# Fraction-free Row Reduction of Matrices of Ore Polynomials

Bernhard Beckermann [a] Howard Cheng [b] George Labahn [c]

[a] *Laboratoire Painlevé UMR 8524 (ANO-EDP), UFR Mathématiques – M3, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France*

[b] *Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, Alberta, Canada, T1K 3M4*

[c] *School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1*

**Abstract**

In this paper we give formulas for performing row reduction of a matrix of Ore polynomials in a fraction-free way. The reductions can be used for finding the rank and left nullspace of such matrices. When specialized to matrices of skew polynomials our reduction can be used for computing a weak Popov form of such matrices and for computing a GCRD and an LCLM of skew polynomials or matrices of skew polynomials. The algorithm is suitable for computation in exact arithmetic domains where the growth of coefficients in intermediate computations is a concern. This coefficient growth is controlled by using fraction-free methods. The known factor can be predicted and removed efficiently.

## 1 Introduction

Ore rings provide a general setting for describing linear differential, recurrence, difference and $q$-difference operators. Formally these are given by $\mathbb{K}[Z;\sigma,\delta]$ with $\mathbb{K}$ a field of coefficients, $Z$ an indeterminate, $\sigma$ an injective homomorphism, $\delta$ a derivation and with the multiplication rule $Za = \sigma(a)Z + \delta(a)$ for all $a \in \mathbb{K}$. In this paper we are interested in matrices of Ore polynomials and look at the problem of transforming such matrices into "simpler" ones using

*Email addresses:* bbecker@math.univ-lille1.fr (Bernhard Beckermann), cheng@cs.uleth.ca (Howard Cheng), glabahn@scg.uwaterloo.ca (George Labahn).

only certain row operations. Examples of such transformations include conversion to special forms, such as row-reduced, Popov or weak Popov normal forms. In our case we are primarily interested in transformations which allow for easy determination of rank and left nullspaces.

For a given $m \times s$ matrix $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$ we are interested in applying two types of elementary row operations. The first type includes

(a) interchange two rows;
(b) multiply a row by a nonzero element in $\mathbb{K}[Z;\sigma,\delta]$ on the left;
(c) add a polynomial left multiple of one row to another.

In the second type of elementary row operations we include (a), (b) and (c) but require that the row multiplier in (b) comes from $\mathbb{K}$. The second set of row operations is useful, for example, when computing a Greatest Common Right Divisor (GCRD) or a Least Common Left Multiple (LCLM) of Ore polynomials.

Formally, in the first instance we can view a sequence of elementary row operations as a matrix $\mathbf{U}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ with the result of these row operations given by $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$. In the second case, $\mathbf{U}(Z)$ would have the additional property that there exists a left inverse $\mathbf{V}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ such that $\mathbf{V}(Z)\,\mathbf{U}(Z) = \mathbf{I}_m$. In the commutative case, such a transformation matrix is called unimodular [Kailath, 1980].

In many cases it is possible to transform via row operations a matrix of Ore polynomials into one whose rank is completely determined by the rank of its leading or trailing coefficient. In the commutative case, this can be done via an algorithm of Beckermann and Labahn [1997] while in the noncommutative case of skew polynomials (i.e. where $\delta = 0$) this can be done using either the EG-elimination method of Abramov [1999] or the algorithm of Abramov and Bronstein [2001]. In the commutative case, examples of applications for such transformations include matrix polynomial division, inversion of matrix polynomials, finding matrix GCDs of two matrix polynomials and finding all solutions to various rational approximation problems. For the skew polynomial case, it was shown by Abramov and Bronstein [2001] that such transformations can be used to find polynomial and rational solutions of linear functional systems.

The algorithm given by Abramov and Bronstein [2001] improves on the EG-elimination method of Abramov [1999] and extends a method of Beckermann and Labahn [1997] to the noncommutative case. While these algorithms have good arithmetic complexity, coefficient growth may occur and can only be controlled through coefficient GCD computations. Without such GCD computations the coefficient growth can be exponential. Examples of such growth can be found in Section 8.

In this paper we consider the problem of determining the rank and left nullspace of a matrix of Ore polynomials for problems where coefficient growth is an issue. Our aim is to give a fraction-free algorithm for finding these quantities when working over the domain $\mathbb{D}[Z;\sigma,\delta]$ with $\mathbb{D}$ an integral domain, and $\sigma(\mathbb{D}) \subset \mathbb{D}$, $\delta(\mathbb{D}) \subset \mathbb{D}$. Examples of such domains include $\mathbb{D} = \mathbb{F}[n]$ for some field $\mathbb{F}$ with $Z$ the shift operator and $\mathbb{D} = \mathbb{F}[x]$ and where $Z$ is the differential operator. By fraction-free we mean that we can work entirely in the domain $\mathbb{D}[Z;\sigma,\delta]$ but that coefficient growth is controlled without any need for costly coefficient GCD computations. In addition we want to ensure that all intermediate results can be bounded in size which allows for a precise analysis of the growth of coefficients of our computation.

Our results extend the algorithm of Beckermann and Labahn [2000] in the commutative case and Beckermann et al. [2002] in the case of matrices of skew polynomials. This extension has considerable technical challenges. For example, unlike the skew and commutative polynomial case, the rank is no longer necessarily determined by the rank of the leading or trailing coefficient matrix. As a result, a different termination criterion is required for matrices of Ore polynomials. We also show how to obtain a row-reduced basis of the left nullspace of matrices of Ore polynomials.

In the common special case of matrices of skew polynomials, we can say more. Our methods can be used to give a fraction-free algorithm to compute a weak Popov form for such matrices with negligible additional computations, which is an improvement over the row-reduced form obtained in our previous algorithm [Beckermann et al., 2002]. In addition, the methods can be used to compute, in a fraction-free way, a GCRD and an LCLM of skew polynomials or matrices of skew polynomials. Finally, we show how the quantities produced during such a GCRD computation relate to the subresultants of two skew polynomials [Li, 1996, 1998], the classical tools used for fraction-free GCRD computations. Therefore, we can view our algorithm as a generalization of the subresultant algorithm. Although previous algorithms (e.g. Abramov and Bronstein [2001]) may be faster in some cases, our algorithms have polynomial time and space complexities in the worst case. In particular, when coefficient growth is significant our algorithm is faster. As our methods for skew polynomials require the coefficients be reversed, we restrict our attention to the case where $\sigma$ is an automorphism when dealing with matrices of skew polynomials.

The remainder of this paper is organized as follows. In Section 2 we discuss classical concepts such as rank and left nullspace of matrices of Ore polynomials and extend some well known facts from matrix polynomial theory to matrix Ore domains. In Section 3 we give a brief overview of our approach. In Section 4 we define *order bases*, the principal tool used for our reduction while in Section 5 we place these bases into a linear algebra setting. A fraction-free recursion formula for computing order bases is given in Section 6 followed by a

discussion of the termination criterion along with the complexity of the algorithm in the following section. Section 8 gives some examples where coefficient growth is an important issue. We also compare the requirements for our algorithm and that of Abramov and Bronstein in these cases. Matrices of skew polynomials are handled in Section 9 where we show that our algorithm can be used to find a weak Popov form of such matrices. In this section we also show how the algorithm can be used to compute a GCRD and LCLM of two skew polynomials and relate order bases to subresultants in the special case of $2 \times 1$ matrices of skew polynomials. The paper ends with a conclusion along with a discussion of directions for future work. Finally, we include an appendix which gives a number of technical facts about matrices of Ore polynomials that are necessary for our results.

**Notation.** We shall adapt the following conventions for the remainder of this paper. We assume that $\mathbf{F}(Z) \in \mathbb{D}[Z;\sigma,\delta]^{m\times s}$. Let $N = \deg \mathbf{F}(Z)$, and write

$$\mathbf{F}(Z) = \sum_{j=0}^{N} F^{(j)} Z^j, \text{ with } F^{(j)} \in \mathbb{D}^{m\times s}.$$

We denote the elements of $\mathbf{F}(Z)$ by $\mathbf{F}(Z)_{k,\ell}$, and the elements of $F^{(j)}$ by $F^{(j)}_{k,\ell}$. The $j$th row of $\mathbf{F}(Z)$ is denoted $\mathbf{F}(Z)_{j,*}$. If $J \subset \{1,\dots,m\}$, the submatrix formed by the rows indexed by the elements of $J$ is denoted $\mathbf{F}(Z)_{J,*}$. For a scalar polynomial, however, we will write $f(Z)$ as $f(Z) = \sum_{j=0}^{N} f_j Z^j$. For any vector of integers (also called multi-index) $\vec{\omega} = (\omega_1,\dots,\omega_p)$, we denote by $|\vec{\omega}| = \sum_{i=1}^{p} \omega_i$. We also denote by $Z^{\vec{\omega}}$ the matrix of Ore polynomials having $Z^{\omega_i}$ on the diagonal and 0 everywhere else. A matrix of Ore polynomials $\mathbf{F}(Z)$ is said to have row degree $\vec{\nu} = \text{row-deg}\,\mathbf{F}(Z)$ (and column degree $\vec{\mu} = \text{col-deg}\,\mathbf{F}(Z)$, respectively) if the $i$th row has degree $\nu_i$ (and the $j$th column has degree $\mu_j$). The vector $\vec{e}_i$ denotes the vector having 1 in component $i$ and 0 elsewhere and $\vec{e} = (1,\dots,1)$.

## 2 Row-reduced Matrices of Ore polynomials

In this section we will generalize some classical notions such as rank, unimodular matrices, and the transformation to row-reduced matrices (see for instance Kailath [1980]) to the case of Ore matrix polynomials. For the sake of completeness, generalizations of other well known classical properties for matrix polynomials such as the invariance of the rank under row operations, the predictable degree property and minimal indices are included in the appendix.

With $\vec{\nu} = \text{row-deg}\,\mathbf{F}(Z)$ and $N = \max_j \nu_j = \deg \mathbf{F}(Z)$, we may write

$$Z^{N\vec{e}-\vec{\nu}}\,\mathbf{F}(Z) = L\,Z^N + \text{ lower degree terms},$$

where the matrix $L(\mathbf{F}(Z)) := L \in \mathbb{K}^{m\times s}$ is called the *leading coefficient matrix* of $\mathbf{F}(Z)$. In analogy with the case of ordinary matrix polynomials $\mathbf{F}(Z)$ is *row-reduced* if $\text{rank}\, L = m$.

**Definition 2.1 (Rank, Unimodular)**

*(a)* *For $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$, the quantity* $\text{rank}\,\mathbf{F}(Z)$ *is defined to be the maximum number of $\mathbb{K}[Z;\sigma,\delta]$-linearly independent rows of $\mathbf{F}(Z)$.*

*(b)* *A matrix $\mathbf{U}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ is* unimodular *if there exists a $\mathbf{V}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ such that $\mathbf{V}(Z)\,\mathbf{U}(Z) = \mathbf{U}(Z)\,\mathbf{V}(Z) = \mathbf{I}_m$.*

□

We remark that our definition of rank is different from (and perhaps simpler than) that of Cohn [1971] or Abramov and Bronstein [2001] who considers the rank of the module of rows of $\mathbf{F}(Z)$ (or the rank of the matrix over the skew-field $\mathbb{K}(Z;\sigma,\delta)$ of left fractions). This definition is more convenient for our purposes. We show in the appendix that these quantities are in fact the same.

For the main result of this section we will show that any matrix of Ore polynomials can be transformed to one whose nonzero rows form a row-reduced matrix by means of elementary row operations of the second type given in the introduction.

**Theorem 2.2** *For any $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$ there exists a unimodular matrix $\mathbf{U}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$, with $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{F}(Z)$ having $r \le \min\{m,s\}$ nonzero rows,* $\text{row-deg}\,\mathbf{T}(Z) \le \text{row-deg}\,\mathbf{F}(Z)$*, and where the submatrix consisting of the $r$ nonzero rows of $\mathbf{T}(Z)$ are row-reduced.*

*Moreover, the unimodular multiplier satisfies the degree bound*

$$\text{row-deg}\,\mathbf{U}(Z) \le \vec{\nu} + (|\vec{\mu}| - |\vec{\nu}| - \min_j\{\mu_j\})\vec{e},$$

*where $\vec{\mu} := \max(\vec{0}, \text{row-deg}\,\mathbf{F}(Z))$ and $\vec{\nu} := \max(\vec{0}, \text{row-deg}\,\mathbf{T}(Z))$.*

*Proof:* We will give a constructive proof of this theorem. Starting with $\mathbf{U}(Z) = \mathbf{I}_m$ and $\mathbf{T}(Z) = \mathbf{F}(Z)$, we construct a sequence of unimodular matrices $\mathbf{U}(Z)$ and $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{F}(Z)$, with $\text{row-deg}\,\mathbf{U}(Z) \le \vec{\nu} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}|)\vec{e}$, $\vec{\nu} = \max(\vec{0}, \text{row-deg}\,\mathbf{T}(Z))$, and the final $\mathbf{T}(Z)$ having the desired additional properties. In one step of this procedure, we will update one row of the previ-

ously computed $\mathbf{U}(Z), \mathbf{T}(Z)$ (and hence one component of $\vec{\nu}$), and obtain the new quantities $\mathbf{U}(Z)^{new}, \mathbf{T}(Z)^{new}$ with $\vec{\nu}^{new} = \max(\vec{0}, \text{row-deg}\,\mathbf{T}(Z)^{new})$.

Denote by $J$ the set of indices of zero rows of $\mathbf{T}(Z)$, and $L = L(\mathbf{T}(Z))$. If the matrix formed by the nontrivial rows of $\mathbf{T}(Z)$ is not yet row-reduced, then we can find a $\vec{v} \in \mathbb{K}^{1\times m}$ with $\vec{v} \neq \vec{0}$, $\vec{v}L = 0$, and $v_j = 0$ for $j \in J$. Choose an index $k$ with $v_k \neq 0$ (the index of the updated row) and

$$\nu_k = \max\{\nu_j : v_j \neq 0\},$$

and define $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ by $\mathbf{Q}(Z)_{1,j} = \sigma^{\nu_k - t}(v_j) Z^{\nu_k - \nu_j}$ if $v_j \neq 0$, and $\mathbf{Q}(Z)_{1,j} = 0$ otherwise, where $t = \deg \mathbf{T}(Z)$. Then

$$\begin{aligned}
\mathbf{T}(Z)^{new}_{k,*} &:= \mathbf{Q}(Z)\,\mathbf{T}(Z) \\
&= \sum_{v_j \neq 0} \sigma^{\nu_k - t}(v_j) Z^{\nu_k - \nu_j} T^{(\nu_j)}_{j,*} Z^{\nu_j} + \text{ lower degree terms} \\
&= \sum_{j=1}^{m} \sigma^{\nu_k - t}(v_j) \sigma^{\nu_k - \nu_j}(T^{(\nu_j)}_{j,*}) Z^{\nu_k} + \text{ lower degree terms} \\
&= \sigma^{\nu_k - t}(vL) Z^{\nu_k} + \text{ lower degree terms.}
\end{aligned}$$

Hence $\deg \mathbf{T}(Z)^{new}_{k,*} \leq \nu_k - 1$, showing that $\text{row-deg}\,\mathbf{T}(Z)^{new} \leq \text{row-deg}\,\mathbf{T}(Z)$. Notice that $\mathbf{U}(Z)^{new} = \mathbf{V}(Z)\,\mathbf{U}(Z)$, where $\mathbf{V}(Z)$ is obtained from $\mathbf{I}_m$ by replacing its $k$th row by $\mathbf{Q}(Z)$. Since $\mathbf{Q}(Z)_{1,k} \in \mathbb{K} \setminus \{0\}$ by construction, we may consider $\mathbf{W}(Z)$ obtained from $\mathbf{I}_m$ by replacing its $(k,j)$ entry by $-(\mathbf{Q}(Z)_{1,k})^{-1}\mathbf{Q}(Z)_{1,j}$ for $j \neq k$, and by $(\mathbf{Q}(Z)_{1,k})^{-1}$ for $j = k$. The reader may easily verify that $\mathbf{W}(Z)\,\mathbf{V}(Z) = \mathbf{V}(Z)\,\mathbf{W}(Z) = \mathbf{I}_m$. Thus, as with $\mathbf{U}(Z)$, $\mathbf{U}(Z)^{new}$ is also unimodular. Making use of the degree bounds for $\mathbf{U}(Z)$, we also get that $\deg(\mathbf{Q}(Z)\,\mathbf{U}(Z)) \leq \nu_k - \mu_k + |\vec{\mu}| - |\vec{\nu}|$. Hence the degree bounds for $\mathbf{U}(Z)^{new}$ are obtained by observing that

$$\text{row-deg}\,\mathbf{U}(Z)^{new} \leq \vec{\nu} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}|)\vec{e} \leq \vec{\nu}^{new} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}^{new}|)\vec{e}.$$

Finally, we notice that, in each step of the algorithm, we either produce a new zero row in $\mathbf{T}(Z)$, or else decrease $|\vec{\nu}|$, the sum of the row degrees of nontrivial rows of $\mathbf{T}(Z)$, by at least one. Hence the procedure terminates, which implies that the nonzero rows of $\mathbf{T}(Z)$ are row-reduced.

**Remark 2.3** *The algorithm given in the proof of Theorem 2.2 closely follows the one in Beckermann and Labahn [1997], Eqn. (12), for ordinary matrix polynomials, and is similar to that of Abramov and Bronstein [2001] in case of skew polynomials. However, we prefer to perform our computations with skew polynomials instead of Laurent skew polynomials (e.g. when $Z$ is the differentiation operator). The degree bounds given in Theorem 2.2 for the multiplier matrix $\mathbf{U}(Z)$ appear to be new.*

**Remark 2.4** *In the case of commutative polynomials there is an example in [Beckermann et al., 2001, Example 5.6] of a* $\mathbf{F}(Z)$ *which is unimodular (and hence* $\mathbf{T}(Z) = \mathbf{I}$*), has row degree* $N\vec{e}$ *and where its multiplier satisfies* $\text{row-deg}\,\mathbf{U}(Z) = (m-1)N\vec{e}$*. Hence the worst case estimate of Theorem 2.2 for the degree of* $\mathbf{U}(Z)$ *is sharp.*

In Theorem A.2 of the appendix we will prove that the quantity $r$ of Theorem 2.2 in fact equals the rank of $\mathbf{F}(Z)$. In addition, this theorem will also show that the matrix $\mathbf{U}(Z)$ of Theorem 2.2 gives some important properties about a basis for the left nullspace of $\mathbf{F}(Z)$ given by

$$\mathcal{N}_{\mathbf{F}(Z)} = \{\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m} : \mathbf{Q}(Z)\,\mathbf{F}(Z) = \mathbf{0}\}.$$

Furthermore, various other properties are included in the appendix. In particular we prove in Lemma A.3 that the rank does not change after performing elementary row operations of the first or second kind.

## 3 Overview

Theorem 2.2 shows that one way to compute a row-reduced form is to repeatedly eliminate unwanted high-order coefficients, until the leading coefficient matrix has the appropriate rank. Instead of eliminating high-order coefficients, our approach is to eliminate low-order coefficients. In the case of skew polynomials a suitable substitution (see Section 9) can be made to reverse the coefficients to eliminate high-order coefficients. By performing elimination until the trailing coefficient has a certain rank (or in triangular form), we can reverse the coefficients to obtain a row-reduced form (or a weak Popov form).

We introduce the notion of order and order bases for the elimination of low-order coefficients. Roughly, the order of an Ore polynomial is the smallest power of $Z$ with a nonzero coefficient; an order basis is a basis of the module of all left polynomial combinations of the rows of $\mathbf{F}(Z)$ such that the combinations have a certain number of low-order coefficients being zero. One can, in fact, view an order basis as a rank-preserving transformation which results in an Ore matrix with a particular order. If the basis element corresponds to a left polynomial combination which is identically zero, then it is also an element in the left nullspace of $\mathbf{F}(Z)$. If we obtain the appropriate number of left polynomial combinations which are identically zero, we get a basis for the left nullspace of $\mathbf{F}(Z)$ because the elements in an order basis are linearly independent.

From degree bounds on the elements in the order basis, we obtain linear systems of equations for the unknown coefficients in an order basis. By studying

the linear systems we obtain results on uniqueness as well as a bound on the sizes of the coefficients in the solutions. The coefficient matrices (called striped Krylov matrices) of these linear systems have a striped structure, so that each stripe consists of the coefficients of $Z^k$ multiplied by a row of $\mathbf{F}(Z)$ for some $k$. One may apply any technique for solving systems of linear equations to obtain an order basis. However, the structure inherent in striped Krylov matrices of the linear systems are not exploited.

Our algorithm exploits the structure by performing elimination on only one row for each stripe. The recursion formulas given in Section 6 are equivalent to performing fraction-free Gaussian elimination [Bareiss, 1968] on the striped Krylov matrix to incrementally eliminate the columns. By performing elimination on the matrix of Ore polynomials directly, our algorithm controls coefficient growth without having to perform elimination on the much larger Krylov matrix. The relationship with fraction-free Gaussian elimination is also used to show that our algorithm can be considered a generalization of the subresultant algorithm (cf. Section 9.4).

## 4 Order Basis

In this section we introduce the notion of *order* and *order bases* for a given matrix of Ore polynomials $\mathbf{F}(Z)$. These are the primary tools which will be used for our algorithm. Informally, we are interested in taking left linear combinations of rows of our input matrix $\mathbf{F}(Z)$ in order to eliminate low order terms, with the elimination differing for various columns. Formally such an elimination is captured using the concept of order.

**Definition 4.1 (Order)** *Let $\mathbf{P}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ be a vector of Ore polynomials and $\vec{\omega}$ a multi-index. Then $\mathbf{P}(Z)$ is said to have* order $\vec{\omega}$ *if*

$$\mathbf{P}(Z)\,\mathbf{F}(Z) = \mathbf{R}(Z)\,Z^{\vec{\omega}} \tag{1}$$

*with $\mathbf{R}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times s}$. The matrix $\mathbf{R}(Z)$ in (1) is called a* residual. □

We are interested in *all* possible row operations which eliminate lower order terms of $\mathbf{F}(Z)$. Using our formalism, this corresponds to finding all left linear combinations (over $\mathbb{K}[Z;\sigma,\delta]$) of elements of a given order. This in turn is captured in the definition of an order basis, which gives a basis of the module of all vectors of Ore polynomials having a particular order.

**Definition 4.2 (Order Basis)** *Let $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$, and $\vec{\omega}$ be a multi-index. A matrix of Ore polynomials $\mathbf{M}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ is said to be an*

order basis *of order* $\vec{\omega}$ *and column degree* $\vec{\mu}$ *if there exists a multi-index* $\vec{\mu} = (\mu_1, ..., \mu_m)$ *such that*

*(a) every row of* $\mathbf{M}(Z)$ *has order* $\vec{\omega}$*,*

*(b) for every* $\mathbf{P}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ *of order* $\vec{\omega}$ *there exists a* $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ *such that*

$$\mathbf{P}(Z) = \mathbf{Q}(Z)\,\mathbf{M}(Z),$$

*(c) there exists a nonzero* $d \in \mathbb{K}$ *such that*

$$\mathbf{M}(Z) = d\,Z^{\vec{\mu}} + \mathbf{L}(Z)$$

*where* $\deg \mathbf{L}(Z)_{k,\ell} \leq \mu_\ell - 1$*.*

*If in addition* $\mathbf{M}(Z)$ *is row-reduced, with* row-deg $\mathbf{M}(Z) = \vec{\mu}$*, then we refer to* $\mathbf{M}(Z)$ *as a* reduced order basis*.* □

Part (a) of Definition 4.2 states that every row of an order basis eliminates rows of $\mathbf{F}(Z)$ up to a certain order while part (b) implies that the rows describe all eliminates of the order. The intuition of part (c) is that $\mu_i$ gives the number of times row $i$ has been used as a pivot row in a row elimination process. A reduced order basis has added degree constraints, which can be thought of as fixing the pivots.

By the Predictable Degree Property for matrices of Ore polynomials shown in Lemma A.1(a) of the appendix we can show that an order basis will be a reduced order basis if and only if row-deg $\mathbf{M}(Z) \leq \vec{\mu}$, and we have the added degree constraint in part (b) that, for all $j = 1, ..., m$,

$$\deg \mathbf{Q}(Z)_{1,j} \leq \deg \mathbf{P}(Z) - \mu_j. \tag{2}$$

**Example 4.3** *Let* $\mathbb{D} = \mathbb{Z}[x]$*,* $\sigma(a(x)) = a(x)$*, and* $\delta(a(x)) = \frac{d}{dx}a(x)$ *for all* $a(x) \in \mathbb{D}$ *and*

$$\mathbf{F}(Z) = \begin{bmatrix} 2Z^2 + 2xZ + x^2 & Z^2 - Z + 2 \\ xZ + 2 & 3xZ + 1 \end{bmatrix}. \tag{3}$$

*Then an order basis for* $\mathbf{F}(Z)$ *of order* $(1,1)$ *and degree* $(1,1)$ *is given by*

$$\mathbf{M}(Z) = \begin{bmatrix} (x^2-4)Z - 2x & 4x \\ 0 & (x^2-4)Z \end{bmatrix}.$$

*Note that* $\mathbf{M}(Z)$ *is a reduced order basis.* □

We remark that the definition of order basis given in Beckermann et al. [2002] is slightly more restrictive than our definition of reduced order basis given

here. We use the more general definition in order to gain more flexibility with our pivoting.

A key theorem for proving the correctness of the fraction-free algorithm deals with the uniqueness of order bases. The proof in Beckermann et al. [2002] is not applicable for the new definition of order bases and so we give a new proof here for this result.

**Theorem 4.4** *Let $\mathbf{M}(Z)$ be an order basis of order $\vec{\omega}$ and degree $\vec{\mu}$.*

*(a) There exists only the trivial row vector $\mathbf{P}(Z) = \vec{0}$ with column degree $\leq \vec{\mu} - \vec{e}$ and order $\geq \vec{\omega}$.*
*(b) For any $k$, a row vector with column degree $\leq \vec{\mu} - \vec{e} + \vec{e}_k$ and order $\geq \vec{\omega}$ is unique up to multiplication with an element from $\mathbb{K}$.*
*(c) An order basis of a particular order and degree is unique up to multiplication by constants from $\mathbb{K}$.*

*Proof:* We only need to show part (a) as (b) and (c) follow directly from (a). Suppose that $\mathbf{P}(Z) \neq \vec{0}$ has order $\vec{\omega}$ and column degree $\vec{\mu} - \vec{e}$. By Definition 4.2(b), there exists $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ such that $\mathbf{P}(Z) = \mathbf{Q}(Z)\,\mathbf{M}(Z)$. Let $j$ be an index such that $\deg \mathbf{Q}(Z)_{1,j}$ is maximum. Since $\mathbf{P}(Z) \neq \vec{0}$, it follows that $\deg \mathbf{Q}(Z)_{1,j} \geq 0$. Now,

$$\deg \mathbf{P}(Z)_{1,j} = \deg \left( \sum_{k=1}^{m} \mathbf{Q}(Z)_{1,k}\, \mathbf{M}(Z)_{k,j} \right).$$

Note that if $k \neq j$, then

$$\begin{aligned} \deg \mathbf{Q}(Z)_{1,k}\, \mathbf{M}(Z)_{k,j} &= \deg \mathbf{Q}(Z)_{1,k} + \deg \mathbf{M}(Z)_{k,j} \\ &\leq \deg \mathbf{Q}(Z)_{1,j} + \deg \mathbf{M}(Z)_{k,j} \\ &\leq \deg \mathbf{Q}(Z)_{1,j} + \mu_j - 1. \end{aligned}$$

Also,

$$\deg \mathbf{Q}(Z)_{1,j}\, \mathbf{M}(Z)_{j,j} = \deg \mathbf{Q}(Z)_{1,j} + \mu_j,$$

so that

$$\deg \mathbf{P}(Z)_{1,j} = \deg \mathbf{Q}(Z)_{1,j} + \mu_j \geq \mu_j.$$

This contradicts the assumption that $\deg \mathbf{P}(Z)_{1,j} \leq \mu_j - 1$.

In the commutative case there are a number of characterizations of order bases. For example in Beckermann and Labahn [1997] order bases are characterized by properties on its determinant.

**Example 4.5** *Let $a(Z), b(Z) \in \mathbb{D}[Z;\sigma,0]$ with degrees $d_a, d_b$, respectively, with $d_a \geq d_b$. Set $t = d_a - d_b$, $\gamma := \prod_{i=0}^{t} \sigma^i(b_0)$ and solve*

$$\gamma\, a(Z) = q(Z)\, b(Z) + r(Z)\, Z^{t+1} \tag{4}$$

*with* $\deg q(Z) = t$ *and* $\deg r(Z) < d_b$. *Equation (4) corresponds to solving the linear system of equations*

$$\gamma\,[a_0, \ldots, a_t] = [q_0, \ldots, q_t] \begin{bmatrix} b_0 & \sigma(b_1) & \cdots & \sigma^t(b_t) \\ & \sigma(b_0) & & \vdots \\ & & \ddots & \vdots \\ & & & \sigma^t(b_0) \end{bmatrix}, \tag{5}$$

*an equation similar to that encountered in performing right pseudo-division of skew polynomials. Setting*

$$\mathbf{M}(Z) = \begin{bmatrix} \gamma & -q(Z) \\ 0 & \gamma Z^{t+1} \end{bmatrix}$$

*we see that*

$$\mathbf{M}(Z) \begin{bmatrix} a(Z) \\ b(Z) \end{bmatrix} = \begin{bmatrix} r(Z) \\ w(Z) \end{bmatrix} Z^{t+1}$$

*where* $w(Z) = \gamma\,\sigma^{t+1}(b(Z)) = \gamma \sum_{i=0}^{d_b} \sigma^{t+1}(b_i) Z^i$. *Properties (a) and (c) of Definition 4.2 are trivially satisfied by* $\mathbf{M}(Z)$. *Property (b) follows from the linear equations given in the next section.* □

## 5 Determinantal Representations

Assume now that the entries of $\mathbf{F}(Z)$ come from $\mathbb{D}[Z; \sigma, \delta]$. We are interested in constructing an algorithm for recursively computing order bases $\mathbf{M}(Z) \in \mathbb{K}^{m \times m}[Z; \sigma, \delta]$ for increasing orders, where $\mathbb{K} = \mathbb{Q}_{\mathbb{D}}$, the quotient field of $\mathbb{D}$. In order to predict the size of these objects and predict common factors, we derive in this section a determinantal representation together with a particular choice of the constant $d$ arising in Definition 4.2(c).

Because the order condition in Definition 4.1 is on the right, we observe that if

$$\mathbf{F}(Z) = \sum_j F^{(j)} Z^j, \quad \mathbf{P}(Z) = \sum_k P^{(k)} Z^k,$$

then we have

$$\mathbf{P}(Z)\,\mathbf{F}(Z) = \sum_j S^{(j)} Z^j \tag{6}$$

with the unknowns $P^{(k)}$ obtained by constructing a system of linear equations by setting the undesired coefficients of $S^{(j)}$ equal to zero.

Let us examine the underlying system of linear equations. Notice first that for any $\mathbf{A}(Z) \in \mathbb{K}[Z;\sigma,\delta]$ we may write

$$c_k(Z\,\mathbf{A}(Z)) = \sigma(c_{k-1}(\mathbf{A}(Z))) + \delta(c_k(\mathbf{A}(Z))) \tag{7}$$

where $c_k$ denotes the $k$th coefficient of a polynomial (with $c_{-1} = 0$). We may write (7) in terms of linear algebra. Denote by $\mathbf{C} = (c_{u,v})_{u,v=0,1,\ldots}$ the lower triangular infinite matrix of operators defined by $c_{u,u} = \delta$, $c_{u+1,u} = \sigma$ and 0 otherwise, and by $\mathbf{C}_\mu$ ($\mu \geq 0$) its principal submatrix of order $\mu$. Furthermore, for each $\mathbf{A}(Z) \in \mathbb{K}[Z;\sigma,\delta]$ and nonnegative integer $\mu$ we associate vectors of coefficients

$$\mathbf{A}^{(\mu)} = [c_0(\mathbf{A}(Z)), \ldots, c_{\mu-1}(\mathbf{A}(Z))]^T = [A^{(0)}, \ldots, A^{(\mu-1)}]^T, \tag{8}$$

$$\mathbf{A} \;= [c_0(\mathbf{A}(Z)), c_1(\mathbf{A}(Z)), \ldots]^T \quad = [A^{(0)}, A^{(1)}, \quad \ldots \quad]^T. \tag{9}$$

Note that we begin our row and column enumeration at 0. We can interpret (7) in terms of matrices by

$$\mathbf{C}_\mu\,\mathbf{A}^{(\mu)} = [c_0(Z\,\mathbf{A}(Z)), \ldots, c_{\mu-1}(Z\,\mathbf{A}(Z))]^T.$$

Comparing with (6), we know that $\mathbf{P}(Z)$ has order $\vec{\omega}$ if and only if for each $\ell = 1, ..., s, j = 0, ..., \omega_\ell - 1$ we have

$$\sum_{k=1}^{m} c_j(\mathbf{P}(Z)_{1,k}\,\mathbf{F}(Z)_{k,\ell}) = 0.$$

If we wish to find solutions $\mathbf{P}(Z)$ such that $\deg \mathbf{P}(Z)_{1,k} \leq \nu_k$ for some multi-index $\vec{\nu}$, then we obtain a system of linear equations of the form

$$(P_{1,1}^{(0)}, ..., P_{1,1}^{(\nu_1)}, ..., P_{1,m}^{(0)}, ..., P_{1,m}^{(\nu_m)})\,K(\vec{\nu}+\vec{e},\vec{\omega}) = 0, \tag{10}$$

where the coefficient matrix has the form

$$K(\vec{\nu}+\vec{e},\vec{\omega}) = (K_{k,\ell}(\nu_k+1,\omega_\ell))_{k=1,\ldots,m}^{\ell=1,\ldots,s},$$

and $K_{k,\ell}(\nu_k+1,\omega_\ell)^T$ may be written as

$$\begin{bmatrix} \mathbf{F}_{k,\ell}^{(\omega_\ell)} & \mathbf{C}_{\omega_\ell}\,\mathbf{F}_{k,\ell}^{(\omega_\ell)} & \cdots & \mathbf{C}_{\omega_\ell}^{\nu_k}\,\mathbf{F}_{k,\ell}^{(\omega_\ell)} \end{bmatrix}. \tag{11}$$

Thus, the matrix $K(\vec{\nu}+\vec{e},\vec{\omega})^T$ is in the form of a striped Krylov matrix [Beckermann and Labahn, 2000], except that by stepping from one column to the next we not only multiply with a lower shift matrix but also apply the functions $\sigma$ and $\delta$. Thus, in contrast to Beckermann and Labahn [2000], here

we obtain a striped Krylov matrix with a matrix $\mathbf{C}$ having operator-valued elements.

**Example 5.1** *Let $\mathbf{F}(Z)$ be as in Example 4.3 with $Z$ a differential operator. Then we have*

$$K((3,3),(3,3)) = \left[\begin{array}{ccc|ccc} x^2 & 2x & 2 & 2 & -1 & 1 \\ 2x & x^2+2 & 2x & 0 & 2 & -1 \\ 2 & 4x & x^2+4 & 0 & 0 & 2 \\ \hline 2 & x & 0 & 1 & 3x & 0 \\ 0 & 3 & x & 0 & 4 & 3x \\ 0 & 0 & 4 & 0 & 0 & 7 \end{array}\right].$$

□

**Example 5.2** *In the case of matrices of skew polynomials, the $\nu\times\omega$ submatrix $K_{k,\ell}(\nu,\omega)$ is simply*

$$\left[\begin{array}{ccccccc} \sigma^0(F_{k,\ell}^{(0)}) & \sigma^0(F_{k,\ell}^{(1)}) & \sigma^0(F_{k,\ell}^{(2)}) & \cdots & \cdots & \sigma^0(F_{k,\ell}^{(\omega-1)}) \\ 0 & \sigma^1(F_{k,\ell}^{(0)}) & \sigma^1(F_{k,\ell}^{(1)}) & \cdots & \cdots & \sigma^1(F_{k,\ell}^{(\omega-2)}) \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & \sigma^{\nu-1}(F_{k,\ell}^{(0)}) & \cdots & \sigma^{\nu-1}(F_{k,\ell}^{(\omega-\nu)}) \end{array}\right].$$

*Thus with $\mathbf{F}(Z)$ as in (3) but with $\sigma(a(x)) = a(x+1)$ and $\delta = 0$ we have*

$$K((3,3),(3,3)) = \left[\begin{array}{ccc|ccc} x^2 & 2x & 2 & 2 & -1 & 1 \\ 0 & (x+1)^2 & 2(x+1) & 0 & 2 & -1 \\ 0 & 0 & (x+2)^2 & 0 & 0 & 2 \\ \hline 2 & x & 0 & 1 & 3x & 0 \\ 0 & 2 & x+1 & 0 & 1 & 3(x+1) \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right].$$

□

According to (10), it follows from Theorem 4.4 that if there exists an order basis $\mathbf{M}(Z)$ of order $\vec{\omega}$ and degree $\vec{\mu}$ then $K(\vec{\mu},\vec{\omega})$ has full row rank, and more precisely

$$k = 1, ..., m: \quad \text{rank } K(\vec{\mu},\vec{\omega}) = \text{rank } K(\vec{\mu}+\vec{e}_k,\vec{\omega}) = |\vec{\mu}|. \tag{12}$$

Suppose more generally that $\vec{\mu}$ and $\vec{\omega}$ are multi-indices verifying (12). We call a *multigradient* $d = d(\vec{\mu}, \vec{\omega})$ any constant $\pm 1$ times the determinant of a regular submatrix $K_*(\vec{\mu}, \vec{\omega})$ of maximal order of $K(\vec{\mu}, \vec{\omega})$, and a *Mahler system* corresponding to $(\vec{\mu}, \vec{\omega})$ a matrix of Ore polynomial $\mathbf{M}(Z)$ with rows having order $\vec{\omega}$ and degree structure

$$\mathbf{M}(z) = d \cdot Z^{\vec{\mu}} + \text{ lower order column degrees.}$$

In order to show that such a system exists, we state explicitly the linear system of equations needed to compute the unknown coefficients of the $k$th row of $\mathbf{M}(Z)$: denote by $b^k(\vec{\mu}, \vec{\omega})$ the row added while passing from $K(\vec{\mu}, \vec{\omega})$ to $K(\vec{\mu} + \vec{e}_k, \vec{\omega})$. Then, by (10), the vector of coefficients is a solution of the (overdetermined) system

$$x \cdot K(\vec{\mu}, \vec{\omega}) = d \cdot b^k(\vec{\mu}, \vec{\omega})$$

which by (12) is equivalent to the system

$$x \cdot K_*(\vec{\mu}, \vec{\omega}) = d \cdot b^k_*(\vec{\mu}, \vec{\omega}), \tag{13}$$

where in $b^k_*(\vec{\mu}, \vec{\omega})$ and in $K_*(\vec{\mu}+\vec{e}_k, \vec{\omega})$ we keep the same columns as in $K_*(\vec{\mu}, \vec{\omega})$. Notice that by Cramer's rule, (13) leads to a solution with coefficients in $\mathbb{D}$. Moreover, we may formally write down a determinantal representation of the elements of a determinantal order basis, namely

$$\mathbf{M}(Z)_{k,\ell} = \pm \det \left[ K_*(\vec{\mu} + \vec{e}_k, \vec{\omega}) \middle| \mathbf{E}_{\ell, \mu_\ell - 1 + \delta_{\ell,k}}(Z) \right] \tag{14}$$

with

$$\mathbf{E}_{\ell,\nu}(Z) = [0, \ldots, 0 | 1, Z, \ldots, Z^\nu | 0, \ldots, 0]^T, \tag{15}$$

the nonzero entries in $\mathbf{E}_{\ell,\nu}(Z)$ occurring in the $\ell$th stripe. In addition, we have that

$$\mathbf{R}(Z)_{k,\ell}\, Z^{\vec{\omega}} = \sum_j \mathbf{M}(Z)_{k,j} \mathbf{F}(Z)_{j,\ell} = \pm \det \left[ K_*(\vec{\mu} + \vec{e}_k, \vec{\omega}) \middle| \mathbf{E}_{\ell, \vec{\mu} + \vec{e}_k}(Z) \right], \tag{16}$$

where

$$\mathbf{E}_{\vec{\nu}}(Z) = [\mathbf{F}(Z)_{1,\ell}, \ldots, Z^{\nu_1 - 1} \mathbf{F}(Z)_{1,\ell} | \ldots\ldots | \mathbf{F}(Z)_{m,\ell}, \ldots, Z^{\nu_m - 1} \mathbf{F}(Z)_{m,\ell}]^T.$$

In both (14) and (16) the matrices have commutative entries in all but the last column. It is understood that the determinant in both cases is expanded along this column.

Finally we mention that, by the uniqueness result of Theorem 4.4, any order basis of degree $\vec{\mu}$ and order $\vec{\omega}$ coincides up to multiplication with some element in $\mathbb{K}$ with an Mahler system associated to $(\vec{\mu}, \vec{\omega})$, which therefore itself is an order basis of the same degree and order. By a particular pivoting technique we get a reduced order basis by computing Mahler systems.

## 6 Fraction-free Recursion Formulas for Order Bases

In this section we show how to recursively compute order bases in a fraction-free way. This can also be thought of as constructing a sequence of eliminations of lower order terms of $\mathbf{F}(Z)$. In terms of linear algebra, the recursion can be viewed as a type of fraction-free Gaussian elimination which takes into consideration the special structure of the coefficient matrix of the linear system associated to the "elimination of lower order terms" problem.

For an order basis $\mathbf{M}(Z)$ of order $\vec{\omega}$ and degree $\vec{\mu}$ having a Mahler system normalization, we look at the first terms of the residuals. If they are all equal to zero then we have an order basis of a higher order. Otherwise, we give a recursive formula for building an order basis of higher order and degree. However, a priori this new order basis has coefficients from $\mathbb{K} = \mathbb{Q}_{\mathbb{D}}$, the quotient field of $\mathbb{D}$, since we divide through some factors. In our case, however, the new system will be a Mahler system according to the existence and uniqueness results established by the determinantal representations, and hence we will keep objects with coefficients in $\mathbb{D}$.

In the following theorem we give a recurrence relation which closely follows the case of skew polynomials [Beckermann et al., 2002] and the commutative case [Beckermann and Labahn, 2000, Theorem 6.1(c)]. The resulting order bases have properties similar to those cited by Beckermann and Labahn [2000, Theorems 7.2 and 7.3].

**Theorem 6.1** *Let $\mathbf{M}(Z)$ be an order basis of order $\vec{\omega}$ and degree $\vec{\mu}$, and $\lambda \in \{1, ..., s\}$. Denote by $r_j = c_{\omega_\lambda}((\mathbf{M}(Z)\,\mathbf{F}(Z))_{j,\lambda})$, the $(j, \lambda)$ entry of the first term of the residual of $\mathbf{M}(Z)$. Finally, set $\widetilde{\vec{\omega}} := \vec{\omega} + \vec{e}_\lambda$.*

*(a) If $r_1 = ... = r_m = 0$ then $\widetilde{\mathbf{M}}(Z) := \mathbf{M}(Z)$ is an order basis of degree $\vec{\nu} := \vec{\mu}$ and order $\widetilde{\vec{\omega}}$.*

*(b) Otherwise, let $\pi$ be an index such that $r_\pi \neq 0$. Then an order basis $\widetilde{\mathbf{M}}(Z)$ of degree $\vec{\nu} := \vec{\mu} + \vec{e}_\pi$ and order $\widetilde{\vec{\omega}}$ with coefficients in $\mathbb{D}$ is obtained via the formulas*

$$p_\pi\, \widetilde{\mathbf{M}}(Z)_{\ell,k} = r_\pi\, \mathbf{M}(Z)_{\ell,k} - r_\ell\, \mathbf{M}(Z)_{\pi,k} \tag{17}$$

*for $\ell, k = 1, 2, ..., m$, $\ell \neq \pi$, and*

$$\sigma(p_\pi)\, \widetilde{\mathbf{M}}(Z)_{\pi,k} = (r_\pi\, Z - \delta(r_\pi))\, \mathbf{M}(Z)_{\pi,k} - \sum_{\ell \neq \pi} \sigma(p_\ell)\, \widetilde{\mathbf{M}}(Z)_{\ell,k} \tag{18}$$

*for $k = 1, 2, ..., m$, where $p_j = c_{\mu_j + \delta_{\pi,j} - 1}(\mathbf{M}(Z)_{\pi,j})$.*

*(c) If in addition $\mathbf{M}(z)$ is a Mahler system with respect to $(\vec{\mu}, \vec{\omega})$, then $\widetilde{\mathbf{M}}(Z)$ is also a Mahler system with respect to $(\vec{\nu}, \widetilde{\vec{\omega}})$. In particular, $\widetilde{\mathbf{M}}(Z)$ has*

*coefficients in* $\mathbb{D}$.

*Proof:* Part (a) is clear from the fact that the rows of $\mathbf{M}(Z)$ have order $\widetilde{\vec{\omega}}$ when $r_1 = \ldots = r_m = 0$.

For part (b) notice first that rows $\widetilde{\mathbf{M}}(Z)_{\ell,*}$ for $\ell \neq \pi$ have order $\widetilde{\vec{\omega}}$ by construction, as required in Definition 4.2(a). In addition row $(r_\pi Z - \delta(r_\pi))\,\mathbf{M}(Z)_{\pi,*}$ also has order $\widetilde{\vec{\omega}}$ since $(r_\pi Z - \delta(r_\pi))(r_\pi) = r_\pi \sigma(r_\pi) Z$. By construction therefore row $\widetilde{\mathbf{M}}(Z)_{\pi,*}$ has order $\widetilde{\vec{\omega}}$.

The verification of the new degree constraints of Definition 4.2(c) (with $\vec{\mu}$ being replaced by $\vec{\nu}$) for the matrix $\widetilde{\mathbf{M}}(Z)$ is straightforward and is the same as in the commutative case [Beckermann and Labahn, 2000, Theorem 7.2]. In addition, notice that $p_\pi$ is the leading coefficient of $\mathbf{M}(Z)_{\ell,\ell}$, so the leading coefficient of $\widetilde{\mathbf{M}}(Z)_{\ell,\ell}$ equals $r_\pi$ for all $\ell$ by construction. However it still remains to show that we obtain a new order basis with coefficients in $\mathbb{D}$.

We now focus on the properties of Definition 4.2(b). If $\mathbf{P}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ has order $\widetilde{\vec{\omega}}$ then it has order $\vec{\omega}$ and so there exists a $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ such that

$$\mathbf{P}(Z) = \sum_{j=1}^{m} \mathbf{Q}(Z)_{1,j}\,\mathbf{M}(Z)_{j,*}.$$

Applying the first set of row operations in (17) to rows $\ell \neq \pi$ results in

$$\mathbf{P}(Z) = \sum_{j\neq\pi} \hat{\mathbf{Q}}(Z)_{1,j}\,\widetilde{\mathbf{M}}(Z)_{j,*} + \hat{\mathbf{Q}}(Z)_{1,\pi}\,\mathbf{M}(Z)_{\pi,*} \tag{19}$$

where

$$\hat{\mathbf{Q}}(Z)_{1,j} = \mathbf{Q}(Z)_{1,j}\,\frac{p_\pi}{r_\pi} \text{ for all } j \neq \pi \text{ and } \hat{\mathbf{Q}}(Z)_{1,\pi} = \sum_{i=1}^{m} \mathbf{Q}(Z)_{1,i}\,\frac{r_i}{r_\pi}. \tag{20}$$

Since $\mathbf{P}(Z)$ and all the $\widetilde{\mathbf{M}}(Z)_{j,*}$ terms have order $\widetilde{\vec{\omega}}$, this must also be the case for $\hat{\mathbf{Q}}(Z)_{1,\pi}\,\mathbf{M}(Z)_{\pi,*}$. Let $\rho$ be the degree of $\hat{\mathbf{Q}}(Z)$ and write $\hat{\mathbf{Q}}(Z)_{1,\pi} = \sum_{k=0}^{\rho} \hat{Q}^{(k)}_{1,\pi}(r_\pi Z - \delta(r_\pi))^k$. Since $(r_\pi Z - \delta(r_\pi))r_\pi = r_\pi\sigma(r_\pi)Z$, we see that $\hat{Q}^{(0)}_{1,\pi}\,r_\pi = 0$. Therefore, by assumption on $\pi$ we have that $\hat{Q}^{(0)}_{1,\pi} = 0$. Writing $\hat{\mathbf{Q}}(Z)_{1,\pi} = \bar{\mathbf{Q}}(Z)_{1,\pi}\,(r_\pi Z - \delta(r_\pi))$ gives

$$\mathbf{P}(Z) = \sum_{j\neq\pi} \hat{\mathbf{Q}}(Z)_{1,j}\,\widetilde{\mathbf{M}}(Z)_{j,*} + \bar{\mathbf{Q}}(Z)_{1,\pi}\,(r_\pi Z - \delta(r_\pi))\,\mathbf{M}(Z)_{\pi,*}. \tag{21}$$

Completing the row operations which normalize the degrees of $\widetilde{\mathbf{M}}(Z)$ in (18) gives a $\widetilde{\mathbf{Q}}(Z)$ with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z)\,\widetilde{\mathbf{M}}(Z)$. Consequently, the property of Definition 4.2(b) holds.

Finally, in order to establish part (c) we know already from Section 5 and the existence of order bases of a specified degree and order that both $(\vec{\mu}, \vec{\omega})$ and $(\vec{\nu}, \widetilde{\vec{\omega}})$ satisfy (12). By the uniqueness result of Theorem 4.4 we only need to show that the "leading coefficient" $\widetilde{d}$ of $\widetilde{\mathbf{M}}(Z)$ in Definition 4.2(c) is a multigradient of $(\vec{\nu}, \widetilde{\vec{\omega}})$, the latter implying that $\widetilde{\mathbf{M}}(Z)$ is a Mahler system and in particular has coefficients from $\mathbb{D}$.

Denote by $d$ the corresponding "leading coefficient" of $\mathbf{M}(Z)$. In the case discussed in part (a), we do not increase the rank by going from $K(\vec{\mu}, \vec{\omega})$ to $K(\vec{\nu}, \widetilde{\vec{\omega}})$ since we just add one column and keep full row rank. Hence $d = \widetilde{d}$ being a multigradient with respect to $(\vec{\mu}, \vec{\omega})$ is also a multigradient with respect to $(\vec{\nu}, \widetilde{\vec{\omega}})$. In the final case described in part (b) we have $\widetilde{d} = r_\pi$. Using formula (16) for the residual of the $\pi$th row of $\mathbf{M}(Z)$ we learn that $r_\pi$ coincides (up to a sign) with the determinant of a submatrix of order $|\vec{\nu}|$ of $K(\vec{\nu}, \widetilde{\vec{\omega}})$. Since $r_\pi \neq 0$ by construction, it follows that $\widetilde{d} = r_\pi$ is a new multigradient, as required for the conclusion.

**Corollary 6.2** *If $\mathbf{M}(Z)$ is a reduced order basis then the order basis $\widetilde{\mathbf{M}}(Z)$ computed by (17) and (18) in Theorem 6.1 is also a reduced order basis of degree $\vec{\nu}$, provided that the pivot $\pi$ is chosen such that*

$$\mu_\pi = \min_j \{\mu_j : r_j \neq 0\}. \tag{22}$$

*Proof:* It is straightforward to check that row-deg $\widetilde{\mathbf{M}}(Z) = \vec{\nu}$. Hence, by Lemma A.1(a), it is sufficient to show that col-deg $\widetilde{\mathbf{Q}}(Z) \leq \deg(\mathbf{P}(Z))\vec{e} - \vec{\nu}$, with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z)\,\widetilde{\mathbf{M}}(Z)$ as in the proof of Theorem 6.1.

We see in (20) that $\deg \hat{\mathbf{Q}}(Z)_{1,j} \leq \deg \mathbf{P}(Z) - \mu_j = \deg \mathbf{P}(Z) - \nu_j$ for all $j \neq \pi$ while $\deg \hat{\mathbf{Q}}(Z)_{1,\pi} \leq \deg \mathbf{P}(Z) - \mu_\pi$ because of the minimality of $\mu_\pi$. In (21), $\deg \bar{\mathbf{Q}}(Z)_{1,\pi} \leq \deg \mathbf{P}(Z) - (\mu_\pi + 1) = \deg \mathbf{P}(Z) - \nu_\pi$. Completing the row operations which normalize the degrees of $\widetilde{\mathbf{M}}(Z)$ in (18) gives a $\widetilde{\mathbf{Q}}(Z)$ with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z)\,\widetilde{\mathbf{M}}(Z)$ having the correct degree bounds.

**Example 6.3** *Let $\mathbf{F}(Z)$ be defined as in Example 5.1. Starting from $\mathbf{M}(Z) = \mathbf{I}_m$ as an order basis of order $(0,0)$ and degree $(0,0)$, we can compute an order basis $\mathbf{M}_1(Z)$ of order $(1,0)$ and degree $(1,0)$ by choosing $\pi = 1$. Then $r_1 = x^2$ and $r_2 = 2$, so that*

$$\mathbf{M}_1(Z) = \begin{bmatrix} x^2 Z - 2x & 0 \\ -2 & x^2 \end{bmatrix}$$

*by (17) and (18).*

*In the next step, we note that $r_1 = -4x$ and $r_2 = x^2 - 4$. Choosing $\pi = 2$ allows us to compute an order basis of order $(1,1)$ and degree $(1,1)$. Noting*

*that the previous pivot $x^2$ is a common factor, (17) and (18) gives the order basis $\mathbf{M}(Z)$ found in Example 4.3.* □

## 7 The FFreduce Algorithm

Theorem 6.1 gives a computational procedure that results in the FFreduce algorithm given in Table 1, where the superscript $[k]$ denotes the value of a variable at iteration $k$. In this section we consider the termination criterion for this algorithm and discuss its complexity.

**Theorem 7.1 (Termination of Algorithm FFreduce)**
*Let $r = \operatorname{rank} \mathbf{F}(Z)$. The final residual $\mathbf{R}(Z)$ has rank $r$ and $m - r$ zero rows. Moreover, if $J \subset \{1, \ldots, m\}$ is the set of row indices corresponding to the zero rows of $\mathbf{R}(Z)$, then the rows $\mathbf{M}(Z)_{j,*}$ for $j \in J$ form a row-reduced basis of the left nullspace $\mathcal{N}_{\mathbf{F}(Z)}$ of $\mathbf{F}(Z)$.*

*Proof:* Recall that the last computed Mahler system $\mathbf{M}(Z)$ results from iteration $k = s\kappa$, $\kappa = mN + 1$, and has order $\kappa\vec{e}$ and degree $\vec{\mu}$.

The statement $\operatorname{rank} \mathbf{F}(Z) = \operatorname{rank} \mathbf{R}(Z)$ follows from Lemma A.3 since $\mathbf{R}(Z)Z^{\kappa}$ is obtained from $\mathbf{F}(Z)$ by applying row operations of the first type.

In order to show that $\mathbf{R}(Z)$ has $m - r$ zero rows, let $\mathbf{W}(Z)$ be as in Theorem A.2, with $\vec{\alpha} = \text{row-deg}\, \mathbf{W}(Z)$. Recall from Theorem A.2 that $\mathbf{W}(Z)$ is row-reduced, and that $\vec{\alpha} \leq (m-1) \cdot N\vec{e}$. Since the rows of $\mathbf{W}(Z)$ have order $\kappa\vec{e}$, there exists $\mathbf{Q}(Z) \in \mathbb{K}[Z; \sigma, \delta]^{(m-r)\times m}$ such that $\mathbf{W}(Z) = \mathbf{Q}(Z)\,\mathbf{M}(Z)$. By construction, $\mathbf{M}(Z)$ is a reduced order basis, and therefore row-reduced, with row degree $\vec{\mu}$. Lemma A.1(c) then implies that there is some permutation $p : \{1, \ldots, m-r\} \mapsto \{1, \ldots, m\}$, with $\alpha_j \geq \mu_{p(j)}$ for $j = 1, ..., m - r$. Hence, for $j = 1, ..., m - r$,

$$\begin{aligned}
\deg \mathbf{R}(Z)_{p(j),*} &= -\kappa + \deg(\mathbf{R}(Z)_{p(j),*} Z^{\kappa\vec{e}}) = -\kappa + \deg(\mathbf{M}(Z)_{p(j),*}\, \mathbf{F}(Z)) \\
&\leq -\kappa + N + \deg(\mathbf{M}(Z)_{p(j),*}) = -\kappa + N + \mu_{p(j)} \\
&\leq -\kappa + N + \alpha_j \leq -\kappa + mN = -1,
\end{aligned}$$

showing that these $m - r$ rows $\mathbf{R}(Z)_{p(j),*}$ are indeed zero rows.

It remains to show the part on the rows $\mathbf{M}(Z)_{j,*}$ for $j \in J$. Clearly, with $\mathbf{M}(Z)$, also the submatrix $\mathbf{M}(Z)_{J,*}$ is row-reduced. Any $\mathbf{P}(Z) \in \mathcal{N}_{\mathbf{F}(Z)}$ has order $\kappa\vec{e}$, so there exists $\mathbf{Q}(Z) \in \mathbb{K}[Z; \sigma, \delta]^{1\times m}$ such that $\mathbf{P}(Z) = \mathbf{Q}(Z)\,\mathbf{M}(Z)$. Thus,

$$\mathbf{Q}(Z)\,\mathbf{R}(Z)Z^{\kappa} = \mathbf{Q}(Z)\,\mathbf{M}(Z)\,\mathbf{F}(Z) = \mathbf{P}(Z)\,\mathbf{F}(Z) = \mathbf{0}.$$

Table 1
*The FFreduce Algorithm*

ALGORITHM FFreduce

INPUT: Matrix of Ore polynomials $\mathbf{F} \in \mathbb{D}[Z;\sigma,\delta]^{m\times s}$.

OUTPUT: Mahler system $\mathbf{M} \in \mathbb{D}[Z;\sigma,\delta]^{m\times m}$,

Residual $\mathbf{R} \in \mathbb{D}[Z;\sigma,\delta]^{m\times s}$ with $\operatorname{rank}\mathbf{F}$ nonzero rows,

Degree $\vec{\mu}$, order $\vec{\omega}$.

INITIALIZATION: $\mathbf{M}^{[0]} \leftarrow \mathbf{I}_m$, $\mathbf{R}^{[0]} \leftarrow \mathbf{F}$, $d^{[0]} \leftarrow 1$, $\vec{\mu}^{[0]} \leftarrow \vec{0}, \vec{\omega}^{[0]} \leftarrow \vec{0}$,

$N \leftarrow \deg(\mathbf{F}(Z))$, $\rho \leftarrow 0$, $k \leftarrow 0$

**While** $k < (mN+1)s$ **do**

$\rho^{[k]} \leftarrow \rho$, $\rho \leftarrow 0$

**For** $\lambda = 1,..,s$ **do**

Calculate for $\ell = 1,..,m$: first term of residuals $r_\ell \leftarrow \mathbf{R}^{[k]}(0)_{\ell,\lambda}$

Define set $\Lambda = \{\ell \in \{1,..,m\} : r_\ell \neq 0\}$.

**If** $\Lambda = \{\}$ **then** $\mathbf{M}^{[k+1]} \leftarrow \mathbf{M}^{[k]}$, $\mathbf{R}^{[k+1]} \leftarrow \mathbf{R}^{[k]}$, $d^{[k+1]} \leftarrow d^{[k]}$, $\vec{\mu}^{[k+1]} \leftarrow \vec{\mu}^{[k]}$

**else**

Choose pivot $\pi^{[k]} \leftarrow \min\{\ell \in \Lambda : \mu_\ell^{[k]} = \min_j\{\mu_j^{[k]} : j \in \Lambda\}\}$.

Calculate for $\ell = 1,..,m$, $\ell \neq \pi^{[k]}$: $p_\ell \leftarrow c_{\mu_\ell^{[k]}-1}(\mathbf{M}^{[k]}_{\pi^{[k]},\ell})$.

Increase order for $\ell = 1,..,m$, $\ell \neq \pi^{[k]}$:

$$\mathbf{M}^{[k+1]}_{\ell,*} \leftarrow \tfrac{1}{d^{[k]}}[r_{\pi^{[k]}}\,\mathbf{M}^{[k]}_{\ell,*} - r_\ell\,\mathbf{M}^{[k]}_{\pi^{[k]},*}]$$
$$\mathbf{R}^{[k+1]}_{\ell,*} \leftarrow \tfrac{1}{d^{[k]}}[r_{\pi^{[k]}}\,\mathbf{R}^{[k]}_{\ell,*} - r_\ell\,\mathbf{R}^{[k]}_{\pi^{[k]},*}]$$

Increase order and adjust degree constraints for row $\pi^{[k]}$:

$$\mathbf{M}^{[k+1]}_{\pi^{[k]},*} \leftarrow \tfrac{1}{\sigma(d^{[k]})}[(r_{\pi^{[k]}}\,Z - \delta(r_{\pi^{[k]}}))\,\mathbf{M}^{[k]}_{\pi^{[k]},*} - \textstyle\sum_{\ell\neq\pi^{[k]}}\sigma(p_\ell)\,\mathbf{M}^{[k+1]}_{\ell,*}]$$
$$\mathbf{R}^{[k+1]}_{\pi^{[k]},*} \leftarrow \tfrac{1}{\sigma(d^{[k]})}[(r_{\pi^{[k]}}\,Z - \delta(r_{\pi^{[k]}}))\,\mathbf{R}^{[k]}_{\pi^{[k]},*} - \textstyle\sum_{\ell\neq\pi^{[k]}}\sigma(p_\ell)\,\mathbf{R}^{[k+1]}_{\ell,*}]$$

Update multigradient, degree and $\rho$:

$d^{[k+1]} \leftarrow r_{\pi^{[k]}}$, $\vec{\mu}^{[k+1]} \leftarrow \vec{\mu}^{[k]} + \vec{e}_{\pi^{[k]}}$, $\rho \leftarrow \rho + 1$

**end if**

Adjust residual in column $\lambda$: for $\ell = 1,...,m$

$\mathbf{R}^{[k+1]}_{\ell,\lambda} \leftarrow \mathbf{R}^{[k+1]}_{\ell,\lambda}/Z$ (formally)

$\vec{\omega}^{[k+1]} \leftarrow \vec{\omega}^{[k]} + \vec{e}_\lambda$, $k \leftarrow k+1$

**end for**

**end while**

$\mathbf{M} \leftarrow \mathbf{M}^{[k]}$, $\mathbf{R} \leftarrow \mathbf{R}^{[k]}$, $\vec{\mu} \leftarrow \vec{\mu}^{[k]}$, $\vec{\omega} \leftarrow \vec{\omega}^{[k]}$

The relation $r = \operatorname{rank} \mathbf{R}(Z)$ implies that the nonzero rows of $\mathbf{R}(Z)$ are $\mathbb{Q}_{\mathbb{D}}[Z;\sigma,\delta]$-linearly independent, and hence $\mathbf{Q}(Z)_{1,j} = 0$ for $j \notin J$. Consequently, the rows of $\mathbf{M}(Z)_{J,*}$ form a basis of $\mathcal{N}_{\mathbf{F}(Z)}$, as claimed in Theorem 7.1.

In what follows we denote by cycle the set of iterations $k = \kappa s, \kappa s+1, ..., (\kappa+1)s-1$ in algorithm FFreduce for some integer $\kappa$ (that is, the execution of the inner loop).

Let us comment on possible improvements of our termination criterion. In all examples given in the remainder of this section, we choose as $\mathbb{D}$ the set of polynomials in $x$ with rational coefficients, with $Z = \frac{d}{dx}$, and thus $\sigma(a(x)) = a(x)$, $\delta(a(x)) = \frac{d}{dx}a(x)$.

**Remark 7.2** *The above proof was based on the estimate $\alpha_j \leq (m-1)N$ for the left minimal indices of the left nullspace $\mathcal{N}_{\mathbf{F}(Z)}$, which for general matrix polynomials is quite pessimistic, but can be attained, as shown in [Beckermann et al., 2001, Example 5.6] for ordinary matrix polynomials. For applications where a lower bound $\gamma$ is available for $|\vec{\nu}|$, the sum of the row degrees of the nontrivial rows of the row-reduced counterpart of $\mathbf{F}(Z)$ (compare with Theorem 2.2), it would be sufficient to compute Mahler systems up to the final order $(mN+1-\gamma)\vec{e}$, since then we get from Theorem 2.2 and Theorem A.2 the improved estimate $\alpha_j \leq (m-1)N - \gamma$.*

**Remark 7.3** *In contrast to the special case of skew polynomials (compare with [Beckermann et al., 2002, Lemma 5.2]), the pivots $\pi^{[k]}$ in one cycle are not necessarily distinct. In case $s > m$, there might be even up to $s$ nontrivial steps in one cycle of the algorithm. Thus $|\vec{\mu}^{[k]}|$ may be as large as $k$ (all iterations are nontrivial). As an example, consider*

$$\mathbf{F}(Z) = [1, x+Z],$$

*leading to $\pi^{[0]} = \pi^{[1]} = 1$.*

**Remark 7.4** *In the special case of skew polynomials ($\delta = 0$), the rank of any matrix polynomial $\mathbf{F}(Z)$ (over $\mathbf{Q}[Z;\sigma,\delta]$) is bounded below by the rank of its trailing coefficient $\mathbf{F}(0)$ (over $\mathbf{Q}$). This property is no longer true for general Ore domains, as it becomes clear from the example*

$$\mathbf{F}(Z) = \begin{bmatrix} 1 & x \\ Z & 1+xZ \end{bmatrix}.$$

*Here the rank of $\mathbf{F}(0)$ is 2, whereas the second row of $\mathbf{F}(Z)$ equals $Z$ times the first row of $\mathbf{F}(Z)$, and hence* $\operatorname{rank} \mathbf{F}(Z) = 1$.

**Remark 7.5** *If in the cycle starting at $k = \kappa s$ there are only distinct pivots, following [Beckermann et al., 2002, Lemma 5.1] we may still prove that the*

*rank of* $\mathbf{R}^{[\kappa s]}(0)$ *coincides with the number of pivots used in this cycle. However, in contrast to [Beckermann et al., 2002, Lemma 5.2], it is no longer true in general that the number of pivots (or distinct pivots) in a cycle is increasing. Indeed, for the example*

$$\mathbf{F}(Z) = \begin{bmatrix} 1 - xZ & 0 \\ 0 & 1 - \epsilon xZ \end{bmatrix}$$

*we have in the first cycle* $\pi^{[0]} = 1$, $\pi^{[1]} = 2$, *giving rise to*

$$\mathbf{R}^{[2]}(Z)Z = \begin{bmatrix} -xZ^2 & 0 \\ 0 & (1-\epsilon)xZ - \epsilon xZ^2 \end{bmatrix}.$$

*Then* $k = 2$ *is a trivial iteration, and there is either one (for* $\epsilon \neq 1$*) or no pivot (for* $\epsilon = 1$*) in the second cycle. Moreover, if* $\epsilon$ *is a positive integer, then we have* 2 *pivots in all further cycles up to the* $\epsilon$*th one. Thus, the trailing coefficients of the residuals after a cycle do not remain nonsingular.*

For the above reasons, we believe that it is quite unlikely that there exists an early termination criterion for FFreduce in Ore domains such as (26) below based on the number of pivots in one cycle which insures that one has found $\operatorname{rank} \mathbf{F}(Z)$. The situation is different for the special case of skew polynomials discussed in Beckermann et al. [2002] which will be further studied in the next section.

Let us now examine bounds on the sizes of the intermediate results in the FFreduce algorithm, leading to a bound on the complexity of the algorithm. For our analysis, we assume that the coefficient domain $\mathbb{ID}$ satisfies

$$\begin{aligned} \operatorname{size}(a+b) &= \mathcal{O}(\max(\operatorname{size}(a), \operatorname{size}(b))) \\ \operatorname{size}(a\,b) &= \mathcal{O}(\operatorname{size}(a) + \operatorname{size}(b)) \\ \operatorname{cost}(a+b) &= \mathcal{O}(\max(\operatorname{size}(a), \operatorname{size}(b))) \\ \operatorname{cost}(a\,b) &= \mathcal{O}(\operatorname{size}(a)\operatorname{size}(b)), \end{aligned}$$

where the function "size" measures the total storage required for its arguments and the function "cost" estimates the number of bit operations required to perform the indicated arithmetic. These assumptions are justified for large operands where, for example, the cost of addition is negligible in comparison to the cost of multiplication.

In a first step, let us examine the size of the coefficients and the complexity of one iteration of algorithm FFreduce.

**Lemma 7.6** *Let* $N = \deg \mathbf{F}(Z)$, *and let* $K$ *be a bound on the size of the coefficients appearing in* $\mathbf{F}(Z)_{j,*}$, $Z\,\mathbf{F}(Z)_{j,*}, \ldots, Z^{\mu_j}\,\mathbf{F}(Z)_{j,*}$ *for* $j = 1, \ldots, m$,

*where $\vec{\mu} = \vec{\mu}^{[k]}$. Then the size of the coefficients in $\mathbf{M}^{[k]}$ and $\mathbf{R}^{[k]}$ is bounded by $\mathcal{O}(|\vec{\mu}|K)$. Moreover, the cost at iteration $k$ is bounded by $\mathcal{O}((msN|\vec{\mu}|^2 + (m + s)|\vec{\mu}|^3)K^2)$.*

*Proof:* Equations (14) and (16) show that both the Mahler system and the residual can be represented as determinants of a square matrix of order $|\vec{\mu}|$. The coefficients in this matrix are coefficients of $\mathbf{F}(Z)_{k,*}, Z\,\mathbf{F}(Z)_{k,*}, \ldots, Z^{\mu_k}\,\mathbf{F}(Z)_{k,*}$. Hence the well-known Hadamard inequality gives the above bound for the size of the coefficients.

In order to obtain the cost, we have to take into account essentially only the multiplication of each row of $(\mathbf{M}^{[k]}, \mathbf{R}^{[k]})$ by two scalars and the multiplication of the pivot row by at most $m + 1$ scalars. It remains to count the number of coefficients, and to take into account that each multiplication with a coefficient has a cost bounded by $\mathcal{O}(|\vec{\mu}|^2K^2)$.

By slightly generalizing [Beckermann and Labahn, 2000, Theorem 6.2], we deduce the following complexity bound (compare also with [Beckermann et al., 2002, Theorem 5.5]).

**Corollary 7.7** *Let $N = \deg \mathbf{F}(Z)$, and let $K$ be a bound on the size of the coefficients appearing in $\mathbf{F}(Z)_{j,*}, Z\,\mathbf{F}(Z)_{j,*}, \ldots, Z^{\mu_j}\,\mathbf{F}(Z)_{j,*}$ for $j = 1, \ldots, m$, where $\vec{\mu} = \vec{\mu}^{[k]}$ of iteration $k$ of FFreduce. Then the total cost for computing $\mathbf{M}^{[k]}$ and $\mathbf{R}^{[k]}$ by algorithm FFreduce is bounded by $\mathcal{O}((msN|\vec{\mu}|^3 + (m + s)|\vec{\mu}|^4)K^2)$.*

*In the general Ore case, we obtain for FFreduce a worst case bit complexity of $\mathcal{O}((m+s)m^4s^4N^4K^2)$, whereas in the case of skew polynomials we may obtain the slightly sharper worst case bound $\mathcal{O}((m + s)m^4 \min(m, s)^4N^4K^2)$.*

*Proof:* The first part of the Corollary is an immediate consequence of Lemma 7.6 and of the fact that the number of iterations in the FFreduce algorithm in which any reduction is done equals $|\vec{\mu}|$. In order to show the second part, we use the bound $|\vec{\mu}| \leq |\vec{\omega}|$ with the final order vector $\vec{\omega} = (mN + 1)\vec{e}$, and $|\vec{\omega}| = s(mN + 1)$. In case of skew polynomials, pivots are distinct, and hence their number in a cycle is bounded by $\min(m, s)$ (in fact by the rank of $\mathbf{F}(Z)$), leading to the bound $|\vec{\mu}| \leq \min(m, s)(mN + 1)$.

**Remark 7.8** *The complexity model proposed before Lemma 7.6 is reasonable not only for $\mathbb{D} = \mathbb{Z}$, but also for $\mathbb{D} = \mathbb{K}[x]$ as long as we measure the size of elements only in terms of $x$-degrees and ignore growth of coefficients. However, the latter simplification is no longer acceptable for domains such as $\mathbb{D} = \mathbb{Z}[x]$, and we have to adapt our complexity analysis.*

*For $a \in \mathbb{Z}[x]$, let $\deg_x(a)$ denote the degree of $a$ with respect to $x$, and $\|a\|$ be the maximal absolute value of the integer coefficients of $a$. A good measure for*

*size for a nonzero* $a \in \mathbb{Z}[x]$ *seems to be*

$$size(a) = \mathcal{O}((1 + \deg_x(a))(1 + \log \|a\|)),$$

*since it reflects worst case memory requirements. In addition the two rules*

$$cost(a + b) = \mathcal{O}(\max(size(a), size(b)))$$
$$cost(a\, b) = \mathcal{O}(size(a)\ size(b)).$$

*continue to hold. However, it is easy to construct polynomials where the rules for* $size(a+b)$ *and* $size(ab)$ *given before Lemma 7.6 are no longer true because of cross products between degrees and the bit lengths of the coefficients. The essential ingredient in the proof of Lemma 7.6 (and thus of Corollary 7.7) was to predict the size of a coefficient* $c^{[k]} \in \mathbb{Z}[x]$ *in* $\mathbf{M}^{[k]}$ *or in* $\mathbf{R}^{[k]}$*, by means of its determinant representation in terms of a matrix of order* $|\vec{\mu}^{[k]}|$ *containing suitable coefficients of* $Z^j\mathbf{F}(Z)$ *for suitable* $j$*. Here we propose to estimate separately the* $x$*-degree and the norm of* $c^{[k]}$*. In our three examples below the applications* $\sigma, \delta : \mathbb{Z}[x] \mapsto \mathbb{Z}[x]$ *will not increase the degree, and thus one easily checks that*

$$\deg_x c^{[k]} \leq |\vec{\mu}^{[k]}|\, K_{deg},$$

*with* $K_{deg}$ *being the maximal degree of a coefficient occurring in* $\mathbf{F}(Z)$*. Define also* $K_{bit}$ *to be the logarithm of the largest norm of a coefficient occurring in* $\mathbf{F}(Z)$*. We will show below that the logarithm of the norm of an entry of the above-mentioned matrix is bounded for our three examples by*

$$K_{bit} + (\max_{\ell} \mu_{\ell}^{[k]})\, f(K_{deg}) \tag{23}$$

*for a suitable function* $f$ *depending only on* $\sigma, \delta$*, and hence*

$$size(c^{[k]}) = \mathcal{O}((1 + |\vec{\mu}^{[k]}|\, K_{deg})(1 + |\vec{\mu}^{[k]}|\, K_{bit} + |\vec{\mu}^{[k]}|\, (\max_{\ell} \mu_{\ell}^{[k]})\, f(K_{deg})))$$

*or*

$$size(c^{[k]}) = \mathcal{O}(K_{deg}\, K_{bit}|\vec{\mu}^{[k]}|^2 + K_{deg} f(K_{deg})|\vec{\mu}^{[k]}|^3),$$

*in contrast to* $size(c^{[k]}) = \mathcal{O}(K|\vec{\mu}^{[k]}|)$ *derived in Lemma 7.6. As a consequence, we may directly generalize both Lemma 7.6 and Corollary 7.7, but now higher powers will be involved. Notice that a tighter estimate could be obtained if we specify the size and cost of the sums and products in two components (*$\deg_x(a)$ *and* $\|a\|$*) separately [Li, 2003].*

*Let us first consider the skew-symmetric case* $\sigma(a(x)) = a(\alpha x)$*,* $\delta(a) = 0$*, for an integer* $\alpha \neq 0$*. Since for the norm of the coefficients of* $Z^k x^j$ *we get* $\log(||\sigma^k(x^j)||) = j\, k\, \log(|\alpha|)$*, we observe that (23) holds with* $f(K_{deg}) = K_{deg}\, \log(|\alpha|)$*.*

*More generally, for the skew-symmetric case* $\sigma(a(x)) = a(\alpha x + \beta)$*,* $\delta(a) = 0$ *with integers* $\alpha \neq 0$ *and* $\beta$*, we have* $\log(||\sigma^k(x^j)||) \leq j\, k\, \log(2\, \max(|\alpha|, |\beta|))$*. Thus here (23) holds with* $f(K_{deg}) = K_{deg}\, \log(2 \max(|\alpha|, |\beta|))$*.*

*We finally consider the differential case in which $\sigma$ is the identity and $\delta(a) = \frac{d}{dx}a$ for all $a \in \mathbb{Z}[x]$. Then $\sigma$ does not increase the norm, and $||\delta(a)|| \leq \deg_x(a)\,||a||$, implying that (23) holds with $f(K_{deg}) = \log(K_{deg})$.*

## 8 Comparisons and Examples

In this section we give some examples which allow us to make some simple comparisons with the algorithm in Abramov and Bronstein [2001]. We make no claims that our algorithm performs better than theirs in general. Indeed for examples where coefficient growth does not enter into the problem, the algorithm of Abramov and Bronstein is typically faster than the one presented in this paper. However, there are instances where the growth of coefficients does become a significant factor and in such cases the near linear growth of our algorithm does allow us to solve larger problems.

The Abramov-Bronstein algorithm uses the constructive approach outlined in Theorem 2.2. It also incorporates a number of additional improvements, for example making use of a basis of elements from the nullspace of the leading or trailing coefficients (rather than just a single element) in order to reduce the number of iterations [Abramov and Bronstein, 2002]. We also note that since the row-reduced form is not unique, the results computed by the Abramov-Bronstein algorithm are typically different from the ones obtained by FFreduce.

It is possible, as suggested in Abramov and Bronstein [2001], to compute the basis for the nullspace by using fraction-free Gaussian elimination on the leading or trailing coefficient matrix, see Bareiss [1968]. This also results in a fraction-free algorithm for row-reducing a matrix of skew polynomials. However it is not the case that this guarantees a reasonable growth of coefficient size. For example, one step of such a method could result in an increased size of coefficients by a factor of $r + 1$ where $r$ is the rank of the actual trailing or leading coefficient matrix. This occurs because the nullspace obtained by Bareiss's method could be as large as $r$ times the original input size. Even removing the contents of the nullspace elements afterwards will not guarantee good coefficient growth as our examples below illustrate.

The implementation of the Abramov-Bronstein algorithm used for our comparisons is that programmed in Maple given in the routine *LinearFunctionalSystems[MatrixTriangularization]*. This implementation finds a basis for the nullspace by working over a field and then clearing denominators. Notice that this approach is mathematically equivalent to using fraction-free Gaussian elimination and then removing the contents from individual basis elements. Note that the contents are only removed from the basis elements used to per-

form the elimination. The contents in the intermediate results are not removed, so that exponential growth may still occur. This implementation performs additional optimizations when the trailing coefficient has a zero row or a zero column. This reduces the number of iterations required to obtain the final result. Our fraction-free algorithm can be adopted to perform such shifts as well. In our comparison, such optimizations are performed in the Abramov-Bronstein algorithm but not in the fraction-free one. Finally, we have done a slight modification to ensure that it works in the case when the rank is not full.

We have run several examples, including those of [Abramov and Bronstein, 2002], in which the dimensions of the matrices, as well as the degree, are varied. For the measure of size we have used the sum of Maple's length of all the coefficients over $\mathbb{Q}[n]$, namely the coefficients of the residuals for the AB algorithm and the coefficients of both the Mahler system and the residuals for FFreduce.

For examples in which coefficient growth is not significant, the Abramov-Bronstein algorithm is in general faster, sometimes by more than a factor of 1000. For these examples, the cost of GCD computations required for removing the content (or for clearing fractions) was negligible.

In contrast, consider the matrix

$$\mathbf{F}(Z) = \begin{bmatrix} \sum_{i=0}^{N} p_i Z^i & \sum_{i=0}^{N-1} p_i Z^i \\ \sum_{i=0}^{N} p_{i+N+1} Z^i & \sum_{i=0}^{N-1} p_{i+N+1} Z^i \end{bmatrix} \tag{24}$$

where $p_i$ is the $(i+1)$-th prime and where we are working over the commutative polynomial domain $\mathbb{Z}[Z]$. The storage and running time requirements for this matrix using the two algorithms is given in Figure 1. In particular we see that the growth in the Abramov-Bronstein algorithm is exponential (varying between 48 for $N = 5$ and 58685030 for $N = 300$) while that of FFreduce is essentially linear for this case (varying between 97 and 880154). This of course impacts the timings of the two algorithms for this example.

Similarly such growth is also possible in the noncommutative case of skew polynomials. For example, one can construct matrices similar to that of (24) but using a noncommutative $Z$ and get comparable behaviour. This is the case with

$$\mathbf{F}(Z) = \begin{bmatrix} q_{0,N}(Z) & q_{0,N-1}(Z) & q_{0,N-2}(Z) \\ q_{2N+2,N}(Z) & q_{2N+2,N-1}(Z) & q_{4N+4,N-2}(Z) \\ q_{4N+4,N}(Z) & q_{4N+4,N-1}(Z) & q_{2N+2,N-2}(Z) \end{bmatrix} \tag{25}$$

Fig. 1. Plots for timings and size for FFreduce and the AB algorithm on the matrices defined in (24).

where $q_{j,k}(Z) = \sum_{i=0}^{k}(p_{2i+j+1}n + p_{2i+j+2})Z^i$ and $Z$ is the forward shift operator acting on $n$. The experimental results are shown in Table 2.

Finally, in Table 3 we show experimental results on larger matrices, in this case of skew polynomials which are generated by applying random transformations to the final result in reverse.

## 9 Applications for Skew Polynomials

In this section we show how the FFreduce algorithm can be used to solve a number of different problems in the special case when the input is a matrix of skew polynomials. Of course when $\sigma$ is the identity then this also gives fraction-free algorithms for ordinary matrix polynomials. We note again that $\sigma$ is assumed to be an automorphism on $\mathbb{Q}_{\mathbf{D}}$.

In the case of skew polynomials [Beckermann et al., 2002], the termination criterion

$$\rho^{[\kappa s]} + \text{the number of zero rows in } \mathbf{R}^{[\kappa s]}(Z) = m \tag{26}$$

| | AB | | FFreduce | |
|---|---|---|---|---|
| $d$ | Time (sec) | Size | Time (sec) | Size |
| 2 | 0.123 | 654 | 0.101 | 1488 |
| 3 | 0.125 | 2606 | 0.239 | 4589 |
| 4 | 0.287 | 7920 | 0.455 | 8621 |
| 5 | 0.691 | 27972 | 0.900 | 17267 |
| 6 | 1.582 | 84523 | 1.867 | 27208 |
| 7 | 4.656 | 265003 | 2.717 | 44369 |
| 8 | 19.342 | 714330 | 6.334 | 62900 |
| 9 | 331.509 | 1948947 | 20.334 | 92194 |
| 10 | 1943.193 | 4770766 | 148.652 | 122964 |
| 11 | 5821.765 | 12177824 | 516.682 | 169323 |
| 12 | 10144.400 | 27971967 | 631.781 | 213626 |
| 13 | ? | ? | 1528.602 | 280124 |
| 14 | ? | ? | 1660.289 | 340995 |
| 15 | ? | ? | 2403.154 | 432665 |

Table 2
Timings and storage for the AB algorithm and FFreduce on input matrices (25). An entry of ? means that no result was obtained within the time limit of 3 hours.

| | AB | | FFreduce | |
|---|---|---|---|---|
| $m, s$ | Time (sec) | Size | Time (sec) | Size |
| 2 | 32.609 | 365188 | 1.600 | 26295 |
| 3 | 542.440 | 2004249 | 145.799 | 430330 |
| 4 | 1996.640 | 1343010 | 546.931 | 950614 |
| 5 | ? | ? | 1480.871 | 1830960 |
| 6 | ? | ? | 2837.691 | 1959785 |
| 7 | 8955.809 | 25525731 | 3851.930 | 2353846 |
| 8 | ? | ? | 5132.750 | 2732281 |

Table 3
Timings and storage for the AB algorithm and FFreduce on input matrices generated by random transformations. An entry of ? means that no result was obtained within the time limit of 4 hours.

allows us to prove [Beckermann et al., 2002, Theorem 5.3] that

$$\operatorname{rank} \mathbf{R}^{[\kappa s]}(0) = \operatorname{rank} \mathbf{R}^{[\kappa s]}(Z)) = \operatorname{rank} \mathbf{F}(Z), \tag{27}$$

the rank of the trailing coefficient matrix $\mathbf{R}^{[\kappa s]}(0)$ being defined over the quotient field $\mathbb{Q}_{\mathbb{D}}$. Moreover [Beckermann et al., 2002, Lemma 5.2],

$$\text{the pivots } \pi^{[k]} \text{ for } \kappa s - s \le k < \kappa s \text{ are distinct,} \tag{28}$$

and hence [Beckermann et al., 2002, Lemma 5.1 and Lemma 5.2]

$$\rho^{[\kappa s]} = \operatorname{rank} \mathbf{R}^{[\kappa s]}(0) = \operatorname{rank} \mathbf{R}^{[\kappa s - s]}(0). \tag{29}$$

It is also shown implicitly in the proof of [Beckermann et al., 2002, Theorem 5.4] that $\kappa \le m(N+1)$ which has to be compared with the number of cycles, $mN + 1$, required by FFreduce. Thus the new termination property (26) essentially does not increase the complexity of algorithm FFreduce, but for many examples may improve the run time.

### *9.1 Full Rank Decomposition and Solutions of Linear Functional Systems*

When $\mathbf{F}(Z)$ represents a system of linear recurrence equations, one can show that an equivalent system with a leading (or trailing) coefficient with full row rank allows one to obtain bounds on the degrees of the numerator and the denominator of all rational solutions. This has been used by Abramov and Bronstein [2001] to compute rational solutions of linear functional systems.

In [Beckermann et al., 2002] it is shown that the output of FFreduce applied to $\mathbf{F}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times s}$ can be used to construct $\mathbf{T}(Z^{-1}) \in \mathbb{D}[Z^{-1};\sigma^{-1},0]^{m\times m}$ and implicitly $\mathbf{S}(Z) \in \mathbb{Q}_{\mathbb{D}}[Z;\sigma,0]^{m\times m}$ such that

$$\mathbf{T}(Z^{-1})\,\mathbf{F}(Z) = \mathbf{W}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times s}, \qquad \mathbf{S}(Z)\mathbf{T}(Z^{-1}) = \mathbf{I}_m,$$

with the number of nonzero rows of $\mathbf{W}(Z)$ coinciding with the rank of the trailing coefficient $\mathbf{W}(0)$, and hence with the rank of $\mathbf{W}(Z)$. The existence of a left inverse $\mathbf{S}(Z)$ shows that the reduction process is invertible in the algebra of Laurent skew polynomials, moreover, we obtain a *full rank decomposition* $\mathbf{F}(Z) = \mathbf{S}(Z)\mathbf{W}(Z)$ in $\mathbb{Q}_{\mathbb{D}}[Z;\sigma,0]$.

In this context we should mention that an exact arithmetic method involving coefficient GCD computations for the computation of $\mathbf{T}(Z^{-1})\,\mathbf{F}(Z) = \mathbf{W}(Z)$ with $\mathbf{W}(Z)$ as above has already been given in Abramov and Bronstein [2001].

The FFreduce algorithm as described above has been used to eliminate low-order coefficients, such that the rank of the trailing coefficient matrix is the same as the rank of the original matrix of skew polynomials. In the case of matrices of commutative polynomials, we can reverse the coefficients so that the high-order coefficients are eliminated [Beckermann and Labahn, 2000]. This allows us to obtain a row-reduced form of the input matrix polynomial.

In this section we show that a similar technique can be used to obtain a row-reduced form for a matrix of skew polynomials. Furthermore, we note that the FFreduce algorithm essentially performs fraction-free Gaussian elimination on the striped Krylov matrix. If we collect the rows used as pivots during the last cycle, we obtain a trailing coefficient that is triangular up to row permutations. As a result, reversing the coefficients gives a weak Popov form. One may reverse the coefficients in the input, invoke the FFreduce algorithm, and reverse the coefficients in the output to obtain the final results. Instead, we will modify the recursion formulas to directly eliminate the high-order coefficients.

Given $\mathbf{F}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times s}$ we can compute $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ such that the nonzero rows of $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{F}(Z)$ form a row-reduced matrix. Since we wish to eliminate high-order coefficients, we perform the substitution $\hat{Z} = Z^{-1}$, $\hat{\sigma} = \sigma^{-1}$ and perform the reduction over $\mathbb{D}[\hat{Z};\hat{\sigma},0]$. We further assume that $\sigma^{-1}$ does not introduce fractions, so that $\sigma^{-1}(a) \in \mathbb{D}$ for all $a \in \mathbb{D}$. We write $\hat{\mathbf{F}}(\hat{Z}) := \mathbf{F}(\hat{Z}^{-1})\,\hat{Z}^N$, and let $\hat{\mathbf{M}}^{[k]}(\hat{Z})$, $\hat{\mathbf{R}}^{[k]}(\hat{Z})$, $\vec{\mu}^{[k]}$, and $\vec{\omega}^{[k]}$ be the intermediate results obtained from the FFreduce algorithm with the input $\hat{\mathbf{F}}(\hat{Z})$. If we define

$$\mathbf{U}^{[k]}(Z) = Z^{\mu_k}\,\hat{\mathbf{M}}^{[k]}(\hat{Z}), \qquad \mathbf{T}^{[k]}(Z) = Z^{\mu_k}\,\hat{\mathbf{R}}^{[k]}(\hat{Z})\,\hat{Z}^{\omega_k - N\,\vec{e}}, \tag{30}$$

then $\mathbf{U}^{[k]}(Z)\,\mathbf{F}(Z) = \mathbf{T}^{[k]}(Z)$. In this case simple algebra shows that the recursion formulas for $\mathbf{U}^{[k]}(Z)$ obtained from (17) and (18) become

$$\sigma^{\mu_\ell^{[k]}}(p_{\pi^{[k]}})\,\mathbf{U}^{[k+1]}(Z)_{\ell,*} = \sigma^{\mu_\ell^{[k]}}(r_{\pi^{[k]}})\,\mathbf{U}^{[k]}(Z)_{\ell,*} - \sigma^{\mu_\ell^{[k]}}(r_\ell)\,Z^{\mu_\ell^{[k]}-\mu_{\pi^{[k]}}^{[k]}}\,\mathbf{U}^{[k]}(Z)_{\pi^{[k]},*} \tag{31}$$

for $\ell \neq \pi^{[k]}$ and

$$\begin{aligned}&\sigma^{\mu_{\pi^{[k]}}^{[k]}+2}(p_{\pi^{[k]}})\,\mathbf{U}^{[k+1]}(Z)_{\pi^{[k]},*}\\ &= \sigma^{\mu_{\pi^{[k]}}^{[k]}+1}(r_{\pi^{[k]}})\,\mathbf{U}^{[k]}(Z)_{\pi^{[k]},*} - \sum_{\ell\neq\pi^{[k]}} \sigma^{\mu_{\pi^{[k]}}^{[k]}+2}(p_\ell)\,Z^{\mu_{\pi^{[k]}}^{[k]}-\mu_\ell^{[k]}+1}\,\mathbf{U}^{[k+1]}(Z)_{\ell,*},\end{aligned} \tag{32}$$

where

$$r_\ell = \sigma^{-\mu_\ell^{[k]}}(c_{N+\mu_\ell^{[k]}-\lfloor k/s \rfloor}(\mathbf{T}^{[k]}(Z)_{\ell,(k \bmod m)+1})),$$
$$p_\ell = \sigma^{-\mu_{\pi^{[k]}}^{[k]}}(c_{\mu_{\pi^{[k]}}^{[k]}-\mu_\ell^{[k]}-\delta_{\pi^{[k]},\ell}+1}(\mathbf{U}^{[k]}(Z)_{\pi^{[k]},\ell})).$$

Since $\mu_{\pi^{[k]}}^{[k]} \leq \mu_\ell^{[k]}$ whenever $r_\ell \neq 0$, and that $p_\ell = 0$ whenever $\mu_{\pi^{[k]}}^{[k]} < \mu_\ell^{[k]} - 1$ by the definition of a reduced order basis, it follows that $\mathbf{U}^{[k+1]}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times m}$. Moreover, $[\mathbf{U}^{[k+1]}(Z), \mathbf{T}^{[k+1]}(Z)]$ is obtained from $[\mathbf{U}^{[k]}(Z), \mathbf{T}^{[k]}(Z)]$ by elementary row operations of the second type, so if $\mathbf{U}^{[k]}(Z)$ is unimodular then $\mathbf{U}^{[k+1]}(Z)$ is also unimodular.

**Theorem 9.1** *Let $\hat{\mathbf{M}}^{[k]}(\hat{Z})$, $\hat{\mathbf{R}}^{[k]}(\hat{Z})$, $\vec{\mu}^{[k]}$, and $\vec{\omega}^{[k]} = \kappa \cdot \vec{e}$ be the final output obtained from the FFreduce algorithm with the input $\hat{\mathbf{F}}(\hat{Z})$. Then*

*(a)* $\mathbf{U}^{[k]}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times m}$ *and* $\mathbf{T}^{[k]}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times s}$*;*
*(b)* $\mathbf{U}^{[k]}(Z)$ *is unimodular;*
*(c)* $\mathbf{U}^{[k]}(Z)\,\mathbf{F}(Z) = \mathbf{T}^{[k]}(Z)$*;*
*(d) the nonzero rows of* $\mathbf{T}^{[k]}(Z)$ *form a row-reduced matrix.*

*Proof:* Parts (a), (b), and (c) have already been shown above. By (27), we see that rank $\hat{\mathbf{R}}^{[k]}(0)$ = rank $\hat{\mathbf{F}}(\hat{Z})$ = rank $\hat{\mathbf{R}}^{[k]}(\hat{Z})$, which is also the number of nonzero rows in $\hat{\mathbf{R}}^{[k]}(\hat{Z})$. Therefore, the nonzero rows of $\hat{\mathbf{R}}^{[k]}(\hat{Z})$ form a matrix with trailing coefficient of full row rank. It is easy to see that row-deg $\mathbf{T}^{[k]}(Z) = \mu_k + (N-\kappa)\cdot\vec{e}$ and that

$$\mathbf{T}^{[k]}(Z)_{i,*} = \sigma^{\mu_i^{[k]}}(\hat{\mathbf{R}}^{[k]}(0)_{i,*})\, Z^{\mu_i^{[k]}+N-\kappa} + \text{lower degree terms.}$$

Therefore, $L(\mathbf{T}^{[k]}(Z)) = \sigma^{\deg \mathbf{T}^{[k]}(Z)-N+\kappa}(\hat{\mathbf{R}}(0))$. Since $\sigma$ is an automorphism on $\mathbb{Q}_\mathbf{D}$, it follows that rank $L(\mathbf{T}^{[k]}(Z))$ = rank $\hat{\mathbf{R}}^{[k]}(0)$, and hence the nonzero rows of $\mathbf{T}^{[k]}(Z)$ form a row-reduced matrix.

In fact, the FFreduce algorithm can be modified to obtain $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ such that $\mathbf{T}(Z)$ is in weak Popov form [Mulders and Storjohann, 2003] (also known as quasi-Popov form [Beckermann et al., 2001]). The weak Popov form is defined as follows.

**Definition 9.2 (Weak Popov Form)** *A matrix of skew polynomials $\mathbf{F}(Z)$ is said to be in* weak Popov Form *if the leading coefficient of the submatrix formed from the nonzero rows of $\mathbf{F}(Z)$ is in upper echelon form (up to row permutation).* □

Formally, if $\vec{\omega} = \kappa\cdot\vec{e}$ is the order obtained at the end of the FFreduce algorithm,

we form the matrices $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ by

$$[\mathbf{U}(Z)_{i,j}, \mathbf{T}(Z)_{i,j}] = \begin{cases} [\mathbf{U}^{[k]}(Z)_{i,j}, \mathbf{T}^{[k]}(Z)_{i,j}] & \text{if } \pi^{[k]} = i \text{ for some } \kappa s - s \leq k < \kappa s, \\ [\mathbf{U}^{[\kappa s]}(Z)_{i,j}, \mathbf{T}^{[\kappa s]}(Z)_{i,j}] & \text{otherwise;} \end{cases}$$

We note that $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ are well-defined because the pivots $\pi^{[k]}$ are distinct for $\kappa s - s \leq k < \kappa s$ by (28). We now show that $\mathbf{T}(Z)$ is in weak Popov form.

**Theorem 9.3** *Let $\vec{\omega} = \kappa \cdot \vec{e}$ be the order obtained from the FFreduce algorithm with the input $\hat{\mathbf{F}}(\hat{Z})$. Then*

*(a)* $\mathbf{U}(Z) \in \mathbb{D}[Z; \sigma, 0]^{m \times m}$ *and* $\mathbf{T}(Z) \in \mathbb{D}[Z; \sigma, 0]^{m \times s}$*;*
*(b)* $\mathbf{U}(Z)$ *is unimodular;*
*(c)* $\mathbf{U}(Z)\,\mathbf{F}(Z) = \mathbf{T}(Z)$*;*
*(d)* $\mathbf{T}(Z)$ *is in weak Popov form.*

*Proof:* Part (a) is clear, and (b) follows from the fact that $\mathbf{U}(Z)$ can be obtained from $\mathbf{U}^{[\kappa s - s]}(Z)$ by applying elementary row operations of the second type on each row until it has been chosen as a pivot. Moreover, we have that for all $k$ and $\ell$, $\mathbf{U}^{[k]}(Z)_{\ell,*}\,\mathbf{F}(Z) = \mathbf{T}^{[k]}(Z)_{\ell,*}$ and hence (c) is true.

Let $H^{[k]}$ be the coefficient of $\hat{Z}^{(\kappa-1)\cdot\vec{e}}$ of $\hat{\mathbf{M}}^{[k]}(\hat{Z})\,\hat{\mathbf{F}}(\hat{Z})$ for $\kappa s - s \leq k \leq \kappa s$. Since $\hat{\mathbf{M}}^{[k]}(\hat{Z})$ is an order basis, it follows that the first $k - (\kappa s - s)$ columns of $H^{[k]}$ are zero. If $\pi^{[k]} = i$, then we have $H^{[k]}_{i,k-(\kappa s-s)+1} \neq 0$. Furthermore, if $i \neq \pi^{[k]}$ for any $\kappa s - s \leq k < \kappa s$ , $H^{[\kappa s]}_{i,*}$ must be zero. Therefore, if we form the matrix $H$ by

$$H_{i,j} = \begin{cases} H^{[k]}_{i,j} & \text{if } \pi^{[k]} = i \text{ for some } \kappa s - s \leq k < \kappa s \\ H^{[\kappa s]}_{i,j} & \text{otherwise,} \end{cases} \tag{33}$$

then the nonzero rows of $H$ form a matrix in upper echelon form (up to a permutation of rows). By reversing the coefficients of $\mathbf{T}(Z)$ we see that

$$\mathbf{T}(Z)_{i,*} = \sigma^{\mu_i^{[\kappa s - s]}}(H_{i,*})\, Z^{\mu_i^{[\kappa s - s]} + N - \kappa} + \text{lower degree terms.}$$

Thus, $L(\mathbf{T}(Z)) = \sigma^{\deg \mathbf{T}(Z) - N + \kappa}(H)$. Since $\sigma$ is an automorphism on $\mathbb{Q}_{\mathbb{D}}$ it follows that the nonzero rows of $L(\mathbf{T}(Z))$ is in upper echelon form and hence $\mathbf{T}(Z)$ is in weak Popov form.

Recall from Theorem A.2 that the multipliers of Theorem 9.1 and of Theorem 9.3 both provide a basis of the left nullspace of $\mathbf{F}(Z)$.

Using the preceding algorithm for row reduction allows us to compute a greatest common right divisor (GCRD) and a least common left multiple (LCLM) of matrices of skew polynomials in the same way it is done in the case of matrices of polynomials [Beckermann and Labahn, 2000, Kailath, 1980]. Let $\mathbf{A}(Z) \in \mathbb{D}[Z;\sigma,0]^{m_1\times s}$ and $\mathbf{B}(Z) \in \mathbb{D}[Z;\sigma,0]^{m_2\times s}$, such that the matrix

$$\mathbf{F}(Z) = \begin{bmatrix}\mathbf{A}(Z)\\ \mathbf{B}(Z)\end{bmatrix}$$

has rank $s$. Such an assumption is natural since otherwise we may have GCRDs of arbitrarily high degree [Kailath, 1980, page 376]. After row reduction and possibly a permutation of the rows, we obtain

$$\mathbf{U}(Z)\,\mathbf{F}(Z) = \begin{bmatrix}\mathbf{U}_{11}(Z)\ \mathbf{U}_{12}(Z)\\ \mathbf{U}_{21}(Z)\ \mathbf{U}_{22}(Z)\end{bmatrix}\cdot\begin{bmatrix}\mathbf{A}(Z)\\ \mathbf{B}(Z)\end{bmatrix} = \begin{bmatrix}\mathbf{G}(Z)\\ 0\end{bmatrix} \tag{34}$$

with $\mathbf{G}(Z) \in \mathbb{D}[Z;\sigma,0]^{s\times s}$, and $\mathbf{U}_{1,j}(Z)$, $\mathbf{U}_{2,j}(Z)$ matrices of skew polynomials of size $s\times m_j$, and $(m_1+m_2-s)\times m_j$, respectively. Standard arguments (see, for example, Kailath [1980]) show that $\mathbf{G}(Z)$ is a GCRD of $\mathbf{A}(Z)$ and $\mathbf{B}(Z)$. Furthermore, for any common left multiple $\mathbf{V}_1(Z)\,\mathbf{A}(Z) = \mathbf{V}_2(Z)\,\mathbf{B}(Z)$ we see that the rows of $\begin{bmatrix}\mathbf{V}_1(Z)\ -\mathbf{V}_2(Z)\end{bmatrix}$ belong to the left nullspace $\mathcal{N}_{\mathbf{F}(Z)}$. Since $\begin{bmatrix}\mathbf{U}_{21}(Z)\ \mathbf{U}_{22}(Z)\end{bmatrix}$ is a basis of $\mathcal{N}_{\mathbf{F}(Z)}$ by Theorem A.2, there exists $\mathbf{Q}(Z) \in \mathbb{Q}_{\mathbb{D}}[Z;\sigma,0]^{(m_1+m_2-s)\times(m_1+m_2-s)}$ such that

$$\begin{bmatrix}\mathbf{V}_1(Z)\ -\mathbf{V}_2(Z)\end{bmatrix} = \mathbf{Q}(Z)\ \begin{bmatrix}\mathbf{U}_{21}(Z)\ \mathbf{U}_{22}(Z)\end{bmatrix},$$

implying that $\mathbf{U}_{21}(Z)\,\mathbf{A}(Z)$ and $-\mathbf{U}_{22}(Z)\,\mathbf{B}(Z)$ are LCLMs of $\mathbf{A}(Z)$ and $\mathbf{B}(Z)$.

In contrast to the method proposed in Beckermann and Labahn [2000], our GCRD has the additional property of being row-reduced or being in weak Popov form.

## *9.4 Computation of Subresultants*

The method of Section 9.3, applied to two $1\times 1$ matrices, gives the GCRD and LCLM of two skew polynomials $a(Z)$ and $b(Z)$. In this subsection we examine the relationship of the intermediate results obtained during our algorithm to

the subresultants of skew polynomials defined by Li [1996, 1998]. Denoting the degrees of $a(Z), b(Z)$ by $d_a \geq d_b$, the $j$-th subresultant $\operatorname{sres}_j(a,b)$ for skew polynomials is defined by taking the determinant of the matrix

$$\left[\begin{array}{ccccccc}
\sigma^{d_b-j-1}(a_{d_a}) & \cdots & \cdots & \cdots & \sigma^{d_b-j-1}(a_{2j+2-d_b}) & Z^{d_b-j-1}a(Z) \\
& \ddots & & & \vdots & \vdots \\
& & \sigma(a_{d_a}) & \cdots & \sigma(a_j) & Za(Z) \\
& & a_{d_a} & \cdots & a_{j+1} & a(Z) \\
\hline
\sigma^{d_a-j-1}(b_{d_b}) & \cdots & \cdots & \cdots & \sigma^{d_a-j-1}(b_{2j+2-d_a}) & Z^{d_a-j-1}b(Z) \\
& \ddots & & & \vdots & \vdots \\
& & \sigma(b_{d_b}) & \cdots & \sigma(b_j) & Zb(Z) \\
& & b_{d_b} & \cdots & b_{j+1} & b(Z)
\end{array}\right].$$

**Theorem 9.4** *Let $a(Z)$ and $b(Z)$ be two skew polynomials of degrees $d_a$ and $d_b$, respectively, such that $d_a \geq d_b$. Then $\operatorname{sres}_j(a,b) \neq 0$ if and only if there exists an $\ell = \ell_j$ with $\vec{\mu}^{[2d_a-2j-1]} = (d_a - j, d_a - j) - \vec{e}_\ell$. In this case,*

$$\mathbf{T}^{[2d_a-2j-1]}(Z)_{\ell,1} = \pm\gamma \operatorname{sres}_j(a,b), \qquad \gamma = \prod_{i=0}^{d_a-d_b-1} \sigma^{d_b-j+i}(a_{d_a}).$$

*In other words, $\operatorname{sres}_j(a,b) \neq 0$ if and only if the FFreduce algorithm computes an order basis of degree $(d_a - j - 1, d_a - j)$ or $(d_a - j, d_a - j - 1)$ as an intermediate result.*

*Proof:* After expanding with respect to the first $d_a - d_b$ columns of the matrix

$$\left[\begin{array}{ccccccc}
\sigma^{d_a-j-1}(a_{d_a}) & & & \cdots & \cdots & \sigma^{d_a-j-1}(a_{2j+2-d_a}) & Z^{d_a-j-1}a(Z) \\
& \ddots & & & & \vdots & \vdots \\
& & \sigma(a_{d_a}) & \cdots & \cdots & \sigma(a_j) & Za(Z) \\
& & & a_{d_a} & \cdots & a_{j+1} & a(Z) \\
\hline
& \sigma^{d_a-j-1}(b_{d_b}) & & \cdots & \cdots & \sigma^{d_a-j-1}(b_{2j+2-d_a}) & Z^{d_a-j-1}b(Z) \\
& & \ddots & & & \vdots & \vdots \\
& & & \sigma(b_{d_a}) & \cdots & \sigma(b_j) & Zb(Z) \\
& & & & b_{d_b} \cdots & b_{j+1} & b(Z)
\end{array}\right],$$

we see that the determinant coincides with the quantity $\gamma \operatorname{sres}_j(a,b)$. Denote by $S_j$ the matrix of size $(2d_a - 2j) \times (2d_a - 2j - 1)$ obtained by dropping the

last column, and notice that

$$\sigma^{-(d_a-j-1)}(S_j) = K((d_a-j, d_a-j), (2d_a-2j-1)), \tag{35}$$

the Krylov matrix associated to $\hat{\mathbf{F}}(\hat{Z}) = (\hat{a}(\hat{Z}), \hat{b}(\hat{Z}))^T$, $\hat{a}(\hat{Z}) = a(\hat{Z}^{-1})\,\hat{Z}^{d_a}$, and $\hat{b}(\hat{Z}) = b(\hat{Z}^{-1})\,\hat{Z}^{d_a}$. Thus $\operatorname{sres}_j(a,b) \neq 0$ if and only if the dimension (over $\mathbb{Q}_{\mathbf{D}}$) of the left nullspace of $S_j$ is equal to one, which in turn is true if and only if there is a unique $\mathbf{P} \in \mathbb{Q}_{\mathbf{D}}[Z;\sigma,0]$ (up to multiplication with an element from $\mathbb{Q}_{\mathbf{D}}$) of order $\vec{\omega} = (2d_a-2j-1)$ and $\deg \mathbf{P} \leq d_a-j-1$.

One verifies using [Beckermann et al., 2002, Lemma 5.2] and the relation $d_a \neq 0$ that $|\vec{\omega}^{[k]}| = k = |\vec{\mu}^{[k]}|$ for all $k$ in algorithm FFreduce. Let $k = 2d_a-2j-1$, then from (2) we conclude that $\operatorname{sres}_j(a,b) \neq 0$ if and only if $\vec{\mu}^{[k]}$ has one component being equal to $d_a-j-1$ and the other one being at least as large as $d_a-j$, that is, $\vec{\mu}^{[k]} = (d_a-j, d_a-j) - \vec{e}_\ell$ for some $\ell \in \{1,2\}$.

Finally, if $\operatorname{sres}_j(a,b) \neq 0$, then we use (35) and the determinant representations of Section 5 together with the uniqueness of Mahler systems in order to conclude that

$$\gamma \operatorname{sres}_j(a,b) = \pm Z^{\mu_\ell}\, \hat{\mathbf{R}}^{[k]}(\hat{Z})_{\ell,*}\, \hat{Z}^{\vec{\omega}-d_a\cdot\vec{e}} = \mathbf{T}^{[k]}(Z)_{\ell,1}.$$

Thus, whenever $\vec{\mu}^{[2k-1]}$ is of the form $(k,k) - \vec{e}_\ell$ for some $\ell \in \{1,2\}$ during the execution of our algorithm, we can recover the nonzero $\operatorname{sres}_{d_a-k}(a,b)$ from $\hat{\mathbf{R}}^{[2k-1]}(\hat{Z})\, Z^{\vec{\omega}-d_a\cdot\vec{e}}$ after multiplying by $Z^k$ and dividing by the extra factor of $\gamma$ (or by dividing $\mathbf{T}^{[2k-1]}(Z)_{\ell,1}$ by $\gamma$).

Notice that the extra factor of $\gamma$ is introduced at the beginning of the algorithm, before any step with $|\Lambda| > 1$. There is no reduction performed in these first $d_a - d_b$ steps. Thus, we may modify our algorithm so that no reduction is done until $|\Lambda| = 2$ for the first time, except the updating of $\vec{\mu}^{[k]}$. Then

$$\operatorname{sres}_{d_a-k}(a,b) = \begin{cases} \pm Z^{\mu_1^{[2k-1]}-d_a+d_b}\, \hat{\mathbf{R}}^{[2k-1]}(\hat{Z})_{1,1}\, \hat{Z}^{2k-1-d_a} & \text{if } \vec{\mu}^{[2k-1]} = (k-1,k), \\ \pm Z^{\mu_2^{[2k-1]}}\, \hat{\mathbf{R}}^{[2k-1]}(\hat{Z})_{2,1}\, \hat{Z}^{2k-1-d_a} & \text{if } \vec{\mu}^{[2k-1]} = (k,k-1). \end{cases}$$

## 10 Conclusion

In this paper we have given a fraction-free algorithm for transforming a given matrix of Ore polynomials into one where both the rank and the left nullspace is easily determined. The algorithm is a modification of the FFFG algorithm of Beckermann and Labahn [2000] in the commutative case. In the case of skew polynomials we also show how our approach can be used to find a weak Popov

form of a matrix of skew polynomials. In addition, in the special case of $2 \times 1$ skew polynomial matrices we relate our algorithm along with the intermediate quantities to the classical subresultants typically used for one sided GCD and LCM computations.

There are a number of topics for future research. In this paper we have given a fraction-free method for elimination of low order terms of a matrix of Ore polynomials. However for general Ore domains it appears to be more useful to work with leading coefficients, something not possible using our methods except for the case of skew-polynomial domains. We note that this is possible to do using the approach of Abramov and Bronstein simply by using Theorem 2.2. In our case we would like to find a fraction-free method for such a reduction over Ore domains. We will look at combining the procedure in Theorem 2.2 along with modified Schur complements [Beckermann et al., 1997] of Krylov matrices to help us develop such an algorithm.

In a recent work Abramov and Bronstein [2002] extend their results to handle the case of nested skew Ore domains, allowing for computations for example in Weyl algebras. We would like to extend our methods to this important class of matrices again with the idea of controlling the growth of the resulting matrices. This is a difficult extension to do using the methods described in our paper since the corresponding associated linear systems do not have commutative elements. As such the standard tools that we use from linear algebra, namely determinants and Cramer's rule, do not exist in the classical sense.

Finally, it is well known that modular algorithms improve on fraction-free methods by an order of magnitude. We plan to investigate such algorithms for our rank and left nullspace computations. We note that the determinantal representations gives a first step in this direction since it provides an upper bound for the sizes of the objects which need to be computed. As in the modular algorithm for computing a GCRD of Ore polynomials [Li, 1996, Li and Nemes, 1997], we expect that the fraction-free algorithm would be a basis for the modular algorithm.

## References


S. Abramov. EG-eliminations. *Journal of Difference Equations and Applications*, 5(393–433), 1999.

S. Abramov and M. Bronstein. On solutions of linear functional systems. In *Proceedings of the 2001 International Symposium on Symbolic and Algebraic Computation*, pages 1–6. ACM, 2001.

S. Abramov and M. Bronstein. Linear algebra for skew-polynomial matrices. Technical Report RR-4420, Rapport de Recherche INRIA, 2002.

E. Bareiss. Sylvester's identity and multistep integer-preserving Gaussian elimination. *Mathematics of Computation*, 22(103):565–578, 1968.

B. Beckermann, S. Cabay, and G. Labahn. Fraction-free computation of matrix Padé systems. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, pages 125–132. ACM, 1997.

B. Beckermann, H. Cheng, and G. Labahn. Fraction-free row reduction of matrices of skew polynomials. In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, pages 8–15. ACM, 2002.

B. Beckermann and G. Labahn. Recursiveness in matrix rational interpolation problems. *Journal of Computational and Applied Mathematics*, 77:5–34, 1997.

B. Beckermann and G. Labahn. Fraction-free computation of matrix GCD's and rational interpolants. *SIAM Journal on Matrix Analysis and Applications*, 22(1):114–144, 2000.

B. Beckermann, G. Labahn, and G. Villard. Normal forms for general polynomial matrices. *To appear in Journal of Symbolic Computation*, 2001.

P. M. Cohn. *Free Rings and Their Relations*. Academic Press, 1971.

T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

Z. Li. *A Subresultant Theory for Linear Differential, Linear Difference and Ore Polynomials, with Applications*. PhD thesis, Johannes Kepler University Linz, Austria, 1996.

Z. Li. A subresultant theory for Ore polynomials with applications. In *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation*, pages 132–139. ACM, 1998.

Z. Li. Private communication, 2003.

Z. Li and I. Nemes. A modular algorithm for computing greatest common right divisors of Ore polynomials. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, pages 282–289. ACM, 1997.

T. Mulders and A. Storjohann. On lattice reduction for polynomial matrices. *Journal of Symbolic Computation*, 35(4):377–401, 2003.

## A Appendix: Further Facts on Matrices of Ore Polynomials

In this appendix we will present a number of technical results that are needed in our paper. These results are typically well understood in the context of commutative matrix polynomials but are not at all obvious for the case of noncommutative matrices of Ore polynomials.

Consider first the notion of the rank of a matrix of Ore polynomials, $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$. Denote by $\mathcal{M}_{\mathbf{F}(Z)} = \{\mathbf{Q}(Z)\mathbf{F}(Z) : \mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}\}$ the submodule of the (left) $\mathbb{K}[Z;\sigma,\delta]$-module $\subset \mathbb{K}[Z;\sigma,\delta]^{1\times s}$ obtained by forming left linear combinations of the rows of $\mathbf{F}(Z)$. Then following [Cohn,

1971, p. 28, Section 0.6], the rank of a module $\mathcal{M}$ over $\mathbb{K}[Z;\sigma,\delta]$ is defined to be the cardinality of a maximal $\mathbb{K}[Z;\sigma,\delta]$-linearly independent subset of $\mathcal{M}$. Comparing with our Definition 2.1, we see that $\operatorname{rank}\mathbf{F}(Z) \leq \operatorname{rank}\mathcal{M}_{\mathbf{F}(Z)}$. Theorem A.2 below shows that in fact we have equality.

Notice that for any $\mathbf{A}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ we have that $\mathcal{M}_{\mathbf{A}(Z)\mathbf{F}(Z)} \subset \mathcal{M}_{\mathbf{F}(Z)}$. If now $\mathbf{A}(Z)$ has a left inverse $\mathbf{V}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$, then we also have the inclusions $\mathcal{M}_{\mathbf{F}(Z)} = \mathcal{M}_{\mathbf{V}(Z)\mathbf{A}(Z)\mathbf{F}(Z)} \subset \mathcal{M}_{\mathbf{A}(Z)\mathbf{F}(Z)}$, showing that in this case $\mathcal{M}_{\mathbf{A}(Z)\mathbf{F}(Z)} = \mathcal{M}_{\mathbf{F}(Z)}$.

For identifying the different concepts of rank, it will be useful to show that the rows of a row-reduced matrix of Ore polynomials are linearly independent over $\mathbb{K}[Z;\sigma,\delta]$. This however is an immediate consequence of Lemma A.1(a) below which in case of ordinary matrix polynomials is referred to as the *predictable degree property* (see Kailath [1980], Theorem 6.3.13).

**Lemma A.1** *Let $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$, with $\vec{\mu} = \text{row-deg}\,\mathbf{F}(Z)$.*

***(a)*** *$\mathbf{F}(Z)$ is row-reduced if and only if, for all $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$,*

$$\deg \mathbf{Q}(Z)\mathbf{F}(Z) = \max_j(\mu_j + \deg \mathbf{Q}(Z)_{1,j}).$$

***(b)*** *Let $\mathbf{A}(Z) = \mathbf{B}(Z)\,\mathbf{C}(Z)$ be matrices of Ore polynomials of sizes $m \times s$, $m \times r$, and $r \times s$, respectively. Then $\operatorname{rank}\mathbf{A}(Z) \leq r$.*

***(c)*** *Let $\mathbf{A}(Z) = \mathbf{B}(Z)\,\mathbf{C}(Z)$ be as in part (b), with $\mathbf{A}(Z)$ and $\mathbf{C}(Z)$ row-reduced with row degrees $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_m$ and $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_r$, respectively. Then $m \leq r$, and $\alpha_j \geq \gamma_j$ for $j = 1, \ldots, m$.*

***(d)*** *Let $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{S}(Z)$, with $\mathbf{U}(Z)$ unimodular and with both $\mathbf{S}(Z)$ and $\mathbf{T}(Z)$ row-reduced. Then, up to permutation, the row degrees of $\mathbf{S}(Z)$ and $\mathbf{T}(Z)$ coincide.*

*Proof:* For any $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ let $N' := \max_j\left(\mu_j + \deg \mathbf{Q}(Z)_{1,j}\right)$ and define the vector $\vec{h} \in \mathbb{K}^{1\times m}$, $\vec{h} \neq \vec{0}$, by

$$\mathbf{Q}(Z)_{1,j} = h_j Z^{N'-\mu_j} + \text{ lower degree terms.}$$

Clearly, $\deg \mathbf{Q}(Z)\,\mathbf{F}(Z) \leq N'$, with the coefficient at $Z^{N'}$ being given by

$$\sum_{j=1}^{m} h_j \sigma^{N'-\mu_j}(F_{j,*}^{(\mu_j)}) = \vec{h}\,\sigma^{N'-N}(L(\mathbf{F}(Z))).$$

Since $\sigma$ is injective, we have that $\mathbf{F}(Z)$ is row-reduced if and only if $\sigma^j(L(\mathbf{F}(Z)))$ is of full row rank for any integer $j$ that is, if and only if $h\sigma^j(L(\mathbf{F}(Z))) \neq 0$ for all $h \neq 0$ and all integers $j$. This in turn holds true if and only if $\deg \mathbf{Q}(Z)\mathbf{F}(Z) = N'$ for any $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$.

In order to show (b), we may suppose by eliminating a suitable number of rows of $\mathbf{A}(Z)$ and $\mathbf{B}(Z)$ that $\operatorname{rank}\mathbf{A}(Z) = m$. If $r < m$, then $\mathcal{M}_{\mathbf{B}(Z)} \subset \mathbb{K}[Z;\sigma,\delta]^{1\times r}$, the latter $\mathbb{K}[Z;\sigma,\delta]$-module being of rank $r$. Hence $r \geq \operatorname{rank}\mathcal{M}_{\mathbf{B}(Z)} \geq \operatorname{rank}\mathbf{B}(Z)$. On the other hand, $\mathbf{B}(Z)$ has more rows than columns. Thus, by definition of $\operatorname{rank}\mathbf{B}(Z)$ there exists a nontrivial $\mathbf{Q}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ with $\mathbf{Q}(Z)\mathbf{B}(Z) = \mathbf{0}$. Thus $\mathbf{Q}(Z)\mathbf{A}(Z) = \mathbf{0}$, a contradiction to the fact that $\mathbf{A}(Z)$ has full row rank $m$. Therefore $r \geq m$, as claimed in part (b).

For a proof of part (c), recall first that the rows of the row-reduced $\mathbf{A}(Z)$ are $\mathbb{K}[Z;\sigma,\delta]$-linearly independent by part (a), and hence $m = \operatorname{rank}\mathbf{A}(Z) \leq r$ by part (b). Suppose that $\alpha_j \geq \gamma_j$ for $j < k$, but $\alpha_k < \gamma_k$. Part (a) tells us that $\deg\mathbf{B}(Z)_{j,\ell} \leq \alpha_j - \gamma_\ell$. Since $\alpha_j < \gamma_k \leq \gamma_\ell$ for $j \leq k \leq \ell$, we may conclude that $\mathbf{B}(Z)_{j,\ell} = 0$ for $j \leq k \leq \ell$, in other words, the first $k$ rows of $\mathbf{A}(Z)$ are left polynomial combinations of the first $k-1$ rows of $\mathbf{C}(Z)$. Again from part (b) it follows that the first $k$ rows of $\mathbf{A}(Z)$ are $\mathbb{K}[Z;\sigma,\delta]$-linearly dependent, a contradiction. Hence the assertion of part (c) holds.

Finally, part (d) is obtained by twice applying part (c) (compare with [Kailath, 1980, Lemma 6.3.14, p.388] for the case of ordinary matrix polynomials).

Consider now the left nullspace $\mathcal{N}_{\mathbf{F}(Z)}$ of a $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$. Clearly, $\mathcal{N}_{\mathbf{F}(Z)}$ is a $\mathbb{K}[Z;\sigma,\delta]$-module. We want to construct a row-reduced basis of this space, and obtain information about the degrees of such a basis.

**Theorem A.2** *Let $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$, and $\mathbf{U}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ be unimodular, with $\mathbf{T}(Z) = \mathbf{U}(Z)\,\mathbf{F}(Z)$ having $r$ nonzero rows, where the submatrix consisting of the $r$ nonzero rows of $\mathbf{T}(Z)$ are row-reduced. Then*

$$r = \operatorname{rank}\mathcal{M}_{\mathbf{F}(Z)} = \operatorname{rank}\mathbf{F}(Z) = m - \operatorname{rank}\mathcal{N}_{\mathbf{F}(Z)}, \tag{A.1}$$

*with a basis over $\mathbb{K}[Z;\sigma,\delta]$ of $\mathcal{N}_{\mathbf{F}(Z)}$ given by those rows of $\mathbf{U}(Z)$ corresponding to the zero rows of $\mathbf{T}(Z)$.*

*Moreover, there exists a row-reduced $\mathbf{W}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{(m-r)\times m}$ with rows forming a basis of the left nullspace $\mathcal{N}_{\mathbf{F}(Z)}$, and*

$$\text{row-deg}\,\mathbf{W}(Z) \leq (m-1)N\vec{e}, \qquad N = \deg\mathbf{F}(Z).$$

*Proof:* Denote by $J$ the set of indices of zero rows of $\mathbf{T}(Z)$, and define the matrix $\mathbf{U}(Z)_{J,*}$ by extracting from $\mathbf{U}(Z)$ the rows with indices in $J$. In a first step, let us determine the left nullspace of $\mathbf{T}(Z)$, and establish equality (A.1) for the matrix $\mathbf{T}(Z)$. For some $\mathbf{P}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{1\times m}$ we have

$$\mathbf{P}(Z)\mathbf{T}(Z) = \sum_{j\notin J} \mathbf{P}(Z)_{1,j}\mathbf{T}(Z)_{j,*}.$$

We have shown implicitly in Lemma A.1(a) that the rows $\mathbf{T}(Z)_{j,*}$ for $j \notin J$ are linearly independent over $\mathbb{K}[Z;\sigma,\delta]$. Therefore $\mathbf{P}(Z) \in \mathcal{N}_{\mathbf{T}(Z)}$ if and only if $\mathbf{P}(Z)_{1,j} = 0$ for all $j \notin J$, and in addition

$$r = \operatorname{rank} \mathbf{T}(Z) = m - \operatorname{rank} \mathcal{N}_{\mathbf{T}(Z)}.$$

As mentioned before, we also have that $\operatorname{rank} \mathbf{T}(Z) \leq \operatorname{rank} \mathcal{M}_{\mathbf{T}(Z)} =: \rho$. Suppose that there is strict inequality. Then there exist $\rho$ elements of $\mathcal{M}_{\mathbf{T}(Z)}$ which are $\mathbb{K}[Z;\sigma,\delta]$-linearly independent and which can be written as rows of the matrix $\mathbf{B}(Z)\mathbf{T}(Z)$ for some $\mathbf{B}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{\rho\times m}$. Then $\operatorname{rank} \mathbf{B}(Z)\mathbf{T}(Z) = \rho$ by construction of $\mathbf{B}(Z)$. However $\mathbf{T}(Z)$ contains only $r$ rows different from zero, and hence $\operatorname{rank} \mathbf{B}(Z)\mathbf{T}(Z) \leq r$ by Lemma A.1(b), a contradiction. Consequently, (A.1) holds for the matrix $\mathbf{F}(Z)$ being replaced by $\mathbf{T}(Z)$.

We now use the fact that $\mathbf{U}(Z)$ is unimodular, that is, there exists a $\mathbf{V}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times m}$ with $\mathbf{V}(Z)\,\mathbf{U}(Z) = \mathbf{U}(Z)\,\mathbf{V}(Z) = \mathbf{I}$. Consequently, $\mathbf{Q}(Z) \in \mathcal{N}_{\mathbf{F}(Z)}$ if and only if $\mathbf{P}(Z) = \mathbf{Q}(Z)\,\mathbf{V}(Z) \in \mathcal{N}_{\mathbf{T}(Z)}$, that is,

$$\mathcal{N}_{\mathbf{F}(Z)} = \{\mathbf{P}(Z)\,\mathbf{U}(Z) : \mathbf{P}(Z)_{1,j} = 0 \text{ for } j \notin J\} = \mathcal{M}_{\mathbf{U}(Z)_{J,*}}.$$

Since $\mathbf{U}(Z)$ has a right inverse, we may conclude that $\mathcal{N}_{\mathbf{U}(Z)} = \{0\}$, showing that rows of unimodular matrices are linearly independent over $\mathbb{K}[Z;\sigma,\delta]$. Thus the rows of $\mathbf{U}(Z)_{J,*}$ form a basis of $\mathcal{N}_{\mathbf{F}(Z)}$, and

$$m - \operatorname{rank} \mathcal{M}_{\mathbf{F}(Z)} = m - \operatorname{rank} \mathcal{M}_{\mathbf{T}(Z)} = m - r = \operatorname{rank} \mathcal{N}_{\mathbf{F}(Z)}.$$

Since again the relation $\rho := \operatorname{rank} \mathbf{F}(Z) \leq \operatorname{rank} \mathcal{M}_{\mathbf{F}(Z)}$ is trivial, for a proof of the first part of the assertion of Theorem A.2 it only remains to show that $\rho < r$ leads to a contradiction. Suppose without loss of generality that the first $\rho$ rows of $\mathbf{F}(Z)$ are linearly independent. Then, by maximality of $\rho$, we find for any $j = \rho+1, ..., m$ quantities $\mathbf{Q}(Z)_{j,k} \in \mathbb{K}[Z;\sigma,\delta]$ with

$$\mathbf{Q}(Z)_{j,j} \neq 0, \qquad \mathbf{Q}(Z)_{j,j}\mathbf{F}(Z)_{j,*} + \sum_{k=1}^{\rho} \mathbf{Q}(Z)_{j,k}\mathbf{F}(Z)_{k,*} = 0,$$

that is, we have found $m - \rho > m - r$ many $\mathbb{K}[Z;\sigma,\delta]$-linearly independent elements of $\mathcal{N}_{\mathbf{F}(Z)}$, in contradiction to our previous findings on $\operatorname{rank} \mathcal{N}_{\mathbf{F}(Z)}$.

In order to show the second part of Theorem A.2, suppose that $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ are those defined in Theorem 2.2. Let $\mathbf{W}(Z)$ be the row-reduced counterpart of $\mathbf{U}(Z)_{J,*}$ obtained by applying Theorem 2.2. Since one is obtained from the other by multiplying on the left by some unimodular factor, the rows of $\mathbf{W}(Z)$ form a row-reduced basis of $\mathcal{N}_{\mathbf{F}(Z)}$, with $\text{row-deg}\, \mathbf{W}(Z) \leq \text{row-deg}\, \mathbf{U}(Z)_{J,*}$. Hence it only remains to recall the bound for the row-degree of the multiplier $\mathbf{U}(Z)$ of Theorem 2.2: we have for $j \in J$

$$\deg \mathbf{U}(Z)_{j,*} \leq \nu_j - \mu_j + (|\vec{\mu}| - |\vec{\nu}|) \leq |\vec{\mu}| - \vec{\mu}_j \leq (m-1)N.$$

We should mention that the quantity row-deg $\mathbf{W}(Z)$ of Theorem A.2 is an invariant of $\mathbf{F}(Z)$ since by Lemma A.1(d), we obtain the same degrees (up to permutation) for any row-reduced basis of the left nullspace of $\mathbf{F}(Z)$. In the case of ordinary matrix polynomials, the components of row-deg $\mathbf{W}(Z)$ are usually referred to as *left minimal indices* or *left Kronecker indices*, (see §6.5.4, p. 456 of Kailath [1980]).

We conclude this appendix by showing that a certain number of elementary properties of the rank remain equally valid for matrices of Ore polynomials.

**Lemma A.3** *For any $\mathbf{F}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{m\times s}$, the rank of $\mathbf{F}(Z)$ does not change by applying any of the row operations of first or second type described in the introduction, or by multiplying $\mathbf{F}(Z)$ on the right by a full rank square matrix of Ore polynomials.*

*Proof:* Suppose that $\mathbf{A}(Z) \in \mathbb{K}[Z;\sigma,\delta]^{s\times s}$ is of rank $s$. Then $\mathcal{N}_{\mathbf{A}(Z)} = \{0\}$ by (A.1), implying that $\mathcal{N}_{\mathbf{F}(Z)\mathbf{A}(Z)} = \mathcal{N}_{\mathbf{F}(Z)}$. Hence $\mathbf{F}(Z)\mathbf{A}(Z)$ and $\mathbf{F}(Z)$ have the same rank by (A.1). If $\mathbf{U}(Z)$ is unimodular, then $\mathcal{M}_{\mathbf{U}(Z)\,\mathbf{F}(Z)} = \mathcal{M}_{\mathbf{F}(Z)}$, showing that the rank remains the same. Finally we need to examine the row operation of multiplying one row of $\mathbf{F}(Z)$ with a nonzero element of $\mathbb{K}[Z;\sigma,\delta]$. Since $\mathbb{K}[Z;\sigma,\delta]$ contains no zero divisors, it is easy to check that $\mathbf{F}(Z)$ and the new matrix will have the same number of $\mathbb{K}[Z;\sigma,\delta]$-linearly independent rows, and hence the same rank.