Article Submitted to Journal of Symbolic Computation

Fraction-free Row Reduction of Matrices of Ore Polynomials

Bernhard Beckermann[1], Howard Cheng[2] and George Labahn[2]

[1] *Laboratoire de Mathématiques Appliquées FRE 2222 (AN0), Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France. bbecker@ano.univ-lille1.fr*

[2] *School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1. {hchcheng,glabahn}@scg.uwaterloo.ca*

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 central 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 $\mathbb{Q}[Z;\sigma,\delta]$ is a ring of Ore polynomials over a field $\mathbb{Q}$ with σ an automorphism, δ a derivation and where the elements of $\mathbb{Q}$ interact with Z via $Za = \sigma(a)Z + \delta(a)$ (Ore, 1933). Classic examples of such domains include $\mathbb{Q} = \mathbb{K}(n)$ for some field $\mathbb{K}$ with Z the shift operator and $\sigma(a(n)) = a(n+1)$, $\delta = 0$, and $\mathbb{Q} = \mathbb{K}(x)$ with Z the differential operator and $\sigma(a(x)) = a(x)$, $\delta(a(x)) = \frac{d}{dx}a(x)$. It is well known that all Ore rings can be transformed into essentially these two cases (Cohn, 1971, Theorem 8.3.1). These transformations may however introduce fractions.

In this paper we look at matrices over Ore rings and look for methods to easily determine their ranks and their left nullspaces. For a given $m \times s$ matrix $\mathbf{F}(z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$ 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 $\mathbb{Q}[Z;\sigma,\delta]$;

(c) add a polynomial 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 $\mathbb{Q}$. The second set of row operations is useful, for example, when computing a GCRD or a LCLM of Ore polynomials.

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

In the commutative case, the algorithm of Beckermann and Labahn (1997) transforms via row operations a matrix of polynomials into one whose rank is completely determined by the rank of its leading or trailing coefficient. 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.

In the noncommutative case of skew polynomials (i.e. where $\delta = 0$) both the EG-elimination method of Abramov (1999) and the algorithm in Abramov and Bronstein (2001) transform a matrix of skew polynomials into one whose rank is completely determined by the rank of its leading or trailing coefficient. 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 the method of Beckermann and Labahn (1997) to the noncommutative case. However, while these algorithms have good arithmetic complexity, coefficient growth is controlled through coefficient GCD computations. Without the GCD computations the coefficient growth can be exponential. Controlling the coefficient growth in intermediate computations is a central concern in computer algebra computations. This is particularly the case when $\mathbb{Q}$ is the quotient field of an integral domain $\mathbb{D}$.

In this paper we consider the problem of determining the rank and left nullspace of a matrix of Ore polynomials. Our object 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 $\sigma(\mathbb{D}) \subset \mathbb{D}$, $\delta(\mathbb{D}) \subset \mathbb{D}$. 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

extends the algorithm of Beckermann and Labahn (2000b) in the commutative case and Beckermann, Cheng and Labahn (2002) in the case of matrices of skew polynomials. Unlike the skew and commutative polynomial case, the rank is no longer necessarily determined by the rank of the leading or trailing coefficient matrix. The approach taken in this paper follows the previous papers by considering the required transformation matrix as an *order basis*, which represents a basis for the module of vectors of a given order. Equivalently, if we represent row operations as a row vector, an order basis gives invertible row operations that eliminate a certain number of trailing coefficients of our Ore matrix polynomial. By examining the underlying systems of linear equations on the coefficients, we reduce the problem to linear algebra and obtain a fraction-free recursion formula which efficiently predicts divisors in order to compute an order basis of a higher order from one of a lower order.

In the 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. In addition, the methods can be used to compute, in a fraction-free way, a greatest common right divisor (GCRD) and a least common left multiple (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.

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 define order bases while in Section 4 we place such bases in a linear algebra setting. A fraction-free recursion formula for computing order bases is given in Section 5 followed by a discussion of the termination criteria along with the complexity of the algorithm in Section 6. Matrices of skew polynomials are handled in Section 7 where we show that our algorithm can be used to find a weak Popov form of such matrices, 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 2×1 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 $F_j^{k,\ell}$. For any vector of integers (also called multi-index) $\vec{\omega} = (\vec{\omega}^1, \ldots, \vec{\omega}^p)$, we denote

by $|\vec{\omega}| = \sum_{i=1}^{p} \vec{\omega}^i$. We also denote by $Z^{\vec{\omega}}$ the matrix of Ore polynomials having $Z^{\vec{\omega}^i}$ on the diagonal and 0 everywhere else. A matrix of Ore polynomials $\mathbf{F}(Z)$ is said to have row degree $\vec{\nu} = rdeg\mathbf{F}(Z)$ (and column degree $\vec{\mu} = cdeg\mathbf{F}(Z)$, respectively) if the ith row has degree $\vec{\nu}^i$ (and the jth column has degree $\vec{\mu}^j$). The vector $\vec{e}_i$ denotes the vector having 1 in component i and 0 elsewhere and $\vec{e} = (1, \ldots, 1)$.

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 $\vec{\nu} = rdeg\mathbf{F}(Z)$ and $N = \max_j \vec{\nu}^j = \deg \mathbf{F}(Z)$, we may write

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

where the matrix $L(\mathbf{F}(Z)) := L \in \mathbb{Q}^{m\times s}$ is called the *leading coefficient matrix* of $\mathbf{F}(Z)$. In analogy with the case of ordinary matrix polynomials (see for instance (Kailath, 1980, Section 6.3)), $\mathbf{F}(Z)$ is *row-reduced* if $\operatorname{rank} L = m$.

Definition 2.1 (Rank, Unimodular):

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

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

□

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 $\mathbb{Q}(Z;\sigma,\delta)$ of left fractions). 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 a row-reduced one by means of elementary row operations of the second type given in the introduction.

Theorem 2.2: *For any $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$ there exists a unimodular matrix $\mathbf{U}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$, with $\mathbf{T}(Z) = \mathbf{U}(Z)\cdot\mathbf{F}(Z)$ having $r \le \min\{m,s\}$ nonzero rows, $\operatorname{rdeg}\mathbf{T}(Z) \le \operatorname{rdeg}\mathbf{F}(Z)$, 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

$$\mathbf{U}(Z) \leq \vec{\nu} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}|)\vec{e},$$

where $\vec{\mu} := \max(\vec{0}, \operatorname{rdeg} \mathbf{F}(Z))$ *and* $\vec{\nu} := \max(\vec{0}, \operatorname{rdeg} \mathbf{T}(Z))$.

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

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 $v \in \mathbb{Q}^{1\times m}$ with $v \neq 0$, $vL = 0$, and $v^j = 0$ for $j \in J$. Choose an index k with $v^k \neq 0$ (the index of the updated row) and

$$\vec{\nu}^k = \max\{\vec{\nu}^j : v^j \neq 0\},$$

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

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

Hence $\deg \mathbf{T}(Z)^{k,\cdot}_{new} \leq \vec{\nu}^k - 1$, showing that $rdeg\mathbf{T}(Z)_{new} \leq rdeg\mathbf{T}(Z)$. Notice that $\mathbf{U}(Z)_{new} = \mathbf{V}(Z) \cdot \mathbf{U}(Z)$, where $\mathbf{V}(Z)$ is obtained from $\mathbf{I}_m$ by replacing its kth row by $\mathbf{Q}(Z)$. Since $\mathbf{Q}(Z)^{1,k} \in \mathbb{Q} \setminus \{0\}$ 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 $\mathbf{W}(Z) \cdot \mathbf{V}(Z) = \mathbf{V}(Z) \cdot \mathbf{W}(Z) = \mathbf{I}_m$. 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 $\deg(\mathbf{Q}(Z) \cdot \mathbf{U}(Z)) \leq \vec{\nu}^k - \vec{\mu}^k + |\vec{\mu}| - |\vec{\nu}|$. Hence the degree bounds for $\mathbf{U}(Z)_{new}$ are obtained by observing that

$$rdeg\mathbf{U}(Z)_{new} \leq \vec{\nu} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}|)\vec{e} \leq \vec{\nu}_{new} - \vec{\mu} + (|\vec{\mu}| - |\vec{\nu}_{new}|)\vec{e}.$$

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. $\square$

Remark: 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. Unlike the latter work, however, our computations are not done in the algebra of Laurent skew polynomials (which does not seem to be a natural object in case of general Ore domains), and we also give additional degree bounds for the multiplier matrix $\mathbf{U}(Z)$.

Of course, as suggested in Abramov and Bronstein (2001), the vector v in the above algorithm could be chosen in $\mathbb{D}^{1\times m}$ by performing fraction-free Gaussian elimination on L, (see Bareiss (1968)), leading to a fraction-free algorithm for row-reducing a matrix of Ore polynomials. In this case, in order to prevent an exponential growth of coefficients, it would be necessary to remove the content of rows of $(\mathbf{U}(Z), \mathbf{T}(Z))$ during the computations, an operation which could be very expensive. In our case we wish to control coefficient growth by using only predicted factors which require no content computations.

Remark: In the case of commutative polynomials there is an example in (Beckermann, Labahn and Villard, 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 $rdeg\,\mathbf{U}(Z) = (m-1)N\vec{e}$. 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

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

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

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

with $\mathbf{R}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times s}$. *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 linear combinations (over $\mathbb{Q}[Z;\sigma,\delta]$) 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 3.2 (Order Basis): *Let* $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$, *and* $\vec{\omega}$ *be a multi-index. A matrix of Ore polynomials* $\mathbf{M}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$ *is said to be an* order basis *of order* $\vec{\omega}$ *and column degree* $\vec{\mu}$ *if there exists a multi-index* $\vec{\mu} = (\vec{\mu}^1, ..., \vec{\mu}^m)$ *such that*

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

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

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

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

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

where $\deg \mathbf{L}(Z)^{k,\ell} \le \vec{\mu}^\ell - 1$.

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

Part (a) of Definition 3.2 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 $\vec{\mu}^i$ 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 $rdeg\,\mathbf{M}(Z) \le \vec{\mu}$, and we have the added degree constraint in part (b) that, for all $j = 1, ..., m$,

$$\deg \mathbf{Q}(Z)^{1,j} \le \deg \mathbf{P}(Z) - \vec{\mu}^j. \tag{2}$$

We remark that the definition of order basis given in Beckermann, Cheng and Labahn (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, Cheng and Labahn (2002) is not applicable for the new definition of order bases and so we give a new proof here for this result.

THEOREM 3.3: *Let* $\mathbf{M}(Z)$ *be an order basis of order* $\vec{\omega}$ *and degree* $\vec{\mu}$.

(a) There exists only the trivial row vector $\mathbf{P}(Z) = 0$ *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* $\mathbb{Q}$.

(c) An order basis of a particular order and degree is unique up to multiplication by constants from $\mathbb{Q}$.

Proof: We only need to show part (a) as (b) and (c) follow directly from (a). Suppose that $\mathbf{P}(Z) \neq 0$ has order $\vec{\omega}$ and column degree $\vec{\mu} - \vec{e}$. By Definition 3.2(b), there exists $\mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}$ such that $\mathbf{P}(Z) = \mathbf{Q}(Z) \cdot \mathbf{M}(Z)$. Let j be the index such that $\deg \mathbf{Q}(Z)^{1,j}$ is maximum. Since $\mathbf{P}(Z) \neq 0$, it follows that $\deg \mathbf{Q}(Z)^{1,j} \geq 0$. Now,

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

Note that if $k \neq j$, then

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

Also,

$$\deg \mathbf{Q}(Z)^{1,j} \cdot \mathbf{M}(Z)^{j,j} = \deg \mathbf{Q}(Z)^{1,j} + \vec{\mu}^j,$$

so that

$$\deg \mathbf{P}(Z)^{1,j} = \deg \mathbf{Q}(Z)^{1,j} + \vec{\mu}^j \geq \vec{\mu}^j.$$

This contradicts the assumption that $\deg \mathbf{P}(Z)^{1,j} \leq \vec{\mu}^j - 1$. □

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 3.4: *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* $t = d_a - d_b$*,* $b_0^{[t+1]} := \prod_{i=0}^{t} \sigma^i(b_0)$ *and solve*

$$b_0^{[t+1]} \cdot a(Z) = q(Z) \cdot b(Z) + r(Z) \cdot Z^{t+1} \tag{3}$$

with $\deg q(Z) = t$ *and* $\deg r(Z) < d_b$*. Equation (3) corresponds to solving the linear system of equations*

$$b_0^{[t+1]} \cdot [a_0, \ldots, a_t] = [q_0, \ldots, q_t] \begin{bmatrix} b_0 & \sigma(b_1) & \cdots & \sigma^t(b_t) \\ & \sigma(b_0) & & \vdots \\ & & \ddots & \vdots \\ & & & \sigma^t(b_0) \end{bmatrix}, \tag{4}$$

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

$$\mathbf{M}(Z) = \begin{bmatrix} b_0^{[t+1]} & -q(Z) \\ 0 & b_0^{[t+1]} Z^{t+1} \end{bmatrix}$$

we see that

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

where $w(Z) = b_0^{[t+1]} \cdot \sigma^{t+1}(b(Z)) = b_0^{[t+1]} \cdot \sum_{i=0}^{d_b} \sigma^{t+1}(b_i) Z^i$*. Properties (a) and (c) of Definition 3.2 are trivially satisfied by* $\mathbf{M}(Z)$*. Property follows from the linear equations given in the next section.* □

4. Determinantal Representations

We are interested in constructing an algorithm for computing recursively order bases $\mathbf{M}(Z)$ for increasing orders. 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 3.2(c).

Because the order condition in Definition 3.1 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

$$\mathbf{P}(Z) \cdot \mathbf{F}(Z) = \sum_j S_j Z^j \tag{5}$$

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 $\mathbf{P}(Z) \in \mathbb{Q}[Z;\sigma,\delta]$ we may write

$$c_k(Z \cdot \mathbf{P}(Z)) = \sigma(c_{k-1}(\mathbf{P}(Z))) + \delta(c_k(\mathbf{P}(Z))) \tag{6}$$

where c_k denotes the kth coefficient of a polynomial (with $c_{-1} = 0$). We may write (6) 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 μ. Furthermore, for each $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]$ and nonnegative integer μ we associate vectors of coefficients

$$\mathbf{F}_\mu = [c_0(\mathbf{F}(Z)), \ldots, c_{\mu-1}(\mathbf{F}(Z))]^T = [F_0, \ldots, F_{\mu-1}]^T, \tag{7}$$

$$\mathbf{F} = [c_0(\mathbf{F}(Z)), c_1(\mathbf{F}(Z)), \ldots]^T = [F_0, F_1, \ldots]^T. \tag{8}$$

Note that we begin our row and column enumeration at 0. We can interpret (6) in terms of matrices by

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

Comparing with (5), we know that $\mathbf{P}(Z)$ has order $\vec{\omega}$ if and only if for each $\ell = 1, ..., s$, $j = 0, ..., \vec{\omega}^\ell - 1$ we have

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

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

$$(P_0^{1,1}, ..., P_{\vec{\nu}^1}^{1,1}, ..., P_0^{1,m}, ..., P_{\vec{\nu}^m}^{1,m}) \cdot K(\vec{\nu} + \vec{e}, \vec{\omega}) = 0, \tag{9}$$

where the coefficient matrix has the form

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

and $K^{k,\ell}(\vec{\nu}^k + 1, \vec{\omega}^\ell)^T$ may be written as

$$\begin{bmatrix} \mathbf{F}_{\vec{\omega}^\ell}^{k,\ell} & \mathbf{C}_{\vec{\omega}^\ell} \cdot \mathbf{F}_{\vec{\omega}^\ell}^{k,\ell} & \cdots & \mathbf{C}_{\vec{\omega}^\ell}^{\vec{\nu}^k} \cdot \mathbf{F}_{\vec{\omega}^\ell}^{k,\ell} \end{bmatrix}. \tag{10}$$

Thus, the matrix $K(\vec{\nu} + \vec{e}, \vec{\omega})^T$ is in the form of a striped Krylov matrix (Beckermann and Labahn, 2000b), except that by stepping from one column to the next we not only multiply with a lower shift matrix but also apply the functions σ and δ. Thus, in contrast to Beckermann and Labahn (2000b), here we obtain a striped Krylov matrix with a matrix $\mathbf{C}$ having operator-valued elements.

EXAMPLE 4.1: *In the case of matrices of skew polynomials, the $\nu \times \omega$ submatrix $K^{k,\ell}(\nu,\omega)$ is*

$$\begin{bmatrix} \sigma^0(F_0^{k,\ell}) & \sigma^0(F_1^{k,\ell}) & \sigma^0(F_2^{k,\ell}) & \cdots & \cdots & \sigma^0(F_{\omega-1}^{k,\ell}) \\ 0 & \sigma^1(F_0^{k,\ell}) & \sigma^1(F_1^{k,\ell}) & \cdots & \cdots & \sigma^1(F_{\omega-2}^{k,\ell}) \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & \sigma^{\nu-1}(F_0^{k,\ell}) & \cdots & \sigma^{\nu-1}(F_{\omega-\nu}^{k,\ell}) \end{bmatrix}.$$

□

According to (9), it follows from Theorem 3.3 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

$$k = 1, ..., m: \quad \text{rank } K(\vec{\mu},\vec{\omega}) = \text{rank } K(\vec{\mu}+\vec{e}_k,\vec{\omega}) = |\vec{\mu}|. \tag{11}$$

Suppose more generally that $\vec{\mu}$ and $\vec{\omega}$ are multi-indices verifying (11). We call a *multigradient* $d = d(\vec{\mu},\vec{\omega})$ any constant ± 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 kth 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 (9), 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 (11) is equivalent to the system

$$x \cdot K_*(\vec{\mu},\vec{\omega}) = d \cdot b_*^k(\vec{\mu},\vec{\omega}), \tag{12}$$

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, (12) 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

$$\mathbf{M}(Z)^{k,\ell} = \pm \det \left[\; K_*(\vec{\mu}+\vec{e}_k,\vec{\omega}) \;\middle|\; \mathbf{E}_{\ell,\vec{\mu}^\ell-1+\delta_{\ell,k}}(Z) \; \right] \tag{13}$$

with

$$\mathbf{E}_{\ell,\nu}(Z) = [0,\ldots,0|1,Z,\ldots,Z^\nu|0,\ldots,0]^T, \tag{14}$$

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

$$\mathbf{R}(Z)\cdot Z^{\vec{\omega}} = \sum_j \mathbf{M}(Z)^{k,j}\mathbf{F}(Z)^{j,\ell} = \pm\det\left[\ K_*(\vec{\mu}+\vec{e}_k,\vec{\omega})\ \middle|\ \mathbf{E}_{\ell,\vec{\mu}+\vec{e}_k}(Z)\ \right], \quad (15)$$

where

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

In both (13) and (15) the matrices have commutative entries in all but the last column. It is understood that the determinant in both cases is expanded along the last column.

We finally mention that, by the uniqueness result of Theorem 3.3, any order basis of degree $\vec{\mu}$ and order $\vec{\omega}$ coincides up to multiplication with some element in $\mathbb{Q}$ 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.

5. 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 system has coefficients from $\mathbb{Q}$ 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 before, 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, Cheng and Labahn, 2002) and the commutative case (Beckermann and Labahn, 2000b, Theorem 6.1(c)). The resulting order bases have properties similar to those cited by Beckermann and Labahn (2000b, Theorems 7.2 and 7.3).

THEOREM 5.1: *Let $\mathbf{M}(Z)$ be an order basis corresponding of order $\vec{\omega}$ and degree $\vec{\mu}$, and $\lambda\in\{1,...,s\}$. Denote by $r_j = c_{\vec{\omega}^\lambda}((\mathbf{M}(Z)\cdot\mathbf{F}(Z))^{j,\lambda})$, the (j,λ) 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* π *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* $\mathbb{Q}$ *is obtained via the formulas*

$$p_\pi \cdot \widetilde{\mathbf{M}}(Z)^{\ell,k} = r_\pi \cdot \mathbf{M}(Z)^{\ell,k} - r_\ell \cdot \mathbf{M}(Z)^{\pi,k} \tag{16}$$

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

$$\sigma(p_\pi) \cdot \widetilde{\mathbf{M}}(Z)^{\pi,k} = (r_\pi \cdot Z - \delta(r_\pi)) \cdot \mathbf{M}(Z)^{\pi,k} - \sum_{\ell \neq \pi} \sigma(p_\ell) \cdot \widetilde{\mathbf{M}}(Z)^{\ell,k} \tag{17}$$

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

(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 $r_1 = \ldots = r_m = 0$.

Let $\mathbf{M}(Z)^{j,\cdot}$ denote the jth row of $\mathbf{M}(Z)$. For part (b) notice first that rows $\widetilde{\mathbf{M}}(Z)^{\ell,\cdot}$ for $\ell \neq \pi$ have order $\widetilde{\vec{\omega}}$ by construction, as required in Definition 3.2(a). In addition row $(r_\pi \cdot Z - \delta(r_\pi)) \cdot \mathbf{M}(Z)^{\pi,\cdot}$ also has order $\widetilde{\vec{\omega}}$ since $(r_\pi \cdot Z - \delta(r_\pi))(r_\pi) = r_\pi \sigma(r_\pi) Z$. By construction therefore row $\widetilde{\mathbf{M}}(Z)^{\pi,\cdot}$ has order $\widetilde{\vec{\omega}}$.

The verification of the new degree constraints of Definition 3.2(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, 2000b, Theorem 7.2). In addition, notice that p_π is the leading coefficient of $\mathbf{M}(Z)^{\ell,\ell}$, so the leading coefficient of $\widetilde{\mathbf{M}}(Z)^{\ell,\ell}$ equals r_π for all ℓ 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 3.2(b). If $\mathbf{P}(Z) \in \mathbb{Q}[Z; \sigma, \delta]^{1 \times m}$ has order $\widetilde{\vec{\omega}}$ then it has order $\vec{\omega}$ and so there exists a $\mathbf{Q}(Z) \in \mathbb{Q}[Z; \sigma, \delta]^{1 \times m}$ such that

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

Applying the first set of row operations in (16) to rows $\ell \neq \pi$ results in

$$\mathbf{P}(Z) = \sum_{j \neq \pi}^{m} \hat{\mathbf{Q}}(Z)^{1,j} \cdot \widetilde{\mathbf{M}}(Z)^{j,\cdot} + \hat{\mathbf{Q}}(Z)^{1,\pi} \cdot \mathbf{M}(Z)^{\pi,\cdot} \tag{18}$$

where

$$\hat{\mathbf{Q}}(Z)^{1,j} = \mathbf{Q}(Z)^{1,j} \cdot \frac{p_\pi}{r_\pi} \text{ for all } j \neq \pi \text{ and } \hat{\mathbf{Q}}(Z)^{1,\pi} = \sum_{i=0}^{m} \mathbf{Q}(Z)^{1,i} \cdot \frac{r_i}{r_\pi}. \tag{19}$$

Since $\mathbf{P}(Z)$ and all the $\widetilde{\mathbf{M}}(Z)^{j,\cdot}$ terms have order $\widetilde{\vec{\omega}}$ this must also be the case for $\hat{\mathbf{Q}}(Z)^{1,\pi} \cdot \mathbf{M}(Z)^{\pi,\cdot}$. Let ρ be the degree of $\hat{\mathbf{Q}}(Z)$ and write $\hat{\mathbf{Q}}(Z)^{1,\pi} = \sum_{k=0}^{\rho} \hat{Q}_k^{1,\pi}(r_\pi \cdot Z - \delta(r_\pi))^k$. Since $(r_\pi \cdot Z - \delta(r_\pi))r_\pi = r_\pi \sigma(r_\pi) Z$, we see that $\hat{Q}_0^{1,\pi} \cdot r_\pi = 0$. Therefore, by assumption on π we have that $\hat{Q}_0^{1,\pi} = 0$. Writing $\hat{\mathbf{Q}}(Z)^{1,\pi} = \bar{\mathbf{Q}}(Z)^{1,\pi} \cdot (r_\pi \cdot Z - \delta(r_\pi))$ gives

$$\mathbf{P}(Z) = \sum_{j \neq \pi}^{m} \hat{\mathbf{Q}}(Z)^{1,j} \cdot \widetilde{\mathbf{M}}(Z)^{j,\cdot} + \bar{\mathbf{Q}}(Z)^{1,\pi} \cdot (r_\pi \cdot Z - \delta(r_\pi)) \cdot \mathbf{M}(Z)^{\pi,\cdot}. \tag{20}$$

Completing the row operations which normalize the degrees of $\widetilde{\mathbf{M}}(Z)$ in (17) gives a $\widetilde{\mathbf{Q}}(Z)$ with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z) \cdot \widetilde{\mathbf{M}}(Z)$. Consequently, the property of Definition 3.2(b) holds.

Finally, in order to establish part (c) we know already from Section 4 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 (11). By the uniqueness result of Theorem 3.3 we only need to show that the "leading coefficient" $\widetilde{d}$ of $\widetilde{\mathbf{M}}(Z)$ in Definition 3.2(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 (15) for the residual of the πth row of $\mathbf{M}(Z)$ we learn that r_π 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 5.2: *If $\mathbf{M}(Z)$ is a reduced order basis then the order basis $\widetilde{\mathbf{M}}(Z)$ computed by (16) and (17) in Theorem 5.1 is also a reduced order basis of degree $\vec{\nu}$, provided that the pivot π is chosen such that*

$$\vec{\mu}^\pi = \min_j \{\vec{\mu}^j : r_j \neq 0\}. \tag{21}$$

Proof: It is straightforward to check that $rdeg \widetilde{\mathbf{M}}(Z) = \vec{\nu}$. Hence, by Lemma A.1(a), it is sufficient to show that $cdeg \widetilde{\mathbf{Q}}(Z) \leq \deg(\mathbf{P}(Z))\vec{e} - \vec{\nu}$, with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z) \cdot \widetilde{\mathbf{M}}(Z)$ as in the proof of Theorem 5.1.

We see in (19) that $\deg\ \hat{\mathbf{Q}}(Z)^{1,j} \leq \deg\ \mathbf{P}(Z) - \vec{\mu}^j = \deg\ \mathbf{P}(Z) - \vec{\nu}^j$ for all $j \neq \pi$ while $\deg\ \hat{\mathbf{Q}}(Z)^{1,\pi} \leq \deg\ \mathbf{P}(Z) - \vec{\mu}^\pi$ because of the minimality of $\vec{\mu}^\pi$. In (20), $\deg\ \bar{\mathbf{Q}}(Z)^{1,\pi} \leq \deg\ \mathbf{P}(Z) - (\vec{\mu}^\pi + 1) = \deg\ \mathbf{P}(Z) - \vec{\nu}^\pi$. Completing the row operations which normalize the degrees of $\widetilde{\mathbf{M}}(Z)$ in (17) gives a $\widetilde{\mathbf{Q}}(Z)$ with $\mathbf{P}(Z) = \widetilde{\mathbf{Q}}(Z) \cdot \widetilde{\mathbf{M}}(Z)$ having the correct degree bounds. □

6. The FFreduce Algorithm

Theorem 5.1 gives a computational procedure that results in the FFreduce algorithm given in Table 1. In this section we consider the termination criterion for this algorithm and discuss its complexity.

THEOREM 6.1 (TERMINATION OF ALGORITHM FFREDUCE):
Let $r = rank\,\mathbf{F}(Z)$. 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 $\mathbf{M}(Z)^{j,\cdot}$ 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 $\vec{\alpha} = rdeg\,\mathbf{W}(Z)$. Recall from Theorem A.2 that $\mathbf{W}(Z)$ is row-reduced, and that $\vec{\alpha} \le (m-1)\cdot N\vec{e}$. Since the rows of $\mathbf{W}(Z)$ have order $\kappa\vec{e}$, there exists $\mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{(m-r)\times m}$ such that $\mathbf{W}(Z) = \mathbf{Q}(Z)\cdot\mathbf{M}(Z)$. 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 $\vec{\alpha}^j \ge \vec{\mu}^{p(j)}$ for $j = 1,...,m-r$. Hence, for $j = 1,...,m-r$,

$$\begin{aligned}
\deg \mathbf{R}(Z)^{p(j),\cdot} &= -\kappa + \deg(\mathbf{R}(Z)^{p(j),\cdot}Z^{\kappa\vec{e}}) = -\kappa + \deg(\mathbf{M}(Z)^{p(j),\cdot}\cdot\mathbf{F}(Z)) \\
&\le -\kappa + N + \deg(\mathbf{M}(Z)^{p(j),\cdot}) = -\kappa + N + \vec{\mu}^{p(j)} \\
&\le -\kappa + N + \vec{\alpha}^j \le \kappa + mN = -1,
\end{aligned}$$

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

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

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

The relation $r = \operatorname{rank}\mathbf{R}(Z)$ implies that the nonzero rows of $\mathbf{R}(Z)$ are $\mathbb{Q}[Z;\sigma,\delta]$-linearly independent, and hence $\mathbf{Q}(Z)^{1,j} = 0$ for $j \notin J$. Consequently, the rows of $\mathbf{M}(Z)^{J,\cdot}$ form a basis of $\mathcal{N}_{\mathbf{F}(Z)}$, as claimed in Theorem 6.1. □

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 κ (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 $\sigma(a(x)) = a(x)$, $\delta(a(x)) = \frac{d}{dx}a(x)$.

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 $\mathbf{R} \in \mathbb{D}[Z;\sigma,\delta]^{m\times s}$
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 $\pi_k \leftarrow \min\{\ell \in \Lambda : \vec{\mu}_k^\ell = \min_j\{\vec{\mu}_k^j : j \in \Lambda\}\}$.

Calculate for $\ell = 1,..,m$, $\ell \neq \pi_k$: $p_\ell \leftarrow c_{\vec{\mu}_k^\ell - 1}(\mathbf{M}_k^{\pi_k,\ell})$.

Increase order for $\ell = 1,..,m$, $\ell \neq \pi_k$:
$\mathbf{M}_{k+1}^{\ell,\cdot} \leftarrow \frac{1}{d_k}[r_{\pi_k} \cdot \mathbf{M}_k^{\ell,\cdot} - r_\ell \cdot \mathbf{M}_k^{\pi_k,\cdot}]$
$\mathbf{R}_{k+1}^{\ell,\cdot} \leftarrow \frac{1}{d_k}[r_{\pi_k} \cdot \mathbf{R}_k^{\ell,\cdot} - r_\ell \cdot \mathbf{R}_k^{\pi_k,\cdot}]$

Increase order and adjust degree constraints for row π_k:
$\mathbf{M}_{k+1}^{\pi_k,\cdot} \leftarrow \frac{1}{\sigma(d_k)}[(r_{\pi_k} \cdot Z - \delta(r_{\pi_k})) \cdot \mathbf{M}_k^{\pi_k,\cdot} - \sum_{\ell\neq\pi_k} \sigma(p_\ell) \cdot \mathbf{M}_k^{\ell,\cdot}]$
$\mathbf{R}_{k+1}^{\pi_k,\cdot} \leftarrow \frac{1}{\sigma(d_k)}[(r_{\pi_k} \cdot Z - \delta(r_{\pi_k})) \cdot \mathbf{R}_k^{\pi_k,\cdot} - \sum_{\ell\neq\pi_k} \sigma(p_\ell) \cdot \mathbf{R}_k^{\ell,\cdot}]$

Update multigradient, degree and ρ:
$d_{k+1} \leftarrow r_{\pi_k}$, $\vec{\mu}_{k+1} \leftarrow \vec{\mu}_k + \vec{e}_{\pi_k}$, $\rho \leftarrow \rho + 1$
end if

Adjust residual in column λ: 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
$\mathbf{M} \leftarrow \mathbf{M}_k$, $\mathbf{R} \leftarrow \mathbf{R}_k$, $\vec{\mu} \leftarrow \vec{\mu}_k$, $\vec{\omega} \leftarrow \vec{\omega}_k$

Remark: The above proof was based on the estimate $\vec{\alpha}^j \leq (m-1)N$ 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, Labahn and Villard, 2001, Example 5.6) for ordinary matrix polynomials. For applications where a lower bound γ 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 $\vec{\alpha}^j \leq (m-1)N - \gamma$.

Remark: In contrast to the special case of skew polynomials (compare with (Beckermann, Cheng and Labahn, 2002, Lemma 5.2)), the pivots π_k in one cyle 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: 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: If in the cycle starting at $k = \kappa s$ there are only distinct pivots, following (Beckermann, Cheng and Labahn, 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, Cheng and Labahn, 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 ϵ is a positive integer, then we have 2 pivots in all further cycles up to the ϵth 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 (22) 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, Cheng and Labahn (2002) which will be further studied in the next section.

In the remainder of this section we 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 $\mathbb{D}$ satisfies

$$\begin{aligned}
\operatorname{size}(a+b) &= \mathcal{O}(\max(\operatorname{size}(a), \operatorname{size}(b))) \\
\operatorname{size}(a \cdot b) &= \mathcal{O}(\operatorname{size}(a) + \operatorname{size}(b)) \\
\operatorname{cost}(a+b) &= \mathcal{O}(1) \\
\operatorname{cost}(a \cdot b) &= \mathcal{O}(\operatorname{size}(a) \cdot \operatorname{size}(b)),
\end{aligned}$$

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 6.2: *Let K be a bound on the size of the coefficients appearing in $\mathbf{F}(Z)^{j,\cdot}$, $Z \cdot \mathbf{F}(Z)^{j,\cdot}, \ldots, Z^{\vec{\mu}^j} \cdot \mathbf{F}(Z)^{j,\cdot}$ 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 (13) and (15) show that both the Mahler system and the residual can be represented as determinants of a square matrix of order $|\vec{\mu}|$. The coefficient in this matrix are coefficients of $\mathbf{F}(Z)^{k,\cdot}, Z \cdot \mathbf{F}(Z)^{k,\cdot}, \ldots, Z^{\vec{\mu}^k} \cdot \mathbf{F}(Z)^{k,\cdot}$. 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 $\mathcal{O}(|\vec{\mu}|^2 K^2)$. □

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

COROLLARY 6.3: *Let K be a bound on the size of the coefficients appearing in $\mathbf{F}(Z)^{j,\cdot}, Z\cdot\mathbf{F}(Z)^{j,\cdot}, \ldots, Z^{\vec{\mu}^j}\cdot\mathbf{F}(Z)^{j,\cdot}$ 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 Theorem is an immediate consequence of Lemma 6.2 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 $|\vec{\mu}| \leq \min(m,s)(mM+1)$. □

7. 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 σ is the identity then this also gives fraction-free algorithms for ordinary matrix polynomials.

In the case of skew polynomials (Beckermann, Cheng and Labahn, 2002), the termination criterion

$$\rho_{\kappa s} + \text{the number of zero rows in } \mathbf{R}_{\kappa s}(Z) = m \tag{22}$$

allows us to prove (Beckermann, Cheng and Labahn, 2002, Theorem 5.3) that

$$\operatorname{rank}\mathbf{R}_{\kappa s}(0) = \operatorname{rank}\mathbf{R}_{\kappa s}(Z)) = \operatorname{rank}\mathbf{F}(Z), \tag{23}$$

the rank of the trailing coefficient matrix $\mathbf{R}_{\kappa s}(0)$ being defined over the field $\mathbb{Q}$. Moreover (Beckermann, Cheng and Labahn, 2002, Lemma 5.2),

$$\text{the pivots } \pi_k \text{ for } \kappa s - s \leq k < \kappa s \text{ are distinct,} \tag{24}$$

and hence (Beckermann, Cheng and Labahn, 2002, Lemma 5.1 and Lemma 5.2)

$$\rho_{\kappa s} = \operatorname{rank}\mathbf{R}_{\kappa s}(0) = \operatorname{rank}\mathbf{R}_{\kappa s - s}(0). \tag{25}$$

It is also shown implicitly in the proof of (Beckermann, Cheng and Labahn, 2002, Theorem 5.4) that $\kappa \leq m(N+1)$ which has to be compared with the number of cycles, $mN+1$, required by FFreduce. Thus the new termination property (22) essentially does not increase the complexity of algorithm FFreduce, but for many examples may improve the complexity.

7.1. 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, Cheng and Labahn, 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 $\mathbf{S}(Z) \in \mathbb{Q}[Z;\sigma,0]^{m\times m}$ such that

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

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 $\mathbb{Q}[Z;\sigma,0]$.

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

7.2. Row-reduced Form and Weak Popov Form

The FFreduce algorithm can be used for row reduction in the case of matrices of skew polynomials. In particular, given $\mathbf{F}(Z) \in \mathbb{D}[Z;\sigma,0]^{m\times s}$ we can compute $\mathbf{U}(Z)$ and $\mathbf{T}(Z)$ satisfying Theorem 2.2. 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 σ^{-1} does not introduce fractions, so that $\sigma^{-1}(a) \in \mathbb{D}$ for all $a \in \mathbb{D}$. We write $\hat{\mathbf{F}}(\hat{Z}) := \mathbf{F}(\hat{Z}^{-1})\cdot\hat{Z}^N$, 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

$$\mathbf{U}_k(Z) = Z^{\vec{\mu}_k}\cdot\hat{\mathbf{M}}_k(\hat{Z}), \qquad \mathbf{T}_k(Z) = Z^{\vec{\mu}_k}\cdot\hat{\mathbf{R}}_k(\hat{Z})\cdot\hat{Z}^{\vec{\omega}_k - N\cdot\vec{e}}, \tag{26}$$

then $\mathbf{U}_k(Z)\cdot\mathbf{F}(Z) = \mathbf{T}_k(Z)$. In this case simple algebra shows that the recursion formulas for $\mathbf{U}_k(Z)$ obtained from (16) and (17) become

$$\sigma^{\vec{\mu}_k^\ell}(p_{\pi_k})\cdot\mathbf{U}_{k+1}(Z)^{\ell,\cdot} = \sigma^{\vec{\mu}_k^\ell}(r_{\pi_k})\cdot\mathbf{U}_k(Z)^{\ell,\cdot} - \sigma^{\vec{\mu}_k^\ell}(r_\ell)\cdot Z^{\vec{\mu}_k^\ell - \vec{\mu}_k^{\pi_k}}\cdot\mathbf{U}_k(Z)^{\pi_k,\cdot} \tag{27}$$

for $\ell \neq k$ and

$$\begin{aligned}&\sigma^{\vec{\mu}_k^{\pi_k}+2}(p_{\pi_k})\cdot\mathbf{U}_{k+1}(Z)^{\pi_k,\cdot}\\ &= \sigma^{\vec{\mu}_k^{\pi_k}+1}(r_{\pi_k})\cdot\mathbf{U}_k(Z)^{\pi_k,\cdot} - \sum_{\ell\neq\pi_k}\sigma^{\vec{\mu}_k^{\pi_k}+2}(p_\ell)\cdot Z^{\vec{\mu}_k^{\pi_k}-\vec{\mu}_k^\ell+1}\cdot\mathbf{U}_{k+1}(Z)^{\ell,\cdot},\end{aligned} \tag{28}$$

where

$$r_\ell = \sigma^{-\vec{\mu}_k^\ell}(c_{N+\vec{\mu}_k^\ell-\lfloor k/s\rfloor}(\mathbf{T}_k(Z)^{\ell,(k \bmod m)+1})),$$
$$p_\ell = \sigma^{-\vec{\mu}_k^{\pi_k}}(c_{\vec{\mu}_k^{\pi_k}-\vec{\mu}_k^\ell-\delta_{\pi_k,\ell}+1}(\mathbf{U}_k(Z)^{\pi_k,\ell})).$$

Since $\vec{\mu}_k^{\pi_k} \leq \vec{\mu}_k^\ell$ whenever $r_\ell \neq 0$, and that $p_\ell = 0$ whenever $\vec{\mu}_k^{\pi_k} < \vec{\mu}_k^\ell - 1$ 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 7.1: *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) $\mathbf{U}_k(Z)\cdot\mathbf{F}(Z) = \mathbf{T}_k(Z)$*;*

(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 (23), we see that rank $\hat{\mathbf{R}}(0)$ = rank $\hat{\mathbf{F}}(\hat{Z})$ = rank $\hat{\mathbf{R}}(\hat{Z})$, which is also the number of nonzero rows in $\hat{\mathbf{R}}(\hat{Z})$. Therefore, the nonzero rows of $\hat{\mathbf{R}}(\hat{Z})$ form a matrix with trailing coefficient of full row rank. It is easy to see that $rdeg\,\mathbf{T}(Z) = \vec{\mu}+(N-\kappa)\cdot\vec{e}$ and that

$$\mathbf{T}(Z)^{i,\cdot} = \sigma^{\vec{\mu}^i}(\hat{\mathbf{R}}(0)^{i,\cdot})\cdot Z^{\vec{\mu}^i+N-\kappa} + \text{lower degree terms.}$$

Therefore, $L(\mathbf{T}(Z)) = \sigma^{\deg\mathbf{T}(Z)}(\hat{\mathbf{R}}(0))$. Since σ is an automorphism on $\mathbb{Q}$, it follows that rank $L(\mathbf{T}(Z))$ = rank $\hat{\mathbf{R}}(0)$, and hence the nonzero rows of $\mathbf{T}(Z)$ 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, 2002) (also known as quasi-Popov form (Beckermann, Labahn and Villard, 2001)). The weak Popov form is defined as follows.

DEFINITION 7.2 (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 π_k are distinct for $\kappa s - s \leq k < \kappa s$ by (24). We now show that $\mathbf{T}(Z)$ is in weak Popov form.

THEOREM 7.3: *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) $\mathbf{U}(Z) \cdot \mathbf{F}(Z) = \mathbf{T}(Z)$*;*

(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 operation of the second type on each row until it has been chosen as a pivot. Moreover, we have that for all k and ℓ, $\mathbf{U}_k(Z)^{\ell,\cdot} \cdot \mathbf{F}(Z) = \mathbf{T}_k(Z)^{\ell,\cdot}$ and hence (c) is true.

Let H_k be the coefficient of $\hat{Z}^{(\kappa-1)\cdot\vec{e}}$ of $\hat{\mathbf{M}}_k(\hat{Z}) \cdot \hat{\mathbf{F}}(\hat{Z})$ 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 $\kappa s - s \leq k < \kappa s$, $H_{\kappa s}^{i,\cdot}$ must be zero. Therefore, if we form the matrix H by

$$H^{i,j} = \begin{cases} H_k^{i,j} & \text{if } \pi_k = i \text{ for some } \kappa s - s \leq k < \kappa s \\ H_{\kappa s}^{i,j} & \text{otherwise,} \end{cases} \tag{29}$$

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

$$\mathbf{T}(Z)^{i,\cdot} = \sigma^{\vec{\mu}^i_{\kappa s-s}}(H^{i,\cdot}) \cdot Z^{\vec{\mu}^i_{\kappa s-s}+N-\kappa} + \text{lower degree terms.}$$

Thus, $L(\mathbf{T}(Z)) = \sigma^{\deg \mathbf{T}(Z)}(H)$. Since σ is an automorphism on $\mathbb{Q}$ 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 7.1 and of Theorem 7.3 both provide a basis of the left nullspace of $\mathbf{F}(Z)$.

7.3. Computing GCRD and LCLM of Matrices Skew Polynomials

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, 2000b; 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

$$\mathbf{U}(Z)\cdot\mathbf{F}(Z)=\begin{bmatrix}\mathbf{U}_{11}(Z) & \mathbf{U}_{12}(Z)\\ \mathbf{U}_{21}(Z) & \mathbf{U}_{22}(Z)\end{bmatrix}\cdot\begin{bmatrix}\mathbf{A}(Z)\\ \mathbf{B}(Z)\end{bmatrix}=\begin{bmatrix}\mathbf{G}(Z)\\ 0\end{bmatrix} \tag{30}$$

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 $\mathbf{V}_1(Z)\cdot\mathbf{A}(Z)=\mathbf{V}_2(Z)\cdot\mathbf{B}(Z)$ 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 $\mathbf{Q}(Z)\in\mathbb{Q}[Z;\sigma,0]^{(m_1+m_2-s)\times(m_1+m_2-s)}$ such that

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

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

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

7.4. Computation of Subresultants

The method of Section 7.3, applied to two 1×1 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 $\mathrm{sres}_j(a,b)$ for skew polynomials is defined by taking the determinant of the matrix

$$\begin{bmatrix}
\sigma^{d_b-j-1}(a_{d_a}) & \sigma^{d_b-j-1}(a_{d_a-1}) & \cdots & \cdots & \cdots & \sigma^{d_b-j-1}(a_{2j+2-d_b}) & Z^{d_b-j-1}a(Z)\\
 & \ddots & & & & \vdots & \vdots\\
 & & \sigma(a_{d_a}) & \cdots & \cdots & \sigma(a_j) & Za(Z)\\
 & & & a_{d_a} & \cdots & a_{j+1} & a(Z)\\
\sigma^{d_a-j-1}(b_{d_b}) & \sigma^{d_a-j-1}(b_{d_b-1}) & \cdots & \cdots & \cdots & \sigma^{d_a-j-1}(b_{2j+2-d_a}) & Z^{d_a-j-1}b(Z)\\
 & \ddots & & & & \vdots & \vdots\\
 & & \sigma(b_{d_b}) & \cdots & \cdots & \sigma(b_j) & Zb(Z)\\
 & & & b_{d_b} & \cdots & b_{j+1} & b(Z)
\end{bmatrix}.$$

THEOREM 7.4: *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 $sres_j(a,b)\neq 0$ 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,*

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

Proof: After expanding with respect to the first columns, we see that the quantity $\gamma \cdot \operatorname{sres}_j(a,b)$ coincides with the determinant of the matrix

$$\begin{bmatrix} \sigma^{d_a-j-1}(a_{d_a}) & \sigma^{d_a-j-1}(a_{d_a-1}) & \cdots & \cdots & \cdots & \sigma^{d_a-j-1}(a_{2j+2-d_a}) & Z^{d_a-j-1}a(Z) \\ & \ddots & & & & \vdots & \vdots \\ & & \sigma(a_{d_a}) & \cdots & \cdots & \sigma(a_j) & Za(Z) \\ & & & a_{d_a} & \cdots & a_{j+1} & a(Z) \\ \sigma^{d_a-j-1}(b_{d_a}) & \sigma^{d_a-j-1}(b_{d_a-1}) & \cdots & \cdots & \cdots & \sigma^{d_a-j-1}(b_{2j+2-d_a}) & Z^{d_a-j-1}b(Z) \\ & \ddots & & & & \vdots & \vdots \\ & & \sigma(b_{d_a}) & \cdots & \cdots & \sigma(b_j) & Zb(Z) \\ & & & b_{d_b} & \cdots & b_{j+1} & b(Z) \end{bmatrix}.$$

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

$$\sigma^{-(d_a-j-1)}(S_j) = K((d_a-j, d_a-j), (2d_a-2j-1)), \tag{31}$$

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

One verifies using (Beckermann, Cheng and Labahn, 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 (31) and the determinant representations of Section 4 together with the uniqueness of Mahler systems in order to conclude that

$$\gamma\cdot\operatorname{sres}_j(a,b) = \pm Z^{\vec{\mu}^\ell}\cdot\hat{\mathbf{R}}_k(\hat{Z})^{\ell,\cdot}\cdot\hat{Z}^{\vec{\omega}-d_a\cdot\vec{e}} = \mathbf{T}_k(Z)^{\ell,1}.$$

□

Therefore, 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 $\hat{\mathbf{R}}_{2k-1}(\hat{Z})\cdot Z^{\vec{\omega}-d_a\cdot\vec{e}}$ after multiplying by Z^k and dividing by the extra factor of γ (or by dividing $\mathbf{T}_{2k-1}(Z)^{\ell,1}$ by γ).

Notice that the extra factor of γ 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

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

8. 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 (2000b) 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 2×1 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 would like to find a fraction-free method for reduction of leading coefficients over Ore domains. We will look at combining the procedure in Theorem 2.2 along with modified Schur complements (Beckermann, Cabay and Labahn, 1997) of Krylov matrices to help us develop such an algorithm.

We are also interested in extending our results to nested Ore polynomial domains, allowing for computations for example in Weyl algebras. This is a difficult extension to do in a fraction-free way 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 null-space computations.

References

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

Abramov, S., Bronstein, M., (2001). On solutions of linear functional systems, *Proceedings of ISSAC 2001, London, Canada*, 1–6.

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

Beckermann, B., Cabay, S., Labahn, G., (1997). Fraction-free Computation of Matrix Padé Systems, *Proceedings of ISSAC 1997, Maui*, 125–132.

Beckermann, B., Cheng, H., Labahn, G., (2002). Fraction-free row reduction of matrices of skew polynomials, *Proceedings of ISSAC 2002, Lille, France*, 8–15.

Beckermann, B., Labahn, G., (1992). A uniform approach for Hermite Padé and simultaneous Padé approximants and their matrix generalization, *Numerical Algorithms*, **3**, 45–54.

Beckermann, B., Labahn, G., (1994). A uniform approach for the fast, reliable computation of matrix-type Padé approximants, *SIAM J. Matrix Anal. Appl.*, **15**, 804–823.

Beckermann, B., Labahn, G., Recursiveness in matrix rational interpolation problems, *J. Comput. Appl. Math.*, (1997). **77**, 5–34.

Beckermann, B., Labahn, G., (2000a). Effective computation of rational approximants and interpolants, *Reliable Computing*, **6**, 365–390.

Beckermann, B., Labahn, G., (2000b). Fraction-free computation of matrix GCD's and rational interpolants, *SIAM J. Matrix Anal. Appl.*, **22(1)**, 114–144.

Beckermann, B., Labahn, G., Villard, G., (2001). Normal Forms for General Polynomial Matrices, *Publication ANO435, Université de Lille 1, France*,

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

Cohn, P. M., (1971). *Free Rings and Their Relations*, Academic Press, London & New York,

Collins, G., (1967). Subresultant and reduced polynomial remainder sequences, *J. ACM*, **14**, 128–142.

Geddes, K., Czapor, S., Labahn, G., (1992). *Algorithms for Computer Algebra*, Kluwer, Boston, MA,

Kailath, T., (1980). *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ,

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

Li, Z., (1998). A subresultant theory for Ore polynomials with applications, *Proceedings of ISSAC 1998, Rostock, Germany*, 132–139.

Mulders, T., Storjohann, A., (2002). On lattice reduction for polynomial matrices, *To appear in Journal of Symbolic Computation*,

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

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, $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$. Denote by $\mathcal{M}_{\mathbf{F}(Z)} = \{\mathbf{Q}(Z)\mathbf{F}(Z) : \mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}\}$ the submodule of the (left) $\mathbb{Q}[Z;\sigma,\delta]$-module $\mathbb{Q}[Z;\sigma,\delta]^{1\times s}$ obtained by forming 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 $\mathbb{Q}[Z;\sigma,\delta]$ is defined to be the cardinality of a maximal $\mathbb{Q}[Z;\sigma,\delta]$-linearly independent subset of $\mathcal{M}$. Comparing with our Definition 2.1, we see that $\operatorname{rank}\mathbf{F}(Z) \le \operatorname{rank}\mathcal{M}_{\mathbf{F}(Z)}$. In Theorem A.2 below show that in fact we have equality.

Notice that for any $\mathbf{A}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$ 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 $\mathbf{V}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$, 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 $\mathbb{Q}[Z;\sigma,\delta]$. 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* $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$, *with* $\vec{\mu} = \operatorname{rdeg}\mathbf{F}(Z)$.

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

$$\deg \mathbf{Q}(Z)\mathbf{F}(Z) = \max_j(\vec{\mu}^j + \deg \mathbf{Q}(Z)^{1,j}).$$

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

(c) *Let* $\mathbf{A}(Z) = \mathbf{B}(Z)\cdot\mathbf{C}(Z)$ *be as in part (b), with* $\mathbf{A}(Z)$ *and* $\mathbf{C}(Z)$ *row-reduced with row degrees* $\vec{\alpha}^1 \le \vec{\alpha}^2 \le \ldots \le \vec{\alpha}^m$ *and* $\vec{\gamma}^1 \le \vec{\gamma}^2 \le \ldots \le \vec{\gamma}^r$, *respectively. Then* $m \le r$, *and* $\vec{\alpha}^j \ge \vec{\gamma}^j$ *for* $j = 1,\ldots,m$.

(d) *Let* $\mathbf{T}(Z) = \mathbf{U}(Z)\cdot\mathbf{S}(Z)$, *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 $\mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}$ let $N' := \max_j \vec{\mu}^j + \deg \mathbf{Q}(Z)^{1,j}$ and define the vector $h \in \mathbb{Q}^{1\times m}$, $h \neq 0$, by

$$\mathbf{Q}(Z)^{1,j} = h^j Z^{N'-\vec{\mu}^j} + \text{ lower degree terms.}$$

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

$$\sum_{j=1}^{m} h^j \sigma^{N'-\vec{\mu}^j}(F^{j,\cdot}_{\vec{\mu}^j}) = h \cdot \sigma^{N'-N}(L(\mathbf{F}(Z))).$$

Since σ is an automorphism, 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 $\mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}$.

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 $\mathcal{M}_{\mathbf{B}(Z)} \subset \mathbb{Q}[Z;\sigma,\delta]^{1\times r}$, the latter $\mathbb{Q}[Z;\sigma,\delta]$-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 $\mathbf{Q}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}$ 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 $\mathbb{Q}[Z;\sigma,\delta]$-linearly independent by part (a), and hence $s = \operatorname{rank} \mathbf{A}(Z) \leq r$ by part (b). Suppose that $\vec{\alpha}^j \geq \vec{\gamma}^j$ for $j < k$, but $\vec{\alpha}^k < \vec{\gamma}^k$. Part (a) tells us that $\deg \mathbf{B}(Z)^{j,\ell} \leq \vec{\alpha}^j - \vec{\gamma}^\ell$. Since $\vec{\alpha}^j < \vec{\gamma}^k \leq \vec{\gamma}^\ell$ 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 polynomial linear 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 $\mathbb{Q}[Z;\sigma,\delta]$-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 $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$. Clearly, $\mathcal{N}_{\mathbf{F}(Z)}$ is a $\mathbb{Q}[Z;\sigma,\delta]$-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 $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$, and $\mathbf{U}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$ be unimodular, with $\mathbf{T}(Z) = \mathbf{U}(Z) \cdot \mathbf{F}(Z)$ having r nonzero rows, where the submatrix consisting of the r nonzero rows of $\mathbf{T}(Z)$ are row-reduced. Then*

$$r = \mathit{rank}\, \mathcal{M}_{\mathbf{F}(Z)} = \mathit{rank}\, \mathbf{F}(Z) = m - \mathit{rank}\, \mathcal{N}_{\mathbf{F}(Z)}, \tag{32}$$

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

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

Proof: Denote by J the set of indices of zero rows of $\mathbf{T}(Z)$, and define the matrix $\mathbf{U}(Z)^{J,\cdot}$ 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 (32) for the matrix $\mathbf{T}(Z)$. For some $\mathbf{P}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{1\times m}$ we have

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

We have shown implicitly in Lemma A.1(a) that the rows $\mathbf{T}(Z)^{j,\cdot}$ for $j \notin J$ are linearly independent over $\mathbb{Q}[Z;\sigma,\delta]$. Therefore $\mathbf{P}(Z) \in \mathcal{N}_{\mathbf{T}(Z)}$ if and only if $\mathbf{P}(Z)^{1,j} = \mathbf{0}$ 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 $\operatorname{rank} \mathbf{T}(Z) \le \operatorname{rank} \mathcal{M}_{\mathbf{T}(Z)} =: \rho$. Suppose that there is strict inequality. Then there exist ρ elements of $\mathcal{M}_{\mathbf{T}(Z)}$ which are $\mathbb{Q}[Z;\sigma,\delta]$-linearly independent and which can be written as rows of the matrix $\mathbf{B}(Z)\mathbf{T}(Z)$ for some $\mathbf{B}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{\rho\times m}$. 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 $\operatorname{rank} \mathbf{B}(Z)\mathbf{T}(Z) \le r$ by Lemma A.1(b), a contradiction. Consequently, (32) 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 $\mathbf{V}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times m}$ with $\mathbf{V}(Z)\cdot\mathbf{U}(Z) = \mathbf{U}(Z)\cdot\mathbf{V}(Z) = \mathbf{I}$. Consequently, $\mathbf{Q}(Z) \in \mathcal{N}_{\mathbf{F}(Z)}$ if and only if $\mathbf{P}(Z) = \mathbf{Q}(Z)\cdot\mathbf{V}(Z) \in \mathcal{N}_{\mathbf{T}(Z)}$, that is,

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

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 $\mathbb{Q}[Z;\sigma,\delta]$. Thus the rows of $\mathbf{U}(Z)^{J,\cdot}$ 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 $\rho := \operatorname{rank} \mathbf{F}(Z) \le \operatorname{rank} \mathcal{M}_{\mathbf{F}(Z)}$ 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 ρ rows of $\mathbf{F}(Z)$ are linearly independent. Then, by maximality of ρ, we find for any $j = \rho+1, ..., m$ quantities $\mathbf{Q}(Z)^{j,k} \in \mathbb{Q}[Z;\sigma,\delta]$ with

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

that is, we have found $m - \rho > m - r$ many $\mathbb{Q}[Z;\sigma,\delta]$-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 $\mathbf{U}(Z)^{J,\cdot}$ 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 $rdeg\,\mathbf{W}(Z) \leq rdeg\,\mathbf{U}(Z)^{J,\cdot}$. Hence it only remains to recall the bound for the row-degree of the multiplier $\mathbf{U}(Z)$ of Theorem A.2: we have for $j \in J$

$$\deg \mathbf{U}(Z)^{j,\cdot} \leq \vec{\nu}^j - \vec{\mu}^j + (|\vec{\mu}| - |\vec{\nu}|) \leq |\vec{\mu}| - \vec{\mu}^j \leq (m-1)N.$$

□

We should mention that the quantity $rdeg\,\mathbf{W}(Z)$ 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 $rdeg\,\mathbf{W}(Z)$ 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 $\mathbf{F}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{m\times s}$, 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 $\mathbf{A}(Z) \in \mathbb{Q}[Z;\sigma,\delta]^{s\times s}$ is of rank s. Then $\mathcal{N}_{\mathbf{A}(Z)} = \{0\}$ by (32), 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 (32). If $\mathbf{U}(Z)$ is unimodular, then $\mathcal{M}_{\mathbf{U}(Z)\cdot\mathbf{F}(Z)} = \mathcal{M}_{\mathbf{F}(Z)}$, 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 $\mathbb{Q}[Z;\sigma,\delta]$. Since $\mathbb{Q}[Z;\sigma,\delta]$ contains no zero divisors, it is easy to check that $\mathbf{F}(Z)$ and the new matrix will have the same number of $\mathbb{Q}[Z;\sigma,\delta]$-linearly independent rows, and hence the same rank. □