

Quantitative Fundamental Theorem of Algebra

Daniel Perrucci^{b*} Marie-Françoise Roy[#]

^b Departamento de Matemática, FCEN, Universidad de Buenos Aires
and IMAS UBA-CONICET,

Ciudad Universitaria, 1428 Buenos Aires, Argentina

[#] IRMAR (UMR CNRS 6625), Université de Rennes 1,

Campus de Beaulieu, 35042 Rennes Cedex, France

March 8, 2019

Abstract

Using subresultants, we modify a real-algebraic proof due to Eisermann of the Fundamental Theorem of Algebra ([FTA]) to obtain the following quantitative information: in order to prove the [FTA] for polynomials of degree d , the Intermediate Value Theorem ([IVT]) is required to hold only for real polynomials of degree at most d^2 . We also explain that the classical proof due to Laplace requires [IVT] for real polynomials of exponential degree. These quantitative results highlight the difference in nature of these two proofs.

Keywords: Fundamental Theorem of Algebra, Intermediate Value Theorem, Cauchy Index, Winding Number, Subresultant Polynomials, Sturm Chains.

AMS subject classifications: 14P99, 12D10, 12D15.

1 Introduction

Let (\mathbf{R}, \leq) be an ordered field. The fact that \mathbf{R} admits an order compatible with the field structure implies that $\text{char}(\mathbf{R}) = 0$ and therefore \mathbf{R} has an infinite number of elements. It also implies that -1 is not a square in \mathbf{R} and, consequently, $\mathbf{R}[T]/\langle T^2 + 1 \rangle = \mathbf{R}[i] = \mathbf{C}$ is an algebraic field extension of \mathbf{R} of degree 2.

We consider the following properties on (\mathbf{R}, \leq) .

*Partially supported by the Argentinian grants UBACYT 20020160100039BA and PIP 11220130100527CO CONICET

- [IVT] (*Intermediate Value Theorem*): for every polynomial $F \in \mathbf{R}[X]$ and every a, b in \mathbf{R} with $a < b$ and $F(a)F(b) < 0$, there exists $c \in \mathbf{R}$ with $a < c < b$ such that $F(c) = 0$.
- [NnS] (*Non-negative elements are Squares*): for every $a \in \mathbf{R}$ with $a \geq 0$, there exists $c \in \mathbf{R}$ such that $a = c^2$.
- [OD] (*An odd degree polynomial has a root*): for every polynomial $F \in \mathbf{R}[X]$ of odd degree, there exists $c \in \mathbf{R}$ such that $F(c) = 0$.
- [FTA] (*Fundamental Theorem of Algebra*): for every polynomial $F \in \mathbf{C}[Z] \setminus \mathbf{C}$, there exists $z \in \mathbf{C}$ such that $F(z) = 0$ (i.e., \mathbf{C} is an algebraically closed field).

If \leq denotes the usual order over the real numbers, (\mathbb{R}, \leq) and $(\mathbb{R}_{\text{alg}}, \leq)$ are typical examples of ordered fields satisfying all the properties above (where \mathbb{R}_{alg} is the set of real algebraic numbers), whereas (\mathbb{Q}, \leq) is a typical example of an ordered field satisfying none of the properties above.

The next theorem is a classical result in real algebraic geometry (see for instance [2, Chapter 1]).

Theorem 1 *Let (\mathbf{R}, \leq) be an ordered field. The following conditions are equivalent:*

- (\mathbf{R}, \leq) satisfies [IVT].
- (\mathbf{R}, \leq) satisfies [NnS] and [OD].
- (\mathbf{R}, \leq) satisfies [FTA].

If (\mathbf{R}, \leq) satisfies these conditions, then it is easy to see that the field order \leq on \mathbf{R} is unique and \mathbf{R} is said to be a **real closed field**.

We sketch briefly a proof of Theorem 1, which is essentially Laplace's proof [6].

Sketch of the proof of Theorem 1: Proving that [IVT] implies [NnS] is very simple: for $a = 0$ we take $c = 0$; and for $a > 0$ we consider the polynomial $F := X^2 - a \in \mathbf{R}[X]$ and notice that $F(0) < 0$ and $F(a + 1) > 0$, then [IVT] ensures the existence of a $c \in \mathbf{R}$ such that $F(c) = 0$, or equivalently, $a = c^2$. In fact, adding the condition $c \geq 0$, it is easy to prove the uniqueness of such c .

In a similar way, in order to prove that [IVT] implies [OD] we only need to note that an odd degree polynomial necessarily changes its sign when evaluated at a and $-a$ with $a \in \mathbf{R}$ big enough.

The proof that [NnS] and [OD] imply [FTA] is much more sophisticated. To prove that $F \in \mathbf{C}[Z] \setminus \mathbf{C}$ has a root in \mathbf{C} , it is enough to prove that the polynomial $F\bar{F} \in \mathbf{R}[Z] \setminus \mathbf{R}$ has a root w in \mathbf{C} (where \bar{F} means the polynomial obtained from F by usual conjugation in \mathbf{C} of the coefficients of F); in this case either w or \bar{w} is a root of F . Now, in order to show that an arbitrary polynomial $G \in \mathbf{R}[Z] \setminus \mathbf{R}$ of degree d has a root in \mathbf{C} , the proof proceeds by induction on the highest value of k such that 2^k divides d . In the base case, which is $k = 0$ (and therefore odd d), the existence of a root of G in $\mathbf{R} \subset \mathbf{C}$ is ensured by [OD]. For $k \geq 1$ (and therefore even d), the existence of a root of G in \mathbf{C} is

ensured by a clever argument involving [NnS] and the fact that every polynomial in $\mathbf{R}[Z]$ of degree $\binom{d}{2}$ has a root in \mathbf{C} . Note that the highest power of 2 dividing $\binom{d}{2} = \frac{1}{2}d(d-1)$ is 2^{k-1} and then the inductive hypothesis holds.

Finally, assuming [FTA], it is possible to prove that the irreducible elements in the unique factorization domain $\mathbf{R}[X]$ have degree 1 or 2 and that the irreducible monic elements in $\mathbf{R}[X]$ of degree 2 are positive when evaluated at any $r \in \mathbf{R}$. From these facts, [IVT] holds easily. \square

The main concern in the present work is the following question: assuming that [IVT] holds for (\mathbf{R}, \leq) , if we take a fixed value of $d \in \mathbb{Z}_{\geq 1}$ and we only want to prove that every polynomial in $\mathbf{C}[Z] \setminus \mathbf{C}$ of degree less than or equal to d has a root in \mathbf{C} , which is the highest degree of a polynomial in $\mathbf{R}[X]$ for which *we need* the Intermediate Value Theorem to hold?

With the aim of stating our problem precisely, we consider for each $d \in \mathbb{Z}_{\geq 1}$, the following properties on (\mathbf{R}, \leq) .

- [IVT] $_d$: for every polynomial $F \in \mathbf{R}[X]$ with $\deg F \leq d$ and every a, b in \mathbf{R} with $a < b$ and $F(a)F(b) < 0$, there exists $c \in \mathbf{R}$ with $a < c < b$ such that $F(c) = 0$.
- [FTA] $_d$: for every polynomial $F \in \mathbf{C}[Z] \setminus \mathbf{C}$ with $\deg F \leq d$, there exists $z \in \mathbf{C}$ such that $F(z) = 0$.

We can now restate our main concern as follows:

Given $d \in \mathbb{Z}_{\geq 1}$, which is the lowest value of $\alpha(d) \in \mathbb{Z}_{\geq 1}$ for which [IVT] $_{\alpha(d)}$ implies [FTA] $_d$?

In order to evaluate from this new quantitative point of view the proof of Theorem 1 we sketched, we define the following functions:

Notation 2 Let $\beta, \gamma : \mathbb{Z}_{\geq 1} \rightarrow \mathbb{Z}_{\geq 1}$ defined as follows:

$$\begin{aligned} \beta(d) &:= \begin{cases} d & \text{if } d \text{ is odd,} \\ \beta\left(\binom{d}{2}\right) & \text{if } d \text{ is even,} \end{cases} \\ \gamma(d) &:= \max_{1 \leq e \leq d} \{\beta(2e)\}. \end{aligned}$$

Note that $\gamma(1) = \beta(2) = 1$ and for $d \geq 2$ we have that $\gamma(d) \geq \beta(4) = 15$. Note also that γ is a non-decreasing function, whereas the behavior of β is rather chaotic.

First, we have that [FTA] $_1$ holds even under no assumptions on (\mathbf{R}, \leq) . Then, for a fixed $d \geq 2$ and a polynomial $F \in \mathbf{C}[Z]$ with $e = \deg F \leq d$, in order to be able to apply the proof of Theorem 1 we need to ensure [NnS] and the fact that the Intermediate Value Theorem holds for polynomials in $\mathbf{R}[X]$ of degree $\beta(2e)$. Since [IVT] $_2$ implies [NnS], we have that

$$[\text{IVT}]_{\gamma(d)} \text{ implies } [\text{FTA}]_d.$$

The final conclusion is that $\alpha(d) \leq \gamma(d)$.

Now we want to exhibit explicit bounds for γ . It is possible to prove that for $d \in \mathbb{Z}, d \geq 4$, if $d = 2^k s$ with $k \in \mathbb{Z}_{\geq 0}$ and odd $s \in \mathbb{Z}_{\geq 1}$ then

$$\frac{8}{3} \left(\frac{3}{4} 2^{k-1} s \right)^{2^k} \leq \beta(d) \leq 2 \left(2^{k-1} s \right)^{2^k} \leq 2 \left(\frac{d}{2} \right)^d.$$

Then, for $d \in \mathbb{Z}, d \geq 4$, we have

$$\gamma(d) \leq 2d^{2d}.$$

Also, by taking $k' := \lfloor \log_2 d \rfloor$, since $2^{k'} \leq d < 2^{k'+1}$,

$$\left(\frac{3}{8} \right)^{d-1} d^d = \frac{8}{3} \left(\frac{3d}{4} \right)^d < \frac{8}{3} \left(\frac{3}{4} 2^{k'} \right)^{2^{k'+1}} \leq \beta(2^{k'+1}) \leq \gamma(d).$$

In this way, we know that γ is bounded from below and above by exponential functions. This leads to an exponential upper bound for α , which cannot be avoided as long as we keep attached to the proof we sketched of Theorem 1.

The exponential value of $\gamma(d)$ plays a significant role in the bounds obtained in a recent joint work of the authors with Henri Lombardi, giving a new constructive proof for Hilbert 17-th problem and Positivstellensatz and providing elementary recursive degree bounds [8]. Exploring other algebraic proofs of [FTA] from a quantitative point of view might be a first step in the improvement of the results of [8]. This hope is part of our motivation in this paper.

In [3], Michael Eisermann found a proof of the Fundamental Theorem of Algebra which is also valid in any real closed field, but in opposition to Laplace's proof which is purely algebraic, has a large real-algebraic geometry flavor. This proof of Eisermann can be seen as a real-algebraic adaptation of one of the classical proofs of the Fundamental Theorem of Algebra using winding numbers and homotopy (see [4, Chapter 8]). One of the main ingredients of Eisermann's proof is the *Cauchy index* of two polynomials which, roughly speaking, is the number of jumps from $-\infty$ to $+\infty$ minus the number of jumps from $+\infty$ to $-\infty$ that the function associated to their quotient has in a given interval. From the fact that the base ordered field (\mathbf{R}, \leq) satisfies [IVT], it follows an *inversion formula* which implies that Cauchy indices can be computed by counting sign variations on Sturm chains. Another of the main ingredients of Eisermann's proof is the fact that the *winding number* of a complex function on a rectangle, which counts the number of zeros of the function in the given rectangle, can be computed in a completely real-algebraic way by means of Cauchy indices on the boundary of the rectangle. One of the most intricate steps in Eisermann's proof is to prove the Main Lemma (see [3, Lemma 5.3]): if a polynomial does not vanish in a rectangle, then the associated winding number is zero. This is achieved by a clever cancellation of terms for a suitable division of the rectangle under consideration. A crucial property for this cancellation is that, considering in the bivariate case one variable as the main variable and the second variable as a parameter, the (pseudo-)remainder sequence produces Sturm chains when specializing the parameter but also when specializing the main variable. Finally, by means of algebraic homotopy-like tools, the proof follows by computing the winding number in a well-known special case. Then the conclusion follows.

Our strategy is similar to the one by Eisermann, and also uses the *Cauchy index* of two polynomials and the *winding number* of a complex function on a rectangle. The main new ingredient is that we use subresultants rather than (pesudo)-remainder sequences to compute the Cauchy index. In order to be able to do this, we introduce the notion of (σ, τ) -chain, which is a generalization of the notion of Sturm chain and prove that it can be used to compute Cauchy indices. In case the base ordered field (\mathbf{R}, \leq) satisfies $[\text{IVT}]_{d^2}$, it follows a refinement of the *inversion formula* which implies that Cauchy indices can be computed by counting sign variations on (σ, τ) -chains, when the degrees of the polynomials in the (σ, τ) -chains are all bounded by d^2 . Again, the most intricate step in our proof of the Quantitative Fundamental Theorem of Algebra is to prove the Quantitative Main Lemma (Lemma 38): if $[\text{IVT}]_{d^2}$ holds and if a polynomial of degree d does not vanish in a rectangle, then the associated winding number is zero. This is also achieved by a cancellation of terms for a suitable division of the rectangle under consideration. As before, a crucial property for this cancellation is that, considering in the bivariate case one variable as the main variable and the second variable as a parameter, the subresultant sequence produces (σ, τ) -Sturm chains when specializing the parameter but also when specializing the main variable. By using degree bounds on subresultant polynomials, we obtain that intermediate polynomials relevant to the proof are all of degree bounded by d^2 .

Using this strategy, we prove the following theorem which is our main result.

Theorem 3 For $d \in \mathbb{Z}_{\geq 1}$

$$[\text{IVT}]_{d^2} \text{ implies } [\text{FTA}]_d.$$

In other words, Theorem 3 is equivalent to saying that $\alpha(d) \leq d^2$. Since $d \mapsto d^2$ is a polynomial function (actually $d^2 \leq \gamma(d)$ for all $d \in \mathbb{Z}_{\geq 1}$), our result highlights the difference in nature between Laplace's proof and our modification of Eisermann's proof.

2 Preliminaries

In subsection 2.1 and subsection 2.2 we introduce Cauchy indices and winding numbers. In subsection 2.3 we extend results by Eisermann on Cauchy indices and winding numbers [3] to the case where the Intermediate Value Theorem holds only for polynomials of bounded degrees. Finally in subsection 2.4 we give the needed preliminaries about subresultants.

2.1 Cauchy index

As said in the introduction, the Cauchy index of two polynomials Q and P on an interval is, roughly speaking, the number of jumps from $-\infty$ to $+\infty$ minus the number of jumps from $+\infty$ to $-\infty$ that the function associated to their quotient $\frac{Q}{P}$ has in this interval. We recall now the precise definition of Cauchy index following [3, Section 3].

Notation 4 Let $x \in \mathbf{R}$, we denote the sign of x by

$$\text{sign}(x) := \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Definition 5 Let $x \in \mathbf{R}$ and $P, Q \in \mathbf{R}[X]$.

- If $P, Q \in \mathbf{R}[X] \setminus \{0\}$, the polynomials P and Q can be written uniquely as

$$\begin{aligned} P &= (X - x)^{\mu(x)} \tilde{P}, \\ Q &= (X - x)^{\nu(x)} \tilde{Q}, \end{aligned}$$

with $\mu(x), \nu(x) \in \mathbb{Z}_{\geq 0}$ and $\tilde{P}(x) \neq 0, \tilde{Q}(x) \neq 0$.

For $\varepsilon \in \{+1, -1\}$, define

$$\text{Ind}_x^\varepsilon(Q, P) := \begin{cases} \frac{1}{2} \text{sign}(\tilde{Q}(x) \tilde{P}(x)) & \text{if } \varepsilon = +1 \text{ and } \mu(x) > \nu(x), \\ \frac{1}{2} (-1)^{\mu(x) - \nu(x)} \text{sign}(\tilde{Q}(x) \tilde{P}(x)) & \text{if } \varepsilon = -1 \text{ and } \mu(x) > \nu(x), \\ 0 & \text{otherwise.} \end{cases}$$

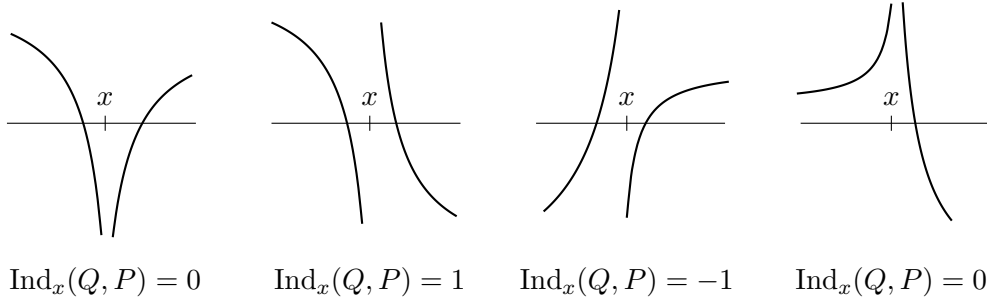
- If $P = 0$ or $Q = 0$, define

$$\text{Ind}_x^\varepsilon(Q, P) := 0.$$

- The Cauchy index of (Q, P) at x is

$$\text{Ind}_x(Q, P) := \text{Ind}_x^+(Q, P) - \text{Ind}_x^-(Q, P).$$

We illustrate this notion considering the graph of the function $\frac{Q}{P}$ around x in each different case.



Definition 6 Let $a, b \in \mathbf{R}$ and $P, Q \in \mathbf{R}[X]$.

- If $a < b$ and $P, Q \neq 0$, the Cauchy index of (Q, P) on the interval $[a, b]$ is

$$\text{Ind}_a^b(Q, P) := \text{Ind}_a^+(Q, P) + \sum_{x \in (a, b)} \text{Ind}_x(Q, P) - \text{Ind}_b^-(Q, P),$$

where the sum is well-defined since only roots x of P in (a, b) contribute.

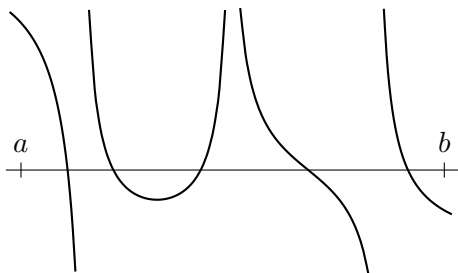
- If $a > b$ and $P, Q \neq 0$,

$$\text{Ind}_a^b(Q, P) := -\text{Ind}_b^a(Q, P).$$

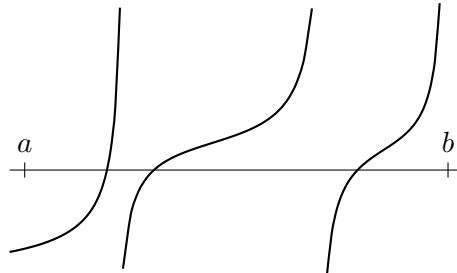
- In every other case,

$$\text{Ind}_a^b(Q, P) := 0.$$

In the following picture we consider again the graph of the function $\frac{Q}{P}$, this time in $[a, b]$.



$$\text{Ind}_a^b(Q, P) = 1 + 0 + 1 = 2$$



$$\text{Ind}_a^b(Q, P) = -1 - 1 - \frac{1}{2} = -\frac{5}{2}$$

Note that the Cauchy index of a pair of polynomials on an interval belongs to $\frac{1}{2}\mathbb{Z}$ and it is not necessarily an integer number.

Remark 7 If both P and Q are multiplied by $S \in \mathbf{R}[X] \setminus \{0\}$, it is clear that $\text{Ind}_a^b(Q, P) = \text{Ind}_a^b(QS, PS)$, so when $P \neq 0$ the Cauchy index is associated to the rational function $\frac{Q}{P}$ rather than to the pair of polynomials (Q, P) . However, when $P = 0$, it is convenient for us to define also the Cauchy index, even if the rational function $\frac{Q}{P}$ does not make sense. This is the reason why we use the notation $\text{Ind}_a^b(Q, P)$ in all cases.

Remark 8 Even though it is not reflected in the notation, the field \mathbf{R} plays a fundamental role in the definition of the Cauchy index. For instance, consider $P := X^2 - 2, Q := 1 \in \mathbb{Q}[X] \subset \mathbb{R}[X]$. If we take $\mathbf{R} = \mathbb{Q}$ we have

$$\text{Ind}_1^2(Q, P) = 0,$$

whereas if we take $\mathbf{R} = \mathbb{R}$ we have

$$\text{Ind}_1^2(Q, P) = 1.$$

Remark 9 Cauchy index is invariant by affine change of variables: given any affine function $\ell : [a, b] \rightarrow \mathbf{R}$, and $P, Q \in \mathbf{R}[X]$,

$$\text{Ind}_{\ell(a)}^{\ell(b)}(Q, P) = \text{Ind}_a^b(Q \circ \ell, P \circ \ell).$$

(By an affine function, we mean a function of type $\ell(X) = cX + d$ with $c, d \in \mathbf{R}$.)

Remark 10 Cauchy index is additive on intervals: given any $a, c_1, \dots, c_k, b \in \mathbf{R}$ and $P, Q \in \mathbf{R}[X]$,

$$\text{Ind}_a^b(Q, P) = \text{Ind}_a^{c_1}(Q, P) + \sum_{1 \leq i \leq k-1} \text{Ind}_{c_i}^{c_{i+1}}(Q, P) + \text{Ind}_{c_k}^b(Q, P).$$

2.2 Winding number

From now on, we consider the usual identification $\mathbf{C} \sim \mathbf{R}^2$.

The *winding number* of a closed curve in \mathbf{C} is a classical object which counts, by means of an analytic expression, the number of counterclockwise turns of the curve around the origin. In this paper, we will always restrict to curves which are the image of a polynomial function $F \in \mathbf{C}[X, Y]$ on the border of a rectangle $\Gamma \subset \mathbf{R}^2$ whose sides are parallel to the axis. For curves of this type, we recall the algebraic definition of *winding number* following [3]. Note that the border of Γ , denoted by $\partial\Gamma$, is simply the union of four segments.

Notation 11 For $F \in \mathbf{C}[X, Y]$, we denote F_{re} and F_{im} the real and imaginary parts of F , i.e. the unique polynomials in $\mathbf{R}[X, Y]$ such that the identity

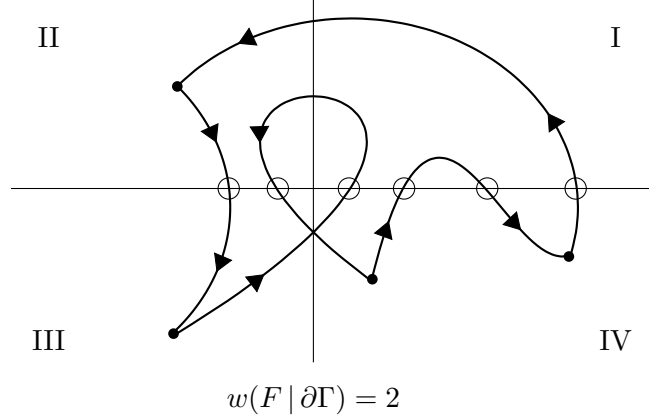
$$F(X, Y) = F_{\text{re}}(X, Y) + iF_{\text{im}}(X, Y)$$

in $\mathbf{C}[X, Y]$ holds.

Definition 12 Let $x_0, x_1, y_0, y_1 \in \mathbf{R}$ with $x_0 < x_1$ and $y_0 < y_1$ and let $\Gamma \subset \mathbf{R}^2$ be the rectangle $\Gamma := [x_0, x_1] \times [y_0, y_1]$. For $F \in \mathbf{C}[X, Y]$ the winding number of F on $\partial\Gamma$ is defined as

$$\begin{aligned} w(F | \partial\Gamma) &:= \frac{1}{2} \left(\text{Ind}_{x_0}^{x_1}(F_{\text{re}}(X, y_0), F_{\text{im}}(X, y_0)) + \text{Ind}_{y_0}^{y_1}(F_{\text{re}}(x_1, Y), F_{\text{im}}(x_1, Y)) \right. \\ &\quad \left. + \text{Ind}_{x_1}^{x_0}(F_{\text{re}}(X, y_1), F_{\text{im}}(X, y_1)) + \text{Ind}_{y_1}^{y_0}(F_{\text{re}}(x_0, Y), F_{\text{im}}(x_0, Y)) \right). \end{aligned}$$

Notice that it follows from the definition of winding number that we are going through $\partial\Gamma$ following the counterclock sense. The idea behind this algebraic definition is to count one half of a turn each time this curve crosses the X -axis from quadrant IV to I or from quadrant II to III, and minus one half of a turn each time it crosses the X -axis from quadrant I to IV or from quadrant III to II. Since these crossings coincide with jumps of the rational function $\frac{F_{\text{re}}}{F_{\text{im}}}$ from $-\infty$ to $+\infty$ and from $+\infty$ to $-\infty$ respectively, the Cauchy index is an appropriate algebraic tool to count the number of turns counterclockwise, which is (when F does not vanish on $\partial\Gamma$) the classical definition of the winding number.



Along the paper we will follow the convention of using X, Y and T for real variables, i.e. variables that will only be eventually evaluated at elements of \mathbf{R} , and Z for a complex variable, i.e. a variable that will be eventually evaluated at arbitrary elements of \mathbf{C} .

To $F \in \mathbf{C}[Z]$ we associate $\bar{F}(X, Y) := F(X + iY) \in \mathbf{C}[X, Y]$. Abusing slightly notation, we denote $F_{\text{re}}, F_{\text{im}}, w(F | \partial\Gamma)$ for $\bar{F}_{\text{re}}, \bar{F}_{\text{im}}, w(\bar{F} | \partial\Gamma)$.

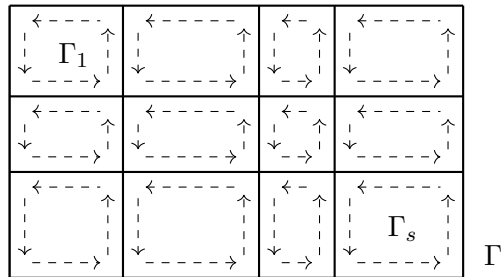
Example 13 (See [3, Proposition 4.4]) Let $\Gamma := [x_0, x_1] \times [y_0, y_1] \subset \mathbf{R}^2$. For $z \in \mathbf{C}$, we have

$$w(Z - z | \partial\Gamma) = \begin{cases} 1 & \text{if } z \text{ is in the interior of } \Gamma, \\ 1/2 & \text{if } z \text{ is in one of the edges of } \Gamma, \\ 1/4 & \text{if } z \text{ is a vertex of } \Gamma, \\ 0 & \text{if } z \text{ is in the exterior of } \Gamma. \end{cases}$$

Lemma 14 Let $\Gamma := [x_0, x_1] \times [y_0, y_1] \subset \mathbf{R}^2$ and consider a grid partition of Γ into a finite number of rectangles $\Gamma_1, \dots, \Gamma_s$. For $F \in \mathbf{C}[X, Y]$, we have

$$w(F | \partial\Gamma) = \sum_{1 \leq i \leq s} w(F | \partial\Gamma_i).$$

Proof: After replacing the winding number of F on $\partial\Gamma_1, \dots, \partial\Gamma_s$ by its definition, along each edge in the interior of Γ we have to add and subtract the Cauchy index of the same pair of polynomials, which adds up to zero. On the other hand, using the additivity of Cauchy index on intervals (Remark 10), adding on the remaining edges we obtain the winding number of F on $\partial\Gamma$.



□

To finish this subsection, we prove the following lemma which will play an important role at the end of Section 3 when applying homotopy tools.

Lemma 15 *Let $x_0, x_1, y_0, y_1, t_0, t_1 \in \mathbf{R}$ with $x_0 < x_1, y_0 < y_1$ and $t_0 < t_1$. Let $\Gamma_T := [x_0, x_1] \times [y_0, y_1], \Gamma_Y := [x_0, x_1] \times [t_0, t_1], \Gamma_X := [y_0, y_1] \times [t_0, t_1] \subset \mathbf{R}^2$. For $F \in \mathbf{C}[X, Y, T]$, we have*

$$\begin{aligned} & - w(F(X, Y, t_0) | \partial\Gamma_T) + w(F(X, y_0, T) | \partial\Gamma_Y) - w(F(x_0, Y, T) | \partial\Gamma_X) \\ & + w(F(X, Y, t_1) | \partial\Gamma_T) - w(F(X, y_1, T) | \partial\Gamma_Y) + w(F(x_1, Y, T) | \partial\Gamma_X) = 0. \end{aligned}$$

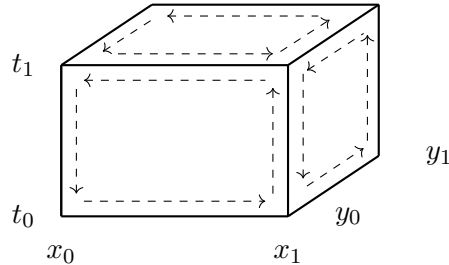
Therefore, if

$$w(F(X, y_0, T) | \partial\Gamma_Y) = w(F(x_0, Y, T) | \partial\Gamma_X) = w(F(X, y_1, T) | \partial\Gamma_Y) = w(F(x_1, Y, T) | \partial\Gamma_X) = 0$$

then

$$w(F(X, Y, t_0) | \partial\Gamma_T) = w(F(X, Y, t_1) | \partial\Gamma_T).$$

Proof: Consider the rectangular parallelepiped $[x_0, x_1] \times [y_0, y_1] \times [t_0, t_1] \subset \mathbf{R}^3$. After replacing each winding number by its definition, along each edge of this parallelepiped we have to add and subtract the Cauchy index of the same pair of polynomials; therefore obtaining 0 as the final result.



□

2.3 The intermediate value property for polynomials of bounded degree

Our main goal in this paper is to prove that $[\text{IVT}]_{d^2}$ implies $[\text{FTA}]_d$. So, from now, we take a fixed value of $d \in \mathbb{Z}_{\geq 1}$ and we suppose that (\mathbf{R}, \leq) is an ordered field satisfying $[\text{IVT}]_{d^2}$ but not necessarily $[\text{IVT}]$. Since $[\text{FTA}]_1$ holds even under no assumptions on (\mathbf{R}, \leq) , we suppose $d \geq 2$.

Note that the current assumption on (\mathbf{R}, \leq) is rather subtle, since for instance, it is only for $P \in \mathbf{R}[X]$ with $\deg P \leq d^2$ that we can claim that if P has no roots on an interval $I \subset \mathbf{R}$, then P has constant sign (different from 0) on I .

The purpose of this section is to reexamine some results from [3] concerning the Cauchy index and the winding number as well as to prove that they still hold in the present setting, despite the fact

that our hypotheses are weaker than in [3]. More explicitly, in [3] the assumption is that \mathbf{R} is a real closed field, and therefore it satisfies [IVT], whereas we only suppose $[\text{IVT}]_{d^2}$. Nevertheless, in the results reviewed in this section, following the proofs in [3] or a slight variation of it, it turns out that the Intermediate Value Theorem is applied to polynomials of degree less than or equal to d^2 , and this is enough to ensure that these results are still valid. For completeness, we include anyway full proofs of the statements in this section. We will use many times the following easy remark.

Remark 16 For $x, y \in \{-1, 0, 1\}$ with $(x, y) \neq (0, 0)$,

$$\text{sign}(xy) = 1 - |x - y|.$$

We introduce the following useful notation.

Notation 17 Let $x \in \mathbf{R}$ and $P, Q \in \mathbf{R}[X]$, we denote the sign variation of (P, Q) at x by

$$\text{Var}_x(P, Q) := \frac{1}{2} \left| \text{sign}(P(x)) - \text{sign}(Q(x)) \right|.$$

For $a, b \in \mathbf{R}$, we denote by $\text{Var}_a^b(P, Q)$ the sign variation of (P, Q) at a minus the sign variation of (P, Q) at b ; namely,

$$\text{Var}_a^b(P, Q) := \text{Var}_a(P, Q) - \text{Var}_b(P, Q).$$

We first prove the following property, which is a refinement of the well known *the inversion formula*.

Proposition 18 Let $a, b \in \mathbf{R}$ and $P, Q \in \mathbf{R}[X]$ with $\deg P, \deg Q \leq d^2$ and such that P and Q have no common root in $[a, b]$. Then

$$\text{Ind}_a^b(Q, P) + \text{Ind}_a^b(P, Q) = \text{Var}_a^b(P, Q)$$

Proof: We follow the arguments from [3, Theorem 3.9]. If $P = 0$ or $Q = 0$, since P and Q have no common root in $[a, b]$ we have that $\text{Var}_a^b(P, Q) = 0$ and the result holds. Suppose now that $P \neq 0$ and $Q \neq 0$. From the invariance by affine change of variables and the additivity on intervals of Cauchy index (Remarks 9 and 10) we suppose that a is the only possible root of P or Q on $[a, b]$ and that a is indeed a root of P .

Write

$$P = (X - a)^{\mu(a)} \tilde{P},$$

with $\mu(a) \in \mathbb{Z}_{>0}$ and $\tilde{P}(a) \neq 0$. Then using Remark 16 we have

$$\begin{aligned}
\text{Ind}_a^b(Q, P) + \text{Ind}_a^b(P, Q) &= \frac{1}{2} \text{sign}(Q(a)\tilde{P}(a)) + 0 \\
&= \frac{1}{2} - \frac{1}{2} \left| \text{sign}(\tilde{P}(a)) - \text{sign}(Q(a)) \right| \\
&= \frac{1}{2} - \frac{1}{2} \left| \text{sign}(P(b)) - \text{sign}(Q(b)) \right| \\
&= \frac{1}{2} - \text{Var}_b(P, Q) \\
&= \text{Var}_a(P, Q) - \text{Var}_b(P, Q) \\
&= \text{Var}_a^b(P, Q).
\end{aligned}$$

□

Next proposition shows the additivity of the winding number with respect to the product of complex polynomials.

Proposition 19 *Let $\Gamma := [x_0, x_1] \times [y_0, y_1] \subset \mathbf{R}^2$ and $F, G \in \mathbf{C}[X, Y]$ with $\deg FG \leq d^2$ and such that F and G do not vanish in $\partial\Gamma$. Then*

$$w(FG | \partial\Gamma) = w(F | \partial\Gamma) + w(G | \partial\Gamma).$$

The proof of Proposition 19 uses the next lemma as an auxiliary result.

Lemma 20 *Let $a, b \in \mathbf{R}$ and $P, Q, R, S \in \mathbf{R}[X]$ with $\deg(PR - QS), \deg(PS + QR) \leq d^2$ and such that P and Q have no common root in $[a, b]$ and R and S have no common root in $[a, b]$. Then*

$$\begin{aligned}
&\text{Ind}_a^b(PR - QS, PS + QR) = \\
&\text{Ind}_a^b(P, Q) + \text{Ind}_a^b(R, S) + \frac{1}{2} \text{sign}\left(\left((PS + QR)QS\right)(a)\right) - \frac{1}{2} \text{sign}\left(\left((PS + QR)QS\right)(b)\right).
\end{aligned}$$

Remark 21 In [3, Theorem 4.5] there is a statement with a slightly different formula and no assumption on polynomials $P, Q, R, S \in \mathbf{R}[X]$. We observed that this formula does not hold for the case $a := 0, b := 1, P := 1, Q := X, R := X - 1, S := X$. Notice that in this example, P, Q, R, S actually meet our extra assumptions, but since $PS + QR = QS = X^2$, if we deal with the rational function $PS + QR/QS$ as in [3, Theorem 4.5], there is a simplification which is the cause of the trouble. For this reason, we work with polynomials and not rational functions; but then the extra assumptions of not having common roots are necessary, since common factors would not modify the Cauchy indices but could modify the signs involved in the formula in Lemma 20. To illustrate this situation, $a := 0, b := 1, P := X - 1, Q := X(X - 1), R := X - 1, S := X$ would be a counterexample if we made no assumptions.

Proof of Lemma 20: First, we prove that the condition $\deg(PR - QS), \deg(PS + QR) \leq d^2$ implies $\deg P, \deg Q, \deg R, \deg S \leq d^2$. If at least one of the polynomials P, Q, R or S equals 0, then the degree bound on the other three polynomials clearly holds. If none of these polynomials equals 0, suppose that the claim does not hold and let us look for a contradiction. Respectively denote by $p, q, r, s \in \mathbf{R}$ their leading coefficients. The fact that there is a degree drop in $PR - QS$ and $PS + QR$ with respect to one of the polynomials P, Q, R or S implies that $pr - qs = 0$ and $ps + qr = 0$. Since $p, q, r, s \neq 0$, we deduce that $p^2 + q^2 = r^2 + s^2 = 0$, and this is not possible since \mathbf{R} is an ordered field.

Now we continue to prove the lemma following the ideas in [3, Theorem 4.5]. If $Q = 0, S = 0$ or $PS + QR = 0$, the result is immediate. If $P = 0$ or $R = 0$, the result follows from Proposition 18 using Remark 16. In every other case, from the invariance by affine change of variables and the additivity on intervals of Cauchy index (Remarks 9 and 10) we suppose that a is the only possible root of P, Q, R, S or $PS + QR$ on $[a, b]$. We consider several cases as follows.

- If $Q(a) \neq 0, S(a) \neq 0$ and $(PS + QR)(a) \neq 0$, then

$$\text{Ind}_a^b(PR - QS, PS + QR) = \text{Ind}_a^b(P, Q) = \text{Ind}_a^b(R, S) = 0$$

and

$$\text{sign}\left(\left((PS + QR)QS\right)(a)\right) = \text{sign}\left(\left((PS + QR)QS\right)(b)\right)$$

so the identity holds.

- If $Q(a) \neq 0, S(a) \neq 0$ and $(PS + QR)(a) = 0$, then

$$\text{Ind}_a^b(P, Q) = \text{Ind}_a^b(R, S) = \text{sign}\left(\left((PS + QR)QS\right)(a)\right) = 0.$$

On the other hand

$$\frac{P(a)}{Q(a)} = -\frac{R(a)}{S(a)},$$

so

$$(PR - QS)(a) = Q(a)S(a) \underbrace{\left(\frac{P(a)}{Q(a)} \frac{R(a)}{S(a)} - 1\right)}_{<0} \neq 0$$

Write $PS + QR = (X - a)^\mu T$ with $\mu \in \mathbb{Z}_{>0}$ and $T(a) \neq 0$. Note that $\text{sign}(T(a)) = \text{sign}(T(b)) = \text{sign}((PS + QR)(b))$. Then we have

$$\text{Ind}_a^b(PR - QS, PS + QR) = -\frac{1}{2} \text{sign}(Q(a)S(a)T(a)) = -\frac{1}{2} \text{sign}\left(\left((PS + QR)QS\right)(b)\right)$$

so the identity holds.

- If $Q(a) = 0$ and $S(a) \neq 0$, since P and Q have no common root in $[a, b]$ then $(PS + QR)(a) \neq 0$ and we have that

$$\text{Ind}_a^b(PR - QS, PS + QR) = \text{Ind}_a^b(R, S) = \text{sign}\left(\left((PS + QR)QS\right)(a)\right) = 0.$$

Write $Q = (X - a)^\mu \tilde{Q}$ with $\mu \in \mathbb{Z}_{>0}$ and $\tilde{Q}(a) \neq 0$. Then

$$\text{Ind}_a^b(P, Q) = \frac{1}{2} \text{sign}(P(a)\tilde{Q}(a)) = \frac{1}{2} \text{sign}\left(\left((PS + QR)\tilde{Q}S\right)(a)\right) = \frac{1}{2} \text{sign}\left(\left((PS + QR)QS\right)(b)\right)$$

so the identity holds.

- If $Q(a) \neq 0$ and $S(a) = 0$ we proceed in a similar way to the previous case.
- If $Q(a) = 0$ and $S(a) = 0$, then $(PS + QR)(a) = 0$, and since P and Q have no common root in $[a, b]$ and R and S have no common root in $[a, b]$, $P(a) \neq 0$, $R(a) \neq 0$.

Write $PS + QR = (X - a)^{\mu_0} T$ with $\mu_0 \in \mathbb{Z}_{>0}$ and $T(a) \neq 0$, $Q = (X - a)^{\mu_1} \tilde{Q}$ with $\mu_1 \in \mathbb{Z}_{>0}$ and $\tilde{Q}(a) \neq 0$ and $S = (X - a)^{\mu_2} \tilde{S}$ with $\mu_2 \in \mathbb{Z}_{>0}$ and $\tilde{S}(a) \neq 0$. We denote

$$\begin{aligned} \sigma_1 &:= \text{sign}(P(a)) \in \{-1, 1\}, \\ \sigma_2 &:= \text{sign}(R(a)) \in \{-1, 1\}, \\ \sigma_3 &:= \text{sign}(T(a)) \in \{-1, 1\}, \\ \sigma_4 &:= \text{sign}(\tilde{Q}(a)) \in \{-1, 1\}, \\ \sigma_5 &:= \text{sign}(\tilde{S}(a)) \in \{-1, 1\}. \end{aligned}$$

We need to prove that

$$\sigma_1 \sigma_2 \sigma_3 = \sigma_1 \sigma_4 + \sigma_2 \sigma_5 - \sigma_3 \sigma_4 \sigma_5$$

or, equivalently,

$$(\sigma_1 \sigma_2 + \sigma_4 \sigma_5) \sigma_3 = \sigma_1 \sigma_4 + \sigma_2 \sigma_5 \tag{1}$$

We take into account that $\sigma_1 = \text{sign}(P(b))$, $\sigma_2 = \text{sign}(R(b))$, $\sigma_3 = \text{sign}((PS + QR)(b))$, $\sigma_4 = \text{sign}(Q(b))$ and $\sigma_5 = \text{sign}(S(b))$ and we divide in cases as follows.

- If $\sigma_1 = \sigma_5$ and $\sigma_2 = \sigma_4$, then $\sigma_3 = 1$ and equation (1) holds.
- If $\sigma_1 = -\sigma_5$ and $\sigma_2 = -\sigma_4$ then $\sigma_3 = -1$ and equation (1) holds.
- In every other case, exactly three elements in the set $\{\sigma_1, \sigma_2, \sigma_4, \sigma_5\}$ are equal and the remaining one is different. Then

$$\sigma_1 \sigma_2 + \sigma_4 \sigma_5 = \sigma_1 \sigma_4 + \sigma_2 \sigma_5 = 0$$

and equation (1) holds. □

Proof of Proposition 19: The proof is done in as in [3, Corollary 4.6 and Corollary 4.7]. After replacing each winding number by its definition, we apply Lemma 20 once on each side of $\partial\Gamma$. For instance, on the bottom side, we take $a := x_0$, $b := x_1$, $P := F_{\text{re}}(X, y_0)$, $Q := F_{\text{im}}(X, y_0)$, $R := G_{\text{re}}(X, y_0)$ and $S := G_{\text{im}}(X, y_0)$. The identity in the lemma is obtained after checking that on each vertex of $\partial\Gamma$, signs cancel after being added on one side and subtracted on the other side. □

From Example 13 and Proposition 19 the following result is easily deduced.

Example 22 For $e \in \mathbb{Z}_{\geq 1}$ with $e \leq d^2$,

$$w(Z^e | \partial\Gamma) = e.$$

if $\Gamma \subset \mathbf{R}^2$ is a rectangle containing 0 in its interior.

Finally, we recall the property saying that the winding number vanishes in a *small* rectangle around a non-zero of a polynomial.

Proposition 23 Let $(x, y) \in \mathbf{R}^2$ and $F \in \mathbf{C}[X, Y]$ with $\deg F \leq d^2$ and such that $F(x, y) \neq 0$. Then there exists $\delta \in \mathbf{R}, \delta > 0$ such that for every rectangle $\Gamma \subset [x - \delta, x + \delta] \times [y - \delta, y + \delta] \subset \mathbf{R}^2$, F does not vanish in Γ and $w(F | \partial\Gamma) = 0$.

Proof: First, since $[\text{IVT}]_{d^2}$ holds, it is easy to see that for every $a \in \mathbf{R}$ with $a \geq 0$ and every $n \in \mathbb{Z}_{\geq 1}, n \leq d^2$, there is a unique $c \in \mathbf{R}$ such that $c \geq 0$ and $c^n = a$, which we note as $c = a^{1/n}$. It is clear that if $a > 0$ then $a^{1/n} > 0$.

Then we just follow the arguments from [3, Lemma 5.2]. We take

$$G := \frac{i}{F(x, y)} F \in \mathbf{C}[X, Y]$$

and we need to prove that there exists $\delta \in \mathbf{R}, \delta > 0$ such that $G_{\text{im}} \in \mathbf{R}[X, Y]$ does not vanish in $[x - \delta, x + \delta] \times [y - \delta, y + \delta] \subset \mathbf{R}^2$. In this case, for every rectangle $\Gamma \subset [x - \delta, x + \delta] \times [y - \delta, y + \delta]$, we have that $G \in \mathbf{C}[X, Y]$ does not vanish in Γ and $w(G | \partial\Gamma) = 0$. Then the lemma follows using Proposition 19.

Suppose now

$$G_{\text{im}} = \sum_{\substack{j=(j_1, j_2) \\ j_1+j_2 \leq d^2}} c_j (X-x)^{j_1} (Y-y)^{j_2}.$$

Since $G(x, y) = i$ we know that $G_{\text{im}}(x, y) = 1$. If G_{im} is constant, then any positive value of δ works. Otherwise, taking

$$\Delta = \frac{1}{2}(d^2 + 1)(d^2 + 2)$$

and

$$\delta := \min \left\{ \left(\frac{1}{\Delta |c_j|} \right)^{\frac{1}{j_1+j_2}} \mid j = (j_1, j_2), 1 \leq j_1 + j_2 \leq d^2, c_j \neq 0 \right\} > 0,$$

for every $(z, w) \in [-\delta, \delta] \times [-\delta, \delta]$ we have

$$G_{\text{im}}(x+z, y+w) = 1 + \sum_{\substack{j=(j_1, j_2) \\ 1 \leq j_1+j_2 \leq d^2}} c_j z^{j_1} w^{j_2} \geq 1 - \sum_{\substack{j=(j_1, j_2) \\ 1 \leq j_1+j_2 \leq d^2}} |c_j| \delta^{j_1+j_2} \geq 1 - \sum_{\substack{j=(j_1, j_2) \\ 1 \leq j_1+j_2 \leq d^2}} \frac{1}{\Delta} = \frac{1}{\Delta} > 0.$$

□

2.4 Subresultant polynomials

Let \mathbf{D} be an integral domain. The subresultant polynomial sequence of two polynomials $P, Q \in \mathbf{D}[X]$ is a sequence of polynomials in $\mathbf{D}[X]$ which contains the classical Sylvester resultant of P, Q ; more specifically, the last subresultant polynomial, which actually belongs to \mathbf{D} , coincides up to sign with the Sylvester resultant. Even though the subresultant polynomials of P and Q are defined in a completely different way, they are closely related to the polynomials appearing in the remainder sequence of P and Q , as reflected in the Structure Theorem of Subresultants (Theorem 26). It can be proved that the behavior of their coefficients is better controlled than the behavior of the coefficients of the polynomials in the remainder sequence, and for this reason, they constitute a widely used tool in gcd computation, real root counting and many other problems in computational algebra. In Section 3, we will use subresultants in the particular case of $\mathbf{D} = \mathbf{R}[Y]$ and the good behavior of their coefficients implies a good control of the degree in Y (Proposition 27), which will be a key point to obtain our main result.

We include now some definitions and properties concerning subresultants. We refer the reader to [1] for proofs and details.

Definition 24 *Let $P, Q \in \mathbf{D}[X] \setminus \{0\}$ with $p := \deg P \geq 1$ and $q := \deg Q < p$.*

- *For $0 \leq j \leq q$, the Sylvester-Habicht matrix $\text{SyHa}_j(P, Q) \in \mathbf{D}^{(p+q-2j) \times (p+q-j)}$ is the matrix whose rows are the polynomials*

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

expressed in the monomial basis $X^{p+q-j-1}, \dots, X, 1$.

- *For $0 \leq j \leq q$, the j -th subresultant polynomial of P and Q , $\text{sResP}_j(P, Q) \in \mathbf{D}[X]$ is the polynomial determinant of $\text{SyHa}_j(P, Q)$, i.e.*

$$\text{sResP}_j(P, Q) := \sum_{0 \leq i \leq j} \det(\text{SyHa}_{j,i}(P, Q)) \cdot X^i \in \mathbf{D}[X]$$

where $\text{SyHa}_{j,i}(P, Q) \in \mathbf{D}^{(p+q-2j) \times (p+q-2j)}$ is the matrix obtained by taking the $p+q-2j-1$ first columns and the $(p+q-j-i)$ -th column of $\text{SyHa}_j(P, Q)$. By convention, we extend this definition with

$$\begin{aligned} \text{sResP}_p(P, Q) &:= P \in \mathbf{D}[X], \\ \text{sResP}_{p-1}(P, Q) &:= Q \in \mathbf{D}[X], \\ \text{sResP}_j(P, Q) &:= 0 \in \mathbf{D}[X] \quad \text{for } q < j < p-1. \end{aligned}$$

- *For $0 \leq j \leq q$, the j -th signed subresultant coefficient of P and Q , $\text{sR}_j(P, Q) \in \mathbf{D}$ is the coefficient of X^j in $\text{sResP}_j(P, Q)$. By convention, we extend this definition with*

$$\begin{aligned} \text{sR}_p(P, Q) &:= 1 \in \mathbf{D} \quad (\text{even if } P \text{ is not monic}), \\ \text{sR}_j(P, Q) &:= 0 \in \mathbf{D} \quad \text{for } q < j \leq p-1. \end{aligned}$$

- For $0 \leq j \leq p$, $\text{sResP}_j(P, Q)$ is said to be defective if $\deg \text{sResP}_j(P, Q) < j$ or, equivalently, if $\text{sR}_j(P, Q) = 0$.

We will also use the following notation.

Notation 25 Let $P, Q \in \mathbf{D}[X] \setminus \{0\}$ with $p := \deg P \geq 1$ and $q := \deg Q < p$. Let (d_0, \dots, d_s) be the sequence of degrees of the non-defective subresultant polynomials of P and Q in decreasing order (note that $d_0 = p$ and $d_1 = q$). For $1 \leq i \leq s$,

$$T_{d_{i-1}-1}(P, Q) := \text{lcoeff}(\text{sResP}_{d_{i-1}-1}(P, Q)) \in \mathbf{D} \setminus \{0\}.$$

We extend this notation with $T_p(P, Q) := 1 \in \mathbf{D} \setminus \{0\}$.

The following theorem is one of the most important results in the theory of subresultants. This result has a long history [5]. We quote its more recent form in [1], which is a slight improvement of [7].

Theorem 26 (Structure Theorem of Subresultants) Let $P, Q \in \mathbf{D}[X] \setminus \{0\}$ with $p := \deg P \geq 1$ and $q := \deg Q < p$. Let (d_0, \dots, d_s) be the sequence of degrees of the non-defective subresultant polynomials of P and Q in decreasing order and let $d_{-1} := p + 1$. Then

- for $1 \leq i \leq s$,

$$\text{sResP}_{d_{i-1}-2}(P, Q) = \dots = \text{sResP}_{d_{i+1}}(P, Q) = 0 \in \mathbf{D}[X]$$

and $\text{sResP}_{d_{i-1}-1}(P, Q)$ and $\text{sResP}_{d_i}(P, Q)$ are proportional. More precisely,

$$\text{sR}_{d_i}(P, Q) \cdot \text{sResP}_{d_{i-1}-1}(P, Q) = T_{d_{i-1}-1}(P, Q) \cdot \text{sResP}_{d_i}(P, Q) \in \mathbf{D}[X]$$

with

$$\text{sR}_{d_i}(P, Q) = (-1)^{\frac{1}{2}(d_{i-1}-d_i)(d_{i-1}-d_i-1)} \frac{T_{d_{i-1}-1}(P, Q)^{d_{i-1}-d_i}}{\text{sR}_{d_{i-1}}(P, Q)^{d_{i-1}-d_i-1}} \in \mathbf{D}.$$

This implies $\deg \text{sResP}_{d_{i-1}-1}(P, Q) = d_i$.

- for $1 \leq i \leq s$,

$$\begin{aligned} & T_{d_{i-2}-1}(P, Q) \cdot \text{sR}_{d_{i-1}}(P, Q) \cdot \text{sResP}_{d_{i-1}}(P, Q) \\ &= -\text{Rem} \left(T_{d_{i-1}-1}(P, Q) \cdot \text{sR}_{d_i}(P, Q) \cdot \text{sResP}_{d_{i-2}-1}(P, Q), \text{sResP}_{d_{i-1}-1}(P, Q) \right) \in \mathbf{D}[X] \end{aligned}$$

and

$$\text{Quot} \left(T_{d_{i-1}-1}(P, Q) \cdot \text{sR}_{d_i}(P, Q) \cdot \text{sResP}_{d_{i-2}-1}(P, Q), \text{sResP}_{d_{i-1}-1}(P, Q) \right) \in \mathbf{D}[X]$$

(where Rem and Quot means the remainder and quotient in the euclidean division in $\text{qf}(\mathbf{D})[X]$ of the first polynomial by the second polynomial).

- $\text{sResP}_{d_{s-1}-1}(P, Q) \in \mathbf{D}[X]$ and $\text{sResP}_{d_s}(P, Q) \in \mathbf{D}[X]$ are the greatest common divisor of P and Q in $\text{qf}(\mathbf{D})[X]$ multiplied by elements in \mathbf{D} . In addition, if $d_s > 0$ then

$$\text{sResP}_{d_s-1}(P, Q) = \cdots = \text{sResP}_0(P, Q) = 0 \in \mathbf{D}[X].$$

$$\begin{array}{r}
\text{-----} \quad \text{sResP}_{d_0} = \text{sResP}_p = P \\
\text{-----} \quad \text{sResP}_{d_0-1} = \text{sResP}_{p-1} = Q \\
0 \\
\vdots \\
0 \\
\text{-----} \quad \text{sResP}_{d_1} = \text{sResP}_q \\
\text{-----} \quad \text{sResP}_{d_1-1} \\
0 \\
\vdots \\
\vdots \\
0 \\
\text{-----} \quad \text{sResP}_{d_2} \\
\vdots \\
\vdots \\
\text{-----} \quad \text{sResP}_{d_{s-1}-1} \\
0 \\
\vdots \\
\vdots \\
0 \\
\text{-----} \quad \text{sResP}_{d_s} \\
0 \\
\vdots \\
0
\end{array}$$

Proof: See [1, Chapter 8]. □

As said before, in Section 3 we will use subresultants in the particular case of $\mathbf{D} = \mathbf{R}[Y]$ and we will need some degree bounds which we develop here.

Proposition 27 *Let $P, Q \in \mathbf{R}[X, Y] \setminus \{0\}$ with $p := \deg_X P \geq 1$, $q := \deg_X Q < p$ and total degree $\deg P, \deg Q \leq d$ (with $d \geq p$). We consider subresultants with respect to variable X (this is to say, following Definition 24 we take $\mathbf{D} = \mathbf{R}[Y]$). For $0 \leq j \leq q$ and $0 \leq i \leq j$, the degree in Y of the coefficient of X^i in $\text{sResP}_j(P, Q) \in \mathbf{R}[Y][X]$ is bounded by d^2 .*

Proof: Let $P = \sum_{0 \leq i \leq p} a_i(Y)X^i$ and $Q = \sum_{0 \leq i \leq q} b_i(Y)X^i$; then $\deg_Y a_i(Y), \deg_Y b_i(Y) \leq d - i$. By definition,

$$\text{sResP}_j(P, Q) = \sum_{0 \leq i \leq j} \det(\text{SyHa}_{j,i}(P, Q)) \cdot X^i \in \mathbf{R}[Y][X],$$

where $\text{SyHa}_{j,i}(P, Q) \in \mathbf{R}[Y]^{(p+q-2j) \times (p+q-2j)}$ is the matrix obtained by taking the $p+q-2j-1$ first columns and the $(p+q-j-i)$ -th column of $\text{SyHa}_j(P, Q) \in \mathbf{R}[Y]^{(p+q-2j) \times (p+q-j)}$. By defining $a_i(Y) = 0$ if $i \geq p+1$ or $i \leq -1$ and $b_i(Y) = 0$ if $i \geq q+1$ or $i \leq -1$ we have that for $1 \leq k \leq p+q-2j$ and $1 \leq \ell \leq p+q-j$,

$$(\text{SyHa}_{j,i}(P, Q))_{k\ell} = \begin{cases} a_{p+k-\ell}(Y) & \text{if } k \leq q-j, \\ b_{p+2q-2j+1-k-\ell}(Y) & \text{if } k \geq q-j+1. \end{cases}$$

The proof can be completed by bounding the degree of any possible nonzero product of entries of $\text{SyHa}_{j,i}(P, Q)$ with one element per row and column. We obtain that the degree in Y of the coefficient of X^i in $\text{sResP}_j(P, Q) \in \mathbf{R}[Y][X]$ is bounded by

$$d(p+q-2j) - pq + j^2 + j - i \leq d(p+q) - pq \leq d^2.$$

□

3 Counting complex roots

In this section we introduce (σ, τ) -chains, develop suitable generalizations of results from [3] and prove Theorem 3. As said before, till the end of the paper, we take a fixed value of $d \in \mathbb{Z}_{\geq 2}$ and we suppose that (\mathbf{R}, \leq) is an ordered field satisfying $[\text{IVT}]_{d^2}$ but not necessarily $[\text{IVT}]$.

3.1 (σ, τ) -chains and Cauchy index

A Sturm chain with respect to I is a finite sequence of univariate polynomials $(P_0, \dots, P_n) \in \mathbf{R}[X]$ such that for every $x \in I$ and $0 < i < n$, if $P_i(x) = 0$ then $P_{i-1}(x)P_{i+1}(x) < 0$ (see [3, Definition 3.10]).

An important property of Sturm chains is its connection to Cauchy indices, given by Proposition 28 ([3, Theorem 3.11])

Proposition 28 *Let $a, b \in \mathbf{R}$ with $a < b$, $I := [a, b]$, $n \in \mathbb{Z}_{\geq 1}$. If (P_0, \dots, P_n) is a Sturm chain with respect to I , then*

$$\text{Ind}_a^b(P_1, P_0) + \text{Ind}_a^b(P_{n-1}, P_n) = \sum_{1 \leq i \leq n} \text{Var}_a^b(P_{i-1}, P_i)$$

Example 29 *Here are important examples illustrating the definition of Sturm chains. Item b) plays a key role in the proof of the Main Lemma (see [3, Lemma 5.3]) stating that if a polynomial does not vanish in a rectangle, then the associated winding number is zero, which is a key step of the algebraic-geometric proof of the Fundamental Theorem of Algebra in [3].*

- a) *Let $P_0, P_1 \in \mathbf{R}[X] \setminus \{0\}$ with $d_0 := \deg P_0 \geq 1$, $d_1 := \deg P_1 < d_0$ and P_0, P_1 coprime. Let (d_0, d_1, \dots, d_s) be the sequence of degrees of the remainder sequence of P_0, P_1 in decreasing order.*

We consider the classical Sturm sequence of P_0, P_1 : for $2 \leq i \leq s$, we define

$$P_i = -\text{Rem}(P_{i-2}, P_{i-1}).$$

It is clear that the Sturm sequence (P_0, P_1, \dots, P_s) is a Sturm chain.

- b) Given $P, Q \in \mathbf{R}[Y][X] \setminus \{0\}$, $p := \deg_X P \geq 1$, $q := \deg_X Q < p$, e the smallest even natural number greater than or equal to $p - q$ and $C \in \mathbf{R}[Y]$ the leading coefficient of Q . We define

$$\text{Prem}(P, Q) = \text{Rem}(C^e P, Q) \in \mathbf{R}[Y][X].$$

Let $P_0, P_1 \in \mathbf{R}[Y][X] \setminus \{0\}$ with $d_0 := \deg_X P_0 \geq 1$, $d_1 := \deg_X P_1 < d_0$ and P_0, P_1 coprime in the unique factorization domain $\mathbf{R}[X, Y]$. Let (d_0, d_1, \dots, d_s) be the sequence of degrees in X of the pseudo-remainders in decreasing order. For $2 \leq i \leq s$, we define

$$P_i := -\text{Prem}(P_{i-2}, P_{i-1}) \in \mathbf{R}[Y][X]$$

and for $1 \leq i \leq s$, $C_i \in \mathbf{R}[Y]$ the leading coefficient of P_i . Since P_0 and P_1 are coprime in $\mathbf{R}[X, Y]$, we have that $P_s = C_s \in \mathbf{R}[Y]$. Take an interval $[b, b']$ such that C_1, \dots, C_s have no zero on $[b, b']$. We have that

- for any $y \in [b, b']$, $(P_0(X, y), \dots, P_s(X, y)) \in \mathbf{R}[X]$ is a Sturm chain with respect to \mathbf{R} ,
- for any $x \in \mathbf{R}$, $(P_0(x, Y), \dots, P_s(x, Y)) \in \mathbf{R}[Y]$ is Sturm chain with respect to $[b, b']$.

We wish to use subresultants rather than (pseudo)-remainder sequences to prove the Quantitative Main Lemma (Lemma 38), taking advantage of good degree bounds for subresultants (Proposition 27). Unfortunately subresultants are not necessarily Sturm chains.

This is our motivation to introduce now the notion of (σ, τ) -chain, which is a generalization of the notion of Sturm chain. Then, in Proposition 34 and Corollary 35, we develop a modified sign changing counting rule, so that we can still use (σ, τ) -chains to compute Cauchy indices.

The benefit of this generalization is that the subresultant polynomial sequence will fit in this definition for some pair (σ, τ) , which is an essential ingredient for the proof of the Quantitative Main Lemma (Lemma 38) where we use the good degree bounds on subresultants obtained in Proposition 27.

Definition 30 Let I be an interval of \mathbf{R} , $n \in \mathbb{Z}_{\geq 1}$ and $\sigma, \tau \in \{-1, 1\}^{n-1}$ with $\sigma = (\sigma_1, \dots, \sigma_{n-1})$ and $\tau = (\tau_1, \dots, \tau_{n-1})$.

A sequence of polynomials $(S_0, \dots, S_n) \in \mathbf{R}[X]$ is a (σ, τ) -chain with respect to I if for $1 \leq i \leq n - 1$ there exists polynomials $A_i, B_i, C_i \in \mathbf{R}[X]$ such that

1. $A_i S_{i+1} + B_i S_i + C_i S_{i-1} = 0$,
2. for every $x \in I$, $\text{sign}(A_i(x)) = \sigma_i$,
3. for every $x \in I$, $\text{sign}(C_i(x)) = \tau_i$.

A sequence of polynomials $(S_0, \dots, S_n) \in \mathbf{R}[X]$ is a Sturm (σ, τ) -chain with respect to I if it is a (σ, τ) -chain with respect to I and S_{n-1} and S_n have no common root on I .

A sequence of polynomials $(S_0, \dots, S_n) \in \mathbf{R}[X]$ is a good Sturm (σ, τ) -chain with respect to I if it is a (σ, τ) -chain with respect to I and S_n has no root on I .

Note that for $n = 1$, taking $\{-1, 1\}^0 = \{\bullet\}$, any sequence (S_0, S_1) in $\mathbf{R}[X]$ is a (\bullet, \bullet) -chain with respect to I .

Note also that if a sequence of polynomials (S_0, \dots, S_n) in $\mathbf{R}[X]$ is a (σ, τ) -chain with respect to I , then for every $0 \leq m \leq n-1$, (S_m, \dots, S_n) is a (σ', τ') -chain with respect to I , with $\sigma' := (\sigma_{m+1}, \dots, \sigma_{n-1})$ and $\tau' := (\tau_{m+1}, \dots, \tau_{n-1})$. The analogous statements hold also for Sturm (σ, τ) -chains and good Sturm (σ, τ) -chains.

Example 31 Here are important examples illustrating the definition of (σ, τ) -chains. Item b) plays a key role in the proof of our Quantitative Main (Lemma 38) which is a crucial step in the proof of our Quantitative Fundamental Theorem of Algebra (Theorem 3).

a) Let $S_0, S_1 \in \mathbf{R}[X] \setminus \{0\}$ with $d_0 := \deg S_0 \geq 1$, $d_1 := \deg S_1 < d_0$ and S_0, S_1 coprime. Let (d_0, d_1, \dots, d_s) be the sequence of degrees of the non-defective subresultant polynomials of S_0, S_1 in decreasing order (note that $d_0 = p$, $d_1 = q$), and $d_{-1} := p + 1$. Finally, for $2 \leq i \leq s$, we define

$$S_i := \text{sResP}_{d_{i-1}-1}(S_0, S_1) \in \mathbf{R}[X]$$

(note that the above identity also holds for $i = 0, 1$). Since S_0 and S_1 are coprime in $\mathbf{R}[X]$, by the Structure Theorem of Subresultants (Theorem 26) we have that $S_s \in \mathbf{R}$ and $d_s = 0$. Also, defining for $1 \leq i \leq s-1$

$$A_i := T_{d_{i-2}-1}(S_0, S_1) \cdot \text{sR}_{d_{i-1}}(S_0, S_1) \in \mathbf{R} \setminus \{0\}$$

$$B_i := -\text{Quot}(T_{d_{i-1}-1}(S_0, S_1) \cdot \text{sR}_{d_i}(S_0, S_1) \cdot S_{i-1}, S_i) \in \mathbf{R}[X],$$

$$C_i := T_{d_{i-1}-1}(S_0, S_1) \cdot \text{sR}_{d_i}(S_0, S_1) \in \mathbf{R} \setminus \{0\},$$

we have

$$A_i S_{i+1} + B_i S_i + C_i S_{i-1} = 0.$$

We define $\sigma_i := \text{sign}(A_i)$ and $\tau_i := \text{sign}(C_i)$ for $1 \leq i \leq s-1$ and we have that (S_0, S_1, \dots, S_s) is a good Sturm (σ, τ) -chain.

b) Let $S_0, S_1 \in \mathbf{R}[X, Y] \setminus \{0\}$ with $p := \deg_X S_0 \geq 1$, $q := \deg_X S_1 < p$ and S_0, S_1 coprime in the unique factorization domain $\mathbf{R}[X, Y]$. Let (d_0, d_1, \dots, d_s) be the sequence of degrees of the non-defective subresultant polynomials in decreasing order (note that $d_0 = p$ and $d_1 = q$), and $d_{-1} := p + 1$; where all the subresultants are defined considering X as the main variable (this is to say, $\mathbf{D} = \mathbf{R}[Y]$ in Definition 24). Finally, for $2 \leq i \leq s$, we define

$$S_i := \text{sResP}_{d_{i-1}-1}(S_0, S_1) \in \mathbf{R}[Y][X]$$

(note that the above identity also holds for $i = 0, 1$). Since S_0 and S_1 are coprime in $\mathbf{R}[X, Y]$, by the Structure Theorem of Subresultants (Theorem 26) we have that $S_s \in \mathbf{R}[Y]$. Also, defining for $1 \leq i \leq s - 1$

$$\begin{aligned} A_i &:= T_{d_{i-2}-1}(S_0, S_1) \cdot \text{sR}_{d_{i-1}}(S_0, S_1) \in \mathbf{R}[Y] \setminus \{0\} \\ B_i &:= -\text{Quot}(T_{d_{i-1}-1}(S_0, S_1) \cdot \text{sR}_{d_i}(S_0, S_1) \cdot S_{i-1}, S_i) \in \mathbf{R}[Y][X], \\ C_i &:= T_{d_{i-1}-1}(S_0, S_1) \cdot \text{sR}_{d_i}(S_0, S_1) \in \mathbf{R}[Y] \setminus \{0\}, \end{aligned}$$

we have

$$A_i S_{i+1} + B_i S_i + C_i S_{i-1} = 0.$$

Take an interval $[b, b']$ such that S_s, A_i and C_i , for $i = 1, \dots, s - 1$, have constant sign different from 0 on $[b, b']$. We define then $\sigma = (\sigma_1, \dots, \sigma_{s-1}), \tau = (\tau_1, \dots, \tau_{s-1}) \in \{-1, 1\}^{s-1}$ by choosing any $c \in [b, b']$ and taking $\sigma_i := \text{sign}(A_i(c))$ and $\tau_i := \text{sign}(C_i(c))$ for $1 \leq i \leq s - 1$. In this way, since also $S_s \in \mathbf{R}[Y]$ does not vanish on $[b, b']$, we have that

- for any $y \in [b, b']$, $(S_0(X, y), \dots, S_s(X, y)) \in \mathbf{R}[X]$ is a good Sturm (σ, τ) -chain with respect to \mathbf{R} ,
- for any $x \in \mathbf{R}$, $(S_0(x, Y), \dots, S_s(x, Y)) \in \mathbf{R}[Y]$ is a good Sturm (σ, τ) -chain with respect to $[b, b']$.

Lemma 32 *Let I be an interval of \mathbf{R} , $n \in \mathbb{Z}_{\geq 1}$ and $\sigma, \tau \in \{-1, 1\}^{n-1}$. If a sequence of polynomials (S_0, \dots, S_n) in $\mathbf{R}[X]$ is a Sturm (σ, τ) -chain with respect to I , then for every $1 \leq m \leq n$, S_{m-1} and S_m have no common root on I .*

Proof: The proof can be easily done by reverse induction on $m = n, \dots, 1$, taking into account that conditions 1 and 2 from Definition 30 imply that for $m < n$, any common root of S_{m-1} and S_m would also be a root of S_{m+1} . \square

We introduce some more useful notation.

Notation 33 *Let $a, b \in \mathbf{R}$, $n \in \mathbb{Z}_{\geq 1}$, (S_0, \dots, S_n) in $\mathbf{R}[X]$ and $\sigma, \tau \in \{-1, 1\}^{n-1}$. We define*

$$\epsilon(\sigma, \tau)_i := \prod_{1 \leq j \leq i-1} \sigma_j \tau_j$$

for $1 \leq i \leq n$ and

$$\text{Var}(\sigma, \tau)_a^b(S_0, \dots, S_n) := \sum_{1 \leq i \leq n} \epsilon(\sigma, \tau)_i \text{Var}_a^b(S_{i-1}, S_i).$$

Note that it is always the case that $\epsilon(\sigma, \tau)_1 = 1$.

Proposition 34 *Let $a, b \in \mathbf{R}$ with $a < b$, $I := [a, b]$, $n \in \mathbb{Z}_{\geq 1}$ and $\sigma, \tau \in \{-1, 1\}^{n-1}$. If (S_0, \dots, S_n) is a Sturm (σ, τ) -chain with respect to I and $\deg S_0, \dots, \deg S_n \leq d^2$, then*

$$\text{Ind}_a^b(S_1, S_0) + \epsilon(\sigma, \tau)_n \text{Ind}_a^b(S_{n-1}, S_n) = \text{Var}(\sigma, \tau)_a^b(S_0, \dots, S_n).$$

Note that the identity in Proposition 28 ([3, Theorem 3.11]) is exactly the identity in Proposition 34 in the particular case $\sigma = \tau = (1, 1, \dots, 1)$

Proof of Proposition 34: By Lemma 32, we know that for every $1 \leq m \leq n$, S_{m-1} and S_m have no common root on I . We proceed then by induction on n . For $n = 1$, the result holds by Proposition 18.

Now we take $n \geq 2$. Let x be a root of S_1 on I (and therefore x is not a root neither of S_0 nor of S_2). The identity

$$A_1 S_2 + B_1 S_1 + C_1 S_0 = 0$$

implies that $C_1(x)S_0(x) = -A_1(x)S_2(x) \neq 0$ and then $\text{sign}(S_0(x)) = -\sigma_1 \tau_1 \text{sign}(S_2(x))$. From this we deduce

$$\text{Ind}_a^b(S_0, S_1) = -\sigma_1 \tau_1 \text{Ind}_a^b(S_2, S_1).$$

We consider $\sigma' := (\sigma_2, \dots, \sigma_{n-1})$, $\tau' := (\tau_2, \dots, \tau_{n-1})$ and we apply the inductive hypothesis to the Sturm (σ', τ') -chain (S_1, \dots, S_n) . For $2 \leq i \leq n$ we have that $\epsilon(\sigma, \tau)_i = \sigma_1 \tau_1 \epsilon(\sigma', \tau')_{i-1}$.

Finally, using Proposition 18,

$$\begin{aligned} & \text{Ind}_a^b(S_1, S_0) + \epsilon(\sigma, \tau)_n \text{Ind}_a^b(S_{n-1}, S_n) \\ &= \text{Ind}_a^b(S_1, S_0) + \text{Ind}_a^b(S_0, S_1) + \sigma_1 \tau_1 \text{Ind}_a^b(S_2, S_1) + \sigma_1 \tau_1 \epsilon(\sigma', \tau')_{n-1} \text{Ind}_a^b(S_{n-1}, S_n) \\ &= \text{Var}_a^b(S_0, S_1) + \sigma_1 \tau_1 \text{Var}(\sigma', \tau')_a^b(S_1, \dots, S_n) \\ &= \text{Var}(\sigma, \tau)_a^b(S_0, \dots, S_n) \end{aligned}$$

as we wanted to prove. □

Corollary 35 *Let $a, b \in \mathbf{R}$ with $a < b$, $I := [a, b]$, $n \in \mathbb{Z}_{\geq 1}$ and $\sigma, \tau \in \{-1, 1\}^{n-1}$. If (S_0, \dots, S_n) is a good Sturm (σ, τ) -chain with respect to I and $\deg S_0, \dots, \deg S_n \leq d^2$, then*

$$\text{Ind}_a^b(S_1, S_0) = \text{Var}(\sigma, \tau)_a^b(S_0, \dots, S_n).$$

Proof: Since (S_0, \dots, S_n) is a good Sturm (σ, τ) -chain with respect to I , S_n has no roots on I and

$$\text{Ind}_a^b(S_{n-1}, S_n) = 0,$$

therefore the claim holds by Proposition 34. □

3.2 Quantitative Main Lemma

Our next goal is to prove a quantitative adaptation of the Main Lemma (see [3, Lemma 5.3]): using $[\text{IVT}]_{d^2}$ and subresultants we want to prove that if $F \in \mathbf{C}[X, Y]$ with $\deg F \leq d$ does not vanish on Γ , then $w(F | \partial\Gamma) = 0$.

In order to be able to work with subresultants, we need to consider separately for $F \in \mathbf{C}[X, Y]$ the degrees with respect to X and Y of F_{re} and F_{im} , and we wish each of these two degrees of F_{re} to drop with respect to the respective degree of F_{im} . Since

$$(iF)_{\text{re}} = -F_{\text{im}} \quad \text{and} \quad (iF)_{\text{im}} = F_{\text{re}},$$

up to multiplication by i , it will be enough for our purposes if these degrees are different. We will also need some degree control on some auxiliary subresultant polynomials which will play a key role in our proof. For these reasons, we introduce the following definition.

Definition 36 *Let $F \in \mathbf{C}[X, Y]$. We say that F is well-controlled if the following conditions are satisfied:*

1. $F_{\text{re}}, F_{\text{im}} \neq 0$,
2. $\deg_X F_{\text{im}} \neq \deg_X F_{\text{re}}$ and $\deg_Y F_{\text{im}} \neq \deg_Y F_{\text{re}}$.

For a well-controlled F , we denote by F^X the unique polynomial in $\{F, iF\} \subset \mathbf{C}[X, Y]$ such that $\deg_X F_{\text{im}}^X > \deg_X F_{\text{re}}^X$. Similarly, we denote by F^Y the unique polynomial in $\{F, iF\} \subset \mathbf{C}[X, Y]$ such that $\deg_Y F_{\text{im}}^Y > \deg_Y F_{\text{re}}^Y$.

Example 37 *Let $F \in \mathbf{C}[Z] \setminus \mathbf{C}$ with F monic. Then F is well-controlled.*

Now we are ready to prove our quantitative version of the Main Lemma [3, Lemma 5.3]. We stress the fact that it is in the proof of the Quantitative Main Lemma (Lemma 38) that subresultant polynomials play a key role to keep control of the degree of intermediate auxiliary polynomials. We need to use the hypothesis $[\text{IVT}]_{d^2}$ for proving the result for a polynomial of degree $\leq d$.

Lemma 38 *(Quantitative Main Lemma) Suppose that (\mathbf{R}, \leq) is an ordered field satisfying $[\text{IVT}]_{d^2}$. Let $\Gamma := [x_0, x_1] \times [y_0, y_1] \subset \mathbf{R}^2$ and $F \in \mathbf{C}[X, Y]$ with $\deg F \leq d$ and such that F does not vanish on Γ . If F is well-controlled, then $w(F | \partial\Gamma) = 0$.*

Proof: We will produce in several steps a suitable grid partition of Γ into a finite number of rectangles $\Gamma_1, \dots, \Gamma_s$ and we will prove that $w(F | \partial\Gamma_i) = 0$ for $1 \leq i \leq s$. Then the result will follow from Lemma 14. Let G be a greatest common divisor of F_{re} and F_{im} in the unique factorization domain $\mathbf{R}[X, Y]$ and let $F^X, F^Y \in \mathbf{C}[X, Y]$ be as in Definition 36.

First step. We decompose $F_{\text{im}}^X = GS_0$ and $F_{\text{re}}^X = GS_1$ and note that we have $\deg S_0, \deg S_1 \leq d$ and

$$d \geq p := \deg_X S_0 > q := \deg_X S_1.$$

We consider S_0 and S_1 as elements of $\mathbf{R}[Y][X]$ and we take the subresultant polynomial sequence with respect to the variable X (as in Section 2.4, taking $\mathbf{D} = \mathbf{R}[Y]$)

$$\text{sResP}_p(S_0, S_1) = S_0, \text{sResP}_{p-1}(S_0, S_1) = S_1, \dots, \text{sResP}_0(S_0, S_1) \in \mathbf{R}[Y][X].$$

For $0 \leq j \leq p$ and $0 \leq i \leq j$, the degree in Y of the coefficient of X^i in $\text{sResP}_j(S_0, S_1) \in \mathbf{R}[Y][X]$ is bounded by d^2 . This is so by Proposition 27 for $0 \leq j \leq q$ and by definition of subresultant polynomials for $q < j \leq p$.

We take (d_0, d_1, \dots, d_s) as the sequence of degrees of the non-defective subresultant polynomials in decreasing order (note that $d_0 = p$ and $d_1 = q$), and $d_{-1} := p + 1$, and we define \mathcal{Y}_1 as the union of the sets of roots in $[y_0, y_1]$ of the polynomials $S_s, T_{d_{-1}-1}(S_0, S_1), \dots, T_{d_{s-2}-1}(S_0, S_1), \text{sR}_{d_0}(S_0, S_1), \dots, \text{sR}_{d_{s-1}}(S_0, S_1) \in \mathbf{R}[Y] \setminus \{0\}$ (note that $T_{d_{-1}-1}(S_0, S_1) = \text{sR}_{d_0}(S_0, S_1) = 1$, therefore these two polynomials actually add no roots to the set \mathcal{Y}_1 ; note also that in the particular case $s = 1$, \mathcal{Y}_1 is just the set of roots in $[y_0, y_1]$ of the polynomial S_1).

For uniformity reasons in exposition we define $\mathcal{Y}_2 := \{y_0, y_1\}$ and we also define

$$\mathcal{Y}_3 := \{y \in [y_0, y_1] \mid F_{\text{re}}^X(X, y) = 0 \in \mathbf{R}[X] \text{ or } F_{\text{im}}^X(X, y) = 0 \in \mathbf{R}[X]\}.$$

Finally, we define

$$\mathcal{Y} := \mathcal{Y}_1 \cup \mathcal{Y}_2 \cup \mathcal{Y}_3.$$

We think of \mathcal{Y} as the set of the Y -coordinates of bad behaving points in Γ . Suppose $\mathcal{Y} = \{b_1, \dots, b_\ell\}$ with $y_0 = b_1 < \dots < b_\ell = y_1$.

Second step. We proceed as in the first step, but replacing polynomial F^X by F^Y and the roles of variables X and Y , to produce a set $\mathcal{X} \subset [x_0, x_1]$, which we think of as the set of the X -coordinates of bad behaving points in Γ . Suppose $\mathcal{X} = \{a_1, \dots, a_k\}$ with $x_0 = a_1 < \dots < a_k = x_1$.

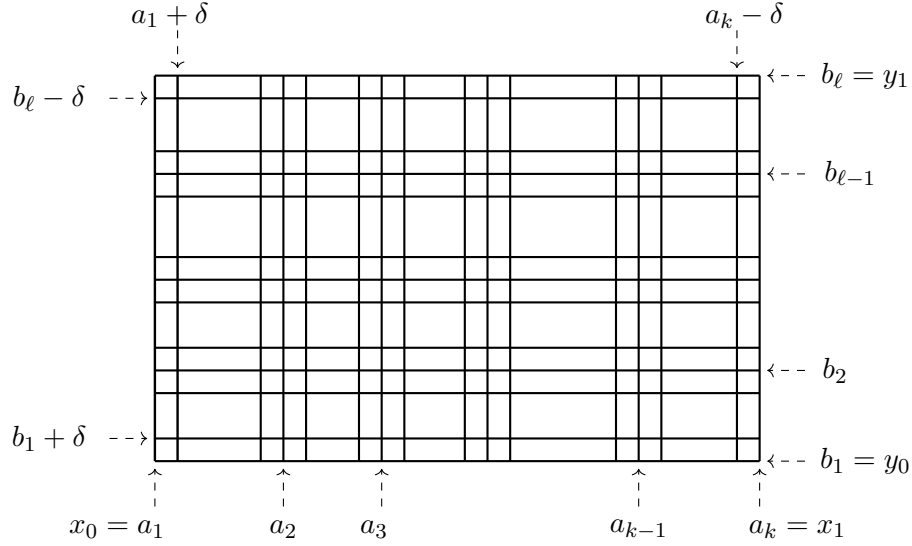
Third step. We take $\mathcal{Z} := \mathcal{X} \times \mathcal{Y} \subset \Gamma$. For each $z = (a, b) \in \mathcal{Z}$, since $F(a, b) \neq 0$, by Proposition 23 there exist $\delta_z > 0$ such that the winding number of F vanishes on any rectangle contained in $[a - \delta_z, a + \delta_z] \times [b - \delta_z, b + \delta_z]$. So we take $\delta > 0$, with $\delta \leq \delta_z$ for every $z \in \mathcal{Z}$ and such that

$$x_0 = a_1 < a_1 + \delta < a_2 - \delta < a_2 + \delta < a_3 - \delta < \dots < a_{k-1} + \delta < a_k - \delta < a_k = x_1$$

and

$$y_0 = b_1 < b_1 + \delta < b_2 - \delta < b_2 + \delta < b_3 - \delta < \dots < b_{\ell-1} + \delta < b_\ell - \delta < b_\ell = y_1.$$

We divide intervals $[x_0, x_1]$ and $[y_0, y_1]$ using all these numbers above, and finally we use these divisions of these intervals to obtain a grid partition of $\Gamma = [x_0, x_1] \times [y_0, y_1]$.



Now that the grid partition is defined, we have to prove that the winding number of F vanishes on each rectangle in the grid. Take $\Gamma' = [a, a'] \times [b, b'] \subset \Gamma$ as one of this rectangles. Then either there is a single point of \mathcal{Z} in Γ' or there is no point of \mathcal{Z} in Γ' . In the first case, $w(F | \partial\Gamma') = 0$ by Proposition 23 and the choice of δ . In the second case, then either $[a, a'] \cap \mathcal{X} = \emptyset$ or $[b, b'] \cap \mathcal{Y} = \emptyset$.

Let us suppose first that $[b, b'] \cap \mathcal{Y} = \emptyset$ holds and prove that $w(F^X | \partial\Gamma') = 0$; then $w(F | \partial\Gamma') = 0$ as well either because $F^X = F$ or because $F^X = iF$ and by Proposition 19.

Since F is well-controlled of degree at most d and the polynomials $T_{d-1-1}(S_0, S_1), \dots, T_{d_s-2-1}(S_0, S_1)$, $sR_{d_0}(S_0, S_1), \dots, sR_{d_{s-1}}(S_0, S_1) \in \mathbf{R}[Y]$ are coefficients of subresultant polynomials of S_0 and S_1 with respect to variable X , their degree in Y is bounded by d^2 . Since none of these polynomials vanishes on $[b, b']$, using the notation and results from Example 31 b), for $1 \leq i \leq s-1$ we have that A_i and C_i have constant sign different from 0 on $[b, b']$ and

- for any $y \in [b, b']$, $(S_0(X, y), \dots, S_s(X, y)) \in \mathbf{R}[X]$ is a good Sturm (σ, τ) -chain with respect to $[a, a']$ with all its elements with degree bounded by $d \leq d^2$,
- for any $x \in [a, a']$, $(S_0(x, Y), \dots, S_s(x, Y)) \in \mathbf{R}[Y]$ is a good Sturm (σ, τ) -chain with respect to $[b, b']$ with all its elements with degree bounded by d^2 .

Taking into account that

$$0 \neq F_{\text{re}}^X(X, b), F_{\text{im}}^X(X, b), F_{\text{re}}^X(X, b'), F_{\text{im}}^X(X, b') \in \mathbf{R}[X],$$

$$0 \neq F_{\text{re}}^X(a, Y), F_{\text{im}}^X(a, Y), F_{\text{re}}^X(a', Y), F_{\text{im}}^X(a', Y) \in \mathbf{R}[Y],$$

we conclude that

$$\begin{aligned}
2w(F^X | \partial\Gamma') &= \text{Ind}_a^{a'}(F_{\text{re}}^X(X, b), F_{\text{im}}^X(X, b)) + \text{Ind}_b^{b'}(F_{\text{re}}^X(a', Y), F_{\text{im}}^X(a', Y)) \\
&\quad + \text{Ind}_{a'}^a(F_{\text{re}}^X(X, b'), F_{\text{im}}^X(X, b')) + \text{Ind}_{b'}^b(F_{\text{re}}^X(a, Y), F_{\text{im}}^X(a, Y)) \\
&= \text{Ind}_a^{a'}(S_1(X, b), S_0(X, b)) + \text{Ind}_b^{b'}(S_1(a', Y), S_0(a', Y)) \\
&\quad + \text{Ind}_{a'}^a(S_1(X, b'), S_0(X, b')) + \text{Ind}_{b'}^b(S_1(a, Y), S_0(a, Y)) \\
&= \text{Var}(\sigma, \tau)_a^{a'}(S_0(X, b), \dots, S_s(X, b)) + \text{Var}(\sigma, \tau)_b^{b'}(S_0(a', Y), \dots, S_s(a', Y)) \\
&\quad + \text{Var}(\sigma, \tau)_{a'}^a(S_0(X, b'), \dots, S_s(X, b')) + \text{Var}(\sigma, \tau)_{b'}^b(S_0(a, Y), \dots, S_s(a, Y)) \\
&= 0
\end{aligned}$$

using Corollary 35.

In case that $[a, a'] \cap \mathcal{X} = \emptyset$ holds, we proceed in a similar way exchanging the roles of X and Y , to prove that $w(F^Y | \partial\Gamma') = 0$, and then we have that $w(F | \partial\Gamma') = 0$ again either because $F^Y = F$ or because $F^Y = iF$ and by Proposition 19. \square

3.3 The winding number counts the complex roots

From Example 13, Proposition 19 and the Quantitative Main Lemma 38 we deduce the following result.

Theorem 39 *Let $\Gamma \subset \mathbf{R}^2$ be a rectangle and $F \in \mathbf{C}[Z]$ with $\deg F \leq d$ and such that F does not vanish in $\partial\Gamma$. Then $w(F | \partial\Gamma)$ counts the number of zeros of F in the interior of Γ with multiplicity.*

Proof: Factorize $F = a \cdot (Z - z_1) \cdot \dots \cdot (Z - z_r) \cdot \tilde{F}$ with $a \in \mathbf{C}$, $z_1, \dots, z_r \in \Gamma \setminus \partial\Gamma$ and monic $\tilde{F} \in \mathbf{C}[Z]$ with no roots in Γ . If $\tilde{F} = 1$ the result follows from Example 13 and Proposition 19. Otherwise, since \tilde{F} is well-controlled (Example 37) the result follows from Example 13 and Proposition 19 and the Main Lemma 38. \square

3.4 Quantitative Homotopy

The last ingredient for the proof of Theorem 3 is a quantitative homotopy tool similar to [3, Theorem 5.4 and Corollary 5.5, Proposition 5.8 and Theorem 5.9]. Since we need to deal with well-controlled polynomials, we have to divide the homotopy in two steps, one for the real part and one for the imaginary part.

Theorem 40 *Let $F \in \mathbf{C}[Z]$, with $F \neq 0$ and $\deg F = e \leq d$. There exists $r \in \mathbf{R}$, $r > 0$ such that if $m \geq r$ and $\Gamma := [-m, m] \times [-m, m]$, then $w(F | \partial\Gamma) = e$.*

Proof: If $e = 0$ there is nothing to prove, so we suppose $e > 0$. By Proposition 19, we can also suppose that F is monic. Let

$$F = Z^e + \sum_{j=0}^{e-1} (a_j + ib_j) Z^j$$

with $a_j, b_j \in \mathbf{R}$ for $0 \leq j \leq e-1$ and take $G := F - Z^e \in \mathbf{C}[Z]$ collecting all the terms of degree less than e in F .

We take the auxiliary polynomial

$$K := (X + iY)^e + G_{\text{re}} \in \mathbf{C}[X, Y].$$

Note that in general K does not come from a polynomial in $\mathbf{C}[Z]$ by means of the substitution $Z = X + iY$. The idea of the proof is to obtain $r \in \mathbf{R}$, $r > 0$ such that if $m \geq r$ and $\Gamma = [-m, m] \times [-m, m]$, then $e = w(Z^e | \partial\Gamma) = w(K | \partial\Gamma) = w(F | \partial\Gamma)$ (see Example 22).

We suppose $G_{\text{re}}, G_{\text{im}} \neq 0$, and if this is not the case, the rest of the proof can be simplified. Actually, the only case where $G_{\text{re}} = 0$ or $G_{\text{im}} = 0$ is $G \in \mathbf{R} \cup i\mathbf{R}$.

We define $H_0, H_1 \in \mathbf{C}[X, Y, T]$ by

$$H_0(X, Y, T) := (1 - T)(X + iY)^e + TK(X, Y) = (X + iY)^e + TG_{\text{re}}(X, Y)$$

and

$$H_1(X, Y, T) := (1 - T)K(X, Y) + TF(X + iY) = (X + iY)^e + G_{\text{re}}(X, Y) + iTG_{\text{im}}(X, Y).$$

Take $r := 1 + 2 \max\{|a_j + ib_j| \mid 0 \leq j \leq e-1\}$. We proceed similarly to [3, Proposition 5.8] to prove that both H_0, H_1 do not vanish on $\partial\Gamma \times [0, 1]$. For $k = 0, 1$, and $(x, y, t) \in \partial\Gamma \times [0, 1]$, we have

$$\begin{aligned} & |H_k(x, y, t) - (x + iy)^e| \\ & \leq 2 \left| \sum_{j=0}^{e-1} (a_j + ib_j)(x + iy)^j \right| \\ & \leq 2 \sum_{j=0}^{e-1} |a_j + ib_j| |x + iy|^j \\ & \leq (r - 1) \sum_{j=0}^{e-1} |x + iy|^j \\ & \leq |x + iy|^e - 1. \end{aligned}$$

Therefore

$$|H_k(x, y, t)| \geq |(x + iy)^e| - |H_k(x, y, t) - (x + iy)^e| \geq |x + iy|^e - |x + iy|^e + 1 = 1.$$

Now enlarge r if necessary so that $X \pm m$ and $Y \pm m$ are not factors of $F_{\text{re}}(X, Y), G_{\text{re}}(X, Y)$ and $G_{\text{im}}(X, Y)$. Then it can be verified that the polynomials $H_0(X, -m, T), H_0(X, m, T), H_1(X, -m, T),$

$H_1(X, m, T) \in \mathbf{C}[X, T]$ and $H_0(-m, Y, T), H_0(m, Y, T), H_1(-m, Y, T), H_1(m, Y, T) \in \mathbf{C}[Y, T]$ are of degree at most d and well-controlled.

Finally, take $\Gamma_X := [-m, m] \times [0, 1]$, $\Gamma_Y := [-m, m] \times [0, 1]$ and $\Gamma_T := \Gamma$. By the Main Lemma 38,

$$w(H_0(X, -m, T) | \partial\Gamma_Y) = w(H_0(X, m, T) | \partial\Gamma_Y) = w(H_1(X, -m, T) | \partial\Gamma_Y) = w(H_1(X, m, T) | \partial\Gamma_Y) = 0$$

and

$$w(H_0(-m, Y, T) | \partial\Gamma_X) = w(H_0(m, Y, T) | \partial\Gamma_X) = w(H_1(-m, Y, T) | \partial\Gamma_X) = w(H_1(m, Y, T) | \partial\Gamma_X) = 0.$$

Therefore, by Lemma 15 applied to H_0 and H_1 and the cube $[-m, m] \times [-m, m] \times [0, 1] \subset \mathbf{R}^3$, we have

$$e = w(Z^e | \partial\Gamma) = w(H_0(X, Y, 0) | \partial\Gamma_T) = w(H_0(X, Y, 1) | \partial\Gamma_T) = w(K | \partial\Gamma)$$

and

$$w(K | \partial\Gamma) = w(H_1(X, Y, 0) | \partial\Gamma_T) = w(H_1(X, Y, 1) | \partial\Gamma_T) = w(F | \partial\Gamma)$$

as we wanted to prove. □

3.5 Proof of Theorem 3

We are now ready to deduce our main result.

Proof of Theorem 3: As mentioned before, since $[\text{FTA}]_1$ holds even under no assumptions on (\mathbf{R}, \leq) we suppose $d \geq 2$. Take $F \in \mathbf{C}[Z] \setminus \mathbf{C}$ with $\deg F = e \leq d$. By Theorem 40, there exists $r \in \mathbf{R}$, $r > 0$ such that if $m \geq r$ and $\Gamma := [-m, m] \times [-m, m]$, then $w(F | \partial\Gamma) = e \geq 1$. By Theorem 39, $w(F | \partial\Gamma)$ counts the number of zeros of F in the interior of Γ with multiplicity. This implies that there exists at least one $z \in \Gamma \subset \mathbf{R}^2 \sim \mathbf{C}$ such that $F(z) = 0$. □

Acknowledgement: We would like to express our gratitude to the anonymous referees whose suggestions helped us to improve the readability of the paper.

References

- [1] S. Basu, R. Pollack R, M.-F. Roy, Algorithms in real algebraic geometry, Algorithms and Computation in Mathematics, 10, Second edition. *Springer-Verlag, Berlin*, 2006.
- [2] J. Bochnak, M. Coste, M.-F. Roy, Real algebraic geometry. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*, 36. *Springer-Verlag, Berlin*, 1998.
- [3] M. Eisermann, The fundamental theorem of algebra made effective: an elementary real-algebraic proof via Sturm chains. *Amer. Math. Monthly* 119 (2012), no. 9, 715–752.
- [4] B. Fine, G. Rosenberger, The fundamental theorem of algebra. Undergraduate Texts in Mathematics. *Springer-Verlag, New York*, 1997.

- [5] W. Habicht, Eine Verallgemeinerung des Sturmschen Wurzelzählverfahrens. *Comm. Math. Helvetici* 21 (1948), 99-116.
- [6] P.-S. Laplace, Leçons de mathématiques données à l'École normale en 1795, Oeuvres complètes, tome 14, 10–177, 1795.
- [7] T. Lickteig, M.-F. Roy, Cauchy Index computation. *Calcolo* 33 (1996), 337-351.
- [8] H. Lombardi, D. Perrucci, M.-F. Roy, An elementary recursive bound for effective Positivstellensatz and Hilbert 17-th problem. Accepted for publication at *Mem. Amer. Math. Soc.*