7 Real Roots of Polynomials

There are many applications which require the solution in real numbers for equations of the form

$$p(x) = 0$$

where $p(x) \in \mathbb{Q}[x]$ (for example ray tracing in Computer Graphics). If p has degree at most 4 then its roots can be expressed by means of a formula. However for degree 4 the formula is not much use because roots which are real can appear to be complex until a significant amount of simplification is carried out (the same holds for degree 3). For polynomials of degree 5 and higher it has been known since the early 1800's that no formula exists. (By formula we mean an expression involving only the coefficients of p in which we can use the basic arithmetic operations as well as taking powers, square roots, cube roots etc.) We are therefore forced to use an algorithmic approach. An important division of the problem was made by Fourier (of the eponymous transforms):

Isolation: we find disjoint open intervals such that each one contains exactly one real root of p and each real root of p is contained in one of the intervals,

Approximation: we shrink each interval so as to approximate the root it contains to the desired degree of accuracy.

Approximation is by far the easier of the two problems so we deal with it first. Throughout this section we are concerned with methods which are *guaranteed* to produce the required results. This rules out methods such as Newton–Raphson iteration. In order to maintain exactness with relative ease we always use rational arithmetic and so all our intervals are of the form (a, b) with a, b rational numbers.

7.1 Real Root Approximation

It is well known that if $p(x)$ does not have a root in the interval (a, b) then it does not change sign in this interval (i.e., either $p(\xi) > 0$ for all $a < \xi < b$ or $p(\xi) < 0$ for all $a < \xi < b$) . This fact is intuitively obvious and is a consequence of the continuity of polynomial functions over the real numbers (think of the graph of $p(x)$). If, on the other hand, $p(x)$ happens to have a root in (a, b) then it does not necessarily change sign. For example x^2 has the root 0 in $(-1, 1)$ but clearly it does not change sign. Note that 0 has multiplicity 2 as a root of x^2. Furthermore the polynomial x has precisely the same roots as x^2 and x *does* change sign in $(-1, 1)$. This simple example is not an isolated case. We define $p(x)$ to be *square free* if it cannot be written as

$$p(x) = f(x)^2 g(x)$$

where $f(x) \in \mathbb{Q}[x]$ is non-constant. Take care to understand this definition in full: what is really says is that $p(x)$ has no repeated non-constant factors (if $f(x)$ occurs $e > 1$ times as a factor then $f(x)^2$ is a factor of $p(x)$). Clearly the roots of $f(x)^2 g(x)$ are the same as those of $f(x)g(x)$ and we could proceed further to remove any other squares in $f(x)g(x)$. Thus every polynomial has precisely the same roots as some square free polynomial (called its *square free part*). Now if $p(x)$ is square free and has a root α then for a sufficiently small ϵ the sign of $p(x)$ in $(\alpha - \epsilon, \epsilon)$ is opposite to its sign in $(\alpha, \alpha + \epsilon)$. This is easy to see since, the fact that α is a root of $p(x)$ means that

$$p(x) = (x - \alpha)q(x)$$

for some non-zero polynomial $q(x)$. Furthermore since $p(x)$ is square free it follows that α is not a root of $q(x)$. Since $q(x)$ has only finitely many roots, none of which is α, it follows that there is an ϵ such that the sign of $q(x)$ in $(\alpha - \epsilon, \alpha + \epsilon)$ does not change. The claim now follows. It is now easy to see that if $p(x)$ is square free then it has a root in (a, b) if and only if it changes sign in this interval.

Suppose that (a, b) isolates a real root of $p(x)$. Then the following algorithm can be used to shrink (a, b):

Algorithm: $APPROX(p(x), a, b, \epsilon) \mapsto A$

(A is either the exact root of p contained in (a, b) or an interval (c, d) which isolates the same root and satisfies $d - c < \epsilon$.)

1. $q(x) := SQFPART(p)$;

2. **if** $q(a) = 0$ **then** $q(x) := q(x)/(x - a)$ **fi**;

3. **if** $q(b) = 0$ **then** $q(x) := q(x)/(x - b)$ **fi**;

4. $c := a$;
$d := b$;
$m := (c + d)/2$;

5. **while** $d - c \geq \epsilon$ **do**
if $q(m) = 0$ **then return** m **fi**;
if $sign(q(c)) \neq sign(q(m))$ **then** $d := m$; $m := (c + d)/2$
else $c := m$; $m := (c + d)/2$
fi
od;

6. $A := (c, d)$

We make some remarks concerning the algorithm.

1. *SQFPART* returns the square free part of its argument, we discuss an algorithm for this below.

2. The second and third steps may look a little strange. They are a consequence of our decision to use *open* intervals for root isolation. This means that, for instance, $(1, 3)$ is an isolating interval for the root 2 of $(x - 1)(x - 2)(x - 3)$. This difficulty is avoided if we use *closed* intervals but then we have difficulties elsewhere!

   Now α is a root of q if and only if $x - \alpha$ divides q (see §4.7.7) and so $q/(x - \alpha)$ is a polynomial. Since q is square free it follows that α is not a root of $q/(x - \alpha)$; all other roots of q are of course also roots of $q/(x - \alpha)$.

3. Obviously *sign* tells us if a number is positive or negative. (It doesn't matter here what we do about 0 here because we never use *sign* with 0 as argument, but a typical definition is to define $sign(r)$ to be $+1$ if $r > 0$, 0 if $r = 0$ and -1 if $r < 0$.)

Exercise 7.1 *Let*

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$$

where $a_m \neq 0$. *Show that*

1. *if* m *is odd then* $p(x)$ *must have at least one real root.*
2. *If* m *is even and* $a_m a_0 < 0$ *then* $p(x)$ *must have at least two real roots (one positive and the other negative).*

Hint: think in terms of the graph of p.

7.1.1 Square Free Decomposition

Given $p(x) \in \mathbb{Q}[x]$ put

$$p(x) = \prod_{i=1}^{r} p_i(x)^i.$$

where the $p_i(x)$ are pairwise coprime and square free. (Of course some of the p_i may be 1, e.g., $x^4 + 3x^3 + 3x^2 + x = x^1 \cdot 1^2 \cdot (x+1)^3$.) We wish to compute the square free part of $p(x)$, i.e. $p_1(x)p_2(x)\cdots p_r(x)$. We proceed as follows. Differentiating $p(x)$ we have:

$$p'(x) = \sum_{i=1}^{r} i p_i(x)^{i-1} p_i'(x) \prod_{j \neq i} p_j(x)^j,$$

so that

$$r(x) = \gcd(p(x), p'(x)) = \prod_{i=2}^{r} p_i(x)^{i-1}.$$

Thus

$$t(x) = p(x)/r(x) = \prod_{i=1}^{r} p_i(x),$$

which is the square free part as required.

In many applications it is useful to know the p_i themselves (e.g., this enables us to know the multiplicity of a root). Put

$$v(x) = \gcd(r(x), t(x)) = \prod_{i=2}^{r} p_i(x),$$

so that

$$p_1(x) = t(x)/v(x).$$

Repeating the process with $r(x)$ in place of $p(x)$ gives us $p_2(x)$ etc.

We have therefore obtained a very efficient method of computing the square free part of $p(x)$ as well as its square free decomposition. Before moving on we must address a potential difficulty. Our algorithm works entirely in $\mathbb{Q}[x]$ whereas we require the square free part of $p(x)$ as an element of $\mathbb{R}[x]$. At first sight there is no reason to believe that the answer in $\mathbb{Q}[x]$ is the same as that in $\mathbb{R}[x]$. After all there are plenty of properties of polynomials which change as we move from $\mathbb{Q}$ to $\mathbb{R}$, e.g., $x^2 - 2$ has no roots in $\mathbb{Q}$ but has two roots in $\mathbb{R}$. Fortunately we have:

Theorem 7.1 *Let D, D' be integral domains with D a subdomain of D' (i.e., D is a subset of D' and the binary operations on D are just those of D' restricted to D). Suppose that $f, g \in D[x]$ have a non-constant common factor in $D'[x]$, then they have a non-constant common factor in $D[x]$.*

Proof Thinking of f, g as elements of $D'[x]$ we see from Theorem 5.5 that $\text{Res}_x(f,g)$ must vanish. But $\text{Res}_x(f,g)$ is an element of $D[x]$ since it is a polynomial in the coefficients of f, g which are all in D. It now follows from Theorem 5.5 that f, g have a non-constant common factor in $D[x]$ as claimed[9]. □

Exercise 7.2 *Use the preceding theorem to show that the square free decompositions of $p(x)$ in $\mathbb{Q}[x]$ and $\mathbb{R}[x]$ are the same up to a constant factor.*

Exercise 7.3 *Show that $p(x), q(x) \in \mathbb{Q}[x]$ have a common root (real or complex) if and only if* gcd$(p(x), q(x))$ *has such a root where the* gcd *is computed in $\mathbb{Q}[x]$.*

7.2 Real Root Isolation

Suppose we have an algorithm *RCOUNT* which takes a polynomial $p(x)$, the endpoints of an interval (a,b) and returns the number of real roots of $p(x)$ in the interval. We can then use this to isolate the roots of $p(x)$ which lie in (a,b) as follows:

Algorithm: $ISOL(p(x), a, b) \mapsto [E, A]$

($p(x)$ is square free. E is a list of (some of the) exact roots of $p(x)$ which lie in (a,b) and A is a list of isolating intervals for the rest of the roots of $p(x)$ in (a,b).)

1. $E := [\,]$;
$A := [\,]$;
$r := RCOUNT(p(x), a, b)$;

2. **if** $r = 0$ **then return** $[[\,], [\,]]$
elif $r = 1$ **then return** $[[\,], [(a,b)]]$
fi;

3. $W := [[a, b, r]]$; (this holds intervals to be explored further)

4. **while** $W \neq [\,]$ **do**
remove the first element $[c, d, r]$ from W;
$m := (c + d)/2$;
if $p(m) = 0$ **then**
$E := append(E, [m])$;
$p(x) := p(x)/(x - m)$ (more efficient to work with lower degree polynomials!)
fi;
$r := RCOUNT(p(x), c, m)$;

[9] The proof of Theorem 5.5 assumed that we were working over a UFD. It is not hard to show that the result holds when the coefficients of the polynomial come from an integral domain.

if $r = 1$ **then** $A := append(A, [(c, m)])$
elif $r > 1$ **then** $W := append(W, [[c, m, r]])$
fi;
$r := RCOUNT(p(x), m, d)$;
if $r = 1$ **then** $A := append(A, [(m, d)]$
elif $r > 1$ **then** $W := append(W, [[m, d, r]])$
fi;
od

It is worth noting that the algorithm presented above is biased towards clarity of presentation rather than efficiency. In practice we would not make all those calls to $RCOUNT$ but do the computations in place and so avoid unnecessary recomputations (some of which can be very expensive).

In order to isolate all the roots of $p(x)$ we need to be able to find an interval which contains all of them. For this it suffices to find a number which is strictly bigger than the absolute values of the roots of $p(x)$.

Theorem 7.2 (Cauchy) *Let*

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$$

be a polynomial with complex coefficients where $m \geq 1$ and $a_m \neq 0$. Then any root α of p satisfies

$$|\alpha| < 1 + \frac{\max\{\,|a_0|, \ldots, |a_{m-1}|\,\}}{|a_m|}.$$

Proof If $|\alpha| \leq 1$ then the result is trivially true. So suppose that $|\alpha| > 1$ and put

$$M = \max\{\,|a_0|, \ldots, |a_{m-1}|\,\}.$$

Since α is a root of p we have

$$a_m\alpha^m = -a_{m-1}\alpha^{m-1} - \cdots - a_0$$

so that

$$|a_m||\alpha|^m \leq M(|\alpha|^{m-1} + \cdots + |\alpha| + 1) < \frac{M|\alpha|^m}{|\alpha| - 1},$$

and

$$|a_m|(|\alpha| - 1) < M$$

which proves the result. □

One unattractive feature of this is that if we replace x by $x/2$, which changes the roots by a factor of 2 only, the bound might change by as much as 2^m. Note also that the bound is for *all* the roots not only the real ones (recall that $|a + b\sqrt{-1}| = \sqrt{a^2 + b^2}$. The following avoids the defect:

Theorem 7.3 (Cauchy) *Let*

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$$

be a polynomial with real coefficients where $m \geq 1$, $a_m > 0$ and which has $\lambda > 0$ strictly negative coefficients. Put

$$B = \max\left\{ \left| \lambda \frac{a_{m-i}}{a_m} \right|^{1/i} \;\middle|\; 1 \leq i \leq m \;\&\; a_{m-i} < 0 \right\}.$$

Then every positive real root of p is no larger than B.

Proof Suppose $b > B$. We have

$$b^i > \lambda \left| \frac{a_{m-i}}{a_m} \right|$$

for each i such that $a_{m-i} < 0$. The inequality can also be written as

$$b^m > \lambda \left| \frac{a_{m-i}}{a_m} \right| b^{m-i}$$

and so

$$\lambda a_m b^m > \lambda \sum_{\substack{1 \leq i \leq m \\ a_{m-i} < 0}} |a_{m-i}| b^{m-i}.$$

or

$$a_m b^m > \sum_{\substack{1 \leq i \leq m \\ a_{m-i} < 0}} |a_{m-i}| b^{m-i}.$$

It follows from this that $p(b) > 0$ and so b is not a root of $p(x)$. □

Of course we can obtain a bound on the absolute values of the negative roots of $p(x)$ by applying the result to $p(-x)$ (if $p(x)$ has odd degree then $p(-x)$ will have negative leading coefficient—we rectify this by using $-p(-x)$). In this way we can obtain bounds B_N, B_P such that every root of $p(x)$ is in the closed interval $[B_N, B_P]$ and so it must be in the open interval $(B_N - 1, B_P + 1)$.

Exercise 7.4 *At first sight the quantity B of the preceding theorem seems to require the extraction of roots and so would be costly to implement. Show how to compute efficiently a good upper bound to B using only rational arithmetic. (For a solution see Akritas* [2, **pp. 350–352**]*.)*

7.2.1 Counting Real Roots of Univariate Polynomials

We start with some useful definitions. Let $y_0, y_1, \ldots, y_m$ be a sequence of real numbers. We define the *variation* in the sequence to be the number of times the sign changes from one number to the next, ignoring any occurrences of 0. Here are some examples:

Sequence	Variation
2, −1, −2, 3, 3	2
2, 0, −1, 0, 0, −2, 3, 3,	2
−1, −10, 0, −3, 0, 1, 2, 0, 3	1
0, 3, −1, 2, 0, −4, 5, 0, −6	5

Now let $S = p_0(x), p_1(x), \ldots, p_m(x)$ be a sequence of polynomials and a a real number. We define the *variation* of S at a to be the variation of the sequence $p_0(a), p_1(a), \ldots, p_m(a)$. This number is denoted by $V_S(a)$. For example let

$$S = x^3 - 7x + 7,\ 3x^2 - 7,\ 2x - 3,\ 1.$$

Then:

x	Sequence				$V_S(x)$
-1	13,	-4,	-5,	1	2
0	7,	-7,	-3,	1	2
1/2	29/8,	$-25/4$,	-2,	1	2
1	1,	-4,	-1,	1	2
3/2	$-1/8$,	$-1/4$,	0,	1	1
2	1,	5,	1,	1	0

Note that $V_S(x)$ cannot change as x varies unless x passes through a root of some polynomial of S. Clearly this important observation is true for any sequence of polynomials.

Recall that a non-zero polynomial can have only finitely many roots. Suppose from now on that all the polynomials in the sequence S are non-zero. It follows that only finitely many real numbers can be the roots of any one of the polynomials in S. In particular any interval (a, b) can contain only finitely many numbers which are the roots of some polynomial in S. We may therefore subdivide the interval (a, b) into finitely many subintervals $(a_1, a_2), (a_2, a_3), \ldots, (a_{n-1}, a_n)$ such that each subinterval contains at most one root and each root is in one of the intervals. Note that

$$V_S(a) - V_S(b) = \sum_{i=1}^{n-1} V_S(a_i) - V_S(a_{i+1}).$$

Suppose that

$$V_S(a_i) - V_S(a_{i+1}) \geq 1$$

whenever (a_i, a_{i+1}) contains a root of some polynomial in S. It then follows that the number of such roots in (a, b) is *at most* $V_S(a) - V_S(b)$. Furthermore suppose that for a given polynomial $p(x)$ we can find a sequence S with the stronger property that

$$V_S(a_i) - V_S(a_{i+1}) = \begin{cases} 1 & \text{if } p(x) \text{ has a root in } (a_i, a_{i+1}), \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

It then follows that $p(x)$ has *exactly* $V_S(a) - V_S(b)$ real roots in (a, b). It is a remarkable fact that such a sequence can be found for any square free polynomial $p(x)$.

We now assume that $p(x)$ is square free and consider the sequence

$$\text{Sturm}(p) = p_0(x), p_1(x), \ldots, p_n(x)$$

defined by

$$\begin{aligned} p_0(x) &= p(x), \\ p_1(x) &= p'(x), \\ p_i(x) &= -\,\text{remainder}(p_{i-2}(x), p_{i-1}(x)), \quad \text{for } i \geq 2. \end{aligned}$$

where 'remainder' means the remainder from polynomial division. Thus we simply keep applying the Euclidean algorithm but take negative remainders. We terminate the sequence when a non-zero

constant is reached (clearly the sequence must stop after at most $\deg(p(x))$ steps). In order to see that we must eventually reach a non-zero constant note that, by the definition of the sequence, we have

$$\begin{aligned} p_0(x) &= p_1(x)q_1(x) - p_2(x) \\ p_1(x) &= p_2(x)q_2(x) - p_3(x) \\ &\vdots \\ p_{n-2}(x) &= p_{n-1}(x)q_{n-1}(x) - p_n(x) \end{aligned} \tag{12}$$

and it follows from this that $p_n(x)$ is a gcd of $p_0(x)$ and $p_1(x)$, i.e., of $p(x)$ and $p'(x)$. But by assumption $p(x)$ and $p'(x)$ do not have a non-constant common factor. In fact this argument shows something more general: we could define a sequence like Sturm(p) for an arbitrary polynomial $p(x)$—the termination condition is that we stop just before a zero remainder is obtained. We then have that $p_n(x)$ is a gcd of $p(x)$ and $p'(x)$ so that $p(x)/p_n(x)$ is the square free part of $p(x)$ up to a non-zero constant multiple. Moreover if we put

$$q_i(x) = \frac{p_i(x)}{p_n(x)}, \quad \text{for } 1 \le i \le n,$$

then the sequence $q_1(x), q_2(x), \ldots, q_n(x)$ can be used as the Sturm sequence of the square free part of $p(x)$.

The sequence just defined has the following properties.

1. If α is a real root of $p(x)$ then for a sufficiently small ϵ we have that $p(x)$ and $p'(x)$ have opposite sign for all values of x in $(\alpha - \epsilon, \alpha)$ and the same sign in $(\alpha, \alpha + \epsilon)$. For a formal proof we write

   $$p(x) = (x - \alpha)^m q(x)$$

   where $m > 0$ and α is not a root of $q(x)$. Then

   $$p'(x) = m(x - \alpha)^{m-1} q(x) + (x - \alpha)^m q'(x).$$

   Thus

   $$\begin{aligned} p(\alpha + \delta) &= \delta^m q(\alpha + \delta), \\ p'(\alpha + \delta) &= m\delta^{m-1} q(\alpha + \delta) + \delta^m q'(\alpha + \delta). \end{aligned}$$

   The claim now follows from the observation that δ^m converges to 0 much faster than δ^{m-1} for sufficiently small δ so that the sign of $p'(\alpha + \delta)$ is determined by that of δ^{m-1}. However the sign of $p(\alpha + \delta)$ is determined by that of δ^m. (The fact that α is not a root of $q(x)$ means that $q(\alpha + \delta)$ does not change sign for sufficiently small values of δ .)

   It is worth noting that this argument does not rely on $p(x)$ being square free. (An intuitive explanation can be found by drawing the possible shapes of the graph $y = p(x)$ near places where it cuts or touches the x-axis.)

2. Two consecutive elements of the sequence cannot vanish at the same point. For if $p_i(\alpha) = p_{i+1}(\alpha) = 0$ then it follows immediately from (12) that $p_n(\alpha) = 0$ but this is impossible since $p_n(x)$ is a non-zero constant.

3. If $p_i(\alpha) = 0$ for some i with $0 < i < n$ then $p_{i-1}(\alpha)$ and $p_{i+1}(\alpha)$ have opposite signs. From (12) we have
$$p_{i-1}(\alpha) = p_i(\alpha)q_i(\alpha) - p_{i+1}(\alpha)$$
so that
$$p_{i-1}(\alpha) = -p_{i+1}(\alpha)$$
and we know from the preceding part that neither of these two numbers can be zero.

Now let (a, b) be any interval such that neither a nor b is a root of $p(x)$ with $a \leq b$. If $a = b$ the claim ios trivial so we may assume that $a < b$. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the (possibly empty) sequence of all real roots of the polynomials in the sequence Sturm(p) which are in the interval (a, b). As above we may partition (a, b) into finitely many subintervals $(a_1, a_2), (a_2, a_3), \ldots, (a_{s-1}, a_s)$ such that each subinterval contains at most one root α_j (and none of the a_i is one of these roots). Since variations do not change unless we pass through a root we may assume that when an interval (a_i, a_{i+1}) contains a root α_j then a_i, a_{i+1} are as close to α_j as we please. Now if an interval (a_i, a_{i+1}) has a root of $p(x)$ which is not a root of any other polynomial in Sturm(p) then it follows from the first property given above that
$$V_S(a_i) - V_S(a_{i+1}) = 1$$
where $S = \text{Sturm}(p)$. If (a_i, a_{i+1}) does not have a root α_j then it is clear that
$$V_S(a_i) - V_S(a_{i+1}) = 0.$$
We now examine the case when (a_i, a_{i+1}) contains a root α_j of a polynomial p_k with $k \geq 1$. It follows from the third property that $p_{k-1}(\alpha_j)$ and $p_{k+1}(\alpha_j)$ have opposite signs and in particular $p_{k-1}(x)$, $p_{k+1}(x)$ do not change sign in the interval (a_i, a_{i+1}) (why?). We may now construct the following table of possible sign changes:

x	p_{k-1}	p_k	p_{k+1}	p_{k-1}	p_k	p_{k+1}
$a_i \leq x < \alpha_j$	$+$	$\pm$	$-$	$-$	$\pm$	$+$
α_j	$+$	0	$-$	$-$	0	$+$
$\alpha_j < x \leq a_{i+1}$	$+$	$\mp(\pm)$	$-$	$-$	$\mp(\pm)$	$+$

Now if α_j is not a root of $p(x)$ then it follows from this table that
$$V_S(a_i) - V_S(a_{i+1}) = 0.$$
On the other hand if α_j is also a root of $p(x)$ then the first property given above and the table together show that
$$V_S(a_i) - V_S(a_{i+1}) = 1.$$

We have therefore verified the following: let $p(x)$ ibe square free and a, b two real numbers which are not roots of $p(x)$ with $a \leq b$. Then the number of real roots of $p(x)$ in the interval (a, b) is $V_S(a) - V_S(b)$ where $S = \text{Sturm}(p)$.

In fact we can generalise this a little, we can drop the assumption that $p(a) \neq 0$ and $p(b) \neq 0$ in which case $V_S(a) - V_S(b)$ counts the number of roots of p in the half closed interval $(a, b]$. To see this note that for a sufficiently small ϵ the number of roots of p in $(a, b]$ is the same the number of them in $(a + \epsilon, b + \epsilon)$. If a is not a root of p then for all sufficiently small ϵ it is clear that $\text{V}_S(a+\epsilon) = \text{V}(a)$. If, on the other hand $p(a) = 0$ then the same fact follows form the first property above. Of course the same holds for $V(b + \epsilon)$ and $\text{V}(b)$. We have thus proved so we have proved

Theorem 7.4 (Sturm, 1835) *Let $p(x)$ be square free and let a, b be two real numbers with $a \leq b$. Then the number of real roots of $p(x)$ in the interval $(a, b]$ is $V_S(a) - V_S(b)$ where $S = \text{Sturm}(p)$.*

We can use this theorem to find the total number of real roots of $p(x)$ as follows. Consider an interval $(-a, a)$ where $a > 0$. For all sufficiently large a, the interval contains all the real roots of $p(x)$ so that the number of them is given by $V_S(-a) - V_S(a)$. Now for any polynomial $f(x)$ and all sufficiently large a the sign of $f(a)$ is just the sign of $\text{lc}(f)$ while the sign of $f(-a)$ is that of $\text{lc}(f)$ if $f(x)$ has even degree and that of $-\text{lc}(f)$ if $f(x)$ has odd degree. So to find $V_S(a)$ for a sufficiently large a we just find the variation of the sequence of leading coefficients of the Sturm sequence. To find $V_S(-a)$ we look at the same sequence but negate those leading coefficients which come from polynomials of odd degree. We summarize this result as follows: we define $V_S(-\infty)$ to be the sign variation of $V_S(v)$ as $v \to -\infty$. We know that this definition makes sense from the preceding discussion, that is, $V_S(v)$ has the same value for all large enough negative values of v. We define $V_S(\infty)$ similarly. Our discussion then shows that the number of roots of f is $V_S(-\infty) - V_S(\infty)$ and moreover we can find this quantity just from the leading terms of the Sturm sequence (indeed all we need is the sign of each leading coefficient and the degree).

If we examine the above discussion carefully we see that it depends on having a sequence $S = p, p_1, \ldots, p_n$ with the following properties:

1. If α is a real root of $p(x)$ then for a sufficiently small ϵ we have that $p(x)$ and $p_1(x)$ have opposite sign for all values of x in $(\alpha - \epsilon, \alpha)$ and the same sign in $(\alpha, \alpha + \epsilon)$.

2. Two consecutive members of the sequence cannot have a common root.

3. If some $p_i(\alpha) = 0$ for some i with $0 < i < n$ then $p_{i-1}(\alpha)$ and $p_{i+1}(\alpha)$ have opposite signs.

4. The last polynomial p_n does not vanish in the interval (a, b) and so it keeps constant sign.

Of course $\text{Sturm}(p)$ satisfies these conditions. Surprisingly it turns out that many other sequences which have the property can be constructed, see Akritas [2] (actually this is not really as surprising as it might seem—why?).

Exercise 7.5 *Find, by hand calculations only, the Sturm sequence of $x^5 - x + 1$ and isolate its real root(s). Use Axiom to check your answer (the relevant functions are* `sturmSequence` *and* `realZeros`*).*

Exercise 7.6 *Let $p(x)$ be a polynomial of degree $n > 0$, not necessarily square-free. Put*

$$\text{Fourier}(p) = p(x), p^{(1)}(x), p^{(2)}(x), \ldots, p^{(n)}(x),$$

where $p^{(i)}(x)$ denotes the i^{th} derivative of $p(x)$ with respect to x. Let a, b be two real numbers which are not roots of $p(x)$. Let r be the number of real roots of $p(x)$ in the interval (a, b). Then

$$V_F(a) - V_F(b) = r + 2s$$

where $F = \text{Fourier}(p)$ and s is a non-negative integer.

(This result is often attributed to Boudin but it seems that Fourier was the first to discover it. Sturm discovered his exact result after studying Fourier's treatise on the solution of numerical equations.)

Exercise 7.7 *Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be any polynomial with real coefficients and $a_n \neq 0$. Deduce Descartes' rule of signs which states that*

$$v = r + 2s$$

where v is the variation of $a_n, a_{n-1}, \ldots, a_0$, the sequence of coefficients, r is the number of real roots of $p(x)$ in the interval $(0, \infty)$ and s is a non-negative integer.

It is worth noting that we can state Sturm's theorem in a form which does not assume that the polynomial is square free. Using the comments made just after the definition of Sturm sequences we obtain:

Theorem 7.5 *Let $p(x)$ be a non-zero polynomial with real coefficients which is not necessarily square free. Let $S = p_0(x), p_1(x), \ldots, p_n(x)$ be the sequence obtained in the same way as a Sturm sequence. Assume that a, b aer real numbers with $a \leq b$. Then the number of distinct roots of $p(x)$ in $(a, b]$ is exactly $V_S(a) - V_S(b)$.*

This version of the result is useful for investigating 'parametrized' polynomials (i.e., some of the coefficients are unknown).

Exercise 7.8 *Use the last version of Sturm's theorem to find conditions on b, c such that the polynomial $x^2 + bx + c$ has two real roots. Also find conditions for the cubic $x^3 + cx + d$ to have three real roots (note that every cubic equation can be transformed to one of this form by a simple substitution). You will find it helpful to use Maple for this exercise!*

Finally we might ask about the isolation of *all* the roots of a polynomial, complex as well as real ones. For non-real roots we use isolating rectangles rather than intervals (think of complex numbers in the Argand diagram representation). It is possible devise a method which ultimately depends on Sturm sequences and uses exact rational arithmetic (e.g., see H. S. Wilf [64]). We look at a related method in §7.4.

7.3 Real Root Isolation by Continued Fractions

Sturm sequences are an impressive achievement, but how efficiently can we isolate roots with them? Unfortunately the length of the coefficients in a sequence tends to grow alarmingly. (Do not confuse *length* with *absolute value*. For example $1/123458998435902435$ is small in absolute value but large in length.) Collins and Loos [17] report that for a certain polynomial of degree 25 with 25 bit coefficients the Sturm sequence has a coefficient which is 1595 bits long. It might be tempting to approximate coefficients with reals but this would destroy the reliability of the method. We proceed to outline another method for root isolation which is much more efficient.

First of all observe that if we can isolate the positive roots of $p(x)$ then the problem is solved since the negative roots of $p(x)$ are just the positive roots of $p(-x)$ (of course we can easily decide if 0 is a root). From now on we focus entirely on positive roots. If $p(x)$ is to have such a root then it is either in the interval $(0, 1)$ or it is equal to 1 or it is in the range $(1, \infty)$. If the first possibility holds then the root can be expressed as $1/(1 + y)$ for some $y > 0$. We can follow up this possibility by substituting $1/(1 + y)$ for x in $p(x)$ and clearing denominators to obtain a polynomial $p_I(y)$ whose positive roots give us all the possible values for y.

Similarly if the third possibility holds then the root can be expressed as $1 + y$ with $y > 0$. Here we can follow up this possibility by substituting $1 + y$ for x in $p(x)$ to obtain a polynomial $p_T(y)$ whose positive roots give us all the possible values for y.

Let us illustrate this process by using

$$p(x) = x^3 - 7x + 7.$$

By inspection we see that 1 is not a root. For roots in $(0, 1)$ we wish to obtain $p_I(y)$ as defined above. This can be achieved by first replacing x with $1/x$, clearing the denominator power of x and finally substituting $1 + y$ for x. Note that the first stage of this just gives us a polynomial which is $p(x)$ but with the coefficients taken in reverse order, i.e., $7x^3 - 7x^2 + 1$. Note that we must take into account any zero coefficients which occur: $x^3 - 7x + 7 = x^3 + 0x^2 - 7x + 7$ so reversing the coefficients results in $7x^3 - 7x^2 + 0x + 1 = 7x^3 - 7x^2 + 1$. This then yields

$$p_I(y) = 7y^3 + 14y^2 + 7y + 1.$$

Now this polynomial cannot possibly have a positive root since all of its coefficients are positive. So $p(x)$ has no roots in $(0, 1)$.

For roots in $(1, \infty)$ we replace x by $1 + y$ in $p(x)$ to obtain

$$p_T(y) = y^3 + 3y^2 - 4y + 1.$$

By inspection 1 is not a root of $p_T(y)$. We proceed to investigate the possible roots in $(0, 1)$ and $(1, \infty)$. The first possibility leads us to

$$p_{TI}(z) = z^3 - z^2 - 2z + 1,$$

while the second leads to

$$p_{TT}(z) = z^3 + 6z^2 + 5z + 1.$$

Note that $p_{TT}(z)$ has no positive roots so we can conclude that $p_T(y)$ has no roots in $(1, \infty)$. We now investigate the positive roots of $p_{TI}(z)$. Again by inspection we see that 1 is not a root. For roots in $(0, 1)$ we obtain

$$p_{TII}(w) = w^3 + w^2 - 2w - 1$$

while for roots in $(1, \infty)$ we obtain

$$p_{TIT} = w^3 + 2w^2 - w - 1.$$

We can cut the investigation short by making use of the important fact (to be proved below) that if a polynomial has real coefficients which present exactly one sign variation then the polynomial has exactly one positive root. Thus both $p_{TII}(w)$ and $p_{TIT}(w)$ have exactly one positive root. It follows that $(0, \infty)$ is an isolating interval for the roots of both polynomials (if we don't like ∞ as an endpoint then we can use Cauchy's bound given in theorem 7.3; ∞ is a good choice here as we shall see shortly). Note that $p_{TII}(w)$ was obtained from $p(x)$ by the sequence of transformations

$$x \mapsto 1 + y, \quad y \mapsto \frac{1}{1+z}, \quad z \mapsto \frac{1}{1+w},$$

which can be combined into the single transformation

$$x \mapsto 1 + \frac{1}{1 + \dfrac{1}{1+w}},$$

and this simplifies to

$$x \mapsto \frac{3+2w}{2+w}.$$

As a function of w this is strictly increasing and so if (a, b) isolates the positive root of $p_{TII}(w)$ then $((3+2a)/(2+a), (3+2b)/(2+b))$ isolates the corresponding root of $p(x)$. In our case we have $(0, \infty)$ which yields the interval $(3/2, 2)$ (here we have used the fact that $(3+2b)/(2+b) = (3/b+2)/(2/b+1) \to 2$ as $b \to \infty$). The sequence of transformations for $p_{TIT}(w)$ is

$$x \mapsto 1+y, \quad y \mapsto \frac{1}{1+z}, \quad z \mapsto 1+w,$$

which can be combined into the single transformation

$$x \mapsto 1 + \frac{1}{2+w},$$

which simplifies to

$$x \mapsto \frac{3+w}{2+w}.$$

This is a strictly decreasing function of w and so it gives us $(1, 3/2)$ as the isolating interval for the root of $p(x)$.

Summing up we have found that $p(x)$ has exactly two positive roots which are isolated by $(1, 3/2)$ and $(3/2, 2)$ (see figure 4).

The preceding example illustrates an easy consequence of a remarkable fact:

Theorem 7.6 (Vincent, 1836) *Let $p(x)$ be a square free polynomial with rational coefficients. Consider a sequence of transformations of $p(x)$ by*

$$x \mapsto a_1 + \frac{1}{x}, \quad x \mapsto a_2 + \frac{1}{x}, \quad x \mapsto a_3 + \frac{1}{x}, \quad \ldots$$

where a_1 is an arbitrary non-negative integer and $a_2, a_3, \ldots$ are arbitrary positive integers. Then after finitely many steps the sequence of coefficients of the transformed polynomial has either zero or one sign variation.

For a proof see Akritas [2] (where an upper bound on the number of substitutions is also provided).

Corollary 7.1 *Let $p(x)$ be a square free polynomial with rational coefficients. Consider a sequence of transformations of $p(x)$ where each transformation is either of the form $x \mapsto 1+x$ or $x \mapsto 1/(1+x)$. Then after finitely many steps the sequence of coefficients of the transformed polynomial has either zero or one sign variation.*

Proof Exercise □

Note that the preceding theorem is not guaranteed to hold if p is not square free. For example take $p = (2x^2-1)^2 = 4x^4 - 4x^2 + 1$ then $p_{IIT} = p_I$ so that the sequence repeats for ever. It is however true that after finitely many steps the resulting polynomial either has no sign variations in its coefficients (and hence no positive root) or has exactly one positive root (but might have more than one sign variation). Of course the example p given has just one positive root already.

A polynomial whose coefficients present no sign variations clearly cannot have a positive root. Let us now settle the other possibility.

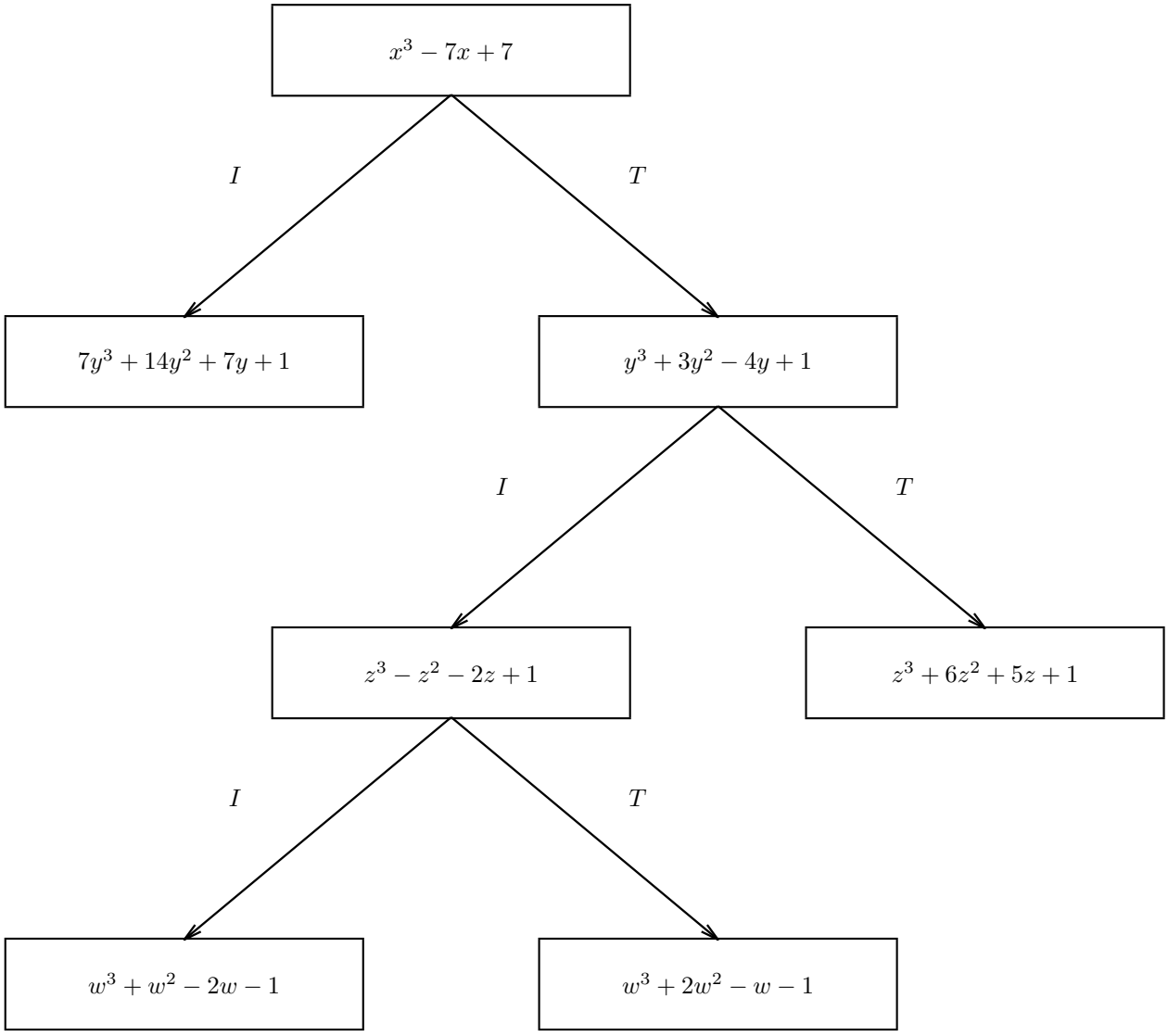

Figure 4: Positive roots of $x^3 - 7x + 7$ isolated by $(1, 3/2)$ and $(3/2, 2)$.

Theorem 7.7 *Let $p(x)$ be a polynomial with real coefficients which have exactly one sign variation. Then $p(x)$ has exactly one positive real root (any other real roots will either be 0 or negative).*

Proof The result follows from Descarte's rule (see Exercise 7.7) but we give a simple proof of this particular case using induction on $m = \deg(p(x))$. Note that m must be at least 1 (otherwise there can be no sign variation in the coefficients!). The result is obvious for $m = 1$. For the induction step assume $m > 1$ and observe that, without loss of generality, we may put

$$p(x) = (a_m x^m + \cdots + a_{n+1}x^{n+1}) - (a_n x^n + \cdots + a_0),$$

where $m > n$, the a_i are all non-negative and $a_m > 0$, $a_0 > 0$. Let

$$\begin{aligned} h(x) &= a_m x^m + \cdots + a_{n+1}x^{n+1}, \\ t(x) &= a_n x^n + \cdots + a_0. \end{aligned}$$

First of all $p(x)$ must have at least one positive root because it changes sign in the interval $(0, \infty)$: we have $p(0) = -a_0 < 0$ while the sign of $p(b)$ as $b \to \infty$ is that of a_m which is positive.

It remains to show that $p(x)$ has at most one positive root. Consider the derivative of $p(x)$:

$$p'(x) = h'(x) - t'(x).$$

This has degree $m - 1$ and its coefficients have either zero or one sign variation. If they have no sign variation then $p'(x)$ is strictly positive for all $x > 0$, i.e., $p(x)$ is an increasing function of x for $x > 0$ and so it can have at most one root satisfying $x > 0$. Otherwise we know from the induction hypothesis that $p'(x)$ has exactly one positive root which means that the graph of $p(x)$ has exactly one turning point for $x > 0$. Now $p'(0) < 0$ and $p'(b) > 0$ for all sufficiently large b. These facts taken together mean that $p(x)$ starts by decreasing in some range $0 < x < d$, then has a turning point at $x = d$ (the unique positive root of $p'(x)$) and then increases in the range $d < x < \infty$. It follows that $p(x)$ cannot have more than one positive root (draw pictures). □

Exercise 7.9 *Give another proof of the preceding theorem by showing that the graphs of the polynomials $h(x)$ and $t(x)$ (defined in the proof) intersect exactly once for $x \geq 0$.*

The meaning of Vincent's theorem is now clearer: any sequence of transformations as described in the statement of the theorem leads to a polynomial which has at most one positive root *and* whose coefficients reveal this fact. The first part of this statements is not at all surprising: transformations of the form $x \mapsto a + x$ with $a > 0$ serve to pull the roots towards the negative part of $\mathbb{R}$ (if the roots of $p(x)$ are $\alpha_1, \ldots, \alpha_s$ then the roots of $p(a + x)$ are $\alpha_1 - a, \ldots, \alpha_s - a$). Transformations of the form $x \mapsto 1/x$ send those roots in the interval $(0, 1)$ to roots in the interval $(1, \infty)$ while roots which are in $(1, \infty)$ are sent to ones in $(0, 1)$ while 1 is kept fixed (if the roots of $p(x)$ are $\alpha_1, \ldots, \alpha_s$ then the roots of $x^m p(1/x)$ are $1/\alpha_1, \ldots, 1/\alpha_s$). The second part really is surprising. For example $x^2 - 2x + 2$ has no real roots even though its coefficients have at least one sign variation. On the other hand the coefficients of $x^3 - x^2 + x - 1 = (x-1)(x-\sqrt{-1})(x+\sqrt{-1})$ have two sign variations even though the polynomial has exactly one positive real root.

The two theorems give us the basis of an algorithm for the isolation of positive roots. There is one more important observation to be made in order to obtain an efficient algorithm. In the example we used transformations of the form $x \mapsto 1/(1 + x)$ and $x \mapsto 1 + x$. Consider a polynomial with more than one positive root and whose smallest such root is $\epsilon + 10^6$ where $\epsilon \geq 0$. Then we

will need 10^6 transformations of the second kind before having a chance of separating the smallest root from the others. It would be much better to have a reasonable integer estimate b, say, of the smallest root so that we can get near it faster by a single transformation of the form $x \mapsto b + x$. How can we find such a bound? The positive roots of $x^{\deg(p)}p(1/x)$ are are of the form $1/\alpha$ where α is a positive root of $p(x)$. Thus if B is an upper bound on the positive roots of $p_I(x)$ then each positive root α of $p(x)$ satisfies $1/\alpha \leq B$ so that $1/B$ truncated to an integer is a lower bound on the roots as required. Of course there is no guarantee that the lower bound obtained really is the integer part of the smallest root, in practice it will be smaller and so we have to carry out at least one substitution of the form $x \mapsto 1+x$ after the substitution $x \mapsto b+x$ applied to some $q(x)$ (which has been obtained from $p(x)$ by a series of transformations). However we must also check for the possibility that $q(x)$ has roots in $(0, 1)$ otherwise they will be lost. For this we could simply use the transformation $x \mapsto 1/(1 + x)$ applied to $q(x)$. However we can save effort by using the following

Theorem 7.8 (Budan,1807) *Let $p(x)$ be square free of degree $m > 0$ and let a, b be two real numbers with $a < b$. Let V_a, V_b be the variation of the coefficients of $p(a+x)$, $p(b+x)$ respectively. Let r be the number of real roots of $p(x)$ in (a, b). Then*

1. *$V_a \geq V_b$.*
2. *$r \leq V_a - V_b$.*
3. *$(V_a - V_b) - r$ is an even number.*

(The proof is similar to that of Fourier's theorem of Exercise 7.6.) Budan's theorem tells us that if the coefficients of $q(x)$ and $q(1+x)$ have the same number of sign variations then $q(x)$ has no roots in $(0, 1)$ and so there is no need to use a transformation of the form $x \mapsto 1/(1 + x)$.

The algorithm can be viewed as an evolving tree. At each vertex of the tree we keep a polynomial $q(x)$, the transformation used to obtain $q(x)$ from the original input polynomial $p(x)$ and the variation of the coefficients of $q(x)$. At the root of the tree we have $p(x)$, the identity transformation and the variation of the coefficients of $p(x)$. Consider now an arbitrary vertex of the tree with polynomial $q(x)$, transformation $M(x)$ and variation v. If $v > 1$ then we must apply further transformations to $q(x)$ (hence to $p(x)$). These are as described above and lead to either one or two children of the current vertex. If $v = 0$ then the vertex leads to no positive roots and so we can drop it. If the $v = 1$ then $q(x)$ has exactly one positive root which corresponds to a positive root of $p(x)$. Here we can isolate the root of $q(x)$ by $(0, \infty)$. Now it is easy to see that

$$M(x) = \frac{a_1x + a_0}{b_1x + b_0}$$

where $a_1b_0 - a_0b_1 \neq 0$ and $a_0, a_1, b_0, b_1 \geq 0$. Clearly $M(x)$ is continuous for all x unless $b_1 \neq 0$ in which case there is a discontinuity at $x = -b_0/b_1$. By differentiating $M(x)$ we see that it is strictly increasing if $a_1b_0 - a_0b_1 > 0$ and otherwise it is strictly decreasing. Moreover if r is the unique positive root of $q(x)$ then the corresponding root of $p(x)$ is $M(r)$. It follows that if (a, b) isolates r then $M(r)$ is isolated by the open interval whose *unordered* endpoints are $M(a)$, $M(b)$. In our case we have $a = 0$ which gives us $M(a) = a_0/b_0$ and $b = \infty$ which we can view as a limit operation so that $M(b) = a_1/b_1$. We must allow for two possible 'degenerate' cases. If $b_1 = 0$ then $M(x)$ does not have a limit as $x \to \infty$ so instead of $(0, \infty)$ we isolate the positive root of $q(x)$ by $(0, B)$ where B is an upper bound on the positive root of $q(x)$, e.g., as given by Theorem 7.3 but see Theorem 7.9

below. This gives us the endpoints a_0/b_0 and $M(B)$. The other possible degenerate case is $b_0 = 0$ but this cannot happen (prove this).

In fact we can take advantage of the special nature of q and can use a simpler upper bound for its unique positive root than than the one given by Theorem 7.3.

Theorem 7.9

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_{n+1}x^{n+1} - a_n x^n - \cdots - a_0$$

be a polynomial of degree m whose coefficients have only one sign variation (so $a_i \geq 0$, for $0 \leq i \leq m$ while $a_m > 0$ and $a_0 > 0$). Then

$$1 + \frac{\max_{0\leq j\leq n} a_j}{\sum_{i=n+1}^{m} a_i}$$

is a strict upper bound on the unique positive root of $p(x)$.

Proof Let α be the positive root of p and set $a = \max_{0\leq j\leq n} a_j$. If $\alpha \leq 1$ the claim is trivial so assume that $\alpha > 1$. Then we have

$$\begin{aligned}(a_m + a_{m-1} + \cdots + a_{n+1})\alpha^{n+1} &< (a_m\alpha^{m-n-1} + a_{m-1}\alpha^{m-n-2} + \cdots + a_{n+1})\alpha^{n+1} \\ &= a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0 \\ &\leq a(\alpha^n + \alpha^{n-1} + \cdots + 1) \\ &= a(\alpha^{n+1} - 1)/(\alpha - 1) \\ &< a\alpha^{n+1}/(\alpha - 1)\end{aligned}$$

Thus $(a_m + \cdots + a_{n+1})(\alpha - 1) < a$, the claim now follows. □

Exercise 7.10 *Prove the claim made above about $M(x)$, i.e., it has the form $(a_1x + a_0)/(b_1x + b_0)$ with $a_1b_0 - a_0b_1 \neq 0$. (Hint: the transforms applied at individual steps are all invertible hence so is any transform composed of them.)*

We now present the algorithm more formally Throughout the algorithm we check for the possibility that we have hit on an exact root and take advantage of the situation whenever it occurs.

Algorithm: *CFISOL*$(p(x)) \mapsto [E, A]$

($p(x)$ is square free. E is a list of (some of the) exact non-negative roots of $p(x)$ and A is a list of isolating intervals for the rest of the non-negative roots of $p(x)$.)

1. **if** $p(0) = 0$ **then** $E := [0]$; $A := [\,]$; $p_w(x) := p(x)/x$
else $E := [\,]$; $A := [\,]$; $p_w(x) := p(x)$
fi;

2. $v := VAR(p_w(x))$;
if $v > 1$ **then** $T := [[p_w(x), x, v]]$ (work to be done)
elif $v = 1$ **then** (exactly one positive root)
$T := [\,]$;

$u := UBPR(p_w(x))$;
if $p_w(u) = 0$ **then** $E := cons(u, E)$
else $A := [(0, u)]$
fi
fi;

3. **while** $T \neq [\,]$ **do**
remove the first element $[p_M(x), M(x), v_M]$ from T;
$b := LBPR(p_M(x))$;
if $b \geq 1$ **then**
$p_M(x) := Moebius(p_M(x), b + x)$;
$M(x) := M(b + x)$;
if $p_M(0) = 0$ **then** $E := cons(M(0), E)$; $p_M(x) := p_M(x)/x$ **fi**;
fi;
$v_1 := v_M$;
$p_{M_1} := Moebius(p_M(x), 1 + x)$;
$M_1 := M(1 + x)$;
if $p_{M_1}(0) = 0$ **then** $E := cons(M_1(0), E)$; $p_{M_1}(x) := p_{M_1}(x)/x$ **fi**;
$v_{M_1} := VAR(p_{M_1})$;
if $v_{M_1} > 1$ **then** $T := cons([p_{M_1}(x), M_1(x), v_{M_1}], T)$
elif $v_{M_1} = 1$ **then** $A := append(A, [make\_interval(p_{M_1}(x), M_1(x))])$
fi;
if $v_1 \neq v_{M_1}$ **then** ($p_M(x)$ might have roots in $(0, 1)$)
$p_I(x) := Moebius(p_M(x), 1/x)$;
if $\mathrm{lc}(p_I(x)) < 0$ **then** $p_I(x) := -p_I(x)$ **fi**;
$M_I := M(1/x)$;
$p_{M_1} := Moebius(p_I(x), 1 + x)$;
$M_1 := M_I(1 + x)$;
$v_{M_1} := VAR(p_{M_1}(x)$;
if $v_{M_1} > 1$ **then** $T := cons([p_{M_1}(x), M_1(x), v_{M_1}], T)$
elif $v_{M_1} = 1$ **then** $A := append(A, [make\_interval(p_{M_1}(x), M_1(x))])$
fi
fi
od

The meaning of the various function calls is as follows:

cons(a, L)**:** inserts the element a at the head of the list L.

VAR$(p(x))$**:** returns the variation in the sequence of coefficients of $p(x)$.

UBPR$(p(x))$**:** produces an upper bound on the positive roots of $p(x)$ using Theorem 7.3.

LBPR$(p(x))$**:** produces a lower bound on positive roots of $p(x)$ as discussed above.

make_interval$(p_M(x), M(x))$**:** produces an isolating interval for a positive root of $p(x)$ as described above.

Moebius$(q(x), M(x))$**:** transforms $q(x)$ by using $M(x)$. In general a transformation of the form

$$M(x) = \frac{a_1 x + a_0}{b_1 x + b_0}$$

where $a_1 b_0 - a_0 b_1 \neq 0$ is called a *Moebius* transformation. To transform $q(x)$ according to $M(x)$ we send it to $(b_1 x + b_0)^n q(M(x))$ where $n = \deg(q(x))$.

Exercise 7.11 *In some applications we are simply interested in knowing whether or not a polynomial has a real root in an interval* $[a, b]$*. Modify the algorithm so that it answers this question directly* without *computing any isolating intervals for roots.*

We have one final but important efficiency consideration. The transforms used by the algorithm are either of form $x \mapsto c + x$ or $x \mapsto 1/x$. The second type are easy to implement efficiently since all we have to do is reverse the order of the coefficients of the polynomial. For transforms of the first type we require the expansion of $p(c+x)$. If this is carried out naïvely then efficiency is greatly impaired.

7.3.1 The Ruffini–Horner Method

We are given $p(x)$ and c and wish to find the coefficients of the expansion of $p(x+c)$. Observe that if we can express $p(x)$ in terms of powers of $x - c$ then we are done for we have:

$$p(x) = a_0'(x-c)^m + a_1'(x-c)^{m-1} + \cdots + a_m'$$

so that

$$p(x+c) = a_0' x^m + a_1' x^{m-1} + \cdots + a_m'$$

i.e., the coefficients we are seeking are $b_0, b_1, \ldots, b_m$. This suggests the following approach: put $p_n(x) = p(x)$ and then

$$\begin{aligned} p_m(x) &= (x-c)p_{m-1}(x) + r_m, \\ p_{m-1} &= (x-c)p_{m-2}(x) + r_{m-1}, \\ &\vdots \\ p_1 &= (x-c)p_0 + r_1 \\ p_0 &= (x-c)\cdot 0 + r_0. \end{aligned}$$

Of course $r_m, r_{m-1}, \ldots, r_0$ are all constant and indeed $r_i = p_i(c)$ for each i. This scheme gives us

$$\begin{aligned} p(x) &= (x-c)p_{m-1}(x) + r_m \\ &= (x-c)^2 p_{m-2}(x) + (x-c)r_{m-1} + r_m \\ &\vdots \\ &= (x-c)^m r_0 + (x-c)^{m-1} r_1 + \cdots + r_m, \end{aligned}$$

which is the required expansion. Let us put

$$p(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_m,$$

and

$$p_{m-1}(x) = b_0 x^{m-1} + b_1 x^{m-2} + \cdots + b_{m-1}.$$

Then it is easy to see that

$$b_0 = a_0, \quad b_i = a_i + c b_{i-1}, \quad 1 \leq i \leq m-1.$$

This method of obtaining $p_{m-1}(x)$ from $p(x)$ is called *synthetic division*. Clearly we can then repeat the process to obtain p_{m-2} from p_{m-1} etc. One important observation is that if $c = 1$ then no multiplications are needed at all.

Exercise 7.12 *How many multiplications and additions/subtractions are needed in the worst case for the Ruffini–Horner method? Consider also the special case when $c = 1$.*

We note also that in the root isolation algorithm we often evaluate a polynomial at some number. Again if this is carried out naïvely then efficiency is impaired. Put

$$p(x) = p_m x^m + p_{m-1} x^{m-1} + \cdots + p_0.$$

Then $p(c)$ can be computed as

$$p(c) = (\cdots((p_m c + p_{m-1})c + p_{m-2})c + \cdots + p_1)c + p_0$$

which involves at most m multiplications and m additions/subtractions. This is frequently called Horner's rule and is a rare example of an algorithm which is optimal: any method which can take an arbitrary polynomial and real number and return the value of the polynomial at the number must use at least as many multiplications and additions/subtractions as Horner's rule does.

Finally we note that the continued fractions method can also be used to approximate roots. To be specific given an isolating interval for a positive root of $p(x)$ and the Moebius transform which was used to obtain the interval then we can approximate the root to any desired degree of accuracy by applying appropriate Moebius transforms till we either obtain an exact root or the isolating interval is small enough; see [2] for details. Of course the bisection method (i.e., the algorithm *APPROX* of p.91) is very efficient as it halves the size of the isolating interval with each iteration.

Exercise 7.13 *Consider a system of finitely many polynomial equations and inequalities (e.g., $p(x) \geq (x)$) where each polynomial comes from $\mathbb{Q}[x]$. The problem is to find all the solutions to the*

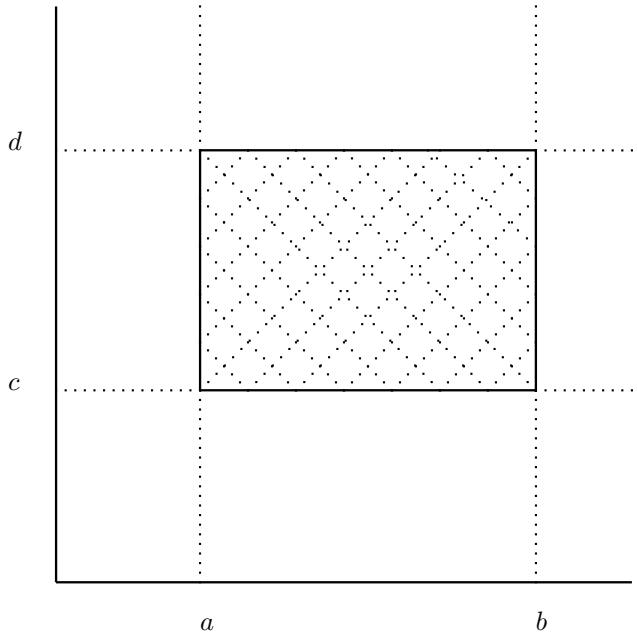

Figure 5: The open rectangle $(a, b) + i(c, d)$ in the Argand plane.

system. Assuming that there is at least one non-trivial equation, show how to reduce the problem to solving finitely many systems of the form

$$p(x) = 0; \quad q_1(x) > 0, \ldots, q_r(x) > 0, \tag{$*$}$$

where the solutions to the original system are obtained by taking the union of the solutions of the finitely many new systems $(*)$.

Give an algorithm for producing all the solutions to a system $(*)$*; the output consists of either exact roots of* $p(x)$ *or open isolating intervals for them.*

7.4 Complex Roots

Although our main concern is with real roots we now give an outline of a method for isolating the complex roots of a polynomial. Here we produce open rectangles of the form $(a, b) + i(c, d)$, see Figure 5.

The outline of the method is quite simple: suppose that given a non-zero polynomial $p(z)$ and a rectangle R of the complex plane we have a method of counting how many roots of $p(z)$ (real or complex) lie inside R. Using Theorem 7.2 we can produce a square which includes all the roots of $p(z)$. Now we can cut this into four equal smaller squares and count the number of roots in each square. If some square has only one root then it forms part of the output. If it has no roots then it can be discarded otherwise we recurse. (In fact, for reasons which will become clear later on, we cannot always subdivide into squares but have to use rectangles instead.)

The first problem we face is how to count roots in a rectangle. This has a very beautiful solution which comes from the basic theory of functions of a complex variable. The method is explained more easily if we consider counting the number of zeros of $p(z)$ inside a simple continuous closed curve C (by *simple* we mean that the curve does not intersect itself). Putting $z = x + iy$ we can write

$$p(z) = u(x, y) + iv(x, y)$$

where $u(x, y)$, $v(x, y)$ have rational coefficients. We think of C as lying in the (x, y)-plane and the image of C, which is another curve (not necessarily simple), as lying in the (u, v)-plane. Note that the image curve $P(C)$ passes through the origin of the (u, v)-plane if and only if there is a root of $p(z)$ on C. From now on we assume that no root of $p(z)$ lies on C. If C is deformed continuously to another curve C' then of course $p(C)$ is also deformed continuously to $p(C')$. Moreover a curve used during the deformation has a root of $p(z)$ if and only if one of the image curves passes through the origin of the (u, v)-plane. Moreover if C is contracted to a point then the same happens to the image. It is now clear that the interior of C has a root of $p(z)$ if and only if when we try to contract $p(C)$ continuously to a point some intermediate curve passes through the origin of the (u, v)-plane. As C is contracted the image curve passes through the origin at least once for each root of $p(z)$ which is in the interior of C. In fact the image curve is wound around the origin at least once for each root. For a better understanding of this suppose first of all that the interior of C has only one root of $p(z)$. Consider shrinking C to smaller and smaller curves around the root (none of which passes through the root). Then none of the image curves passes through the origin of the (u, v)-plane but as C shrinks the image curves also shrink and enclose the origin. Since no intermediate curve has passed through the origin it follows that all curves including C enclose the origin. Another way of phrasing this is to imagine that we are told to walk around $p(C)$ holding a piece of elastic whose other end is tied to a pole stuck into the origin of the (u, v)-plane. We stop upon arriving at the starting point. If by the end of the walk the elastic has wound round the pole at least once then there must be a root of $p(z)$ in the interior of C otherwise the interior of C has no such roots. (At the start the elastic is not wound around the pole at all.) This gives us a test for detecting the presence of roots but we also want to count them.

Let us be more precise about the process of walking around $p(C)$. We do not simply view this as following a path which has already been traced out. Imagine that C is parametrized by a time parameter t, i.e., we have a continuous function $\tau(t)$ with the property that $\tau(0) = \tau(1)$ and which traces out C as t varies from 0 to 1 (the choice of 0 and 1 as start and end times is arbitrary but has become standard—in fact 0 is an excellent choice for computational reasons). Now As t varies the points of the image of C are given by $p(\tau(t))$ and our walk consists of following these points in the order of generation. Note that this could involve us in going round $p(C)$ several times, going faster for some parts rather than others etc. At the end of the walk we see how many times the elastic has wound around the pole and call this the *winding number* of $p(z)$ on C. What we have now is that the interior of C contains at least as many roots as the winding number. Further investigation shows that a root of multiplicity d contributes precisely d to the winding number. Let us verify this for

$$p(z) = (z - \alpha)^d.$$

We can take C to be a circle of unit radius and centre α. Thus C is given by

$$\tau(t) = \alpha + \sin(2\pi t) + i\cos(2\pi t).$$

The image of C is thus

$$(\sin(2\pi t) + i\cos(2\pi t))^d$$

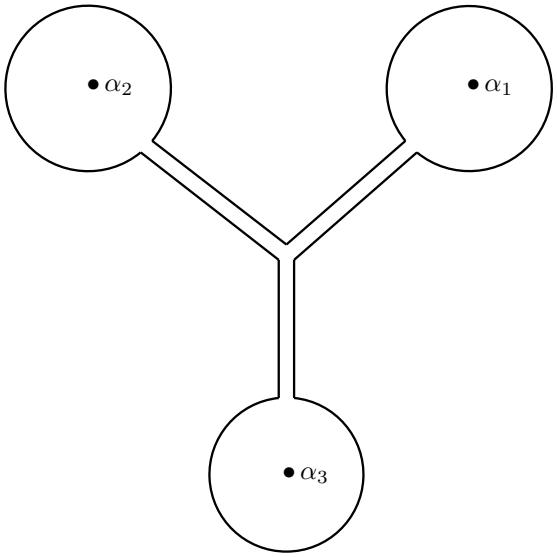

Figure 6: Deformed closed curve around the roots α_1,α_2, α_3 of $p(z)$.

and by de Moivre's theorem this is the same as

$$\sin(2d\pi t) + i\cos(2d\pi t).$$

which traces the unit circle with centre at the origin exactly d times as t ranges from 0 to 1. For the general case you might like to contemplate Figure 6. Finally as an illustration let $p(z) = z^3 - 1$. The roots of this polynomial are all included in the square $(-2, 2) + i(-2, 2)$ whose image under $p(z)$ is shown in Figure 7

Exercise 7.14 *Write an Axiom function* `contour(p,z,a,b,c,d)` *which takes a polynomial* $p \in \mathbb{Q}[z]$ *and plots the image of the rectangle* $(a, b) + i(c, d)$.

The preceding argument is an intuitive explanation of a rigorous theorem which holds for functions which are more general than polynomials. For details see Henrici [**31**].

Note that in order for the bisection algorithm to work we must ensure that each root has multiplicity 1 by working with the square free part of $p(z)$. We shall assume from now on the $p(z)$ is square free.

How can we compute the winding number? First of all we label the quadrants of the plane in the usual way (see Figure 8). Let w_0 be the number of times that we cross from quadrant I to quadrant II or from quadrant III to quadrant IV. Let w_1 be the number of times that we cross from quadrant II to quadrant I or from quadrant VI to quadrant III. Then it is clear that the winding number is given by $(w_0 - w_1)/2$ (assuming that we traverse the curve in an anticlockwise direction, otherwise we obtain the negation of the winding number). In particular we are interested in the winding number when C is a rectangle $(a, b) + i(c, d)$. We will always consider traversing rectangles starting at the lower left hand corner, i.e., the point $a + ic$, and going round anti-clockwise. We can break this up into the traversal of the four edges that make up the rectangle—minus the start-point

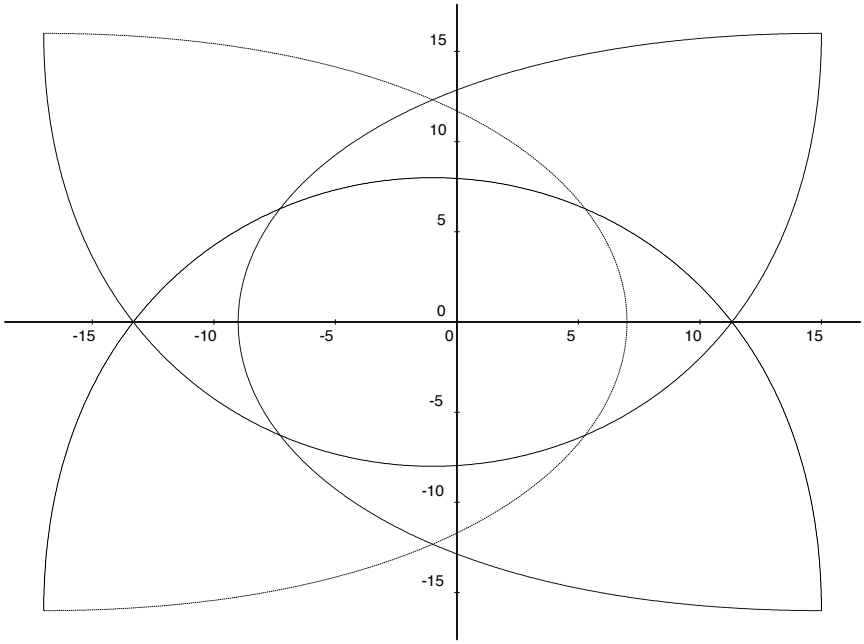

Figure 7: The image of $(-2,2)+i(-2,2)$ under z^3-1.

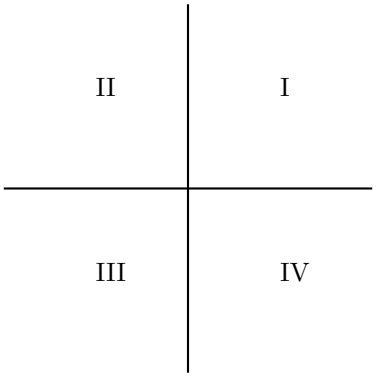

Figure 8: The four quadrants of the plane.

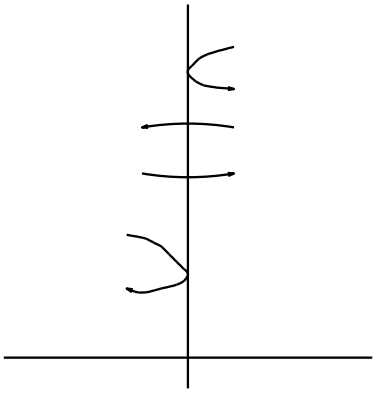

Figure 9: Some of the possible behaviours of the image curve near the roots of $u_{z_1 z_2}(t)$.

of each edge in order to avoid duplicating it. We count the contribution that each segment makes to the winding number and add up. The problem therefore reduces to considering the image of a line segment.

The line segment joining a point z_1 to z_2 is given by $z_1 + (z_2 - z_1)t$ and we can easily compute polynomials $u_{z_1 z_2}, v_{z_1 z_2} \in \mathbb{Q}[t]$ such that the image of the segment under $p(z)$ is $u_{z_1 z_2}(t) + i v_{z_1 z_2}(t)$. The times at which the image curve can cross from one quadrant to another are given by the real roots of $u_{z_1 z_2}(t)$ which lie in $(0, 1]$. These can be isolated by using the continued fractions method. For each root we have to decide which quadrants, if any, are crossed by the curve as t passes through the root. Figure 9 shows some of the possibilities that must be considered. The essential idea is to find isolating intervals for the roots of $u_{z_1 z_2}(t)$ in which $v_{z_1 z_2}(t)$ has constant sign. Moreover we ensure that the endpoints of such an interval are not roots of $u_{z_1 z_2}(t)$. With these properties we can then detect which quadrants, if any, are being crossed around each root.

Exercise 7.15 *The real root isolation algorithm might return some exact roots as well as isolating*

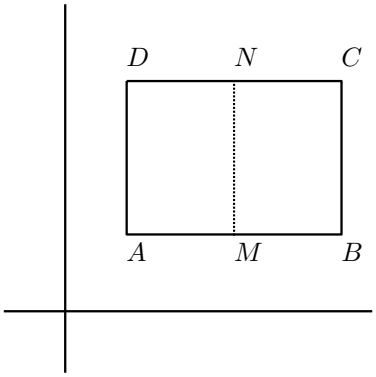

Figure 10: Subdividing a rectangle.

.

intervals. How do we deal with these? (Avoid the obvious answer that we can replace them with isolating intervals—it is correct but boring.)

Exercise 7.16 *There is an extra complication when the end of a segment is on the vertical axis (i.e., the full curve seems to be just about to cross from one quadrant to another). How do we deal with this?*

Computing winding numbers is an expensive process and so it is worth avoiding this whenever possible. We proceed to outline some improvements to the original naïve algorithm which are extremely effective. First of all the subdivision into four parts is better carried out in two stages. First we subdivide horizontally and then vertically. So what? The idea is that for each rectangle we do not just remember its winding number but rather the contributions to the winding number of the four edges which make it up. Consider Figure 10. Upon subdivision by MN we need only work out the winding number contributions of AN, MN and ND, the others can be deduced from the old data and these new numbers. (Observe that the winding number is just negated if we traverse the same edge but in the opposite direction. Of course the contribution from traversing an edge need not be an integer.) If one of the smaller rectangles has winding number 1 then it can be output. If it has winding number equal to 0 then it can be discarded. Otherwise it goes through to be subdivided horizontally.

Before making any further improvements we must address one difficulty. The method of winding numbers requires that no root of $p(z)$ lies on the edges of the rectangle. At the outset this is fine because Theorem 7.2 gives us a square which properly encloses all the roots of $p(z)$. But as we subdivide rectangles it is possible to hit a root. This can be detected as follows. If the image curve of a segment is $u(t) + iv(t)$ then the segment contains a root of $p(z)$ if and only if $u(t)$, $v(t)$ have a common root in $[0, 1]$. Thus we just test $w(t) = \gcd(u(t), v(t))$ for roots in $[0, 1]$. Actually we need only test for roots in $(0, 1)$ (why?) and by using the transformation $t \mapsto 1/(1+t)$ we can reduce this to testing for positive roots. Now if it so happens that a proposed subdivision segment contains a root of $p(z)$ then we simply move it and try again. This process is bound to terminate since in any rectangle there are only finitely many roots of $p(z)$.

All the methods which have been described so far work even if the coefficients of $p(z)$ are complex numbers (with rational real and imaginary parts). However for $p(z)$ with rational coefficients we can exploit a well known property of the roots: if α is a root of $p(z)$ then so is its complex conjugate $\bar{\alpha}$. Thus the roots of $p(z)$ are either real or they come in pairs $x+iy$, $x-iy$ with $y \neq 0$. This suggests that we isolate the real roots by the continued fractions method and then isolate only the roots with positive imaginary part—for if $(a,b)+i(c,d)$ isolates $x+iy$ then $(a,b)+i(-d,-c)$ isolates $x-iy$. Unfortunately there is a snag: what rectangle do we use at the start? (this rectangle must enclose only those roots of $p(z)$ with a strictly positive imaginary part). It would be natural to take the same rectangle as in the original algorithm but use only its upper half (i.e., cut it along the x-axis). Alas this is no good if $p(z)$ happens to have any real roots at all, otherwise it is fine (of course we will know if there are any real roots from the use of the continued fractions method).

Exercise 7.17 *Actually if we are a little lucky with the real root finding process we can still proceed as though $p(z)$ had no real roots. Explain.*

So we have now reduced to the case when $p(z)$ has at least one real root. We could look for a lower bound δ on the distance between the roots of $p(z)$ because if b is the bound given by Theorem 7.2 then we can use the rectangle $(\delta/2, b) + i(-b, b)$. Such a bound can be obtained from

Theorem 7.10 (Mahler) *Let $p(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_m$ be a polynomial with integer coefficients of degree at least 2. Let Δ be the minimum distance between its roots. Then*

$$\Delta \geq \sqrt{3} m^{-(m+1)/2} (|a_0| + |a_1| + \cdots + |a_m|)^{-(m-1)}.$$

Proof See Akritas [2] □

Unfortunately the bounds obtained from this theorem tend to be extremely pessimistic and involve us in working with numbers of large representation size. Their use would therefore slow the algorithm down! (The theorem is extremely important in analyzing the worst case behaviour of root finding algorithms.) We are thus forced to seek an alternative solution. Let $(a,b) + i(c,d)$ be a rectangle which isolates some roots of $p(z)$ and has no roots on its edges. We make two observations.

1. If $d < 0$ then the rectangle isolates only roots with negative imaginary parts. We can just discard this because by the end of the algorithm we will have isolated the conjugate roots (with positive imaginary parts) and we can then recover the 'lost' roots.

2. If $c < 0$, $d > 0$ and $-d < c$ then it is safe to split the rectangle into $(a,b) + i(-c,d)$ and $(a,b) + i(c,-c)$. The first of these isolates only roots with positive imaginary parts. We have also made a gain because the winding number of the splitting segment is just the negation of the known number of the segment which joins $a + ic$ to $b + ic$ (why?).

With these modifications the algorithm tends to discard many rectangles and focuses on real roots and complex roots with positive imaginary part. At the start of the algorithm it is a good idea not to use a square but a rectangle of the form $(-b, b) + i(-1, b)$. Note that in the output from this modified version those rectangles which straddle the x-axis isolate real roots.

Exercise 7.18 *The root finding algorithm has to carry out a great deal of complex number arithmetic. The naïve method for mutliplying two complex numbers uses four multiplications, one*

addition and one subtraction. The number of multiplications can be reduced to three as follows: given $x_1 + iy_1$ *and* $x_2 + iy_2$ *compute*

$$s_1 := (x_2 + y_2) * x_1;$$
$$s_2 := -y_2 * (x_1 + y_1);$$
$$s_3 := x_2 * (y_1 - x_1);$$

then the real part of $(x_1 + iy_1)(x_2 + iy_2)$ *is* $s_1 + s_2$ *and the imaginary part is* $s_1 + s_3$*. Do you think this is likely to help in speeding up the algorithm? Discuss pros and cons.*

8 Bibilography

1. S. S. Abhyankar, *Albebraic Geometry for Scientists and Engineers*, Mathematical Surveys and Monographs, 35, AMS (1990).

2. A. G. Akritas, *Elements of Computer Algebra With Applications*, Wiley, (1989).

3. W. M. Anderson, *A Survey of Polynomial Factorisation Algorithms*, M.Phil. Thesis, Department of Computer Algebra, University of Edinburgh, Scotland.

4. D. Bayer and M. Stillman, A theorem on refining division orders by the reverse lexicographic orders. *Duke J. Math.* **55**, (1987) 321–328.

5. T. Becker and V. Weispfenning, *Gröbner Bases. A Computational Approach to Commutative Algebra* (in cooperation with H. Kreidel). Springer, New York (1993).

6. E. R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill (1968).

7. M. Bronstein, The Transcendental Risch Differential Equation, *JSC*, **9**, 1 (1990) 49–60.

8. B. Buchberger, An Algorithm for Finding a Basis for the Residue Class Ring of a Zero-dimensional Polynomial Ideal (German), PhD Thesis, University of Insbruck (Austria), Math. Inst. (1965).

9. B. Buchberger, An Algorithmic Criterion for the Solvability of Algebraic Systems of Equations (German), *Aequationes Mathematicae*, **4**, 3, (1970) 374–383.

10. B. Buchberger, A Criterion for Detecting Unnecessary Reductions in the Construction of Gröbner Bases, *EUROSAM79, Lecture Notes in COmputer Science 72*, E. W. Ng (editor), Springer (1979) 3–21.

11. B. Buchberger, G. E. Collins and R. Loos (editors), *Computer Algebra—Symbolic and Algebraic Computation. Computing Supplementum 4*, Springer (1983).

12. B. Buchberger, Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory, in *Multidimensional Systems Theory*, N. K. Bose (editor), D. Reidel Publishing Company (1985).

13. B. W. Char, K. O. Geddes, W. M. Gentleman and G. H. Gonnet, The design of Maple: A compact, portable, and powerful computer algebra system, *Proceedings of Eurocal '83*, pp. 101–115 (April 1983). Springer Lecture Notes in Computer Science no. 162.

14. B. W. Char, K. O. Geddes, G. H. Gonnet, M. B. Monagan and S. M. Watt, *First Leaves: A Tutorial Introduction to Maple*, WATCOM Publications Ltd., Waterloo, Ontario, Canada (1988).

15. B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, and S. M. Watt, *Maple V Language Reference Manual*, Springer (1992).

16. B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monaan, and S. M. Watt, *Maple V Library Reference Manual*, Springer (1992).

17. G. E. Collins and R. Loos, Real Zeros of Polynomials, in [11].

18. T. H. Cormen, C. E. Leiserson and R. L. Rivest, *Introduction to Algorithms*, MIT Press (1990).

19. D. Cox, J. Little and D. O'Shea, *Ideals, Varieties and Algorithms: an Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer (1992).

20. J. H. Davenport, *On the Integration of Algebraic Functions*, Springer Lecture Notes in Computer Science 102, Springer (1981).

21. J. H. Davenport, Y. Siret and E. Tournier, *Computer Algebra: systems and algorithms for algebraic computation*, Academic Press (1988).

22. P. J. Davis and R. Hersh, *The Mathematical Experience*, Penguin Books (1981).

23. D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, New York (1995).

24. J. C. Faugère, P. Gianni, D. Lazard and T. Mora, Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Computation* **16**, (1993) 329–344.

25. J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, Cambridge University Press (1999).

26. K. O. Geddes, S. R. Czapor and G. Labahn, *Algorithms for Computer Algebra*, Kluwer Academic Publishers (1992).

27. R. W. Gosper, Decision Procedure for Indefinite Hypergeometric Summation, *Proc. National Academy of Sciences, USA*, **75**, 1 (1978) 40–42.

28. D. Harper, C. Woof and D. Hodgkinson, *A Guide to Computer Algebra Systems*, Computer Algebra Support Project, Computer Laboratory, University of Liverpool, (1989).

29. D. Harper, C. Woof and D. Hodgkinson, *A Guide to Computer Algebra*, Wiley, (1991).

30. E. Horowitz, Algorithms for partial fraction decomposition and rational function integration, *Proc. 2^{nd} Symp. on Symbolic and Algebraic Computation*, S. K. Petrich (ed), ACM (1971).

31. P. Henrici, *Applied and Computational Complex Analysis*, Vol. I, Wiley–Interscience, (1974).

32. C. M. Hoffmann, *Geometric and Solid Modeling, an Introduction*, Morgan Kaufmann, (1989).

33. D. T. Huynh, A superexponential lower bound for Gröbner bases and Church-Rosser commutative Thue systems. *Inf. and Control* **68**, (1986) 196–206.

34. C. Jordan, *Calculus of Finite Differences*, Sopron, Röttig u. Romwalter (1939).

35. E. Kaltofen, Factorization of Polynomials, in [**11**].

36. A. Karatsuba and Yu. Offman, *Dokl. Akad. Nauk SSSR* **145** (1962) 293–294. (English translation: Multiplication of multi-digit numbers on automata, *Soviet Phys. Dokl.* **7** (1963) 595–596).

37. M. Karr, Summation in Finite Terms, *JACM*, **28**, 2 (1981) 305–350.

38. M. Karr, Theory of Summation in Finite Terms, *JSC*, **1**, 3 (1985) 303–315.

39. D. E. Knuth, *Seminumerical Algorithms*, (Second Edition), Addison-Wesley (1981).

40. J. C. Lafon, Summation in Finite Terms, in [11].

41. A. K. Lenstra, H. W. Lenstra and L. Lovász, Factoring Polynomials with Rational Coefficients, *Math. Ann.*, 261 (1982) 515–534.

42. Y. N. Lakshman, A single exponential bound on the complexity of computing Gröbner bases of zero dimensional ideals. In *Effective Methods in Algebraic Geometry*, Mora, T., Traverso, C. (eds.). Birkhäser, Boston (1991).

43. E. W. Mayr and A. R. Meyer, The Complexity of the Word Problem for Commutative Semigroups and Polynomial Ideals, *Advances in Math.*, **46**, (1982) 305–329.

44. M. Mignotte, An inequality about factors of polynomials, *Mathematics of Computation*, **28** (1974) 1153–1157.

45. B. Mishra and C. Yap, Notes on Gröbner Bases, *Information Science*, **48** (1989) 219–252.

46. M. Mignotte, *Mathematics for Computer Algebra*, Springer (1992).

47. H. M. Möller and T. Mora, Upper and lower bounds for the degree of Gröbner bases. In *EUROSAM'84, International Symposium on Symbolic and Algebraic Computation*, Fitch, J. (ed.), LNCS 162. Springer (1984).

48. F. Pauer and M. Pfeifhofer, The Theory of Gröbner Bases, *L'Enseignment Mathématic*, **34**, (1988) 215–232.

49. M. Pohst and H. Zassenhaus, *Algorithmic Algebraic Number Theory*, Cambridge University Press, (1990).

50. C. H. Papadimitriou, *Computational Complexity*, Addison-Wesley, Reading, Mass. (1994).

51. M. Reid, *Undergraduate Algebraic Geometry*, LMS Student Texts 12, Cambridge University Press (1990).

52. D. Richardson, Some Unsolvable Problems Involving Elementary Functions of a Real Variable, *J. Symbolic Logic*, **33** (1968) 511–520.

53. R. H. Risch, The Problem of Integration in Finite Terms, *Trans. AMS*, 139 (1969) 167–189.

54. R. H. Risch, The Solution of the Problem of Integration in Finite Terms, *Bull AMS*, **76** (1970) 605–608.

55. M. Rosenlicht, On Liouville's Theory of Elementary Functions, *Pacific J. of Math.*, **65**, 2 (1976) 485–492.

56. A. Schönhage and V. Strassen, Schnelle Multiplikation großer Zahlen, *Computing*, **7**, (1971) 281–292.

57. I. Stewart, *Galois Theory*, Second Edition, Chapman and Hall (1989).

58. D. R. Stoutemyer, Crimes and misdemeanors in the computer algebra trade, *Notices of the AMS*, Sept. 1991, 701-705.

59. W. Trinks, On Improving Approximate Results of Buchberger's Algorithm by Newton's Method. *SIGSAM* **8** (3), (1984) 7–11.

60. B. M. Trager, *On the Integration of Algebraic Functions*, Ph.D. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., (1985).

61. L. G. Valiant, The complexity of enumeration and reliability problems, *SIAM J. Comput.* **8**, (1979) 410–421.

62. B. L. van der Waerden, *Modern Algebra*, (2 volumes), Ungar (1964).

63. R. Walker, *Algebraic Curves*, Springer (1978).

64. H. S. Wilf, A Global Bisection Algorithm for Computing the Zeros of Polynomials in the Complex Plane, *J. ACM*, **25**, 3 (1978) 415–420).

65. F. Winkler, *Polynomial Algorithms in Computer Algebra*, Springer (1996).

66. H. Zassenhaus, On Hensel Factorization, *I. J. Number Theory*, **1**, (1969) 291–311.

67. R. Zippel, *Effective Polynomial Computation*, Kluwer Academic Publishers (1993).