# Lecture 9

# Root Isolation

The determination of the roots of a univariate real polynomial is ubiquitous in geometric computing. Let $f$ be such a polynomial and let $n$ be its degree, i.e.,

$$f = \sum_{0 \le i \le n} f_i x^i \in \mathbb{R}[x].$$

A polynomial of degree $n$ has exactly $n$ complex roots[1] if roots are counted with multiplicities. We can write $f$ as a product of linear factors, i.e.,

$$f(x) = f_n \prod_i (x - \xi_i)^{k_i},$$

where the $\xi_i$'s are the distinct roots of $f$ and $k_i$ is the multiplicity of $\xi_i$. Then $\sum_i k_i = n$. A root $\xi_i \in \mathbb{R}$ is called a real root. The nonreal roots come in pairs of conjugate roots. More precisely, if $\xi = a + b\mathbf{i}$ with $b > 0$ is a root of $f$, then the complex conjugate $\overline{\xi} = a - b\mathbf{i}$ of $\xi$ is also a root of $f$. Indeed $f(\xi) = 0$ implies[2]

$$0 = \overline{0} = \overline{f(\xi)} = \overline{\sum_{0 \le i \le n} f_i \xi^i} = \sum_{0 \le i \le n} \overline{f_i \xi^i} = \sum_{0 \le i \le n} \overline{f_i}\,\overline{\xi^i} = \sum_{0 \le i \le n} f_i \overline{\xi}^i,$$

where the last equality holds since the coefficients $f_i$ are real numbers. Thus $\overline{\xi}$ is a root of $f$. We ask the reader to show that $\xi$ and $\overline{\xi}$ have the same multiplicity.

**Exercise 0.1:** Let $\xi$ be a nonreal root of $f$. Show that $\xi$ and $\overline{\xi}$ have the same multiplicity as roots of $f$. Hint: Let $\xi = a + \mathbf{i}b$. Then $g(x) = (x - \xi)(x - \overline{\xi}) = x^2 - 2ax + a^2 + b^2$ is a real polynomial that divides $f$. Thus $f/g$ is also a real polynomial. Now apply induction. ♢

How can we determine the roots of a polynomial? For polynomials of degree one, the task is trivial.

$x - 5 = 0$ has a single root, namely $x = 5$.

For polynomials of degree two, we learned the solution in high school.

$x^2 + bx + c = 0$ has two roots, namely $x = (-b \pm \sqrt{b^2 - 4c})/2$.

The polynomial has two distinct real roots if $b^2 - 4c > 0$, it has a double real root if $b^2 - 4c = 0$, and its has two complex roots if $b^2 - 4c < 0$. Explicit solutions are also known for polynomials of degree three and four.

[1] This is called the Fundamental Theorem of Algebra.

[2] Conjugation commutes with addition and multiplication, i.e., for $x, y \in \mathbb{C}$, $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$ and $\overline{x + y} = \overline{x} + \overline{y}$.

**Exercise 0.2:** Look up the solution method for polynomials of degree three and apply it to $x^3 - 2x^2 + 7x + 19$. ♢

**Exercise 0.3:** Look up the solution method for polynomials of degree four and apply it to $x^4 + x^3 - 2x^2 + 7x + 19$. ♢

For polynomials of degree five and higher, explicit solutions are not available. What does it mean then to compute the roots of a univariate polynomial $f$? It can mean different things.

(1) Determine the number of roots of $f$. This is easy. The number of roots of $f$ in $\mathbb{C}$ counted with multiplicities is precisely the degree of $f$. All formulations to follow are non-trivial computational tasks.

(2) Determine the number of distinct roots of $f$ (and their multiplicities).

(3) Determine the number of real roots of $f$.

(4) Determine the number of distinct real roots of $f$ (and their multiplicities).

(5) Isolate the complex roots of $f$, i.e., determine triples $(\xi_i, r_i, k_i)$ with $\xi \in \mathbb{C}$, $r_i \in \mathbb{R}$, and $k_i \in \mathbb{N}$ with the following properties.

- The disk (in the complex plane) with center $\xi_i$ and radius $r_i$ contains a $k_i$-fold root of $f$.
- $\sum_i k_i = \deg f$
- The disks $D(\xi_i, r_i)$ are disjoint.

(6) Isolate the real roots of $f$, i.e., determine triples $(\xi_i, r_i, k_i)$ with $\xi \in \mathbb{R}$, $r_i \in \mathbb{R}$, and $k_i \in \mathbb{N}$ with the following properties.

- The interval (of the real axis) with center $\xi$ and radius $r_i$ contains a $k_i$-fold real root of $f$.
- $\sum_i k_i$ is equal to the number of real roots of $f$ counted with multiplicities.
- The intervals $I(\xi_i, r_i)$ are disjoint.

(7) As (5) or (6), but guarantee in addition that the $r_i$ are smaller than some presribed $\varepsilon$.

It is also interesting to solve these problems for polynomials with special properties, in particular, square-free polynomials. A polynomial is square-free if all roots (complex or otherwise) of $f$ are distinct. We will see in Lecture **??** how to factor a polynomial $f$ into a product $f_1 f_2 \dots f_n$, $n = \deg f$, of square-free polynomials such that the roots of $f_k$ are precisely the $k$-fold roots of $f$. Hence, restricting attention to square-free polynomials is justified.

In this section, we will concentrate on (6) for square-free polynomials. We will start with polynomials with rational coefficients ($f \in \mathbb{Q}[x]$) and then generalize to polynomials with real coefficients. (7) will be the topic of Lecture **??**.

## 9.1 Root Isolation for Polynomials with Integer Coefficients

Throughout this section, $f = \sum_{0 \le i \le n} f_i x^i$ is a polynomial with integer coefficients. This includes the case that the coefficients are rational numbers; multiplying $f$ with (a multiple of) the least common multiple of the denominators of the $f_i$'s converts $f$ into a polynomial with integer coefficients.

### 9.1.1 Root Bounds

We derive an upper bound on the absolute value of any root of $f$. Such a bound, call it $B$, is very useful. It allows us to restrict the search for complex roots to the disk (in the complex plane) of radius $B$ centered at 0 and the search for real roots to the interval of radius $R$ centered at 0. The following result is due to Cauchy[3].

THEOREM 1. *Let $f \in \mathbb{R}[x]$, and let $\xi$ be a root of $f$. Then*

$$|\xi| < B := 2 \max_{0 \le i \le n} \frac{|f_i|}{|f_n|}.$$

*Proof.* Observe first that $B \ge 2$ since the fraction $|f_n|/|f_n|$ is included in the maximization. For the sake of a contradiction, assume $f$ has a root $\xi$ with $|\xi| \ge B$. Then $0 = f_n\xi^n + \sum_{0 \le i \le n-1} f_i\xi^i$ and hence $|f_n\xi^n| = \left|\sum_{0 \le i \le n-1} f_i\xi^i\right|$. Thus

$$B^n \le |\xi^n| = \frac{\left|\sum_{0 \le i \le n-1} f_i\xi^i\right|}{|f_n|} \le \sum_{0 \le i < n} \frac{|f_i|}{|f_n|} |\xi|^i \le \sum_{0 \le i < n} \frac{B}{2} B^i = \frac{B}{2}\frac{B^n - 1}{B-1} \le B^n - 1,$$

a contradiction. The last inequality follows from $B/2 \le B-1$ for $B \ge 2$. □

### 9.1.2 Descartes' Rule of Sign

Descartes[4] established a simple rule for bounding the number of positive real roots of a polynomial. Let $f = \sum_{0 \le i \le n} f_i x^i \in \mathbb{R}[x]$ be a univariate polynomial of degree $n$. We define the *number of sign changes* $\mathtt{Var}(f)$ *in the coefficient sequence* of $f$ as the number of pairs $(i, j)$ with $i < j$, $f_i f_j < 0$ and $f_{i+1} = \ldots = f_{j-1} = 0$. The sequence $(-2, 0, +2, +2, -1)$ has two sign changes.

THEOREM 2 (Descartes). *Let $f = \sum_{0 \le i \le n} f_i x^i \in \mathbb{R}[x]$, $f \not\equiv 0$, be a univariate polynomial with real coefficients. Let $PZ(f)$ be the number of positive real roots of $f$ counted with multiplicities. Then*

$$\mathtt{Var}(f) - PZ(f) \text{ is an even nonnegative integer.}$$

*Proof.* We may assume that $f_0 \neq 0$. Otherwise, we consider $f/x$ instead of $f$. $\mathtt{Var}(f)$ is even if $f_0$ and $f_n$ have the same sign and is odd otherwise. The sign of $f(0)$ is equal to the sign of the constant coefficient and the sign of $f(x)$ for sufficiently large $x$ is equal to the sign of the leading coefficient. Thus the number of real zeros counted with multiplicities is even iff $f_0$ and $f_n$ have the same sign. We have now established that $\mathtt{Var}(f)$ and $PZ(f)$ have the same parity.

[3] Augustin-Louis Cauchy (21 August 1789 – 23 May 1857) was a French mathematician who was an early pioneer of analysis. He started the project of formulating and proving the theorems of infinitesimal calculus in a rigorous manner. He also gave several important theorems in complex analysis and initiated the study of permutation groups in abstract algebra. A profound mathematician, Cauchy exercised a great influence over his contemporaries and successors. His writings cover the entire range of mathematics and mathematical physics. Quote from Wikipedia (January 5, 2010).

[4] René Descartes (31 March 1596 – 11 February 1650) was a French philosopher, mathematician, physicist, and writer who spent most of his adult life in the Dutch Republic. He has been dubbed the "Father of Modern Philosophy", and much of subsequent Western philosophy is a response to his writings, which continue to be studied closely to this day. In particular, his Meditations on First Philosophy continues to be a standard text at most university philosophy departments. Descartes' influence in mathematics is also apparent, the Cartesian coordinate system–allowing geometric shapes to be expressed in algebraic equations being named for him. He is credited as the father of analytical geometry. Descartes was also one of the key figures in the Scientific Revolution. (Quote from Wikipedia, January 5, 2010)

It remains to establish $\mathtt{Var}(f) \geq PZ(f)$. We do so by induction on the degree of $f$. If the degree of $f$ is zero, $\mathtt{Var}(f) = 0 = PZ(f)$. For $\deg f > 0$, consider the derivative $f'$ of $f$. By induction hypothesis, $\mathtt{Var}(f') \geq PZ(f')$. Also $\mathtt{Var}(f) \geq \mathtt{Var}(f')$ and $PZ(f) \leq PZ(f') + 1$. Thus

$$\mathtt{Var}(f) - PZ(f) \geq \mathtt{Var}(f') - (PZ(f') + 1)) = \mathtt{Var}(f') - PZ(f') - 1 \geq -1.$$

Since $\mathtt{Var}(f) - PZ(f)$ is even, we conclude $\mathtt{Var}(f) \geq PZ(f)$. □

The case of zero or one sign change deserves special mentioning.

COROLLARY 3. *If $\mathtt{Var}(f) = 0$, $f$ has no positive real root, and if $\mathtt{Var}(f) = 1$, $f$ has exactly one positive real root.*

The rule is easily extended to arbitrary open intervals by a suitable coordinate transformation. Let $I = (a,b)$ be an open interval. The mapping

$$x \mapsto \frac{ax+b}{x+1}$$

maps $(0,\infty)$ bijectively onto $(a,b)$ and hence the positive real roots of

$$f_I(x) := (1+x)^n \cdot f\left(\frac{ax+b}{x+1}\right)$$

correspond bijectively to the real roots of $f$ in $I$. We define $\mathtt{Var}(f,I)$ as $\mathtt{Var}(f_I)$. The factor $(1+x)^n$ in the definition of $f_I$ clears denominators and guarantees that $f_I$ is a polynomial. We thus have the following extension of Descartes' rule to intervals.

THEOREM 4. *Let $f$ be a polynomial, let $I = (a,b)$ be an interval and let $k$ be the number of zeros of $f$ (counted with multiplities) in $I$. Then*

$$\mathtt{Var}(f,I) - k \text{ is an even nonnegative integer}.$$

The number of sign changes $\mathtt{Var}(f,I)$ of $f$ with respect to an interval is an upper bound on the number of roots of $f$ in $I$. It may and, in general, will overestimate the number of roots. The Bulgarian mathematician Obreschkoff has established tighter bounds on $\mathtt{Var}(f,I)$.

THEOREM 5 ([**?**, **?**]). *Let $f$ be a polynomial of degree $n$, $I$ an open interval, and $v = \mathtt{Var}(f,I)$. If the Obreshkoff lens $L_{n-q}$ (see Figure 9.1) contains at least $q$ roots (counted with multiplicity) of $f$, then $v \geq q$. If the Obreshkoff area $A_q$ (see Figure 9.1) contains at most $q$ roots (counted with multiplicity) of $f$, then $v \leq q$. In particular,*

$$\#\text{ of roots of } f \text{ in } L_n \leq \mathtt{Var}(f,I) \leq \#\text{ of roots of } f \text{ in } A_n.$$

The cases $q = 0$ and $q = 1$ deserve special mentioning. They run under the names one-circle theorem and two-circle theorem, respectively.

THEOREM 6 ([**?**, **?**]). *Consider a real polynomial $f(x)$ and an interval $I = (a,b)$ with midpoint $m_I = (a+b)/2$ and let $v = \mathtt{Var}(f,I)$.*

- *(One-Circle Theorem) If the open disc bounded by the circle $C_0$ centered at $m_I$ and passing through the endpoints of $I$ contains no root of $f(x)$, then $v = 0$.*

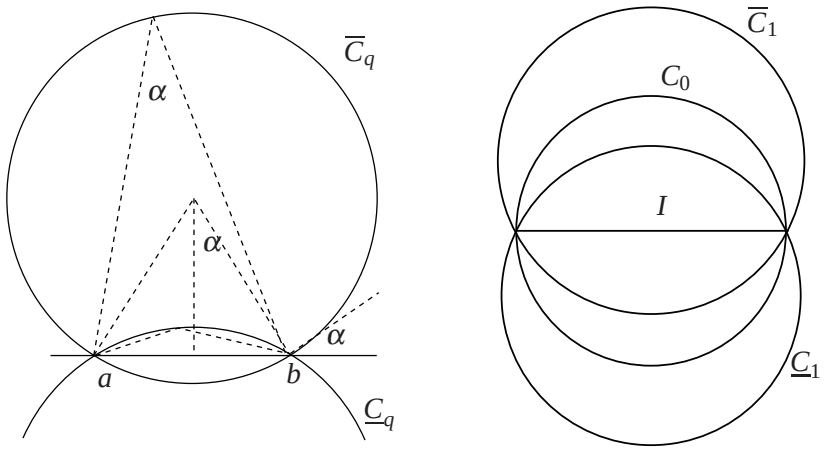


Figure 9.1: For any $q$ with $0 \le q \le n$, the Obreshkoff disks $\overline{C}_q$ and $\underline{C}_q$ for $I$ have the endpoints of $I$ on their boundary; their centers see the line segment $[a,b]$ under the angle $2\alpha = 2\pi/(q+2)$. The Obreshkoff lens $L_q$ is the interior of $\overline{C}_q \cap \underline{C}_q$ and the Obreshkoff area $A_q$ is the interior of $\overline{C}_q \cup \underline{C}_q$. Any point (except for $a$ and $b$) on the boundary of $A_q$ sees $[a,b]$ under an angle $\pi/(q+2)$ (= half the angle at the center) and any point (except for $a$ and $b$) on the boundary of $L_q$ sees $[a,b]$ under angle $\pi - \pi/(q+2)$ (= half the complementary angle at the center). We have $L_n \subset L_{n-1} \subset \ldots \subset L_1 \subset L_0$ and $A_0 \subset A_1 \subset \ldots \subset A_{n-1} \subset A_n$. The circles $\overline{C}_0$ and $\underline{C}_0$ coincide. They have their center at the midpoint of $I$. The circles $\overline{C}_1$ and $\underline{C}_1$ are the circumcircles of the two equilateral triangles having $I$ as one of their edges. We call $A_1$ the *two-circle region* of $I$.

- *(Two-Circle Theorem) If the union of the open discs bounded by the circles $\underline{C}_1$ and $\overline{C}_1$ centered at $m_I \pm i(1/(2\sqrt{3}))w(I)$ and passing through the endpoints of $I$ contains precisely one root of $f(x)$, then $v = 1$.*

We would expect the number of sign variations $\mathtt{Var}(f,I)$ to be a monotone function in $I$, i.e., if $I \subseteq J$ then $\mathtt{Var}(f,I) \le \mathtt{Var}(f,J)$. In fact, much more is true. The function $\mathtt{Var}(f,I)$ is subadditive. For a simple self-contained proof, we refer the reader to [**?**, Corollary 2.27].

THEOREM 7 ([**?**]). *Let $f$ be a real polynomial. If the pairwise disjoint open intervals $J_1, \ldots, J_l$ are subsets of the open interval $I$, then*

$$\sum_{1 \le i \le \ell} \mathtt{Var}(f,J_i) \le \mathtt{Var}(f,I).$$

### 9.1.3 Proofs of the One-Circle and Two-Circle Theorems*

The goal of this section is to prove the one-circle and two-circle theorems. We proceed in two steps. In the first step, we derive conditions under which $\mathtt{Var}(f) = 0$ and $\mathtt{Var}(f) = 1$, respectively. In the second step, we transform these conditions to intervals.

LEMMA 8. *If all roots of a real polynomial $f$ lie in the closed left halfplane of the complex plane (if $\xi = a + b\mathbf{i}$ is a root of $f$ then $a \le 0$), $\mathtt{Var}(f) = 0$.*

*Proof.* Since the nonreal roots of $f$ come in conjugate pairs, we have

$$f = f_n \prod_{\xi \in \mathbb{R};\ f(\xi)=0} (x-\xi) \prod_{\xi \in \mathbb{C};\ f(\xi)=0;\ \mathtt{Im}(\xi)>0} (x-\xi)(x-\overline{\xi}).$$

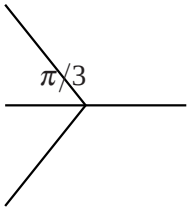


Figure 9.2: The cone mentioned in Lemma 9. [[Ugly figure. Make it an inline figure.]]

If $\xi$ is a real root of $f$, $\xi \leq 0$ and hence both coefficients of $x-\xi$ are nonnegative. If $\xi = a+b\mathbf{i}$ is a nonreal root of $f$ with $a \leq 0$, all coefficients of $(x-\xi)(x-\overline{\xi}) = x^2 - 2ax + a^2 + b^2$ are nonnegative. Thus all coefficients of $f/f_n$ are nonnegative and hence $\mathtt{Var}(f) = 0$. □

LEMMA 9. *Let $C = \{\xi \in \mathbb{C} \mid \xi = |\xi| e^{\mathbf{i}\varphi}$ and $2\pi/3 \leq \varphi \leq 4\pi/3\}$ be the cone with opening angle $2\pi/3$ and centered at the negative real axis, see Figure 9.2. If a real polynomial $f$ has exactly one positive real root and all other roots in $C$, $\mathtt{Var}(f) = 1$.*

*Proof.* Since $f$ has a positive real root, $\mathtt{Var}(f) \geq 1$. It remains to show $\mathtt{Var}(f) \leq 1$. As in the proof of the preceding lemma, we write

$$f = f_n \prod_{\xi\in\mathbb{R};\ f(\xi)=0} (x-\xi) \prod_{\xi\in C;\ f(\xi)=0;\ \mathtt{Im}(\xi)>0} (x-\xi)(x-\overline{\xi}).$$

We now build the product $f/f_n$ inductively. We start with $x-\xi$ where $\xi$ is the positive real root of $f$. Then we have one sign change. Assume now that we have $h(x) = \sum_{0\leq i\leq m} a_i x^i$ with $\mathtt{Var}(h) \leq 1$, say $a_0,\ldots,a_k \leq 0$ and $a_{k+1},\ldots,a_m \geq 0$. We will show $\mathtt{Var}(h\cdot(x-\xi)) \leq 1$, whenever $\xi \leq 0$ and $\mathtt{Var}(h\cdot(x-\xi)(x-\overline{\xi})) \leq 1$, whenever $\xi \in C\setminus\mathbb{R}$.

Consider first $h\cdot(x-\xi))$ with $\xi \leq 0$. Write $h\cdot(x-\xi) = \sum_{0\leq i\leq m+1} c_i x^i$. Then $c_{m+1} = a_m \geq 0$, $c_0 = -a_0\xi \leq 0$ and $c_i = a_{i-1} - \xi a_i$ for $1 \leq i \leq m$. Thus $c_i \geq 0$ for $i \geq k+2$ and $c_i \leq 0$ for $i \leq k$. Whatever the sign of $c_{k+1}$, $\mathtt{Var}(h\cdot(x-\xi)) \leq 1$.

Consider next $g := h\cdot(x-\xi)(x-\overline{\xi})$ with $\xi \in C\setminus\mathbb{R}$. Let $\xi = a+\mathbf{i}b$. Then $a \leq 0$ and $b^2 \leq 3a^2$ and $(x-\xi)(x-\overline{\xi}) = x^2 - 2ax + a^2 + b^2$. Since $b > 0$, we have $a < 0$. We may substitute $-2ax$ for $x$ without changing the number of sign changes in either $h$ or $g$; this holds since $-2a > 0$. The substitution changes $x^2 - 2ax + a^2 + b^2$ into $4a^2(x^2 + x + \lambda)$ where $\lambda = (a^2+b^2)/(4a^2)$ and hence $1/4 \leq \lambda \leq 1$.

Write $g = \sum_{0\leq i\leq m+2} c_i x^i$. Then $c_i = a_{i-2} + a_{i-1} + \lambda a_i$ for $0 \leq i \leq m+2$ with the convention $a_{-2} = a-1 = a_{m+1} = a_{m+2} = 0$. Thus $c_i \geq 0$ for $i \geq k+3$ and $c_i \leq 0$ for $i \leq k$ and hence $\mathtt{Var}(g) \leq 3$. Since $g$ has exactly one positive real root, $\mathtt{Var}(g)$ is odd. It remains to exclude the case $\mathtt{Var}(g) = 3$. If $\mathtt{Var}(g) = 3$, necessarily, $c_{k+2} < c_{k+1}$. However,

$$\begin{aligned}
c_{k+1} &= a_{k-1} + a_k + \lambda a_{k+1} && \text{definition of } c_{k+1}\\
&\leq a_k + a_{k+1} && \text{since } \lambda \leq 1,\ a_{k-1} \leq 0,\ a_{k+1} \geq 0\\
&\leq a_k + a_{k+1} + \lambda a_{k+2} && \text{since } \lambda > 0 \text{ and } a_{k+2} \geq 0\\
&= c_{k+2} && \text{definition of } c_{k+2}.
\end{aligned}$$

□

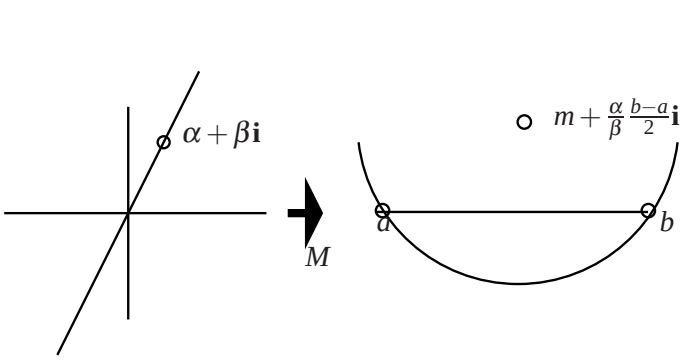


Figure 9.3: $M$ maps the line with slope $\beta/\alpha$ through the origin onto the circle with center $m(I)+(\alpha/\beta)w(I)/2\mathbf{i}$ and passing through the endpoints of $I=(a,b)$. [[This figure looks ugly. I should learn how to use IPE.

Recall, the definition of $\mathtt{Var}(f,I)$. We defined $\mathtt{Var}(f,I)$ as $\mathtt{Var}(f_i)$, where

$$f_I(x) := (1+x)^n \cdot f\left(\frac{ax+b}{x+1}\right).$$

The mapping

$$M:\ x \mapsto \frac{ax+b}{x+1}$$

maps $(0,\infty)$ bijectively onto $(a,b)$ and hence the positive real roots of $f_I$ correspond bijectively to the real roots of $f$ in $I$. We have $\mathtt{Var}(f_I)=0$ if all roots of $f_I$ lie in the closed left halfplane of the complex plane (equivalently: $f_I$ has no root in the open right halfplane) and $\mathtt{Var}(f_I)=1$ if $f_I$ has one positive real root and all other roots in the cone $C$ (equivalently: has exactly one root outside the cone $C$). So we need to ask ourselves: into which region does $M$ map the open right halfplane and the complement of cone $C$, respectively. These are exactly the regions mentioned in the one- and two-circle theorems as we show next.

LEMMA 10. *The mapping $M$ maps lines through the origin into circles passing through the endpoints of $I$. More precisely, a line with slope $\gamma$ is mapped into the circle with center $(a+b)/2+((b-a)/(2\gamma))\mathbf{i}$ and passing through the endpoints of $I$, see Figure 9.3 for an illustration.*

*Proof.* The proof is a straightforward but somewhat tedious calculation. We observe first that

$$\frac{ax+1}{x+1} = \frac{a+b}{2} + \frac{b-a}{2}\cdot\frac{1-x}{1+x}.$$

Therefore is suffices to prove that the mapping $\widetilde{M}:\ x\mapsto(1-x)/(1+x)$ maps the line $\ell$ with slope $\gamma$ into the circle $C_\gamma$ with center $0+(1/\gamma)\mathbf{i}$ and passing through the points $\pm1+0\mathbf{i}$.

Let $\alpha+\beta\mathbf{i}$ be an intersection of $\ell$ with the unit circle. Then $\gamma=\beta/\alpha$ and $\alpha^2+\beta^2=1$. The points on $\ell$ are parameterized as $(\alpha+\beta\mathbf{i})t$ with $t\in\mathbb{R}$. $\widetilde{M}$ maps $(\alpha+\beta\mathbf{i})t$ into

$$\frac{1-\alpha t-\beta t\mathbf{i}}{1+\alpha t+\beta t\mathbf{i}} = \frac{(1-\alpha t-\beta t\mathbf{i})(1+\alpha t-\beta t\mathbf{i})}{(1+\alpha t+\beta t\mathbf{i})(1+\alpha t-\beta t\mathbf{i})} = \frac{1-2\beta t\mathbf{i}-t^2}{1+2\alpha t+t^2} = \frac{1-t^2}{1+2\alpha t+t^2} + \frac{-2\beta t}{1+2\alpha t+t^2}\mathbf{i}.$$

The parameter value 0 is mapped to $1+0\mathbf{i}$ and the parameter value $\infty$ is mapped to $-1+0\mathbf{i}$. The squared distance of the generic image from the center of $C_\gamma$ is

$$\begin{aligned}\left(\frac{1-t^2}{1+2\alpha t+t^2}\right)^2+\left(\frac{-2\beta}{1+2\alpha t+t^2}-\frac{\alpha}{\beta}\right)^2 &= \frac{\beta^2(1-t^2)+(-2\beta^2 t-\alpha(1+2\alpha t+t^2))^2}{\beta^2(1+2\alpha t+t^2)^2}\\ &= \frac{\beta^2(1-t^2)+(-\alpha-2t-\alpha t^2)^2}{\beta^2(1+2\alpha t+t^2)^2}\\ &= \frac{\beta^2-2\beta^2t^2+t^4+\alpha^2+4\alpha t+(4+2\alpha^2)t^2+4\alpha t^3+\alpha^4t^4}{\beta^2(1+2\alpha t+t^2)^2}\\ &= \frac{1+4\alpha t+(-2\beta^2+4+2\alpha^2)t^2+4\alpha t^3+t^4}{\beta^2(1+4\alpha t+(4\alpha^2+2)t^2+4\alpha t^3+t^4)}\\ &= \frac{1}{\beta^2},\end{aligned}$$

since $-2\beta^2+4+2\alpha^2=4\alpha^2+2$. We have now shown that $\widetilde{M}$ maps $\ell$ onto the circle with center $0+(1/\gamma)\mathbf{i}$ and passing through $\pm1+0\mathbf{i}$. □

It is now easy to prove the one- and two-circle theorem. We have $\mathtt{Var}(f_I)=0$ if $f_I$ has no root in the open right halfplane. By Lemma 10, $M$ maps the imaginary axis (this is a line with slope $\gamma=\infty$) onto the circle $C_0$ centered at the midpoint of $I$ and passing through the endpoints of $I$. Continuity tells us that the two open halfplanes defined by the imaginary axis are mapped into the interior and exterior of the circle, respectively. The right open halfplane is mapped into the interior since the positive real axis is mapped onto $I$. Thus $f_I$ has no root in the open right halfplane if and only if $f$ has no roots inside $C_0$. This establishes the one-circle theorem.

We come to the two-circle theorem. We have $\mathtt{Var}(f_I)=1$, if $f_I$ has exactly one root outside the cone $C$. This cone is bounded by the rays $t\mapsto t\pm\sqrt{3}t\mathbf{i}$, $t\in\mathbb{R}_{\geq 0}$. By Lemma 10, $M$ maps the supporting lines of these rays into the circles $\underline{C}_1$ and $\overline{C}_1$ centered at $m_I\pm i(1/(2\sqrt{3}))w(I)$ and passing through the endpoints of $I$. The complement of the cone $C$ is the union of two open halfplanes. These open halfplanes are mapped into the interior of $\underline{C}_1$ and $\overline{C}_1$, respectively. Thus $f_I$ has exactly one root outside $C$ if and only if $f$ has exactly one root in the union of the interiors of $\underline{C}_1$ and $\overline{C}_1$. This establishes the two-circle theorem.

[[ drawings would be helpful]]

### 9.1.4 A Bisection Algorithm for Root Isolation

[[**The polynomial is now called** $p$]]

For a real root $z$ of $p$, let $\sigma(z,p)$ be the minimal distance of $z$ to another root of $p$. For a nonreal root $z$ of $p$, let $\sigma(z,p)$ be the absolute value of its imaginary coordinate. Let $\sigma(p)$ be the minimal value of $\sigma(z,p)$ over all roots of $p$. For an interval $I=(a,b)$, let $w(I):=b-a$ be its length or width.

Algorithm 1 shows a bisection algorithm for isolating the roots of a real polynomial $p$ in an open interval $I_0$ based on Descartes' rule of sign. The algorithm requires that the real roots of $p$ in $I_0$ are simple. If the requirement is not met, the algorithm diverges. It maintains a set $A$ of active intervals. Initially, $A$ contains $I_0$, and the algorithm stops as soon as $A$ is empty. In each iteration, some interval $I\in A$ is processed. The action taken depends on the integer $\mathtt{Var}(p,I)$, the outcome of Descartes' rule of signs applied to $p$ and $I$.

**Algorithm 1** Bisection Algorithm for Isolating Real Roots

**Require:** $p = \sum_{0 \le i \le n} p_i x^i$ is a real polynomial and $I_0$ is an open interval. The real roots of $p$ in $I_0$ are simple.

**Ensure:** returns a list $O$ of isolating intervals for the real roots of $p$ in $I$.

```
A := {I_0}                                   {list of active intervals}
O := ∅                                       {list of isolating intervals}
repeat
  I := some interval in A; delete I from A;
  if Var(p,I) = 0 do nothing;
  if Var(p,I) = 1 add I to O;
  if Var(p,I) ≥ 2 then
    let I = (a,b) and set m := (a+b)/2;
    if p(m) = 0 add [m,m] to O;
    add (a,m) and (m,b) to A;
  end if
until A is empty
return O
```

If there is no sign change, $I$ contains no root of $p$ and we discard it. If there is exactly one sign change, $I$ contains exactly one root of $p$ and hence is an isolating interval for it. We add $I$ to the list $O$ of isolating intervals. If there is more than one sign change, we divide $I$ at its midpoint and add the subintervals to the set of active intervals. If the midpoint $m$ is a zero of $p$, we add the trivial interval $[m,m]$ to the list of isolating intervals.

Correctness of the algorithm is obvious. Termination and complexity analysis rest on the one- and two-circle theorems.

LEMMA 11. *No interval of length $\sigma(p)$ or less is split.*

*Proof.* Such an interval, recall that is is open, cannot contain two real roots and its two-circle region cannot contain any nonreal root. Thus $\mathtt{Var}(p,I) \le 1$ by Theorem 6. □

THEOREM 12. *The depth of the recursion tree is at most* $\log(w(I_0)/\sigma(p))$. *The total size of the recursion tree is* $O(n\log(w(I_0)/\sigma(p)))$.

*Proof.* The root of the recursion tree has an associated interval of length $w(I_0)$, every internal node has an associated interval of length at least $\sigma(p)$, and the interval associated with a node has half the length of the interval associated with the parent. Thus the depth of any internal node (the depth of the root is zero) is at most $\log(w(I_0)/\sigma(p))$.

At any level of the tree, we can have at most $n/2$ internal nodes. This holdss since the intervals are associated with the internal nodes at any level are disjoint and hence their sign variations add to at most $n$ by Theorem 7. Each internal node contributes at least two sign variations. Thus the number of internal nodes is at most $(1+\log(w(I_0)/\sigma(p)))n/2$.

The recursion tree is binary, i.e., each nonleaf has exactly two children. In such a tree the number of leaves is equal to the number of internal nodes plus one. Thus the total size of the tree is at most $1+(1+\log(w(I_0)/\sigma(p))n$. □

[[[*what was called $f_I$ before is now called $q_I$*]]]

The computation of $q_I$ from $p$ at every node of the recursion is costly. It is better to store with every interval $I = (a,b)$ the polynomial $p_I(x) := p(a+x(b-a))$, whose roots in $(0,1)$ correspond to the roots of $p$ in $I$. If $I$ is split at $m = (a+b)/2$ into $I_\ell = (a,m)$ and $I_r = (m,b)$, the polynomials associated with the subintervals are

$$p_{I_\ell}(x) = 2^n p_I(x/2) \quad \text{and} \quad p_{I_r}(x) = 2^n p_I((1+x)/2) = p_{I_\ell}(1+x).$$

Also, $q_I(x) = (1+x)^n p_I(1/(1+x))$. The polynomials $p_{I_\ell}$, $p_{I_r}$, and $q_I$ can be obtained from $p_I$ by $n^2$ additions. Also, if the coefficients are integral, the coefficients grow by $O(n)$ bits in every node.

[[todo: give more details and do complexity analysis]]

### 9.1.5 The Continued Fraction Algorithm for Root Isolation*

[[give a brief account of the continued fraction method.]]

## 9.2 Root Isolation for Polynomials with Real Coefficients

We extend the findings of the preceding section to polynomials with real coefficients. In principle, there is no need for extending. Algorithm 1 works perfectly for polynomials with real coefficients. There is a problem however. We need to determine the number of sign variatations in sequences of real numbers, e.g., in

$$(\pi, -\sqrt{2}, \sqrt{2}, \pi).$$

This is computationally hard. The sign determination of algebraic expressions is the topic of Lecture **??**.

We take a different route in this section. We assume that we can approximate the coefficients with any desired accuracy, i.e., for any coefficient $f_i$ and any integer $L$, we can compute a binary fraction $\tilde{f}_i = F_i/2^L$ with $F_i \in \mathbb{Z}$ and $\left|f_i - \tilde{f}_i\right| \leq 2^{-L}$, e.g., $F_i = \lfloor f_i 2^L \rfloor$ or $F_i = \lceil f_i 2^L \rceil$. Alternatively, we view the coefficients as binary numbers with potentially infinite binary places after the binary point; $\tilde{f}_i$ is then obtained by keeping the first $L$ digits after the binary point.

We pursue the following idea. In order to isolate the roots of $f = \sum_{0\leq i\leq n} f_i x^i$, we perform the following three steps:

(1) Approximate $f$ by a polynomial $\tilde{f} = \sum_{0\leq i\leq n} \tilde{f}_i x^i$, where each $\tilde{f}_i$ is a binary fraction approximating $f_i$.

(2) Isolate the roots of $\tilde{f}$ by means of the Algorithm 1.

(3) return the isolating intervals for the roots of $\tilde{f}$ (after a suitable widening) as isolating intervals for the roots of $f$.

Since the roots of a polynomial depend continuously on its coefficients such an approach might work; if $\tilde{f}$ is sufficiently close to $f$, the roots of $\tilde{f}$ should be good approximations for the roots of $f$. Thus if $I$ is an isolating interval for a root of $\tilde{f}$, a slightly widened $I$ might be an isolating interval for the corresponding root of $f$. In order to turn the idea into an algorithm, we need to overcome some obstacles.

- How well do we have to approximate the coefficients of $f$?
- Algorithm 1 may return isolating intervals sharing an endpoint. If we widen such intervals, we loose disjointness.

We overcome the second problem by modifying the algorithm slightly. Instead of recursing only for intervals $I$ with $\mathtt{Var}(p,I) \geq 2$, we recurse for all intervals with $\mathtt{Var}(p,I^+) \geq 2$, where $I^+$ is the interval of length $5w(I)$ enlarging $I$ by $2w(I)$ on either side, i.e., if $I=(a,b)$, $I^+ = (a-2(b-a), b+2(b-a))$. We call $I^+$ the *extension* of $I$ or an *extended interval*.

The small change ensures that isolated intervals are well-separated, see Lemma 16, without increasing the depth of the recursion by much, see Lemma 15. It has the nice side effect that the algorithm also computes an estimate of the root separation of the input polynomial, see Lemma 17. Before, we state and prove these Lemmas, we address item (1). We remarked above that the roots of a polynomial are continuous functions of the coefficients. Schönhage proved a quantitative version of this fact. For a polynomial $p = \sum_{0\leq i\leq n} p_i x^i$, $|p| = \sum_{0\leq i\leq n} |p_i|$ denotes the 1-norm of $p$.

THEOREM 13 ([**?**]). *Let $p = \sum_{0\leq i\leq n} p_i x^i = p_n \prod_{1\leq i\leq n}(x-z_i)$ be a polynomial of degree $n$ with $|z_i| < 1$ for all $i$. Let $\mu$ be a positive real with $\mu \leq 2^{-7n}$ and let $p^*(x) = \sum_{0\leq i\leq n} p_i^* x^i = p_n^* \prod_{1\leq i\leq n}(x-z_i^*)$ be such that*

$$|p-p^*| < \mu\,|p|\,.$$

*Then up to a permutation of the indices of the $z_i^*$*

$$|z_i^* - z_i| < 9\sqrt[n]{\mu}.$$

*Proof.* We prove the stronger claim $|z_i^* - z_i| \leq \sqrt[n]{\mu}$ under the additional assumptions $\sqrt[n]{\mu} < \sigma(p)/2$ and $|p-p^*| \leq \mu\,|p_n^*|$.

Let $z_i$ be a root of $p$, let $g(x) = p^*(z_i+x)/p_n^*$, and let $x_i$ be a root of $g$ of smallest modulus. Then $z_i + x_i$ is a root of $p^*$ and $g$ is monic. Since $g$ is monic, the product of its roots is equal (up to sign) to $g(0)$. Since $x_i$ is a root of $g$ of smallest modulus, we have $|x_i|^n \leq |g(0)|$. Thus

$$\begin{aligned}
|x_i| &\leq \sqrt[n]{|g(0)|} \\
&= \sqrt[n]{|p^*(z_i) - p(z_i)|\,/\,|p_n^*|} && \text{since } g(0) = p^*(z_i) \text{ and } p(z_i) = 0 \\
&\leq \sqrt[n]{\sum_{0\leq k\leq n} |p_k - p_k^*|\,/\,|p_n^*|\,|z_i|^k} \\
&\leq \sqrt[n]{\sum_{0\leq k\leq n} |p_k - p_k^*|\,/\,|p_n^*|} && \text{since } |z_i| \leq 1 \\
&\leq \sqrt[n]{\mu} && \text{since } |p-p^*| \leq \mu\,|p_n^*|
\end{aligned}$$

So for any root $z_i$ of $p$ there is a root $z_i^*$ of $p^*$ of distance at most $\sqrt[n]{\mu}$ and hence of distance less than $\sigma(p)/2$. Since any two roots of $p$ are at least $\sigma(p)$ apart, we have a bijection between the $z_i$ and the $z_i^*$. □

[[*my notation is not consistent: In the proof above, $\sigma(p)$ is the smallest distance between any two roots of $p$. The definition given in Subsection 9.1.4 is different. Either introduce two concepts or change the definition given in Subsection 9.1.4.*
*Also, what was called $\tilde{f}$ before is now called $p^*$.*]]

Theorem 13 requires all roots of $p$ to lie in the unit circle. It is easy to guarantee this condition. Let $P = \sum_{0\leq i\leq n} P_i x^i \in \mathbb{R}[x]$. By Theorem 1, the absolute value of all roots of $P$ is strictly bounded by

$$B := 2 \max_{0\leq i\leq n} \frac{|P_i|}{|P_n|}.$$

Let $p(x) := P(2Bx)$. Then all roots of $p$ have modulus less than 1/2. Thus Theorem 13 applies to $p$.

What is a good choice for $\mu$ and how can we determine a $p^*$ with $|p - p^*| \leq \mu |p|$? We will address these questions in two steps. We first assume that we know the root separation of $p$ (more precisely, have a lower bound for it) and then we show how to do without this assumption.

### 9.2.1 The Case of Known Root Separation

We choose $\mu$ such that the following three properties hold:

- $\mu \leq 2^{-7n}$; this is required by the Theorem.
- $9\sqrt[n]{\mu} \leq 1/2$; this guarantees that roots move by at most 1/2 and hence all roots of $p^*$ have modulus at most 1. Thus it is safe to start root isolation for $p^*$ with the start interval $(-1,+1)$.
- $9\sqrt[n]{\mu} \leq \sigma(p)/12$; this makes sure that roots move by at most $\sigma(p)/12$. Since the imaginary part of a nonreal root of $p$ is at least $\sigma(p)/2$, nonreal roots of $p$ become nonreal roots of $p^*$. Since real roots of $p$ have distance at least $\sigma(p)$ from each other, real roots of $p$ become real roots of $p^*$. Thus $p$ and $p^*$ have the same number of real roots. Also,

$$\sigma(p^*) \geq \sigma(p) - 2\sigma(p)/12 = 5\sigma(p)/6.$$

Similarly $\sigma(p^*) \leq \sigma(p) + 2\sigma(p)/12 = 7\sigma(p)/6$.

**Exercise 0.4:** Show that $\mu \leq 2^{-7n}$ implies $9\sqrt[n]{\mu} \leq 1/2$. ♢

LEMMA 14. *Assume $\sigma \leq \min(1/2, \sigma(p)$ and let $\mu = (\sigma/108)^n$. Then*

- $\mu \leq 2^{-7n}$, $9\sqrt[n]{\mu} \leq 1/2$, *and* $9\sqrt[n]{\mu} \leq \sigma(p)/12$.
- $\sigma(p^*) \geq 5\sigma(p)/6 \geq 90\sqrt[n]{\mu}$.

Now that we know how to choose $\mu$, we come to the choice of $p^*$. For each $i$, we determine a binary fraction $p_i^*$ with $|p_i^* - p_i| \leq 2^{-L}$ for a still to be determined $L$. Then $|p^* - p| \leq (n+1)2^{-L}$. We need $|p^* - p| \leq \mu |p|$ and therefore choose $L$ such that

$$2^{-L} \leq \frac{\mu |p|}{n+1}.$$

We can now state the algorithm, see Algorithm 2. The only difference to Algorithm 1 is that we now recurse whenever $\texttt{Var}(p^*, I^+) \geq 2$.

LEMMA 15. *Algorithm 2 generates no interval of length less than $\sigma(p^*)/10$.*

*Proof.* Consider any interval $I$ with $w(I) \leq \sigma(p^*)/5$. Then $w(I^+) \leq \sigma(p^*)$ and hence either the one- or the two-circle theorem applies to $I^+$. Thus $\texttt{Var}(p^*, I^+) \leq 1$ and $I$ is not split. Thus only intervals with length $\geq \sigma(p^*)/5$ are split and hence no interval has length less than $\sigma(p^*)/10$. □

Let $O^*$ be the list of isolating intervals computed for $p^*$. Any interval in $O^*$ is either a singleton or has length at least $\sigma(p^*)/10$. For an interval $I$, let $\tilde{I}$ be its expansion by $9\sqrt[n]{\mu}$ on both sides, i.e, if $I = (a,b)$, then $\tilde{I} = (a - 9\sqrt[n]{\mu}, b + 9\sqrt[n]{\mu})$ and if $I = [m,m]$, then $\tilde{I} = (m - 9\sqrt[n]{\mu}, m + 9\sqrt[n]{\mu})$. If $I$ is an isolating interval for a real root of $p^*$, $\tilde{I}$ contains the corresponding root of $p$.

**Algorithm 2** Bisection Algorithm for a Real Polynomial $p$ with Root Separation Estimate $\sigma$

**Require:** $p$ is a real polynomial with roots in the disc of radius 1/2 centered at 0, $\sigma \leq \min(1/2, \sigma(p))$, $\mu = (\sigma/108)^n$, $|p_i^* - p_i| \leq \mu |p| /(n+1)$ for all $i$.

**Ensure:** returns a list $O^*$ of well-separated isolating intervals for the real roots of $p^*$.

$A := \{(-1,1)\}$ {list of active intervals}
$O^* := \emptyset$ {list of isolating intervals}
**repeat**
  $I :=$ some interval in $A$; delete $I$ from $A$;
  $I^+ = (a-2(b-a), b+2(b-a))$, where $I = (a,b)$;
  **if** $\mathtt{Var}(p^*, I^+) > 1$ **then**
    add $(a,m)$ and $(m,b)$ to $A$ where $m = (a+b)/2$;
    **if** $p^*(m) = 0$ add $[m,m]$ to $O^*$;
  **else**
    **if** $\mathtt{Var}(p^*, I) = 0$ do nothing;
    **if** $\mathtt{Var}(p^*, I) = 1$ add $I$ to $O^*$;
  **end if**
**until** $A$ is empty
**return** $O^*$

LEMMA 16. *$O := \{\tilde{I} \mid I \in O^*\}$ is a set of isolating intervals for $p$.*

*Proof.* By our choice of $\mu$, $p$ and $p^*$ have the same number of real roots and each expanded interval contains a real root of $p$. We need to argue disjointness.

Let $I$ and $J$ be two intervals in $O^*$. If $I$ and $J$ are singletons, they have distance at least $\sigma(p^*)$ from each other. Since $\sigma(p^*) \geq 90\sqrt[n]{\mu}$, disjointness is preserved after expanding both intervals.

So assume, that at least one of the intervals is not a singleton, say $I$. We may also assume $w(I) \geq w(J)$. Since $I$ and $J$ are in $O^*$, both contain a real root of $p^*$. If $I^+$ would contain $J$, it would contain two real roots, and we would have $\mathtt{Var}(p^*, I^+) \geq 2$. So $I$ would be split. Thus $I^+$ does not contain $J$ and hence is disjoint from $J$ (since $w(I) \geq w(J)$). Thus the distance of $I$ and $J$ is at least $2w(I)$. Also,

$$2w(I) \geq \frac{\sigma(p^*)}{5} \geq 18\sqrt[n]{\mu}$$

by our choice of $\mu$ and hence $\tilde{I}$ and $\tilde{J}$ are disjoint. □

This concludes the analysis of modified algorithm. It computes isolating intervals for $p$ and its recursion depth is not much larger than for the original algorithm. Its drawback is that it needs an estimate $\sigma$ with $\sigma \leq \sigma(p)$.

### 9.2.2 The Case of Unknown Root Separation

The nice thing is that the algorithm computes such an estimate $\sigma$.

LEMMA 17. *Algorithm 2 refines at least one interval to a length less than $n\sigma(p^*)$.*

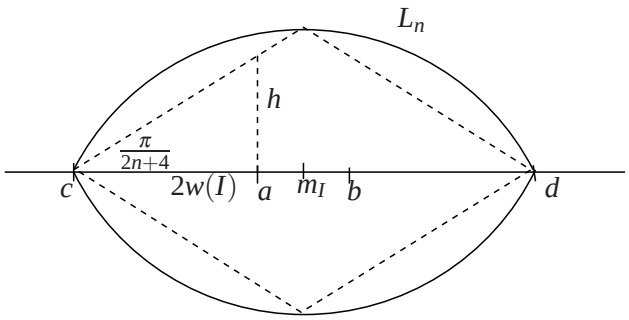


Figure 9.4: $I=(a,b)$, $I^+=(c,d)$, and $L_n$ denotes the Obreshkoff lens $L_n(I^+)$. The height of $L_n$ at endpoint of $I$ is at least $h$, where $h=2w(I)\tan(\pi/(2(n+2)))\geq 2w(I)/(n+2)\geq w(I)/n$. By Theorem 5, $\mathtt{Var}(p^*,I^+)$ is at least the number of roots of $p^*$ in the rectangle $I\times[-hi,+hi]$.

*Proof.* We distinguish cases. If $\sigma(p^*)$ is equal to the distance of two real roots, let $I$ be the separating interval computed for one of them. Then $w(I)\leq\sigma(p^*)/2$ because otherwise $I^+$ would contain both roots and $I$ would be split.

If $\sigma(p^*)$ is equal to the imaginary coordinate of a nonreal root, consider a leaf $I$ of the Descartes tree with the property that the real part of the root is contained in the closure of $I$. Since $I$ is a leaf, $\mathtt{Var}(p^*,I^+)\leq 1$ and hence $\sigma(p^*)\geq w(I)/n$, see Figure 9.4. □

**Exercise 0.5:** Show that Lemma 17 is not true for Algorithm 1). Consider $p(x)=x^2+\delta^2=(x-\mathbf{i}\delta)(x+\mathbf{i}\delta)$ with $\delta\approx 0$. This polynomial has a pair of conjugate complex roots at $\pm\mathbf{i}\delta$ and hence separation $2\delta$. However, $\mathtt{Var}(p,(-1,1))=2$ and $\mathtt{Var}(p,(-1,0))=\mathtt{Var}(p,(0,1))=0$. Verify these statements. Thus the algorithm ends with intervals of length $1/2$, although the separation may be arbitrarily small. ♢

**Exercise 0.6:** In Algorithm 2 we split all intervals $I$ satisfying $\mathtt{Var}(p,I^+)\geq 2$. Is there a less agressive rule, which still guarantees a variant of Lemma 17. The question is how to handle intervals with $\mathtt{Var}(p^*,I)\leq 1$ and $\mathtt{Var}(p^*,I^+)\geq 2$. ♢

Lemma 17 yields a simple method for verifying whether a guess $\sigma$ is no larger than $\sigma(p)$.

LEMMA 18. *If Algorithm 2 produces no interval of length less than $2n\sigma$ then $\sigma\leq\sigma(p)$.*

*Proof.* The algorithm produces an interval of length at most $n\sigma(p^*)$ and hence $2n\sigma\leq n\sigma(p^*)$. Also, $\sigma(p^*)\leq\sigma(p)+18\sqrt[n]{\mu}=\sigma(p)+\sigma/6$. Thus

$$2\sigma\leq\sigma(p^*)\leq\sigma(p)+\sigma/6$$

and hence $\sigma\leq\sigma(p)$. □

We can now state the complete algorithm for isolating the roots of a polynomial with bitstream coefficients. We start with an initial guess $\sigma=1/2$; then $(\sigma/108)^n\leq 2^{-7n}$. We compute $\mu$ and $p^*$ and run Algorithm 2 on $p^*$. If no interval of length less than $2n\sigma$ is produced, we have $\sigma\leq\sigma(p)$ and the algorithm returns isolating intervals for $p$. On the other hand, if an interval of length less than $2n\sigma$ is produced, we take this as an indication that our current guess is too large. We replace $\sigma$ by $\sigma^2$ and repeat. We obtain Algorithm 3. It remains to estimate how small $\sigma$ can become.

**Algorithm 3** Bisection Algorithm for Real Polynomials

**Require:** $p = \sum_{0 \le i \le n} p_i x^i$ and all roots of $p$ lie in a disc of radius 1/2 centered at 0. Real roots are distinct.

**Ensure:** returns isolating intervals for the real roots of $p$.

$\sigma = 1/2$;
$\mu = (\sigma/108)^n$;
**while** (**true**) **do**
  choose $\varepsilon \le \mu\,|p|\,/(n+2)$ and let $p^*$ be such that $|p_i^* - p_i| \le \mu\,|p|\,/(n+1)$ for all $i$;
  run Algorithm 2 on $p^*$ and start interval $I = (-1,1)$; // we do not guarantee $\mu \le (\sigma(p)/108)^n$
  **if** the algorithm does not produce an interval of length less than $2n\sigma$ **then**
    exit from the loop;
  **else**
    $\sigma = \sigma^2$;
  **end if**
**end while**
return $O := \{\tilde{I} \mid I \in O^*\}$

LEMMA 19. *Algorithm 3 stops with*

$$\sigma \ge \min\left(\frac{1}{2}, \left(\frac{\sigma(p)}{20n+1}\right)^2\right).$$

*Proof.* If the algorithm stops in the first iteration, it stops with $\sigma = 1/2$. If the algorithm performs more than one iteration, consider the next to last iteration. An interval of length less than $2n\sigma$ is produced. On the other hand, by Lemma 15, no interval of length less than $\sigma(p^*)/10$ is generated. Thus $2n\sigma \ge \sigma(p^*)/10$ and hence $20n\sigma \ge \sigma(p^*) \ge \sigma(p) - 18\sqrt[n]{\sigma} \ge \sigma(p) - \sigma/6$ and finally $\sigma \ge \sigma(p)/(20n+1)$.

Since $\sigma$ is squared from one iteration to the next, we have $\sigma \ge (\sigma(p)/(20n+1))^2$ in the last iteration. □

## 9.3 Further Reading

See [**?**, **?**] for extensive treatments and references.