Course "Polynomials: Their Power and How to Use Them",
JASS'07

Basics about Polynomials

Maximilian Butz

Fakultät für Mathematik
TU München

March 20, 2007

Definition 1

A ring is an algebraic system $(R, +, \cdot)$ satisfying the following:

Definition 1

A ring is an algebraic system $(R, +, \cdot)$ satisfying the following:

- The set R with the addition $+$ is an abelian group.

Definition 1

A ring is an algebraic system $(R, +, \cdot)$ satisfying the following:

- The set R with the addition $+$ is an abelian group.
- The multiplication $\cdot$ is associative.

Definition 1

A ring is an algebraic system $(R, +, \cdot)$ satisfying the following:

- The set R with the addition $+$ is an abelian group.
- The multiplication $\cdot$ is associative.
- Multiplication distributes over addition:

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ and } (b + c) \cdot a = b \cdot a + c \cdot a$$

for all $a, b, c \in R$.

Definition 1

A ring is an algebraic system $(R, +, \cdot)$ satisfying the following:

- The set R with the addition $+$ is an abelian group.
- The multiplication $\cdot$ is associative.
- Multiplication distributes over addition:

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ and } (b + c) \cdot a = b \cdot a + c \cdot a$$

  for all $a, b, c \in R$.

We say that $(R, +, \cdot)$ is a ring with unity, if R contains an multiplicative identity, denoted by 1. For commutative rings, multiplication has to be commutative, too.

For a commutative ring $(R, +, \cdot)$ with unity we define for the elements of R:

For a commutative ring $(R, +, \cdot)$ with unity we define for the elements of R:

- $a \in R$ divides $c \in R$, if there exists $b \in R$, so that $c = a \cdot b$.

For a commutative ring $(R, +, \cdot)$ with unity we define for the elements of R:

- $a \in R$ divides $c \in R$, if there exists $b \in R$, so that $c = a \cdot b$.
- In particular, $a \in R$, $a \neq 0$ is called a zerodivisor, if there exists $b \in R$, $b \neq 0$ with $a \cdot b = 0$.

For a commutative ring $(R, +, \cdot)$ with unity we define for the elements of R:

- $a \in R$ divides $c \in R$, if there exists $b \in R$, so that $c = a \cdot b$.
- In particular, $a \in R$, $a \neq 0$ is called a zerodivisor, if there exists $b \in R$, $b \neq 0$ with $a \cdot b = 0$.
- $u \in R$ is called a unit if there is an multiplicative inverse $v \in R$ so that $u \cdot v = 1$.

For a commutative ring $(R, +, \cdot)$ with unity we define for the elements of R:

- $a \in R$ divides $c \in R$, if there exists $b \in R$, so that $c = a \cdot b$.
- In particular, $a \in R$, $a \neq 0$ is called a zerodivisor, if there exists $b \in R$, $b \neq 0$ with $a \cdot b = 0$.
- $u \in R$ is called a unit if there is an multiplicative inverse $v \in R$ so that $u \cdot v = 1$.

Example 2

The residue classes $\mathbb{Z}/8\mathbb{Z}$ with the usual addition and multiplication form a ring. The equivalence classes of odd numbers are units, the equivalence classes [2], [4] and [6] are zerodivisors.

Definition 3

A nontrivial ring (a ring that contains more than one element), with unity and without zero divisors is called domain. If multiplication is commutative, we call it integral domain.

Definition 3

A nontrivial ring (a ring that contains more than one element), with unity and without zero divisors is called domain. If multiplication is commutative, we call it integral domain.

Example 4

The ring of integers $(\mathbb{Z}, +, \cdot)$ is an integral domain with units 1 and -1.

Definition 3

A nontrivial ring (a ring that contains more than one element), with unity and without zero divisors is called domain. If multiplication is commutative, we call it integral domain.

Example 4

The ring of integers $(\mathbb{Z}, +, \cdot)$ is an integral domain with units 1 and -1.

Definition 5

A field is a commutative, nontrivial ring with unity, in which every nonzero element is a unit.

Definition 3

A nontrivial ring (a ring that contains more than one element), with unity and without zero divisors is called domain. If multiplication is commutative, we call it integral domain.

Example 4

The ring of integers $(\mathbb{Z}, +, \cdot)$ is an integral domain with units 1 and -1.

Definition 5

A field is a commutative, nontrivial ring with unity, in which every nonzero element is a unit.

Wellknown examples for fields are the rationals $\mathbb{Q}$, the reals $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

Definition 6

Let $(R, +, \cdot)$ be a ring and S be the set of sequences

$$\{a_0, a_1, ...\} \text{ with } a_i \in R \text{ for all } i \in \mathbb{N}_0$$

such that $a_i = 0$ for all but a finite number of $i \in \mathbb{N}_0$.

Definition 6

Let $(R, +, \cdot)$ be a ring and S be the set of sequences

$$\{a_0, a_1, ...\} \text{ with } a_i \in R \text{ for all } i \in \mathbb{N}_0$$

such that $a_i = 0$ for all but a finite number of $i \in \mathbb{N}_0$.
If we define addition and multiplication on S by:

$$\{a_0, a_1, ...\} + \{b_0, b_1, ...\} := \{a_0 + b_0, a_1 + b_1, ...\}$$

$$\{a_0, a_1, ...\} \cdot \{b_0, b_1, ...\} := \{a_0 \cdot b_0, a_1 \cdot b_0 + a_0 \cdot b_1, ...\}$$

then $(S, +, \cdot)$ is the ring $R[X]$ of univariate polynomials over R.

Definition 7

For a polynomial $P = \{a_0, a_1, ...\} \in R[X]$, the degree $deg(P)$ is defined as the maximal number n so, that $a_n \neq 0$. In this case, $lc(P) := a_n$ is called the leading coefficient of P.

Definition 7

For a polynomial $P = \{a_0, a_1, ...\} \in R[X]$, the degree $deg(P)$ is defined as the maximal number n so, that $a_n \neq 0$. In this case, $lc(P) := a_n$ is called the leading coefficient of P.

By definition of addition and multiplication on $R[X]$ we have for two polynomials P, Q:

Definition 7

For a polynomial $P = \{a_0, a_1, ...\} \in R[X]$, the degree $deg(P)$ is defined as the maximal number n so, that $a_n \neq 0$. In this case, $lc(P) := a_n$ is called the leading coefficient of P.

By definition of addition and multiplication on $R[X]$ we have for two polynomials P, Q:

- $deg(P + Q) \leq max\{deg(P), deg(Q)\}$

Definition 7

For a polynomial $P = \{a_0, a_1, ...\} \in R[X]$, the degree $deg(P)$ is defined as the maximal number n so, that $a_n \neq 0$. In this case, $lc(P) := a_n$ is called the leading coefficient of P.

By definition of addition and multiplication on $R[X]$ we have for two polynomials P, Q:

- $deg(P + Q) \leq max\{deg(P), deg(Q)\}$
- $deg(P \cdot Q) \leq deg(P) + deg(Q)$

Definition 7

For a polynomial $P = \{a_0, a_1, ...\} \in R[X]$, the degree $deg(P)$ is defined as the maximal number n so, that $a_n \neq 0$. In this case, $lc(P) := a_n$ is called the leading coefficient of P.

By definition of addition and multiplication on $R[X]$ we have for two polynomials P, Q:

- $deg(P + Q) \leq max\{deg(P), deg(Q)\}$
- $deg(P \cdot Q) \leq deg(P) + deg(Q)$
- if R contains no zerodivisors, its even $deg(P \cdot Q) = deg(P) + deg(Q)$.

For a ring with unity, we can define the variable

$$X := \{0, 1, 0, 0, ...\}$$

For a ring with unity, we can define the variable

$$X := \{0, 1, 0, 0, ...\}$$

With this definition we have

$$X^n := \{\underbrace{0, ..., 0}_{\text{n zeroes}}, 1, 0, 0, ...\}$$

For a ring with unity, we can define the variable

$$X := \{0, 1, 0, 0, ...\}$$

With this definition we have

$$X^n := \{\underbrace{0, ..., 0}_{\text{n zeroes}}, 1, 0, 0, ...\}$$

Now we can write a polynomial of degree n like this:

$$\{a_0, a_1, ...\} = \sum_{k=0}^{n} a_k X^k$$

For any polynomial $P(X) = \sum_{k=0}^{n} a_k X^k$ in $R[X]$ we can define a function

$$P : R \to R, \text{ with } P(z) := \sum_{k=0}^{n} a_k z^k$$

by substituting the formal symbol X by elements of R.

For any polynomial $P(X) = \sum_{k=0}^{n} a_k X^k$ in $R[X]$ we can define a function

$$P : R \to R, \text{ with } P(z) := \sum_{k=0}^{n} a_k z^k$$

by substituting the formal symbol X by elements of R.
Note that, for different *polynomials* $P(X)$ and $Q(X)$, the *functions* P and Q can be equal.

For any polynomial $P(X) = \sum_{k=0}^{n} a_k X^k$ in $R[X]$ we can define a function

$$P : R \to R, \text{ with } P(z) := \sum_{k=0}^{n} a_k z^k$$

by substituting the formal symbol X by elements of R.
Note that, for different *polynomials* $P(X)$ and $Q(X)$, the *functions* P and Q can be equal.

Example 8

For p prime, $z^p - z = 0$ for all elements of $\mathbb{Z}/p\mathbb{Z}$, but $X^p - X$ is obviously *not* the zero polynomial (the polynomial with zero coefficients).

The definition of multivariate polynomials follows from the univariate case:

Definition 9

Let R be a ring. For $m \in \mathbb{N}$ we define the ring of multivariate polynomials in m variables $\{X_1, ..., X_m\}$ over R by

$$R[X_1, ..., X_m] = R[X_1, ..., X_{m-1}][X_m]$$

To store and to represent a polynomial $P(X)$ of degree n, we can use a dense representation like

$$P = \{X, n, a_n, ..., a_1, a_0\},$$

where we mention *all* coefficients of P.

To store and to represent a polynomial $P(X)$ of degree n, we can use a dense representation like

$$P = \{X, n, a_n, ..., a_1, a_0\},$$

where we mention *all* coefficients of P.
However, for a polynomial with many zero coefficients it is enough to store the nonzero coefficients in a sparse representation:

$$P = \{X, a_s, m_s, ..., a_2, m_2, a_1, m_1\},$$

where a_i are the nonzero coefficients and m_i are the exponents in decreasing order.

Now, we want to take a look on the computational complexity of addition and multiplication in $R[X]$.

Now, we want to take a look on the computational complexity of addition and multiplication in $R[X]$.
Assume that operations in R can be done in time $O(1)$, and let $P(X)$ and $Q(X)$ two Polynomials, with $deg(P) = m$, $deg(Q) = n$, and let s and t be the numbers of nonzero coefficients.
It is obvious that the calculation of $P + Q$ is done in a time of $O(max\{m, n\})$ in dense representation, while sparse representation leads to a computing time of $O(max\{s, t\})$.

Theorem 10

In dense representation, the calculation of $P \cdot Q$ is done in $O(mn)$, while in sparse representation the calculation is done in $O(st \cdot log_2(t))$

Theorem 10

In dense representation, the calculation of $P \cdot Q$ is done in $O(mn)$, while in sparse representation the calculation is done in $O(st \cdot log_2(t))$

Proof.

(dense case)

$$P(X) \cdot Q(X) = \left(\sum_{j=0}^{m} a_j X^j \right) \cdot \left(\sum_{k=0}^{n} b_k X^k \right) = \sum_{l=0}^{m+n} c_l X^l$$

with $c_l = \sum_{s=0}^{l} a_s b_{l-s}$. Thus, we are doing $(m+1) \cdot (n+1)$ multiplications and mn additions. □

The sparse algorithm is illustrated by an (not very sparse) example:
For $s = 3, t = 4$, we want to multiply $X^3 + 7X + 9$ and $X^4 + X^2 + 3X + 2$ over the integers.
First, we calculate all $s \cdot t$ monomials:

The sparse algorithm is illustrated by an (not very sparse) example: For $s = 3, t = 4$, we want to multiply $X^3 + 7X + 9$ and $X^4 + X^2 + 3X + 2$ over the integers.
First, we calculate all $s \cdot t$ monomials:

$$\begin{array}{ccc} X^7 & 7X^5 & 9X^4 \\ X^5 & 7X^3 & 9X^2 \\ 3X^4 & 21X^2 & 27X \\ 2X^3 & 14X & 18 \end{array}$$

The sparse algorithm is illustrated by an (not very sparse) example: For $s = 3, t = 4$, we want to multiply $X^3 + 7X + 9$ and $X^4 + X^2 + 3X + 2$ over the integers.
First, we calculate all $s \cdot t$ monomials:

$$\begin{array}{ccc} X^7 & 7X^5 & 9X^4 \\ X^5 & 7X^3 & 9X^2 \\ 3X^4 & 21X^2 & 27X \\ 2X^3 & 14X & 18 \end{array}$$

Then we fuse, sort, and where possible, add, neighbouring rows:

$$\begin{array}{ccccc} X^7 & 8X^5 & 9X^4 & 7X^3 & 9X^2 \\ 3X^4 & 2X^3 & 21X^2 & 41X & 18 \end{array}$$

Sorting again, we have:

$$X^7 \quad 8X^5 \quad 12X^4 \quad 9X^3 \quad 30X^2 \quad 41X \quad 18$$

Theorem 11

Let R be an integral domain and $P_1(X)$ and $P_2(X)$ two polynomials over R with $lc(P_2)$ a unit in R. Then there exist unique $Q(X), R(X)$, so that:

$$P_1(X) = Q(X) \cdot P_2(X) + R(X) \text{ and } deg(R(X)) < deg(P_2(X)).$$

Theorem 11

Let R be an integral domain and $P_1(X)$ and $P_2(X)$ two polynomials over R with $lc(P_2)$ a unit in R. Then there exist unique $Q(X), R(X)$, so that:

$$P_1(X) = Q(X) \cdot P_2(X) + R(X) \text{ and } deg(R(X)) < deg(P_2(X)).$$

Definition 12

In this situation, we call $Q(X) =: quo(P_1(X), P_2(X))$ the quotient and $R(X) =: rem(P_1(X), P_2(X))$ the remainder of $P_1(X), P_2(X)$.

Let $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$, $m \geq n \geq 0$, and b_n be a unit. The algorithm for polynomial division is:

Let $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$, $m \geq n \geq 0$, and b_n be a unit. The algorithm for polynomial division is:

$$
\begin{array}{l}
\underline{\texttt{for}}\ i = m - n\ \underline{\texttt{down to}}\ 0\ \underline{\texttt{do}} \\
\quad q_i := a_{n+i} b_n^{-1} \\
\quad \underline{\texttt{for}}\ l = n + i - 1\ \underline{\texttt{down to}}\ i\ \underline{\texttt{do}} \\
\quad\quad a_l := a_l - q_i b_{l-i} \\
\quad \underline{\texttt{od}} \\
\underline{\texttt{od}}
\end{array}
$$

Let $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$, $m \geq n \geq 0$, and b_n be a unit. The algorithm for polynomial division is:

```
for i = m − n down to 0 do
   q_i := a_{n+i} b_n^{-1}
   for l = n + i − 1 down to i do
      a_l := a_l − q_i b_{l−i}
   od
od
```

Then, $Q(X) = \sum_{i=0}^{m-n} q_i X^i$, and $R(X) = \sum_{l=0}^{n-1} a_l X^l$.

Let $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$, $m \geq n \geq 0$, and b_n be a unit. The algorithm for polynomial division is:

$$
\begin{array}{l}
\underline{\texttt{for}}\ i = m - n\ \underline{\texttt{down to}}\ 0\ \underline{\texttt{do}} \\
\quad q_i := a_{n+i} b_n^{-1} \\
\quad \underline{\texttt{for}}\ l = n + i - 1\ \underline{\texttt{down to}}\ i\ \underline{\texttt{do}} \\
\quad\quad a_l := a_l - q_i b_{l-i} \\
\quad \underline{\texttt{od}} \\
\underline{\texttt{od}}
\end{array}
$$

Then, $Q(X) = \sum_{i=0}^{m-n} q_i X^i$, and $R(X) = \sum_{l=0}^{n-1} a_l X^l$.
Computing time: Assuming that operations in R take $O(1)$, the whole algorithm is done in $O(n(m - n + 1))$.

Let R be an integral domain.

Definition 13

For $P(X) \in R[X]$, $\alpha \in R$ is called a root of $P(X)$, if $P(\alpha)=0$.

Let R be an integral domain.

Definition 13

For $P(X) \in R[X]$, $\alpha \in R$ is called a root of $P(X)$, if $P(\alpha)=0$.

Theorem 14

$\alpha \in R$ is a root of $P(X)$ if $(X - \alpha)$ divides $P(X)$.

Let R be an integral domain.

Definition 13

For $P(X) \in R[X]$, $\alpha \in R$ is called a root of $P(X)$, if $P(\alpha)=0$.

Theorem 14

$\alpha \in R$ is a root of $P(X)$ if $(X - \alpha)$ divides $P(X)$.

Proof.

Observe that $rem(P(X), (X - \alpha)) = P(\alpha)$. □

Let R be an integral domain.

Definition 13

For $P(X) \in R[X]$, $\alpha \in R$ is called a root of $P(X)$, if $P(\alpha)=0$.

Theorem 14

$\alpha \in R$ is a root of $P(X)$ if $(X - \alpha)$ divides $P(X)$.

Proof.

Observe that $rem(P(X),(X - \alpha)) = P(\alpha)$. □

Definition 15

$\alpha \in R$ is a root with multiplicity m, if $(X - \alpha)^m$ divides $P(X)$.

Let R be an integral domain.

Definition 13

For $P(X) \in R[X]$, $\alpha \in R$ is called a root of $P(X)$, if $P(\alpha)=0$.

Theorem 14

$\alpha \in R$ is a root of $P(X)$ if $(X - \alpha)$ divides $P(X)$.

Proof.

Observe that $rem(P(X), (X - \alpha)) = P(\alpha)$. □

Definition 15

$\alpha \in R$ is a root with multiplicity m, if $(X - \alpha)^m$ divides $P(X)$.

Theorem 16

If $P(X) \neq 0$, $P(X)$ can have at most $deg(P(X))$ roots, counting multiplicities.

Definition 17

Let K be a field, and $M(X) \in K[X]$ with $deg(M(X)) > 0$. Then we can define the equivalence relation $\equiv_{M(X)}$ on $K[X]$:

$$P(X) \equiv_{M(X)} Q(X) \text{ if } rem(P(X), M(X)) = rem(Q(X), M(X)).$$

Definition 17

Let K be a field, and $M(X) \in K[X]$ with $deg(M(X)) > 0$. Then we can define the equivalence relation $\equiv_{M(X)}$ on $K[X]$:

$$P(X) \equiv_{M(X)} Q(X) \text{ if } rem(P(X), M(X)) = rem(Q(X), M(X)).$$

The set of equivalence classes, denoted by $K[X]_{M(X)}$, together with the operations

$$[P(X)]_{M(X)} + [Q(X)]_{M(X)} := [P(X) + Q(X)]_{M(X)}$$

$$[P(X)]_{M(X)} \cdot [Q(X)]_{M(X)} := [P(X) \cdot Q(X)]_{M(X)}$$

is a commutative ring with unity.

Definition 18

A polynomial $P(X) \in R[X]$, R an integral domain, is called irreducible, if, whenever $P(X) = P_1(X) \cdot P_2(X)$, $P_1(X)$ or $P_2(X)$ is a unit of $R[X]$.

Definition 18

A polynomial $P(X) \in R[X]$, R an integral domain, is called irreducible, if, whenever $P(X) = P_1(X) \cdot P_2(X)$, $P_1(X)$ or $P_2(X)$ is a unit of $R[X]$.

Example 19

$2X^2 + 4$ is reducible both over $\mathbb{Z}$ and $\mathbb{C}$, but not over $\mathbb{R}$.

Definition 18

A polynomial $P(X) \in R[X]$, R an integral domain, is called irreducible, if, whenever $P(X) = P_1(X) \cdot P_2(X)$, $P_1(X)$ or $P_2(X)$ is a unit of $R[X]$.

Example 19

$2X^2 + 4$ is reducible both over $\mathbb{Z}$ and $\mathbb{C}$, but not over $\mathbb{R}$.

Theorem 20

For K a field and $M(X) \in K[X]$ with $deg(M(X)) > 0$, $K[X]_{M(X)}$ is a field if and only if $M(X)$ is irreducible over K.

Definition 18

A polynomial $P(X) \in R[X]$, R an integral domain, is called irreducible, if, whenever $P(X) = P_1(X) \cdot P_2(X)$, $P_1(X)$ or $P_2(X)$ is a unit of $R[X]$.

Example 19

$2X^2 + 4$ is reducible both over $\mathbb{Z}$ and $\mathbb{C}$, but not over $\mathbb{R}$.

Theorem 20

For K a field and $M(X) \in K[X]$ with deg$(M(X)) > 0$, $K[X]_{M(X)}$ is a field if and only if $M(X)$ is irreducible over K.

In this case, $K[X]_{M(X)}$ contains a subfield isomorphic to K and is therefore a field extension of K.

Example 21

Let $K = \mathbb{R}$ and $M(X) = X^2 + 1$. Then all elements of $\mathbb{R}[X]_{X^2+1}$ are of the form $a \cdot [1] + b \cdot [X]$ with $a, b \in \mathbb{R}$. Addition and multiplication are given by:

Example 21

Let $K = \mathbb{R}$ and $M(X) = X^2 + 1$. Then all elements of $\mathbb{R}[X]_{X^2+1}$ are of the form $a \cdot [1] + b \cdot [X]$ with $a, b \in \mathbb{R}$. Addition and multiplication are given by:

$$(a[1] + b[X]) + (c[1] + d[X]) = (a + c)[1] + (b + d)[X]$$

$$(a[1] + b[X]) \cdot (c[1] + d[X]) = ac[1] + bd[X^2] + ad[X] + bc[X]$$

$$= (ac - bd)[1] + (ad + bc)[X]$$

Example 21

Let $K = \mathbb{R}$ and $M(X) = X^2 + 1$. Then all elements of $\mathbb{R}[X]_{X^2+1}$ are of the form $a \cdot [1] + b \cdot [X]$ with $a, b \in \mathbb{R}$. Addition and multiplication are given by:

$$(a[1] + b[X]) + (c[1] + d[X]) = (a + c)[1] + (b + d)[X]$$

$$\begin{aligned}(a[1] + b[X]) \cdot (c[1] + d[X]) &= ac[1] + bd[X^2] + ad[X] + bc[X] \\ &= (ac - bd)[1] + (ad + bc)[X]\end{aligned}$$

Therefore, $\mathbb{R}[X]_{X^2+1}$ is isomorphic to $\mathbb{C}$.

Definition 22

A field K is algebraically closed, if every nonconstant polynomial with coefficients in K has a root in K.

Definition 22

A field K is algebraically closed, if every nonconstant polynomial with coefficients in K has a root in K.

Theorem 23

Every field J has an algebraic closure, i.e. a field extension K that is algebraically closed.

Definition 22

A field K is algebraically closed, if every nonconstant polynomial with coefficients in K has a root in K.

Theorem 23

Every field J has an algebraic closure, i.e. a field extension K that is algebraically closed.

For us, it is important to know the

Theorem 24

(Fundamental Theorem of Algebra) $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$.

Definition 25

For $a, b \in R$, R an integral domain, $d \in R$ is called a greatest common divisor of a and b, $d = gcd(a, b)$, if d divides a and b, and every $t \in R$ dividing a and b divides d, too.

Definition 25

For $a, b \in R$, R an integral domain, $d \in R$ is called a greatest common divisor of a and b, $d = gcd(a, b)$, if d divides a and b, and every $t \in R$ dividing a and b divides d, too.

If a $gcd(a, b)$ exists, it is unique up to units, and thus it makes sense to speak of *the gcd* of a and b.

Theorem 26

Let K be a field, and $P_1(X)$, $P_2(X) \neq 0$ polynomials from $K[X]$. Then there exists $gcd(P_1(X), P_2(X)) \in K[X]$, and there are $A(X), B(X) \in K[X]$, with $deg(A(X)) < deg(P_2(X))$ and $deg(B(X)) < deg(P_1(X))$ with

$$gcd(P_1(X), P_2(X)) = A(X) \cdot P_1(X) + B(X) \cdot P_2(X).$$

Theorem 26

Let K be a field, and $P_1(X)$, $P_2(X) \neq 0$ polynomials from $K[X]$. Then there exists $gcd(P_1(X), P_2(X)) \in K[X]$, and there are $A(X), B(X) \in K[X]$, with $deg(A(X)) < deg(P_2(X))$ and $deg(B(X)) < deg(P_1(X))$ with

$$gcd(P_1(X), P_2(X)) = A(X) \cdot P_1(X) + B(X) \cdot P_2(X).$$

Proof.

We construct both $gcd(P_1(X), P_2(X))$ and $A(X), B(X)$ by the extended Euclidean Algorithm over a field. □

Extended Euclidean Algorithm

Extended Euclidean Algorithm

$$[A(X), B(X)] := [1, 0]$$
$$[a(X), b(X)] := [0, 1]$$

Extended Euclidean Algorithm

$[A(X), B(X)] := [1, 0]$
$[a(X), b(X)] := [0, 1]$
$\underline{\texttt{while}}\ P_2(X) \neq 0\ \underline{\texttt{do}}$
$\quad [Q(X), R(X)] := [quo(P_1(X), P_2(X)), rem(P_1(X), P_2(X))]$
$\quad [P_1(X), P_2(X)] := [P_2(X), R(X)]$

Extended Euclidean Algorithm

$[A(X), B(X)] := [1, 0]$
$[a(X), b(X)] := [0, 1]$
while $P_2(X) \neq 0$ do
  $[Q(X), R(X)] := [quo(P_1(X), P_2(X)), rem(P_1(X), P_2(X))]$
  $[P_1(X), P_2(X)] := [P_2(X), R(X)]$
  $[A(X), a(X)] := [a(X), A(X) - Q(X)a(X)]$
  $[B(X), b(X)] := [b(X), B(X) - Q(X)b(X)]$
od

Extended Euclidean Algorithm

$[A(X), B(X)] := [1, 0]$
$[a(X), b(X)] := [0, 1]$
while $P_2(X) \neq 0$ **do**
$\quad [Q(X), R(X)] := [quo(P_1(X), P_2(X)), rem(P_1(X), P_2(X))]$
$\quad [P_1(X), P_2(X)] := [P_2(X), R(X)]$
$\quad [A(X), a(X)] := [a(X), A(X) - Q(X)a(X)]$
$\quad [B(X), b(X)] := [b(X), B(X) - Q(X)b(X)]$
od
return $[P_1(X), A(X), B(X)]$

Polynomial division in the Euclidean algorithm works, because $lc(P_2(X))$ is a unit throughout the algorithm, because K is a field. From now on, we consider Polynomials over $\mathbb{Z}$, and the Euclidean algorithm will not work in general.

Polynomial division in the Euclidean algorithm works, because $lc(P_2(X))$ is a unit throughout the algorithm, because K is a field. From now on, we consider Polynomials over $\mathbb{Z}$, and the Euclidean algorithm will not work in general.

Let $P_1(X) = \sum_{i=0}^{m} a_i X^i$, $P_2(X) = \sum_{j=0}^{n} b_j X^j \neq 0$, $m \geq n$. For a pseudodivision in $\mathbb{Z}[X]$, premultiply $P_1(X)$ by b_n^{m-n+1}, and define pseudoquotient and pseudoremainder by

$$pquo(P_1(X), P_2(X)) = quo(b_n^{m-n+1} \cdot P_1(X), P_2(X))$$

$$prem(P_1(X), P_2(X)) = rem(b_n^{m-n+1} \cdot P_1(X), P_2(X)).$$

Definition 27

For $P(X) \in \mathbb{Z}[X]$, define the content $cont(P(X))$ as gcd of the coefficients of $P(X)$, and the primitive part $pp(P(X)) = \frac{P(X)}{cont(P(X))}$.

Definition 27

For $P(X) \in \mathbb{Z}[X]$, define the content $cont(P(X))$ as gcd of the coefficients of $P(X)$, and the primitive part $pp(P(X)) = \frac{P(X)}{cont(P(X))}$.

Theorem 28

$\mathbb{Z}[X]$ is a unique factorization domain, and therefore, a gcd exists for all pairs of nonzero Polynomials over Z.

Definition 27

For $P(X) \in \mathbb{Z}[X]$, define the content $cont(P(X))$ as gcd of the coefficients of $P(X)$, and the primitive part $pp(P(X)) = \frac{P(X)}{cont(P(X))}$.

Theorem 28

$\mathbb{Z}[X]$ is a unique factorization domain, and therefore, a gcd exists for all pairs of nonzero Polynomials over Z.

For $P_1(X), P_2(X) \in \mathbb{Z}[X]$, we have

$$\begin{aligned} cont(gcd(P_1(X), P_2(X))) &= gcd(cont(P_1(X)), cont(P_2(X))) \\ pp(gcd(P_1(X), P_2(X))) &= gcd(pp(P_1(X)), pp(P_2(X))). \end{aligned}$$

Generalized Euclidean Algorithm

$$c := gcd(cont(P_1(X)), cont(P_2(X)))$$
$$[P_1(X), P_2(X)] := [pp(P_1(X)), pp(P_2(X))]$$

Generalized Euclidean Algorithm

$c := gcd(cont(P_1(X)), cont(P_2(X)))$
$[P_1(X), P_2(X)] := [pp(P_1(X)), pp(P_2(X))]$
<u>while</u> $P_2(X) \neq 0$ <u>do</u>
   $[P_1(X), P_2(X)] := [P_2(X), prem(P_1(X), P_2(X))]$
<u>od</u>

Generalized Euclidean Algorithm

$c := gcd(cont(P_1(X)), cont(P_2(X)))$
$[P_1(X), P_2(X)] := [pp(P_1(X)), pp(P_2(X))]$
while $P_2(X) \neq 0$ do
$\quad [P_1(X), P_2(X)] := [P_2(X), prem(P_1(X), P_2(X))]$
od
return $c \cdot pp(P_1(X))$

Example 29

$P_1(X) = X^3 - 2X^2 + 3 + 1,\ P_2(X) = 2X^2 + 1$

Example 29

$P_1(X) = X^3 - 2X^2 + 3 + 1,\ P_2(X) = 2X^2 + 1$

$$2^2 \cdot (X^3 - 2X^2 + 3 + 1) \quad = \quad (2X - 4) \cdot (2X^2 + 1) + (10X + 8)$$

Example 29

$P_1(X) = X^3 - 2X^2 + 3 + 1,\ P_2(X) = 2X^2 + 1$

$$\begin{aligned} 2^2 \cdot (X^3 - 2X^2 + 3 + 1) &= (2X - 4) \cdot (2X^2 + 1) + (10X + 8) \\ 10^2 \cdot (2X^2 + 1) &= (20X - 16) \cdot (10X + 8) + 228 \end{aligned}$$

Example 29

$P_1(X) = X^3 - 2X^2 + 3 + 1,\ P_2(X) = 2X^2 + 1$

$$
\begin{aligned}
2^2 \cdot (X^3 - 2X^2 + 3 + 1) &= (2X - 4) \cdot (2X^2 + 1) + (10X + 8) \\
10^2 \cdot (2X^2 + 1) &= (20X - 16) \cdot (10X + 8) + 228 \\
228^2 \cdot (10X - 8) &= (2280X - 1824) \cdot 228 + 0
\end{aligned}
$$

Example 29

$P_1(X) = X^3 - 2X^2 + 3 + 1,\ P_2(X) = 2X^2 + 1$

$$\begin{aligned} 2^2 \cdot (X^3 - 2X^2 + 3 + 1) &= (2X - 4) \cdot (2X^2 + 1) + (10X + 8) \\ 10^2 \cdot (2X^2 + 1) &= (20X - 16) \cdot (10X + 8) + 228 \\ 228^2 \cdot (10X - 8) &= (2280X - 1824) \cdot 228 + 0 \end{aligned}$$

$gcd(cont(P_1(X)), cont(P_2(X))) = 1$, and therefore, $gcd(P_1(X), P_2(X)) = 1$.
Premultiplication leads to an exponential growth of coefficients, the greatest number in our calculation was 519840.

One possiblility to reduce the the coefficient growth, is to divide every pseudoremainder by its content.

One possiblility to reduce the the coefficient growth, is to divide every pseudoremainder by its content.
The problem is,that we would have to do a *gcd* calculation in $\mathbb{Z}$ at every step of our algorithm.

One possiblility to reduce the the coefficient growth, is to divide every pseudoremainder by its content.
The problem is,that we would have to do a *gcd* calculation in $\mathbb{Z}$ at every step of our algorithm.
Before we go on with *gcd* computations, we ask what it means, if two Polynomials in $\mathbb{Z}[X]$ have a common root in $\mathbb{C}$.

Definition 30

For two Polynomials $P_1(X) = \sum_{j=0}^{m} a_j X^j, P_2(X) = \sum_{k=0}^{n} b_k X^k$ in $\mathbb{Z}[X]$, we define the resultant

$$res[P_1(X), P_2(X)] := a_m^n b_n^m \prod_{j=0}^{m} \prod_{k=0}^{n} (\alpha_j - \beta_k).$$

where α_j are the roots of P_1, β_k of P_2.

Definition 30

For two Polynomials $P_1(X) = \sum_{j=0}^{m} a_j X^j, P_2(X) = \sum_{k=0}^{n} b_k X^k$ in $\mathbb{Z}[X]$, we define the resultant

$$res[P_1(X), P_2(X)] := a_m^n b_n^m \prod_{j=0}^{m} \prod_{k=0}^{n} (\alpha_j - \beta_k).$$

where α_j are the roots of P_1, β_k of P_2.

Theorem 31

Definition 30

For two Polynomials $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$ in $\mathbb{Z}[X]$, we define the resultant

$$res[P_1(X), P_2(X)] := a_m^n b_n^m \prod_{j=0}^{m} \prod_{k=0}^{n} (\alpha_j - \beta_k).$$

where α_j are the roots of P_1, β_k of P_2.

Theorem 31

1. *$res[P_1(X), P_2(X)] = 0$ if $P_1(X)$ and $P_2(X)$ have a common root.*

Definition 30

For two Polynomials $P_1(X) = \sum_{j=0}^{m} a_j X^j$, $P_2(X) = \sum_{k=0}^{n} b_k X^k$ in $\mathbb{Z}[X]$, we define the resultant

$$res[P_1(X), P_2(X)] := a_m^n b_n^m \prod_{j=0}^{m} \prod_{k=0}^{n} (\alpha_j - \beta_k).$$

where α_j are the roots of P_1, β_k of P_2.

Theorem 31

1. *$res[P_1(X), P_2(X)] = 0$ if $P_1(X)$ and $P_2(X)$ have a common root.*
2. $res[P_1(X), P_2(X)] = (-1)^{mn} b_n^m \prod_{k=1}^{n} P_1(\beta_k) = a_m^n \prod_{j=1}^{m} P_2(\alpha_j)$

Theorem 32

$$res(P_1(X), P_2(X)) = det \begin{pmatrix} a_m & & \cdots & \cdots & & a_0 & & \mathbf{0} \\ & \ddots & & & & & \ddots & \\ \mathbf{0} & & a_m & & \cdots & \cdots & & a_0 \\ b_n & \cdots & \cdots & b_0 & & & & \\ & & & & & & \mathbf{0} & \\ & & \ddots & & & \ddots & & \\ & \mathbf{0} & & & & & & \\ & & & & b_n & \cdots & \cdots & b_n \end{pmatrix}$$

Theorem 32

$$res(P_1(X), P_2(X)) \quad = det \begin{pmatrix} a_m & & \cdots & \cdots & & a_0 & & \mathbf{0} \\ & \ddots & & & & & \ddots & \\ \mathbf{0} & & a_m & & \cdots & \cdots & & a_0 \\ b_n & \cdots & \cdots & b_0 & & & & \\ & & & & & & \mathbf{0} & \\ & & \ddots & & & \ddots & & \\ & \mathbf{0} & & & & & & \\ & & & & b_n & \cdots & \cdots & b_n \end{pmatrix}$$

The Matrix is $(m+n) \times (m+n)$ *and contains* n *"a" rows and* m *"b" rows.*

Proof.

Consider the polynomial

$$q(\lambda) \quad := det \begin{pmatrix} a_m & & \cdots & \cdots & & a_0 - \lambda & & \mathbf{0} \\ & \ddots & & & & & \ddots & \\ \mathbf{0} & & a_m & & \cdots & \cdots & & a_0 - \lambda \\ b_n & \cdots & \cdots & b_0 & & & & \\ & & & & & & \mathbf{0} & \\ & & \ddots & & & \ddots & & \\ & \mathbf{0} & & & & & & \\ & & & & b_n & \cdots & \cdots & b_n \end{pmatrix},$$

and assume the simple case, that $P_2(X)$ has only single roots β_i, and that all $P_1(\beta_i)$ are different.

Then, for all $1 \leq i \leq n$, $\lambda = P_1(\beta_i)$ is a root of $q(\lambda)$, and because $q(\lambda)$ has at most n different roots, $P_1(\beta_i)$ are all roots.

Then, for all $1 \leq i \leq n$, $\lambda = P_1(\beta_i)$ is a root of $q(\lambda)$, and because $q(\lambda)$ has at most n different roots, $P_1(\beta_i)$ are all roots.
Defining $q_n = lc(q(\lambda))$ and $q_0 = q(0)$, we have:

$$(-1)^n q_n \prod_{k=1}^{n} P_1(\beta_k) = q_0.$$

Then, for all $1 \leq i \leq n$, $\lambda = P_1(\beta_i)$ is a root of $q(\lambda)$, and because $q(\lambda)$ has at most n different roots, $P_1(\beta_i)$ are all roots.
Defining $q_n = lc(q(\lambda))$ and $q_0 = q(0)$, we have:

$$(-1)^n q_n \prod_{k=1}^{n} P_1(\beta_k) = q_0.$$

And, by the structure of the matrix:

$$q(\lambda) = (-1)^{mn} b_n^m \cdot (-\lambda)^n + ...$$

Then, for all $1 \leq i \leq n$, $\lambda = P_1(\beta_i)$ is a root of $q(\lambda)$, and because $q(\lambda)$ has at most n different roots, $P_1(\beta_i)$ are all roots.
Defining $q_n = lc(q(\lambda))$ and $q_0 = q(0)$, we have:

$$(-1)^n q_n \prod_{k=1}^{n} P_1(\beta_k) = q_0.$$

And, by the structure of the matrix:

$$q(\lambda) = (-1)^{mn} b_n^m \cdot (-\lambda)^n + ...$$

Therefore,

$$(-1)^{mn} b_n^m \prod_{k=1}^{n} P_1(\beta_k) = q_0.$$

Instead of

$$(lc(P_{i+1}(X)))^{n_i - n_i + 1 + 1} P_i(X) = P_{i+1}(X) Q_i(X) + P_{i+2}(X),$$

calculate

$$(lc(P_{i+1}(X)))^{n_i - n_i + 1 + 1} P_i(X) = P_{i+1}(X) Q_i(X) + \beta_{\mathbf{i}} P_{i+2}(X).$$

Where

$$\beta_1 = (-1)^{n_1 - n_2 + 1}, \; \beta_i = (-1)^{n_i - n_{i+1} + 1} lc(P_i(X)) \cdot H_i^{n_i - n_{i+1}},$$

and

$$H_2 = (lc(P_2(X))^{n_1 - n_2}, \; H_i = (lc(P_i(X))^{n_{i-1} - n_i} H_{i-1}^{1 + n_i - n_{i-1}}.$$

Now we will consider the real roots of polynomials in $\mathbb{Z}[X]$. We are interested in methods to count them, in order to isolate and finally approximate them.

Theorem 33

Let $p(x)$ be a polynomial in $\mathbb{R}[x]$. Then, for a real root y of multiplicity m, we have, that the sequence

$$[p(y-\epsilon), p'(y-\epsilon), ..., p^{(m)}(y-\epsilon)]$$

has alternating sign, while the elements of

$$[p(y+\epsilon), p'(y+\epsilon), ..., p^{(m)}(y+\epsilon)]$$

have the same sign, for ϵ sufficiently small.

Definition 34

For a polynomial $p(x)$, with $n = deg(p(x)) > 0$, the Fourier sequence is defined as $fseq(x) := [p(x), p^{(1)}(x), ..., p^{(n)}(x)]$.

Definition 34

For a polynomial $p(x)$, with $n = deg(p(x)) > 0$, the Fourier sequence is defined as $fseq(x) := [p(x), p^{(1)}(x), ..., p^{(n)}(x)]$.

Definition 35

For a sequence of real numbers $S = [a_0, ..., a_n]$, we say that there is a sign variation between a_i and a_j, if a_i and a_j have opposite sign, and all members between (if there are any) are zero.
The number of sign variations is denoted by $Var(S)$.

Definition 34

For a polynomial $p(x)$, with $n = deg(p(x)) > 0$, the Fourier sequence is defined as $fseq(x) := [p(x), p^{(1)}(x), ..., p^{(n)}(x)]$.

Definition 35

For a sequence of real numbers $S = [a_0, ..., a_n]$, we say that there is a sign variation between a_i and a_j, if a_i and a_j have opposite sign, and all members between (if there are any) are zero.
The number of sign variations is denoted by $Var(S)$.

Theorem 36

(Fourier) For real numbers $a < b$, we have: The number N of roots in $(a, b]$, counting multiplicities is bounded by:

$$N = V(fseq(a)) - V(fseq(b)) - 2 \cdot \lambda, \lambda \geq 0.$$

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

Definition 37

For $p(x) \in \mathbb{R}[x]$ a generalized Sturm sequence is a sequence of polynomials $gsseq(x) := [p(x), p_1(x), ..., p_{k+1}(x)]$, so that:

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

Definition 37

For $p(x) \in \mathbb{R}[x]$ a generalized Sturm sequence is a sequence of polynomials $gsseq(x) := [p(x), p_1(x), ..., p_{k+1}(x)]$, so that:

- In a sufficiently small neighbourhood of every zero y of $p(x)$, $p(x)$ and $p_1(x)$ have opposite signs for $x < y$, and same signs for $x \geq y$.

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

Definition 37

For $p(x) \in \mathbb{R}[x]$ a generalized Sturm sequence is a sequence of polynomials $gsseq(x) := [p(x), p_1(x), \ldots, p_{k+1}(x)]$, so that:

- In a sufficiently small neighbourhood of every zero y of $p(x)$, $p(x)$ and $p_1(x)$ have opposite signs for $x < y$, and same signs for $x \geq y$.
- Consecutive members do not vanish simultaneously.

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

Definition 37

For $p(x) \in \mathbb{R}[x]$ a generalized Sturm sequence is a sequence of polynomials $gsseq(x) := [p(x), p_1(x), ..., p_{k+1}(x)]$, so that:

- In a sufficiently small neighbourhood of every zero y of $p(x)$, $p(x)$ and $p_1(x)$ have opposite signs for $x < y$, and same signs for $x \geq y$.
- Consecutive members do not vanish simultaneously.
- The two neighbours of a vanishing member have opposite sign.

With his theorem, Fourier could only give an upper bound. Sturm gave a method for exact counting:

Definition 37

For $p(x) \in \mathbb{R}[x]$ a generalized Sturm sequence is a sequence of polynomials $gsseq(x) := [p(x), p_1(x), ..., p_{k+1}(x)]$, so that:

- In a sufficiently small neighbourhood of every zero y of $p(x)$, $p(x)$ and $p_1(x)$ have opposite signs for $x < y$, and same signs for $x \geq y$.
- Consecutive members do not vanish simultaneously.
- The two neighbours of a vanishing member have opposite sign.
- $p_{k+1}(x)$ has no real roots, and thus always the same sign.

For $p(x)$ without multiple roots in $\mathbb{R}$, one possible *gsseq* is

$$sseq(x) = [p(x), p'(x), r_1(x), ..., r_k(x)],$$

with

$$r_{j-2}(x) := r_{j-1}(x) q_k(x) - r_j(x).$$

For $p(x)$ without multiple roots in $\mathbb{R}$, one possible *gsseq* is

$$sseq(x) = [p(x), p'(x), r_1(x), ..., r_k(x)],$$

with

$$r_{j-2}(x) := r_{j-1}(x) q_k(x) - r_j(x).$$

Theorem 38

(Sturm) For real numbers $a < b$, we have:

$$|\{a < x \leq b : p(x) = 0\}| = V(gsseq(a)) - V(gsseq(b)).$$

Now we also have a method for counting the complex roots of $p(x)$:

Theorem 39

Let $p(x)$ be a polynomial of degree n, and let $gsseq(x)$ be a complete sequence (i.e. it contains $n+1$ members). Then $p(x)$ has as many pairs of complex roots as there are sign variations in the sequence of leading coefficients in $gsseq(x)$.

Theorem 40

(Cauchy) If $p(x) = \sum_{j=0}^{n} c_j x^j$ with $c_n > 0$ has got $\lambda \geq 0$ negative coefficients,

$$b := max_{\{1 \leq k < n : c_{n-k} < 0\}} \{ | \frac{\lambda c_{n-k}}{c_n} |^{\frac{1}{k}} \}$$

is an upper bound for the positive roots of $p(x)$.

Theorem 40

(Cauchy) If $p(x) = \sum_{j=0}^{n} c_j x^j$ with $c_n > 0$ has got $\lambda \geq 0$ negative coefficients,

$$b := max_{\{1 \leq k < n: c_{n-k} < 0\}}\{|\frac{\lambda c_{n-k}}{c_n}|^{\frac{1}{k}}\}$$

is an upper bound for the positive roots of $p(x)$.

Now, we are ready to understand Sturms Bisection algorithm for isolation of real roots. For $p(x) \in \mathbb{Z}[x]$ with only single roots the algorithm will

- determine, whether 0 is a root,

Theorem 40

(Cauchy) If $p(x) = \sum_{j=0}^{n} c_j x^j$ with $c_n > 0$ has got $\lambda \geq 0$ negative coefficients,

$$b := max_{\{1 \leq k < n : c_{n-k} < 0\}} \{ | \frac{\lambda c_{n-k}}{c_n} |^{\frac{1}{k}} \}$$

is an upper bound for the positive roots of $p(x)$.

Now, we are ready to understand Sturms Bisection algorithm for isolation of real roots. For $p(x) \in \mathbb{Z}[x]$ with only single roots the algorithm will

- determine, whether 0 is a root,
- calculate a bound on the positive roots, obtain isolation intervals using bisection and Sturms theorem,

Theorem 40

(Cauchy) If $p(x) = \sum_{j=0}^{n} c_j x^j$ *with* $c_n > 0$ *has got* $\lambda \geq 0$ *negative coefficients,*

$$b := max_{\{1 \leq k < n : c_{n-k} < 0\}} \{ | \frac{\lambda c_{n-k}}{c_n} |^{\frac{1}{k}} \}$$

is an upper bound for the positive roots of $p(x)$.

Now, we are ready to understand Sturms Bisection algorithm for isolation of real roots. For $p(x) \in \mathbb{Z}[x]$ with only single roots the algorithm will

- determine, whether 0 is a root,
- calculate a bound on the positive roots, obtain isolation intervals using bisection and Sturms theorem,
- do the same for the negative roots.

Without a derivation: The Sturm bisection method is performed in $O(n^7 L^3[|p(x)|_\infty])$, where $L[m] := \lfloor log_2(|m|) \rfloor + 1$.

Without a derivation: The Sturm bisection method is performed in $O(n^7 L^3[\|p(x)|_\infty])$, where $L[m] := \lfloor log_2(|m|)\rfloor + 1$.

Example 41

For $p(x) = x^3 + 2x^2 - x - 2$ we have:
$sseq(x) = [x^3 + 2x^2 - x - 2, 3x^2 + 4x - 1, 7x + 8, 1]$.

Without a derivation: The Sturm bisection method is performed in $O(n^7 L^3[|p(x)|_\infty])$, where $L[m] := \lfloor log_2(|m|) \rfloor + 1$.

Example 41

For $p(x) = x^3 + 2x^2 - x - 2$ we have:
$sseq(x) = [x^3 + 2x^2 - x - 2, 3x^2 + 4x - 1, 7x + 8, 1]$.
$b_p = 2$ is a bound for positive roots, $b_n = -4$ is a bound for negative roots.

Without a derivation: The Sturm bisection method is performed in $O(n^7 L^3[|p(x)|_\infty])$, where $L[m] := \lfloor log_2(|m|) \rfloor + 1$.

Example 41

For $p(x) = x^3 + 2x^2 - x - 2$ we have:
$sseq(x) = [x^3 + 2x^2 - x - 2, 3x^2 + 4x - 1, 7x + 8, 1]$.
$b_p = 2$ is a bound for positive roots, $b_n = -4$ is a bound for negative roots.
The algorithm directly finds the root -2, and returns $(-2, 0)$ and $(0, 2)$ as isolation intervals for the roots -1 and 1.

Theorem 42

(Budan, equivalent to Fourier) Let $a < b$ be real and consider $p(x) \in \mathbb{R}[x]$. The number of roots that $p(x)$ has in $(a, b]$ is never greater than the loss of sign variations in the coefficient sequence of $p(x + b)$ compared to $p(x + a)$.

Theorem 42

(Budan, equivalent to Fourier) Let $a < b$ be real and consider $p(x) \in \mathbb{R}[x]$. The number of roots that $p(x)$ has in $(a, b]$ is never greater than the loss of sign variations in the coefficient sequence of $p(x + b)$ compared to $p(x + a)$.

Definition 43

For a nonsingular matrix $\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, define the Möbius substitution by $y := \mathbf{M}(x) = \frac{a \cdot x + b}{c \cdot x + d}$.

Theorem 44

For a polynomial $p(x)$ with rational coefficients and without multiple roots, and for $a_1 \geq 0$, $a_i > 0, i > 1$ there is always $m \in \mathbb{N}$ and the corresponding transformation

$$x := a_1 + \cfrac{1}{a_2 + \cfrac{}{\ddots \cfrac{1}{a_m + \frac{1}{y}}}}$$

so that the transformed polynomial $p_{ti}(y)$ has at most one sign variation in its coefficient sequence.

The continued fraction transformation can be written as Möbius substitution:

$$x = \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_m & 1 \\ 1 & 0 \end{bmatrix} (y).$$

The continued fraction transformation can be written as Möbius substitution:

$$x = \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_m & 1 \\ 1 & 0 \end{bmatrix} (y).$$

Theorem 45

(Cardano-Descartes) A polynomial with no or exactly one sign variation in its coefficient sequence has no or exactly one positive root, respectivly.

The continued fraction isolation algorithm returns isolation intervals for polynomials $p(x)$ in $\mathbb{Z}[x]$ without multiple roots. It will:

The continued fraction isolation algorithm returns isolation intervals for polynomials $p(x)$ in $\mathbb{Z}[x]$ without multiple roots. It will:

- Calculate lower bounds for the positive zeroes of $p(x)$ and transformed polynomials.

The continued fraction isolation algorithm returns isolation intervals for polynomials $p(x)$ in $\mathbb{Z}[x]$ without multiple roots. It will:

- Calculate lower bounds for the positive zeroes of $p(x)$ and transformed polynomials.
- Use Möbius substitutions to transform every positve root of $p(x)$ to the only positive root of some $p_{ti}(y)$, and calculate isolation intervals from the transformation formula.

The continued fraction isolation algorithm returns isolation intervals for polynomials $p(x)$ in $\mathbb{Z}[x]$ without multiple roots. It will:

- Calculate lower bounds for the positive zeroes of $p(x)$ and transformed polynomials.
- Use Möbius substitutions to transform every positve root of $p(x)$ to the only positive root of some $p_{ti}(y)$, and calculate isolation intervals from the transformation formula.
- Treat the negative roots in the same way by substituting $p(x) := \pm p(-x)$.

The continued fraction isolation algorithm returns isolation intervals for polynomials $p(x)$ in $\mathbb{Z}[x]$ without multiple roots. It will:

- Calculate lower bounds for the positive zeroes of $p(x)$ and transformed polynomials.
- Use Möbius substitutions to transform every positve root of $p(x)$ to the only positive root of some $p_{ti}(y)$, and calculate isolation intervals from the transformation formula.
- Treat the negative roots in the same way by substituting $p(x) := \pm p(-x)$.

Complexity: $O(n^5 L^3 [|p(x)|_\infty])$.

Now we are left with a single root of $p(x)$ inside an open isolation interval (a, b). To approximate it with a precision of ϵ, we can use the bisection algorithm.

```
while b − a > ϵ do
  if p((a+b)/2) = 0
    return (a+b)/2
  else
    if sgn(p((a+b)/2)) = sgn(p(a))
      a := (a+b)/2
    else
      b := (a+b)/2
    endif
  endif
od return (a, b)
```

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

1. Compute the integer part a of the positive root of $p_M(y)$.

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

1. Compute the integer part a of the positive root of $p_M(y)$.
2. Update $p_M(y) := p_M(y + a)$ and $\mathbf{M}(y) := \mathbf{M}(y + a)$.

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

1. Compute the integer part a of the positive root of $p_M(y)$.
2. Update $p_M(y) := p_M(y + a)$ and $\mathbf{M}(y) := \mathbf{M}(y + a)$.
3. Test, whether $p_M(0) = 0$. Then <u>return</u> $\frac{M_{12}}{M_{22}}$ as exact value for the root.

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

1. Compute the integer part a of the positive root of $p_M(y)$.
2. Update $p_M(y) := p_M(y + a)$ and $\mathbf{M}(y) := \mathbf{M}(y + a)$.
3. Test, whether $p_M(0) = 0$. Then return $\frac{M_{12}}{M_{22}}$ as exact value for the root.
4. Test, whether $|\frac{M_{11}}{M_{21}} - \frac{M_{12}}{M_{22}}| \leq \epsilon$. If so, return $\left(\frac{M_{11}}{M_{21}}, \frac{M_{12}}{M_{22}}\right)$.

If the root was isolated by the continued fraction method, we already know a transformation $x = \mathbf{M}(y)$ and a polynomial $p_M(y)$ which has only one positive root. Then our algorithm looks like this:

1. Compute the integer part a of the positive root of $p_M(y)$.
2. Update $p_M(y) := p_M(y+a)$ and $\mathbf{M}(y) := \mathbf{M}(y+a)$.
3. Test, whether $p_M(0) = 0$. Then <u>return</u> $\frac{M_{12}}{M_{22}}$ as exact value for the root.
4. Test, whether $\left|\frac{M_{11}}{M_{21}} - \frac{M_{12}}{M_{22}}\right| \leq \epsilon$. If so, <u>return</u> $\left(\frac{M_{11}}{M_{21}}, \frac{M_{12}}{M_{22}}\right)$.
5. Set $p_M(y) := p_M(\frac{1}{y})$ and $\mathbf{M}(y) := \mathbf{M}(\frac{1}{y})$, and return to 1.