Converting between the Popov and the Hermite form of matrices of differential operators using an FGLM-like algorithm

Johannes Middeke[1]

Research Institute for Symbolic Computation (RISC)
Johannes Kepler University Linz
Austria
Email: jmiddeke@risc.jku.at

May 2010

RISC Technical Report

[1] This work was supported by the Austrian Science Foundation (FWF) under the project DIFFOP (P20 336–N18).

Abstract

We consider matrices over a ring $K[\partial; \sigma, \vartheta]$ of Ore polynomials over a skew field K. Since the Popov and Hermite normal forms are both Gröbner bases (for term over position and position over term ordering resp.), the classical FGLM-algorithm provides a method of converting one into the other. In this report we investigate the exact formulation of the FGLM algorithm for not necessarily "zero-dimensional" modules and give an illustrating implementation in Maple$^{\text{TM}}$. In an additional section, we will introduce a second notion of Gröbner bases roughly following [Pau07]. We will show that these vectorial Gröbner bases correspond to row-reduced matrices.

Contents

1

Introduction

When considering systems of ordinary linear differential or difference equations it is natural to write them in matrix form. For example would the system

$$\begin{aligned} -2f_1 + \frac{df_2}{dx} + \frac{3}{2}xf_2 - 3f_3 - \frac{3}{2}x^2 f_3 &= 0 \\ \frac{1}{2}f_2 + \frac{df_3}{dx} - \frac{1}{2}xf_3 &= 0 \end{aligned}$$

be written as

$$\begin{pmatrix} -2 & d/dx + \frac{3}{2}x & -3 - \frac{3}{2}x^2 \\ 0 & \frac{1}{2} & d/dx - \frac{1}{2}x \end{pmatrix} \bullet \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = 0$$

where the bullet indicates application. We will study the normal forms of matrices as the one above. We will not explicitly consider how to solve such a system.

We will model derivations algebraically by using so-called Ore polynomials. These have first been considered by Øystein Ore in [Ore33]. They are a generalisation of ordinary polynomials which offers a unified way of describing linear differential and difference operators. They share many properties with the ordinary polynomials but since they model operators, their multiplication is not commutative. We will introduce Ore polynomials in Section 2.1 and summarise those properties there that are important for the rest of this report.

We will introduce matrices over Ore polynomials in Section 2.2. That section contains all the basic definitions and notations that will be used later.

As in the commutative theory, normal forms are an important tool for the analysis of matrices and the systems they represent. This report will concentrate on two normal forms, namely the Hermite form and the Popov form. Both are normal forms for left equivalence, i. e., for matrix M there are unimodular matrices U and V such that the product UM is in Hermite form and VM is in Popov form. (This will be made precise later).

The Hermite normal form was first introduced and proved to exist by Charles Hermite in [Her51] for non-singular square matrices over the integers. It was later extended to more general matrices. Its main application is solving of Diophantine equations. We will introduce the Hermite form in Definition 12 in Section 2.3.

The Popov normal form was first described by Vasile Mihai Popov in [Pop70, Pop72]. Together with similar concepts like row-reduction that is sometimes also called row-properness it is widely used in control theory—see for example [Zer07]. We will treat the Popov normal form in Definition 13.

Gröbner bases have been first mentioned in [Buc65]. Devised originally for ideals of (commutative) multivariate polynomials they have since then been extended to non-commutative domains and also to (free) modules over them. See for example [CS98] for a treatise and an application of Gröbner bases for Ore polynomials. An introduction to extensions of Gröbner bases to modules may be found in the textbooks [AL94] for the commutative case and similarly in [BGTV03] for non-commutative domains. We will use the later book extensively in this report.

Gröbner bases are important in modern day computational algebra because they provide two useful features: First, they solve the *ideal membership problem*. That is, they provide an algorithmic way to check whether a given polynomial is contained in a given ideal. Second, Gröbner bases have an *elimination property* which makes them useful for solving systems of polynomial equations. The same properties carry over to the non-commutative and to the module case.

In [KRT07] it was shown that Popov and Hermite normal forms of matrices with (univariate) polynomial entries are actually Gröbner bases for their row span. This result can easily be generalised to Ore polynomial rings. We do this in Theorems 25 and 27.

The *FGLM algorithm* is an efficient method to convert Gröbner bases of zero-dimensional ideals from one admissible ordering to another. It was presented in [FGLM93] for ideals of commutative polynomials. The main reason for its efficiency is that it manages to translate the problem from polynomial to linear algebra. The FGLM algorithm may lead to a speed-up of Gröbner basis computations for slow orderings like the lexicographic ordering by first computing a Gröbner basis with respect to a faster ordering like a degree ordering and then converting it to the desired ordering. A short introduction to the FGLM algorithm is given in Section 3.3.1

We will translate the original algorithm from the case of commutative ideals to modules over non-commutative domains in Section 3.3.4. There, we will in particular deal with the problem that the modules we consider are not "zero-dimensional", i. e., that their quotient spaces can be of infinite dimension. We will be able to solve this using a degree bound on Hermite and Popov normal forms that is similar to the results in [GK09].

A different though less powerful approach of defining Gröbner bases for modules may be derived from the theory of Gröbner bases for rings. We will investigate a possible translation of [Pau07] to modules in Section 3.2. It will turn out that this kind of Gröbner bases—which we will call vectorial Gröbner bases in order to avoid confusion—naturally correspond to row-reduction.

In [Vil96] a method for converting Popov forms into Hermite forms for non-singular square matrices over the commutative polynomials $K[x]$ is presented. There, the Popov form is used as an intermediate step for the computation of the Hermite form. The paper uses *Hankel matrices* and *matrix fraction representations* that are also mentioned in [Pop70] and are a standard tool in control theory. A polynomial time algorithm for computing a Hermite normal form directly without intermediate use of a Popov form can be found for example in [GK09].

Furthermore we will implement those parts of this report in the computer algebra system Maple$^{\text{TM}}$ that are necessary for the conversion of Popov forms to Hermite forms. This is just an illustration for the ideas used here; and we do not claim particular efficiency or correctness. The code will always follow the corresponding definitions. Thus, these implementations are scattered throughout this report, but we believe that it makes the individual procedures easier to check. A complete version of the programme may be obtain via email to the author.

2

Basic definitions

2.1 Ore polynomials

Ore polynomials—that are also called skew polynomials by some authors—were first investigated by Øystein Ore in [Ore33]. They are a generalisation of the usual polynomials with which they share most properties except that multiplication needs not to be commutative. As shown in the examples below, Ore polynomials contain important classes of operators such as differential operators and shift (time-delay) operators.

Informally, Ore polynomials can be defined as follows: Let K be a skew field and consider polynomial expressions in the variable ∂ of the form

$$a_n \partial^n + \ldots + a_1 \partial + a_0 \qquad \text{where} \quad a_0, \ldots, a_n \in K.$$

Note, that we write the coefficients on the left hand side of the variable. Put in other words, the Ore polynomials over K are the free left K-module generated by the powers of ∂. Addition and (left) scalar multiplication are just the usual coefficient-wise addition and scalar multiplication of polynomials. But the multiplication is more general, since we do not want to enforce commutativity. To make things a little easier, however, we will require that the degree formula

$$\deg(pq) = \deg p + \deg q$$

holds for all Ore polynomials p and q, where the degree is defined in the usual way as the exponent of the highest power of ∂ occurring with a non-zero coefficient. In order to eliminate case distinctions, we set the degree of the zero Ore polynomial to $\deg 0 = -\infty$. We define the i^{th} coefficient $\operatorname{coeff}(i, p)$ of $p = a_n \partial^n + \ldots + a_1 \partial + a_0$ to be a_i if $0 \leqslant i \leqslant n$ and to be 0 otherwise. Using the degree and the coefficient functions, we also define the *leading coefficient* of an Ore polynomial p to be $\operatorname{lc}(p) = \operatorname{coeff}(\deg p, p)$.

The degree formula implies in particular that for all $a \in K$ there exist $\sigma(a)$ and $\vartheta(a) \in K$ such that

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

Sometimes we will call the above formula the *multiplication rule*. Since the powers of ∂ form a module basis for the Ore polynomials, for each $a \in K$ the elements $\sigma(a)$ and $\vartheta(a)$ are unique, i. e.,

$\sigma\colon K \to K$ and $\vartheta\colon K \to K$ are mappings. Equating the coefficients of $\partial 1 = 1\partial$, $\partial(a+b) = \partial a + \partial b$ and $\partial(ab) = (\partial a)b$ for a and $b \in K$, one sees that σ is an endomorphism and that ϑ is additive and fulfils a kind of Leibniz rule. This leads to the following definition:

Definition 1 (Derivation)**.** Let K be a skew field and let $\sigma\colon K \to K$ be an endomorphism. An additive mapping $\vartheta\colon K \to K$ is called a *σ-derivation* if ϑ fulfils the *σ-Leibniz rule*

$$\vartheta(ab) = \sigma(a)\vartheta(b) + \vartheta(a)b$$

for all a and $b \in K$.

If $\sigma = \mathrm{id}$ is the identity function, then ϑ is usually just called a *derivation*: Indeed, for $\sigma = \mathrm{id}$ we regain the Leibniz rule $\vartheta(ab) = a\vartheta(b) + \vartheta(a)b$.

Now, we can give a definition of Ore polynomials. We still remain informal here; a precise description of Ore polynomials may be found for example in [Coh85, Section 0.10].

Definition 2 (Ore polynomials)**.** Let K be a ring with endomorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$. The *(left) Ore polynomials* are the set

$$K[\partial;\sigma,\vartheta] = \{a_n\partial^n + \ldots + a_1\partial + a_0 \mid a_0,\ldots,a_n \in K\}$$

together with the usual addition and scalar multiplication and with the multiplication given by the multiplication rule $\partial a = \sigma(a)\partial + \vartheta(a)$ for all $a \in K$.

It remains to ensure that such a structure may indeed be constructed and that it does satisfy the ring axioms. For this we merely cite a result from the literature.

Theorem 3 ([Coh00, Theorem 5.7])**.** *Let K be an integral domain*[1] *with endomorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$. Then there is a Ore polynomial ring $K[\partial;\sigma,\vartheta]$ with a degree function* deg *satisfying the conditions*

1. $\deg f \in \mathbb{N}$ *for* $f \neq 0$*, while* $\deg 0 = -\infty$*;*
2. $\deg(f-g) \leqslant \max\{\deg f, \deg g\}$*;*
3. $\deg(fg) = \deg f + \deg g$*.*

Moreover, $K[\partial;\sigma,\vartheta]$ is an integral domain.

It is easy to see that the multiplication rule extends to Ore polynomials—i. e., that $\partial f = \sigma(f)\partial + \vartheta(f)$ for $f \in K[\partial;\sigma,\vartheta]$—if one applies σ and ϑ coefficient-wise. In particular, $\mathrm{lc}(\partial f) = \sigma(\mathrm{lc}(f))$. Using this, one may inductively prove that for all $f, g \in K[\partial;\sigma,\vartheta] \setminus \{0\}$ the *leading coefficient formula*

$$\mathrm{lc}(fg) = \mathrm{lc}(f)\sigma^{\deg f}(\mathrm{lc}(g))$$

holds.

We will close this section with some examples of Ore polynomials. There are more examples in the literature. For instance, the reader might have a look at [Che03] and [CS98].

Example 4. 1. Let K be arbitrary. Using $\sigma = \mathrm{id}$ and $\vartheta = 0$ (which is obviously a derivation) the multiplication rule becomes just $\partial a = a\partial$. Thus

$$K[\partial;\mathrm{id},0] \cong K[x],$$

and the usual commutative polynomials are just a special case of Ore polynomials.

[1] Integral domains are not necessarily commutative in [Coh00].

2. For a differential skew field[2] (K, ϑ) the ring $K[\partial; \mathrm{id}, \vartheta]$ is called the ring of *differential operators* over K. In this case the multiplication rule $\partial a = a\partial + \vartheta(a)$ is sometimes also called Leibniz rule. The name differential operator arises from the following special case:

   Take $K = \mathbb{Q}(x)$ and $\vartheta = d/dx$, then for a fixed $f \in \mathbb{Q}(x)$ and arbitrary $g \in \mathbb{Q}(x)$ we obtain
   $$(\tfrac{d}{dx}f)g = \tfrac{d}{dx}(fg) = f\tfrac{d}{dx}g + \tfrac{df}{dx}g = (f\tfrac{d}{dx} + \tfrac{df}{dx})g,$$
   and hence we obtain $(d/dx)f = (fd/dx + df/dx)$ as operators.

   If we take $\mathbb{Q}[x]$ instead of a skew field then the ring $\mathbb{Q}[x]\llbracket\partial; \mathrm{id}, d/dx]$ is called the *first Weyl algebra*.

3. In the case that $\sigma \neq \mathrm{id}$ but $\vartheta = 0$ we obtain the *shift operators* (which are also called *time delay operators*, twisted polynomials or skew polynomials by some authors). The Leibniz rule becomes $\partial a = \sigma(a)\partial$.

   The name shift polynomials comes from the special case of $K = \mathbb{Q}(x)$ where σ is the substitution automorphism $p \mapsto p(x + 1)$.

As we show below in Lemma 5, the three examples given above are in a sense all examples one can find. Let K be a skew field, $q \in K$ and let $\sigma \colon K \to K$ be an endomorphism. Then the map
$$\vartheta_q : x \mapsto \sigma(x)q - qx$$
is a σ-derivation, the so called *inner σ-derivation induced by q*. Indeed, ϑ_q is obviously additive and for all $a, b \in K$ we have
$$\sigma(a)\vartheta_q(b) + \vartheta_q(a)b = \sigma(a)(\sigma(b)q - qb) + (\sigma(a)q - qa)b = \sigma(ab)q - qab = \vartheta_q(ab).$$
There is also the notion of an *inner endomorphism*: An endomorphism $\varphi \colon K \to K$ is called inner, if there is a unit $u \in K^*$ such that
$$\varphi(a) = uau^{-1}$$
for all $a \in K$.

Lemma 5 ([Coh85, Proposition 8.3.1]). *Let K be a skew field[3] with endomorphism $\sigma \colon K \to K$ and σ-derivation $\vartheta \colon K \to K$. For the Ore polynomial ring $K[\partial; \sigma, \vartheta]$ there are three possibilities:*

1. *Either ϑ is an inner derivation and after a suitable change of the variable ∂ we may assume $\vartheta = 0$, or*

2. *σ is an inner endomorphism and by a suitable choice of ∂ we may assume $\sigma = \mathrm{id}$, or*

3. *σ leaves the centre C of K fixed and δ maps it to 0; in that case C is contained in the centre of R.*

[2] A differential ring (A, δ) is just a ring A together with a derivation $\delta \colon A \to A$.

[3] The formulation in the reference is just "field" instead of "skew field", but as noted in on page xviii, in [Coh85] fields are not necessarily considered to be commutative.

2.2 Matrices over Ore polynomials

Let K be a skew field with automorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$, and let $R = K[\partial; \sigma, \vartheta]$ be the corresponding ring of Ore polynomials. For positive integers $m, n \in \mathbb{N} \setminus \{0\}$ we denote by ${}^{m}R^{n}$ the set of $m \times n$-matrices with entries in R. As special cases of matrices we have the free (left) R-module of row vectors ${}^{1}R^{n} = R^{n}$ and the free (right) R-module of column vectors ${}^{m}R^{1} = {}^{m}R$. A square matrix $U \in {}^{m}R^{m}$ that has a two-sided inverse $U^{-1} \in {}^{m}R^{m}$ is called *unimodular*. The set of all unimodular $m \times m$-matrices is denoted by $({}^{m}R^{m})^{*}$. Furthermore, we will denote the $m \times m$ identity matrix by $\mathbf{1}_{m}$ and the $m \times n$ zero matrix by ${}_{m}\mathbf{0}_{n}$.

Let $M \in {}^{m}R^{n}$ be arbitrary. We denote the element at position (i, j) of M by $M_{i,j}$. We use the notation $M_{j,\bullet}$ for the j^{th} row of M where $1 \leqslant j \leqslant m$, and analogously we write $M_{\bullet,k}$ for the k^{th} column of M for $1 \leqslant k \leqslant n$. With $R^{m}M$ we denote the *row space* of M which we consider as left R-module. On the other hand, $M\,{}^{n}R$ denotes the *column space* of M which is a right R-module.

We need a few more definitions which are used below for defining the main object of our study, the Hermite and the Popov normal forms. First, we extend the notation of degree to matrices: Let $M \in {}^{m}R^{n}$ be a matrix; then we set

$$\deg M = \max\{\deg M_{i,j} \mid 1 \leqslant i \leqslant m \text{ and } 1 \leqslant j \leqslant n\}.$$

In particular, for $1 \leqslant i \leqslant n$ we define the *i^{th} row degree* to be $\operatorname{rdeg}_{i} M = \deg M_{i,\bullet}$. Similarly, we extend the notion of coefficients and leading coefficients. That is, we set

$$\operatorname{coeff}(i, M) = \Big(\operatorname{coeff}(i, M_{j,k})\Big)_{j,k} \in {}^{m}K^{n}$$

and we define the *leading vector* of M to be $\operatorname{lv}(M) = \operatorname{coeff}(\deg M, M)$.[4] A special case of the last formula gives $\operatorname{lv}(M_{i,\bullet}) = \operatorname{coeff}(\operatorname{rdeg}_{i} M, M_{i,\bullet})$.

We shortly give an implementation of these definitions in the computer algebra system Maple™. Since Maple™ forces us to distinguish between vectors and matrices we will start with a procedure for the degree $\deg v$ of a (row) vector v. It expects as parameters a symbol Q that will denote how the variable ∂ is represented in Maple™.[5] The second argument is the vector. The procedure may produce erroneous output if the entries of v are not expanded.

```
VectorDegree := proc(Q::symbol, v::Vector) :: extended_numeric:
description "Computes the degree of v w.r.t. the variable Q.":
    return max(map(p → degree(p,Q), v)):
end proc:
```

The definition of vector degrees can be expanded easily to matrices by first computing all row degrees and than selecting the maximum value. The parameters of this procedure are similar to the preceding one.

```
MatrixDegree := proc(Q::symbol, M::Matrix) :: extended_numeric:
description "Computes the degree of M w.r.t. the variable Q.":
local m,n:
    m,n := LinearAlgebra:-Dimension(M):
    return max(seq(RowDegree(Q,j,M),j=1..m)):
end proc:
```

[4] We chose the name leading vector since it will mostly be applied to vectors (which we consider as matrices with only one row or column) and since we will also need a concept of *leading terms* in Section 3.1 that is defined differently.

[5] We would have liked to call the parameter D—but this causes trouble as D is a built-in function in Maple™.

Specialising the above code a little bit, we may also compute the j^{th} row degree $\operatorname{rdeg}_j M$ of a matrix M using the following procedure. The parameters are analogous to the former two procedures. The ASSERT statement will check the validity of the input if the `assertlevel` in the Maple™ `kernelopts` is set to 1.

```
RowDegree := proc(Q::symbol, j::posint, M::Matrix) :: extended_numeric:
description "Computes the j-th row degree of M w.r.t. the variable Q.":
local m,n:
    m,n := LinearAlgebra:-Dimension(M):
    ASSERT(j ⩽ m):
    return VectorDegree(Q,M[j,1..n]):
end proc:
```

While the coefficients and degrees we just defined are sometimes useful, we will need another notation to refer to a special kind of matrix obtained from M. For this we take the

Definition 6 (Leading row coefficient matrix)**.** Let $M \in {}^mR^n$ be a matrix. We define

$$\operatorname{LC}_{\text{row}}(M) = \Big(\sigma^{\deg M - \operatorname{rdeg}_i M}(\operatorname{coeff}(\operatorname{rdeg}_i M, M_{i,j}))\Big)_{i,j} \in {}^mK^n.$$

In other words, the i^{th} row of $\operatorname{LC}_{\text{row}}(M)$ will be $\sigma^{\deg M - \operatorname{rdeg}_i M} \operatorname{lv}(M_{i,\bullet})$.

One might also consider a row-wise definition of $\operatorname{LC}_{\text{row}}(M)$: The j^{th} row will just be $\operatorname{LC}_{\text{row}}(M)_{j,\bullet} = \sigma^{\deg M - \operatorname{rdeg}_j M}(\operatorname{lv}(M_{j,\bullet}))$. Just another way of defining the leading row coefficient matrix is to bring all rows of M to the same degree $\deg M$ by multiplying it from the left with the diagonal matrix $E = \operatorname{diag}(\partial^{\deg M - \operatorname{rdeg}_1 M}, \ldots, \partial^{\deg M - \operatorname{rdeg}_m M})$—where we have to ignore the zero rows of M—and to define $\operatorname{LC}_{\text{row}}(M) = \operatorname{lv}(EM)$. This last definition can be found in [Che03, Page 25]. As Definition 6 is a little bit intimidating at first sight, we will give an example of its computation.

Example 7. Let $K = \mathbb{Q}(x)$, σ the substitution automorphism $x \mapsto x+1$ and $\vartheta = 0$. Let $R = K[\partial; \sigma, \vartheta]$ be the shift operators, and consider the matrix

$$M = \begin{pmatrix} x\partial^2 - \partial + 1 & (x+1)\partial & \partial - x \\ x\partial - 1 & \partial + 1 & x + 1 \\ 0 & 0 & 0 \end{pmatrix} \in {}^3R^3.$$

The degree of M is $\deg M = 2$ since the highest power of ∂ occurring is ∂^2 at position $(1,1)$. The row degrees are $\operatorname{rdeg}_1 M = 2$, $\operatorname{rdeg}_2 M = 1$ and $\operatorname{rdeg}_3 M = -\infty$. Hence the first row of $\operatorname{LC}_{\text{row}}(M)$ will be the leading vector of the first row of M with σ applied zero times, and the second row is the leading vector of the second row of M with σ applied once. The third row of $\operatorname{LC}_{\text{row}}(M)$ is zero since $\operatorname{coeff}(-\infty, p) = 0$ for every $p \in R$ by our convention. Thus, we obtain

$$\operatorname{LC}_{\text{row}}(M) = \begin{pmatrix} x & 0 & 0 \\ x+1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in {}^mK^n.$$

The definition of leading row coefficient matrices is important in order to define row-reducedness. Being row-reduced is an requirement for being in Popov normal form—see Definition 13 below. Furthermore, row-reduced matrices have the so-called predictable degree property which we explain in Lemma 10. Since we are working over a non-commutative field, we have to decide whether we want to consider the left or the right row rank; and because we will always consider the row space to be a left R-module, we choose the *left row rank* which equals the right column rank—see [Art71, Theorem 4].

Definition 8 (Row-reducedness)**.** A matrix $M \in {}^{m}R^{n}$ is called *row-reduced* if $\mathrm{LC}_{\mathrm{row}}(M)$ has (left row) rank m.

The matrix in Example 7 was not row-reduced since its leading row coefficient matrix had zero rows. However, every matrix can be brought to a row-reduced form if we omit zero rows. More precisely, [BCL06, Theorem 2.2] holds:[6]

Theorem 9. *Let $M \in {}^{m}R^{n}$. Then there is a unimodular matrix $U \in ({}^{m}R^{m})^{*}$ such that*

$$UM = \begin{pmatrix} T \\ \hdashline \mathbf{0} \end{pmatrix}$$

where T is row-reduced. Furthermore, $\mathrm{rdeg}_j \left(\begin{smallmatrix} T \\ 0 \end{smallmatrix}\right) \leqslant \mathrm{rdeg}_j M$ for all j up to permutation of the rows and

$$\mathrm{rdeg}_j U \leqslant \nu_j + \sum_{k=1}^{m} (\mu_k - \nu_k) - \min\{\mu_k \mid k = 1, \ldots, m\}$$

where $\mu_j = \max\{0, \mathrm{rdeg}_j M\}$ and $\nu_j = \max\{0, \mathrm{rdeg}_j UM\}$.

We include a short summary of the algorithmic proof of [BCL06] since we want to point out the similarities to the division in the sense of vectorial Gröbner bases which we will define in Section 3.2. Because we need only to consider the non-zero rows of M, we will without loss of generality assume that M does not have zero rows. We will show that if M is not row-reduced then there exist a unimodular matrix $U \in ({}^{m}R^{m})^{*}$ and a row index j such that the jth row-degree of UM is smaller than that of M.

If M is not row-reduced then there must exist a vector $v \in K^{m} \setminus \{0\}$ such that $v\,\mathrm{LC}_{\mathrm{row}}(M) = 0$. We will lift the relation v of the rows of $\mathrm{LC}_{\mathrm{row}}(M)$ to those of M. For this we first choose a row index j such that $v_j \neq 0$ and $\mathrm{rdeg}_j M$ is maximal. Then we define the vector $\tilde{v} \in R^{m}$ via $\tilde{v}_i = \sigma^{\mathrm{rdeg}_j M - \deg M}(v_j)\partial^{\mathrm{rdeg}_j M - \mathrm{rdeg}_i M}$ if $M_{j,\bullet} \neq 0$ and $\tilde{v}_j = 0$ otherwise. The jth row of the product $\tilde{v}M$ will then have the leading vector $\mathrm{lv}((vM)_{j,\bullet}) = \sigma^{\mathrm{rdeg}_j - \deg M}(v_j) \cdot \sigma^{\mathrm{rdeg}_j M - \mathrm{rdeg}_i M}(\mathrm{lv}(M_{j,\bullet}))$; and hence if we apply $\sigma^{\deg M - \mathrm{rdeg}_j M}$ to $\tilde{v}M$ we see that the coefficient of $\partial^{\mathrm{rdeg}_j M}$ must vanish. Hence, $\mathrm{rdeg}_j M > \mathrm{rdeg}_j \tilde{v}M$.

Since $\tilde{v}_j = v_j \in K \setminus \{0\}$, we may extend v to the invertible matrix

$$\left(\begin{array}{ccc:c:ccc} & & & 0 & & & \\ & \mathbf{1}_{j-1} & & \vdots & & {}_{j-1}\mathbf{0}_{j-1} & \\ & & & 0 & & & \\ \hdashline \tilde{v}_1 & \cdots & \tilde{v}_{j-1} & \tilde{v}_j & \tilde{v}_{j+1} & \cdots & \tilde{v}_m \\ \hdashline & & & 0 & & & \\ & {}_{m-j-1}\mathbf{0}_{m-j-1} & & \vdots & & \mathbf{1}_{m-j-1} & \\ & & & 0 & & & \end{array}\right) \in ({}^{m}R^{m})^{*}.$$

The product UM will have a smaller degree than M in the jth row. (In the language of Section 3.2 we will say that the jth row of M got reduced with respect to to the other rows.) Obviously this cannot be repeated infinitely often; and hence we must eventually reach a row-reduced matrix.

We want to close this section with some important properties of row-reduced matrices. The following lemma is well-known for matrices of ordinary (commutative) polynomials—see for example

[6] Although the authors of [BCL06] consider only commutative fields, it is easy to see that the proof carries over to non-commutative coefficient domains.

[For75] for a treatment. For a reference in the Ore polynomial case including a proof see [BCL06, Lemma A.1(a)].

Lemma 10 (Predictable degree property). *A matrix $M \in {}^mR^n$ is row-reduced if and only if for all $v \in R^m$ we have*

$$\deg(vM) = \max\{\deg v_i + \operatorname{rdeg}_i M \mid 1 \leqslant i \leqslant m\}.$$

The following statement is a direct consequence of the predictable degree property.

Corollary 11. *If $M \in {}^mR^n$ is row-reduced then the rows of M are linearly independent.*

Proof. Assume $vM = 0$ for some $v \in R^m$. Since a row-reduced matrix cannot have zero rows, the predictable degree property (Lemma 10) implies that $v = 0$. □

2.3 Hermite and Popov normal forms

We now introduce the main objects we will be working with in this paper. Let in this whole section K be a skew field with automorphism $\sigma\colon K \to K$ and σ-derivation $\vartheta\colon K \to K$; and let $R = K[\partial; \sigma, \vartheta]$ be the corresponding Ore polynomial ring. The following definitions are adopted from [GK09, Definition 2.3] and [KRT07, Definition 1].

The Hermite normal form is essentially an upper triangular form with some additional restrictions that are similar to the integer case—see for example [Coh93] or [GK09, Definition 3.2]. In order to make the definitions more symmetric and to avoid case distinctions later, in contrast to some other authors we disallow zero rows in both Hermite and Popov normal forms.

Definition 12 (Hermite form). A matrix $M \in {}^mR^n$ is in *Hermite form* if there exist column indices $j_1 > j_2 > \ldots > j_m$ which we call *pivot indices* such that

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

The Popov normal form is essentially a row-reduced matrix with some additional requirements on the degrees of the entries that will make it unique.

Definition 13 (Popov form). A matrix $M \in {}^mR^n$ is said to be in *Popov 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) M_{i,j_i} is monic and $\deg M_{i,j_i} = \operatorname{rdeg}_i M$;
   (b) $\deg M_{i,k} < \operatorname{rdeg}_i M$ if $k < j_i$;
   (c) $\deg M_{k,j_i} < \operatorname{rdeg}_i M$ if $k \neq i$; and
   (d) if $\operatorname{rdeg}_i M = \operatorname{rdeg}_k M$ and $i < k$ then $j_i < j_k$ (i. e., pivot indices are ordered increasingly).

The Popov form is sometimes also called *row-echelon form*. It may also be described with the help of the leading row coefficient matrix that was defined above:

Lemma 14. *If $M \in {}^mR^n$ is in Popov form then—up to permutation of the rows—its leading row coefficient matrix $\operatorname{LC}_{\text{row}}(M)$ is in row echelon form.*

Proof. Let M be in Popov form. Since we are allowed to permute the rows, in this proof we will assume that the rows are ordered with respect to to the pivot indices. For the k^{th} row of M the pivot index j_k corresponds to an entry M_{k,j_k} of degree $\deg M_{k,j_k} = \operatorname{rdeg}_k M$ by condition 2 (a) of Definition 13. Hence, its coefficient contributes to the leading row coefficient matrix of M. Furthermore, by 2 (b) all elements left of the pivot have a smaller degree and do hence not contribute to $\mathrm{LC}_{\text{row}}(M)$—meaning that the pivot contributes the left-most entry of $\mathrm{LC}_{\text{row}}(M)_{k,\bullet}$. Finally, condition 2 (c) assures that there cannot be two pivots in the same column: Assume that $j_i = j_k$ then condition 2 (c) applied to the i^{th} row implies that $\deg M_{k,j_k} = \operatorname{rdeg}_k M < \operatorname{rdeg}_i M$ and the same condition applied to the k^{th} row yields $\operatorname{rdeg}_i M < \operatorname{rdeg}_k M$—a contradiction. □

On the other hand, having a leading row coefficient matrix in row echelon form is not a sufficient condition for being in Popov form. Some authors call a matrix M with a leading row coefficient matrix $\mathrm{LC}_{\text{row}}(M)$ in row echelon form to be in *weak Popov form*—see for example [Che03, Definition 2.7].

Each matrix $M \in {}^mR^n$ can be transformed into a matrix whose non-zero rows are in Hermite normal form or in Popov form. Below in Corollary 29 we will also show that the Popov and Hermite form of a matrix are unique. We will sketch algorithms for doing the transformations here. We do not claim that the algorithms are efficient but want merely to show that they exist.

To convert M to Popov normal form we first apply row-reduction as in [BCL06, Theorem 2.2]. Then we have to bring the leading row coefficient matrix into reduced row-echelon form which can be done using the same method. This will of course preserve row-reducedness. Finally, operations of the same kind are used to ensure the remaining conditions of Definition 13. See for example [MS03] or [Che03, Section 2.5.1].

Conversion into Hermite form on the other hand uses the Euclidean algorithm—see for example [BP96, Section 3] for a version of the Euclidean in the Ore polynomial case. Given M we apply Euclid transforming the first non-zero column of M to the form $(d, 0, \ldots, 0)$ where d is the greatest common right divisor of the entries of this column. That means, we have a unimodular matrix $Q \in ({}^mR^m)^*$ such that

$$QM = \begin{pmatrix} 0 & \cdots & 0 & d & * & \cdots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & \tilde{M} & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}.$$

Applying the same procedure recursively to $\tilde{M}$ we obtain a matrix in row echelon form. We just need to apply division with remainder in order to decrease the degrees above the pivot entries in order to fulfil condition 3 of Definition 12.

Together with the uniqueness in Corollary 29 this construction yields the following corollary:

Corollary 15. *Whenever $UA = H$ is in Hermite form for $A \in {}^mR^n$ and $U \in ({}^mR^m)^*$ then the first pivot of H is the greatest common right divisor of the elements in the $j_1{}^{th}$ column $A_{\bullet,j_1}$ of A.*

3 Converting between Hermite and Popov

3.1 Gröbner bases

Gröbner bases have been a useful tool in (ordinary) polynomial algebra ever since their invention by Bruno Buchberger in 1965 in his Ph. D. thesis [Buc65]. Since then, they have been extended to various other kinds of rings including Ore polynomials [CS98]. In parallel, there have been extensions for the computation of module bases over polynomial rings [MM86].

For our Gröbner bases, we use the definitions of [BGTV03]. They consider Gröbner bases for so-called *Poincaré-Birkhoff-Witt rings*, a class of rings more general than Ore polynomials. We refer the reader to [BGTV03, Definition 2.2.5] for the definition of Poincaré-Birkhoff-Witt rings; and [BGTV03, Corollary 3.3] shows that univariate Ore polynomials are indeed Poincaré-Birkhoff-Witt rings with respect to to the term order defined below.

3.1.1 Gröbner bases in [BGTV03]

In this section, we summarise the most important definitions from [BGTV03]. We do this for two reasons: First, since we are considering only a special case of the theory developed there, everything can be simplified a little bit; and second, we can bring everything to the notation used in this report.

Let K be a skew field, and let $\sigma\colon K \to K$ be an automorphism and $\vartheta\colon K \to K$ be a σ-derivation. The terms in $R = K[\partial; \sigma, \vartheta]$ in the sense of Gröbner bases are elements of the form $c\partial^\alpha \in R$ where $c \in K$ and $\alpha \in \mathbb{N}$. We consider row vectors in R^n. First, we fix the canonical basis $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$. That means that $\mathfrak{e}_i$ is the vector with the i^{th} entry equal to 1 and all other entries equal to 0. Then the *terms* in R^n take the form $c\partial^\alpha \mathfrak{e}_j$ where $c \in K$, $\alpha \in \mathbb{N}$ and $1 \leqslant j \leqslant n$. We will now introduce admissible orderings on these terms.

Definition 16 (Admissible ordering)**.** An *admissible term ordering* is a total ordering $<$ of the terms satisfying the additional conditions

1. $\partial^\alpha \mathfrak{e}_i < \partial^{\alpha+\beta} \mathfrak{e}_i$ for all α and $\beta \in \mathbb{N}$; and
2. $\partial^\alpha \mathfrak{e}_i < \partial^\beta \mathfrak{e}_k$ implies $\partial^{\alpha+\gamma} \mathfrak{e}_i < \partial^{\beta+\gamma} \mathfrak{e}_k$ for all α, β and $\gamma \in \mathbb{N}$.

The only admissible ordering in R is

$$\partial^\alpha < \partial^\beta \iff \alpha < \beta$$

making R a Poincaré-Birkhoff-Witt ring. We can extent this ordering to an admissible ordering on R^n. There are two standard ways for doing this: the term over position and the position over term ordering. The definitions can be found in [BGTV03, Definition 5.3.8] and [BGTV03, Definition 5.3.9], but we will repeat them here for the convenience of the reader. (In both cases there is actually more freedom involved as the positions $1, \ldots, n$ may be ordered arbitrarily—we will fix a particular order though that fits our needs).

Definition 17 (Position over term ordering). Let α and $\beta \in \mathbb{N}$. The *position over term ordering* $<_{\mathrm{POT}}$ is defined as

$$\partial^\alpha \mathfrak{e}_j <_{\mathrm{POT}} \partial^\beta \mathfrak{e}_k \quad :\iff \quad j > k \vee (j = k \wedge \alpha < \beta).$$

Note, that we consider the left-most position to be the largest. That means, that we obtain the following sequence

$$(0, \ldots, 0, 1) <_{\mathrm{POT}} (0, \ldots, 0, \partial) <_{\mathrm{POT}} (0, \ldots, 0, \partial^2) <_{\mathrm{POT}} \ldots$$
$$<_{\mathrm{POT}} (1, 0, \ldots, 0) <_{\mathrm{POT}} (\partial, 0, \ldots, 0) <_{\mathrm{POT}} (\partial^2, 0, \ldots, 0) <_{\mathrm{POT}} \ldots.$$

As can be seen from the example, the position over term ordering reminds of the lexicographic ordering in the usual polynomial case.

Definition 18 (Term over position ordering). Let α and $\beta \in \mathbb{N}$. The *term over position ordering* $<_{\mathrm{TOP}}$ is defined as

$$\partial^\alpha \mathfrak{e}_j <_{\mathrm{TOP}} \partial^\beta \mathfrak{e}_k \quad :\iff \quad \alpha < \beta \vee (\alpha = \beta \wedge j > k).$$

Note, that also here we order the positions with the left-most as the largest one. In this case we obtain

$$(0, \ldots, 0, 1) <_{\mathrm{TOP}} (0, \ldots, 0, 1, 0) <_{\mathrm{TOP}} \ldots <_{\mathrm{TOP}} (1, 0, \ldots, 0)$$
$$<_{\mathrm{TOP}} (0, \ldots, 0, \partial) <_{\mathrm{TOP}} \ldots <_{\mathrm{TOP}} (\partial, 0, \ldots, 0) <_{\mathrm{TOP}} \ldots.$$

The term over position ordering is reminiscent to graded orderings in the commutative Gröbner basis theory.

Also at this point we would like to do an implementation in MAPLE™. The first procedure computes the leading term of the non-zero vector v with respect to position over term ordering where the first argument Q tells us how the variable ∂ is denoted in MAPLE™. The procedure does not return a vector but a pair i, d such that $\mathrm{lt}(v) = \partial^d \mathfrak{e}_i$. If the elements of v are not expanded then the output might not be correctly computed.

```
POTlterm := proc(Q::symbol, v::Vector) :: list(nonnegint):
description "POT leading term of v as pair of position and degree.":
local p,i:
    p := LinearAlgebra:-Dimension(v):
    for i to p do
        if v[i] ≠ 0 then return i, degree(v[i],Q) fi:
    od:
    error "Leading term of zero is undefined!":
end proc:
```

Analogously, the next procedure computes the leading term of the non-zero vector v with respect to term over position ordering. Parameters and output are as above. In fact, we use the trick that $\mathrm{lt}(v)$ with respect to term over position ordering equals $\partial^{\deg v} \mathrm{lt}(\mathrm{lv}(v))$ where the last leading term is taken with respect to position over term ordering.

```
TOPlterm := proc(Q::symbol, v::Vector) :: list(nonnegint):
description "TOP leading term of v as pair of position and degree.":
local d,p:
    d := VectorDegree(Q,v):
    assert(d ⩾ 0):
    POTlterm(Q, map(p → coeff(p,Q,d)·Q^d, v)):
end proc:
```

Given an element $v \in R^n \setminus \{0\}$ and an admissible ordering $<$ we may write v as a finite sum of terms

$$v = c_1 \partial^{\alpha_1} \mathfrak{e}_{j_1} + \ldots + c_s \partial^{\alpha_s} \mathfrak{e}_{j_s}$$

where $\partial^{\alpha_1} \mathfrak{e}_{j_1} > \ldots > \partial^{\alpha_s} \mathfrak{e}_{j_s}$ and where $c_1 \neq 0$. Extending our previous definitions, we define the *leading coefficient* of v with respect to $<$ to be $\mathrm{lc}_<(v) = c_1$ and the *leading monomial* of v with respect to $<$ to be $\mathrm{lm}_<(v) = \partial^{\alpha_1} \mathfrak{e}_{j_1}$. Furthermore, we let the *leading term* of v with respect to $<$ be $\mathrm{lt}_<(v) = \mathrm{lc}_<(v) \cdot \mathrm{lm}_<(v)$. Since most of the time we will work with a fixed ordering $<$, we will usually just write $\mathrm{lc}(v)$, $\mathrm{lm}(v)$ and $\mathrm{lt}(v)$ instead of $\mathrm{lc}_<(v)$, $\mathrm{lm}_<(v)$ and $\mathrm{lt}_<(v)$.

The next step in Gröbner basis theory is reduction. We will be content with giving the definitions and citing the necessary theorems from [BGTV03]. We fix an admissible ordering $<$. We start with the definition of reducibility.

Definition 19 (Reducibility)**.** An element $v = c_1 \partial^{\alpha_1} \mathfrak{e}_{j_1} + \ldots + c_s \partial^{\alpha_s} \mathfrak{e}_{j_s} \in R^n$ is said to be *reducible* by a subset $F \subseteq R^n$ if there exists $w \in F$, $1 \leqslant k \leqslant s$ and $\beta \in \mathbb{N}$ such that $\partial^{\alpha_k} \mathfrak{e}_{j_k} = \partial^{\beta} \mathrm{lm}(w)$.

Given an element v and a finite subset W of R^n, it is always possible to *reduce* v by W in such a way that the remainder is irreducible.

Theorem 20. *Given $v \in R^n$ and $\{w_1, \ldots, w_s\} \subseteq R^n$ there are $u_1, \ldots, u_s \in R^n$ and $r \in R^n$ such that*

$$v = u_1 w_1 + \ldots + u_s w_s + r$$

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 [BGTV03, Theorem 5.4.3]. Directly after the theorem—in [BGTV03, Algorithm 10]—the division method is explained in detail. □

Now we have all the necessary prerequisites for the definition of Gröbner bases. Instead of copying [BGTV03, Definition 5.4.7] we use a formulation which is shown to be equivalent in [BGTV03, Theorem 5.4.9].

Definition 21 (Gröbner basis)**.** Let $\mathfrak{M}$ be an R-submodule of R^n. A finite set $G \subseteq R^n$ is a *Gröbner basis* for $\mathfrak{M}$ if for all $v \in R^n$ the conditions that $v \in \mathfrak{M}$ and that the remainder of v of the division by G is 0 are equivalent.

Lemma 22. *Every non-zero submodule $\mathfrak{M}$ of R^n has a Gröbner basis G, the elements of G generate $\mathfrak{M}$ and the remainder of the division of an element $v \in R^n$ by G does not dependent on the order of the elements in G.*

Proof. These statements are found in [BGTV03, Proposition 5.4.8, Corollary 4.10 and Theorem 5.4.9]. □

The following definition is [BGTV03, Definition 4.17]:

Definition 23 (Reduced Gröbner bases)**.** A Gröbner basis G of $\mathfrak{M} \subseteq R^n$ is *reduced* if for all $g \in G$ we have $\operatorname{lc}(g) = 1$ and there is no $h \in G \setminus \{g\}$ such that $\operatorname{lm}(h)$ divides a term in g.

As in the usual, commutative Gröbner basis theory there exist a notion of *S-polynomials* and a *Buchberger criterion* for Gröbner bases in R^n. They are stated in [BGTV03, Definition 5.4.11 and Theorem 5.4.13]. But we will not need the full Buchberger criterion here and will be content with stating a merely sufficient condition for being a Gröbner basis.

Theorem 24. *Let $G = \{g_1, \ldots, g_s\} \subseteq R^n$ with leading monomials $\operatorname{lm}(g_k) = \partial^{\alpha_k} \mathfrak{e}_{j_k}$ 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 $Rg_1 + \ldots + Rg_s \subseteq R^n$ generated by its elements.*

Proof. This is [BGTV03, Corollary 5.4.14]. □

3.1.2 Hermite and Popov forms are Gröbner bases

As mentioned above, in [KRT07, Proposition 2 and 4] it is shown that the rows Hermite and Popov forms of matrices of ordinary (commutative) polynomials are Gröbner bases. Using the results of [BGTV03] it is easy to extend this to general Ore polynomials.

Theorem 25. *Let $M \in {}^mR^n$ with the rows sorted in ascending order with respect to term over position ordering. Then the rows of M form a reduced Gröbner basis for $R^m M$ with respect to term over position ordering if and only if M is in Popov form.*

Proof. When considering term over position ordering, the leading term of a vector $v \in R^m \setminus \{0\}$ is the left-most term of highest degree. This is exactly that term whose leading coefficient is the left-most entry of $\operatorname{LC}_{\text{row}}(v)$ or—in other words—the leading term of the pivot in Definition 13.

Let first M be in Popov form. Then the leading row coefficient matrix of M is in row echelon form by Lemma 14. Using the considerations above that means that the leading terms of the rows of M are in different columns. Hence, by Theorem 24 they form a Gröbner basis.

The leading monomials are all monic by the definition of Popov forms—see condition 2 (a) of Definition 13. Furthermore, condition 2 (c) of the definition implies that the leading terms do not divide the other rows. Hence, M is a reduced Gröbner basis according to Definition 23.

Suppose on the other hand, that the rows of M form a reduced Gröbner basis with respect to term over position ordering. For each row index i let j_i be the column index of the leading term of $M_{i,\bullet}$. We will show that these j_i's are precisely those of Definition 13. Choose a row index i. By our considerations above, $\deg M_{i,j_i} = \operatorname{rdeg}_i M$ and $\deg M_{i,k} < \operatorname{rdeg}_i M$ for $k < j_i$. This is condition 2 (b) of Definition 13. Since the rows of M are a reduced Gröbner basis, the rows are monic and hence also condition 2 (a) holds. Additionally, for each $k \neq i$ the degree of M_{k,j_i} must be less than $\operatorname{rdeg}_i M$ since otherwise $M_{k,\bullet}$ could be reduced by $M_{i,\bullet}$. This means that condition 2 (c) is fulfilled. Furthermore there can be no other leading term in column j_i because otherwise $M_{i,\bullet}$ could be reduced by $M_{k,\bullet}$. That means that—up to permutation of the rows—the leading row coefficient matrix of M is in row echelon form. Since the reducedness implies that there are no zero rows in M, this means that $\operatorname{LC}_{\text{row}}(M)$ must have full rank. Together with the fact that by assumption the rows of M are sorted with respect to to their degree, this yields condition 1. Condition 2 (d) follows again from the assumption that the rows of M are sorted. □

We remark, that row-reducedness alone is not sufficient for being a Gröbner basis in the sense of [BGTV03]:

Example 26. Consider $R = \mathbb{Q}[x]$ and the (obviously row-reduced) matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in {}^3R^3.$$

Then A is not a Gröbner basis with respect to term over position ordering since $(0, 0, 1) \in R^3A$ is not divisible by the leading terms $(1, 0, 0)$ and $(0, 1, 0)$ of A.

If we take the position over term ordering, we find a correspondence to Hermite normal forms:

Theorem 27. *Let $M \in {}^mR^n$ with the rows sorted in descending order with respect to position over term ordering putting all zero rows at the end. Then the rows of M form a reduced Gröbner basis for R^mM with respect to position over term ordering if and only if M is in Hermite form.*

Proof. Since we use position over term ordering, the leading term of a vector $v \in R^m$ is the leading term of the left-most non-zero entry of v.

Assume M is in Hermite form. Using the notation from Definition 12, the leading terms of the rows are the leading terms of M_{i,j_i} by condition 1. Since the indices $j_1, \ldots, j_r$ are different, the leading terms are all in different columns, and hence by Theorem 24 they form a Gröbner basis. Furthermore, condition 3 ensures that no leading term divides a term in another row. Since the leading terms are also monic by condition 2, the rows of M form a reduced Gröbner basis according to Definition 23.

Contrarily, assume now that the rows M form a reduced Gröbner basis with respect to position over term ordering. Choose a row index i and assume that the leading coefficient of $M_{i,\bullet}$ is in column j_i. Then the degrees of the other entries in the same column must be strictly lower since the Gröbner basis formed by the rows of M is reduced. In particular there cannot be another leading term in the same column since otherwise one would reduce the other. Hence the indices $j_1, \ldots, j_m$ are all different. That implies that the rows are sorted with respect to the position of the leading term only and hence $j_1 > \ldots > j_m$. Using these indices in Definition 12, we see that conditions 2 and 3 are fulfilled. Since the leading terms are the left-most non-zero entries in each row, also condition 1 holds. □

We also state the important [BGTV03, Theorem 5.4.18]:

Theorem 28. *Fix an admissible ordering. Then, every submodule of R^n has a unique reduced Gröbner basis with respect to this ordering.*

From this we derive the corollary:

Corollary 29. *The Hermite and the Popov normal form of a matrix are unique.*

3.2 Gröbner bases à la [Pau07]

Above we considered the extension of Gröbner bases theory to modules using the methods from [BGTV03]. Another possibility of defining Gröbner bases would be to take an analogous approach to [Pau07]. This shall be examined in this section. The motivation for this is that—as shall be seen below—the division algorithm proposed in [Pau07] is quite reminiscent of the row-reduction algorithm.

We follow the definitions in [Pau07]—but instead of taking a ring as coefficient domain we use a vector space, namely K^n. That means that *terms* in this section will be of the form $y\partial^k$ where $y \in K^n$. Since we are dealing with univariate Ore polynomials, there is only one possible term order. Hence, in comparison with [Pau07, Definition 1] our degree function remains the same, but what the "leading coefficient" is in [Pau07] will be our leading vector.

We define an alternative divisibility criterion using leading vectors instead of leading monomials. Let $v \in R^n$ be given and let $W = \{w_1, \ldots, w_s\} \subseteq R^n$ be a finite set of vectors. We say that v is *divisible* by W if

$$\operatorname{lv}(v) \in \left\langle \sigma^{\deg v - \deg w}(\operatorname{lv}(w)) \;\middle|\; \deg w \leqslant \deg v \right\rangle$$

where the angle braces denote the K-vector space spanned by the leading vectors and where σ is applied to the vectors component-wise. If v and $w_1, \ldots, w_n$ were the rows of a matrix then divisibility would imply that this matrix is not row-reduced. If the above inclusion holds then there are $\alpha_w \in K$ for each $w \in W$ such that

$$\operatorname{lv}(v) = \sum_{\substack{w \in W \\ \deg w \leqslant \deg v}} \alpha_w \sigma^{\deg v - \deg w}(\operatorname{lv}(w))$$

where $\alpha_w = 0$ whenever $\deg w > \deg v$. Since $\operatorname{lv}(\partial^\beta(u)) = \sigma^\beta(\operatorname{lv}(u))$ for every $\beta \in \mathbb{N}$ and $u \in R^n$ and since $\deg(pu) = \deg p + \deg u$ for every $p \in R$, we conclude that

$$\tilde{v} = v - \sum_{\substack{w \in W \\ \deg w \leqslant \deg v}} \alpha_w \partial^{\deg v - \deg w} w$$

has a smaller degree than v. Note that the right hand side is exactly the reduction we used in the row-reduction algorithm.

Using the definition of divisibility given above we may formulate the following theorem. This is a reformulation of [Pau07, Proposition 2].

Theorem 30 (Division). *Let $W \subseteq R^n$ be a finite subset and let $v \in R^n$. Then there are an element $r \in R^n$ and a family $(p_w)_{w \in W}$ in R such that*

1. *we have the identity*
$$v = \sum_{w \in W} p_w w + r;$$
2. *for all $w \in W$ we have $p_w = 0$ or $\deg p_w + \deg w \leqslant \deg v$; and*
3. *$r = 0$ or no term in r is divisible by W.*

We call r the remainder *of the division of v by W.*

Proof. We do induction on the degree of v. Assume first that $v = 0$. Then with $r = 0$ and the zero family $(0)_{w \in W}$ clearly the conditions of the theorem are fulfilled.

Otherwise let $\deg v = \ell$. There are two possible cases: If v is divisible by W then there are $(\alpha_w)_{w \in W}$ such that $\tilde{v} = v - \sum_{w \in W} \alpha_w \partial^{\deg v - \deg w} w$ has a degree lower than that of v where we ignore the terms with $\deg v < \deg w$ since the corresponding α_w are zero anyway. By induction there are a remainder r and a family $(q_w)_{w \in W}$ such that the conditions of the theorem hold. Set $p_w = \alpha_w \partial^{\deg v - \deg w} + q_w$. Then

$$v = \sum_{w \in W} \alpha_w \partial^{\deg v - \deg w} w + \tilde{v} = \sum_{w \in W} \alpha_w \partial^{\deg v - \deg w} w + \sum_{w \in W} q_w w + r = \sum_{w \in W} p_w w + r.$$

Since $\deg w \leqslant \deg v$ for all w with $\alpha_w \neq 0$, we have also $\deg p_w + \deg w \leqslant \deg v$. Hence, the conditions hold for v with r and $(p_w)_{w\in W}$.

On the other hand, if v is not divisible by W then consider $\tilde{v} = v - \mathrm{lv}(v)\partial^{\deg v}$. Since $\tilde{v}$ has a degree lower than that of v, there are again a remainder s and a family $(p_w)_{w\in W}$ such that the conditions of the theorem hold. We have

$$v = \tilde{v} + \mathrm{lv}(v)\partial^{\deg v} = \sum_{w\in W} p_w w + s + \mathrm{lv}(v)\partial^{\deg v} = \sum_{w\in W} p_w w + r$$

where we set $r = \mathrm{lv}(v)\partial^{\deg v} + s$. Since the term $\mathrm{lv}(v)\partial^{\deg v}$ is not divisible by W, by induction no term of r is divisible by r. Thus, also in this case the conditions hold for v with r and $(p_w)_{w\in W}$. □

In the following we will usually regard the set $W = \{w_1, \ldots, w_m\} \subseteq R^n$ as a matrix with rows $w_1, \ldots, w_m$. If we define the *k^{th} leading row coefficient matrix* $\mathrm{LC}^k_{\mathrm{row}}(W) \in {}^mK^n$ to have the rows

$$\mathrm{LC}^k_{\mathrm{row}}(W)_{j,\bullet} = \begin{cases} \sigma^{k-\mathrm{rdeg}_j W}(W_{j,\bullet}), & \text{if } \mathrm{rdeg}_j W \leqslant k \\ 0, & \text{otherwise} \end{cases},$$

then $v \in R^n$ is divisible by W if and only if $\mathrm{lv}(v) \in K^m\, \mathrm{LC}^{\deg v}_{\mathrm{row}}(W)$. (This definition brings out the similarity to row-reduction even more). The above theorem then reads as

Corollary 31. *Let $W \in {}^mR^n$ and let $v \in R^n$. Then there are an element $r \in R^n$ and a vector $p \in R^m$such that*

1. *we have the identity $v = pW + r$;*

2. *for all $1 \leqslant j \leqslant m$ we have $p_j = 0$ or $\deg p_j + \mathrm{rdeg}_j W \leqslant \deg v$; and*

3. *$r = 0$ or no term in r is divisible by W.*

We call r the remainder *of the division of v by W.*

Let $W \subseteq R^n \setminus \{0\}$. If there is no $w \in W$ such that w is divisible by $W \setminus \{w\}$ then we call W *weakly auto-reduced*.[1] If we regard W as matrix in ${}^mR^n$, then we claim that W is weakly auto-reduced if and only if W is row-reduced. To prove this statement, we first state that a matrix W with non-zero rows is row-reduced if and only if $\mathrm{LC}^k_{\mathrm{row}}(W)$ has linear independent non-zero rows for every $k \in \mathbb{N}$. The reason is that—because σ is an automorphism—we could construct a linear dependency in $\mathrm{LC}_{\mathrm{row}}(W) = \mathrm{LC}^{\deg W}_{\mathrm{row}}(W)$ if we had dependencies in $\mathrm{LC}^k_{\mathrm{row}}(W)$ for some k. To prove the equivalence of weak auto-reducedness and row-reducedness, assume that for a row index i the row $W_{i,\bullet}$ was divisible by rows $W_{j_1,\bullet}, \ldots, W_{j_s,\bullet}$ with $j_k \neq i$ for all k. Because of the definition of divisibility we can omit those indices j_k with $\mathrm{rdeg}_{j_k} W > \mathrm{rdeg}_i W$. But that means that $\mathrm{lv}(W_{i,\bullet})$ depends on the rows of $\mathrm{LC}^{\mathrm{rdeg}_i W}_{\mathrm{row}}(W)$ with row index unequal to i; and hence that the non-zero rows of $\mathrm{LC}^{\mathrm{rdeg}_i W}_{\mathrm{row}}(W)$ are linearly dependent. Contrarily, assume there is a dependency of the rows of $\mathrm{LC}^k_{\mathrm{row}}(W)$ for some $k \in \mathbb{N}$. Since σ is an automorphism, if we choose k minimal with this property then there must be a row index i such that $k = \mathrm{rdeg}_i W$ and such that $\mathrm{LC}^k_{\mathrm{row}}(W)_{i,\bullet}$ is linear dependent on the other rows of $\mathrm{LC}^k_{\mathrm{row}}(W)$. But this means that $W_{i,\bullet}$ is reducible by the other rows of W. Hence, W is not auto-reduced.

[1] To be auto-reduced—as opposed to weakly auto-reduced—every $w \in W$ would have to be a remainder of division of w by $W \setminus \{w\}$, i. e., no term in w were allowed to be divisible by $W \setminus \{w\}$.

Definition 32 (Leading vector space). Let $\mathfrak{M} \in R^n$ be a submodule and let $G \subseteq \mathfrak{M} \setminus \{0\}$ be a finite subset. For $s \in \mathbb{N}$ we define the *leading vector space*

$$\mathsf{LV}(s, \mathfrak{M}) = \{\mathsf{lv}(v) \mid v \in \mathfrak{M} \text{ and } \deg v = s\} \cup \{0\}.$$

Then G is a *vectorial Gröbner basis* of $\mathfrak{M}$ if and only if for all $i \in \mathbb{N}$ the vector space $\mathsf{LV}(i, \mathfrak{M})$ is generated by

$$\{\sigma^{i-\deg g}(\mathsf{lv}(g)) \mid g \in G \text{ and } \deg g \leqslant i\}.$$

In other words, since the vectors $\sigma^{i-\deg g}(\mathsf{lv}(g))$ are just the non-zero rows of $\mathsf{LC}^i_{\mathrm{row}}(G)$, a finite subset $G \subseteq \mathfrak{M}$ is a vectorial Gröbner basis for $\mathfrak{M}$ if and only if for all $i \in \mathbb{N}$

$$\mathsf{LV}(i, \mathfrak{M}) = K^{|G|}\,\mathsf{LC}^i_{\mathrm{row}}(G).$$

Theorem 33. *Let $\mathfrak{M} \subseteq R^n$ be a submodule and G be a vectorial Gröbner basis of $\mathfrak{M}$. Then $v \in \mathfrak{M}$ if and only if the remainder of v after division by G is zero.*

Proof. This follows from Theorem 30: If $v \in \mathfrak{M}$ then in every division step the remainder is in $\mathfrak{M}$ meaning that it is always divisible. Contrarily, if $v \notin \mathfrak{M}$ then there must be a non-zero remainder r. □

We will now draw a connection between vectorial Gröbner bases and row-reduced matrices. First, we note that a vectorial Gröbner basis needs not to be row-reduced since it could contain the zero vector or otherwise linear dependent elements. We call a vectorial Gröbner basis G for a submodule $\mathfrak{M} \subseteq R^n$ *minimal* if for all $i \in \mathbb{N}$ the non-rows of $\mathsf{LC}^i_{\mathrm{row}}(G)$ form a minimal set of generators of $\mathsf{LV}(i, \mathfrak{M})$.[2]

Theorem 34. *A matrix $M \in {}^mR^n$ is row-reduced if and only if the rows of M are a minimal vectorial Gröbner basis of $R^m M$.*

Proof. Assume first that M is row-reduced. For a vector $v = yM \in R^m M$ we must show that $\mathsf{lv}(v)$ depends linearly on the rows of $\mathsf{LC}^{\deg v}_{\mathrm{row}}(M)$. Using the predictable degree property (Lemma 10) we obtain

$$v = \sum_{\substack{1 \leqslant i \leqslant m \\ \mathrm{rdeg}_i M \leqslant \deg v}} y_i M_{i,\bullet}.$$

This yields for the leading vectors

$$\mathsf{lv}(v) = \sum_{\substack{1 \leqslant i \leqslant m \\ \deg y_i + \mathrm{rdeg}_i M = \deg v}} \mathsf{lc}(p_i)\sigma^{\deg p_i}\,\mathsf{lv}(M_{i,\bullet}) \; = \sum_{\substack{1 \leqslant i \leqslant m \\ \deg y_i + \mathrm{rdeg}_i M = \deg v}} \mathsf{lc}(p_i)\sigma^{\deg v - \mathrm{rdeg}_i M}\,\mathsf{lv}(M_{i,\bullet}).$$

Hence, $\mathsf{lv}(v) \in K^m\,\mathsf{LC}^{\deg v}_{\mathrm{row}}(M)$, and M is a vectorial Gröbner basis.

To prove minimality it is sufficient to note that for each k the non-zero rows of $\mathsf{LV}(k, R^m M)$ are linearly independent.

Assume now that M is a minimal vectorial Gröbner basis. Obviously M cannot have zero rows. We will show that M is weakly auto-reduced. If this was not the case then there was a row index i such that $M_{i,\bullet}$ was divisible by the remaining rows meaning that there was a linear dependency between the non-zero rows of $\mathsf{LC}^{\mathrm{rdeg}_i M}_{\mathrm{row}}(M)$ meaning that these were not a minimal set of generators for $\mathsf{LV}(\mathrm{rdeg}_i M, R^m M,)$. Hence, a minimal vectorial Gröbner basis must be weakly auto-reduced and thus row-reduced. □

[2] A minimal set of generators is a basis.

The next lemma gives just another condition for a matrix to be a vectorial Gröbner basis. It can be seen as a version of [For75, Main Theorem] for Ore polynomials. Further properties of vectorial Gröbner basis may—via Theorem 34— be found in the appendix of [BCL06].

Lemma 35. *A matrix $M \in {}^{m}R^{n}$ is a minimal vectorial Gröbner basis for $R^m M$ if and only if its rows are linearly independent over R and if using the notation $\mathfrak{M}_{<d} = \{v \in R^m M \mid \deg v < d\}$ we have*

$$\dim_K \mathfrak{M}_{<d} = \sum_{\substack{j=1 \\ \operatorname{rdeg}_j M \leqslant d}}^{m} d - \operatorname{rdeg}_j M.$$

Proof. By Theorem 34 above we know that a minimal vectorial Gröbner basis M is row-reduced. Thus, M has the predictable degree property (Lemma 10) and the rows of M are linearly independent. Furthermore, the predictable degree property implies that $v = yM \in R^m M$ is in $\mathfrak{M}_{<d}$ if and only if $\max\{\deg y_k + \operatorname{rdeg}_k M \mid 1 \leqslant k \leqslant m\}$ is strictly less than d, or equivalently if $\deg y_k < d - \operatorname{rdeg}_k M$ for all k. The K-space of all Ore polynomials y_k with a degree strictly less than $d - \operatorname{rdeg}_k M$ has K-dimension $d - \operatorname{rdeg}_k M$ if $d \geqslant \operatorname{rdeg}_k M$ and 0 otherwise.

Assume now that the rows of M are linearly independent—implying in particular that M does not have zero rows—and that the identity $\dim_K \mathfrak{M}_{<d} = \sum_{j=1}^{m} \max\{0, d - \operatorname{rdeg}_j M\}$ holds for all d. We will show that the rows of $\operatorname{LC}_{\text{row}}^{d-1}(M)$ are a basis for $\operatorname{LV}(d-1, R^m M)$ for all $d \geqslant 1$. We have $\operatorname{LV}(d-1, R^m M) = \{\operatorname{lv}(v) \mid v \in \mathfrak{M}_{<d} \setminus \mathfrak{M}_{<d-1}\}$. Since taking the $(d-1)^{\text{th}}$ coefficient of an element in $\mathfrak{M}_{<d}$ is a K-linear map to $\operatorname{LV}(d-1, R^m M)$ with kernel $\mathfrak{M}_{<d-1}$, we see that

$$\operatorname{LV}(d-1, R^m M) \cong \frac{\mathfrak{M}_{<d}}{\mathfrak{M}_{<d-1}}$$

as K-spaces. Hence, by assumption we obtain

$$\begin{aligned}
\dim_K \operatorname{LV}(d-1, R^m M) &= \dim_K \mathfrak{M}_{<d} \setminus \mathfrak{M}_{<d-1} \\
&= \sum_{\substack{j=1 \\ \operatorname{rdeg}_j M \leqslant d}}^{m} d - \operatorname{rdeg}_j M - \sum_{\substack{j=1 \\ \operatorname{rdeg}_j M \leqslant d-1}}^{m} d - 1 - \operatorname{rdeg}_j M \\
&= \sum_{\substack{j=1 \\ \operatorname{rdeg}_j M = d}} 0 + \sum_{\substack{j=1 \\ \operatorname{rdeg}_j M \leqslant d-1}} 1 = |\{i \mid \operatorname{rdeg}_i M \leqslant d-1\}|,
\end{aligned}$$

i. e., $\dim_K \operatorname{LV}(d-1, R^m M)$ equals the number of non-zero rows of $\operatorname{LC}_{\text{row}}^{d-1}(M)$. We will now prove by induction on d that the non-zero rows of $\operatorname{LC}_{\text{row}}^{d-1}(M)$ are linearly independent. For $d = 1$ this is obvious as the non-zero rows of $\operatorname{LC}_{\text{row}}^{0}(M)$ are just the rows of M of degree 0 which are linearly independent. Let now that $d \geqslant 2$, and assume the rows of $\operatorname{LC}_{\text{row}}^{d-1}(M)$ where linearly dependent. Since the rows of $\operatorname{LC}_{\text{row}}^{d-2}(M)$ are linearly independent and since σ is an automorphism, that means that there is a row index i such that $\operatorname{rdeg}_i M = d - 1$ and such that $\operatorname{lv}(M_{i,\bullet})$ is a linear combination of the other rows of $\operatorname{LC}_{\text{row}}^{d-1}(M)$. Hence, we may apply reduction yielding a representation

$$M_{i,\bullet} = \sum_{\substack{j \neq i \\ \operatorname{rdeg}_j M \leqslant d-1}} M_{j,\bullet} + r$$

where $\deg r$ must be strictly smaller than $d-1$ and no term in r is divisible by M. Since $r \in R^n M$ and the spaces $\operatorname{LV}(s, R^m M)$ for $s < d-1$ are generated by the non-zero rows of $\operatorname{LC}_{\text{row}}^{s}(M)$ this implies $r = 0$. That means that $M_{i,\bullet}$ is a linear combination of the other rows of M contradicting the assumption that the rows of M are linearly independent. Thus, the rows of $\mathrm{LC}_{\mathrm{row}}^{d-1}(M)$ must be linearly independent and hence M must be a minimal vectorial Gröbner basis. □

At the end of this section we want to mention that reduction in the sense of vectorial Gröbner bases has an important disadvantage compared to division in the sense of usual Gröbner bases: The sum of two irreducible elements might be reducible meaning that vectorial Gröbner basis division is (in general) not K-linear. As an example simply consider the matrix

$$M = \begin{pmatrix} 1 & 1 \end{pmatrix} \in {}^{1}R^{2}$$

and the vectors $v = (0, 1)$ and $w = (1, 0) \in R^2$. Although neither v nor w are divisible by M the sum $v + w = (1, 1)$ clearly is.

3.3 FGLM

3.3.1 Introduction to FGLM

The FGLM algorithm—named after the initials of its inventors—was first presented in [FGLM93]. In this section we give a rough overview about the algorithm following the original paper. Since the FGLM algorithm there was aimed at commutative domains, Gröbner basis and division in this section are to be understood in the usual sense, i. e., as for example in [AL94]. Although the algorithm does not in particular need commutativity it seems not to have been ported to Ore polynomials and their Gröbner bases yet. The only source which is known to the author is a short mention in [Kou09].

Let F be a commutative field. We consider the commutative polynomial ring $P = F[x_1, \ldots, x_n]$. Assume that we are given an ideal $I \subseteq P$ with the property that $0 < \dim_F P/I < \infty$. Let $<$ be an admissible ordering and let G be the reduced Gröbner basis of I with respect to $<$. Then, an F-basis for P/I is given by those monomials that are not divisible by the leading terms in G as stated in [AL94, Proposition 2.1.6]. Following [FGLM93] we will call these monomials the *natural basis* of P/I with respect to G and denote it by $\mathfrak{B}_G$. Also, we define the *degree* of I to be $\deg I = \dim_F P/I$. By [FGLM93, Proposition 2.1] the leading monomials of the elements in G are of the form $x_j b$ for some $b \in \mathfrak{B}_G$ and some $1 \leqslant j \leqslant n$.

For each j, the multiplication by x_j is an F-linear map $P/I \to P/I$. Let T_j^G be the matrix of this map with respect to the natural basis corresponding to G. In [FGLM93, Procedure 3.1] a method for the computation of $T_1^G, \ldots, T_n^G$ is given. These matrices allow computation in P/I without the need of dividing by G after every multiplication. That is, they translate modular arithmetic to linear algebra.

The heart of [FGLM93] are [FGLM93, Proposition 4.1] and [FGLM93, Procedure 4.1] where the FGLM-algorithm is explained. Roughly, it proceeds as follows: The procedure expects as input an admissible ordering $<_1$, a reduced Gröbner basis G_1 of I with respect to $<_1$ and an admissible ordering $<_2$. It computes the reduced Gröbner basis G_2 with respect to $<_2$. For this, it considers two sets G and B that are to become the new Gröbner basis and its natural basis. Initially we set $G = \emptyset$ and $B = \{1\}$ where we know that $1 \in \mathfrak{B}_{G_2}$ since $\deg I > 0$. The algorithm now considers the with respect to $<_2$ smallest monomial $\mu = bx_j$ for $b \in B$ and $1 \leqslant j \leqslant n$ that is not in B and not a multiple of a leading term in G. Using G_1—or more precisely the representation of μ in $\mathfrak{B}_{G_1}$—we check whether there is an F-linear dependency modulo I between the monomials in B and μ. If this is the case then μ is the leading monomial of a polynomial in G_2 whose coefficients are determined by the coefficients of the dependency and which we add to G. Otherwise, $\mu \in \mathfrak{B}_{G_2}$ and we add μ

to B. If there are no monomials left that are not in B or not divisible by a leading term in G then the algorithm terminates returning $G_2 = G$ and $\mathfrak{B}_{G_2} = B$.

The linear dependencies which are mentioned in the algorithm are found in the following way. We claim that we always know the coordinates of the elements in B. This is true initially, since we start with $B = 1$ and $1 \in \mathfrak{B}_{G_1}$. Later we consider only monomials μ of the form $\mu = x_j b$ where $b \in B$. But here the coordinates may easily be computed by the multiplication $T_j^{G_1} \cdot b_{\mathfrak{B}_{G_1}}$ of the (known) coordinate vector of b by the corresponding multiplication matrix. Since division by G_1 is F-linear any linear dependency the coordinate vectors of $\mu \cup B$ corresponds to a linear dependency of $\mu \cup B$ modulo I.

3.3.2 The quotient module and pseudo-linear maps

If we are to extend this considerations to submodules of R^n then we first need to study the quotient module and the multiplication map therein. In general the quotient space will not be finite dimensional, but we will be able to give a condition for avoiding infinite computations below.

For a matrix $M \in {}^mR^n$ in Hermite or in Popov normal form we need to study the vector space structure of the quotient module $R^n/R^m M$. The multiplication by ∂ is a K-pseudo-linear map in the sense of [Jac37, Section 1] or [BP96, Definition 4]. There, given a K-space V, a map $\varphi\colon V \to V$ is called *K-pseudo-linear* if for all $v, w \in V$ and for all $a \in K$ the identities

$$\varphi(v + w) = \varphi(v) + \varphi(w) \qquad \text{and} \qquad \varphi(av) = \sigma(a)\varphi(v) + \vartheta(a)v$$

hold.

Assume that V has a finite K-dimension. If we fix a basis $\mathfrak{B}$ of V then there is a unique matrix $T \in {}^{|\mathfrak{B}|}K^{|\mathfrak{B}|}$ such that for $v \in V$

$$\varphi(v)_{\mathfrak{B}} = T\sigma(v_{\mathfrak{B}}) + \vartheta(v_{\mathfrak{B}})$$

where $w_{\mathfrak{B}}$ denotes the coordinate vector of $w \in V$ with respect to the basis $\mathfrak{B}$—see [Jac37, Section 2] or [BP96, Page 4]. Contrarily, each matrix $S \in {}^{|\mathfrak{B}|}K^{|\mathfrak{B}|}$ yields a K-pseudo-linear map via the above formula, i. e., there is a one-to-one correspondence between matrices and K-pseudo-linear maps just as with usual linear maps.

In order to study the K-pseudo-linear maps of $R^n/R^m M$ we need to study bases first. For this we use the fact that that Popov and Hermite forms are Gröbner bases. In the following we will denote the residue class of $v \in R^n$ in $R^n/R^m M$ by $\overline{v}$. The following lemma uses the fact that only those terms $\partial^i \mathfrak{e}_k R^n$ are divisible by the rows of M where k is the column index of the pivot of the j^{th} row and $\deg M_{j,k} \leqslant i$.

Lemma 36. *Let $M \in {}^mR^n$ be in Hermite or in Popov form and let $J = \{j_1, \ldots, j_m\}$ be the corresponding set of pivot indices.*[3] *Then a basis of $R^n/R^m M$ is given by the elements*

$$\{\overline{\partial^k \mathfrak{e}_{j_i}} \mid j_i \in J \text{ and } k < \deg M_{i,j_i}\} \cup \{\overline{\partial^k \mathfrak{e}_i} \mid i \notin J \text{ and } k \in \mathbb{N}\}.$$

Moreover is the Gröbner basis division by M a linear map from R^n to R^n.

Proof. This follows immediately from [BGTV03, Proposition 5.6.2 and Proposition 5.6.3] and the correspondence of the pivot indices and the leading monomials. □

[3] Please note that the definition of “pivot” is different for the two normal forms.

We will call $\{\overline{\partial^k \mathfrak{e}_i} \mid i \notin J \text{ and } k \in \mathbb{N}\}$ the *infinite part* of the basis while the set $\{\overline{\partial^k \mathfrak{e}_{j_i}} \mid j_i \in J \text{ and } k < \deg M_{i,j_i}\}$ will be called the *finite part*.

We return to the multiplication map $R^n/R^m M \to R^n/R^m M$ defined by $\overline{v} \mapsto \partial \overline{v} = \overline{\partial v}$. This is obviously K-pseudo-linear as $\partial(ap) = \sigma(a)\partial p + \vartheta(a)p$ for all $a \in K$ and all $p \in R$. The basis given in the lemma is in general not finite. Still we can easily compute an (infinite) matrix T for the multiplication map. In fact, if we assume M to be an irreducible Gröbner basis then the elements in the infinite part of the basis and their multiples are not reducible. Hence the part of the multiplication matrix corresponding to the infinite part will just be a matrix with 1's on the upper secondary diagonal. Also, the part of the multiplication matrix corresponding to the finite part of the basis has a structure. Since only the ∂-multiples of those elements of highest degree can be reduced, we obtain a form somewhat similar to a companion matrix.

Remark 37. In [CK02] multiplication matrices are called *∂-connections* or *defining ∂-matrices*. Change of basis is done via so-called *gauge transformations*—see [Jac37] or [CK02].

We will now give an indication of how finite subparts of the multiplication matrices can be computed. For this we consider only a truncated basis consisting of those modular basis elements with representatives of degree strictly less than a given bound d. That is, we define the *truncated (canonical) basis* $\mathfrak{B}_{\leqslant d}$ with respect to $d \in \mathbb{N}$ to be

$$\mathfrak{B}_{\leqslant d} = \{\overline{b} \in \mathfrak{B}_1 \mid \deg b \leqslant d\}$$

where $\mathfrak{B}$ is the canonical basis with respect to an ordering $<$ as in Lemma 36. Accordingly, we will also speak of the *truncated multiplication matrix* with respect to d as the portion of the multiplication matrix corresponding to this basis. In other words, the truncated multiplication matrix is the matrix of the projection to the truncated space after the multiplication by ∂. In the next subsection we will show that these truncated matrices are indeed sufficient for our needs.

Remark 38. Let $M \in {}^m R^n$ be in Hermite or in Popov normal form with pivot indices $j_1, \ldots, j_m$. Let a degree bound d be given. We define

$$\tau_k = \begin{cases} \min\{\deg M_{s,j_s}, d\}, & \text{if } k = j_s \\ \\ d & \text{otherwise.} \end{cases}$$

Then the truncated basis with respect to d will be just

$$\overline{\mathfrak{e}_1}, \ldots, \overline{\partial^{\tau_1} \mathfrak{e}_1}, \quad \ldots, \quad \overline{\mathfrak{e}_n}, \ldots, \overline{\partial^{\tau_n} \mathfrak{e}_n}.$$

Furthermore, the k^{th} coordinate of the product of ∂ with the i^{th} basis vector will be either

1. 0, if $0 < i - \tau_1 - \ldots - \tau_s < \tau_{s+1}$ for some s and $k \neq i + 1$, or
2. 1, if $0 < i - \tau_1 - \ldots - \tau_s < \tau_{s+1}$ for some s and $k = i + 1$, or
3. $-\operatorname{coeff}(\mu, M_{s,t})$ if $i = \tau_1 + \ldots + \tau_s$ for some s and where $\mu = k - \tau_1 - \ldots - \tau_t$ for some k such that $0 \leqslant \mu < \tau_{t+1}$.

This follows from the fact that M is a reduced Gröbner basis and hence for each row index i of M the terms in $M_{i,\bullet} - \operatorname{lt}(M_{i,\bullet})$ are irreducible.

Example 39. Let $R = \mathbb{Q}(x)[\partial; \mathrm{id}, d/dx]$ and

$$M = \begin{pmatrix} \partial^2 + x & x-1 & \partial - x \\ x+1 & \partial + 1 & \partial - x \end{pmatrix} \in {}^2R^3$$

being in Popov normal form. The pivot indices are $j_1 = 1$ and $j_2 = 2$. Let $d = 5$. Then $\tau_1 = 2$, $\tau_2 = 1$ and $\tau_3 = 5$ and the truncated basis will be

$$\mathfrak{e}_1, \partial\mathfrak{e}_1,\ \mathfrak{e}_2,\ \mathfrak{e}_3, \partial\mathfrak{e}_3, \partial^2\mathfrak{e}_3, \partial^3\mathfrak{e}_3, \partial^4\mathfrak{e}_3.$$

The truncated multiplication matrix with respect to d is then

$$T = \begin{array}{c|cccccccc} & \mathfrak{e}_1 & \partial\mathfrak{e}_1 & \mathfrak{e}_2 & \mathfrak{e}_3 & \partial\mathfrak{e}_3 & \partial^2\mathfrak{e}_3 & \partial^3\mathfrak{e}_3 & \partial^4\mathfrak{e}_3 \\ \hline \mathfrak{e}_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \partial\mathfrak{e}_1 & -x & 0 & 1-x & x & -1 & 0 & 0 & 0 \\ \mathfrak{e}_2 & -x-1 & 0 & -1 & x & -1 & 0 & 0 & 0 \\ \mathfrak{e}_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \partial\mathfrak{e}_3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \partial^2\mathfrak{e}_3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \partial^3\mathfrak{e}_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \partial^3\mathfrak{e}_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \in {}^8K^8.$$

Remark 40. In order to be able to compute in the truncated modular space we are still missing the representation of the canonical basis vectors $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ of R^n in the truncated basis. Taking the definitions from Remark 38, if $\tau_k \geqslant 1$ for a column index k then $\mathfrak{e}_k$ is not reducible by M and the representative will just be $\overline{\mathfrak{e}_k}$ corresponding to the $(1 + \sum_{s<k} \tau_k)^{\text{th}}$ element of the truncated basis.

If $\tau_k = 0$ then $\mathfrak{e}_k$ is reducible by M. If τ_k corresponds to the pivot in the i^{th} row of M then the remainder of the reduction of $\mathfrak{e}_k$ by M will just be $\mathfrak{e}_k - M_{i,\bullet}$. The terms in $\mathfrak{e}_k - M_{i,\bullet}$ are irreducible and thus we can compute the basis of this vector easily: The coefficient of the s^{th} truncated basis element in $\mathfrak{e}_k - M_{i,\bullet}$ will be just $-\operatorname{coeff}(\mu, M_{i,t})$ where $\mu = k - \tau_1 - \ldots - \tau_t$ for some k such that $0 \leqslant \mu < \tau_{t+1}$. This is similar to case 3 of Remark 38.

Again, we would like to implement the topics discussed in this section in Maple™. The resulting procedure expects as input a symbol Q that tells Maple™ how the variable ∂ is represented, a degree bound d, a procedure lt computing the leading term of a vector and finally the matrix $M \in {}^mR^n$ itself which must be a Gröbner basis for the ordering that is used in lt. The procedure returns the truncated multiplication matrix with respect to d, the coordinate vectors of the residue classes of the canonical basis vectors $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ of R^n and the dimension of truncated space.

```
ModularSpace := proc(Q::symbol, d::posint, lt::procedure, M::Matrix)
    :: list:
description "Compute the truncated modular structure of R^n/R^m M.":
```

We start by computing the positions and degrees of the pivot elements in M: The meaning of a tuple (i, j, k) in the list ϱ is that there is a pivot at position (i, j) in M with degree $\deg M_{i,j} = k$. The list σ will contain information about the columns of M. If $\sigma_j = \infty$ then there is no pivot in the j^{th} column of M. Else, if $\sigma_j = (k, i)$ then there is a pivot of degree k in the i^{th} row. The list τ contains information about the elements of the truncated basis: These are precisely the vectors $\overline{\partial^k \mathfrak{e}_j}$ where $k < \tau_j$. Finally, e contains the dimension of the truncated space.

```
local ϱ,σ,m,n,e,a,T,τ,j,z,r,c,i,E,k:
    m,n := LinearAlgebra:-Dimension(M):
```

```
ϱ:= [seq([j,lt(M[j,1..n])], j=1..m)]:
σ:= [∞ $n]:
for j in ϱ do
    σ[j[2]] := [j[3], j[1]]:
od:
τ:= map(a → if a = ∞ then d else min(a[1],d) fi, σ):
e := add(a, a in τ):
```

Initially, the truncated multiplication matrix T is just the zero matrix. We will fill in its entries later in the procedure. Also the list of coordinate vectors of the canonical basis elements is initially set to just empty entries. The variable r will contain the sum $\sum_{s<j}\tau_s$ with j being the control variable of the outmost loop. This will be used to handle cases 1 and 2 of Remark 38. Initially, r is of course zero.

```
T := Matrix(e,e,0):
E := [empty$n]:
r := 0:
for j to n do
```

For the j^{th} column of M we have to check whether it contains a pivot of degree zero. This is the case if $\tau_j = 0$. That means that $\mathfrak{e}_j$ is reducible by M and we have to compute its coordinates as described in Remark 40. We start with a zero vector filling in the corresponding entries in a loop. The variable c corresponds to the sum $\sum_{s<k}\tau_k$ in the remark and i is μ. Furthermore, $\sigma_{j,2}$ corresponds to i in Remark 40.

```
    if τ[j] = 0 then
        z := Vector[row](e,0):
        c := 0:
        for k to n do
            for i from 0 to τ[k]-1 do
                z[c+i+1] := -coeff(M[σ[j][2],k],Q,i):
            od:
            c := c+τ[k]:
        od:
        E[j] := z:
```

If the j^{th} column does not contain a pivot of degree zero then the representation of $\mathfrak{e}_i$ is just the $(r+1)^{\text{th}}$ unit vector according to Remark 40. Additionally, we have to treat the r^{th} through $(r+\tau_j-1)^{\text{th}}$ row of the truncated multiplication matrix corresponding to the truncated basis elements $\overline{\mathfrak{e}_j}$ through $\overline{\partial^{\tau_j-2}\mathfrak{e}_j}$. We first fill in the ones on the upper secondary diagonal according to case 2 of Remark 38. If the row does not contain a pivot then we are already done. Otherwise, we treat $(r+\tau_j)^{\text{th}}$ row analogously to the computation of the coordinates in the case of a zero-degree pivot.

```
    else
        E[j] := Vector[row](e, shape=unit[r+1]):
        #
        for r from r+1 to r+τ[j]-1 do
            T[r,r+1] := 1:
        od:
        if σ[j] ≠ ∞ then
            c := 0:
            for k to n do
                for i from 0 to τ[k]-1 do
                    T[r,c+i+1] := -coeff(M[σ[j][2],k],Q,i):
                od:
                    c := c+τ[k]:
                od:
            fi:
            #
        fi:
    od:
    return T,E,e:
end proc:
```

3.3.3 Degree bounds

In order to be able to apply the module FGLM algorithm that we will introduce below, we need to derive degree bounds. We start with a bound for the Popov normal form. The following result was proved for commutative polynomials as part of [For75, Main Theorem].

Lemma 41. *A matrix $M \in {}^{m}R^{n}$ is row-reduced if and only if $\sum_{j=1,M_{j,\bullet}\neq 0}^{m} \operatorname{rdeg}_j M$ is minimal among all matrices with the same row space.*

Proof. Assume that M is row-reduced. Then by Theorem 34 M is a minimal vectorial Gröbner basis meaning that the non-zero rows of $\mathrm{LC}_{\mathrm{row}}^{i}(M)$ are a basis for $\mathrm{LV}(i, R^{m}M)$ for every i. For every other matrix $W \in {}^{s}R^{n}$ with $R^{s}W = R^{m}M$ it follows that $\mathrm{LC}_{\mathrm{row}}^{i}(W)$ is a multiple of $\mathrm{LC}_{\mathrm{row}}^{i}(M)$.

Assume that $\sum_{j=1}^{m} \operatorname{rdeg}_j M > \sum_{k=1}^{s} \operatorname{rdeg}_j W$. As row-reduction only reduces the degrees of the rows by Theorem 9, we may without loss of generality assume that W is row-reduced, too. The inequality implies that either there was a degree d such that $\mathrm{LC}_{\mathrm{row}}^{d}(M)$ has less non-zero rows than $\mathrm{LC}_{\mathrm{row}}^{d}(W)$ or $s < m$ and $\mathrm{LC}_{\mathrm{row}}(M)$ has more non-zero rows than $\mathrm{LC}_{\mathrm{row}}(\deg M)W$. Both are not possible since the non-zero rows of both $\mathrm{LC}_{\mathrm{row}}^{d}(M)$ and $\mathrm{LC}_{\mathrm{row}}^{d}(W)$ in the first case or $\mathrm{LC}_{\mathrm{row}}(M)$ and $\mathrm{LC}_{\mathrm{row}}^{\deg M}(W)$ in the second case are bases for the same space.

Contrarily, by [BCL06, Theorem 2.2], row-reduction only reduces the row degrees of a matrix. Hence, if M was not row-reduced then could the sum of the row degrees not be minimal. By contraposition minimality of $\sum_j \operatorname{rdeg}_j M$ implies row-reducedness. □

Remark 42. In [For75], the sum $\sum_{j=1,M_{j,\bullet}\neq 0}^{m} \operatorname{rdeg}_j M$ of the row degrees is called the *order* of M. A matrix with minimal order is called a *minimal basis*. Thus, we have just proved that minimal bases are row-reduced.

We derive a corollary from the above lemma.

Corollary 43. *A row-reduced matrix $M \in {}^{m}R^{m}$ is unimodular if and only if we have $M \in ({}^{m}K^{m})^{*}$.*

Proof. If $M \in {}^{m}R^{m}$ is unimodular then $R^{m}M = R^{m}$ having the minimal basis $\mathbf{1}_m$. Hence must the sum of the row degrees of M be zero. Since M is row-reduced from the predictable degree property (Lemma 10) we further conclude that also its inverse must be in $({}^{m}K^{m})^{*}$.

On the other hand being in $({}^{m}K^{m})^{*}$ means that M is unimodular. □

Lemma 44. *Let $A \in {}^{m}R^{n}$ be any matrix and M its Popov form. Then $\deg M \leqslant \deg A$.*

Proof. Applying row-reduction to A can only lower the row degrees and hence the total degree by [BCL06, Theorem 2.2]. Further transformation into Popov form does not change the row degrees according to [BCL06, Lemma A.1 (d)]. □

Now, we want to derive a degree bound for the Hermite normal form as well. We follow an approach found in [GK09]. Since in the reference only commutative fields, trivial automorphisms and square matrices are considered, we will redo all the proofs here. We will start with [GK09, Theorem 3.3].

Theorem 45. *Let $A \in {}^{m}R^{n}$ with $\deg A \leqslant d$ and full row rank. Let $UA = H$ for a unimodular matrix $U \in ({}^{m}R^{m})^{*}$ and $H \in {}^{m}R^{n}$ in Hermite normal form having no zero rows. Then there exists a unimodular $V \in ({}^{m}R^{m})^{*}$ such that $A = VH$, $UV = \mathbf{1}_m$ and $\deg V \leqslant d$.*

Proof. Since U is unimodular, the inverse V trivially exists. Let $j_1 > \ldots > j_m$ be the pivot indices of H. We prove the claim about the degree of V by induction on the column index s of V. For $s = 1$ from Corollary 15 we obtain that $0 \neq H_{1,j_1} = \operatorname{gcrd}(A_{\bullet,j_1})$. Thus $\deg H_{1,j_1} \leqslant d$, and from $V_{\bullet,1} H_{1,j_1} = A_{\bullet,j_1}$ we conclude that $\deg V_{\bullet,1} \leqslant d$.

Assume now that $s \leqslant n$ and that for $1 \leqslant k < s \leqslant n$ we know that $\operatorname{rdeg}_k V \leqslant d$. We will now prove that also $\operatorname{rdeg}_s V \leqslant d$. For this we need to distinguish two cases: If $\operatorname{rdeg}_s V \leqslant \max\{\operatorname{rdeg}_k V \mid k < s\} \leqslant d$ then there is nothing to do. Otherwise, if $\operatorname{rdeg}_s V > \max\{\operatorname{rdeg}_k V \mid k < s\}$ then since by condition 3 of Definition 12 $\deg H_{s,j_s} > \max\{\deg H_{k,j_s} \mid 1 \leqslant k < s\}$ and thus for $1 \leqslant i \leqslant n$ we obtain $\deg A_{i,j_s} = \deg V_{i,s} H_{s,j_s}$ because $A_{i,j_s} = \sum_{t=1}^{s} V_{i,t} H_{t,j_s}$ and all the other terms are of lower degree by our assumptions. Since i was arbitrary, we conclude $\operatorname{rdeg}_s V \leqslant d$. By induction the claim follows. □

The next corollaries are [GK09, Corollary 3.3] and [GK09, Corollary 3.4].

Corollary 46. *Let A, U and V be as in Theorem 45. Then $\deg U \leqslant (m-1)\deg A$.*

Proof. By Corollary 43 we know that row-reduction applied to V yields a matrix in $({}^{m}K^{m})^{*}$. W. l. o. g., we may assume that the row-reduced form of V is $\mathbf{1}_m$. Moreover, since $\mathbf{1}_m = UV$ using the uniqueness of the inverse we can compute the degree bound on U by Theorem 9. Since—with the notation of Theorem 9—it is $\nu_j = \operatorname{rdeg}_j \mathbf{1} = 0$ for all j, we obtain

$$\operatorname{rdeg}_j U \leqslant \nu_j + \sum_{k=1}^{m} (\mu_j - \nu_j) - \min\{\mu_k \mid k = 1, \ldots, m\}$$

$$\leqslant \sum_{k=1}^{m} \mu_j - \min\{\mu_k \mid k = 1, \ldots, m\}.$$

Now, the bound on $d \geqslant \deg V \geqslant \mu_j$ for all j that was obtained in Theorem 45 implies $\operatorname{rdeg}_j U \leqslant (n-1)d$. □

Corollary 47. *Let A and H be as in Theorem 45. Then $\deg H \leqslant m \deg A$.*

Proof. We have $\deg H = \deg(UA) \leqslant \deg U + \deg A = m \deg A$ by the usual rules for matrix degrees and the previous corollary. □

We summarise this section:

Theorem 48. *Let M and $H \in {}^{m}R^{n}$ be similar and assume that M is in Popov form and that H is in Hermite form. Then*

$$\deg H \leqslant m \deg M \qquad \text{and} \qquad \deg M \leqslant \deg H.$$

3.3.4 Converting between Hermite and Popov forms

Assume we are given a matrix $M \in {}^mR^n$ that is in Popov normal form. If we compute the corresponding Hermite form $H = SM \in {}^mR^n$ where $S \in ({}^mR^m)^*$ is unimodular then $R^mM = R^mSM = R^mH$ and H must be the unique position over term Gröbner basis for R^mM. Hence, applying a change of basis to M must yield H. The same is true if we start with H and aim at computing the term over position Gröbner basis M.

Contrarily, if we apply change of basis to a Gröbner basis $M \in {}^mR^n$ in term over position ordering we will get a Gröbner basis $H \in {}^sR^n$. Since every row of M is an R-linear combination of rows of H and vice versa we obtain two matrices $A \in {}^sR^m$ and $B \in {}^mR^s$ such that $H = AM$ and $M = BH$. Hence, $H = ABH$ and $M = BAM$. Since H and M have linearly independent rows, we conclude that $AB = \mathbf{1}_s$ and $BA = \mathbf{1}_m$. But this can only be true if $m = s$. We conclude that H must be a Hermite form for M. Again, the same holds with Hermite and Popov interchanged.

In this section we will assume that we know a degree bound for the elements of the Gröbner basis we are trying to compute via the FGLM algorithm. More precisely, let $M \in {}^mR^n$ be a matrix that is a reduced Gröbner basis with respect to the admissible ordering $<_1$ and we assume that we know that the reduced Gröbner basis of R^mM with respect to the admissible ordering $<_2$ has rows of degree less than or equal to d. The goal is to extend the FGLM algorithm in such a way that the reduced Gröbner basis with respect to $<_2$ is computed. This approach is roughly comparable to the technique of *Hilbert driven Gröbner basis computation* in the commutative case where knowledge of the Hilbert function gives a criterion to decide whether certain S-polynomials reduce to zero—see for example [Tra97].

The problem is certainly to decide whether we are in the finite or in the infinite part of the basis. For this we use the degree bound. During the FGLM computation a considered monomial is either added to the canonical basis $\mathfrak{B}_2$ with respect to $<_2$ or it is identified as the leading monomial of an element of the new Gröbner basis. In the new Gröbner basis all monomials must be of degree less or equal to d. Hence, we do not need to consider monomials from the infinite part of the basis with degree greater than d.

For $v \in R^n$ we define the *support* to be the set of all monomials in v with non-zero coefficient, i. e.,

$$\operatorname{supp}(v) = \{\partial^a \mathfrak{e}_j \in R^n \mid \operatorname{coeff}(a, v)_j \neq 0\}.$$

Furthermore, we denote the vector space generated by all monomials of degree less or equal to a certain bound $d \in \mathbb{N}$ by

$$R^n_{\leqslant d} = \{v \in R^n \mid \deg v \leqslant d\}.$$

With this preparations we may state a lemma that is useful for working with truncated spaces:

Lemma 49. *Let $M \in {}^mR^n$ be a Gröbner basis with respect to the ordering $<$, and let $d \in \mathbb{N}$ such that $\deg M \leqslant d$. Then $v \in R^n_{\leqslant d}$ is zero modulo M if and only if the coordinate vector in the truncated basis with respect to d of the remainder $\tilde{v}$ of reduction of v by M is zero.*

Proof. Let $\mathfrak{B}$ be the canonical basis with respect to M and denote by $\mathfrak{B}_{\leqslant d}$ the truncated basis with respect to d. Since d is a degree bound for all rows of M, the remainder of an element from $R^n_{\leqslant d}$ by division with M will be in $R^n_{\leqslant d} \cap \mathfrak{B} = \mathfrak{B}_{\leqslant d}$. Thus will the projection to the truncated space yields zero if and only if the remainder is zero. Using Definition 21 we see that $v \in R^mM$ if and only if $\tilde{v} = 0$ if and only if the coordinate vector of $\tilde{v}$ with respect to $\mathfrak{B}_{\leqslant d}$ is zero. □

We are now prepared to give the algorithm. A statement about its correctness and finiteness may be found in Theorem 51.

Algorithm 50 (FGLM with degree bound).

Input: The reduced Gröbner basis $M \in {}^mR^n$ with respect to to the admissible ordering $<_1$, another admissible ordering $<_2$ and a bound d on the degrees of the reduced Gröbner basis with respect to $<_2$.

Output: The reduced Gröbner basis of $R^m M$ with respect to $<_2$.

Procedure:

```
1. Let $\mathfrak{B}_1$ be the truncated canonical basis with respect to $<_1$ and $d$, and let $T$ be the corresponding truncated multiplication matrix.
2. Initialise $C \leftarrow \varnothing$, $\mathfrak{B}_2 \leftarrow \varnothing$ and $G_2 \leftarrow \varnothing$.
3. If it exists, choose the smallest monomial $\mu = \partial^a \mathfrak{e}_i$ with respect to $<_2$ such that $\mu \notin \mathfrak{B}_2$ and $\mu$ is not divisible by the elements of $G_2$ and $a \leqslant d$.
   (a) Compute the coordinate vector $\mu_{\mathfrak{B}_1}$ of $\mu$ with respect to $\mathfrak{B}_1$.
       Here, $T$ can be used.
   (b) If $C \cup \{\mu_{\mathfrak{B}_1}\}$ is $K$-linearly independent, then set $C \leftarrow C \cup \mu_{\mathfrak{B}_1}$ and $\mathfrak{B}_2 \leftarrow \mathfrak{B}_2 \cup \{\mu\}$.
   (c) Else let $\mu_{\mathfrak{B}_1} = \sum_{\beta \in \mathfrak{B}_2} a_\beta \beta_{\mathfrak{B}_1}$ where all $a_\beta \in K$, and set $G_2 \leftarrow G_2 \cup \{\mu - \sum_{\beta \in \mathfrak{B}_2} a_\beta \beta\}$.
   (d) Go to step 3.
4. If no such $\mu$ exists then stop. The output is $G_2$.
```

Theorem 51. *Algorithm 50 is correct and terminates.*

Proof. In Algorithm 50 only monomials with degree lower than or equal to d are considered. This is a finite set and hence does the algorithm terminate after finitely many steps.

We turn ourselves now to the correctness. We will use the notations of Algorithm 50 in the proof. In each iteration of the loop in step 3 the elements of $\mathfrak{B}_2$ are linearly independent modulo $R^m M$: This is clearly true in step 2 and remains true since only those elements are added to $\mathfrak{B}_2$ which do not destroy this property by Lemma 49.

The elements of G_2 are in $R^m M$, since by the linearity of Gröbner basis division (Lemma 36) and the linearity of the coordinate map (and of the projection onto the truncated quotient space which is injective since the degrees of the elements of G_2 are small enough) we obtain

$$\overline{\mu - \sum_{\beta \in \mathfrak{B}_2} a_\beta \beta} = \mu_{\mathfrak{B}_1} - \sum_{\beta \in \mathfrak{B}_2} a_\beta \beta_{\mathfrak{B}_1} = 0$$

which is just the condition that $\mu - \sum_\beta a_\beta \beta$ is added to G_2 in step 3 using again Lemma 49. Since the monomials are chosen in ascending order, the leading term of $\mu - \sum_\beta a_\beta \beta$ with respect to the ordering $<_2$ will be just μ.

We consider the set of all multiples of leading monomials of G denoted by $\mathrm{LM}(G_2) = \{\mu \operatorname{lm}(g) \mid \mu \in R \text{ and } g \in G_2\}$. If we set $\mathrm{LM}_{\leqslant d}(G_2) = \mathrm{LM}(G_2) \cap R^n_{\leqslant d}$ then we obtain

$$\{\partial^a \mathfrak{e}_i \in R^n \mid a \leqslant d\} = \mathrm{LM}_{\leqslant d}(G_2) \,\dot{\cup}\, \mathfrak{B}_2.$$

This follows from the choice of monomials in step 3: Only those monomials can be added to $\mathfrak{B}_2$ that are not multiples of leading monomials in G_2; and leading monomials added to G_2 cannot divide anything in that is already in $\mathfrak{B}_2$ because the monomials are selected in ascending order and a monomial can never be divided by a bigger one—this is exactly the definition of admissible orderings in [BGTV03, Definition 5.3.7 (1)]. Since $G_2 \in R^m M$, this implies in particular that $\mathfrak{B}_{\leqslant d} \subseteq \mathfrak{B}_2$ for the truncated canonical basis $\mathfrak{B}_{\leqslant d}$ with respect to $<_2$.

Let $\tilde{G}$ denote the unique reduced Gröbner basis of $R^m M$ with respect to to $<_2$ which exists by Theorem 28. We claim that for every $g \in \tilde{G}$ we have $\operatorname{lm}(g) \in \operatorname{LM}_{\leqslant d}(G_2)$. Since we know a degree bound, we must have $\operatorname{lm}(g) \in R^m_{\leqslant d}$. Assume $\operatorname{lm}(g)$ was in $\mathfrak{B}_2$. Then, since $\tilde{G}$ is a Gröbner basis, we could reduce an element of $\mathfrak{B}_2$ contradicting the linear independence of $\mathfrak{B}_2$ modulo $R^m M$. Hence, every time a vector is reducible by $\tilde{G}$ it must also be reducible by G_2. In particular is the remainder of an element in R^n by division with $\tilde{G}$ zero if and only if it is zero by division with G_2. But using Definition 21 this means that G_2 must be a Gröbner basis since $\tilde{G}$ is one.

By construction are the leading terms of the elements of G_2 monic and do not divide each other. Since the other monomials of the elements in G_2 are in $\mathfrak{B}_2$ we further see that G_2 is autoreduced. Hence, G_2 must be the (unique) reduced Gröbner basis of $R^m M$ with respect to $<_2$ by Definition 23. □

Remark 52. It is possible to compute the transformation matrices using Gröbner basis division.

Algorithm 50 may be refined in case that the input M is in Popov normal form and $<_2$ is the position over term ordering, i. e., when we are aiming at computing the Hermite normal form. In step 1 of Algorithm 50 the truncated basis and multiplication matrix can be computed via Remark 38; the representations of the residue classes canonical basis vectors of R^n is easily computed with Remark 40. The sorted sequence of monomials with respect to position over term ordering is obtained as follows. Start with the column index $j = n$ and the degree $a = 0$; and set μ in step 3 to $\partial^0 \mathfrak{e}_j$. While there is no linear dependency detected, increase a until $a > d$. All the while, the next μ will just be $\partial^a \mathfrak{e}_j$ and its coordinate vector $\mu_{\mathfrak{B}_1}$ can be computed by multiplication with T. If we can find a linear dependency then all we have to do for the next iteration is decreasing j.

In Section 3.3.5 we do a complexity analysis for this case.

Combining the above Theorems 48 and 51 we obtain.

Corollary 53 (Main result). *A matrix $M \in {}^m R^n$ can be converted from Popov to Hermite form using Algorithm 50 with a degree bound of $m \deg M$.*

Example 54. As an example we would like to consider the system from the introduction of this report. Let $R = \mathbb{Q}(x)[\partial; \mathrm{id}, d/dx]$ be the ring of differential operators with rational coefficients. We are considering the matrix

$$M = \begin{pmatrix} -2 & \partial + \frac{3}{2}x & -3 - \frac{3}{2}x^2 \\ 0 & \frac{1}{2} & \partial - \frac{1}{2}x \end{pmatrix} \in {}^2R^3.$$

Since the leading terms with respect to term over position ordering are in different columns, M is a term over position Gröbner basis by Theorem 24 and hence in Popov form by Theorem 25. The degree of M is $\deg M = 1$ yielding a degree bound of 2 for the Hermite form of M by Theorem 48. The truncated basis with respect to 2 is thus

$$\mathfrak{B} = \overline{\mathfrak{e}_1}, \overline{\partial\mathfrak{e}_1}, \overline{\mathfrak{e}_2}, \overline{\mathfrak{e}_3}$$

and the corresponding truncated multiplication matrix is

$$T = \begin{array}{c} \\ \mathfrak{e}_1 \\ \partial\mathfrak{e}_1 \\ \mathfrak{e}_2 \\ \mathfrak{e}_3 \end{array} \begin{array}{c} \begin{array}{cccc} \mathfrak{e}_1 & \partial\mathfrak{e}_1 & \mathfrak{e}_2 & \mathfrak{e}_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & -\frac{3}{2}x & 3 + \frac{3}{2}x^2 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2}x \end{pmatrix} \end{array} \in {}^4\mathbb{Q}(x)^4.$$

The smallest monomial with respect to position over term ordering is $\mathfrak{e}_3$ which has the coordinates $(0,0,0,1)$ in the truncated basis since it is irreducible. The next vector is $\partial\mathfrak{e}_3$ which can be computed as

$$(\partial\overline{\mathfrak{e}_3})_{\mathfrak{B}} = (0,0,0,1)T + \frac{d}{dx}(0,0,0,1) = (0,0,-\tfrac{1}{2},\tfrac{1}{2}x)$$

One sees that the coordinate vectors of $\mathfrak{e}_3$ and $\partial\mathfrak{e}_3$ are linearly independent. The next monomial is $\partial^2\mathfrak{e}_3$ which we compute as

$$(\partial^2\overline{\mathfrak{e}_3})_{\mathfrak{B}} = (0,0,-\tfrac{1}{2},\tfrac{1}{2}x)T + \frac{d}{dx}(0,0,-\tfrac{1}{2},\tfrac{1}{2}x) = (-1,0,\tfrac{1}{2}x,-1-\tfrac{1}{2}x^2).$$

The three coordinate vectors are still linearly independent. But we have reached the bound 2 for the degrees of the monomials to consider. Hence, we continue with the next smallest monomial $\mathfrak{e}_2$ having $(\overline{\mathfrak{e}_2})_{\mathfrak{B}} = (0,0,1,0)$. This depends linearly on the previous coordinates:

$$(0,0,1,0) = (x,-2,0)\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2}x \\ -1 & 0 & \frac{1}{2}x & -1-\frac{1}{2}x^2 \end{pmatrix}.$$

Thus, the last row of the Hermite form is

$$\mathfrak{e}_2 - x\mathfrak{e}_3 + 2\partial\mathfrak{e}_3 = (0,1,2\partial - x).$$

We have to continue with the next smallest monomial not divisible by $\mathfrak{e}_2$ which is $\mathfrak{e}_1$ having the coordinates $(\overline{\mathfrak{e}_1})_{\mathfrak{B}} = (1,0,0,0)$. Again this is linear dependent on the previous coordinates:

$$(1,0,0,0) = (-1,-x,-1)\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2}x \\ -1 & 0 & \frac{1}{2}x & -1-\frac{1}{2}x^2 \end{pmatrix}.$$

This yields the second last row of the Hermite form

$$\mathfrak{e}_1 + \mathfrak{e}_3 + x\partial\mathfrak{e}_3 + \partial^2\mathfrak{e}_3 = (1,0,\partial^2 + x\partial + 1).$$

There are no monomials left that are not divisible by $\mathfrak{e}_1$ or $\mathfrak{e}_2$ and that are of degree less than 2. Hence, the algorithm stops returning the Hermite normal form of M

$$\begin{pmatrix} 1 & 0 & \partial^2 + x\partial + 1 \\ 0 & 1 & 2\partial - x \end{pmatrix}$$

In the example we may instantly compute solutions from the Hermite form. Namely,

$$M\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = 0 \iff f_1 = -\frac{d^2f_3}{dx^2} - x\frac{df_3}{dx} - f_3 \text{ and } f_2 = xf_3 - 2\frac{df_3}{dx}.$$

Furthermore, in this example it is easy to read of the transformation matrix from

$$M = QH \iff Q = \begin{pmatrix} -2 & \partial + \frac{3}{2}x \\ 0 & \frac{1}{2} \end{pmatrix}.$$

In general one would have to do Gröbner basis division in order to obtain this result.

More examples are in Appendix A.

Finally, we would like to present a Maple$^{\text{TM}}$ procedure implementing the conversion algorithm for the special case of converting a matrix $M \in {}^{m}R^{n}$ in Popov form into a matrix H in Hermite form. The input parameters are Q denoting how the variable ∂ is represented in Maple$^{\text{TM}}$, the automorphism σ as map of vectors, the derivation ϑ as map of vectors and the matrix M which must be in Popov form. The output will be the Hermite form of M.

```
Convert := proc(Q::symbol, σ::appliable, ϑ::appliable, M::Matrix)
    :: Matrix:
description "Convert M in Popov form into Hermite form.":
```

The procedure starts by setting some constants: As usual, with m and n we denote the dimensions of the matrix, $d-1$ is the bound for the degrees in H. We use the procedure `ModularSpace` defined on page 24 to compute the truncated multiplication matrix T and a list E containing the coordinate vectors of the residue classes of the canonical basis vectors $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ of R^n in the truncated basis. The variable B is just $\mathfrak{B}$ from Algorithm 50, H is G_2 and C will hold the coordinate vectors of the entries in B. The variable r contains the number of linear independent elements in B.

```
local T,m,n,d,H,C,B,j,w,k,v,S,E,F,e:
    m,n := LinearAlgebra:-Dimension(M):
    d := m·MatrixDegree(Q,M)+1:
    T,E,e := ModularSpace(Q,d,v → TOPlterm(Q,v),M):
    C := []:
    B := []:
    H := []:
```

Now, we iterate over the monomials in R^n. Since we are using a fixed term ordering—namely the position over term ordering—we have the procedure already specialised for this. The outer loop iterates over all column indices j and inner loop iterates over the exponents k from 0 through d. All the time, we have $w = \partial^k \mathfrak{e}_j$ and $v = w_{\mathfrak{B}_1}$. If v is linear independent of the previous coordinate vectors that have been stored in C then we add v to C and w to B and continue the loop with ∂w. Else do we compute a linear combination S such that $v = S^T C$.[4] Then we add $w - S^T B$ to H and break the inner loop continuing with the next column index. We are adding $w - S^T B$ to the top of H since the rows of the result should be sorted in descending order by Theorem 27.

```
    for j from n by -1 to 1 do
        w := Vector[row](n, shape=unit[j]):
        v := E[j]:
        for k to d do
            if nops(C) < LinearAlgebra:-Rank(<op(C),v>) then
                C := [op(C),v]:
                B := [op(B),w]:
                v := σ(v).T + ϑ(v):
                w := Q·w:
            else
                F := LinearAlgebra:-Transpose(<op(C),v>):
                S := LinearAlgebra:-LinearSolve(F):
                H := [w - LinearAlgebra:-Transpose(S).<op(B)>, op(H)]:
                break:
            fi:
        od: # Inner loop
    od: # Outer loop
    return <op(H)>:
end proc:
```

[4] We have to transpose C since Maple$^{\text{TM}}$'s `LinearAlgebra:-LinearSolve` expects this.

Examples of the application of this procedure can be found in Appendix B.

3.3.5 Complexity

We briefly want to reason about the complexity of Algorithm 50 applied to the conversion from a matrix $M \in R^m n$ in Popov form to Hermite normal form. The first step is to compute the truncated multiplication matrix T and the representations of the residues of the canonical basis vectors with respect to the truncated basis. Using Remark 38 and 40 this can be done by just copying the corresponding coefficients into the appropriate places. By Theorem 48 the degree bound is $\mathscr{O}(md)$ where $d = \deg M$. Hence there are at most $\mathscr{O}(md + (n-m)md) = \mathscr{O}(nmd)$ vectors in the truncated basis since in m columns—those containing the pivots—the degree of the basis elements can be at most d and since the $n-m$ other columns are truncated at degree md. For the special case of a square matrix M this means that the truncated basis contains $\mathscr{O}(md)$ elements. Thus we need at most $\mathscr{O}((mnd)^2)$ copy operations to compute T and $\mathscr{O}(mn^2d)$ for the basis representations in the general and $\mathscr{O}(m^2d^2)$ and $\mathscr{O}(m^2d)$ operations in the square case.

We must go through at most $\mathscr{O}(mnd)$ iterations of the main loop since in n columns we must go up to the bound md. Using the notations from Algorithm 50, in each iteration we multiply the vector $\mu_{\mathfrak{B}_1}$ of length $\mathscr{O}(mnd)$ by T needing $\mathscr{O}((mnd)^2)$ operations in K. Trying to solve the system $\mu_{\mathfrak{B}_1} = \sum_{\beta \in \mathfrak{B}_2} a_\beta \beta_{\mathfrak{B}_1}$ needs at most another $\mathscr{O}((mnd)^3)$ operations in K. Thus the maximum number of iterations of the loop and also of the total algorithm is $\mathscr{O}((mnd)^4)$. For square matrices this drops down to $\mathscr{O}(m^4d^4)$. This result is not too far from [Vil96].

4
Conclusion

In this report we have transfered the result of [KRT07] that Hermite and Popov normal forms are Gröbner bases with respect to different admissible orderings to Ore polynomial matrices. We have adapted the classic FGLM algorithm in order to convert a matrix from Popov form into Hermite form and vice versa. Using degree bounds for the Hermite and the Popov form, we were able to extend the FGLM algorithm to non-square matrices, i. e., to the case of infinite dimensional quotient spaces.

We also did a complexity analysis. Our results show that non-square matrices $M \in {}^{m}R^{n}$ in Popov form over Ore polynomials can be transformed into Popov form using $\mathcal{O}((mnd)^4)$ operations in K which drops down to $\mathcal{O}(m^4 d^4)$ for square matrices. The reason for the different complexities is that square matrices correspond to zero-dimensional ideals.

We have also provided an implementation in the computer algebra system Maple™ that is included in this report. Being but a toy implementation it still shows that the formulation of the FGLM algorithm for the conversion of Popov into Hermite forms is quite simple.

We have formulated our algorithm for left normal forms, i. e., for row operations. But all results should be easily translatable to right normal forms and column operations as well.

A possible extension is to try to make some kind of FGLM work with matrices that are only row-reduced. That is, we want to extend FGLM to vectorial Gröbner bases. The main problem here is that—as outlined the end of Section 3.2—division in the sense of vectorial Gröbner bases is not K-linear.

There is not much hope of extending the methods used in this report to multivariate operators, e. g., to partial differential operators. The reason for this is simply that for multivariate operators the property that a Gröbner basis has not more elements than original set of generators does in general not hold. Hence, defining normal forms of operator matrices via Gröbner bases would raise the problem that the normal form had a different (and unpredictable) size compared to the original matrix. Consequently, the transformation matrices would not be square anymore and could thus not be unimodular.

It might still be interesting to look at applications of Hermite and Popov forms in areas of mathematics like control theory or coding. It is well possible that the result in this areas can be reformulated to use Gröbner bases instead of normal forms. Then, the results might be extendable to the multivariate case.

Acknowledgements

The author would like to thank Franz Winkler, Günter Landsmann, George Labahn, Anja Korporal and Georg Regensburger for their comments and critics. The representation of Algorithm 50 was inspired by a lecture of Manuel Kauers.

Bibliography

[AL94] William Wells Adams and Philippe Loustaunau, *An introduction to Gröbner bases*, Graduate studies in mathematics, AMS, 1994.

[Art71] Emil Artin, *Galois theory*, second edition (sixth printing) ed., Notre Dame Mathematical Lectures, University of Notre Dame, Notre Dame, 1971.

[BCL06] Bernhard Beckermann, Howard Cheng, and George Labahn, *Fraction-free row reduction of matrices of Ore polynomials*, Journal of Symbolic Computation **41** (2006), 513 – 543.

[BGTV03] José L. Bueso, José Gómez-Torrecillas, and Alain Verschoren, *Algorithmic methods in non-commutative algebra*, Mathematical modelling: Theory and applications, vol. 17, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003.

[BP96] Manuel Bronstein and Marko Petkovšek, *An introduction to pseudo-linear algebra*, Theoretical Computer Science 157, 3-33 **157** (1996), 3–33.

[Buc65] Bruno Buchberger, *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)*, Ph.D. thesis, Mathematical Institute, University of Innsbruck, Austria, 1965, (English translation to appear in Journal of Symbolic Computation, 2004).

[Che03] Howard Cheng, *Algorithms for normal forms for matrices of polynomials and Ore polynomials*, Ph.D. thesis, University of Waterloo, 2003, Adviser: George Labahn.

[CK02] Richard C. Churchill and Jerald J. Kovacic, *Cyclic vectors*, Differential algebra and related topics (Li Guo, Phyllis J. Cassidy, William F. Keigher, and William Y. Sit, eds.), World Scientific Publishing Co. Pte. Ltd., 2002, pp. 191–218.

[Coh85] Paul Moritz Cohn, *Free rings and their relations*, 2nd edition ed., Academic press inc. (London) Ltd, 1985.

[Coh93] Henri Cohen, *A course in computational algebraic number theory*, Springer-Verlag New York, Inc., New York, NY, USA, 1993.

[Coh00] Paul Moritz Cohn, *An introduction to ring theory*, Springer, Berlin Heidelberg New York, 2000.

[CS98] Frédéric Chyzak and Bruno Salvy, *Non-commutative elimination in Ore algebras proves multivariate identities*, Journal of Symbolic Computation **26** (1998), no. 2, 187–227.

[FGLM93] Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard, and Teo Mora, *Efficient computation of zero-dimensional Gröbner bases by change of ordering*, J. Symb. Comput. **16** (1993), no. 4, 329–344.

[For75] G. David Forney jr., *Minimal bases of rational vector spaces with applications to multivariable linear systems*, SIAM J. Control **13** (1975), 493 – 520.

[GK09] Mark Giesbrecht and Myung Sub Kim, *Computer algebra in scientific computing*, Lecture Notes in Computer Science, vol. 5743, ch. On Computing the Hermite Form of a Matrix of Differential Polynomials, pp. 118–129, Springer, Berlin / Heidelberg, 2009.

[Her51] Charles Hermite, *Sur l'introduction des variables continues dans la théorie des nombres*, Journal der reinen und angewandten Mathematik (1851), no. 41, 191–216.

[Jac37] N. Jacobson, *Pseudo-linear transformations*, The Annals of Mathematics **38** (1937), no. 2, 484–507.

[Kou09] Christoph Koutschan, *Advanced applications of the holonomic systems approach*, Ph.D. thesis, RISC-Linz, Johannes Kepler University, September 2009.

[KRT07] Chiaki Kojima, Paolo Rapisarda, and Kiyotsugu Takaba, *Canonical forms for polynomial and quadratic differential operators*, System & Control Letters (2007), 678–684.

[MM86] F. Mora and H. Möller, *New constructive methods in classical ideal theory*, Journal of Algebra **100** (1986), no. 1, 138–178 (English).

[MS03] Thom Mulders and Arne Storjohann, *On lattice reductions for polynomial matrices*, Journal of Symbolic Computation (2003), no. 35, 377–401.

[Ore33] Oystein Ore, *Theory of non-commutative polynomials*, Annals of Mathematics **34** (1933), 480 – 508.

[Pau07] Franz Pauer, *Gröbner bases with coefficients in rings*, J. Symb. Comput. **42** (2007), no. 11-12, 1003–1011.

[Pop70] Vasile Mihai Popov, *Some properties of the control systems with irreducible matrix-transfer functions*, Seminar on Differential Equations and Dynamical Systems, II, Lecture Notes in Mathematics, Springer, Berlin / Heidelberg, 1970, pp. 169–180.

[Pop72] ______, *Invariant description of linear, time-invariant controllable systems*, SIAM Journal on Controll (1972), no. 2, 252–264.

[Tra97] Carlo Traverso, *Hilbert functions and the Buchberger algorithm*, Journal of Symbolic Computation (1997), no. 22, 355–376.

[Vil96] Gilles Villard, *Computing Popov and Hermite forms of polynomial matrices*, ISSAC, 1996, pp. 250–258.

[Zer07] Eva Zerz, *State representations of time-varying linear systems*, Gröbner bases in control theory and signal processing (Hyungju Alan Park and Georg Regensburger, eds.), Radon Series on Computational and Applied Mathematics, de Gruyter, Berlin, 2007, pp. 235–251.

A
Examples

A.1 A small example over $\mathbb{Q}[x]$

Let $R = \mathbb{Q}[x]$. We consider

$$M = \begin{pmatrix} 1 & x & 1 \\ 1 & 0 & x \end{pmatrix} \in {}^{2}R^{3}.$$

The leading monomials with respect to term over position ordering are $(0, x, 0)$ and $(0, 0, x)$ making M a Gröbner basis by Theorem 24. The degree bound for the Hermite form in $2 \deg M = 2$ by Theorem 48. The truncated canonical basis is

$$\mathfrak{B} = \mathfrak{e}_1, x\mathfrak{e}_1, x^2\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$$

and the multiplication matrix is

$$T = \begin{array}{c} \\ \mathfrak{e}_1 \\ x\mathfrak{e}_1 \\ x^2\mathfrak{e}_1 \\ \mathfrak{e}_2 \\ \mathfrak{e}_3 \end{array} \begin{array}{c} \begin{array}{ccccc} \mathfrak{e}_1 & x\mathfrak{e}_1 & x^2\mathfrak{e}_1 & \mathfrak{e}_2 & \mathfrak{e}_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

since $x\mathfrak{e}_2 \equiv -\mathfrak{e}_1 - \mathfrak{e}_2 \pmod{R^2M}$ and $x\mathfrak{e}_3 \equiv -\mathfrak{e}_1 \pmod{R^2M}$.

The first monomial with respect to position over term ordering is $\mathfrak{e}_3$ having the coordinates $(0, 0, 0, 0, 1)$. This is unequal to zero. The next monomials are $x\mathfrak{e}_3$ and $x^2\mathfrak{e}_3$ with coordinates

$$(0, 0, 0, 0, 1)T = (-1, 0, 0, 0, 0) \qquad \text{and} \qquad (-1, 0, 0, 0, 0)T = (0, -1, 0, 0, 0).$$

All three coordinate vectors are linearly independent. Since we have reached the degree bound, the next monomial to consider is $\mathfrak{e}_2$ with coordinates $(0, 0, 0, 1, 0)$. The coordinate vectors are still linearly independent. But for $x\mathfrak{e}_2$ with coordinates

$$(0, 0, 0, 1, 0)T = (-1, 0, 0, 0, -1)$$

we obtain

$$(-1,0,0,0,-1) = (-1,1,0,0)\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and hence must the last row of the Hermite form be

$$x\mathfrak{e}_2 + \mathfrak{e}_3 + x\mathfrak{e}_3 = (0, x, x+1).$$

The next monomial we must consider is $\mathfrak{e}_1$ with coordinates $(1,0,0,0,0)$. Again we have a linear dependency

$$(1,0,0,0,0) = (0,-1,0,0)\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and hence the next row of the Hermite for is

$$\mathfrak{e}_1 + x\mathfrak{e}_3 = (1, 0, x).$$

Since there are no more monomials to be considered, we have the Hermite form

$$H = \begin{pmatrix} 1 & 0 & x \\ 0 & x & 1-x \end{pmatrix}$$

which is confirmed by MAPLE$^{\text{TM}}$'s built-in procedure.

A.2 Another example over $\mathbb{Q}[x]$

The matrix

$$M = \begin{pmatrix} 0 & x & 1 & 1 \\ 0 & 1 & x+1 & x \end{pmatrix} \in {}^{2}\mathbb{Q}[x]^{4}$$

has leading monomials $(0, x, 0, 0)$ and $(0, 0, x, 0)$ and is thus a reduced Gröbner basis. The bound for the degrees in the Hermite form is 2. We choose the truncated basis

$$\mathfrak{B} = (\mathfrak{e}_1, x\mathfrak{e}_1, x^2\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3, \mathfrak{e}_4, x\mathfrak{e}_4, x^2\mathfrak{e}_4)$$

with the multiplication matrix

$$T = \begin{array}{c|cccccccc} & \mathfrak{e}_1 & x\mathfrak{e}_1 & x^2\mathfrak{e}_1 & \mathfrak{e}_2 & \mathfrak{e}_3 & \mathfrak{e}_4 & x\mathfrak{e}_4 & x^2\mathfrak{e}_4 \\ \mathfrak{e}_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ x\mathfrak{e}_1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ x^2\mathfrak{e}_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathfrak{e}_2 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ \mathfrak{e}_3 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \\ \mathfrak{e}_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ x\mathfrak{e}_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ x^2\mathfrak{e}_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

The smallest monomial is $\mathfrak{e}_4$ having coordinates $(0,0,0,0,0,1,0,0)$ in the truncated basis. Computation of the coordinates shows that $\mathfrak{e}$, $x\mathfrak{e}_4$, $x^2\mathfrak{e}_4$, $\mathfrak{e}_3$ and $x\mathfrak{e}_3$ are linearly independent. (Note that we reached the degree bound 2 in this sequence and had thus to start a new iteration with $\mathfrak{e}_3$.) But, if we compute $x^2\mathfrak{e}_3$

$$(0,0,0,0,1,0,0,0)T^2 = (0,0,0,1,2,1,1,-1)$$

then this dependes on the previous coordinate vectors: It equals

$$(1,0,-1,1,-1)\left(\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \end{array}\right)\begin{array}{c} \mathfrak{e}_4 \\ x\mathfrak{e}_4 \\ .x^2\mathfrak{e}_4 \\ \mathfrak{e}_3 \\ x\mathfrak{e}_3 \end{array}$$

This means that the last row of the Hermite form is

$$x^2\mathfrak{e}_3 - \mathfrak{e}_4 + x^2\mathfrak{e}_4 - \mathfrak{e}_3 + x\mathfrak{e}_3 = (0,0,x^2+x-1,x^2-1).$$

Also the next monomial $\mathfrak{e}_2$ with coordinates $(0,0,0,1,0,0,0,0)$ depends on the previous coordinate vectors having the representation

$$(0,-1,0,-1,-1)\left(\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \end{array}\right)\begin{array}{c} \mathfrak{e}_4 \\ x\mathfrak{e}_4 \\ .x^2\mathfrak{e}_4 \\ \mathfrak{e}_3 \\ x\mathfrak{e}_3 \end{array}$$

This leads to the next row of the Hermite normal form, namely

$$\mathfrak{e}_2 + x\mathfrak{e}_4 + \mathfrak{e}_3 + x\mathfrak{e}_3 = (0,1,x+1,x).$$

We already know the Hermite form because the number of rows must be two. But, following Algorithm 50 we continue considering now $\mathfrak{e}_1$, $x\mathfrak{e}_1$ and $x^2\mathfrak{e}_1$ with coordinate vectors $(1,0,0,0,0,0,0,0)$, $(0,1,0,0,0,0,0,0)$ and $(0,0,1,0,0,0,0,0)$. These three and the previously computed coordinate vectors are linearly independent and we stop the algorithm because we reached the degree bound. The result is

$$\begin{pmatrix} 0 & 1 & x+1 & x \\ 0 & 0 & x^2+x-1 & x^2-1 \end{pmatrix}$$

which is also computed by MAPLE$^{\text{TM}}$'s built-in procedure.

A.3 Converting from Hermite to Popov form

This time we consider the differential operators $R = \mathbb{Q}(x)[\partial; \mathrm{id}, d/dx]$ with rational coefficients. We want to convert the matrix

$$H = \begin{pmatrix} 1 & 0 & \partial^2 + x\partial + 1 \\ 0 & 1 & 2\partial - x \end{pmatrix} \in {}^2R^3$$

in Hermite normal form that was obtained in Example 54 back into Popov form. The canonical basis is $\mathfrak{e}_3, \partial\mathfrak{e}_3, \ldots$. From this we derive the truncated basis $\mathfrak{B} = \mathfrak{e}_3, \partial\mathfrak{e}_3, \partial^2\mathfrak{e}_3$ since $\deg H = 2$. The truncated multiplication matrix is just

$$T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in {}^3\mathbb{Q}(x)^3.$$

The coordinates of the canonical basis vectors are

$$(\overline{\mathfrak{e}_1})_{\mathfrak{B}} = (-1, -x, -1), \quad (\overline{\mathfrak{e}_2})_{\mathfrak{B}} = (x, -2, 0) \quad \text{and} \quad (\overline{\mathfrak{e}_3})_{\mathfrak{B}} = (1, 0, 0)$$

because $\mathfrak{e}_1 \equiv (0, 0, -\partial^2 - x\partial - 1) \pmod{R^2H}$ and $\mathfrak{e}_2 \equiv (0, 0, x - 2\partial) \pmod{R^2H}$.

One sees easily that the three smallest vectors with respect to term over position ordering—namely $\mathfrak{e}_3$, $\mathfrak{e}_2$ and $\mathfrak{e}_1$—are linearly independent modulo R^2H. We compute the coordinates of the next vector $\partial\mathfrak{e}_3$ using

$$(\overline{\partial\mathfrak{e}_3})_{\mathfrak{B}} = (1, 0, 0)T + \frac{d}{dx}(1, 0, 0) = (0, 1, 0).$$

We obtain that

$$(0, 1, 0) = (\tfrac{1}{2}x, -\tfrac{1}{2}, 0) \begin{pmatrix} 1 & 0 & 0 \\ x & -2 & 0 \\ -1 & -x & -1 \end{pmatrix}.$$

Hence, the first term over position Gröbner basis element is

$$\partial\mathfrak{e}_3 - \frac{1}{2}x\mathfrak{e}_3 + \frac{1}{2}\mathfrak{e}_2 = (0, \tfrac{1}{2}, \partial - \tfrac{1}{2}x).$$

The next vector which we must consider is $\partial\mathfrak{e}_2$. It is

$$(\overline{\partial\mathfrak{e}_2})_{\mathfrak{B}} = (x, -2, 0)T + \frac{d}{dx}(x, -2, 0) = (1, x, -2)$$

and

$$(1, x, -2) = (3 + \tfrac{3}{2}x^2, -\tfrac{3}{2}x, 2) \begin{pmatrix} 1 & 0 & 0 \\ x & -2 & 0 \\ -1 & -x & -1 \end{pmatrix}.$$

Thus, the second vector in the term over position Gröbner basis is

$$\partial\mathfrak{e}_2 - (3 + \frac{3}{2}x^2)\mathfrak{e}_3 + \frac{3}{2}x\mathfrak{e}_2 - 2\mathfrak{e}_1 = (-2, \partial + \frac{3}{2}x, -3 - \frac{3}{2}x^2).$$

Thus, the Popov form of H is

$$\begin{pmatrix} -2 & \partial + \frac{3}{2}x & -3 - \frac{3}{2}x^2 \\ 0 & \frac{1}{2} & \partial - \frac{1}{2}x \end{pmatrix}$$

precisely as in Example 54.

B

Example session in Maple$^{\text{TM}}$

B.1 Computing in $\mathbb{Q}[x]$

In order to work with commutative polynomials we set $\sigma = \text{id}$ and ϑ to be the zero map. The variable is denoted by x.

```
Q := 'x': σ:= v → v: ϑ:= v → map(0,v):
```

The example from Section A.1 is then defined as follows

```
M := <<1|x|1>,<1|0|x>>;
```

$$M := \begin{bmatrix} 1 & x & 1 \\ 1 & 0 & x \end{bmatrix}$$

The Hermite form is computed by our procedure using

```
H := Convert(Q,σ,ϑ,M);
```

$$H := \begin{bmatrix} 1 & 0 & x \\ 0 & x & 1-x \end{bmatrix}$$

We may check this result using Maple$^{\text{TM}}$'s built-in procedure for Hermite form computation:

```
H = LinearAlgebra:-HermiteForm(M);
```

$$\begin{bmatrix} 1 & 0 & x \\ 0 & x & 1-x \end{bmatrix} = \begin{bmatrix} 1 & 0 & x \\ 0 & x & 1-x \end{bmatrix}$$

The example from Section A.1 is checked by:

```
M := <<0|x|1|1>,<0|1|x+1|x>>:
H := Convert(Q,σ,ϑ,M):
H = LinearAlgebra:-HermiteForm(M);
```

$$\begin{bmatrix} 0 & 1 & x+1 & x \\ 0 & 0 & x^2-1+x & -1+x^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & x+1 & x \\ 0 & 0 & x^2-1+x & -1+x^2 \end{bmatrix}$$

B.2 Computing in $\mathbb{Q}(x)[\partial; \mathrm{id}, d/dx]$

For checking Example 54 using our implementation we first define σ and τ. Since Maple™ wont let us use `D` as symbol we will denote the variable ∂ by `Q`.

```
Q := 'Q': σ := v → v: ϑ:= v → map(p → diff(p,'x'), v):
```

The matrix M from Example 54 is defined in Maple™ by

```
M := <<-2|Q+3/2·x|-3-3/2·x^2>,<0|1/2|Q-1/2·x>>;
```

$$M := \begin{bmatrix} -2 & Q + 3/2\,x & -3 - 3/2\,x^2 \\ 0 & 1/2 & Q - 1/2\,x \end{bmatrix}$$

and its Hermite form is computed via

```
H := Convert(Q,σ,ϑ,M);
```

$$H := \begin{bmatrix} 1 & 0 & 1 + x\,Q + Q^2 \\ 0 & 1 & -x + 2\,Q \end{bmatrix}$$

Since there seems to be no built-in Hermite form procedure for general Ore polynomials we cannot check this result in Maple™.