Almost Tight Recursion Tree Bounds for the Descartes Method

Arno Eigenwillig
Max-Planck-Institut für Informatik
Saarbrücken, Germany
arno@mpi-inf.mpg.de

Vikram Sharma
Dept. of Computer Science, NYU
New York, USA
sharma@cs.nyu.edu

Chee K. Yap*
Dept. of Computer Science, NYU
New York, USA
yap@cs.nyu.edu

ABSTRACT

We give a unified ("basis free") framework for the Descartes method for real root isolation of square-free real polynomials. This framework encompasses the usual Descartes' rule of sign method for polynomials in the power basis as well as its analog in the Bernstein basis. We then give a new bound on the size of the recursion tree in the Descartes method for polynomials with real coefficients. Applied to polynomials $A(X) = \sum_{i=0}^{n} a_i X^i$ with integer coefficients $|a_i| < 2^L$, this yields a bound of $O(n(L + \log n))$ on the size of recursion trees. We show that this bound is tight for $L = \Omega(\log n)$, and we use it to derive the best known bit complexity bound for the integer case.

Categories and Subject Descriptors

F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems—*Computations on polynomials*; G.1.5 [**Numerical Analysis**]: Roots of Nonlinear Equations—*Methods for polynomials*

General Terms

Algorithms, Theory.

Keywords

Polynomial real root isolation, Descartes method, Descartes rule of signs, Bernstein basis, Davenport-Mahler bound.

1. INTRODUCTION

Let $A(X)$ be a polynomial of degree $n > 1$ with real coefficients. A fundamental task in computer algebra is **real root isolation**, that is, to assign an enclosing interval to each real root of $A(X)$ such that distinct roots are assigned disjoint intervals. We assume that $A(X)$ is **square free** in this paper.

The classic approach to real root isolation starts from an open interval I_0 containing all real roots of $A(X)$ and bisects it recursively as follows: Given an interval J, test for the number $\#(J)$ of real roots in it. If $\#(J) = 0$ is known, stop. If $\#(J) = 1$ is known, report J as an isolating interval and stop. Otherwise, subdivide $J = (c, d)$ at its midpoint $m = (c + d)/2$; report $[m, m]$ if $f(m) = 0$; recur on (c, m) and (m, d).

To carry out this approach, we need a method for estimating the number of roots in an interval. The two choices here are **Sturm sequences** (e.g., [27, chap. 7]) that give an exact count of distinct real roots in an interval, and **Descartes' rule of signs** (e.g., Proposition 2.1 below) that counts real roots with multiplicity and may overestimate this number by an even positive integer. Despite the apparent inferiority of Descartes' rule as compared to Sturm sequences, there is considerable recent interest in the Descartes approach because of its excellent performance in practice [9, 24, 19, 25].

This paper shows that the asymptotic worst case bound on recursion tree size for the Descartes method (Theorem 3.4) is no worse than the best known bound for Sturm's method (Theorem 6 of [6]). For the particular case of polynomials with integer coefficients of magnitude less than L, the recursion tree is $O(n(L + \log n))$ both for Sturm's method [5, 6] and the Descartes method (Corollary 3.5); and the work at each node of this tree can be done with $\widetilde{O}(n^3 L)$ bit operations (using asymptotically fast basic operations), where $\widetilde{O}$ indicates that we are omitting logarithmic factors (see [23, 14, 6] or Theorem 4.2, respectively).

The connection between root isolation in the power basis using the Descartes method, and in the Bernstein basis using de Casteljau's algorithm and the variation-diminishing property of Bézier curves was already pointed out by Lane and Riesenfeld [13], but this connection is often unclear in the literature. In Section 2, we provide a general framework for viewing both as a form of the Descartes method. In Section 3, we present the main result, which is a new upper bound on the size of the recursion tree in the Descartes method. Up to that point, our analysis holds for all square-free polynomials with real coefficients. We then restrict to the case of integer polynomials with L-bit coefficients to show that this new bound on tree size is optimal under the assumption $L = \Omega(\log n)$ (Section 3.3) and allows a straightforward derivation of the best known bit complexity bound (Section 4).

*This author's work is supported in part by NSF Grant #CCF-0430836.

1.1 Previous work

Root isolation using Descartes' rule of signs was cast into its modern form by Collins and Akritas [3], using a representation of polynomials in the usual power basis. Rouillier and Zimmermann [25] summarize various improvements of this method until 2004.

The algorithm's equivalent formulation using the Bernstein basis was first described by Lane and Riesenfeld [13] and more recently by Mourrain, Rouillier and Roy [19] and Mourrain, Vrahatis and Yakoubsohn [20]; see also [1, §10.2].

The crucial tool for our bound on the size of the recursion tree is Davenport's generalization [5] of Mahler's bound [15] on root separation. Davenport used his bound for an analysis of Sturm's method (see [6]). He mentioned a relation to the Descartes method but did not work it out. This has been done later by Johnson [9] and, filling a gap in Johnson's argument, by Krandick [11]. However, they bound the number of internal nodes at each level of the recursion tree separately. This leads to bounds that imply[1] a tree size of $O(n \log n\,(\log n + L))$ and a bit complexity of $O(n^5(\log n + L)^2)$ for a polynomial of degree n with L-bit integer coefficients. Their argument uses a termination criterion for the Descartes method due to Collins and Johnson [4].

Krandick and Mehlhorn [12] employ a theorem by Ostrowski [21] that yields a sharper termination criterion. However, they just use it to improve on the constants of the bounds in [11][2]. We will show that Ostrowski's result allows an immediate bound on the number of *all* internal nodes of the recursion tree. This bound is better by a factor of $\log n$ and leads to the same bit complexity bound in a simpler fashion.

2. THE DESCARTES METHOD

2.1 A Basis-free Framework

The Descartes method is based on the following theorem about sign variations. A **sign variation** in a sequence $(a_0, \ldots, a_n)$ of real numbers is a pair $i < j$ of indices such that $a_i a_j < 0$ and $a_{i+1} = \cdots = a_{j-1} = 0$. The number of sign variations in a sequence $(a_0, \ldots, a_n)$ is denoted $\mathrm{Var}(a_0, \ldots, a_n)$.

PROPOSITION 2.1. **[Descartes' rule of signs]** *Let $A(X) = \sum_{i=0}^{n} a_i X^i$ be a polynomial with real coefficients that has exactly p positive real roots, counted with multiplicities. Let $v = \mathrm{Var}(a_0, \ldots, a_n)$ be the number of sign variations in its coefficient sequence. Then $v \geq p$, and $v - p$ is even.*

See [12] for a proof with careful historic references. Already Jacobi [8, IV] made the "little observation" that this extends to estimating the number of real roots of a real polynomial $A(X)$ of degree n over an arbitrary open interval (c, d) by applying Descartes' rule to $(X+1)^n A((cX+d)/(X+1)) = \sum_{i=0}^{n} a_i^* X^i$, because the Möbius transformation $X \mapsto (cX+d)/(X+1)$ puts $(0, \infty)$ in one-to-one correspondence to (c, d). So we define $\mathit{DescartesTest}(A, (c, d)) := \mathrm{Var}(a_0^*, \ldots, a_n^*)$. Since $v - p$ is non-negative and even, the Descartes test yields the exact number of roots whenever its result is 0 or 1.

The Descartes method for isolating the real roots of an input polynomial $A_{\mathrm{in}}(X)$ in an open interval J consists of a recursive procedure $\mathit{Descartes}(A, J)$ operating on a polynomial $A(X)$ and an interval J where the roots of $A(X)$ in $(0, 1)$ correspond to the roots of $A_{\mathrm{in}}(X)$ in J as follows:

(*) There is a constant $\lambda \neq 0$ and an affine transformation $\phi\colon \mathbb{R} \to \mathbb{R}$ such that $J = \phi((0, 1))$ and $\lambda A = A_{\mathrm{in}} \circ \phi$.

To isolate all the roots of $A_{\mathrm{in}}(X)$, we choose an interval $I_0 = (-B_1, +B_2)$ enclosing all real roots of A_{in} (see, e.g., [27, §6.2]). The recursion begins with $\mathit{Descartes}(A, I_0)$, where $A(X) := A_{\mathrm{in}}((B_1 + B_2)X - B_1)$; thus initially the roots of $A(X)$ in $(0, 1)$ correspond to the real roots of $A_{\mathrm{in}}(X)$ in I_0 via the affine transformation $\phi(X) = (B_1 + B_2)X - B_1$. The procedure goes as follows:

procedure $\mathit{Descartes}(A, (c, d))$
  {Assert: Invariant (*) holds with $J = (c, d)$.}
  $v := \mathit{DescartesTest}(A, (0, 1))$;
  if $v = 0$ **then return**; **fi**;
  if $v = 1$ **then** report (c, d); **return**; **fi**;
  $m := (c + d)/2$;
  $(A_L, A_R) := (H(A), TH(A))$;
  if $A_R(0) = 0$ **then** report $[m, m]$; **fi**;
  $\mathit{Descartes}(A_L, (c, m))$; $\mathit{Descartes}(A_R, (m, d))$;
  return;

The polynomials A_L and A_R are defined using the homothetic transformation $H(A)(X) := 2^n A(X/2)$ and the translation transformation $T(A)(X) := A(X + 1)$. For later use, we also introduce the reversal transformation $R(A)(X) := X^n A(1/X)$.

Note that in the initial invocation of $\mathit{Descartes}(A, (c, d))$, one has $\mathit{DescartesTest}(A, (0, 1)) = \mathit{DescartesTest}(A_{\mathrm{in}}, (c, d))$. In its recursive calls, one has $\mathit{DescartesTest}(A_L, (0, 1)) = \mathit{DescartesTest}(A_{\mathrm{in}}, (c, m))$ and $\mathit{DescartesTest}(A_R, (0, 1)) = \mathit{DescartesTest}(A_{\mathrm{in}}, (m, d))$, and so on.

The above description of $\mathit{Descartes}()$ does not refer to any basis in the vector space of polynomials of degree at most n. However, an implementation needs to represent polynomials by coefficients with respect to some specific basis.

The classical choice of basis for $\mathit{Descartes}()$ is the usual power basis $(1, X, X^2, \ldots, X^n)$. The transformations H, T and R are carried out literally. $\mathit{DescartesTest}(A, (0, 1))$ consists in counting the number of sign changes in the coefficient sequence of $TR(A)$. The test whether $A_R(0) = 0$ amounts to inspection of the constant term. We call the resulting algorithm the *power basis variant* of the Descartes method.

An alternative choice of basis is the $[0, 1]$-Bernstein basis

$$(B_0^n(X), B_1^n(X), \ldots, B_n^n(X)),$$

with $B_i^n(X) := B_i^n[0, 1](X)$ where

$$B_i^n[c, d](X) := \binom{n}{i} \frac{(X - c)^i (d - X)^{n-i}}{(d - c)^n}, \quad 0 \leq i \leq n.$$

Its usefulness for the Descartes method lies in the following: Since

$$TR(B_i^n)(X) = \binom{n}{i} X^{n-i}, \tag{1}$$

[1] Personal communication, Krandick and Mehlhorn.

[2] This potential use of Ostrowski's result is mentioned but not carried out in the 1999 Ph.D. thesis of P. Batra [2].

for $A(X) = \sum_{i=0}^{n} b_i B_i^n(X)$ one has that

$$DescartesTest(A,(0,1)) = \mathrm{Var}(b_0,\ldots,b_n),$$

without any additional transformation.

To obtain A_L and A_R from $A(X) = \sum_{i=0}^{n} b_i B_i^n(X)$, we use a fraction-free variant of de Casteljau's algorithm [22]: For $0 \le i \le n$ set $b_{0,i} := b_i$. For $1 \le j \le n$ and $0 \le i \le n-j$ set $b_{j,i} := b_{j-1,i} + b_{j-1,i+1}$. From this, one obtains coefficients of $2^n A(X) = \sum_{i=0}^{n} b'_i B_i^n[0,\frac{1}{2}](X) = \sum_{i=0}^{n} b''_i B_i^n[\frac{1}{2},1](X)$ by setting $b'_i := 2^{n-i} b_{i,0}$ and $b''_i := 2^i b_{n-i,i}$. Since

$$\begin{aligned} H(2^{-n} B_i^n[0,\tfrac{1}{2}])(X) &= B_i^n[0,1] \\ TH(2^{-n} B_i^n[\tfrac{1}{2},1])(X) &= B_i^n[0,1], \end{aligned}$$

one has $A_L(X) = H(A)(X) = \sum_{i=0}^{n} b'_i B_i^n(X)$ and $A_R(X) = TH(A)(X) = \sum_{i=0}^{n} b''_i B_i^n(X)$. Finally, the test whether $A_R(0) = 0$ amounts to inspection of b''_0, since $B_i^n(0) = 0$ for $i > 0$. We call the resulting algorithm the *Bernstein basis variant* of the Descartes method.

For consistency with the power basis variant, we have described the Bernstein basis variant as passing transformed polynomials A_L and A_R expressed in a globally fixed basis $(B_n^i[0,1])_i$ in recursive calls. Equivalently, one can think of it as passing (a constant multiple of) the same polynomial all the time, but converting it to the Bernstein basis w.r.t. the interval under consideration.

Both variants of the Descartes method as presented above work for polynomials with arbitrary real coefficients. However, if the initial coefficients are integers, then integrality is preserved. If this is not needed, one can leave out the factor 2^n in the definition of $H(A)$ and, for the Bernstein basis variant, apply the ordinary instead of the fraction-free de Casteljau algorithm.

2.2 Termination

Since the Descartes test only gives an upper bound on the number of real roots in an interval, an extra argument is needed that each path in the recursion tree of the Descartes method eventually reaches an interval for which it counts 0 or 1 and thus terminates. We use a result from Krandick and Mehlhorn [12] based on a theorem by Ostrowski [21].

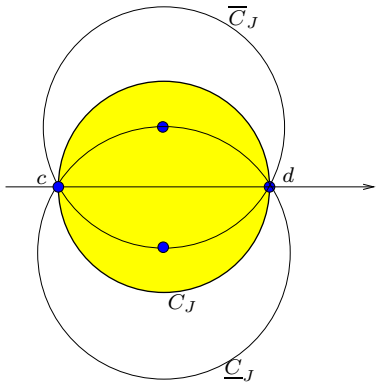

Figure 1: Three circles associated with the interval $J = (c,d)$.

Consider a real polynomial $A(X)$ and its roots in the complex plane. Let $J = (c,d)$ be an open interval with midpoint $m = (c+d)/2$ and **width** $w(J) = d - c$, and let $v = DescartesTest(A,J)$.

Proposition 2.2. **[One-Circle Theorem]** *If the open disc bounded by the circle C_J centered at m passing through the endpoints of J does not contain any root of $A(X)$, then $v = 0$.*

Proposition 2.3. **[Two-Circle Theorem]** *If the union of the open discs bounded by the circles $\underline{C}_J$ and $\overline{C}_J$ centered at $m \pm i(\sqrt{3}/6)w(J)$ and passing through the endpoints of J contains precisely one simple root of $A(X)$ (which is then necessarily a real root), then $v = 1$.*

See [12] for proofs. The circles $\underline{C}_J$ and $\overline{C}_J$ are characterized by being the circumcircles of the two equilateral triangles that have J as one of their edges. In the sequel, we call the union of discs bounded by $\underline{C}_J$ and $\overline{C}_J$ (as defined above in Proposition 2.3) the **two-circles figure** around interval J. Notice that the two-circles figure contains the disc bounded by C_J.

3. THE SIZE OF THE RECURSION TREE

3.1 The Davenport-Mahler Bound

The Davenport-Mahler theorem gives a lower bound on the product of differences of certain pairs of roots of a polynomial $A(X) = a_n \prod_{i=1}^{n}(X-\alpha_i)$ in terms of its **discriminant** $\mathrm{discr}(A) = a_n^{2n-2} \prod_{1\le i<j\le n}(\alpha_i-\alpha_j)^2$ and **Mahler measure** $\mathrm{M}(A) = |a_n| \prod_{i=1}^{n} \max\{1,|\alpha_i|\}$, see [27, §6.6, §4.5] [18, §1.5, §2.1]. This theorem appears in the literature in several variants that all use the same proof but formulate different conditions on how roots may be paired so that the proof works. We give the most general condition supported by the proof. It is equivalent to Johnson's formulation [9] and generalizes Davenport's original formulation [5, Prop. I.5.8].

Theorem 3.1. *Let $A(X) = a_n \prod_{i=1}^{n}(X-\alpha_i)$ be a square-free complex polynomial of degree n. Let $G = (V,E)$ be a directed graph whose nodes $\{v_1,\ldots,v_k\}$ are a subset of the roots of $A(X)$ such that*

(i) If $(v_i,v_j) \in E$ then $|v_i| \le |v_j|$.
(ii) G is acyclic.
(iii) The in-degree of any node is at most 1.

If exactly m of the nodes have in-degree 1, then

$$\prod_{(v_i,v_j)\in E} |v_i - v_j| \ge \sqrt{|\mathrm{discr}(A)|} \cdot \mathrm{M}(A)^{-(n-1)} \cdot (n/\sqrt{3})^{-m} \cdot n^{-n/2}.$$

Proof. This proof is not self-contained, but refers to the standard argument from Davenport [5, 27]. Let $(v_1,\ldots,v_k)$ be the topologically sorted list of the vertices of G, where $(v_i,v_j) \in E$ implies $j < i$. Given such an ordering we modify the $n \times n$ Vandermonde matrix $W_A = (\alpha_i^{j-1})_{j,i}$ as follows: For $j = 1$ to k in turn, we process v_j. If there exists an $i > j$ such that $(v_i,v_j) \in E$ then in W_A we subtract the column of v_i from the column of v_j; if no such i exists then the column of v_j remains unchanged. This finally yields a transformed matrix M such that $\det W_A = \det M$. Note that exactly m columns of M are modified from W_A. Moreover, $\det M = \prod_{(v_i,v_j)\in E}(v_j - v_i) \cdot \det M'$, where M' is a matrix similar to the one in [27, Theorem 6.28, Eqn. (19)]. As in the proof in

[27], we conclude:

$$|\det(W_A)| \le \left(\prod_{(v_i,v_j)\in E} |v_i - v_j| \right) \cdot \mathrm{M}(A)^{(n-1)} \left(\frac{n}{\sqrt{3}} \right)^m n^{n/2}.$$

But $\sqrt{|\mathrm{discr}(A)|} = |\det W_A|$, thus giving us the desired result. □

REMARK. The bound in Theorem 3.1 is invariant under replacing $A(X)$ by a non-zero scalar multiple $\lambda A(X)$.

REMARK. A bound similar to Theorem 3.1 appears in [17]. Instead of $\mathrm{M}(A)^{n-1}$, it uses a product of root magnitudes with varying exponents of $n-1$ or less.

3.2 The Recursion Tree

Our application of the Davenport-Mahler theorem rests on the following lemma. It reflects an important structural advantage of Proposition 2.3 over the weaker two-circle theorem by Collins and Johnson [4]: An intersection of the two-circles figures of two non-overlapping intervals can only occur if the intervals are adjacent, even if they reside on very different levels of the recursion tree.

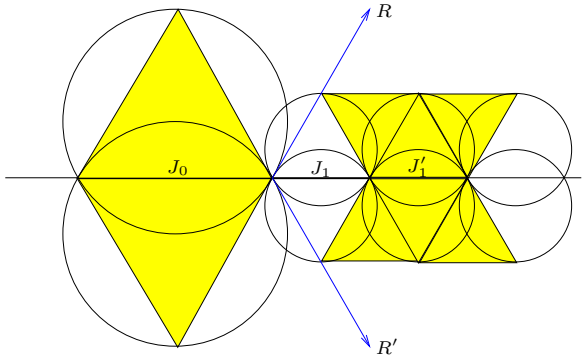

Figure 2: The two-circles figure around J_0 can overlap with that of J_1 but not with any two-circles figure further right.

LEMMA 3.2. *Let J_0 and J_1 be any two open intervals appearing in the recursive subdivision of some initial interval I_0. If the two-circles figures of Proposition 2.3 around J_0 and J_1 intersect, then J_0 and J_1 overlap or have a common endpoint.*

PROOF. We show that non-overlapping intervals with intersecting two-circles figures have a common endpoint. Let us choose indices such that $w(J_0) \ge w(J_1)$. Assume J_0 lies to the left of J_1 (the opposite case is symmetric). All intervals right of J_0 that have width $w(J_1)$ and appear in the recursive subdivision of I_0 have distance $k \cdot w(J_1)$ from J_0 for a non-negative integer k. They are depicted in Figure 2. The interval with $k = 0$ has a two-circles figure intersecting the two-circles figure of J_0. For $k > 0$, we claim that the two-circles figure of J_0 is disjoint from the two-circles figure of J_1. To see this, consider the convex cone delimited by the two tangent rays (R, R') of the two-circles figure of J_0 at its right endpoint. The two-circles figure of J_0 lies outside that cone, but if $k > 0$, then the two-circles figure of J_1 lies inside the cone. Figure 2 illustrates this for the case $k = 1$: the corresponding interval is J_1', and the two-circles figure of J_1' is covered by six equilateral triangles. Since the rays R, R' meet the x-axis at $60°$, this shows that the six equilateral triangles lie within the cone. Hence there is no intersection. □

The recursion tree T of the Descartes method in Section 2 is a binary tree. With each node $u \in T$ we can associate an interval I_u; the root is associated with I_0. A leaf u of T is said to be of **type-i** if the open interval I_u contains exactly i real roots; the termination condition of the algorithm implies i is either 0 or 1.

Our aim is to bound the number of nodes in T, denoted by $\#(T)$. We next introduce a subtree T' of T by pruning certain leaves from T:

- If a leaf u has a sibling that is a non-leaf, we prune u.
- If u, v are both leaves and siblings of each other, then we prune exactly one of them; the choice to prune can be arbitrary except that we prefer to prune a type-0 leaf over a type-1.

Clearly, $\#(T) < 2\#(T')$; hence it is enough to bound $\#(T')$. Let U be the set of leaves in T'. Then the number of nodes along the path from any $u \in U$ to the root of T' is exactly $\log \frac{w(I_0)}{w(I_u)}$. Thus

$$\#(T') \le \sum_{u\in U} \log \frac{w(I_0)}{w(I_u)}. \tag{2}$$

Our next goal is to reduce this bound to the Davenport-Mahler type bound shown in Theorem 3.1.

Two cases.
Let u be a leaf of T', and v be its parent. We will define two roots α_u, β_u such that the number of nodes along the path from u to the root is

$$O\left(\log \frac{w(I_0)}{|\alpha_u - \beta_u|} \right).$$

Furthermore, we will show that if u, u' are two leaves of the same type (both type-0 or both type-1), then $\{\alpha_u, \beta_u\}$ and $\{\alpha_{u'}, \beta_{u'}\}$ are disjoint.

In the following arguments, we will overload the notation C_I, $\overline{C}_I$ and $\underline{C}_I$ to represent the three open discs that have one of the circles as their boundary.

1. If u is type-1 then its interval I_u contains a real root α. Consider its parent v. By Proposition 2.3, $\overline{C}_{I_v} \cup \underline{C}_{I_v}$ must contain a root apart from α_u; let β_u be any root in this region. Then it follows that

$$|\alpha_u - \beta_u| < \frac{2}{\sqrt{3}} w(I_v) = \frac{4}{\sqrt{3}} w(I_u). \tag{3}$$

Thus the number of nodes in the path from u to the root of T' is

$$\log \frac{w(I_0)}{w(I_u)} < \log \frac{4w(I_0)}{\sqrt{3}|\alpha_u - \beta_u|}. \tag{4}$$

Let u' be another type-1 leaf different from u. Clearly, $\alpha_u \ne \alpha_{u'}$. We claim that β_u and $\beta_{u'}$ can be chosen such that $\beta_u \ne \beta_{u'}$. From Lemma 3.2 it is clear that we only need to consider the case when I_v and $I_{v'}$ are adjacent to each other. Moreover, assume β_u and $\overline{\beta_u}$ are the only non-real roots in $\overline{C}_{I_v} \cup \underline{C}_{I_v}$ and $\overline{C}_{I_{v'}} \cup \underline{C}_{I_{v'}}$. Then it must be that either $\beta_u \in \overline{C}_{I_v} \cap \overline{C}_{I_{v'}}$ or $\beta_u \in$

$\underline{C}_{I_v} \cap \underline{C}_{I_{v'}}$. In either case we can choose $\beta_{u'} = \overline{\beta_u}$ distinct from β_u.

2. If u is type-0, it had a type-0 sibling that was pruned. Consider their parent node v and let I_v be the interval associated with it. There are two cases to consider:
   - I_v does not contain a real root. Thus Proposition 2.2 implies that C_{I_v} must contain some non-real root α_u and its conjugate $\beta_u := \overline{\alpha_u}$. Moreover,
   $$|\alpha_u - \beta_u| \leq w(I_v) = 2w(I_u). \tag{5}$$
   - The midpoint of I_v is a real root, say α. Since the sign variations for I_v is greater than one, there is a pair of non-real roots $(\beta, \overline{\beta})$ in $\overline{C}_{I_v} \cup \underline{C}_{I_v}$. If $\beta \in C_{I_v}$ then let $\alpha_u := \beta$ and $\beta_u := \overline{\beta}$; otherwise, let $\alpha_u = \alpha$ and $\beta_u = \beta$. It can be verified that (5) still holds.

   Hence the number of nodes on the path from u to root of T' is
   $$\log \frac{w(I_0)}{w(I_u)} \leq \log \frac{2w(I_0)}{|\alpha_u - \beta_u|}. \tag{6}$$
   Again, if u' is another type-0 leaf different from u, then $\alpha_u \neq \alpha_{u'}$, since $\alpha_u \in C_{I_u}$, $\alpha_{u'} \in C_{I'_u}$ and $C_{I_u} \cap C_{I_{u'}} = \emptyset$. Furthermore, we can choose β_u and $\beta_{u'}$ such that $\beta_u \neq \beta_{u'}$. This is clear if both α_u and $\alpha_{u'}$ are not real, since then $\beta_w = \overline{\alpha_w}$, $w = u, u'$; if both are real then β_u and $\beta_{u'}$ can be chosen as in the argument of type-1 leaves; otherwise, say α_u is real and $\alpha_{u'}$ is not, we can choose $\beta_u = \alpha_{u'}$ and $\beta_{u'} = \overline{\alpha_{u'}}$ without affecting (6).

Let $U_0 \subseteq U$ and $U_1 \subseteq U$ denote the set of type-0 and type-1 leaves respectively. Then substituting (4) and (6) in (2) we get

$$\#(T') \leq \sum_{u \in U_0} \log \frac{2w(I_0)}{|\alpha_u - \beta_u|} + \sum_{u \in U_1} \log \frac{4w(I_0)}{\sqrt{3}|\alpha_u - \beta_u|}. \tag{7}$$

We obtain a bound on the number of type-0 and type-1 leaves:

LEMMA 3.3. *For U_0 and U_1 defined as above we have:*
(i) $|U_0|$ is at most the number of non-real roots of $A(X)$.
(ii) $|U_1|$ is at most the number of real roots of $A(X)$.

PROOF. As shown above, with each $u \in U_0$ we can associate a unique pair of roots (α_u, β_u), where at least one of them is complex and uniquely chosen thus implying the upper bound on $|U_0|$.

Again by the arguments given earlier, for each $u \in U_1$ we can associate a unique real root α_u, and hence the upper bound on $|U_1|$. □

Now we can show our main result:

THEOREM 3.4. *Let $A(X) \in \mathbb{R}[X]$ be a square-free polynomial of degree n. Let T be the recursion tree of the Descartes method run on (A, I_0). Then the number of nodes in T is $O(\log(\frac{1}{|\mathrm{discr}(A)|}) + n(\log \mathrm{M}(A) + \log n + \log w(I_0)))$.*

PROOF. From (7), we know that the number of nodes in T' is bounded by

$$\#(T') \leq |U| \log 4w(I_0) - \sum_{u \in U} \log(|\alpha_u - \beta_u|). \tag{8}$$

Consider the graph G whose edge set is $E_1 \cup E_0$, where $E_0 := \{(\alpha_u, \beta_u) | u \in U_0\}$ and $E_1 := \{(\alpha_u, \beta_u) | u \in U_1\}$. We want to show that G satisfies the conditions of Theorem 3.1. First of all, for any $u \in U$ we can reorder the pair (α_u, β_u) to ensure that $|\alpha_u| \leq |\beta_u|$ without affecting (7).

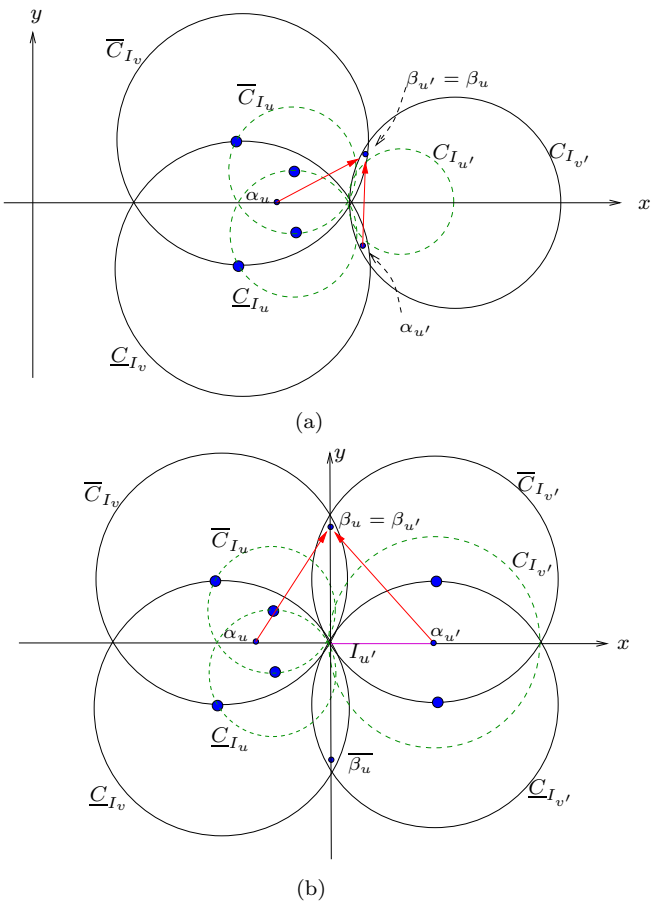

Figure 3: A type-0 and type-1 leaf sharing the same root.

Now we show that the in-degree of G may be assumed to be at most one. Clearly, the edge sets E_0 and E_1 have in-degree one. However, in $E_0 \cup E_1$ cases like that illustrated in Figure 3 may occur. But we can reduce the in-degree of β_u to one in both cases: in (a), we can always re-order the edge $(\alpha_{u'}, \beta_{u'})$ to $(\beta_{u'}, \alpha_{u'})$, since $\beta_{u'} = \overline{\alpha_{u'}}$; in (b), we can choose $\beta_{u'} = \overline{\beta_u}$.

Applying Theorem 3.1 to G we get:

$$\prod_{u \in U} |\alpha_u - \beta_u| \geq \sqrt{|\mathrm{discr}(A)|} \cdot \mathrm{M}(A)^{-(n-1)} \cdot \left(\frac{n}{\sqrt{3}}\right)^{-|U|} n^{-n/2}. \tag{9}$$

Taking logarithm on both sides yields:

$$\sum_{u \in U} \log |\alpha_u - \beta_u| \geq \frac{1}{2} \log(|\mathrm{discr}(A)|) - (n-1) \log \mathrm{M}(A) - n \log \frac{n}{\sqrt{3}} - \frac{n}{2} \log n; \tag{10}$$

since $|U| \leq n$ (by Lemma 3.3). Plugging this into (8) gives

us:

$$\begin{aligned}\#(T') &\leq |U|\log w(I_0) + 2|U| + n\log \mathrm{M}(A) \\ &\quad + \frac{1}{2}\log\frac{1}{|\mathrm{discr}(A)|} + 2n\log n\end{aligned}$$

Using $|U| \leq n$ again, the claim follows. □

REMARKS. (i) There exist intervals I_0 enclosing all real roots of $A(X)$ such that $w(I_0) \leq 2\mathrm{M}(A)/|a_n|$, because $\mathrm{M}(A)/|a_n|$ is an upper bound on the magnitude of all roots.
(ii) Landau's inequality $\mathrm{M}(A) \leq \|A\|_2$ (e.g., [27, Lem. 4.14(i)]) and the obvious estimate $\|A\|_2 \leq \sqrt{n+1}\|A\|_\infty$ immediately yield bounds on the number of nodes in T in terms of these norms of $A(X)$.

COROLLARY 3.5. *Let $A(X)$ be a square-free polynomial of degree n with integer coefficients of magnitude less than 2^L. Let I_0 be an open interval enclosing all real roots of $A(X)$ such that $\log w(I_0) = O(L)$. Let T be the recursion tree of the Descartes method run on (A, I_0). Then the number of nodes in T is $O(n(L + \log n))$.*

PROOF. Since $A(X)$ is a square-free integer polynomial, $|\mathrm{discr}(A)|$ is at least one. From the remark above, we have $\mathrm{M}(A) < 2^L\sqrt{n+1}$. Finally, $\log w(I_0) \leq L + 1$. □

The condition $\log w(I_0) = O(L)$ is no restriction, as 2^L is an upper bound on the absolute value of all roots of $A(X)$ (e.g., [27, Cor. 6.8]).

3.3 Almost Tight Lower Bound

We show that our tree size bound $O(n(L + \log n))$ for integer polynomials is optimal under the assumption $L = \Omega(\log n)$. To do so, we construct a family of inputs of unbounded degree n and coefficient length L for which the height of the recursion tree is $\Omega(nL)$.

Mignotte [16] gave a family of polynomials $P(X) = X^n - 2(aX-1)^2$ parameterized by integers $n \geq 3$ and $a \geq 3$. By Eisenstein's criterion, $P(X)$ is irreducible (use the prime number 2). Let $h = a^{-n/2-1}$. Since $P(a^{-1}) > 0$ and $P(a^{-1} \pm h) = (a^{-1} \pm h)^n - 2a^{-n} < 0$, there exist two distinct roots α and β of $P(X)$ in $(a^{-1} - h,\, a^{-1} + h)$. Clearly, $|\alpha - \beta| < 2h$. In the sequel, we shall restrict to the case that the degree n is even. This allows us to conclude that any interval I_0 enclosing all roots of $P(X)$ is a superset of $(0, 1)$, because the sign of $P(X)$ is positive for $X \to \pm\infty$ but negative for $X = 0$ and $X = 1$.

If one is willing to accept certain assumptions on the choice of the initial interval $I_0 = (-B_1, +B_2)$, such as integrality of B_1 and B_2, the input $P(X)$ can be used to demonstrate the necessity of $\Omega(nL)$ bisections before α and β are separated. However, less customary choices of I_0 could cause some bisection to separate α and β much earlier.

We shall avoid this problem. Let us consider the closely related polynomial $P_2(X) = X^n - (aX-1)^2$ which appears in a later work of Mignotte [17] on complex roots. Again, we see that $P_2(a^{-1}) > 0$, and furthermore $P_2(a^{-1} - h) = (a^{-1} - h)^n - a^{-n} < 0$. Hence there is a root γ of $P_2(X)$ in $(a^{-1} - h,\, a^{-1})$. By irreducibility of $P(X)$, the product $Q(X) = P(X) \cdot P_2(X)$ is square free and has three distinct roots α, β, and γ in $(a^{-1} - h,\, a^{-1} + h)$.

THEOREM 3.6. *Let $a \geq 3$ be an L-bit integer. and let $n \geq 4$ be an even integer. Consider the square-free polynomial $Q(X) = P(X) \cdot P_2(X)$ of degree $2n$. Its coefficients are integers of at most $O(L)$ bits. The Descartes method executed for $Q(X)$ and any initial interval I_0 enclosing all roots of $Q(X)$ has a recursion tree of height $\Omega(nL)$.*

PROOF. As discussed above, I_0 is a superset of $(0, 1)$ and thus has width $w(I_0) > 1$. Let I_1 be the isolating interval reported by the Descartes method for the median of $\alpha, \beta, \gamma \in (a^{-1} - h,\, a^{-1} + h)$. Clearly, $w(I_1) < 2h$. The number of bisections needed to obtain I_1 from I_0 is $\log w(I_0)/w(I_1) > \log(1/2h) \geq (n/2+1)(L-1) - 1 = \Omega(nL)$. □

Clearly, the same argument applies to any form of root isolation by repeated bisection, including Sturm's method.

4. THE BIT COMPLEXITY

We derive the bit complexity of the Descartes method for a square-free polynomial $A_{\mathrm{in}}(X)$ with integer coefficients of magnitude less than 2^L in the power basis. We can enclose all its real roots in an interval $(-B_1, +B_2)$ such that B_1 and B_2 are positive integers of magnitude less than 2^{L+1} (e.g., [27, Cor. 6.8]).

We discuss the bit complexity of the power basis and Bernstein basis variants of the Descartes method applied to the scaled polynomial $A(X) := \sum_{i=0}^n a_i X^i := A_{\mathrm{in}}((B_1 + B_2)X - B_1)$. We can bound the bit length of its coefficients as follows. The power basis coefficients a_i of $A(X)$ have bit lengths $O(nL)$. For conversion from power basis to Bernstein basis, one has [22, §2.8]

$$n!A(X) = \sum_{i=0}^{n} B_i^n(X) \sum_{k=0}^{i} i(i-1)\cdots(i-k+1)(n-k)!a_k. \tag{11}$$

To avoid fractions, we use $n!A(X)$ for the Bernstein basis variant. Observe that $l(l-1)\cdots(l-k+1)(n-k)! \leq n! \leq n^n$, so that the Bernstein coefficients of $n!A(X)$ have bit length $O(nL + n\log n)$.

From Corollary 3.5 we know that the size of the recursion tree is $O(n(L + \log n))$. Note that the transformation from $A_{\mathrm{in}}(X)$ to $A(X)$ does not affect the size of the recursion tree, i.e., the size does not increase to $O(n(L' + \log n))$ where L' bounds the bit size of the coefficients of $A(X)$ or $n!A(X)$.

Let us now bound coefficient length at depth $h > 0$. For the power basis variant, we start with coefficients of length $O(nL)$. Both the H and TH transformations increase the length of the coefficients by $O(n)$ bits on each level. It is known that we can perform the T-transformation in $O(n^2)$ additions [11, 10, 26]; the H-transformation needs $O(n)$ shift operations. Hence a node at recursion depth h has bit cost $O(n^2(nL + nh))$ for the power basis. In the Bernstein basis, we need $O(n^2)$ additions and $O(n)$ shifts for the fraction-free de Casteljau algorithm, which also increases the length of the coefficients by $O(n)$ bits on each level. This gives us a bit cost of $O(n^2(nL + n\log n + nh))$. Since $h = O(n(L + \log n))$, the worst-case cost in any node is $O(n^4(L + \log n))$ for both variants. Multiplied with the tree size, this yields an overall bit complexity of $O(n^5(L + \log n)^2)$, cf. [9, Thm. 13] [11, Thm. 50]. To summarize:

THEOREM 4.1. *Let $A(X)$ be a square-free polynomial of degree n with integer coefficients of magnitude less than 2^L. Then the bit complexity of isolating all real roots of $A(X)$ using the Descartes method (in either power basis or Bernstein basis variant) is $O(n^5(L + \log n)^2)$ using only classical*

arithmetic. Except for the initial transformation, only additions and shifts are used.

For the Bernstein basis variant, this result is an improvement by a factor of n on the result in [19]. For the power basis variant, this bound was already achieved by Krandick [11]. Theorem 4.1 can be improved using a fast Taylor shift algorithm [26, Method F]:

THEOREM 4.2. *Let $A(X)$ be a square-free polynomial of degree n with integer coefficients of magnitude less than 2^L. Then the bit complexity of isolating the real roots of $A(X)$ using the Descartes method in the power basis with a fast Taylor shift is $O(nM(n^3(L+\log n))(L+\log n))$. Here, $M(n)$ is the bit complexity of multiplying two n-bit integers.*

PROOF. The work at a node at depth h of the recursion tree has bit cost $O(M(n^2 \log n + n^2 L + n^2 h)$ [26]. Substituting $h = O(n(L + \log n))$, we get the bound $O(M(n^3(L + \log n))$. Multiplied by tree size $O(n(L + \log n))$, we obtain the theorem. □

REMARK.[3] Emiris, Mourrain, and Tsigaridas [7] describe the following approach to obtain a similar speedup for the Bernstein basis variant: Suppose the vector $(b_i)_i$ of Bernstein coefficients of $A(X) = \sum_{i=0}^n b_i B_i^n(X)$ is given and the Bernstein coefficients $(b'_i)_i$ of $A_L(X) = H(A)(X) = \sum_{i=0}^n b'_i B_i^n(X)$ are wanted. Define the auxiliary polynomial $Q(X) = \sum_{i=0}^n b_{n-i}\binom{n}{i}X^i$ $(= TR(A(X)))$ and transform it by substituting $2X + 1$ for X. It is straightforward to verify that $Q_L(X) := Q(2X+1) = \sum_{i=0}^n b'_{n-i}\binom{n}{i}X^i$; thus one can compute the Bernstein coefficients of $A_L(X)$ from the Bernstein coefficients of $A(X)$ using one asymptotically fast Taylor shift and scalings of coefficients. By symmetry, the same holds for the Bernstein coefficients of $A_R(X)$. More precisely, define[4] $Q_R(X) := (2+X)^n Q(X/(2+X)) = \sum_{i=0}^n b''_{n-i}\binom{n}{i}X^i$. Then the b''_i's are Bernstein coefficients of $A_R(X)$. Together with bounds on the size of the recursion tree (Cor. 3.5) and the lengths of coefficients, this leads [7] to a bit complexity of $\widetilde{O}(n^4L^2)$ for the Bernstein basis variant of the Descartes method.

However, repeatedly putting in and taking out the extra factor $\binom{n}{i}$ in the i-th coefficient is an unnecessary artifact of insisting on the Bernstein basis. A more natural formulation of this approach avoids this extra scaling and the reversal of the coefficient sequence by representing polynomials in the scaled and reversed Bernstein basis $\widetilde{B}_i^n(X) = \binom{n}{i}^{-1} B_{n-i}^n(X) = (1-X)^i X^{n-i}$. Now the steps from $A(X)$ to $Q(X)$ and back from $Q(2X+1)$ to $A_L(X)$ are purely conceptual: reinterpret the coefficients of $\widetilde{B}_i^n(X)$ as coefficients of X^i and vice versa. The resulting algorithm is the **scaled Bernstein basis variant** of the Descartes method.

An alternative view on this variant is to regard it as an optimization of the power basis variant: By Eq. (1), the reinterpretation of coefficients is equivalent to the transformation TR. Recall that each recursive invocation of the power basis variant handles four polynomials: $A(X)$ is received from the parent, the Descartes test constructs $TR(A)(X)$, and subdivision computes $A_L(X)$ and $A_R(X)$. In these terms, the scaled Bernstein basis variant receives $TR(A)(X)$ instead of $A(X)$, eliminating the need for a separate transformation in the Descartes test, and it subdivides $TR(A)(X)$ into $TR(A_L)(X)$ and $TR(A_R)(X)$ directly, without explicitly constructing $A_L(X)$ and $A_R(X)$. Over the entire recursion tree, this saves one third of the T transformations in the power basis formulation.

[3]We thank an anonymous referee for pointing out the necessity of a remark on this aspect.

[4]Let $Q_L(X)$ be expressed as $H_2(T(Q(X)))$ where $H_2(Q(X)) := Q(2X)$. Then $Q_R(X)$ is $R(H_2(T(R(Q(X)))))$.

5. CONCLUSION

Our work aims to achieve the best possible complexity bounds for the Descartes method (either power basis or Bernstein basis), and to match similar bounds for Sturm's method. We achieve matching bounds for two measures: (1) the size of the recursion tree, and (2) the bit complexity of the overall algorithm. Moreover, we show that the tree size bound is the best possible under the assumption that $L = \Omega(\log n)$. It would be of some interest to completely resolve this optimality question.

Another direction of interest is to extend these algorithms and results to the non-squarefree case. The standard way to achieve such extensions is to apply the above results to the square-free part $A/\gcd(A, A')$ of a given polynomial A (see, e.g., [1, Algo. 10.41] [7]) – but the real challenge is to provide an algorithm based on the Descartes method that works directly on non-squarefree polynomials.

Acknowledgements

The authors thank Werner Krandick and Kurt Mehlhorn for useful comments on the subject matter of this paper.

6. REFERENCES

[1] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2003.

[2] P. Batra. *Abschätzungen und Iterationsverfahren für Polynom-Nullstellen*. PhD thesis, Technical University Hamburg-Harburg, 1999.

[3] 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.

[4] G. E. Collins and J. R. Johnson. Quantifier elimination and the sign variation method for real root isolation. In *Proc. ACM-SIGSAM Symposium on Symbolic and Algebraic Computation*, pages 264–271, 1989.

[5] J. H. Davenport. Computer algebra for cylindrical algebraic decomposition. Tech. Rep., Royal Inst. of Technology, Dept. of Numer. Analysis and Computing Science, Stockholm, Sweden, 1985. Reprinted as Tech. Rep. 88-10, U. of Bath, School of Math. Sciences, Bath, England. `http://www.bath.ac.uk/~masjhd/TRITA.pdf`.

[6] Z. Du, V. Sharma, and C. K. Yap. Amortized bound for root isolation via Sturm sequences. In D. Wang and L. Zhi, editors, *Proc. Internat. Workshop on Symbolic-Numeric Computation*, pages 81–93, 2005. Int'l Workshop on Symbolic-Numeric Computation, Xi'an, China, Jul 19–21, 2005.

[7] I. Z. Emiris, B. Mourrain, and E. P. Tsigaridas. Real algebraic numbers: Complexity analysis and

experimentations. Research Report 5897, INRIA, April 2006. http://www.inria.fr/rrrt/rr-5897.html.

[8] C. G. J. Jacobi. Observatiunculae ad theoriam aequationum pertinentes. *Journal für die reine und angewandte Mathematik*, 13:340–352, 1835. Available from http://gdz.sub.uni-goettingen.de.

[9] J. R. Johnson. Algorithms for polynomial real root isolation. In B. F. Caviness and J. R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, pages 269–299. Springer, 1998.

[10] 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.

[11] W. Krandick. Isolierung reeller Nullstellen von Polynomen. In J. Herzberger, editor, *Wissenschaftliches Rechnen*, pages 105–154. Akademie-Verlag, Berlin, 1995.

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

[13] J. M. Lane and R. F. Riesenfeld. Bounds on a polynomial. *BIT*, 21:112–117, 1981.

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

[15] K. Mahler. An inequality for the discriminant of a polynomial. *Michigan Mathematical Journal*, 11:257–262, 1964.

[16] M. Mignotte. Some inequalities about univariate polynomials. In *Proc. 1981 ACM Symposium on Symbolic and Algebraic Computation (SYMSAC 1981)*, pages 195–199. ACM, 1981.

[17] M. Mignotte. On the distance between the roots of a polynomial. *Applicable Algebra in Engineering, Commun., and Comput.*, 6:327–332, 1995.

[18] M. Mignotte and D. Ştefănescu. *Polynomials: An Algorithmic Approach*. Springer, Singapore, 1999.

[19] 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.

[20] 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.

[21] A. M. Ostrowski. Note on Vincent's theorem. *Annals of Mathematics, 2nd Ser.*, 52:702–707, 1950. Reprinted in: A. Ostrowski, Collected Mathematical Papers, vol. 1, 728–733, Birkhäuser, 1983.

[22] H. Prautzsch, W. Boehm, and M. Paluszny. *Bézier and B-Spline Techniques*. Springer, 2002.

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

[24] F. Rouillier and P. Zimmermann. Efficient isolation of a polynomial['s] real roots. Rapport de Recherche 4113, INRIA, 2001. http://www.inria.fr/rrrt/rr-4113.html.

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

[26] 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.

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