Improved Polynomial Remainder Sequences for Ore Polynomials

Maximilian Jaroschek[1]

Research Institute for Symbolic Computation
Johannes Kepler University
A4040 Linz, Austria

Abstract

Polynomial remainder sequences contain the intermediate results of the Euclidean algorithm when applied to (non-)commutative polynomials. The running time of the algorithm is dependent on the size of the coefficients of the remainders. Different ways have been studied to make these as small as possible. The subresultant sequence of two polynomials is a polynomial remainder sequence in which the size of the coefficients is optimal in the generic case, but when taking the input from applications, the coefficients are often larger than necessary. We generalize two improvements of the subresultant sequence to Ore polynomials and derive a new bound for the minimal coefficient size. Our approach also yields a new proof for the results in the commutative case, providing a new point of view on the origin of the extraneous factors of the coefficients.

Keywords: Ore polynomials, greatest common right divisor, polynomial remainder sequences, subresultants
11A05, 68W30

1. Introduction

When given a system of differential equations, one might be interested in finding the common solutions of these equations. In order to do so, one can compute another differential equation whose solution space is the intersection of the solution spaces of the equations in the original system. One way to do this is to translate the equations into operators and use the Euclidean algorithm to compute their greatest common right divisor. The solution space of the greatest common right divisor then consists of the desired elements.

Similarly, given a sequence of numbers $(t_n)_{n\in\{0,1,\dots\}}$ that satisfies two different recurrence equations, the Euclidean algorithm is used in applications to find a reasonable candidate for the least order equation of which $(t_n)_{n\in\{0,1,\dots\}}$ is a solution.

Carrying out Euclid's algorithm applied to two polynomials over a domain $\mathbb{D}$ usually requires a prediction of the denominators that might appear in the coefficients of the remainders in order to bypass costly computations in the quotient field of $\mathbb{D}$. While such a prediction can be done easily, the growth of the coefficients of the remainders can be tremendous, which might result in an unnecessary high running time. This can be avoided by dividing out possible content

Email address: mjarosch@risc.jku.at (Maximilian Jaroschek)
[1]Supported by the Austrian Science Fund (FWF) grant Y464-N18

of the remainders to make their coefficients as small as possible. For commutative polynomials as well as for non-commutative operators, different ways have been extensively studied to find factors of the content in the sequence of remainders without computing the GCD of the coefficients of each element of the sequence. Most notably in this respect are subresultant sequences, where the growth of the coefficients can be reduced from exponential to linear in the number of reduction steps in the Euclidean algorithm. When taking generic, randomly generated input, the coefficient size in the subresultant sequence is usually optimal, but when taking the input from applications in e.g. combinatorics or physics, the remainders still have non-trivial content in many cases.

For commutative polynomials, some ways are known to improve on subresultants. In this article we generalize two of these results to Ore polynomials and we also give a new proof for the commutative case that is based on the structure of subresultants as matrix determinants. Furthermore, we use these results to derive a new bound for the coefficient size of the content-free remainders.

In Section 2 the basic notions of Ore polynomial rings are stated. A precise definition and examples of polynomial remainder sequences are given in Section 3 and further details on the subresultant sequence are then presented in Section 4. The main results of this article can be found in Sections 5 and 6, where we first describe how additional content in the subresultant sequence can emerge and then use these results to improve on the Euclidean algorithm and to get a new bound for the size of the coefficients.

2. Preliminaries

The algebraic framework for different kinds of operators that we consider here are Ore polynomial rings, which were introduced by Øystein Ore in the 1930's. We provide an overview of some basic facts that suffice our needs and that can be found in Ore (1933) and Bronstein and Petkovšek (1996).

Definition 1. Let $\mathbb{D}$ be a commutative domain, $\mathbb{D}[x]$ the set of univariate polynomials over $\mathbb{D}$ and let $\sigma\colon \mathbb{D} \to \mathbb{D}$ be an injective endomorphism.

1. A map $\delta\colon \mathbb{D} \to \mathbb{D}$ is called pseudo-derivation w.r.t. σ, if for any $a, b \in \mathbb{D}$

$$\delta(a+b) = \delta(a) + \delta(b) \quad \text{and} \quad \delta(ab) = \sigma(a)\delta(b) + \delta(a)b.$$

2. Suppose that δ is a pseudo-derivation w.r.t. σ. We define the Ore polynomial ring $(\mathbb{D}[x], +, \cdot)$ with componentwise addition and the unique distributive and associative extension of the multiplication rule

$$xa = \sigma(a)x + \delta(a) \qquad \text{for any } a \in \mathbb{D},$$

to arbitrary polynomials in $\mathbb{D}[x]$. To clearly distinguish this ring from the standard polynomial ring over $\mathbb{D}$, we denote it by $\mathbb{D}[x; \sigma, \delta]$.

Elements of an Ore polynomial ring are called operators and are denoted by capital letters. We refer to the leading coefficient of an operator A as $\operatorname{lc}(A)$, to the coefficient of x^0 in A as $\operatorname{tc}(A)$ and to the polynomial degree of A in x as the order d_A of A.

Example 1. Commonly used Ore polynomial rings are:

1. $\mathbb{D}[x] = \mathbb{D}[x; 1, 0]$, the ring of commutative polynomials over $\mathbb{D}$.

2. $\mathbb{C}(y)[D; 1, \frac{d}{dy}]$, the ring of linear ordinary differential operators.
3. If $s_n \colon \mathbb{C}(n) \to \mathbb{C}(n)$ is the forward shift in n, i.e. $s_n(a(n)) = a(n+1)$, then $\mathbb{C}(n)[S; s_n, 0]$ is the ring of linear ordinary recurrence operators.
4. If $\sigma \colon \mathbb{C}(q)(y) \to \mathbb{C}(q)(y)$ is the q-shift in y, i.e. $\sigma(a(y)) = a(qy)$, then $\mathbb{C}(q)(y)[J; \sigma, \frac{d}{dy}]$ is the ring of Jackson's q-derivative operators.

In this article, we consider the following situation: Let $\mathbb{D}$ be a Euclidean domain with degree function deg and let $\mathbb{D}[x; \sigma, \delta]$ be an Ore polynomial ring where σ is an automorphism. For any operator $A \in \mathbb{D}[x; \sigma, \delta]$, we define $\|A\|$ to be the maximal coefficient degree of A. The content $\operatorname{cont}(A)$ of A is the greatest common divisor of all the coefficients of A and it is defined to be $\operatorname{lc}(A)$ if $\mathbb{D}$ is a field. It is possible to extend $\mathbb{D}[x; \sigma, \delta]$ to an Ore polynomial ring over the quotient field $\mathbb{K}$ of $\mathbb{D}$ by setting $\sigma(a^{-1}) = \sigma(a)^{-1}$ and $\delta(a/b) = (b\delta(a) - a\delta(b))/(b\sigma(b))$ for $a, b \in \mathbb{D}$, $b \neq 0$ (see Li (1996), Proposition 2.2.1). We will denote this ring by $\mathbb{K}[x; \sigma, \delta]$ without making it explicit that the automorphism and the pseudo-derivation are extensions of the functions used in $\mathbb{D}[x; \sigma, \delta]$. It is well known that for any two operators $A, B \in \mathbb{K}[x; \sigma, \delta]$, there exists a greatest common right divisor (GCRD) and it can be made unique (up to units in $\mathbb{D}$) by setting $\operatorname{gcrd}(A, B)$ to a nonzero $\mathbb{K}$- left multiple of any GCRD of A and B that has coefficients in $\mathbb{D}$ but does not have any content in $\mathbb{D}$.

Throughout this article, we let $A, B, G \in \mathbb{D}[x; \sigma, \delta]$, $B \neq 0$ be such that $d_A \geq d_B$ and G is the GCRD of A and B.

Definition 2. For $a \in \mathbb{D}$ and $n \in \mathbb{N}$, $\sigma^n(a)$ is obtained by applying n times σ to a and $\sigma^{-n}(a) := (\sigma^{-1})^n(a)$, where σ^{-1} is the inverse map of σ. The nth σ-factorial of $a \in \mathbb{D}$ is defined as the product

$$a^{[n]} := \prod_{i=0}^{n-1} \sigma^i(a).$$

3. Polynomial Remainder Sequences for Ore Polynomials

The greatest common right divisor of A and B can be computed by using the Euclidean algorithm. If we multiply any intermediate result that appears during the execution of the algorithm by an element of $\mathbb{K} \setminus \{0\}$, the final output will be a $\mathbb{K}$-left multiple of G. This amount of freedom allows us to optimize the running time by choosing these factors appropriately. In order to be able to formulate improvements of this kind, the notion of polynomial remainder sequences has been introduced. Each element of such a sequence corresponds to a remainder computed in one iteration of the Euclidean algorithm.

Definition 3. Let $(R_i)_{i \in \{0,\ldots,\ell+1\}}$ and $(Q_i)_{i \in \{1,\ldots,\ell\}}$ be sequences in $\mathbb{K}[x; \sigma, \delta]$, $(d_i)_{i \in \{0,\ldots,\ell\}}$ a sequence in $\mathbb{N}$ and let $(\alpha_i)_{i \in \{1,\ldots,\ell\}}$ and $(\beta_i)_{i \in \{1,\ldots,\ell\}}$ be sequences in $\mathbb{K}$ such that

$$\begin{aligned} &R_0 = A, \quad R_1 = B, \quad d_i = d_{R_i}, \\ &\alpha_i R_{i-1} = Q_i R_i + \beta_i R_{i+1}, \quad d_{i+1} < d_i, \end{aligned}$$

and all R_i are nonzero except for $R_{\ell+1}$. We call the sequence $(R_i)_{i \in \{0,\ldots,\ell+1\}}$ a polynomial remainder sequence (PRS) of A and B.

A PRS of A and B is uniquely determined by specifying the α_i and β_i. Whenever we talk about a PRS $(R_i)_{i\in\{0,\dots,\ell+1\}}$, we allow ourselves to refer to the related sequences $(Q_i)_{i\in\{1,\dots,\ell\}}$, $(d_i)_{i\in\{0,\dots,\ell\}}$ etc. as in the above definition without explicitly introducing them.

In order to efficiently compute G, one wants to make sure that all the remainders are elements of $\mathbb{D}[x;\sigma,\delta]$ rather than $\mathbb{K}[x;\sigma,\delta]$. This can be achieved by choosing the α_i in a way such that the quotient of any two consecutive remainders has coefficients in $\mathbb{D}$. To this extent, for $1 \le i \le \ell$ set $\alpha_i := \operatorname{lc}(R_i)^{[d_{i-1}-d_i+1]}$ and division with remainder yields Q_i and R_{i+1} in $\mathbb{D}[x;\sigma,\delta]$ with:

$$\alpha_i R_{i-1} = Q_i R_i + R_{i+1}, \qquad d_{i+1} < d_i. \tag{1}$$

We call $\operatorname{pquo}(R_{i-1}, R_i) := Q_i$ the pseudo-quotient of R_{i-1} and R_i and $\operatorname{prem}(R_{i-1}, R_i) := R_{i+1}$ the pseudo-remainder of R_{i-1} and R_i.

The α_i are used to make sure that computations can be done in $\mathbb{D}[x;\sigma,\delta]$ and the β_i control the coefficient growth in a PRS. We want β_i to contain as many factors of the content of R_{i+1} as possible without much computational overhead needed to obtain these factors.

Example 2. Set $\alpha_i = \operatorname{lc}(R_i)^{[d_{i-1}-d_i+1]}$ and

1. $\beta_i = 1$. This is called the pseudo PRS of A and B. Here, no content will be divided out.
2. $\beta_i = \operatorname{cont}(R_{i+1})$. This is called the primitive PRS of A and B. The coefficients of the remainders will be as small as possible, but it is necessary to compute the GCD of the coefficients of each remainder in order to get the β_i.
3. The subresultant PRS of A and B (see Section 4) is given by

$$\beta_i = \begin{cases} -\sigma(\psi_1)^{[d_0-d_1]}, & \text{if } i = 1, \\ -\operatorname{lc}(R_{i-1})\sigma(\psi_i)^{[d_{i-1}-d_i]}, & \text{if } 2 \le i \le \ell, \end{cases}$$

where

$$\psi_i = \begin{cases} -1, & \text{if } i = 1, \\ \dfrac{(-\operatorname{lc}(R_{i-1}))^{[d_{i-2}-d_{i-1}]}}{\sigma(\psi_{i-1})^{[d_{i-2}-d_{i-1}-1]}}, & \text{if } 2 \le i \le \ell. \end{cases}$$

In this PRS, the content that is generated systematically by pseudo-remaindering will be cleared from the remainders.

While in all of the above PRSs the remainders are elements of $\mathbb{D}[x;\sigma,\delta]$, the degrees of the coefficients differ drastically, as illustrated in the following example. It can be shown that the degrees of the coefficients in the pseudo PRS grow exponentially with i, which renders this PRS practically useless. The growth in the subresultant and primitive PRS is linear in i.

Example 3. Assume we are given a finite sequence of rational numbers that comes from a sequence $(t_n)_{n\in\{0,1,\dots\}}$ which admits a linear recurrence equation with polynomial coefficients. If the amount of data is sufficiently large, we are able to guess recurrence operators of some fixed order and maximal coefficient degree that annihilate $(t_n)_{n\in\{0,1,\dots\}}$, i.e. the operators applied to the sequence give zero. (For details on guessing and a Mathematica implementation of the method, see Kauers (2009).) For example, consider

$$t_n = \sum_{k=0}^{n} \binom{2n+4}{k} + (2n-k)! + k^3.$$

Given the first 300 terms of this sequence, we can find two operators A and B in $\mathbb{Q}[n][S; s_n, 0]$ with $d_A = 14$, $d_B = 13$ and maximal coefficient degree $\|A\| = 5$, $\|B\| = 6$ resp. Both operators annihilate the given sequence, but none of them is of minimal order. To get an annihilating minimal order operator, we compute the GCRD of A and B in $\mathbb{Q}(n)[S; s_n, 0]$. Table 1 shows the maximal coefficient degrees of the remainders for different PRSs of A and B.

PRS	R_2	R_3	R_4	R_5	R_6	R_7	R_8
pseudo	11	22	49	114	271	650	1565
subresultant	11	16	21	26	31	36	41
primitive	9	12	15	18	21	24	21

Table 1: Maximal coefficient degrees for different PRSs.

The example confirms that the degrees in the pseudo PRS grow exponentially, whereas the subresultant PRS and the primitive PRS show linear growth. At the same time, the degrees in the subresultant PRS are not as small as possible. This behavior is typical not only for this pair A and B, but in general for operators coming from applications. For randomly generated operators, the subresultant PRS and the primitive PRS usually coincide. Our goal is to understand the difference between randomly generated input and the operators A and B as above and to identify the source of some (and most often all) of the additional content in the subresultant PRS. To make use of this knowledge, we will then adjust the formulas for α_i and β_i from Example 2.3 so that we get a PRS with smaller degrees without having to compute the content of every remainder.

4. Subresultant Theory for Ore Polynomials

For commutative polynomials, the theory of subresultants was intensively studied by Brown (1978), Brown and Traub (1971), Collins (1967) and Loos (1982). The main idea is to translate relations between the elements of a PRS like the Bézout relation or the (pseudo-)remainder formula into linear algebra. A central tool in this context is the Sylvester matrix, which, roughly speaking, contains the coefficients of all the monomial multiples of the input polynomials that are necessary to compute remainders of any possible degree. The remainders in the subresultant sequence turn out to be polynomials whose coefficients are determinants of certain submatrices of this matrix. Li (1998) generalized these results to Ore polynomials.

$$\begin{pmatrix}
\mathrm{lc}(x^{d_B-1}A) & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \mathrm{tc}(x^{d_B-1}A) \\
& \ddots & & & & & & \vdots \\
& & \mathrm{lc}(A) & \cdots & \cdots & \cdots & \cdots & \mathrm{tc}(A) \\
\mathrm{lc}(x^{d_A-1}B) & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \mathrm{tc}(x^{d_A-1}B) \\
& \ddots & & & & & & \vdots \\
& & \ddots & & & & & \vdots \\
& & & \mathrm{lc}(B) & \cdots & \cdots & \cdots & \mathrm{tc}(B)
\end{pmatrix}
\begin{matrix} \\ \\ \} d_B \\ \\ \\ \\ \} d_A \end{matrix}$$

Figure 1: The form of the Sylvester matrix of A and B. Entries outside of the gray area are zero.

The Sylvester matrix $\mathrm{Syl}(A, B)$ is defined to be the matrix of size $(d_A + d_B) \times (d_A + d_B)$ with the following entries: If $1 \leq i \leq d_B$ and $1 \leq j \leq d_A + d_B$, the entry in the ith row and jth column is the $(d_A + d_B - j)$th coefficient of $x^{d_B - i}A$. If $d_B + 1 \leq i \leq d_A + d_B$ and $1 \leq j \leq d_A + d_B$, the entry in the ith row and jth column is the $(d_A + d_B - j)$th coefficient of $x^{d_A - (i - d_B)}B$.

For $i, j \in \mathbb{N}$ with $0 \leq j \leq i \leq d_B$, the matrix $\mathrm{Syl}_{i,j}(A, B)$ is obtained from $\mathrm{Syl}(A, B)$ by removing the rows 1 to i, the rows $d_B + 1$ to $d_B + i$, the columns 1 to i and the last $i + 1$ columns except for the column $d_A + d_B - j$.

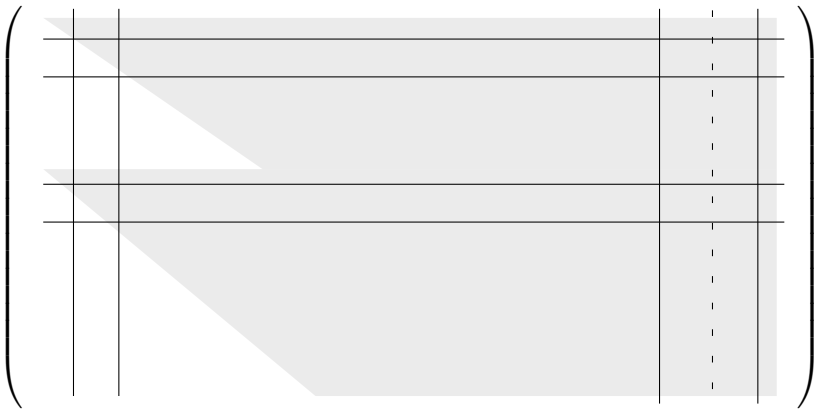

Figure 2: Sketch of $\mathrm{Syl}_{2,1}(A, B)$. The lines indicate the removed rows and columns. The column under the dotted line is added again.

Definition 4. For $0 \leq i \leq d_B$, the polynomial

$$\mathrm{sres}_i(A, B) := \sum_{j=0}^{i} \det(\mathrm{Syl}_{i,j}(A, B))x^j$$

is called the ith (polynomial) subresultant of A and B. If the order of $\mathrm{sres}_i(A, B)$ is strictly less than i, the ith subresultant of A and B is called defective, otherwise it is called regular. The subresultant sequence of A and B of the first kind is the subsequence of

$$(A, B, \mathrm{sres}_{d_B - 1}(A, B), \mathrm{sres}_{d_B - 2}(A, B), \ldots, \mathrm{sres}_0(A, B), 0)$$

that contains A, B, the trailing zero and all nonzero $\mathrm{sres}_i(A, B)$ for which $\mathrm{sres}_{i+1}(A, B)$ is regular.

Theorem 1 (Li (1998)). *The polynomial remainder sequence given by α_i and β_i as in Example 2.3, the subresultant PRS, is equal to the subresultant sequence of A and B of the first kind.*

5. Identifying Content of Polynomial Subresultants

The representation of subresultants in terms of determinants of the matrices $\mathrm{Syl}_{i,j}(A, B)$ makes it possible to identify content by exploiting the special form of these matrices as well as the correspondence between rows of the Sylvester matrix and monomial multiples of A and B. For the case of commutative polynomials, some results are known for detecting such additional content. We generalize two results to the Ore setting. The first (Theorem 2) is a generalization of an observation mentioned in Brown (1978), which carries over quite easily to the Ore case. The second (Theorem 4) usually performs better in terms of coefficient size of the remainders, but a heuristic argument is necessary to use it algorithmically (see Section 6).

Theorem 2. *With* $t := \gcd(\sigma^{d_B-1}(\mathrm{lc}(A)), \sigma^{d_A-1}(\mathrm{lc}(B)))$ *and* $\gamma_i := \sigma^{-i}(t)$ *for* $0 \leq i \leq d_B - 1$*, we get:* $\gamma_i \mid \mathrm{cont}(\mathrm{sres}_i(A, B))$.

Proof. Let i be fixed. The coefficients of $\mathrm{sres}_i(A, B)$ are the determinants of the matrices $\mathrm{Syl}_{i,j}(A, B)$ for $0 \leq j \leq i$. The first column of all of these matrices is

$$(\sigma^{d_B-1-i}(\mathrm{lc}(A)), 0, \ldots, 0, \sigma^{d_A-1-i}(\mathrm{lc}(B)), 0, \ldots, 0)^{\mathsf{T}}.$$

Laplace expansion along this column proves the claim. □

Not all of the subresultants of A and B are in the subresultant PRS of A and B. To make use of Theorem 2 for a new PRS, we need a minor specialisation of the statement:

Corollary 1. *Let* $(R_i)_{i\in\{0,\ldots,\ell+1\}}$ *be the subresultant PRS of* A *and* B *(not necessarily normal). If we choose*

$$t = \gcd(\sigma^{d_B-1}(\mathrm{lc}(A)), \sigma^{d_A-1}(\mathrm{lc}(B))), \quad \gamma_2 = \sigma^{-d_B+1}(t) \text{ and } \gamma_i = \sigma^{d_{i-2}-d_{i-1}}(\gamma_{i-1}) \text{ for } 2 < i \leq \ell,$$

then $\gamma_i \mid \mathrm{cont}(R_i)$ *for* $2 \leq i \leq \ell$.

Proof. Suppose R_i is the jth subresultant of A and B. Then, by the definition of the subresultant sequence of the first kind and Theorem 1, the $(j+1)$st subresultant of A and B is regular. Because of this and the subresultant block structure (see Li (1998)), R_{i-1} is of order $j+1$ and so j is equal to $d_{i-1} - 1$. By Theorem 2, the content of R_i is divisible by $\sigma^{-d_{i-1}+1}(t)$. It is easy to see that $\sigma^{-d_{i-1}+1}(t)$ is equal to γ_i. □

In the commutative case, a second source of additional content was determined, although this result is not widely known. The following theorem can be found in Knuth (1981):

Theorem 3. *Let* $A, B \in \mathbb{D}[x]$ *be such that the subresultant PRS of* A *and* B *is normal, i.e.* $d_{i-1} = d_i + 1$ *for* $1 \leq i \leq \ell$*, and let* G *be the GCD of* A *and* B*. Then* $\mathrm{lc}(G)^{2(i-1)} \mid \mathrm{cont}(R_i)$ *for* $2 \leq i \leq \ell$.

A generalization of Theorem 3 to Ore polynomials is not straightforward, as Example 4 shows.

Example 4 (Example 3 cont.). If we take A and B as in Example 3, then the leading coefficient of the GCRD of A and B is $(n+9)p(n)$, where $p(n)$ is a polynomial of degree 17. The subresultant PRS of A and B turns out to be normal and R_2 is of order $d_2 = 12$. By Theorem 3, if the polynomials were elements of $\mathbb{D}[x]$, $\mathrm{cont}(R_2)$ would be divisible by $\mathrm{lc}(G)^2$ and a naive translation of the theorem to the non-commutative case suggests divisibility by a polynomial of degree at least 36. The (monic) content of R_2, however, is only $(n+16)(n+17)$, which is contained in, but not equal to, $\sigma^7(\mathrm{lc}(G))^{[2]}$.

Again in the commutative case, let Q_A, $Q_B \in \mathbb{D}[x]$ be such that $A = Q_A G$ and $B = Q_B G$. Knuth (1981) proves Theorem 3 by showing that if $(R_i)_{i\in\{0,\ldots,\ell+1\}}$ is the subresultant PRS of A and B and $(\tilde{R}_i)_{i\in\{0,\ldots,\ell+1\}}$ is the subresultant PRS of Q_A, Q_B, then $Q_i = \mathrm{lc}(G)^{2(i-1)}\tilde{R}_i$. This approach is problematic for Ore polynomials, because there the Q_i's and the $\tilde{R}_i$'s have coefficients in $\mathbb{K}$ and not necessarily in $\mathbb{D}$. This means that even after showing that a quotient Q_i is a

$\mathbb{D}$-left multiple of some subresultant $\tilde{R}_i$ of Q_A and Q_B, the left factor and the denominators in the coefficients of $\tilde{R}_i$ might not be coprime and thus lead to cancellation. Therefore we will not only describe why in the non-commutative case only some factors of $\mathrm{lc}(G)$ appear as content, but we also present a new proof of Theorem 3 that makes it more explicit where the additional content comes from. Moreover, we won't require the remainder sequence to be normal.

In $\mathbb{D}[x]$, if A is a multiple of the primitive polynomial G, then their quotient will always have coefficients in $\mathbb{D}$, and therefore, the leading coefficient of A contains all the factors of the leading coefficient of G. For Ore polynomials, this is not necessarily true, since the quotient of A and G might be an element of $\mathbb{K}[x;\sigma,\delta] \setminus \mathbb{D}[x;\sigma,\delta]$. Still, different left multiples of G in $\mathbb{D}[x;\sigma,\delta]$ may share some common factors in their leading coefficients, as described in Lemma 1.

Lemma 1. *Let $d_T \in \mathbb{N}$ be fixed, let $\mathcal{I} \lhd \mathbb{D}[x;\sigma,\delta]$ be a left ideal and let T be any element of $\mathcal{I}$ of order d_T such that, among all the operators of order d_T in $\mathcal{I}$, its leading coefficient t is minimal with respect to the degree. Then t is independent of the choice of T (up to multiplication by units in $\mathbb{D}$) and for any $L \in \mathcal{I}$ with $d_L \leq d_T$ we have $\sigma^{d_L-d_T}(t) \mid \mathrm{lc}(L)$.*

Proof. Assume there are $T, L \in \mathcal{I}$ for which the claim $\sigma^{d_L-d_T}(t) \mid \mathrm{lc}(L)$ does not hold. We let $L' = x^{d_T-d_L}L$ and get $\mathrm{lc}(L') = \sigma^{d_T-d_L}(\mathrm{lc}(L))$, thus $t \nmid \mathrm{lc}(L')$ by assumption. Division with remainder yields nonzero $q, r \in \mathbb{D}$ such that

$$\mathrm{lc}(L') = qt + r, \quad \deg(r) < \deg(t).$$

Hence the operator $L' - qT$ is an element of $\mathcal{I}$ whose leading coefficient has degree less than $\deg(t)$. This contradicts the choice of T.
For the uniqueness, let $T' \in \mathcal{I}$ be any other operator of order d_T with minimal leading coefficient degree. By what was just shown above, we get $\mathrm{lc}(T') \mid t$ and $t \mid \mathrm{lc}(T')$, so t and $\mathrm{lc}(T')$ are associates. □

Definition 5. Consider $\mathcal{I}$, T and t from Lemma 1. The shift $\sigma^{-d_T}(t)$ of the leading coefficient of T is called the essential part of $\mathcal{I}$ at order d_T. If there is no operator in $\mathcal{I}$ for some order n, the essential part of $\mathcal{I}$ at order n is defined to be 1.

Let $L \in \mathbb{C}[y][D;1,\frac{d}{dy}]$ and $\mathcal{I} = \mathcal{I}' \cap \mathbb{C}[y][D;1,\frac{d}{dy}]$ where $\mathcal{I}' \lhd \mathbb{C}(y)[D;1,\frac{d}{dy}]$ is the left ideal generated by L. We give an informal explanation of essential parts of $\mathcal{I}$ in terms of solutions of L, i.e. functions that are annihilated by L. Any non-removable singularity of a solution of L corresponds to a root of the leading coefficient of L, but not for any root of $\mathrm{lc}(L)$ there has to be a solution with a non-removable singularity at that point. Any solution of L is also a solution of every operator in $\mathcal{I}$ and it can happen that there are nonzero $\mathbb{K}$-left multiples of L in $\mathcal{I}$ that have strictly smaller leading coefficient degree than L. If such a *desingularized* operator exists, it means that some of the roots of $\mathrm{lc}(L)$ can be removed by multiplying L with another operator from the left. These removable roots are called the *apparent singularities* of L. It is shown in Jaroschek (2013) that there exists a unique minimal (w.r.t. degree) essential part of $\mathcal{I}$ that appears in the essential parts of $\mathcal{I}$ at every order greater than d_L. This minimal essential part of $\mathcal{I}$ is a polynomial whose roots are exactly the non-apparent singularities of L, and it turns out that for each root of the essential part of $\mathcal{I}$, there is at least one solution of L that does not admit an analytic continuation at that point. A more detailed description of desingularization and apparent

singularities of differential equations can be found in Ince (1926). Further references and recent results on desingularization of Ore operators can be found in Chen et al. (2013).

Note that for commutative polynomials, by Gauß' Lemma, the essential part of a nonzero ideal at any order is equal to the leading coefficient of the primitive greatest common divisor of the ideal elements.

For the remaining part of this article, let $\mathcal{I} \lhd \mathbb{D}[x;\sigma,\delta]$ be the left ideal generaed by A and B. We formulate our Ore generalization of Theorem 3, where now some of the essential parts of $\mathcal{I}$ play the role of the leading coefficient of the GCRD of A and B.

Theorem 4. *Let $i \in \{0,\dots,d_B-1\}$ and $\Delta := d_A + d_B - 2i$. If t_k is the essential part of $\mathcal{I}$ at order k for $i < k \leq \Delta + i - 1$, then*

$$\left(\prod_{k=i+1}^{\Delta+i-1} t_k\right) \mid \operatorname{cont}(\operatorname{sres}_i(A,B)).$$

Proof. For any $j \in \{0,\dots,i\}$, $\operatorname{Syl}_{i,j}(A,B)$ is of size $\Delta \times \Delta$ and if the last column is removed, the resulting matrix does not depend on j anymore. For $n \in \{1,\dots,\Delta-1\}$, let $\mathcal{M}_{i,n}$ be the set of all $n \times n$ matrices obtained by removing the last $\Delta - n$ columns and any $\Delta - n$ rows from $\operatorname{Syl}_{i,j}(A,B)$. The jth coefficient of $\operatorname{sres}_i(A,B)$ is the determinant of $\operatorname{Syl}_{i,j}(A,B)$ and Laplace expansion along the last column shows that it is a $\mathbb{D}$-linear combination of the elements of $\mathcal{M}_{i,\Delta-1}$. By induction on n we show that the determinant of any element of $\mathcal{M}_{i,n}$ is divisible by $t_{\Delta+i-n}t_{\Delta+i-(n-1)}\cdots t_{\Delta+i-1}$. The theorem is then proven by setting $n = \Delta - 1$.

For $n = 1$, the only entry in a matrix in $\mathcal{M}_{i,1}$ is either zero or the leading coefficient of a monomial left multiple of A or B of order $\Delta + i - 1$, so the claim follows from Lemma 1.

Now suppose the claim is true for $1 \leq n < \Delta - 1$ and let M be any element of $\mathcal{M}_{i,n+1}$. If the determinant of M is zero, then there is nothing to show. Consider the case where $\det(M) \neq 0$. Then there is a $v \in \mathbb{K}^{n+1}$ such that $M^{\mathsf{T}}v = (0,\dots,0,1)^{\mathsf{T}}$. By Cramer's rule, the jth component v_j of v is of the form $p_j/\det(M)$ where $p_j \in \mathbb{D}$ is the determinant of some element of $\mathcal{M}_{i,n}$. By induction hypothesis it is divisible by $t_{\Delta+i-n}t_{\Delta+i-(n-1)}\cdots t_{\Delta+i-1}$. Every row in M corresponds to an operator of the form x^kA or x^kB for $k \in \mathbb{N}$, minus some of the lower order terms. For the jth row, $1 \leq j \leq n+1$, we denote the corresponding operator by L_j. By the definition of v, the operator $\sum_{j=0}^{n+1} v_jL_j \in \mathbb{K}[x;\sigma,\delta]$ will have order $\Delta + i - (n+1)$ and leading coefficient 1. So if we set

$$v' := \frac{\det(M)}{t_{\Delta+i-n}t_{\Delta+i-(n-1)}\cdots t_{\Delta+i-1}} v \in \mathbb{D}^{n+1}$$

and $L = \sum_{j=0}^{n+1} v'_jL_j$, then L is an element in $\mathcal{I}$ of order $\Delta + i - (n+1)$ and its leading coefficient is $\det(M)/(t_{\Delta+i-n}t_{\Delta+i-(n-1)}\cdots t_{\Delta+i-1}) \in \mathbb{D}$. Lemma 1 yields that $\operatorname{lc}(L)$ is divisible by $t_{\Delta+i-(n+1)}$, so we get in total $t_{\Delta+i-(n+1)}t_{\Delta+i-n}\cdots t_{\Delta+i-1} \mid \det(M)$. □

In practice, the essential parts of $\mathcal{I}$ will most likely be the same at every order n with $d_G \leq n \leq d_A + d_B$. In that case, Theorem 4 is equivalent to the following simplification, where only the essential part of $\mathcal{I}$ at order $d_A + d_B$ needs to be known.

Corollary 2. *Let $i \in \{0,\dots,d_B-1\}$ and $\Delta := d_A + d_B - 2i$. If t is the essential part of $\mathcal{I}$ at order $d_A + d_B$, then*

$$\sigma^{i+1}(t)^{[\Delta-1]} \mid \operatorname{cont}(\operatorname{sres}_i(A,B)).$$

PROOF. According to Lemma 1, $\sigma^j(t)$ divides the essential part of $\mathcal{I}$ at order j for any $d_G \leq j \leq d_A + d_B$. If $i < d_G$, then the ith subresultnat of A and B is zero. Otherwise, Theorem 4 yields that $\operatorname{cont}(\operatorname{sres}_i(A,B))$ is divisible by

$$\sigma^{i+1}(t)\sigma^{i+1}(t)\ldots\sigma^{\Delta+i-1}(t) = \sigma^{i+1}(t)^{[\Delta-1]}. \qquad \square$$

Like for Theorem 2, an adjustment of Corollary 2 to the block structure of the subresultant sequence of the first kind is needed in order to construct a new PRS.

Corollary 3. *Let $(R_i)_{i\in\{0,\ldots,\ell+1\}}$ be the subresultant PRS of A and B (not necessarily normal) and let t be the essential part of $\mathcal{I}$ at order $d_A + d_B$. If we set $\gamma_2 = \sigma^{d_B}(t)^{[d_A-d_B+1]}$ and*

$$\gamma_i = \sigma^{d_{i-1}}(t)^{[d_{i-2}-d_{i-1}]}\gamma_{i-1}\sigma^{d_A+d_B-d_{i-2}+1}(t)^{[d_{i-2}-d_{i-1}]} \text{ for } 2 < i \leq \ell,$$

then $\gamma_i \mid \operatorname{cont}(R_i)$ for $2 \leq i \leq \ell$.

PROOF. Suppose R_i is the jth subresultant of A and B. As in the proof of Corollary 1, we have that j is equal to $d_{i-1}-1$. So by Corollary 2, the content of R_i is divisible by $\sigma^{d_{i-1}}(t)^{[d_A+d_B-2d_{i-1}+1]}$. Simple hand calculation shows that this is equal to γ_i. $\square$

6. Improved Polynomial Remainder Sequence

We now derive formulas for the α_i and β_i that take into account the potential additional content characterized by Theorems 2 and 4. For this we need the following lemma:

Lemma 2. *For $\gamma_1, \gamma_2 \in \mathbb{K}$: $\operatorname{pquo}(\gamma_1 A, \gamma_2 B)\gamma_2 = \gamma_1\gamma_2^{[d_A-d_B+1]}\operatorname{pquo}(A,B)$.*

PROOF. By Lemma 2.3 in Li (1998), the pseudo-remainder of $\gamma_1 A$ and $\gamma_2 B$ is the (d_B-1)st subresultant of $\gamma_1 A$ and $\gamma_2 B$ (up to sign). Consequently, its coefficients are determinants of submatrices of $\operatorname{Syl}(\gamma_1 A, \gamma_2 B)$ that contain one row corresponding to the operator $\gamma_1 A$ and $d_A - d_B + 1$ rows corresponding to operators of the form $x^i\gamma_2 B$, $0 \leq i \leq d_A - d_B$. Thus, by Lemma 2.2 in Li (1998), it follows that (up to sign)

$$\operatorname{prem}(\gamma_1 A, \gamma_2 B) = \gamma_1\gamma_2^{[d_A-d_B+1]}\operatorname{prem}(A,B). \tag{2}$$

The pseudo-remainder formula (1) applied to $\gamma_1 A$ and $\gamma_2 B$ is

$$\operatorname{lc}(\gamma_2 B)^{[d_A-d_B+1]}\gamma_1 A = \operatorname{pquo}(\gamma_1 A, \gamma_2 B)\gamma_2 B + \operatorname{prem}(\gamma_1 A, \gamma_2 B).$$

Combining this with (2) and dividing the resulting equation by $\gamma_1\gamma_2^{[d_A-d_B+1]}$ from the left gives the desired result. $\square$

This now allows us to state α_i and β_i for improved polynomial remainder sequences:

Theorem 5. *Suppose* $(R_i)_{i\in\{0,\dots,\ell+1\}}$ *is the subresultant PRS of* A *and* B *and* $(\gamma_i)_{i\in\{0,\dots,\ell+1\}}$ *is any sequence in* $\mathbb{K}\setminus\{0\}$ *with* $\gamma_0=\gamma_1=1$. *Set* $\tilde{R}_i=\frac{1}{\gamma_i}R_i$. *Then* $(\tilde{R}_i)_{i\in\{0,\dots,\ell+1\}}$ *is a PRS of* A *and* B *with:*

$$\tilde{\alpha}_i=\operatorname{lc}(\tilde{R}_i)^{[d_{i-1}-d_i+1]},$$
$$\tilde{\beta}_i=\begin{cases}-\sigma(\tilde{\psi}_1)^{[d_0-d_1]}\gamma_2, & \text{if } i=1,\\ \dfrac{-\operatorname{lc}(\tilde{R}_{i-1})\sigma(\tilde{\psi}_i)^{[d_{i-1}-d_i]}}{\gamma_i{}^{[d_{i-1}-d_i+1]}}\gamma_{i+1}, & \text{if } 2\le i\le\ell,\end{cases}$$

where

$$\tilde{\psi}_i=\begin{cases}-1, & \text{if } i=1,\\ \dfrac{(-\gamma_{i-1}\operatorname{lc}(\tilde{R}_{i-1}))^{[d_{i-2}-d_{i-1}]}}{\sigma(\tilde{\psi}_{i-1})^{[d_{i-2}-d_{i-1}-1]}}, & \text{if } 2\le i\le\ell.\end{cases}$$

Proof. From the definition of $\tilde{R}_i$ and the equations

$$\alpha_iR_{i-1}=Q_iR_i+\beta_iR_{i+1}\qquad\text{and}\qquad\alpha_i=\gamma_i^{[d_{i-1}-d_i+1]}\tilde{\alpha}_i,$$

it follows that

$$\gamma_i^{[d_{i-1}-d_i+1]}\gamma_{i-1}\tilde{\alpha}_i\tilde{R}_{i-1}=Q_i\gamma_i\tilde{R}_i+\beta_i\gamma_{i+1}\tilde{R}_{i+1}. \tag{3}$$

For the first summand on the right hand side, Lemma 2 yields

$$Q_i\gamma_i=\gamma_i^{[d_{i-1}-d_i+1]}\gamma_{i-1}\tilde{Q}_i. \tag{4}$$

For the second summand, observe that since $\gamma_i\operatorname{lc}(\tilde{R}_i)$ equals $\operatorname{lc}(R_i)$, we have that ψ_i equals $\tilde{\psi}_i$ for all $1\le i\le\ell$. Thus

$$\beta_i\gamma_{i+1}=\gamma_i^{[d_{i-1}-d_i+1]}\gamma_{i-1}\tilde{\beta}_i. \tag{5}$$

The proof is concluded by combining (3), (4) and (5) and dividing the resulting equation by $\gamma_i^{[d_{i-1}-d_i+1]}\gamma_{i-1}$ from the left. □

Two possible choices for $(\gamma_i)_{i\in\{i,\dots,\ell+1\}}$ were presented in Corollary 1 and 3. The computation of γ_i in Corollary 1 is straightforward, but in Corollary 3, the essential part of $\mathcal{I}$ (the ideal generated by A and B) at order d_A+d_B is usually not known. A simple heuristic can solve this problem in most cases: As was shown in Lemma 1, the essential part of $\mathcal{I}$ at order d_A+d_B appears in a shifted version in the leading coefficient of every nonzero ideal element with order less than or equal to d_A+d_B. In particular it is contained in $\operatorname{lc}(A)$ and $\operatorname{lc}(B)$. Thus, if t is the essential part of $\mathcal{I}$ at order d_A+d_B, we have

$$\sigma^{d_A}(t)\mid\gcd(\operatorname{lc}(A),\sigma^{d_A-d_B}(\operatorname{lc}(B))) \tag{6}$$

and in most cases, we not only have divisibility but equality. In fact, in all the examples we looked at that came from combinatorics or physics, this guess for the essential part turned out to be correct.

Example 5 (Example 4 cont.). We now use Theorem 5 and Corollaries 1 and 3 to compute new PRSs of A and B as in Example 3. The essential part of $\mathcal{I}$ at order d_A+d_B is $(n+3)$,

so $\sigma^{d_A}(n+3) = (n+17)$, which is also the guess given by the right hand side of (6). Applying Corollary 1 yields the factors

$$\gamma_2 = n+17, \quad \gamma_3 = n+18, \quad \ldots \quad \gamma_i = n+16+i-1, \quad \ldots$$

whereas Corollary 3 gives

$$\gamma_2 = (n+16)^{[2]}, \quad \gamma_3 = (n+15)^{[4]}, \quad \ldots \quad \gamma_i = (n+16-i+2)^{[2(i-1)]}, \quad \ldots$$

The improvements from Corollary 1 are marginal, while the degrees in the improved PRS with the results from Corollary 3 are equal to the degrees in the primitive PRS, except for the very last step:

PRS	R_2	R_3	R_4	R_5	R_6	R_7	R_8
subresultant	11	16	21	26	31	36	41
improved (Cor. 1)	10	15	20	25	30	35	40
improved (Cor. 3)	9	12	15	18	21	24	27
primitive	9	12	15	18	21	24	21

Table 2: Maximal coefficient degrees for the subresultant, improved and primitive PRS.

Example 6. Although the remainders in the PRS based on Corollary 3 are usually primitive when starting from randomly generated operators or operators that come from some applications, it is not guaranteed that this is always the case. As an example, consider

$$\begin{aligned} &A, B \in \mathbb{Q}[y][x], \\ &A = x^4 + yx^2 + yx + y, \\ &B = x^3 + yx^2. \end{aligned}$$

The second subresultant of A and B is $\operatorname{sres}_2(A,B) = (y+y^2)x^2 + yx + y$, so $\operatorname{cont}(\operatorname{sres}_2(A,B)) = y$, but in the improved PRS, no content will be found.

As mentioned, it may also happen that the guess for the essential part of $\mathcal{I}$ at order $d_A + d_B$ is too large, for example:

$$\begin{aligned} &A, B \in \mathbb{Q}(y)[D, 1, \tfrac{d}{dx}], \\ &A = (y+1)D^4 + D^3 + D^2 + yD + 1, \\ &B = (y+1)D^3 + D^2 + 1. \end{aligned}$$

Here, $\operatorname{cont}(R_3)$ in the subresultant PRS is $(y+1)$, but a factor $(y+1)^2$ is predicted. The mistake in predicting the essential part can be noticed on the fly during the execution of the algorithm as soon as a remainder with coefficients in $\mathbb{Q}(y)$ appears. It is then possible to either switch to another PRS or to refine the guess of the essential part. One strategy to do so is to remove all the factors from the guess that could be responsible for the appearance of denominators. Let t be the guess for the essential part of $\mathcal{I}$ at $d_A + d_B$ and let c be the non-trivial common denominator of the coefficients of a remainder R_i in the improved PRS. Furthermore let M be the set of all integers m

such that $\gcd(\sigma^m(c), t) \neq 1$. Update R_i, γ_i and t with

$$\begin{aligned} R_i &\leftarrow cR_i, \\ \gamma_i &\leftarrow \frac{\gamma_i}{c}, \\ t &\leftarrow \frac{t}{\gcd(t, \prod_{m\in M} \sigma^m(c))}, \\ \gamma_{i+1} &\leftarrow \sigma^{d_i-d_B}(t)^{[d_A+d_B-2d_i+1]}, \end{aligned}$$

and continue the computation with these new values. For differential operators in $\mathbb{C}(y)[D; 1, \frac{d}{dy}]$, we have $M = \{0\}$ and for recurrence operators in $\mathbb{C}(n)[S_n; s_n, 0]$, M contains all the integer roots of the polynomial $\mathrm{res}_n(c(n+m), t)$.

Example 7. We can guess two operators A and B in $\mathbb{Q}[n][S; s_n, 0]$ of order $d_A = 16$, $d_B = 14$, resp. that annihilate the sequence

$$t_n = (7n^3 + 5n^2 + n + 1)^7((n + 1/7)^{\overline{12}})^7 \frac{(2n)!^3}{(3n)!^2}.$$

The GCRD of A and B is of order 1 and the essential part of $\mathcal{I}$ at $d_A + d_B$ is of degree 4. The essential part of $\mathcal{I}$ at order 11, however, is of degree 11, so here we are in the rare case where the essential part of $\mathcal{I}$ at order $d_A + d_B$ is only contained but not equal to the essential part at lower orders. Formula (6) only predicts the essential part of $\mathcal{I}$ at order $d_A + d_B$ and during the GCRD computation, content that comes from lower order essential parts emerges

PRS	R_2	R_3	R_4	R_5	R_6	R_7	R_8
improved (Cor. 3)	31	44	57	70	83	96	109
primitive	31	44	50	56	62	68	74

Table 3: Maximal coefficient degrees for the first few remainders in the improved and primitive PRS.

It is possible to guess the essential part of $\mathcal{I}$ at lower orders and then use Theorem 4 to get the primitive remainders, but like in the direct computation of the primitive PRS, GCD computations in the base ring would be necessary after each division step.

As another consequence of Theorem 4, we can give a new bound for the coefficient degrees of the primitive PRS in terms of the essential parts of the left ideal generated by A and B.

Theorem 6. *Let $(R_i)_{i\in\{0,\dots,\ell+1\}}$ be the primitive PRS of A and B. Fix $i \in \{0, \dots, \ell\}$ and let $b \in \mathbb{N}$ be such that $\max_{k\in\{0,\dots,d_B-d_{i-1}\}}(\|x^k A\|) \leq b$ and $\max_{k\in\{0,\dots,d_A-d_{i-1}\}}(\|x^k B\|) \leq b$. If t_j denotes the essential part of $\mathcal{I}$ at order $j \in \mathbb{N}$, then*

$$\|R_i\| \leq (d_A + d_B - 2(d_{i-1} - 1))b - \sum_{j=d_{i-1}}^{d_A+d_B-d_{i-1}+1} \deg(t_j)$$

Proof. The bound follows directly from Hadamard's inequality, the subresultant block structure and Corollary 3. □

Acknowledgements

I would like to thank Ziming Li and Manuel Kauers for their helpful comments and support during our personal communication. Also, I thank the reviewers for their careful reading of this article and for pointing out mistakes and shortcomings in the draft.

References

Bronstein, M., Petkovšek, M., 1996. An introduction to pseudo-linear algebra. Theoretical Computer Science 157, 3–33.

Brown, W.S., 1978. The subresultant PRS algorithm. ACM Trans. Math. Softw. 4(3), 237–249.

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

Chen, S., Jaroschek, M., Kauers, M., Singer, M., 2013. Desingularization explains order-degree curves for Ore operators, in: Proc. ISSAC 2013. To appear.

Collins, G.E., 1967. Subresultants and reduced polynomial remainder sequences. J. ACM 14(1), 128–142.

Ince, E.L., 1926. Ordinary Differential Equations. Dover.

Jaroschek, M., 2013. The Role of (Non-)Apparent Singularities in the Algorithmic Treatment of Ore Operators (title pending). Ph.D. thesis. Research Institute for Symbolic Computation, Johannes Kepler University, Linz, A-4040, Austria. To appear.

Kauers, M., 2009. Guessing handbook. Technical Report, RISC, Johannes Kepler University Linz 09-07.

Knuth, D.E., 1981. The Art of Computer Programming, Volume II: Seminumerical Algorithms, 2nd Edition. Addison-Wesley.

Li, Z., 1996. A Subresultant Theory for Linear Differential, Linear Difference, and Ore Polynomials with Applications. Ph.D. thesis. Research Institute for Symbolic Computation, Johannes Kepler University, Linz, A-4040, Austria.

Li, Z., 1998. A subresultant theory for Ore polynomials with applications, in: Proc. ISSAC 98, pp. 132–139.

Loos, R., 1982. Generalized polynomial remainder sequences, in: Computer Algebra, Symbolic and Algebraic Computation. B. Buchberger, G. Collins, R. Loos (eds). Springer-Verlag, Wien-New York, pp. 115–137.

Ore, Ø., 1933. Theory of non-commutative polynomials. Annals of Mathematics 34(3), 480–508.