Near Optimal Subdivision Algorithms for Real Root Isolation

Vikram Sharma
Institute of Mathematical Sciences
Chennai, India 600113
vikram@imsc.res.in

Prashant Batra
TU Hamburg
Hamburg, Germany
batra@tuhh.de

January 22, 2015

Abstract

The problem of isolating real roots of a square-free polynomial inside a given interval I_0 is a fundamental problem. Subdivision based algorithms are a standard approach to solve this problem. Given an interval I, such algorithms rely on two predicates: an exclusion predicate, which if true means I has no roots, and an inclusion predicate, which if true, reports an isolated root in I. If neither predicate holds, then we subdivide the interval and proceed recursively, starting from I_0. Example algorithms are Sturm's method (predicates based on Sturm sequences), the Descartes method (using Descartes's rule of signs), and `Eval` (using interval-arithmetic). For the canonical problem of isolating all real roots of a degree n polynomial with integer coefficients of bit-length L, the subdivision tree size of (almost all) these algorithms is bounded by $O(n(L + \log n))$. This is known to be optimal for subdivision based algorithms.

We describe a subroutine that improves the running time of any subdivision algorithm for real root isolation. The subroutine first detects clusters of roots using a result of Ostrowski, and then uses Newton iteration to converge to them. Near a cluster, we switch to subdivision, and proceed recursively. The subroutine has the advantage that it is independent of the predicates used to terminate the subdivision. This gives us an alternative and simpler approach to recent developments of Sagraloff (2012) and Sagraloff-Mehlhorn (2013), assuming exact arithmetic.

The subdivision tree size of our algorithm using predicates based on Descartes's rule of signs is bounded by $O(n \log n)$, which is better by $O(n \log L)$ compared to known results. Our analysis differs in two key aspects. First, we use the general technique of continuous amortization from Burr-Krahmer-Yap (2009), and second, we use the geometry of clusters of roots instead of the Davenport-Mahler bound. The analysis naturally extends to other predicates.

1 Introduction

Given a polynomial $f \in \mathbb{R}[x]$ of degree n, the problem is to isolate the real roots of f in an input interval I_0, i.e., compute disjoint intervals which contain exactly one real root of f, and together contain all roots of f in $I_0 \cap \mathbb{R}$. Subdivision based algorithms have been successful in addressing the problem. A general subdivision algorithm uses two predicates, given an interval I: the exclusion predicate $C_0(I)$, which if true means I has no roots; the inclusion predicate, $C_1(I)$, which if true means I has exactly one root. The algorithm outputs a **root-partition** $\mathcal{P}$ of I_0, i.e., a set of pairwise disjoint open intervals such that for each interval either C_0 or C_1 holds, and $I_0 \setminus \mathcal{P}$ contains no roots of f. To compute isolating intervals for roots of f, check the sign of f at the endpoints of the intervals in $\mathcal{P}$ (this works if f is square-free). The following generic subdivision algorithm constructs a root-partition:

```
Isolate(I_0)
INPUT: f ∈ ℝ[x] and an interval I_0 ⊆ ℝ.
OUTPUT: A root-partition 𝒫 of f in I_0.
0.  Preprocessing step.
1.  Initialize a queue Q with I_0, and 𝒫 ← ∅.
2.  While Q is not empty
        Remove an interval I = (a, b) from Q.
        If C_0(I) ∨ C_1(I) then add I to 𝒫.
        else  ◁ Subdivide I
            Let m ← (a + b)/2.
            Push (a, m) and (m, b) into Q.
3.  Output 𝒫.
```

The algorithm is guaranteed to terminate for square-free polynomials; otherwise we get an infinite sequence of intervals converging to a root of multiplicity greater than one. Some standard choices of the predicates and the corresponding algorithms are:

(i) Sturm sequences and Sturm's method [5],

(ii) Descartes's rule of signs and the Descartes method [4],

(iii) Interval-arithmetic based approaches and `Eval` [2].

The complexity of these algorithms is well understood for the benchmark problem of isolating all real roots of a square-free integer polynomial with coefficients of bit-length L. One measure of complexity is the size of the subdivision tree constructed by the algorithm. For the first two algorithms a bound of $O(n(L + \log n))$ was shown in [5] and [6], respectively. For `Eval` a weaker bound of $O(n(L + n))$ was established in [16]. It is also known [6] that the bound $O(n(L + \log n))$ is essentially tight for any algorithm doing uniform subdivision, i.e., reduces the width at every step by some constant (in our case by half).

Uniform subdivision cannot improve on the bound mentioned above because it only gives linear convergence to a "root cluster", i.e., roots which are relatively closer to each other than to any other root. But it is known that from points sufficiently far away from the cluster, Newton iteration (more precisely, its variants for multiple roots) converges quadratically to the cluster. This has been an underlying idea in improving the linear convergence of subdivision algorithms for root isolation [9, 11, 12], and has also been combined with homotopy based approaches [17, 14].

We follow the same idea with some key differences. Given C_0 and C_1, our algorithm can be described as follows (we only give the inner loop, see Section 3 for complete details):

```
Newton-Isol(I_0)
    ...
    If C_0(I) ∨ C_1(I) then add I to 𝒫.
    else if a cluster 𝒞 of roots is detected in I then
        Apply Newton iteration to approximate 𝒞
        while quadratic convergence holds.
        Estimate an interval J containing 𝒞.
        Push J into Q.
    else  ◁ Subdivide I
        ...
```

For detecting root clusters, we use a result of Ostrowski based on Newton diagram of a polynomial [8]; other choices are based on a generalization of Smale's α-theory (see [7] and the references therein); the details can be found in Section 2. These tools and approaches have been used earlier [9], however, our approach has the following differences:

(i) The tools used to detect and estimate the size of a cluster are independent of the particular choice of the exclusion-inclusion predicates (cf. [11]). This way we obtain a general approach to improve any subdivision algorithm.

(ii) Another difference is the method that is combined with bisection to improve convergence. In [11] Abbott's QIR method is combined with the Schröder operator [7], whereas we apply standard Newton iteration to a suitable derivative of f. The former combination is a backtracking approach to get quadratic convergence; the latter gives quadratic convergence right away (but perhaps increasing subdivisions). This has the advantage of separating the Newton iteration steps from the subdivision tree, which is reflected in the bounds on the subdivision tree size for the two approaches: for the former we have $O(n \log(nL))$, and for the latter we have $O(n \log n)$. The number of quadratically converging steps remains the same in both cases.

(iii) Our approach can be modified to isolate complex roots; replace binary subdivision with a quad-tree subdivision, and choose appropriate predicates (e.g., Ostrowski's result mentioned above, or Pellet's test). This avoids Graeffe iteration (cf. [9]), and yet the modified algorithm can be shown to attain a near optimal bound on subdivision tree size.

In this paper, we focus on bounding the size of subdivision tree of `Newton-Isol`. For this purpose, we use the general framework of continuous amortization [2, 3]. The key idea here is to bound the tree size by $\int_{I_0} G(x)dx$, where G is a suitable "charging" function corresponding to the predicates used in the algorithm (e.g., see [3]). Our key contributions are the following:

(i) We derive a near optimal bound of $O(n \log n)$ on the size of the subdivision tree of `Newton-Isol` when C_0, C_1 are based on Descartes's rule of signs (see Theorem 10). This is the first application of the continuous amortization framework to a non-uniform subdivision algorithm.

(ii) We show that if the distance of the cluster center to the nearest root outside the cluster exceeds roughly n^3 times the diameter of the cluster, then Ostrowski's criterion for cluster detection works, and we obtain quadratic convergence to the cluster center (see Lemma 6).

(iii) Our analysis crucially uses the cluster tree of the polynomial (see Proposition 1). We derive an integral bound on the size of the subdivision tree (see Theorem 9). The usual approach to upper bound this integral is to break it over the (real) Voronoi regions of the roots [3]. We instead break the integral over the Voronoi regions corresponding to the clusters in an inductive manner based on the cluster tree. The integral over the portion of the region outside the cluster is bounded using known techniques. However, for the portion inside the cluster, we devise an amortized bound on the integral (see Lemma 12), which is of independent interest, and is analogous to the improvement given by Davenport-Mahler bound over repeated applications of the root separation bound. It is this result that underlies the $O(n \log n)$ bound. A simple argument extends these bounds to Sturm's method and the `Eval` algorithm. The details are in Section 4.

2 Notation and Basic Results

Let $f \in \mathbb{R}[x]$ be a square-free polynomial of degree $n \geq 2$ and $Z(f) \subset \mathbb{C}$ be its set of roots. Given a finite pointset $S \subseteq \mathbb{R}^2$, let D_S be the disc $D(m_S, r_S)$ such that m_S is the centroid of the points in S, and r_S is the least radius such that all the points in S are contained in $D(m_S, r_S)$. Given a $\lambda \in \mathbb{R}_{>0}$, define $\lambda D_S := D(m_S, \lambda r_S)$. We borrow the following definition from [13]: A subset $\mathcal{C} \subseteq Z(f)$ of size at least two is called a (root) **cluster** if the only roots in $3D_{\mathcal{C}}$ are from $\mathcal{C}$. We treat individual roots as (trivial) clusters. In this paper, the non-real roots in $\mathcal{C}$ come in conjugate pairs. Therefore, the center of $D_{\mathcal{C}}$ will always be in $\mathbb{R}$. Define $R_{\mathcal{C}}$ as the distance from $m_{\mathcal{C}}$ to the nearest point in the set $Z(f) \setminus \mathcal{C}$. From the definition it follows that $Z(f)$ trivially forms a cluster and $R_{Z(f)} = \infty$. Given an interval I, let $m(I)$ denote its midpoint and $w(I)$ its width. We will often use the shorthand $I = [m(I) \pm w(I)/2]$, and for $\lambda > 0$, $\lambda I := [m(I) \pm \lambda w(I)/2]$. An **interval I contains a cluster** $\mathcal{C}$ if $\mathcal{C} \subseteq D(m(I), w(I)/2)$.

We use the following convenient notation in the subsequent definitions: for $x, y \in \mathbb{R}$, '$x \gg y$' if there is a constant $c \geq 1$ such that $x \geq cy$; similarly define $x \ll y$.

A **strongly-separated cluster (ssc)** is a cluster $\mathcal{C}$ for which $R_{\mathcal{C}}/r_{\mathcal{C}} \gg n^3$; the exact constant can be found in Corollary 7. For a ssc $\mathcal{C}$, define the following three quantities:

(i) The interval $I_{\mathcal{C}} := [m_{\mathcal{C}} \pm c \cdot k r_{\mathcal{C}}]$, for some constant $c \geq 1$.

(ii) The interval $\mathcal{I}_{\mathcal{C}} := \left\{x : |x - m_{\mathcal{C}}| \ll R_{\mathcal{C}}/n^2\right\}$.

(iii) The annulus $\mathcal{A}_{\mathcal{C}} := \mathcal{I}_{\mathcal{C}} \setminus I_{\mathcal{C}} = \left\{z \in \mathbb{C} : |\mathcal{C}| r_{\mathcal{C}} \ll |z - m_{\mathcal{C}}| \ll R_{\mathcal{C}}/n^2\right\}$.

The exact constants in these definitions are given in Lemma 6. See Figure 1 for an illustration of these concepts. If $\mathcal{C}$ is not a ssc, then we define $I_{\mathcal{C}} := [m_{\mathcal{C}} \pm r_{\mathcal{C}}]$ and $\mathcal{I}_{\mathcal{C}} := 2I_{\mathcal{C}}$. Note that for all clusters $\mathcal{C}$, $I_{\mathcal{C}} \subseteq \mathcal{I}_{\mathcal{C}}$. We will need the following result later in our analysis [13, Lemma 2.1]:

Proposition 1. *Given a root cluster $\mathcal{C}$ of f. There is a unique unordered tree $T_{\mathcal{C}}$ rooted at $\mathcal{C}$ whose set of nodes are the clusters contained in $\mathcal{C}$, and the parent-child relation is subset inclusion. Let T_f be the tree where the parent is the cluster $Z(f)$ of all roots.*

The result originally is stated for root clusters of $f \in \mathbb{C}[x]$. However, for $f \in \mathbb{R}[x]$ the clusters come in conjugate pairs, and by taking the union of such pairs the result still holds. The tree $T_{\mathcal{C}}$ is called the **cluster tree of** $\mathcal{C}$. The leaves of this tree are the roots in $\mathcal{C}$.

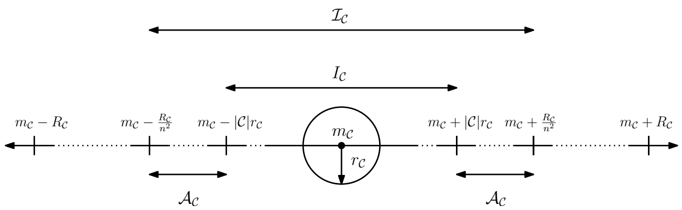

Figure 1: Geometry of a ssc $\mathcal{C}$. We focus on the the relative geometry, overlooking the exact constants involved in the definition of the intervals.

2.1 Cluster Detection and Approximation

The literature on detection and approximation of root clusters is vast (see [7] and the references therein). One approach is based on Pellet's test: if for a complex polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ there is an $r > 0$ such that $|a_k|r^k > \sum_{i \neq k} |a_i| r^i$ then the disc $D(0, r)$ contains exactly k roots of f. A point $z \in \mathbb{C}$ is said to satisfy Pellet's test, if there is a k and r for which the test holds with the coefficients of $f(x + z)$. Results in [17, 7] generalize Smale's α-theory and relate it to Pellet's test; an alternative derivation based on tropical algebra is given in [15]. We instead use a result by Ostrowski [8].

We need the following definitions. Let $f(x) = \sum_{i=0}^{n} a_i x^i$, where $a_i \in \mathbb{C}$. With each index i, $a_i \neq 0$, associate the point $P_i := (i, -\log|a_i|) \in \mathbb{R}^2$. The lower-hull of the convex-hull of these points is called the **Newton diagram** of f. Given an index $k \in \{0, \ldots, n\}$, let y_k be the point such that (k, y_k) is on the diagram. Define $\rho_k := e^{y_k - y_{k-1}}$, for $1 \leq k \leq n$, $\rho_{n+1} := \infty$, and the kth **deviation** $\Delta_k := \rho_{k+1}/\rho_k$, for $0 < k < n$.

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be the roots of f ordered such that $|\alpha_1| \leq |\alpha_2| \leq \cdots \leq |\alpha_n|$. Ostrowski showed the following fundamental relation between the absolute values of the roots and ρ_k's [8, p. 143]:

$$\frac{1}{2k} < \frac{|\alpha_k|}{\rho_k} < 2(n - k + 1). \tag{1}$$

Given $z \in \mathbb{C}$, we will be interested in the Newton diagram of $f(x + z)$. If $f_j(z) := f^{(j)}(z)/j!$, then from a result of Ostrowski [8, p. 128] we get:

$$\rho_k(z) = \max_{j<k} \left| \frac{f_j(z)}{f_k(z)} \right|^{\frac{1}{(k-j)}}, \text{ and } \rho_{k+1}(z) = \min_{j>k} \left| \frac{f_k(z)}{f_j(z)} \right|^{\frac{1}{(j-k)}}. \tag{2}$$

The RHS is defined for any k such that $f_k(z) \neq 0$; however, we are only interested in those k for which P_k is on the diagram. The kth deviation $\Delta_k(z) := \rho_{k+1}(z)/\rho_k(z)$. We have the following result for detecting clusters:

Lemma 2. *If $\Delta_k(z) \geq 27$, for some index $0 < k < n$, then there are exactly k roots in $D(z, 3\rho_k(z))$ and $D(z, \rho_{k+1}(z)/3)$. Moreover, as $\rho_{k+1}(z)/3 \geq 9\rho_k(z)$, these roots form a cluster.*

The proof shows that the inequality $\Delta_k(z) \geq 27$ implies that Pellet's test holds for $D(z, r)$, $3\rho_k(z) \leq r \leq D(z, \rho_{k+1}(z)/3)$ (see [7, Thm. 1.5]). Since the P_i's are sorted by x-coordinate, all the ρ_k's can be computed in $O(n)$ steps using, e.g., Graham's scan for convex hull computation.

Once we have detected a cluster $\mathcal{C}$ near z, we want a good approximation to $m_{\mathcal{C}}$. A standard way is to do the iteration $z_{i+1} = z_i - kf(z_i)/f'(z_i)$, starting from z, but this may not be numerically desirable, as both f and f' are small near $\mathcal{C}$. Another option is to use the standard Newton iteration applied to $f^{(k-1)}$. We show that if $\Delta_k(z) \geq 27$, then z is an approximate zero, in the sense of Smale et al. [1, p. 160, Thm. 2], to the root of $f^{(k-1)}$ in $D(z, \frac{3\rho_k(z)}{2k})$. Subsequently we show that if $\Delta_k(z) \geq c_0$, for some constant $c_0 \geq 27$, then for all z' in this disc $\Delta_k(z') \geq 27$, and hence there is a cluster of k roots in $D(z', 3\rho_k(z'))$. Moreover, the cluster is exactly $\mathcal{C}$. These results are summarized in the following:

Lemma 3. *Let $z \in \mathbb{C}$ be such that $\Delta_k(z) \geq c_0$, for some $k \geq 2$, $\mathcal{C}$ be the cluster in $D(z, 3\rho_k(z))$, and $D' := D(z, \frac{3\rho_k(z)}{2k})$. Then the following hold:*

(i) z is an approximate zero to the root z^* of $f^{(k-1)}$ in D' and the Newton iterates starting from z are in D'.

(ii) For all $z' \in D'$, $\Delta_k(z') \geq 27$, and $\mathcal{C}$ is the cluster in $D(z', 3\rho_k(z'))$.

(iii) If z, w are such that $\Delta_k(z), \Delta_k(w) \geq 27$ and $D(z, 3\rho_k(z))$, $D(w, 3\rho_k(w))$ intersect then the discs have the same cluster.

The proof is given in the appendix. We choose $c_0 := 27 \times 6e^6$. Given $z \in \mathbb{C}$, a value of k satisfying the condition $\Delta_k(z) \geq c_0$ is called an **admissible value** for z, with the corresponding **inclusion disc** $D(z, 3\rho_k(z))$. Note that there can be more than one admissible value for a point z corresponding to clusters of different sizes.

3 The Algorithm

Let C_0 and C_1 be some exclusion and inclusion predicate respectively. The following algorithm takes as input f and an interval I_0 and outputs a root partition of I_0.

`Newton-Isol`(f, I_0)
1 Initialize $\mathcal{P} \leftarrow \emptyset$, $\Phi \leftarrow \emptyset$; let Q be an empty queue.
1.a. If this is a recursive call then subdivide I_0 and push the two halves into Q; else $Q \leftarrow \{I_0\}$.
2. While Q is not empty do
    Remove an interval I from Q.
2.a.   If $C_0(I) \vee C_1(I)$ then add I to $\mathcal{P}$.
    else if `Newton-Incl-Exc`(I) is successful then
        Let (J, k) be the pair returned.
2.b.   If $\forall J' \in \Phi$, $J \cap J' = \emptyset$ and $J \cap I_0 \neq \emptyset$ then
2.c. $\forall\ I' \in Q$, $I' \leftarrow I' \setminus D(m(J), \frac{\rho_{k+1}(m(J))}{3})$.
2.d.   Add $J \cap I_0$ to Φ.
    else subdivide I and push the two halves into Q.
3. Return $\mathcal{P} \cup_{J \in \Phi}$ `Newton-Isol`(f, J).

The input to `Newton-Incl-Exc` is an interval $I = (a, b)$. If the predicate is successful then it returns an interval J containing a cluster such that $w(J) < w(I)/2$, and an admissible value k for $m(J)$; otherwise it returns failure.

`Newton-Incl-Exc` (f, I)

1. Let $m := (a+b)/2$.
2. For $p \in \{a, m, b\}$, let $k_p \geq 2$ be the *smallest* admissible value k for p such that $I \subseteq D\left(p, \frac{\rho_{k+1}(p)}{3}\right)$.
3. If the three admissible values are equal and the three inclusion discs are contained in $D\left(m, \frac{\rho_{k_m+1}(m)}{3}\right)$ then:

3.a. $\quad z_0 := m$, $k := k_m$, $g := f^{(k-1)}$, $i := 0$.

4. $\quad$ While $\rho_k(z_i) \leq 2^{5-2^i}\rho_k(z_0)$
   $\qquad z_{i+1} := z_i - g(z_i)/g'(z_i)$; $i := i+1$.
   $\quad J := [z_{i-1} \pm 3\rho_k(z_{i-1})]$
5. $\quad$ If $w(J) \geq w(I)/2$ then return failure
6. $\quad$ else return (J, k).
7. Return failure.

We first explain some steps in the predicate above:

Step 2. A point p in I can have more than one admissible value associated with it. The right admissible value is governed by $w(I)$, since we should only consider those clusters $\mathcal{C}$ for which $r_{\mathcal{C}} \ll w(I) \ll R_{\mathcal{C}}$.

Step 3. As $D(m, \rho_{k_m+1}(m)/3)$ contains all the three inclusion discs, they all contain the same cluster $\mathcal{C}$. Otherwise, it is possible that the three inclusion discs contain different clusters but of the same size.

Step 4. This ensures that as z_i converges to the root of $f^{(k-1)}$, the distance to $\mathcal{C}$ decreases quadratically; this fails when we are near $\mathcal{C}$, or the root of $f^{(k-1)}$ is not near $\mathcal{C}$.

Step 5. Required to ensure linear convergence to $\mathcal{C}$.

Step 6. The interval J contains the cluster $\mathcal{C}$. Moreover, as $I \subseteq D(m, \rho_{k+1}(m)/3)$, we know that if the roots in I are a subset of $\mathcal{C}$, and hence are inside J. By now $w(J) < w(I)/2$, therefore, it suffices to return J.

We now comment on some steps in `Newton-Isol`:

Step 1.a. Ensures that a successful call to `Newton-Incl-Exc` is followed by a subdivision step. Thus the recursion tree is a binary tree. The predicate can still be successful on an interval J returned by an earlier successful call. But the convergence in this case would only be linear, and so we prefer subdivision, though in practice one can go ahead with the linear convergence.

Step 2.b. Checks if $\mathcal{C}$ has not been found before (see Lemma 3(iii)), and that J is inside I_0; if either of this test fails, then I contains no roots and can be excluded.

Step 2.c. As the only cluster in $D(m(J), \frac{\rho_{k+1}(m(J))}{3})$ is $\mathcal{C}$, we can remove this disc from the intervals in Q. It is this exclusion step that significantly contributes to the improvement of the subdivision algorithm.

Step 2.d. This step adds the interval $J \cap I_0$ containing the newly discovered cluster $\mathcal{C}$ to the set Φ.

There are only two loops in the algorithm: first, the while-loop in step 2 of the algorithm, and second, the Newton iteration in step 4 of `Newton-Incl-Exc`. The argument for the termination of the first loop is the same as for `Isolate`. The termination of the second loop is guaranteed, because if z_i's are such that $\rho_k(z_i)$ keeps on decreasing, then in the limit ρ_k's converge to zero; but the disc $D(z_i, 3\rho_k(z_i))$ contains exactly k roots; since, in the limit z_i's tend to a root z^* of $f^{(k-1)}$, this implies that z^* is a k-fold root of f, which is a contradiction as f is square-free.

The following is a proof of correctness of the algorithm.

THEOREM 4. *Given a polynomial f and an interval I_0, `Newton-Isol`(f, I_0) outputs a root partition $\mathcal{P}$ of I_0.*

Proof. We need to show the following claims:

1. $I_0 \setminus \mathcal{P}$ contains no real roots of f.
2. $\mathcal{P}$ contains (interior) pairwise disjoint intervals.
3. For all $I \in \mathcal{P}$, C_0 or C_1 holds (follows from step 2.a.).

Lemma 3 gives us the correctness of `Newton-Incl-Exc`(I), i.e., if the test is successful then it returns an interval J such that any roots in I are contained in J. We only argue for the first claim. For every interval J returned by a successful call of the predicate, define

$$A_J := D\left(m(J), \rho_{k+1}(m(J))/3\right) \setminus D(m(J), w(J)/2), \tag{3}$$

i.e., the annulus around J that does not contain any roots. We exclude intervals if step (2.b) fails for the interval J, or a portion of an interval is removed in step (2.c.). In the former case, either the cluster contained in J was already detected, or it is outside I_0. In the latter case, we do not loose any roots since A_J has no roots. So $I_0 \setminus \mathcal{P}$ contains no roots. **Q.E.D.**

4 Complexity Analysis

The main result is that `Newton-Incl-Exc` will be successful near a ssc $\mathcal{C}$. Let $c_0 > 20$ be the constant in Lemma 3, and $\mathcal{C}$ a ssc throughout this section. Our first claim is that $|\mathcal{C}|$ is an admissible value for all points in $\mathcal{I}_\mathcal{C}$.

Lemma 5. *If* $|z - m_\mathcal{C}| \leq R_\mathcal{C}/(8c_0n^2)$ *then* $\Delta_k(z) \geq c_0$.

Proof. Let $\alpha_1, \ldots, \alpha_k \in \mathcal{C}$ and $\alpha_{k+1}, \ldots, \alpha_n \in Z(f) \setminus \mathcal{C}$. Moreover, assume that they are ordered in increasing distance from z. From (1), we know that $2k|z - \alpha_{k+1}| > \rho_{k+1}(z) > |z - \alpha_{k+1}|/(2(n - k + 1))$. Moreover, $|z - \alpha_{k+1}| > R_\mathcal{C} - |z - m_\mathcal{C}| \geq R_\mathcal{C}/2$; similarly, $|z - \alpha_{k+1}| < 3R_\mathcal{C}/2$. Therefore,

$$\frac{R_\mathcal{C}}{4n} \leq \rho_{k+1}(z) \leq 3|\mathcal{C}|R_\mathcal{C}. \tag{4}$$

From (1), we again have $\rho_k(z) < 2k|z - \alpha_k|$. But $|z - \alpha_k| \leq |z - m_\mathcal{C}| + r_\mathcal{C}$, which gives us

$$\rho_k(z) \leq 2k(|z - m_\mathcal{C}| + r_\mathcal{C}). \tag{5}$$

Since $|z - m_\mathcal{C}|, r_\mathcal{C} \leq R_\mathcal{C}/(8c_0n^2)$, we get $\rho_k(z) \leq kR_\mathcal{C}/(2c_0n^2)$. Combining this with (4), and the observation that $(n - k)k \leq n^2/4$, we obtain that $\Delta_k \geq 2c_0n^2/(8(n - k)k) \geq c_0$. **Q.E.D.**

Recall the definition of the intervals $I_\mathcal{C}$, $\mathcal{I}_\mathcal{C}$ and the annulus $\mathcal{A}_\mathcal{C}$ from Section 2, and A_J from (3).

Lemma 6. *If an interval* I *is such that*

$$I \subseteq \mathcal{I}_\mathcal{C} = [m_\mathcal{C} \pm R_\mathcal{C}/(8c_0n^2)] \text{ and } w(I) > 72|\mathcal{C}|r_\mathcal{C}$$

then the pair (J, k) *returned by* `Newton-Incl-Exc`(I) *is such that* $k = |\mathcal{C}|$, $J \subseteq I_\mathcal{C} = [m_\mathcal{C} \pm 20kr_\mathcal{C}]$, *and* $A_J \supseteq \mathcal{A}_\mathcal{C}$.

Proof. We show that the conditions on I above imply that `Newton-Incl-Exc`(I) reaches step 6 of `Newton-Incl-Exc` (all the steps below refer to the steps in the predicate). This requires showing the following: (i) all the conditions in step 3 are met; (ii) Newton-iteration in step 4 converges quadratically terminating with an interval J with $w(J) < w(I)/2$, and (iii) $J \subseteq I_\mathcal{C}$. The following claims provide the proof. Let $I = [a, b]$ and $m = m(I)$.

Claim 1: For all $p \in \{a, m, b\}$, $k_p = |\mathcal{C}|$. Recall from Step 2 that k_p is defined as the *smallest* admissible value k for which $I \subset D(p, \rho_{k+1}(p)/3)$. From Lemma 5, we have $k_p \leq |\mathcal{C}|$. Since the roots in I can only come from $\mathcal{C}$, any smaller admissible value corresponds to a subcluster $\mathcal{C}'$ of $\mathcal{C}$, which implies $R_{\mathcal{C}'} \leq r_\mathcal{C}$. From (4) we know that $\rho_{|\mathcal{C}'|+1}(p) \leq 3(|\mathcal{C}'| + 1)R_{\mathcal{C}'} \leq 3|\mathcal{C}|r_\mathcal{C}$. Since $w(I) \geq 72|\mathcal{C}|r_\mathcal{C}$, clearly $I \not\subseteq D(p, \rho_{|\mathcal{C}'|+1}(p)/3)$ for any subcluster $\mathcal{C}' \subset \mathcal{C}$. Thus $k_p \geq |\mathcal{C}|$.

Claim 2: For all $p \in I$, $I \subseteq D(p, \rho_{k+1}(p)/3)$. This will follow from the more general claim that

$$D_1 := D(m_\mathcal{C}, R_\mathcal{C}/(8c_0n^2)) \subseteq D(z, \rho_{|\mathcal{C}|+1}(z)/3) =: D_2,$$

for all $z \in D_1$; since $a, m, n \in I \subseteq D_1$, the claim holds. But for any $z \in D_1$, we know from (4) that $\frac{\rho_{|\mathcal{C}|+1}(z)}{3} \geq \frac{R_\mathcal{C}}{12n}$ which is greater than $\frac{R_\mathcal{C}}{4c_0n^2}$, the diameter of D_1, for $c_0 \geq 3$.

Claim 3: For all $z, w \in D_1$, the inclusion disc $D(z, 3\rho_k(z)) \subseteq D(w, \frac{\rho_{k+1}(w)}{3})$. This follows if

$$|z - w| + 3\rho_k(z) \leq \frac{\rho_{k+1}(w)}{3}. \tag{6}$$

But $|z-w|, r_{\mathcal{C}} \le R_{\mathcal{C}}/(8c_0n^2)$, which along with (5) implies that $3\rho_k(z) \le 6kR_{\mathcal{C}}/(4c_0n^2)$. Therefore, LHS of (6) is smaller than $13kR_{\mathcal{C}}/(8c_0n^2)$, which is smaller than $R_{\mathcal{C}}/(12n)$ for $c_0 \ge 20$, but from (4) we know that the latter is smaller than the RHS of (6).

Claim 4: Let z_i be the sequence of iterates computed in the while-loop in Step 4. If $z_i \in D(m_{\mathcal{C}}, \frac{R_{\mathcal{C}}}{8c_0n^2}) \setminus D(m_{\mathcal{C}}, 2r_{\mathcal{C}})$, then $\rho_k(z_i) < 2^{5-2^i}\rho_k(z_0)$. Since $z_i \notin D(m_{\mathcal{C}}, 2r_{\mathcal{C}})$, $r_{\mathcal{C}} \le |z_i - m_{\mathcal{C}}|$, and hence from (5) we obtain $\rho_k(z_i) \le 4k|z_i - m_{\mathcal{C}}|$. From [10, Thm. 2.2] we know that there is a unique root z^* of $f^{(k-1)}$ in $D(m_{\mathcal{C}}, r_{\mathcal{C}})$. Therefore, $|z_i - m_{\mathcal{C}}| \le |z_i - z^*| + r_{\mathcal{C}}$. But as $z_i \notin D(m_{\mathcal{C}}, 2r_{\mathcal{C}})$ and $z^* \in D(m_{\mathcal{C}}, r_{\mathcal{C}})$, we have $r_{\mathcal{C}} \le |z_i - z^*|$, and hence $|z_i - m_{\mathcal{C}}| \le 2|z_i - z^*|$. Thus, $\rho_k(z_i) \le 8k|z_i - z^*|$. As z_0 is an approximate zero to z^* (see Lemma 3(i)), we know $|z_i - z^*| \le 2^{1-2^i}|z_0 - z^*|$, which implies that $\rho_k(z_i) \le 2^{4-2^i}k|z_0 - z^*|$. Furthermore, from Lemma 3(i) we know $k|z_0 - z^*| < 2\rho_k(z_0)$. Hence $\rho_k(z_i) < 2^{5-2^i}\rho_k(z_0)$.

Claim 5: The interval $J \subseteq I_{\mathcal{C}}$ and $w(J) < w(I)/2$. The previous claim shows that if $z_i \notin D(m_{\mathcal{C}}, 2r_{\mathcal{C}})$, then we will obtain quadratically decreasing values of $\rho_k(z_i)$. Thus when the iteration stops $z_i \in D(m_{\mathcal{C}}, 2r_{\mathcal{C}})$, and it follows from (5) that $\rho_k(z_i) \le 6kr_{\mathcal{C}}$. Hence the interval $J = z_i \pm 3\rho_k(z_i)$ is contained in $I_{\mathcal{C}}$, for $k \ge 2$. Moreover, $w(J) \le 36kr_{\mathcal{C}} < w(I)/2$, and hence the condition in Step 5 fails and we return J. The claim on the annulus follows from (4).

Q.E.D.

The following result translates the result above in terms of the subdivision tree:

Corollary 7. *Let $\mathcal{C}$ be a ssc such that $\mathcal{I}_{\mathcal{C}} \subseteq I_0$. If I is the first interval such that* `Newton-Incl-Exc`(I) *is successful and the interval returned contains $\mathcal{C}$, then $\mathcal{I}_{\mathcal{C}} \subseteq I' \cup I''$, where I' is the parent-interval of I and I'' is one of I''s neighbors.*

Proof. In the worst case, $\mathcal{C}$ will be detected the first time in the subdivision tree an interval $I \subseteq \mathcal{I}_{\mathcal{C}}$. For such an I, we show $w(I) \gg kr_{\mathcal{C}}$. Since I is the first interval to fall in $\mathcal{I}_{\mathcal{C}}$, both I' and I'' have endpoints outside $\mathcal{I}_{\mathcal{C}}$, thus $\mathcal{I}_{\mathcal{C}} \subseteq I' \cup I''$. So $2w(I) \ge R_{\mathcal{C}}/(16c_0n^2) > 72kr_{\mathcal{C}}$, as $\mathcal{C}$ is ssc. The claim clearly holds if $\mathcal{C}$ is detected at an ancestor of I. **Q.E.D.**

Remark: The proof above gives us the explicit constant in the definition of ssc, namely, we require $R_{\mathcal{C}}/r_{\mathcal{C}} > 16c_0 \times 72n^3$. A careful working out of the proofs shows that the weaker inequality $R_{\mathcal{C}}/r_{\mathcal{C}} > 4c_0 \times 72(n - |\mathcal{C}|)|\mathcal{C}|^2$, (or even $50c_0n^3$) is sufficient.

Recall that the set of all roots $Z(f)$ is a cluster. As a consequence of Lemma 6, we assume that $I_0 \subseteq nI_{Z(f)}$; otherwise `Newton-Incl-Exc` will be successful right away and the interval returned will satisfy the property.

4.1 An Integral Bound on the Subdivision Tree

Let $\mathcal{N}(I_0)$ be the set of leaves in the subdivision tree of `Newton-Isol`(f, I_0). Step 1.a. of the algorithm ensures that the subdivision tree is a binary tree. Therefore, it suffices to bound $|\mathcal{N}(I_0)|$. For this purpose, we use the general framework of continuous amortization developed in [2] and generalized in [3]. The idea is to bound $|\mathcal{N}(I_0)|$ by an integral and then derive an upper bound on this integral. For this purpose we need the following notion: Given a choice of predicates C_0, C_1, a function $G : \mathbb{R} \to \mathbb{R}_{\ge 0}$ is called a **stopping function** corresponding to C_0 and C_1 if for every interval I, if there is an $x \in I$ such that $w(I)G(x) \le 1$, then either $C_0(I)$ or $C_1(I)$ holds. Stopping functions, corresponding to different predicates, are provided in [3]. The crucial property of $G(x)$ is the following:

Lemma 8. *If $C_0(I)$ and $C_1(I)$ fail for an interval I, then for all $J \subseteq I$, such that $2w(J) \ge w(I)$, $2\int_J G(x)dx \ge 1$.*

Proof. From the definition of $G(x)$, we have for all $x \in I$, $G(x)w(I) \ge 1$. As $J \subseteq I$, $\forall x \in J$, $2G(x)w(J) \ge G(x)w(I) \ge 1$. Thus $2\int_J G(x)dx \ge 2w(J)\min_{x \in J} G(x) \ge 1$. **Q.E.D.**

The main result of this section is the following:

Theorem 9.

$$|\mathcal{N}(I_0)| \le 4n + 2\int_{I_0 \setminus \cup_{\mathcal{C}} \mathcal{A}_{\mathcal{C}}} G(x)dx,$$

where the union is over all ssc $\mathcal{C}$ in T_f.

We bound $\mathcal{N}(I_0)$ recursively. The leaves in $\mathcal{N}(I_0)$ correspond to three types of intervals:

(i) intervals in the root partition $\mathcal{P}$,

(ii) intervals that were discarded in step 2.c., and

(iii) intervals for which condition 2.b fails to hold (either cluster already found, or $J \cap I_0 = \emptyset$).

We will bound each of these three types. We analyse what happens before the first set of recursive calls.

Let Φ be the set of intervals collected in Step 2.d. of the algorithm, A_J be as defined in (3), and $\mathcal{I}_J := J \cup A_J$. From the construction of Φ, we know that all intervals $J \in \Phi$ are contained in I_0 and each contains a unique cluster. For each $J \in \Phi$, let L_J be the set of *parent-intervals* of intervals in Q that intersect $\mathcal{I}_J$; the type (ii) intervals are children of intervals in L_J. Let M_J be the set of intervals that do not intersect $\mathcal{I}_J$ and are of type (iii). See Figure 2 for an illustration of these types. Note that if $I \in L_J$ contains an endpoint of $\mathcal{I}_J$, then $I \setminus \mathcal{I}_J$ can be of type (i) or (iii); but there can be at most two such intervals for each J on either side of $\mathcal{I}_J$. We abuse notation and use L_J to represent a set as well as the union of the intervals in it; same for M_J.

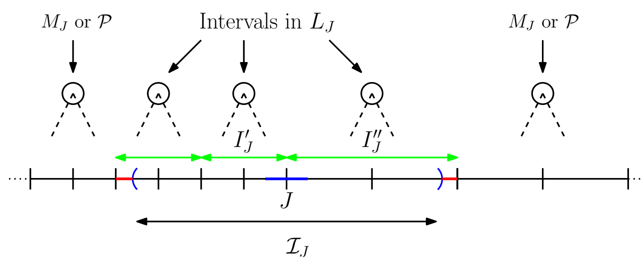

Figure 2: The three types of intervals in $\mathcal{N}(I_0)$. Intervals in L_J are shown in green. The remaining intervals could be in M_J or $\mathcal{P}$. The width of the red colored intervals can be much smaller than their parents. But there are at most two such intervals.

For an $I \in M_J$, both C_0 and C_1 failed. Therefore, from Lemma 8 we get $|M_J| \le 2\sum_{I \in M_J} \int_I G(x)dx = 2\int_{M_J} G(x)dx$. As the predicates C_0 and C_1 also fail for the intervals in L_J, we can similarly bound $|L_J|$. But this effectively amounts to doing subdivision on J. To avoid this we do the following: since the width of the intervals in L_J is more than $w(J)$, we know that there are at most two neighboring intervals I'_J and I''_J that contain J. We count them separately, and for the rest we use Lemma 8 to get $|L_J| \le 2 + 2\int_{L_J \setminus (I'_J \cup I''_J)} G(x)dx$. For an interval $I \in \mathcal{P}$, we expect $2\int_I G(x)dx \ge 1$, as the predicates must have failed for the parent I' of I. However, Lemma 8 requires that $w(I') \le 2w(I)$. This can fail to happen near the boundary of $\mathcal{I}_J$, as noted earlier. But then there are at most two such intervals. Therefore, the number of intervals in $\mathcal{P}$ coming from the non-recursive calls is at most $2|\Phi| + 2\int_{I_0 \setminus \cup_J (L_J \cup M_J)} G(x)dx$. Combining this with the bounds on $|L_J|$ and $|M_J|$ we get

$$|\mathcal{N}(I_0)| \le 4|\Phi| + 2\int_{I_0 \setminus \cup_J (I'_J \cup I''_J)} G(x)dx + \sum_{J \in \Phi} |\mathcal{N}(J)|. \tag{7}$$

To open the RHS recursively, we introduce the notion of **cluster tree T_{I_0} with respect to an interval** I_0: It is the smallest subtree $T_{\mathcal{C}}$ of T_f rooted at a cluster $\mathcal{C}$ such that $I_0 \subseteq I_{\mathcal{C}}$; since by assumption $I_0 \subseteq nI_{Z(f)}$, in the worst case, T_{I_0} is T_f. Moreover, as enlarging I_0 increases the integral in (7), we further make the simplifying assumption that $I_0 = 2I_{\mathcal{C}_0}$, where $\mathcal{C}_0$ is the root of T_{I_0}.

Let $\mathcal{C}$ be the cluster associated with a node u in T_{I_0}. Let $J_u \in \Phi$ be the interval returned the first time $\mathcal{C}$ is detected by `Newton-Incl-Exc` . Define $A_u := (I'_{J_u} \cup I''_{J_u}) \setminus J_u$; if $\mathcal{C}$ is not detected, let $A_u = J_u = \emptyset$. Using this notation, the following bound can be derived from (7) by induction:

$$|\mathcal{N}(I_0)| \le 4|T_{I_0}| + 2\int_{I_0 \setminus \bigcup_{u \in T_{I_0}} A_u} G(x)dx. \tag{8}$$

For a ssc $\mathcal{C} \in T_{I_0}$, the assumption $I_0 = 2I_{\mathcal{C}_0}$ ensures that $\mathcal{I}_{\mathcal{C}} \subseteq I_0$. So Corollary 7 implies that $I'_u \cup I''_u \supseteq \mathcal{I}_{\mathcal{C}}$, and Lemma 6 implies that $J_u \subseteq I_{\mathcal{C}}$; hence, $A_u \supseteq \mathcal{A}_{\mathcal{C}}$. Considering only the ssc in T_{I_0} on the RHS of (8) we obtain

$$|\mathcal{N}(I_0)| \leq 4|T_{I_0}| + 2\int_{I_0 \setminus \cup_{\mathcal{C}} \mathcal{A}_{\mathcal{C}}} G(x)dx, \tag{9}$$

As $|T_{I_0}| \leq n$, we get Theorem 9.

4.2 Bound for the Descartes's rule of signs

In this section, we derive the following bound:

THEOREM 10. *Given a square-free polynomial $f \in \mathbb{R}[x]$ of degree n, the size of the subdivision tree constructed by `Newton-Isol`(f, I_0) using predicates based on the Descartes's rule of signs is bounded by $O(n \ln n)$.*

We bound the RHS of (9), where the stopping function corresponds to the Descartes's rule of signs. We use the same stopping function as described in [3], but explain why the argument there fails to give us the bound above.

Let $V := Z(f)$, the set of roots of f. Define $d(x, V)$ as the distance from x to the closest point in V, and $d_2(x, V)$ as the distance to the second closest point in V. The crucial idea in [3] is to partition the integral over the (real) Voronoi interval I_α of each root α (for the moment suppose $\alpha \in \mathbb{R}$). Define $J_\alpha := [\alpha \pm \frac{d_2(\alpha,V)}{2}]$. Then for $x \in J_\alpha$, $G(x) := 2/d_2(x, V)$, and for $x \in I_\alpha \setminus J_\alpha$, $G(x) := 1/|x - \alpha|$. Break $\int_{I_\alpha} G(x)dx$ as $\int_{J_\alpha} G(x)dx + \int_{I_\alpha \setminus J_\alpha} G(x)dx$. In [3] it is shown that the first integral is $O(1)$, and the second integral is $O(\log w(I_\alpha)/d_2(\alpha, V))$; from Cauchy's bound we can assume that $w(I_\alpha) = 2^{O(L)}$. The problem is that in the worst case this ratio can be $\Omega(n(L + \log n))$. E.g., if all the other roots are of the form $\alpha \pm it$, for increasing values of t, then I_α is the x-axis. Therefore, $\int_{I_\alpha \setminus J_\alpha} G(x)dx = \Omega(L - \log d_2(\alpha, V))$; in the worst case $d_2(\alpha, V)$ can be the root separation bound.

Our idea is based on the observation that roots with very small separation give rise to root clusters. For clusters that are not ssc, the ratio $R_{\mathcal{C}}/r_{\mathcal{C}} = O(n^3)$, therefore, the number of subdivisions needed to bridge this gap is $O(\log n)$. For a ssc, the gap is bridged by `Newton-Incl-Exc` so that the subdivision is restricted to the ranges $R_{\mathcal{C}}$ to roughly $R_{\mathcal{C}}/n^2$ and $|\mathcal{C}|r_{\mathcal{C}}$ to $r_{\mathcal{C}}$, both of which take $O(\log n)$ subdivisions. Doing this for all clusters basically gives the bound in Theorem 10.

Let $P \subseteq \mathbb{C}$ be a pointset such that any non-real point in P also has its complex conjugate in P. Such a set of points is called **dense** if no proper subset of P forms a non-trivial cluster, i.e., for all $S \subset P$, such that $|S| \geq 2$, the disc $3D_S$ contains a point from $P \setminus S$. This structure plays a fundamental role in our arguments, as do the following two integrals (see [3, 16]):

LEMMA 11. *Let $\gamma \in \mathbb{C}$ and $J = [r, s]$.*
(Re) *If $\gamma \in \mathbb{R} \setminus J$, then*

$$\int_J \frac{dx}{|\gamma - x|} = \ln\left|\frac{\gamma - s}{\gamma - r}\right|^{\delta(J>\gamma)}, \tag{10}$$

where $\delta(J > \gamma) = +1$ if $r > \gamma$ and -1 if $s < \gamma$.
(Im) *If $\gamma \in \mathbb{C} \setminus \mathbb{R}$ then*

$$\int_J \frac{dx}{|\gamma - x|} \leq O\left(\ln \frac{\max\{|s - \gamma|, |r - \gamma|\}}{|\mathtt{Im}(\gamma)|}\right). \tag{11}$$

We now give the proof of Theorem 10.
Proof. The proof is by induction on $|T_{I_0}|$. We claim that

$$\int_{I_0 \setminus \cup_{\mathcal{C}} \mathcal{A}_{\mathcal{C}}} G(x)dx = O(|T_{I_0}| \ln n). \tag{12}$$

Let $\mathcal{C}_0$ be the root of T_{I_0}; by assumption we have $I_0 = 2I_{\mathcal{C}_0}$. Let $\mathcal{M}_0$ be the children of $\mathcal{C}_0$ in T_{I_0}. Consider a ssc $\mathcal{C} \in \mathcal{M}_0$. Then $I_0 \setminus \mathcal{A}_{\mathcal{C}} \subseteq (I_0 \setminus \mathcal{I}_{\mathcal{C}}) \cup 2I_{\mathcal{C}}$. If $\mathcal{C}'$ is a ssc contained in $\mathcal{C}$, we can inductively remove $\mathcal{A}_{\mathcal{C}'}$ from $I_{\mathcal{C}}$. This also works for clusters that are not ssc in $\mathcal{M}_0$, since by definition $\mathcal{I}_{\mathcal{C}} = 2I_{\mathcal{C}}$. Therefore,

$$I_0 \setminus \cup_{\mathcal{C}} \mathcal{A}_{\mathcal{C}} \subseteq (I_0 \setminus \cup_{\mathcal{C} \in \mathcal{M}_0} \mathcal{I}_{\mathcal{C}}) \cup \left(\cup_{\mathcal{C} \in \mathcal{M}_0} (2I_{\mathcal{C}} \setminus \cup_{\mathcal{C}' \subset \mathcal{C}} \mathcal{A}_{\mathcal{C}'})\right).$$

We claim that

$$\int_{I_0 \setminus \cup_{\mathcal{C} \in \mathcal{M}_0} \mathcal{I}_\mathcal{C}} G(x)dx = O(|\mathcal{M}_0| \ln n). \tag{13}$$

As $|T_\mathcal{C}| < |T_{I_0}|$, for $\mathcal{C} \in \mathcal{M}_0$, by induction we obtain

$$\int_{2I_\mathcal{C} \setminus \cup_{\mathcal{C}' \subset \mathcal{C}} \mathcal{A}_{\mathcal{C}'}} G(x)dx = O(|T_\mathcal{C}| \ln n).$$

This bound along with (13) and the observation that $|\mathcal{M}_0| + \sum_{\mathcal{C} \in \mathcal{M}_0} |T_\mathcal{C}| < |T_{I_0}|$ gives us (12). The base case is when $\mathcal{M}_0$ contains only leaves, in which case (12) reduces to (13).

We next claim that

$$\int_{I_0 \setminus \cup_{\mathcal{C} \in \mathcal{M}_0} \mathcal{I}_\mathcal{C}} G(x)dx = O(\ln n) + \int_{I_0' \setminus \cup_{\mathcal{C} \in \mathcal{M}_0} \mathcal{I}_\mathcal{C}} G(x)dx,$$

where $I_0' := [m_{\mathcal{C}_0} \pm 2r_{\mathcal{C}_0}]$. If $\mathcal{C}_0$ is not a ssc, then this is clear as $I_0' = 2I_{\mathcal{C}_0} = I_0$. If $\mathcal{C}_0$ is a ssc, then $I_0 = [m_{\mathcal{C}_0} \pm |\mathcal{C}_0| r_{\mathcal{C}_0}]$. Break I_0 as I_0', $[m_{\mathcal{C}_0} + 2r_{\mathcal{C}_0}, m_{\mathcal{C}_0} + |\mathcal{C}_0| r_{\mathcal{C}_0}]$ and $[m_{\mathcal{C}_0} - 2r_{\mathcal{C}_0}, m_{\mathcal{C}_0} - |\mathcal{C}_0| r_{\mathcal{C}_0}]$. The closest root to any x in these intervals is from $\mathcal{C}_0$. Moreover, as $|x - m_{\mathcal{C}_0}| \geq 2r_{\mathcal{C}_0}$, we get $G(x) := 1/d(x,V) \leq 2/|x - m_{\mathcal{C}_0}|$. Therefore, from Lemma 11(Re) it follows that $\int_{m_{\mathcal{C}_0}+2r_{\mathcal{C}_0}}^{m_{\mathcal{C}_0}+|\mathcal{C}_0||r_{\mathcal{C}_0}} \frac{2}{|x - m_{\mathcal{C}_0}|} = O(\ln |\mathcal{C}_0|)$. Similarly for the other interval. Hence to prove (13), it suffices to show

$$\int_{I_0' \setminus \cup_{\mathcal{C} \in \mathcal{M}_0} \mathcal{I}_\mathcal{C}} G(x)dx = O(|\mathcal{M}_0| \ln n). \tag{14}$$

Replace each $\mathcal{C} \in \mathcal{M}_0$ by its center $m_\mathcal{C}$. Let $\mathcal{M}_0$ also denote this pointset. We will use Lemma 12 to prove (14). As no subset of $\mathcal{M}_0$ forms a cluster, $\mathcal{M}_0$ is a dense pointset, and Lemma 12 is applicable. However, we first remove some region around every $p \in \mathcal{M}_0 \cap \mathbb{R}$ to be able to invoke Lemma 12. For each such p, define $J_p := [p \pm d_2(p, \mathcal{M}_0)/2]$. If $p = m_\mathcal{C}$, for $\mathcal{C} \in \mathcal{M}_0$, then $\mathcal{I}_p := \mathcal{I}_\mathcal{C} \subseteq J_p$. We claim

$$\int_{\cup_{p \in \mathcal{M}_0} (J_p \setminus \mathcal{I}_p)} G(x)dx = O(|\mathcal{M}_0| \ln n). \tag{15}$$

From Lemma 12 we get

$$\int_{I_0' \setminus (\cup_{p \in \mathcal{M}_0} J_p)} G(x)dx = O(|\mathcal{M}_0| \ln n).$$

Combining these two bounds, along with the observation that the union of the sets $I_0' \setminus (\cup_{p \in \mathcal{M}_0} J_p)$ and $\cup_{p \in \mathcal{M}_0} (J_p \setminus \mathcal{I}_p)$ is the set $I_0' \setminus \cup_{p \in \mathcal{M}_0} \mathcal{I}_p$, completes the proof of (14).

To prove (15), we show that $\int_{J_p \setminus \mathcal{I}_p} G(x)dx = O(\ln n)$, and then sum over all $p \in \mathcal{M}_0$. There are three cases to consider:

(i) $p = m_\mathcal{C}$ for some normal cluster $\mathcal{C} \in \mathcal{M}_0$. Then $J_p = [m_\mathcal{C} \pm R_\mathcal{C}/2]$ and $\mathcal{I}_p = \mathcal{I}_\mathcal{C} = [m_\mathcal{C} \pm 2r_\mathcal{C}]$. Therefore, $J_p \setminus \mathcal{I}_p$ contains $[m_\mathcal{C} + 2r_\mathcal{C}, m_\mathcal{C} + R_\mathcal{C}/2]$ and $[m_\mathcal{C} - R_\mathcal{C}/2, m_\mathcal{C} - 2r_\mathcal{C}]$. The nearest root to any x in these two intervals is from $\mathcal{C}$. Since x is outside $2I_\mathcal{C}$, it follows that $d(x,V) \geq |x - m_\mathcal{C}|/2$. Therefore, $G(x) := \frac{1}{d(x,V)} \leq 2/|x - m_\mathcal{C}|$. From Lemma 11(Re), we obtain $\int_{m_\mathcal{C}+2r_\mathcal{C}}^{m_\mathcal{C}+R_\mathcal{C}/2} G(x)dx = O(\ln R_\mathcal{C}/r_\mathcal{C})$. Since $\mathcal{C}$ is not a ssc, $R_\mathcal{C}/r_\mathcal{C} = O(n^3)$, which gives us the desired bound. The same applies to the other interval.

(ii) Suppose $p = m_\mathcal{C}$, where $\mathcal{C} \in \mathcal{M}_0$ is a ssc. Then $J_p = [m_\mathcal{C} \pm R_\mathcal{C}/2]$ and $\mathcal{I}_p = \mathcal{I}_\mathcal{C} = [m_\mathcal{C} \pm R_\mathcal{C}/n^2]$. Let $I_\mathcal{C}' := [m_\mathcal{C} + \frac{R_\mathcal{C}}{n^2}, m_\mathcal{C} + \frac{R_\mathcal{C}}{2}]$ be one of the intervals $J_p \setminus \mathcal{I}_p$. The nearest root to any $x \in I_\mathcal{C}'$ is from $\mathcal{C}$. Since $x \notin 2I_\mathcal{C}$, it follows that $d(x,V) \geq |x - m_\mathcal{C}|/2$. Therefore, $G(x) := \frac{1}{d(x,V)} \leq 2/|x - m_\mathcal{C}|$. Applying Lemma 11(Re), we get $\int_{I_\mathcal{C}'} G(x)dx \leq 4 \ln n$. Similarly, for the other interval.

(iii) p is a real root then $\mathcal{I}_p = \emptyset$. For $x \in J_p$, our stopping function $G(x) = 2/d_2(x,P)$, i.e., corresponding to the inclusion predicate. Suppose $q \in P$ is such that $d_2(p,P) = |p - q|$. Then for all $x \in J_p$, $d_2(x,P) \geq |p - q| - |p - x| \geq \sigma_p/2$, and hence $\int_{J_p} \frac{2dx}{d_2(x,P)} \leq 4 \int_{p-\sigma_p/2}^{p+\sigma_p/2} \frac{dx}{\sigma_p} = O(1)$.

Q.E.D.

The proof above can carried out with the exact constants involved in the definitions of $I_{\mathcal{C}}$, $\mathcal{I}_{\mathcal{C}}$ and $\mathcal{A}_{\mathcal{C}}$ (see Lemma 6), but they will be absorbed by the big-O notation. Note that the $O(n \ln n)$ bounds the number of calls to the C_0 predicate. The specialization of $G(x)$ for C_0 is $1/d(x, V)$. The corresponding specialization for Sturm sequences is $1/d(x, V \cap \mathbb{R}) \leq 1/d(x, V)$. Therefore, $O(n \ln n)$ holds for `Newton-Isol` combined with Sturm sequences. For `Eval`, one specialization of the stopping function for the C_0 predicate is $n/d(x, V)$, which immediately gives an $O(n^2 \ln n)$ bound for `Newton-Isol` combined with `Eval`. Whether it can be improved using the more precise specialization $\sum_{\alpha \in V} \frac{1}{|x-\alpha|}$ remains open.

Let P be a dense pointset P. Given a point $p \in P$, define $\sigma_p := \min |p - q|$, where $q \in P \setminus \{p\}$, and $J_p := [p \pm \sigma_p/2]$. We want to bound $\int_{(2D_P \cap \mathbb{R}) \setminus \cup_p J_p} dx/d(x, P)$. We first show an $O(|P|^2)$ bound, essentially following [3]. Let $\mathcal{V}_p$ be the set of points in $2D_P \cap \mathbb{R}$ closer to p than to any other point in P. It is clear that $J_p \subseteq \mathcal{V}_p$. The intervals $\mathcal{V}_p$ partition $2D_P \cap \mathbb{R}$. Then $\int_{\mathcal{V}_p \setminus J_p} dx/d(x, P)$ can be shown to be bounded by $O(\ln(r(D_p)/\sigma_p))$. Using the density of P, it can be shown that if p, q are such that $\sigma_p = |p - q|$ then $P \subseteq 3^{O(|P|)} D_{\{p,q\}}$, which implies that $r(D_P) \leq 3^{O(|P|)} \sigma_p$, for all $p \in P$. This gives an $O(|P|^2)$ bound instead of the bound in Theorem 10. To obtain that we need to amortize the integral carefully. The intuition is that if σ_p is very small then there there must a lot of other points close to p, and hence the width of $\mathcal{V}_p$ cannot be very large compared to σ_p . The challenge is to get an "almost cluster-like" decomposition of P. We construct a tree on P that gives us this decomposition.

We describe an iterative bottom-up procedure to construct a tree $\mathcal{T}_P$ with leaves from P. Let $\sigma := \min_{p \in P} \sigma_p$. For all points $p \in P$, draw a disc of radius $\sigma/2$ centered at p. As σ is the smallest distance between any pair of points, two such discs can at most touch each other. The discs touching each other form a connected component. The collection of the largest connected components partitions P (leaves are considered as components). Moreover, there is at least one component $G \subseteq P$ that has cardinality strictly greater than one; the components with cardinality one are the leaves. For each such component G, we introduce an internal node u in $\mathcal{T}_P$ with children as the leaves p, where $p \in G$; let $G_u := G$, the associated component, and $\sigma_u := \sigma$. Now redefine σ as the minimum separation between the components constructed so far, draw a disc of radius $\sigma/2$ centered at each $p \in P$, and continue as above. Let $\mathcal{T}_P$ be the tree constructed in this bottom-up manner; see Figure 3. Further define the following quantities for each $u \in \mathcal{T}_P$:

(i) ν_u as the number of children of u,
(ii) m_u be the center and r_u be the radius of $D(G_u)$.

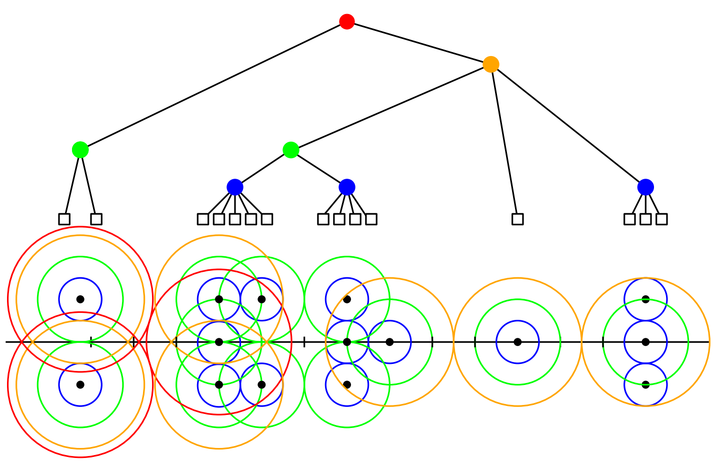

Figure 3: A dense pointset P and construction of $\mathcal{T}_P$. Circles of different colors correspond to different σ's. The first choice of σ corresponds to blue colored circles, followed by green, orange and red. The components formed are shown in the corresponding colors. We only draw some of the relevant circles to give an idea.

Let $u, v \in \mathcal{T}_P$ be such that v is a child of u. We have the following properties of $\mathcal{T}_P$:

(P1) $\sigma_u \leq \min_{p \in G_v, q \in P \setminus G_v} |p - q| \leq 3r_v$. The upper bound follows from the density of P. The lower bound follows from the observation that the discs with radius $\sigma/2$ centered at $p \in G_v$, where $\sigma_v < \sigma < \sigma_u$, do not touch the discs of any other component, except when the radius is $\sigma_u/2$.

(P2) $r_u \le |G_u|\sigma_u$. Consider the graph $\mathcal{G}$ with the vertices as G_u and edges between two vertices p, q if $D(p, \sigma_u/2) \cap D(q, \sigma_u/2) \neq \emptyset$. As G_u is a connected component of these discs, we know that $\mathcal{G}$ is connected. Therefore, if m is the number of vertices on the path joining p, q in $\mathcal{G}$, then by triangular inequality $|p-q| \le m\sigma_u \le |G_u|\sigma_u$.

(P3) If p is a leaf-child of u then $\sigma_u = \sigma_p$. It is clear that any disc $D(p, r)$, with $r < \sigma_p/2$, cannot touch $D(q, r)$, for any other point q. The first time they touch is when $\sigma_u = \sigma_p$. If $p \in \mathbb{C} \setminus \mathbb{R}$, then we further obtain that $|\texttt{Im}(p)| \ge \sigma_p \ge \sigma_u$.

(P4) The size of $\mathcal{T}_P = O(|P|)$. Every level has a node with more than one child, as there are pairs of components with separation exactly σ.

(P5) P is the component associated with the root of $\mathcal{T}_P$.

The next result is an amortization analogous to that of the Davenport-Mahler bound over the root separation bound.

Lemma 12. *If P is a dense pointset then*

$$\int_{(2D_P \cap \mathbb{R}) \setminus \cup_{p \in P} J_p} \frac{dx}{d(x, P)} = O(|P| \ln |P|), \tag{16}$$

where for $p \in P \cap \mathbb{R}$, $J_p := [p \pm \sigma_p/2]$, and $J_p = \emptyset$ otherwise.

Proof. We break the integral recursively over the nodes of $\mathcal{T}_P$. For an internal node u of $\mathcal{T}_P$, we will show the following claim:

$$\int_{(2D(G_u) \cap \mathbb{R}) \setminus \cup_{p \in G_u} J_p} \frac{dx}{d(x, P)} = O(\nu_u \ln |G_u|). \tag{17}$$

We take the sum over all internal nodes u. From (P4) we know that $|\mathcal{T}_P| = O(|P|)$, and hence $\sum_u \nu_u = O(|P|)$; moreover, from (P5) we know that the component associated with the root of $\mathcal{T}_P$ is P. These observations then give us (16).

For a point $p \in P$, recall that $\mathcal{V}_p$ is the set of points in $2D_P \cap \mathbb{R}$ closer to p than to any other point of P; by definition $J_p \subseteq \mathcal{V}_p$. Suppose u is the parent of p. We will bound the integral over $\mathcal{V}_p$ in two steps: the portion of $\mathcal{V}_p$ inside $2D(G_u)$ and the portion outside $2D(G_u)$. The latter portion is where amortization occurs, as for an $x \notin 2D(G_u)$, the distance of x to $p \in G_u$ is roughly $|x - m_u|$. Let v be a child of u. There are three cases to consider:

Case 1. v is a leaf $p \in \mathbb{R}$. We first bound the portion I_p of $\mathcal{V}_p$ inside $2D(G_u)$; the portion outside will be handled collectively for all points in the third case. For all $x \in I_p \setminus J_p$, it is clear that $d(x, P) = |x - p|$. From Lemma 11(Re) we obtain that $\int_{I_p \setminus J_p} \frac{dx}{|x-p|} = O\left(\ln \frac{w(I_p)}{\sigma_p}\right)$. But as $I_p \subseteq 2D(G_u) \cap \mathbb{R}$, we know that $w(I_p) \le 4r_u$. From (P3), we know that $\sigma_p = \sigma_u$. Therefore, $\int_{I_p \setminus J_p} \frac{dx}{|x-p|} = O(\ln r_u/\sigma_u) = O(\ln |G_u|)$, from (P2).

Case 2. v is a leaf $p = \texttt{Re}(p) + i\texttt{Im}(p) \in \mathbb{C} \setminus \mathbb{R}$. Again consider the interval $I_p := \mathcal{V}_p \cap 2D(G_u)$; in this case $J_p = \emptyset$. As p is the closest point to any $x \in I_p$, $d(x, P) = |x - p|$. Moreover, p and both the endpoints of I_p are in $2D(G_u)$, so the maximum distance of an endpoint of I_p from p is $\le 2r_u$. Therefore, from Lemma 11(Im) we have

$$\int_{I_p} \frac{dx}{d(x, P)} = O\left(\ln \frac{r_u}{|\texttt{Im}(p)|}\right).$$

But recall from (P3) that $|\texttt{Im}(p)| \ge \sigma_u$, hence $r_u/|\texttt{Im}(p)| \le r_u/\sigma_u \le |G_u|$, where the last inequality follows from (P2). Therefore, $\int_{I_p} \frac{dx}{d(x,P)} = O(\ln |G_u|)$.

Case 3. v is an internal node. Inductively, we have already bounded the integral $\int_{(2D(G_v) \cap \mathbb{R}) \setminus \cup_{p \in G_v} J_p} dx/d(x, P)$. However, it is possible that $\mathcal{V}_p$, for some point $p \in G_v$ extends beyond $2D(G_v) \cap \mathbb{R}$. Suppose p is such a point, and $x \in W_p := \mathcal{V}_p \cap (2D(G_u) \setminus 2D(G_v))$. Then we know that $|x - p| \ge |x - m_v|/2$, where m_v is the center of $D(G_v)$. Therefore,

$$\sum_{p \in G_v} \int_{W_p} \frac{dx}{|x - p|} \le \int_{(2D(G_u) \setminus 2D(G_v)) \cap \mathbb{R}} \frac{2dx}{|x - m_v|}.$$

As $2w(I_u) = 4r_u$, from Lemma 11(Re), it follows that the integral on the RHS is bounded by $O(\ln r_u/r_v)$. But from (P2) we have $r_u \leq |G_u|\sigma_u$, and $\sigma_u \leq 3r_v$ from (P1). Therefore, we obtain

$$\sum_{p \in G_v} \int_{W_p} \frac{dx}{d(x,P)} = O(|\ln |G_u||).$$

This is the case where the amortization of the integral over the Voronoi regions takes place.

Summing the bounds for all children v of u gives (17). **Q.E.D.**

The following is the analogue of Lemma 12 in $\mathbb{C}$: define $D_p := D(p, \sigma_p/2)$, then

$$\int_{2D_P \setminus \cup_p D_p} \frac{dz}{d(z,P)} = O(|P| \ln |P|).$$

4.3 Bound for Newton+EVAL

In this section, we derive an upper bound on the integral in the RHS of (9) where the stopping function corresponds to the centered form interval arithmetic based predicates used in the EVAL algorithm [3, 16]. More precisely, the stopping function in this case is $G(x) := 1.5 \min\{S_0(x), S_1(x)\}$, where

$$S_0(x) := \sum_{\alpha \in Z(f)} \frac{1}{|x-\alpha|}, \text{ and } S_1(x) := \sum_{\alpha' \in Z(f')} \frac{1}{|x-\alpha'|}. \tag{18}$$

The idea is similar to what was done in Section 4.2, to charge S_0 on the region between clusters, and S_1 on the roots.

5 Concluding Remarks

Our aim has been to devise a general approach to improve any subdivision based algorithm for real root isolation. This is achieved by the `Newton-Incl-Exc` predicate, which detects strongly separated clusters, and hence reduces the number of subdivisions from $O(\log R_\mathcal{C}/r_\mathcal{C})$ to $O(n \log n)$. The crucial ingredient is Ostrowski's criterion based on deviations of the Newton diagram of a polynomial. The criterion works for complex polynomials, so we expect an analogue of `Newton-Isol` for isolating complex roots that is conceptually simpler than the existing approaches. We have not explored the practical aspects of the algorithm, nevertheless, we think that the analysis based on the geometry of cluster provides tools and techniques for an alternate approach to understand existing algorithms.

We can bound the arithmetic complexity of `Newton-Isol` as follows. The Newton diagram computation takes $O(n)$, and the Taylor shift $O(n \log n)$ operations. The number of Newton iterations to approximate $\mathcal{C}$ is bounded by $O(\log\log \frac{R_\mathcal{C}}{r_\mathcal{C}})$, which is $O(\log(nL))$ from root separation bounds. Therefore, the arithmetic complexity, ignoring poly-log factors, is bounded by $\widetilde{O}(n^2)$. The extension to the bitstream model involves deriving a robust version of Ostrowski's result and bounding precision requirements. The latter will be governed by perturbation bounds for clusters. For a cluster of size k, we expect these bounds to be $O(\epsilon^{1/k})$, for ϵ-perturbation in the coefficients. In the worst case, this would give an $O(n(L + \log n))$ bound on the precision.

References

[1] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation.* Springer, 1998.

[2] M. Burr, F. Krahmer, and C. Yap. Continuous amortization: A non-probabilistic adaptive analysis technique. *Elec. Colloq. Comp. Compl.*, 16:136, 2009.

[3] M. Burr. Applications of continuous amortization to bisection-based root isolation. *CoRR*, abs/1309.5991 (2013).

[4] G. E. Collins and A. G. Akritas. Polynomial real root isolation using Descartes' rule of signs. In R. D. Jenks, editor, *Proc. 1976 ACM Symp. Symb. and Alg. Comp.*, pp. 272–275. 1976.

[5] J. H. Davenport. Computer algebra for cylindrical algebraic decomposition. Tech. Report 88-10 , School of Math. Sci., U. of Bath, Bath, England.

[6] A. Eigenwillig, V. Sharma, and C. Yap. Almost tight complexity bounds for the Descartes method. In *Proc. 31st ISSAC*, pp. 71–78, 2006.

[7] M. Giusti, G. Lecerf, B. Salvy, and J.-C. Yakoubsohn. On location & approximation of clusters of zeros of analytic functns. *Fnd. Comp. Math.*, 5(3):257–311, '05.

[8] A. Ostrowski. Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent. *Acta Mathematica*, 72:99–155, 1940.

[9] V. Y. Pan. Approximating complex polynomial zeros: Modified Weyl's quadtree construction and improved Newton's iteration. *J. Complexity*, 16(1):213–264, '00.

[10] P. Pawlowski. The location of the zeros of the higher order derivatives of a polynomial. *Proc. AMS*, 127(5):1493–1497, 1999.

[11] M. Sagraloff. When Newton meets Descartes: a simple and fast algorithm to isolate the real roots of a polynomial. In *Proc. Intl. Symp. Symb. & Algb. Comp (ISSAC)*, pp. 297–304, 2012.

[12] M. Sagraloff and K. Mehlhorn. Computing real roots of real polynomials - an efficient method based on Descartes' rule of signs and Newton iteration. *CoRR*, abs/1308.4088 (2013).

[13] M. Sagraloff, V. Sharma, and C. Yap. Analytic root clustering: A complete algorithm using soft zero tests. In P. Bonizzoni *et al.*, eds., *In Proc. CiE 2013.*, vol. 7921 of *LNCS*, pp. 434–444. Springer, 2013.

[14] T. Sasaki and A. Terui. Computing clustered close-roots of univariate polynomials. In *Symb. Num. Comp. (SNC)*, pp. 177–184. ACM, 2009.

[15] M. Sharify. *Scaling Algorithms and Tropical Methods in Numerical Matrix Analysis.* PhD thesis, Ecole Polytechnique Saclay, 2011.

[16] V. Sharma and C.-K. Yap. Near optimal tree size bounds on a simple real root isolation algorithm. In *Proc. 37th ISSAC*, pp. 319–326. ACM, 2012.

[17] J.-C. Yakoubsohn. Simultaneous computation of all the zero-clusters of a univariate polynomial. In F. Cucker and J. M. Rojas, eds., *Proc. of the Smalefest, Found. of Comp. Math.*, pp. 433–455, 2000.

Appendix

We give the proof of Lemma 3; the arguments are based standard manipulations with Taylor series in alpha-theory of Smale et al. We first prove Lemma 3(i), for which we need the following functions from [1] defined for $f^{(k-1)}$:

$$\beta_k(z) = \left|\frac{f^{(k-1)}(z)}{f^{(k)}(z)}\right|, \ \gamma_k(z) = \max_{j\geq 1}\left|\frac{f^{(k+j)}(z)}{(j+1)!f^{(k)}(z)}\right|^{\frac{1}{j}} \tag{19}$$

and $\alpha_k(z) = \beta_k(z)\gamma_k(z)$. We derive relations between these quantities and $\rho_k(z)$'s given in (2). Considering the RHS of $\rho_k(z)$ for $j = k-1$, we immediately have

$$\rho_k(z) \geq \left|\frac{f_{k-1}(z)}{f_k(z)}\right| = k\beta_k(z). \tag{20}$$

Multiplying and dividing the inner term on the RHS of $\gamma_k(z)$ in (19) by $(k+j)!/k!$ we obtain that

$$\begin{aligned}\gamma_k(z) &\leq \max_{j\geq 1}\left(\frac{(k+j)!}{k!(j+1)!}\right)^{1/j} \max_{j>k}\left|\frac{f_j(z)}{f_k(z)}\right|^{1/(j-k)} \\ &= \max_{j\geq 1}\left(\frac{(k+j)!}{k!(j+1)!}\right)^{1/j} \frac{1}{\rho_{k+1}(z)}.\end{aligned}$$

The max-term is bounded by $(k+1)$, which implies that

$$\gamma_k(z)\rho_{k+1}(z) \leq (k+1). \tag{21}$$

Multiplying (20) and (21) we obtain that $\alpha_k(z)\Delta_k(z) \leq 2$. Therefore, if $\Delta_k(z) \geq 12$ then z is an approximate zero of $f^{(k-1)}$ with associated root in $D(z, 1.5\beta_k(z)) \subseteq D(z, \frac{3\rho_k(z)}{2k})$, where the inclusion follows from (20). The claim on Newton iterates follows from [1, p. 160,Thm.2].

We now prove Lemma 3(ii). We will need the following result [1, p. 161, Lem. 3]: for a $u \in [0,1)$

$$\sum_{j\geq 0}\binom{k+j}{j}u^j = \frac{1}{(1-u)^{k+1}}. \tag{22}$$

Let $\epsilon := 1.5$, $\delta := \epsilon\rho_k/k$, and $u := \frac{\delta}{\rho_{k+1}} = \frac{\epsilon}{k\Delta_k}$; here we express $\rho_k(z)$ by ρ_k (similarly, for the other quantities).

Lemma 13. *If z is such that $\Delta_k(z) \geq 16$ then for $z' \in D(z,\delta)$, we have $|f_k(z')| \geq |f_k(z)|(1-u)^{-(k+1)}/2$.*

Proof. Take the absolute values in the Taylor expansion of $f^{(k)}(z')$ and apply the triangular inequality to obtain

$$|f_k(z')| \geq \frac{1}{k!}\left(|f^{(k)}(z)| - \sum_{j\geq 1}\left|\frac{f^{(k+j)}(z)}{j!}\right|\delta^j\right).$$

Dividing both sides by $|f_k(z)|$, and multiplying and dividing the summation term on the RHS by $k!$ and $(k+j)!$ we obtain that

$$\left|\frac{f_k(z')}{f_k(z)}\right| \geq \left(1 - \sum_{j\geq 1}\binom{k+j}{j}\left|\frac{f_{k+j}(z)}{f_k(z)}\right|\delta^j\right).$$

From the expression of $\rho_{k+1}(z)$ in (2) and definition of u, we deduce that

$$\left|\frac{f_k(z')}{f_k(z)}\right| \geq \left(2 - \sum_{j\geq 0}\binom{k+j}{j}u^j\right).$$

Using (22), and the bound on Δ_k, the RHS can be simplified to $(1-u)^{-(k+1)}/2$. **Q.E.D.**

Lemma 14. *If $\Delta_k(z) \geq 16$ then for all $z' \in D(z,\delta)$ we have $\rho_k(z') < 2e^6\rho_k(z)$ and $\rho_{k+1}(z') \geq \rho_{k+1}(z)/3$. Therefore, $\Delta_k(z') \geq \Delta_k(z)/(6e^6)$.*

Proof. For $j < k$, take absolute values in the Taylor expansion of $f_j(z')$, apply triangular inequality, and split the summation up to k and beyond k, to get

$$|f_j(z')| = |f_j(z)| + \sum_{\ell=1}^{k-j} \binom{\ell+j}{\ell} |f_{\ell+j}(z)|\,\delta^\ell + \sum_{\ell>k-j} \binom{\ell+j}{\ell} |f_{\ell+j}(z)|\,\delta^\ell.$$

Divide by $|f_k(z)|$ and use the expressions in (2) to obtain

$$\left|\frac{f_j(z')}{f_k(z)}\right| \leq \rho_k^{k-j} + \sum_{\ell=1}^{k-j} \binom{\ell+j}{\ell} \rho_k^{k-\ell-j}\delta^\ell + \sum_{\ell>k-j} \binom{\ell+j}{\ell} \frac{\delta^\ell}{\rho_{k+1}^{\ell+j-k}}.$$

Since $\delta = \epsilon\rho_k/k$, we can pull out ρ_k^{k-j} from the RHS (and since $\Delta_k > 1$), we get that

$$\left|\frac{f_j(z')}{f_k(z)}\right| \leq \rho_k^{k-j} \left(1 + \sum_{\ell=1}^{k-j} \binom{\ell+j}{\ell} \left(\frac{\epsilon}{k}\right)^\ell + \sum_{\ell>k-j} \binom{\ell+j}{\ell} \left(\frac{\epsilon}{k}\right)^\ell\right).$$

Assuming $k \geq 2$, from (22) we obtain that

$$\left|\frac{f_j(z')}{f_k(z)}\right| \leq \rho_k^{k-j} \left(1 - \frac{\epsilon}{k}\right)^{-(j+1)}.$$

Combining this bound with Lemma 13, and doing some further simplifications we obtain the upper bound on $\rho_k(z')$. Note that we require $k \geq 2 > \epsilon$. To derive a lower bound on $\rho_{k+1}(z')$ in terms of $\rho_{k+1}(z)$, we take absolute values in the Taylor expansion of $f_j(z')$, for $j > k$, apply the triangular inequality, and divide both sides by $|f_k(z)|$, to get

$$\left|\frac{f_j(z')}{f_k(z)}\right| \leq \left|\frac{f_j(z)}{f_k(z)}\right| + \sum_{\ell\geq 1} \binom{\ell+j}{\ell} \left|\frac{f_{\ell+j}(z)}{f_k(z)}\right| \delta^\ell.$$

From the expression for ρ_{k+1} in (2) and (22) it follows that

$$\left|\frac{f_j(z')}{f_k(z)}\right| \leq \frac{1}{\rho_{k+1}^{j-k}} (1-u)^{-(j+1)}.$$

Combining this with the lower bound in Lemma 13, and using the lower bound on Δ_k we further obtain that

$$\left|\frac{f_j(z')}{f_k(z')}\right|^{1/(j-k)} \leq \frac{2^{j-k}}{(1-u)\rho_{k+1}}.$$

Since $u \leq \epsilon/(k\Delta_k)$ and $\Delta_k \geq 16$, we get that

$$\left|\frac{f_j(z')}{f_k(z')}\right|^{1/(j-k)} \leq \frac{3}{\rho_{k+1}},$$

which implies the desired lower bound on $\rho_{k+1}(z')$. **Q.E.D.**

To show Lemma 3(iii), we suppose that $\rho_k(w) \leq \rho_k(z)$. As the two inclusion discs intersect it follows that

$$|w \pm 3\rho_k(w) - z| \leq |z - w| + 3\rho_k(w) \leq 9\rho_k(z) \leq \frac{\rho_{k+1}(z)}{3},$$

where the last inequality follows from $\Delta_k(z) \geq 27$. This implies that $D(w, 3\rho_k(w)) \subseteq D(z, \frac{\rho_{k+1}(z)}{3})$, and hence both the discs have the same cluster.