions of a certain size and thus to approximate the roots faster than by simple bisection. But to exploit this, we need to leave the comfortable framework of quadtrees. In future work, we will explore such techniques. We remark that all these techniques directly correspond to the additional steps done in the continued fraction algorithm in comparison to the bisection algorithm by Vincent, Collins and Akritas.

We observe that the computation of $R(m, r)$ requires general division operations even if m and r are dyadic. However, we could compute the dyadic value $d!R(m, r)$ where $d = \deg f(z)$, or simply compute a dyadic upper bound for $R(m, r)$.

7 Conclusion

This paper continues a line of recent work to develop exact subdivision algorithms based on the Bolzano principle. The primitives in such algorithms are based on numerical function evaluation, and hence simple to implement, extendible to analytic functions, and quite practical.

Here we introduce a new complex root isolation algorithm whose asymptotic complexity is shown to be competitive (up to logarithmic factors) with known exact practical algorithms. It is somewhat unexpected that algorithms based on such simple primitives can match those based on more sophisticated primitives such as found in Descartes, Continued Fraction or Sturm methods. Another surprise is that the complex case has (up to logarithmic terms) the same bit complexity as the real case.

Our complexity analysis introduces new ideas including a technique of root clusters which is expected to have other applications. One open problem is to sharpen our complexity estimates (only improvements in logarithmic terms can be expected).

Another direction is to develop practical techniques for implementing our method. We expect the competitiveness of Bolzano to be observed in practice. Although our theoretical bounds are based on asymptotically fast Taylor shifts, experiments by various authors [42,16,10] have shown that the straightforward implementation is probably better for degrees up to 1000. Various trade offs arise in their use in the algorithm, and it is clear that we can exploit partial information obtained while evaluating such predicates.

The Descartes method had been successfully extended to the so called bitstream model [11,24] in which the coefficients of the input polynomial are given by a bitstream on-demand. It has useful applications in situations where the coefficients are algebraic numbers (e.g., in cylindrical algebraic decomposition). We plan to extend EVAL and CEVAL algorithms in this direction. This is justified by the following simple observation: CEVAL does not only isolate the roots of f but it also comes with a lower bound on the separation $\sigma(z_i)$ of each root z_i: If B_i is an isolating box for z_i in the output then $\sigma(z_i) > r(B_i)$ (see also 3.5 (ii)). In [24] this was the critical property. We claim that, in a complete analogous manner, our algorithm also extends to the bitstream model. This will be in future work.

References

[1] A. G. Akritas and A. Strzeboński. A comparative study of two real root isolation methods. *Nonlinear Analysis:Modelling and Control*, 10(4):297–304, 2005.

[2] M. Burr, S. Choi, B. Galehouse, and C. Yap. Complete subdivision algorithms, II: Isotopic meshing of singular algebraic curves. In *Proc. Int'l Symp. Symbolic and Algebraic Computation (ISSAC'08)*, pages 87–94, 2008. Hagenberg, Austria. Jul 20-23, 2008.

[3] M. Burr, F. Krahmer, and C. Yap. Integral analysis of evaluation-based real root isolation. Submitted, 2009. Download from http://cs.nyu.edu/exact/papers/.

[4] M. Burr, V. Sharma, and C. Yap. Evaluation-based root isolation, Feb. 2009. In preparation.

[5] G. E. Collins and A. G. Akritas. Polynomial real root isolation using Descartes' rule of signs. In R. D. Jenks, editor, *Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation*, pages 272–275. ACM Press, 1976.

[6] G. E. Collins, J. R. Johnson, and W. Krandick. Interval arithmetic in cylindrical algebraic decomposition. *J. of Symbolic Computation*, 34:145–157, 2002.

[7] G. E. Collins and R. Loos. Real zeros of polynomials. In B. Buchberger, G. E. Collins, and R. Loos, editors, *Computer Algebra*, pages 83–94. Springer-Verlag, 2nd edition, 1983.

[8] J. H. Davenport. Computer algebra for cylindrical algebraic decomposition. Tech. Rep., The Royal Inst. of Technology, Dept. of Numerical Analysis and Computing Science, S-100 44, Stockholm, Sweden, 1985. Reprinted as Tech. Report 88-10, School of Mathematical Sci., U. of Bath, Claverton Down, Bath BA2 7AY, England. URL http://www.bath.ac.uk/ masjhd/TRITA.pdf.

[9] Z. Du, V. Sharma, and C. Yap. Amortized bounds for root isolation via Sturm sequences. In D. Wang and L. Zhi, editors, *Symbolic-Numeric Computation*, Trends in Mathematics, pages 113–130. Birkhäuser Verlag AG, Basel, 2007. Proc. Int'l Workshop on Symbolic-Numeric Computation, Xi'an, China, Jul 19–21, 2005.

[10] A. Eigenwillig. *Real Root Isolation for Exact and Approximate Polynomials using Descartes' Rule of Signs*. PhD thesis, Universität des Saarlandes, May 2008.

[11] A. Eigenwillig, L. Kettner, W. Krandick, K. Mehlhorn, S. Schmitt, and N. Wolpert. A Descartes algorithm for polynomials with bit stream coefficients. In *8th Int'l Workshop on Comp.Algebra in Sci.Computing (CASC 2005)*, pages 138–149. Springer, 2005. LNCS 3718.

[12] A. Eigenwillig, V. Sharma, and C. Yap. Almost tight complexity bounds for the Descartes method. In *Proc. Int'l Symp. Symbolic and Algebraic Computation (ISSAC'06)*, pages 71–78, 2006. Genova, Italy. Jul 9-12, 2006.

[13] I. Z. Emiris and E. P. Tsigaridas. Univariate polynomial real root isolation: Continued fractions revisited. In Y. Azar and T. Erlebach, editors, *Proc. 13th European Symp. on Algorithms (ESA)*, volume 4168 of *Lecture Notes in Computer Science*, pages 817–828. Springerl-Verlag, 2006.

[14] J. Gerhard. Modular algorithms in symbolic summation and symbolic integration. *LNCS, Springer*, 3218, 2004.

[15] M. Hemmer, E. P. Tsigaridas, Z. Zafeirakopoulos, I. Z. Emiris, M. I. Karavelas, and B. Mourrain. Experimental evaluation and cross-benchmarking of univariate real solvers. In *SNC '09: Proceedings of the 2009 conference on Symbolic numeric computation*, pages 45–54, New York, NY, USA, 2009. ACM.

[16] J. Johnson. Algorithms for polynomial real root isolation. In B. Caviness and J. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, Texts and monographs in Symbolic Computation, pages 269–299. Springer, 1998.

[17] J. R. Johnson, W. Krandick, and A. D. Ruslanov. Architecture-aware classical Taylor shift by 1. In *Proc. 2005 International Symposium on Symbolic and Algebraic Computation (ISSAC 2005)*, pages 200–207. ACM, 2005.

[18] W. Krandick and G. E. Collins. An efficient algorithm for infallible polynomial complex root isolation. In *ISSAC 97*, pages 189–194, 1992.

[19] W. Krandick and K. Mehlhorn. New bounds for the Descartes method. *J. Symbolic Computation*, 41(1):49–66, 2006.

[20] T. Lickteig and M.-F. Roy. Sylvester-Habicht sequences and fast Cauchy index computation. *J. of Symbolic Computation*, 31:315–341, 2001.

[21] L. Lin and C. Yap. Adaptive isotopic approximation of nonsingular curves: the parametrizability and non-local isotopy approach. In *Proc. 25th ACM Symp. on Comp. Geometry*, page to appear, June 2009. Aarhus, Denmark, Jun 8-10, 2009.

[22] J. McNamee. A bibliography on roots of polynomials. *Journal of Computing and Applied Mathematics*, 47:391–394, 1993. Available online at http://www.elsevier.com/homepage/sac/cam/mcnamee.

[23] K. Mehlhorn and S. Ray. Faster algorithms for computing hong's bound on absolute positiveness. *Journal of Symbolic Computation*, 2009. submitted.

[24] K. Mehlhorn and M. Sagraloff. Isolating real roots of real polynomials. In *ISSAC 09*, 2009.

[25] D. P. Mitchell. Robust ray intersection with interval arithmetic. In *Graphics Interface'90*, pages 68–74, 1990.

[26] R. E. Moore. *Interval Analysis.* Prentice Hall, Englewood Cliffs, NJ, 1966.

[27] B. Mourrain, F. Rouillier, and M.-F. Roy. The Bernstein basis and real root isolation. In J. E. Goodman, J. Pach, and E. Welzl, editors, *Combinatorial and Computational Geometry*, number 52 in MSRI Publications, pages 459–478. Cambridge University Press, 2005.

[28] B. Mourrain, M. N. Vrahatis, and J. C. Yakoubsohn. On the complexity of isolating real roots and computing with certainty the topological degree. *J. Complexity*, 18:612–640, 2002.

[29] V. Y. Pan. Sequential and parallel complexity of approximate evaluation of polynomial zeros. *Comput. Math. Applic.*, 14(8):591–622, 1987.

[30] V. Y. Pan. New techniques for approximating complex polynomial zeros. *Proc. 5th ACM-SIAM Symp. on Discrete Algorithms (SODA94)*, pages 260–270, 1994.

[31] V. Y. Pan. Optimal (up to polylog factors) sequential and parallel algorithms for approximating complex polynomial zeros. *Proc. 27th STOC*, pages 741–750, 1995.

[32] V. Y. Pan. On approximating polynomial zeros: Modified quadtree (weyl's) construction and improved newton's iteration. Research report 2894, INRIA, Sophia-Antipolis, 1996.

[33] V. Y. Pan. Solving a polynomial equation: some history and recent progress. *SIAM Review*, 39(2):187–220, 1997.

[34] J. R. Pinkert. An exact method for finding the roots of a complex polynomial. *ACM Trans. on Math. Software*, 2:351–363, 1976.

[35] S. Plantinga and G. Vegter. Isotopic approximation of implicit curves and surfaces. In *Proc. Eurographics Symposium on Geometry Processing*, pages 245–254, New York, 2004. ACM Press.

[36] H. Ratschek and J. Rokne. *Computer Methods for the Range of Functions.* Horwood Publishing Limited, Chichester, West Sussex, UK, 1984.

[37] D. Reischert. Asymptotically fast computation of subresultants. In *ISSAC 97*, pages 233–240, 1997. Maui, Hawaii.

[38] F. Rouillier and P. Zimmermann. Efficient isolation of [a] polynomial's real roots. *J. Computational and Applied Mathematics*, 162:33–50, 2004.

[39] A. Schönhage. The fundamental theorem of algebra in terms of computational complexity, 1982. Manuscript, Department of Mathematics, University of Tübingen. Updated 2004.

[40] V. Sharma. Complexity of real root isolation using continued fractions. *Theor. Computer Science*, 409(2), 2008. Also: proceedings ISSAC'07.

[41] S. Smale. The fundamental theorem of algebra and complexity theory. *Bulletin (N.S.) of the AMS*, 4(1):1–36, 1981.

[42] J. von zur Gathen and J. Gerhard. Fast algorithms for Taylor shifts and certain difference equations. In *Proc. 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC 1997)*, pages 40–47. ACM, 1997.

[43] H. S. Wilf. A global bisection algorithm for computing the zeros of polynomials in the complex plane. *J. of the ACM*, 25(3):415–420, 1978.

[44] C. K. Yap. *Fundamental Problems of Algorithmic Algebra.* Oxford University Press, 2000.

[45] C. K. Yap. Theory of real computation according to EGC. In P. Hertling, C. Hoffmann, W. Luther, and N.Revol, editors, *Reliable Implementation of Real Number Algorithms: Theory and Practice*, number 5045 in Lecture Notes in Computer Science, pages 193–237. Springer, 2008.