ALBANIAN JOURNAL
OF MATHEMATICS
Volume 4, Number 4, Pages 181–193
ISSN 1930-1235: (2010)

CONVERSION BETWEEN HERMITE AND POPOV NORMAL FORMS USING AN FGLM-LIKE APPROACH

JOHANNES MIDDEKE

Abstract. We are working with matrices over a ring $K[\partial;\sigma,\vartheta]$ of Ore polynomials over a skew field K. Extending a result of [18] for usual polynomials it is shown that in this setting the Hermite and Popov normal forms correspond to Gröbner bases with respect to certain orders. The FGLM algorithm is adapted to this setting and used for converting Popov forms into Hermite forms and vice versa. The approach works for arbitrary, that is, not necessarily square matrices where we establish termination criteria to deal with infinitely dimensional factor spaces.

1. Introduction

Since long, normal forms have played a prominent rôle in those branches of mathematics that involve the study of equational systems. Among these, polynomial systems form a major subclass. But also systems of ordinary linear equations are important for applications. Usually, these systems are modelled by matrices. The computation of normal forms can answer some important questions about the structure of the underlying system.

Linear systems can be represented by matrices with entries being linear operators. In this paper we will consider Ore polynomials—which some authors also call skew polynomials. This is a class of non-commutative polynomials that was introduced by Øystein Ore in [21]. They are a generalisation of the ordinary (commutative) polynomials that includes linear differential operators and shift operators.

We will tread Ore polynomials in section 2. There we will also introduce some notations for matrices that are used in later sections.

Among those normal forms that are used in practise, we will concentrate on the Hermite and Popov forms. These are both one-sided normal forms, that is, normal forms with respect to elementary row operations. Invented by Charles Hermite in [14], the Hermite form was originally a row echelon form for square matrices over the integers. It has later been extended to non-square matrices and other domains.

The Popov normal form was introduced by Vasile Mihai Popov in [22, 23]. It is related to row-reduction—a concept that has been described by [12] for commutative polynomials. We will give definitions for this forms in section 3.

Gröbner bases were first considered in Bruno Buchberger's PhD thesis [4]—named after his advisor Wolfgang Gröbner. They are very useful to solve problems

Key words and phrases. 15B33, 34M03, 47B39 .

This work was supported by the Austrian Science Foundation (FWF) under the project DIFFOP (P20 336-N18).

that are related to polynomial ideals algorithmically, most importantly the solving of polynomial equations and the ideal membership problem.

It is possible to define Gröbner bases for modules—see, for example, [1]. In [18] it was shown that matrices in Hermite or Popov form are in fact Gröbner bases with respect to this definition. In section 4 we will introduce Gröbner bases over Ore polynomials based on [5]. We will extend the result of [18] to this case.

Gröbner bases usually suffer from high computational complexity. In [11] the authors Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard and Teo Mora therefor attacked a special problem: Compute the Gröbner basis of an zero-dimensional ideal fast—provided that a Gröbner basis with respect to a different monomial ordering is already known. Breaking down the problem to linear algebra, they managed to obtain an efficient algorithm for that task.

In section 5 of this paper we will adapt the FGLM algorithm to modules over Ore polynomials. We will modify it in such a way that it also handles sub-modules that are not "zero-dimensional". We also will give an estimation of the complexity of this algorithm in the special case of converting Popov and Hermite forms.

There are other approaches for converting matrices in Popov and Hermite normal form into each other. One, for example, may be found in [24]. To our best knowledge, this paper is the first though that explores the connection of normal forms and Gröbner bases to complete this task.

We also compiled a technical report about this topic that contains a MAPLE™ implementation for the conversion of Popov into Hermite forms as well as detailed examples.

2. Basic notations

Ore polynomials—also called skew polynomials by some authors—are a generalisation of the usual polynomials with a non-commutative multiplication. They are named after Øystein Ore who was the first to describe them in [21]. We will only give an informal description of Ore polynomials here. A more rigid description may be found in [9, Chapter 0.10] or [10, Chapter 5.2].

Let K be a (computable) skew field, and let $\sigma\colon K \to K$ be an automorphism. A map $\vartheta\colon K \to K$ such that

$$\vartheta(a+b) = \vartheta(a) + \vartheta(b) \qquad \text{and} \qquad \vartheta(ab) = \sigma(a)\vartheta(b) + \vartheta(a)b$$

for all $a, b \in K$ is called a *σ-derivation* of K. (The second identity is sometimes referred to as *σ-Leibniz rule*.) Let now ∂ be a variable. An *Ore polynomial* is just a polynomial expression

$$a_n\partial^n + a_{n-1}\partial^{n-1} + \ldots + a_2\partial^2 + a_1\partial + a_0$$

where $n \geqslant 0$ and where the coefficients $a_0, \ldots, a_n$ are in K. The set of all Ore polynomials is denoted by $K[\partial; \sigma, \vartheta]$. Two Ore polynomials in $K[\partial; \sigma, \vartheta]$ are added in the same way as usual polynomials. The multiplication of Ore polynomials is given by extending the *commutation rule*

$$\partial a = \sigma(a)\partial + \vartheta(a)$$

with $a \in K$ assuming associativity and distributivity. This defines a ring structure on $K[\partial; \sigma, \vartheta]$. For a proof see [9, Theorem 10.1].

Example 1. The typical examples of Ore polynomials are the following. See also [10, Page 186] or [8, Table 2] for further examples.

(1) For $K = \mathbb{Q}$, $\sigma = \mathrm{id}$ and $\vartheta = 0$ (that is, the identity and the constant zero function respectively) we obtain just the usual commutative polynomials with the commutation rule $\partial a = a\partial$.
(2) For $K = \mathbb{Q}(x)$ (or K being the meromorphic functions in x), $\sigma = \mathrm{id}$ and $\vartheta = d/dx$ we obtain *differential operators* with the commutation rule $\partial f = f\partial + \frac{df}{dx}$ reflecting the composition of linear differential operators.
(3) For $K = \mathbb{Q}(n)$, $\sigma(a(n)) = a(n+1)$ and $\vartheta = 0$ we obtain the *shift operators* having the commutation rule $\partial a(n) = a(n+1)\partial$.

Obviously, the multiplication of Ore polynomials needs not to be commutative. (Thus it is also important that we write coefficients always on the left hand side.) Still, they retain a lot of the usual properties of ordinary polynomials. Given $f = a_n\partial^n + \ldots + a_1\partial + a_0$ where $a_0, \ldots, a_n \in K$ and $a_n \neq 0$, we define the *degree* of f as $\deg f = n$. We refer to $\operatorname{lcoeff}(f) = a_n$ as the *leading coefficient* of f. For convenience, we set $\deg 0 = -\infty$. Degree and leading coefficient fulfill the identities

$$\deg(fg) = \deg f + \deg g \qquad \text{and} \qquad \operatorname{lcoeff}(fg) = \operatorname{lcoeff}(f)\sigma^{\deg f}(\operatorname{lcoeff}(g))$$

for all Ore polynomials f and g.

Using this degree function we can do polynomial long division almost as in the commutative case. We have to distinguish between division from the left and from the right, though. Furthermore, we can compute left greatest common divisors and right greatest common divisors. See [10, Theorem 5.8] or [3] for the algorithms and their proofs of correctness.

For any ring R, we denote the set of $m \times n$ matrices over R by $R^{m\times n}$. The $n \times n$ identity matrix is denoted by $\mathbf{1}_n$ and the $m \times n$ zero matrix is written as $\mathbf{0}_{m\times n}$. A square matrix $M \in R^{n\times n}$ that has a two-sided inverse $M^{-1} \in R^{n\times n}$ is called *unimodular*. The set of $n \times n$ unimodular matrices is denoted by $\mathrm{GL}(R, n)$.

We will need to extract certain rows or columns from our matrices. For $M = (a_{ij}) \in R^{m\times n}$ and $1 \leqslant i \leqslant m$ we denote the i^{th} row by $M_{i,\bullet} = (a_{i,1}, \ldots, a_{i,n})$. Similarly, for $1 \leqslant j \leqslant n$ the j^{th} column is denoted by $M_{\bullet,j} = {}^t(a_{1,j}, \ldots, a_{m,j})$, where ${}^t\bullet$ denotes transposition.

The set of row vectors with entries in R of size n will be written as $R^{1\times n}$ and the set of column vectors of size m is denoted by R^m. In this paper, row vectors are treated as left module over R and column vectors form a right module over R. We will often regard vectors as matrices with only one row or column respectively.

If in particular $R = K[\partial; \sigma, \vartheta]$ is a ring of Ore polynomials, then for a matrix $M = (a_{i,j}) \in R^{m\times n}$ we define $\deg M = \max\{\deg a_{i,j} \mid i = 1, \ldots, m \text{ and } j = 1, \ldots, n\}$. As a further abbreviation we also define the i^{th} row degree for $1 \leqslant i \leqslant m$ as $\operatorname{rdeg}_i M = \deg M_{i,\bullet} = \max\{\deg a_{i,j} \mid j = 1, \ldots, n\}$. Finally, we will need the *leading vector* $\operatorname{lvec}(M) = \big(\operatorname{coeff}_\partial(\deg M, a_{i,j})\big)_{i,j} \in K^{m\times n}$. (We use the name "leading vector" instead of "leading matrix" because it will mostly be applied to vectors.)

3. Hermite and Popov normal forms

We will now define the main concepts we are dealing with in this paper, namely Hermite and Popov normal forms. Let again K be any skew field with automorphism $\sigma\colon K \to K$, σ-derivation $\vartheta\colon K \to K$, and let $R = K[\partial; \sigma, \vartheta]$.

We will start with the Hermite normal form. Our definition is taken from [13, Definition 3.2]. A matrix in Hermite form is basically just in row echelon form with some additional properties of the degrees of the entries.

Definition 2 (Hermite normal form)**.** A matrix $M = (a_{i,j}) \in R^{m\times n}$ is in *Hermite normal form* if and only if there exists indices $j_1 > j_2 > \ldots > j_m$ that are called *pivot indices* such that

(1) $a_{i,k} = 0$ if $k < j_i$,
(2) the entries a_{i,j_i} are monic, and
(3) $\deg a_{i,j_i} > \deg a_{k,j_i}$ for $k \neq i$.

Every matrix $N \in R^{m\times n}$ can be transformed using elementary row operations into a matrix M whose non-zero rows form a matrix in Hermite normal form. That is, for every such N there exists an invertible matrix $S \in \mathrm{GL}(R, m)$ such that

$$SN = \begin{pmatrix} M \\ \mathbf{0}_{m-s\times n} \end{pmatrix}$$

where $M \in R^{s\times n}$ is in Hermite form. Sometimes we will a little bit sloppily also refer to the whole right hand side—that is, with zero rows included—as the Hermite form of N. The computations can be done applying the (matrix form of the) Euclidean algorithm to the columns of N to achieve a row echelon form and then using polynomial division to enforce the degree restrictions. See [13, Theorem 3.2] for a more detailed description in the case of square matrices. We will show later in corollary 15 that the Hermite form of N is actually uniquely determined.

The definition of the Popov normal form is slightly more involved. We need to proceed in two steps. First we will introduce the concept of row-reducedness and afterwards as second step we define the Popov as a row-reduced matrix with additional properties.

Row-reducedness was first introduced in [12] for commutative polynomials. A presentation for Ore polynomials can be found in [2]. We repeat the definitions here for the convenience of the reader. Let $M \in R^{m\times n}$. When we multiply M from the left by the matrix $D = \mathrm{diag}(\partial^{\deg M - \mathrm{rdeg}_1 M}, \ldots, \partial^{\deg M - \mathrm{rdeg}_m M})$, we obtain a matrix DM with all rows having the same degree $\deg M$. Its leading vector

$$\mathrm{lvec}(DM) = \begin{pmatrix} \sigma^{\deg M - \mathrm{rdeg}_1 M}(\mathrm{lvec}(M_{1,\bullet})) \\ \vdots \\ \sigma^{\deg M - \mathrm{rdeg}_m M}(\mathrm{lvec}(M_{m,\bullet})) \end{pmatrix} \in K^{m\times n}$$

is called the *leading (row) coefficient matrix* of M. We denote it by $\mathrm{LC}(M)$.

Definition 3 (Row-reducedness)**.** A matrix $M \in R^{m\times n}$ is called *row reduced* if $\mathrm{LC}(M)$ has full left row rank.

Being row-reduced is the most important requirement for being in Popov normal form. The other points in the following definition basically just make sure that the matrix is uniquely determined. One can show that a matrix in Popov form has a leading coefficient matrix in row-echelon form. See for example [19, Lemma 14] for a proof. The definition is taken from [18, Definition 1].

Definition 4 (Popov normal form)**.** A matrix $M = (a_{i,j})ij \in R^{m\times n}$ is said to be in *Popov normal form*, if

(1) M is row-reduced and $\operatorname{rdeg}_i M \leqslant \operatorname{rdeg}_{i+1} M$ for all i;
(2) for the i^{th} row there exists a column index j_i (the *pivot index*) such that
   (a) a_{i,j_i} is monic and $\deg a_{i,j_i} = \operatorname{rdeg}_i M$;
   (b) $\deg a_{i,k} < \operatorname{rdeg}_i M$ if $k < j_i$;
   (c) $\deg a_{k,j_i} < \operatorname{rdeg}_i M$ if $k \neq i$; and
   (d) if $\operatorname{rdeg}_i M = \operatorname{rdeg}_k M$ and $i < k$ then $j_i < j_k$ (that is, pivot indices are ordered increasingly).

Also Popov forms are normal forms in the sense that each matrix can be transformed by elementary row operations into a matrix whose non-zero rows are in Popov form. Again this later matrix is sometimes simply referred to as the Popov form. The conversion can be done by first applying row-reduction—which is described, for example, in [2, Theorem 2.2]—and then using similar operations to achieve the degree constraints in the definition. See [6, Section 2.5.1] for a more detailed description.

Remark 5. Hermite forms clearly have independent rows since they are in row echelon form. But also the rows of matrices in Popov form are linearly independent—actually row-reducedness is already sufficient for that: By the so-called *predictable degree property* [2, Lemma A.1 (a)], if $v \in R^{1\times m}$ and $M \in R^{m\times n}$, then $vM = 0$ is only possible if $\deg v_i + \operatorname{rdeg}_i M < 0$ for all $1 \leqslant i \leqslant m$. This implies immediately that $v = 0$, since the rows of M are all non-zero.

4. Gröbner bases

Gröbner bases have been invented by Bruno Buchberger in [4]. Though initially defined for multivariate commutative polynomials, the concept has since been extended to more general domains such as Ore polynomials (see, for example, [8]) or modules over polynomial rings (see, for example, [20]). In this paper we will use the nice description of Gröbner bases of modules over Poincaré-Birkhoff-Witt rings given in [5, Chapter 5]. Poincaré-Birkhoff-Witt rings are a more general class of non-commutative domains that includes Ore polynomials. See [5, Definition 2.2.5] for the definition of Poincaré-Birkhoff-Witt rings and [5, Corollary 2.3.3] for the proof that Ore polynomials are included. An approach exclusively for Ore polynomials may be found in [8]—but there seems to be no extension to modules.

We include some of the results of [5] here for completeness and in order to adapt them to our notation. Let once more K be any skew field with automorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$. Also, let $R = K[\partial; \sigma, \vartheta]$. We will just briefly skip through the most important definitions and provide pointers to the corresponding sections of [5]. Readers who are familiar with commutative Gröbner bases will find that everything translates well to the non-commutative case.

For $i = 1, \ldots, n$, let $\mathfrak{e}_i$ denote the i^{th} unit vector in $R^{1\times n}$. A *monomial* is a product $\partial^\alpha \mathfrak{e}_i$ of a power of ∂ and a unit vector where $\alpha \geqslant 0$ and $1 \leqslant i \leqslant n$. A *term* is the product of a scalar (that is, an element in K) and a monomial. There are two obvious ways of introducing a total ordering on monomials. See also [5, Definitions 5.3.8 and 5.3.9] and the definition of admissible orderings [5, Definition 5.3.7].

Definition 6 (Position over term/term over position ordering)**.** Let $\partial^\alpha \mathfrak{e}_i$ and $\partial^\beta \mathfrak{e}_j$ be monomials in $R^{1\times n}$ with $\alpha, \beta \geqslant 0$ and $1 \leqslant i, j \leqslant n$.

(1) The *position over term (POT)* ordering is defined by

$$\partial^\alpha \mathfrak{e}_i <_{\text{POT}} \partial^\beta \mathfrak{e}_j \quad :\Longleftrightarrow \quad i > j \vee (i = j \wedge \alpha < \beta);$$

(2) the *term over position (TOP)* ordering is given by

$$\partial^\alpha \mathfrak{e}_i <_{\text{TOP}} \partial^\beta \mathfrak{e}_j \quad :\Longleftrightarrow \quad \alpha < \beta \vee (\alpha = \beta \wedge i > j).$$

We will use $\leqslant_{\text{POT}}$, $\geqslant_{\text{POT}}$, $>_{\text{POT}}$, and $\leqslant_{\text{TOP}}$, $\geqslant_{\text{TOP}}$, $>_{\text{TOP}}$ in the usual way.

It is important to note here that we fixed an ordering on the indices (positions). The reason is that, although for Gröbner basis theory any ordering of the indices would be fine, for our application to Hermite and Popov forms this particular ordering is crucial.

If, for example, $n = 3$ then the smallest monomials with respect to the position over term ordering are

$$\begin{aligned}(0,0,1) <_{\text{POT}} (0,0,\partial) <_{\text{POT}} (0,0,\partial^2) <_{\text{POT}} \ldots <_{\text{POT}} (0,1,0) <_{\text{POT}} (0,\partial,0)\\ <_{\text{POT}} (0,\partial^2,0) <_{\text{POT}} \ldots <_{\text{POT}} (1,0,0) <_{\text{POT}} (\partial,0,0) <_{\text{POT}} (\partial^2,0,0) <_{\text{POT}} \ldots\end{aligned}$$

while with respect to the term over position ordering we obtain the chain

$$\begin{aligned}(0,0,1) <_{\text{TOP}} (0,1,0) <_{\text{TOP}} (1,0,0) <_{\text{TOP}} (0,0,\partial) <_{\text{TOP}} (0,\partial,0)\\ <_{\text{TOP}} (\partial,0,0) <_{\text{TOP}} (0,0,\partial^2) <_{\text{TOP}} (0,\partial^2,0) <_{\text{TOP}} (\partial^2,0,0) <_{\text{TOP}} \ldots.\end{aligned}$$

Thus, the position over term ordering has similarities to the lexicographic ordering in the usual commutative Gröbner basis theory while the term over position ordering corresponds to the degree lexicographic ordering.

Let now for a while $<$ denote either $<_{\text{POT}}$ or $<_{\text{TOP}}$. Any vector in $R^{1\times n}$ may be written as K-linear combination of monomials. That is, taking $v \in R^{1\times n}$ there are $k \geqslant 0$, $c_1, \ldots, c_k \in K$ and monomials $\mathfrak{m}_1, \ldots, \mathfrak{m}_k$ such that $v = c_1\mathfrak{m}_1 + \ldots + c_k\mathfrak{m}_k$. If $c_1 \neq 0$ and $\mathfrak{m}_1 > \mathfrak{m}_j$ for $2 \leqslant j \leqslant k$, then we call $\mathfrak{m}_1 = \operatorname{lmonom}_<(v)$ the *leading monomial* of v with respect to $<$. In this case, $c_1 = \operatorname{lcoeff}_<(v)$ is the *leading coefficient* and $c_1\mathfrak{m}_1 = \operatorname{lterm}_<(v)$ is the *leading term*. (Note the difference between leading coefficient and leading vector). If no confusion about to which order we confer may arise, then we just write $\operatorname{lmonom}(v)$ instead of $\operatorname{lmonom}_<(v)$ and the same for $\operatorname{lcoeff}(v)$ and $\operatorname{lterm}(v)$. Leading monomial, term and coefficient of the zero vector remain undefined.

Example 7. With respect to the position over term ordering, the leading monomial of a non-zero vector $v \in R^{1\times n}$ corresponds to the term of highest degree of the left-most non-zero entry of v. With respect to the term over position ordering, the leading term corresponds to the left-most of the entries of highest degree.

Using the above definition of leading term, reduction is defined as in the commutative case. That is, if $v = c_1\mathfrak{m}_1 + \ldots + c_k\mathfrak{m}_k$ is as above and if $W \subseteq R^{1\times n} \setminus \{0\}$ is given, then v is said to be *reducible* by W if there are $w \in W$, $1 \leqslant i \leqslant k$ and $\alpha \geqslant 0$ such that $\mathfrak{m}_i = \partial^\alpha \operatorname{lmonom}(w)$. Otherwise, v is called *irreducible*.

Theorem 8. *Given $v \in R^{1\times n}$ and $\{w_1, \ldots, w_s\} \subseteq R^{1\times n} \setminus \{0\}$, there are elements $u_1, \ldots, u_s \in R$ and $r \in R^{1\times n}$ such that*

$$v = u_1w_1 + \ldots + u_sw_s + r$$

where r is not reducible by $\{w_1, \ldots, w_s\}$.

We will call r the remainder *of the division of v by $\{w_1, \ldots, w_n\}$.*

Proof. This is [5, Theorem 5.4.3]. Directly after the theorem—namely in [5, Algorithm 10]—the division method is explained in detail. □

We have now everything set in order to define Gröbner bases. We start with [5, Definition 5.4.7].

Definition 9 (Gröbner basis)**.** Let $\mathfrak{M}$ be an R-submodule of $R^{1\times n}$. A finite set $G \subseteq \mathfrak{M}$ is a *Gröbner basis* for $\mathfrak{M}$ if for all $v \in \mathfrak{M}$ there is $\alpha \geqslant 0$ and $g \in G$ such that $\operatorname{lmonom}(v) = \partial^\alpha \operatorname{lmonom}(g)$.

Lemma 10. *Every non-zero submodule $\mathfrak{M} \subseteq R^{1\times n}$ has a Gröbner basis G, $\mathfrak{M}$ is generated by G as a left R-module and the remainder of the division of an element $v \in R^{1\times n}$ by G does not dependent on the order of the elements in G.*

Furthermore, $v \in \mathfrak{M}$ if and only the remainder by division with G is zero.

Proof. These statements are found in [5, Proposition 5.4.8, Corollary 4.10 and Theorem 5.4.9]. □

The following definition is [5, Definition 4.17].

Definition 11 (Reduced Gröbner bases)**.** A Gröbner basis G of $\mathfrak{M} \subseteq R^{1\times n}$ is *reduced* if for all $g \in G$ we have $\operatorname{lcoeff}(g) = 1$ and there is no $h \in G \setminus \{g\}$ such that $\operatorname{lmonom}(h)$ divides a term in g.

As in the usual, commutative Gröbner basis theory one may define *S-polynomials* and prove a *Buchberger criterion* for Gröbner bases in $R^{1\times n}$. This can be found in [5, Definition 5.4.11 and Theorem 5.4.13]. But since we will not need the full Buchberger criterion in our proofs, we will be content with stating a corollary here.

Theorem 12. *Let $G = \{g_1, \ldots, g_s\} \subseteq R^{1\times n}$ with leading monomials $\operatorname{lmonom}(g_k) = \partial^{\alpha_k}\mathfrak{e}_{j_k}$ for $1 \leqslant k \leqslant s$. If $j_i \neq j_k$ whenever $i \neq k$ then G is a Gröbner basis for the submodule $Rg_1 + \ldots + Rg_s \subseteq R^{1\times n}$ generated by its elements.*

Proof. This is [5, Corollary 5.4.14]. □

We will now draw the connection from Gröbner bases to normal forms. For this we have to make the transition between matrices and sets of row vectors. We will say that a matrix $M \in R^{m\times n}$ is a (reduced) Gröbner basis with respect to a certain term ordering if the set of its rows $\{M_{1,\bullet}, \ldots, M_{m,\bullet}\}$ is a (reduced) Gröbner basis for its row space $R^{1\times m}M$.

The following two theorems are generalisations of [18, Proposition 2 and 4] to Ore polynomials.

Theorem 13. *Let $M \in R^{m\times n}$ with the rows sorted in descending order with respect to position over term ordering. Then M is in Hermite form if and only if the non-zero rows of M form a reduced Gröbner basis for $R^{1\times m}M$ with respect to position over term ordering.*

Proof. By example 7, with respect to position over term ordering, the leading terms of the rows are exactly those corresponding to the pivot indices in the sense of definition 2. Since the pivot indices are all different, M is a Gröbner basis by theorem 12, and since the corresponding entries are monic and the entries in the rows above are of lower degree, we even have a reduced Gröbner bases.

Conversely, one easily sees, that for a reduced Gröbner bases the leading terms must be in different positions. Setting these as the pivot indices, from this observation one deduces all properties listed in definition 2. □

Theorem 14. *Let $M \in R^{m\times n}$ with the rows sorted in ascending order with respect to term over position ordering. Then M is in Popov form if and only if the non-zero rows of M form a reduced Gröbner basis for $R^{1\times m}M$ with respect to term over position ordering.*

Proof. Analogously to the Hermite form, here the leading terms with respect to position over term ordering are those corresponding to the pivot indices—this time in the sense of definition 4. Again, they are in different positions and thus we obtain a Gröbner basis. As before, the properties listed in definition 4 make sure that the Gröbner basis is reduced. Also the converse is easily proven by letting the pivot indices be the positions of the leading terms and checking the properties in the definition. (For the row-reducedness note that the pivot indices are in different columns and hence the leading coefficient matrix must be in row echelon form.) □

Since reduced Gröbner bases for submodules by [5, Theorem 5.4.18] are unique, from the previous theorems we obtain (together with the existence considerations from section 3)

Corollary 15. *Every matrix has exactly one Hermite form and exactly one Popov form.*

5. FGLM

The first version of the FGLM algorithm—named after its inventors—was presented in [11]. It solves the following problem: Given a Gröbner basis of a zero-dimensional ideal I in a ring $F[x]$ of commutative polynomials over a field F with respect to a certain term order, compute the Gröbner basis of I with respect to another term order. That is, the FGLM algorithm allows to convert Gröbner bases between different term orderings. Since it does so quite efficiently, it is thus possible to compute a Gröbner basis for a "slow" term ordering by first computing it with respect to a "fast" term ordering and then using FGLM for conversion.

The main achievement of [11] is, that they managed to break this problem down to a linear algebra problem: Instead of calculating in $F[x]$ they solve the task in $F[x]/I$ which is a finite dimensional vector space over F. In this space they iterate over all (representatives of) monomials deciding whether they are leading monomials of an element of the new Gröbner basis or not.

Let again K be a skew field with automorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$. As before we abbreviate $K[\partial;\sigma,\vartheta]$ by R. Let $M \in R^{m\times n}$ be a Gröbner bases for the term over position or for the position over term ordering. It will turn out that the FGLM algorithm translates quite nicely to this setting. There is one problem, though, namely that $R^{1\times n}/R^{1\times m}M$ needs not to be finite dimensional. That is, we possibly have to traverse over infinitely many monomials.

Our first goal is thus to limit the number of monomials we have to search. For this we will need the next two lemmata that will give an estimate on the degrees of Popov and Hermite forms of a given matrix.

Lemma 16. *Let $A \in R^{m\times n}$ be any matrix and $M \in R^{m\times n}$ its Popov form. Then* $\deg M \leqslant \deg A$.

Proof. By [2, Theorem 2.2] does row-reduction applied to A at most lower the degree. Furthermore, since the Popov form M is by definition also row-reduced, by [2, Lemma A.1 (d)] we may conclude that its degree is the same as that of the result of the row-reduction and thus not larger than the degree of A, too. □

The next lemma is [13, Corollary 3.4]. Although in the reference the result is only stated for square matrices over rings of differential operators (see example 1), following the proofs one easily sees that they generalise to arbitrary Ore polynomials and to matrices that are not necessarily square.

Lemma 17. *Let $A \in R^{m\times n}$ be a matrix of full left row-rank, and let $M \in R^{m\times n}$ be its Hermite form. Then* $\deg M \leqslant m \deg A$.

Proof. See [13, Corollary 3.4]. □

Having thus established degree bounds for Hermite and Popov forms, we may use them to limit our search space. The correctness of this statement is proven below in theorem 21. But we first need to introduce a few notations and definitions which are necessary for the formulation of the algorithm.

For any set $\mathfrak{S} \subseteq R^{1\times n}$ we denote the set of elements of degree at most $d \geqslant 0$ in $\mathfrak{S}$ by $\mathfrak{S}_{\leqslant d} = \{v \in \mathfrak{S} \mid \deg v \leqslant d\}$. Let M be in Hermite or Popov form. We write the set of all those monomials which are not reducible by M as $\mathfrak{B}$. By [5, Proposition 5.6.3] $\overline{\mathfrak{B}} = \{\overline{\mathfrak{m}} \mid \mathfrak{m} \in \mathfrak{B}\}$ is a K-basis of $R^{1\times n}/R^{1\times m}M$ where the bar denotes residue classes modulo M. We would like to emphasise here that $\mathfrak{B}$ depends on the monomial ordering in respect to which M is a Gröbner basis. For any $u \in R^{1\times n}$ we will write the coordinate vector of u in $R^{1\times n}/R^{1\times m}M$ with respect to $\overline{\mathfrak{B}}$ as $u_{\overline{\mathfrak{B}}}$.

The factor module $R^{1\times n}/R^{1\times m}M$ is not only a vector space but also a left R-module. Hence, the multiplication by ∂ induces a map of $R^{1\times n}/R^{1\times m}M$ into itself that we will call $\partial\bullet$. It has the properties that

$$\partial(\overline{v} + \overline{w}) = \partial\overline{v} + \partial\overline{w} \qquad \text{and} \qquad \partial(a\overline{v}) = \sigma(a)\partial\overline{v} + \vartheta(a)\overline{v}.$$

for all v and $w \in R^{1\times n}$ and $a \in K$. Such a map is called *pseudo-linear* in [15].

Fix a degree bound. We will consider the *truncated basis* $\overline{\mathfrak{B}_{\leqslant d}}$. Let π be the projection of $R^{1\times n}/R^{1\times m}M$ onto the K-span $\langle\overline{\mathfrak{B}_{\leqslant d}}\rangle$ of the truncated basis. We introduce the *truncated ∂-multiplication* $\tau = \pi\circ(\partial\bullet)|_{\langle\overline{\mathfrak{B}_{\leqslant d}}\rangle}$ as a map of $\langle\overline{\mathfrak{B}_{\leqslant d}}\rangle$. (The composition with π lets us ignore products which are not in $\langle\overline{\mathfrak{B}_{\leqslant d}}\rangle$ any more.) Let $v = \pi v \in \langle\overline{\mathfrak{B}_{\leqslant d}}\rangle$. Then $\tau(a\pi(v)) = \pi\circ(\partial\bullet)(a\pi(v)) = \sigma(a)\big(\pi\circ(\partial\bullet)\big)(\pi v) + \vartheta(a)\pi^2(v) = \sigma(a)\tau(\pi(v)) + \vartheta(a)\pi(v)$ since $\pi^2 = \pi$. Thus, τ is also a pseudo-linear map. By [15, Section 2] we may construct a matrix $T \in K^{|\mathfrak{B}_{\leqslant d}|\times|\mathfrak{B}_{\leqslant d}|}$ such that

$$\tau(u)_{\overline{\mathfrak{B}_{\leqslant d}}} = \sigma(u_{\overline{\mathfrak{B}_{\leqslant d}}})T + \vartheta(u_{\overline{\mathfrak{B}_{\leqslant d}}})$$

where σ and ϑ are applied to vectors component-wise. The truncated multiplication matrix T is called a *τ-connection* in [7].

Remark 18. Computing T is actually quite easy. If for $\mathfrak{m} \in \mathfrak{B}_{\leqslant d}$ also $\partial\mathfrak{m} \in \mathfrak{B}_{\leqslant d}$, then the row corresponding to $\overline{\mathfrak{m}}$ in T is a unit vector. If otherwise $\partial\mathfrak{m} \notin \mathfrak{B}_{\leqslant d}$, then there are two possibilities. Either $\partial\mathfrak{m} \in \mathfrak{B}$ or $\partial\mathfrak{m}$ is divisible by a row in M.

In the first case the row of $\overline{\mathfrak{m}}$ in T will just be zero. In the second case, there is a leading monomial $\mathfrak{n}$ of a row $M_{i,\bullet}$ of M and $\alpha \geqslant 0$ such that $\partial^\alpha \mathfrak{n} = \partial\mathfrak{m}$. Since $\mathfrak{m}$ is irreducible, we may conclude that $\alpha = 0$, that is, that $\mathfrak{n} = \partial\mathfrak{m}$. Thus, the remainder is $\mathfrak{n} - M_{i,\bullet} \in \mathfrak{B}_{\leqslant d}$ which is irreducible since M is a reduced Gröbner basis. The corresponding row in T is then just $(\overline{\mathfrak{n} - M_{i,\bullet}})_{\overline{\mathfrak{B}_{\leqslant d}}}$. The coordinates may hence be plainly read off from the coefficients in $M_{i,\bullet}$.

In example 7 we already established the correspondence between the pivot indices and the leading monomials in M. This allows us to write down $\mathfrak{B}_{\leqslant d}$ quite easily as

$$\mathfrak{B}_{\leqslant d} = \left\{ \partial^\alpha \mathfrak{e}_j \;\middle|\; j = j_i \in J \wedge \alpha < \operatorname{rdeg}_i M \right\} \cup \left\{ \partial^\alpha \mathfrak{e}_j \middle| j \notin J \wedge \alpha \leqslant d \right\}$$

where $J = \{j_1, \dots, j_m\}$ is the set of all pivot indices. We may compute the coordinates of the residue classes of the unit vectors $\overline{\mathfrak{e}_1}, \dots, \overline{\mathfrak{e}_n}$ in the same way as the we computed T. From them we can compute the residue class of any $\partial^\alpha \mathfrak{e}_k \in \mathfrak{B}_{\leqslant d}$ just by using T.

We are now ready to state the algorithm. Although the only admissible orderings we have considered are the position over term and the term over position ordering, the algorithm would also work for other orderings. We denote lists (that is, ordered sets) by enclosing their elements in square brackets, that is, we write $L = [L_1, \dots, L_k]$. If ℓ is an element, then $\ell : L$ denote the list with its first element being ℓ and then the elements of L following in order, that is, $\ell : L = [\ell, L_1, \dots, L_k]$.

Algorithm 19 (FGLM with degree bound)**.**

Input: A reduced Gröbner basis $M \in R^{m \times n}$ with respect to the admissible ordering $<_1$ and an admissible ordering $<_2$ as well as a degree bound d for the reduced Gröbner basis with respect to $<_2$.

Output: The reduced Gröbner basis with respect to $<_2$.

Procedure:

(1) Let $\mathfrak{B}_1$ be the truncated basis with respect to $<_1$ and d, and let T be the corresponding multiplication matrix.

(2) Initialise $C \leftarrow [\,]$, $\mathfrak{B}_2 \leftarrow [\,]$ and $G_2 \leftarrow \varnothing$.
Upon termination, G_2 will be the reduced Gröbner basis, $\mathfrak{B}_2$ will be the truncated basis with respect to $<_2$ and d, and C will contain the coordinate vectors of the elements of $\mathfrak{B}_2$ with respect to $\overline{\mathfrak{B}_1}$.

(3) If there are monomials of degree less or equal to d that are not divisible by G_2, then:

  (a) Choose the smallest such monomial $\mathfrak{m}$ with respect to $<_2$ and compute its coordinate vector $w = \overline{\mathfrak{m}}_{\overline{\mathfrak{B}_1}}$ using T.

  (b) If $w : C$ is K-linear independent, then set $C \leftarrow w : C$ and $\mathfrak{B}_2 \leftarrow \mathfrak{m} : \mathfrak{B}_2$.

  (c) Else there are $a_j \in K$ such that $w = \sum_j a_j C_j$. Set $G_2 \leftarrow G_2 \cup \{\mathfrak{m} - \sum_j a_j (\mathfrak{B}_2)_j\}$.

  (d) Go to step 3.

(4) Else stop and return G_2 as a matrix with the rows sorted with respect to $<_2$.

Remark 20. If $<_2 = <_{\text{POT}}$, then the sequence of monomials that are chosen in step 3a can be computed as follows. Set $j \leftarrow n$ and start with $\mathfrak{m} \leftarrow \mathfrak{e}_j = \mathfrak{e}_n$ which

is the smallest element. If in step 3b w does not depend on C, then set $\mu \leftarrow \partial\mu$, $w \leftarrow \sigma(w)T + \vartheta(w)$ and go to step 3b. Else, set $j \leftarrow j-1$ and continue with the next $\mathfrak{m}$ being $\mathfrak{e}_j$. The sorting in step 4 can be omitted if G_2 is maintained as a list with new elements added in front.

If $<_2 = <_{\mathrm{TOP}}$, then we maintain a list $\mathfrak{M}$ of monomials which initially is just $[\mathfrak{e}_n, \ldots, \mathfrak{e}_1]$ and a corresponding list of coordinate vectors W. We iterate over $(\mathfrak{m}, w)$ in the zipped list $(\mathfrak{M}, W)$. If in step 3b we find a linear dependence, then we remove $(\mathfrak{m}, w)$ from $(\mathfrak{M}, W)$. Once we are through the list, if $\mathfrak{M} \neq [\,]$ we set $\mathfrak{M} \leftarrow [\partial\mathfrak{m} \mid \mathfrak{m} \in \mathfrak{M}]$ and $W \leftarrow [\sigma(w)T + \vartheta(w) \mid w \in W]$ and continue. Also here, the sorting in step 4 is not necessary if G_2 is a list with the elements added at the end.

Theorem 21. *Algorithm 19 is correct and terminates.*

Proof. The iteration considers only monomials up to certain degree. Since there are only finitely many of them, the algorithm clearly terminates.

It remains to prove the correctness. We will use the notations from the algorithm. First, we note that the elements in C are always linearly independent by construction. Since they are just the $\overline{\mathfrak{B}_1}$-coordinate vectors of the elements in $\mathfrak{B}_2$—and since the coordinate map is K-linear—also $\mathfrak{B}_2$ is linear independent modulo $R^{1\times m}M$.

Moreover, we claim that the elements of G_2 are in $R^{1\times m}M$. Let in step 3c $g = \mathfrak{m} - \sum_j a_j(\mathfrak{B}_2)_j$. Let $r = g - uM$ be the remainder of g by division with M where $u \in R^{1\times m}$ contains the coefficients from theorem 8. We have $\overline{r} = \overline{g - uM} = \overline{g} = w - \sum_j a_j C_j = 0$. Since r is irreducible, this implies $r = 0$, that is, $g \in R^{1\times m}M$.

Let $\mathrm{LM}(G_2) = \{\partial^\alpha \mathfrak{m} \mid \mathfrak{m} \in G_2 \text{ and } \alpha \geqslant 0\}$. We claim that $\mathfrak{B}_2 \cap \mathrm{LM}(G_2) = \varnothing$. This holds in step 2 and cannot be destroyed if we add elements to $\mathfrak{B}_2$ in step 3b. In step 3c, if an element is added to G_2 it is bigger than all elements in $\mathfrak{B}_2$ with respect to $<_2$ since we iterate over all monomials in order. Using the definition of admissible orderings in [5, Defintion 2], we see that it cannot divide any monomial in $\mathfrak{B}_2$. Since we consider all monomials of degree at most d, we obtain

$$\mathfrak{M}_{\leqslant d} := \{\mathfrak{m} \text{ monomial} \mid \deg \mathfrak{m} \leqslant d\} = \mathrm{LM}(G_2)_{\leqslant d} \dot\cup \mathfrak{B}_2.$$

Let $\tilde{G}$ be the Gröbner basis of $R^{1\times m}M$ with respect to $<_2$ and let $\tilde{\mathfrak{B}} \subseteq \mathfrak{M}_{\leqslant d}$ denote the corresponding truncated basis. Since $G_2 \subseteq R^{1\times m}M$, we must have $\tilde{\mathfrak{B}} \subseteq \mathfrak{B}_2$. We claim that $\mathrm{lmonom}(g) \in \mathrm{LM}(G_2)$ for any $g \in \tilde{G}$. By our degree bound, we know that $\mathrm{lmonom}(g) \in \mathfrak{M}_{\leqslant d}$. Assume $\mathrm{lmonom}(g)$ was in $\mathfrak{B}_2$. This meant that we could reduce an element of $\mathfrak{B}_2$ contradicting the linear independence of $\mathfrak{B}_2$ modulo $R^{1\times m}M$. Thus $\mathrm{LM}(\tilde{G}) \subseteq \mathrm{LM}(G_2)$. Hence, by definition 9, G_2 must be a Gröbner basis.

By construction, the leading monomials of G_2 are monic and do not divide each other. Further more, since for each $g \in G_2$ we have $g - \mathrm{lmonom}(g) \in \langle \mathfrak{B}_2 \rangle$, we see that g is irreducible by $G_2 \setminus \{g\}$. Thus, G_2 is the unique reduced Gröbner basis of $R^{1\times m}M$ with respect to $<_2$. □

Corollary 22 (Main theorem)**.** *Because of the degree bound in the lemmata 16 and 17, we may use algorithm 19 to convert Hermite forms into Popov form and vice versa.*

Proof. Let $H \in R^{m\times n}$ be in Hermite form and assume $P \in R^{s\times n}$ is the output of algorithm 19. Then P is in Popov form and using Gröbner basis division we may compute $A \in R^{m\times s}$ such that $H = AP$. Since also H is a Gröbner basis we can find $B \in R^{s\times m}$ such that $BH = P$. Now, since H and P have linearly independent rows by remark 5, we conclude $AB = \mathbf{1}_m$ and $BA = \mathbf{1}_s$. By [16, p. 32] (applicable since by [10, Theorem 5.8] Ore polynomials can be embedded in skew fields) this implies $m = s$ and hence $A = B^{-1} \in \mathrm{GL}(R, s)$. Thus, P really is the Popov form of H. Analogously, also for a Popov form as input we receive the corresponding Hermite form. □

Finally, we would like to reason about the complexity of algorithm 19. We will consider only the conversion from Popov to Hermite form. In the steps 1 and 2 there is not much to do, since the computation of T involves just the copying of the coefficients of M by remark 18. The real work is done in step 3. Here, we have to compute all the candidates for leading monomials and their coordinate vectors, and we have to check sets of monomials for linear dependence. Let $d = \deg M$. The degree bound is md in this case. The number of monomials generated (and also the size of $\mathfrak{B}_1$) does thus not exceed $\mathcal{O}(nmd)$. To generate a monomial we either look it up from a list containing the unit vectors and their coordinates (as can be precomputed analogously to T) or by remark 20 we compute it as a product with ∂ and the previous monomial. In the later case to compute the coordinates we need $\mathcal{O}(mnd)$ applications of σ and ϑ and $\mathcal{O}((mnd)^2)$ multiplications and additions in K for the multiplication by T. The most expensive step is to solve the $\mathcal{O}(nmd)$ variables system in step 3b which needs $\mathcal{O}((nmd)^3)$ operations in K by [17, Bemerkung 2.19 (2)]. Since $\mathfrak{B}_2$ contains only (different) monomials, computation of $\mathfrak{m} - \sum_j a_j(\mathfrak{B}_2)_j$ is again just copying coefficients.

The estimate becomes tighter if M is a square matrix. Then, the degree bound is never needed because there will be a pivot in every row of M. Hence, $R^{1\times n}/R^{1\times m}M$ is finite. This corresponds to the case of zero-dimensional ideals in the theory of commutative polynomials. We need to consider at most $\mathcal{O}(md)$ monomials. This bound can even be lowered using the *index* of M which is $\operatorname{ind} M = \sum_i \operatorname{rdeg}_i M$ as introduced in [12]. This yields a total complexity of $\mathcal{O}((\operatorname{ind} M)^4)$.

Remark 23 (Complexity)**.** For the conversion of a Hermite form in $M \in R^{m\times n}$ into Popov form one needs $\mathcal{O}((mnd)^4)$ operations in K where $d = \deg M$. If M is square, then $\mathcal{O}((\operatorname{ind} M)^4) \leqslant \mathcal{O}((md)^4)$ operations are sufficient.

6. Conclusion

In this paper we have extended the result of [18] that Hermite and Popov forms are Gröbner bases to a general Ore polynomial setting. We adapted the classical FGLM algorithm for this case and used it to convert matrices from Hermite form into Popov form and vice versa. The complexity of this is polynomial and not to far away from other approaches as for example [24]. The version presented here is slightly more general though as it works with arbitrary Ore polynomials.

References

[1] Adams, W. W., and Loustaunau, P. *An introduction to Gröbner bases.* Graduate studies in mathematics. AMS, 1994.

[2] Beckermann, B., Cheng, H., and Labahn, G. Fraction-free row reduction of matrices of Ore polynomials. *Journal of Symbolic Computation 41* (2006), 513 – 543.

[3] Bronstein, M., and Petkovšek, M. An introduction to pseudo-linear algebra. *Theoretical Computer Science 157, 3-33 157* (1996), 3–33.

[4] Buchberger, B. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal (An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal)*. PhD thesis, Mathematical Institute, University of Innsbruck, Austria, 1965. (English translation to appear in Journal of Symbolic Computation, 2004).

[5] Bueso, J. L., Gómez-Torrecillas, J., and Verschoren, A. *Algorithmic methods in non-commutative algebra*, vol. 17 of *Mathematical modelling: Theory and applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003.

[6] Cheng, H. *Algorithms for normal forms for matrices of polynomials and Ore polynomials*. PhD thesis, University of Waterloo, 2003. Advisor: George Labahn.

[7] Churchill, R. C., and Kovacic, J. J. Cyclic vectors. In *Differential algebra and related topics* (2002), L. Guo, P. J. Cassidy, W. F. Keigher, and W. Y. Sit, Eds., World Scientific Publishing Co. Pte. Ltd., pp. 191–218.

[8] Chyzak, F., and Salvy, B. Non-commutative elimination in Ore algebras proves multivariate identities. *Journal of Symbolic Computation 26*, 2 (1998), 187–227.

[9] Cohn, P. M. *Free rings and their relations*, 2nd edition ed. Academic press inc. (London) Ltd, 1985.

[10] Cohn, P. M. *An introduction to ring theory*. Springer, Berlin Heidelberg New York, 2000.

[11] Faugère, J.-C., Gianni, P. M., Lazard, D., and Mora, T. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symb. Comput. 16*, 4 (1993), 329–344.

[12] Forney jr., G. D. Minimal bases of rational vector spaces with applications to multivariable linear systems. *SIAM J. Control 13* (May 1975), 493 – 520.

[13] Giesbrecht, M., and Kim, M. S. *Computer Algebra in Scientific Computing*, vol. 5743 of *Lecture Notes in Computer Science*. Springer, Berlin / Heidelberg, 2009, ch. On Computing the Hermite Form of a Matrix of Differential Polynomials, pp. 118–129.

[14] Hermite, C. Sur l'introduction des variables continues dans la théorie des nombres. *Journal der reinen und angewandten Mathematik*, 41 (1851), 191–216.

[15] Jacobson, N. Pseudo-linear transformations. *The Annals of Mathematics 38*, 2 (1937), 484–507.

[16] Jacobson, N. *The theory of rings*, vol. 2 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1943.

[17] Kiyek, K.-H., and Schwarz, F. *Mathematik für Informatiker*, vol. 1. Teubner, 1989.

[18] Kojima, C., Rapisarda, P., and Takaba, K. Canonical forms for polynomial and quadratic differential operators. *System & Control Letters* (2007), 678–684.

[19] Middeke, J. Converting between the Popov and the Hermite form of matrices of differential operators using an FGLM-like algorithm. Tech. Rep. 10-16, RISC Report Series, University of Linz, Austria, 2010.

[20] Mora, F., and Möller, H. New constructive methods classical ideal theory. *Journal of Algebra 100*, 1 (1986), 138–178.

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

[22] Popov, V. M. Some properties of the control systems with irreducible matrix-transfer functions. In *Seminar on Differential Equations and Dynamical Systems, II*, Lecture Notes in Mathematics. Springer, Berlin / Heidelberg, 1970, pp. 169–180.

[23] Popov, V. M. Invariant description of linear, time-invariant controllable systems. *SIAM Journal on Controll*, 2 (May 1972), 252–264.

[24] Villard, G. Computing Popov and Hermite forms of polynomial matrices. In *ISSAC* (1996), pp. 250–258.

Research Institute for Symbolic Computation (RISC), Linz, Austria
E-mail address: jmiddeke@risc.jku.at