A new general formula to compute the Cauchy Index with Subresultants in an interval

Daniel Perrucci\* Marie-Françoise Roy\*

\* 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

December 7, 2018

Abstract
We present a new formula to compute the Cauchy index of a rational function in an interval using subresultant polynomials. There is no condition on the endpoints of the interval and the formula also involves in some cases less subresultant polynomials.

Keywords: Cauchy Index, Subresultant Polynomials.

AMS subject classifications: 14P99, 13P15, 12D10, 26C15

1 Introduction

Let \((R, \leq)\) be a real closed field. The Cauchy index is, classically, an integer value associated to a rational function with coefficients in \(R\), which, roughly speaking, counts its number of jumps from \(-\infty\) to \(+\infty\) minus its number of jumps from \(+\infty\) to \(-\infty\) in a given interval. This value is closely related with the computation of Tarski queries, and plays a significant role in many algorithms for resolution of polynomial equations and inequalities systems over \(R\) (see [1]).

Let \(P, Q \in R[X] \setminus \{0\}\). The usual definition of the Cauchy index of \(\frac{Q}{P}\) is made directly on intervals whose extremities are not roots of \(P\). In this paper we use the extended definition of the Cauchy index introduced in [2] Section 3], which is made first at roots of \(P\) and then on intervals without restriction.
Definition 1 Let $x \in \mathbb{R}$ and $P, Q \in \mathbb{R}[X] \setminus \{0\}$.

- The rational fraction $\frac{Q}{P}$ can be written uniquely

$$
\frac{Q}{P} = (X - x)^m \frac{\bar{Q}}{\bar{P}}
$$

with $m \in \mathbb{Z}$ and $\bar{P}(x) \neq 0, \bar{Q}(x) \neq 0$. For $\varepsilon \in \{+,-\}$, define

$$
\text{Ind}_x^\varepsilon \left( \frac{Q}{P} \right) = \begin{cases} 
\frac{1}{2} \cdot \varepsilon^m \cdot \text{sign} \left( \frac{\bar{Q}(x)}{\bar{P}(x)} \right) & \text{if } m < 0, \\
0 & \text{otherwise}.
\end{cases}
$$

- The Cauchy index of $\frac{Q}{P}$ at $x$ is

$$
\text{Ind}_x \left( \frac{Q}{P} \right) = \text{Ind}_x^+ \left( \frac{Q}{P} \right) - \text{Ind}_x^- \left( \frac{Q}{P} \right).
$$

Definition 2 Let $a, b \in \mathbb{R}$ with $a < b$ and $P, Q \in \mathbb{R}[X] \setminus \{0\}$. The Cauchy index of $\frac{Q}{P}$ between $a$ and $b$ is

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

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

Note that with this extended definition of the Cauchy index, the Cauchy index of a rational function on an interval belongs to $\frac{1}{2} \mathbb{Z}$ and is not necessarily an integer number.

In order to state our main result, we first need to extend the notion of sign of a rational function to degenerate cases, following [2].

Notation 3 Using the same notation as before, we denote

$$
\text{sign} \left( \frac{Q}{P} \right) = \begin{cases} 
\text{sign} \left( \bar{Q}(x) \bar{P}(x) \right) & \text{if } m = 0, \\
0 & \text{otherwise}.
\end{cases}
$$

It is also unavoidable to include definitions and properties concerning subresultant polynomials. We refer the reader to [1] for proofs and details.

Let $D$ be a domain and let $\text{ff}(D)$ be its fraction field.

Definition 4 Let $P, Q \in D[X] \setminus \{0\}$ with $\deg P = p \geq 1$ and $\deg Q = q < p$. 

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

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

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

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

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

where $\text{SyHa}_{j,i}(P, Q) \in \mathbb{R}^{(p+q-2j) \times (p+q-2j)}$ 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

$$\text{sRes}_p(P, Q) = P \in D[X],$$
$$\text{sRes}_{p-1}(P, Q) = Q \in D[X],$$
$$\text{sRes}_j(P, Q) = 0 \in D[X] \quad \text{for } q < j < p - 1.$$

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

$$\text{sRes}_p(P, Q) = 1 \in D \quad \text{(even if } P \text{ is not monic)},$$
$$\text{sRes}_j(P, Q) = 0 \in D \quad \text{for } q < j \leq p - 1.$$

For $0 \leq j \leq p$, $\text{sRes}_j(P, Q)$ is said to be

- **defective** if $\deg \text{sRes}_j(P, Q) < j$ or, equivalently, if $\text{sRes}_j(P, Q) = 0$,
- **non-defective** if $\deg \text{sRes}_j(P, Q) = j$ or, equivalently, if $\text{sRes}_j(P, Q) \neq 0$.

The following Structure Theorem is a key result in the theory of subresultants. To state it, we need to introduce a notation.

**Notation 5** For $n \in \mathbb{Z}$, we denote $\epsilon_n = (-1)^{\frac{1}{2}n(n-1)}$.

**Theorem 6 (Structure Theorem of Subresultants)** Let $P, Q \in D[X] \setminus \{0\}$ with $\deg P = p \geq 1$ and $\deg Q = q < p$. Let $(d_0, \ldots, d_s)$ be the sequence of degrees of the non-defective subresultant polynomials of $P$ and $Q$ in decreasing order and let $d_{s-1} = p + 1$ (note that $d_0 = p$ and $d_1 = q$).

- For $1 \leq i \leq s$,
  $$\text{sRes}_{d_{s-1} - 2}(P, Q) = \cdots = \text{sRes}_{d_{i+1}}(P, Q) = 0 \in D[X]$$
and $\text{sRes}_{d_i-1}(P, Q)$ and $\text{sRes}_{d_i}(P, Q)$ are proportional. More precisely, for $1 \leq i \leq s$, denote

$$T_i = \text{sRes}_{d_i-1}(P, Q) \in D[X],$$

$$t_i = \text{lc}(T_i) \in D$$

(note that $T_1 = Q$), and extend this notation with $T_0 = P$ and $t_0 = 1 \in D$. Then

$$\text{sRes}_{d_i}(P, Q) \cdot T_i = t_i \cdot \text{sRes}_{d_i}(P, Q) \in D[X]$$

with

$$\text{sRes}_{d_i}(P, Q) = \epsilon_{d_i-1-d_i} \cdot \frac{t_i^{d_i-1-d_i}}{\text{sRes}_{d_i-1}(P, Q)^{d_i-1-d_i}} \in D.$$

This implies $\deg T_i = d_i \leq d_{i-1} - 1$.

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

$$t_{i-1} \cdot \text{sRes}_{d_{i-1}}(P, Q) \cdot T_{i+1} = -\text{Rem}(t_i \cdot \text{sRes}_{d_i}(P, Q) \cdot T_{i-1}, T_i) \in D[X]$$

where Rem is the remainder in the euclidean division in $\text{ff}(D)[X]$ of the first polynomial by the second polynomial, and the quotient belongs to $D[X]$.

- Both $T_s \in D[X]$ and $\text{sRes}_{d_s}(P, Q) \in D[X]$ are greatest common divisors of $P$ and $Q$ in $\text{ff}(D)[X]$ and they divide $\text{sRes}_{p}(P, Q)$ for $0 \leq j \leq p$. In addition, if $d_s > 0$ then

$$\text{sRes}_{d_{s-1}}(P, Q) = \cdots = \text{sRes}_{0}(P, Q) = 0 \in D[X].$$
Proof: See [1, Chapter 8].

Finally, in order to state our main result we introduce the following notation.

**Notation 7** Using the same notation as before, for \(0 \leq i \leq s\), let

\[
p(i) = \max\{j \mid 0 \leq j \leq i, \ d_{j-1} - d_j \text{ is odd}\}
\]

\((p(i)\) is well-defined since \(d_{-1} - d_0 = 1\) is odd).

We are ready now to state our main result, which is a new formula to compute \(\text{Ind}_b^a\(\frac{Q}{P}\)\) using subresultants.

**Theorem 8** Let \(a, b \in \mathbb{R}\) with \(a < b\) and \(P, Q \in \mathbb{R}[X] \setminus \{0\}\) with \(\deg P = p \geq 1\) and \(\deg Q = q < p\). Then

\[
\text{Ind}_b^a\(\frac{Q}{P}\) = \frac{1}{2} \sum_{0 \leq i \leq s-1} \epsilon_{d_{p(i)-1} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i) \cdot \left(\text{sign}\left(\frac{T_i}{T_{i+1}}, b\right) - \text{sign}\left(\frac{T_i}{T_{i+1}}, a\right)\right).
\]

This paper is organized as follows. In Section 2 we show the consequences of Theorem 8 in some particular cases and we comment the difference between Theorem 8 and previously known results. In Section 3 we include some preliminary results about the Cauchy index. Finally, in Section 4 we prove Theorem 8 using the key notion of \((\sigma, \tau)\)-chain.

## 2 Consequences and comparisons with previously known results

### 2.1 Cauchy index with sign variations

Let \(a, b \in \mathbb{R}\) with \(a < b\) and \(P, Q \in \mathbb{R}[X] \setminus \{0\}\). If we add the condition that \(a\) and \(b\) are no roots of \(P\) and \(Q\), from Theorem 8 we obtain sign-variation-counting-like formulas which give a uniform treatment to the case of sequences which involves \(0\) as a sign.

We introduce the following useful notation.

**Notation 9** Let \(x \in \mathbb{R}\) and \(P, Q \in \mathbb{R}[X]\), we denote the sign variation of \((P, Q)\) at \(x\) by

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

If \(a, b \in \mathbb{R}\) with \(a < b\), 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,

\[
\text{Var}_a^b(P, Q) = \text{Var}_a(P, Q) - \text{Var}_b(P, Q).
\]
Note that for $x \in \mathbb{R}$,

$$\text{Var}_x(P, Q) = \begin{cases} 
0 & \text{if } P(x) \text{ and } Q(x) \text{ have same sign}, \\
1 & \text{if } P(x) \text{ and } Q(x) \text{ have opposite non-zero sign}, \\
\frac{1}{2} & \text{if exactly one of } P(x) \text{ and } Q(x) \text{ has zero sign}. 
\end{cases}$$

Theorem 6 clearly implies that if for some $0 \leq i \leq s - 1$, two consecutive polynomials $T_i$ and $T_{i+1}$ have a common root, then every polynomial in the subresultant sequence shares this root. So, suppose now that $a$ and $b$ are not common roots of $P$ and $Q$, and therefore they are not common roots of $T_i$ and $T_{i+1}$ for any $0 \leq i \leq s - 1$.

Under this assumption, we can use the following rule for the sign of a quotient.

**Remark 10** Let $x \in \mathbb{R}$ and $P, Q \in \mathbb{R}[X] \setminus \{0\}$ such that $x$ is not a common root of $P$ and $Q$. Then

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

We can then deduce from Theorem 6 the following sign-variation-counting-like formula for the Cauchy index.

**Corollary 11** Under the assumption that $a$ and $b$ are not common roots of $P$ and $Q$,

$$\text{Ind}_a^b\left(\frac{Q}{P}\right) = \sum_{0 \leq i \leq s - 1} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i) \cdot \text{Var}_a^b(T_i, T_{i+1}).$$

### 2.2 Previously known formula

There is a previously known formula to compute the Cauchy index $\text{Ind}_a^b\left(\frac{Q}{P}\right)$ by means of subresultant polynomials which is as follows (see [1, Chapter 9]).

**Notation 12** Let $s = s_n, 0, \ldots, 0, s'$, be a finite sequence of elements in $\mathbb{R}$ such that $s_n \neq 0$ and $s' = \emptyset$ or $s' = s_m, \ldots, s_0$ with $s_m \neq 0$. The **modified number of sign variations** in $s$ is defined inductively as follows

$$\text{MVar}(s) = \begin{cases} 
0 & \text{if } s' = \emptyset, \\
\text{MVar}(s') + 1 & \text{if } s_n s_m < 0, \\
\text{MVar}(s') + 2 & \text{if } s_n s_m > 0 \text{ and } n - m = 3, \\
\text{MVar}(s') & \text{if } s_n s_m > 0 \text{ and } n - m \neq 3.
\end{cases}$$

In other words, we modify the usual definition of the number of sign variations by counting 2 sign variations for the groups: $+, 0, 0, +$ and $-, 0, 0, -$. If there are no zeros in the sequence $s$, $\text{MVar}(s)$ is just the classical number of sign variations in the sequence.
Let $\mathcal{P} = P_0, P_1, \ldots, P_d$ be a sequence of polynomials in $\mathbb{R}[X]$ and let $x$ be an element of $\mathbb{R}$ which is not a root of $\gcd(\mathcal{P})$. Then $\text{MVar}(\mathcal{P}; x)$, the modified number of sign variations of $\mathcal{P}$ at $x$, is the number defined as follows:

- delete from $\mathcal{P}$ those polynomials that are identically 0 to obtain the sequence of polynomials $\mathcal{Q} = Q_0, \ldots, Q_s$ in $\mathbb{D}[X]$,
- define $\text{MVar}(\mathcal{P}; x)$ as $\text{MVar}(Q_0(x), \ldots, Q_s(x))$.

Let $a$ and $b$ be elements of $\mathbb{R}$ which are not roots of $\gcd(\mathcal{P})$. The difference between the number of modified sign variations in $\mathcal{P}$ at $a$ and $b$ is denoted by

$$\text{MVar}(\mathcal{P}; a, b) = \text{MVar}(\mathcal{P}; a) - \text{MVar}(\mathcal{P}; b).$$

Denoting by $\text{SResP}(P, Q)$ the list of subresultant polynomials of $P$ and $Q$, we have the following result.

**Proposition 13** Let $a, b \in \mathbb{R}$ with $a < b$ and $P, Q \in \mathbb{R}[X] \setminus \{0\}$ with $\deg P = p \geq 1$ and $\deg Q = q < p$. If $a$ and $b$ are not roots of $P$, then

$$\text{Ind}^b_a\left(\frac{Q}{P}\right) = \text{MVar}(\text{SResP}(P, Q); a, b).$$

The new formula to compute $\text{Ind}^b_a\left(\frac{Q}{P}\right)$ given in Theorem \S improves on Proposition \[3\] in several aspects:

- the first one is that Theorem \S is general and there are no restrictions on $a$ and $b$.
- the second and most important one is that potentially less subresultant polynomials are involved in this new formula. More precisely, the Structure Theorem of Subresultants (Theorem \[6\]), states that in the subresultant polynomial sequence, some polynomials appear only once and other polynomials appear exactly twice, always considering appearances up to scalar multiples. In addition to this, if a polynomial appears twice, its first appearance is more suitable for computation, since it can be defined as the polynomial determinant of a matrix of smaller size (in comparison with its second appearance). In respect to this, our formula involves only one appearance of each polynomial in the subresultant polynomial sequence, which is always the first one for polynomials which appear twice (as said before, always considering appearances up to scalar multiples).

In the special case that $a$ and $b$ are not a common root of $P$ and $Q$, Corollary \[11\] gives a sign-variation-counting-like formula better than Proposition \[3\] since:

- it is more general, since it may happen that $a$ and $b$ are roots of $P$. 

7
• it also more natural since the sign-variation-counting in Corollary \ref{cor:sign-var} is local and needs to consider
  the sign of two consecutive elements only, contrarily to the modified number of sign variations.
• as explained before, potentially less subresultant polynomials are involved.

Last but not least, the proofs of our results are also less technically involved than the proof of Propo-

\section{2.3 The non-defective case}

In the particular case that every subresultant polynomial is non-defective, then \( s = p \), and for \( 0 \leq i \leq p \), \( p(i) = i \) and \( T_i = \text{sRes}_P p_{-i} \). In this case, Theorem \ref{thm:main} simplifies as follows.

\begin{corollary}
 If every subresultant polynomial is non-defective, then
 \[ \text{Ind}_b^a \left( \frac{Q}{P} \right) = \frac{1}{2} \sum_{0 \leq j \leq p-1} \left( \text{sign} \left( \frac{\text{sRes}_P j + 1}{\text{sRes}_P j}, b \right) - \text{sign} \left( \frac{\text{sRes}_P j + 1}{\text{sRes}_P j}, a \right) \right). \]
\end{corollary}

Under the extra assumption that \( a \) and \( b \) are not common roots of \( P \) and \( Q \), we get the simplified

\begin{corollary}
 If every subresultant polynomial is non-defective and \( a \) and \( b \) are not common roots of \( P \) and \( Q \), then
 \[ \text{Ind}_b^a \left( \frac{Q}{P} \right) = \sum_{0 \leq j \leq p-1} \text{Var}_b^a(\text{sRes}_P j, \text{sRes}_P j + 1). \]
\end{corollary}

So in the non-defective case, Corollary \ref{cor:ind} generalizes the previously known result (see \cite{1} Chapters 2 and 9]), which was exactly the same formula for \( \text{Ind}_b^a \left( \frac{Q}{P} \right) \) under the extra assumption that \( a \) and \( b \)

\section{2.4 Cauchy index between \(-\infty\) and \(+\infty\)}

One final remark to be done is how to interpret our main formula to compute the Cauchy index of a

\[ \text{Ind}_R \left( \frac{Q}{P} \right) = \sum_{x \in \mathbb{R}} \text{Ind}_x \left( \frac{Q}{P} \right). \]

As usual, this can be computed taking \( a = -r \) and \( b = r \) with \( r \) big enough. We introduce the notation

\[ \text{Var}_{-\infty}(P, Q) = \frac{1}{2} \left| (-1)^{\deg(P)} \text{sign}(\text{lct}(P)) - (-1)^{\deg(Q)} \text{sign}(\text{lct}(Q)) \right|, \]

\[ \text{Var}_{+\infty}(P, Q) = \frac{1}{2} \left| \text{sign}(\text{lct}(P)) - \text{sign}(\text{lct}(Q)) \right|, \]

\[ \text{Var}^{+\infty}_{-\infty}(P, Q) = \text{Var}_{-\infty}(P, Q) - \text{Var}_{+\infty}(P, Q). \]
Note that, if \( \deg(P) - \deg(Q) \) is even, then \( \text{Var}^{\infty}_{\infty}(P, Q) = 0 \), and if \( \deg(P) - \deg(Q) \) is odd, then \( \text{Var}^{\infty}_{\infty}(P, Q) = \text{sign} \circ (\text{lc}(P)) \cdot \text{sign} \circ (\text{lc}(Q)) \).

Since there is an ad-hoc definition of \( t_0 = 1 \) (and not as the leading coefficient of \( T_0 = P \)), we obtain the following formula.

**Corollary 16** If the leading coefficient of \( P \) is positive or if \( d_0 - d_1 = p - q \) is even, then
\[
\text{Ind}_{\mathbb{R}} \left( \frac{Q}{P} \right) = \sum_{0 \leq i \leq s-1, \ d_i - d_{i+1} \text{ odd}} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_{i+1}).
\]

If the leading coefficient of \( P \) is negative and \( d_0 - d_1 = p - q \) is odd, then
\[
\text{Ind}_{\mathbb{R}} \left( \frac{Q}{P} \right) = -\text{sign}(t_1) + \sum_{1 \leq i \leq s-1, \ d_i - d_{i+1} \text{ odd}} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_{i+1}).
\]

**Proof:** Choosing \( r \in \mathbb{R} \) big enough and applying Corollary 11
\[
\text{Ind}_{\mathbb{R}} \left( \frac{Q}{P} \right) = \text{Ind}^{-r}_{\mathbb{R}} \left( \frac{Q}{P} \right) = \sum_{0 \leq i \leq s-1} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i) \cdot \text{Var}^{-r}_{\mathbb{R}}(T_i, T_{i+1})
\]
\[
= \sum_{0 \leq i \leq s-1} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i) \cdot \text{Var}^{\infty}_{\infty}(T_i, T_{i+1})
\]
\[
= \sum_{0 \leq i \leq s-1, \ d_i - d_{i+1} \text{ odd}} \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i) \cdot \text{sign}(\text{lc}(T_i)) \cdot \text{sign}(\text{lc}(T_{i+1})).
\]

From this identity the result can be easily proved. \( \square \)

Even in this special case, the new formula to compute \( \text{Ind}^{b}_{\mathbb{R}} \left( \frac{Q}{P} \right) \) given in Corollary 16 improves on the previously known one, which we introduce below (see [1, Chapter 4]).

**Proposition 17** For \( 0 \leq i \leq s \), let \( s_i = s\text{Res}_{d_i}(P, Q) \) be the leading coefficient of the non-defective subresultant \( s\text{Res}_{d_i}(P, Q) \) (which is proportional to \( T_i \)). Then
\[
\text{Ind}_{\mathbb{R}} \left( \frac{Q}{P} \right) = \sum_{0 \leq i \leq s-1, \ d_i - d_{i+1} \text{ odd}} \epsilon_{d_i - d_{i+1}} \cdot \text{sign}(s_i) \cdot \text{sign}(s_{i+1}).
\]

The main difference between the two formulas is that the \( t_i \) are, in the defective cases, defined as determinants of matrices of smaller sizes than the \( s_i \) and therefore can be computed more efficiently.

On the other hand, one advantage of Proposition 17 is that it can be proved directly, using minors extracted from the Hermite matrix and does not use the definition of the subresultant polynomials and the Structure Theorem of subresultants (see [1, Chapter 4]).
3 Preliminaries on Cauchy index

In this section we include some useful properties of Cauchy index.

Lemma 18 Let $a, b \in \mathbb{R}$ with $a < b$, $P, Q \in \mathbb{R}[X] \setminus \{0\}$ and $c \in \mathbb{R} \setminus \{0\}$. Then
\[ \text{Ind}_{b}^{a} \left( \frac{c \cdot Q}{P} \right) = \text{sign}(c) \cdot \text{Ind}_{a}^{b} \left( \frac{Q}{P} \right). \]

Proof: Follows immediately from the definition of Cauchy index. \hfill \square

Lemma 19 Let $a, b \in \mathbb{R}$ with $a < b$, $P, Q, R \in \mathbb{R}[X] \setminus \{0\}$ and $T \in \mathbb{R}[X]$ such that
\[ Q = PT + R. \]
Then
\[ \text{Ind}_{b}^{a} \left( \frac{Q}{P} \right) = \text{Ind}_{a}^{b} \left( \frac{R}{P} \right). \]

Proof: For each $x \in [a, b]$, we first note that if
\[ \frac{Q}{P} = (X - x)^{m} \frac{\tilde{Q}}{P} \]
with $m \in \mathbb{Z}$, $\tilde{P}(x) \neq 0$, $\tilde{Q}(x) \neq 0$ and $m < 0$, then defining
\[ \tilde{R} = \tilde{Q} - (X - x)^{-m} \tilde{P}T, \]
we have
\[ \frac{R}{P} = (X - x)^{m} \frac{\tilde{R}}{P} \]
with $\tilde{P}(x) \neq 0$ and $\tilde{R}(x) = \tilde{Q}(x) \neq 0$. This proves that $\text{Ind}_{x}^{\epsilon} \left( \frac{Q}{P} \right) = \text{Ind}_{x}^{\epsilon} \left( \frac{R}{P} \right)$ for every $\epsilon \in \{−1, 1\}$. The claim follows from the definition of the Cauchy index. \hfill \square

The following property is known as the inversion formula.

Proposition 20 Let $a, b \in \mathbb{R}$ with $a < b$ and $P, Q \in \mathbb{R}[X] \setminus \{0\}$. Then
\[ \text{Ind}_{a}^{b} \left( \frac{Q}{P} \right) + \text{Ind}_{b}^{a} \left( \frac{P}{Q} \right) = \frac{1}{2} \text{sign} \left( \frac{P}{Q}, b \right) - \frac{1}{2} \text{sign} \left( \frac{P}{Q}, a \right). \]

Proof: See [2, Theorem 3.9]. \hfill \square
4 Main result

4.1 \((\sigma, \tau)\)-chains and Cauchy index

The notion of \((\sigma, \tau)\)-chain was introduced in [3]. Here, we need to introduce a slight variation of this notion.

**Definition 21** Let \(n \in \mathbb{Z}_{\geq 1}\) and \(\sigma, \tau \in \{-1, 1\}^{n-1}\) with \(\sigma = (\sigma_1, \ldots, \sigma_{n-1})\) and \(\tau = (\tau_1, \ldots, \tau_{n-1})\). A sequence of polynomials \((S_0, \ldots, S_n) \in \mathbb{R}[X]\) is a special \((\sigma, \tau)\)-chain if for \(1 \leq i \leq n - 1\) there exist \(a_i, c_i \in \mathbb{R} \setminus \{0\}\) and \(B_i \in \mathbb{R}[X]\) such that

1. \(a_i S_{i+1} + B_i S_i + c_i S_{i-1} = 0\),
2. \(\text{sign}(a_i) = \sigma_i\),
3. \(\text{sign}(c_i) = \tau_i\).

As in [3], note that for \(n = 1\), taking \(\{-1, 1\}^0 = \{\bullet\}\), any sequence \((S_0, S_1)\) in \(\mathbb{R}[X]\) is a special \((\bullet, \bullet)\)-chain.

We introduce some more useful notation.

**Notation 22** Let \(a, b \in \mathbb{R}, n \in \mathbb{Z}_{\geq 1}\), \((S_0, \ldots, S_n) \in \mathbb{R}[X]\) and \(\sigma, \tau \in \{-1, 1\}^{n-1}\). We define for \(0 \leq i \leq n - 1\),

\[
\theta(\sigma, \tau)_i = \prod_{1 \leq j \leq i} \sigma_j \tau_j
\]

and

\[
W(\sigma, \tau)_a^b(S_0, \ldots, S_n) = \frac{1}{2} \sum_{0 \leq i \leq n-1} \theta(\sigma, \tau)_i \cdot \left( \text{sign}\left(\frac{S_i}{S_{i+1}}, b\right) - \text{sign}\left(\frac{S_i}{S_{i+1}}, a\right) \right).
\]

Note that it is always the case that \(\theta(\sigma, \tau)_0 = 1\).

The following result is an extension of [2, Theorem 3.11].

**Proposition 23** Let \(a, b \in \mathbb{R}\) with \(a < b\), \(n \in \mathbb{Z}_{\geq 1}\) and \(\sigma, \tau \in \{-1, 1\}^{n-1}\). If \((S_0, \ldots, S_n)\) is a special \((\sigma, \tau)\)-chain then

\[
\text{Ind}^b_a\left(\frac{S_1}{S_0}\right) + \theta(\sigma, \tau)_{n-1} \cdot \text{Ind}^b_a\left(\frac{S_{n-1}}{S_n}\right) = W(\sigma, \tau)_a^b(S_0, \ldots, S_n).
\]

**Proof:** We proceed by induction in \(n\). If \(n = 1\), the result follows from Proposition [20] (Inversion Formula).

Suppose now that \(n \geq 2\). By Lemmas [18] and [19] it is easy to see that

\[
\text{Ind}^b_a\left(\frac{S_0}{S_1}\right) + \sigma_1 \cdot \tau_1 \cdot \text{Ind}^b_a\left(\frac{S_2}{S_1}\right) = 0.
\]
We consider \( \sigma' = (\sigma_2, \ldots, \sigma_{n-1}) \), \( \tau' = (\tau_2, \ldots, \tau_{n-1}) \) and we apply the inductive hypothesis to the special \((\sigma', \tau')\)-chain \((S_1, \ldots, S_n)\). For \( 1 \leq i \leq n - 1 \) we have that \( \theta(\sigma, \tau)_i = \sigma_1 \cdot \tau_1 \cdot \theta(\sigma', \tau')_{i-1} \). Finally, using Proposition 20 (Inversion Formula) and the inductive hypothesis, we consider

\[
\Ind_a^b \left( \frac{S_1}{S_0} \right) + \theta(\sigma, \tau)_{n-1} \cdot \Ind_a^b \left( \frac{S_{n-1}}{S_n} \right)
\]

\[
= \Ind_a^b \left( \frac{S_1}{S_0} \right) + \Ind_a^b \left( \frac{S_0}{S_1} \right) + \sigma_1 \cdot \tau_1 \cdot \Ind_a^b \left( \frac{S_2}{S_1} \right) + \sigma_1 \cdot \tau_1 \cdot \theta(\sigma', \tau')_{n-2} \cdot \Ind_a^b \left( \frac{S_{n-1}}{S_n} \right)
\]

\[
= \frac{1}{2} \sign \left( \frac{S_0}{S_1}, a \right) + \frac{1}{2} \sign \left( \frac{S_0}{S_1}, b \right) + \sigma_1 \cdot \tau_1 \cdot W(\sigma', \tau')_a(S_1, \ldots, S_n)
\]

as we wanted to prove. \( \square \)

**Corollary 24** Let \( a, b \in \mathbb{R} \) with \( a < b \), \( n \in \mathbb{Z}_{\geq 1} \) and \( \sigma, \tau \in \{-1, 1\}^{n-1} \). If \((S_0, \ldots, S_n)\) is a special \((\sigma, \tau)\)-chain and \( S_n \) divides \( S_{n-1} \), then

\[
\Ind_a^b \left( \frac{S_1}{S_0} \right) = W(\sigma, \tau)_a(S_0, \ldots, S_n).
\]

### 4.2 Proof of Theorem \[8\]

We fix the notation we will use from this point.

**Notation 25** Let \( P, Q \in \mathbb{R}[X] \setminus \{0\} \) with \( \deg P = p \geq 1 \) and \( \deg Q = q < p \). Let \((d_0, \ldots, d_n)\) be the sequence of degrees of the non-defective subresultant polynomials of \( P \) and \( Q \) in decreasing order and let \( d_{-1} = p + 1 \).

- **Using Notation \[7\]** for \( 1 \leq i \leq s - 1 \), let
  
  \[
  a_i = t_{i-1} \cdot \text{sRes}_{d_{i-1}}(P, Q) \in \mathbb{R},
  
  B_i = -\text{Quot}(t_i \cdot \text{sRes}_{d_i}(P, Q) \cdot T_{i-1}, T_i) \in \mathbb{R}[X],
  
  c_i = t_i \cdot \text{sRes}_{d_i}(P, Q) \in \mathbb{R}.
  
- **For** \( 1 \leq i \leq s - 1 \), let
  
  \[
  \sigma_i = \sign(a_i) \in \{-1, 1\},
  
  \tau_i = \sign(c_i) \in \{-1, 1\},
  
  \]

and let \( \sigma = (\sigma_1, \ldots, \sigma_{s-1}) \) and \( \tau = (\tau_1, \ldots, \tau_{s-1}) \).

**Lemma 26** \((T_0, \ldots, T_s)\) is a special \((\sigma, \tau)\)-chain satisfying, in addition, that \( T_s \) divides all its elements.
Proof: Recall that \( T_0 = P \) and \( T_1 = Q \). Also, by the Structure Theorem of Subresultants (Theorem 6), we have that for \( 1 \leq i \leq s - 1 \),

\[
a_i T_{i+1} + B_i T_i + c_i T_{i-1} = 0.
\]

The claim follows from the definition of \( \sigma, \tau \).

The following lemma explores the relation between the signs of the leading coefficients of the subresultants polynomials.

Lemma 27 Let \( P, Q \in R[X] \setminus \{0\} \) with \( \deg P = p \geq 1 \) and \( \deg Q = q < p \). Following Notation \( [7] \) and \( [4] \) for \( 0 \leq i \leq s \),

\[
\text{sign}(sRes_{d_i}(P, Q)) = \epsilon_{d_{p(i)}-1-d_i} \cdot \text{sign}(t_{p(i)}).
\]

Before proving the lemma, note that \( \epsilon_n = 1 \) if the remainder of \( n \) in the division by 4 is 0 or 1 and \( \epsilon_n = -1 \) if the remainder of \( n \) in the division by 4 is 2 or 3; this implies that for \( k \in \mathbb{Z} \)

\[
\epsilon_{2k+n} = (-1)^k \epsilon_n = \epsilon_{2k} \epsilon_n.
\]

Proof of Lemma 27: For \( i = 0 \) the result is clear. For \( 1 \leq i \leq s \), by the Structure Theorem of Subresultants (Theorem 6),

\[
\text{sign}(sRes_{d_i}(P, Q)) = \epsilon_{d_{i-1}-d_i} \cdot \text{sign}(t_{i})^{d_{i-1}-d_i} \cdot \text{sign}(sRes_{d_{i-1}}(P, Q))^{d_{i-1}-d_i-1}.
\]

We proceed then by induction on \( i - p(i) \). If \( i = p(i) \), then \( d_{i-1} - d_i \) is odd and

\[
\text{sign}(sRes_{d_i}(P, Q)) = \epsilon_{d_{p(i)-1}-d_i} \cdot \text{sign}(t_{p(i)}).
\]

If \( i > p(i) \), then \( d_{i-1} - d_i \) is even, \( p(i) = p(i-1) \) and \( i - 1 - p(i-1) < i - p(i) \); therefore by the inductive hypothesis,

\[
\text{sign}(sRes_{d_i}(P, Q)) = \epsilon_{d_{i-1}-d_i} \cdot \text{sign}(sRes_{d_{i-1}}(P, Q)) = \epsilon_{d_{i-1}-d_i} \cdot \epsilon_{d_{p(i-1)-1}-d_{i-1}} \cdot \text{sign}(t_{p(i-1)}) = \epsilon_{d_{i-1}-d_i} \cdot \epsilon_{d_{p(i)-1}-d_{i-1}} \cdot \text{sign}(t_{p(i)}) = \epsilon_{d_{p(i)-1}-d_i} \cdot \text{sign}(t_{p(i)})
\]

using equation (1).

Now we are ready to prove our main result.

Proof of Theorem 8: By Corollary 24 since \( (T_0, \ldots, T_s) \) is a special \((\sigma, \tau)\)-chain and \( T_s \) divides \( T_{s-1} \),

\[
\text{Ind}_a^b \left( \frac{Q}{P} \right) = W(\sigma, \tau)_a(T_0, \ldots, T_s) = \frac{1}{2} \sum_{0 \leq i \leq s-1} \theta(\sigma, \tau)_i \cdot \left( \text{sign} \left( \frac{T_i}{T_{i+1}}, b \right) - \text{sign} \left( \frac{T_i}{T_{i+1}}, a \right) \right).
\]
So, we only need to prove that for $0 \leq i \leq s - 1$,

$$\theta(\sigma, \tau)_i = \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i).$$

Indeed, using Lemma 27,

$$\theta(\sigma, \tau)_i = \prod_{1 \leq j \leq i} \sigma_j \cdot \tau_j$$

$$= \prod_{1 \leq j \leq i} \text{sign}(t_j - 1) \cdot \text{sign}(\text{Res}_{d_j - 1}(P, Q)) \cdot \text{sign}(t_j) \cdot \text{sign}(\text{Res}_{d_j}(P, Q))$$

$$= \text{sign}(\text{Res}_{d_i}(P, Q)) \cdot \text{sign}(t_i)$$

$$= \epsilon_{d_{p(i)} - d_i} \cdot \text{sign}(t_{p(i)}) \cdot \text{sign}(t_i)$$

and we are done. □

References

[1] S. Basu, R. Pollack, M.-F. Roy, Algorithms in real algebraic geometry, Algorithms and Computation in Mathematics, 10, Second edition. Springer-Verlag, Berlin, 2006.

[2] 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.

[3] D. Perrucci, M.-F. Roy, Quantitative Fundamental Theorem of Algebra. Preprint. Available as arxiv:1803.04358.