Almost Tight Recursion Tree Bounds for the Descartes Method

Arno Eigenwillig†  Vikram Sharma*  Chee K. Yap*

†Max-Planck-Institut für Informatik
Saarbrücken, Germany

*Courant Institute
Dept. of Computer Science
New York University, NY, USA

ISSAC 2006 at Genoa, Italy
11th July 2006
The Descartes Method

What is the Descartes Method?
Real root isolation by recursive interval bisection using Descartes’ Rule of Signs to test for roots.

What makes the Descartes Method interesting?
- It performs very well in practice.
- It is simple to implement.
- It is used a lot.
The Descartes Test for roots in an interval

**Descartes Test (classical form) [Jacobi, 1835]**

Consider the real polynomial $A(X)$ and an interval $(c, d)$. Let $A^*(X) = \sum_{i=0}^{n} a_i^* X^i = A((cX + d)/(X + 1)) \cdot (X + 1)^n$ and define

$$\text{DescartesTest}(A, (c, d)) := \text{var}(a_0^*, \ldots, a_n^*).$$

**Descartes Test (Bernstein form) [Pólya/Schoenberg, 1958]**

Let $A(X) = \sum_{i=0}^{n} b_i B_i^n(X)$, where $B_i^n(X) = {n \choose i} \frac{(X-c)^i(d-X)^{n-i}}{(d-c)^n}$. Then

$$\text{DescartesTest}(A, (c, d)) = \text{var}(b_0, \ldots, b_n).$$
The Descartes Test for roots in an interval

Properties

Let \( v = \text{DescartesTest}(A, (c, d)) \).

- If \( v = 0 \), then \( A(X) \) has no roots in \((c, d)\).
- If \( v = 1 \), then \( A(X) \) has exactly one root in \((c, d)\), which is simple.
- If \( v \geq 2 \), then \( A(X) \) has two or more roots (or a multiple root) in or near \((c, d)\) in the complex plane.
The Descartes Test for roots in an interval

Properties

Let \( v = \text{DescartesTest}(A, (c, d)) \).

- If \( v = 0 \), then \( A(X) \) has no roots in \( (c, d) \).
- If \( v = 1 \), then \( A(X) \) has exactly one root in \( (c, d) \), which is simple.
- If \( v \geq 2 \), then \( A(X) \) has two or more roots (or a multiple root) in or near \( (c, d) \) in the complex plane.

From now on, let \( A(X) \) be square free.
The initial interval $I_0$ is chosen to enclose all real roots.
The Descartes Method

\[ \text{DescartesTest}(A, I_0) \geq 2 \implies \text{subdivide } I_0. \]
The Descartes Method

Continue recursively.
The Descartes Method

Continue recursively.
The Descartes Method

\[ \text{DescartesTest}(\ldots) = 0 \implies \text{no roots found, return.} \]
The Descartes Method

\[ \text{DescartesTest}(\ldots) = 1 \implies \text{report isolating interval, return.} \]
Continue recursion.
The Descartes Method
The Descartes Method

4 isolating intervals have been found.
Related Work (selection)

### Description of the algorithm

- Classical / power basis variant: [Collins/Akritas, 1976]
- Bernstein basis variant: [Lane/Riesenfeld, 1981]
  (later: e.g., [Mourr./Vrah./Yakoubs., 2002] [Mourr./Rouillier/Roy, 2005])
## Related Work (selection)

### Description of the algorithm

- Classical / power basis variant: [Collins/Akritas, 1976]
- Bernstein basis variant: [Lane/Riesenfeld, 1981]  
  (later: e.g., [Mourr./Vrah./Yakoubs., 2002] [Mourr./Rouillier/Roy, 2005])

### Tools from previous analyses

- [Krandick/Mehlhorn, 2006] used a Theorem of [Ostrowski, 1950]  
  (also mentioned by [Batra, 1999]).

We use the same tools, but in a more direct way.
Tool #1: A partial converse of Descartes’ Rule
Tool #1: A partial converse of Descartes’ Rule
Tool #1: A partial converse of Descartes’ Rule
Two-circle Theorem (contrapositive)
([Ostrowski, 1950], see [Kra./Meh., 2006])

If $\text{DescartesTest}(A, (c, d)) \geq 2$, then the two-circles figure in $\mathbb{C}$ around interval $(c, d)$ contains two roots $\alpha, \beta$ of $A(X)$.  

Tool #1: A partial converse of Descartes’ Rule
Tool #1: A partial converse of Descartes’ Rule

Two-circle Theorem (contrapositive)
([Ostrowski, 1950], see [Kra./Meh., 2006])

If $\text{DescartesTest}(A, (c, d)) \geq 2$, then the two-circles figure in $\mathbb{C}$ around interval $(c, d)$ contains two roots $\alpha, \beta$ of $A(X)$.

Corollary

We can choose $\alpha, \beta$ to be complex conjugate or adjacent real roots.
Tool #1: A partial converse of Descartes’ Rule

Two-circle Theorem (contrapositive)
([Ostrowski, 1950], see [Kra./Meh., 2006])

If \( \text{DescartesTest}(A, (c, d)) \geq 2 \), then the two-circles figure in \( \mathbb{C} \) around interval \((c, d)\) contains two roots \( \alpha, \beta \) of \( A(X) \).

Corollary

We can choose \( \alpha, \beta \) to be complex conjugate or adjacent real roots.
Tool #1: A partial converse of Descartes’ Rule

Two-circle Theorem (contrapositive)
([Ostrowski, 1950], see [Kra./Meh., 2006])

If \( \text{DescartesTest}(A, (c, d)) \geq 2 \), then the two-circles figure in \( \mathbb{C} \) around interval \((c, d)\) contains two roots \( \alpha, \beta \) of \( A(X) \).

Corollary

We can choose \( \alpha, \beta \) to be complex conjugate or adjacent real roots.
Tool #1: A partial converse of Descartes’ Rule

Two-circle Theorem (contrapositive)
([Ostrowski, 1950], see [Kra./Meh., 2006])

If $\text{DescartesTest}(A, (c, d)) \geq 2$, then the two-circles figure in $\mathbb{C}$ around interval $(c, d)$ contains two roots $\alpha, \beta$ of $A(X)$.

Corollary

We can choose $\alpha, \beta$ to be complex conjugate or adjacent real roots. It holds that $|\beta - \alpha| < \sqrt{3}(d - c)$; i.e., $(d - c) > |\beta - \alpha|/\sqrt{3}$. 
A tree bound in terms of roots (1)

A bound on path length

Consider any path in the recursion tree from $I_0$ to a parent $J$ of two leaves.
A tree bound in terms of roots (1)

Consider any path in the recursion tree from $I_0$ to a parent $J$ of two leaves.

At depth $d$, interval width is $2^{-d}|I_0|$. Hence $J$ is at depth $d = \log |I_0|/|J|$. 

A bound on path length
A tree bound in terms of roots (1)

Consider any path in the recursion tree from $I_0$ to a parent $J$ of two leaves.

1. At depth $d$, interval width is $2^{-d}|I_0|$. Hence $J$ is at depth $d = \log \frac{|I_0|}{|J|}$.

2. The whole path consists of $d + 1$ internal nodes.
A tree bound in terms of roots (1)

Consider any path in the recursion tree from $I_0$ to a parent $J$ of two leaves.

At depth $d$, interval width is $2^{-d}|I_0|$. Hence $J$ is at depth $d = \log |I_0|/|J|$.

The whole path consists of $d + 1$ internal nodes.

There is a pair of roots $(\alpha_J, \beta_J)$ such that $|J| > |\beta_J - \alpha_J|/\sqrt{3}$; hence $d + 1 < \log |I_0| - \log |\beta_J - \alpha_J| + 2$. 

A bound on path length

- Consider any path in the recursion tree from $I_0$ to a parent $J$ of two leaves.
- At depth $d$, interval width is $2^{-d}|I_0|$. Hence $J$ is at depth $d = \log |I_0|/|J|$.
- The whole path consists of $d + 1$ internal nodes.
- There is a pair of roots $(\alpha_J, \beta_J)$ such that $|J| > |\beta_J - \alpha_J|/\sqrt{3}$; hence $d + 1 < \log |I_0| - \log |\beta_J - \alpha_J| + 2$. 
A tree bound in terms of roots (2)

\[
\text{#(internal nodes on path)} < \log |I_0| - \log |\beta_J - \alpha_J| + 2
\]
A tree bound in terms of roots (2)

\[(\text{internal nodes on path}) < \log |I_0| - \log |\beta_J - \alpha_J| + 2\]

\[\sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)\]

\[(\text{internal nodes in tree}) < \sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)\]
A tree bound in terms of roots (2)

$$\#(\text{internal nodes on path}) < \log |I_0| - \log |\beta_J - \alpha_J| + 2$$

$$\#(\text{internal nodes in tree}) < \sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)$$

$$\#(\text{all nodes in tree}) < 1 + 2 \cdot \sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)$$
A tree bound in terms of roots (2)

Proposition

The size of the recursion tree is bounded by

\[-2 \log \prod_{J} |\beta_{J} - \alpha_{J}| + n \log |I_{0}| + 2n + 1\]

#(internal nodes on path) < \[\log |I_{0}| - \log |\beta_{J} - \alpha_{J}| + 2\]

#(internal nodes in tree) < \[\sum_{J} (\log |I_{0}| - \log |\beta_{J} - \alpha_{J}| + 2)\]

#(all nodes in tree) < \[1 + 2 \cdot \sum_{J} (\log |I_{0}| - \log |\beta_{J} - \alpha_{J}| + 2)\]
Proposition

The size of the recursion tree is bounded by

$$-2 \log \prod_J |\beta_J - \alpha_J| + n \log |I_0| + 2n + 1$$

#(internal nodes on path) < $\log |I_0| - \log |\beta_J - \alpha_J| + 2$

#(internal nodes in tree) < $\sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)$

#(all nodes in tree) < $1 + 2 \cdot \sum_J (\log |I_0| - \log |\beta_J - \alpha_J| + 2)$
Tool #2: The Davenport–Mahler bound

**Theorem (Davenport–Mahler [Dav., 1985] [Johnson, 1991/98])**

Consider a polynomial \( A(X) \in \mathbb{C}[X] \) of degree \( n \). Let \( G = (V, E) \) be a digraph whose node set \( V \) consists of the roots \( \vartheta_1, \ldots, \vartheta_n \) of \( A(X) \). If

(i) \( (\alpha, \beta) \in E \implies |\alpha| \leq |\beta| \),

(ii) \( \beta \in V \implies \text{indeg}(\beta) \leq 1 \), and

(iii) \( G \) is acyclic,

then

\[
\prod_{(\alpha, \beta) \in E} |\beta - \alpha| \geq \frac{\sqrt{|\text{discr}(A)|}}{M(A)^{n-1}} \cdot 2^{-O(n \log n)},
\]

where

\[
\text{discr}(A) := a_n^{2^{n-2}} \prod_{i>j} (\vartheta_i - \vartheta_j)^2 \quad \text{and} \quad M(A) := |a_n| \prod_i \max\{1, |\vartheta_i|\}.
\]
Turning our product into an admissible graph

We want to rewrite

\[ \prod_{J} |\beta_{J} - \alpha_{J}| \quad \text{as} \quad \prod_{(\alpha, \beta) \in E} |\beta - \alpha|. \]
Turning our product into an admissible graph

We want to rewrite

$$\prod_{J} |\beta_{J} - \alpha_{J}| \text{ as } \prod_{(\alpha, \beta) \in E} |\beta - \alpha|.$$  

How often does $|\beta_{J} - \alpha_{J}|$ appear?

- adjacent real: $\leq 1$
Turning our product into an admissible graph

We want to rewrite
\[ \prod_{J} |\beta_J - \alpha_J| \text{ as } \prod_{(\alpha, \beta) \in E} |\beta - \alpha|. \]

How often does \( |\beta_J - \alpha_J| \) appear?
- adjacent real: \( \leq 1 \)
- complex conjugate:
Turning our product into an admissible graph

We want to rewrite

$$\prod_{J} |\beta_{J} - \alpha_{J}| \quad \text{as} \quad \prod_{(\alpha, \beta) \in E} |\beta - \alpha|.$$ 

How often does $|\beta_{J} - \alpha_{J}|$ appear?

- adjacent real: $\leq 1$
- complex conjugate:
Turning our product into an admissible graph

We want to rewrite
\[
\prod_{J} |\beta_J - \alpha_J| \quad \text{as} \quad \prod_{(\alpha, \beta) \in E} |\beta - \alpha|.
\]

How often does \( |\beta_J - \alpha_J| \) appear?
- adjacent real: \( \leq 1 \)
- complex conjugate: \( \leq 2 \)

We need two graphs. (Paper: just 1.)
Turning our product into an admissible graph

We want to rewrite
\[ \prod_{J} |\beta_J - \alpha_J| \] as
\[ \prod_{(\alpha, \beta) \in E} |\beta - \alpha|. \]

How often does \(|\beta_J - \alpha_J|\) appear?
- adjacent real: \(\leq 1\)
- complex conjugate: \(\leq 2\)

We need two graphs. (Paper: just 1.)

Conditions on \(G = (V, E)\)

(i) \((\alpha, \beta) \in E \implies |\alpha| \leq |\beta|\)
(ii) \(\beta \in V \implies \text{indeg}(\beta) \leq 1\)
(iii) \(G\) is acyclic
Turning our product into an admissible graph

We want to rewrite
\[
\prod_J |\beta_J - \alpha_J| \quad \text{as} \quad \prod_{(\alpha, \beta) \in E} |\beta - \alpha|.
\]

How often does \(|\beta_J - \alpha_J|\) appear?
- adjacent real: \(\leq 1\)
- complex conjugate: \(\leq 2\)

We need two graphs. (Paper: just 1.)

**Conditions on** \(G = (V, E)\)

1. \((\alpha, \beta) \in E \implies |\alpha| \leq |\beta|\)
2. \(\beta \in V \implies \text{indeg}(\beta) \leq 1\)
3. \(G\) is acyclic
Turning our product into an admissible graph

We want to rewrite
\[ \prod_{J} |\beta_J - \alpha_J| \] as
\[ \prod_{(\alpha, \beta) \in E} |\beta - \alpha|. \]

How often does \( |\beta_J - \alpha_J| \) appear?

- adjacent real: \( \leq 1 \)
- complex conjugate: \( \leq 2 \)

We need two graphs. (Paper: just 1.)

Conditions on \( G = (V, E) \)

(i) \( (\alpha, \beta) \in E \implies |\alpha| \leq |\beta| \)
(ii) \( \beta \in V \implies \text{indeg}(\beta) \leq 1 \)
(iii) \( G \) is acyclic
Main Result

Theorem

Let \( A(X) \in \mathbb{R}[X] \) be a square-free polynomial of degree \( n \).

The Descartes Method run on \( A(X) \) starting from interval \( I_0 \)

has a recursion tree \( T \) bounded in size by

\[
|T| = O\left(\log \frac{1}{|\text{discr}(A)|} + n(\log M(A) + \log n + \log |I_0|)\right)
\]

Eigenwillig, Sharma, Yap (MPII + NYU)
Main Result

**Theorem**

Let \( A(X) \in \mathbb{R}[X] \) be a square-free polynomial of degree \( n \). The Descartes Method run on \( A(X) \) starting from interval \( I_0 \) has a recursion tree \( T \) bounded in size by

\[
|T| = O\left(\log \frac{1}{|\text{discr}(A)|} + n(\log M(A) + \log n + \log |I_0|)\right)
\]

**Corollary**

If \( A(X) \in \mathbb{Z}[X] \) and \( |a_i| < 2^L \), then easily \( \log |I_0| = O(L) \), and one has

\[
|T| = O(n(L + \log n)).
\]

Argument of [Krandick/Mehlhorn, 2006]: \( |T| = O(n \log n (L + \log n)) \).
Almost tightness of the bound

Choose integers $n \geq 3$ and $a \geq 3$. Let $h = a^{-n/2} - 1$. Consider

\[
P(X) = X^n - 2(aX - 1)^2 \quad (\text{irreducible}) \quad \text{[Mignotte, 1981]}
\]
\[
P_2(X) = X^n - (aX - 1)^2 \quad \text{[Mignotte, 1995]}
\]

The interval $(a^{-1} - h, a^{-1} + h)$ contains two roots of $P(X)$ and one root of $P_2(X)$ and thus three roots of $Q(X) = P(X) \cdot P_2(X)$. Their median has an isolating interval of width less than $2h$, but $Q(X)$ has real roots outside $(0, 1)$, so $|I_0| > 1$.

Hence recursion depth is more than $\log(1/(2h)) = \Omega(n \log a)$. $Q(X)$ has degree $2n = \Theta(n)$ and coefficient length $L = \Theta(\log a)$.

Lower bound $\Omega(nL)$ matching $O(n(L + \log n))$ if $\log n = O(L)$. 

Eigenwillig, Sharma, Yap  (MPII + NYU)  
Tree Bounds for the Descartes Method  
ISSAC 2006  ●  Genoa, Italy  
13 / 17
Bit complexity for integer polynomials

<table>
<thead>
<tr>
<th>Bit complexity depends on . . .</th>
</tr>
</thead>
<tbody>
<tr>
<td>• the basis chosen to represent polynomials</td>
</tr>
<tr>
<td>• Power basis $(x^i)_i = (1, x, x^2, \ldots, x^n)$</td>
</tr>
<tr>
<td>• $[0, 1]$-Bernstein basis $(\binom{n}{i} x^i (1-x)^{n-i})_i$</td>
</tr>
<tr>
<td>• scaled $[0, 1]$-Bernstein basis $(x^i (1-x)^{n-i})_i$</td>
</tr>
<tr>
<td>(NB: Coefficient length $L$ always refers to power basis.)</td>
</tr>
<tr>
<td>• the implementation of basic operations, esp. transformation of $A(X)$ to $A_L(X) = 2^n A(X/2)$ and $A_R(X) = 2^n A((X+1)/2)$.</td>
</tr>
</tbody>
</table>
Bit complexity for integer polynomials

Bit complexity depends on...

- the basis chosen to represent polynomials
  - Power basis \((x^i)_i = (1, x, x^2, \ldots, x^n)\)
  - \([0, 1]\)-Bernstein basis \(\binom{n}{i} x^i (1-x)^{n-i}_i\)
  - scaled \([0, 1]\)-Bernstein basis \(x^i (1-x)^{n-i}_i\)

  (NB: Coefficient length \(L\) always refers to power basis.)

- the implementation of basic operations, esp. transformation of \(A(X)\) to \(A_L(X) = 2^n A(X/2)\) and \(A_R(X) = 2^n A((X + 1)/2)\).

Classical subdivision

- Power basis + classical Taylor shift: \(O(n^5 (L + \log n)^2)\).
  (Same bound as Johnson/Krandick/Mehlhorn, but simpler proof.)

- Bernstein basis + de Casteljau subdivision: \(O(n^5 (L + \log n)^2)\).
Bit complexity for integer polynomials

Classical subdivision

- Power basis + classical Taylor shift: \( O(n^5(L + \log n)^2) \).
  (Same bound as Johnson/Krandick/Mehlhorn, but simpler proof.)
- Bernstein basis + de Casteljau subdivision: \( O(n^5(L + \log n)^2) \).

Asymptotically fast subdivision

- Power basis + fast Taylor shift [vzGathen/Gerhard, 1997]:
  \[
  O(n(L + \log n)M(n^3(L + \log n))) = \tilde{O}(n^4L^2).
  \]
  Same bound as [Du/Sharma/Yap, 2005] for Sturm’s method.
- Bernstein basis: How to subdivide fast?
- A detour through the scaled Bernstein basis (“dual algorithm” of [Johnson, 1991]) makes it possible to apply a fast Taylor shift.
  Our tree bound \( \sim \tilde{O}(n^4L^2) \) [Emiris/Mourrain/Tsigaridas, 2006].
Summary

What have we done?

- Our paper gives a basis-free description of the Descartes Method for a uniform treatment of its power and Bernstein basis variants.
- We have recombined
  - tool #1: Ostrowski’s partial converse of Descartes’ rule
  - tool #2: the Davenport–Mahler bound in a new and simpler way.
- This gives a new and almost tight bound on the recursion tree.
- Bounds on bit complexity follow directly (some old, some new). Asymptotically fast variant attains $\tilde{O}(n^4 L^2)$ like Sturm’s method.
- Replacing $A$ by $A / \gcd(A, A')$ removes squarefreeness condition. Standard arguments show that our bounds remain valid.
Thank you!