Journal of Symbolic Computation 35 (2003) 281–292

Journal of
Symbolic
Computation

www.elsevier.com/locate/jsc

An elementary approach to subresultants theory

M´hammed El Kahoui

Department of Mathematics, Faculty of Sciences Semlalia, Cadi Ayyad University, Boulevard My Abdellah, P.O. Box 2390, Marrakech, Morocco

Received 22 February 2001; accepted 31 October 2002

Abstract

In this paper we give an elementary approach to univariate polynomial subresultants theory. Most of the known results of subresultants are recovered, some with more precision, without using Euclidean divisions or existence of roots for univariate polynomials. The main contributions of this paper are not new results on subresultants, but rather extensions of the main results over integral rings to arbitrary commutative rings. © 2003 Elsevier Science Ltd. All rights reserved.

Keywords: Polynomial determinant; Subresultant sequence

1. Introduction

The Euclidean remainder sequence played a fundamental role for computing gcds in Euclidean domains such as $\mathbf{K}[y]$, where $\mathbf{K}$ is a commutative field (see Gathen and Luking, 2000; Loos, 1982 for a historical note). Also, in 1835 Sturm (see Sturm, 1835) found out a method to compute the number of real roots of a polynomial P using the Euclidean remainders of P and its derivative modulo some sign changes. Sturm's solution of the real root counting led later to a solution of the quantifier elimination problem over the reals (Tarski, 1951). In contrast to the algorithmic possibilities it offers, the Euclidean remainder sequence has a relatively bad numerical behaviour (see Lickteig and Roy, 2001). Moreover, because of the denominators appearing in its coefficients, when the input coefficients are parameter dependent, the Euclidean remainder sequence has bad specialization properties.

In Collins (1967) Collins studied the connection between subresultants and Euclidean remainders (see also Loos, 1982; Gathen and Luking, 2000 for further precisions). He proved in particular that the polynomials in the two sequences are pair-wise proportional, and thus they accomplish the same algorithmic tasks. Contrary to the Euclidean remainder sequence, the subresultant sequence has a good behaviour under specialization and a

E-mail address: elkahoui@ucam.ac.ma (M. El Kahoui).

0014-5793/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0747-7171(02)00135-9

well controlled growth of coefficient size (see Lombardi et al., 2000 for optimal complexity bounds). Moreover, there are actually many efficient algorithms to compute subresultants (see Lombardi et al., 2000; Ducos, 2000; Reischert, 1997; Gathen and Luking, 2000; Gathen and Gerhard, 1999; Lickteig and Roy, 1996a, b, 2001). The connection between subresultants and the Euclidean remainders offers as well a tool for proving facts about subresultants. Indeed, to prove any fact about subresultants one can first do it for the Euclidean remainder sequence and then transfer it to subresultants via the established connection. A typical instance of this way of reasoning is given in Hong (1997) where the behaviour of subresultants under composition is studied. Recently, Hong developed in (Hong, 1999) an alternative method for proving facts about subresultants. His method is based on an explicit expression of subresultants in terms of the roots of the input polynomials, and hence it offers the possibility of geometric reasoning.

These two methods are hard to generalize to other graded rings, the first one is based on a division process and the second one uses the existence of roots for univariate polynomials. These two concepts are closely related to univariate polynomials over integral rings.

In this paper we give an elementary approach to subresultants theory. By “elementary” we mean that every thing will be deduced from algebraic identities, and hence holds over arbitrary commutative rings. The paper is structured as follows: in Section 2 we study polynomial determinants and their basic properties. In Section 3 we recall some fundamental properties of subresultants. Section 4 is devoted to a systematic study of the algebraic identities fulfilled by subresultants. We give in this section new algebraic identities from which we deduce a new proof of the gap structure theorem (Lickteig and Roy, 1996a). In Section 5 we give elementary proofs of some well-known facts on the behaviour of subresultants under operations on polynomials.

2. Polynomial determinants and their basic properties

In the remainder of this paper $\mathbf{A}$ will be a commutative ring and $m \leq n$ will be two positive integers. We denote by $\mathcal{M}_{m,n}(\mathbf{A})$ the $\mathbf{A}$-module of $m \times n$ matrices with coefficients in $\mathbf{A}$. Consider the free $\mathbf{A}$-module $\mathcal{P}_n$ of polynomials with coefficients in $\mathbf{A}$ of degree at most $n-1$ equipped with the basis $\mathcal{B}_n = [y^{n-1}, \ldots, y, 1]$. A sequence of polynomials $[P_1, \ldots, P_m]$ in $\mathcal{P}_n$ will be identified with the $m \times n$ matrix whose row coefficients are the coordinates of the P_i’s in the basis $\mathcal{B}_n$.

Definition. Let $M = (a_{i,j})$ be a matrix in $\mathcal{M}_{m,n}(\mathbf{A})$. For $0 \leq j \leq n-m$ let d_j be the $m \times m$ minor of M extracted on the columns $1, \ldots, m-1, n-j$. The polynomial $\mathrm{DetPol}(M) = \sum_j d_j y^j$ is called the polynomial determinant of M.

The following well-known lemma shows that the polynomial $\mathrm{DetPol}(M)$ is the determinant of a matrix with coefficients in the ring $\mathbf{A}[y]$.

Lemma 2.1. *Let M be a matrix in $\mathcal{M}_{m,n}(\mathbf{A})$. For any $i = 1, \ldots, m$ let $P_i = \sum_j a_{i,j} y^{n-j}$ and M' be the $m \times m$ matrix whose $m-1$ first columns are the $m-1$ first columns of M and the coefficients of the last column are $P_1, \ldots, P_m$. Then* $\mathrm{DetPol}(M) = \mathrm{Det}(M')$.

2.1. Row and column operations

A row (respectively column) operation on a matrix $M \in \mathcal{M}_{m,n}(\mathbf{A})$ consists in multiplying to the left (respectively to the right) M by an $m \times m$ (respectively $n \times n$) matrix. It is relatively obvious to see that polynomial determinants behave nicely under row operations. More precisely we have:

Lemma 2.2. *Let M be a matrix in $\mathcal{M}_{m,n}(\mathbf{A})$. Then for any $m \times m$ matrix U one has* $\mathrm{DetPol}(UM) = \mathrm{Det}(U)\mathrm{DetPol}(M)$.

In contrast to row operations, polynomial determinants do not behave in a "nice" way under arbitrary column operations. Nevertheless, some results on this behaviour can be stated for specific classes of column operations. For our purpose, see Section 5, we consider column operations given by upper triangular matrices. Since we identify matrices with lists of polynomials, a column operation on a matrix can be viewed as applying an endomorphism of the $\mathbf{A}$-module $\mathbf{A}[y]$ to its rows. Column operations given by upper triangular matrices correspond to a special class of endomorphisms that we precise in the following definition.

Definition. We say that an endomorphism ϕ of the $\mathbf{A}$-module $\mathbf{A}[y]$ preserves degrees if for any integer d it sends any polynomial of degree d to a polynomial of degree $\leq d$. If moreover, ϕ sends any monic degree d polynomial to a monic degree d polynomial then we say that ϕ preserves degrees and volumes.

Given an endomorphism ϕ preserving degrees and n a positive integer, the restriction ϕ_n of ϕ to $\mathcal{P}_n$ is an endomorphism whose matrix in $\mathcal{B}_n$ is lower triangular. If moreover ϕ preserves volumes, the diagonal coefficients of this matrix are equal to 1 and so $\mathrm{Det}(\phi_n) = 1$. The following lemma tells how polynomial determinants behave under transformations by endomorphisms preserving degrees.

Lemma 2.3. *Let ϕ be an endomorphism of $\mathbf{A}[y]$ preserving degrees. Then for any $m \times n$ matrix $[P_1, \ldots, P_m]$ one has*

$$\mathrm{DetPol}([\phi(P_1), \ldots, \phi(P_m)]) = \alpha_n \phi(\mathrm{DetPol}([P_1, \ldots, P_m]))$$

where $\alpha_n \in \mathbf{A}$ depends only on ϕ, n and m. If ϕ is one to one then $\alpha_n = \mathrm{Det}(\phi_n)$ $\mathrm{Det}(\phi_{n-m+1})^{-1}$, and $\alpha_n = 1$ if moreover ϕ preserves degrees and volumes.

Proof. The matrix $U = (u_{i,j})$ of ϕ_n in the basis $\mathcal{B}_n$ is lower triangular. Let c_j (respectively c'_j) be the jth column of $[P_1, \ldots, P_m]$ (respectively $[\phi(P_1), \ldots, \phi(P_m)]$). Since $[\phi(P_1), \ldots, \phi(P_m)] = [P_1, \ldots, P_m]U^T$ and U^T is upper triangular one has

$$c'_j = u_{j,j}c_j + \sum_{i=1}^{j-1} u_{j,i}c_i \qquad j = 1, \ldots, n. \tag{1}$$

According to Lemma 2.1 the polynomial $\mathrm{DetPol}([\phi(P_1), \ldots, \phi(P_m)])$ is the determinant of the matrix K built with the columns $c'_1, \ldots, c'_{m-1}$ and the column of coefficients $\phi(P_1), \ldots, \phi(P_m)$. Using multilinearity of determinants and the relation (1) one gets $\mathrm{Det}(K) = \left(\prod_{j=1}^{m-1} u_{j,j}\right)\mathrm{Det}(K')$ where K' is the matrix whose $m-1$ first columns are $c_1, \dots, c_{m-1}$ and the coefficients of the last column are $\phi(P_1), \dots, \phi(P_m)$. According to the linearity of $\mathrm{Det}(K')$ with respect to the last column of K' one gets $\mathrm{DetPol}([\phi(P_1), \dots, \phi(P_m)]) = \alpha_n \phi(\mathrm{DetPol}([P_1, \dots, P_m]))$ with $\alpha_n = \prod_{j=1}^{m-1} u_{j,j}$. The quantity α_n obviously depends only on ϕ, n and m. Moreover, if ϕ is one to one then $\alpha_n = \mathrm{Det}(\phi_n)\mathrm{Det}(\phi_{n-m+1})^{-1}$. □

3. Subresultants and their basic properties

Let p, q be nonnegative integers and $P, Q \in \mathbf{A}[y]$ be two polynomials with $\deg(P) \le p$ and $\deg(Q) \le q$. Let $\delta(p,q) = \min(p,q)$ if $p \ne q$ and $\delta(p,q) = q-1$ if $p = q$ and $p > 0$ (note here that we exclude the case $p = q = 0$). Let us write

$$P = a_0 y^p + a_1 y^{p-1} + \cdots + a_p$$
$$Q = b_0 y^q + b_1 y^{q-1} + \cdots + b_q\,.$$

For $0 \le i \le \min(p,q) - 1$ we let the ith Sylvester matrix of P, p and Q, q to be $\mathrm{Sylv}_i(P, p, Q, q) = [y^{q-i-1}P, \dots, P, y^{p-i-1}Q, \dots, Q]$. When $p \ne q$ we let the $\delta(p,q)$th Sylvester matrix of P, p and Q, q to be $[y^{q-p-1}P, \dots, P]$ if $p < q$ and $[y^{p-q-1}Q, \dots, Q]$ if $q < p$. The qth Sylvester matrix is not defined when $p = q$.[1]

Definition. Let $P, Q \in \mathbf{A}[y]$ be two polynomials, with $\deg(P) \le p$ and $\deg(Q) \le q$. For any $i \le \delta(p,q)$ the polynomial determinant of the matrix $\mathrm{Sylv}_i(P, p, Q, q)$, denoted by $\mathrm{Sr}_i(P, p, Q, q)$, is called the ith subresultant of P, p and Q, q. The coefficient of degree i of the polynomial $\mathrm{Sr}_i(P, p, Q, q)$, denoted by $\mathrm{sr}_i(P, p, Q, q)$, is called the ith principal subresultant coefficient of P, p and Q, q.

When $\deg(P) = p$ and $\deg(Q) = q$ we write $\mathrm{Sylv}_i(P, Q)$, $\mathrm{Sr}_i(P, Q)$ and $\mathrm{sr}_i(P, Q)$ for short instead of $\mathrm{Sylv}_i(P, p, Q, q)$, $\mathrm{Sr}_i(P, p, Q, q)$ and $\mathrm{sr}_i(P, p, Q, q)$. The polynomial $\mathrm{Sr}_i(P, p, Q, q)$ is of degree at most i, in particular $\mathrm{Sr}_0(P, p, Q, q)$ is constant and is nothing but the resultant of P and Q provided that $\deg(P) = p$ and $\deg(Q) = q$. Let us note on the other hand that the matrix $\mathrm{Sylv}_i(Q, q, P, p)$ is obtained from $\mathrm{Sylv}_i(P, p, Q, q)$ by row exchanges in such a way that

$$\mathrm{Sr}_i(Q, q, P, p) = (-1)^{(p-i)(q-i)}\mathrm{Sr}_i(P, p, Q, q). \qquad (2)$$

Following this fact one can assume without loss of generality that $q \le p$.

3.1. Specialization of subresultants

In this subsection we give a fundamental result concerning specialization of subresultants. A detailed study of the question, together with a proof of the result we give here, can be found in González-Vega et al. (1990, 1994).

[1] When $p = q$ and the ring $\mathbf{A}$ is integral then $\mathrm{Sylv}_q(P, p, Q, q)$ is actually defined as $[b_0^{-1}Q]$. In our case we do not define it for this case because we do not assume $\mathbf{A}$ to be integral.

Lemma 3.1. *Let $p \geq q \geq 0$ and $P, Q \in \mathbf{A}[y]$ be two polynomials with $\deg(P) \leq p$ and $\deg(Q) \leq q$. Then:*

(i) *if* $\deg(P) < p$ *and* $\deg(Q) < q$ *then*

$$\mathrm{Sr}_i(P, p, Q, q) = 0 \qquad \textit{for } i = 0, \ldots, \delta(p, q),$$

(ii) *if* $\deg(P) = p$ *and* $\deg(Q) \leq s < q$ *then*

$$\mathrm{Sr}_i(P, p, Q, q) = \begin{cases} a_0^{q-s}\mathrm{Sr}_i(P, p, Q, s) & \textit{for } i = 0, \ldots, s \\ 0 & \textit{for } i = s + 1, \ldots, \delta(p, q), \end{cases}$$

(iii) *if* $\deg(P) \leq s < p$, $\deg(Q) = q$ *and* $s \geq q$ *then*

$$\mathrm{Sr}_i(P, p, Q, q) = (-1)^{\mu_i} b_0^{p-s}\mathrm{Sr}_i(P, s, Q, q) \qquad \textit{for } i = 0, \ldots, \delta(s, q)$$

where $\mu_i = (q - i)(p - s)$,

(iv) *if* $\deg(P) \leq s < p$, $\deg(Q) = q$ *and* $s < q$ *then*

$$\mathrm{Sr}_i(P, p, Q, q) = \begin{cases} (-1)^{\mu_i} b_0^{p-s}\mathrm{Sr}_i(P, s, Q, q) & \textit{for } i = 0, \ldots, s \\ 0 & \textit{for } i = s + 1, \ldots, \delta(p, q). \end{cases}$$

3.2. Bézout identities

Let $p \geq q \geq 0$ and $P, Q \in \mathbf{A}[y]$ be two polynomials with $\deg(P) \leq p$ and $\deg(Q) \leq q$. Let $0 \leq i \leq \delta(p, q)$ and let M be the matrix whose $p + q - 2i - 1$ first columns are the $p + q - 2i - 1$ first columns of $\mathrm{Sylv}_i(P, p, Q, q)$ and the coefficients of the last column are $y^{q-i-1}P, \ldots, P, y^{p-i-1}Q, \ldots, Q$. Following Lemma 2.1 one has $\mathrm{Sr}_i(P, p, Q, q) = \mathrm{Det}(M)$. Moreover, by expanding $\mathrm{Det}(M)$ with respect to the last column of M one gets

$$\mathrm{Sr}_i(P, p, Q, q) = \mathrm{U}_i(P, p, Q, q)P + \mathrm{V}_i(P, p, Q, q)Q \tag{3}$$

where $\deg(\mathrm{U}_i(P, p, Q, q)) \leq q - i - 1$ and $\deg(\mathrm{V}_i(P, p, Q, q)) \leq p - i - 1$ (see Habicht, 1948). This last identity is called the ith *Bézout identity* of P, p and Q, q. Let us note that the coefficients of U_i and V_i are, up to signs, minors of order $p + q - 2i - 1$ extracted on $\mathrm{Sylv}_i(P, p, Q, q)$, and so they belong to the ring generated by the coefficients of P and Q.

Proposition 3.1. *If one at least of the coefficients of* $\mathrm{Sr}_i(P, Q)$ *is regular, i.e. not a zero divisor in* $\mathbf{A}$ *then the polynomials* $\mathrm{U}_i(P, p, Q, q)$ *and* $\mathrm{V}_i(P, p, Q, q)$ *are uniquely determined over* $\mathbf{A}$ *by the conditions* $\deg(\mathrm{U}_i(P, p, Q, q)) \leq q - i - 1$, $\deg(\mathrm{V}_i(P, p, Q, q)) \leq p - i - 1$ *and the relation* (3)*.*

4. The chain rule of subresultants

In this section we will be concerned with some algebraic identities fulfilled by the subresultant sequence. For this reason we assume that the coefficients of P and Q are indeterminates and we let $\mathbb{Z}[\mathbf{a}, \mathbf{b}]$ be the ring generated by these coefficients. If the formal degrees p and q of P and Q are understood from the context then we denote by Sr_i the ith subresultant of P, p and Q, q and write $\mathrm{Sr}_i = \mathrm{sr}_i y^i + \mathrm{sr}_{i,i-1} y^{i-1} + \cdots + \mathrm{sr}_{i,0}$. Let us first recall the generic chain rule of subresultants, usually known as the Habicht theorem (Habicht, 1948; Loos, 1982; Collins, 1967; Brown and Traub, 1971; Ho and Yap, 1996).

Theorem 4.1. *Let $p \geq q > 0$ and $P, Q \in \mathbb{Z}[\mathbf{a}, \mathbf{b}][y]$ be two polynomials.*

(i) *If* $\deg(P) \leq q + 1$ *and* $\deg(Q) \leq q$ *then for any* $j < i \leq q - 1$ *one has*

$$\mathrm{sr}_{i+1}^{2(i-j)} \mathrm{Sr}_j = \mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i + 1, \mathrm{Sr}_i, i),$$

(ii) *if* $\deg(P) = p$ *and* $\deg(Q) = q$ *then for any* $j < i \leq \delta(p, q) - 1$ *one has*

$$\mathrm{sr}_{i+1}^{2(i-j)} \mathrm{Sr}_j = \mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i + 1, \mathrm{Sr}_i, i).$$

The original proof of this result is due to Habicht and consists in using induction on i starting from $q - 1$. The initialization step is achieved by using the fact that Sr_{q-1} is the pseudo-remainder of P by Q. We can hide this pseudo-division by using the Bézout identity corresponding to Sr_{q-1}. The rest of the proof consists in using suitable row operations on Sylvester matrices so that the chain rule of subresultants can be seen as a consequence of the behaviour of polynomial determinants under row operations and the existence of Bézout identities.

4.1. New algebraic identities of subresultants

The generic chain rule of subresultants is not enough to handle in a precise way the cases where some polynomials in the subresultant sequence drop down in degree. In this subsection we give some algebraic identities fulfilled by the subresultant sequence which, to our knowledge, are not known. These algebraic identities, which are interesting in their own right, permit to build a new proof of the *gap structure theorem*. They also allow, as we shall see in Corollary 5.1, to give precisions on gcds computations over integrally closed domains. First we start by proving some irreducibility results concerning the coefficients in the subresultant sequence. For this aim the following elementary lemma will be used.

Lemma 4.1. *Let* $\mathbf{A}$ *be a UFD and* $S \in \mathbf{A}[x_1, \ldots, x_r, y_1, \ldots, y_s]$ *be a homogeneous polynomial such that* $S(x_1, \ldots, x_r, 0, \ldots, 0)$ *is irreducible over* $\mathbf{A}$. *Then* S *is also irreducible over* $\mathbf{A}$.

We can now state the irreducibility results concerning the principal subresultant coefficients.

Lemma 4.2. *Let $p \geq q > 0$ and $P, Q \in \mathbb{Z}[\mathbf{a}, \mathbf{b}][y]$ be two polynomials with* $\deg(P) = p$ *and* $\deg(Q) = q$. *Then the principal subresultants coefficients* sr_i *are irreducible, and pair-wise distinct in* $\mathbb{Z}[\mathbf{a}, \mathbf{b}]$. *In particular, for any* $i \leq \delta(p, q) - 1$ *the polynomial* Sr_i *is primitive over* $\mathbb{Z}[\mathbf{a}, \mathbf{b}]$.

Proof. Let us note that the $\mathrm{sr}_{i,j}$'s are homogeneous polynomials in terms of the a_k's and b_l's. Also, it is a classical fact that sr_0 is irreducible over $\mathbb{Z}$.

Next we prove by induction on q that the sr_i's are irreducible and pair-wise distinct in $\mathbb{Z}[\mathbf{a}, \mathbf{b}]$. For $q = 1$ and any $p \geq 1$ the coefficients in question are $\mathrm{sr}_1 = b_0$ and sr_0, and they are irreducible over $\mathbb{Z}$ and distinct. Thus the result is true for $q = 1$ and any $p \geq q$. Assume that the result is still true for a given $q > 1$ and any $p \geq q$, and let us prove it for $q + 1$ and any $p \geq q + 1$. For this let P and Q be polynomials of degrees respectively p and $q + 1$ such that $p \geq q + 1$. Then using Taylor expansions one can write each sr_i, $i \geq 1$, in the form $\mathrm{sr}_i = \mathrm{sr}_i(P_1, Q_1) + a_p C_i + b_{q+1} D_i$ with $P_1 = a_0 y^p + \cdots + a_{p-1} y = y P_2$ and $Q_1 = b_0 y^{q+1} + \cdots + a_q y = y Q_2$. A direct computation (as we shall see in Proposition 5.2) shows that $\mathrm{sr}_i(P_1, Q_1) = \mathrm{sr}_{i-1}(P_2, Q_2)$ for any $i \geq 1$. Thus $\mathrm{sr}_{i-1}(P_2, Q_2)$ is the constant term of sr_i with respect to a_p and b_{q+1}. According to the induction hypothesis the $\mathrm{sr}_{i-1}(P_2, Q_2)$'s are irreducible over $\mathbb{Z}$ and pair-wise distinct. Since the sr_i's are homogeneous they are irreducible and pair-wise distinct over $\mathbb{Z}$ for $i \geq 1$. The fact that $\mathrm{sr}_0 = 0$ for $a_p = b_{q+1} = 0$ while the other sr_i's do not implies that there is no divisibility relation between sr_0 and the other coefficients.

Let us now prove that Sr_i is primitive. Let c be its content and assume that $i \geq 1$ (the case $i = 0$ is obvious). Then c divides sr_i, and according to the relation $\mathrm{sr}_{i+1}^2 \mathrm{Sr}_{i-1} = \mathrm{Sr}_{i-1}(\mathrm{Sr}_{i+1}, i + 1, \mathrm{Sr}_i, i)$ given by Theorem 4.1 c divides $\mathrm{sr}_{i+1}^2 \mathrm{sr}_{i-1}$. Therefore c is a unit in $\mathbb{Z}$ since sr_i and $\mathrm{sr}_{i+1}^2 \mathrm{sr}_{i-1}$ are co-primes. □

Theorem 4.2. *Let $p \geq q > 0$ and $P, Q \in \mathbb{Z}[\mathbf{a}, \mathbf{b}][y]$ be two polynomials with $\deg(P) = p$ and $\deg(Q) = q$. Then:*

(i) *for any $0 \leq i \leq q - 2$*

$$\mathrm{sr}_{i+1}^2 P = U_i \mathrm{Sr}_{i+1} + V_i \mathrm{Sr}_i$$
$$\mathrm{sr}_{i+1}^2 Q = U_i' \mathrm{Sr}_{i+1} + V_i' \mathrm{Sr}_i$$

with $\deg(U_i) \leq p - i - 1$, $\deg(V_i) \leq p - i - 2$, $\deg(U_i') \leq q - i - 1$ and $\deg(V_i') \leq q - i - 2$,

(ii) *for any $0 \leq i \leq \delta(p, q) - 1$ and $j < i$*

$$\mathrm{sr}_{i+1}^2 \mathrm{Sr}_j = U_{i,j} \mathrm{Sr}_{i+1} + V_{i,j} \mathrm{Sr}_i$$

with $\deg(U_{i,j}) \leq i - j - 1$ and $\deg(V_{i,j}) \leq i - j$,

(iii) *for any $0 \leq i \leq \delta(p, q) - 1$ and $j < i$ the identity*

$$\mathrm{sr}_{i+1}^{2(i-j-1)} U_{i,j} \mathrm{Sr}_{i+1} + \mathrm{sr}_{i+1}^{2(i-j-1)} V_{i,j} \mathrm{Sr}_i$$

is the Bézout identity corresponding to $\mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i + 1, \mathrm{Sr}_i, i)$.

Proof. (i) For these identities we use a descending induction on i. To start induction we shall distinguish two cases.

- If $p > q$ then one has $\mathrm{sr}_q^2 \mathrm{Sr}_{q-2} = \mathrm{Sr}_{q-2}(\mathrm{Sr}_q, q, \mathrm{Sr}_{q-1}, q-1) = \mathrm{sr}_{q-1}^2 \mathrm{Sr}_q + V_1 \mathrm{Sr}_{q-1}$, with $\deg(V_1) \leq 1$, according to Theorem 4.1 and Bézout identities. On the other hand, the $(q - 1)$th Bézout identity of P and Q gives $b_0^{p-q+1} P = UQ + \mathrm{Sr}_{q-1}$ with $\deg(U) \leq p - q$. Multiplying this last relation by b_0^{p-q-1} and using the relation $\mathrm{Sr}_q = b_0^{p-q-1}Q$ one gets $\mathrm{sr}_q^2 P = U'\mathrm{Sr}_q + V'\mathrm{Sr}_{q-1}$, and multiplying this last equality by sr_{q-1}^2 and replacing $\mathrm{sr}_{q-1}^2\mathrm{Sr}_q$ by $\mathrm{sr}_q^2\mathrm{Sr}_{q-2} - V_1\mathrm{Sr}_{q-1}$ one gets $\mathrm{sr}_q^2\mathrm{sr}_{q-1}^2 P = U'\mathrm{sr}_q^2\mathrm{Sr}_{q-2} + (V'\mathrm{sr}_{q-1}^2 - U'V_1)\mathrm{Sr}_{q-1}$. This last relation shows that sr_q^2 divides the content of $(V'\mathrm{sr}_{q-1}^2 - U'V_1)\mathrm{Sr}_{q-1}$. Since Sr_{q-1} is primitive sr_q^2 divides the content of $V'\mathrm{sr}_{q-1}^2 - U'V_1$, and simplifying by it one gets a relation $\mathrm{sr}_{q-1}^2 P = U_{q-2}\mathrm{Sr}_{q-1} + V_{q-2}\mathrm{Sr}_{q-2}$, with $\deg(U_{q-2}) \leq p-q+1$ and $\deg(V_{q-2}) \leq p-q$.

For the polynomial Q we apply Theorem 4.1 and Bézout identities to get the relation $\mathrm{sr}_q^2\mathrm{Sr}_{q-2} = \mathrm{Sr}_{q-2}(\mathrm{Sr}_q, q, \mathrm{Sr}_{q-1}, q-1) = \mathrm{sr}_{q-1}^2\mathrm{Sr}_q + V_1\mathrm{Sr}_{q-1}$ with $\deg(V_1) \leq 1$. Since $\mathrm{Sr}_q = b_0^{p-q-1}Q$, $\mathrm{sr}_q = b_0^{p-q}$ and Sr_{q-1} is primitive the term b_0^{p-q-1} divides the content of V_1. Simplifying by it we get the desired relation.

- If $p = q$ then let us simplify in two different ways $\mathrm{Sr}_{q-2}(\mathrm{Sr}_{q-1}, q, Q, q)$. On the first hand, according to the relation $\mathrm{Sr}_{q-1} = -b_0P + a_0Q$ and Lemma 2.2 one has the relation $\mathrm{Sr}_{q-2}(\mathrm{Sr}_{q-1}, q, Q, q) = \mathrm{Sr}_{q-2}(-b_0P, q, Q, q) = b_0^2\mathrm{Sr}_{q-2}$. On the other hand, using Lemma 3.1 (iv) and the relation (2) one gets the relation $\mathrm{Sr}_{q-2}(\mathrm{Sr}_{q-1}, q, Q, q) = b_0\mathrm{Sr}_{q-2}(Q, q, \mathrm{Sr}_{q-1}, q-1)$. The Bézout identity of $\mathrm{Sr}_{q-2}(Q, q, \mathrm{Sr}_{q-1}, q-1)$ writes as $\mathrm{Sr}_{q-2}(Q, q, \mathrm{Sr}_{q-1}, q-1) = \mathrm{sr}_{q-1}^2 Q + V_1\mathrm{Sr}_{q-1}$, with $\deg(V_1) \leq 1$. Thus $b_0^2\mathrm{Sr}_{q-2} = b_0(\mathrm{sr}_{q-1}^2 Q + V_1\mathrm{Sr}_{q-1})$, and simplifying by b_0 one finally gets $\mathrm{sr}_{q-1}^2 Q = U'_{q-2}\mathrm{Sr}_{q-1} + V'_{q-2}\mathrm{Sr}_{q-2}$, with $\deg(U'_{q-1}) \leq 1$ and $V'_{q-2} = b_0$. To prove the relation corresponding to P we multiply the last relation by a_0 and replace a_0Q by $\mathrm{Sr}_{q-1} + b_0P$ to get $b_0\mathrm{sr}_{q-1}^2 P = (a_0U'_{q-2} - \mathrm{sr}_{q-1}^2)\mathrm{Sr}_{q-1} + a_0b_0\mathrm{Sr}_{q-2}$. Thus b_0 divides the content of $a_0U'_{q-2} - \mathrm{sr}_{q-1}^2$ since Sr_{q-1} is primitive, and simplifying by b_0 one gets the desired relation.

At the close of the cases studied above we conclude that the identities of (i) are fulfilled for $i = q-2$. Now assume that the same holds for a given $1 \leq i < q-2$ and let us prove it for $i-1$. According to the induction hypothesis one has

$$\mathrm{sr}_{i+1}^2 P = U_i\mathrm{Sr}_{i+1} + V_i\mathrm{Sr}_i \tag{4}$$

with $\deg(U_i) \leq p-i-1$ and $\deg(V_i) \leq p-i-2$. On the other hand, one has $\mathrm{sr}_{i+1}^2\mathrm{Sr}_{i-1} = \mathrm{Sr}_{i-1}(\mathrm{Sr}_{i+1}, i+1, \mathrm{Sr}_i, i) = \mathrm{sr}_i^2\mathrm{Sr}_{i+1} + V_1\mathrm{Sr}_i$, with $\deg(V_1) \leq 1$ according to Theorem 4.1 and Bézout identities. Multiplying (4) by sr_i^2 and replacing $\mathrm{sr}_i^2\mathrm{Sr}_{i+1}$ by $\mathrm{sr}_{i+1}^2\mathrm{Sr}_{i-1} - V_1\mathrm{Sr}_i$ one gets $\mathrm{sr}_{i+1}^2\mathrm{sr}_i^2 P = (V_i - U_iV_1)\mathrm{Sr}_i + U_i\mathrm{sr}_{i+1}^2\mathrm{Sr}_{i-1}$. This last relation shows that sr_{i+1}^2 divides the content of $V_i - U_iV_1$ since Sr_i is primitive. Simplifying by sr_{i+1}^2 one finally gets a relation of the form $\mathrm{sr}_i^2 P = U_{i-1}\mathrm{Sr}_i + V_i\mathrm{Sr}_{i-1}$ with $\deg(U_{i-1}) \leq p-i$ and $\deg(V_{i-1}) \leq p-i-1$. Using similar arguments one obtains the relation corresponding to Q.

(ii) For the proof of the identity we use an ascending induction on i starting from the case $j+1$ which is given by Theorem 4.1. Assume now that for a given $j+1 < i < q-1$ one has $\mathrm{sr}_{i+1}^2\mathrm{Sr}_j = U_{i,j}\mathrm{Sr}_{i+1} + V_{i,j}\mathrm{Sr}_i$ with $\deg(U_{i,j}) \leq i-j-1$ and $\deg(V_{i,j}) \leq i-j$. Theorem 4.1 and Bézout identities give another identity of the form $\mathrm{sr}_{i+2}^2\mathrm{Sr}_i = \mathrm{sr}_{i+1}^2\mathrm{Sr}_{i+2} + V_1\mathrm{Sr}_{i+1}$ with $\deg(V_1) \leq 1$. Combining these two last identities after multiplying the first one by sr_{i+2}^2 and taking into account the fact that the Sr_i's are primitive one gets the desired relation.

(iii) Let us write the Bézout identity corresponding to $\mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i+1, \mathrm{Sr}_i, i)$ in the form $\mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i+1, \mathrm{Sr}_i, i) = U\mathrm{Sr}_{i+1} + V\mathrm{Sr}_i$ with $\deg(U) \le i-j-1$ and $\deg(V) \le i-j$. According to the identity of (ii) and Theorem 4.1 one also has $\mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i+1, \mathrm{Sr}_i, i) = \mathrm{sr}_{i+1}^{2(i-j-1)} U_{i,j}\mathrm{Sr}_{i+1} + \mathrm{sr}_{i+1}^{2(i-j-1)} V_{i,j}\mathrm{Sr}_i$, with $\deg(U_{i,j}) \le i-j-1$ and $\deg(V_{i,j}) \le i-j$. Since $\mathbb{Z}[\mathbf{a}, \mathbf{b}]$ is integral and $\mathrm{Sr}_j(\mathrm{Sr}_{i+1}, i+1, \mathrm{Sr}_i, i) \ne 0$ one has $U = \mathrm{sr}_{i+1}^{2(i-j-1)} U_{i,j}$ and $V = \mathrm{sr}_{i+1}^{2(i-j-1)} V_{i,j}$ according to Lemma 3.1. □

4.2. The gap structure theorem

In this subsection we give a new proof of the gap structure theorem of subresultants (Lickteig and Roy, 1996a), which is a refinement of the subresultant theorem (Habicht, 1948; Collins, 1967; Brown and Traub, 1971; Loos, 1982; González-Vega et al., 1990; Ducos, 2000). Actually this theorem is formulated over an integral ring. Here we give a version where the ring is arbitrary.

Theorem 4.3. *Let* $\mathbf{A}$ *be a commutative ring,* $p \ge q$ *be positive integers and* $P, Q \in \mathbf{A}[y]$ *be polynomials with* $\deg(P) = p$ *and* $\deg(Q) = q$. *Let* $0 \le j \le \delta(p,q)$, *and assume that* sr_j *is regular in* $\mathbf{A}$ *and* $\mathrm{Sr}_{j-1} \ne 0$ *is of degree* $k < j-1$. *Then:*

(i) $\mathrm{Sr}_{j-2} = \cdots = \mathrm{Sr}_{k+1} = 0$,

(ii) $\mathrm{sr}_j^{j-k-1}\mathrm{Sr}_k = \mathrm{sr}_{j-1,k}^{j-k-1}\mathrm{Sr}_{j-1}$,

(iii) $\mathrm{sr}_j^2\mathrm{Sr}_{k-1} = (-1)^{j-k}\mathrm{sr}_{j-1,k}\mathrm{sr}_k\mathrm{Sr}_j + C_j\mathrm{Sr}_{j-1}$, *with* $C_j \in \mathbf{A}[y]$.

Proof. The facts (i) and (ii) follow obviously from Theorem 4.1 and Lemma 3.1(ii). (iii) Let us write the Bézout identity of $\mathrm{Sr}_{k-1}(\mathrm{Sr}_j, j, \mathrm{Sr}_{j-1}, j-1)$ in the form $\mathrm{Sr}_{k-1}(\mathrm{Sr}_j, j, \mathrm{Sr}_{j-1}, j-1) = U\mathrm{Sr}_j + V\mathrm{Sr}_{j-1}$ with $\deg(U) \le j-k-1$ and $\deg(V) \le j-k$. Using Lemma 3.1 one gets $U = (-1)^{j-k}\mathrm{sr}_j^{j-k-1}\mathrm{sr}_{j-1,k}^{j-k+1}$, and using the relation (ii) one gets $U = (-1)^{j-k}\mathrm{sr}_j^{2(j-k-1)}\mathrm{sr}_{j-1,k}\mathrm{sr}_k$. Now using Theorem 4.4 (ii) and (iii) one has $\mathrm{sr}_j^2\mathrm{Sr}_{k-1} = U_j\mathrm{Sr}_j + V_j\mathrm{Sr}_{j-1}$ and $U = \mathrm{sr}_j^{2(j-k-1)}U_j$. Since sr_j is regular in $\mathbf{A}$ we have $U_j = (-1)^{j-k}\mathrm{sr}_{j-1,k}\mathrm{sr}_k$ and this proves the desired relation. □

5. Behaviour with respect to operations on polynomials

In this section we study how subresultants behave with respect to some elementary operations on polynomials. The results of this section are classical and can for instance be found in Chardin (1991), Cohen et al. (1999), Hong (1999) and Cheng (2001). The proofs we give for the results of this section are elementary in so far as only properties of polynomial determinants are used.

It is a classical fact that the resultant is invariant under translation. The following proposition shows that subresultants "commute" with translation.

Proposition 5.1. *Let $p \geq q$ be two positive integers and $P, Q \in \mathbf{A}[y]$ be two polynomials with $\deg(P) \leq p$ and $\deg(Q) \leq q$. Then for any $\alpha \in \mathbf{A}$ one has:*

$$\mathrm{Sr}_i(P(y+\alpha), Q(y+\alpha)) = \mathrm{Sr}_i(P, Q)(y+\alpha) \qquad \textit{for } i = 0, 1, \ldots, \delta(p, q).$$

Proof. For $0 \leq i \leq \delta(p,q)$ one has $\mathrm{Sr}_i(P(y+\alpha), Q(y+\alpha)) = \mathrm{DetPol}(M_i)$ where $M_i = [y^{q-i-1}P(y+\alpha), \ldots, P(y+\alpha), y^{p-i-1}Q(y+\alpha), \ldots, Q(y+\alpha)]$. By using a suitable row operation one can transform the matrix M_i into the matrix $M'_i = [(y+\alpha)^{q-i-1}P(y+\alpha), \ldots, P(y+\alpha), (y+\alpha)^{p-i-1}Q(y+\alpha), \ldots, Q(y+\alpha)]$, so that $\mathrm{Sr}_i(P(y+\alpha), Q(y+\alpha)) = \mathrm{DetPol}(M'_i)$. On the other hand, if we let ϕ be the automorphism of $\mathbf{A}$-algebras defined by $\phi(y) = y + \alpha$ then we have the relation $M'_i = [\phi(y^{q-i-1}P(y)), \ldots, \phi(P(y)), \phi(y^{p-i-1}Q(y)), \ldots, \phi(Q(y))]$, and using Lemma 2.3 we get $\mathrm{DetPol}(M'_i) = \phi(\mathrm{DetPol}(\mathrm{Sylv}_i(P, Q)))$, which gives $\mathrm{Sr}_i(P(y+\alpha), Q(y+\alpha)) = \mathrm{Sr}_i(P, Q)(y+\alpha)$. □

The following proposition concerns the behaviour of the subresultant sequence when the polynomials P and Q have a common factor.

Proposition 5.2. *Let $p \geq q$ be two positive integers and $P, Q, R \in \mathbf{A}[y]$ be polynomials with $\deg(P) \leq p$, $\deg(Q) \leq q$ and $\deg(R) = r$. Then:*

(i) $\mathrm{Sr}_i(PR, p+r, QR, q+r) = 0$ *for* $i = 0, \ldots, r-1$,
(ii) $\mathrm{Sr}_i(PR, p+r, QR, q+r) = a^{\mu_i}\mathrm{Sr}_{i-r}(P, p, Q, q)R$ *for* $i = r, \ldots, \delta(p,q)+r$, *where* $\mu_i = p+q+2r-2i-1$ *and a is the leading coefficient of R.*

Proof. First we prove the result in the case where $R = y^r$. For any $0 \leq i \leq \delta(p,q)+r$ the r last columns of $S_i = \mathrm{Sylv}_i(y^rP, p+r, y^rQ, q+r)$ are zero.

- If $i \leq r-1$ then any coefficient of $\mathrm{Sr}_i(y^rP, p+r, y^rQ, q+r)$ is the determinant of a sub-matrix of S_i involving at least one of its r last columns. Therefore $\mathrm{Sr}_i(y^rP, p+r, y^rQ, q+r) = 0$.
- If $i \geq r$ then by deleting the r last columns of S_i one gets the matrix $\mathrm{Sylv}_{i-r}(P, p, Q, q)$. Therefore $\mathrm{Sr}_i(y^rP, p+r, y^rQ, q+r) = y^r\mathrm{Sr}_{i-r}(P, p, Q, q)$.

Now consider a degree r polynomial R and let ϕ be the homomorphism of $\mathbf{A}[y]$ defined by $\phi(y^i) = y^i$ if $i \leq r-1$ and $\phi(y^i) = y^{i-r}R$ if $i \geq r$. It is clear that ϕ preserves degrees and that $\phi(S_i) = \mathrm{Sylv}_i(PR, p+r, QR, q+r)$. This gives the relation $\mathrm{Sr}_i(PR, p+r, QR, q+r) = a^{p+q+2r-2i-1}\phi(\mathrm{Sr}_i(y^rP, p+r, y^rQ, q+r))$ according to Lemma 2.3. The desired relation is deduced from this last one and the relation corresponding to the case $R = y^r$. □

Corollary 5.1. *Let $\mathbf{A}$ be a commutative ring and $p \geq q$ be two positive integers. Let $P, Q \in \mathbf{A}[y]$ be two polynomials with $\deg(P) = p$ and $\deg(Q) = q$.*

(i) *If the polynomials P and Q have a common divisor of degree k then*

$$\mathrm{sr}_i = 0, \qquad i = 0, 1, \ldots, k-1. \tag{5}$$

(ii) *If $\mathbf{A}$ is integral, the sr_i's satisfy Eq. (5) and $\mathrm{sr}_k \neq 0$ then Sr_k is a* gcd *of P and Q over the fractions field of* $\mathbf{A}$.

(iii) *If* $\mathbf{A}$ *is a UFD and* P *or* Q *is primitive, the* sr_i*'s satisfy Eq. (5) and* $\mathrm{sr}_k \neq 0$ *then the primitive part of* Sr_k *is a* gcd *of* P *and* Q *over* $\mathbf{A}$.

(iv) *If* $\mathbf{A}$ *is integrally closed and* P *or* Q *is monic, the* sr_i*'s satisfy Eq. (5) and* $\mathrm{sr}_k \neq 0$ *then* $\mathrm{sr}_k^{-1}\mathrm{Sr}_k \in \mathbf{A}[y]$ *and is a* gcd *of* P *and* Q *over* $\mathbf{A}$.

Proof. (i) and (ii) are direct consequences of Proposition 5.2.

(iii) Assume that $\mathbf{A}$ is a UFD and for example that P is primitive. Then the gcd of P and Q over $\mathbf{A}$ is primitive. Since the primitive part S' of Sr_k is a gcd of P and Q over $\mathbf{K}$, it is also a gcd of P and Q over $\mathbf{A}$.

(iv) Assume that $\mathbf{A}$ is integrally closed and for example that P is monic. We have $P = \mathrm{sr}_k^{-1}U_k\mathrm{sr}_k^{-1}\mathrm{Sr}_k$ with $\mathrm{sr}_k^{-1}U_k$ and $\mathrm{sr}_k^{-1}\mathrm{Sr}_k$ monic with coefficients in $\mathbf{K}$. Since $\mathbf{A}$ is integrally closed and P has its coefficients in $\mathbf{A}$ both of $\mathrm{sr}_k^{-1}U_k$ and $\mathrm{sr}_k^{-1}\mathrm{Sr}_k$ have their coefficients in $\mathbf{A}$. On the other hand, one has $Q = \mathrm{sr}_k^{-1}U_k'\mathrm{sr}_k^{-1}\mathrm{Sr}_k$ and Q and $\mathrm{sr}_k^{-1}\mathrm{Sr}_k$ have their coefficients in $\mathbf{A}$. Thus $\mathrm{sr}_k^{-1}U_k' \in \mathbf{A}[y]$. □

References

Brown, W.S., Traub, J.F., 1971. On Euclid's algorithm and the theory of subresultants. J. ACM 18 (4), 505–514.

Chardin, M., 1991. Differential resultants and resultants, Lecture Notes in Computer Science, vol. 529, pp. 180–189.

Cheng, C.C., 2001. A chain rule for subresultants. J. Pure Appl. Algebra 157 (1), 33–39.

Cohen, A.M., Cuypers, H., Sterk, H., 1999. Some Tapas of Computer Algebra, Algorithms and Computations in Mathematics, vol. 4, Springer-Verlag, Heidelberg.

Collins, G.E., 1967. Subresultants and reduced polynomial remainder sequences. J. ACM 14, 128–142.

Ducos, L., 2000. Optimization of the subresultant algorithm. J. Pure Appl. Abgebra 145, 149–163.

von zur Gathen, J., Gerhard, J., 1999. Modern Computer Algebra, Cambridge University Press, Cambridge.

von zur Gathen, J., Luking, T., 2000. Subresultants revisited. Proceedings of Latin 2000: Theoretical Informatics, Punta del Este, Uruguay, Springer Lecture Notes in Computer Science, vol. 1776. Heidelberg, pp. 318–342.

González-Vega, L., Lombardi, H., Recio, T., Roy, M.-F., 1990. Spécialisation de la suite de Sturm et sous-résultants. RAIRO Inform. Théor. Appl. 24 (6), 561–588.

González-Vega, L., Lombardi, H., Recio, T., Roy, M.-F., 1994. Spécialisation de la suite de Sturm. RAIRO Inform. Théor. Appl. 28 (1), 1–24.

Habicht, W., 1948. Eine verallgemeinerung des sturmschen wurzelzahlverfahrens. Commun. Math. Halvetici 21, 99–116.

Ho, C.J., Yap, C.K., 1996. The Habicht approach to subresultants. J. Symb. Comput. 21 (1), 1–14.

Hong, H., 1997. Subresultants under composition. J. Symb. Comput. 23 (4), 355–365.

Hong, H., 1999. Subresultants in Roots (preprint).

Lickteig, T., Roy, M.-F., 1996a. Cauchy index computation. Calcolo 33, 337–351.

Lickteig, T., Roy, M.-F., 1996b. Semi-algebraic complexity of quotients and sign determination of remainders. Special issue for the Foundations of Computational Mathematics Conference, Rio de Janeiro, 1997. J. Complexity 12 (4), 545–571.

Lickteig, T., Roy, M.-F., 2001. Sylvester–Habicht sequences and fast Cauchy index computation. J. Symb. Comput. 31 (3), 315–341.

Lombardi, H., Roy, M.-F., Safey El Din, M., 2000. New structure theorem for subresultants. J. Symb. Comput. 29, 663–690.

Loos, R., 1982. Generalized polynomial remainder sequences. In: Computer Algebra Symbolic and Algebraic Computation, Springer-Verlag, pp. 115–138.

Reischert, D., 1997. Asymptotically fast computation of subresultants. In: Proceedings of ISSAC' 97 (Kihei, Hi), ACM, New York, pp. 233–240.

Sturm, C., 1835. Mémoire sur la résolution des équations numériques. Inst. France Sc. Math. Phys. Vol. VI, pp. 273–318.

Tarski, A., 1951. A Decision Method for Elementary Algebra and Geometry, second edn, University of California Press, Berkeley.