NEW BOUNDS FOR THE DESCARTES METHOD

WERNER KRANDICK AND KURT MEHLHORN

Abstract. We give a new bound for the number of recursive subdivisions in the Descartes method for polynomial real root isolation. Our proof uses Ostrowski's theory of normal power series from 1950 which has so far been overlooked in the literature. We combine Ostrowski's results with a theorem of Davenport from 1985 to obtain our bound. We also characterize normality of cubic polynomials by explicit conditions on their roots and derive a generalization of one of Ostrowski's theorems.

1. Introduction

Polynomial real root isolation is the task of computing disjoint intervals, each containing a single root, for all the real roots of a given univariate polynomial with real coefficients. In the 1830s, Vincent [36] showed that polynomial real root isolation can be performed using a test based on the Descartes Rule of Signs. The test evaluates a condition that implies that a given interval contains a single root, and another condition that implies that the interval does not contain any roots. If neither condition is satisfied, the interval is bisected and each subinterval is tested recursively. It is not obvious that Vincent's method terminates.

In 1976, Collins and Akritas [10] proposed a method with a much better worst-case computing time than Vincent's method. We will refer to the improved method as "Descartes method". A study by Johnson [21] shows that the Descartes method typically outperforms Sturm's method and other methods for real root isolation. Johnson's findings are confirmed in experiments by Rouillier and Zimmermann [32, Figures 2,3]. Recent versions of the Descartes method use floating point arithmetic [22, 12, 33], parallel computation [15, 16], or they minimize space requirements [33].

We give a new bound (Theorem 5.5) for the number of recursive subdivisions in the Descartes method. Our proof uses Ostrowski's theory [29] of normal power series from 1950 which has so far been overlooked in the literature. We combine Ostrowski's results with a theorem of Davenport [14] from 1985 to obtain our bound. We also characterize normality of cubic polynomials by explicit conditions on their roots and derive a generalization (Theorem 6.4) of one of Ostrowski's theorems.

The history of termination proofs starts in the 1830s with Vincent [36]. Alesina and Galuzzi [2] present Vincent's original proof in modern mathematical language and provide extensive historical information on related earlier and later results. It seems that Vincent's method was forgotten until 1948 when Uspensky [34] modified Vincent's proof and bounded the number of recursive steps required by the method. In 1950, Ostrowski [29] used a result from his earlier work [31] to improve Uspensky's bound. Ostrowski's contribution, though summarized in *Mathematical Reviews* [26], was completely overlooked in later literature until it became accessible through an electronic database [3]. When Collins and Akritas [10] improved

Vincent's algorithm in 1976 they based their analysis, later elaborated by Collins and Loos [7], on Uspensky's work. Collins and Johnson [6] improved the analysis significantly, but also their result is strictly weaker than Ostrowski's. Eventually, one of Ostrowski's theorems, the present Theorem 3.9, was independently rediscovered by Alesina and Galuzzi [2, Corollary 8.2]. These authors give a concise and direct proof, but their approach cannot be used to prove the stronger Theorem 6.4 of this paper.

In Section 2 we review the Descartes method. In Section 3 we present Ostrowski's theory of normal power series and strengthen one of his results that links normality of polynomials and termination of the Descartes method (Theorem 3.3). We also present Ostrowski's sufficient condition on the roots of a polynomial to guarantee normality (Theorem 3.8). We use these results in Section 4 to prove Theorem 4.6 on the proximity of complex roots to those intervals on which the Descartes method recurs. In Section 5 we combine Theorem 4.6 with Davenport's root separation theorem to obtain new bounds for the recursion tree of the Descartes method. In Section 6 we use Theorem 3.3 to characterize the normal cubic polynomials by explicit conditions on their roots. We gauge the extent of the improvement by applying the Descartes method to 2.3 billion cubic polynomials. We use the new result to prove Theorem 6.4—thus strengthening Theorem 3.9.

2. Review of the Descartes Method

Definition 2.1. Let $a = (a_0, \ldots, a_n)$ be a finite sequence of real numbers. The *number of sign variations* in a, $\mathrm{var}(a)$, is the number of pairs (i, j) with $0 \le i < j \le n$ and $a_i a_j < 0$ and $a_{i+1} = \cdots = a_{j-1} = 0$. Let A be the polynomial $a_0 + a_1 x + \cdots + a_n x^n$. The *number of coefficient sign variations* in A, $\mathrm{var}(A)$, is $\mathrm{var}(a)$.

Theorem 2.2. [Descartes Rule of Signs] *For any non-zero real polynomial the number of coefficient sign variations exceeds the number of positive real roots—counting multiplicities—by a non-negative, even integer.*

Proof. Let $A(x)$ be a non-zero real polynomial. If x^k is the highest power of x that divides A, the polynomial A/x^k has the same number of coefficient sign variations and positive real roots as A, and its constant term is non-zero. Hence, we may assume that the constant term of A is non-zero. Let a_0 be this constant term, let n be the degree of A, and let a_n be the leading coefficient. Let $v = \mathrm{var}(A)$, and let p be the number of positive real roots of A, counting multiplicities.

To show that v and p have the same parity we use an argument given by Conkwright [13]. Let $z_1, \ldots, z_n \in \mathbb{C}$ be the roots of A. Then

$$A(x) = a_n(x - z_1)\cdots(x - z_n), \tag{2.1}$$

and hence $a_0 = A(0) = (-1)^n a_n z_1 \ldots z_n$. Since the non-real roots occur in complex conjugate pairs, their product is positive. The product of the positive roots is likewise positive, no root is zero since a_0 is non-zero, and the product of the negative real roots has the sign $(-1)^{n-p}$. It follows that the sign of a_0/a_n is $(-1)^p$. Hence v and p have the same parity.

Gauss [18] proves $v \ge p$ by showing that, for any non-zero real polynomial $B(x)$ and any positive real number a,

$$\mathrm{var}(B) < \mathrm{var}((x - a) \cdot B). \tag{2.2}$$

So, in equation (2.1), every positive root of A contributes at least one sign variation.

To show inequality (2.2) let $B = b_m x^m + \cdots + b_0$, let $a > 0$, and let $C = (x-a)B = c_{m+1}x^{m+1} + \cdots + c_0$. If $\operatorname{var}(B) > 0$ let (i,j) be an index pair that contributes to $\operatorname{var}(B)$. Then $0 \le i < j \le m$ and $b_i b_j < 0$ and either $j = i+1$ or $b_{i+1} = 0$. If $\sigma : \mathbb{R} \longrightarrow \{-1, 0, 1\}$ denotes the sign function then

$$\sigma(c_{i+1}) = \sigma(b_i - ab_{i+1}) = \sigma(b_i).$$

So, if $(i_1, j_1), \ldots, (i_k, j_k)$ are all the index pairs that contribute to $\operatorname{var}(B)$, and if $0 \le i_1 < j_1 \le \cdots \le i_k < j_k \le m$, then

$$\operatorname{var}(c_{i_1+1}, \ldots, c_{i_k+1}, c_{m+1}) = \operatorname{var}(b_{i_1}, \ldots, b_{i_k}, b_m) = \operatorname{var}(B).$$

Now let i be the smallest index for which $b_i \neq 0$. Then $0 \le i \le i_1$ and $\sigma(c_i) = \sigma(-ab_i) = -\sigma(b_i) = -\sigma(b_{i_1}) = -\sigma(c_{i_1+1})$, and so

$$\operatorname{var}(C) \ge \operatorname{var}(c_i, c_{i_1+1}, \ldots, c_{i_k+1}, c_{m+1}) = 1 + \operatorname{var}(B).$$

If $\operatorname{var}(B) = 0$ then $\operatorname{var}(C) \ge \operatorname{var}(c_i, c_{m+1}) = \operatorname{var}(-ab_i, b_m) \ge 1$. $\square$

Theorem 2.2 is named after Descartes although he merely stated that there can be as many positive real roots as there are coefficient sign variations [17]. The assertion that there are at least as many sign variations as there are positive roots was first stated and proved by Gauss [5]. Some modern authors [1, 37] seem to be unaware of Gauss's contribution.

Theorem 2.3. *Let A be a non-zero real polynomial. If $\operatorname{var}(A) = 0$ then A does not have any positive real root; if $\operatorname{var}(A) = 1$ then A has exactly one positive real root.*

Definition 2.4. Let S be a subring of $\mathbb{R}$ with $1 \in S$. We define three polynomial transformations $S[x] \longrightarrow S[x]$. Let $A = a_n x^n + \cdots + a_1 x + a_0$ be an element of $S[x]$.

(1) The *homothetic transformation* of A is the polynomial

$$H(A) = a_n x^n + 2a_{n-1}x^{n-1} + \cdots + 2^{n-1}a_1 x + 2^n a_0.$$

(2) The *Taylor shift by 1* of A is the polynomial

$$T(A) = b_n x^n + \cdots + b_1 x + b_0$$

where $b_k = \sum_{j=k}^{n} \binom{j}{k} a_j$ for $k \in \{0, \ldots, n\}$.

(3) The *reciprocal transformation* of A is the polynomial

$$R(A) = a_0 x^n + \cdots + a_{n-1}x + a_n.$$

Note that $R(A) = 0$ if and only if $A = 0$, and that $x \mid A$ implies $R(A) = R(A/x)$.

The Descartes method can now be stated as Algorithm 1.

To show that Algorithm 1 is partially correct we relate the roots of transformed real polynomials to the roots of the untransformed polynomials. Since we want to use bijective mappings we add the point ∞ to $\mathbb{C}$.

Algorithm 1 [Descartes method] This version is specialized to root counting in $I = (0,1)$. The algorithm can easily be modified to perform real root isolation.

```
int  roots_in_I (A ∈ S[x], A non-zero and squarefree, S ⊂ ℝ subring, 1 ∈ S)
     d ← var(TR(A));
     if d ≤ 1  return d;
     B ← H(A); C ← T(B);
     if x | C  m ← 1; else m ← 0;     Note: m = 1 if and only if A(1/2) = 0.
     return roots_in_I (B) + m + roots_in_I (C);
```

Definition 2.5. Let $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ be the Riemann sphere. We define three functions $\overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$.

$$h(z) = \begin{cases} z/2, & \text{if } z \in \mathbb{C}\,; \\ \infty, & \text{if } z = \infty. \end{cases}$$

$$t(z) = \begin{cases} z+1, & \text{if } z \in \mathbb{C}\,; \\ \infty, & \text{if } z = \infty. \end{cases}$$

$$r(z) = \begin{cases} 1/z, & \text{if } z \in \mathbb{C} - \{0\}\,; \\ \infty, & \text{if } z = 0\,; \\ 0, & \text{if } z = \infty. \end{cases}$$

The functions h, t, and r are elements of the group of *Möbius transformations*. These are all functions $\overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$ given by

$$z \longmapsto \frac{az+b}{cz+d} \tag{2.3}$$

with $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$. Anderson [4] explains how formula (2.3) handles division by 0 and evaluation at ∞; he also discusses the properties of Möbius transformations.

Remark 2.6. Let $A \in \mathbb{R}[x]$, and let $n = \deg(A)$; we adopt the convention that $\deg(0) = 0$ and $\mathrm{ldcf}(0) = 0$. Then, for all $z \in \mathbb{C}$,

$$\begin{aligned} H(A)(z) &= 2^n A(h(z)), \\ T(A)(z) &= A(t(z)), \\ R(A)(z) &= \begin{cases} z^n A(r(z)), & \text{if } z \neq 0\,; \\ \mathrm{ldcf}(A), & \text{if } z = 0\,. \end{cases} \end{aligned}$$

So, for all $z \in \mathbb{C}$,

$$\begin{aligned} TH(A)(z) &= 2^n A((h \circ t)(z)), \\ TR(A)(z) &= \begin{cases} (t(z))^n A((r \circ t)(z)), & \text{if } z \neq -1\,; \\ \mathrm{ldcf}(A), & \text{if } z = -1\,. \end{cases} \end{aligned}$$

Remark 2.7. By Remark 2.6, the following statements hold for all polynomials $A \in \mathbb{R}[x]$.

(1) The function h maps the roots of $H(A)$ one-to-one onto the roots of A; in particular, the roots of $H(A)$ in $(0,1)$ correspond to the roots of A in $(0,1/2)$.
(2) The function t maps the roots of $T(A)$ one-to-one onto the roots of A.
(3) The function r maps the non-zero roots of $R(A)$ one-to-one onto the non-zero roots of A; the roots of $R(A)$ are non-zero unless $A=0$.
(4) The function $h\circ t$ maps the roots of $TH(A)$ one-to-one onto the roots of A; in particular, the roots of $TH(A)$ in $(0,1)$ correspond to the roots of A in $(1/2,1)$.
(5) The function $r\circ t$ maps those roots of $TR(A)$ that are different from -1 one-to-one onto the non-zero roots of A; the roots of $TR(A)$ are different from -1 unless $A=0$. The positive real roots of $TR(A)$ correspond to the roots of A in $(0,1)$.

Observations (1), (4), and (5) of Remark 2.7 combined with Theorem 2.3 prove the partial correctness of Algorithm 1.

3. Ostrowski's Theory

Definition 3.1. A power series

$$\sum_{k=-\infty}^{+\infty} a_k z^k$$

with non-negative real coefficients is *normal* [31] if

(1) $a_k^2 \geq a_{k-1}a_{k+1}$ for all indices k, and
(2) $a_h>0$ and $a_j>0$ for indices $h<j$ implies $a_{h+1},\dots,a_{j-1}>0$.

In 1950, Ostrowski linked the normality of a polynomial and the Descartes rule. He stated his result [29, Lemma 1] for polynomials all of whose coefficients are positive. Generalizing slightly we show in Theorem 3.3 that it suffices to require that the leading coefficient be positive.

Definition 3.2. A polynomial with real coefficients is *positive* if its leading coefficient is positive.

Theorem 3.3. *A positive polynomial $A(x)$ is normal if and only if* $\operatorname{var}((x-\alpha)A(x))=1$ *for all positive real numbers α.*

Proof. (i) Let $A(x)$ be positive and normal, and let α be a positive real number. There is a non-negative integer m such that $A(x)=B(x)\cdot x^m$ where $B(x)$ is normal and all the coefficients of $B(x)$ are positive. Let $B(x)=b_nx^n+\cdots+b_1x+b_0$. Then

$$\frac{b_{n-1}}{b_n} \geq \frac{b_{n-2}}{b_{n-1}} \geq \cdots \geq \frac{b_0}{b_1}$$

and hence

$$\frac{b_{n-1}}{b_n}-\alpha \geq \frac{b_{n-2}}{b_{n-1}}-\alpha \geq \cdots \geq \frac{b_0}{b_1}-\alpha.$$

Since also $b_n>0$ and $-\alpha b_0<0$, the polynomial

$$(x-\alpha)B(x) = b_nx^{n+1} + b_n\Big(\frac{b_{n-1}}{b_n}-\alpha\Big)x^n + \cdots + b_1\Big(\frac{b_0}{b_1}-\alpha\Big)x - \alpha b_0$$

has exactly 1 coefficient sign variation. And so,

$$1 = \mathrm{var}((x-\alpha)B(x)) = \mathrm{var}((x-\alpha)B(x)\cdot x^m) = \mathrm{var}((x-\alpha)A(x)).$$

(ii) Conversely, let $A(x)$ be positive but not normal. There is a non-negative integer m such that $A = B(x)\cdot x^m$ where $B(x)$ has a non-zero constant term. Moreover, the polynomial $B(x)$ is positive and not normal—and hence non-constant. For any real number α let $C^{(\alpha)}(x) = (x-\alpha)B(x)$. Then $\mathrm{var}((x-\alpha)A(x)) = \mathrm{var}(C^{(\alpha)}(x))$, and it suffices to find a positive number α such that $\mathrm{var}(C^{(\alpha)}(x)) \neq 1$.

Let $B(x) = b_n x^n + \cdots + b_1 x + b_0$. Then $n \geq 1$ and $b_n > 0$ and $b_0 \neq 0$. Let $C^{(\alpha)}(x) = c_{n+1}^{(\alpha)} x^{n+1} + \cdots + c_1^{(\alpha)} x + c_0^{(\alpha)}$. Then $c_0^{(\alpha)} = -\alpha b_0$, $c_k^{(\alpha)} = b_{k-1} - \alpha b_k$ for $1 \leq k \leq n$, and $c_{n+1}^{(\alpha)} = b_n$.

If $\mathrm{var}(B(x)) \geq 2$ choose α so small that, for all k with $1 \leq k \leq n$, the signs of $c_k^{(\alpha)}$ and b_{k-1} are equal whenever $b_{k-1} \neq 0$; then $\mathrm{var}(C^{(\alpha)}(x)) \geq \mathrm{var}(B(x)) \geq 2$.

If $\mathrm{var}(B(x)) = 1$ the polynomial $B(x)$ has exactly one positive real root by the Descartes rule. So, for any $\alpha > 0$, the polynomial $C^{(\alpha)}(x)$ has two positive real roots, and, again by the Descartes rule, $\mathrm{var}(C^{(\alpha)}(x)) \geq 2$.

Finally, assume $\mathrm{var}(B(x)) = 0$. Then, since $b_n > 0$, all the coefficients of $B(x)$ are non-negative. If all the coefficients of $B(x)$ are positive, then, since $B(x)$ is not normal, there is an index k with $1 \leq k \leq n-1$ such that $0 < b_k/b_{k+1} < b_{k-1}/b_k$. Choose α such that $b_k/b_{k+1} < \alpha < b_{k-1}/b_k$. Now $\alpha > 0$ and $c_{n+1}^{(\alpha)} = b_n > 0$, $c_{k+1}^{(\alpha)} = b_k - \alpha b_{k+1} < 0$ and $c_k^{(\alpha)} = b_{k-1} - \alpha b_k > 0$, and hence $\mathrm{var}(C^{(\alpha)}(x)) \geq 2$. If not all the coefficients of $B(x)$ are positive, there is a zero-coefficient. Let b_k be the zero-coefficient with the highest index; then $c_{k+1}^{(\alpha)} < 0$. Since $b_0 > 0$ there is an index $j < k$ such that $b_{j+1} = 0$ and $b_j > 0$; then $c_{j+1}^{(\alpha)} > 0$. Now $c_0^{(\alpha)} < 0$ implies $\mathrm{var}(C^{(\alpha)}(x)) \geq 2$ also in this case. □

By Theorem 3.3, the Descartes rule will reveal the existence of a single positive root of a positive polynomial if the other—possibly non-real—roots $\alpha_1, \ldots, \alpha_{n-1}$ are such that

$$(x-\alpha_1)\cdots(x-\alpha_{n-1})$$

is a normal polynomial.

Theorem 3.4. *A positive linear polynomial is normal if and only if its root is negative or zero.*

Proof. Let A be a positive linear polynomial, and let $\alpha \in \mathbb{R}$ be its root. Then there is a positive real number a such that $A(x) = a(x-\alpha) = ax - a\alpha$. Now A is normal if and only if $-a\alpha \geq 0$, that is, if and only if $\alpha \leq 0$. □

Definition 3.5. Let

$$\mathcal{C} = \left\{ a + ib \,\middle|\, a \leq 0 \text{ and } |b| \leq |a|\sqrt{3} \right\}.$$

For an illustration see Figure 4.1(a); the cone contains its bordering rays and the vertex 0.

Theorem 3.6. *A positive quadratic polynomial is normal if and only if its roots are elements of the cone $\mathcal{C}$.*

Proof. Let A be a positive quadratic polynomial, and let $c > 0$ be its leading coefficient.

If the roots of A are complex conjugates $a + ib$ and $a - ib$ with real numbers a, b then $A(x) = c(x - (a + ib))(x - (a - ib))$. Now $A(x) = cx^2 - 2acx + c\left(a^2 + b^2\right)$ is normal if and only if $-2ac \geq 0$ and $c(a^2 + b^2) \geq 0$ and $(-2ac)^2 \geq c \cdot c(a^2 + b^2)$, that is, if and only if $a \leq 0$ and $4a^2 \geq a^2 + b^2$, or, equivalently, if and only if $a \pm ib \in \mathcal{C}$.

Otherwise, the roots of A are real numbers α and β, and we have $A(x) = c(x-\alpha)(x-\beta) = cx^2 - c(\alpha+\beta)x + c\alpha\beta$. Now A is normal if and only if $-c(\alpha+\beta) \geq 0$ and $c\alpha\beta \geq 0$ and $(-c(\alpha + \beta))^2 \geq c \cdot c\alpha\beta$, that is, if and only if $\alpha + \beta \leq 0$ and $\alpha\beta \geq 0$ and $(\alpha + \beta)^2 \geq \alpha\beta$, or, equivalently, if and only if $\alpha, \beta \leq 0$. □

In Section 6 we will characterize normal cubic polynomials. The "if"-direction of Theorems 3.4 and 3.6 can be generalized to polynomials of any degree using an earlier result of Ostrowski. Ostrowski showed in 1939 that the product of two normal series, if it exists, is normal [31]. In 1950, he gave a simpler proof for the case of polynomials [29].

Theorem 3.7. *The product of two normal polynomials is normal.*

Proof. Let $A = \sum_{h=0}^{m} a_h x^h$ and $B = \sum_{j=0}^{n} b_j x^j$ be normal polynomials. Any normal polynomial can be written as $P \cdot x^k$ where k is a non-negative integer and P is a normal polynomial and all the coefficients of P are positive. Hence it suffices to consider the case where all the coefficients of A and B are positive.

Let $C = A \cdot B = \sum_{k=0}^{m+n} c_k x^k$. Write $c_k = \sum_h a_h b_{k-h}$ where h and k range over the set of all integers and all a_h with $h \notin \{0, \ldots, m\}$, all b_j with $j \notin \{0, \ldots, n\}$, and all c_k with $k \notin \{0, \ldots, m + n\}$ are taken as zero. Clearly, all the coefficients of C are positive; it remains to show that $c_k^2 - c_{k-1}c_{k+1} \geq 0$ for all k.

Using the following decomposition of the set of summation indices

$$\left\{(h, j) \in \mathbb{Z}^2 \;\middle|\; h > j\right\} = \left\{(j + 1, h - 1) \in \mathbb{Z}^2 \;\middle|\; h \leq j\right\} \cup \left\{(h, h - 1) \in \mathbb{Z}^2\right\}$$

we obtain, for any index k,

$$\begin{aligned}
&c_k^2 - c_{k-1}c_{k+1} \\
&= \sum_{h \leq j} a_h a_j b_{k-h} b_{k-j} + \sum_{h > j} a_h a_j b_{k-h} b_{k-j} \\
&\quad - \sum_{h \leq j} a_h a_j b_{k-h+1} b_{k-j-1} - \sum_{h > j} a_h a_j b_{k-h+1} b_{k-j-1} \\
&= \sum_{h \leq j} a_h a_j b_{k-h} b_{k-j} + \sum_{h \leq j} a_{j+1} a_{h-1} b_{k-j-1} b_{k-h+1} + \sum_{h} a_h a_{h-1} b_{k-h} b_{k-h+1} \\
&\quad - \sum_{h \leq j} a_h a_j b_{k-h+1} b_{k-j-1} - \sum_{h \leq j} a_{j+1} a_{h-1} b_{k-j} b_{k-h} - \sum_{h} a_h a_{h-1} b_{k-h+1} b_{k-h} \\
&= \sum_{h \leq j} (a_h a_j - a_{h-1} a_{j+1})(b_{k-j} b_{k-h} - b_{k-j-1} b_{k-h+1}),
\end{aligned}$$

that is,

$$c_k^2 - c_{k-1}c_{k+1} = \sum_{h \leq j} (a_h a_j - a_{h-1} a_{j+1})(b_{k-j} b_{k-h} - b_{k-j-1} b_{k-h+1}). \tag{3.1}$$

Since A is normal and $a_0, \ldots, a_m$ are positive, one has

$$\frac{a_{m-1}}{a_m} \geq \frac{a_{m-2}}{a_{m-1}} \geq \cdots \geq \frac{a_0}{a_1},$$

and hence $a_h a_j - a_{h-1}a_{j+1} \geq 0$ for all $h \leq j$; the analogous statement holds for the coefficients of B. Hence each summand on the right hand side of equation (3.1) is non-negative, and thus $c_k^2 - c_{k-1}c_{k+1} \geq 0$ for all k. □

Theorem 3.8. *If the roots of a positive polynomial are in the cone $\mathcal{C}$ then the polynomial is normal.*

Proof. Let A be a positive polynomial all of whose roots are elements of the cone $\mathcal{C}$. The complete factorization of A over the field of real numbers is a product of linear and quadratic factors. We may assume that all these factors are positive. Since all the roots are in the cone $\mathcal{C}$, Theorems 3.4 and 3.6 apply, and each factor is normal. Thus, by Theorem 3.7, the polynomial A is normal. □

Of all the theorems in this section, we will invoke only Theorem 3.9 in Sections 4 and 5.

Theorem 3.9. *If the roots of a non-zero polynomial $A(x)$ are in the cone $\mathcal{C}$ then* $\operatorname{var}((x-\alpha)A(x)) = 1$ *for all positive real numbers α.*

Proof. Theorems 3.8 and 3.3. □

4. Three Circles

By Theorem 3.9, Algorithm 1 will stop calling itself when it encounters a polynomial $TR(A)$ that has exactly one positive root and whose other roots are elements of the cone $\mathcal{C}$. We want to state this condition in terms of the roots of the polynomial A. Since A is non-zero, Remark 2.7 (5) implies that the function $r \circ t$ maps the roots of $TR(A)$ one-to-one onto the non-zero roots of A. But much more is true since $r \circ t$ is a Möbius transformation.

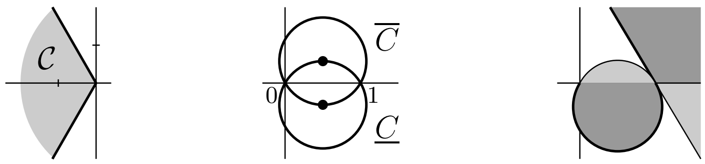
Remark 4.1. Anderson [4] reviews some properties of Möbius transformations. These transformations are homeomorphisms of the Riemann sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ that map circles in $\overline{\mathbb{C}}$ to circles. In particular, circles and lines in $\mathbb{C}$ are mapped to circles and lines. To identify the image of a given circle or line K under a given Möbius transformation it suffices to select three distinct points on K, to compute their images under the transformation, and to determine the unique circle or line L that contains those images. The sets $\mathbb{C} - K$ and $\mathbb{C} - L$ each have exactly two connected components. Each component of $\mathbb{C} - K$ is mapped to a different component of $\mathbb{C} - L$ since Möbius transformations are homeomorphisms of $\overline{\mathbb{C}}$. By applying the transformation to a single point in $\mathbb{C} - K$ one can determine the image of each component of $\mathbb{C} - K$.

Definition 4.2. We define three circular disks.

$$\begin{aligned}
\underline{C} &= \left\{ z \in \mathbb{C} \;\middle|\; \left|z - \left(1/2 - i\sqrt{3}/6\right)\right| < \sqrt{3}/3 \right\}, \\
\overline{C} &= \left\{ z \in \mathbb{C} \;\middle|\; \left|z - \left(1/2 + i\sqrt{3}/6\right)\right| < \sqrt{3}/3 \right\}, \\
C &= \left\{ z \in \mathbb{C} \;\middle|\; |z - 1/2| < 1/2 \right\}.
\end{aligned}$$

Remark 4.3. The Möbius transformation $r \circ t$ maps the cone $\mathcal{C}$ one-to-one onto $\overline{\mathbb{C}} - (\underline{C} \cup \overline{C})$ and the half-plane $\{z \in \mathbb{C} \,|\, \mathrm{Re}(z) \leq 0\}$ one-to-one onto $\mathbb{C} - C$. Both statements can be verified using the method described in Remark 4.1.

Figure 4.1(a) shows the cone $\mathcal{C}$. Figure 4.1(b) shows the boundaries of the open disks $\underline{C}$ and $\overline{C}$. Figure 4.1(c) shows how the Möbius transformation r operates on the boundary of $\underline{C}$. If z traverses the boundary of $\underline{C}$ clockwise from 1 towards 0, the reciprocal $r(z)$ traverses the ray $\{1 - s + \sqrt{3}si \,|\, s \geq 0\}$ upwards starting at 1. Similarly, if z traverses the boundary of $\overline{C}$ counterclockwise from 1 towards 0, the reciprocal $r(z)$ traverses the ray $\{1 - s - \sqrt{3}si \,|\, s \geq 0\}$ downwards starting at 1. The point $z = 0$ is mapped to $r(0) = \infty \notin \mathbb{C}$. Thus the figure illustrates how the function $t^{-1} \circ r = (r \circ t)^{-1}$ maps $\overline{\mathbb{C}} - (\underline{C} \cup \overline{C})$ one-to-one onto $\mathcal{C}$.



(a) (b) (c)

FIGURE 4.1. (a) A positive quadratic polynomial is normal if and only if its roots are in the cone $\mathcal{C}$. (b) If a polynomial A has a simple root in the interval $(0, 1)$ and no other real or non-real roots in $\underline{C} \cup \overline{C}$ then $\mathrm{var}(TR(A)) = 1$. (c) The image of $\underline{C}$ under r.

Theorem 4.4. [Two–Circle Theorem] *Let A be a real polynomial with a single, simple root in the interval $(0, 1)$ and no other real or non-real roots in the open disks $\underline{C}$ and $\overline{C}$. Then $\mathrm{var}(TR(A)) = 1$.*

Proof. Let A be as described. Then $A \neq 0$ and, by Remark 2.7 (5), the roots of $B = TR(A)$ are all different from -1. Therefore, the function $(r \circ t)^{-1}$ maps the non-zero roots of A one-to-one onto the roots of B. Hence, B has a single, simple root in $(r \circ t)^{-1}((0, 1)) = (0, \infty)$, and its other roots are in $(r \circ t)^{-1}(\overline{\mathbb{C}} - (\underline{C} \cup \overline{C}))$ which equals $\mathcal{C}$ by Remark 4.3. Now Theorem 3.9 yields $\mathrm{var}(B) = 1$. □

The two-circle condition is not necessary for the termination of the Descartes method. Indeed, the polynomial $A = 32x^3 - 16x^2 + 2x - 1$ has the single, simple root $1/2$ in the interval $(0, 1)$, the pair of complex conjugate roots $\pm i/4$ inside the open disks $\underline{C}$ and $\overline{C}$, and $\mathrm{var}(TR(A)) = 1$.

Our two-circle theorem improves upon a two-circle theorem of Collins and Johnson [6]. They use the disks $D_1 = \{z \in \mathbb{C} \mid |z| < 1\}$ and $D_2 = \{z \in \mathbb{C} \mid |z - 1| < 1\}$ instead of $\underline{C}$ and $\overline{C}$. But $\underline{C} \cup \overline{C}$ is a proper subset of $D_1 \cup D_2$, and the area of $\underline{C} \cup \overline{C}$ is exactly one-third of the area of $D_1 \cup D_2$. Indeed, the Möbius transformation

$$z \longmapsto \frac{i\sqrt{3}}{3} z + \left(\frac{1}{2} - i\frac{\sqrt{3}}{6}\right)$$

maps $D_1 \cup D_2$ onto $\underline{C} \cup \overline{C}$.

The following well-known theorem completes our converse of Theorem 2.3.

Theorem 4.5. *If a polynomial A does not have any roots in the open disk C then $\mathrm{var}(TR(A)) = 0$.*

Proof. Let A be as described. Then $A \neq 0$ and, by Remark 2.7 (5), the roots of $B = TR(A)$ are all different from -1. Therefore, the function $(r \circ t)^{-1}$ maps the non-zero roots of A one-to-one onto the roots of B. But since the roots of A are all in $\mathbb{C} - C$, the roots of B have non-positive real parts by Remark 4.3. Hence, in the decomposition of B into a product of a constant and monic linear and quadratic factors, every linear factor is of the form $x - \alpha$ where $\alpha \leq 0$, and every quadratic factor is of the form $(x-(a+ib))(x-(a-ib)) = x^2 - 2ax + (a^2+b^2)$ where $a \leq 0$. Since all the non-zero coefficients of all the linear and quadratic factors of B have the same sign, the non-zero coefficients of B all have the same sign. $\square$

When we bound the recursion depth of the Descartes method we will use Theorem 4.6 which summarizes the preceding results.

Theorem 4.6. *Let A be a real polynomial with* $\operatorname{var}(TR(A)) \geq 2$. *Then either the open disk C contains at least two roots of A, or the interval $(0,1)$ contains exactly one real root and the union of the open disks $\underline{C}$ and $\overline{C}$ contains a pair of complex conjugate roots.*

Proof. If A has no root in C then $\operatorname{var}(TR(A)) = 0$ by Theorem 4.5. Thus, A has at least one root in C. If this is the only root in C, the root is real and it is, in fact, the only real root in the interval $(0,1)$. Then $\underline{C} \cup \overline{C}$ must contain a pair of complex conjugate roots because otherwise $\operatorname{var}(TR(A)) = 1$ by Theorem 4.4. $\square$

5. Bounds for the Recursion Tree

For any input polynomial A the recursion tree of Algorithm 1 is a full binary tree; Figure 5.1 shows an example. With every node of the tree we associate a pair (B, I) consisting of a polynomial B and an interval I. With the root of the tree we associate the pair $(A, (0,1))$. If an internal node is associated with the pair (B, I) we associate one child with the pair (B_L, I_L) where $B_L = H(B)$ and I_L is the open left half of I, and we associate the other child with the pair (B_R, I_R) where $B_R = TH(B)$ and I_R is the open right half of I.

Remark 5.1. By Remarks 2.7 (1) and (4), the function h maps the roots of B_L in $(0,1)$ onto the roots of B in I_L, and the function $h \circ t$ maps the roots of B_R in $(0,1)$ onto the roots of B in I_R. Thus, there is a sequence of elements of $\{h, t\}$ whose composition m maps the roots of B in $(0,1)$ onto the roots of the input polynomial A in I. When m maps the interval $(0,1)$ onto the interval I it transforms at the same time the disks C, $\underline{C}$ and $\overline{C}$ of Section 4. These disks are the circumscribing disks of isosceles triangles with base $(0,1)$ and base angles $45°$, $-60°$ and $60°$, respectively, as shown in Figure 5.1. But h, t and, hence, m are Möbius transformations and thus preserve angles [4]. Moreover, the transformations h, t and, hence, m map straight lines in $\mathbb{C}$ onto straight lines in $\mathbb{C}$ and circles in $\mathbb{C}$ onto circles in $\mathbb{C}$. Therefore, the images $m(C)$, $m(\underline{C})$ and $m(\overline{C})$ are the circumscribing disks of the isosceles triangles with base I and base angles $45°$, $-60°$ and $60°$, respectively. Figure 5.1 shows the disks that are considered at the leaf nodes of a particular recursion tree.

The depth of the recursion tree can be bounded using Mahler's root separation theorem [25]. To obtain a bound that also covers the width of the tree we use Davenport's generalization [14] of Mahler's theorem in a form due to Johnson [21].

Definition 5.2. Let $A = a_n x^n + \cdots + a_1 x + a_0$ be a non-zero polynomial of degree n with complex coefficients and the complex roots $\alpha_1, \ldots, \alpha_n$. The *Euclidean norm*

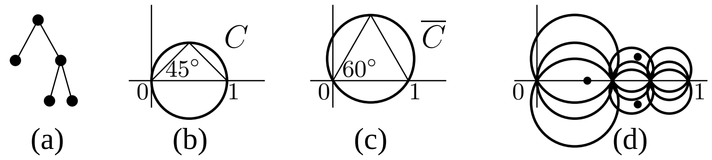


FIGURE 5.1. (a) Recursion tree for $A = 27648x^3 - 46080x^2 + 25251x - 4321$. (b),(c) Triangles with circumscribing disks C, $\overline{C}$. (d) Circumscribing disks for the intervals at the leaf nodes of the tree in (a). Also shown are $1/3$ and $2/3 \pm i \cdot 5/32$, the roots of A.

of A is $|A|_2 = (a_n^2 + \cdots + a_0^2)^{1/2}$, the *measure* of A is $M(A) = |a_n| \cdot \prod_{i=1}^n \max(1, |\alpha_i|)$, and the *discriminant* of A is $D(A) = a_n^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2$.

Remark 5.3. A theorem of Landau [24] implies $M(A) \leq |A|_2$. The inequality was independently rediscovered more than once. Ostrowski [30] summarizes its history until 1961 and proves a generalization. Mignotte [27, 28] gives a short elementary proof. The discriminant $D(A)$ is known [35] to be a polynomial in the coefficients of A; hence $D(A) \geq 1$ if A is a squarefree integer polynomial.

Theorem 5.4. *Let A be a non-zero complex polynomial of degree n with the roots $\alpha_1, \ldots, \alpha_n$. Let k be an integer, $1 \leq k \leq n$, and let $(\beta_1, \ldots, \beta_k)$ be a sequence of roots of A such that*

$$\beta_i \notin \{\alpha_1, \ldots, \alpha_i\} \quad \text{and} \quad |\beta_i| \leq |\alpha_i| \quad \text{for all} \quad i \in \{1, \ldots, k\}.$$

Then

$$\prod_{i=1}^k |\alpha_i - \beta_i| \quad \geq \quad 3^{k/2} D(A)^{1/2} M(A)^{-n+1} n^{-k-n/2}.$$

Proof. See [21]. □

Theorem 5.5. *Let A be a non-zero real polynomial of degree n, measure M, and discriminant D. Let the integers $h \geq 0$ and $k \geq 1$ be such that k is the number of internal nodes of depth h in the recursion tree of Algorithm 1 with input A where* depth *is the distance from the root. Then*

(1) *$k \leq n$, and*
(2) *$2^{(1-h)k} > 3^k D^{1/2} M^{-n+1} n^{-k-n/2}$.*

Proof. Let $I_1 < \ldots < I_k$ be the open subintervals of $(0, 1)$ that are associated with the internal nodes of depth h, and let $A_1, \ldots, A_k$ be the corresponding polynomials. The intervals have width 2^{-h}. For every index $i \in \{1, \ldots, k\}$ let C_i, $\underline{C}_i$ and $\overline{C}_i$ be the circumscribing disks of the isosceles triangles with base I_i and base angles 45°, −60° and 60°, respectively. By Remark 5.1 the roots of A_i in the disks C, $\underline{C}$ and $\overline{C}$, correspond, respectively, to the roots of A in the disks C_i, $\underline{C}_i$ and $\overline{C}_i$. But the polynomials A_i are at internal nodes of the recursion tree, so $\mathrm{var}(TR(A_i)) \geq 2$, and hence, by Theorem 4.6, either C_i contains at least two roots of A, or I_i contains exactly one real root of A and $\underline{C}_i \cup \overline{C}_i$ contains a pair of complex conjugate roots of A.

Assertion (1) holds since every disk C_i contains at least one root of A, and the disks $C_1, \ldots, C_k$ are pairwise disjoint.

Assertion (2) holds if A has a multiple root since $D = 0$ in that case. If all roots are simple, define, for every index $i \in \{1, \ldots, k\}$, a set R_i of roots of A in $\underline{C}_i \cup \overline{C}_i$. If C_i contains at least two roots of A, let $R_i = \{s, t\}$ where s and t are either two arbitrary distinct real roots in I_i or two arbitrary non-real complex conjugate roots in C_i; otherwise, let $R_i = \{r, s, t\}$ where r is the unique real root in I_i, and s and t are two arbitrary non-real complex conjugate roots in $\underline{C}_i \cup \overline{C}_i$. For notational convenience let $R_0 = R_{k+1} = \emptyset$. Note that, for all distinct indices $i, j \in \{1, \ldots, k\}$, the intersection of R_i and R_j is either empty or it consists of two non-real complex conjugate roots and $j = i - 1$ or $j = i + 1$. Moreover, if $R_i \cap R_{i+1} \neq \emptyset$ then $R_{i-1} \cap R_i = \emptyset$ and $R_{i+1} \cap R_{i+2} = \emptyset$. So, for all indices $i \in \{1, \ldots, k\}$, the set R_i is either disjoint from all sets R_j, $j \neq i$, or there is exactly one set R_j such that $j \neq i$ and $R_i \cap R_j \neq \emptyset$.

Let $i \in \{1, \ldots, k\}$. If R_i is disjoint from all sets R_j, $j \neq i$, select two distinct elements from R_i that are both in C_i or both in $\underline{C}_i$ or both in $\overline{C}_i$, and label them α_i and β_i so that $|\beta_i| \leq |\alpha_i|$. If there is exactly one set R_j such that $j \neq i$ and $R_i \cap R_j \neq \emptyset$ then select $\alpha_i, \beta_i, \alpha_j, \beta_j \in R_i \cup R_j$ as described in Figure 5.2 for the case $j = i+1$. Since $R_i \cap R_j \neq \emptyset$, at least one of the sets R_i and R_j has 3 elements, and the figure shows how the roots are selected depending on whether only R_i has 3 elements or only R_j or both R_i and R_j.

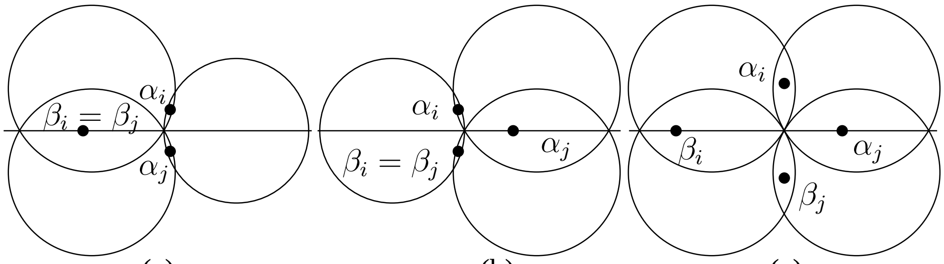


FIGURE 5.2. Adjacent intervals with coinciding roots. Here, $j = i + 1$. (a) $|R_i| = 3$ and $|R_j| = 2$. Note that $|\beta_i| \leq |\alpha_i|$ and $|\beta_j| \leq |\alpha_j|$ and $\alpha_i, \beta_i \in \overline{C}_i$ and $\alpha_j, \beta_j \in \underline{C}_i$. (b) $|R_i| = 2$ and $|R_j| = 3$. (c) $|R_i| = 3$ and $|R_j| = 3$. In $\overline{C}_i$ the root with the smaller modulus is labeled β_i and the other root α_i; likewise for $\underline{C}_j$, β_j and α_j.

By construction, the selected roots $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_k$ not only satisfy $\beta_i \notin \{\alpha_1, \ldots, \alpha_i\}$ and $|\beta_i| \leq |\alpha_i|$ for all $i \in \{1, \ldots, k\}$ but also, for all $i \in \{1, \ldots, k\}$, both roots α_i and β_i are in one of the disks C_i, $\underline{C}_i$, $\overline{C}_i$, or, if $i > 1$, in the disk $\overline{C}_{i-1}$, so $|\alpha_i - \beta_i| < 2^{1-h}/\sqrt{3}$. Now Theorem 5.4 implies

$$2^{(1-h)k} 3^{-k/2} > \prod_{i=1}^{k} |\alpha_i - \beta_i| \geq 3^{k/2} D^{1/2} M^{-n+1} n^{-k-n/2}.$$

□

Theorem 5.6. *Let A be a non-zero squarefree integer polynomial of degree $n \geq 2$ with Euclidean norm d. Let h and k be as in Theorem 5.5, and let $\log = \log_2$. Then*

(1) *$k \leq n$, and*

(2) $(h-1)k < (n-1)\log d + (k+n/2)\log n - k\log 3$, *and*
(3) $h \leq (n-1)\log d + (n/2+1)\log n - \log 3$.

Proof. Assertion (1) holds due to assertion (1) of Theorem 5.5. To show assertion (2), consider assertion (2) of Theorem 5.5, apply Remark 5.3, take logarithms, and multiply by -1. To show assertion (3), consider assertion (2) and collect all terms involving k on one side to obtain $k(h-1-\log n+\log 3) < (n-1)\log d + n/2\log n$. If $h-1-\log n+\log 3 < 0$ then assertion (3) clearly holds. If, on the other hand, $h-1-\log n+\log 3 \geq 0$ then $k \geq 1$ implies $h-1-\log n+\log 3 < (n-1)\log d + n/2\log n$, and hence assertion (3) holds also in this case. □

Remark 5.7. Theorem 5.6 is stronger than an earlier result of Krandick [23, Satz 47], and the proof is shorter. The theorem implies the dominance relations $hk \preceq n\log(nd)$ and $h \preceq n\log(nd)$ which can be used in an asymptotic computing time analysis of Algorithm 1 when the ring S of coefficients is $\mathbb{Z}$; the notation $\preceq$ is due to Collins [8].

6. Normal Cubics

By Theorem 3.8 any positive polynomial whose roots are in the cone $\mathcal{C}$ is normal. By Theorems 3.4 and 3.6 the converse holds for linear and quadratic polynomials. For cubic polynomials, however, the converse is false. Indeed, the normal polynomial $x^3+5x^2+16x+30$ has roots $-1\pm 3i \notin \mathcal{C}$. Theorems 6.1 and 6.2 together completely characterize the normal cubic polynomials.

Theorem 6.1. *Let A be a positive polynomial all of whose roots are real. Then A is normal if and only if the roots are in the cone $\mathcal{C}$.*

Proof. If the roots of A are in the cone $\mathcal{C}$ then Theorem 3.8 implies that A is normal. Otherwise, A has a positive root. In this case, $\mathrm{var}((x-1)A(x)) > 1$ by Theorem 2.2, and A is not normal by Theorem 3.3. □

Theorem 6.2. *Let A be a positive cubic polynomial whose roots are a and $b\pm ic$ where a,b,c are real numbers. Then A is normal if and only if*

$$a \leq 0 \quad \textit{and} \tag{6.1}$$
$$b \leq 0 \quad \textit{and} \tag{6.2}$$
$$c^2-3b^2-2ab-a^2 \leq 0 \quad \textit{and} \tag{6.3}$$
$$c^4+2b^2c^2+2abc^2-a^2c^2+b^4+2ab^3+3a^2b^2 \geq 0. \tag{6.4}$$

Proof. We may assume that A is monic since A is normal if and only if $A/\mathrm{ldcf}(A)$ is normal. Hence,

$$A = (x-a)\cdot(x-(b+ic))\cdot(x-(b-ic))$$

and thus

$$A = x^3 + a_2x^2 + a_1x + a_0$$

where

$$\begin{aligned} a_2 &= -a-2b, \\ a_1 &= 2ab+b^2+c^2, \\ a_0 &= -ab^2-ac^2. \end{aligned}$$

By definition, A is normal if and only if all of the following hold.

$$a_2 \geq 0, \tag{6.5}$$
$$a_1 \geq 0, \tag{6.6}$$
$$a_0 \geq 0, \tag{6.7}$$
$$a_2^2 \geq a_1, \tag{6.8}$$
$$a_1^2 \geq a_2 a_0, \tag{6.9}$$
$$a_2 = 0 \Rightarrow a_1 = a_0 = 0, \tag{6.10}$$
$$a_1 = 0 \Rightarrow a_0 = 0. \tag{6.11}$$

Implication (6.11) is redundant since it follows from (6.9), (6.5) and (6.7). Also the implication ($a_2 = 0 \Rightarrow a_1 = 0$) in (6.10) is redundant since it follows from (6.8) and (6.6). We note the pairwise equivalence of (6.1) and (6.7), (6.3) and (6.8), and (6.4) and (6.9). We will show that the conjunction of (6.1)–(6.4) is equivalent to the conjunction of (6.5)–(6.11).

Assume (6.1)–(6.4). Clearly, (6.1) and (6.2) imply (6.5) and (6.6). The pairwise equivalences yield (6.7), (6.8) and (6.9). The implication ($a_2 = 0 \Rightarrow a_0 = 0$) in (6.10) holds since $a_2 = 0$ together with (6.1) and (6.2) implies $a = 0$.

Assume now (6.5)–(6.11). The pairwise equivalences yield (6.1), (6.3), and (6.4). To complete the proof we have to show (6.2). By (6.1) we have $a \leq 0$. If $a = 0$ then (6.2) follows from (6.5), so we may assume $a < 0$. Next observe that if (a, b, c) satisfies (6.5)–(6.11) then, for any $t > 0$, (ta, tb, tc) satisfies (6.5)–(6.11). So we may assume $a = -1$. Now (6.5) implies that $b \leq 1/2$, and we need to show that $b \leq 0$. Figure 6.1 illustrates the situation. If $b = 1/2$ then, by (6.5), $a_2 = 0$, hence, by (6.10), $a_0 = 0$, and thus $a = 0$, a contradiction. So, $b < 1/2$ and we need to show $b \leq 0$. Multiplying (6.3) and (6.6), and combining the result with (6.4) we obtain the inequalities

$$(c^2 - 3b^2 + 2b - 1)(-2b + b^2 + c^2) \leq 0 \leq c^4 + 2b^2c^2 + 2abc^2 - a^2c^2 + b^4 + 2ab^3 + 3a^2b^2.$$

Collecting all the terms on the left hand side and factoring yields

$$-2b(2b-1)((b-1)^2 + c^2) \leq 0,$$

so $0 < b < 1/2$ is impossible, and we have $b \leq 0$ as desired. □

Figure 6.1 supports the notion that Theorem 6.2 recognizes more normal cubics than Theorem 3.8. In an attempt to quantify the improvement we perform extensive experiments that use Algorithm 1.

Definition 6.3. The *max-norm* of a complex polynomial $A = a_n x^n + \cdots + a_1 x + a_0$ is $|A|_\infty = \max(|a_n|, \ldots, |a_0|)$.

Let m be a positive integer. The set of all normal cubic integer polynomials of max-norm m can be efficiently enumerated. For each such polynomial A,

$$A = a_3x^3 + a_2x^2 + a_1x + a_0,$$

we want to decide whether all of its roots are in the cone $\mathcal{C}$. Since A is cubic, either A has one real root and two non-real complex conjugate roots, or all the roots of A are real. In particular, if A has a multiple root then all the roots of A are real. Since all the coefficients of A are non-negative, all the real roots of A are non-positive and, hence, in $\mathcal{C}$. Using polynomial factorization and Algorithm 1 we thus reduce

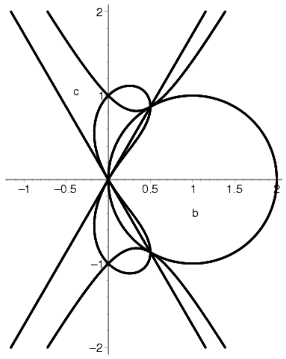


FIGURE 6.1. For $a = -1$ the points (b, c) that satisfy (6.1)–(6.4) are precisely the points in the left half-plane (6.2) between the two branches of the hyperbola (6.3) and outside of the figure "8" (6.4). For $a = 0$ the solution set coincides with the cone $\mathcal{C}$ which is delimited by the curve $c^2 - 3b^2 = 0$. The solutions of inequality (6.6) are precisely the points outside the circle.

the decision problem to the case where A is irreducible and has a single real root $\alpha \in \mathcal{C}$. The other roots of A are the roots of the polynomial

$$B = A(x)/(x - \alpha) = a_3x^2 + (a_3\alpha + a_2)x + (a_3\alpha^2 + a_2\alpha + a_1).$$

By Theorem 3.6, these roots are in $\mathcal{C}$ if and only if B is normal. We decide the latter by performing arithmetic in $\mathbb{Z}[\alpha]$ on the coefficients of B.

The computing time of the decision method can be reduced by a factor of about 3.5 by using floating point computations instead of exact arithmetic. Indeed, we use the floating point interval arithmetic techniques described by Collins, Johnson, and Krandick [12], and we fall back to exact arithmetic just in case the floating point results are inconclusive. In our experiments we represent α by an isolating interval of width 2^{-40}, and we use IEEE-double precision arithmetic [20]. For all our inputs, the floating point method is inconclusive only in case the roots of B lie on the boundary of $\mathcal{C}$; this situation occurs when B is normal and $(a_3\alpha + a_2)^2 = a_3 \cdot (a_3\alpha^2 + a_2\alpha + a_1)$.

Table 1 shows that only about 57 percent of the $2,353,361,850$ normal cubic polynomials we examined have all of their roots in the cone $\mathcal{C}$. It seems reasonable to expect smaller ratios when the experiment is carried out for polynomials of higher degrees. The table also shows that we had to use exact arithmetic for relatively few polynomials.

We can now generalize Theorem 3.9.

Theorem 6.4. *Let $A(x)$ be a non-zero polynomial such that $A(x) = B(x) \cdot C(x)$ where all the roots of $B(x)$ are in the cone $\mathcal{C}$ and $C(x)$ is a product of cubic polynomials each of whose roots are as described in Theorem 6.2 then*

$$\operatorname{var}((x - \alpha)A(x)) = 1 \quad \textit{for all real} \quad \alpha > 0.$$

Proof. Theorems 6.1, 6.2, 3.7, and 3.3. □

It is easy to state higher-degree analogues of Theorem 6.2. The analogous theorems result in additional improvements of Theorem 3.9, but it is not clear how

m	$N(m)$	$C(m)$	$C(m)/N(m)$	boundary
100	780708	445288	.57036	122
200	6232898	3558002	.57084	277
300	21019770	12004290	.57110	453
400	49814320	28450698	.57113	640
500	97252440	55564678	.57134	807
600	168075834	96011988	.57124	996
700	266842438	152459384	.57135	1140
800	398334336	227573618	.57131	1355
900	567119096	324020078	.57134	1766
1000	777890010	444469060	.57138	1695

TABLE 1. For any positive integer m, let $N(m)$ be the number of normal cubic integer polynomials with max-norm m, and let $C(m)$ be the number of those normal cubic integer polynomials of max-norm m that have all roots in the cone $\mathcal{C}$. The ratios $C(m)/N(m)$ are rounded to five decimal digits. The last column lists the number of polynomials that have non-real roots on the boundary of $\mathcal{C}$.

the improvements can be used to obtain better general bounds for the Descartes method.

Acknowledgement. We thank G. E. Collins for references [13] and [37], A. Eigenwillig for improvements of the presentation, M. Mignotte for reference [19] which led us to reference [30], and D. G. Richardson for developing a leak-free memory management system for saclib that allowed us to perform the experiments in Section 6. The program qepcad developed by H. Hong and others was used to construct the example polynomials in Sections 4, 5 and 6 from sample points in a cylindrical algebraic decomposition [9, 11], and to obtain the quantifier-free formula in Theorem 6.2. The second author acknowledges partial support by the IST Programme of the EU under contract IST-2000-26473 (ECG—Effective Computational Geometry for Curves and Surfaces).

REFERENCES

[1] A. A. Albert. An inductive proof of Descartes' rule of signs. *The American Mathematical Monthly*, 50(3):178–180, 1943.

[2] Alberto Alesina and Massimo Galuzzi. A new proof of Vincent's theorem. *L'Enseignement Mathématique*, 44:219–256, 1998.

[3] Alberto Alesina and Massimo Galuzzi. Addendum to the paper "A new proof of Vincent's theorem". *L'Enseignement Mathématique*, 45:379–380, 1999.

[4] James W. Anderson. *Hyperbolic Geometry.* Springer-Verlag London, 1999.

[5] Margherita Bartolozzi and Raffaella Franci. La regola dei segni dall' enunciato di R. Descartes (1637) alla dimostrazione di C. F. Gauss (1828). *Archive for History of Exact Sciences*, 45(4):335–374, 1993.

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

[7] G. E. Collins and R. Loos. Real zeros of polynomials. In B. Buchberger, G. E. Collins, and R. Loos, editors, *Computer Algebra: Symbolic and Algebraic Computation*, pages 83–94. Springer-Verlag, 2nd edition, 1982.

[8] George E. Collins. The computing time of the Euclidean algorithm. *SIAM Journal on Computing*, 3(1):1–10, 1974.

[9] George E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In H. Brakhage, editor, *Automata Theory and Formal Languages*, volume 33 of *Lecture Notes in Computer Science*, pages 134–183. Springer-Verlag, Berlin, 1975. Reprinted (with corrections by the author) in: B. F. Caviness and J. R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, Springer-Verlag, 1998, pages 85–121.

[10] George E. Collins and Alkiviadis 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.

[11] George E. Collins and Hoon Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, 1991. Reprinted in: B. F. Caviness, J. R. Johnson, editors, Quantifier Elimination and Cylindrical Algebraic Decomposition, Springer-Verlag, 1998, pages 174–200.

[12] George E. Collins, Jeremy R. Johnson, and Werner Krandick. Interval arithmetic in cylindrical algebraic decomposition. *Journal of Symbolic Computation*, 34(2):143–155, 2002.

[13] N. B. Conkwright. *Introduction to the Theory of Equations*. Ginn and Co., 1941.

[14] J. H. Davenport. Computer algebra for cylindrical algebraic decomposition. Technical report, The Royal Institute of Technology, Department of Numerical Analysis and Computing Science, S-100 44, Stockholm, Sweden, 1985. Reprinted as: Technical Report 88-10, School of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, England.

[15] Thomas Decker and Werner Krandick. Parallel real root isolation using the Descartes method. In P. Banerjee, V. K. Prasanna, and B. P. Sinha, editors, *High Performance Computing – HiPC'99*, volume 1745 of *Lecture Notes in Computer Science*, pages 261–268. Springer-Verlag, 1999.

[16] Thomas Decker and Werner Krandick. Isoefficiency and the parallel Descartes method. In G. Alefeld, J. Rohn, S. Rump, and T. Yamamoto, editors, *Symbolic Algebraic Methods and Verification Methods*, Springer Mathematics, pages 55–67. Springer-Verlag, 2001.

[17] René Descartes. *The Geometry.* Dover Publications, New York, 1954. Translated from the French and Latin by D. E. Smith and M. L. Latham. With a facsimile of the first edition, 1637.

[18] Carl Friedrich Gauss. Beweis eines algebraischen Lehrsatzes. *Journal für die reine und angewandte Mathematik*, 3(1):1–4, 1828. Reprinted in: Carl Friedrich Gauss, Werke, vol. 3, Königliche Gesellschaft der Wissenschaften, Dieterich, Göttingen, 1866, pages 65–70.

[19] Vicente Gonçalves. L'inégalité de W. Specht. *Revista da Faculdade de Ciências da Universidade de Lisboa A—2nd series*, 1:167–171, 1950.

[20] The Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, NY 10017, USA. *ANSI/IEEE Std 754-1985. An American National Standard: IEEE Standard for Binary Floating-Point Arithmetic*, 1985. Reprinted as: ANSI/IEEE Standard 754-1985 for binary floating-point arithmetic. *ACM SIGPLAN Notices*, 22(2):9–25, 1987.

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

[22] J. R. Johnson and W. Krandick. Polynomial real root isolation using approximate arithmetic. In W. W. Küchlin, editor, *International Symposium on Symbolic and Algebraic Computation*, pages 225–232. ACM Press, 1997.

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

[24] M[onsieur] E. Landau. Sur quelques théorèmes de M. Petrovitch relatifs aux zéros des fonctions analytiques. *Bulletin de la Société Mathématique de France*, 33:251–261, 1905. Reprinted in: Edmund Landau: Collected Works, L. Mirsky, I. J. Schoenberg, W. Schwarz, H. Wefelscheid, editors, vol. 2, Thales Verlag, Essen, 1986, pages 180–190.

[25] K. Mahler. An inequality for the discriminant of a polynomial. *The Michigan Mathematical Journal*, 11(3):257–262, 1964.

[26] M. Marden. Ostrowski A. M.: Note on Vincent's theorem. *Mathematical Reviews*, 12(6):408–409, 1951.

[27] M. Mignotte. An inequality about factors of polynomials. *Mathematics of Computation*, 28(128):1153–1157, 1974.

[28] M. Mignotte. Some useful bounds. In B. Buchberger, G. E. Collins, and R. Loos, editors, *Computer Algebra: Symbolic and Algebraic Computation*, pages 259–263. Springer-Verlag, 2nd edition, 1982.
[29] A. M. Ostrowski. Note on Vincent's theorem. *Annals of Mathematics, Second Series*, 52(3):702–707, 1950. Reprinted in: Alexander Ostrowski, Collected Mathematical Papers, vol. 1, Birkhäuser Verlag, 1983, pages 728–733.
[30] A. M. Ostrowski. On an inequality of J. Vicente Gonçalves. *Revista da Faculdade de Ciências da Universidade de Lisboa A—2nd series*, 8:115–119, 1961. Reprinted in: Alexander Ostrowski, Collected Mathematical Papers, vol. 1, Birkhäuser Verlag, 1983, pages 785–789.
[31] M[onsieur] Alexandre Ostrowski. Note sur les produits de séries normales. *Bulletin de la Société Royale des Sciences de Liège*, 8:458–468, 1939. Reprinted in: Alexander Ostrowski, Collected Mathematical Papers, vol. 3, Birkhäuser Verlag, 1984, pages 414–424.
[32] Fabrice Rouillier and Paul Zimmermann. Efficient isolation of a polynomial real roots. Rapport de recherche 4113, Institut National de Recherche en Informatique et en Automatique, 2001. 16 pages.
[33] Fabrice Rouillier and Paul Zimmermann. Efficient isolation of a polynomial's real roots. *Journal of Computational and Applied Mathematics*, 162:33–50, 2004.
[34] J. V. Uspensky. *Theory of Equations*. McGraw-Hill Book Company, Inc., 1948.
[35] B. L. van der Waerden. *Modern Algebra*, volume I. Frederick Ungar Publishing Co., New York, 1949.
[36] M[onsieur] Vincent. Sur la résolution des équations numériques. *Journal de Mathématiques pures et appliquées*, 1:341–372, 1836.
[37] Xiaoshen Wang. A simple proof of Descartes's rule of signs. *The American Mathematical Monthly*, 111(6):525–526, 2004.