united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

SMR1537/3

Summer School and Conference on Real Algebraic Geometry and its Applications

(4 - 22 August 2003)

Some Algorithms in Real Algebraic Geometry

Marie-Françoise Roy
IRMAR
Université de Rennes I
Campus de Beaulieu
F-35042 Rennes
France

Some Algorithms in Real Algebraic Geometry

Marie-Françoise Roy

August 2, 2003

Contents

This course is based on the book Algorithms in Real Algebraic Geometry by S. Basu, R. Pollack and M.-F. Roy [1]. Whenever a proof is ommitted in the notes, it can be found in [1]. Note that a quotation to [1] does not mean that the result is new or the proof is original, it only means that it appears there. The bibliography given in these notes is very incomplete and better references appear in [1].

Since a real univariate polynomial does not always have real roots, a very natural algorithmic problem, is to design a method to count the number of real roots of a given polynomial (and thus decide whether it has any). The "real root counting problem" plays a key role in nearly all the "algorithms in real algebraic geometry" and is studied in the first chapter of this course.

The second chapter is devoted to quantifier elimination. The basic geometric objects are the semi-algebraic sets. These are the subsets of $\mathbb{R}^n$ that are defined by a finite number of polynomial equations $(P = 0)$ and inequalities $(P > 0)$. We prove that the projection of a semi-algebraic set is semi-algebraic. The proof is based on a parametric version of real root counting techniques explained in the first chapter. The geometric statement "the projection of a semi-algebraic set is semi-algebraic" yields, after introducing the necessary terminology, the theorem of Tarski that "the theory of real closed fields admits quantifier elimination". A consequence of this last result is the decidability of elementary algebra and geometry, which was Tarski's initial motivation.

Then we describe an algorithm for computing the cylindrical decomposition. The basic idea of this algorithm is to successively eliminate variables. Cylindrical decomposition has numerous applications among which are: deciding the truth of a sentence, and eliminating quantifiers. The huge degree bounds (doubly exponential in the number of variables) output by the cylindrical decomposition method make it desirable to improve these results.

We finallty present an algorithm for the existential theory of the reals whose complexity is singly exponential in the number of variables. Using an algorithm for computing a point in every connected component of an algebraic set and perturbation methods to obtain polynomials in general position, we can compute the realizable sign conditions.

Chapter 1

Real roots

1.1 Descartes's Law of Signs and the Budan-Fourier Theorem

The first result in the direction of real root counting was found more than 350 years ago by Descartes [4].

Notation 1.1.1 The **number of sign changes**, $\mathrm{V}(a)$, in a sequence, $a = a_0, \cdots, a_p$, of elements in $\mathrm{R} \setminus \{0\}$ is defined by induction on p by:

$$\mathrm{V}(a_0) = 0$$

$$\mathrm{V}(a_0, \cdots, a_p) = \begin{cases} \mathrm{V}(a_1, \cdots, a_p) + 1 & \text{if } a_0 a_1 < 0 \\ \\ \mathrm{V}(a_1, \cdots, a_p) & \text{if } a_0 a_1 > 0 \end{cases}$$

This definition extends to any finite sequence a of elements in R by considering the finite sequence b obtained by dropping the zeros in a and defining $\mathrm{V}(a) = \mathrm{V}(b)$, stipulating that V of the empty sequence is 0.

Let $\mathcal{P} = P_0, P_1, \ldots, P_d$ be a sequence of polynomials and let a be an element of $\mathrm{R} \cup \{-\infty, +\infty\}$. The **number of sign changes** of $\mathcal{P}$ at a, denoted by $\mathrm{V}(\mathcal{P}; a)$, is $\mathrm{V}(P_0(a), \ldots, P_d(a))$ (at $-\infty$ and $+\infty$ the signs to consider are the signs of the leading monomials.

Given a and b in $\mathrm{R} \cup \{-\infty, +\infty\}$, we write $\mathrm{V}(\mathcal{P}; a, b)$ for $\mathrm{V}(\mathcal{P}; a) - \mathrm{V}(\mathcal{P}; b)$.

For example $\mathrm{V}(1,-1,2,0,0,3,4,-5,-2,0,3) = 4$. If

$$\mathcal{P} = X^5, X^2-1, 0, X^2-1, X+2, 1,$$

$\mathrm{V}(\mathcal{P};1) = 0$.

Let $P = a_pX^p + \ldots + a_0$ be a univariate polynomial in $\mathrm{R}[X]$. We write $\mathrm{V}(P)$ for the number of sign changes in $a_0, \ldots, a_p$ and $\mathrm{pos}(P)$ for the number of positive real roots of P, counted with multiplicity.

Theorem 1.1.2 (Descartes' law of signs) $\mathrm{pos}(P) \leq \mathrm{V}(P)$,

$\mathrm{V}(P) - \mathrm{pos}(P)$ *is even.*

We will prove the following generalization of Descartes's law of signs due to Budan and Fourier.

Notation 1.1.3 Let P be a univariate polynomial of degree p in $\mathrm{R}[X]$. We denote by $\mathrm{Der}(P)$ the list $P, P', \ldots, P^{(p)}$.

We denote by $n(P;(a,b])$ for the number of roots of P in $(a,b]$ counted with multiplicities.

Theorem 1.1.4 (Budan-Fourier theorem) *Let P be a univariate polynomial of degree p in* $\mathrm{R}[X]$. *Given a and b in* $\mathrm{R} \cup \{-\infty, +\infty\}$

$n(P;(a,b]) \leq \mathrm{V}(\mathrm{Der}(P);a,b)$,

$\mathrm{V}(\mathrm{Der}(P);a,b) - n(P;(a,b])$ *is even.*

Descartes's law of signs is a particular case of Budan-Fourier theorem [2, 6], since the coefficients of the polynomial have the same signs as the derivatives evaluated at 0 and there are no sign changes in the signs of the derivatives at $+\infty$, so that

$$\mathrm{V}(P) = \mathrm{V}(\mathrm{Der}(P);0,+\infty).$$

The following lemma is the key to the proof of Theorem

Lemma 1.1.5 *Let c be a root of P of multiplicity $\mu \geq 0$. If no $P^{(k)}$, $0 \leq k \leq p$, has a root in $[d,c) \cup (c,d']$, then*

$\mathrm{V}(\mathrm{Der}(P); d, c) - \mu$ *is non-negative and even, and*

$\mathrm{V}(\mathrm{Der}(P); c, d') = 0.$

Proof: We prove the claim by induction on the degree of P. The claim is true if the degree of P is 1.
Suppose first that $P(c) = 0$, and hence $\mu > 0$. By induction hypothesis applied to P',

$\mathrm{V}(\mathrm{Der}(P'); d, c) - \mu - 1$ is non-negative and even

$\mathrm{V}(\mathrm{Der}(P'); c, d') = 0.$

The sign of P at the left of c is the opposite of the sign of P' at the left of c and the sign of P at the right of c is the sign of P' at the right of c. Thus

$\mathrm{V}(\mathrm{Der}(P); d) = \mathrm{V}(\mathrm{Der}(P'); d) + 1,$

$\mathrm{V}(\mathrm{Der}(P); c) = \mathrm{V}(\mathrm{Der}(P'); c),$

$\mathrm{V}(\mathrm{Der}(P); d') = \mathrm{V}(\mathrm{Der}(P'); d'),$

and the claim follows.
Suppose now that $P(c) \neq 0$, and hence $\mu = 0$. Let ν be the multiplicity of c as a root of P'. By induction hypothesis applied to P'

$\mathrm{V}(\mathrm{Der}(P'); d, c) - \nu$ is non-negative and even, and

$\mathrm{V}(\mathrm{Der}(P'); c, d') = 0.$

There are four cases to consider.
If ν is odd, and $\mathrm{sign}(P^{\nu+1}(c)P(c)) > 0$,

$\mathrm{V}(\mathrm{Der}(P); d) = \mathrm{V}(\mathrm{Der}(P'); d) + 1,$

$\mathrm{V}(\mathrm{Der}(P); c) = \mathrm{V}(\mathrm{Der}(P'); c),$

$\mathrm{V}(\mathrm{Der}(P); d') = \mathrm{V}(\mathrm{Der}(P'); d').$

If ν is odd, and $\mathrm{sign}(P^{\nu+1}(c)P(c)) < 0$,

$$\mathrm{V}(\mathrm{Der}(P);d) = \mathrm{V}(\mathrm{Der}(P');d),$$

$$\mathrm{V}(\mathrm{Der}(P);c) = \mathrm{V}(\mathrm{Der}(P');c) + 1,$$

$$\mathrm{V}(\mathrm{Der}(P);d') = \mathrm{V}(\mathrm{Der}(P');d') + 1.$$

If ν is even, and $\mathrm{sign}(P^{\nu+1}(c)P(c)) > 0$,

$$\mathrm{V}(\mathrm{Der}(P);d) = \mathrm{V}(\mathrm{Der}(P');d),$$

$$\mathrm{V}(\mathrm{Der}(P);c) = \mathrm{V}(\mathrm{Der}(P');c),$$

$$\mathrm{V}(\mathrm{Der}(P);d') = \mathrm{V}(\mathrm{Der}(P');d').$$

If ν is even, and $\mathrm{sign}(P^{\nu+1}(c)P(c)) < 0$,

$$\mathrm{V}(\mathrm{Der}(P);d) = \mathrm{V}(\mathrm{Der}(P');d) + 1,$$

$$\mathrm{V}(\mathrm{Der}(P);c) = \mathrm{V}(\mathrm{Der}(P');c) + 1,$$

$$\mathrm{V}(\mathrm{Der}(P);d') = \mathrm{V}(\mathrm{Der}(P');d') + 1.$$

The claim is true in each of these four cases. □

Proof of Theorem 1.1.4: It is clear that, for every $c \in (a,b)$,

$$n(P;(a,b]) = n(P;(a,c]) + n(P;(c,b])$$
$$\mathrm{V}(\mathrm{Der}(P);a,b) = \mathrm{V}(\mathrm{Der}(P);a,c) + \mathrm{V}(\mathrm{Der}(P);c,b).$$

Let $c_1 < \cdots < c_r$ be the roots of all the polynomials $P^{(j)}$, $0 \le j \le p-1$, in the interval (a,b) and let $a = c_0, b = c_{r+1}$, $d_i \in (c_i, c_{i+1})$ so that $a = c_0 < d_0 < c_1 < \cdots < c_r < d_r < c_{r+1} = b$.

Since

$$n(P;(a,b]) = \sum_{i=0}^{r} n(P;(c_i,d_i]) + n(P;(d_i,c_{i+1}]),$$

$$\mathrm{V}(\mathrm{Der}(P);a,b) = \sum_{i=0}^{r} \mathrm{V}(\mathrm{Der}(P);c_i,d_i) + \mathrm{V}(\mathrm{Der}(P);d_i,c_{i+1}),$$

the claim follows immediately from Lemma 1.1.5. □

Remark 1.1.6 Let $P \in \mathrm{R}[X]$ be of degree p, and let

$$c_0 < c_1 < \ldots < c_N < c_{N+1},$$

with $c_0 \in \mathrm{R} \cup \{-\infty\}$, $c_{N+1} \in \mathrm{R} \cup \{+\infty\}$. The number of i for which $\mathrm{V}(\mathrm{Der}(P); c_i, c_{i+1})$ is non-zero is bounded by $\mathrm{V}(\mathrm{Der}(P); c_0, c_{N+1})$. Indeed,

$$\sum_{i=0}^{N} \mathrm{V}(\mathrm{Der}(P); c_i, c_{i+1}) = \mathrm{V}(\mathrm{Der}(P); c_0, c_{N+1}) \leq p.$$

There are particular cases where the number of roots on an interval can be obtained using only Theorem 1.1.4:

Exercise 1.1.7 *Prove that*

1. *If* $\mathrm{V}(\mathrm{Der}(P); a, b) = 0$, *then* P *has no root in* $(a, b]$.
2. *If* $\mathrm{V}(\mathrm{Der}(P); a, b) = 1$, *then* P *has exactly one root in* $(a, b]$.

In general it is not possible to conclude much about the number of roots on an interval using only Theorem 1.1.4.

Remark 1.1.8 An important instance, where Descartes's law of signs permits a sharp conclusion is the following. When we know in advance that all the roots of a polynomial are real, i.e. when

$$n(P; (-\infty, +\infty)) = p,$$

$\mathrm{V}(\mathrm{Der}(P); a, b)$ is exactly the number of roots counted with multiplicities in $(a, b]$. Indeed the number $\mathrm{V}(\mathrm{Der}(P); -\infty, +\infty)$, which is always at most p, is here equal to p, hence

$$\begin{aligned} n(P; (-\infty, a]) &\leq \mathrm{V}(\mathrm{Der}(P); -\infty, a) \\ n(P; (a, b]) &\leq \mathrm{V}(\mathrm{Der}(P); a, b) \\ n(P; (b, +\infty)) &\leq \mathrm{V}(\mathrm{Der}(P); b, +\infty) \end{aligned}$$

imply $n(P, (a, b]) = \mathrm{V}(\mathrm{Der}(P); a, b)$.

A last instance where Descartes's law of sign permits a sharp conclusion is the following.

Theorem 1.1.9 *Let*

$$\mathcal{D} = \{(x+iy) \mid x < -\frac{1}{2}, (x+1)^2 + y^2 < 1\}$$

be the part of the disk with center $(-1,0)$ *and radius* 1 *which is to the left of the line* $x = -\frac{1}{2}$ *in* $\mathrm{R}^2 = \mathrm{R}[i]$. *If* $P \in \mathrm{R}[X]$ *has either no roots or exactly one simple root in* $(0,+\infty)$, *and all its complex roots in* $\mathcal{D}$, *then* $\mathrm{V}(P) = 0$ *or* $\mathrm{V}(P) = 1$ *and*

P *has one root in* $(0,+\infty)$ *if and only if* $\mathrm{V}(P) = 1$,

P *has no root in* $(0,+\infty)$ *if and only if* $\mathrm{V}(P) = 0$.

The proof of the theorem relies on the following lemmas.

Lemma 1.1.10 *For* $A, B \in \mathrm{R}[X]$

$$\mathrm{V}(A) = 0, \mathrm{V}(B) = 0 \Rightarrow \mathrm{V}(AB) = 0.$$

Proof: Obvious. □

Lemma 1.1.11 *For* $A, B \in \mathrm{R}[X]$

$$\mathrm{V}(A) = 1, B = X + b, b \geq 0 \Rightarrow \mathrm{V}(AB) = 1.$$

Proof: If $b = 0$, $\mathrm{V}(AB) = \mathrm{V}(A) = 1$. Now, let $b > 0$. Let

$$A = a_d X^d + a_{d-1} X^{d-1} + \ldots + a_0,$$

and suppose, without loss of generality, that $a_d = 1$. Since $\mathrm{V}(A) = 1$ and $a_d = 1$, there exists k such that

$$\begin{cases} a_i \geq 0 & \text{if } i > k, \\ a_k < 0, & \\ a_i \leq 0 & \text{if } i < k. \end{cases}$$

Letting c_i be the coefficient of X^i in AB and making the convention that $a_{d+1} = a_{-1} = 0$, we have

$$\begin{cases} c_i = a_{i-1} + a_i b \geq 0 & \text{if } k+1 < i \leq d, \\ c_k = a_{k-1} + a_k b < 0, & \\ c_i = a_{i-1} + a_i b \leq 0, & \text{if } i < k, \end{cases}$$

and $c_{d+1} = a_d > 0$. So, whatever the sign of c_{k+1}, $\mathrm{V}(AB) = 1$. □

Lemma 1.1.12 *If* $\mathrm{V}(A) = 1, B = X^2 + bX + c$ *with* $b > 1, b > c > 0$, *then* $\mathrm{V}(AB) = 1$.

Proof: Let

$$A = a_d X^d + a_{d-1} X^{d-1} + \ldots + a_0,$$

and suppose without loss of generality that $a_d = 1$. Since $\mathrm{V}(P) = 1$ and $a_d = 1$, there exists k such that

$$\begin{cases} a_i \geq 0, & \text{if } i > k, \\ a_k < 0, & \\ a_i \leq 0, & \text{if } i < k. \end{cases}$$

Letting c_i be the coefficient of X^i in AB and making the convention that $a_{d+2} = a_{d+1} = a_{-1} = a_{-2} = 0$, we have

$$\begin{cases} c_i = a_{i-2} + a_{i-1}b + a_i c \geq 0, & \text{for } k+2 < i \leq d+2 \\ c_k = a_{k-2} + a_{k-1}b + a_k c < 0, & \\ c_i = a_{i-2} + a_{i-1}b + a_i c \leq 0, & \text{for } i < k. \end{cases}$$

The only way to have $\mathrm{V}(AB) > 1$ would be to have $c_{k+1} > 0, c_{k+2} < 0$, but this is impossible since

$$c_{k+2} - c_{k+1} = a_{k+2}c + a_{k+1}(b - c) + a_k(1 - b) - a_{k-1} > 0.$$

□

Proof of Theorem 1.1.9: Notice first that

$\mathrm{V}(P) = 1$ implies P has one root in $(0, +\infty)$ and

$\mathrm{V}(P) = 0$ implies P has no root in $(0, +\infty)$,

using Theorem 1.1.2.

Decompose P into irreducible factors of degree 1 and 2 over R, and note that

if $X + a$ has its root in $(0, +\infty)$, then $a < 0$ and $\mathrm{V}(X + a) = 1$,

if $X + b$ has its root in $(-\infty, 0]$, then $b \geq 0$ and $\mathrm{V}(X + b) = 0$,

if $X^2 + bX + c$ has its roots in $\mathcal{D}$, then $b > 1, b > c > 0$ and $\mathrm{V}(X^2 + bX + c) = 0$.

If P has one root a in $(0, +\infty)$, $\mathrm{V}(X + a) = 1$. Starting from $X + a$ and multiplying successively by the other irreducible factors of P, we get polynomials with sign variations equal to 1, using Lemma 1.1.11 and Lemma 1.1.12. Finally, $\mathrm{V}(P) = 1$.

If P has no root in $(0, +\infty)$, starting from 1 and multiplying successively by the irreducible factors of P, we get polynomials with sign variations equal to 0, using Lemma 1.1.10. Finally, $\mathrm{V}(P) = 0$. □

1.2 Isolating Real Roots

Throughout this section, R is an archimedean real closed field. Let P be a polynomial of degree p in $\mathrm{R}[X]$. We are going to explain how to perform exact computations for determining several properties of the roots of P in R: characterization of a root, sign of another polynomial at a root, and comparisons between roots of two polynomials.

We consider a polynomial

$$P = a_p X^p + \cdots + a_q X^q, p > q, a_q a_p \neq 0,$$

with coefficients in an ordered field K, a real closed field R containing K, and $\mathrm{C} = \mathrm{R}[i]$.

Lemma 1.2.1 (Cauchy) *The absolute value of any root of P in* R *is smaller than*

$$C(P) = \sum_{q \leq i \leq p} \left| \frac{a_i}{a_p} \right|.$$

Proof: Let $x \in \mathrm{R}$ be a root of $P = a_p X^p + \cdots + a_q X^q, p > q$. Then

$$a_p x = - \sum_{q \leq i \leq p-1} a_i x^{i-p+1}.$$

If $|x| \geq 1$ this gives

$$|a_p||x| \leq \sum_{q \leq i \leq p-1} |a_i||x|^{i-p+1} \leq \sum_{q \leq i \leq p-1} |a_i|.$$

Thus it is clear that $|x| \leq C(P)$.

If $|x| \leq 1$, we have $|x| \leq 1 \leq C(P)$, since $C(P) \geq 1$. □

The characterization of the roots of P in R will be performed by finding intervals with rational end points. Our method will be based on Descartes's law of signs (Theorem 1.1.2) and the properties of the Bernstein basis defined below.

Notation 1.2.2 Let P be a polynomial of degree $\leq p$. The **Bernstein polynomials** of degree p for c, d are the

$$B_{p,i}(c,d) = \binom{p}{i} \frac{(X-c)^i (d-X)^{p-i}}{(d-c)^p},$$

for $i = 0, \ldots, p$.

Remark 1.2.3 Note that $B_{p,i}(c,d) = B_{p,p-i}(d,c)$ and that

$$B_{p,i}(c,d) = \frac{(X-c)}{d-c}\frac{p}{i} B_{p-1,i-1}(c,d) = \frac{(d-X)}{d-c}\frac{p}{p-i} B_{p-1,i}(c,d).$$

In order to prove that the Bernstein polynomials form a basis of polynomials of degree $\leq p$, we are going to need three simple transformations of P.

Reciprocal polynomial in degree p: $\mathrm{Rec}_p(P(X)) = X^p P(1/X)$. The non-zero roots of P are the inverses of the non-zero roots of $\mathrm{Rec}(P)$.

Contraction by ratio λ: for every non-zero λ, $\mathrm{C}_\lambda(P(X)) = P(\lambda X)$. The roots of $\mathrm{C}_\lambda(P)$ are of the form $\frac{x}{\lambda}$, where x is a root of P.

Translation by c: for every c, $\mathrm{T}_c(P(X)) = P(X-c)$. The roots of $\mathrm{T}_c(P(X))$ are of the form $x + c$ where x is a root of P.

These three transformations clearly define bijections from the set of polynomials of degree at most p into itself.

Proposition 1.2.4 *Let* $P = \sum_{i=0}^{p} b_i B_{p,i}(d,c) \in \mathrm{R}[X]$ *be of degree* $\leq p$. *Let*

$$\mathrm{T}_{-1}\left(\mathrm{Rec}_p(\mathrm{C}_{d-c}(\mathrm{T}_{-c}(P)))\right) = \sum_{i=0}^{p} c_i X^i.$$

Then

$$\binom{p}{i} b_i = c_{p-i}.$$

Proof: Performing the contraction of ratio $d-c$ after translating by $-c$ transforms $\binom{p}{i}\dfrac{(X-c)^i(d-X)^{p-i}}{(d-c)^p}$ into $\binom{p}{i}X^i(1-X)^{p-i}$. Translating by -1 after taking the reciprocal polynomial in degree p transforms $\binom{p}{i}X^i(1-X)^{p-i}$ into $\binom{p}{i}X^{p-i}$. □

Corollary 1.2.5 *The Bernstein polynomials for* c, d *form a basis of the vector space of polynomials of degree* $\leq p$.

We denote as usual by $\mathrm{V}(b)$ the number of sign changes in a list b.

Proposition 1.2.6 *Let* P *be of degree* p. *We denote by* $b = b_0, \ldots, b_p$ *the coefficients of* P *in the Bernstein basis of* c, d. *Let* $n(P;(c,d))$ *be the number of roots of* P *in* (c,d) *counted with multiplicities. Then*

$\mathrm{V}(b) \geq n(P;(c,d))$,

$\mathrm{V}(b) - n(P;(c,d))$ *is even.*

Proof: The claim follows immediately from Descartes's law of signs (Theorem 1.1.2), using Proposition 1.2.4. Indeed, the image of (c,d) under the translation by $-c$ followed by the contraction of ratio $d-c$ is $(0,1)$. The image of $(0,1)$ under the inversion $z \mapsto 1/z$ is $(1,+\infty)$. Finally, translating by -1 gives $(0,+\infty)$. □

The coefficients $b = b_0, \ldots, b_p$ of P in the Bernstein basis of c, d give a rough idea of the shape of the polynomial P on the interval c, d. The

control line of P **on** $[c,d]$ is the union of the segments $[M_i, M_{i+1}]$ for $i = 0, \ldots, p-1$, with

$$M_i = \left(\frac{id + (p-i)c}{p}, b_i \right).$$

It is clear from the definitions that the graph of P goes through M_0 and M_p and that the line M_0, M_1 (resp M_{p-1}, M_p) is tangent to the graph of P at M_0 (respectively M_p).

The **control polygon of** P **on** $[c,d]$ is the convex hull of the points M_i for $i = 0, \ldots, p$.

Proposition 1.2.7 *The graph of P on $[c,d]$ is contained in the control polygon of P on $[c,d]$.*

Proof: In order to prove the proposition, it is enough to prove that any line L above (respectively under) all the points in the control polygon of P on $[c,d]$ is above (respectively under) the graph of P on $[c,d]$. If L is defined by $Y = aX + b$, let us express the polynomial $aX + b$ in the Bernstein basis. Since

$$1 = \left(\frac{X-c}{d-c} + \frac{d-X}{d-c} \right)^p,$$

the binomial formula gives

$$1 = \sum_{i=0}^{p} \binom{p}{i} \left(\frac{X-c}{d-c} \right)^i \left(\frac{d-X}{d-c} \right)^{p-i} = \sum_{i=0}^{p} B_{p,i}(c,d).$$

Since

$$X = \left(d \left(\frac{X-c}{d-c} \right) + c \left(\frac{d-X}{d-c} \right) \right) \left(\frac{X-c}{d-c} + \frac{d-X}{d-c} \right)^{p-1},$$

the binomial formula together with Remark 1.2.3 gives

$$\begin{aligned} X &= \sum_{i=0}^{p-1} \left(d \left(\frac{X-c}{d-c} \right) + c \left(\frac{d-X}{d-c} \right) \right) B_{p-1,i}(c,d) \\ &= \sum_{i=0}^{p} \left(\frac{id + (p-i)c}{p} \right) B_{p,i}(c,d). \end{aligned}$$

Thus,

$$aX + b = \sum_{i=0}^{p} \left(a \left(\frac{id + (p-i)c}{p} \right) + b \right) B_{p,i}(c,d).$$

It follows immediately that if L is above every M_i, i.e. if

$$a \left(\frac{id + (p-i)c}{p} \right) + b \geq b_i$$

for every i, then L is above the graph of P on $[c,d]$, since $P = \sum_{i=0}^{p} b_i B_{p,i}(c,d)$ and the Bernstein basis of c,d is non-negative on $[c,d]$. A similar argument holds for L under every M_i. □

The following remarkable algorithm due to De Casteljau [5] computes the coefficients of P in the Bernstein bases of c,e and e,d from the coefficients of P in the Bernstein basis of c,d.

Algorithm 1.2.8 (Bernstein Coefficients)

Input: *a list $b = b_0, \ldots, b_p$ representing a polynomial P of degree $\leq p$ in the Bernstein basis of c,d, and a number $e \in \mathrm{R}$.*

Output: *the list $b' = b'_0, \ldots, b'_p$ representing P in the Bernstein basis of c,e and the list $b'' = b''_0, \ldots, b''_p$ representing P in the Bernstein basis of e,d.*

Procedure:

Define $\alpha = \dfrac{d-e}{d-c}, \beta = \dfrac{e-c}{d-c}$.

Initialization: $b_j^{(0)} := b_j$, $j = 0, \ldots, p$.

For $i = 1, \ldots, p$,

For $j = 0, \ldots, p-i$, *compute*

$$b_j^{(i)} := \alpha b_j^{(i-1)} + \beta b_{j+1}^{(i-1)}$$

Define

$$b^{(p)} = b_0^{(0)}, \ldots, b_0^{(j)}, \ldots, b_0^{(p)}, \ldots, b_{p-j}^{(j)}, \ldots, b_p^{(0)},$$

and output

$$b' = b_0^{(0)}, \ldots, b_0^{(j)}, \ldots, b_0^{(p)}$$

and

$$b'' = b_0^{(p)}, \ldots, b_j^{(p-j)}, \ldots, b_p^{(0)}.$$

Algorithm 1.2.8 (Bernstein Coefficients) can be visualized with the following triangle.

$$\begin{array}{ccccccccccccc}
b_0^{(0)} & & \cdots & & \cdots & & \cdots & & \cdots & & & & b_p^{(0)} \\
& b_0^{(1)} & & \cdots & & \cdots & & \cdots & & & b_{p-1}^{(1)} & & \\
& & \cdots & & \cdots & & \cdots & & \cdots & & & & \\
& & & b_0^{(i)} & & \cdots & & b_{p-i}^{(i)} & & & & & \\
& & & & b_0^{(p-1)} & & b_1^{(p)} & & & & & & \\
& & & & & b_0^{(p)} & & & & & & &
\end{array}$$

with $b_j^{(i)} := \alpha b_j^{(i-1)} + \beta b_{j+1}^{(i-1)}$, $\alpha = \dfrac{d-e}{d-c}, \beta = \dfrac{e-c}{d-c}$.

The coefficients of P in the Bernstein basis of c, d appear in the top side of the triangle and the coefficients of P in the Bernstein basis of c, e and e, d appear in the two other sides of the triangle.

Notation 1.2.9 We denote by $\tilde{a}$ the list obtained by reversing the list a.

Proof of correctness: It is enough to prove the part of the claim concerning c, e. Indeed, by Remark 1.2.3, $\tilde{b}$ represents P in the Bernstein basis of d, c, and the claim is obtained by applying Algorithm 1.2.8 (Bernstein Coefficients) to $\tilde{b}$ at e. The output is $\tilde{b''}$ and $\tilde{b}$ and the conclusion follows using again Remark 1.2.3.

Let $\delta_{p,i}$ be the list of length $p+1$ consisting all zeroes except a 1 at the $i+1$-th place. Note that $\delta_{p,i}$ is the list of coefficients of $B_{p,i}(c, d)$ in the Bernstein basis of c, d. We will prove that the coefficients of $B_{p,i}(c, d)$ in the Bernstein basis of c, e coincide with the result of Algorithm 1.2.8 (Bernstein Coefficients) performed with input $\delta_{p,i}$. The correctness of Algorithm 1.2.8 (Bernstein Coefficients) for c, e then follows by linearity.

First notice that, since $\alpha = \dfrac{d-e}{d-c}, \beta = \dfrac{e-c}{d-c}$,

$$\frac{X-c}{d-c} = \beta \frac{X-c}{e-c},$$
$$\frac{d-X}{d-c} = \alpha \frac{X-c}{e-c} + \frac{e-X}{e-c}.$$

Thus

$$\left(\frac{X-c}{e-c}\right)^i = \beta^i \left(\frac{X-c}{d-c}\right)^i,$$
$$\left(\frac{d-X}{d-c}\right)^{p-i} = \sum_{k=0}^{p-i} \binom{p-i}{k} \alpha^k \left(\frac{X-c}{e-c}\right)^k \left(\frac{e-X}{e-c}\right)^{p-i-k}.$$

It follows that

$$B_{p,i}(c,d) = \binom{p}{i} \sum_{j=i}^{p} \binom{p-i}{j-i} \alpha^{j-i} \beta^i \left(\frac{X-c}{e-c}\right)^j \left(\frac{e-X}{e-c}\right)^{p-j}.$$

Since

$$\binom{p}{i}\binom{p-i}{j-i} = \binom{j}{i}\binom{p}{j},$$
$$B_{p,i}(c,d) = \sum_{j=i}^{p} \binom{j}{i} \alpha^{j-i} \beta^i \binom{p}{j} \left(\frac{X-c}{e-c}\right)^j \left(\frac{e-X}{e-c}\right)^{p-j}.$$

Finally,

$$B_{p,i}(c,d) = \sum_{j=i}^{p} \binom{j}{i} \alpha^{j-i} \beta^i B_{p,j}(c,e).$$

On the other hand, we prove by induction on p that Algorithm 1.2.8 (Bernstein Coefficients) with input $\delta_{p,i}$ outputs the list $\delta'_{p,i}$ starting with i zeroes and with $(j+1)$-th element $\binom{j}{i}\alpha^{j-i}\beta^i$ for $j = i, \ldots, p$.

The result is clear for $p = i = 0$. If Algorithm 1.2.8 (Bernstein Coefficients) applied to $\delta_{p-1,i-1}$ outputs $\delta'_{p-1,i-1}$, the equality

$$\binom{j}{i}\alpha^{j-i}\beta^i = \alpha\binom{j-1}{i}\alpha^{j-i-1}\beta^i + \beta\binom{j-1}{i-1}\alpha^{j-i}\beta^{i-1}$$

proves by induction on j that Algorithm 1.2.8 (Bernstein Coefficients) applied to $\delta_{p,i}$ outputs $\delta'_{p,i}$. So the coefficients of $B_{p,i}(c,d)$ in the Bernstein basis of e,d coincide with the output of Algorithm 1.2.8 (Bernstein Coefficients) with input $\delta_{p,i}$.

□

Algorithm 1.2.8 (Bernstein Coefficients) works both ways.

Corollary 1.2.10 *Let b, b' and b'' be the lists of coefficients of P in the Bernstein basis of c,d, c,e, and e,d respectively.*

Algorithm 1.2.8 (Bernstein Coefficients) applied to b with weights

$$\alpha = \frac{d-e}{d-c}, \beta = \frac{e-c}{d-c}$$

outputs b' and b''.

Algorithm 1.2.8 (Bernstein Coefficients) applied to b' with weights

$$\alpha' = \frac{e-d}{e-c}, \beta' = \frac{d-c}{e-c}$$

outputs b and $\tilde{b}''$.

Algorithm 1.2.8 (Bernstein Coefficients) applied to b'' with weights

$$\alpha'' = \frac{d-c}{d-e}, \beta'' = \frac{c-e}{d-e}$$

outputs $\tilde{b}'$ and b.

Complexity analysis of Algorithm 1.2.8: The number of multiplications in the algorithm is $2\frac{p(p+1)}{2}$, the number of additions is $\frac{p(p+1)}{2}$.
□

Algorithm 1.2.8 (Bernstein Coefficients) gives a geometric construction of the control polygon of P on $[c,e]$ and on $[e,d]$ from the control polygon of P on $[c,d]$. The points of the new control polygons are constructed by taking iterated barycenters with weights α and β.

Example 1.2.11 We take $p = 3$, and consider the polynomial P with coefficients $4, -6, 7, 10$ in the Bernstein basis for $0, 1$

$$(1-X)^3, 3X(1-X)^2, 3X^2(1-X), X^3.$$

Algorithm 1.2.8 (Bernstein Coefficients) gives the following results.

$$\begin{array}{ccccccc} 4 & & -6 & & 7 & & 10 \\ & -1 & & 1/2 & & 17/2 & \\ & & -1/4 & & 9/2 & & \\ & & & 17/8 & & & \end{array}$$

We denote as usual by $\mathrm{V}(b)$ the number of sign changes in a list b.

Proposition 1.2.12 *Let b, b' and b'' be the lists of coefficients of P in the Bernstein basis of c, d, c, e, and e, d. If $c < e < d$,then*

$$\mathrm{V}(b') + \mathrm{V}(b'') \leq \mathrm{V}(b).$$

Moreover $\mathrm{V}(b) - \mathrm{V}(b') - \mathrm{V}(b'')$ is even.

Proof: The proof of the proposition is based on the following easy observations:

Inserting in a list $a = a_0, \ldots, a_n$ a value x in $[a_i, a_{i+1}]$ if $a_{i+1} \geq a_i$ (respectively in $[a_{i+1}, a_i]$ if $a_{i+1} < a_i$) between a_i and a_{i+1} does not modify the number of sign variations.

Removing from a list $a = a_0, \ldots, a_n$ with first non-zero $a_k, k \geq 0$, and last non-zero $a_\ell, k \leq \ell \leq n$, an element a_i, $i \neq k, i \neq \ell$ decreases the number of sign variation by an even (possibly zero) natural number.

Indeed the lists

$$\begin{array}{c} b = b_0^{(0)}, \ldots, \ldots, \ldots, \ldots, \ldots, b_p^{(0)} \\ b^{(1)} = b_0^{(0)}, b_0^{(1)}, \ldots, \ldots, \ldots, \ldots, b_{p-1}^{(1)}, b_p^{(0)} \\ \ldots \\ b^{(i)} = b_0^{(0)}, \ldots, \ldots, b_0^{(i)}, \ldots, \ldots, b_{p-i}^{(i)}, \ldots, \ldots, b_p^{(0)} \\ \ldots \\ b^{(p-1)} = b_0^{(0)}, \ldots, \ldots, \ldots, b_0^{(p-1)}, b_1^{(p)}, \ldots, \ldots, \ldots, b_p^{(0)} \\ b^{(p)} = b_0^{(0)}, \ldots, \ldots, \ldots, \quad b_0^{(p)} \quad , \ldots, \ldots, \ldots, \ldots, b_p^{(0)} \end{array}$$

are successively obtained by inserting intermediate values and removing elements that are not end points, since when $c < e < d$, $b_j^{(i)}$ is between $b_j^{(i-1)}$ and $b_{j+1}^{(i-1)}$, for $i = 1, \ldots, p, j = 0, \ldots, p-i-1$. Thus $\mathrm{V}(b^{(p)}) \leq \mathrm{V}(b)$ and the difference is even. It is clear that $\mathrm{V}(b^{(p)}) = \mathrm{V}(b') + \mathrm{V}(b'')$.

□

Example 1.2.13 Continuing Example 1.2.11, we observe, denoting by b, b' and b'', the lists of coefficients of P in the Bernstein basis of $0, 1$, $0, 1/2$, and $1/2, 1$, that $\mathrm{V}(b) = 2$. This is visible on the figure: the control line for $[0, 1]$ cuts twice the X-axis. Similarly, $\mathrm{V}(b') = 2$. This is visible on the figure: the control line for $[0, 1/2]$ also cuts twice the X-axis. Similarly, it is easy to check that $\mathrm{V}(b'') = 0$.

We cannot decide from this information whether P has two roots on $[0, 1/2]$ or no root on $[0, 1/2]$.

Let $b(P, c, d)$ be the list of coefficients of P in the Bernstein basis of c, d, $d > c$. The interval (c, d) is **active** if $\mathrm{V}(b(P, c, d)) > 0$.

Remark 1.2.14 It is clear from Proposition 1.2.12 that if

$$c_0 < \ldots < c_N,$$

the number of active intervals among (c_i, c_{i+1}) is at most p.

Let $P \in \mathrm{R}[X]$ and let b be the list of coefficients of P in the Bernstein basis of c, d. We now describe a special case where the number $\mathrm{V}(b)$ coincides with the number of roots of P on (c, d). Let $d > c$, let $\mathcal{C}(c, d)_0$ be the closed disk with center $(c, 0)$ and radius $d - c$, and let $\mathcal{C}(c, d)_1$ be the closed disk with center $(d, 0)$ and radius $d - c$.

Theorem 1.2.15 (Theorem of two circles) *If P has either no root or exactly one simple root in (c, d) and P has no complex root in $\mathcal{C}(c, d)_0 \cup \mathcal{C}(c, d)_1$, then*

P has one root in (c, d) if and only if $\mathrm{V}(b) = 1$,

P has no root in (c, d) if and only if $\mathrm{V}(b) = 0$.

Proof: We identify R^2 with $\mathrm{C} = \mathrm{R}[i]$. The image of the complement of $\mathcal{C}(c,d)_0$ (resp $\mathcal{C}(c,d)_1$) under the translation by $-c$ followed by the contraction by ratio $d-c$ is the complement of $\mathcal{C}(0,1)_0$ (resp $\mathcal{C}(0,1)_1$). The image of the complement of $\mathcal{C}(0,1)_0$ under the inversion $z \mapsto 1/z$ is

$$\{(x+iy) \in \mathrm{R}[i] \mid 0 < x^2 + y^2 < 1\}.$$

The image of the complement of $\mathcal{C}(0,1)_1$ under the inversion $z \mapsto 1/z$ is

$$\{(x+iy) \in \mathrm{R}[i] \mid x < \frac{1}{2}\}.$$

The image of the complement of $\mathcal{C}(0,1)_0 \cup \mathcal{C}(0,1)_1$ under the inversion $z \mapsto 1/z$ is

$$\{(x+iy) \in \mathrm{R}[i] \mid 0 < x^2 + y^2 < 1, x < \frac{1}{2}\}.$$

Translating this region by -1, we get the region

$$\mathcal{D} = \{(x+iy) \mid x < -\frac{1}{2}, (x+1)^2 + y^2 < 1\}$$

defined in Theorem 1.1.9.

The statement then follows from Theorem 1.1.9 and Proposition 1.2.4. □

Suppose that $P \in \mathrm{R}[X]$ is a polynomial of degree p with all its real zeroes in $(-2^\ell, 2^\ell)$ and let $\bar{P}$ be the squarefree part of P. Consider natural numbers k and c such that $0 \leq c \leq 2^k$ and define

$$a_{c,k} = \frac{-2^{\ell+k} + c2^{\ell+1}}{2^k}.$$

It is clear that, for k big enough, the polynomial $\bar{P}$ has at most one root in $(a_{c,k}, a_{c+1,k})$ and has no other complex root in $\mathcal{C}(a_{c,k}, a_{c+1,k})_0 \cup \mathcal{C}(a_{c,k}, a_{c+1,k})_1$.

Let $b(\bar{P}, c, k)$ denote the list of coefficients of $\bar{P}$ in the Bernstein basis of $(a_{c,k}, a_{c+1,k})$ and $\mathrm{V}(b(\bar{P}, c, k))$ its sign variations.

Using Theorem 1.2.15, it is possible to decide, for k big enough, whether $\bar{P}$ has exactly one root in $(a_{c,k}, a_{c+1,k})$ or has no root on $(a_{c,k}, a_{c+1,k})$ by testing whether $\mathrm{V}(b(\bar{P}, c, k)$ is zero or one.

An **isolating list for** P is a finite list L of rational points and closed intervals with rational end points of R such that each point or interval of L contains exactly one root of P in R and every root of P in R belongs to an element of L.

Example 1.2.16 Continuing Example 1.2.13, let us study the roots of P on $[0,1]$, as a preparation to a more formal description of Algorithm Algorithm 1.2.17 (Real Root Isolation).

The Bernstein coefficients of P on $[0,1]$ are $4, -6, 7, 10$. There maybe roots of P on $[0,1]$ as there are sign variations in its Bernstein coefficients.

As already seen in Example 1.2.13,a first application of Algorithm 1.2.8 (Bernstein Coefficients) gives

4 -6 7 10
-1 1/2 17/2
-1/4 9/2
17/8

There maybe roots of P on $[0,1/2]$ as there are as there are sign variations in the Bernstein coefficients of $8P$ which are $32, -8, -2, 17$. There are no roots of P on $[1/2,1]$.

Let us apply once more Algorithm 1.2.8 (Bernstein Coefficients):

4 -1 -1/4 17/8
3/2 -5/8 15/16
7/16 5/32
19/64

There are no sign variations on the sides of the triangle so there are no roots of P on $[0,1/4]$ and on $[1/4,1/2]$.

Algorithm 1.2.17 (Real Root Isolation)

Input: *a square free non-zero polynomial* $P \in \mathrm{R}[X]$.

Output: *a list* $L(P)$ *isolating for* P.

Procedure:

Compute ℓ *such that* $(-2^\ell, 2^\ell)$ *contains the roots of* P *in* R *using Lemma 1.2.1.*

Compute $b(P,0,0)$, the Bernstein coefficients of P, using Proposition 1.2.4 on $(-2^\ell, 2^\ell)$.

Initialization: $Pos := \{(b(P,0,0)\}$ and $L(P)$ is the empty list.

While Pos is non-empty,

Remove $b(P,c,k)$ from Pos.

If $\mathrm{V}(b(P,c,k)) = 1$ and $P(a_{c,k})P(a_{c+1,k}) \neq 0$, add $[a_{c,k}, a_{c+1,k}]$ to $L(P)$.

If $\mathrm{V}(b(P,c,k)) = 0$ do nothing.

If $\mathrm{V}(b(P,c,k)) > 1$ or $P(a_{c,k})P(a_{c+1,k}) = 0$, compute $b(P,2c,k+1)$ and $b(P,2c+1,k+1)$ using Algorithm 1.2.8 Bernstein Coefficients) and add them to Pos. If $P(a_{2c+1,k+1}) = 0$, add $\{a_{2c+1,k+1}\}$ to $L(P)$.

Proof of correctness: [9] The algorithm terminates since R is archimedean. Its correctness follows from Theorem 1.2.15. Note that, since there is only one root of P on each interval $[a,b]$ of $L(P)$, $P(a)P(b) < 0$. □

Remark 1.2.18 It is clear that Algorithm 1.2.17 (Real Root Isolation) also provides a method for counting real roots.

1.3 Sturm and Sylvester's theorems

Let P be a non-zero polynomial with coefficients in a real closed field R. Not only would we like to determine whether P has a root in R but also to determine whether P has a root at which another polynomial Q is positive.

With this goal in mind, it is profitable to look at the jumps (discontinuities) of the rational function $\frac{P'Q}{P}$. Clearly, these occur only at points c for which $P(c) = 0, Q(c) \neq 0$. If c occurs as a root of P with multiplicity μ then $\frac{P'Q}{P} = \frac{\mu Q(c)}{X-c} + R_c$, where R_c is a rational function defined at c. It is now obvious that if $Q(c) > 0$, then $\frac{P'Q}{P}$ jumps from

$-\infty$ to $+\infty$ at c, and if $Q(c) < 0$, then $\frac{P'Q}{P}$ jumps from $+\infty$ to $-\infty$ at c. Thus the number of jumps of $\frac{P'Q}{P}$ from $-\infty$ to $+\infty$ minus the number of jumps of $\frac{P'Q}{P}$ from $+\infty$ to $-\infty$ is equal to the number of roots of P at which Q is positive minus the number of roots of P at which Q is negative. This observation leads us to the following definition. We need first what we mean by a jump from $-\infty$ to $+\infty$.

Notation 1.3.1 Let x be a root of P. The function $\frac{Q}{P}$ **jumps from** $-\infty$ **to** $+\infty$ **at** x if the multiplicity μ of x as a root of P is bigger than the multiplicity ν of x as a root of Q, $\mu - \nu$ is odd and the sign of $\frac{Q}{P}$ at the right of x is positive. Similarly, the function $\frac{Q}{P}$ **jumps from** $+\infty$ **to** $-\infty$ **at** x if if the multiplicity μ of x as a root of P is bigger than the multiplicity ν of x as a root of Q, $\mu - \nu$ is odd and the sign of $\frac{Q}{P}$ at the right of x is negative.

Given $a < b$ in $\mathrm{R} \cup \{-\infty, +\infty\}$ and $P, Q \in \mathrm{R}[X]$, we define the **Cauchy index** of $\frac{Q}{P}$ on (a, b), $\mathrm{Ind}\left(\frac{Q}{P}; a, b\right)$, to be the number of jumps of the function $\frac{Q}{P}$ from $-\infty$ to $+\infty$ minus the number of jumps of the function $\frac{Q}{P}$ from $+\infty$ to $-\infty$ on the open interval (a, b). The **Cauchy index** of $\frac{Q}{P}$ on R is simply called the **Cauchy index** of $\frac{Q}{P}$ and it is denoted by $\mathrm{Ind}\left(\frac{Q}{P}\right)$, rather than by $\mathrm{Ind}\left(\frac{Q}{P}; -\infty, +\infty\right)$.

Remark 1.3.2

a) Suppose that $\deg(P) = p$ and $\deg(Q) = q < p$. The Cauchy index $\mathrm{Ind}\left(\frac{Q}{P}; a, b\right)$ is equal to p if and only if $q = p - 1$, the signs of the leading coefficients of P and Q are equal, all the roots of P and Q are simple and belong to (a, b), and there is exactly a root of Q between two roots of P.

b) If $R = \mathrm{Rem}(Q, P)$, it follows clearly from the definition that

$$\mathrm{Ind}\left(\frac{Q}{P}; a, b\right) = \mathrm{Ind}\left(\frac{R}{P}; a, b\right).$$

With this definition we can reformulate our observation, using the following notation.

Notation 1.3.3 The **Sturm-query** of Q for P in (a, b) is the number

$$\begin{gathered}\mathrm{SQ}(Q, P; a, b) = \\ \#(\{x \in (a, b) \mid P(x) = 0 \wedge Q(x) > 0\}) - \\ \#(\{x \in (a, b) \mid P(x) = 0 \wedge Q(x) < 0\}),\end{gathered}$$

where $\#(S)$ is the number of elements in the finite set S.

The Sturm-query of Q for P on R is simply called the **Sturm-query** of Q for P, and is denoted by $\mathrm{SQ}(Q, P)$, rather than by $\mathrm{SQ}(Q, P; -\infty, +\infty)$.

The preceding discussion implies:

Proposition 1.3.4

$$\mathrm{SQ}(Q, P; a, b) = \mathrm{Ind}\left(\frac{P'Q}{P}; a, b\right).$$

In particular the number of roots of P *in* (a, b) *is* $\mathrm{Ind}\left(\dfrac{P'}{P}; a, b\right)$.

We now describe how to compute $\mathrm{Ind}\left(\dfrac{Q}{P}; a, b\right)$. We will see that the Cauchy index is the difference in the number of sign changes (Definition page 3) in the signed remainder sequence $\mathrm{S}(P, Q)$ evaluated at a and b, defined above.

If $Q \neq 0$, the **remainder** in the **euclidean division of** P **by** Q, denoted $\mathrm{Rem}(P, Q)$, is the unique polynomial $R \in \mathrm{K}[X]$ of degree smaller than the degree of Q such that $P = AQ + R$ with $A \in \mathrm{K}[X]$. The **quotient** in the euclidean division of P by Q, denoted $\mathrm{Quo}(P, Q)$, is A.

Algorithm 1.3.5 (Euclidean Division)

Input: *two univariate polynomials* $P = a_p X^p + \cdots + a_0$ *and* $Q = b_q X^q + \cdots + b_0$ *in* $\mathrm{K}[X]$ *with* $b_q \neq 0$.

Output: $\mathrm{Quo}(P,Q)$ *and* $\mathrm{Rem}(P,Q)$, *the quotient and remainder in the Euclidean division of* P *by* Q.

Procedure: *Initialization:* $C := 0$, $R := P$.

For every j *from* p *to* q,

$$C := C + \frac{\mathrm{cof}_j(R)}{b_q} X^{j-q},$$
$$R := R - \frac{\mathrm{cof}_j(R)}{b_q} X^{j-q} Q.$$

Output C, R.

A **greatest common divisor of** P **and** Q, denoted $\gcd(P,Q)$, is a polynomial $G \in \mathrm{K}[X]$ such that G is a divisor of both P and Q, and any divisor of both P and Q is a divisor of G.

We now prove that greatest common divisors exist by using euclidean division repeatedly. Given $P, Q \in K[X]$, not both 0, we define the **signed remainder sequence** $\mathrm{S}(P,Q) = \mathrm{S}_0, \mathrm{S}_1, \ldots, \mathrm{S}_k$ **of** P **and** Q by,

$$\begin{gathered}
\mathrm{S}_0 = P, \\
\mathrm{S}_1 = Q, \\
\mathrm{S}_2 = -\mathrm{S}_0 + A_0 \mathrm{S}_1, A_0 = \mathrm{Quo}(\mathrm{S}_0, \mathrm{S}_1), \\
\mathrm{S}_3 = -\mathrm{S}_1 + A_1 \mathrm{S}_2, A_1 = \mathrm{Quo}(\mathrm{S}_1, \mathrm{S}_2), \\
\vdots \\
\mathrm{S}_k = -\mathrm{S}_{k-2} + A_{k-2} \mathrm{S}_{k-1}, A_{k-2} = \mathrm{Quo}(\mathrm{S}_{k-2}, \mathrm{S}_{k-1}), \\
\mathrm{S}_{k+1} = -\mathrm{S}_{k-1} + A_{k-1} \mathrm{S}_k, A_{k-1} = \mathrm{Quo}(\mathrm{S}_{k-1}, \mathrm{S}_k), \\
\mathrm{S}_k \neq 0, \mathrm{S}_{k+1} = 0.
\end{gathered}$$

In the above, each S_i is the negative of the remainder in the euclidean division of S_{i-2} by S_{i-1} for $2 \leq i \leq k+1$, and the sequence ends when $S_{k+1} = 0$, for $k \geq 0$. We claim that S_k is a greatest common divisor of P and Q. Observe that if a polynomial A divides two polynomials B, C then it also divides $UB + VC$ for arbitrary polynomials U, V. Since $S_{k+1} = -\mathrm{Rem}(S_{k-1}, S_k) = 0$, S_k divides S_{k-1} and since, $S_{k-2} = -S_k + A_{k-2}S_{k-1}$, S_k divides S_{k-2} using the above observation. Continuing this process one obtains that S_k divides $S_0 = P$. Also, if any polynomial divides S_0, S_1 (that is P, Q) then it divides S_2 and hence S_3 and so on. Hence, it divides S_k, proving that S_k is a greatest common divisor of P, Q.

Note that the signed remainder sequence of P and 0 is P and when Q is not 0, the signed remainder sequence of 0 and Q is $0, Q, 0$.

Also, note that by unwinding the definitions of the S_i's, we can express $S_k = \gcd(P, Q)$ as $UP + VQ$ for some polynomials U, V in $\mathrm{K}[X]$. We prove bounds on the degrees of U, V by elucidating the preceding remark.

Proposition 1.3.6 *If G is a greatest common divisor of P and Q, then there exist U and V with*

$$UP + VQ = G.$$

Moreover, if $G \neq Q$ and $\deg(G) = g$, U and V can be chosen so that $\deg(U) < q - g$, $\deg(V) < p - g$.

Theorem 1.3.7 *Let P, $P \neq 0$, and Q be two polynomials with coefficients in a real closed field R, and let a and b (with $a < b$) be elements of $\mathrm{R} \cup \{-\infty, +\infty\}$ that are not roots of P. Then*

$$\mathrm{V}(\mathrm{S}(P, Q); a, b) = \mathrm{Ind}\left(\frac{Q}{P}; a, b\right).$$

Proof: We can assume without loss of generality that a and b are not roots of a polynomial in the signed remainder sequence. Indeed if $a < a' < b' < b$ with $(a, a']$ and $[b', b)$ containing no root of the polynomials in the signed remainder sequence, it is clear that

$$\mathrm{Ind}\left(\frac{Q}{P}; a, b\right) = \mathrm{Ind}\left(\frac{Q}{P}; a', b'\right).$$

We prove now that

$$\mathrm{V}(\mathrm{S}(P,Q);a,b) = \mathrm{V}(\mathrm{S}(P,Q);a',b').$$

First notice that since a is not a root of P, a is not a root of the greatest common divisor of P and Q, and hence a is not simultaneously a root of S_j and S_{j+1} (respectively S_{j-1} and S_j). So, if a is a root of of S_j, $j \neq 0$, $\mathrm{S}_{j-1}(a)\mathrm{S}_{j+1}(a) < 0$, since

$$\mathrm{S}_{j+1} = -\mathrm{S}_{j-1} + \mathrm{Quo}(\mathrm{S}_j, \mathrm{S}_{j-1}) \cdot \mathrm{S}_j$$

so that

$$\mathrm{V}(\mathrm{S}_{j-1}, \mathrm{S}_j, \mathrm{S}_{j+1}; a) = \mathrm{V}(\mathrm{S}_{j-1}, \mathrm{S}_j, \mathrm{S}_{j+1}; a') = 1.$$

This implies $\mathrm{V}(\mathrm{S}(P,Q);a) = \mathrm{V}(\mathrm{S}(P,Q);a')$, and similarly $\mathrm{V}(\mathrm{S}(P,Q);b) = \mathrm{V}(\mathrm{S}(P,Q);b')$.

Let $R = \mathrm{Rem}(P,Q)$ and let $\sigma(a)$ be the sign of PQ at a and $\sigma(b)$ be the sign of PQ at b.

Lemma 1.3.8

$$\mathrm{V}(\mathrm{S}(P,Q);a,b) = \begin{cases} \mathrm{V}(\mathrm{S}(Q,-R);a,b) + \sigma(b) & \text{if } \sigma(a)\sigma(b) = -1, \\ \mathrm{V}(\mathrm{S}(Q,-R);a,b) & \text{if } \sigma(a)\sigma(b) = 1. \end{cases}$$

Proof: The claim follows from the fact that at any x which is not a root of P and Q (and in particular at a and b)

$$\mathrm{V}(\mathrm{S}(P,Q);x) = \begin{cases} \mathrm{V}(\mathrm{S}(Q,-R);x) + 1 & \text{if } P(x)Q(x) < 0, \\ \mathrm{V}(\mathrm{S}(Q,-R);x) & \text{if } P(x)Q(x) > 0, \end{cases}$$

looking at all possible cases. □

Lemma 1.3.9

$$\mathrm{Ind}\left(\frac{Q}{P};a,b\right) = \begin{cases} \mathrm{Ind}\left(\dfrac{-R}{Q};a,b\right) + \sigma(b) & \text{if } \sigma(a)\sigma(b) = -1, \\ \mathrm{Ind}\left(\dfrac{-R}{Q};a,b\right) & \text{if } \sigma(a)\sigma(b) = 1. \end{cases}$$

Proof: We can suppose without loss of generality that Q and P are coprime. Indeed if D be a greatest common divisor of P and Q and

$$P_1 = \frac{P}{D}, Q_1 = \frac{Q}{D}, R_1 = \mathrm{Rem}(P_1, Q_1) = \frac{R}{D},$$

then P_1 and Q_1 are coprime,

$$\mathrm{Ind}\left(\frac{Q}{P}; a, b\right) = \mathrm{Ind}\left(\frac{Q_1}{P_1}; a, b\right), \mathrm{Ind}\left(\frac{-R}{Q}; a, b\right) = \mathrm{Ind}\left(\frac{-R_1}{Q_1}; a, b\right),$$

and the signs of $P(x)Q(x)$ and $P_1(x)Q_1(x)$ coincide at any point which is not a root of PQ.

Let n_{-+} (respectively n_{+-}) denote the number of sign changes from -1 to 1 (respectively from 1 to -1) of PQ when x varies from a to b. Noting that

$$\mathrm{Ind}\left(\frac{R}{Q}; a, b\right) = \mathrm{Ind}\left(\frac{P}{Q}; a, b\right),$$

it follows from the definition of Cauchy index that

$$\mathrm{Ind}\left(\frac{Q}{P}; a, b\right) + \mathrm{Ind}\left(\frac{R}{Q}; a, b\right) = n_{-+} - n_{+-}.$$

The claim of the lemma is now clear, since

$$n_{-+} - n_{+-} = \begin{cases} 0 & \text{if } \sigma(a)\sigma(b) = 1 \\ \sigma(b) & \text{if } \sigma(a)\sigma(b) = -1. \end{cases}$$

□

The proof of the theorem now proceeds by induction on the number $n \geq 2$ of elements in the signed remainder sequence. The base case $n = 2$ corresponds to $R = 0$ and follows from Lemma 1.3.8 and Lemma 1.3.9. Let us suppose that the Theorem holds for $n - 1$ and consider P and Q such that their signed remainder sequence has n elements. The signed remainder sequence of Q and $-R$ has $n - 1$ elements and, by the induction hypothesis,

$$\mathrm{V}(\mathrm{S}(Q, -R); a, b) = \mathrm{Ind}\left(\frac{-R}{Q}; a, b\right).$$

So, by Lemma 1.3.8 and Lemma 1.3.9,

$$\mathrm{V}(\mathrm{S}(P,Q);a,b) = \mathrm{Ind}\left(\frac{Q}{P};a,b\right).$$

□

As a consequence of the above we derive the following theorem due to Sylvester.

Theorem 1.3.10 (Sylvester's theorem) *If $a < b$ are elements of $\mathrm{R} \cup \{-\infty, +\infty\}$ that are not roots of P, with $P, Q \in \mathrm{R}[X]$, then*

$$\mathrm{V}(\mathrm{S}(P,P'Q);a,b) = \mathrm{SQ}(Q,P;a,b).$$

Proof: This is immediate from Theorem 1.3.7 and Proposition 1.3.4. □

The sequence of signed remainders of P and P', $\mathrm{S}(P,P')$, is the **Sturm sequence** of P.

As an easy consequence of Theorem 1.3.10 we have the following theorem.

Theorem 1.3.11 (Sturm's theorem) *With the same hypothesis and notations used in Theorem 1.3.10, $\mathrm{V}(\mathrm{S}(P,P');a,b)$ is the number of roots of P in the interval (a,b).*

Proof: The proof is immediate by take $Q = 1$ in the previous corollary. □

Example 1.3.12 Consider the polynomial $P = X^4 - 5X^2 + 4$. The Sturm sequence of P is

$$\begin{aligned}
\mathrm{S}_0(P,P') &= P = X^4 - 5X^2 + 4,\\
\mathrm{S}_1(P,P') &= P' = 4X^3 - 10X,\\
S_2(P,P') &= \frac{5}{2}X^2 - 4,\\
\mathrm{S}_3(P,P') &= \frac{18}{5}X,\\
\mathrm{S}_4(P,P') &= 4.
\end{aligned}$$

The leading coefficients of the Sturm sequence are $1, 1, \frac{5}{2}, \frac{18}{5}, 4$, and the degrees of the polynomials in the Sturm sequence are $4, 3, 2, 1, 0$. The signs of the polynomials of the Sturm sequence at $-\infty$ are $+-+-+$, and the signs of the polynomials of the Sturm sequence at $+\infty$ are $+++++$, so $\mathrm{V}(\mathrm{S}(P, P'); -\infty, +\infty) = 4$. There are indeed 4 real roots: $(1, -1, 2, -2)$.

We have the following method for computing a Sturm-query. Recall that the Sturm-query of Q for P is the number

$$\begin{gathered}\mathrm{SQ}(Q, P) = \\ \#(\{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) > 0\}) - \\ \#(\{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) < 0\}).\end{gathered}$$

Algorithm 1.3.13 (Sylvester Univariate Sturm-query)

Input: *a non-zero univariate polynomial P and a univariate polynomial Q, both with coefficients in* K.

Output: *the Sturm-query* $\mathrm{SQ}(Q, P)$.

Procedure:

Initialization: $\mathrm{S}_0 := P$, $\mathrm{S}_1 := P'Q$, $i := 1$.

While $\mathrm{S}_i \neq 0$

$$\mathrm{S}_{i+1} = -\mathrm{Rem}(\mathrm{S}_{i-1}, \mathrm{S}_i),$$
$$i := i + 1.$$

Compute the difference in sign variations at $-\infty$ and $+\infty$ from the degrees and signs of leading coefficients of the polynomials in this sequence.

Proof of correctness: The correctness follows from Theorem 1.3.10 (Sylvester's theorem). □

The bitsizes of the coefficients in the signed remainder sequence can indeed increase dramatically as we see in the next example.

Example 1.3.14 Consider the following numerical example:

$$P := 9X^{13} - 18X^{11} - 33X^{10} + 102X^8 + 7X^7 - 36X^6$$
$$-122X^5 + 49X^4 + 93X^3 - 42X^2 - 18X + 9.$$

The greatest common divisor of P and P' is of degree 5. The leading coefficients of the signed remainder sequence of P and P' are:

$$\frac{36}{13},$$
$$-\frac{10989}{16},$$
$$-\frac{2228672}{165649},$$
$$-\frac{900202097355}{4850565316},$$
$$-\frac{3841677139249510908}{543561530761725025},$$
$$-\frac{664885490073994444789496725}{67614035252757953531569 6712},$$
$$-\frac{2001176705547816993081646924785441 84}{1807309302290980501324553958871415645}.$$

1.4 Signed Subresultant Polynomials

Now we define and study the subresultant polynomials [8] which provide another real root counting method. Their coefficients are determinants extracted from the Sylvester matrix, and they are closely related to the remainder sequence[7].

1.4.1 Resultant and Subresultant Coefficients

Let P and Q be two non-zero polynomials of degree p and q in D[X], where D is a ring. When D is a domain, its fraction field is denoted by K. Let

$$\begin{aligned} P &= a_p X^p + a_{p-1} X^{p-1} + \cdots + a_0, \\ Q &= b_q X^q + b_{q-1} X^{q-1} + \cdots + b_0. \end{aligned}$$

We define the Sylvester matrix associated to P and Q and the resultant of P and Q.

Notation 1.4.1 The **Sylvester matrix** of P and Q, $\mathrm{Syl}(P,Q)$, is the matrix

$$\underbrace{\begin{pmatrix} a_p & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_p & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 \\ b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 \cdots & \cdots & \cdots & 0 & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 \end{pmatrix}}_{p+q} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} \begin{matrix} q \\ \\ \\ \\ \\ \\ p \\ \\ \\ \end{matrix}.$$

It has $p+q$ columns and $p+q$ rows. Note that its rows are

$$X^{q-1}P, \ldots, P, X^{p-1}Q, \ldots, Q$$

considered as vectors in the basis $X^{p+q-1}, \ldots, X, 1$.

The **resultant** of P and Q, denoted $\mathrm{Res}(P,Q)$, is the determinant of $\mathrm{Syl}(P,Q)$.

Remark 1.4.2 This matrix comes about quite naturally since it is the transpose of the matrix of the linear mapping

$$U, V \mapsto UP + VQ,$$

where (U,V) is identified with

$$(u_{q-1}, \ldots, u_0, v_{p-1}, \ldots, v_0),$$

and

$$U = u_{q-1}X^{q-1} + \cdots + u_0$$

and

$$V = v_{p-1}X^{p-1} + \cdots + v_0.$$

The following lemma is clear from this remark.

Lemma 1.4.3 *Let* D *be a domain.*

$$\mathrm{Res}(P,Q) = 0$$

if and only if there exist non-zero polynomials $U \in \mathrm{K}[X]$ *and* $V \in \mathrm{K}[X]$, *with* $\deg(U) < q$ *and* $\deg(V) < p$, *such that* $UP + VQ = 0$.

We can now prove the well-known proposition.

Proposition 1.4.4 *Let* D *be a domain. Then* $\mathrm{Res}(P,Q) = 0$ *if and only if* P *and* Q *have a common factor in* $\mathrm{K}[X]$.

Proof: The proposition is an immediate consequence of the preceding lemma, since the least common multiple of P and Q has degree $< p+q$ if and only if there exist non-zero polynomials U and V with $\deg(U) < q$ and $\deg(V) < p$ such that $UP + VQ = 0$. □

If D is a domain, with fraction field K, $a_p \neq 0$ and $b_q \neq 0$, the resultant can be expressed as a function of the roots of P and Q in an algebraically closed field C containing K.

Theorem 1.4.5 *Let*

$$P = a_p \prod_{i=1}^{p} (X - x_i)$$

$$Q = b_q \prod_{j=1}^{q} (X - y_j),$$

in other words $x_1, \ldots, x_p$ *are the roots of* P *(counted with multiplicities) and* $y_1, \ldots, y_q$ *are the roots of* Q *(counted with multiplicities).*

$$\mathrm{Res}(P,Q) = a_p^q b_q^p \prod_{i=1}^{p} \prod_{j=1}^{q} (x_i - y_j).$$

Proof: Let

$$\Theta(P,Q) = a_p^q b_q^p \prod_{i=1}^{p} \prod_{j=1}^{q} (x_i - y_j).$$

If P and Q have a root in common,

$$\mathrm{Res}(P,Q) = \Theta(P,Q) = 0,$$

and the theorem holds. So we suppose now that P and Q are coprime. The theorem is proved by induction on the length n of the remainder sequence of P and Q.

When $n = 2$, Q is a constant b, and

$$\mathrm{Res}(P,Q) = \Theta(P,Q) = b^p.$$

The induction step uses the following lemma.

Lemma 1.4.6 *Let R be the remainder of the Euclidean division of P by Q and let r be the degree of R. Then,*

$$\mathrm{Res}(P,Q) = (-1)^{pq} b_q^{p-r} \mathrm{Res}(Q,R),$$
$$\Theta(P,Q) = (-1)^{pq} b_q^{p-r} \Theta(Q,R).$$

Proof : Let $R = c_r X^r + \ldots + c_0$. Replacing the rows of coefficients of the polynomials $X^{q-1}P, \ldots, P$ by the rows of coefficients of the polynomials $X^{q-1}R, \ldots, R$ in the Sylvester matrix of P and Q gives the matrix

$$M = \underbrace{\begin{pmatrix}
0 & \cdots & c_r & \cdots & \cdots & \cdots & \cdots & c_0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & \cdots & c_r & \cdots & \cdots & \cdots & \cdots & c_0 \\
b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & 0 & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\
\vdots & \ddots & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0
\end{pmatrix}}_{p+q} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} q \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} p-r \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} r \end{matrix}$$

such that

$$\det(M) = \mathrm{Res}(P,Q).$$

Indeed,

$$R = P - \sum_{i=0}^{p-q} d_i(X^i Q),$$

where $C = \sum_{i=0}^{p-q} d_i X^i$ is the quotient of P in the euclidean division of P by Q, and adding to a row a multiple of other rows does not change the determinant. Denoting by N the matrix whose rows are

$$X^{p-1}Q, \ldots, X^{r-1}Q, \ldots, Q, X^{q-1}R, \ldots, R,$$

we note that

$$N = \underbrace{\begin{pmatrix}
b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & 0 & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\
\vdots & \ddots & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 \\
0 & \cdots & c_r & \cdots & \cdots & \cdots & \cdots & c_0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & \cdots & c_r & \cdots & \cdots & \cdots & \cdots & c_0
\end{pmatrix}}_{p-r+q+r} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} \end{matrix} \begin{matrix} p-r \\ r \\ q \end{matrix}$$

is obtained from M by exchanging the order of rows, so that

$$\det(N) = (-1)^{pq} \det(M).$$

It is clear, developing the determinant of N by its $p-r$ first columns, that

$$\det(N) = b_q^{p-r} \mathrm{Res}(Q, R).$$

On the other hand, since $P = CQ + R$, $P(y_j) = R(y_j)$ and

$$\Theta(P, Q) = a_p^q \prod_{i=1}^{p} Q(x_i) = (-1)^{pq} b_q^p \prod_{j=1}^{q} P(y_j),$$

we have

$$\Theta(P,Q) = (-1)^{pq} b_q^p \prod_{j=1}^{q} P(y_j)$$

$$= (-1)^{pq} b_q^p \prod_{j=1}^{q} R(y_j) = (-1)^{pq} b_q^{p-r} \Theta(Q,R).$$

□

□

For any ring D, the following holds:

Proposition 1.4.7 *If $P, Q \in \mathrm{D}[X]$, then there exist $U, V \in \mathrm{D}[X]$ such that $\deg(U) < q$, $\deg(V) < p$, and*

$$\mathrm{Res}(P,Q) = UP + VQ.$$

Proof: Let $\mathrm{Syl}(P,Q)^*$ be the matrix whose first $p+q-1$ columns are the first $p+q-1$ first columns of $\mathrm{Syl}(P,Q)$ and such that the elements of the last column are the polynomials $X^{q-1}P, \ldots, P, X^{p-1}Q, \ldots, Q$. Using the linearity of $\det(\mathrm{Syl}(P,Q)^*)$ as a function of its last column it is clear that

$$\det(\mathrm{Syl}(P,Q)^*) = \mathrm{Res}(P,Q) + \sum_{j=1}^{p+q-1} d_j X^j,$$

where d_j is the determinant of the matrix $\mathrm{Syl}(P,Q)_j$ whose first $p+q-1$ columns are the first $p+q-1$ columns of $\mathrm{Syl}(P,Q)$ and such that the last column is the $p+q-j$-th column of $\mathrm{Syl}(P,Q)$. Since $\mathrm{Syl}(P,Q)_j$ has two identical columns, $d_j = 0$ for $j = 1, \ldots, p+q-1$ and

$$\det(\mathrm{Syl}(P,Q)^*) = \mathrm{Res}(P,Q).$$

Expanding the determinant of $\mathrm{Syl}(P,Q)^*$ by its last column, we obtain the claimed identity. □

The Sylvester matrix and the resultant also have the following useful interpretation. Let C be an algebraically closed field. Identify a monic polynomial $X^q + b_{q-1}X^{q-1} + \cdots + b_0 \in \mathrm{C}[X]$ of degree q with the point $(b_{q-1}, \ldots, b_0) \in \mathrm{C}^q$. Let

$$\begin{array}{rcl} m : \mathrm{C}^q \times \mathrm{C}^p & \longrightarrow & \mathrm{C}^{q+p} \\ (Q, P) & \longmapsto & QP \end{array}$$

be the mapping defined by the multiplication of monic polynomials. The map m sends

$$(b_{q-1}, \ldots, b_0, a_{p-1}, \ldots, a_0)$$

to the vector whose entries are $(m_{p+q-1}, \ldots, m_0)$, where

$$m_j = \sum_{q-i+p-k=j} b_{q-i}a_{p-k} \text{ for } j = p+q-1, \ldots, 0$$

(with $b_q = a_p = 1$). The following proposition is thus clear:

Proposition 1.4.8 *The Jacobian matrix of m is the Sylvester matrix of P and Q and the Jacobian of m is the resultant.*

Finally, the definition of resultants as determinants implies that:

Proposition 1.4.9 *If P is monic, $\deg(Q) \leq \deg(P)$, and $f\colon \mathrm{D} \to \mathrm{D}'$ is a ring homomorphism, then*

$$f(\mathrm{Res}(P, Q)) = \mathrm{Res}(f(P), f(Q))$$

(denoting by f the induced homomorphism from $\mathrm{D}[X]$ to $\mathrm{D}'[X]$).

We now define the Sylvester-Habicht matrices and the signed subresultant coefficients of P and Q.

Notation 1.4.10 Let $0 \leq j \leq \min(p, q)$ if $p \neq q$ (respectively $0 \leq j \leq p-1$ if $p = q$). The **j-th Sylvester-Habicht matrix** of P and Q,

denoted $\mathrm{SH}_j(P,Q)$, is the matrix

$$\underbrace{\begin{pmatrix} a_p & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_p & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 \\ 0 & \cdots & \cdots & 0 & b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 \\ \vdots & & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & 0 \\ \vdots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \vdots \\ 0 & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & & \vdots \\ b_q & \cdots & \cdots & \cdots & \cdots & \cdots & b_0 & 0 & \cdots & \cdots & 0 \end{pmatrix}}_{p+q-2j} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} \begin{matrix} q-j \\ \\ \\ \\ \\ p-j \\ \\ \\ \\ \end{matrix} .$$

It has $p+q-j$ columns and $p+q-2j$ rows. Note that its rows are

$$X^{q-j-1}P, \ldots, P, Q, \ldots, X^{p-j-1}Q$$

considered as vectors in the basis $X^{p+q-j-1}, \ldots, X, 1$.

The **j-th signed subresultant coefficient** denoted $\mathrm{sr}_j(P,Q)$ or sr_j is the determinant of the square matrix obtained by taking the first $p+q-2j$ columns of $\mathrm{SH}_j(P,Q)$.

Remark 1.4.11 This matrix comes about quite naturally since it is the transpose of the matrix of the mapping

$$U, V \mapsto UP + VQ,$$

where (U,V) is identified with

$$(u_{q-j-1}, \ldots, u_0, v_0, \ldots, v_{p-j-1}),$$

with

$$U = u_{q-j-1}X^{q-j-1} + \cdots + u_0$$

and

$$V = v_{p-j-1}X^{p-j-1} + \cdots + v_0.$$

The peculiar order of rows is adapted to the real root counting results presented later.

The following lemma is clear from this remark:

Lemma 1.4.12 *Let* D *be a domain and* $0 \leq j \leq \min(p,q)$ *if* $p \neq q$ *(respectively* $0 \leq j \leq p-1$ *if* $p = q$*). Then,*

$$\mathrm{sr}_j(P,Q) = 0$$

if and only if there exist non-zero polynomials $U \in \mathrm{K}[X]$ *and* $V \in \mathrm{K}[X]$*, with* $\deg(U) < q-j$ *and* $\deg(V) < p-j$*, such that* $\deg(UP+VQ) < j$*.*

Proposition 1.4.13 *Let* D *be a domain and* $0 \leq j \leq \min(p,q)$ *if* $p \neq q$ *(respectively* $0 \leq j \leq p-1$ *if* $p = q$*). Then* $\deg(\gcd(P,Q)) \geq j$ *if and only if*

$$\mathrm{sr}_0(P,Q) = \cdots = \mathrm{sr}_{j-1}(P,Q) = 0.$$

Proof: Suppose that $\deg(\gcd(P,Q)) \geq j$. Then, the least common multiple of P and Q,

$$\mathrm{lcm}(P,Q) = \frac{PQ}{\gcd(P,Q)}$$

has degree $\leq p+q-j$. This is clearly equivalent to the existence of polynomials U and V, with $\deg(U) \leq q-j$ and $\deg(V) \leq p-j$, such that $UP = -VQ = \mathrm{lcm}(P,Q)$. Or, equivalently, that there exist polynomials U and V with $\deg(U) \leq q-j$ and $\deg(V) \leq p-j$ such that $UP+VQ = 0$. This implies that $\mathrm{sr}_0 = \ldots = \mathrm{sr}_{j-1} = 0$ using Lemma 1.4.12.

The reverse implication is proved by induction on j. If $j-1 = 0$, $\mathrm{sr}_0 = 0$ implies, using Lemma 1.4.12, that there exist U and V with $\deg(U) < q$ and $\deg(V) < p$ satisfying $UP+VQ = 0$. Hence $\deg(\gcd(P,Q)) > 0$. If

$$\mathrm{sr}_0(P,Q) = \cdots = \mathrm{sr}_{j-2}(P,Q) = 0,$$

the induction hypothesis implies that $\deg(\gcd(P,Q)) \geq j-1$. If in addition $\mathrm{sr}_{j-1} = 0$ then, by Lemma 1.4.12, there exist U and V with $\deg(U) \leq q-j$ and $\deg(V) \leq p-j$ such that $\deg(UP+VQ) < j-1$. Since the greatest common divisor of P and Q divides $UP+VQ$ and has degree $\geq j-1$, we have $UP+VQ = 0$, which implies that $\deg(\mathrm{lcm}(P,Q)) \leq p+q-j$ and hence $\deg(\gcd(P,Q)) \geq j$. □

The following corollary is clear, using Lemma 1.4.12 and Proposition 1.4.13.

Corollary 1.4.14 *Let* D *be a domain and* $0 \leq j \leq \min(p,q)$ *if* $p \neq q$ *(respectively* $0 \leq j \leq p-1$ *if* $p = q$*). Then* $\deg(\gcd(P,Q)) = j$ *if and only if*

$$\mathrm{sr}_0(P,Q) = \cdots = \mathrm{sr}_{j-1}(P,Q) = 0, \mathrm{sr}_j(P,Q) \neq 0.$$

Remark 1.4.15 Writing $\varepsilon_i = (-1)^{i(i-1)/2}$, we note that ε_i is the signature of the permutation reversing the order of i consecutive rows in a matrix. For every natural number $i \geq 1$,

$$\varepsilon_{4i} = 1, \varepsilon_{4i-1} = -1, \varepsilon_{4i-2} = -1, \varepsilon_{4i-3} = 1. \tag{1.1}$$

In particular, $\varepsilon_{i-2j} = (-1)^j \varepsilon_i$.

Thus, it is clear from the definitions that

$$\mathrm{sr}_0(P,Q) = \varepsilon_p \mathrm{Res}(P,Q). \tag{1.2}$$

Note that, as a consequence Proposition 1.4.4 is a special case of Proposition 1.4.13.

If P is monic,we define the discriminant as

$$\mathrm{Disc}(P) = \mathrm{sr}_0(P,P') = \varepsilon_p \mathrm{Res}(P,P'). \tag{1.3}$$

If P is not monic, we define

$$\mathrm{Disc}(P) = \frac{1}{a_p}\mathrm{sr}_0(P,P') = \frac{1}{a_p}\varepsilon_p \mathrm{Res}(P,P'). \tag{1.4}$$

since

$$\mathrm{Disc}(P) = a_p^{2p-2}\mathrm{Disc}\left(\frac{P}{a_p}\right),$$
$$\mathrm{sr}_0(P,P') = a_p^{2p-1}\mathrm{sr}_0\left(\frac{P}{a_p}, \frac{P'}{a_p}\right).$$

Exercise 1.4.16 *a)*

$$\mathrm{Disc}(P) = \prod_{p \geq i > j \geq 1} (x_i - x_j)^2.$$

b) If $P \in \mathrm{R}[X]$ *is monic with* R *real closed, of degree* p*, and with* p *distinct roots in* C*, and* t *is the number of roots of* P *in* R*.*

$\mathrm{Disc}(P) > 0$ *if and only if* $t \equiv p$ modulo 4*,*

$\mathrm{Disc}(P) < 0$ *if and only if* $t \equiv p-2$ modulo 4*.*

Remark 1.4.17 Note that if $P \in \mathrm{D}[X]$, then $\mathrm{Disc}(P) \in \mathrm{D}$, since $\mathrm{Syl}(P, P')$ has coefficients in D and the only non zero elements of its first column are a_p and pa_p which are both multiple of a_p.

Definition 1.4.18 *The discriminant of a polynomial* P *with coefficients in a ring is defined as the determinant of the matrix*

$$\underbrace{\begin{pmatrix}
1 & a_{p-1} & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 & 0 & \cdots & 0 \\
0 & a_p & a_{p-1} & \ddots & \ddots & \ddots & \ddots & \ddots & a_0 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & a_p & a_{p-1} & \cdots & \cdots & \cdots & \cdots & \cdots & a_0 \\
0 & \cdots & \cdots & 0 & pa_p & \cdots & \cdots & \cdots & \cdots & 2a_2 & a_1 \\
\vdots & & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & 0 \\
\vdots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \iddots & \vdots \\
0 & pa_p & \iddots & \iddots & \iddots & \iddots & 2a_2 & a_1 & \iddots & & \vdots \\
p & (p-1)a_{p-1} & \cdots & \cdots & \cdots & 2a_2 & a_1 & 0 & \cdots & \cdots & 0
\end{pmatrix}}_{2p-1} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} p-1 \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} p .$$

1.4.2 Polynomial Determinants

We define polynomial determinants, which will be useful in the study of subresultant polynomials.

Let K be a field of characteristic 0. Consider the K-vector space $\mathcal{F}_n$, consisting of polynomials whose degrees are less than n, equipped with the basis

$$\mathcal{B} = X^{n-1}, \ldots, X, 1.$$

We associate to a list of polynomials $\mathcal{P} = P_1, \ldots, P_m$, with $m \leq n$ a matrix $\mathrm{Mat}(\mathcal{P})$ whose rows are the coordinates of the P_i's in the basis $\mathcal{B}$. Note that $\mathrm{Mat}(\mathcal{B})$ is the identity matrix of size n.

Let $0 < m \leq n$. A mapping Φ from $(\mathcal{F}_n)^m$ to $\mathcal{F}_{n-m+1}$ is **multilinear** if for $\lambda \in \mathrm{K}, \mu \in \mathrm{K}$

$$\Phi(\ldots, \lambda A_i + \mu B_i, \ldots) = \lambda\Phi(\ldots, A_i, \ldots) + \mu\Phi(\ldots, B_i, \ldots).$$

A mapping Φ from $(\mathcal{F}_n)^m$ to $\mathcal{F}_{n-m+1}$ is **alternating** if

$$\Phi(\ldots, A, \ldots, A, \ldots) = 0.$$

A mapping Φ from $(\mathcal{F}_n)^m$ to $\mathcal{F}_{n-m+1}$ is **antisymmetric** if

$$\Phi(\ldots, A, \ldots, B, \ldots) = -\Phi(\ldots, B, \ldots, A, \ldots).$$

Lemma 1.4.19 *A mapping from $(\mathcal{F}_n)^m$ to $\mathcal{F}_{n-m+1}$ which is multilinear and alternating is antisymmetric.*

Proof: Since Φ is alternating,

$$\Phi(\ldots, A + B, \ldots, A + B, \ldots) = 0$$
$$\Phi(\ldots, A, \ldots, A, \ldots) = \Phi(\ldots, B, \ldots, B, \ldots) = 0;$$

Using multilinearity, we get easily

$$\Phi(\ldots, A, \ldots, B, \ldots) + \Phi(\ldots, B, \ldots, A, \ldots) = 0.$$

□

Proposition 1.4.20 *There exists a unique multilinear alternating mapping Φ from $(\mathcal{F}_n)^m$ to $\mathcal{F}_{n-m+1}$ satisfying*

$$\begin{cases} \Phi(X^{n-1}, \ldots, X^{n-m+1}, X^i) = X^i & \text{if } i \leq n - m \\ \Phi(X^{\ell(1)}, \ldots, X^{\ell(m-1)}, X^{\ell(m)}) = 0 & \text{if } n > \ell(1) > \ldots > \ell(m), \\ & \text{and there exists } i < m \\ & \text{with } \ell(i) \neq n - i. \end{cases}$$

Proof: Decomposing each P_i in the basis $\mathcal{B}$ of monomials and using multilinearity and antisymmetry, it is clear that a multilinear and alternating mapping Φ from $\mathcal{F}_n^m$ to $\mathcal{F}_{n-m+1}$ depends only on the values $\Phi(X^{\ell(1)}, \ldots, X^{\ell(m-1)}, X^{\ell(m)})$ for $\ell(1) > \ldots > \ell(m)$. This proves the uniqueness.

In order to prove existence, let m_i, $i \leq n$, be the $m \times m$ minor of $\mathrm{Mat}(\mathcal{P})$ based on the columns $1, \ldots, m-1, n-i$, then

$$\Phi(\mathcal{P}) = \sum_{i \leq n-m} m_i X^i \tag{1.5}$$

satisfies all the properties required. □

The (n,m)-**polynomial determinant** mapping, denoted $\mathrm{pdet}_{n,m}$, is the unique multilinear alternating mapping from $\mathcal{F}_n^m$ to $\mathcal{F}_{n-m+1}$ satisfying the properties of Proposition 1.4.20.

When $n = m$, it is clear that $\mathrm{pdet}_{n,n}(\mathcal{P}) = \det(\mathrm{Mat}(\mathcal{P}))$, since det is known to be the unique multilinear alternating map sending the identity matrix to 1.

On the other hand, when $m = 1$, $\mathrm{pdet}(P)_{n,1}(X^i) = X^i$ and, by linearity, $\mathrm{pdet}(P)_{n,1} = P$.

If follows immediately from the definition that

Lemma 1.4.21 *Let $\mathcal{P} = P_1, \ldots, P_m$.*

If $\mathcal{Q} = Q_1, \ldots, Q_m$ is such that $Q_i = P_i$, $i \neq j$, $Q_j = P_j + \sum_{j \neq i} \lambda_j P_j$, then $\mathrm{pdet}_{n,m}(\mathcal{Q}) = \mathrm{pdet}_{n,m}(\mathcal{P})$.

If $\mathcal{Q} = P_m, \ldots, P_1$, then $\mathrm{pdet}_{n,m}(\mathcal{Q}) = \varepsilon_m \mathrm{pdet}_{n,m}(\mathcal{P})$, *where $\varepsilon_m = (-1)^{m(m-1)/2}$ (see Notation 1.4.15).*

We consider now a sequence $\mathcal{P}$ of polynomials with coefficients in a ring D. Equation (1.5) provides a definition of the (n,m)-polynomial determinant $\mathrm{pdet}_{n,m}(\mathcal{P})$ of $\mathcal{P}$. Note that $\mathrm{pdet}_{n,m}(\mathcal{P}) \in \mathrm{D}[X]$.

We can express the polynomial determinant as the classical determinant of a matrix whose last column has polynomial entries in the following way:

If $\mathcal{P} = P_1, \ldots, P_m$ we let $\mathrm{Mat}(\mathcal{P})^*$ be the $m \times m$ matrix whose first $m-1$ columns are the first $m-1$ columns of $\mathrm{Mat}(\mathcal{P})$ and such that the elements of the last column are the polynomials $P_1, \ldots, P_m$.

With this notation, we have

Lemma 1.4.22

$$\mathrm{pdet}_{n,m}(\mathcal{P}) = \det(\mathrm{Mat}(\mathcal{P})^*).$$

Proof: Using the linearity of $\det(\mathrm{Mat}(\mathcal{P})^*)$ as a function of its last column, it is clear that $\det(\mathrm{Mat}(\mathcal{P})^*) = \sum_{i\leq n} m_i X^i$, using the notation of Proposition 1.4.20. For $i > n-m$, $m_i = \bar{0}$ since it is the determinant of a matrix with two equal columns. □

Remark 1.4.23 Expanding $\det(\mathrm{Mat}(\mathcal{P})^*)$ by its last column we observe that $\mathrm{pdet}_{n,j}(\mathcal{P})$ is a linear combination of the P_i with coefficients equal (up to sign) $(m-1)\times(m-1)$ to minors extracted on the $m-1$ first columns of $\mathcal{P}$. It is thus a linear combination with coefficients in D of the P_i's.

The following immediate consequences of Lemma 1.4.22 will be useful.

Lemma 1.4.24 *Let* $\mathcal{P} = P_1, \ldots, P_\ell, P_{\ell+1}, \ldots, P_m$ *be such that*

$$\deg(P_i) = n-i, i \leq \ell, \deg(P_i) < n-1-\ell, \ell < i \leq m.$$

Then

$$\mathrm{pdet}_{n,m}(\mathcal{P}) = \prod_{i=1}^{\ell} \mathrm{lcof}(P_i)\mathrm{pdet}_{n-\ell,m-\ell}(\mathcal{Q}),$$

where $\mathcal{Q} = P_{\ell+1}, \ldots, P_m$.

Proof: The shape of the matrix $\mathrm{Mat}(\mathcal{P})$ is as follows

$$\underbrace{\begin{pmatrix} \mathrm{lcof}(P_1) & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \mathrm{lcof}(P_\ell) & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \mathrm{cof}_{n-\ell-1}(P_{\ell+1}) & \cdots & \mathrm{cof}_0(P_{\ell+1}) \\ \vdots & & & \vdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \mathrm{cof}_{n-\ell-1}(P_m) & \cdots & \mathrm{cof}_0(P_m) \end{pmatrix}}_{n} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} \begin{matrix} \ell \\ \\ \\ \\ \\ m-\ell \\ \\ \end{matrix}.$$

Using Lemma 1.4.22, develop the determinant $\det(\mathrm{Mat}(\mathcal{P})^*)$ by its first ℓ columns. □

Lemma 1.4.25 *Let $\mathcal{P} = P_1, , \ldots, P_m$ be such that for every i, $1 \leq i \leq m$, $\deg(P_i) < n-1$. Then*

$$\mathrm{pdet}_{n,m}(\mathcal{P}) = 0.$$

Proof: Using Lemma 1.4.22, develop the determinant $\det(\mathrm{Mat}(\mathcal{P})^*)$ by its first column which is zero. □

1.4.3 Definition of Signed Subresultants

For the remainder of this chapter, let P and Q be two non-zero polynomials of degrees p and q and with $q < p$ with coefficients in an integral domain D. The fraction field of D is denoted by K. Let

$$\begin{aligned} P &= a_p X^p + a_{p-1} X^{p-1} + a_{p-2} X^{p-2} + \cdots + a_0, \\ Q &= b_q X^q + b_{q-1} X^{q-1} + \cdots + b_0. \end{aligned}$$

We define the signed subresultants of P and Q and some related notions.

Notation 1.4.26 For $0 \leq j \leq q$, the **j-th signed subresultant of P and Q**, denoted $\mathrm{SR}_j(P,Q)$, is the $(p+q-j, p+q-2j)$-polynomial determinant of the sequence of polynomials

$$\mathcal{H}_j(P,Q) = X^{q-j-1}P, \ldots, P, Q, \ldots, X^{p-j-1}Q.$$

The Sylvester-Habicht matrix $\mathrm{SH}_j(P,Q)$ (Notation 1.4.1) is $\mathrm{Mat}(\mathcal{H}_j(P,Q))$. Clearly, $\deg(\mathrm{SR}_j(P,Q)) \leq j$. By convention, we extend these definitions for $q < j \leq p$ by

$$\begin{aligned} \mathrm{SR}_p(P,Q) &= \mathrm{sign}(a_p^{p-q-1})P, \\ \mathrm{SR}_{p-1}(P,Q) &= \mathrm{sign}(a_p^{p-q+1})Q, \\ \mathrm{SR}_j(P,Q) &= 0, \ q < j < p-1. \end{aligned}$$

Also by convention $\mathrm{SR}_{-1}(P,Q) = 0$. Note that $\mathrm{SR}_q(P,Q) = \varepsilon_{p-q} b_q^{p-q-1} Q$.

The j-th signed subresultant coefficient of P and Q, denoted $\mathrm{sr}_j(P,Q)$, is the coefficient of X^j in $\mathrm{SR}_j(P,Q)$, $j < p$, and by convention

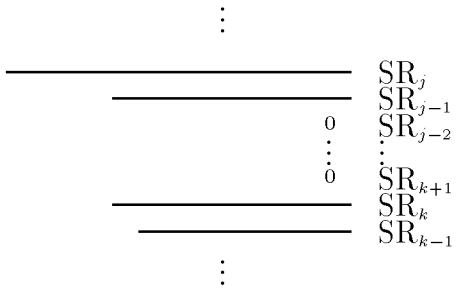
$$\mathrm{sr}_p(P,Q) = \mathrm{sign}(a_p^{p-q}).$$

Note that $\mathrm{sr}_q(P,Q) = \varepsilon_{p-q} b_q^{p-q}$. The j-th signed subresultant coefficient was already considered in Section 1.4.1.

If $\deg(\mathrm{SR}_j(P,Q)) = j$ (equivalently if $\mathrm{sr}_j(P,Q) \neq 0$) we say that $\mathrm{SR}_j(P,Q)$ is **non-defective**. If $\deg(\mathrm{SR}_j(P,Q)) = k < j$ we say that $\mathrm{SR}_j(P,Q)$ is **defective** of degree k.

1.4.4 Structure Theorem for Signed Subresultants

We are going to see that the non-zero signed subresultants are proportional to the polynomials in the signed remainder sequence. Moreover, the signed subresultant polynomials present the gap structure, graphically displayed by the following diagram: when SR_{j-1} is defective of degree k, SR_{j-1} and SR_k are proportional, $\mathrm{SR}_{j-2}, \ldots, \mathrm{SR}_{k+1}$ are zero.



The structure theorem for signed subresultants describes precisely this situation. We omit P and Q in the notation of the theorem. We write $\overline{\mathrm{sr}}_j$ for $\mathrm{lcof}(\mathrm{SR}_j)$. Note that if $\deg(\mathrm{SR}_j) = j$, $\overline{\mathrm{sr}}_j = \mathrm{sr}_j$.

Theorem 1.4.27 (Structure theorem) *Let $0 \leq j < i \leq p$. Suppose that SR_{i-1} is non-zero and of degree j.*

a) If SR_{j-1} is zero, then $\mathrm{SR}_{i-1} = \gcd(P,Q)$, and SR_ℓ is zero, $\ell \leq j-1$.

b) If $\mathrm{SR}_{j-1} \neq 0$ has degree k then

$$\mathrm{sr}_j \overline{\mathrm{sr}}_{i-1} \mathrm{SR}_{k-1} = -\mathrm{Rem}(\mathrm{sr}_k \overline{\mathrm{sr}}_{j-1} \mathrm{SR}_{i-1}, \mathrm{SR}_{j-1}).$$

In fact, the quotient is in $\mathrm{D}[X]$. That is

$$\mathrm{sr}_j \overline{sr}_{i-1} \mathrm{SR}_{k-1} = -\mathrm{sr}_k \overline{\mathrm{sr}}_{j-1} \mathrm{SR}_{i-1} + C_{k-1} \mathrm{SR}_{j-1} \qquad (1.6)$$

with $\deg(C_{k-1}) = j - k$ *and* $C_{k-1} \in \mathrm{D}[X]$.

Moreover if $j \leq q$, $k < j-1$, SR_k *is proportional to* SR_{j-1}. *More precisely*

i) $\mathrm{SR}_{j-2} = \ldots = \mathrm{SR}_{k+1} = 0$,

ii) $\mathrm{sr}_k = \varepsilon_{j-k} \dfrac{\overline{\mathrm{sr}}_{j-1}^{j-k}}{\mathrm{sr}_j^{j-k-1}}$, *(where* $\varepsilon_i = (-1)^{i(i-1)/2}$*)*

iii) $\overline{\mathrm{sr}}_{j-1}\mathrm{SR}_k = \mathrm{sr}_k\mathrm{SR}_{j-1}$.

Note that Theorem 1.4.27 implies that SR_{i-1} and SR_j are proportional. The following corollary of Theorem 1.4.27 will be used later in this chapter.

Corollary 1.4.28 *If* SR_{j-1} *is defective of degree* k,

$$\mathrm{sr}_j^2\mathrm{SR}_{k-1} = -\mathrm{Rem}(\mathrm{sr}_k\overline{\mathrm{sr}}_{j-1}\mathrm{SR}_j, \mathrm{SR}_{j-1}).$$

Proof: Immediate, using

$$\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}\mathrm{SR}_{k-1} = -\mathrm{Rem}(\mathrm{sr}_k\overline{\mathrm{sr}}_{j-1}\mathrm{SR}_{i-1}, \mathrm{SR}_{j-1})$$

and the proportionality between SR_{i-1} and SR_j. □

Note that we have seen in Proposition 1.4.13 that $\deg(\gcd(P,Q))$ is the smallest j such that $\mathrm{sr}_j \neq 0$. The Structure Theorem 1.4.27 makes this statement more precise:

Corollary 1.4.29 *The last non-zero signed subresultant of* P *and* Q *is non-defective and a greatest common divisor of* P *and* Q.

Proof: Suppose that $\mathrm{SR}_j \neq 0$, and $\forall\ \ell < k$ $\mathrm{SR}_\ell = 0$. By Theorem 1.4.27 b) there exists i such that $\deg(\mathrm{SR}_{i-1}) = j$, and SR_{i-1} and SR_j are proportional. So SR_j is non-defective and by a) SR_{i-1} is a greatest common divisor of P and Q. □

Moreover, a consequence of the Structure Theorem 1.4.27 is that signed subresultant polynomials are closely related to the polynomials in the signed remainder sequence.

In the non-defective case, we have:

Corollary 1.4.30 *When all* $\mathrm{SR}_j(P,Q)$ *are non-defective,* $j = p, \ldots, 0$, *the signed subresultant polynomials are proportional up to a square to the polynomials in the signed remainder sequence.*

Proof: We consider the signed remainder sequence

$$\begin{aligned} \mathrm{S}_0 &= P, \\ \mathrm{S}_1 &= Q, \\ &\vdots \\ \mathrm{S}_{\ell+1} &= -\mathrm{Rem}(\mathrm{S}_{\ell-1}, \mathrm{S}_\ell), \\ &\vdots \\ \mathrm{S}_k &= -\mathrm{Rem}(\mathrm{S}_{k-2}, \mathrm{S}_{k-1}), \\ \mathrm{S}_{k+1} &= 0, \end{aligned}$$

and prove by induction on ℓ that $\mathrm{SR}_{p-\ell}$ is proportional to S_ℓ.

The claim is true for $\ell = 0$ and $\ell = 1$ by definition of SR_p and SR_{p-1}.

Suppose that the claim is true up to ℓ. In the non-defective case, the Structure Theorem 1.4.27 b) implies

$$\mathrm{sr}^2_{p-\ell+1}\mathrm{SR}_{p-\ell-1} = -\mathrm{Rem}(\mathrm{sr}^2_{p-\ell}\mathrm{SR}_{p-\ell+1}, \mathrm{SR}_{p-\ell}). \tag{1.7}$$

By induction hypothesis, $\mathrm{SR}_{p-\ell+1}$ and $\mathrm{SR}_{p-\ell}$ are proportional to $\mathrm{S}_{\ell-1}$ and S_ℓ. Thus, by definition of the signed remainder sequence and by equation (1.7) $\mathrm{SR}_{p-\ell-1}$, is proportional to $\mathrm{S}_{\ell+1}$. □

More generally, the signed subresultants are either proportional to polynomials in the remainder sequence or zero.

Corollary 1.4.31 *If* $\mathrm{S}_{\ell-1}$ *and* S_ℓ *are two successive polynomials in the signed remainder sequence of* P *and* Q, *of degrees* $d(\ell-1)$ *and* $d(\ell)$, *then* $\mathrm{SR}_{d(\ell-1)-1}$ *and* $\mathrm{SR}_{d(\ell)}$ *are proportional to* S_ℓ.

Proof: The proof if by induction on ℓ. Note first that $P = \mathrm{S}_0$ is proportional to SR_p. The claim is true for $\ell = 1$ by definition of SR_p, SR_{p-1}, and SR_q. Suppose that the claim is true up to ℓ. The Structure

Theorem 1.4.27 b) implies (with $i = d(\ell-2)$, $j = d(\ell-1)$,$k = d(\ell)$) that $\mathrm{SR}_{d(\ell)-1}$ is proportional to $\mathrm{Rem}(\mathrm{SR}_{d(\ell-2)-1}, \mathrm{SR}_{d(\ell-1)-1})$. By the induction hypothesis, $\mathrm{SR}_{d(\ell-2)-1}$ and $\mathrm{SR}_{d(\ell-1)-1}$ are proportional to $\mathrm{S}_{\ell-1}$ and S_ℓ. Thus, $\mathrm{SR}_{d(\ell)-1}$ is proportional to $\mathrm{S}_{\ell+1}$. Moreover $\mathrm{SR}_{d(\ell)-1}$ and $\mathrm{SR}_{d(\ell+1)}$ are proportional by the Structure Theorem 1.4.27. □

The proof of the structure theorem relies on the following proposition relating the signed subresultants of P and Q and of Q and $-R$, with $R = \mathrm{Rem}(P, Q)$.

We recall that P and Q are two non-zero polynomials of degrees p and q, $q < p$, with coefficients in an integral domain D, with

$$\begin{aligned} P &= a_p X^p + a_{p-1} X^{p-1} + a_{p-2} X^{p-2} + \cdots + a_0, \\ Q &= b_q X^q + b_{q-1} X^{q-1} + \cdots + b_0. \end{aligned}$$

Proposition 1.4.32 *Let r be the degree of $R = \mathrm{Rem}(P, Q)$.*

$$\begin{cases} \mathrm{SR}_j(P,Q) = \varepsilon_{p-q} b_q^{p-r} \mathrm{SR}_j(Q,-R) & \text{if } j \le r, \\ \mathrm{SR}_j(P,Q) = \varepsilon_{p-q} b_q^{p-r} \mathrm{SR}_j(Q,-R) = 0 & \text{if } r < j < q-1, \\ \mathrm{SR}_{q-1}(P,Q) = \varepsilon_{p-q} \mathrm{sign}(b_q^{q-r-1}) b_q^{p-q+1} \mathrm{SR}_{q-1}(Q,-R), & \\ \mathrm{SR}_q(P,Q) = \varepsilon_{p-q} \mathrm{sign}(b_q^{q-r+1}) b_q^{p-q-1} \mathrm{SR}_q(Q,-R), & \end{cases}$$

where $\varepsilon_i = (-1)^{i(i-1)/2}$.

Proof: Replacing the polynomials $X^{q-j-1}P, \ldots, P$ by the polynomials $X^{q-j-1}R, \ldots, R$ in $\mathcal{H}_j(P,Q)$ does not modify the polynomial determinant. Indeed,

$$R = P - \sum_{i=0}^{p-q} c_i (X^i Q),$$

where $C = \sum_{i=0}^{p-q} c_i X^i$ is the quotient of P in the euclidean division of P by Q, and adding to a polynomial of a sequence a multiple of another polynomial of the sequence does not change the polynomial determinant, by Lemma 1.4.21.

Reversing the order of the polynomials multiplies the polynomial determinant by ε_{p+q-2j} using again Lemma 1.4.21. Replacing R by -R

multiplies the polynomial determinant by $(-1)^{q-j}$, by Lemma 1.4.21, and $(-1)^{q-j}\varepsilon_{p+q-2j} = \varepsilon_{p-q}$ (see Notation 1.4.15). So

$$\mathrm{SR}_j(P,Q) = \varepsilon_{p-q}\mathrm{pdet}_{p+q-j,p+q-2j}(X^{p-j-1}Q,\ldots,Q,-R\ldots,-X^{q-j-1}R).$$

If $j \le r$,

$$\begin{aligned}&\mathrm{pdet}_{p+q-j,p+q-2j}(X^{p-j-1}Q,\ldots,Q,-R\ldots,-X^{q-j-1}R)\\ &= b_q^{p-r}\mathrm{pdet}_{q+r-j,q+r-2j}(X^{r-j-1}Q,\ldots,Q,-R\ldots,-X^{q-j-1}R)\\ &= b_q^{p-r}\mathrm{SR}_j(Q,-R),\end{aligned}$$

using Lemma 1.4.24.

If $r < j < q-1$,

$$\mathrm{pdet}_{p+q-j,p+q-2j}(X^{p-j-1}Q,\ldots,Q,-R\ldots,-X^{q-j-1}R) = 0,$$

using Lemma 1.4.24 and Lemma 1.4.25, since $\deg(-X^{q-j-1}R) < q-1$.

If $j = q-1$,

$$\begin{aligned}&\mathrm{pdet}_{p+1,p-q+2}(X^{p-q}Q,\ldots,Q,-R) = -b_q^{p-q+1}R\\ &= \mathrm{sign}(b_q^{q-r-1})b_q^{p-q+1}\mathrm{SR}_{q-1}(Q,-R),\end{aligned}$$

using Lemma 1.4.24 and the convention in Notation 1.4.26.

If $j = q$,

$$\begin{aligned}&\mathrm{pdet}_{p,p-q}(X^{p-q-1}Q,\ldots,Q) = b_q^{p-q-1}Q\\ &= \mathrm{sign}(b_q^{q-r-1})b_q^{p-q-1}\mathrm{SR}_q(Q,-R),\end{aligned}$$

using Lemma 1.4.24 and the convention in Notation 1.4.26. □

Proof of Theorem 1.4.27: For $q < j \le p$, the only thing to check is that

$$\mathrm{SR}_{q-1} = -\mathrm{Rem}(\mathrm{sr}_q\overline{\mathrm{sr}}_{p-1}\mathrm{SR}_p,\mathrm{SR}_{p-1}),$$

which follows from

$$-\varepsilon_{p-q}b_q^{p-q+1}R = -\mathrm{Rem}(\varepsilon_{p-q}b_q^{p-q+1}P,Q)$$

since $\mathrm{sr}_q = \varepsilon_{p-q} b_q^{p-q}$, $\overline{\mathrm{sr}}_{p-1} = \mathrm{sign}(a_p)^{p-q-1} b_q$, and $\mathrm{SR}_p = \mathrm{sign}(a_p)^{p-q-1} P$.

The remainder of the proof is by induction on the length of the remainder sequence of P and Q.

Suppose that the theorem is true for $Q, -R$. The fact that the theorem holds for P, Q for $j \leq r$ is clear by Proposition 1.4.32, since $\mathrm{SR}_j(P,Q)$ and $\mathrm{SR}_j(Q,-R)$, $j \leq r$, are proportional, with the same factor of proportionality $\varepsilon_{p-q} b_q^{p-r}$.

For $r < j \leq q$, the only thing to check is that

$$\overline{\mathrm{sr}}_{p-1}\mathrm{sr}_q\mathrm{SR}_{r-1} = -\mathrm{Rem}(\mathrm{sr}_r\overline{\mathrm{sr}}_{q-1}\mathrm{SR}_{p-1}, \mathrm{SR}_{q-1}),$$

which follows from

$$\mathrm{SR}_{r-1}(Q,-R) = -\mathrm{Rem}(\mathrm{sr}_r(Q,-R)\overline{\mathrm{sr}}_{q-1}(Q,-R)\mathrm{SR}_q(Q,-R), \mathrm{SR}_{q-1}(Q,-R))$$

since

$$\begin{aligned}
\overline{\mathrm{sr}}_{p-1}\mathrm{sr}_q\mathrm{SR}_{r-1} &= \mathrm{sign}(a_p)^{p-q} b_q \cdot b_q^{p-q} \cdot b_q^{p-r}\mathrm{SR}_{r-1}(Q,-R) \\
\mathrm{sr}_r &= \varepsilon_{p-q} b_q^{p-r}\mathrm{sr}_r(Q,-R) \\
\overline{\mathrm{sr}}_{q-1} &= \varepsilon_{p-q}\mathrm{sign}(b_q)^{q-r-1} b_q^{p-q+1}\overline{\mathrm{sr}}_{q-1}(Q,-R) \\
\mathrm{SR}_{p-1} &= \mathrm{sign}(a_p)^{p-q+1} Q, \\
\mathrm{SR}_q(Q,-R) &= \mathrm{sign}(b_q)^{q-r-1} Q.
\end{aligned}$$

The fact that $C_{k-1} \in \mathrm{D}[X]$ is proved later in this section (Lemma 1.4.38). □

Before proving the last part of Theorem 1.4.27, we need to prove an analogue of Proposition 1.3.6 for subresultants.

Notation 1.4.33 Define $\mathrm{SU}_j(P,Q)$ (respectively $\mathrm{SV}_j(P,Q)$) as $\det(M_i)$ (respectively $\det(N_i)$), where M_i (respectively N_i) is the square matrix obtained by taking the first $p+q-2j-1$ columns of $\mathrm{SH}_j(P,Q)$ and with last column equal to $(X^{q-1-j}, \ldots, X, 1, 0, \ldots, 0)^t$ (respectively $(0, \ldots, 0, 1, X, \ldots,$
$X^{p-1-j})^t$). Note that if $P, Q \in \mathrm{D}[X]$, then $\mathrm{SU}_j(P,Q), \mathrm{SV}_j(P,Q) \in \mathrm{D}[X]$.

Proposition 1.4.34 *Let $j \leq q$. Then,*

a) $\deg(\mathrm{SU}_{j-1}(P,Q)) \leq q-j, \deg(\mathrm{SV}_{j-1}(P,Q)) \leq p-j,$

$$\mathrm{SR}_j(P,Q) = \mathrm{SU}_j(P,Q)P + \mathrm{SV}_j(P,Q)Q.$$

b) If $\mathrm{SR}_j(P,Q)$ *is not* 0 *and if* U *and* V *are such that*

$$UP + VQ = \mathrm{SR}_j(P,Q),$$

$\deg(U) \leq q-j-1$, *and* $\deg(V) \leq p-j-1$, *then* $U = \mathrm{SU}_j(P,Q)$ *and* $V = \mathrm{SV}_j(P,Q)$.

c) If $\mathrm{SR}_j(P,Q)$ *is non-defective, then*

$$\deg(\mathrm{SU}_{j-1}(P,Q)) = q-j, \deg(\mathrm{SV}_{j-1}(P,Q)) = p-j,$$

and $\mathrm{lcof}(\mathrm{SV}_{j-1}(P,Q)) = a_p \mathrm{sr}_j(P,Q)$.

Proof: a) The conditions

$$\deg(\mathrm{SU}_{j-1}(P,Q)) = q-j, \deg(\mathrm{SV}_{j-1}(P,Q)) = p-j$$

follow from the definitions of $\mathrm{SU}_{j-1}(P,Q)$ and $\mathrm{SV}_{j-1}(P,Q)$. By Lemma 1.4.22, $\mathrm{SR}_j(P,Q) = \det(\mathrm{SH}_j(P,Q)^*)$, where $\mathrm{SH}_j(P,Q)^*$ is the square matrix obtained by taking the first $p+q-2j-1$ columns of $\mathrm{SH}_j(P,Q)$ and with last column equal to

$$(X^{q-1-j}P, \ldots, XP, P, Q, \ldots, X^{p-j-1}Q)^t.$$

Expanding the determinant by its last column, we obtain the claimed identity.

b) Suppose $\deg(U) \leq q-j-1$, $\deg(V) \leq p-j-1$, and $\mathrm{SR}_j(P,Q) = UP + VQ$ so that $(\mathrm{SU}_j(P,Q) - U)P + (\mathrm{SV}_j(P,Q) - V)Q = 0$. If $\mathrm{SU}_j(P,Q) - U$ is not 0, then $\mathrm{SV}_j(P,Q) - V$ cannot be 0, and $\deg(\gcd(P,Q)) > j$. But this is impossible since $\mathrm{SR}_j(P,Q)$ is a non-zero polynomial of degree $\leq j$ belonging to the ideal generated by P and Q.

c) Since $\mathrm{SR}_j(P,Q)$ is non-defective, it follows that $\mathrm{sr}_j(P,Q) \neq 0$. By considering the determinant of the matrix $\mathrm{SH}_{j-1}(P,Q)^*$, it is clear that the coefficient of X^{p-j} in $\mathrm{SV}_{j-1}(P,Q)$ is $a_p \mathrm{sr}_j(P,Q)$. Moreover, $\deg(\mathrm{SV}_{j-1}) = p-j$ and $\deg(\mathrm{SU}_{j-1}(P,Q)) = q-j$.

□

We omit P and Q in the notation in the next paragraphs. For SR_{i-1} non-zero of degree j, we define

$$B_{j,i} = \begin{pmatrix} \mathrm{SU}_{i-1} & \mathrm{SV}_{i-1} \\ \mathrm{SU}_{j-1} & \mathrm{SV}_{j-1} \end{pmatrix},$$

where $\mathrm{SU}_{i-1}, \mathrm{SV}_{i-1}, \mathrm{SU}_{j-1}, \mathrm{SV}_{j-1} \in \mathrm{D}[X]$ are the polynomials of the $(i-1)$-th and $(j-1)$-th relations of Proposition 1.4.34, whence

$$\begin{pmatrix} \mathrm{SR}_{i-1} \\ \mathrm{SR}_{j-1} \end{pmatrix} = B_{j,i} \cdot \begin{pmatrix} P \\ Q \end{pmatrix}. \tag{1.8}$$

Lemma 1.4.35 *If* SR_{i-1} *is non-zero of degree* j*, then*

$$\det(B_{j,i}) = \mathrm{sr}_j \overline{\mathrm{sr}}_{i-1}.$$

Proof: Eliminating Q from the system (1.8), we have

$$(\mathrm{SU}_{i-1}\mathrm{SV}_{j-1} - \mathrm{SU}_{j-1}\mathrm{SV}_{i-1})P = \mathrm{SV}_{j-1}\mathrm{SR}_{i-1} - \mathrm{SV}_{i-1}\mathrm{SR}_{j-1}. \tag{1.9}$$

Since $\deg(SR_{i-1}) = j$, $\deg(SR_j) = j$ by the first part of the Structure Theorem 1.4.27, which is already proved, and $\deg(\mathrm{SV}_{j-1}) = p - j$. Using $\deg(SR_{j-1}) \leq j-1$ and $\deg(\mathrm{SV}_{i-1}) \leq p - i < p - j$, we see that the right hand side of equation (1.9) has degree p . The leading coefficient of SV_{j-1} is $a_p \mathrm{sr}_j$ by Proposition 1.4.34. Hence $\mathrm{SU}_{i-1}\mathrm{SV}_{j-1} - \mathrm{SU}_{j-1}\mathrm{SV}_{i-1} = \mathrm{sr}_j \overline{\mathrm{sr}}_{i-1} \neq 0$. □

Corollary 1.4.36 *If* SR_{i-1} *is non-zero of degree* j*, then*

$$B_{j,i}^{-1} = \frac{1}{\mathrm{sr}_j \overline{\mathrm{sr}}_{i-1}} \begin{pmatrix} \mathrm{SV}_{j-1} & -\mathrm{SV}_{i-1} \\ -\mathrm{SU}_{j-1} & \mathrm{SU}_{i-1} \end{pmatrix},$$

whence $\mathrm{sr}_j \overline{\mathrm{sr}}_{i-1} B_{j,i}^{-1} \in \mathrm{D}[X]$.

Now we study the transition between two consecutive couples of signed subresultant polynomials $\mathrm{SR}_{i-1}, \mathrm{SR}_{j-1}$ and $\mathrm{SR}_{j-1}, \mathrm{SR}_{k-1}$, where SR_{i-1} is of degree j, SR_{j-1} is of degree k, and $0 \leq k < j \leq p$.

The **signed subresultant transition matrix** is

$$T_j = \begin{pmatrix} 0 & 1 \\ -\frac{\mathrm{sr}_k \overline{\mathrm{sr}}_{j-1}}{\mathrm{sr}_j \overline{\mathrm{sr}}_{i-1}} & \frac{C_{k-1}}{\mathrm{sr}_j \overline{\mathrm{sr}}_{i-1}} \end{pmatrix} \in \mathrm{K}[X]^{2\times 2},$$

so that

$$\mathrm{SR}_{k-1} = -\frac{\mathrm{sr}_k\overline{\mathrm{sr}}_{j-1}}{\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}}\mathrm{SR}_{i-1} + \frac{C_{k-1}}{\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}}\mathrm{SR}_{j-1} \tag{1.10}$$

and

$$\begin{pmatrix} \mathrm{SR}_{j-1} \\ \mathrm{SR}_{k-1} \end{pmatrix} = T_j \cdot \begin{pmatrix} \mathrm{SR}_{i-1} \\ \mathrm{SR}_{j-1} \end{pmatrix} \tag{1.11}$$

by the Structure Theorem 1.4.27.

Lemma 1.4.37 *If* SR_{i-1} *is non-zero of degree* j *and* SR_{j-1} *is non-zero of degree* k*, then*

$$B_{k,j} = T_j B_{j,i}.$$

Proof: Let

$$T_j B_{j,i} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

A simple degree calculation shows that $\deg(A) \leq q-j, \deg(B) \leq p-j, \deg(C) = q-k$, and $\deg(D) = p-k$. From equations (1.11) and (1.8) we see that

$$\mathrm{SR}_{j-1} = AP + BQ$$
$$\mathrm{SR}_{k-1} = CP + DQ.$$

The conclusion then follows from the uniqueness asserted in Proposition 1.4.34 b). □

We can now prove the fact that C_{k-1} is in $\mathrm{D}[X]$ as is claimed in the Structure Theorem 1.4.27.

Lemma 1.4.38 *If* SR_{i-1} *is non-zero of degree* j *and* SR_{j-1} *is non-zero of degree* k*, then* $C_{k-1} \in \mathrm{D}[X]$.

Proof: From Lemma 1.4.37, we see that $T_j = B_{k,j}B_{j,i}^{-1}$, which together with the definition of $B_{k,j}$ and Corollary 1.4.36 shows that

$$\frac{C_{k-1}}{\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}} = \frac{1}{\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}} \cdot (-\mathrm{SU}_{k-1} \cdot \mathrm{SV}_{i-1} + \mathrm{SV}_{k-1} \cdot \mathrm{SU}_{i-1}),$$

whence $C_{k-1} = \mathrm{SU}_{k-1} \cdot \mathrm{SV}_{i-1} - \mathrm{SV}_{k-1} \cdot \mathrm{SU}_{i-1} \in \mathrm{D}[X]$. □

The following proposition elaborates more on the proportionality between SR_{j-1} and SR_k.

Proposition 1.4.39 *Let $j \leq q$ with SR_j non-defective and $\deg(\mathrm{SR}_{j-1}) = k \leq j-1$. Define*

$$\overline{\mathrm{SR}}_{j-2} = -\frac{\overline{\mathrm{sr}}_{j-1} \cdot \mathrm{SR}_{j-1}}{\mathrm{sr}_j},$$

$$\overline{\mathrm{SR}}_{j-\delta-1} = (-1)^{\delta} \frac{\overline{\mathrm{sr}}_{j-1} \cdot \overline{\mathrm{SR}}_{j-\delta}}{\mathrm{sr}_j}, \text{for } \delta = 2, \ldots, j-k-1.$$

Then all of these polynomials are in $\mathrm{D}[X]$ and $\mathrm{SR}_k = \overline{\mathrm{SR}}_k$.

Proof: Add the $j-k-1-\delta$ polynomials $X^{k+\delta+1}, \ldots, X^j$ to $\mathcal{H}_{j-1}$ to obtain $\overline{\mathcal{H}}_{j-1-\delta}$. It is easy to see that the polynomial determinant of $\overline{\mathcal{H}}_{j-1-\delta}$ is $\overline{\mathrm{SR}}_{j-1-\delta}$. □

Size of Remainders and Subresultants Observe, comparing the following example with Example 1.3.14, that the bitsizes of coefficients in the signed subresultant sequence can be much smaller than in the signed remainder sequence.

Example 1.4.40 We consider, as in Example 1.3.14,

$$P := 9X^{13} - 18X^{11} - 33X^{10} + 102X^8 + 7X^7 - 36X^6$$
$$-122X^5 + 49X^4 + 93X^3 - 42X^2 - 18X + 9.$$

The subresultant coefficients of P and P' for j from 11 to 5 are:

$$37908$$
$$-72098829$$
$$-666229317948$$
$$-1663522740400320$$
$$-2181968897553243072$$
$$-151645911413926622112$$
$$-165117711302736225120,$$

the remaining subresultants being 0.

The difference in bitsizes of coefficients between signed remainder and signed subresultant sequences observed in Examples 1.3.14 and 1.4.40 is a general fact (see [1]).

1.4.5 Subresultant Computation

We now describe an algorithm for computing the subresultant sequence, based upon the preceding results.

The **principal signed subresultant polynomial sequence** is the sequence of signed subresultant polynomials SR_{p-d-1} where d is the degree of a polynomial in the signed remainder sequence.

Algorithm 1.4.41 (Signed Subresultant)

Input: *two univariate polynomials* $P = a_pX^p + \ldots + a_0$ *and* $Q = b_qX^q + \ldots + b_0$ *with coefficients* D *of respective degrees* p *and* q, $p > q$.

Output: *the sequence of principal signed subresultant polynomials.*

Procedure:

Initialization : $\mathrm{SR}_p := \mathrm{sign}(a_p^{p-q-1})P$, $\mathrm{sr}_p = \overline{\mathrm{sr}}_p := \mathrm{sign}(a_p^{p-q})$, $\mathrm{SR}_{p-1} := \mathrm{sign}(a_p^{p-q+1})Q$, $\overline{\mathrm{sr}}_{p-1} := \mathrm{sign}(a_p^{p-q+1})b_q$, $i := p+1$, $j := p$.

While $\mathrm{SR}_{j-1} \neq 0$,

$k := \deg(\mathrm{SR}_{j-1})$,

If $k = j-1$,

$\mathrm{sr}_{j-1} := \overline{\mathrm{sr}}_{j-1}$,

$\mathrm{SR}_{k-1} := -\mathrm{Rem}(\mathrm{sr}_{j-1}^2\mathrm{SR}_{i-1}, \mathrm{SR}_{j-1})/(\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1})$.

If $k < j-1$,

Computation of sr_k*:*

For δ *from* 1 *to* $j-k-1$*:*

$\overline{\mathrm{sr}}_{j-\delta-1} := (-1)^{\delta}(\overline{\mathrm{sr}}_{j-1} \cdot \overline{\mathrm{sr}}_{j-\delta})/\mathrm{sr}_j$,

$\mathrm{sr}_k := \overline{\mathrm{sr}}_k$.

Computation of SR_{k-1}*:*

$$\mathrm{SR}_{k-1} := -\mathrm{Rem}(\overline{\mathrm{sr}}_{j-1}\mathrm{sr}_k\mathrm{SR}_{i-1}, \mathrm{SR}_{j-1})/(\mathrm{sr}_j\overline{\mathrm{sr}}_{i-1}),$$
$$\overline{\mathrm{sr}}_{k-1} := \mathrm{lcof}(\mathrm{SR}_{k-1}).$$
$$i := j, j := k.$$

Proof of correctness: The correctness of the algorithm follows from Theorem 1.4.27. □

1.4.6 Signed Subresultant Coefficients and Cauchy Index

We indicate how to compute the Cauchy index by using only the signed subresultant coefficients. We need a definition:

Notation 1.4.42 Let $s = s_p, \ldots, s_0$ be a finite list of elements in an ordered field K such that $s_p \neq 0$, $s_{p-1} = \ldots = s_{q+1} = 0$, and $s_q \neq 0$. Let $s' = s_q, \ldots, s_0$ (if $q = 0$, s' is the empty list). We define inductively

$$D(s) = \begin{cases} 0 & \text{if } s' = \emptyset, \\ D(s') + \varepsilon_{p-q}\mathrm{sign}(s_p s_q) & \text{if } p-q \text{ odd}, \\ D(s') & \text{if } p-q \text{ even}. \end{cases}$$

where $\varepsilon_{p-q} = (-1)^{(p-q)(p-q-1)/2}$.

Note that when all elements of s are non-zero, $D(s)$ is the difference between the number of sign permanencies and the number of sign changes in $s_p, \ldots, s_0$. Note also that when s is the sequence of leading coefficients of polynomials $\mathcal{P} = P_p, \ldots, P_0$ with $\deg(P_i) = i$, then

$$D(s) = \mathrm{V}(\mathcal{P}; -\infty, +\infty)$$

(see Notation 1.1.1).

Let P and Q be two polynomials with:

$$P = a_p X^p + a_{p-1}X^{p-1} + \cdots + a_0$$
$$Q = b_{p-1}X^{p-1} + \cdots + b_0,$$

$\deg(P) = p, \deg(Q) = q \leq p-1$.

We denote by $\mathrm{sr}(P,Q)$ the sequence of $\mathrm{sr}_j(P,Q)$, $j = p, \ldots, 0$. Note that $\mathrm{sr}_j(P,Q) \neq 0$ if only if $\mathrm{SR}_j(P,Q)$ is non-defective.

Theorem 1.4.43

$$D(\mathrm{sr}(P,Q)) = \mathrm{Ind}\left(\frac{Q}{P}\right).$$

Note that in the non-defective case ($\deg(\mathrm{SR}_j(P,Q)) = j$ for every $j = p,\dots,0$) Theorem 1.4.43 is a consequence of Theorem 1.3.7 (with $a = -\infty, b = +\infty$) and Corollary 1.4.30.

Proof of Theorem 1.4.43: The proof of the theorem will use the following two lemmas.

Lemma 1.4.44

$$\mathrm{Ind}\left(\frac{Q}{P}\right) = \begin{cases} \mathrm{Ind}\left(\dfrac{-R}{Q}\right) + \mathrm{sign}(a_p b_q) & \text{if } p-q \text{ is odd,} \\ \mathrm{Ind}\left(\dfrac{-R}{Q}\right) & \text{if } p-q \text{ is even.} \end{cases}$$

Proof: We can suppose without loss of generality that P and Q are coprime. The claim is an immediate consequence of Lemma 1.3.9. □

Lemma 1.4.45

$$D(\mathrm{sr}(P,Q)) = \begin{cases} D(\mathrm{sr}(Q,-R)) + \mathrm{sign}(a_p b_q) & \text{if } p-q \text{ is odd,} \\ D(\mathrm{sr}(Q,-R)) & \text{if } p-q \text{ is even.} \end{cases}$$

Proof: By Proposition 1.4.32, for $j \le r$,

$$\mathrm{SR}_j(P,Q) = \varepsilon_{p-q} b_q^{p-r} \mathrm{SR}_j(Q,-R).$$

Using the convention in Notation 1.4.26 and the definition of $\mathrm{SR}_q(P,Q)$,

$$\mathrm{SR}_q(Q,-R) = \mathrm{sign}(b_q^{q-r-1})Q,$$
$$\mathrm{SR}_q(P,Q) = \varepsilon_{p-q} b_q^{p-q-1} Q.$$

Thus,

$$\mathrm{SR}_q(P,Q) = \varepsilon_{p-q}\mathrm{sign}(b_q^{q-r-1}) b_q^{p-q-1} \mathrm{SR}_q(Q,-R).$$

So,

$$D(\mathrm{sr}_q(P,Q),\dots,\mathrm{sr}_0(P,Q)) = D(\mathrm{sr}(Q,-R)).$$

Noticing that $\varepsilon_{p-q}\mathrm{sr}_p(P,Q)\mathrm{sr}_q(P,Q) = (\mathrm{sign}(a_p)b_q)^{p-q}$, the conclusion follows by definition of D. □

The proof of the theorem proceeds by induction on the number n of elements with distinct degrees in the signed subresultant sequence.

If $n = 2$, Q divides P. We have

$$\mathrm{Ind}\left(\frac{Q}{P}\right) = \begin{cases} \mathrm{sign}(a_p b_q) & \text{if } p-q \text{ is odd,} \\ 0 & \text{if } p-q \text{ is even,} \end{cases}$$

by Lemma 1.4.44 and

$$D(\mathrm{sr}(P,Q)) = \begin{cases} \mathrm{sign}(a_p b_q) & \text{if } p-q \text{ is odd,} \\ 0 & \text{if } p-q \text{ is even,} \end{cases}$$

by Lemma 1.4.45.

Let us suppose that the theorem holds for $n-1$ and consider P and Q such that their signed subresultant sequence has n elements with distinct degrees. The signed subresultant sequence of Q and $-R$ has $n-1$ elements with distinct degrees. By the induction hypothesis,

$$D(\mathrm{sr}(Q,-R)) = \mathrm{Ind}\left(\frac{-R}{Q}\right).$$

So, by Lemma 1.4.44 and Lemma 1.4.45 ,

$$D(\mathrm{sr}(P,Q)) = \mathrm{Ind}\left(\frac{Q}{P}\right).$$

□

Denoting as before

$$\begin{gathered} \mathrm{SQ}(Q,P) = \\ \#(\{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) > 0\}) - \\ \#(\{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) < 0\}). \end{gathered}$$

Corollary 1.4.46 *Let P and Q be polynomials in $\mathrm{D}[X]$ and R the remainder of $P'Q$ and P. Then $D(\mathrm{sr}(P,R)) = \mathrm{SQ}(Q,P)$.*

Proof: Apply Theorem 1.4.43 and Proposition 1.3.4, since

$$\mathrm{Ind}\left(\frac{P'Q}{P}\right) = \mathrm{Ind}\left(\frac{R}{P}\right)$$

by Remark 1.3.2. □

Corollary 1.4.47 *Let P be a polynomial in* $\mathrm{D}[X]$. *Then $D(\mathrm{sr}(P, P'))$ is the number of roots of P in* R.

Algorithm 1.4.48 (Univariate Sturm-query)

Input: *a non-zero univariate polynomial P and a univariate polynomial Q, both with coefficients in* D *of respective degree p and q.*

Output: *the Sturm-query* $\mathrm{SQ}(Q, P)$.

Procedure:

if $\deg(Q) = 0$, $Q = b_0$, *compute the sequence* $\mathrm{sr}(P, P')$ *of signed subresultant coefficient of P and P' using Algorithm 1.4.41 (Signed Subresultant), and compute $D(\mathrm{sr}(P, P'))$ (Definition 1.4.42). Output*

$$\begin{cases} D(\mathrm{sr}(P, P')) & \text{if } b_0 > 0 \\ -D(\mathrm{sr}(P, P')) & \text{if } b_0 < 0. \end{cases}$$

if $\deg(Q) = 1$, $Q = b_1 X + b_0$, *compute* $\bar{R} := \mathrm{sign}(b_1)(pb_1 P - P'Q)$, *the sequence* $\mathrm{sr}(P, \bar{R})$ *of signed subresultant coefficient of P and $\bar{R}$, use Algorithm 1.4.41 (Signed Subresultant) to compute $D(\mathrm{sr}(P, \bar{R}))$ (Definition 1.4.42).*

if $\deg(Q) > 1$ *use Algorithm 1.4.41 (Signed Subresultant) to compute the sequence* $\mathrm{sr}(-P'Q, P)$ *of signed subresultant coefficient of $-P'Q$ and Q, and compute $D(\mathrm{sr}(-P'Q, P))$ (Definition 1.4.42). Output*

$$\begin{cases} D(\mathrm{sr}(-P'Q, P)) + \mathrm{sign}(b_q) & \text{if } q-1 \text{ is odd,} \\ D(\mathrm{sr}(-P'Q, P)) & \text{if } q-1 \text{ is even.} \end{cases}$$

Proof of correctness: The correctness follows from Corollary 1.4.46 and Lemma 1.4.45. □

Chapter 2

Quantifier elimination

2.1 Tarski-Seidenberg theorem

2.1.1 Sign Determination

We consider a $P \in \mathrm{R}[X]$ with P not identically zero, $\mathcal{Q}$ a finite subset of $\mathrm{R}[X]$, and the finite set $Z = \mathrm{Z}(P, \mathrm{R}) = \{x \in \mathrm{R} \mid P(x) = 0\}$.

We will give an expression for the number of elements of Z at which $\mathcal{Q}$ satisfies a given sign condition σ.

Let σ be a sign condition on $\mathcal{Q}$ i.e. an element of $\{0, 1, -1\}^{\mathcal{Q}}$. The **realization of the sign condition σ over Z** is

$$\mathcal{R}(\sigma, P = 0) = \{x \in \mathrm{R} \mid P(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} \mathrm{sign}(Q(x)) = \sigma(Q)\}.$$

Its cardinality is denoted $c(\sigma, P = 0)$. We denote

$$\begin{aligned}
\mathcal{R}(Q = 0, P = 0) &= \{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) = 0\}, \\
\mathcal{R}(Q > 0, P = 0) &= \{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) > 0\}, \\
\mathcal{R}(Q < 0, P = 0) &= \{x \in \mathrm{R} \mid P(x) = 0 \wedge Q(x) < 0\},
\end{aligned}$$

and $c(Q = 0, P = 0), c(Q > 0, P = 0), c(Q < 0, P = 0)$ are the cardinalities of the corresponding sets.

Given $\alpha \in \{0, 1, 2\}^{\mathcal{Q}}$, we write $\mathcal{Q}^\alpha$ for $\prod_{Q \in \mathcal{Q}} Q^{\alpha(Q)}$. When $\mathcal{R}(\sigma, P = 0) \neq \emptyset$, the sign of $\mathcal{Q}^\alpha$ is fixed on $\mathcal{R}(\sigma, P = 0)$ and is equal to

$\prod_{Q\in\mathcal{Q}} \sigma(Q)^{\alpha(Q)}$, with the convention that $0^0 = 1$. Hence, we define the sign of $\mathcal{Q}^\alpha$ on σ, $\mathrm{sign}(\mathcal{Q}^\alpha, \sigma)$, to be $\prod_{Q\in\mathcal{Q}} \sigma(Q)^{\alpha(Q)}$, whether or not $\mathcal{R}(\sigma, P = 0)$ is empty.

We number the elements of $\mathcal{Q}$ so that $\mathcal{Q} = \{Q_1, \ldots, Q_s\}$.

The **lexicographical ordering** on $\{0, 1, 2\}^{\mathcal{Q}}$ is defined by $\alpha <_{\mathrm{lex}} \beta$ if and only if

$$\begin{cases} \alpha(Q_s) < \beta(Q_s) \\ \vee\ (\alpha(Q_s) = \beta(Q_s)\ \wedge\ \alpha(Q_{s-1}) < \beta(Q_{s-1})) \\ \ldots \\ \vee\ (\alpha(Q_s) = \beta(Q_s) \wedge\ \ldots \wedge\ \alpha(Q_3) = \beta(Q_3)\ \wedge\ \alpha(Q_2) < \beta(Q_2)), \\ \vee\ (\alpha(Q_s) = \beta(Q_s) \wedge\ \ldots \wedge\ \alpha(Q_2) = \beta(Q_2) \wedge\ \alpha(Q_1) < \beta(Q_1)). \end{cases}$$

Given a list of elements $A = \alpha_1, \ldots, \alpha_m$ of $\{0, 1, 2\}^{\mathcal{Q}}$ ordered lexicographically by

$$\alpha_1 <_{\mathrm{lex}} \ldots <_{\mathrm{lex}} \alpha_m,$$

we write $\mathcal{Q}^A$ for $\mathcal{Q}^{\alpha_1}, \ldots, \mathcal{Q}^{\alpha_m}$ and $\mathrm{SQ}(\mathcal{Q}^A, P)$ for

$$\mathrm{SQ}(\mathcal{Q}^{\alpha_1}, P), \ldots, \mathrm{SQ}(\mathcal{Q}^{\alpha_m}, P).$$

The **lexicographical ordering** on $\{0, 1, -1\}^{\mathcal{Q}}$ is defined by $\sigma <_{\mathrm{lex}} \tau$ if and only if

$$\begin{cases} \sigma(Q_s) \prec \tau(Q_s) \\ \vee\ (\sigma(Q_s) = \tau(Q_s)\ \wedge\ \sigma(Q_{s-1}) \prec \tau(Q_{s-1})) \\ \ldots \\ \vee\ (\sigma(Q_s) = \tau(Q_s) \wedge\ \ldots \wedge\ \sigma(Q_3) = \tau(Q_3)\ \wedge\ \sigma(Q_2) \prec \tau(Q_2)), \\ \vee\ (\sigma(Q_s) = \tau(Q_s) \wedge\ \ldots \wedge\ \sigma(Q_2) = \tau(Q_2) \wedge\ \sigma(Q_1) \prec \tau(Q_1)), \end{cases}$$

where $0 \prec 1 \prec -1$.

Given a list of elements $\Sigma = \sigma_1, \ldots, \sigma_n$ of $\{0, 1, -1\}^{\mathcal{Q}}$, with

$$\sigma_1 <_{\mathrm{lex}} \ldots <_{\mathrm{lex}} \sigma_n,$$

we write $\mathcal{R}(\Sigma, P = 0)$ for

$$\mathcal{R}(\sigma_1, P = 0), \ldots, \mathcal{R}(\sigma_n, P = 0)$$

and $c(\Sigma, P = 0)$ for

$$c(\sigma_1, P = 0), \ldots, c(\sigma_n, P = 0).$$

Definition 2.1.1 The **matrix of signs of $\mathcal{Q}^A$ on** Σ is the $m \times n$ matrix $M(\mathcal{Q}^A, \Sigma)$ whose i, j-th entry is $\mathrm{sign}(\mathcal{Q}^{\alpha_i}, \sigma_j)$.

Example 2.1.2 If $\mathcal{Q} = \{Q_1, Q_2\}$ and $A = \{0, 1, 2\}^{\{Q\}}$, $\{Q_1, Q_2\}^A$ is the list $1, Q_1, Q_1^2, Q_2, Q_1Q_2, Q_1^2Q_2, Q_2^2, Q_1Q_2^2, Q_1^2Q_2^2$. Taking for Σ

$$\begin{aligned}
&Q_1 = 0 \wedge Q_2 = 0, Q_1 > 0 \wedge Q_2 = 0, Q_1 < 0 \wedge Q_2 = 0,\\
&Q_1 = 0 \wedge Q_2 > 0, Q_1 > 0 \wedge Q_2 > 0, Q_1 > 0 \wedge Q_2 < 0,\\
&Q_1 = 0 \wedge Q_2 < 0, Q_1 < 0 \wedge Q_2 > 0, Q_1 < 0 \wedge Q_2 < 0,
\end{aligned}$$

the matrix of signs of these nine polynomials on these nine sign conditions is

$$M(\{Q_1, Q_2\}^A, \Sigma) = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\
0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1\\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1\\
0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1\\
0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1\\
0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & -1\\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1\\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1\\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1
\end{bmatrix}.$$

For example, the 5-th row of the matrix reads as follows: the signs of the 5-th polynomial of $\{Q_1, Q_2\}^A$ which is Q_1Q_2 on the 9 sign conditions Σ are

$$[\,0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1 \quad 0 \quad -1 \quad 1\,].$$

Proposition 2.1.3 *If* $\bigcup_{\sigma \in \Sigma} \mathcal{R}(\sigma, P = 0) = Z$ *then*

$$M(\mathcal{Q}^A, \Sigma) \cdot c(\Sigma, P = 0) = \mathrm{SQ}(\mathcal{Q}^A, P).$$

Proof: It is obvious since the $(i, j)-$th entry of $M(\mathcal{Q}^A, \Sigma)$ is the sign of the polynomial $\mathcal{Q}^{\alpha_i}$ of $\mathcal{Q}^A$ on the sign condition σ_j of Σ. □

Note that when $\mathcal{Q} = \{Q\}$, $A = \{0,1,2\}^{\{Q\}}$ and $\Sigma = \{0,1,-1\}^{\{Q\}}$ the conclusion of Proposition 2.1.3 is

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} c(Q=0, P=0) \\ c(Q>0, P=0) \\ c(Q<0, P=0) \end{bmatrix} = \begin{bmatrix} \mathrm{SQ}(1,P) \\ \mathrm{SQ}(Q,P) \\ \mathrm{SQ}(Q^2,P) \end{bmatrix}. \tag{2.1}$$

It follows from Proposition 2.1.3 that when the matrix $M(\mathcal{Q}^A, \Sigma)$ is invertible, we can express $c(\Sigma, P=0)$ in terms of $\mathrm{SQ}(\mathcal{Q}^A, P)$. This is the case when $A = \{0,1,2\}^{\mathcal{Q}}$ and $\Sigma = \{0,1,-1\}^{\mathcal{Q}}$, as we will see now.

Notation 2.1.4 Let M and $M' = [\, m'_{ij} \,]$ be two matrices with respective dimensions $n \times m$ and $n' \times m'$. The matrix $M \otimes M'$ is the $nn' \times mm'$ matrix

$$[\, m'_{ij} M \,].$$

The matrix $M \otimes M'$ is the **tensor product** of M and M'.

Example 2.1.5 If

$$M = M' = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix},$$

$$M \otimes M' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Notice that $M \otimes M'$ coincides with the matrix of signs of $A = \{0,1,2\}^{\{Q_1,Q_2\}}$ on $\Sigma = \{0,1,-1\}^{\{Q_1,Q_2\}}$.

Notation 2.1.6 Let M_s be the $3^s \times 3^s$ matrix defined inductively by

$$M_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

and

$$M_{t+1} = M_t \otimes M_1.$$

Exercise 2.1.7 *Prove that M_s is invertible using induction on s.*

Proposition 2.1.8 *Let $\mathcal{Q}$ be a set of polynomials with s elements, $A = \{0,1,2\}^{\mathcal{Q}}$ and $\Sigma = \{0,1,-1\}^{\mathcal{Q}}$, ordered lexicographically. Then*

$$M(\mathcal{Q}^A, \Sigma) = M_s.$$

Proof: The proof is by induction on s. If s=1, the claim is Equation (2.1). If the claim holds for s, it holds also for $s+1$ given the definitions of M_{s+1}, $M(\mathcal{Q}^A, \Sigma)$, and the orderings on $A = \{0,1,2\}^{\mathcal{Q}}$ and $\Sigma = \{0,1,-1\}^{\mathcal{Q}}$. □

So, Proposition 2.1.3 and Proposition 2.1.8 imply

Corollary 2.1.9

$$M_s \cdot c(\Sigma, P = 0) = \mathrm{SQ}(\mathcal{Q}^A, P).$$

We have all the ingredients needed to decide whether a subset of R defined by a sign condition is empty or not, with the following two lemmas.

Lemma 2.1.10 *Consider the finite set $Z = \mathrm{Z}(P, \mathrm{R})$ and a sign condition σ on $\mathcal{Q}$. Whether or not $\mathcal{R}(\sigma, P = 0) = \emptyset$ is determined by the degrees of the polynomials in the signed pseudo-remainder sequences of $P, P'\mathcal{Q}^\alpha$ and the signs of their leading coefficients for all $\alpha \in A = \{0,1,2\}^{\mathcal{Q}}$.*

Proof: For each $\alpha \in \{0,1,2\}^{\mathcal{Q}}$, the degrees and the signs of the leading coefficients of all of the polynomials of the signed pseudo-remainder sequences $\mathrm{S}(P, P'\mathcal{Q}^{\alpha})$ clearly determine the number of sign changes of $\mathrm{S}(P, P'\mathcal{Q}^{\alpha})$ at $-\infty$ and $+\infty$, i.e. $\mathrm{V}(\mathrm{S}(P, P'\mathcal{Q}^{\alpha}); -\infty)$ and $\mathrm{V}(\mathrm{S}(P, P'\mathcal{Q}^{\alpha}); +\infty)$, and their difference is $\mathrm{SQ}(\mathcal{Q}^{\alpha}, P)$ by Theorem 1.3.10. Using Propositions 2.1.8, Proposition 2.1.3, and Corollary 2.1.9

$$M_s^{-1} \cdot \mathrm{SQ}(\mathcal{Q}^A, P) = c(\Sigma, P = 0).$$

Denoting the row of M_s^{-1} that corresponds to the row of σ in $c(\Sigma, P = 0)$ by r_σ, we see that $r_\sigma \cdot \mathrm{SQ}(\mathcal{Q}^A, P) = c(\sigma, P = 0)$. Finally,

$$\mathcal{R}(\sigma, P = 0) = \{x \in \mathrm{R} | P(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} \mathrm{sign}(Q(x)) = \sigma(Q)\}$$

is non-empty if and only if $c(\sigma, P = 0) > 0$. □

Lemma 2.1.11 *Let σ be a sign condition on $\mathcal{Q}$. Whether or not $\mathcal{R}(\sigma) = \emptyset$ is determined by the degrees and the signs of the leading coefficients of the polynomials in $\mathrm{V}(\mathrm{S}(C, C'))$ (with $C = \prod_{Q \in \mathcal{Q}} Q$) and the signs of the leading coefficients of the polynomials in $\mathrm{V}(\mathrm{S}(C', C''\mathcal{Q}^{\alpha}))$ for all $\alpha \in A = \{0,1,2\}^{\mathcal{Q}}$.*

Proof: Recall (Theorem 1.3.11) that the number of roots of C is determined by the signs of the leading coefficients of $\mathrm{V}(\mathrm{S}(C, C'))$.

If C has no roots, then each $Q \in \mathcal{Q}$ has constant sign which is the same as the sign of its leading coefficient.

If C has one root, then the possible sign conditions on $\mathcal{Q}$ are determined by the sign conditions on $\mathcal{Q}$ at $+\infty$ and at $-\infty$.

If C has at least two roots, then all intervals between two roots of C contain a root of C' and thus all sign conditions on $\mathcal{Q}$ are determined by the sign conditions on $\mathcal{Q}$ at $+\infty$ and at $-\infty$ and by the sign conditions on $\mathcal{Q}$ at the roots of C' this is covered by Lemma 2.1.10.

□

2.1.2 Projection of semi-algebraic sets

Let R be a real closed field. If $\mathcal{P}$ is a finite subset of $\mathrm{R}[X_1, \ldots, X_k]$, we write the **set of zeros** of $\mathcal{P}$ in R^k as

$$\mathrm{Z}(\mathcal{P}, \mathrm{R}^k) = \{x \in \mathrm{R}^k \mid \bigwedge_{P \in \mathcal{P}} P(x) = 0\}.$$

These are the **algebraic sets** of $\mathrm{R}^k = \mathrm{Z}(\{0\}, \mathrm{R}^k)$.

The smallest family of sets of R^k that contains the algebraic sets and is closed under the boolean operations (complementation, finite unions, and finite intersections) is the **constructible sets**.

We define the **semi-algebraic sets** of R^k as the smallest family of sets in R^k that contains the algebraic sets as well as sets defined by polynomial **inequalities** i.e. sets of the form $\{x \in \mathrm{R}^k | P(x) > 0\}$ for some polynomial $P \in \mathrm{R}[X_1, \ldots, X_k]$, and which is also closed under the boolean operations (complementation, finite unions, and finite intersections). If the coefficients of the polynomials defining S lie in a subring $\mathrm{D} \subset \mathrm{R}$, we say that the semi-algebraic set S is **defined over** D.

It is obvious that any semi-algebraic set in R^k is the finite union of sets of the form $\{x \in \mathrm{R}^k | P(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} Q(x) > 0\}$. These are the **basic semi-algebraic sets**.

Some terminology from logic is useful for the study of semi-algebraic sets.

We define the language of ordered fields by describing the formulas of this language. The formulas are built starting with atoms, which are polynomial equations and inequalities. A formula is written using atoms together with the logical connectives "and", "or", and "negation" ($\wedge$, $\vee$, and $\neg$) and the existential and universal quantifiers ($\exists, \forall$). A formula has free variables, i.e. non-quantified variables, and bounded variables, i.e. quantified variables. More precisely, let D be a subring of R. We define the **language of ordered fields with coefficients in** D as follows. An **atom** is $P = 0$ or $P > 0$, where P is a polynomial in $\mathrm{D}[X_1, \ldots, X_k]$. We define simultaneously the **formulas** and the set $\mathrm{Free}(\Phi)$ of **free variables of a formula** Φ as follows

- an atom $P = 0$ or $P > 0$, where P is a polynomial in $\mathrm{D}[X_1, \ldots, X_k]$ is a formula with free variables $\{X_1, \ldots, X_k\}$,

- if Φ_1 and Φ_2 are formulas, then $\Phi_1 \wedge \Phi_2$ and $\Phi_1 \vee \Phi_2$ are formulas with $\mathrm{Free}(\Phi_1 \wedge \Phi_2) = \mathrm{Free}(\Phi_1 \vee \Phi_2) = \mathrm{Free}(\Phi_1) \cup \mathrm{Free}(\Phi_2)$,

- if Φ is a formula, then $\neg(\Phi)$ is a formula with $\mathrm{Free}(\neg(\Phi)) = \mathrm{Free}(\Phi)$,

- if Φ is a formula and $X \in \mathrm{Free}(\Phi)$, then $(\exists X)\ \Phi$ and $(\forall X)\ \Phi$ are formulas with $\mathrm{Free}((\exists X)\ \Phi) = \mathrm{Free}((\forall X)\ \Phi) = \mathrm{Free}(\Phi) \setminus \{X\}$.

If Φ and Ψ are formulas, $\Phi \Rightarrow \Psi$ is the formula $\neg(\Phi) \vee \Psi$.

A **quantifier free formula** is a formula in which no quantifier appears, neither $\exists$ nor $\forall$.

The R-**realization of a formula** $\Phi(Y_1, \ldots, Y_k)$ with free variables $\{Y_1, \ldots, Y_k\}$ is denoted by $\mathcal{R}(\Phi(Y_1, \ldots, y_k), \mathrm{R}^k)$. It is the subset of elements $(y_1, \ldots, y_k)$ of R^k such that $\Phi(y_1, \ldots, y_k)$ is true:

$$\mathcal{R}(\Phi(Y_1, \ldots, y_k), \mathrm{R}^k) = \{(y_1, \ldots, y_k) \in \mathrm{R}^k \mid \Phi(y_1, \ldots, y_k)\}.$$

Two formulas Φ and Ψ such that $\mathrm{Free}(\Phi) = \mathrm{Free}(\Psi) = \{Y_1, \ldots, Y_k\}$ are R-**equivalent** if $\mathcal{R}(\Phi(Y_1, \ldots, Y_k), \mathrm{R}^k) = \mathcal{R}(\Psi(Y_1, \ldots, Y_k), \mathrm{R}^k)$. If there is no ambiguity, we simply write $\mathcal{R}(\Phi(Y_1, \ldots, Y_k))$ for $\mathcal{R}(\Phi(Y_1, \ldots, Y_k), \mathrm{R}^k)$ and talk about realization and equivalence.

It is clear that a set is semi-algebraic if and only if it can be represented as the realization of a quantifier free formula. It is also easy to see that any formula in the language of fields with coefficients in D is R-equivalent to

$$\Phi(Y) = (\mathrm{Q_1 X_1}) \ldots (\mathrm{Q_m X_m})\ \mathcal{B}(\mathrm{X_1}, \ldots, \mathrm{X_m}, \mathrm{Y_1}, \ldots \mathrm{Y_k})$$

where each $\mathrm{Q_i} \in \{\forall, \exists\}$ and $\mathcal{B}$ is a quantifier free formula involving polynomials in $\mathrm{D}[X_1, \ldots, X_m, Y_1, \ldots Y_k]$. This is called its **prenex normal form** (see Section 10, Chapter 1 of [**?**]). The variables $X_1, \ldots, X_m$ are called **bound variables** . If a formula has no free variables, then it is called a **sentence**, and it is either true or false in R.

When $P \in \mathrm{R}[X], \mathcal{Q} \subset \mathrm{R}[X]$, the realization $\mathcal{R}(\sigma, \mathrm{Z}(P, \mathrm{R}))$ of a sign condition σ on $\mathcal{Q}$ over $\mathrm{Z}(P, \mathrm{R})$ (definition page 61) is exactly the realization of the quantifier free formula

$$P(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} \mathrm{sign}(Q(x)) = \sigma(Q)$$

and is a basic semi-algebraic set of the line R.

The goal of this section is to show that the semi-algebraic sets in R^{k+1} are closed under projection if R is a real closed field.

Now that we know how to decide (see Lemmas 2.1.10 and 2.1.11) whether or not a basic semi-algebraic set in R is empty, we can show that the projection from R^{k+1} to R^k of a basic semi-algebraic set is semi-algebraic. We extend our method from the univariate case to the multivariate case by viewing the univariate case parametrically. The basic semi-algebraic set $S \subset \mathrm{R}^{k+1}$ can be described as

$$S = \{x \in \mathrm{R}^{k+1} \mid \bigwedge_{P \in \mathcal{P}} P(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} Q(x) > 0\}$$

with $\mathcal{P}, \mathcal{Q}$ finite subsets of $\mathrm{R}[X_1, \ldots, X_k, X_{k+1}]$, and its projection $\pi(S)$ (forgetting the last coordinate) is

$$\pi(S) = \{y \in \mathrm{R}^k \mid \exists x \in \mathrm{R}\ (\bigwedge_{P \in \mathcal{P}} P_y(x) = 0 \bigwedge_{Q \in \mathcal{Q}} Q_y(x) > 0)\}.$$

For a particular $y \in \mathrm{R}^k$ we can decide, using Lemmas 2.1.10 and 2.1.11, whether or not

$$\exists x \in \mathrm{R}\ (\bigwedge_{P \in \mathcal{P}} P_y(x) = 0 \bigwedge_{Q \in \mathcal{Q}} Q_y(x) > 0)$$

is true.

What is crucial here is to partition the parameter space R^k into finitely many parts so that for all points y in the same part, the set

$$S_y = \{x \in \mathrm{R} \mid \bigwedge_{P \in \mathcal{P}} P_y(x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} Q_y(x) > 0\}$$

is empty or is not empty. It is important too that $\pi(S)$ is the union of those parts where $S_y \neq \emptyset$. In fact, the decision method is the same (is uniform) for all y in any given part. Thus each part is a semi-algebraic set and consequently $\pi(S)$ is semi-algebraic being the union of finitely many semi-algebraic sets.

We have been able to decide whether a basic semi-algebraic set in R is or is not empty using Sturm's theorem and its extension by Sylvester

(Lemmas 2.1.10 and 2.1.11). Extending this method to the parametric situation will yield the desired theorems.

For a specialization of Y to $y = (y_1, \ldots, y_k) \in \mathrm{R}^k$, we write $P_y(X)$ for $P(y_1, \ldots, y_k, X)$. We next study the signed remainder sequence of P_y and Q_y for all possible specialization of Y to $y \in \mathrm{R}^k$. This cannot be done in a completely uniform way, since denominators appear in the euclidean division process. Nevertheless, fixing the degrees of the polynomials in the signed remainder sequence, it is possible to partition the parameter space, R^k, into a finite number of parts so that the signed remainder sequence is uniform in each part.

Example 2.1.12 We consider a general polynomial of degree 4. Dividing by its leading coefficient, it is not a loss of generality to take P to be monic. So let $P = X^4 + \alpha X^3 + \beta X^2 + \gamma X + \delta$. The translation $X \mapsto X - \alpha/4$ kills the term of degree 3, so we can suppose $P = X^4 + aX^2 + bX + c$.

Consider $P = X^4 + aX^2 + bX + c$ and its derivative $P' = 4X^3 + 2aX + b$. Their signed remainder sequence in $\mathbb{Q}(a, b, c)[X]$ is

$$
\begin{gathered}
P = X^4 + aX^2 + bX + c \\
P' = 4X^3 + 2aX + b \\
S_2 = -\mathrm{Rem}(P, P') = -\frac{1}{2}aX^2 - \frac{3}{4}bX - c \\
S_3 = -\mathrm{Rem}(P', S_2) = \frac{\left(8\,ac - 9\,b^2 - 2\,a^3\right)X}{a^2} - \frac{b\left(12\,c + a^2\right)}{a^2} \\
S_4 = -\mathrm{Rem}(S_2, S_3) = \\
= \frac{1}{4}\,\frac{a^2\left(256\,c^3 - 128\,a^2c^2 + 144\,acb^2 - 16\,a^4c - 27\,b^4 - 4\,b^2a^3\right)}{\left(8\,ac - 9\,b^2 - 2\,a^3\right)^2}
\end{gathered}
$$

Note that when (a, b, c) are specialized to values in $\mathbb{C}^3$ for which $a = 0$ or $8ac - 9b^2 - 2a^3 = 0$, the signed remainder sequence of P and P' for these special values is not obtained by specializing a, b, c in the signed remainder sequence in $\mathbb{Q}(a, b, c)[X]$.

In order to take into account all the possible signed remainder sequences that can appear when we specialize the parameters, we introduce the following definitions and notation.

We get rid of denominators appearing in the remainders through the notion of signed pseudo-remainders. Let

$$P = a_p X^p + \ldots + a_0 \in \mathrm{D}[X], Q = b_q X^q + \ldots + b_0 \in \mathrm{D}[X],$$

where D is a subring of R. Note that the only denominators occurring in the euclidean division of P by Q are b_q^i, $i \leq p - q + 1$.

The **signed pseudo-remainder** denoted $\mathrm{Prem}(P, Q)$, is the remainder in the euclidean division of $b_q^d P$ by Q, where d is the smallest even integer greater than or equal to $p - q + 1$. Note that the euclidean division of $b_q^d P$ by Q can be performed in D and that $\mathrm{Prem}(P, Q) \in \mathrm{D}[X]$. The even exponent is useful in Chapter 2 and later when we deal with signs.

Let $Q = b_q X^q + \ldots + b_0 \in \mathrm{D}[X]$. We define for $0 \leq i \leq q$, the **truncation of Q at i** by

$$\mathrm{Tru}_i(Q) = b_i X^i + \ldots + b_0.$$

The **set of truncations** of a polynomial $Q \in \mathrm{D}[Y_1, \ldots, Y_k][X]$ is a finite subset of $\mathrm{D}[Y_1, \ldots, Y_k][X]$ defined by

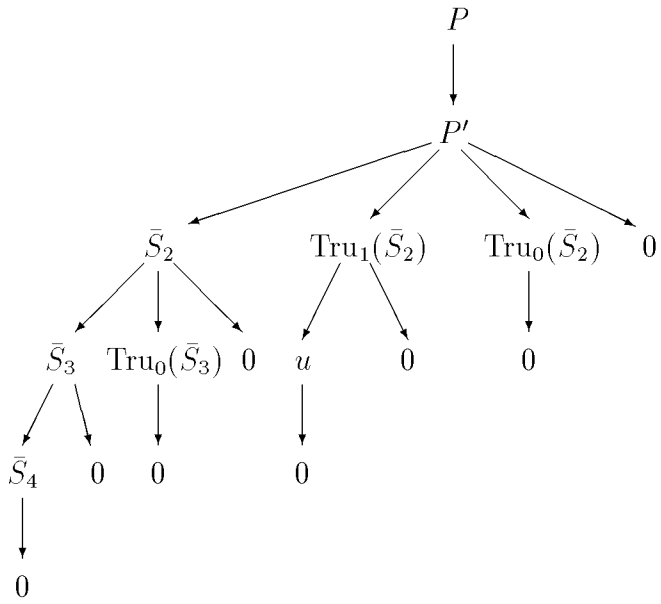
$$\mathrm{Tru}(Q) = \begin{cases} \{Q\} & \text{if } \mathrm{lcof}(Q) \in \mathrm{D} \\ \{Q\} \cup \mathrm{Tru}(\mathrm{Tru}_{\deg(Q)-1}(Q)), & \text{otherwise.} \end{cases}$$

The **tree of possible signed pseudo-remainder sequences of two polynomials** $P, Q \in \mathrm{D}[Y_1, \ldots, Y_k][X]$, denoted $\mathrm{TRems(P, Q)}$, is a tree whose root R contains P. The children of the root contain the elements of the set of truncations of Q. Each node N contains a polynomial $\mathrm{Pol}(N) \in \mathrm{D}[Y_1, \ldots, Y_k][X]$. A node N is a leaf if $\mathrm{Pol}(N) = 0$. If N is not a leaf, the children of N contain the truncations of $-\mathrm{Prem}(\mathrm{Pol}(p(N)), \mathrm{Pol}(N))$ where $p(N)$ is the parent of N.

Example 2.1.13 Continuing Example 2.1.12, we consider $P = X^4 + aX^2 + bX + c$ and its derivative $P' = 4X^3 + 2aX + b$ and write down the tree $\mathrm{TRems(P, P')}$, denoting

$$\begin{aligned} \bar{S}_2 &= -\mathrm{Prem}(P, P') \\ &= -8\,aX^2 - 12\,bX - 16\,c, \\ \bar{S}_3 &= -\mathrm{Prem}(P', \bar{S}_2) \end{aligned}$$

$$= 64((8ac - 9b^2 - 2a^3)X - b(12c + a^2)),$$
$$\bar{S}_4 = -\mathrm{Prem}(\bar{S}_3, \bar{S}_2)$$
$$= 16384\, a^2 \left(256c^3 - 128a^2c^2 + 144ab^2c + 16a^4c - 27b^4 - 4a^3b^2\right),$$
$$u = -\mathrm{Prem}(P', \mathrm{Tru}_1(\bar{S}_2))$$
$$= 768b \left(-27\, b^4 + 72\, acb^2 + 256\, c^3\right).$$



Define

$$s = 8ac - 9b^2 - 2a^3,$$
$$t = -b(12c + a^2)$$
$$\delta = 256c^3 - 128a^2c^2 + 144ab^2c + 16a^4c - 27b^4 - 4a^3b^2.$$

The leftmost path in the tree going from the root to a leaf, namely the path $P, P', \bar{S}_2, \bar{S}_3, \bar{S}_4, 0$ can be understood as follows: if $(a, b, c) \in \mathrm{R}^3$ are such that the degree of the polynomials in the remainder sequence of P and P' are $4, 3, 2, 1, 0$, i.e. when $a \neq 0, s \neq 0, \delta \neq 0$ (getting rid

of obviously irrelevant factors), then the signed remainder sequence of $P = X^4 + aX^2 + bX + c$ and P' is proportional (up to non-zero squares of elements in R) to $P, P', \bar{S}_2, \bar{S}_3, \bar{S}_4$.

Notation 2.1.14 For a specialization of $Y = (Y_1, \ldots, Y_k)$ to $y = (y_1, \ldots, y_k) \in \mathrm{R}^k$, and $Q \in \mathrm{D}[Y_1, \ldots, Y_k][X]$, we denote the polynomial in $\mathrm{R}[X]$ obtained by substituting y for Y by Q_y. Given $\mathcal{Q} \subset \mathrm{D}[Y_1, \ldots, Y_k][X]$, we define $\mathcal{Q}_y \subset \mathrm{R}[X]$ as $\{Q_y \mid Q \in \mathcal{Q}\}$.

Let $Q = b_q X^q + \ldots + b_0 \in \mathrm{D}[Y_1, \ldots, Y_k][X]$. We define the quantifier free formula $\deg_X(Q) = i$ as

$$\begin{cases} b_q = 0 \wedge \ldots \wedge b_{i+1} = 0 \wedge b_i \neq 0 & \text{when} \quad 0 \leq i < q, \\ b_q \neq 0 & \text{when} \quad i = q, \\ b_q = 0 \wedge \ldots \wedge b_0 = 0 & \text{when} \quad i = -\infty, \end{cases}$$

so that the sets $\mathcal{R}(\deg_X(Q) = i)$ partition R^k and $y \in \mathcal{R}(\deg_X(Q) = i)$ if and only if $\deg(Q_y) = i$.

Given a leaf L of $\mathrm{TRems}(\mathrm{P}, \mathrm{Q})$, we denote by $\mathcal{B}_L$ the unique path from the root of $\mathrm{TRems}(\mathrm{P}, \mathrm{Q})$ to the leaf L. If N is a node in $\mathcal{B}_L$ which is not a leaf, we denote by $c(N)$ the unique child of N in $\mathcal{B}_L$. We denote by $\mathcal{C}_L$ the quantifier free formula

$$\deg_X(Q) = \deg_X(\mathrm{Pol}(c(R))) \wedge$$
$$\bigwedge_{N \in \mathcal{B}_L, N \neq R} \deg_X(-\mathrm{Prem}(\mathrm{Pol}(p(N)), \mathrm{Pol}(N))) = \deg_X(\mathrm{Pol}(c(N)))$$

It is clear from the definitions, since the remainder and pseudo-remainder of two polynomials in $C[X]$ are equal up to a square, that

Lemma 2.1.15 *The $\mathcal{R}(\mathcal{C}_L)$ partition R^k. Moreover, $y \in \mathcal{R}(\mathcal{C}_L)$ implies that the signed remainder sequence of P_y and Q_y is proportional (up to a square) to the sequence of polynomials $\mathrm{Pol}(N)_y$ in the nodes along the path $\mathcal{B}_L$ leading to L. In particular, $\mathrm{Pol}(p(L))_y$ is $\gcd(P_y, Q_y)$.*

Example 2.1.16 We start with an example. We describe the projection of the algebraic set

$$\{(a, b, c, X) \in \mathrm{R}^4 \mid X^4 + aX^2 + bX + c = 0\}$$

to R^3, i.e. the set

$$\{(a;b;c) \in \mathrm{R}^3 \mid \exists X \in \mathrm{R},\ X^4 + aX^2 + bX + c = 0\},$$

as a semi-algebraic set.

We look at all leaves of $\mathrm{TRems}(\mathrm{P}, \mathrm{P}')$ and at all possible signs for leading coefficients of all possible signed pseudo-remainders (using Example 2.1.12). We denote by n the difference between the number of sign changes at $-\infty$ and $+\infty$ in the Sturm sequence of $P = X^4 + aX^2 + bX + c$ for each case. We indicate for each leaf L of $\mathrm{TRems}(\mathrm{P}, \mathrm{P}')$ the quantifier free formula $\mathcal{C}_L$ and the degrees occurring in the signed pseudo-remainder sequence of P and P' along the path $\mathcal{B}_L$.

$$(a \neq 0 \wedge s \neq 0 \wedge \delta \neq 0, (4,3,2,1,0))$$

a	$-$	$-$	$-$	$-$	$+$	$+$	$+$	$+$
s	$+$	$+$	$-$	$-$	$+$	$+$	$-$	$-$
δ	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
n	4	2	0	2	0	-2	0	2

The first column can be read as follows: for every polynomial

$$P = X^4 + aX^2 + bX + c$$

satisfying $a < 0, s > 0, \delta > 0$, since the leading coefficients of the signed pseudo-remainder sequence of P and P' are $1, 4, -a, 64s, 16384a^2\delta$ (see Example 2.1.13) and the degrees of the polynomials in the signed pseudo-remainder sequence of P and P' are $4,3,2,1,0$, the signs of the signed pseudo-remainder sequence of P and P' at $-\infty$ are $+-+-+$ and at $+\infty$ are $+++++$. Therefore the number of real roots is 4.

The other columns can be read similarly. Notice that n can be negative (for $a > 0, s > 0, \delta < 0$). Though this looks paradoxical, Sturm's theorem is not violated. This only means that there is no polynomial $P \in \mathrm{R}[X]$ with $P = X^4 + aX^2 + bX + c$ and $a > 0, s > 0, \delta < 0$. Notice that even when n is non-negative, there might be no polynomial $P \in \mathrm{R}[X]$ with $P = X^4 + aX^2 + bX + c$ and (a, s, δ) satisfying the corresponding sign condition.

Similarly, for the other leaves of $\mathrm{TRems}(\mathrm{P}, \mathrm{P}')$

$$(a \neq 0 \wedge s \neq 0 \wedge \delta = 0, (4, 3, 2, 1))$$

a	$-$	$-$	$+$	$+$
s	$+$	$-$	$+$	$-$
n	3	1	-1	1

$$(a \neq 0 \wedge s = 0 \wedge t \neq 0, (4, 3, 2, 0))$$

a	$-$	$-$	$+$	$+$
t	$+$	$-$	$+$	$-$
n	2	2	0	0

$$(a \neq 0 \wedge s = t = 0, (4, 3, 2))$$

a	$-$	$+$
n	2	0

$$(a = 0 \wedge b \neq 0 \wedge u \neq 0, (4, 3, 1, 0))$$

b	$+$	$+$	$-$	$-$
u	$+$	$-$	$+$	$-$
n	2	0	0	2

$$(a = 0 \wedge b \neq 0 \wedge u = 0, (4, 3, 1))$$

b	$+$	$-$
n	1	1

$$(a = b = 0 \wedge c \neq 0, (4, 3, 0))$$

c	$+$	$-$
n	0	2

$$(a = b = c = 0, (4, 3))$$

$$n = 1$$

Finally, the formula $(\exists\ X)\ X^4 + aX^2 + bX + c = 0$ is equivalent to

$$\begin{array}{c}(a < 0 \wedge s > 0)\\ \vee\ (a < 0 \wedge s < 0 \wedge \delta < 0)\\ \vee\ (a > 0 \wedge s < 0 \wedge \delta < 0)\\ \vee\ (a < 0 \wedge s \neq 0 \wedge \delta = 0)\\ \vee\ (a > 0 \wedge s < 0 \wedge \delta = 0)\\ \vee\ (a < 0 \wedge s = 0 \wedge t \neq 0)\\ \vee\ (a < 0 \wedge s = 0 \wedge t = 0)\\ \vee\ (a = 0 \wedge b < 0 \wedge u < 0)\\ \vee\ (a = 0 \wedge b > 0 \wedge u > 0)\\ \vee\ (a = 0 \wedge b \neq 0 \wedge u = 0)\\ \vee\ (a = 0 \wedge b = 0 \wedge c < 0)\\ \vee\ (a = 0 \wedge b = 0 \wedge c = 0),\end{array}$$

by collecting all the sign conditions with $n \geq 1$.

The example should be useful in order to understand the proof of the following theorem.

Theorem 2.1.17 *Given a semi-algebraic set of* R^{k+1} *defined over* D*, its projection to* R^k *is a semi-algebraic set defined over* D*.*

Proof: Since every semi-algebraic set is a finite union of basic semi-algebraic sets it is sufficient to prove that the projection of a basic semi-algebraic set is semi-algebraic. Suppose that the basic semi-algebraic set S in R^{k+1} is

$$\mathcal{R}(\sigma, Z) = \{(y, x) \in \mathrm{R}^k \times \mathrm{R} \mid P(y, x) = 0 \wedge \bigwedge_{Q \in \mathcal{Q}} \mathrm{sign}(\mathcal{Q}(y, x)) = \sigma(Q)\},$$

$Z = \{z \in \mathrm{R}^{k+1} \mid P(z) = 0\}$. Let S' be the intersection of S with the subset of $(y, x) \in \mathrm{R}^{k+1}$ such that P_y is not identically zero.

Let L be a function on $\{0, 1, 2\}^{\mathcal{Q}}$ associating to each $\alpha \in \{0, 1, 2\}^{\mathcal{Q}}$ a leaf L_α of $\mathrm{TRems}(\mathrm{P}, \mathrm{P}'\mathcal{Q}^\alpha)$, and let $\mathcal{A}(L_\alpha)$ be the set of non-zero polynomials of $\mathrm{D}[Y_1, \ldots, Y_k]$ appearing in the quantifier free formula $\mathcal{C}_{L_\alpha}$, (see Notation 2.1.14).

Let $\mathcal{L}$ be the set of all functions L on $\{0,1,2\}^{\mathcal{Q}}$ associating to each α a leaf L_α of $\mathrm{TRems}(\mathrm{P}, \mathcal{Q}^\alpha)$, and

$$\mathcal{A} = \bigcup_{L \in \mathcal{L}} \bigcup_{\alpha \in \{0,1,2\}^{\mathcal{Q}}} \mathcal{A}(L_\alpha) \subset \mathrm{D}[Y_1, \ldots, Y_k].$$

Note that since $\mathcal{A}$ contains the coefficients of P',the signs of the coefficients of P are fixed as soon as the signs of the polynomials of $\mathcal{A}$ are fixed.

We define

$$\Sigma = \{\tau \in \{0,1,-1\}^{\mathcal{A}} | \forall y \in \mathcal{R}(\tau), \mathcal{R}(\sigma_y, Z_y) \neq \emptyset\},$$

where $Z_y = \{x \in \mathrm{R} \mid P(x,y) = 0\}$, $\sigma_y(Q_y) = \sigma(Q)$.

Using Lemma 2.1.10, it is clear that the semi-algebraic set

$$\bigcup_{\tau \in \Sigma} \{y \in \mathrm{R}^k | \mathrm{sign}(\mathcal{A}(y)) = \tau\}$$

coincides with the projection of S'.

The fact that the projection of the intersection of S with the subset of $(y,x) \in \mathrm{R}^{k+1}$ such that P_y is identically zero is semi-algebraic follows in a similar way, using Lemma 2.1.11.

Thus the whole projection $S = S' \cup (S \setminus S')$ is semi-algebraic as a union of semi-algebraic sets. □

Exercise 2.1.18 *Find the conditions on a, b such that $X^3 + aX + b$ has a strictly positive real root.*

The projection theorem (Theorem 2.1.17) implies that the theory of real closed fields admits quantifier elimination in the language of ordered fields, which is the following theorem.

Theorem 2.1.19 *Let $\Phi(Y)$ be a formula in the language of ordered fields with coefficients in an ordered ring* D *contained in the real closed field* R. *Then there is a quantifier free formula $\Psi(Y)$ with coefficients in* D *such that for every $y \in \mathrm{R}^k$, $\Phi(y)$ is true if and only if $\Psi(y)$ is true.*

Proof: Given a formula

$$\Theta(Y) = (\exists X)\ \mathcal{B}(X, Y),$$

where $\mathcal{B}$ is a quantifier free formula whose atoms are equations and inequalities involving polynomials in $\mathrm{D}[X, Y_1, \ldots, Y_k]$, Theorem 2.1.17 shows that there is a quantifier free formula $\Xi(Y)$ whose atoms are equations and inequalities involving polynomials in $\mathrm{D}[X, Y_1, \ldots, Y_k]$ and that is equivalent to $\Theta(Y)$. This is because $\mathcal{R}(\Theta(Y), \mathrm{R}^k)$ which is the projection of the semi-algebraic set $\mathcal{R}(\mathcal{B}(X, Y), \mathrm{R}^{k+1})$ defined over D is semi-algebraic and defined over D, and semi-algebraic sets defined over D are realizations of quantifier free formulas with coefficients in D. Since $(\forall X)\ \Phi$ is equivalent to $\neg((\exists X)\ \neg(\Phi))$, the theorem immediately follows by induction on the number of quantifiers. □

Corollary 2.1.20 *Let $\Phi(Y)$ be a formula in the language of ordered fields with coefficients in* D. *The set $\{y \in \mathrm{R}^k | \Phi(y)\}$ is semi-algebraic.*

Corollary 2.1.21 *A subset of* R *defined by a formula in the language of ordered fields with coefficients in* R *is a finite union of points and intervals.*

Proof: By Theorem 2.1.19 a subset of R defined by a formula in the language of ordered fields with coefficients in R is semi-algebraic and this is clearly a finite union of points and intervals. □

Exercise 2.1.22 *Show that the set $\{(x, y) \in \mathbb{R}^2 | \exists n \in \mathbb{N},\ \ y = nx\}$ is not a semi-algebraic set.*

Theorem 2.1.19 immediately implies the following theorem known as the Tarski-Seidenberg Principle or the Transfer Principle for real closed fields.

Theorem 2.1.23 (Tarski-Seidenberg principle) *Suppose that* R' *is a real closed field that contains the real closed field* R. *If Φ is a sentence in the language of ordered fields with coefficients in* R, *then it is true in* R *if and only if it is true in* R'.

Proof: By Theorem 2.1.19, there is a quantifier free formula Ψ R-equivalent to Φ. It follows from the proof of Theorem 2.1.17 that Ψ is R′-equivalent to Φ as well. Notice, too, that Ψ is a boolean combination of atoms of the form $c = 0, c > 0$, or $c < 0$, where $c \in \mathrm{R}$. Clearly, Ψ is true in R if and only if it is true in R′. □

Since any real closed field contains the real closure of $\mathbb{Q}$, a consequence of Theorem 2.1.23 is

Theorem 2.1.24 *Let* R *be a real closed field. A sentence in the language of fields with coefficients in* $\mathbb{Q}$ *is true in* R *if and only if it is true in any real closed field.*

2.2 Cylindrical algebraic decomposition

Using subresultants, it is possible to simplify description of the projection of a semi-algebraic set.

Example 2.2.1 Consider $P = X^4 + aX^2 + bX + c$,

$$\begin{aligned}
\mathrm{sr}_4(P, P') &= 1, \\
\mathrm{sr}_3(P, P') &= 4, \\
\mathrm{sr}_2(P, P') &= -8a, \\
\mathrm{sr}_1(P, P') &= 4(8ac - 9b^2 - 2a^3) \\
\mathrm{sr}_0(P, P') &= 256c^3 - 128a^2c^2 + 144ab^2c + 16a^4c - 27b^4 - 4a^3b^2.
\end{aligned}$$

Let

$$\begin{aligned}
s &= 8ac - 9b^2 - 2a^3, \\
\delta &= 256c^3 - 128a^2c^2 + 144ab^2c + 16a^4c - 27b^4 - 4a^3b^2.
\end{aligned}$$

Note that $\delta = \mathrm{sr}_0(P, P')$ is the discriminant of P.

We indicate in the following tables the number of real roots of P (computed using Theorem 1.4.43) in the various cases corresponding to

all the possible signs for a, s, δ:

1	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
4	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$-a$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
s	$+$	$+$	$+$	$-$	$-$	$-$	0	0	0
δ	$+$	$-$	0	$+$	$-$	0	$+$	$-$	0
n	4	2	3	0	2	1	2	2	2

1	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
4	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$-a$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$
s	$+$	$+$	$+$	$-$	$-$	$-$	0	0	0
δ	$+$	$-$	0	$+$	$-$	0	$+$	$-$	0
n	0	-2	-1	0	2	1	0	0	0

1	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
4	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$-a$	0	0	0	0	0	0	0	0	0
s	$+$	$+$	$+$	$-$	$-$	$-$	0	0	0
δ	$+$	$-$	0	$+$	$-$	0	$+$	$-$	0
n	2	0	1	0	2	1	0	2	1

Note that when $a = s = 0$, according to the definition of D when there are two consecutive zeroes,

$$\begin{cases} D(\mathrm{sr}(P, P')) = 0 & \text{if } \delta > 0 \\ D(\mathrm{sr}(P, P')) = 2 & \text{if } \delta < 0 \\ D(\mathrm{sr}(P, P')) = 1 & \text{if } \delta = 0. \end{cases}$$

As a consequence, the formula $(\exists\ X)\ X^4 + aX^2 + bX + c = 0$ is equivalent to

$$\begin{gathered} (a < 0 \wedge s \geq 0 \wedge \delta > 0) \vee \\ (a < 0 \wedge \delta \leq 0) \vee \\ (a > 0 \wedge s < 0 \wedge \delta \leq 0) \vee \\ (a = 0 \wedge s > 0 \wedge \delta \geq 0) \vee \\ (a = 0 \wedge s \leq 0 \wedge \delta \leq 0). \end{gathered}$$

collecting all sign conditions giving $n \geq 1$. It can be checked easily that the realization of the sign conditions $(a = 0 \wedge s > 0 \wedge \delta \geq 0)$ and $(a < 0 \wedge s = 0 \wedge \delta > 0)$ are empty. So that $(\exists\, X)\; X^4 + aX^2 + bX + c = 0$ is finally equivalent to

$$\begin{gathered}(a < 0 \wedge s > 0 \wedge \delta > 0) \vee \\ (a < 0 \wedge \delta \leq 0) \vee \\ (a > 0 \wedge s < 0 \wedge \delta \leq 0) \vee \\ (a = 0 \wedge s \leq 0 \wedge \delta \leq 0).\end{gathered}$$

2.2.1 Computing cylindrical decomposition

The cylindrical decomposition method is due to [3]

Definition 2.2.2 A **cylindrical decomposition** of R^k is a sequence $\mathcal{S}_1, \ldots, \mathcal{S}_k$ where, for each $1 \leq i \leq k$, $\mathcal{S}_i$ is a finite partition of R^i into semi-algebraic subsets, called the **cells of level** i, which satisfy the following properties:

Each cell $S \in \mathcal{S}_1$ is either a point or an open interval.

For every $1 \leq i < k$ and every $S \in \mathcal{S}_i$, there are finitely many continuous semi-algebraic functions

$$\xi_{S,1} < \ldots < \xi_{S,\ell_S} : S \longrightarrow \mathrm{R}$$

such that the cylinder $S \times \mathrm{R} \subset \mathrm{R}^{i+1}$ is the disjoint union of cells of $\mathcal{S}_{i+1}$ which are:

either the graph of one of the functions $\xi_{S,j}$, for $j = 1, \ldots, \ell_S$:

$$\{(x', x_{j+1}) \in S \times \mathrm{R} \mid x_{j+1} = \xi_{S,j}(x')\}\ ,$$

or a band of the cylinder bounded from below and from above by the graphs of the functions $\xi_{S,j}$ and $\xi_{S,j+1}$, for $j = 0, \ldots, \ell_S$, where we take $\xi_{S,0} = -\infty$ and $\xi_{i,\ell_S+1} = +\infty$:

$$\{(x', x_{j+1}) \in S \times \mathrm{R} \mid \xi_{S,j}(x') < x_{j+1} < \xi_{S,j+1}(x')\}\ .$$

Remark 2.2.3 Denoting by π_ℓ the canonical projection of R^k to R^ℓ, it follows immediately from the definition that for every cell T of $\mathcal{S}_i$, $i \geq \ell$, $S = \pi_\ell(T)$ is a cell of $\mathcal{S}_\ell$. We say that the cell T lies above the cell S. It is also clear that if S is a cell of $\mathcal{S}_i$, denoting by $T_1, \ldots, T_m$ the cells of $\mathcal{S}_{i+1}$ lying above S, $S \times \mathrm{R} = \bigcup_{j=1}^{m} T_j$.

Example 2.2.4 We illustrate this definition by presenting a cylindrical decomposition of R^3 adapted to the unit sphere.

The decomposition of R consists of five cells of level 1 corresponding to the points -1 and 1 and the three intervals they define.

$$\begin{cases} S_1 = (-\infty, -1) \\ S_2 = \{-1\} \\ S_3 = (-1, 1) \\ S_4 = \{1\} \\ S_5 = (1, \infty). \end{cases}$$

Above S_1 (respectively S_5) in R^2, there are no semi-algebraic functions, and only one cell $S_{1,1} = S_1 \times \mathrm{R}$ (respectively $S_{5,1} = S_5 \times \mathrm{R}$).

Above S_2 (respectively S_4), there is only one semi-algebraic function associating to -1 and 1 the constant value 0, and there are three cells.

$$\begin{cases} S_{2,1} = S_2 \times (-\infty, 0) \\ S_{2,2} = S_2 \times \{0\} \\ S_{2,3} = S_2 \times (0, \infty) \end{cases}$$

$$\left(\text{respectively } \begin{cases} S_{4,1} = S_4 \times (-\infty, 0) \\ S_{4,2} = S_4 \times \{0\} \\ S_{4,3} = S_4 \times (0, \infty) \end{cases} \right).$$

Above S_3, there are two semi-algebraic functions $\xi_{3,1}$ and $\xi_{3,2}$ associating to $x \in S_3$ the values $\xi_{3,1}(x) = -\sqrt{1-x^2}$ and $\xi_{3,2}(x) = \sqrt{1-x^2}$. There are 5 cells above S_3, the graphs of $\xi_{3,1}$ and $\xi_{3,2}$ and the bands they define

$$\begin{cases} S_{3,1} = \{(x,y) \mid -1 < x < 1, y < \xi_{3,1}(x)\} \\ S_{3,1} = \{(x,y) \mid -1 < x < 1, y = \xi_{3,1}(x)\} \\ S_{3,3} = \{(x,y) \mid -1 < x < 1, \xi_{3,1}(x) < y < \xi_{3,2}(x)\} \\ S_{3,4} = \{(x,y) \mid -1 < x < 1, y = \xi_{3,2}(x)\} \\ S_{3,5} = \{(x,y) \mid -1 < x < 1, \xi_{3,2}(x) < y\}. \end{cases}$$

Above $S_{1,1}$ (respectively $S_{2,1}, S_{2,3}, S_{3,1}, S_{3,5}, S_{4,1}, S_{4,3}, S_{5,1}$), there are no semi-algebraic functions, and only one cell:

$$\begin{gathered} S_{1,1,1} = S_{1,1} \times \mathrm{R} \\ (\text{respectively } S_{2,1,1} = S_{2,1} \times \mathrm{R}, \\ S_{2,3,1} = S_{2,3} \times \mathrm{R}, \\ S_{3,1,1} = S_{3,1} \times \mathrm{R}, \\ S_{3,5,1} = S_{3,5} \times \mathrm{R}, \\ S_{4,1,1} = S_{4,1} \times \mathrm{R}, \\ S_{4,3,1} = S_{4,3} \times \mathrm{R}, \\ S_{5,1,1} = S_{5,1} \times \mathrm{R}). \end{gathered}$$

Above $S_{2,2}$ (respectively $S_{3,2}$, $S_{3,4}$, $S_{4,2}$,) there is only one semi-algebraic function, the constant function 0, and three cells:

$$\left\{ \begin{array}{l} S_{2,2,1} = S_{2,2} \times (-\infty, 0) \\ S_{2,2,2} = S_{2,2} \times \{0\} \\ S_{2,2,3} = S_{2,2} \times (0, \infty) \end{array} \right.$$

$$\left(\text{respectively } \left\{ \begin{array}{l} S_{3,2,1} = S_{3,2} \times (-\infty, 0) \\ S_{3,2,2} = S_{3,2} \times \{0\} \\ S_{3,2,3} = S_{3,2} \times (0, \infty), \end{array} \right. , \right.$$

$$\left\{ \begin{array}{l} S_{3,4,1} = S_{3,4} \times (-\infty, 0) \\ S_{3,4,2} = S_{3,4} \times \{0\} \\ S_{3,4,3} = S_{3,4} \times (0, \infty) \end{array} \right. ,$$

$$\left. \left\{ \begin{array}{l} S_{4,2,1} = S_{4,2} \times (-\infty, 0) \\ S_{4,2,2} = S_{4,2} \times \{0\} \\ S_{4,2,3} = S_{4,2} \times (0, \infty) \end{array} \right. \right).$$

Above $S_{3,3}$, there are two semi-algebraic functions $\xi_{3,3,1}$ and $\xi_{3,3,2}$ associating to $(x, y) \in S_{3,3}$ the values $\xi_{3,3,1}(x, y) = -\sqrt{1 - x^2 - y^2}$ and $\xi_{3,3,2}(x, y) = \sqrt{1 - x^2 - y^2}$, and five cells

$$\left\{ \begin{array}{l} S_{3,3,1} = \{(x, y, z) \mid (x, y) \in S_{3,3}, z < \xi_{3,3,1}(x, y)\} \\ S_{3,3,2} = \{(x, y, z) \mid (x, y) \in S_{3,3}, z = \xi_{3,3,1}(x, y)\} \\ S_{3,3,3} = \{(x, y, z) \mid (x, y) \in S_{3,3}, \xi_{3,3,1}(x, y) < z < \xi_{3,3,2}(x, y)\} \\ S_{3,3,4} = \{(x, y, z) \mid (x, y) \in S_{3,3}, z = \xi_{3,3,2}(x, y)\} \\ S_{3,3,5} = \{(x, y, z) \mid (x, y) \in S_{3,3}, \xi_{3,3,2}(x, y) < z\}. \end{array} \right.$$

Definition 2.2.5 Given a finite set $\mathcal{P}$ of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$, a subset S of R^k is **$\mathcal{P}$-semi-algebraic** if S is the realization of a quantifier free formula with atoms $P = 0$, $P > 0$ or $P < 0$ with $P \in \mathcal{P}$. It is clear that for every semi-algebraic subset S of R^k, there exists a finite set $\mathcal{P}$ of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$ such that S is $\mathcal{P}$-semi-algebraic. A subset S of R^k is **$\mathcal{P}$-invariant** if every polynomial $P \in \mathcal{P}$ has a constant sign (> 0, < 0, or $= 0$) on S. A **cylindrical decomposition of R^k adapted to $\mathcal{P}$** is a cylindrical decomposition for which each cell $C \in \mathcal{S}_k$ is $\mathcal{P}$-invariant. It is clear that if S is $\mathcal{P}$-semi-algebraic, a cylindrical decomposition adapted to $\mathcal{P}$ is a cylindrical decomposition adapted to S.

The main result about cylindrical decomposition is the following.

Theorem 2.2.6 *For every finite set $\mathcal{P}$ of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$, there is a cylindrical decomposition of R^k adapted to $\mathcal{P}$.*

Since we intend to construct a cylindrical decomposition adapted to $\mathcal{P}$ it is convenient if for $S \in \mathcal{S}_{k-1}$ we choose each $\xi_{S,j}$ to be a root of a polynomial $P \in \mathcal{P}$, as a function of $(x_1, \ldots, x_{k-1}) \in S$. To this end, we shall prove that the real and complex roots (those in $\mathrm{R}[i] = C$) of a univariate polynomial depend continuously on its coefficients.

Notation 2.2.7 We write $D(z, r) = \{w \in \mathrm{C} \mid |z - w| < r\}$ for the **open disk** centered at z with radius r.

Theorem 2.2.8 (Continuity of Roots) *Let $P \in \mathrm{R}[X_1, \ldots, X_k]$ and let S be a semi-algebraic subset of R^{k-1}. Assume that $\deg(P(x', X_k))$ is constant on S and that for some $a' \in S$, $z_1, \ldots, z_j$ are the distinct roots of $P(a', X_k)$ in $\mathrm{C} = \mathrm{R}[i]$, with multiplicities $\mu_1, \ldots, \mu_j$, respectively. If the open disks $D(z_i, r) \subset \mathrm{C}$ are disjoint then there is an open neighborhood V of a' such that for every $x' \in V \cap S$, the polynomial $P(x', X_k)$ has exactly μ_i roots, counted with multiplicities, in the disk $D(z_i, r)$, for $i = 1, \ldots, j$.*

We next consider the conditions which ensure that the zeros of two polynomials over a connected semi-algebraic set define a cylindrical structure. :

Theorem 2.2.9 *Let $\mathcal{P}$ be a finite subset of $\mathrm{R}[X_1, \ldots X_k]$ and S a semi-algebraically connected semi-algebraic subset of R^{k-1}. Suppose that, for every $P \in \mathcal{P}$, $\deg(P(x', X_k))$ and the number of distinct real roots of P are constant over S and that, for every pair $P, Q \in \mathcal{P}$, $\deg(\gcd(P(x', X_k), Q(x', X_k))$ is also constant for all $x' \in S$. Then there are ℓ continuous semi-algebraic functions $\xi_1 < \ldots < \xi_\ell : S \to \mathrm{R}$ such that, for every $x' \in S$, the set of real roots of $\prod_{P \in \mathcal{P}'} P(x', X_k)$, where $\mathcal{P}'$ is the subset of $\mathcal{P}$ consisting of polynomials not identically 0 over S, is exactly $\{\xi_1(x'), \ldots, \xi_\ell(x')\}$. Moreover, for $i = 1, \ldots, \ell$ and for every $P \in \mathcal{P}'$, the multiplicity of the root $\xi_i(x')$ of $P(x', X_k)$ is constant for $x' \in S$.*

It follows from Proposition 1.4.13 that the number of distinct complex roots of P and Q and the degree of the greatest common divisor of P and Q are determined by whether the signed subresultant coefficients $\mathrm{sr}_i(P, P')$ and $\mathrm{sr}_i(P, Q)$ are zero or not, as long as the degrees (with respect to X_k) of P and Q are fixed.

Notation 2.2.10 Let

$$\mathrm{Tru}(\mathcal{P}) = \{\mathrm{Tru}(P) \mid P \in \mathcal{P}\}.$$

We define $\mathrm{Elim}_{X_k}(\mathcal{P})$ to be the set of polynomials in $\mathrm{R}[X_1, \ldots, X_{k-1}]$ defined as follows:

If $R \in \mathrm{Tru}(\mathcal{P})$, $\deg_{X_k}(R) \geq 2$, $\mathrm{Elim}_{X_k}(\mathcal{P})$ contains all $\mathrm{sr}_j(R, \partial R/\partial X_k)$ which are not in R, $j = 0, \ldots, \deg_{X_k}(R) - 2$.

If $R \in \mathrm{Tru}(\mathcal{P})$, $S \in \mathrm{Tru}(\mathcal{P})$, $\mathrm{Elim}_{X_k}(\mathcal{P})$ contains all $\mathrm{sr}_j(R, S)$ which are not in R, $j = 0, \ldots, \min(\deg_{X_k}(R), \deg_{X_k}(S)) - 1$.

If $R \in \mathrm{Tru}(\mathcal{P})$, and $\mathrm{lcof}(R)$ is not in R, $\mathrm{Elim}_{X_k}(\mathcal{P})$ contains $\mathrm{lcof}(R)$.

Theorem 2.2.11 *Let $\mathcal{P}$ be a set of polynomials in $\mathrm{R}[X_1, \ldots, X_k]$, and let S be a semi-algebraically connected semi-algebraic subset of R^{k-1} which is $\mathrm{Elim}_{X_k}(\mathcal{P})$-invariant. Then there are continuous semi-algebraic functions $\xi_1 < \ldots < \xi_\ell : S \to \mathrm{R}$ such that, for every $x' \in S$, the set*

$\{\xi_1(x'), \ldots, \xi_\ell(x')\}$ is the set of all real roots of all non-zero polynomials $P(x', X_k)$, $P \in \mathcal{P}$. The graph of each ξ_i (respectively each band of the cylinder $S \times \mathrm{R}$ bounded by these graphs) is a semi-algebraically connected semi-algebraic set semi-algebraically homeomorphic to S (respectively $S \times (0,1)$) and is $\mathcal{P}$-invariant.

Proof: Let $R \in \mathrm{Tru}(P)$ and consider the constructible set $A \subset \mathrm{R}^{k-1}$ defined by $\mathrm{lcof}(R) \neq 0, \deg(P) = \deg(R)$. By Proposition 1.4.13, for every $a' \in A$, the vanishing or non-vanishing of the $\mathrm{sr}_j(R, \partial R/\partial X_k)(a')$ determines the number of distinct roots of $P(a', X_k)$ in C, which is $\deg(R(a', X_k)) - \deg(\gcd(R(a', X_k), \partial R/\partial X_k(a', X_k))$.

Similarly, let $R \in \mathrm{Tru}(P)$ and $S \in \mathrm{Tru}(Q)$ and consider the constructible set B defined by

$$\mathrm{lcof}(R) \neq 0, \deg(P) = \deg(R), \mathrm{lcof}(S) \neq 0, \deg(Q) = \deg(S).$$

For every $a' \in B$, the vanishing or non-vanishing of the $\mathrm{sr}_j(R, S)(a')$ determine $\deg(\gcd(P(a', X_k), Q(a', X_k)))$. Thus, the assumption that a connected semi-algebraic subset of R^{k-1} is $\mathrm{Elim}_{X_k}(\mathcal{P})$-invariant implies that the hypotheses of Theorem 2.2.9 are satisfied. □

We are finally ready for the proof of Theorem 2.2.6.

Proof of Theorem 2.2.6 The proof is by induction on the dimension of the ambient space.

Let $\mathcal{Q} \subset \mathrm{R}[X_1]$ be finite. It is clear that there is a cylindrical decomposition of R adapted to $\mathcal{Q}$ since the real roots of the polynomials in $\mathcal{Q}$ decompose the line into finitely many points and open intervals which constitute the cells of a cylindrical decomposition of R adapted to $\mathcal{Q}$.

Let $\mathcal{Q} \subset \mathrm{R}[X_1, \ldots, X_i]$ be finite. Starting from a cylindrical decomposition of R^{i-1} adapted to $\mathrm{Elim}_{X_i}(\mathcal{Q})$, and applying to the cells of this cylindrical decomposition Proposition 2.2.11, yields a cylindrical decomposition of R^i adapted to $\mathcal{Q}$.

This proves the theorem. □

Example 2.2.12 We illustrate this result by presenting a cylindrical decomposition of R^3 adapted to the polynomial $P = X_1^2 + X_2^2 + X_3^2 - 1$. The 0-th Sylvester-Habicht matrix of P and $\frac{\partial P}{\partial X_3}$ is

$$\begin{pmatrix} 1 & 0 & X_1^2 + X_2^2 - 1 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Hence, $\mathrm{sr}_0(P, \frac{\partial P}{\partial X_3}) = -4(X_1^2 + X_2^2 - 1)$ and $\mathrm{sr}_1(P, \frac{\partial P}{\partial X_3}) = 2$. Getting rid of irrelevant constant factors, we obtain

$$\mathrm{Elim}_{X_3}(P) = \{X_1^2 + X_2^2 - 1\}.$$

Similarly,

$$\mathrm{Elim}_{X_2}(\mathrm{Elim}_{X_3}(P)) = \{X_1^2 - 1\}.$$

The associated cylindrical decomposition has already been described in Example 2.2.4.

We denote, for $i = k-1, \ldots, 1$,

$$\mathrm{C}_i(\mathcal{P}) = \mathrm{Elim}_{X_{i+1}}(\mathrm{C}_{i+1}(\mathcal{P})),$$

with $\mathrm{C}_k(\mathcal{P}) = \mathcal{P}$, so that

$$\mathrm{C}_i(\mathcal{P}) \subset \mathrm{R}[X_1, \ldots, X_i].$$

The family $\mathrm{C}(\mathcal{P}) = \bigcup_{i \leq k} \mathrm{C}_i(\mathcal{P})$ is the **cylindrifying family of polynomials associated to** $\mathcal{P}$. It follows from the proof of Theorem 2.2.6 that the semi-algebraically connected components of the sign conditions on $\mathrm{C}(\mathcal{P})$ are the cells of a cylindrical decomposition adapted to $\mathcal{P}$.

The Cylindrical Decomposition Algorithm consists of two phases: in the first phase the cylindrifying family of polynomials associated to $\mathcal{P}$ is computed and in the second phase the cells defined by these polynomials are used to define inductively, starting from $i = 1$, the cylindrical decomposition.

The computation of the cylindrifying family of polynomials associated to $\mathcal{P}$ is based on the following Elimination Algorithm.

The set of truncations of a polynomial $Q \in \mathrm{D}[X_1, \ldots, X_{k-1}][X_k]$ is the finite subset of $\mathrm{D}[X_1, \ldots, X_{k-1}][X_k]$ defined by

$$\mathrm{Tru}(Q) = \begin{cases} \{Q\} & \text{if } \mathrm{lcof}(Q) \in \mathrm{D}, \\ \{Q\} \cup \mathrm{Tru}(Q - \mathrm{lcof}_{X_k}(Q) X_k^{\deg_{X_k}(Q)}), & \text{otherwise.} \end{cases}$$

Algorithm 2.2.13 (Subresultant Elimination)

Input: *a finite ordered list of variables* $X_1, \ldots, X_k$*, a finite set* $\mathcal{P} \subset \mathrm{D}[X_1, \ldots,$
$X_k]$*, and a variable* X_k*.*

Output: *a finite set* $\mathrm{Elim}_{X_k}(\mathcal{P}) \subset \mathrm{D}[X_1, \ldots, X_{k-1}]$*. The set* $\mathrm{Elim}_{X_k}(\mathcal{P})$ *is such that the degree of* $P \in \mathcal{P}$ *with respect to* X_k*, the number of real roots of* $P \in \mathcal{P}$*, and the number of real roots common to* $P \in \mathcal{P}$ *and* $Q \in \mathcal{P}$ *is fixed on every semi-algebraically connected component of the realization of each sign condition on* $\mathrm{Elim}_{X_k}(\mathcal{P})$*.*

Procedure: *Place in* $\mathrm{Elim}_{X_k}(\mathcal{P})$ *the following polynomials when they are not in* D*:*

$\mathrm{sr}_j(R, \dfrac{\partial R}{\partial X_k})$ *for* $P \in \mathcal{P}$, $\deg_{X_k}(P) = p \geq 2$, $R \in \mathrm{Tru}(P)$, $j = 0, \ldots, \deg_{X_k}(R) - 2$.

$\mathrm{sr}_j(R, S)$ *for* $P \in \mathcal{P}$, $Q \in \mathcal{P}$, $R \in \mathrm{Tru}(P)$, $S \in \mathrm{Tru}(Q)$, $j = 0, \ldots, \min(\deg_{X_k}(R), \deg_{X_k}(S)) - 1$.

$\mathrm{lcof}(R)$ *for* $P \in \mathcal{P}$, $R \in \mathrm{Tru}(P)$.

Sketch of complexity analysis: Let

$$\mathrm{D}[X_1, \ldots, X_k] = \mathrm{D}[X_1, \ldots, X_{k-1}][X_k],$$

s a bound on $\#(\mathcal{P})$, and d a bound on the degrees of the elements of $\mathcal{P}$. There are $O(s^2 d^2)$ subresultant sequences to compute, since there are $O(s^2)$ couples of polynomials in $\mathcal{P}$ and $O(d)$ truncations for each polynomial to consider. Each of these subresultant sequence takes $O(d^2)$ arithmetic operations in the integral domain $\mathrm{D}[X_1, \ldots, X_{k-1}]$ according

to the complexity analysis of Algorithm 1.4.41 (Signed Subresultant). The degree with respect to $X_1, \ldots, X_{k-1}$ of the polynomials throughout these computations is bounded by $2d^2$. □

Example 2.2.14 a) Let $P = X_1^2 + X_2^2 + X_3^2 - 1$. The output of Algorithm 2.2.13 (Subresultant Elimination) with input the variable X_3 and the set $\mathcal{P} = \{P\}$ is (getting rid of irrelevant constant factors) the polynomial

$$\mathrm{sr}_0(P, \frac{\partial P}{\partial X_3}) = X_1^2 + X_2^2 - 1$$

(see example 2.2.12).

b) Consider the two polynomials

$$P = X_2^2 - X_1(X_1 + 1)(X_1 - 2)$$

and

$$Q = X_2^2 - (X_1 + 2)(X_1 - 1)(X_1 - 3).$$

The output of Algorithm 2.2.13 (Subresultant Elimination) with input the variable Y and $\mathcal{P} = \{P, Q\}$ contains three polynomials: the discriminant of P with respect to X_2,

$$\mathrm{sr}_0(P, \frac{\partial P}{X_2}) = 4X_1(X_1 + 1)(X_1 - 2),$$

the discriminant of Q with respect to Y,

$$\mathrm{sr}_0(Q, \frac{\partial Q}{X_2}) = 4(X_1 + 2)(X_1 - 1)(X_1 - 3),$$

and the resultant of P and Q with respect to Y,

$$\mathrm{sr}_0(P, Q) = (-X_1^2 - 3X_1 + 6)^2,$$

since $\mathrm{sr}_1(P, Q) = 0$ is a constant.

Now we are ready to describe the two phases of the cylindrical decomposition method.

Let $\mathcal{S} = \mathcal{S}_1, \ldots, \mathcal{S}_k$ be a cylindrical decomposition of R^k. A **cylindrical set of sample points** of $\mathcal{S}$, $\mathcal{A} = \mathcal{A}_1, \ldots, \mathcal{A}_k$, is a list of k sets such that

for every i, $1 \leq i \leq k$, $\mathcal{A}_i$ is a finite subset of R^i which intersects every $S \in \mathcal{S}_i$,

for every i, $1 \leq i \leq k-1$, $\pi_i(\mathcal{A}_{i+1}) = \mathcal{A}_i$, where π_i is the projection from R^{i+1} to R^i forgetting the last coordinate.

Algorithm 2.2.15 (Cylindrical Decomposition)

Input: *a finite ordered list of variables* $X_1, \ldots, X_k$, *and a finite set* $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$.

Output: *a cylindrical set of sample points of a cylindrical decomposition* $\mathcal{S}$ *adapted to* $\mathcal{P}$ *and the sign of the elements of* $\mathcal{P}$ *on each cell of* $\mathcal{S}_k$.

Procedure:

Initialize $\mathrm{C}_k(\mathcal{P}) := \mathcal{P}$.

Elimination phase: *Compute* $\mathrm{C}_i(\mathcal{P}) = \mathrm{Elim}_{X_{i+1}}(\mathrm{C}_{i+1}(\mathcal{P}))$, *for* $i = k-1, \ldots, 1$, *applying repeatedly* $\mathrm{Elim}_{X_{i+1}}$ *using Algorithm 2.2.13 (Subresultant Elimination).*

Lifting phase: *Compute the sample points of the cells in* $\mathcal{S}_1$ *by characterizing the roots of* $\mathrm{C}_1(\mathcal{P})$ *and choosing a point in each interval they determine.*

For every $i = 2, \ldots, k$, *compute the sample points of the cells of* $\mathcal{S}_i$ *from the sample points of the cells in* $\mathcal{S}_{i-1}$ *as follows: Consider, for every sample point* x *of a cell in* $\mathcal{S}_{i-1}$, *the list* L *of non-zero polynomials* $P_i(x, X_i)$ *with* $P_i \in \mathrm{C}_i(\mathcal{P})$. *Characterize the roots of* L *and choose a point in each interval they determine.*

Output the sample points of the cells and the sign of $P \in \mathcal{P}$ *on the corresponding cells of* R^k.

Note that we have not been very precise about how we describe and compute sample points.

Sketch of omplexity analysis: Let s be a bound on $\#(\mathcal{P})$ and let d be a bound on the degrees of the elements of $\mathcal{P}$. Using the complexity analysis of Algorithm 2.2.13 (Subresultant Elimination), if the input polynomials have degree D, the degree of the output is $2(D^2)$ after one application of Algorithm 2.2.13 (Subresultant Elimination). Thus, the degrees of the polynomials output after $k-1$ applications of Algorithm 2.2.13 (Subresultant Elimination) are bounded by $f(d, k-1)$, where f satisfies the recurrence relation

$$f(d,i) = 2f(d,i-1)^2, f(d,0) = d. \tag{2.2}$$

Solving the recurrence we get that $f(d,k) = 2^{1+2+\ldots+2^{k-2}} d^{2^{k-1}}$, and hence the degrees of the polynomials in the intermediate computations and the output are bounded by $2^{1+2+\ldots+2^{k-2}} d^{2^{k-1}} = O(d)^{2^{k-1}}$, which is polynomial in d and doubly exponential in k. A similar analysis shows that the number of polynomials output is bounded by $(sd)^{3^{k-1}}$, which is polynomial in s and d and doubly exponential in k. □

Example 2.2.16 Let $P = X_1^2 + X_2^2 + X_3^2 - 1$. Continuing Example 2.2.12, we describe the output of the Cylindrical Decomposition Algorithm applied to $\mathcal{P} = \{P\}$.

We have

$$\begin{aligned} \mathrm{C}_3(\mathcal{P}) &= \{X_1^2 + X_2^2 + X_3^2 - 1\}, \\ \mathrm{C}_2(\mathcal{P}) &= \{X_1^2 + X_2^2 - 1\}, \\ \mathrm{C}_1(\mathcal{P}) &= \{X_1^2 - 1\}. \end{aligned}$$

The sample points of $\mathbb{R}$ consists of five points, corresponding to the two roots of $X^2 - 1$ and one point in each of the three intervals they define: these are the semi-algebraically connected components of the realization of sign conditions defined by $\mathrm{C}_1(\mathcal{P})$. We choose a sample point in each cell and obtain $\{(S_1, -2), (S_2, -1), (S_3, 0), (S_4, 1), (S_5, 2)\}$.

The cells in $\mathbb{R}^2$ are obtained by taking the semi-algebraically connected components of the realization of sign conditions defined by $\mathrm{C}_1(\mathcal{P}) \cup \mathrm{C}_2(\mathcal{P})$. There are thirteen such cells, listed in Example 2.2.4. The sample points in $\mathbb{R}^2$ consist of thirteen points, one in each cell. The projection of a sample point in a cell of $\mathbb{R}^2$ on its first coordinate

is a point in a cell of $\mathbb{R}$. We choose a sample point in each cell and obtain

$$\begin{gathered}\{(S_{1,1},(-2,0)),\\ (S_{2,1},(-1,-1)),(S_{2,2},(-1,0)),(S_{2,3},(-1,1)),\\ (S_{3,1},(0,-2)),(S_{3,2},(0,-1)),(S_{3,3},(0,0)),(S_{3,4},(0,1)),(S_{3,5},(0,2)),\\ (S_{4,1},(1,-1)),(S_{4,2},(1,0)),(S_{4,3},(1,1)),\\ (S_{5,1},(2,0))\}.\end{gathered}$$

The cells in $\mathbb{R}^3$ are obtained by taking the semi-algebraically connected components of the realization of sign conditions defined by $\mathrm{C}_1(\mathcal{P}) \cup \mathrm{C}_2(\mathcal{P}) \cup \mathrm{C}_3(\mathcal{P})$. There are twenty five such cells, listed in Example 2.2.4. The sample points in $\mathbb{R}^3$ consist of twenty five points, one in each cell. The projection of a sample point in a cell of $\mathbb{R}^3$ is a point in a cell of $\mathbb{R}^2$. We choose the following sample points and obtain, indicating the cell, its sample point and the sign of P at this sample point:

$$\begin{gathered}\{(S_{1,1,1},(-2,0,0),1),\\ (S_{2,1,1},(-1,-1,0),1),\\ (S_{2,2,1},(-1,0,-1),1),(S_{2,2,2},(-1,0,0),0),(S_{2,2,3},(-1,0,1),1),\\ (S_{2,3,1},(-1,1,0),1),\\ (S_{3,1,1},(0,-2,0),1),\\ (S_{3,2,1},(0,-1,-1),1),(S_{3,2,2},(0,-1,0),0),(S_{3,2,3},(0,-1,1),1),\\ (S_{3,3,1},(0,0,-2),1),(S_{3,3,2},(0,0,-1),0),\\ (S_{3,3,3},(0,0,0),-1),\\ (S_{3,3,4},(0,0,1),0),(S_{3,3,5},(0,0,2),1),\\ (S_{3,4,1},(0,1,-1),1),(S_{3,4,2},(0,1,0),0),(S_{3,4,3},(0,1,1),1),\\ (S_{3,5,1},(0,2,0),1),\\ (S_{4,1,1},(1,-1,0),1),\\ (S_{4,2,1},(1,0,-1),1),(S_{4,2,2},(1,0,0),0),(S_{4,2,3},(1,0,1),1),\\ (S_{4,3,1},(1,1,0),1),\\ (S_{5,1,1},(2,0,0),1)\}.\end{gathered}$$

This example is particularly simple because we can choose all sample points with rational coordinates. This will not be the case in general: the coordinates of the sample points will be roots of univariate polynomials above sample points of cells of lower dimension, and the real root isolation technique has to be generalized to deal with the cylindrical situation.

2.2.2 Decision

Now we explain how to decide the truth or falsity of a sentence using the Cylindrical Decomposition Algorithm applied to the family of polynomials used to build the sentence.

Let $\mathcal{P}$ be a finite subset of $\mathrm{R}[X_1, \ldots, X_k]$. A **$\mathcal{P}$-atom** is one of $P = 0, P \neq 0, P > 0, P < 0$, where P is a polynomial in $\mathcal{P}$. A **$\mathcal{P}$-formula** is a formula (Definition page 67) written with $\mathcal{P}$-atoms. A **$\mathcal{P}$-sentence** is a sentence (Definition page 68) written with $\mathcal{P}$-atoms.

Notation 2.2.17 For $z \in \mathrm{R}^k$, we denote by $\mathrm{sign}(\mathcal{P})(z)$ the sign condition on $\mathcal{P}$ mapping $P \in \mathcal{P}$ to $\mathrm{sign}(P)(z) \in \{0, 1, -1\}$.

We are going to define inductively the tree of cylindrical realizable sign conditions, $\mathrm{CSign}(\mathcal{P})$, of $\mathcal{P}$. The importance of this notion is that the truth or falsity of any $\mathcal{P}$-sentence can be decided from $\mathrm{CSign}(\mathcal{P})$.

We denote by π_i the projection from R^{i+1} to R^i forgetting the last coordinate. By convention, $\mathrm{R}^0 = \{0\}$.

For $z \in \mathrm{R}^k$, let $\mathrm{CSign}_k(\mathcal{P})(z) = \mathrm{sign}(\mathcal{P})(z)$.

For i, $0 \leq i < k$, and all $y \in \mathrm{R}^i$, we inductively define

$$\mathrm{CSign}_i(\mathcal{P})(y) = \{\mathrm{CSign}_{i+1}(\mathcal{P})(z) | z \in \mathrm{R}^{i+1}, \pi_i(z) = y\}.$$

Finally, we define the **tree of cylindrical realizable sign conditions of $\mathcal{P}$**, $\mathrm{CSign}(\mathcal{P})$, by

$$\mathrm{CSign}(\mathcal{P}) = \mathrm{CSign}_0(\mathcal{P})(0).$$

Example 2.2.18 Consider two bivariate polynomials $P_1 = X_2, P_2 = X_1^2 + X_2^2 - 1$ and $\mathcal{P} = \{P_1, P_2\}$.

We order the set $\mathcal{P}$ with the order $P_1 < P_2$.

For $y \in \mathrm{R}^2$, $\mathrm{sign}(\mathcal{P})(y)$ is the mapping from $\mathcal{P}$ to $\{0, 1, -1\}$ sending (P_1, P_2) to $(\mathrm{sign}(P_1(y)), \mathrm{sign}(P_2(y)))$. Abusing notation, we denote the mapping $\mathrm{sign}(\mathcal{P})(y)$ by $(\mathrm{sign}(P_1(y)), \mathrm{sign}(P_2(y)))$. For example if $y = (0, 0)$, $\mathrm{sign}(\mathcal{P})(0, 0) = (0, -1)$ since $\mathrm{sign}(P_1(0, 0)) = 0$ and $\mathrm{sign}(P_2(0, 0)) = -1$.

Fixing $x \in \mathrm{R}$, $\mathrm{CSign}_1(\mathcal{P})(x)$ is the set of all possible $\mathrm{sign}(\mathcal{P})(z)$ for $z \in \mathrm{R}^2$ such that $\pi_1(z) = x$. For example if $x = 0$, there are seven possibilities for $\mathrm{sign}(\mathcal{P})(z)$ as z varies in $\{0\} \times \mathrm{R}$:

$$(-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1), (1, 0), (1, 1).$$

So $\mathrm{CSign}_1(\mathcal{P})(0)$ is

$$\{(-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1), (1, 0), (1, 1)\}.$$

Similarly, if $x = 1$, there are three possibilities for $\mathrm{sign}(\mathcal{P})(z)$ as z varies in $\{1\} \times \mathrm{R}$:

$$(-1, 1), (0, 0), (1, 1).$$

So $\mathrm{CSign}_1(\mathcal{P})(1)$ is

$$\{(-1, 1), (0, 0), (1, 1)\}.$$

If $x = 2$, there are three possibilities for $\mathrm{sign}(\mathcal{P})(z)$ as z varies in $\{2\} \times \mathrm{R}$:

$$(-1, 1), (0, 1), (1, 1).$$

So $\mathrm{CSign}_1(\mathcal{P})(2)$ is

$$\{(-1, 1), (0, 1), (1, 1)\}.$$

Finally $\mathrm{CSign}(\mathcal{P})$ is the set of all possible $\mathrm{CSign}_1(\mathcal{P})(x)$ for $x \in \mathrm{R}$. It is easy to check that the three cases we have considered ($x = 0, x = 1, x = 2$) already give all possible $\mathrm{CSign}_1(\mathcal{P})(x)$ for $x \in \mathrm{R}$. So $\mathrm{CSign}(\mathcal{P})$ is the set with three elements

$$\{\{(-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1), (1, 0), (1, 1)\},$$
$$\{(-1, 1), (0, 0), (1, 1)\},$$
$$\{(-1, 1), (0, 1), (1, 1)\}\}.$$

We now explain how $\mathrm{CSign}(\mathcal{P})$ can be determined from a cylindrical set of sample points of a cylindrical decomposition adapted to $\mathcal{P}$ and the signs of $P \in \mathcal{P}$ at these points.

If $\mathcal{A} = \mathcal{A}_1, \ldots, \mathcal{A}_k$, $\mathcal{A}_i \subset \mathrm{R}^k$, $\pi_i(\mathcal{A}_{i+1}) = \mathcal{A}_i$, where π_i is the projection from R^{i+1} to R^i forgetting the last coordinate, we define inductively the **tree of cylindrical realizable sign conditions** $\mathrm{CSign}(\mathcal{P}, \mathcal{A})$ of $\mathcal{P}$ on $\mathcal{A}$.

For $z \in \mathcal{A}_k$, let

$$\mathrm{CSign}_k(\mathcal{P}, \mathcal{A})(z) = \mathrm{sign}(\mathcal{P})(z).$$

For all i, $0 \leq i < k$, and all $y \in \mathcal{A}_i$, we inductively define

$$\mathrm{CSign}_i(\mathcal{P}, \mathcal{A})(y) = \{\mathrm{CSign}_{i+1}(\mathcal{P}, \mathcal{A})(z) | z \in \mathcal{A}_{i+1}, \pi_i(z) = y\}.$$

Finally,

$$\mathrm{CSign}(\mathcal{P}, \mathcal{A}) = \mathrm{CSign}_0(\mathcal{P}, \mathcal{A})(0).$$

Note that $\mathrm{CSign}(\mathcal{P}) = \mathrm{CSign}(\mathcal{P}, \mathrm{R}^k)$. Note also that $\mathrm{CSign}(\mathcal{P}, \mathcal{A})$ is a subtree of $\mathrm{CSign}(\mathcal{P})$.

We are going to prove the following result.

Proposition 2.2.19 *Let $\mathcal{S} = \mathcal{S}_1, \ldots, \mathcal{S}_k$ be a cylindrical decomposition of R^k adapted to $\mathcal{P}$ and let $\mathcal{A} = \mathcal{A}_1, \ldots, \mathcal{A}_k$ be a cylindrical set of sample points for $\mathcal{S}$. Then*

$$\mathrm{CSign}(\mathcal{P}, \mathcal{A}) = \mathrm{CSign}(\mathcal{P}).$$

We first start by explaining how this works on an example.

Example 2.2.20 Let $P = X_1^2 + X_2^2 + X_3^2 - 1$ and $\mathcal{P} = \{P\}$. Since there is only one polynomial in $\mathcal{P}$, we identify $\{0, 1, -1\}^{\mathcal{P}}$ with $\{0, 1, -1\}$.

We use Example 2.2.16, where the cells and sample points of the cylindrical decomposition of $\{P = X_1^2 + X_2^2 + X_3^2 - 1\}$ were described. The sign condition $\mathrm{sign}(\mathcal{P})(u)$ is fixed on each cell of R^3 by the sign of P at the sample point of the cell and thus

$$\mathrm{sign}(\mathcal{P})(z) = \begin{cases} -1 & \text{if } z \in S_{3,3,3} \\ 0 & \text{if } z \in S_{2,2,1} \cup S_{2,2,2} \cup S_{3,2,2} \\ & \cup S_{3,3,2} \cup S_{3,3,4} \cup S_{3,4,2} \cup S_{4,2,2} \\ 1 & \text{otherwise.} \end{cases}$$

The set $\mathrm{CSign}_2(\mathcal{P})(y)$ is fixed on each cell of R^2 by its value at the sample point of the cell and thus

$$\mathrm{CSign}_2(\mathcal{P})(y) = \begin{cases} \{0,1,-1\} & \text{if } y \in S_{3,3} \\ \{0,1\} & \text{if } y \in S_{2,2} \cup S_{3,2} \cup S_{3,4} \cup S_{4,2} \\ \{1\} & \text{otherwise.} \end{cases}$$

The set $\mathrm{CSign}_1(\mathcal{P})(x)$ is fixed on each cell of R by its value at the sample point of the cell and thus

$$\mathrm{CSign}_1(\mathcal{P})(x) = \begin{cases} \{\{1\},\{0,1\},\{0,1,-1\}\} & \text{if } x \in S_3 \\ \{\{1\},\{0,1\}\} & \text{if } x \in S_2 \cup S_4 \\ \{\{1\}\} & \text{if } x \in S_1 \cup S_5\ . \end{cases}$$

Finally the set $\mathrm{CSign}(\mathcal{P})$ has three elements and

$$\mathrm{CSign}(\mathcal{P}) = \{\{\{1\},\{0,1\},\{0,1,-1\}\},\{\{1\},\{0,1\}\},\{\{1\}\}\}.$$

This means that there are three possible cases:

there are values of $x_1 \in \mathrm{R}$ for which

- for some value of $x_2 \in \mathrm{R}$, the only sign taken by $P(x_1, x_2, x_3)$ when x_3 varies in R is 1,
- for some value of $x_2 \in \mathrm{R}$, the only signs taken by $P(x_1, x_2, x_3)$ when x_3 varies in R are 0 or 1,
- for some value of $x_2 \in \mathrm{R}$, the signs taken by $P(x_1, x_2, x_3)$ when x_3 varies in R are 0, 1, or -1,
- and these are the only possibilities,

there are values of x_1 for which

- for some value of $x_2 \in \mathrm{R}$, the only sign taken by $P(x_1, x_2, x_3)$ when x_3 varies in R is 1,
- for some value of $x_2 \in \mathrm{R}$, the only signs taken by $P(x_1, x_2, x_3)$ when x_3 varies in R are 0 or 1,
- and these are the only possibilities,

and there are values of x_1 for which

the only sign taken by $P(x_1, x_2, x_3)$ when (x_2, x_3) varies in R^2 is 1,

and together these three cases exhaust all possible values of $x_1 \in \mathrm{R}$.

Proposition 2.2.21 *Let $\mathcal{S} = \mathcal{S}_1, \ldots, \mathcal{S}_k$ be a cylindrical decomposition of R^k adapted to $\mathcal{P}$. For every $1 \leq i \leq k$ and every $S \in \mathcal{S}_i$, $\mathrm{CSign}_i(y)$ is constant as y varies in S.*

Proof : The proof is by induction on $k - i$.

If $i = k$, the claim is true since the sign of every $P \in \mathcal{P}$ is fixed on $S \in \mathcal{S}_k$.

Suppose that the claim is true for $i + 1$ and consider $S \in \mathcal{S}_i$. Let $T_1, \ldots, T_\ell$ be the cells of $\mathcal{S}_{i+1}$ such that $\pi_i(T_j) = S$. By induction hypothesis, $\mathrm{CSign}_{i+1}(\mathcal{P})(z)$ is constant as z varies in T_j. Since $\mathcal{S}$ is a cylindrical decomposition, $\bigcup_{j=1}^{\ell} T_j = S \times \mathrm{R}$. Thus

$$\mathrm{CSign}_i(\mathcal{P})(y) = \{\mathrm{CSign}_{i+1}(\mathcal{P})(z) | z \in \mathrm{R}^{i+1}, \pi_i(z) = y\}$$

is constant as y varies in S. □

Proof of Proposition 2.2.19: Let $\mathcal{A}_0 = \{0\}$. We are going to prove that for every $y \in \mathcal{A}_i$,

$$\mathrm{CSign}_i(\mathcal{P})(y) = \mathrm{CSign}_i(\mathcal{P}, \mathcal{A})(y).$$

The proof is by induction on $k - i$.

If $i = k$, the claim is true since $\mathcal{A}_k$ meets every cell of $\mathcal{S}_k$.

Suppose that the claim is true for $i + 1$ and consider $y \in \mathcal{A}_i$. Let $S \in \mathcal{S}_i$ be the cell containing y, and let $T_1, \ldots, T_\ell$ be the cells of $\mathcal{S}_{i+1}$ such that $\pi_i(T_j) = S$. Denote by z_j the unique point of $T_j \cap \mathcal{A}_{i+1}$ such that $\pi_i(z_j) = y$. By induction hypothesis,

$$\mathrm{CSign}_{i+1}(\mathcal{P})(z_j) = \mathrm{CSign}_{i+1}(\mathcal{P}, \mathcal{A})(z_j).$$

Since $\mathrm{CSign}_{i+1}(\mathcal{P})(z)$ is constant as z varies in T_j,

$$\mathrm{CSign}_i(\mathcal{P})(y) = \{\mathrm{CSign}_{i+1}(\mathcal{P})(z) | z \in \mathrm{R}^{i+1}, \pi_i(z) = y\}$$

$$= \{\mathrm{CSign}_{i+1}(\mathcal{P}, \mathcal{A})(z) | z \in \mathcal{A}_{i+1}, \pi_i(z) = y\} = \mathrm{CSign}_i(\mathcal{P}, \mathcal{A})(y)$$

□

The Cylindrical Decision Algorithm is based on the following result. We are going to need a notation.

Notation 2.2.22 If $\mathcal{P} \subset \mathrm{K}[X_1, \ldots, X_k]$ is finite, $X = (X_1, \ldots, X_k)$, $F(X)$ is a $\mathcal{P}$-quantifier free formula, and $\sigma \in \mathcal{P}^{\{0,1,-1\}}$ is a sign condition on $\mathcal{P}$, we define $F^\star(\sigma) \in \{\mathrm{True}, \mathrm{False}\}$ as follows :

If F is the atom $P = 0$, $P \in \mathcal{P}$, $F^\star(\sigma) = \mathrm{True}$ if $\sigma(P) = 0$, $F^\star(\sigma) = \mathrm{False}$ otherwise.

If F is the atom $P > 0$, $P \in \mathcal{P}$, $F^\star(\sigma) = \mathrm{True}$ if $\sigma(P) = 1$, $F^\star(\sigma) = \mathrm{False}$ otherwise.

If F is the atom $P < 0$, $P \in \mathcal{P}$, $F^\star(\sigma) = \mathrm{True}$ if $\sigma(P) = -1$, $F^\star(\sigma) = \mathrm{False}$ otherwise.

If $F = F_1 \wedge F_2$, $F^\star(\sigma) = F_1^\star(\sigma) \wedge F_2^\star(\sigma)$.

If $F = F_1 \vee F_2$, $F^\star(\sigma) = F_1^\star(\sigma) \vee F_2^\star(\sigma)$.

If $F = \neg(G)$, $F^\star(\sigma) = \neg(G^\star(\sigma))$.

Example 2.2.23 If

$$F = X_1^2 + X_2^2 + X_3^2 - 1 > 0,$$

then

$$F^\star(\sigma) = \begin{cases} \mathrm{True} & \text{if } \sigma = 1 \\ \mathrm{False} & \text{if } \sigma = 0, -1 \end{cases}$$

Proposition 2.2.24 *The $\mathcal{P}$-sentence*

$$(\mathrm{Q_1 X_1})\ (\mathrm{Q_2 X_2}) \ldots\ (\mathrm{Q_k X_k})\ \mathrm{F}(\mathrm{X_1}, \ldots, \mathrm{X_k}),$$

where $F(X_1, \ldots, X_k)$ is quantifier free, $\mathrm{Q_i} \in \{\exists, \forall\}$, is true if and only if

$$(\mathrm{Q_1}\sigma_1 \in \mathrm{CSign}(\mathcal{P}))\ (\mathrm{Q_2}\sigma_2 \in \sigma_1) \ldots\ (\mathrm{Q_k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\mathrm{F}^\star(\sigma_\mathrm{k})$$

is true.

Example 2.2.25 We illustrate the statement of the proposition by an example. Consider again $\mathcal{P} = \{X_1^2 + X_2^2 + X_3^2 - 1\}$, and recall that

$$\mathrm{CSign}(\mathcal{P}) = \{\{\{1\}, \{0,1\}, \{0,1,-1\}\}, \{\{1\}, \{0,1\}\}, \{\{1\}\}\}$$

by Example 2.2.16.

The sentence

$$(\forall\, X_1)(\forall\, X_2)(\forall\, X_3)\ F,$$

with $F = X_1^2+X_2^2+X_3^2-1 > 0$ is false since taking $(x_1, x_2, x_3) = (0,0,0)$ we get $x_1^2 + x_2^2 + x_3^2 - 1 < 0$. It is also the case that

$$\forall\sigma_1 \mathrm{CSign}(\mathcal{P}))\forall\sigma_2 \in \sigma_1 \forall\sigma_3 \in \sigma_2\ F^\star(\sigma_3)$$

is false since taking $\sigma_1 = \{\{1\}, \{0,1\}, \{0,1,-1\}\}, \sigma_2 = \{0,1,-1\}, \sigma_3 = -1$, the value of $F^\star(\sigma_3)$ is false.

Proof of Proposition 2.2.24 : The proof is by induction on the number k of quantifiers, starting from the one outside.

Since $(\forall\, X)\ \Phi$ is equivalent to $\neg\ (\exists X)\ \neg\Phi$, we can suppose without loss of generality that $\mathrm{Q_1}$ is $\exists$.

The claim is certainly true when there is only one existential quantifier, by definition of sign($\mathcal{P}$.

Suppose that

$$(\exists X_1)\ (\mathrm{Q_2 X_2}) \ldots\ \ (\mathrm{Q_k X_k})\ \mathrm{F(X_1, \ldots, X_k)},$$

is true, and choose $a \in \mathrm{R}$ such that

$$(\mathrm{Q_2 X_2}) \ldots\ \ (\mathrm{Q_k X_k})\ \mathrm{F(a, \ldots, X_k)}$$

is true. Note that, if $\mathcal{P}_a$ is the set of polynomials obtained by substituting $a \in \mathrm{R}$ to X_1 in $\mathcal{P}$,

$$\mathrm{CSign_1}(\mathcal{P})(a) = \mathrm{CSign}(\mathcal{P}_a).$$

By induction hypothesis,

$$(\mathrm{Q_2}\sigma_2 \in \mathrm{CSign}(\mathcal{P}_\mathrm{a})) \ldots\ \ (\mathrm{Q_k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k-1}})\ \mathrm{F}^\star(\sigma_\mathrm{k})$$

is true. So, taking $\sigma_1 = \mathrm{CSign}(\mathcal{P}_a) = \mathrm{CSign}(\mathcal{P})(a) \in \mathrm{CSign}(\mathcal{P})$,

$$(\exists \sigma_1 \in \mathrm{CSign}(\mathcal{P}))(\mathrm{Q}_2\sigma_2 \in \sigma_1)\ldots \ \ (\mathrm{Q}_\mathrm{k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\ \mathrm{F}^\star(\sigma_\mathrm{k})$$

is true.

Conversely suppose

$$(\exists \sigma_1 \in \mathrm{CSign}(\mathcal{P}))(\mathrm{Q}_2\sigma_2 \in \sigma_1)\ldots \ \ (\mathrm{Q}_\mathrm{k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\ \mathrm{F}^\star(\sigma_\mathrm{k})$$

is true and choose $\sigma_1 \in \mathrm{CSign}(\mathcal{P})$ such that

$$(\mathrm{Q}_2\sigma_2 \in \sigma_1)\ldots \ \ (\mathrm{Q}_\mathrm{k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\ \mathrm{F}^\star(\sigma_\mathrm{k})$$

is true. By definition of $\mathrm{CSign}(\mathcal{P})$, $\sigma_1 = \mathrm{CSign}(\mathcal{P})(a)$ for some $a \in \mathrm{R}$, and hence

$$(\mathrm{Q}_2\sigma_2 \in \mathrm{CSign}(\mathcal{P}_\mathrm{a}))\ldots \ \ (\mathrm{Q}_\mathrm{k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\ \mathrm{F}^\star(\sigma_\mathrm{k})$$

is true. By induction hypothesis,

$$(\mathrm{Q}_2\mathrm{X}_2)\ldots \ \ (\mathrm{Q}_\mathrm{k}\mathrm{X}_\mathrm{k})\ \mathrm{F}(\mathrm{a},\ldots,\mathrm{X}_\mathrm{k})$$

is true. Thus

$$(\exists X_1)\ (\mathrm{Q}_2\mathrm{X}_2)\ldots \ \ (\mathrm{Q}_\mathrm{k}\mathrm{X}_\mathrm{k})\ \mathrm{F}(\mathrm{X}_1,\ldots,\mathrm{X}_\mathrm{k})$$

is true. □

Before giving a description of the Cylindrical Decision Algorithm, we explain how it works on the following example.

Example 2.2.26 We continue Example 2.2.20 to illustrate Proposition 2.2.24. We had determined

$$\mathrm{CSign}(\mathcal{P}) = \{\{\{\{1\},\{0,1\},\{0,1,-1\}\},\{\{1\},\{0,1\}\},\{\{1\}\}\}.$$

The formula

$$(\exists X_1)\ (\forall X_2)\ (\forall X_3)\ X_1^2 + X_2^2 + X_3^2 - 1 > 0$$

is certainly true since

$$(\exists \sigma_1 \in \mathrm{CSign}(\mathcal{P}))\ (\forall\ \sigma_2 \in \sigma_1)\ (\forall \sigma_3 \in \sigma_2)\ \sigma_3(P) = 1 :$$

take $\sigma_1 = \{\{1\}\}$. It is also the case that the formula

$$(\forall X_1)\ (\exists X_2)\ (\exists X_3)\ \ X_1^2 + X_2^2 + X_3^2 - 1 > 0$$

is true since

$$(\forall \sigma_1 \in \mathrm{CSign}(\mathcal{P}))\ (\exists \sigma_2 \in \sigma_1)\ (\exists \sigma_3 \in \sigma_2)\ \sigma_3(P) = 1.$$

The formula

$$(\forall X_1)\ (\exists X_2)\ (\exists X_3)\ \ X_1^2 + X_2^2 + X_3^2 - 1 = 0$$

is false since it is not the case that

$$(\forall \sigma_1 \in \mathrm{CSign}(\mathcal{P}))\ (\exists \sigma_2 \in \sigma_1)\ (\exists \sigma_3 \in \sigma_2)\ \sigma_3(P) = 0 :$$

take $\sigma_1 = \{\{1\}\}$ to obtain a counter-example. It is also easy to check that the formula

$$(\exists X_1)\ (\forall X_2)\ (\exists X_3)\ X_1^2 + X_2^2 + X_3^2 - 1 = 0$$

is false since it is not the case that

$$(\exists \sigma_1 \in \mathrm{CSign}(\mathcal{P}))\ (\forall \sigma_2 \in \sigma_1)\ (\exists \sigma_3 \in \sigma_2)\ \sigma_3(P) = 0.$$

We are ready for the Decision Algorithm using cylindrical decomposition. We consider a finite set $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$, where D is an ordered integral domain.

Algorithm 2.2.27 (Cylindrical Decision)

Input: *a finite set* $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$, *a* $\mathcal{P}$*-sentence*

$$\Phi = (\mathrm{Q_1 X_1})\ (\mathrm{Q_2 X_2}) \ldots\ (\mathrm{Q_k X_k})\ \mathrm{F(X_1, \ldots, X_k)},$$

where $F(X_1, \ldots, X_k)$ *is quantifier free,* $\mathrm{Q_i} \in \{\exists, \forall\}$.

Output: True *if* Φ *is true and* False *otherwise.*

Procedure:

Run Algorithm 2.2.15 (Cylindrical Decomposition) with input $X_1, \ldots, X_k$ *and* $\mathcal{P}$.

Extract $\mathrm{CSign}(\mathcal{P})$ *from the set of cylindrical sample points and the signs of the polynomials of* $\mathcal{P}$ *on the cells of* R^k *using Proposition 2.2.19.*

Trying all possibilities, decide whether

$$(\mathrm{Q}_1\sigma_1 \in \mathrm{CSign}(\mathcal{P}))\,(\mathrm{Q}_2\sigma_2 \in \sigma_1)\,\ldots(\mathrm{Q}_\mathrm{k}\sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\,\mathrm{F}^\star(\sigma_\mathrm{k}) = \mathrm{True},$$

which is clearly a finite verification.

Proof of correctness: Follows from Proposition 2.2.24. Note that the two first steps of the computation depend only on $\mathcal{P}$ and not on Φ. As noted before $\mathrm{CSign}(\mathcal{P})$ allows us to decide the truth or falsity of every $\mathcal{P}$-sentence. □

Complexity analysis: According to the complexity analysis of Algorithm 2.2.15 (Cylindrical Decomposition), the number of sample points output is $(sd)^{O(1)^k}$, so the total complexity is $(sd)^{O(1)^k}$ arithmetic operations in D. Note that the evaluations of the boolean formulas are not counted in this model of complexity since we count only arithmetic operations in D. □

2.2.3 Quantifier elimination

Let us consider now quantifier elimination. We start by explaining that the set of points of R^ℓ at which a $\mathcal{P}$-formula Φ with free variables $Y_1, \ldots, Y_\ell$ is true, is a union of cells in R^ℓ of a cylindrical decomposition adapted to $\mathcal{P}$.

Indeed, let $\mathcal{P} \subset \mathrm{R}[Y_1, \ldots, Y_\ell, X_1, \ldots, X_k]$ and let $\mathcal{S}_1, \ldots, \mathcal{S}_{\ell+k}$ a cylindrical decomposition of $\mathrm{R}^{k+\ell}$ adapted to $\mathcal{P}$. Let $S \in \mathcal{S}_i$. We denote $\mathrm{CSign}_i(\mathcal{P})(y)$ for $y \in S$ by $\mathrm{CSign}_i(\mathcal{P})(S)$, using Proposition 2.2.21.

Let $\Phi(Y) = (\mathrm{Q_1X_1})\ (\mathrm{Q_2X_2})\dots\ (\mathrm{Q_kX_k})\mathrm{F}(\mathrm{Y_1},\dots,\mathrm{Y}_\ell,\mathrm{X_1},\dots,\mathrm{X_k})$, where $F(X_1,\dots,X_k)$ is quantifier free, $\mathrm{Q_i} \in \{\exists,\forall\}$, be a $\mathcal{P}$-formula. Let $\mathcal{L}$ is the union of cells S of $\mathcal{S}_\ell$ such that

$$(\mathrm{Q_1}\sigma_1 \in \mathrm{CSign}_\ell(\mathcal{P})(\mathrm{S}))\ (\mathrm{Q_2}\sigma_2 \in \sigma_1)\dots\ (\mathrm{Q_k}\ \sigma_\mathrm{k} \in \sigma_{\mathrm{k}-1})\ \mathrm{F}^\star(\sigma_\mathrm{k}) = \mathrm{True}.$$

Then

$$\mathcal{R}(\Phi,\mathrm{R}^\ell) = \{y \in \mathrm{R}^\ell \mid \Phi(y)\} = \mathcal{L}.$$

So we are not far from quantifier elimination.

However, a union of cells of a cylindrical decomposition in R^ℓ is not necessarily the realization of a $\mathrm{C}_{\leq\ell}(\mathcal{P})$-quantifier free formulas, where $\mathrm{C}_{\leq\ell}(\mathcal{P}) = \bigcup_{i\leq\ell}\mathrm{C}_i(\mathcal{P})$. So a cylindrical decomposition does not always provide a $\mathrm{C}_{\leq\ell}(\mathcal{P})$-quantifier free formula equivalent to Φ. We give an example of this situation:

Example 2.2.28 Continuing Example 2.2.14 b), we consider $\mathcal{P} = \{P,Q\}$ with

$$P = X_2^2 - X_1(X_1+1)(X_1-2)$$

and

$$Q = X_2^2 - (X_1+2)(X_1-1)(X_1-3).$$

We have seen in Example 2.2.14 b) that

$$\mathrm{C}_1(\mathcal{P}) = \{A,B,C\},$$

with

$$A(X_1) = \mathrm{sr}_0(P,\frac{\partial P}{\partial X_2}) = 4X_1(X_1+1)(X_1-2),$$

$$B(X_1) = \mathrm{sr}_0(Q,\frac{\partial Q}{\partial X_2}) = 4(X_1+2)(X_1-1)(X_1-3)$$

and

$$C(X_1) = \mathrm{sr}_0(P,Q) = (-X_1^2 - 3X_1 + 6)^2.$$

The zero sets of P and Q in $\mathbb{R}^2$ are two cubic curves with no intersection.

This can be checked algebraically. The roots of $(-X_1^2 - 3X_1 + 6)^2$, which is the resultant of P and Q, are $a = \frac{-3}{2} + \frac{1}{2}\sqrt{33}$ and $b = \frac{-3}{2} - \frac{1}{2}\sqrt{33}$.

Substituting these values in P and Q gives polynomials of degree 2 without real roots.

The only subset of $\mathbb{R}$ defined by sign conditions on $\mathrm{C}_1(\mathcal{P})$ are

$$\begin{gathered}
\{-1,0\} = \{x \in \mathbb{R} \mid A(x) = 0 \wedge B(x) > 0 \wedge C(x) > 0\},\\
(-1,0) \cup (3,+\infty) = \{x \in \mathbb{R} \mid A(x) > 0 \wedge B(x) > 0 \wedge C(x) > 0\},\\
(-2,-1) \cup (0,1) = \{x \in \mathbb{R} \mid A(x) < 0 \wedge B(x) > 0 \wedge C(x) > 0\},\\
\{3\} = \{x \in \mathbb{R} \mid A(x) > 0 \wedge B(x) = 0 \wedge C(x) > 0\},\\
\{-2,1\} = \{x \in \mathbb{R} \mid A(x) < 0 \wedge B(x) = 0 \wedge C(x) > 0\},\\
\{2\} = \{x \in \mathbb{R} \mid A(x) = 0 \wedge B(x) < 0 \wedge C(x) > 0\},\\
(2,3) = \{x \in \mathbb{R} \mid A(x) > 0 \wedge B(x) < 0 \wedge C(x) > 0\},\\
(-\infty,-2) \cup (1,2) \setminus \{a,b\} = \{x \in \mathbb{R} \mid A(x) < 0 \wedge B(x) < 0 \wedge C(x) > 0\},\\
\{a,b\} = \{x \in \mathbb{R} \mid A(x) < 0 \wedge B(x) < 0 \wedge C(x) = 0\}.
\end{gathered}$$

The set

$$\{x \in \mathbb{R} \mid \exists y \in \mathbb{R}\ P(x,y) < 0 \wedge Q(x,y) > 0\} = (2,+\infty)$$

is the union of semi-algebraically connected components of semi-algebraic sets defined by sign conditions on $\mathrm{C}_1(\mathcal{P})$, but is not defined by any $\mathrm{C}_1(\mathcal{P})$-quantifier free formula. There are $\mathcal{P}$-formulas whose realization set cannot be described by $\mathrm{C}_1(\mathcal{P})$-quantifier free formulas.

Fortunately, closing the set of polynomials under differentiation before each application of elimination of a variable provides an extended cylindrifying family whose realization of sign conditions are the cells of a cylindrical decomposition.

2.3 Existential theory of the reals

The **decision problem for the existential theory of the reals** is to decide the truth or falsity of a sentence

$$(\exists\ X_1)\ \ldots\ (\exists\ X_k)\ F(X_1,\ldots,X_k),$$

where $F(X_1,\ldots,X_k)$ is a quantifier free formula in the language of ordered fields with coefficients in a real closed field R. This problem

is equivalent to deciding whether or not a given semi-algebraic set is empty. It is a special case of the general decision problem.

When done by the Cylindrical Decomposition Algorithm, deciding the existential theory of the reals has complexity doubly exponential in k, the number of variables. But the existential theory of the reals has a special logical structure, since the sentence to decide has a single block of existential quantifiers. We take advantage of this special structure to find an algorithm which is singly exponential in k.

Our method for solving the existential theory of the reals is to compute the set of realizable sign conditions of the set of polynomials $\mathcal{P}$ appearing in the quantifier free formula F.

We first consider the case of a single polynomial.

2.3.1 One polynomial

We are going to describe a method for finding at least one point in every semi-algebraically connected component of an algebraic set.

The field of Puiseux series which os an important example of a non-archimedean real closed field containing R plays a key role in this method.

The collection of Puiseux series in ε with coefficients in R will be a real closed field containing the field $\mathrm{R}(\varepsilon)$ of rational functions in the variable ε ordered by 0_+ (see Notation **??**). In order to include in our field roots of equations such as $X^2 - \varepsilon = 0$, we introduce rational exponents such as $\varepsilon^{1/2}$. This partially motivates the following definitions.

Let ε a variable. The **ring of formal power series in ε with coefficients in** R, denoted $\mathrm{R}[[\varepsilon]]$, consists of series of the form

$$\overline{a} = \sum_{i \geq 0} a_i \varepsilon^i, \tag{2.3}$$

with $i \in \mathbb{N}$, $a_i \in \mathrm{R}$. Its field of quotients, denoted $\mathrm{R}((\varepsilon))$, is called the **field of formal Laurent series in ε with coefficients in** R and consists of series of the form

$$\overline{a} = \sum_{i \geq k} a_i \varepsilon^i, \tag{2.4}$$

with $k \in \mathbb{Z}$, $i \in \mathbb{Z}$, $a_i \in \mathrm{R}$.

A **Puiseux series in ε with coefficients in** R is a series of the form

$$\overline{a} = \sum_{i \geq k} a_i \varepsilon^{i/q}, \tag{2.5}$$

with $k \in \mathbb{Z}$, $i \in \mathbb{Z}$, $a_i \in \mathrm{R}$, q a positive integer. Puiseux series are formal Laurent series in the indeterminate $\varepsilon^{1/q}$ for some positive integer q. The **field of Puiseux series in ε with coefficients in** R is denoted $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$.

These series are formal in the sense that there is no assertion of convergence; ε is simply an indeterminate. We do assume that the different symbols ε^r, $r \in \mathbb{Q}$, satisfy

$\varepsilon^{r_1}\varepsilon^{r_2} = \varepsilon^{r_1+r_2}$,

$(\varepsilon^{r_1})^{r_2} = \varepsilon^{r_1 r_2}$,

$\varepsilon^0 = 1$.

Hence any two Puiseux series,

$$\overline{a} = \sum_{i \geq k_1} a_i \varepsilon^{i/q_1} \quad , \quad \overline{b} = \sum_{j \geq k_2} b_j \varepsilon^{j/q_2}, \tag{2.6}$$

can be written as formal Laurent series in $\varepsilon^{1/q}$, where q is the least common multiple of q_1 and q_2. Thus, it is clear how to add and multiply two Puiseux series. Also, any finite number of Puiseux series can be written as formal Laurent series in $\varepsilon^{1/q}$ with a common q.

If

$$\overline{a} = a_1\varepsilon^{r_1} + a_2\varepsilon^{r_2} + \ldots \in \mathrm{R}\langle\langle\varepsilon\rangle\rangle,$$

(with $a_1 \neq 0$ and $r_1 < r_2 < \ldots$), then the **order** of $\overline{a}$, denoted $o(\overline{a})$, is r_1. By convention, the order of 0 is ∞. This function from $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ to $\mathbb{Q} \cup \{\infty\}$ satisfies

$o(\overline{a}\overline{b}) = o(\overline{a}) + o(\overline{b})$.

$o(\overline{a} + \overline{b}) \geq \min(o(\overline{a}), o(\overline{b}))$, with equality if $o(\overline{a}) \neq o(\overline{b})$.

It is a straightforward exercise to verify that $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ is a field. We make $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ an ordered field by defining a Puiseux series $\overline{a}$ to be positive if the coefficient of $\varepsilon^{o(\overline{a})}$ is positive. It is clear that the field of rational functions $\mathrm{R}(\varepsilon)$ equipped with the order 0_+ is a subfield of the ordered field of Puiseux series $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$, using Laurent's expansions about 0.

In this order, ε is infinitesimal over R, since it is positive and smaller than any positive $r \in \mathrm{R}$, since $r - \varepsilon > 0$. Hence, $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ is a non-archimedean field. This is the reason why we have chosen to name the indeterminate ε rather than some more neutral X.

Theorem 2.3.1 *Let* R *be a real closed field. Then, the field* $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ *is real closed.*

We denote by $\mathrm{R}\langle\varepsilon\rangle$ the subfield of $\mathrm{R}\langle\langle\varepsilon\rangle\rangle$ of **algebraic Puiseux series**, which consists of those elements that are algebraic over $\mathrm{R}(\varepsilon)$, i.e. that satisfy a polynomial equation with coefficients in $\mathrm{R}(\varepsilon)$.

Corollary 2.3.2 *When* R *is real closed,* $\mathrm{R}\langle\varepsilon\rangle$ *is real closed. The field* $\mathrm{R}\langle\varepsilon\rangle$ *is the real closure of* $\mathrm{R}(\varepsilon)$ *equipped with the order* 0_+.

We first explain how to associate to a possibly unbounded algebraic set $Z \subset \mathrm{R}^k$ a bounded algebraic set $Z' \subset \mathrm{R}\langle\varepsilon\rangle^{k+1}$, whose semi-algebraically connected components are closely related to those of Z.

Let $Z = \mathrm{Z}(Q, \mathrm{R}^k)$ and consider

$$Z' = \mathrm{Z}(Q^2 + (\varepsilon^2(X_1^2 + \ldots + X_{k+1}^2) - 1)^2, \mathrm{R}\langle\varepsilon\rangle^{k+1}).$$

The set Z' is the intersection of the sphere S_ε^k of center 0 and radius $\dfrac{1}{\varepsilon}$ with a cylinder based on the extension of Z to $\mathrm{R}\langle\varepsilon\rangle$. The intersection of Z' with the hyperplane $X_{k+1} = 0$ is the intersection of Z with the sphere S_ε^{k-1} of center 0 and radius $\dfrac{1}{\varepsilon}$. Denote by π the projection from $\mathrm{R}\langle\varepsilon\rangle^{k+1}$ to $\mathrm{R}\langle\varepsilon\rangle^k$.

Proposition 2.3.3 *Let* N *be a finite number of points meeting every semi-algebraically connected component of* Z'. *Then* $\pi(N)$ *meets every semi- algebraically connected component of the extension* $\mathrm{Ext}(Z', \mathrm{R}\langle\varepsilon\rangle)$ *of* Z' *to* $\mathrm{R}\langle\varepsilon\rangle$.

Proof : Let D a semi-algebraically connected components of Z. If D is bounded, $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$ does not intersect S_ε^{k-1}, and $\pi^{-1}(\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle))$ is semi-algebraically homeomorphic to two copies of $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$, one in each half-space defined by $X_{k+1} = 0$. Thus, since N intersects every semi-algebraically connected component of Z', N intersects $\pi^{-1}(\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle))$ and $\pi(N)$ intersects $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$.

If D is unbounded, the set A of elements $r \in \mathrm{R}$ such that D intersects the sphere $S^{k-1}(0, r)$ of center 0 and radius r is semi-algebraic and unbounded and contains an open interval $(a, +\infty)$. Thus $\frac{1}{\varepsilon} \in \mathrm{Ext}(A, \mathrm{R}\langle\varepsilon\rangle)$, and $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$ intersects S_ε^{k-1}. Take $z \in \mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle) \cap S_\varepsilon^{k-1}$, and denote by D' be the semi-algebraically connected component of Z' containing $z' = (z, 0) \in Z'$. Take $x \in D' \cap N$ and consider a semi-algebraic path γ connecting z' to x inside D'. Then, $\pi(\gamma)$ is a semi-algebraic path connecting z to $\pi(x)$ inside Z, thus $\pi(x)$ and z belong to the same semi-algebraically connected component of $\mathrm{Ext}(Z, \mathrm{R}\langle\varepsilon\rangle)$. Since $z \in \mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$, $\pi(x) \in \mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$, and $\pi(N)$ intersects $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$. □

Let us illustrate this result. If $Q = X_2^2 - X_1(X_1 - 1)(X_1 + 1)$, then $Z = \mathrm{Z}(Q, \mathrm{R}^2)$ is a cubic curve with one bounded semi-algebraically connected component and one unbounded semi-algebraically connected component.

The corresponding $Z' \subset \mathrm{R}\langle\varepsilon\rangle^3$ has two semi-algebraically connected components above the bounded semi-algebraically connected of the cubic curve, and one semi-algebraically connected component above the unbounded semi-algebraically connected of the cubic curve.

So, if we have a method for finding a point in every semi-algebraically connected component of a bounded algebraic set, we obtain immediately, using Proposition 2.3.3, a method for finding a point in every connected component of an algebraic set. Note that these points have coordinates in the extension $\mathrm{R}\langle\varepsilon\rangle$ rather than in the real closed field R we started with. However, the extension from R to $\mathrm{R}\langle\varepsilon\rangle$ preserves semi-algebraically connected components.

We are going to define X_1-pseudo-critical points of $\mathrm{Z}(Q, \mathrm{R}^k)$ when $\mathrm{Z}(Q, \mathrm{R}^k)$ is a bounded algebraic set. These pseudo-critical points are a finite set of points meeting every semi-algebraically connected com-

ponent of $\mathrm{Z}(Q, \mathrm{R}^k)$. They are the limits of the critical points of the projection to the X_1 coordinate of a bounded nonsingular algebraic hypersurface defined by a particular infinitesimal perturbation of the polynomial Q. Moreover, the equations defining the critical points of the projection on the X_1 coordinate on the perturbed algebraic set have the special algebraic structure considered in Proposition **??**.

Given a polynomial $Q \in \mathrm{R}[X_1, \ldots, X_k]$ we define $\mathrm{tdeg}_{X_i}(Q)$, the **total degree of Q in X_i**, as the maximal total degree of the monomials in Q containing the variable X_i.

Notation 2.3.4 Let $\bar{d} = (\bar{d_1}, \ldots, \bar{d_k})$,

$$G_k(\bar{d}, c) = c^{\bar{d_1}}(X_1^{\bar{d_1}} + \cdots + X_k^{\bar{d_k}} + X_2^2 + \cdots + X_k^2) - (2k-1), \quad (2.7)$$

$$\mathrm{Def}(Q, \bar{d}, c, \zeta) = \zeta G_k(\bar{d}, c) + (1-\zeta)Q. \quad (2.8)$$

Note that $\forall\, x \in B(0, 1/c), G_k(\bar{d}, c)(x) < 0$.

In the next pages, the polynomial $Q \in \mathrm{D}[X_1, \ldots, X_k]$, where D is a ring contained in the real closed field R, and $(d_1, \ldots, d_k)$ satisfy the following conditions:

$Q(x) \geq 0$ for every $x \in \mathrm{R}^k$,

$\mathrm{Z}(Q, \mathrm{R}^k) \subset B(0, 1/c)$ for some $c \leq 1, c \in \mathrm{D}$,

$d_1 \geq d_2 \cdots \geq d_k$,

$\deg(Q) \leq d_1$, $\mathrm{tdeg}_{X_i}(Q) \leq d_i$, for $i = 2, \ldots, k$.

Remark 2.3.5 Note that supposing $Q(x) \geq 0$ for every $x \in \mathrm{R}^k$ is not a big loss of generality since we can always replace Q by Q^2 if it is not the case. Note also that we can always take

$$d_1 = \ldots = d_k = \deg(Q).$$

However considering different d_i will be useful when the degree with respect to some variables is small.

Let $\bar{d_i}$ be an even number $> d_i, i = 1, \ldots, k$, and $\bar{d} = (\bar{d_1}, \ldots, \bar{d_k})$.

Proposition 2.3.6 *The algebraic set* $\mathrm{Z}(\mathrm{Def}(Q,\bar{d},c,\zeta),\mathrm{R}\langle\zeta\rangle^k)$ *is a non-singular algebraic hypersurface bounded over* R *contained in* $B(0,1/c)$*, and*

$$\lim_\zeta(\mathrm{Z}(\mathrm{Def}(Q,\bar{d},c,\zeta),\mathrm{R}\langle\zeta\rangle^k)) = \mathrm{Z}(Q,\mathrm{R}^k).$$

Notation 2.3.7 Let $\bar{d} = (\bar{d}_1,\ldots,\bar{d}_k)$, and using Notation 2.3.4, consider the polynomial system

$$\mathrm{Cr}(Q,\bar{d},c,\zeta) = \left\{\mathrm{Def}(Q,\bar{d},c,\zeta), \frac{\partial \mathrm{Def}(Q,\bar{d},c,\zeta)}{\partial X_2},\ldots,\frac{\partial \mathrm{Def}(Q,\bar{d},c,\zeta)}{\partial X_k}\right\}. \tag{2.9}$$

Definition 2.3.8 An **X_1-pseudo-critical point** on $\mathrm{Z}(Q,\mathrm{R}^k)$ is the $\lim_\zeta$ of an X_1-critical point on $\mathrm{Z}(\mathrm{Def}(Q,\bar{d},c,\zeta),\mathrm{R}\langle\zeta\rangle^k)$.

An **X_1-pseudo-critical value** on $\mathrm{Z}(Q,\mathrm{R}^k)$ is the projection to the X_1-axis of an X_1-pseudo-critical point on $\mathrm{Z}(Q,\mathrm{R}^k)$.

According to Definition 2.3.8, an X_1-pseudo-critical point of $\mathrm{Z}(Q,\mathrm{R}^k)$ is the $\lim_\zeta$ of an X_1-critical point on

$$\mathrm{Z}(\mathrm{Def}(Q,\bar{d},c,\zeta),\mathrm{R}\langle\zeta\rangle^k).$$

Proposition 2.3.9 *The set of* X_1 *-pseudo-critical points on* $\mathrm{Z}(Q,\mathrm{R}^k)$ *meets every semi-algebraically connected component of* $\mathrm{Z}(Q,\mathrm{R}^k)$.

Moreover, the polynomial system $\mathrm{Cr}(Q,\bar{d},c,\zeta)$ has good algebraic properties.

Proposition 2.3.10 *1. The polynomial system* $\mathrm{Cr}(Q,\bar{d},c,\zeta)$ *is a Gröbner basis of* $\mathrm{I}(\mathrm{Cr}(Q,\bar{d},c,\zeta),\mathrm{R}\langle\zeta\rangle)$ *for the graded lexicographical ordering with* $X_1 >_{\mathrm{grlex}} \cdots >_{\mathrm{grlex}} X_k$.

2. The set $\mathrm{Z}(\mathrm{Cr}(Q,\bar{d},c,\zeta),\mathrm{C}\langle\zeta\rangle^k)$ *is finite.*

3. The zeros of the polynomial system $\mathrm{Cr}(Q,\bar{d},c,\zeta)$ *are simple.*

We are now ready to describe an algorithm giving a point in every connected component of a bounded algebraic set. We simply compute pseudo-critical values and their limits.

Algorithm 2.3.11 (Bounded Algebraic Sampling)

Input : *a polynomial* $Q \in \mathrm{D}[X_1, \ldots, X_k]$ *such that* $Q(x) \geq 0$ *for every* $x \in \mathrm{R}^k$ *and such that* $\mathrm{Z}(Q, \mathrm{R}^k)$ *is contained in* $B(0, c)$.

Output : *a set* $\mathcal{U}$ *of real univariate representations of the form*

$$(f, g_0, \ldots, g_k), \sigma,$$

with $\{f, g_0, \ldots, g_k\} \subset \mathrm{D}[T]^{k+2}$. *The set of points associated to these univariate representations meets every semi-algebraically connected component of* $\mathrm{Z}(Q, \mathrm{R}^k)$ *and contains the set of* X_1*-pseudo-critical points on* $\mathrm{Z}(Q, \mathrm{R}^k)$.

Procedure :

Choose $(d_1, \ldots, d_k)$ *such that* $d_1 \geq \ldots \geq d_k$, $\deg(Q) \leq d_1$, $\mathrm{tdeg}_{X_i}(Q) \leq d_i$, *for* $i = 2, \ldots, k$. *Take as* $\bar{d}_i$ *the smallest even number* $> d_i, i = 1, \ldots, k$, $\bar{d} = (\bar{d}_1, \ldots, \bar{d}_k)$.

Compute the multiplication table $\mathcal{M}$ *of* $\mathrm{Cr}(Q, \bar{d}, c, \zeta)$.

Apply the $\lim_\zeta$ *map with input* $\mathcal{M}$, *and obtain a set* $\mathcal{U}$ *of real univariate representations* v *with*

$$v = (f(T), g_0(T), \ldots, g_k(T)), \sigma$$
$$\{f(T), g_0(T), \ldots, g_k(T)\} \subset \mathrm{D}[T]^{k+2}.$$

Sketch of complexity analysis: See detailed analysis in[1]. The complexity is $(d_1 \ldots d_k)^{O(1)}$ in the ring D. The polynomials output are of degree $O(d_1) \ldots O(d_k)$ in T. □

Algorithm 2.3.12 (Algebraic Sampling)

Input : *a polynomial* $Q \in \mathrm{D}[X_1, \ldots, X_k]$.

Output : *a set* $\mathcal{U}$ *of real univariate representations of the form*

$$(f, g_0, \ldots, g_k), \sigma,$$

with $\{f, g_0, \ldots, g_k\} \subset \mathrm{D}[\varepsilon][T]^{k+2}$. *The set of points associated to these univariate representations meets every semi-algebraically connected component of* $\mathrm{Z}(Q, \mathrm{R}\langle\varepsilon\rangle^k)$.

Procedure :

Define

$$R := Q^2 + (\varepsilon(X_1^2 + \ldots + X_{k+1}^2) - 1)^2.$$

Apply Algorithm 2.3.11 (Bounded Algebraic Sampling) to R, and obtain a set $\mathcal{V}$ of real univariate representations v with

$$v = (f(T), g_0(T), \ldots, g_k(T), g_{k+1}(T)), \sigma$$
$$\{f(T), g_0(T), \ldots, g_k(T), g_{k+1}(T)\} \subset \mathrm{D}[\varepsilon][T]^{k+3}.$$

Define $\pi(v)$ by (u), with

$$u = (f(T), g_0(T), \ldots, g_k(T)), \sigma$$
$$\{f(T), g_0(T), \ldots, g_k(T)\} \subset \mathrm{D}[\varepsilon][T]^{k+2},$$

and $\mathcal{U} = \pi(\mathcal{V})$.

Sketch of omplexity analysis: See detailed analysis in [1]. The complexity is $(d_1 \ldots d_k)^{O(1)}$ in the ring $\mathrm{D}[\varepsilon]$. The polynomials output are of degree $O(d_1) \ldots O(d_k)$ in T. Moreover the degrees with respect to ε occurring in the computations of the multiplication table are bounded by

$$O(d_1 + \ldots + d_{k-1})kd_k$$

. □

2.3.2 Several polynomials

Let $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_s\} \subset \mathrm{R}[X_1, \ldots, X_k]$. Recall that we denote by $\mathrm{Sign}(\mathcal{P}) \subset \{0, 1, -1\}^{\mathcal{P}}$ the set of all realizable sign conditions for $\mathcal{P}$. We are now going to present an algorithm which computes $\mathrm{Sign}(\mathcal{P})$.

We first prove that we can reduce the problem of computing a set of sample points meeting the realizations of every realizable sign conditions of a family of polynomials to the problem of finding points in every semi-algebraically connected component of certain algebraic sets.

Proposition 2.3.13 *Let $D \subset \mathrm{R}^k$ be a non-empty semi-algebraically connected component of a basic closed semi-algebraic set defined by*

$$P_1 = \cdots = P_\ell = 0, P_{\ell+1} \geq 0, \cdots, P_s \geq 0.$$

There exists an algebraic set W defined by equations

$$P_1 = \cdots = P_\ell = P_{i_1} = \cdots P_{i_m} = 0,$$

(with $\{i_1, \ldots, i_m\} \subset \{\ell+1, \ldots, s\}$) such that a semi-algebraically connected component D' of W is contained in D.

Proof: Consider a maximal set of polynomials

$$\{P_1, \ldots, P_\ell, P_{i_1}, \ldots, P_{i_m}\},$$

where

$$m = 0 \text{ or } \ell < i_1 < \cdots < i_m \leq s,$$

with the property that there exists a point $p \in D$ where

$$P_1 = \cdots = P_\ell = P_{i_1} = \cdots = P_{i_m} = 0.$$

Consider the semi-algebraically connected component D' of the algebraic set defined by

$$P_1 = \cdots = P_\ell = P_{i_1} = \cdots = P_{i_m} = 0,$$

which contains p. We claim that $D' \subset D$. Suppose that there exists a point $q \in D'$ such that $q \notin D$. Then by Proposition **??**, there exists a semi-algebraic path $\gamma : [0,1] \to D'$ joining p to q in D'. Denote by q' the first point of the path γ on the boundary of D. More precisely, note that

$$A = \{t \in [0,1] \mid \gamma([0,t]) \subset D\}$$

is a closed semi-algebraic subset of $[0,1]$ which does not contain 1. Thus A is the union of a finite number of closed intervals

$$A = [0, b_1] \cup \ldots \cup [a_\ell, b_\ell].$$

Take $q' = \gamma(b_1)$. At least one of the polynomials, say P_j, $j \notin \{1, \ldots, \ell, i_1, \ldots, i_m\}$ must be 0 at q'. This violates the maximality of the set

$$\{P_1, \ldots, P_\ell, P_{i_1}, \ldots, P_{i_m}\}.$$

It is clear that if D is bounded, D' is bounded. □

Corollary 2.3.14 *Let $D \subset \mathrm{R}^k$ be a non-empty semi-algebraically connected component of a semi-algebraic set defined by*

$$P_1 = \cdots = P_\ell = 0, P_{\ell+1} > 0, \cdots, P_s > 0.$$

There exists an algebraic set $W \subset \mathrm{R}\langle\varepsilon\rangle^k$ defined by equations

$$P_1 = \cdots = P_\ell = 0, P_{i_1} = \cdots P_{i_m} = \varepsilon$$

(with $\{i_1, \ldots, i_m\} \subset \{\ell+1, \ldots, s\}$) such that there exists a semi-algebraically connected component D' of W which is contained in $\mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$.

Proof : Consider two points x and y in D. By Proposition **??**, there is a semi-algebraic path γ from x to y inside D. Since γ is closed and bounded, the semi-algebraic and continuous function $\min_{\ell+1 \leq i \leq s}(P_i)$ has a strictly positive minimum on γ. The extension of the path γ to $\mathrm{R}\langle\varepsilon\rangle$ is thus entirely contained inside the subset S of $\mathrm{R}\langle\varepsilon\rangle^k$ defined by

$$P_1 = \cdots = P_\ell = 0, P_{\ell+1} - \varepsilon \geq 0, \cdots, P_s - \varepsilon \geq 0.$$

Thus, there is only one non-empty semi-algebraically connected component $\bar{D}$ of S containing D. Applying Proposition 2.3.13 to $\bar{D}$ and S, we get a semi-algebraically connected component D' of some

$$P_1 = \cdots = P_\ell = 0, P_{i_1} = \cdots P_{i_m} = \varepsilon,$$

contained in $\bar{D}$. Then $D' \subset \mathrm{Ext}(D, \mathrm{R}\langle\varepsilon\rangle)$. □

Remark 2.3.15 Corollary 2.3.14 and Algorithm 2.3.11 (Bounded Algebraic Sampling), provides an algorithm outputting a set of points meeting every semi-algebraically connected component of the realization of a realizable sign condition of a family $\mathcal{P}$ of s polynomials on a bounded algebraic set $\mathrm{Z}(Q, \mathrm{R}^k)$ with complexity $2^s d^{O(k)}$ (where d is a bound on the degree of Q and the $P \in \mathcal{P}$), considering all possible subsets of $\mathcal{P}$. Note that this algorithm does not involve polynomials of degree doubly exponential in k, in contrast to Algorithm 2.2.15 (Cylindrical Decomposition).

When s is bigger than the dimension k of the ambient space, the algorithm proposed in the preceding remark does not give a satisfactory complexity bound, since the complexity is exponential in s. Reduction to general position, using infinitesimal deformations, will be the key for a better complexity result.

Let us define precisely the notion of general position that we consider. Let

$$\mathcal{P}^\star = \{\mathcal{P}_1^\star, \ldots, \mathcal{P}_s^\star\},$$

where for every $i = 1, \ldots, s$, $\mathcal{P}_i^\star \subset \mathrm{R}[X_1, \ldots, X_k]$ is finite, and such that two distinct elements of $\mathcal{P}_i^\star$ have no common zeros in R^k. The family $\mathcal{P}^\star$ is in **ℓ-general position** with respect to $Q \in \mathrm{R}[X_1, \ldots, X_k]$ in R^k if no $\ell+1$ polynomials belonging to different $\mathcal{P}_i^\star$ have a zero in common with Q in R^k.

The family $\mathcal{P}^\star$ is in **strong ℓ-general position** with respect to $Q \in \mathrm{R}[X_1, \ldots, X_k]$ in R^k if moreover any ℓ polynomials belonging to different $\mathcal{P}_i^\star$ have at most a finite number of zeros in common with Q in R^k.

When $Q = 0$, we simply say that $\mathcal{P}^\star \subset \mathrm{R}[X_1, \ldots, X_k]$ is in ℓ-general position (respectively strong ℓ-general position) in R^k.

We also need the notion of a family of homogeneous polynomials in general position in $\mathbb{P}_k(\mathrm{C})$. The reason for considering common zeros in $\mathbb{P}_k(\mathrm{C})$ is that we are going to use in our proofs the fact that, in the context of complex projective geometry, the projection of an algebraic set is algebraic [1].

Let

$$\mathcal{P}^\star = \{\mathcal{P}_1^\star, \ldots, \mathcal{P}_s^\star\},$$

where for every $i = 1, \ldots, s$, $\mathcal{P}_i^\star \in \mathrm{R}[X_0, X_1, \ldots, X_k]$ is homogeneous. The family $\mathcal{P}^\star$ is in **ℓ-general position** with respect to a homogeneous polynomial $Q^h \in \mathrm{R}[X_0, X_1, \ldots, X_k]$ in $\mathbb{P}_k(\mathrm{C})$ if no more than ℓ polynomials of $\mathcal{P}_i^\star$ have a zero in common with Q^h in $\mathbb{P}_k(\mathrm{C})$.

We first give an example of a finite family of polynomials in general position and then explain how to perturb a finite set of polynomials to get a family in strong general position.

Notation 2.3.16 Define

$$H_k(d,i) = 1 + \sum_{1 \leq j \leq k} i^j X_j^d,$$
$$H_k^h(d,i) = X_0^d + \sum_{1 \leq j \leq k} i^j X_j^d.$$

Note that when d is even, $H_k(d,i)(x) > 0$ for every $x \in \mathrm{R}^k$.

Lemma 2.3.17 *For any positive integer d, the polynomials $H_k^h(d,i)$, $0 \leq i \leq s$, are in k-general position in $\mathbb{P}_k(\mathrm{C})$.*

Proof: Take $P(T, X_0, \ldots, X_k) = {X_0}^d + \sum_{1 \leq j \leq k} T^j X_j^d$. If $k+1$ of the $H_k^h(d,i)$ had a common zero $\bar{x}$ in $\mathbb{P}_k(\mathrm{C})$, substituting homogeneous coordinates of this common zero in P would give a nonzero univariate polynomial in T of degree at most k with $k+1$ distinct roots, which is impossible. □

Consider three variables $\varepsilon, \delta, \gamma$ and $\mathrm{R}\langle\varepsilon,\delta,\gamma\rangle$. Note that $\varepsilon, \delta, \gamma$ are three infinitesimals in $\mathrm{R}\langle\varepsilon,\delta,\gamma\rangle$ with $\varepsilon > \delta > \gamma > 0$. The reason for using these three infinitesimals is the following. The variable ε is used to get bounded sets, the variables δ, γ are used to reach general position, and describe sets which are closely related to realizations of sign conditions on the original family.

Let $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{R}[X_1, \ldots, X_k]$ be polynomials of degree bounded by d. With $d' > d$, let $\mathcal{P}^\star$ be the family $\{P_1^\star, \ldots, P_s^\star\}$ with

$$P_i^\star = \{(1-\delta)P_i + \delta H_k(d',i), (1-\delta)P_i - \delta H_k(d',i),$$
$$(1-\delta)P_i + \delta\gamma H_k(d',i), (1-\delta)P_i - \delta\gamma H_k(d',i)\}.$$

We prove

Proposition 2.3.18 *The family $\mathcal{P}^\star$ is in strong k-general position in $\mathrm{R}\langle\epsilon,\delta,\gamma\rangle^k$.*

There is a close relationship between the sign conditions on $\mathcal{P}$ and certain weak sign conditions on the polynomials of $\mathcal{P}^*$ described by the following proposition. The role of the two infinitesimals δ and γ is the following: δ is used to replace strict inequalities by weak inequalities and γ to replace equations by weak inequalities.

Proposition 2.3.19 *Let $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{R}[X_1, \ldots, X_k]$ be such that $\deg P_i \leq d$ for all i, and suppose $d' > d$, d' even. Let $D \subset \mathrm{R}^k$ be a semi-algebraically connected component of the realization of the sign condition*

$$P_i = 0, i \in I \subset \{1, \ldots, s\},$$
$$P_i > 0, i \in \{1, \ldots, s\} \setminus I.$$

Then there exists a semi-algebraically connected component D' of the subset $\bar{D} \subset \mathrm{R}\langle\varepsilon, \delta, \gamma\rangle^k$ defined by the weak sign condition

$$-\gamma\delta H_k(d', i) \leq (1-\delta)P_i \leq \gamma\delta H_k(d', i),\ i \in I,$$

$$(1-\delta)P_i \geq \delta H_k(d', i),\ i \in \{1, \ldots, s\} \setminus I$$

$$\varepsilon^2(X_1^2 + \cdots + X_k^2) \leq 1$$

such that $\lim_\gamma(D')$ is contained in the extension of D to $\mathrm{R}\langle\varepsilon, \delta\rangle$.

Corollary 2.3.20 *Let $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{R}[X_1, \ldots, X_k]$ be a finite subset of polynomials of degree less than d and suppose $d' > d$, d' even. Let D be a semi-algebraically connected component of the realization of the sign condition*

$$P_i = 0, i \in I \subset \{1, \ldots, s\}$$
$$P_i > 0, i \in \{1, \ldots, s\} \setminus I.$$

Then there exists a semi-algebraically connected component E' of the realization $E \subset \mathrm{R}\langle\varepsilon, \delta, \gamma\rangle^{k+1}$ of

$$-\gamma\delta H_k(d', i) \leq (1-\delta)P_i \leq \gamma\delta H_k(d', i),\ 1i \in \{1, \ldots, s\} \setminus I \in I,$$

$$(1-\delta)P_i \geq \delta H_k(d', i),$$

$$\varepsilon^2(X_1^2 + \cdots + X_k^2 + X_{k+1}^2) = 1$$

such that $\Pi(\lim_\gamma(E'))$ is contained in the extension of D to $\mathrm{R}\langle\varepsilon, \delta\rangle$, where Π is the projection of R^{k+1} to R^k forgetting the last coordinate.

As a consequence of Corollary 2.3.20, in order to compute all realizable sign conditions on $\mathcal{P}$ it will be enough, using Proposition 2.3.13 and Proposition 2.3.18, to consider equations of the form

$$Q = {Q_{i_1}}^2 + \cdots + {Q_{i_j}}^2 + (\varepsilon^2(X_1^2 + \cdots + X_k^2 + X_{k+1}^2) - 1)^2 = 0,$$

where $j \leq k, Q_{i_1} \in P_{i_1}^\star, \ldots, 1 \leq i_1 < \ldots < i_j \leq s, Q_{i_j} \in P_{i_j}^\star$, to find a point in each of the semi-algebraically connected components of their zero sets and to take their limit under $\lim_\gamma$.

A finite set $\mathcal{S} \subset \mathrm{R}^k$ is a **set of sample points for** $\mathcal{P}$ in R^k if $\mathcal{S}$ meets the realizations of all $\sigma \in \mathrm{Sign}(\mathcal{P})$ (Notation **??**). Note that the sample points output by Algorithm 2.2.15 (Cylindrical Decomposition) are a set of sample points for $\mathcal{P}$ in R^k, since the cells of a cylindrical decomposition of R^k adapted to $\mathcal{P}$ are $\mathcal{P}$ invariant and partition R^k. We are going to produce a set of sample points much smaller than the one output by Algorithm 2.2.15 (Cylindrical Decomposition), which was doubly exponential in the number of variables.

Algorithm 2.3.21 (Computing Realizable Sign Conditions)

Input: *a set of s polynomials,*

$$\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{D}[X_1, \ldots, X_k],$$

each of degree at most d.

Output: *a set of real univariate representations in* $\mathrm{D}[\varepsilon, \delta, T]^{k+2}$ *such that the associated points form a set of sample points for* $\mathcal{P}$ *in* $\mathrm{R}\langle\varepsilon, \delta\rangle^k$*, meeting every semi-algebraically connected component of* $\mathcal{R}(\sigma)$ *for every* $\sigma \in \mathrm{Sign}(\mathcal{P})$ *and the signs of the elements of* $\mathcal{P}$ *at these points.*

Procedure:

Initialize $\mathcal{U}$ *to the empty set.*

Take as d' *the smallest even natural number* $> d$.

Define

$$P_i^\star = \{(1-\delta)P_i + \delta H_k(d', i), (1-\delta)P_i - \delta H_k(d', i),$$
$$(1-\delta)P_i + \delta\gamma H_k(d', i), (1-\delta)P_i - \delta\gamma H_k(d', i)\}$$
$$\mathcal{P}^\star = \{P_1^\star, \ldots, P_s^\star\}$$

for $0 \leq i \leq s$, *using Notation 2.3.16.*

For every subset of $j \leq k$ *polynomials* $Q_{i_1} \in P_{i_1}^\star, \ldots, Q_{i_j} \in P_{i_j}^\star$,

Let

$$Q = {Q_{i_1}}^2 + \cdots + {Q_{i_j}}^2 + (\varepsilon^2(X_1^2 + \cdots + X_k^2 + X_{k+1}^2) - 1)^2.$$

For $i = 1, \ldots, k$, *let* $\bar{d}_i$ *be the smallest even natural number* $> \deg(Q)$, $i = 1, \ldots, k$, *and let* $\bar{d}_{k+1} = 6$, $\bar{d} = (\bar{d}_1, \ldots, \bar{d}_k, \bar{d}_{k+1})$.

Compute the multiplication table $\mathcal{M}$ *of* $\overline{\mathrm{Cr}}(Q, \bar{d}, \varepsilon, \zeta)$ *(Notation* **??***).*

Apply the $\lim_{\gamma,\zeta}$ *map with input* $\mathcal{M}$, *and obtain a set of real univariate representations* (v, σ) *with*

$$v = (f(T), g_0(T), \ldots, g_k(T), g_{k+1}(T)) \in \mathrm{D}[\varepsilon, \delta][T]^{k+3}.$$

Ignore $g_{k+1}(T)$ *and consider only the real univariate representations* (u, σ)

$$u = (f(T), g_0(T), \ldots, g_k(T)) \in \mathrm{D}[\varepsilon, \delta][T]^{k+2}.$$

Add u *to* $\mathcal{U}$.

Compute the signs of $P \in \mathcal{P}$ *at the points associated to the real univariate representations in* $\mathcal{U}$, *with input* f *and its derivatives and the* P_u, $P \in \mathcal{P}$.

Sketch of complexity analysis : See detailed proof in [1]. The total number of $j \leq k$-tuples examined is $\sum_{j \leq k} 4^j \binom{s}{j}$. Each such call costs

$d^{O(k)}$ arithmetic operations in $\mathrm{D}[\varepsilon, \delta, \gamma, \zeta]$. Since there is a fixed number of infinitesimals appearing with degree one in the input equations, the number of arithmetic operations in D is also $d^{O(k)}$. Thus the total number of real univariate representations produced is bounded by $\sum_{j \leq k} 4^j \binom{s}{j} O(d)^k$, while the number of arithmetic operations performed for outputting sample points in $\mathrm{R}\langle\varepsilon, \delta\rangle^k$, is bounded by

$$\sum_{j \leq k} 4^j \binom{s}{j} d^{O(k)} = s^k d^{O(k)}.$$

The sign determination takes

$$s \sum_{j \leq k} 4^j \binom{s}{j} d^{O(k)} = s^{k+1} d^{O(k)}$$

arithmetic operations. □

Bibliography

[1] S. Basu, R. Pollack, M.-F. Roy, *Algorithms in real algebraic geometry*, Springer, (2003).

[2] F. Budan de Boislaurent, *Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque*, (1807), 2nd edition, Paris (1822).

[3] G. Collins, *Quantifier elimination for real closed fields by cylindric algebraic decomposition*, In Second GI Conference on Automata Theory and Formal Languages. Lecture Notes in Computer Science, vol. 33, pp. 134-183, Springer- Verlag, Berlin (1975).

[4] R. Descartes, *Géométrie* (1636). A source book in Mathematics, 90-31. Harvard University press (1969).

[5] G. Farin, *Curves and surfaces for Computer Aided Design*, Academic Press (1990).

[6] J. Fourier, *Analyse des équations déterminées*, F. Didot, Paris (1831).

[7] L. Gonzalez Vega, *La sucesión de Sturm–Habicht y sus aplicaciones al Algebra Computacional*, Doctoral Thesis, Universidad de Cantabria (1989).

[8] W. Habicht, *Eine Verallgemeinerung des Sturmschen Wurzelzählverfahrens,* Comm. Math. Helvetici 21, 99-116 (1948).

[9] B. Mourrain, M. N. Vrahatis, J.-C. Yakhoubson *On the Complexity of Isolating Real Roots and Computing with Certainty*

the Topological Degree, Journal of Complexity, 182, 612–640 (2002).

[10] A. SEIDENBERG, *A new decision method for elementary algebra*, Annals of Mathematics, 60:365–374, (1954).

[11] C. STURM, *Mémoire sur la résolution des équations numériques.* Inst. France Sc.Math. Phys.6 (1835).

[12] J. J. SYLVESTER, *On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's function.* Trans. Roy. Soc. London (1853).

[13] A. TARSKI, *A Decision method for elementary algebra and geometry*, University of California Press (1951).