Algorithmic aspects of algebraic system theory

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften genehmigte Dissertation vorgelegt von

Diplom-Mathematikerin

Kristina Schindelar

aus Bratislava

Berichter: Universitätsprofessorin Dr. Eva Zerz
Universitätsprofessor Dr. Sebastian Walcher

Tag der mündlichen Prüfung: 16. März 2010

Contents

Preface

The mathematical roots of system and control theory date back to the paper "On Governors" by J. C. Maxwell published 1868 in *Proceedings of the Royal Society of London*. The seminal work of R. E. Kalman established system theory as a mathematical discipline in the 1950s. About thirty years later, J. C. Willems proposed a novel approach to signals and systems, the so-called behavioral approach. This approach offers a very general definition of a dynamical system, a triple consisting of the mathematical model of time, the system-relevant quantities summarized in the so-called signals, and a system law, that is, equations defining the relations between the signals. The contribution of U. Oberst, which appeared in 1990, gives fundamental insight for algebraic system theory. A very important algebraic property of the signal space is realized to be highly copious for signals and systems there, namely the property of the signal space to be an injective cogenerator over the underlying operator ring. The algebraic approach to system theory has been developed among others by B. Malgrange, U. Oberst, J. F. Pommaret, A. Quadrat and E. Zerz.

The goal of algebraic system theory is the structural analysis of dynamical systems using algebraic tools. These systems may arise from various practical problems settled for instance in a scientific, technical or economical area. The systems are mainly described via differential or difference equations. Their solutions are contained in a certain signal space which possesses a module structure over the ring of differential/difference operators. In case the signal space is an injective cogenerator, algebraic properties of the system module are dual to analytic properties of the signals due to Oberst's observation. Then control theoretic characterizations like autonomy, controllability and observability can be translated into algebraic terms.

Classically linear time-invariant systems with field coefficients are studied. In the recent past variations of these systems have proved to be worthy for extended studies. From the applied point of view, there is obviously the interest to consider corresponding generalizations. From the algebraic point of view, some particular settings are very interesting for further investigations since ring theory and homological algebra provide a deep insight. Beyond theoretical studies, the computer algebra machinery allows the enormous benefit of constructive analyses. This thesis elaborates both aspects, the theoretical and the computational, in parallel. It is organized as follows.

Chapter 1 and Chapter 2 serve for an extended introduction. System theoretical aspects are provided in Chapter 1. Basic concepts and definitions are presented and furthermore the following chapters are motivated from the system theoretical point of view. Section 1.3 motivates the subject of study of Chapter 3, Section 1.4 points out the relevance of Chapter 4 and finally Section 1.5 and Section 1.6 give an introduction

to Chapter 5. Chapter 2 is devoted to Gröbner bases theory. Beside the classical case of polynomial modules with field coefficients, we discuss ring coefficients and G-algebras. Connections between G-algebras and Ore algebras are outlined. Furthermore their relevance for system theory is shown and the algorithmic motivation for the following Chapters is composed.
Chapter 3 studies systems with coefficients in a finite ring, in contrast to the classical case. The general motivation for this framework stems mainly from communication theory. However, the extension leads to problems like zero-divisors and the principal ideal domain property is lost. Therefore concepts useful for coding fail to generalize straightforwardly. In the field case the so-called predictable degree property is useful for many areas of system theory, ranging from controller parameterization to minimal realizations of linear systems over fields. This property does not carry over directly to the ring case. The paper "The predictable degree property and row reducedness for systems over a finite ring" by M. Kuijper, R. Pinto, J. W. Polderman and P. Rocha [KPP07] establishes a new framework which allows the adoption of that classical result in a novel setting. Results of that work were presented in the plenary talk of M. Kuijper at the international symposium "Mathematical Theory of Networks and Systems" in 2008. Thereupon J. Rosenthal proposed the conjecture that the presented results are closely connected to the topic of Gröbner bases. This has proved to be correct. By the tool of Gröbner bases the results of [KPP07] are extended to a more general framework which additionally allows concrete calculations. For this purpose the notion of the so-called minimal Gröbner p-basis is established and the connection to known results is pointed out. The application to parametrization of all shortest linear recurrence relations and to minimal state realization are discussed. The results presented in Chapter 3 are based on joint work with M. Kuijper.
Chapter 4 is focused on one-dimensional systems with time-varying rational coefficients. This leads to the non-commutative operator ring called rational Weyl algebra which is a principal ideal domain. Therefore the non-commutative analogon to the Smith form, the so-called Jacobson form, exists. This normal form can be used to obtain a decomposition into a controllable and an autonomous subsystem of the corresponding linear abstract system. Furthermore the order of the underlying ordinary differential equation system is obtained directly. But computational problems known from the commutative counterpart even increase due to the non-commutative structure, namely the explosive growth of the coefficients. A novel approach which can be applied in a completely fraction free framework is presented in this chapter. This approach shows first how to obtain a decoupled form. It should be stressed that this decoupled form may even be interesting by itself. Further we show how to obtain a normal form from the decoupled form. Due to collaboration with V. Levandovskyy the proposed algorithm can even be applied to an extended operator class of certain G-algebras. The implementation is realized as a library called `jacobson.lib` for the computer algebra system Singular::Plural [GPS05, GLH05], which is freely available. This implementation is compared with all implementations which are available to the best of our knowledge.
In [AW93] a behavioral approach to linear exact modeling is formulated for one-dimensional systems with constant coefficients. This problem of system identification is extended to a multi-dimensional setting in [Zer05, Zer08]. In co-operation with V. Levandovskyy and E. Zerz, this modeling concept is developed for polynomial-

exponential signals in a multi-dimensional time-varying model class in Chapter 5. These model classes are summarized in the so-called Ore algebras. The idea of this approach is to derive a model describing the observed data and containing as much information as possible. It turns out that the particular model classes yield a very precise description, as pointed out in the case of continuous systems. Two alternative possibilities to calculate the models will be presented, one of them working in a purely commutative framework. Both rely on annihilator calculations which are constructively tackled using Gröbner bases. All constructions can be realized in SINGULAR::PLURAL.

Acknowledgment

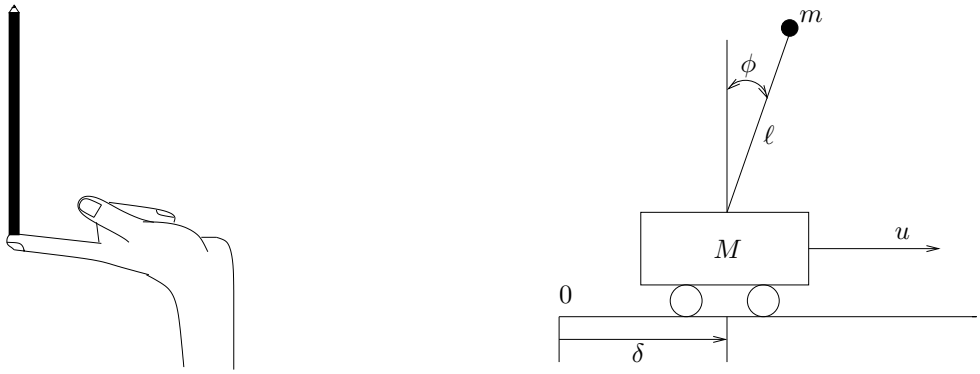
First of all, I would like to express my deepest gratitude to Prof. Dr. Eva Zerz for her guidance and advice which made this thesis possible. For the support and numerous fruitful discussions during the 6 months visit at the Melbourne University I would like to thank Prof. Dr. Margreta Kuijper. Further I am grateful to Prof. Dr. Sebastian Walcher for helpful suggestions, and for an excellent collaboration I want to thank Dr. Viktor Levandovskyy deeply.
Several organizations funded this thesis. I would like to thank the Deutsche Forschungsgemeinschaft for my scholarship in the research training group "Hierarchie und Symmetrie in mathematischen Modellen", the Deutscher Akademischer Austausch Dienst for financial support during my stay in Australia, and the FAZIT Stiftung for supporting me during the last eight months of the thesis project.
Furthermore, I would like to thank my colleagues for the great atmosphere and helpful tips especially Dr. Annika Meyer, Dr. Markus Kirschmer and Moritz Schröer.

Chapter 1

Introduction to algebraic system theory

1.1 Motivation

A demonstrative control theoretical example is the balance of a pencil on the fingertip, see the following left figure. On the right, a simplified but closely related problem is drafted, a so-called inverted pendulum.



The right figure shows a cart driven by a motor of force u along the line. The coefficient of friction on the cart is denoted by F. An inverted pendulum is attached through a frictionless pivot on the top of the cart. All motions are assumed in a plane. We denote the angle between the pendulum of length ℓ and the vertical by ϕ, the gravitation constant by g, the position of the cart by δ, the mass of the cart by M and the mass of the pendulum by m, which is assumed to be concentrated at the tip. The mathematical description of the inverted pendulum is given by:

$$(M+m)\ \frac{d^2}{dt^2}\delta + m\ell\cos(\phi)\ \frac{d^2}{dt^2}\phi - m\ell\sin(\phi)\ (\frac{d}{dt}\phi)^2 + F\ \frac{d}{dt}\delta = u$$

and

$$\ell\frac{d^2}{dt^2}\phi - g\sin(\phi) + \cos(\phi)\ \frac{d^2}{dt^2}\delta = 0.$$

This is a non-linear system of ordinary differential equations. The non-linearity is typical for real life examples. However, it is not convenient for algebraic studies. Therefore we consider exclusively linearized models. A linearization can be applied in

different ways. If just an equilibrium state is relevant, one linearizes the equations at this equilibrium of the system. For this purpose the non-linear parts of the equation are expanded into the first two summands of their Taylor series around the equilibrium. This linearization leads to equations with constant coefficients. The above system is interesting for small angles ϕ, thus we obtain the linearized equations:

$$(M+m)\frac{d^2}{dt^2}\delta + m\ell\frac{d^2}{dt^2}\phi + F\frac{d}{dt}\delta = u \quad \text{and} \quad \ell\frac{d^2}{dt^2}\phi - g\phi + \frac{d^2}{dt^2}\delta = 0.$$

One can easily check that the representation matrix possesses Smith form $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. As we will see later, this means that the underlying system is controllable.

Another possibility is to linearize the system along a trajectory. Think of a non-linear system for which a certain motion is relevant, like for instance the motion of a robot arm. Then the Taylor series needs to be evaluated at the relevant trajectory. This leads to differential equations with time-varying coefficients.

1.2 Basic properties of linear systems

In order to derive a general and formal setting we need to define precisely the term "system" first. This has been done by J.C. Willems [Wil88] giving a definition consisting of three components:

- The set T, a mathematical model of time.
- The signal value set W in which the signals take their values. A signal is a map from T to W.
- The subset $\mathcal{B}$ of the **signal set** W^T called **behavior**, constraining the signals by relevant conditions. Usually the system law will not be satisfied by all elements of W^T, thus to get an adequate description of the behavior, we have to search for all trajectories satisfying the system law, i.e.

$$\mathcal{B} = \left\{\omega \in W^T \mid \omega \text{ satisfies the system law}\right\}.$$

According to Willems, a triple $\Sigma = (T, W, \mathcal{B})$ is called **system**. In the one-dimensional setting one can think of the time set $T = \mathbb{N}$ for the discrete, and $T = \mathbb{R}$ for the continuous framework. The system law may then be given by ordinary difference or by differential equations. We suppose W to be a K-vector space for a field K. This thesis is exclusively devoted to **linear** systems, i.e. systems satisfying that for all $\omega_1, \omega_2 \in \mathcal{B}$ and $k_1, k_2 \in K$ it follows that

$$k_1\omega_1 + k_2\omega_2 \in \mathcal{B}.$$

Linearity is an intrinsic system property. Another important and intrinsic system property is its time-variance or time-invariance. Suppose the time set to be additively closed. Then a system Σ is called **time-invariant** if for all $\omega(\cdot) \in \mathcal{B}$ and $t \in T$ the trajectory $\omega(\cdot + t)$ is contained in $\mathcal{B}$.

In [Obe90] U. Oberst gave a modified definition of a system to permit an extended class of signals, for instance the class of distributions. Let $\mathcal{A}$ denote a set of scalar-valued signals, let q denote the number of signals occurring in the system and suppose $\mathcal{B} \subseteq \mathcal{A}^q$. Then a system is the triple $(\mathcal{A}, q, \mathcal{B})$. Note that the number of signals q is an intrinsic system property whereas a system law can be formulated in several ways. The question under which conditions one can find a kernel representation can not be answered in this thesis. However, in the sequel we will assume the system to possess a kernel representation. Let $\mathcal{D}$ denote a left Noetherian ring with unity acting on the left $\mathcal{D}$-module $\mathcal{A}$. Note that whenever an element $d \in \mathcal{D}$ acts on an element $\omega \in \mathcal{A}$, we write $d \bullet \omega$ to keep the notation accurate. For some positive integer g and $R \in \mathcal{D}^{g\times q}$, we call the behavior

$$\mathcal{B} = \{\omega \in \mathcal{A}^q \mid R \bullet \omega = 0\},$$

a **linear abstract system**. The matrix R is called **representation** of $\mathcal{B}$. Its number of rows g can obviously vary. The set $\mathcal{B}$ is an additive subgroup of $\mathcal{A}^q$. But note that a linear abstract system does not have a $\mathcal{D}$-module structure in general. This structural property depends on the choice of $\mathcal{D}$. If $\mathcal{D}$ is commutative, then $\mathcal{B}$ is a $\mathcal{D}$-module. In many cases of relevance, system classes boil down to one of those listed below.

Example 1.2.1

1. **One-dimensional systems**

   - **Continuous ODE's with constant coefficients**
     Describing a dynamical process given by linear ordinary differential equations with constant coefficients, choose $\mathcal{D} = \mathbb{C}[\partial]$ and $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, the space of smooth functions, or $\mathcal{A} = \mathcal{D}'(\mathbb{R}, \mathbb{C})$, the set of distributions. Apparently $\mathcal{A}$ becomes a $\mathcal{D}$-module with

     $$\partial \bullet \omega = \frac{d\omega}{dt}.$$

     Choosing $R \in \mathcal{D}^{g\times q}$, the corresponding linear abstract system becomes a $\mathcal{D}$-module.
   - **Discrete ODE's with constant coefficients**
     The discrete analogon to the previous item can be modelled by setting $\mathcal{D} = \mathbb{F}[\mathbf{s}]$, where $\mathbb{F}$ denotes a field and $\mathcal{A} = \mathbb{F}^T$, where $T = \mathbb{N}$. Then $\mathbf{s}$ acts on $\mathcal{A}$ as the backward shift, i.e., $(\mathbf{s} \bullet \omega)(t) = \omega(t+1)$ for all $\omega \in \mathcal{A}$. This yields a system representation consisting of linear ordinary difference equations with constant coefficients.
   - **ODE's with polynomial coefficients**
     Lifting the situation to differential equations with polynomial coefficients we need to switch to the so-called first Weyl algebra $\mathrm{W}_1 = \mathbb{C}[t][\partial; \mathrm{id}_{\mathrm{W}_1}, \frac{d}{dt}]$, where

     $$\partial t = t\partial + 1$$

     and

     $$\partial \bullet \omega = \frac{d\omega}{dt} \quad \text{and} \quad t \bullet \omega = t\omega.$$

For a precise and more general definition see Example 2.2.5, where additionally the changeover to the multi-dimensional and discrete case are studied exactly. Let $\mathcal{A} = \mathbb{R}[t]$. Then

$$\mathcal{B} = \{\omega \in \mathcal{A} \mid \partial \bullet \omega = 0\}$$

describes the set of constant functions. For a constant function $c \neq 0$, the W_1-module product tc is obviously not contained in $\mathcal{B}$:

$$\partial \bullet (tc) = c.$$

Thus $\mathcal{B}$ is not a W_1-module.

- **Discrete ODE's with coefficients in finite rings**
  Another crucial system class consists of signals with values in a finite ring, for example $\mathbb{Z}_{p^r}$, for a prime p, and a system law given by difference equations with coefficients in $\mathbb{Z}_{p^r}$. Thus $\mathcal{A} = \mathbb{Z}_{p^r}^{\mathbb{N}}$ and $\mathcal{D} = \mathbb{Z}_{p^r}[\mathbf{s}]$.

2. **Multi-dimensional systems**

   - **PDE's with constant coefficients**
     - Substituting $\mathbb{C}[\partial]$ by $\mathbb{C}[\partial_1, \ldots, \partial_n]$ and suitably $\mathcal{A}$ by $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ gives a generalization of ordinary differential equations with constant coefficients.
     - Substituting $\mathbb{F}[\mathbf{s}]$ by $\mathbb{F}[\mathbf{s}_1, \ldots, \mathbf{s}_n]$ and suitably $\mathcal{A}$ by $\mathbb{F}^{T^n}$ leads to a generalization of ordinary difference equations with constant coefficients.

Note that the shift operator is mostly denoted by σ instead of $\mathbf{s}$ in the literature, but to match the notation common for Ore-algebras, we reserve this symbol for a ring endomorphism, see Definition 2.2.4. Furthermore note that the discussed operator rings are of polynomial nature, a structure that can be tackled efficiently by the computer algebra machinery.

Let $\mathcal{B} = \{\omega \in \mathcal{A}^q \mid R \bullet \omega = 0\}$ be an abstract linear system, where $R \in \mathcal{D}^{g \times q}$. We define the **system module**

$$\mathcal{M} := \mathcal{D}^{1\times q}/\mathcal{D}^{1\times g}R.$$

Due to the Malgrange isomorphism [Mal64], the group isomorphism

$$\mathcal{B} \cong \mathrm{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A})$$

holds. Recall that a left $\mathcal{D}$-module $\mathcal{A}$ is called **injective** if $\mathrm{Hom}_{\mathcal{D}}(\cdot, \mathcal{A})$ is an exact functor, i.e. for the left $\mathcal{D}$-modules $\mathcal{H}$, $\mathcal{N}$ and $\mathcal{P}$, the exactness of the sequence

$$\mathcal{H} \xrightarrow{f} \mathcal{N} \xrightarrow{g} \mathcal{P} \tag{1.1}$$

implies the exactness of

$$\mathrm{Hom}_{\mathcal{D}}(\mathcal{H}, \mathcal{A}) \xleftarrow{\mathrm{Hom}_{\mathcal{D}}(f,\mathcal{A})} \mathrm{Hom}_{\mathcal{D}}(\mathcal{N}, \mathcal{A}) \xleftarrow{\mathrm{Hom}_{\mathcal{D}}(g,\mathcal{A})} \mathrm{Hom}_{\mathcal{D}}(\mathcal{P}, \mathcal{A}), \tag{1.2}$$

where

$$\mathrm{Hom}_{\mathcal{D}}(f,\mathcal{A}) : \mathrm{Hom}_{\mathcal{D}}(\mathcal{N},\mathcal{A}) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{H},\mathcal{A}), \quad \phi \mapsto \phi \circ f.$$

Let $\mathcal{A}$ be injective and suppose our system is given via an image representation, that is, $v \in \mathcal{B}$ if $R \bullet \omega = v$ for some $\omega \in \mathcal{A}^q$. The left kernel $\ker(\cdot R)$ is a finitely generated $\mathcal{D}$-module since $\mathcal{D}$ is Noetherian. Thus there exists a matrix $Z \in \mathcal{D}^{h\times g}$ such that $\ker(\cdot R) = \mathrm{im}(\cdot Z)$ and therefore

$$\mathcal{D}^{1\times h} \xrightarrow{\cdot Z} \mathcal{D}^{1\times g} \xrightarrow{\cdot R} \mathcal{D}^{1\times q}$$

is exact. This implies the exactness of

$$\mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times h},\mathcal{A}) \xleftarrow{\mathrm{Hom}_{\mathcal{D}}(\cdot Z,\mathcal{A})} \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times g},\mathcal{A}) \xleftarrow{\mathrm{Hom}_{\mathcal{D}}(\cdot R,\mathcal{A})} \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times g},\mathcal{A}).$$

Since further $\mathcal{A} \cong \mathrm{Hom}_{\mathcal{D}}(\mathcal{D},\mathcal{A})$, due to the Malgrange isomorphism we obtain the exact sequence

$$\mathcal{A}^h \xleftarrow{Z\bullet} \mathcal{A}^g \xleftarrow{R\bullet} \mathcal{A}^q$$

and thus $\ker(Z\bullet) = \mathrm{im}(R\bullet)$. This yields the so-called **fundamental principle**.

Theorem 1.2.2 *For an injective $\mathcal{D}$-module $\mathcal{A}$, two matrices $R \in \mathcal{D}^{g\times q}$ and $Z \in \mathcal{D}^{h\times g}$ such that* $\ker(\cdot R) = \mathrm{im}(\cdot Z)$ *and $v \in \mathcal{A}^q$ it follows that*

$$\exists \omega \in \mathcal{A}^q : \; R \bullet \omega = v \;\Leftrightarrow\; Z \bullet v = 0.$$

If exactness of (1.1) and (1.2) are equivalent, then $\mathcal{A}$ is called **injective cogenerator**. Interesting examples are for instance:

1. $\mathcal{D} = \mathbb{C}[\partial_1,\ldots,\partial_n]$ and $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}^n,\mathbb{C})$ or $\mathcal{A} = \mathcal{D}'(\mathbb{R}^n,\mathbb{C})$
2. $\mathcal{D} = \mathbb{F}[\mathbf{s}_1,\ldots,\mathbf{s}_n]$ and $\mathcal{A} = \mathbb{F}^{\mathbb{N}^n}$
3. $\mathcal{D} = \mathbb{F}[\mathbf{s}_1,\ldots,\mathbf{s}_n,\mathbf{s}_1^{-1},\ldots,\mathbf{s}_n^{-1}]$ and $\mathcal{A} = \mathbb{F}^{\mathbb{Z}^n}$

In the sequel we discuss the duality between $\mathcal{B}$ and $\mathcal{M}$ which occurs if $\mathcal{A}$ is an injective cogenerator. Define for a linear abstract system $\mathcal{B}$ the left $\mathcal{D}$-module

$$\mathcal{B}^\perp := \{x \in \mathcal{D}^{1\times q} \mid x \bullet \omega = 0 \;\forall \omega \in \mathcal{B}\}.$$

By assumption, $\mathcal{D}$ is Noetherian which yields $\mathcal{B}^\perp = \mathcal{D}^{1\times g_1} R_1$ for some R_1. We claim that

$$\mathcal{B}^\perp = \mathcal{D}^{1\times g} R.$$

Proof: Since every row of R is contained in $\mathcal{B}^\perp$, it follows that

$$\mathcal{D}^{1\times g} R \subseteq \mathcal{D}^{1\times g_1} R_1. \tag{1.3}$$

To show the other inclusion, let $\mathcal{B}_1 := \{\omega \in \mathcal{A}^q \mid R_1 \bullet \omega = 0\}$. Due to (1.3) there exists a matrix X such that $R = XR_1$ and thus $\mathcal{B}_1 \subseteq \mathcal{B}$. Further any $\omega \in \mathcal{B}$ is annihilated

by every $x \in \mathcal{B}^\perp$ and thus it is annihilated by R_1. This yields $\mathcal{B} = \mathcal{B}_1$. Therefore the sequence

$$0 \longleftarrow \mathcal{B} \longleftarrow \mathcal{B}_1$$

is exact. Then due to the Malgrange isomorphism

$$0 \longleftarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A}) \longleftarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{M}_1, \mathcal{A}), \quad \text{where} \quad \mathcal{M}_1 = \mathcal{D}^{1\times q}/\mathcal{D}^{1\times g_1} R_1,$$

is exact. Since $\mathcal{A}$ is an injective cogenerator, it follows that

$$0 \longrightarrow \mathcal{M} \longrightarrow \mathcal{M}_1$$

is exact. This yields the claim. □

Furthermore the results lead to the following relation:

Theorem 1.2.3 *Let $\mathcal{B}_i$ be represented by $R_i \in \mathcal{D}^{g_i \times q}$ for $i = 1, 2$. Then*

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \quad \Leftrightarrow \quad \mathcal{B}_1^\perp \supseteq \mathcal{B}_2^\perp \quad \Leftrightarrow \quad \mathcal{D}^{1\times g_1} R_1 \supseteq \mathcal{D}^{1\times g_2} R_2.$$

We obtain three characteristic properties.

Remark 1.2.4

- Let $\mathcal{B}_i$ be represented by $R_i \in \mathcal{D}^{g_i \times q}$ for $i = 1, 2$. Then $\mathcal{B}_1 = \mathcal{B}_2$ if and only if $\mathcal{D}^{1\times g_1} R_1 = \mathcal{D}^{1\times g_2} R_2$.
- Let $d \in \mathcal{D}$. Then $d \bullet \mathcal{A} = \{0\}$ if and only if $d = 0$.
- $\mathcal{B} = \{0\}$ if and only if there exists a matrix $X \in \mathcal{D}^{q\times g}$ such that $XR = \mathrm{I}_q$.

In contrast to the already listed cogenerators the $\mathbb{Z}[\mathbf{s}, \mathbf{s}^{-1}]$-module $\mathbb{Z}^{\mathbb{Z}}$ is not an injective cogenerator. Consider the system given by

$$\mathcal{B} := \{\omega \in \mathbb{Z}^{\mathbb{Z}} \mid 2 \bullet \omega = 0\}.$$

Then $\mathcal{B}$ obviously equals

$$\{\omega \in \mathbb{Z}^{\mathbb{Z}} \mid \omega = 0\},$$

but

$$_{\mathbb{Z}[\mathbf{s},\mathbf{s}^{-1}]}\langle 2\rangle \subsetneqq \mathcal{B}^\perp = \mathbb{Z}[\mathbf{s}, \mathbf{s}^{-1}].$$

For the rest of this section, let $\mathcal{A}$ be an injective cogenerator. In the sequel we will outline basic system properties and show how these translate into algebraic structure properties of the associated system module. The presented results are based on [Zer06c].

Autonomy

For i contained in $\{1, \ldots, q\}$, let

$$\pi_i : \mathcal{B} \to \mathcal{A}, \quad \omega \mapsto \omega_i$$

be the i-th projection of $\mathcal{B}$.

Definition 1.2.5 We call ω_i a **free variable** if π_i is surjective.

We call a system **autonomous** if it admits no free variables. The interpretation is self-explanatory. If π_i is surjective, then an arbitrary $f \in \mathcal{A}$ can be fixed and there exist $\omega_1, \ldots, \omega_{i-1}, \omega_{i+1}, \ldots, \omega_q \in \mathcal{A}$ such that $(\omega_1, \ldots, \omega_{i-1}, f, \omega_{i+1}, \ldots, \omega_q)^T \in \mathcal{B}$. Autonomy can be characterized via the system module. In the sequel we assume $\mathcal{D}$ to be a domain. How to deal with the operator ring $\mathbb{Z}_{p^r}[\mathbf{s}]$, which possesses zero-divisors, will be discussed in the next Section.

Lemma 1.2.6 *If the system module $\mathcal{M}$ is torsion, then $\mathcal{B}$ is autonomous.*

Proof: Suppose that $\mathcal{B}$ is not autonomous. Then there exists $1 \leq i \leq q$ such that

$$\mathcal{B} \xrightarrow{\pi_i} \mathcal{A} \longrightarrow 0$$

is exact, that is,

$$\mathrm{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A}) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}, \mathcal{A}) \longrightarrow 0$$

is exact and yields the exact sequence

$$\mathcal{M} \longleftarrow \mathcal{D} \longleftarrow 0.$$

That is, there exists an injective homomorphism ι from $\mathcal{D}$ to $\mathcal{M}$. Let m denote the image of 1 under ι. Then $0 = dm = d\iota(1) = \iota(d)$ implies that $d = 0$. Thus m is not a torsion element and thus $\mathcal{M}$ is not torsion. □

The implication of the previous lemma turns out to be an equivalence if $\mathcal{D}$ possesses the so-called left Ore property. The ring $\mathcal{D}$ has the **left Ore property** if any non-zero elements $d_1, d_2 \in \mathcal{D}$ have a left common multiple, that is, there exist $c_1, c_2 \in \mathcal{D}$ not both equal to zero such that $c_1 d_1 = c_2 d_2$. One can show that $\mathcal{D}$ possesses the left Ore property if it is a left Noetherian domain. Since we assume the domain $\mathcal{D}$ to be Noetherian, the following claim holds:

$$\mathcal{B} \text{ autonomous} \quad \Rightarrow \quad \mathcal{M} \text{ torsion.}$$

Assume that $\mathcal{M}$ is not torsion. Let e_i denote the i-th unit vector. Using the left Ore property of $\mathcal{D}$, one can show that there exists $1 \leq i \leq q$ such that $[\mathrm{e}_i]$ is not torsion. Defining the homomorphism $\iota : \mathcal{D} \to \mathcal{M}$ by $\iota(1) := [\mathrm{e}_i]$ yields the exact sequence

$$0 \longrightarrow \mathcal{D} \xrightarrow{\iota} \mathcal{M}.$$

Then

$$0 \longleftarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}, \mathcal{A}) \longleftarrow \mathrm{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A})$$

is exact. The Malgrange isomorphism yields

$$0 \longleftarrow \mathcal{A} \xleftarrow{\pi_i} \mathcal{B},$$

that is, the surjective projection π_i. Thus $\mathcal{B}$ is not autonomous.
Finally we obtain the following characterization of autonomy.

Theorem 1.2.7 *The following assertions are equivalent:*

1. $\mathcal{M}$ *is torsion.*
2. *There exists* $0 \neq d \in \mathcal{D}$ *and* $X \in \mathcal{D}^{q\times g}$ *such that* $d\,\mathrm{I}_q = XR$.
3. $\mathcal{B}$ *is autonomous.*

Proof: Due to the previous observation it is sufficient to show that the first item yields the second and the second implies the third to prove the claim. So let $\mathcal{M}$ be torsion. For all residue classes $[\mathrm{e}_i] \in \mathcal{M}$ there exist non-zero elements $d_i \in \mathcal{D}$ such that $d_i[\mathrm{e}_i] = 0$, that is, $d_i\,\mathrm{e}_i = y_i R$ for $y_i \in \mathcal{D}^{1\times g}$. Due to the Ore property $\{d_1, \ldots, d_q\}$ possesses a left common multiple $0 \neq d = c_i d_i$ for suitable c_i. Thus

$$d\,\mathrm{I}_q = \mathrm{diag}(c_1 d_1, \ldots, c_q d_q) = \underbrace{\begin{bmatrix} c_1 y_1 \\ \vdots \\ c_q y_q \end{bmatrix}}_{=:X} R.$$

Defining X as indicated yields the second item.
Now let $d\,\mathrm{I}_q = XR$ for a non-zero element $d \in \mathcal{D}$. Then $d \bullet \omega_i = 0$ for any component of all $\omega \in \mathcal{B}$. Suppose $\mathcal{B}$ is not autonomous. Then there exists $1 \leq i \leq q$ such that the i-th projection of $\mathcal{B}$ is surjective. Then $d \bullet \pi_i(\mathcal{B}) = d \bullet \mathcal{A} = \{0\}$. Due to Remark 1.2.4 $d \bullet \mathcal{A} = \{0\}$ if and only if $d = 0$, which leads to a contradiction to the assumptions. □

The required assumption on $\mathcal{D}$ to be a domain and left Noetherian permits to define the fraction field $\mathcal{K}$ of $\mathcal{D}$ as

$$\mathcal{K} := \{d^{-1}n \mid d, n \in \mathcal{D},\ d \neq 0\}.$$

Then the column rank of R is defined as the dimension of $V := R\mathcal{K}^q \subseteq \mathcal{K}^g$. One can show that $\dim(R\mathcal{K}^q) = \dim(\mathcal{K}^{1\times g}R)$, that is, the column rank and the row rank of R coincide and we write

$$\mathrm{rank}(R) := \dim(V).$$

One can show that R has full column rank if and only if there exists a non-zero $d \in \mathcal{D}$ such that $XR = d\,\mathrm{I}_q$. This yields the following corollary.

Corollary 1.2.8 *The linear abstract system* $\mathcal{B}$ *is autonomous if and only if* R *has full column rank.*

Input-output structures

For a linear abstract system, there is no a priori classification of free and dependent variables. However, such a partition of variables may be desirable. Let us discuss how to obtain such a representation. The matrices R_1, R_2 denote two representations of $\mathcal{B}$. Due to Remark 1.2.4, there exist suitable matrices X, Y such that

$$R_1 = XR_2 \quad \text{and} \quad R_2 = YR_1. \tag{1.4}$$

This permits to define the so-called **output-dimension** of $\mathcal{B}$ as the rank of the corresponding representation matrix. Let $p := \operatorname{rank}(R)$. Then R possesses p columns $R_{j_1}, \ldots, R_{j_p}$ forming a basis of $V = R\mathcal{K}^q$. Without loss of generality we may assume

$$R = [-Q,\ P], \text{ where } P = [R_{j_1}, \ldots, R_{j_p}].$$

Writing each signal $\omega = [u^T,\ y^T]^T$ corresponding the introduced partition of R, the system $\mathcal{B}$ can be written as

$$\{\begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{A}^q \mid Q \bullet u = P \bullet y\}, \tag{1.5}$$

called an **input-output** structure of $\mathcal{B}$. Since $R_{j_1}, \ldots, R_{j_p}$ form a basis of V, there exists a uniquely determined matrix $H \in \mathcal{K}^{p\times m}$ such that $Q = PH$. We call $m := q-p$ the **input-dimension** and H the **transfer matrix** of $\mathcal{B}$. One can easily show that these definitions are independent of the chosen representation:
Assume $R_1 = [-Q_1,\ P_1]$ and $R_2 = [-Q_2,\ P_2]$ to be two representations of $\mathcal{B}$. Then $Q_1 = P_1H_1$ and $Q_2 = P_2H_2$. Using (1.4) we obtain that $P_1(H_1 - H_2) = 0$ and since P_1 is of full column rank $H_1 = H_2$ holds.
It is still left to show that the chosen terminology is justified. We show that the vector u consists of free variables:
Let the rows of Z generate the left kernel of P, i.e., $\ker(\cdot P) = \operatorname{im}(\cdot Z)$. Due to the fundamental principle,

$$\text{there exists } y \in \mathcal{A}^p \text{ such that } P \bullet y = Q \bullet u \quad \text{if and only if} \quad (ZQ) \bullet u = 0.$$

Since $Q = PH$, we obtain that $ZQ = ZPH = 0$ and hence $(ZQ) \bullet u = 0$ holds for any $u \in \mathcal{A}^m$. That is, u consists of free variables and therefore, it is called an **input**.

Choosing $u = 0$, the associated zero-input system

$$\mathcal{B}_{u=0} = \{y \in \mathcal{A}^p \mid P \bullet y = 0\}$$

is autonomous. Thus y is called an **output**.

Note that input-output structures are not unique. The transfer matrix does not depend on the underlying representation, but on the chosen input-output structure.

Controllability

The problem of controllability addresses the question "Can a system be forced to go from one trajectory to another without violating the system law?" However, in the

abstract setting of this section, we have to use a more algebraic definition based on the fact that controllability coincides with parametrizability for many important system classes. We call a linear abstract system **controllable** if and only if it admits an **image representation**, that is, there exists $L \in \mathcal{D}^{q\times l}$ such that

$$\mathcal{B} = \{\omega \in \mathcal{A}^q \mid \exists \ell \in \mathcal{A}^l \colon \omega = L \bullet \ell\}.$$

In Section 1.3, 1.4 we will point out for particular signal classes how this definition is linked with the proposed interpretation.
The fundamental principle offers a direct characterization of controllability:

$$\mathcal{B} \text{ is controllable if and only if } \operatorname{im}(\cdot R) = \ker(\cdot L) \text{ for some } L. \tag{1.6}$$

Then one calls R a **left syzygy matrix**. Recall that by assumption, $\mathcal{D}$ is a domain. Then controllability implies that the associated system module is torsion-free. This can be easily checked. Let $0 \neq d \in \mathcal{D}$ and $x \in \mathcal{D}^{1\times q}$ such that $d[x] = 0$, that is, $dx \in \operatorname{im}(\cdot R)$. Then by (1.6) $dxL = 0$ and since $\mathcal{D}$ is a domain, $x \in \ker(\cdot L) = \operatorname{im}(\cdot R)$, that is, $[x] = 0$.

One can show that every finitely generated torsion-free module over a Noetherian domain can be embedded into a finitely generated free module. This permits the following characterization.

Theorem 1.2.9 *The following claims are equivalent:*

1. *$\mathcal{B}$ is controllable.*
2. *$\mathcal{M}$ is torsion-free.*
3. *R is a left syzygy matrix.*

Proof: Referring to the previous observations, it is sufficient to show that the second item yields the third. Let $\mathcal{M}$ be torsion-free. Then there exists an embedding $\iota : \mathcal{M} \to \mathcal{D}^{1\times l}$. Define $\pi \colon \mathcal{D}^{1\times q} \to \mathcal{M}$, $x \mapsto [x]$. Then

$$\mathcal{D}^{1\times g} \xrightarrow{\cdot R} \mathcal{D}^{1\times q} \xrightarrow{\iota\circ\pi} \mathcal{D}^{1\times l}$$

is exact. Defining L as the matrix associated to $\iota \circ \pi$ shows that R is a left syzygy matrix. □

Relying on certain structure properties of $\mathcal{D}$, it is possible to decompose $\mathcal{B}$ into a controllable and autonomous subsystem. In Section 1.4, we point out how to obtain such a decomposition for one-dimensional time-varying systems.

Observability

Suppose $R = [R_1, R_2]$ and $\omega = [\omega_1^T, \omega_2^T]^T$ to be partitioned accordingly, that is,

$$\mathcal{B} = \{[\omega_1^T, \omega_2^T]^T \in \mathcal{A}^{q_1+q_2} \mid R_1 \bullet \omega_1 + R_2 \bullet \omega_2 = 0\}.$$

Then ω_1 is called **observable** from ω_2 if and only if ω_1 is uniquely determined by ω_2 and the fact that $[\omega_1^T, \omega_2^T]^T$ satisfies the system law. That is, ω_1 is observable from ω_2 if and only if $\mathcal{B}_1 := \{\omega_1 \in \mathcal{A}^{q_1} \mid R_1 \bullet \omega_1 = 0\} = \{0\}$. Due Theorem 1.2.3, this is equivalent to $\mathcal{D}^{1\times g_1} R_1 = \mathcal{B}_1^\perp = \mathcal{D}^{1\times q_1}$. Summing up, we obtain that ω_1 is observable from ω_2 if and only if R_1 is left invertible.

1.3 One-dimensional systems over rings

The behavioral approach to system theory has been successfully applied to some areas in communication. In this framework problems like the decoding of Reed-Solomon block codes over fields [BF01, LO08], catastrophicity issues on convolutional codes over fields, and the construction of minimal trellis [Fit95] are handled effectively [Kui01, KP04, KvDHO01, KW97, RSY96]. Moreover, interpolation questions can be tackled with the help of the so-called most powerful unfalsified model, see Section 1.6.

The interest in systems over rings stems mainly from the applications in communication theory. As outlined in [KPPR06], one relevant topic is convolutional codes over rings, which are linear, time-invariant behaviors over the underlying ring. Here the ring structure is better suited for phase modulation than the field structure. Further [HKC$^+$94] stresses the importance of codes over $\mathbb{Z}_4$. The impact of these lies in the connection to certain efficient nonlinear binary codes under the Gray map. Beside, the communications literature offers many results for sequences over finite rings [BHK92, US00, KP08b, KWP05]. Thus there are several aspects that motivate to extend the well established theory of one-dimensional systems over fields to those over the rings $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ for a integer m.

In connection to systems over rings we assume the signal set $\mathcal{A} := \{\omega : \mathbb{N} \to \mathbb{Z}_{p^r}\}$ for a prime number p and we consider exclusively linear and $\mathbf{s}$-invariant systems. Recall that $\mathbf{s}$ denotes the backward shift, acting on $\mathcal{A}$ as $\mathbf{s} \bullet \omega(t) = \omega(t+1)$.

Let $\mathcal{D}$ denote $\mathbb{Z}_{p^r}[\mathbf{s}]$. In [LLO04] it is shown that the $\mathcal{D}$-module $\mathcal{A}$ is an injective cogenerator (in fact [LLO04] even considers the multi-dimensional case). This result relies on the fact that the rings $\mathbb{Z}_m$ are Quasi-Frobenius, that is, Noetherian and self-injective.

Kernel representations

From the algebraic point of view, the main difficulties that arise going from the field case to the ring case is the existence of zero-divisors and the fact that $\mathcal{D}$ is no longer a PID. The injective cogenerator property permits to relate the kernel representations of related linear abstract systems. Recall that due to Theorem 1.2.3, we obtain for two linear abstract systems $\mathcal{B}_1, \mathcal{B}_2$ given via $R_1 \in \mathcal{D}^{g_1 \times q}$ and $R_2 \in \mathcal{D}^{g_2 \times q}$ that

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \quad \Leftrightarrow \quad \exists X \in \mathcal{D}^{g_2 \times g_1} : \; R_2 = X R_1. \tag{1.7}$$

But in contrast to the case of a coefficient field, the equality of two abstract linear systems $\mathcal{B}_1, \mathcal{B}_2$ is not equivalent to the requirement that there exists a unimodular matrix X satisfying the right hand side of equivalence (1.7). This is due to the fact that behaviors over rings (rather than fields) need not have full row rank representations, as will be pointed out below. Let us emphasize this effect by [KPPR06, Example 3.1]. Suppose $p = 3$, $r = 2$ and $\mathcal{B}_1, \mathcal{B}_2$ to be given via the kernel representation matrices

$$R_1 = \begin{bmatrix} \mathbf{s}^2 + \mathbf{s} \\ 3(\mathbf{s} - 1) \end{bmatrix} \quad \text{and} \quad R_2 = \begin{bmatrix} \mathbf{s}^2 + \mathbf{s} \\ 3 \end{bmatrix}.$$

Then one can easily check that $\mathcal{B}_1 = \mathcal{B}_2 = \{(a, b, -b, b, -b, \dots) \mid a, b \in \{0, 3, 6\}\}$. Further

$$\begin{bmatrix} 1 & 0 \\ 6 & \mathbf{s}-1 \end{bmatrix} R_1 = R_2 \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{s}-1 \end{bmatrix} R_2 = R_1.$$

Determinant arguments show that the transformation matrices are not invertible. To prove that these matrices can not be chosen to be unimodular, it would be necessary to introduce some basic facts on matrices over $\mathbb{Z}_m[x]$ which will not be needed later, thus we refer to [KPPR06].

Another effect of the domain or rather PID loss can be observed in the size of representations. Consider a linear abstract system given via a kernel representation matrix $R \in \mathcal{D}^{g\times q}$ where $g > q$. In the field case (or even in case R consists of entries in a PID), we can find due to the Smith form a representation of q rows at most. But here consider for instance the matrix given in [KPPR06, Example 3.1] where $p = 3$ and $r = 2$:

$$\begin{bmatrix} \mathbf{s}-1 \\ 3 \end{bmatrix} \bullet \omega = 0. \tag{1.8}$$

The first equation restricts the solution set to constant sequences, the second equation determines the corresponding values to 0, 3 or 6. This supports the assertion that there can not be found a single equation in $\mathcal{D}$ possessing the same solution set. By use of minimal Gröbner bases arguments, the given representation can not be reduced, see Definition 2.1.12 and Section 2.1. Beside we will see later in Lemma 2.1.25 that there can always be found a representation of rq rows at most.
For a commutative ring, we define the rank of a matrix by determinantal ideals. More precisely, let $J_s(R)$ be the ideal generated by all $s \times s$ subdeterminants of R. Then the **rank** of R is given by

$$\operatorname{rank}(R) := \max\{s \in \mathbb{N} \mid J_s(R) \neq 0\}.$$

The example (1.8) shows that there does not exist a full row rank representation for systems over finite rings in general, which is different to the field case.

Predictable Degree property

We define the row degree of a non-zero polynomial row vector as the maximum of the degrees of its components. Then for univariate polynomial matrices over a field $\mathbb{F}$ the concept of the Predictable Degree property (terminology from [For75]) is defined as follows.

Definition 1.3.1 Let R be a matrix in $\mathbb{F}^{m\times q}[x]$ with full row rank and with row degrees $d_1, \dots, d_m$. Then R is said to have the **Predictable Degree property** if for any nonzero polynomial vector

$$a = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \text{ in } \mathbb{F}^m[x]$$

we have that

$$\text{row degree of } aR = \max_{1\leq i\leq m; a_i\neq 0} (d_i + \deg(a_i)).$$

Thus the row degree of aR can be predicted from the degrees in a and the row degrees of R. Suppose for instance

$$R = \begin{bmatrix} x & x^3 \\ 1 & x^3 \end{bmatrix}.$$

Then $d_1 = d_2 = 3$ and R does not possess the Predictable Degree property, since we obtain for $a = [1 \quad -1]$:

$$\text{row degree of } aR = 1 \neq 3 = \max_{1 \leq i \leq 2} \{d_i + \deg(a_i)\}.$$

For the field case, it is proven in [Wed34, For75] and in [Kai80, Thm6.3-13] that the Predictable Degree property is equivalent to the property that the leading row coefficient matrix of R has full row rank, i.e., that R is row reduced. In Section 2.1.1, we outline how to obtain a row reduced representation with the help of Gröbner bases. The concept of the Predictable Degree property is useful for many areas of system theory, ranging from controller parameterization to minimal realizations of linear systems over fields [Buc01, LXB08, Mor03, Obe90, PR07, WRO98, ZL01].

The adaptation to systems over finite rings is not straightforward at all, however it is worth to investigate, see [FZ97, KP09, KP08b]. The major approach is given in [KPP07]. The idea goes like this: Let ${}_{\mathcal{D}}R$ denote the left module generated by the rows of R. Then the coefficients operating on the rows of R should be restricted to ensure the row reducedness. Suppose for instance $R = 3x + 1 \in \mathbb{Z}_{27}[x]$. The polynomial R has row degree 1 whereas $9R$ has row degree zero which does not coincide with $1 = 1 + \deg(9)$. Therefore just certain linear combination in the $\mathcal{D}$-span of R should be permitted. More precisely, exclusively polynomials with coefficients restricted to $\mathrm{A}_3 := \{0, 1, 2\} \subseteq \mathbb{Z}_{27}$ could occur in the linear combination. But evidently not every element contained in the $\mathcal{D}$-span of R could be represented in this way. Thus we would just consider certain generating systems. In the discussed example, we would consider $[R, 3R]^T$ instead of R.
These results rely on a specific setting: The concept of the so-called p-generator sequences, p-linear combination, p-basis and so on. These notions were introduced in [VSR96]. Chapter 3 gives a brief introduction to that framework and outlines the connection to minimal Gröbner bases. Furthermore we give an even more general answer to the question of the Predictable Degree property based on Gröbner bases. In Section 2.1.1, the so-called Predictable Leading Monomial property will be introduced for the field case and Chapter 3 gives the suitable adoption to the finite ring case.

In the sequel we will pick up some system theoretical questions already introduced in the previous section, and adopt those to our underlying framework for which we have assumed that the underlying operator ring is a domain. All results presented below are taken from [Zer07b] even though in a more special setting.

Autonomy

Recall that in the previous section we have defined autonomy as the absence of free variables. In the case of one-dimensional systems over a field, this definition coincides with the requirement that the system's future is uniquely determined by its past,

see [KPPR06]. Considering systems over finite rings, these two interpretations come apart. This observation is made in [Zer07b] by the investigation of the so-called **weak autonomy** and **strong autonomy**.

Definition 1.3.2 [Zer07b] A linear abstract system $\mathcal{B}$ is called **weakly autonomous** if it does not possess any free variables in the sense of Definition 1.2.5.

The following characterization holds.

Theorem 1.3.3 *[Zer07a, Theorem 4] The linear abstract system $\mathcal{B}$ represented by R is weakly autonomous if and only if there exist an element $d \in \mathcal{D}\backslash\{0\}$ and a matrix $X \in \mathcal{D}^{q\times g}$ such that $XR = d\,\mathrm{I}$.*

Note that due to the assumption that p^r is a prime power, we obtain exactly the result of the previous section. Else the proof of Theorem 1.2.7 could be adapted using the fact that any two non-zero elements of $\mathcal{D}$ have a non-zero common multiple.
In Corollary 1.2.8, we have seen that a system over a domain is autonomous if and only if it is represented by a full column rank matrix. The behavior given by $R = \mathrm{diag}(3\mathbf{s}, 3\mathbf{s}) \in \mathbb{Z}_9[\mathbf{s}]^{2\times 2}$ is obviously weakly autonomous, but R has rank one.

By definition

$$\{\omega : \mathbb{N} \to \mathbb{Z}_9 \mid 3 \bullet \omega = 0\} \tag{1.9}$$

is weakly autonomous. But the system's future is not determined by its past. Before specifying, we need to formulate this sloppy proposition in formal setting first.

Definition 1.3.4 [KPPR06, Definition 15] The linear abstract system $\mathcal{B}$ is called **strongly autonomous** if there exists $T \in \mathbb{N}$ such that for all ω and ω' where $\omega_{|[0,T]} = \omega'_{|[0,T]}$, it follows that $\omega = \omega'$.

Note that since we exclusively consider linear systems, the previous definition can be reformulated like this: A system $\mathcal{B}$ is strongly autonomous if and only if there exists $T \in \mathbb{N}$ such that for all ω satisfying $\omega_{|[0,T]} = 0$, it follows that $\omega = 0$.
Let us return to the behavior given in (1.9). Obviously every sequence such that $\omega(k) \in \{0, 3, 6\}$ for all $k \in \mathbb{N}$ is contained in $\mathcal{B}$. Let us define

$$\omega^{(N)}(k) := \begin{cases} 3 & \text{if } k \leq N \\ 0 & \text{else.} \end{cases}$$

Then for every $T \in \mathbb{N}$ the identity $\omega^{(T)}_{[0,T]} = \omega^{(T+1)}_{[0,T]}$ holds but $\omega^{(T)} \neq \omega^{(T+1)}$. Altogether we have demonstrated that $\mathcal{B}$ is weakly autonomous, but not strongly autonomous.

Strong autonomy can be characterized by the so-called **reduced rank**. According to [Zer07b] the reduced rang of a matrix R denoted by red-rank(R) is given by

$$\text{red-rank}(R) := \max\{s \in \mathbb{N} \mid \mathrm{ann}(J_s(R)) = 0\},$$

where $J_s(R)$ denotes the s-th determinantal ideal. Note that the crucial idea behind the reduced rank is that it generalizes an important property from the field coefficient case to our setting. Let $\mathbb{F}$ be a field and $R \in \mathbb{F}[x]^{g\times q}$. Clearly there exists $0 \neq \omega \in \mathbb{F}[x]^q$ such that $R\omega = 0$ if and only if R is not of full column rank. Now assume $R \in \mathcal{D}^{g\times q}$. Then there exists $0 \neq \omega \in \mathcal{D}^q$ such that $R\omega = 0$ if and only if R has reduced rank smaller than q.

Theorem 1.3.5 *[Zer07b, Theorem 2] Any representation R of the linear abstract system $\mathcal{B}$ has reduced rank q if and only if $\mathcal{B}$ is strongly autonomous.*

Our setting provides the equivalence that R is of reduced rank q if and only if there exists a matrix $X \in \mathcal{D}^{q\times g}$ and a non-zero divisor d such that $XR = d\,\mathrm{I}$, see [Zer07a, Theorem 1]. This offers a concrete relation between weak and strong autonomy.
Let R_p denote the matrix we obtain by considering each entry of R in $\mathbb{Z}_p[\mathbf{s}]$. Then one can show that R has reduced rank q if and only if R_p has full column rank in $\mathbb{Z}_p[\mathbf{s}]$. Therefore the characterization given in the theorem above equals the one given in [KPPR06, Proposition 22].

Input-Output representations

The idea of an input-output representation is based on exactly the same idea as discussed in the previous section. In this sense we would like to point out those components of an element $\omega \in \mathcal{B}$ which are free and those which are not. In other words, we want to partition the system in a weakly autonomous part and a free part. As we have seen before, this can not be characterized by the study of the rank alone. Thus here the so-called **input-dimension** is defined as the cardinality of a maximal subset of free components of ω and suitably, the **output-dimension** as q minus the input-dimension. Let m denote the input-dimension of $\mathcal{B}$ and further let its components be permuted such that there exists a surjection

$$\pi : \mathcal{B} \to \mathcal{A}^m, \qquad \omega \mapsto (\omega_1, \ldots, \omega_m).$$

This yields the exact sequence

$$\mathcal{B} \to \mathcal{A}^m \to 0.$$

Using the Malgrange isomorphism and the isomorphism $\mathcal{A}^m \cong \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times m}, \mathcal{A})$ yields the exact sequence

$$\mathrm{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A}) \to \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times m}, \mathcal{A}) \to 0.$$

Since $\mathcal{A}$ is an injective cogenerator, the sequence

$$\mathcal{M} = \mathcal{D}^{1\times q}/\mathcal{D}^{1\times g}R \leftarrow \mathcal{D}^{1\times m} \leftarrow 0$$

is exact. Adding $q - m$ zeros to $\mathcal{D}^{1\times m}$ and re-arranging the columns of R, the input-dimension can be identified as the largest integer such that

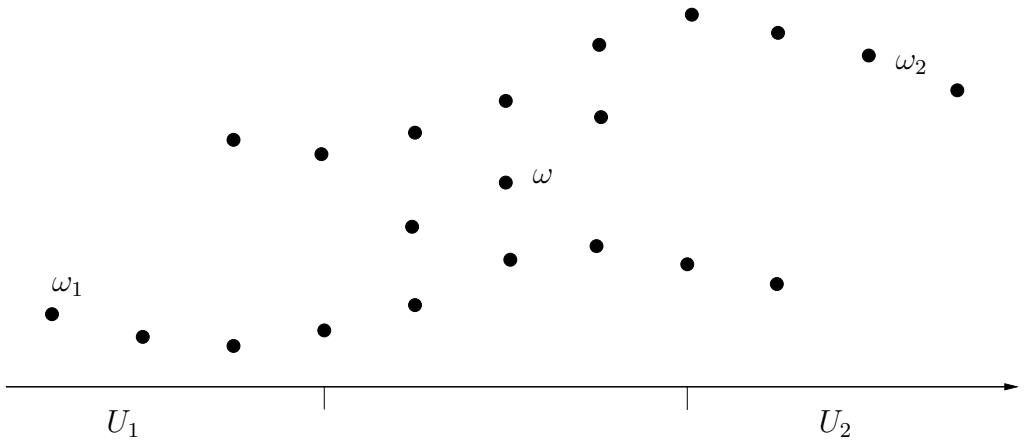
$$\mathcal{D}^{1\times g}R \cap (\mathcal{D}^{1\times m} \times \{0\}) = \{0\}$$

holds.

Theorem 1.3.6 *[Zer07b, Theorem 3] Let m denote the input-dimension of $\mathcal{B}$ and $p = q - m$ the output-dimension. Further suppose the components of $\mathcal{B}$ to be permuted such that π is surjective. Define the input $u := (\omega_1, \ldots, \omega_m)^T$, the output $y := (\omega_{m+1}, \ldots, \omega_q)^T$ and suppose the columns of R to be permuted corresponding, written as $R = [-Q, P]$. Then for all inputs $u \in \mathcal{A}^m$ there exists an output y such that $P \bullet y = Q \bullet u$. Further the linear abstract system given via P is weakly autonomous.*

Controllability

Even though we are not working over a domain anymore, the definition given in the previous section can be adopted. In this sense, we call a behavior **controllable** if it admits an image representation. Referring to [Zer07b] we have the following interpretation for $T = \mathbb{Z}$:
A linear abstract system $\mathcal{B}$ is controllable if and only if for any $\omega', \omega'' \in \mathcal{B}$, and any $U_1, U_2 \subseteq \mathbb{Z}$ sufficiently far apart there exists $\omega \in \mathcal{B}$ such that $\omega_{|U_1} = \omega'_{|U_1}$ and $\omega_{|U_2} = \omega''_{|U_2}$. This can be visualized as follows.



In case of $T = \mathbb{N}$, the previous interpretation requires an adoption, see [WZ99]. Suppose that $dm = 0$ for a non-zero divisor $d \in \mathcal{D}$ and $m \in \mathcal{M}$ implies that $m = 0$. Then $\mathcal{M}$ is called **torsion-free**. This generalizes the classical notion of torsion-freeness as used in Theorem 1.2.7.

Theorem 1.3.7 *[Zer07b, Theorem 4] The following propositions are equivalent:*

1. *$\mathcal{B}$ is controllable.*
2. *$\mathcal{M}$ is torsion-free.*
3. *Any kernel representation R of $\mathcal{B}$ is a left syzygy matrix.*

1.4 One-dimensional time-varying systems

As the name suggests, time-varying systems are systems that are not invariant under time-shifts. Thus systems that are variant under differentiation are time-varying systems as well. Examples are behaviors given via differential equations with polynomial,

rational or meromorphic coefficients. In applications, those systems can occur if a non-linear kinematical model is linearized along a trajectory.
In this section, we focus on the one-dimensional and continuous case with polynomial or rational coefficients. As already suggested in Example 1.2.1, the corresponding operator ring is the so-called one-dimensional Weyl algebra. Let us outline the difference to the constant coefficients case first. Evidently the operator ring that corresponds to ODE's with constant coefficients in the field $K \in \{\mathbb{C}, \mathbb{R}\}$ is the polynomial ring in one variable $K[\partial]$, that operates via differentiation on the signal set $\mathcal{A} \in \{\mathcal{C}^\infty(\mathbb{R}, K), \mathcal{D}'(\mathbb{R}, K)\}$, the space of smooth functions or the space of distributions. Note that this signal set is an injective cogenerator. Defining the action

$$\partial \bullet f = \frac{df}{dt} \quad \text{for all} \quad f \in \mathcal{A},$$

which extends naturally to

$$\partial \bullet f = [\frac{df_1}{dt}, \ldots, \frac{df_q}{dt}]^T \quad \text{for all} \quad f \in \mathcal{A}^q,$$

the signal set $\mathcal{A}^q$ becomes a $K[\partial]$-module. Further, the abstract linear system given via $R \in K[\partial]^{g \times q}$, namely

$$\mathcal{B} = \{\omega \in \mathcal{A}^q \mid R \bullet \omega = 0\},$$

becomes a submodule of $\mathcal{A}^q$. This is easily verified since

$$R \bullet (\partial \bullet \omega) = \partial \bullet (R \bullet \omega) = \partial \bullet 0 = 0$$

holds for all $\omega \in \mathcal{B}$. The algebraic reason for this is the commutative structure of the operator ring. Now let us consider systems given by ODE's with polynomial coefficients, that is, coefficients in $K[t]$. The disparity

$$(t\partial) \bullet \omega = t(\partial \bullet \omega) = t\frac{d\omega}{dt} \neq (\partial t) \bullet \omega = \partial \bullet (t\omega) = \omega + t\frac{d\omega}{dt} = (1 + t\partial) \bullet \omega$$

points out the need of a non-commutative operator ring to describe the model class properly. The suitable candidate is the first Weyl algebra W_1, a certain Ore algebra, for a general definition see Section 2.2. The algebra W_1 is a $\mathbb{C}$-algebra in t and ∂, such that $\partial t = t\partial + 1$, that is, the Leibniz-rule is satisfied. Then $\mathcal{A}$ becomes a W_1-module by the action

$$\partial \bullet \omega = \frac{d\omega}{dt} \quad \text{and} \quad t \bullet \omega = t\omega \qquad \text{for all } \omega \in \mathcal{A}.$$

Note that in this situation $\mathcal{B}$ is not a submodule anymore. Consider for instance a system law given by the equation $t\partial - 1$. Then t is contained in the behavior but 1 is not.
The concept of the polynomial Weyl algebra can be generalized to the so-called rational Weyl algebra B_1. Permitting the function field $K(t)$ instead of $K[t]$, the definition extends analogously. This leads to behaviors given by ODE's with rational coefficients.

The above comparison of time-varying and time-invariant systems outlined exclusively structural differences of the associated operator rings. However, the crucial difference is based on the signals:

- Modeling:
  Let us anticipate the ideas of Section (1.6) and Chapter 5. The aim of that section is to set up a model to describe a set of observed trajectories. In that context a time-varying linear abstract system loses degrees of freedom. More precisely suppose $\mathcal{A}$ to be the space of smooth functions and consider a system law given via $(t\partial - 1) \bullet \omega = 0$. The function $\omega : t \mapsto t$ is obviously a solution, whereas its derivative does not satisfy the system law. Modelling the same trajectory via differential equations with constant coefficients yields the system law $\partial^2 \bullet \omega = 0$ and thus the solution space $\omega : t \mapsto bt + a$ for $a, b \in \mathbb{R}$. Note that this example gives a slight insight in the effect of the missing $\mathcal{D}$-module structure of $\mathcal{B}$ in the time-varying framework.

- Singularities:
  The solution space of autonomous linear ordinary differential equations with constant coefficients is spanned by linear combinations of polynomial-exponential functions. Thus there are no problems with singularities in the signal set at all. However, these difficulties arise for the time-varying case directly. Zeros of the leading coefficient in the describing equations may produce poles in the solution space. The system law $(t\partial + 1) \bullet \omega = 0$ possesses the rational solution $\frac{1}{t}$. Thus a signal space defined on an interval not containing 0 would lead to a non-trivial behavior whereas the signal space in $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ would be $\{0\}$, that is, unlike the time-invariant situation, local solutions may not be extendable.

According to Section 1.2, we are interested in a $\mathcal{D}$-module $\mathcal{A}$ which is an injective cogenerator. It should be stressed that due to analytic arguments the generalization from the constant coefficient case is not obvious at all, see [Woo02]. Following [Zer06a] the space of all functions that are smooth except for a finite number of points is an injective cogenerator for the rational Weyl algebra. Anyway, since we know injective cogenerators for both cases, the results of Section 1.2 provide a system theoretical framework for this time-varying context.

The rational case can be considered as the time-varying counterpart to the well studied time-invariant linear case. There are several approaches to tackle this framework [Bou05, CBSW02, DR95, Fli90, FO98, IM05, PQ98, Zer06a, Zer07c]. The following subsection gives a brief insight into a few of the issues, and stresses mainly the results of Section 1.2 that can be sharpened with the help of the Jacobson form to give a convincing motivation for Chapter 4.

Ordinary differential equations with rational coefficients

For the rest of this section, let $\mathcal{D}$ be the first rational Weyl algebra B_1 and $\mathcal{A}$ the space of all functions that are smooth except for a finite number of points.
For every element in B_1 the highest exponent in ∂ is called the degree (see Section 2.2). Following [WW89, Coh71], the first rational Weyl algebra is simple and furthermore a left and right Euclidean domain.
Therefore every abstract linear system can be decoupled: There exist invertible ma-

trices $U \in \mathrm{B}_1^{g\times g}$ and $V \in \mathrm{B}_1^{q\times q}$ such that

$$URV = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

where $D = \mathrm{diag}(d_1, \ldots, d_r)$ such that $d_i \neq 0$ for all $1 \leq i \leq r$. In Chapter 4, we focus on the computational aspect. We remark that

$$\mathcal{B} = \mathcal{B}_1 \oplus \ \ldots \ \oplus \mathcal{B}_r \oplus \mathcal{A}^{q-\mathrm{rank}(R)}, \qquad \text{where} \quad \mathcal{B}_i \cong \{y \in \mathcal{A} \mid d_i \bullet y = 0\}.$$

One should notice that this representation is not unique. But two invariants come up with the diagonal form.

1. The number of the free subsystems $\mathcal{B}_i = \mathcal{A}$ is uniquely determined by the rank of R.

2. The order of underlying ODE system is uniquely given by $\sum_{i=1}^{r} \deg(d_i)$. For details see Remark 4.2.10.

The strong properties of the operator ring B_1 leads to the following structure theorem.

Theorem 1.4.1 *[Coh71, Ch. 3] [Jac43, Ch. 8.1] Let $R \in \mathrm{B}_1^{q\times q}$. Then there exist invertible matrices $U \in \mathrm{B}_1^{q\times g}$ and $V \in \mathrm{B}_1^{q\times q}$ such that*

$$URV = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} =: J,$$

where $D = \mathrm{diag}(1, \ldots, 1, m_R) \in \mathcal{D}^{r\times r}$ such that $r = \mathrm{rank}(R)$. Then J is called the ***Jacobson form*** *of R.*

Note that the Jacobson form is the non-commutative analogon of the Smith form. One obvious benefit of the Jacobson form is the existence of a full rank representation for every linear abstract system $\mathcal{B}$. Suppose $\mathcal{B}$ to be given by R and let U, V be suitable such that these matrices yields to the Jacobson form $\mathrm{diag}(1, \ldots, 1, m_R, 0 \ldots, 0)$. For the purpose of a **full row rank representation** partition

$$W := V^{-1} = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}. \tag{1.10}$$

Then

$$\{\omega \in \mathcal{A}^q \mid \ \mathrm{diag}(1, \ldots, 1, m_R) W_1 \bullet \omega = 0\} = \mathcal{B}$$

provides the desired representation.
Some results of Section (1.2) can be sharpened due to the Jacobson form as we will see below.

Autonomy

Since $\mathcal{B}$ is autonomous if and only if the representation has full column rank this is equivalent to the conclusion that there exists a representation of $\mathcal{B}$ with a square full rank matrix.
The chosen signal set $\mathcal{A}$ provides an additional interpretation. Referring to [Zer06a, Theorem 5] $\mathcal{B}$ is autonomous if and only if there exists a finite set $E \subseteq \mathbb{R}$ such that for all open intervals $I \subseteq \mathbb{R}\backslash E$ and for all $\omega \in \mathcal{B}$ that are smooth on I we have

$$\omega_{|J} = 0 \quad \Rightarrow \quad \omega_{|I} = 0,$$

for all open intervals $J \subseteq I$.

Controllability

Let $U \in \mathrm{B}_1^{g\times g}$, $V \in \mathrm{B}_1^{q\times q}$ and W be chosen as before. Then

$$\begin{aligned}\mathcal{B} &= \{\omega \in \mathcal{A}^q \mid \operatorname{diag}(1,\dots,1,m_R,0,\dots,0)W \bullet \omega = 0\}\\ &\cong \{\omega \in \mathcal{A}^q \mid \operatorname{diag}(1,\dots,1,m_R,0,\dots,0) \bullet \omega = 0\}\\ &\cong \{\omega \in \mathcal{A}^{\mathrm{rank}(R)} \mid \operatorname{diag}(1,\dots,1,m_R) \bullet \omega = 0\} \oplus \mathcal{A}^{q-\mathrm{rank}(R)}\\ &\cong \{\omega \in \mathcal{A} \mid m_R \bullet \omega = 0\} \oplus \mathcal{A}^{q-\mathrm{rank}(R)},\end{aligned}$$

that is, we obtain a minimal representation, namely a representation given by the single equation m_R. In terms of the system module, this amounts to the group isomorphism

$$\mathcal{M} \cong \mathrm{B}_1 \,/\, \mathrm{B}_1\, m_R \oplus \mathrm{B}_1^{1\times(q-\mathrm{rank}(R))}.$$

Then every linear abstract system decomposes into

$$\mathcal{B} = \mathcal{B}_a \oplus \mathcal{B}_c, \tag{1.11}$$

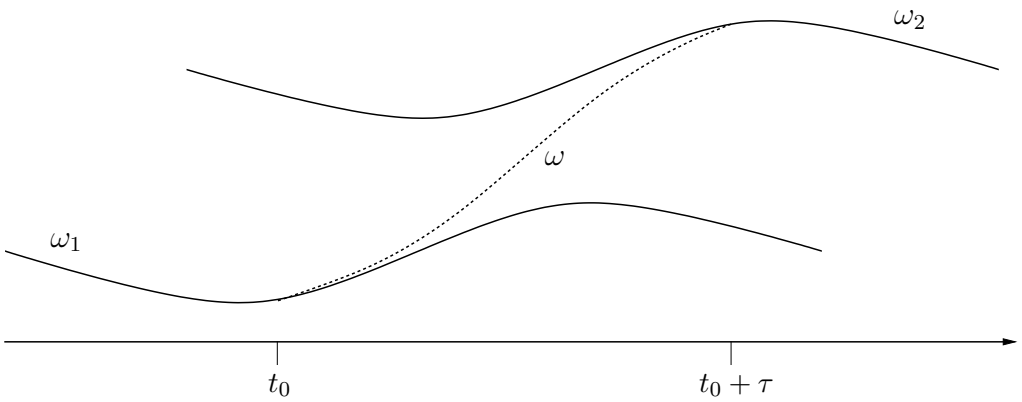
the decoupled controllable subsystem $\mathcal{B}_c \cong \mathcal{A}^{q-\mathrm{rank}(R)}$ and the autonomous subsystem $\mathcal{B}_a \cong \{\omega \in \mathcal{A} \mid m_R \bullet \omega = 0\}$.

Remark 1.4.2 The abstract linear system $\mathcal{B}$ is controllable if and only if $\mathcal{M}$ is free.

Further referring to [Zer06a, Theorem 7], we obtain that $\mathcal{B}$ is controllable if and only if for all $\omega_1, \omega_2 \in \mathcal{B}$ and almost all $t_0 \in \mathbb{R}$, there exists a connecting trajectory $\omega \in \mathcal{B}$, an open interval $t_0 \in I \subseteq \mathbb{R}$ such that $\omega_1, \omega_2, \omega$ are smooth on I, and $\tau > 0$ with $t_0 + \tau \in I$ such that

$$\omega(t) = \begin{cases} \omega_1(t) & \text{if } t < t_0 \\ \omega_2(t) & \text{if } t > t_0 + \tau, \end{cases}$$

for all $t \in I$. This concept can be visualized as follows:



Specifying (1.11) we claim for W defined as in (1.10) and $V=[V_1,V_2]$ that

$$\mathcal{B}_c=\{\omega\in\mathcal{A}^q|\exists\ell\in\mathcal{A}^{q-\mathrm{rank}(R)}:\ \omega=V_2\bullet\ell\}$$

and

$$\mathcal{B}_a=\{\omega\in\mathcal{A}^q\ \mid \mathrm{diag}(1,\dots,1,m_R)W_1\bullet\omega=0 \text{ and } W_2\bullet\omega=0\}.$$

Note that $VW=\mathrm{I}_q=WV$ lead to

$$V_1W_1+V_2W_2=\mathrm{I}_q\quad\text{ and }\quad W_1V_2=0. \tag{1.12}$$

Let us show $\mathcal{B}_a\oplus\mathcal{B}_c\subseteq\mathcal{B}$ of equality (1.11) first.
Due to the first system law condition of $\mathcal{B}_a$, it is easy to see that $\mathcal{B}_a\subseteq\mathcal{B}$, and it is autonomous.
Evidently $\mathcal{B}_c$ is controllable since it is given via an image representation. To see that it is contained in $\mathcal{B}$ we first show that $\mathrm{im}(\cdot W_1)=\ker(\cdot V_2)$:
For every $y\in\mathrm{im}(\cdot W_1)$ there exists an element x such that $xW_1=y$ thus $yV_2=xW_1V_2=0$ due to (1.12) and we can conclude that $y\in\ker(\cdot V_2)$.
Conversely suppose $y\in\ker(\cdot V_2)$. Then due to (1.12) $y=yV_1W_1+yV_2W_2$ and thus $y=yV_1W_1$, which concludes the claim. Using that $\mathrm{im}(\cdot W_1)=\ker(\cdot V_2)$ and the fundamental principle, Theorem 1.2.2, it follows that

$$\begin{aligned}\mathcal{B}_c&=\{\omega\in\mathcal{A}^q\mid\exists\ell\in\mathcal{A}^{q-\mathrm{rank}(R)}:\ \omega=V_2\bullet\ell\}\\&=\{\omega\in\mathcal{A}^q\mid W_1\bullet\omega=0\}.\end{aligned}$$

From this is follows directly that $\mathcal{B}_c\subseteq\mathcal{B}$ and furthermore, since W is invertible, that $\mathcal{B}_a\cap\mathcal{B}_c=\{0\}$.
Still left to show that $\mathcal{B}\subseteq\mathcal{B}_a\oplus\mathcal{B}_c$. Every $\omega\in\mathcal{A}^q$ can be decomposed into $\omega=V_1W_1\bullet\omega+V_2W_2\bullet\omega$ due to (1.12). It is easily verified that $V_1W_1\bullet\omega\in\mathcal{B}_a$ and $V_2W_2\bullet\omega\in\mathcal{B}_c$.

Now let us illustrate the results via the following the dynamic system

$$\begin{aligned}\dot{x_1}(t)+tx_1(t)-x_2(t)+u(t)&=0\\t\dot{x_2}(t)+t^2u(t)&=0.\end{aligned}$$

Here u corresponds to the input and x_1 and x_2 describe the dynamics. This yields the behavior

$$\mathcal{B}=\{[x_1,x_2,u]^T\in\mathcal{A}^3\mid \underbrace{\begin{bmatrix}\partial+t&-1&1\\0&t\partial&t^2\end{bmatrix}}_{=:R}\bullet\begin{bmatrix}x_1\\x_2\\u\end{bmatrix}=0\}.$$

Then R can be transformed to its Jacobson form J, i.e.

$$\underbrace{\begin{bmatrix} -1 & 0 \\ t^2 & -1 \end{bmatrix}}_{=:U} R \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & t \\ -1 & 1 & -\partial \end{bmatrix}}_{=:V} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -t\partial - t^2 & 0 \end{bmatrix} = J.$$

Since $-t\partial - t^2$ does not correspond to a unit in B_1, it defines a non-trivial autonomous subsystem. Hence $\mathcal{B}$ is a non-controllable dynamical system. Since

$$V^{-1} = \begin{bmatrix} -t-\partial & 1 & -1 \\ -t & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

we obtain the decomposition

$$\mathcal{B}_a = \{[x_1, x_2, u]^T \in \mathcal{A}^3 \mid \begin{bmatrix} -t-\partial & 1 & -1 \\ t^3 + t^2\partial & -t^2 - t\partial & 0 \end{bmatrix} \bullet \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = 0 \text{ and } x_1 = 0\}$$

$$\mathcal{B}_c = \{[x_1, x_2, u]^T \in \mathcal{A}^3 \mid \exists \ell \in \mathcal{A} : \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = \begin{bmatrix} 1 \\ t \\ -\partial \end{bmatrix} \bullet \ell\}.$$

In this particular example, which has state space form, there is another convincing argument to see that the system is not controllable. The equivalence $\omega \in \mathcal{B}$ if and only if $JV^{-1} \bullet \omega = 0$ implies that

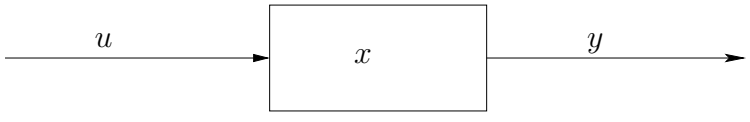
$$(-t\partial - t^2) \bullet (-tx_1 + x_2) = 0.$$

But this condition is evidently completely decoupled from the input u.

To summerize, the Jacobson form reduces every homogeneous system of equations given by ordinary differential equations with rational coefficients to a single equation. However, one has to admit that this equation can become quite "ugly". Apart from general computational aspects, this difficulty will be specified and studied in Chapter 4. Furthermore note that instead of calculating a Jacobson form, it may be sufficient to focus on a decoupled system, which corresponds to a diagonal form of the underlying matrix.

State space representation

The state space representation was not discussed yet, because there is no adequate interpretation in the general framework of Section 1.2. But the concept is well established for the one-dimensional linear time-invariant case. The idea is to introduce the latent variable x, the so-called state, that can be interpreted as an additional information to associate an output y to an input u uniquely.



This idea is essential for Control Theory, because the state comprises all the information about the system's past to determine its future with respect to a known input. Thus it can be used to appoint a certain feedback. In [Zer07c] state space representations for time-varying systems are studied and the work provides the result that any behavior admits an observable state representation. Beyond that it is shown that the interpretation given above extends to time-varying systems.

1.5 Multi-dimensional time-varying systems

Multi-dimensional systems

The motivation to consider multi-dimensional systems is very convincing since many dynamical systems of interest depend on changing of time and space. There are several studies about multi-dimensional time-invariant systems like for instance [RW01, Obe90, PQ99, WRO00, Zer00]. In the most popular cases of partial differential equations with constant coefficients in $\mathbb{C}$ and partial difference equations with constant coefficients in $\mathbb{C}$, the corresponding injective cogenerators are the space of smooth functions $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ and the function space $\mathbb{C}^{\mathbb{N}^n}$. Let $\mathcal{D} \in \{\mathbb{C}[\mathbf{s}_1, \ldots, \mathbf{s}_n], \mathbb{C}[\partial_1, \ldots, \partial_n]\}$ and $\mathcal{A}$ denote $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ or $\mathbb{C}^{\mathbb{N}^n}$ suitably.
From the algebraic point of view, the main difference to the one-dimensional setting is the loss of the PID property and thus no tool like the Smith form is available. Properties which coincide in the one-dimensional case for the operator ring $\mathbb{C}[\mathbf{s}]$ or $\mathbb{C}[\partial]$ need to be refined for systems over $\mathcal{D}$.

For $n = 1$, the behavior $\mathcal{B} \subseteq \mathcal{A}^q$ is autonomous if and only if it is a finitely generated $\mathbb{C}$-vector space. In the continuous case, this is equivalent to $\mathcal{B}$ being an **over-determined** system, which is defined here as foloows: Each smooth function which satisfies the system law locally on a neighborhood of infinity can be extended uniquely to a signal contained in $\mathcal{B}$. For $n \in \mathbb{N}$, the three described equivalent conclusions come apart (for the discrete case consider just the first and the third conclusion):

$$\mathcal{B} \text{ is a finitely generated } \mathbb{C}\text{-vector space.}$$

$$\Downarrow \qquad (\Uparrow \quad \text{if } n = 1)$$

$$\mathcal{B} \text{ is over-determined.}$$

$$\Downarrow \qquad (\Uparrow \quad \text{if } n = 1)$$

$$\mathcal{B} \text{ is autonomous, i.e., it has no free variables.}$$

The so-called autonomy degree gives an insight: Let $\mathcal{M}$ denote the system module of $\mathcal{B}$. Then the dimension of $\mathcal{B}$ is given by the Krull dimension of $\mathcal{D}/\operatorname{ann}(\mathcal{M})$ and $\mathcal{B}$ is said to be of **autonomy degree at least** r if $\dim(\mathcal{B}) < n - r$. One can show that $\dim(\mathcal{B}) = \dim(J_q(R))$, where $J_q(R)$ denotes the q-th determinantal ideal of R.
Then $\mathcal{B}$ is autonomous (see Section 1.2) if $r = 0$. Autonomy degree at least one yields over-determined systems. Furthermore one can show that $\mathcal{B}$ is a finitely generated $\mathbb{C}$-vector space if and only if it is of autonomy degree at least $r = n - 1$.

The following examples, see [Zer06c], show how the above conclusion chain comes apart for $n > 1$:

- Let $n = 2$, $\mathcal{D} = \mathbb{C}[\partial_1, \partial_2]$, $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ and $R = \begin{bmatrix} \partial_1 & -\partial_2 \\ \partial_2 & \partial_1 \end{bmatrix}$.
  Then $J_2(R) = \langle \partial_1^2 + \partial_2^2 \rangle$ and thus the resulting system $\mathcal{B}$ is of dimension one. That is, $\mathcal{B}$ is autonomous, but not over-determined.
- Let $n = 4$, $\mathcal{D} = \mathbb{C}[\partial_1, \partial_2, \partial_3, \partial_4]$, $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}^4, \mathbb{R})$ and
  $$R = \begin{bmatrix} \partial_1 & -\partial_2 \\ \partial_2 & \partial_1 \\ \partial_3 & -\partial_4 \\ \partial_4 & \partial_3 \end{bmatrix}.$$
  Then $J_2(R) = \langle \partial_1^2 + \partial_2^2, \partial_3^2 + \partial_4^2, \partial_1\partial_4 - \partial_2\partial_3, \partial_1\partial_3 + \partial_2\partial_4 \rangle$ and thus the resulting system $\mathcal{B}$ is of dimension two. That is, $\mathcal{B}$ is over-determined, but not a finitely generated $\mathbb{C}$-vector space.

A similar situation can be observed for controllability. Section 1.3 and 1.4 already introduced the interpretation of controllability from the signal perspective, namely for any two trajectories there exists a connecting one contained in the behavior. The dual conclusion from the ring perspective is that the system module is torsion-free. A stronger form of controllability is the so-called **complementability**. A behavior $\mathcal{B}$ is complementable if there exists a behavior $\mathcal{B}_1 \subseteq \mathcal{A}^q$ such that $\mathcal{B} \oplus \mathcal{B}_1 = \mathcal{A}^q$. One can show the following equivalence:

$$\begin{aligned} & \mathcal{B} \text{ is complementable} \\ \Leftrightarrow\ & \mathcal{M} \text{ is projective} \\ \Leftrightarrow\ & \forall\, \mathcal{B}' \subseteq \mathcal{B}\ \ \exists \mathcal{K}\colon\ \mathcal{B} \cap \mathcal{K} = \mathcal{B}' \text{ and } \mathcal{B} + \mathcal{K} = \mathcal{A}^q. \end{aligned}$$

Since $\mathcal{M}$ is finitely generated, the second assertion yields that $\mathcal{M}$ is free due to the Quillen-Suslin theorem. One can show that the system module $\mathcal{M}$ is free if and only if $\mathcal{B}$ possesses an observable image representation, that is, there exists a kernel representation matrix which is right invertible. Controllability and complementability coincidence for $n = 1$, since over a principal ideal domain, a finitely generated module is free if and only if it is torsion-free. Thus we obtain:

$$\begin{array}{c} \mathcal{B} \text{ is complementable, i.e., it has an observable image representation.} \\ \Downarrow \qquad (\Uparrow \ \text{ if } n = 1) \\ \mathcal{B} \text{ is controllable, i.e., it has an image representation.} \end{array}$$

Computational tests for autonomy or controllability can be done by computing the extension modules. The SINGULAR library `control.lib` provides the proper functions.

Delay-differential systems

A system class that is often discussed as a particular case of multi-dimensional systems is given via the so-called **delay-differential** equations, that is, equations involving the shift operator and the differential operator. These equations can be used to

model time-lags. One might think for instance of systems with reaction delay. Let $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ and $\mathcal{D} = \mathbb{R}[\partial, \mathbf{s}]$. Working with this system class leads to fundamental problems, because objects which are independent in the operator ring can cause coupled results in the signal space: The elements ∂ and $\mathbf{s}-1$ are algebraically independent, but

$$\partial \bullet \omega = 0 \quad \Rightarrow \quad (\mathbf{s}-1) \bullet \omega = 0$$

for $\omega \in \mathcal{A}$. This is due to the fact that $\mathcal{A}$ is not an injective $\mathcal{D}$-module. To see this, consider the matrices

$$P = \begin{bmatrix} \mathbf{s}-1 \\ -\partial \end{bmatrix} \quad \text{and} \quad Q = [\partial,\ \mathbf{s}-1]\,.$$

One can easily check that

$$\mathcal{D} \xrightarrow{\cdot Q} \mathcal{D}^{1\times 2} \xrightarrow{\cdot P} \mathcal{D}$$

is exact, but

$$\mathcal{A} \xleftarrow{Q\bullet} \mathcal{A}^2 \xleftarrow{P\bullet} \mathcal{A}$$

is not exact since $\ker(Q\bullet) \ni [0,\ 1]^T \notin \operatorname{im}(P\bullet)$. In [GL00] this problem is fixed by the investigation of the operator ring

$$\{q = \frac{p}{\phi} \in \mathbb{R}(\partial)[\sigma] \mid q^* \text{ is an entire function}\},$$

where $q^*(s)$ denotes the evaluation $\frac{p(s,\mathrm{e}^{-s})}{\phi(s)}$ for all complex non-singularities s of ϕ.

In Chapter 5, the so-called Ore algebras are used to tackle many relevant operator rings in the time-varying framework. From the algebraic point of view, the resulting operator ring is non-commutative due to the time-variance.
It should be stressed that the system class discussed in Section 1.4 is the only case for which a concrete injective cogenerator is known in the time-varying setting. However, due to [Rot79] each module possesses an injective cogenerator:
One can show that each left module can be embedded into an injective left module. And if $\mathcal{A}$ is an injective $\mathcal{D}$-module for a Noetherian ring $\mathcal{D}$, then it is a cogenerator if and only if

$$\operatorname{Hom}_{\mathcal{D}}(\mathcal{M}, \mathcal{A}) = 0 \quad \Rightarrow \quad \mathcal{M} = 0$$

for every finitely generated $\mathcal{D}$-module $\mathcal{M} = \mathcal{D}^{1\times q}/\mathcal{D}^{1\times g}R$. Thus due to the Malgrange isomorphism, $\mathcal{A}$ is a cogenerator if and only if

$$\mathcal{B} = \{\omega \in \mathcal{A}^q \mid R \bullet \omega = 0\} = 0 \quad \Rightarrow \quad \mathcal{D}^{1\times q}/\mathcal{D}^{1\times g}R = 0.$$

If the system module is non-trivial, that is, if R is not left invertible, then it is sufficient to choose a signal space such that $R \bullet \omega = 0$ for at least one non-zero signal ω and to embed this signal space into an injective $\mathcal{D}$-module.

Thus in this setting, one assumes the signal space to be an injective cogenerator even if no concrete injective cogenerator is known. Then basic properties like (strong) autonomy and (strong) controllability are defined exclusively from the module perspective, see for instance [PQ99, Woo00]. As already outlined in Section 1.2, controllability corresponds to a torsion-free system module and autonomy to a torsion system module. The resulting systems can be computationally analyzed with the help of homological algebra elegantly, see for instance [CQR05].

1.6 Most powerful unfalsified model

Linear exact modeling is a problem of system identification. It was formulated for one-dimensional behaviors in [AW93], see also [KP02, KP04]. Starting with an observed set of polynomial-exponential signals, the aim is to find a linear differentiation-invariant model for these. Evidently the whole signal set is a behavior that is not falsified by observation. But such a model has no significance. Making the behavior larger than necessary, the accuracy of the explanation decreases. So in addition to the condition that the desired model should be unfalsified, we are searching for the most powerful one. This means that the model does not admit more data than necessary. A model satisfying all conditions is shortly called MPUM (most powerful unfalsified model). In [Zer05], the modeling was extended to multidimensional behaviors [CQR07, PQ99], and in [Zer08] to the discrete framework, that is, instead of the requirement that the model should contain all derivatives of the signals, it is required that all shifts of the signals are contained. Chapter 5 generalizes the setting to time-varying systems, which leads to interesting algebraic questions as well.

This section gives an overview of the MPUM approaches for time-invariant systems. For the field $\mathbb{C}$ and $T \in \{\mathbb{N}, \mathbb{R}\}$ suppose to observe a set of trajectories $\Omega = \{\omega_1, \ldots, \omega_N\}$ of the form

$$\omega_l : T^n \to \mathbb{C}^q, \quad t \mapsto (p_l \exp_{\lambda_l})(t) = \begin{cases} p_l(t) \exp(\lambda_{l1} t_1 + \cdots + \lambda_{ln} t_n) & \text{if } T = \mathbb{R} \\ p_l(t) \lambda_{l1}^{t_1} \cdots \lambda_{ln}^{t_n} & \text{if } T = \mathbb{N}, \end{cases}$$

with $p_l \in \mathbb{C}[t_1, \cdots, t_n]^q$ and $\lambda_l \in \mathbb{C}^n$. The goal is to find $\mathcal{B}_\Omega$ satisfying:

1. $\mathcal{B}_\Omega$ is unfalsified by Ω, i.e. $\Omega \subseteq \mathcal{B}_\Omega$.

2. $\mathcal{B}_\Omega$ is most powerful, i.e. for every behavior $\mathcal{B}$ with $\Omega \subseteq \mathcal{B}$, it follows that $\mathcal{B}_\Omega \subseteq \mathcal{B}$.

Let $\mathcal{D}$ be $\mathbb{C}[\partial_1, \ldots, \partial_n]$, where ∂_i acts as the i-th derivative if $T = \mathbb{R}$, and $\mathbb{C}[\mathbf{s}_1, \ldots, \mathbf{s}_n]$, where $\mathbf{s}_i$ acts as the i-th shift, else. If $\mathcal{B}_\Omega$ is invariant under the action of $\mathcal{D}$, that is, if we have for all $o \in \mathcal{D}$

$$\omega \in \mathcal{B}_\Omega \quad \Rightarrow \quad o \bullet \omega \in \mathcal{B}_\Omega,$$

it is called (time-invariant) **most powerful unfalsified model**, short MPUM.

Continuous case

Assume $\mathcal{D} = \mathbb{C}[\partial_1, \cdots, \partial_n]$ and let $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ denote the space of complex-valued smooth functions defined on $\mathbb{R}^n$. Each polynomial vector $p \in \mathbb{C}[t_1, \cdots, t_n]^q$ can be written as

$$p(t) = \sum_{\nu \in \mathbb{N}^n,\, |\nu| < d} \frac{1}{\nu!} p_\nu t^\nu, \quad p_\nu \in \mathbb{C}^q, \tag{1.13}$$

where $\nu! = \nu_1! \ldots \nu_n!$ and d is chosen as small as possible. Since p is a polynomial, it has only finitely many nonzero derivatives. It is easy to see that

$$\partial^\mu \bullet p = 0 \quad \text{for all} \quad \mu \in \mathbb{N}^n \text{ with } |\mu| = d.$$

Consider $P := \mathbb{C}[t_1, \ldots, t_n] / \langle t_1, \ldots, t_n \rangle^d$ as a vector space over $\mathbb{C}$, generated by the elements $[t^\nu]$, where $|\nu| < d$. Define $\delta := |\{\nu \in \mathbb{N}^n \mid |\nu| < d\}|$, the number of basis elements in P. Combinatorial arguments yield

$$\delta = \binom{n+d-1}{n}.$$

The multiplication by t_i in P defines a $\mathbb{C}$-linear transformation in $P \cong \mathbb{C}^\delta$. So after fixing a basis of P, one can compute the corresponding matrices. For this purpose we enumerate the elements of $\{\nu \in \mathbb{N}^n \mid |\nu| < d\} =: \{\nu_1, \ldots, \nu_\delta\}$ with respect to a monomial ordering. Then $[t^{\nu_k}] \in P$ can be identified with the k-th element. Let the matrix A_i represent the multiplication with t_i, that is,

$$(A_i)_{kl} = \begin{cases} 1, & \text{if } [t^{\nu_k}] = t_i \cdot [t^{\nu_l}] \\ 0, & \text{else.} \end{cases}$$

Since the multiplication in P is commutative and every element is nilpotent, the matrices A_i commute and are nilpotent as well. Further let $C = (p_{\nu_1}, \ldots, p_{\nu_\delta})$ be defined with respect to representation (1.13).

Now let $p_1, \ldots, p_N \in \mathbb{C}[t_1, \ldots, t_n]^q$. Using the introduced notation, let $A_{l1}, \ldots, A_{ln}$ be associated to p_l. Then we define

$$\mathbf{A}_i = \operatorname{diag}(A_{1i}, \ldots, A_{Ni})$$

and

$$\mathbf{C} = \operatorname{diag}(C_1, \ldots, C_N).$$

Theorem 1.6.1 *[Zer05] Using the above notation, the* MPUM *of* $\Omega = \{\omega_1, \ldots, \omega_N\}$ *with* $\omega_l = p_l \exp_{\lambda^{(l)}}$ *is given by*

$$\mathcal{B}_\Omega = \left\{\omega \in \mathcal{A}^q \mid \exists\, f \in \mathcal{A}^\delta : \partial_i \bullet f = \Lambda_i f + \mathbf{A}_i f \text{ for all } 1 \le i \le n \text{ and } \omega = \mathbf{C} f\right\}$$

with $\delta := \delta_1 + \cdots + \delta_N$ *and* $\Lambda_i := \operatorname{diag}\left(\lambda_i^{(1)} \mathrm{I}_{\delta_1}, \ldots, \lambda_i^{(N)} \mathrm{I}_{\delta_N}\right)$, *where* δ_l *and* $\lambda^{(l)}$ *belong to the* MPUM

$$\mathcal{B}_{\omega_l} = \left\{\omega \in \mathcal{A}^q \mid \exists\, f_l \in \mathcal{A}^{\delta_l} : \partial_i \bullet f_l = \lambda_i^{(l)} f_l + A_{li} f_l \text{ for all } 1 \le i \le n \text{ and } \omega = C_l f_l\right\}$$

of $p_l \exp_{\lambda^{(l)}}$. *The behavior* $\mathcal{B}_{\omega_l}$ *is autonomous for each* $1 \le l \le N$, *that is,* $\mathcal{B}_\Omega$ *is autonomous as well. Moreover,* $\mathcal{B}_\Omega$ *is a finite-dimensional* $\mathbb{C}$*-vector space.*

The behavior constructed in Theorem 1.6.1 is minimal with respect to the number of solutions, but it is not necessarily minimal with respect to the size of its so-called realization. Using the notation of Theorem 1.6.1, we call $(\Lambda_1 + \mathbf{A}_1, \ldots, \Lambda_n + \mathbf{A}_n, \mathbf{C})$ a **realization** of $\mathcal{B}_\Omega$ of size δ. Further a realization is called **minimal** if there exists no realization of strictly smaller size. In [Zer05] is it shown that a realization $(\mathbf{A}_1, \ldots, \mathbf{A}_n, \mathbf{C})$ of $\mathcal{B}_\Omega$ is minimal if and only if it is observable, that is,

$$\bigcap_{\mu \in \mathbb{N}^n} \ker(\mathbf{C}(\Lambda_1 + \mathbf{A}_1)^{\mu_1} \cdots (\Lambda_n + \mathbf{A}_n)^{\mu_n}) = \{0\}.$$

An alternative and recursive way to compute the MPUM of continuous polynomial exponential trajectories is given in [Zer06b]. A recursive computation can obviously be profitable. One could imagine a situation in which a new observation becomes relevant. Suppose $R \in \mathbb{C}[\partial_1, \ldots, \partial_n]^{g\times q}$ to be a kernel representation of the MPUM of $\{\omega_1, \ldots, \omega_N\}$. Note that due to the fundamental principle, a kernel representation can easily be computed from the one given in Theorem 1.6.1. Let ω_{N+1} be an additional observed trajectory. We define the so-called **error signal** as $e = R(\partial_1, \ldots, \partial_n) \bullet \omega_{N+1}$. Further let $\Gamma \in \mathbb{C}[\partial_1, \ldots, \partial_n]^{h\times g}$ be a kernel representation of the MPUM of e. Then $R_{new} := \Gamma R$ represents the MPUM of $\omega_1, \ldots, \omega_{N+1}$.

Discrete case

Let $\mathcal{D} := \mathbb{C}[\mathbf{s}_1, \ldots, \mathbf{s}_n]$ and $\mathcal{A} = \mathbb{C}^{\mathbb{N}^n}$ (instead of $\mathbb{C}$ one could choose an arbitrary field of characteristic zero and obtain the same results). The discrete case can be treated similar to the continuous one. In the continuous case, we have an upper bound for the number of nonzero derivates of a polynomial trajectory. So by constructing a matrix representation of the corresponding behavior, it is not difficult to incorporate the condition that every derivative of the trajectory is contained in the MPUM. In the previous section, this was realized by the matrices A_i.
In the discrete case, the situation differs since shifting a polynomial trajectory does not lead to zero. But in the sequel we show that each shift of a polynomial trajectory can be expressed by taking $\mathbb{C}$-linear combinations of finitely many polynomials, determined by the trajectory. Therefore a special representation of the polynomial trajectories and a little modification of the shift operator have to be chosen.

Define $\Delta_i := \mathbf{s}_i - 1$ for $1 \le i \le n$ and consider the $\mathbb{C}$-algebra isomorphism

$$\mathbb{C}[\mathbf{s}_1, \ldots, \mathbf{s}_n] \cong \mathbb{C}[\Delta_1, \ldots, \Delta_n], \quad \mathbf{s}_i \mapsto \Delta_i + 1.$$

Let $\mathcal{P} \subseteq \mathcal{A}$ denote the set of all polynomial functions from $\mathbb{N}^n$ to $\mathbb{C}$. We identify a polynomial with the corresponding polynomial function. We choose a special representation of the polynomials that is adapted to the action of Δ, see [Zer08]. For $t \in \mathbb{N}^n$ and $\nu = (\nu_1, \ldots, \nu_n)$, we consider the binomial functions

$$p_\nu : \mathbb{N}^n \to \mathbb{C}, \quad t \mapsto \binom{t_1}{\nu_1} \cdots \binom{t_n}{\nu_n},$$

where $\binom{t_i}{0} = 1$ for all $i \in \{1, \ldots, n\}$. Then

$$\nu!\, p_\nu = t_1 \cdots (t_1 - \nu_1 + 1) \cdots t_n \cdots (t_n - \nu_n + 1)$$

and moreover, each element $p \in \mathcal{P}^q$ can be written as

$$p = \sum_{\nu\in\mathbb{N}^n, \nu\le_{cw}\varrho} c_\nu p_\nu \tag{1.14}$$

for $\varrho \in \mathbb{N}^n$, some suitable coefficient vectors $c_\nu \in \mathbb{C}^q$ and $\le_{cw}$ denoting the componentwise order on $\mathbb{N}^n$, that is, $\nu_i \le \varrho_i$ for all $1 \le i \le n$. Let us describe how to find this

representation. We restrict to the scalar and one-dimensional case, where $m = n = 1$. The general case can be treated similarly. For $p \in \mathcal{P} = \mathbb{C}[t]$ we show how to find the introduced representation. Usually, a polynomial p is given in the form

$$p(t) = d_v t^v + d_{v-1} t^{v-1} + \cdots + d_1 t + d_0, \quad \text{where } d_i \in \mathbb{C}.$$

To write p in the form (1.14), the occurring coefficients c_ν have to be determined. We will show how this can be done for a monomial $d_v t^v$. Since $\nu! p_\nu = t \cdot (t-1) \cdots (t-\nu+1)$, we define

$$g^{(\nu)} := t \cdot (t-1) \cdots (t-\nu+1) = t^\nu + g^{(\nu)}_{\nu-1} t^{\nu-1} + \cdots + g^{(\nu)}_1 t.$$

First, the coefficients $g^{(\nu)}_v$ will be determined for $1 \le v \le \nu$ by using the fact that $g^{(\nu)} = g^{(\nu-1)} \cdot (t - \nu + 1)$.

1. Determine $g^{(\nu)}_1$:
   The polynomial $g^{(\nu)}$ is a multiple of t. Recursively, one gets that
   $$g^{(\nu)}_1 = \begin{cases} 1 & \text{for } \nu = 1 \\ (-1)^{\nu-1} \prod_{k=1}^{\nu-1} k & \text{for } \nu > 1. \end{cases}$$

2. Determine $g^{(\nu)}_2$:
   Using $g^{(\nu)} = g^{(\nu-1)} \cdot (t - \nu + 1)$, we get
   $$\begin{aligned} g^{(\nu)}_2 &= g^{(\nu-1)}_1 - (\nu-1) \cdot g^{(\nu-1)}_2 \\ &= (-1)^{\nu-2} \prod_{k=1}^{\nu-2} k \; - \; (\nu-1) g^{(\nu-1)}_2 \end{aligned}$$
   Since $g^{(2)}_2 = 1$, we get a recursive formula.

3. Determine $g^{(\nu)}_j$ for $j \le \nu$:
   A similar consideration as in the previous point yields
   $$g^{(\nu)}_j = g^{(\nu-1)}_{j-1} - (\nu-1) \cdot g^{(\nu-1)}_j.$$

Finally, we observe

$$\begin{aligned} d_v t^v &= d_v \left(g(v) - g^{(v)}_{v-1} \cdot g(v-1) - (g^{(v)}_{v-2} - g^{(v)}_{v-1} \cdot g^{(v-1)}_{v-2}) g(v-2) - \cdots \right) \\ &= d_v \left(g(v) + \sum_{i=1}^{v-1} k_v(i) \cdot g(v-i) \right) \\ &= d_v \left(v! p_v + \sum_{i=1}^{v-1} k_v(i) \cdot (v-i)! \cdot p_{v-i} \right), \end{aligned}$$

where

$$k_v(1) := -g^{(v)}_{v-1}, \quad \text{and} \quad k_v(l) = \begin{cases} -g^{(v)}_{v-l} + \sum_{i=1}^{l-1} k_v(i) \cdot g^{(v-i)}_{v-l}, & \text{if } l < v \\ 0, & \text{if } l \ge v. \end{cases}$$

Consider for example $p(t) = t^3 + t^2 + 1$. The bounding value ϱ equals three, so by using

j	$p_1^{(j)}$	$p_2^{(j)}$	$p_3^{(j)}$	$k_3(j)$	$k_2(j)$
1	1	0	0	3	1
2	-1	1	0	1	0
3	2	-3	1	0	0

we finally get

$$\begin{aligned} t^3 &= 6 \cdot p_3 + 3 \cdot 2 \cdot p_2 + 1 \cdot p_1 \\ t^2 &= 2 \cdot p_2 + 1 \cdot p_1 \\ 1 &= p_0, \end{aligned}$$

that is,

$$p(t) = 6 \cdot p_3 + 8 \cdot p_2 + 2 \cdot p_1 + p_0. \tag{1.15}$$

In the following we show the advantage of this notation. Since

$$\begin{aligned} (\Delta_i \bullet p_{\nu_i})(t_i) &= \binom{t_i+1}{\nu_i} - \binom{t_i}{\nu_i} \\ &= \begin{cases} \frac{((t_i+1)-(t_i-\nu_i+1))\,(t_i \cdots (t_i-\nu_i+2))}{\nu_i!} & \text{if } \nu_i \geq 1 \\ 0 & \text{if } \nu_i = 0 \end{cases} \\ &= \begin{cases} \binom{t_i}{\nu_i-1} & \text{if } \nu_i \geq 1 \\ 0 & \text{if } \nu_i = 0, \end{cases} \end{aligned}$$

one gets, by using the fact that $\Delta^\mu \bullet p_\nu = \Delta_1^{\mu_1} \bullet p_{\nu_1} \cdots \Delta_n^{\mu_n} \bullet p_{\nu_n}$, the equality

$$\Delta^\mu \bullet p_\nu = \begin{cases} p_{\nu-\mu} & \text{if } \mu \leq_{cw} \nu \\ 0 & \text{otherwise.} \end{cases} \tag{1.16}$$

Remark 1.6.2 Let $p = [p_1, \ldots, p_m]^T \in \mathcal{P}^m$ with $p_i(t) = a_{d_i i} t^{\mu_{d_i i}} + \cdots + a_{1i} t^{\mu_{1i}}$, using multi-index notation. Define

$$\varrho_i = \max_{cw} \{(v_1, \ldots, v_n) \in \mathbb{N}^n \mid v_j = (\mu_{ki})_j \text{ for } 1 \leq k \leq d_i\}.$$

Then the bounding multi-index ϱ belonging to the binomial representation (1.14) is given by

$$\varrho = \max_{cw} \{(v_1, \ldots, v_n) \mid v_i = (\varrho_j)_i \text{ for } 1 \leq j \leq m\}.$$

From now on suppose that

$$p = \sum_{\nu \in \mathbb{N}^n,\, \nu \leq_{cw} \varrho} c_\nu p_\nu.$$

Due to (1.16) the identity

$$\Delta^\mu \bullet p = 0 \quad \text{for all } \mu \in \mathbb{N}^n \text{ with } \exists\, i : \mu_i > \varrho_i$$

holds. That is, after choosing representation (1.14) and fixing ϱ, we can determine the matrices A_i similarly to the continuous case. These matrices are responsible for the

condition that for every trajectory belonging to the behavior, its shift is contained as well. The operation of Δ on p can be considered analog to the continuous case as a multiplication in the corresponding module, since we have the upper bound ϱ.

For the ideal $\mathcal{I} := \left\langle \Delta_1^{\varrho_1+1}, \ldots, \Delta_n^{\varrho_n+1} \right\rangle$ in $\mathbb{C}[\Delta_1, \ldots, \Delta_n]$, consider the free $\mathbb{C}$-module $\mathcal{M} := \mathbb{C}[\Delta_1, \ldots, \Delta_n]/\mathcal{I}$. Then $\mathcal{M}$ is generated by the elements $\{[\Delta^\mu] \mid \mu \leq_{cw} \varrho\}$. The basis of $\mathcal{M}$ has $r = \prod_{i=1}^n (\varrho_i + 1)$ elements, i.e., at most r possible operations of Δ on p have to be discussed. Again we have to choose an enumeration of the basis of $\mathcal{M}$. Let $\{\mu_1, \ldots, \mu_r\}$ be an ordering of the exponents such that $\{\Delta^{\mu_1}, \ldots, \Delta^{\mu_r}\}$ is the corresponding ordered basis. Define

$$(\tilde{A}_i)_{k,l} = \begin{cases} 1 & \text{if } \mu_k = \mu_l + \mathrm{e}_i \\ 0 & \text{otherwise} \end{cases}$$

and set $A_i := \mathrm{I}_r + \tilde{A}_i$ to make it compatible with the operation of $\mathbf{s}_i$.
Since the matrices $\tilde{A}_i$ commute, the matrices A_i commute too. According to the coefficients in representation (1.14), define the $q \times r$ matrix C whose j-th column equals c_{μ_j}.
Let $\{p_1, \ldots, p_N\} \subseteq \mathcal{P}^m$. Using the introduced notation, let $A_{l1}, \ldots, A_{ln}$ belong to p_l. Then we define

$$\mathbf{A}_i = \operatorname{diag}(A_{1i}, \ldots, A_{Ni})$$

and

$$\mathbf{C} = \operatorname{diag}(C_1, \ldots, C_N).$$

Theorem 1.6.3 *[Zer08] Let $\Omega = \{p_1 \exp_{\lambda^{(1)}}, \ldots, p_N \exp_{\lambda^{(N)}}\}$ be a set of trajectories. Using the above notation, Ω possesses the* MPUM

$$\mathcal{B}_\Omega = \{\omega \in \mathcal{A}^q \mid \exists f \in \mathcal{A}^r : \ \mathbf{s}_i \bullet f = \mathbf{\Lambda_i} f \text{ for } 1 \leq i \leq n, \omega = \mathbf{C}f\},$$

where $\mathbf{\Lambda_i} = \operatorname{diag}(\lambda_i^{(1)} A_{1i}, \ldots, \lambda_i^{(N)} A_{Ni})$. Moreover, $\mathcal{B}_\Omega$ is a finite-dimensional $\mathbb{C}$-vector space.

Example 1.6.4 Let $p(t) = t^3 + t^2 + 1$. Due to (1.15), we obtain that $r = 4$ and

$$\mathcal{B}_{\{p\}} = \{\omega \in \mathbb{C}^{\mathbb{N}} \mid \exists f \in (\mathbb{C}^{\mathbb{N}})^4 : \ \mathbf{s} \bullet f = Af \text{ and } \omega = Cf\},$$

where

$$C = [1,\ 2,\ 8,\ 6] \quad \text{and} \quad A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Then $\mathcal{B}_{\{p\}}$ is equivalent to

$$\{\omega \in \mathbb{C}^{\mathbb{N}} \mid \exists x_0 \in \mathbb{C}^4 \forall t \in \mathbb{N} : \omega(t) = CA^t x_0\}.$$

Since

$$A^t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ t & 1 & 0 & 0 \\ \frac{t^2-t}{2} & t & 1 & 0 \\ \frac{t^3-3t^2+2t}{6} & \frac{t^2-t}{2} & t & 1 \end{bmatrix},$$

it follows that

$$\begin{aligned} \mathcal{B}_{\{p\}} &= \{x_{01}(t^3+t^2+1) + x_{02}(3t^2+5t+2) + x_{03}(6t+8) + x_{04}6 \mid x_{0i} \in \mathbb{C}\} \\ &= \{x_{01}p(t) + x_{02}(\Delta p)(t) + x_{03}(\Delta^2 p)(t) + x_{04}(\Delta^3 p)(t) \mid x_{0i} \in \mathbb{C}\}. \end{aligned}$$

Remark 1.6.5 In [Zer08] the previous theorem is elaborated for polynomial exponential trajectories over finite rings as well (exclusive of the vector space property). The result differs slightly in case $\lambda_i^{(j)}$ is not a unit. Then an additional requirement needs to be satisfied by the latent variable f.

Remark 1.6.6 Minimality questions can be answered like in the continuous case. But note that this can not be done straightforwardly for systems over finite rings.

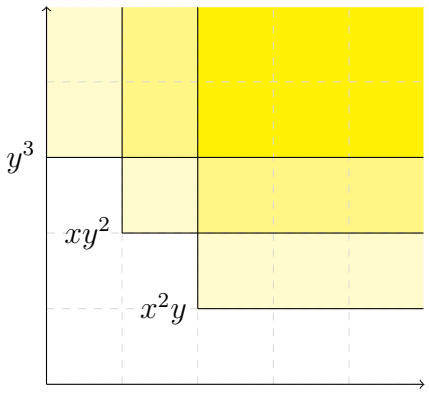
Chapter 2

Gröbner bases

Gröbner bases are a powerful and copious instrument of modern computer algebra. The breakthrough for the computational aspect is the PhD thesis of Bruno Buchberger published 1965, in which Buchberger set up a constructive method to calculate these bases. Gröbner bases can be considered as a generalization of the Gaussian elimination and the Euclidean algorithm. The goal is to find a specific generating system of a polynomial ideal or more generally, of a polynomial module. In the case of a univariate polynomial ring with field coefficients, an ideal can always be generated by a single element. The degree of this element is obviously minimal.
Gröbner bases carry this feature over to the multivariate case in a generalized meaning. The specific generating set satisfies the property of containing the smallest degree elements with respect to a monomial ordering. Consider for instance the ideal generated by the elements of

$$F = \{y^3 + x^2y, xy^2, x^2y^2 + x^2y\} \subseteq K[x, y]$$

for a field K. One can show that $\{y^3, xy^2, x^2y\}$ is a Gröbner basis of $\langle F\rangle_{K[x,y]}$ with respect to the lexicographical ordering.



We will see later that there does not exist an element in $\langle F\rangle_{K[x,y]}$ with leading monomial in the uncolored area. Note that the membership question can then be answered as follows: One reduces the leading term of the potential candidate as long as possible via the Gröbner basis elements. If the procedure terminates with zero, the element is contained in the ideal. Else we obtain an element with leading term in the uncolored area and thus the candidate is not contained in the ideal. This outlined observation is

just one advantage of the Gröbner bases approach. The crucial idea behind the theory is the observation that the so-called leading ideal or rather, the leading submodule may contain a lot of information.
The groundbreaking advantage of Gröbner bases lies in the algorithmical aspect. They provide the theoretical fundament to work constructively with polynomial ideals and modules. Beyond that, there exist various computer algebra systems with a respectable implementation for commutative Gröbner bases with field coefficients. Well known examples are MAPLE, MAGMA, COCOA and SINGULAR::PLURAL. Further we want to stress that the freely available software SINGULAR::PLURAL provides the framework of ring coefficients as well as the non-commutative features.

Gröbner bases are a keystone in this thesis. In Chapter 3, Gröbner bases over univariate polynomial rings with coefficients in a finite ring serve to extend the idea of the predictable degree property from the field case to the ring case. Non-commutative Gröbner bases allow to set up a new algorithm to compute the Jacobson form in Chapter 4. And finally, the exact linear modeling approach proposed in Chapter 5 can explicitly be obtained by using multivariate non-commutative Gröbner bases.
There exist several textbooks giving an extensive introduction to the subject. In this thesis, we will work close to [AL94] when we discuss commutative structures and close to [Lev05a] in connection to the non-commutative ones. In the sequel we will set up the formal framework and point out thesis-relevant applications. We want to stress that [LXB08] gives a comprehensive overview and references of Gröbner bases applications in signal and system theory.

2.1 Commutative Gröbner bases

Let $\mathcal{D}$ denote the polynomial ring $\mathcal{R}[x_1, \ldots, x_n]$ for a commutative Noetherian ring $\mathcal{R}$. Then due to Hilbert's Basis Theorem, the ring $\mathcal{D}$ is Noetherian as well. We denote the set of all **monomials** by

$$\mathrm{Mon}(\mathcal{D}) := \{x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha_i \in \mathbb{N}\}.$$

In order to keep the notation short we denote $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ by x^α, where $\alpha = (\alpha_1, \ldots, \alpha_n)$. In the motivating example we have already indicated the interest to order the monomials. In the univariate case, this is can be handled by the degree. There are several possibilities to order the monomials of a multivariate polynomial ring. We are exclusively interested in monomial orderings. A **monomial order** on $\mathrm{Mon}(\mathcal{D})$ is a relation $<$ satisfying:

1. $<$ is a total ordering, that is, transitive and for all $x^\alpha, x^\beta \in \mathrm{Mon}(\mathcal{D})$ precisely one of the following relations must hold:

$$x^\alpha < x^\beta, \quad x^\alpha = x^\beta \quad \text{or} \quad x^\beta < x^\alpha.$$

2. $1 < x^\alpha$ for all $x^\alpha \in \mathrm{Mon}(\mathcal{D}) \backslash \{1\}$.

3. If $x^\alpha < x^\beta$, then $x^\alpha x^\gamma < x^\beta x^\gamma$, for all $x^\gamma \in \mathrm{Mon}(\mathcal{D})$.

One can show that a monomial order extends the divisibility relation, that is, if x^α divides x^β then $x^\alpha \leq x^\beta$, see [AL94, Proposition 1.4.5]. And furthermore every monomial ordering is a well-ordering, that is, every non-empty set of monomials has a minimal element with respect to $<$, see [AL94, Theorem 1.4.6].
Let $\deg(x^\alpha) = \alpha_1 + \cdots + \alpha_n$ denote the **total degree** of $x^\alpha \in \mathcal{D}$, $\alpha \in \mathbb{N}^n$.

Example 2.1.1 Let $x^\alpha, x^\beta \in \mathcal{D}$ with $\alpha, \beta \in \mathbb{N}^n$.

1. We define the **lexicographical order** as follows:

   $$x^\alpha <_{lex} x^\beta$$

   if and only if there exists $1 \leq i \leq n$ such that $\alpha_1 = \beta_1, \ldots, \alpha_{i-1} = \beta_{i-1}, \alpha_i < \beta_i$.

2. We define the **reverse lexicographical order** as follows:

   $$x^\alpha <_{revlex} x^\beta$$

   if and only if there exists $1 \leq i \leq n$ such that $\alpha_n = \beta_n, \ldots, \alpha_{i+1} = \beta_{i+1}, \alpha_i < \beta_i$.

3. We define the **degree reverse lexicographical order** as follows:

   $$x^\alpha <_{degrevlex} x^\beta$$

   if and only if $\deg(x^\alpha) < \deg(x^\beta)$, or $\deg(x^\alpha) = \deg(x^\beta)$ and $x^\alpha <_{revlex} x^\beta$.

4. We define the **degree lexicographical ordering** as follows:

   $$x^\alpha <_{deglex} x^\beta$$

   if and only if $\deg(x^\alpha) < \deg(x^\beta)$, or $\deg(x^\alpha) = \deg(x^\beta)$ and $x^\alpha <_{lex} x^\beta$.

In the sequel let $<$ denote a monomial ordering. Clearly, whatever ordering is chosen, every nonzero element $f \in \mathcal{D}$ can be written as

$$f = \sum_{i=1}^{L} c_i X_i,$$

where $L \in \mathbb{N}$, the c_i's are nonzero elements of $\mathcal{R}$ for $i = 1, \ldots, L$, and $X_1, \ldots, X_L$ are monomials, ordered as $X_1 > \cdots > X_L$. Using the terminology of [AL94], we define

- $\mathrm{lm}(f) := X_1$ as the **leading monomial** of f
- $\mathrm{lt}(f) := c_1 X_1$ as the **leading term** of f
- $\mathrm{lc}(f) := c_1$ as the **leading coefficient** of f.

There are several ways to define Gröbner bases, here we adopt the definition of [AL94] which requires us to first define the concept of "leading ideal".

Definition 2.1.2 Let G be a non-empty subset of $\mathcal{D}$. Then the submodule $\mathrm{L}(G) \subseteq \mathcal{D}$, defined as

$$\mathrm{L}(G) := \langle \mathrm{lt}(g) \mid g \in G \backslash \{0\} \rangle$$

is called the **leading ideal** of G.

For example, let $G = \{x_1x_2 + x_1, x_1^2 + x_2^2\} \subseteq \mathbb{R}[x_1, x_2]$. Using the lexicographical ordering, we obtain $\mathrm{L}(G) = \langle x_1^2, x_1x_2 \rangle$, whereas using the reverse lexicographical ordering, we get $\mathrm{L}(G) = \langle x_1x_2, x_2^2 \rangle$.

Definition 2.1.3 Let $I \subseteq \mathcal{D}$ be a nonzero ideal and G a non-empty finite subset of I consisting of nonzero elements. Then G is called a **Gröbner basis** of I if

$$\mathrm{L}(G) = \mathrm{L}(I).$$

Referring to [AL94, Corollary 4.1.17], a Gröbner basis does always exist. It is obvious that a Gröbner basis is not unique at all. Further we want to stress that a Gröbner basis may vary with respect to the chosen ordering.

Example 2.1.4 Consider the previous example. Then one can show that corresponding to the lexicographical ordering, we obtain $\{x_2^3 + x_2^2, x_1^2 + x_2^2, x_1x_2 + x_1\}$ as a Gröbner basis, whereas corresponding to the reverse lexicographical ordering, a Gröbner basis is given by $\{x_1^3 + x_1, x_2^2 + x_1^2, x_1x_2 + x_1\}$.

Remark 2.1.5 A Gröbner basis is not a basis in the sense of linear algebra, that is, it does not provide a unique representation of each element in general. This can be easily seen from the simple example $M := \langle x, y \rangle \subseteq \mathbb{R}[x, y]$. The ideal M possesses the Gröbner basis $\{x, y\}$ and the element xy contained in M can be written as a product of x and as a product of y as well. Under certain conditions on the underlying ring and suitable requirements on the Gröbner basis, a unique representation can be achieved. We will pick up this idea later.

The whole introduced notation can be extended easily to the multivariable case. Consider the free module

$$\mathcal{D}^q := \mathcal{D}\,\mathrm{e}_1 \oplus \cdots \oplus \mathcal{D}\,\mathrm{e}_q,$$

where e_i denotes the i-th unit vector, a vector of length q possessing one in the i-th position and zeros else. The elements of $\mathcal{D}^q$ are considered to be row vectors in this and the next chapter. Then

$$\mathrm{Mon}(\mathcal{D}^q) := \{X\,\mathrm{e}_i \mid X \in \mathrm{Mon}(\mathcal{D}) \text{ and } 1 \leq i \leq q\}.$$

According to [AL94], we call a total order $<$ on $\mathrm{Mon}(\mathcal{D}^q)$ satisfying:

1. $X < ZX$, for every $X \in \mathrm{Mon}(\mathcal{D}^q)$ and $Z \in \mathrm{Mon}(\mathcal{D}) \backslash \{1\}$
2. If $X < Y$, then $ZX < ZY$ for all $X, Y \in \mathrm{Mon}(\mathcal{D}^q)$ and every $Z \in \mathrm{Mon}(\mathcal{D})$

a **monomial ordering** on the monomials of $\mathcal{D}^q$.

The notations leading monomial, leading term and leading coefficient given for the skalar case carry over. The leading term of any element $0 \neq f \in \mathcal{D}^q$ can be written as $X_f \, \mathrm{e}_i$ with $X_f \in \mathrm{Mon}(\mathcal{D})$ and we define $\mathrm{lpos}(f) := i$ as the **leading position** of f. Possible orderings here in the multivariable case are listed below. Let $<$ be a monomial ordering on $\mathrm{Mon}(\mathcal{D})$ and $X, Y \in \mathrm{Mon}(\mathcal{D})$.

Definition 2.1.6

- The **Term Over Position (TOP)** ordering is characterized by

$$X \, \mathrm{e}_i <_{\mathrm{TOP}} Y \, \mathrm{e}_j \quad \text{if and only if} \quad X < Y \text{ or } (X = Y \text{ and } i > j).$$

- The **Position Over Term (POT)** ordering is given by

$$X \, \mathrm{e}_i <_{\mathrm{POT}} Y \, \mathrm{e}_j \quad \text{if and only if} \quad i > j \text{ or } (i = j \text{ and } X < Y).$$

Throughout the thesis, a module ordering will be POT or TOP. In Example 2.1.4 we have already seen for the ideal case that a Gröbner basis depends on the chosen ordering. Evidently this extends to the module case. We will show in Chapter 4 how to apply the POT ordering to obtain certain matrix forms. Beyond that, Chapter 3 and the following section demonstrate the benefit of TOP for row-reduced representations.

It is easily verified that the next observation holds irrespective of whether TOP or POT ordering is used.

Observation 2.1.7 Let $f_1, f_2, \ldots, f_m$ be nonzero vectors in $\mathcal{D}^q$ with distinct leading monomials, ordered accordingly as $\mathrm{lm}(f_1) > \mathrm{lm}(f_2) > \cdots > \mathrm{lm}(f_m)$. Then

$$\mathrm{lt}(f_1 + f_2 + \cdots + f_m) = \mathrm{lt}(f_1).$$

As before, for a subset G of $\mathcal{D}^q$, we define the **leading submodule** $\mathrm{L}(G)$ to be the module generated by the leading terms of all nonzero elements contained in G. And we define further for a nonzero module $M \subseteq \mathcal{D}^q$, a non-empty finite subset $G \subseteq M$ consisting of nonzero elements and satisfying $\mathrm{L}(G) = \mathrm{L}(M)$ to be a Gröbner basis of M. Again referring to [AL94, Corollary 4.1.17, Exercise 4.1.14], a Gröbner basis does always exist. The definition provides the next lemma.

Observation 2.1.8 Let M be a submodule of $\mathcal{D}^q$ with Gröbner basis $\{g_1, \ldots, g_m\}$ and let $0 \neq f \in M$. Then there exist a subset $\{g_{j_1}, \ldots, g_{j_s}\}$ of G, $\alpha_1, \ldots, \alpha_s \in \mathbb{N}^n$ and $c_1, \ldots, c_s \in \mathcal{R}$, such that

- $\mathrm{lm}(f) = x^{\alpha_i} \, \mathrm{lm}(g_{j_i})$ for $i = 1, \ldots, s$ and
- $\mathrm{lt}(f) = c_1 x^{\alpha_1} \, \mathrm{lt}(g_{j_1}) + \cdots + c_s x^{\alpha_s} \, \mathrm{lt}(g_{j_s})$.

Note that the g_{j_i}'s of the above observation all satisfy $\mathrm{lpos}(g_{j_i}) = \mathrm{lpos}(f)$ and $\mathrm{lm}(g_{j_i}) \leq \mathrm{lm}(f)$. The above observation inspires the next definition.

Definition 2.1.9 ([AL94]) Let $0 \neq f \in \mathcal{D}^q$ and let $F = \{f_1, \ldots, f_s\} \subseteq \mathcal{D}^q$ be a set of nonzero elements. Let $\alpha_1, \ldots, \alpha_s \in \mathbb{N}^n$ and let $c_1, \ldots, c_s$ be elements of $\mathcal{R}$ such that

1. $\mathrm{lm}(f) = x^{\alpha_i} \mathrm{lm}(f_i)$ for $i = 1, \ldots, s$ and
2. $\mathrm{lt}(f) = c_1 x^{\alpha_1} \mathrm{lt}(f_1) + \cdots + c_s x^{\alpha_s} \mathrm{lt}(f_s)$.

Define

$$h := f - (c_1 x^{\alpha_1} f_1 + \cdots + c_s x^{\alpha_s} f_s).$$

Then we say that f **reduces** to h modulo F and we write

$$f \xrightarrow{F} h.$$

If f cannot be reduced modulo F, we say that f is **minimal** with respect to F.

Lemma 2.1.10 *Let f, h and F be as in the above definition. If $f \xrightarrow{F} h$, then* $\mathrm{lm}(h) < \mathrm{lm}(f)$.

Proof: From property 1. of Definition 2.1.9, it follows that property 2. of Definition 2.1.9 translates into

$$\begin{aligned} \mathrm{lt}(f) &= c_1 x^{\alpha_1} \mathrm{lt}(f_1) + \cdots + c_s x^{\alpha_s} \mathrm{lt}(f_s) \\ &= \mathrm{lt}(c_1 x^{\alpha_1} f_1 + \cdots + c_s x^{\alpha_s} f_s). \end{aligned}$$

From this, it immediately follows that

$$\mathrm{lm}(h) = \mathrm{lm}(f - (c_1 x^{\alpha_1} f_1 + \cdots + c_s x^{\alpha_s} f_s)) < \mathrm{lm}(f).$$

□

The next observation follows by definition and will prove useful in the sequel.

Observation 2.1.11 Let M be a submodule of $\mathcal{D}^q$ with Gröbner basis G and let $0 \neq f \in M$. Then

$$f \in \langle g \in G \mid \mathrm{lm}(g) \leq \mathrm{lm}(f) \rangle.$$

Definition 2.1.12 [AL94] A Gröbner basis G is called **minimal** if all its elements g are minimal with respect to $G \backslash \{g\}$.

Referring to [AL94, Exercise 4.1.9, Exercise 4.1.14], a minimal Gröbner basis does always exist. In certain cases like for instance field coefficients, the so-called reduced Gröbner bases can be introduced, which even provides a canonical form. Later these observations will be analyzed more in detail.

2.1.1 One-dimensional case and applications to signals and systems

We give an extended introduction to the preliminaries outlined in [KSb] to motivate the results introduced in Chapter 3.

Coefficients over fields

In this subsection, we limit our attention to the case that $\mathcal{R}$ is a field and $n = 1$. It is well-known that Gröbner bases are useful for various applications over fields, including univariate applications. In this section, we attribute this usefulness to a particular property of minimal Gröbner bases that we label the "Predictable Leading Monomial (PLM)" property. We consider two particular applications and show how the PLM property is useful for these applications.
As we will see below, the elements of a minimal Gröbner basis G in $\mathcal{D}^q$ can be ordered according to their respective leading monomials. Since $\mathcal{R}$ is a field, this yields even more information.

Remark 2.1.13 Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis G. Then all leading positions of elements of G are distinct. Further G has at most q elements.

Proof: The second claim of the proposition follows directly from the first claim. To see the first claim, assume $g_1, g_2 \in G$ to be such that $\mathrm{lpos}(g_1) = \mathrm{lpos}(g_2)$. Further assume $x^{\alpha_1} =: \mathrm{lm}(g_1) \leq \mathrm{lm}(g_2) := x^{\alpha_2}$. Then g_2 can be reduced to $g_2 - \frac{c_2}{c_1} x^{\alpha_2 - \alpha_1} g_1$, where $c_1 := \mathrm{lc}(g_1)$ and $c_2 := \mathrm{lc}(g_2)$, modulo $G \backslash \{g_2\}$. This yields a contradiction to the minimality of G. □

Corollary 2.1.14 *A minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$ has the convenient property that its elements can be ordered as $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$.*

Focusing on the case that $\mathcal{R}$ is a field, in the next theorem we identify an important property of a minimal Gröbner basis. We first introduce the following terminologies. Note that the degree of a nonzero polynomial vector $f \in \mathcal{D}^q$ is defined as degree of $\mathrm{lm}(f)$.

Definition 2.1.15 Let $\mathcal{R}$ be a field. Further let $F = \{f_1, \ldots, f_s\} \subseteq \mathcal{D}^q$ be a set of nonzero elements. Then F has the **Predictable Degree (PD) property** if for any $0 \neq f \in {}_{\mathcal{D}}\langle F \rangle$, written as

$$f = a_1 f_1 + \cdots + a_s f_s, \tag{2.1}$$

where $a_1, \ldots, a_s \in \mathcal{D}$, we have

$$\deg(f) = \max_{1 \leq i \leq s; a_i \neq 0} (\deg(a_i) + \deg(f_i)).$$

Next, F is said to have the **Predictable Leading Position (PLP) property** if

$$\mathrm{lpos}(f) = \min_{1 \leq i \leq s; a_i \neq 0} \mathrm{lpos}(f_i).$$

Finally, F is said to have the **Predictable Leading Monomial (PLM) property** if

$$\mathrm{lm}(f) = \max_{1\le i\le s; a_i\neq 0}(\mathrm{lm}(a_i)\,\mathrm{lm}(f_i)). \tag{2.2}$$

Note that the PD property is well established in the literature, see [For75] where it was first introduced. Above we defined the PLM property as a more general and stronger concept that is natural for minimal Gröbner bases. It can be easily verified that the PLM property holds if and only if both the PD property and the PLP property hold. As we shall see below, in applications such as minimal state space realization, the PD property suffices, whereas an application such as minimal partial realization/interpolation requires the PLM property.

Remark 2.1.16 Suppose $F \subseteq {}_{\mathcal{D}}\langle F\rangle \subseteq \mathcal{D}^q$ to possess the PLM property. Then F is a basis of ${}_{\mathcal{D}}\langle F\rangle$.

Proof: Let $F = \{f_1, \ldots, f_s\}$. Evidently F generates ${}_{\mathcal{D}}\langle F\rangle$. Thus it is left to show that the elements of F are linearly independent. For this purpose note first that $f_i \neq f_j$ implies $a_i f_i \neq a_j f_j$ for arbitrary nonzero elements $a_i, a_j \in \mathcal{D}$. To see this suppose the claim does not hold, thus let $a_i f_i = a_j f_j$ for $i \neq j$. Without loss of generality we may additionally assume that $a_i \in \mathcal{D}\backslash\mathcal{R}$, because else one could replace a_i, a_j by xa_i, xa_j. Then we can write

$$f_i = a_i f_i - a_j f_j + f_i = (a_i + 1) f_i - a_j f_j,$$

which introduces a contradiction to the PLM property. Further

$$f_i \neq f_j \quad \text{yields} \quad \mathrm{lt}(a_i f_i) \neq \mathrm{lt}(a_j f_j) \tag{2.3}$$

for arbitrary nonzero elements $a_i, a_j \in \mathcal{D}$, because otherwise

$$\mathrm{lm}(\underbrace{a_i f_i - a_j f_j}_{\neq 0}) < \max_{k\in\{i,j\}}(\mathrm{lm}(a_k)\,\mathrm{lm}(f_k)).$$

Choosing $a_i = \mathrm{lc}(f_i)^{-1}$ and $a_j = \mathrm{lc}(f_j)^{-1}$, we obtain by using (2.3) that

$$f_i \neq f_j \quad \text{yields} \quad \mathrm{lm}(a_i f_i) \neq \mathrm{lm}(a_j f_j). \tag{2.4}$$

Finally, suppose that $\sum a_i f_i = 0$ for some $a_i \in \mathcal{D}$. We need to show that $a_i = 0$ for all i. Assume the converse and let k be the first integer such that $a_k \neq 0$. This leads to $a_k f_k = -\sum_{j=k+1} a_j f_j$ and since the PLM property holds, there would exist an element $k \neq j_* \in \{k+1, \ldots, s\}$ such that $\mathrm{lm}(a_k f_k) = \mathrm{lm}(a_{j_*} f_{j_*})$. This is a contradiction to (2.4) and completes the proof. □

Theorem 2.1.17 *Let $\mathcal{R}$ be a field. Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis G. Then G has the Predictable Leading Monomial (PLM) property.*

Proof: Write $G = \{g_1, \ldots, g_m\}$. Since G is minimal, we may assume that $\mathrm{lm}(g_1) > \mathrm{lm}(g_2) > \cdots > \mathrm{lm}(g_m)$. Let $f = a_1 g_1 + \cdots + a_m g_m$. For simplicity of notation, we assume that a_i is nonzero for $1 \leq i \leq m$. Since $\mathcal{R}$ is a field, we have that $\mathrm{lpos}(a_i g_i) = \mathrm{lpos}(g_i)$ for $1 \leq i \leq m$. Also, all leading positions of the g_i's are distinct due to Remark 2.1.13. As a result, all leading monomials of the $a_i g_i$'s are distinct. Thus there exists an ordering

$$\mathrm{lm}(a_{j_1} g_{j_1}) > \mathrm{lm}(a_{j_2} g_{j_2}) > \cdots > \mathrm{lm}(a_{j_m} g_{j_m}).$$

It now follows from Observation 2.1.7 that

$$\mathrm{lm}(f) = \mathrm{lm}(a_{j_1} g_{j_1}) = \mathrm{lm}(a_{j_1}) \, \mathrm{lm}(g_{j_1}) = \max_{1 \leq i \leq m} (\mathrm{lm}(a_i) \, \mathrm{lm}(g_i)),$$

which proves the PLM property. □

Corollary 2.1.18 *Due to Remark 2.1.16 and Theorem 2.1.17, a minimal Gröbner basis is a basis.*

The previous results rely on the fact that the considered ring is univariate and possesses field coefficients. But as already outlined in Remark 2.1.5, a Gröbner basis is not a basis in the sense of linear algebra in general. In this sense, minimal Gröbner bases are not bases for multivariate rings with field coefficients. This difficulty can be solved via the notion of "Janet bases" [GY05, PR05]. In the sequel, we show some applications of the previous results.

We call a matrix $R \in \mathcal{D}^{m \times q}$ **upper triangular** if $R_{ij} = 0$ for $j < i$.

Remark 2.1.19 Suppose $G = \{g_1, \ldots, g_m\}$ to be a minimal Gröbner basis of a module $M \subseteq \mathcal{D}^q$ corresponding the POT ordering. Further let $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Then $[g_1^T, \ldots, g_m^T]^T$ corresponds to a full row rank upper triangular matrix whose rows generate M.

Proof: It is sufficient to show the triangularity of $[g_1^T, \ldots, g_m^T]^T$, since the rank property can be deduced easily. Suppose the claim does not hold. Due to the use of POT, this implies that there exist two elements with equal leading positions. But this introduces a contradiction to Remark 2.1.16. □

Remark 2.1.20 Suppose $G = \{g_1, \ldots, g_m\}$ to be a minimal Gröbner basis of a module $M \subseteq \mathcal{D}^q$ corresponding the TOP ordering. Further let $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Then

$$\begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} = \mathrm{diag}\,(x^{\deg(g_1)}, \ldots, x^{\deg(g_m)}) B(x),$$

where $B(x)$ is a proper rational matrix such that $B(\infty)$ is of full row rank. Further the rows of $B(\infty)$ can be permuted to a upper triangular matrix.

Proof: By the choice of the ordering, it is evident that $B(x)$ is proper. Further it is easy to see that $B(\infty)_{ij} \in \mathcal{R}\backslash\{0\}$ if and only if $\deg(B_{ij}) = \deg(g_i)$ and zero else. Thus $B(\infty)_i$ is a vector with constant coefficient entries and by the choice of the ordering

$$B(\infty)_{ij} = 0 \text{ for all } j < \text{lpos}(g_i). \tag{2.5}$$

Due to Remark 2.1.13, there exists a permutation π of $(g_1, \ldots, g_m) =: G_{seq}$ such that

$$\text{lpos}(\pi(G_{seq})_m) > \cdots > \text{lpos}(\pi(G_{seq})_1).$$

Using (2.5) yields that $[\pi(G_{seq})_1^T, \ldots, \pi(G_{seq})_m^T]^T$ is a upper triangular matrix. Since the row rank is invariant under permutation of rows, $B(\infty)$ is obviously of full rank.□

It should be noted that the upper triangularity in the previous remark is crucial. Without this requirement, the matrix is called **row reduced** in the literature. Clearly, the row vectors of a row reduced matrix do not necessarily constitute a minimal Gröbner basis. For example, for $q = 2$ and $\mathcal{R} = \mathbb{Z}_2$ consider

$$G(x) = \begin{bmatrix} \underline{x^2} & 0 \\ \underline{x} & x \end{bmatrix}.$$

This matrix is clearly row reduced since $G(x) = \text{diag}\,(x^2, x)B(x)$ with

$$B(\infty) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

However, the row vectors of $G(x)$ do not constitute a minimal Gröbner basis for their span, since the first row vector can be reduced modulo the second row vector, yielding

$$\begin{bmatrix} \underline{x^2} & 0 \end{bmatrix} - x \begin{bmatrix} \underline{x} & x \end{bmatrix} = \begin{bmatrix} 0 & \underline{-x^2} \end{bmatrix}.$$

The next two examples show two applications over fields where the PD property and the PLM property are useful.

Example 2.1.21 : Using minimal Gröbner bases for parameterization of all shortest linear recurrence relations
Consider the sequence $S_0, S_1, S_2, S_3, S_4 = 1, 4, 3, 3, 2$ over the field $\mathbb{Z}_5$. A polynomial $d(x)$, written as $d(x) = x^L + d_{L-1}x^{L-1} + \cdots + d_1 x + d_0$, is called a **linear recurrence relation of length** L for S_0, S_1, S_2, S_3, S_4 if

$$S_{L+j} + \sum_{i=1}^{L} d_{L-i} S_{L+j-i} = 0 \qquad \text{for } j = 0, \ldots, 5 - L - 1. \tag{2.6}$$

Defining the **partial impulse response trajectory** $\boldsymbol{b}$ on the time-axis $\mathbb{N}$ as

$$\boldsymbol{b} = \left(\begin{bmatrix} S_0 \\ 0 \end{bmatrix}, \begin{bmatrix} S_1 \\ 0 \end{bmatrix}, \begin{bmatrix} S_2 \\ 0 \end{bmatrix}, \begin{bmatrix} S_3 \\ 0 \end{bmatrix}, \begin{bmatrix} S_4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ldots \right), \tag{2.7}$$

we can reformulate (2.6) as $[d(\mathbf{s}) \quad -h(\mathbf{s})] \bullet \boldsymbol{b} = 0$, where $h(x)$ is a polynomial of degree $\leq L$ and $\mathbf{s}$ is the backward shift operator, acting on trajectories $\boldsymbol{w}$ on $\mathbb{N}$ as

$(\mathbf{s}\boldsymbol{w})(k) = \boldsymbol{w}(k+1)$. A linear recurrence relation for S_0, S_1, S_2, S_3, S_4 thus corresponds to a kernel representation

$$[d(\mathbf{s}) \quad -h(\mathbf{s})] \bullet \boldsymbol{w} = 0$$

whose behavior includes the so-called **partial impulse response behavior**

$$\mathcal{B} := \text{span}\{\boldsymbol{b}, \mathbf{s}\boldsymbol{b}, \mathbf{s}^2\boldsymbol{b}, , \ldots, \mathbf{s}^5\boldsymbol{b}\}, \tag{2.8}$$

where $\boldsymbol{b}$ is defined by (2.7). The search for shortest linear recurrence relations now translates into a search for an annihilator $[d(\mathbf{s}) \quad -h(\mathbf{s})] \bullet \boldsymbol{w} = 0$ for $\mathcal{B}$ that has minimal row degree and satisfies $\deg(h) \leq \deg(d)$. Next, define the polynomial $S(x)$ as

$$S(x) := S_0x^5 + S_1x^4 + S_2x^3 + S_3x^2 + S_4x, \tag{2.9}$$

and consider the module M spanned by $\begin{bmatrix} 1 & -S(x) \end{bmatrix}$ and $\begin{bmatrix} 0 & x^6 \end{bmatrix}$. Clearly, these two polynomial vectors are linearly independent annihilators of $\mathcal{B}$, and thus M essentially consists of all annihilators of $\mathcal{B}$. It is not difficult to see that any minimal Gröbner basis for M must consist of 2 vectors. Exactly one of these vectors has leading position 1. Because of the PLM property, this vector yields a shortest linear recurrence relation. In this example, a minimal Gröbner basis for M is given by $G = \{g_1, g_2\}$, where

$$g_1(x) = \begin{bmatrix} 2x+2 & x^4 - 2x^3 + x \end{bmatrix} \text{ and } g_2(x) = \begin{bmatrix} x^2 - 3x - 1 & 4x^2 - 3x \end{bmatrix}.$$

It follows that $x^2 - 3x - 1$ is a shortest linear recurrence relation for the sequence $S_0, S_1, S_2, S_3, S_4 = 1, 4, 3, 3, 2$ over $\mathbb{Z}_5$. More precisely, we obtain

$$S_{2+j} + 2S_{1+j} + 4S_j = 0, \text{ where } j = 0 \ldots 2.$$

Example 2.1.22 : **Using minimal Gröbner bases for minimal state space realization—convolutional coding application**
According to [RSY96, RS99, GLS07], a finite support binary convolutional code of length n is defined as a submodule of $\mathbb{Z}_2[x]^n$. Consider the finite support binary convolutional code $\mathcal{C}$ of length 3 given by the encoder

$$E(x) = \begin{bmatrix} x^2+1 & 1 & 0 \\ x & 0 & 1 \end{bmatrix}.$$

A Viterbi decoder for $\mathcal{C} = \text{im}\, E(x)$ is based on a so-called "trellis representation" of $\mathcal{C}$, which is essentially a state space realization $E(x) = B(x^{-1}I - A)^{-1}C + D$, see [JW93, JZ99, GLS07]. The need for low complexity decoding motivates the use of a trellis representation, where the matrix A is of minimal size. In this example, a minimal Gröbner basis for the module $\mathcal{C}$ is given by $G = \{g_1, g_2\}$, where

$$g_1(x) = \begin{bmatrix} x & 0 & 1 \end{bmatrix} \text{ and } g_2(x) = \begin{bmatrix} 1 & 1 & x \end{bmatrix}.$$

Thus

$$\tilde{E}(x) = \begin{bmatrix} x & 0 & 1 \\ 1 & 1 & x \end{bmatrix}.$$

is also an encoder for $\mathcal{C}$; its controller canonical realization (A, B, C, D) is given by inspection as

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Note that the size of A equals the sum of the row degrees of $\tilde{E}$. Because of the PLM property of G (or actually the PD property), there exists no encoder of $\mathcal{C}$ whose sum of row degrees is smaller than 2. For this reason, (A, B, C, D) is a minimal state space realization and the corresponding trellis representation is also minimal.

Coefficients over a finite ring

In our univariate context, Theorem 2.1.17 fails when $\mathcal{R}$ is not a field. Indeed, consider the module $M := \langle x+1, 2\rangle$ in $\mathbb{Z}_4[x]$. The set $\{x+1, 2\}$ is a minimal Gröbner basis for M. However, the element $2 \in M$ can be generated in two different ways, namely $2 = 0\cdot(x+1)+1\cdot 2$, but also $2 = 2\cdot(x+1)+x\cdot 2$. Thus, a minimal Gröbner basis is not necessarily a basis in the ring case and does not necessarily have the PLM property. The result given in Remark 2.1.13 fails.
Assuming $\mathcal{R} = \mathbb{Z}_{p^r}$ for p prime, we can formulate another relation according to the number of generators. Before we can give the result, some preliminaries on $\mathbb{Z}_{p^r}$ are required. A set that plays a fundamental role in connection to these rings is the set of "digits", denoted by $\mathrm{A}_p = \{0, 1, \ldots, p-1\} \subset \mathbb{Z}_{p^r}$. Recall that any element $a \in \mathbb{Z}_{p^r}$ can be written uniquely as

$$a = \theta_0 + p\theta_1 + \cdots + p^{r-1}\theta_{r-1},$$

where $\theta_\ell \in \mathrm{A}_p$ for $\ell = 0, \ldots, r-1$ (p-adic expansion).
Next, an element a in $\mathbb{Z}_{p^r}$ is said to have **order** k if the additive subgroup generated by a has p^k elements. Elements of order r are called **units**. Further, two elements a_1 and a_2 are called **associates** if $a_1 = \theta a_2$, for some unit $\theta \in \mathbb{Z}_{p^r}$. One can easily see that each $a \in \mathbb{Z}_{p^r}$ of order k can be written as $a = \theta p^{r-k}$, where θ is a unit. Therefore, a_1 and a_2 are associates if and only if they have the same order. Thus the elements $1, p, p^2, \ldots, p^{r-1}$ have orders $r, r-1, r-2, \ldots, 1$, respectively. We extend the notion of order to polynomial vectors as follows.

Definition 2.1.23 The **order** of a nonzero polynomial vector $f \in \mathcal{R}[x]^q$ is defined as the order of $\mathrm{lc}(f)$, and is denoted by $\mathrm{ord}(f)$.

Unlike the field case, a minimal Gröbner basis of a module in $\mathbb{Z}_{p^r}[x]^q$ is not a basis. In fact, the leading positions of its elements are not necessarily distinct. This is shown, for instance, by the minimal Gröbner basis $G := \{[3x \quad 0], [x^2 \quad 0]\} \in \mathbb{Z}_9[x]^2$.

Remark 2.1.24 Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with minimal Gröbner basis G. Then the elements of G can be ordered as $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$.

Proof: Suppose the claim does not hold. Then one can find unequal indices i, j such that $\mathrm{lm}(g_i) = \mathrm{lm}(g_j) = x^\alpha \mathrm{e}_k$ for $1 \leq k \leq m$ and $\alpha \in \mathbb{N}$. Thus the leading term of g_i and g_j can we written as

$$\mathrm{lt}(g_i) = u_i p^{\beta_i} x^\alpha \mathrm{e}_k \quad \text{and} \quad \mathrm{lt}(g_j) = u_j p^{\beta_j} x^\alpha \mathrm{e}_k,$$

where u_i, u_j are units and $0 \leq \beta_i, \beta_j \leq r-1$. Assume without loss of generality that $\beta_i \leq \beta_j$. Then $\mathrm{lt}(g_j) = \frac{u_j}{u_i} p^{\beta_j - \beta_i} \mathrm{lt}(g_i)$ and thus g_j is not minimal with respect to $G \backslash \{g_j\}$. But this introduces a contradiction to the minimality of G. □

Lemma 2.1.25 *Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered as $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $j < i$ be such that $\mathrm{lpos}(g_j) = \mathrm{lpos}(g_i)$. Then $\deg(g_j) > \deg(g_i)$ and $\mathrm{ord}(g_j) > \mathrm{ord}(g_i)$. In particular, $m \leq qr$.*

Proof: Since $\mathrm{lpos}(g_j) = \mathrm{lpos}(g_i)$ and $\mathrm{lm}(g_j) > \mathrm{lm}(g_i)$, we must have that $\deg(g_j) > \deg(g_i)$, regardless of whether the TOP ordering or the POT ordering of monomials is used. It then follows that $\mathrm{ord}(g_j) > \mathrm{ord}(g_i)$, otherwise g_j could be reduced by g_i and this would contradict the fact that G is a minimal Gröbner basis. This proves the main result of the lemma. Since only r values of $\mathrm{ord}(g_i)$ are possible, it also follows that $m \leq qr$. □

As a result of the previous lemma, we can define a sequence of order differences as follows.

Definition 2.1.26 Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$ ordered as $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. For $1 \leq j \leq m$ define

$$\beta_j := \mathrm{ord}(g_j) - \mathrm{ord}(g_i),$$

where i is the smallest integer $> j$ with $\mathrm{lpos}(g_i) = \mathrm{lpos}(g_j)$. If i does not exist, we define $\beta_j := \mathrm{ord}(g_j)$. The sequence $(\beta_1, \ldots, \beta_m) \in \mathbb{N}^m$ is called the **sequence of order differences** of G.

Generalizations of Definition 2.1.15 and Theorem 2.1.17 will be discussed in Chapter 3.

2.2 Non-commutative Gröbner bases

Let us give an introduction to non-commutative Gröbner basis theory, which has been studied by [Chy98, Kre93, Lev05a]. We refer to [Lev05a] to give a brief overview.
Let K be a field. In describing K-algebras via finite sets of generators G and relations R, we write $\mathcal{D} = K\langle G \mid R \rangle$. This means that $\mathcal{D}$ is a factor algebra of the free associative algebra generated by G, modulo the two-sided ideal generated by R. Hence, yet another notation is $\mathcal{D} = K\langle G \rangle / \langle R \rangle$.

Definition 2.2.1 [LS03] Let $\mathcal{D}$ denote a quotient of the free associative K-algebra $K\langle x_1, \ldots, x_n\rangle$ by the two-sided ideal I, generated by the finite set $\{x_j x_i - x_i x_j - d_{ij}\}$ for all $1 \leq i < j \leq n$, where d_{ij} is a polynomial in standard monomials. Then $\mathcal{D}$ is called a G**-algebra**, if the following conditions hold.

1. For all $1 \leq i < j < k \leq n$ the expression $d_{ij}x_k - x_k d_{ij} + x_j d_{ik} - d_{ik}x_j + d_{jk}x_i - x_i d_{jk}$ is contained in I.

2. There exists a monomial ordering $\prec$ on $K[x_1, \ldots, x_n]$, such that $\mathrm{lm}(d_{ij}) \prec x_i x_j$ for each $i < j$. Here, lm stands for the classical notion of leading monomial of a polynomial from $K[x_1, \ldots, x_n]$.

We call an ordering on a G-algebra **admissible**, if it satisfies the second condition of the definition.

Example 2.2.2 Let K be a field of characteristic 0.

- Then
$$\mathrm{W}_1(K) := K\langle x, \partial \mid \partial x = x\partial + 1\rangle$$
is called the first **polynomial Weyl algebra**.

- The first **polynomial difference algebra** is defined by
$$\mathcal{S}_1 := K\langle x, \Delta \mid \Delta x = x\Delta + \Delta + 1\rangle.$$

- Let $\mathbf{q} \neq 0$ be a unit (a parameter) in the ground field. Then
$$\mathrm{W}_1^{\mathbf{q}}(K) := K\langle x, \partial \mid \partial x = \mathbf{q} \cdot x\partial + 1\rangle$$
is called the first **polynomial q-Weyl algebra**.

- The first **polynomial q-difference algebra** is defined by
$$\mathcal{Q} := K\langle x, \partial \mid \mathbf{q} \cdot x\partial + (\mathbf{q} - 1)x\rangle,$$
where $\sigma(p) = p(\mathbf{q}x)$ and $\delta(p) = p(\mathbf{q}x) - p(x)$.

All the algebras discussed in the previous example can be extended from the first to the n-th case, see Example 2.2.5.
In the sequel let $\mathcal{D} = K\langle x_1, \ldots, x_n \mid \{x_j x_i = c_{ij}x_i x_j + d_{ij}\}_{1 \leq i < j \leq n}\rangle$ denote a G-algebra. The **monomials** of $\mathcal{D}^q$ are given via

$$\mathrm{Mon}(\mathcal{D}^q) := \{x^\alpha \, \mathrm{e}_i \mid \alpha \in \mathbb{N}^n, 1 \leq i \leq q\}.$$

We say that $x^\alpha \, \mathrm{e}_i$ divides $x^\beta \, \mathrm{e}_j$ if and only if $i = j$ and $\alpha_k \leq \beta_k$ for all $k = 1, \ldots, n$. We define a module ordering and all corresponding terms as usual, see the commutative case.

Definition 2.2.3 [Lev05a, Definition 1.6] Let M be a left submodule of $\mathcal{D}^q$. Then a finite subset $G \subseteq M\backslash\{0\}$ is called a **left Gröbner basis** of M if and only if for all $m \in M\backslash\{0\}$, there exists an element $g \in G$ such that $\mathrm{lm}(g)$ divides $\mathrm{lm}(m)$.

In other words, G is a Gröbner basis of M if and only if $\mathrm{L}(G) = \mathrm{L}(M)$ holds. We call a subset $G \subseteq \mathcal{D}^q$ **minimal** if $\mathrm{lm}(g) \notin \mathrm{L}(G\backslash\{g\})$ for all $g \in G$. Further we say $g \in \mathcal{D}^q$ is **reduced with respect to** $G \subseteq \mathcal{D}^q$ if no monomial of g is contained in $\mathrm{L}(G)$. And finally, a subset $G \subseteq \mathcal{D}^q$ is called **reduced** if each $g \in G$ is reduced with respect to $G\backslash\{g\}$ and $g - \mathrm{lc}(g)\,\mathrm{lm}(g)$ is reduced with respect to G.
The computer algebra system SINGULAR::PLURAL can compute Gröbner bases for G-algebras. It is well known that Gröbner bases have proved to be very useful to compute objects of interest like for instance syzygies or kernels. In Subsection 2.2.1, we will elaborate this in more detail.

Preliminaries on Ore extensions

Ore extensions build the framework for most of the problems studied in this thesis. A first indication is pointed out in Section 1.4, where one-dimensional time-varying systems are introduced. Further motivation to consider these rings in the context of system and control theory is given for instance in [Zer06a, IM05, INS84, Rob06, CQ05]. Ore extensions are noncommutative rings possessing a σ-derivation and a certain endomorphism to define the commutation of two elements, thus giving the extension from commutative to non-commutative polynomial rings. This kind of rings are used in analyzing the structure of analytic equations, like linear ordinary or partial differential equations or partial shift or difference equations with rational or polynomial coefficients, see Example 2.2.5. Many of the relevant operator algebras have the structure of an Ore algebra, as studied e.g. in [CQR07, CQR05, CS98]. The name is inspired by Øystein Ore, who introduced and studied this kind of rings. Further studies are given for instance in [CS98] and [MR01]. We give a definition that is motivated by [Chy98, CS98]. Moreover, this simplifies the more general setup of [Kre93].

Definition and Remark 2.2.4 [MR01] Let K be a field and A a K-algebra.

1. Further let $\sigma : A \to A$ be a ring endomorphism. Then the map $\delta : A \to A$ is called σ-**derivation**, if δ is K-linear and satisfies the skew Leibniz rule

$$\delta(ab) = \sigma(a)\delta(b) + \delta(a)b \quad \text{for all} \quad a, b \in A.$$

   For a σ-derivation δ, the ring $A[\partial; \sigma, \delta]$ consisting of all polynomials in ∂ with coefficients in A with the usual addition and a product defined by the commutation rule

$$\partial a = \sigma(a)\partial + \delta(a) \quad \text{for all} \quad a \in A$$

   is called a **skew polynomial ring**, or an **Ore extension** of A with ∂ subject to σ, δ.

2. Let $A = K[x_1, \ldots, x_n]$. An iterated skew polynomial ring

$$O = K[x_1, \ldots, x_n][\partial_1; \sigma_1, \delta_1] \ldots [\partial_s; \sigma_s, \delta_s]$$

is called a (polynomial) **Ore algebra** if the σ_i's and δ_j's commute for $1 \leq i, j \leq s$, the ∂_i's commute with ∂_j's, and further for all $1 \leq i \leq s$, the map $\sigma_i : O \to O$ is an injective K-algebra endomorphism and $\delta_i : O \to O$ is a σ_i-derivation satisfying

$$\sigma_i(\partial_j) = \partial_j \quad \text{and} \quad \delta_i(\partial_j) = 0.$$

Using multi-index notation, every element of an Ore algebra can be expressed in the **normal form**

$$\sum_{\alpha \in \mathbb{N}^s} p_\alpha \partial^\alpha = \sum_{\alpha \in \mathbb{N}^s} p_\alpha {\partial_1}^{\alpha_1} \cdots {\partial_s}^{\alpha_s} \quad \text{where} \quad p_\alpha \in A. \tag{2.10}$$

It is easy to see that any nonzero element $a \in A[\partial; \sigma, \delta]$ can be written as $a = a_d \partial^d + \cdots + a_1 \partial + a_0$, where $d \in \mathbb{N}$ and $a_i \in A$ with $a_d \neq 0$. We call d the **degree** of a, sometimes it is also called the **order** of a.
If A is a domain and σ is injective, the Ore extension $A[\partial; \sigma, \delta]$ is a domain by degree arguments. Then the definition can be iterated and the resulting ring is called **Ore algebra**.
In the next example, we enlist some interesting Ore algebras. These rings are of great interest in applications.

Example 2.2.5

- Let $A = K[x_1, \ldots, x_n]$ for a field K. Further let $\sigma_i := \mathrm{id}_A$ and $\delta_i := \frac{\partial}{\partial x_i}$ for all $1 \leq i \leq n$. Then

  $$\mathrm{W}_n(K) := K[x_1, \ldots, x_n][\partial_1; \sigma_1, \delta_1] \ldots [\partial_n; \sigma_n, \delta_n]$$

  is called the n-th **polynomial Weyl algebra**. We get commutation rule

  $$\partial_i x_j = \begin{cases} x_j \partial_i & \text{if } i \neq j \\ x_i \partial_i + 1 & \text{else.} \end{cases}$$

- Let $A = K[x_1, \ldots, x_n]$ for a field K. Further let $(\sigma_i p)(x) = p(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_n)$ and $\delta_i(p) = \sigma_i(p) - p$ for all $p \in A$. Then

  $$\mathcal{S}_n := A[\Delta_1; \sigma_1, \delta_1] \ldots [\Delta_n; \sigma_n, \delta_n]$$

  is called the n-th **polynomial difference algebra**. We get the commutation rule

  $$\Delta_i x_j = \begin{cases} x_j \Delta_i & \text{if } i \neq j \\ x_i \Delta_i + \Delta_i + 1 & \text{else.} \end{cases}$$

- Let $A = K[x_1, \ldots, x_n]$ for a field K. The following Ore algebra is a combination of the first and second one. Define

  $$\mathcal{SW}_n := A\left[\Delta; \mathbf{s}_1, \mathbf{d}_1\right]\left[\partial; \sigma_1, \delta_1\right] \ldots \left[\Delta; \mathbf{s}_n, \mathbf{d}_n\right]\left[\partial; \sigma_n, \delta_n\right],$$

  where $\sigma_i := \mathrm{id}_{\mathcal{SW}_n}$, $\delta_i := \frac{\partial}{\partial t_i}$ and $(\mathbf{s}_1 p)(x) = p(x+1)$, $\mathbf{d}_i(p) = \mathbf{s}_i(p) - p$ for all $p \in \mathcal{SW}_i$. We get the commutation rules

  $$\partial_i x_j = \begin{cases} x_j \partial_i & \text{if } i \neq j \\ x_i \partial_i + 1 & \text{else} \end{cases} \quad \text{and} \quad \Delta_i x_j = \begin{cases} x_j \Delta_i & \text{if } i \neq j \\ x_i \Delta_i + \Delta_i + 1 & \text{else.} \end{cases}$$

- Let $A = K(x)$ and let $\mathbf{q} \neq 0$ be a unit (a parameter) in the ground field. Set $\sigma(p(x)) = p(\mathbf{q}x)$ and $\delta := \frac{\partial}{\partial x}$. Then

$$W_1^{\mathbf{q}}(K) := A[\partial; \sigma, \frac{\partial}{\partial x}]$$

  is called the first **rational q-Weyl algebra**. We get the commutation rule

$$\partial x = \mathbf{q}x\partial + 1.$$

- Let $A = K(x)$ and let $\mathbf{q} \neq 0$ be a unit (a parameter) in the ground field. The first **continuous q-difference algebra** is defined by

$$\mathcal{Q} := A[\partial; \sigma, \delta],$$

  where $\sigma(p) = p(\mathbf{q}x)$ and $\delta(p) = p(\mathbf{q}x) - p(x)$. We get the commutation rule

$$\partial x = \mathbf{q}x\partial + \mathbf{q}x - x.$$

All these algebras are G-algebras. Thus the proposed theory of Gröbner bases can be applied. Here, it should be stressed that the theory of Gröbner bases over certain polynomial Ore extensions is studied in [CS98]. More precisely, Ore extensions $K[x_1, \ldots, x_n][\partial_1; \sigma_1, \delta_1] \ldots [\partial_n; \sigma_n, \delta_n]$ which satisfy the commutation rule

$$\partial_i x_j = a_{ij} x_j \partial_i + c_{ij}(x),$$

where $a_{ij} \in K \backslash \{0\}$ and $c_{ij}(x) \in K[x_1, \ldots, x_n]$, are tackled in that publication.

Let O be an Ore algebra. We call $a \in O$ a **left divisor** (or just a **divisor**, if no confusion arises) of $b \in O$ if and only if there exists $f \in O$ such that $af = b$ and write $a|b$ shortly. Analogously we define a to be a **right divisor** (or just a **divisor**, if no confusion arises) of b if and only if there exists an element $f \in O$ such that $fa = b$. Again we write $a|b$.

Theorem 2.2.6 *[BGTV03] Let A be a division ring, $\sigma : A \to A$ an endomorphism and $\mathcal{D} = A[\partial; \sigma, \delta]$ an Ore extension with a σ-derivation δ.*

- *(PID) $\mathcal{D}$ is a left principal ideal domain. If σ is an automorphism, then $\mathcal{D}$ is also a right principal ideal domain.*
- *(Bezout's Theorem) For any nonzero $a, b \in \mathcal{D}$ there exists a right greatest common divisor g_r of a, b and there exist $s, t \in \mathcal{D}$, such that $g_r = sa + tb$. If σ is an automorphism, then for any nonzero $a, b \in \mathcal{D}$ there exists the left greatest common divisor g_ℓ of a, b and there exist $s', t' \in \mathcal{D}$, such that $g_\ell = as' + bt'$.*
- *(ED) $\mathcal{D}$ is a right Euclidean domain. If σ is an automorphism, then $\mathcal{D}$ is also a left Euclidean domain.*

Hence, when σ is bijective, there are left and right Euclidean division algorithms. All Ore extensions discussed in Example 2.2.5 possess a bijective σ.

Remark 2.2.7 [Coh71, Chapter 5, Proposition 1.2] If K is of characteristic zero, the first polynomial Weyl algebra is simple, that is, every two-sided ideal equals 0 or W_1.

Note that the previous remark yields, in case K is of characteristic zero, that the first rational Weyl algebra is simple as well. We want to stress that the characteristic restriction on K is crucial for the result. Consider for instance the ideal $\langle x^2 \rangle \subseteq \mathbb{F}_2[x][\partial; \mathrm{id}_{\mathbb{F}_2[x]}, \frac{d}{dx}]$. Since then

$$x^2\partial = \partial x^2,$$

it is easy to see that 1 is not contained in the ideal generated by x^2.

Ore localizations of G-algebras

This subsection addresses the question how to construct a certain Ore extension from a G-algebra. This enables us to extend the framework of Chapter 4 to a larger class of algebras in which the proposed Algorithms 2 and 3 can be implemented.

We propose a new class of univariate skew polynomial rings, which are obtained as Ore localizations of G-algebras. For this purpose, we need to show how localization in the non-commutative framework works. Note that the crossover from the commutative to the non-commutative world can cause complications. However, we choose a suitable setting to avoid these.

Definition 2.2.8 [MR01] Let S be a multiplicatively closed set in a Noetherian domain $\mathcal{D}$, such that $0 \notin S$.

- The set S is called an **Ore set** in $\mathcal{D}$, if for all $s_1 \in S, a_1 \in \mathcal{D}$ there exist $s_2 \in S, a_2 \in \mathcal{D}$, such that $s_2 a_1 = a_2 s_1$.
- One defines a **ring of fractions** or an **Ore localization** of $\mathcal{D}$ with respect to S to be a ring $\mathcal{D}_S$ (often denoted as $S^{-1}\mathcal{D}$) together with an injective homomorphism $\phi : \mathcal{D} \to \mathcal{D}_S$, such that
  (i) for all $s \in S$, $\phi(s)$ is a unit in $\mathcal{D}_S$,
  (ii) for all $f \in \mathcal{D}_S$, $f = \phi(s)^{-1}\phi(a)$ for some $a \in \mathcal{D}, s \in S$.

Remark 2.2.9 Let W_n be the n-th polynomial Weyl algebra. Then defining $S := K[x_1, \ldots, x_n]\backslash\{0\}$ and ϕ as the natural embedding yields the rational n-th Weyl algebra. Note that all rational counterparts of the algebras given in Example 2.2.5 can be constructed like this.

The Ore property of S in $\mathcal{D}$ guarantees that any left-sided fraction can be written as a right-sided fraction. However, this manipulation is not unique.

Due to [LS03] a G-algebra $\mathcal{D}$ is a Noetherian domain. Hence, there exists its total two-sided ring of fractions $\mathrm{Quot}(\mathcal{D}) = \mathcal{D}_{\mathcal{D}\backslash\{0\}}$, which is a division ring. Assume that $\mathcal{D}$

is generated by the variables $x_1, \ldots, x_{n+1}$. Let $\lambda = \{1, \ldots, n\}$, $I_\lambda = \{x_j x_i - x_i x_j - d_{ij} \mid i, j \in \lambda,\ i < j\}$ and suppose $B = K\langle x_1, \ldots, x_n \mid I_\lambda\rangle$ to be a G-algebra. Moreover, define $x_{n+1} =: \partial$ and $B^* := B \setminus \{0\}$.

Theorem 2.2.10 *Suppose there exists an admissible monomial ordering $\prec$ on $\mathcal{D}$, satisfying $x_k \prec \partial$ for all $1 \leq k \leq n$. Using the introduced notation, B^* is a multiplicatively closed Ore set in $\mathcal{D}$. Hence, there exists a Ore localization of $\mathcal{D}$ with respect to B^*. Moreover, it can be presented as an Ore extension of* $\mathrm{Quot}(B)$ *by the variable ∂.*

Proof: Since B is a G-algebra, it is a domain. Hence, B^* is multiplicatively closed and does not contain zero. Since $\mathcal{D}$ and B are G-algebras and $\prec$ is an admissible ordering, the relation $\partial x_j = c_j x_j \partial + d_j$ where $c_j \in K^*$ and $d_j \in \mathcal{D}$ satisfies $\mathrm{lm}(d_j) \prec x_j\partial$. Further $x_j \prec \partial$ yields $x_j\partial \prec \partial^2$ and hence d_j is at most linear in ∂. Therefore we can write $d_j = a_j \cdot \partial + b_j$ for $a_j, b_j \in B$. Thus we obtain a relation $\partial x_j = c_j'\partial + b_j$, where $c_j' = c_j x_j + a_j$ and $x_j, c_j', b_j \in B$.
By defining $\sigma(x_j) = c_j x_j + a_j$ and $\delta(x_j) = b_j$ for all $1 \leq j \leq n$, we see that σ is an automorphism of $\mathrm{Quot}(B)$. Thus an Ore extension $\mathrm{Quot}(B)[\partial; \sigma, \delta]$ is indeed another presentation of $\mathcal{D}_{B^*}$ as soon as B^* is an Ore set in $\mathcal{D}$. Thus let us finally show that B^* is an Ore set in $\mathcal{D}$. Using $\mathrm{lm}(d_j) = \mathrm{lm}(a_j\partial + b_j) \prec x_j\partial$ yields on the one hand $\mathrm{lm}(a_j) \prec x_j$ and on the other hand $\mathrm{lm}(b_j) \prec x_j\partial$. The latter implies that there exist positive weights ω and $w_1, \ldots, w_n$ for the variables $\{\partial, x_1, \ldots, x_n\}$, such that for $\mathrm{lm}(a_j)x^\alpha$ and $\mathrm{lm}(b_j) = x^\alpha$ one has $\sum_i w_i\alpha_i \leq x_j$ and $\sum_i w_i\beta_i \leq w_j + \omega$. In particular, this can be achieved by setting ω large enough. Then we follow the recipe from [BGTV03] and construct a block ordering from this setting. Consider an ordering $\prec_\partial$ on $\mathcal{D}$, which is a block ordering for blocks of variables $\{\partial\}, \{x_1, \ldots, x_n\}$. This means that $\partial \gg x_j$ for all j, that is, the variable ∂ is greater than any power of x_j. The second block is an ordering $\prec_B$ on B, for which $\mathrm{lm}(a_j) \prec_B x_j$ holds. For instance, one can take $\prec_B$ to be the restriction of $\prec$ to B. Then $\mathrm{lm}(d_j) = \max_{\prec_\partial}(a_j\partial, b_j) \prec_\partial x_j\partial$ holds, hence $\prec_\partial$ is an admissible ordering on $\mathcal{D}$. From Proposition 28 of [GML] (which holds for a much more general situation), the existence of such a block ordering as $\prec_\partial$ implies that the set B^* is an Ore set in $\mathcal{D}$. □

Remark 2.2.11 Note that by construction $\mathcal{D}_{B^*}$ is a Euclidean (principal ideal) domain by Theorem 2.2.6. In particular, all but one variables are invertible.

Example 2.2.12 To illustrate Theorem 2.2.10, we consider the difference algebra $\mathcal{S}_1 := K\langle x, \Delta \mid \Delta x = x\Delta + \Delta + 1\rangle$. Since $\Delta \prec x\Delta$ is a consequence of $1 \prec x$ (we assume we are dealing with well-orderings only), $\mathcal{S}_1$ can be localized at both $K[x]^*$ and $K[\Delta]^*$. On the other hand, the algebra associated with the operator of partial integration $\mathcal{I}_1 := K\langle x, I \mid Ix = xI - I^2\rangle$ can be localized only at $K[I]^*$ but not at $K[x]^*$, since $I^2 \prec xI$ is a consequence of $I \prec x$, and any ordering satisfying $x \prec I$ is not admissible for $\mathcal{I}_1$.

2.2.1 Algorithmic computations

For the concrete calculations needed in Chapter 5, we need algorithms for the following computational tasks over (polynomial) Ore algebras:

1. syzygy module of a tuple of vectors
2. elimination of module components from a submodule of a free module
3. annihilator ideal of an element in a finitely presented module
4. kernel of a homomorphism of modules
5. intersection of a finite number of submodules of a free module.

In the sequel let O be a Noetherian Ore algebra. Note that we have the following sufficient condition:

Remark 2.2.13 [MR01, Theorem 1.2.9.] Let A be a Noetherian K-algebra for a field K. Then $\mathcal{D}$ is Noetherian if σ_i is an automorphism for all $1 \leq i \leq s$ on A. (Thus all Ore algebras considered in Example 2.2.5 are Noetherian.)

Let M be a finitely presented left O-module, that is, there exists a matrix $P \in O^{m\times n}$ such that there is the following exact sequence of left O-modules:

$$O^{1\times m} \xrightarrow{P} O^{1\times n} \to M \to 0.$$

Recall that for a tuple $F = (f_1, \ldots, f_s), f_i \in O^{1\times n}$, the set $\mathrm{Syz}(F) := \{[a_1, \ldots, a_s] \in O^{1\times s} \mid \sum_i a_i f_i = 0\}$ carries the structure of a left O-module and is called the **left syzygy module** of F. Since O is Noetherian, $\mathrm{Syz}(F)$ is finitely generated. Computation of syzygies over Noetherian Ore algebras can be accomplished with several algorithms and requires Gröbner basis techniques; see [Kre93] for Ore algebras and [GPS05] for the commutative case.

Let $\{\mathrm{e}_i\}$ be the canonical basis of the free module $O^{1\times\ell} = \bigoplus_{i=1}^{\ell} O\,\mathrm{e}_i$.

Theorem 2.2.14

1. *"Elimination of module components":*
   Let $S \subset O^{1\times\ell}$ be a submodule and G be a Gröbner basis of S with respect to the POT ordering. Then the intersection $G \cap \bigoplus_{i=k}^{\ell} O\,\mathrm{e}_i$ is a Gröbner basis of $S \cap \bigoplus_{i=k}^{\ell} O\,\mathrm{e}_i$ for all $1 \leq k < \ell$.
2. *"Kernel of a module homomorphism of modules":*
   Consider an O-module homomorphism $O^{1\times s} \xrightarrow{\psi} O^{1\times n}/O^{1\times m}P$, $\mathrm{e}_i \mapsto [\Psi_i]$, where $\Psi_i \in O^{1\times n}$. Let P_i be the i-th row of the matrix P. Then

$$\ker\psi = \mathrm{Syz}(\ (\Psi_1, \ldots, \Psi_s, P_1, \ldots, P_m)\ ) \cap \bigoplus_{i=1}^{s} O\,\mathrm{e}_i\,.$$

Proof:

1. Define $W = \bigoplus_{i=k}^{\ell} O\,\mathrm{e}_i$. Since G is a Gröbner basis of S, for any $0 \neq s \in S$ there exists $g \in G$ such that $\mathrm{lm}(g)$ divides $\mathrm{lm}(s)$. If $s \in S \cap W$, then $\mathrm{lm}(g) \in W$ and hence, by definition of the underlying ordering, we have $g \in W$ and $g \in G \cap W$. So, $G \cap W$ is a Gröbner basis of $S \cap W$.

2. We have

$$[b_1, \ldots, b_s] \in \ker \psi \;\Leftrightarrow\; \exists a_k \in O \;:\; \sum_{i=1}^{s} b_i \Psi_i + \sum_{k=1}^{m} a_k P_k = 0$$

$$\Leftrightarrow [b_1, \ldots, b_s] \in \mathrm{Syz}(\ (\Psi_1, \ldots, \Psi_s, P_1, \ldots, P_m)\ ) \cap \bigoplus_{i=1}^{s} O\,\mathrm{e}_i\,.$$

□

Corollary 2.2.15

1. *"Annihilator of a module element":*
   Let $M = O^{1\times n}/O^{1\times m}P$ and let $P_1, \ldots, P_m$ denote the rows of P. Moreover, let $v \in O^{1\times n}$. Then the left ideal $\mathrm{ann}_M^O(v) := \{a \in O \mid a[v] = 0 \in M\} \subseteq O$ can be computed as

$$\mathrm{ann}_M^O(v) = \ker(O \xrightarrow{\cdot[v]} M) = \mathrm{Syz}(\ (v, P_1, \ldots, P_m)\ ) \cap O\,\mathrm{e}_1\,.$$

2. *"Intersection of finitely many submodules":*
   Let $N_1, \ldots, N_m \subset O^{1\times r}$ be submodules. Then

$$\bigcap_{i=1}^{m} N_i = \ker\big(O^{1\times r} \to (O^{1\times r}/N_1) \oplus \cdots \oplus (O^{1\times r}/N_m),\ \ \mathrm{e}_i \mapsto ([\mathrm{e}_i], \ldots, [\mathrm{e}_i])\big).$$

Remark 2.2.16 For an O-module homomorphism $O^{1\times s}/O^{1\times r}Q \xrightarrow{\psi'} O^{1\times n}/O^{1\times m}P$, its kernel is the image of $\ker\psi$ (as in Theorem 2.2.14) under the natural projection $O^{1\times s} \to O^{1\times s}/O^{1\times r}Q$. A left Gröbner basis can be obtained by reducing a left Gröbner basis of $\ker\psi + O^{1\times r}Q$ with a left Gröbner basis of $O^{1\times r}Q$, see [Lev05b].
The algorithms we have discussed are implemented in computer algebra systems like e.g. SINGULAR::PLURAL [GLH05] or MAPLE [CQR07, CS98] with the package OREMODULES. More background on these algorithms can be found in e.g. [Kre93], [Lev05b].

Chapter 3

One-dimensional systems over finite rings

In this Chapter, we address the question how to extend the definition of the so-called Predictable Leading Monomial (PLM) property given in Definition 2.1.15 to the finite ring case. Note that the question of the Predictable Degree (PD) property was studied in [KPP07] in a completely Gröbner bases free context. The results presented here rely on minimal Gröbner bases and are more general, because the PLM property implies the PD property. We use the framework of Gröbner bases, because it is a natural extension of Definition 2.1.15 and further, we can revert to computer algebra packages. An explicit motivation to analyze the PLM property for the ring case is given in Section 1.3. Note that the results of this Chapter can be found in [KSa, KSb]. Let us recall the crucial point about the predictable degree property given in the field case. A set of vectors is said to possess the predictable degree property if there is no cancellation of leading monomials in every possible linear combination of these elements. Since there exist zero-divisors in $\mathbb{Z}_{p^r}$, such a result does not hold there in general.

Example 3.0.17 Consider the submodule $M = \langle g_1, g_2 \rangle = \langle x^3, 9x \rangle$ of $\mathbb{Z}_{27}[x]$. Then the generators of M are already a minimal Gröbner basis, and it is easy to see that $12x^4 \in M$. We have the identity $12x^4 = 12xg_1 = 3xg_1 + x^3g_2 = 12xg_1 + 3x^5g_2$. But then

$$x^4 = \mathrm{lm}(12x^4) \neq \max\{\mathrm{lm}(12x)\,\mathrm{lm}(g_1), \mathrm{lm}(3x^5)\,\mathrm{lm}(g_2)\} = x^6.$$

So we have seen that the PLM property does not hold.

We will see in the sequel how to restrict the coefficients of the linear combination to obtain unique representations. Let $\mathcal{D}$ denote the ring $\mathbb{Z}_{p^r}[x]$ throughout this chapter. Further the set of digits will be denoted by $\mathrm{A}_p = \{0, 1, \ldots, p-1\} \subset \mathbb{Z}_{p^r}$ as introduced in Section 2.1.1. Recall that according to Remark 2.1.24, the leading monomials of a minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$ can always be ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. The next lemma shows how to restrict the coefficients of $\mathcal{D}$ to get a unique representation of the leading term of any element contained in the span of a minimal Gröbner basis.

Lemma 3.0.18 *Let M be a submodule of $\mathcal{D}^q$ with the minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G as per Definition 2.1.26. Assume that G is normalized so that $\mathrm{lc}(g_i) = p^{r-\mathrm{ord}(g_i)}$ for $i = 1, \ldots, m$. Let $0 \neq f \in M$ and j_1 be the largest integer such that $\mathrm{ord}(g_{j_1}) \geq \mathrm{ord}(f)$, $\mathrm{lpos}(g_{j_1}) = \mathrm{lpos}(f)$ and $\deg(g_{j_1}) \leq \deg(f)$. Let $g_{j_2}, \ldots, g_{j_s}$ be all successors of g_{j_1} for which $\mathrm{lpos}(g_{j_i}) = \mathrm{lpos}(f)$ for $i = 2, \ldots, s$. Then $\mathrm{lt}(f)$ can be uniquely written as*

$$\mathrm{lt}(f) = c_{j_1} x^{\alpha_{j_1}} \mathrm{lt}(g_{j_1}) + \cdots + c_{j_s} x^{\alpha_{j_s}} \mathrm{lt}(g_{j_s}), \tag{3.1}$$

where

$$\mathrm{lm}(f) = x^{\alpha_{j_i}} \mathrm{lm}(g_{j_i}),$$

with $\alpha_{j_i} \in \mathbb{N}$ and $c_{j_i} \in \mathrm{A}_p + \cdots + p^{\beta_{j_i}-1}\mathrm{A}_p$ for $i = 1, \ldots, s$.

Proof: Write

$$\mathrm{lc}(f) = p^{\ell}\theta_{\ell} + p^{\ell+1}\theta_{\ell+1} + \cdots + p^{r-1}\theta_{r-1}, \tag{3.2}$$

where $\ell = r - \mathrm{ord}(f)$, $\theta_j \in \mathrm{A}_p$ for $j = \ell, \ldots, r-1$ and $\theta_\ell \neq 0$. Note that the θ_j's are unique. It follows from Observation 2.1.8 that a largest integer j_1 exists such that $\mathrm{ord}(g_{j_1}) \geq \mathrm{ord}(f)$, $\mathrm{lpos}(g_{j_1}) = \mathrm{lpos}(f)$ and $\deg(g_{j_1}) \leq \deg(f)$. Let $g_{j_2}, \ldots, g_{j_s}$ be all successors of g_{j_1} for which $\mathrm{lpos}(g_{j_i}) = \mathrm{lpos}(f)$ for $i = 2, \ldots, s$. Then for $i = 1, \ldots, s$, $\alpha_{j_i} := \deg(f) - \deg(g_{j_i})$ is well-defined by Lemma 2.1.25. Evidently $\mathrm{lm}(f) = x^{\alpha_{j_i}} \mathrm{lm}(g_{j_i})$ for $i = 1, \ldots, s$. Next, let us define

$$d := \mathrm{ord}(g_{j_1}) - \mathrm{ord}(f).$$

Note that, by definition of j_1 and Lemma 2.1.25, we have $d < \beta_{j_1}$. Further define

$$c_{j_1} := p^d\theta_\ell + p^{d+1}\theta_{\ell+1} + \cdots + p^{\beta_{j_1}-1}\theta_{\ell+\beta_{j_1}-1-d},$$

and, for $i = 2, \ldots, s$,

$$\begin{aligned} c_{j_i} = \quad & \theta_{\ell+\sum_{k=1}^{i-1}\beta_{j_k}-d} + p\theta_{\ell+\sum_{k=1}^{i-1}\beta_{j_k}-d+1} \\ & + \cdots + p^{\beta_{j_i}-1}\theta_{\ell+\sum_{k=1}^{i-1}\beta_{j_k}-d+\beta_{j_i}-1}. \end{aligned}$$

Then $c_{j_1} \neq 0$ since $\theta_\ell \neq 0$. Clearly $c_{j_i} \in \mathrm{A}_p + \cdots + p^{\beta_{j_i}-1}\mathrm{A}_p$ and (3.1) holds. The uniqueness of the representation follows from the uniqueness of the θ_j's in (3.2) and the fact that, by definition

$$\beta_{j_i} - 1 + r - \mathrm{ord}(g_{j_i}) \leq r - 1$$

for $i = 1, \ldots, s$. □

Example 3.0.19 Let us return to Example 3.0.17. By definition, the sequence of order differences (β_1, β_2) equals $(2, 1)$. Thus we restrict the coefficients of x^3 to $\mathrm{A}_p + p\,\mathrm{A}_p$ and the coefficients of $9x$ to A_p. Then $12x^4 = 3x \cdot x^3 + x^3 \cdot 9x$ is the desired unique representation.

The next theorem will extend the idea proposed by the previous lemma and show that a minimal Gröbner basis of a module in $\mathcal{D}^q$ has a particular type of Predictable Leading Monomial property (compare Definition 3.1.11 and Theorem 3.2.3).

Theorem 3.0.20 *Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G as per Definition 2.1.26. Then any $f \in M$ is uniquely represented as*

$$f = h_1 g_1 + \cdots + h_m g_m, \tag{3.3}$$

with h_j restricted to the subset $\mathrm{A}_p[x] + p\,\mathrm{A}_p[x] + \cdots + p^{\beta_j - 1}\,\mathrm{A}_p[x]$ of $\mathbb{Z}_{p^r}[x]$. Furthermore, for $f \neq 0$ we have

$$\mathrm{lm}(f) = \max_{1 \leq i \leq m; h_i \neq 0} (\mathrm{lm}(h_i)\,\mathrm{lm}(g_i)). \tag{3.4}$$

Proof: Let

$$r_1 := f - \sum_{l=1}^{s_1} c^1_{j^1_l} x^{\alpha^1_{j^1_l}} g_{j^1_l},$$

where j^1_i, $c^1_{j^1_i}$ and $\alpha^1_{j^1_i}$ are as in the above lemma for $i = 1, \ldots, s_1$. Note that $c^1_{j^1_1} \neq 0$. Then, by Lemma 3.0.18, $f \xrightarrow{G} r_1$, so that, by Lemma 2.1.10, $\mathrm{lm}(r_1) < \mathrm{lm}(f)$. Now repeat this step, defining at each step k

$$f^k_{j^k_\ell} := c^k_{j^k_\ell} x^{\alpha^k_{j^k_\ell}} \quad \text{for } \ell = 1, \ldots, s_k \text{ and}$$

$$r_k := r_{k-1} - \sum_{l=1}^{s_k} f^k_{j^k_l} g_{j^k_l}.$$

Since $\mathrm{lm}(r_k) < \mathrm{lm}(r_{k-1})$, this reduction procedure must stop at $r_t = 0$ for some $t \in \mathbb{N}$. Next, for $i = 1, \ldots, m$ list all steps at which g_i is used as $k_1, \ldots, k_{N_i}$, where $N_i \in \mathbb{N}$. Now define

$$h_i := \sum_{j=1}^{N_i} f^{k_j}_i.$$

Then evidently $f = h_1 g_1 + \cdots + h_m g_m$. In order to show that $h_i \in \mathrm{A}_p[x] + \cdots + p^{\beta_i - 1}\,\mathrm{A}_p[x]$ it is clearly sufficient to prove that $\deg(f^{k_{j+1}}_i) < \deg(f^{k_j}_i)$. For this, it follows from $\mathrm{lm}(r_{k_j - 1}) > \mathrm{lm}(r_{k_j})$ that $\deg(r_{k_j - 1}) > \deg(r_{k_j})$, so that

$$\begin{aligned} \deg(f^{k_{j+1}}_i) &= \deg(r_{k_{j+1}-1}) - \deg(g_i) \\ &< \deg(r_{k_j - 1}) - \deg(g_i) \\ &= \deg(f^{k_j}_i). \end{aligned}$$

This shows the existence of the representation. Its uniqueness follows from the uniqueness of the j^k_ℓ's, the c^k_i's and the α^k_i's by Lemma 3.0.18. Finally, from $\mathrm{lm}(r_k) < \mathrm{lm}(r_{k-1})$ and $c^1_{j^1} \neq 0$ it follows that

$$\mathrm{lm}(f) = \mathrm{lm}(\sum_{l=1}^{s_1} f^1_{j^1_l} g_{j^1_l}) > \mathrm{lm}(f^k_{j^k_l} g_{j^k_l}) \qquad \text{for all } j^k_l \text{ with } 2 \leq k.$$

As a result, $\mathrm{lm}(f) = \mathrm{lm}(f^1_{j^1_1})\,\mathrm{lm}(g_{j^1_1})$, so that (3.4) holds. □

Example 3.0.21 Let M be a submodule of $\mathbb{Z}_9^2[x]$ generated by the rows of the following matrix

$$R = \begin{bmatrix} 1 & 8x^5+5x^4+5x^3+2x^2+2x \\ 0 & x^6 \\ 3 & 6x^5+6x^4+6x^3+6x^2+6x \\ 0 & 3x^6 \end{bmatrix}.$$

Denote the rows of R by R_1, R_2, R_3 and R_4. Note that $R_3 = 3R_1$ and $R_4 = 3R_2$.

- Using the TOP ordering:
  a minimal Gröbner basis $G = \{g_1, \ldots, g_4\}$ of M is given by the rows of

$$\begin{bmatrix} 8 & \underline{x^5}+4x^4+4x^3+7x^2+7x \\ x+5 & \underline{3x^4}+3x^2+x \\ \underline{x^2}+3x+2 & x^2+4x \\ \underline{3x}+6 & 3x \end{bmatrix}.$$

  The sequence of order differences $(\beta_1, \beta_2, \beta_3, \beta_4)$ equals $(1, 1, 1, 1)$. Thus each $f \in M$ can be written uniquely as $h_1g_1 + h_2g_2 + h_3g_3 + h_4g_4$, where $h_i \in \mathrm{A}_p[x]$.

- Using the POT ordering:
  in this case, the vectors R_1 and R_2 form a minimal Gröbner basis. The sequence of order differences (β_1, β_2) equals $(2, 2)$. Thus $f \in M$ can be written uniquely as $h_1R_1 + h_2R_2$, where $h_i \in \mathrm{A}_p[x] + p\,\mathrm{A}_p[x]$.

As the previous theorem shows, the restriction of the coefficients depends on the minimal Gröbner basis or more precisely, on the sequence of order differences of the minimal Gröbner basis. In order to derive a more general notion which is compatible to the results of [KPP07], we introduce the so-called minimal Gröbner p-basis. But this requires some preparation.

3.1 Preliminaries on p-generator sequences

The concepts presented below are extensions of [VSR96], first presented in [KPP07]. The crucial idea behind the whole framework is to get rid of zero-divisors. Thus first the multiplicative variables are restricted to A_p. The resulting structure varies from the original one. But to deal with equal sets, certain conditions will be required from the module generators.

Definition 3.1.1 [KPP07] Let $\{v_1, \ldots, v_N\} \subset \mathcal{D}^q$. A ***p*-linear combination** of $v_1, \ldots, v_N$ is a vector $\sum_{j=1}^{N} a_j v_j$, where $a_j \in \mathcal{D}$ is a polynomial with coefficients in A_p for $j = 1, \ldots, N$. Furthermore, the set of all p-linear combinations of $v_1, \ldots, v_N$ is denoted by ***p*-span**$(v_1, \ldots, v_N)$, whereas the set of all linear combinations of $v_1, \ldots, v_N$ with coefficients in $\mathcal{D}$ is denoted by $\mathrm{span}(v_1, \ldots, v_N)$.

Remark 3.1.2 Note that the p-span and span do not coincide. This is obvious: Suppose $r \geq 2$. Since $px \notin p\text{-span}(x)$, it follows that $\text{span}(x) \neq p\text{-span}(x)$. But there exist specific requirements on the generators such that their span and p-span coincide. The subset needs to be a so-called p-generator sequence.

Definition 3.1.3 [KPP07] An ordered sequence $(v_1, \ldots, v_N)$ of vectors in $\mathcal{D}^q$ is said to be a ***p*-generator sequence** if $p\,v_N = 0$ and $p\,v_i$ is a p-linear combination of $v_{i+1}, \ldots, v_N$ for $i = 1, \ldots, N-1$.

Theorem 3.1.4 [KPP07] *Let $v_1, \ldots, v_N \in \mathcal{D}^q$. If $(v_1, \ldots, v_N)$ is a p-generator sequence, then*

$$p\text{-span}(v_1, \ldots, v_N) = \text{span}(v_1, \ldots, v_N).$$

In particular, p-span$(v_1, \ldots, v_N)$ is a submodule of $\mathcal{D}^q$.

All submodules of $\mathcal{D}^q$ can be written as the p-span of a p-generator sequence. In fact, if $M = \text{span}(g_1, \ldots, g_m)$, then M is the p-span of the p-generator sequence $(g_1, pg_1, \ldots, p^{r-1}g_1, \ldots, g_m, pg_m, \ldots, p^{r-1}g_m)$.
But the restriction to A_p even allows to set up the notion of a specific basis.

Definition 3.1.5 [KPP07] The vectors $v_1, \ldots, v_N \in \mathcal{D}^q$ are said to be ***p*-linearly independent** if the only p-linear combination of $v_1, \ldots, v_N$ that equals zero is the trivial one.

Definition 3.1.6 Let M be a submodule of $\mathcal{D}^q$, written as a p-span of a p-generator sequence $(v_1, \ldots, v_N)$. Then $(v_1, \ldots, v_N)$ is called a ***p*-basis** of M if the vectors $v_1, \ldots, v_N$ are p-linearly independent in $\mathcal{D}^q$.

Lemma 3.1.7 [KPP07] *Let M be a submodule of $\mathcal{D}^q$ and let $(v_1, v_2, \ldots, v_N)$ be a p-basis of M. Then each vector of M is written in a unique way as a p-linear combination of $v_1, \ldots, v_N$.*

Suppose M to be a submodule of $\mathbb{Z}_{p^r}^q$ with a p-basis $(v_1, \ldots, v_N)$. Then in [VSR96], the so-called p-dimension of M is defined as $p\text{-dim}(M) = N$. The adaption to $\mathcal{D}^q$ requires an additional constraint.

Definition 3.1.8 [KPP07] The **row degree** of a nonzero polynomial vector $v \in \mathcal{D}^q$ is defined as the highest degree of its nonzero components in $\mathcal{D}^q$. It is denoted by $\text{rowdeg}(v)$. The coefficient vector in $\mathbb{Z}_{p^r}^q$ of the term $x^{\text{rowdeg}(v)}$ in v is called the **leading row coefficient vector** of v and denoted by v^{lrc}, that is,

$$v = v^{\text{lrc}} x^{\text{rowdeg}(v)} + \text{ lower order terms.}$$

Definition 3.1.9 Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with p-basis $(v_1, \ldots, v_N)$. Then the sequence $(v_1, \ldots, v_N)$ is called a **reduced *p*-basis** of M if the leading coefficient vectors $v_1^{\mathrm{lrc}}, \ldots, v_N^{\mathrm{lrc}}$ are p-linearly independent in $\mathbb{Z}_{p^r}^q$.

By [KPP07, Theorem 3.12], the following Algorithm terminates with the desired result.

Input : Module M spanned by $(v_1, \ldots, v_N)$
Output: A reduced p-basis $(w_1, \ldots, w_k)$ of M

$$W \leftarrow (v_1, pv_1, \ldots, p^{r-1}v_1, \ldots, v_N, \ldots, p^{r-1}v_N)$$

Step 1: Remove zero vectors in W, resulting in

$$W \leftarrow (w_1, \ldots, w_k).$$

Step 2: Re-order W according to nonincreasing degree such that

$$W \leftarrow (w_1, \ldots, w_k),$$

making sure that vectors of equal degree are not swapped.
Step 3: Determine the smallest ℓ such that

1. $(w_{\ell+1}^{\mathrm{lrc}}, \ldots, w_k^{\mathrm{lrc}})$ is a p-generator sequence
2. $p\text{-dim}(\mathrm{span}(w_{\ell+1}^{\mathrm{lrc}}, \ldots, w_k^{\mathrm{lrc}}))) = k - \ell$

If $\ell = 0$ **return** $(w_1, \ldots, w_k)$ else go to Step 4.
Step 4: For $i = 1, \ldots, k-\ell$, let $\alpha_i \in \mathbb{Z}_{p^r}$ be such that

$$\mathrm{lc}(w_\ell) + \alpha_1 \mathrm{lc}(w_{\ell+1}) + \cdots + \alpha_{k-\ell} \mathrm{lc}(w_k) = 0.$$

Replace w_ℓ by

$$w_\ell + \alpha_1 x^{\deg(w_\ell)-\deg(w_{\ell+1})} w_{\ell+1} + \cdots + \alpha_{k-\ell} x^{\deg(w_\ell)-\deg(w_k)} w_k.$$

Go to Step 1.

Algorithm 1: Reduced p-basis Algorithm [KPP07, Algorithm 3.11]

According to [KPP07, Theorem 3.13], the number of elements in a reduced p-basis is uniquely determined, which yields the next definition.

Definition 3.1.10 Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with reduced p-basis $(v_1, \ldots, v_N)$. Then the p-dimension of M is defined as

$$p\text{-dim}(M) := N.$$

The following definition adjusts the PLM property, introduced for the field case in Definition 2.1.15, to the specific structure of $\mathbb{Z}_{p^r}$. It extends the p-predictable degree property introduced in [KPP07] to a stronger property that will prove useful in the sequel.

Definition 3.1.11 Let $F = \{f_1, \ldots, f_s\} \subseteq \mathcal{D}^q$ be a set of nonzero elements. Then F has the ***p*-Predictable Degree (*p*-PD) property** if for any $0 \neq f \in p\text{-span}(f_1, \ldots, f_s)$, written as

$$f = a_1 f_1 + \cdots + a_s f_s, \tag{3.5}$$

where $a_1, \ldots, a_s \in \mathrm{A}_p[x]$, we have

$$\deg(f) = \max_{1 \leq i \leq s; a_i \neq 0} (\deg(a_i) + \deg(f_i)).$$

Next, F is said to have the ***p*-Predictable Leading Position (*p*-PLP) property** if

$$\mathrm{lpos}(f) = \min_{1 \leq i \leq s; a_i \neq 0} \mathrm{lpos}(f_i).$$

Finally, F is said to have the ***p*-Predictable Leading Monomial (*p*-PLM) property** if

$$\mathrm{lm}(f) = \max_{1 \leq i \leq s; a_i \neq 0} (\mathrm{lm}(a_i)\,\mathrm{lm}(f_i)).$$

Note that in the above definition, $a_i \in \mathrm{A}_p[x]$ rather than $a_i \in \mathcal{R}[x]$ as in Definition 2.1.15. In analogy with the field case, it is easily seen that the p-PLM property holds if and only if both the p-PD property and the p-PLP property hold.

3.2 Minimal Gröbner p-basis and the p-predictable degree property

A minimal Gröbner basis is not a p-generator sequence, because due to the minimality, no multiples of a generator are contained in the Gröbner basis. The next theorem shows how to obtain a p-generator sequence from a minimal Gröbner basis.

Theorem 3.2.1 [KSb] *Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G as per Definition 2.1.26. Then*

$$(g_1, pg_1, \ldots, p^{\beta_1 - 1} g_1, g_2, pg_2, \ldots, p^{\beta_2 - 1} g_2, \ldots, g_m, pg_m, \ldots, p^{\beta_m - 1} g_m) \tag{3.6}$$

is a p-generator sequence whose p-span equals M.

Proof: We first prove that (3.6) satisfies Definition 3.1.3. By definition $\beta_m = \mathrm{ord}(g_m)$, so that

$$\mathrm{lm}(p^{\beta_m} g_m) < \mathrm{lm}(g_m). \tag{3.7}$$

Suppose $p^{\beta_m} g_m \neq 0$, then according to Observation 2.1.8, there exists $g_i \in G$ such that $\mathrm{lm}(g_i) \leq \mathrm{lm}(p^{\beta_m} g_m)$. But then (3.7) implies that $\mathrm{lm}(g_i) < \mathrm{lm}(g_m)$, which contradicts $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. We conclude that

$$p^{\beta_m} g_m = 0. \tag{3.8}$$

To prove that (3.6) satisfies Definition 3.1.3, it now obviously remains to prove that $p^{\beta_j} g_j$ is a p-linear combination of

$$g_{j+1}, pg_{j+1}, \ldots, p^{\beta_{j+1}-1} g_{j+1}, g_{j+2}, pg_{j+2}, \ldots, p^{\beta_{j+2}-1} g_{j+2}, \ldots, g_m, \ldots, p^{\beta_m - 1} g_m \quad (3.9)$$

for $1 \leq j \leq m-1$. For this, we first prove that $p^{\beta_j} g_j$ is a linear combination of $g_{j+1}, g_{j+2}, \ldots, g_m$. We distinguish two cases:
case I

> $\beta_j = \mathrm{ord}(g_j)$. Then $\mathrm{lm}(p^{\beta_j} g_j) < \mathrm{lm}(g_j)$, so that, by Observation 2.1.11, $p^{\beta_j} g_j$ is a linear combination of $g_{j+1}, g_{j+2}, \ldots, g_m$.

case II

> $\beta_j < \mathrm{ord}(g_j)$, so that $\mathrm{lm}(p^{\beta_j} g_j) = \mathrm{lm}(g_j)$. By definition, there exists a smallest integer $i > j$ with $\mathrm{lpos}(g_i) = \mathrm{lpos}(g_j)$ and $\beta_j = \mathrm{ord}(g_j) - \mathrm{ord}(g_i)$. Observe that then $\mathrm{ord}(p^{\beta_j} g_j) = \mathrm{ord}(g_i)$ and $\deg(p^{\beta_j} g_j) = \deg(g_j) > \deg(g_i)$ (use Lemma 2.1.25), whereas $\mathrm{lpos}(p^{\beta_j} g_j) = \mathrm{lpos}(g_j) = \mathrm{lpos}(g_i)$. Thus we can find $a \in \mathbb{Z}_{p^r}[x]$ such that $\mathrm{lt}(p^{\beta_j} g_j) = \mathrm{lt}(a g_i)$. As a result, $\mathrm{lm}(p^{\beta_j} g_j - a g_i) < \mathrm{lm}(p^{\beta_j} g_j) = \mathrm{lm}(g_j)$. Consequently, by Observation 2.1.11, $p^{\beta_j} g_j - a g_i$ is a linear combination of $g_{j+1}, g_{j+2}, \ldots, g_m$. Since $i > j$, it follows that $p^{\beta_j} g_j$ is also a linear combination of $g_{j+1}, g_{j+2}, \ldots, g_m$.

Thus for $1 \leq j \leq m-1$

$$p^{\beta_j} g_j \text{ is a linear combination of } g_{j+1}, \ldots, g_m. \quad (3.10)$$

Finally, we prove by induction that (3.9) holds for $1 \leq j \leq m-1$. For $j = m-1$, this follows from (3.8) and the fact that $p^{\beta_{m-1}} g_{m-1}$ is a multiple of g_m because of (3.10). Now suppose that (3.9) holds for $j = j_0 \in \{1, \ldots, m-1\}$. Consider the vector $p^{\beta_{j_0-1}} g_{j_0-1}$. By (3.10), there exist $a_{j_0}, \ldots, a_m \in \mathbb{Z}_{p^r}[x]$ such that

$$p^{\beta_{j_0-1}} g_{j_0-1} = a_{j_0} g_{j_0} + \cdots + a_m g_m.$$

Now use the p-adic decomposition to write

$$a_{j_0} = a_{j_0}^0 + p a_{j_0}^1 + \cdots + p^{r-1} a_{j_0}^{r-1},$$

where $a_{j_0}^i \in \mathrm{A}_p[x]$ for $0 \leq i \leq r-1$. Repeatedly using the induction hypothesis, it follows that

$$p^{\beta_{j_0-1}} g_{j_0-1} = a_{j_0}^0 g_{j_0} + \cdots + p^{\beta_{j_0}-1} a_{j_0}^{\beta_{j_0}-1} g_{j_0} + p\text{-linear combination of } g_{j_0+1}, \ldots, p^{\beta_m - 1} g_m.$$

This proves that (3.9) holds for $j = j_0 - 1$, so that, by induction, (3.6) is a p-generator sequence.
To prove that its p-span equals M, we first note that, by Observation 2.1.11, any element of M can be written as a linear combination of $g_1, g_2, \ldots, g_m$. Using a similar reasoning as above, this can be alternatively be written as a p-linear combination of the vectors in (3.9). □

The next lemma follows immediately from Definition 2.1.26.

Lemma 3.2.2 *Let M be a submodule of $\mathcal{D}^q$ possessing a minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G as per Definition 2.1.26 and let $N = \beta_1 + \beta_2 + \cdots + \beta_m$. Let $(v_1, \ldots, v_N)$ be the p-generator sequence given by (3.6). Then for any $i, j \in \{1, \ldots, N\}$ with $i \neq j$ we have*

$$\mathrm{lpos}(v_i) = \mathrm{lpos}(v_j) \Rightarrow \mathrm{ord}(v_i) \neq \mathrm{ord}(v_j).$$

The next theorem is the ring analogon of Theorem 2.1.17 and presents the main result of this chapter.

Theorem 3.2.3 *Let M, $(\beta_1, \ldots, \beta_m)$ and $\{v_1, \ldots, v_N\}$ be defined as in the previous lemma. Then $\{v_1, \ldots, v_N\}$ has the p-PLM property. In particular, $(v_1, \ldots, v_N)$ is a p-basis of M.*

Proof: Let

$$f = a_1 v_1 + \cdots + a_N v_N \tag{3.11}$$

with $a_1, \ldots, a_N \in \mathrm{A}_p[x]$. For simplicity of notation, we assume that a_i is nonzero for $1 \leq i \leq N$. Let us first examine two special cases:

Special case I

> All g_i's have distinct leading positions. Then the proof is analogous to the field case, i.e., the proof of Theorem 2.1.17.

Special case II

> All g_i's have the same leading position. Then all v_i's also have the same leading position. By Lemma 3.2.2, their orders are all different. Now observe that $\mathrm{ord}(a_i v_i) = \mathrm{ord}(v_i)$ for $1 \leq i \leq N$, since $a_i \in \mathrm{A}_p[x]$. Thus all $a_i v_i$'s have different orders. In particular, all $a_i v_i$'s of largest degree have different orders, so that their leading coefficients add up to a nonzero element of $\mathbb{Z}_{p^r}$ (use the p-adic decomposition). This implies that the p-PLM property holds.

Let us now consider the general case. By grouping together all vectors $a_i v_i$ of the same leading position, we write

$$f = f_1 + f_2 + \cdots + f_q,$$

where $f_i = 0$ if position i is not used in (3.11). As in Special case II above, it can be shown that $\mathrm{lpos}(f_i) = i$ whenever $f_i \neq 0$. As a result, the nonzero f_i's can be ordered and Observation 2.1.7 yields

$$\mathrm{lt}(f) = \mathrm{lt}(f_j) \tag{3.12}$$

for some nonzero f_j with $j \in \{1, \ldots, q\}$. Recall that f_j is defined as the sum of all vectors in the right hand side of (3.11) that have leading position j. It now follows from Special case II above that there exists $\ell \in \{1, \ldots, N\}$ such that $\mathrm{lm}(f_j) = \mathrm{lm}(a_\ell)\,\mathrm{lm}(v_\ell)$. As a result, by equation (3.12),

$$\mathrm{lm}(f) = \mathrm{lm}(a_\ell)\,\mathrm{lm}(v_\ell). \tag{3.13}$$

Evidently $\mathrm{lm}(f) \leq \max_{1\leq i\leq N; a_i\neq 0}(\mathrm{lm}(a_i)\,\mathrm{lm}(v_i))$, so that (3.13) implies that equality holds. This proves the p-PLM property.
Finally, to prove that $(v_1,\ldots,v_N)$ is a p-basis for M, first observe that p-span of $(v_1,\ldots,v_N)$ equals M by Theorem 3.2.1. Also, it follows immediately from the p-PLM property that any nontrivial p-linear combination of vectors in $\{v_1,\ldots,v_N\}$ has to be nonzero. We conclude that $(v_1,\ldots,v_N)$ is a p-basis of M, so that $N = p\text{-dim}(M) = \beta_1+\beta_2+\cdots+\beta_m$. □

Definition 3.2.4 [KSb] Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis $G = \{g_1,\ldots,g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. Let $(\beta_1,\ldots,\beta_m)$ be the sequence of order differences of G as per Definition 2.1.26. Let $(v_1,v_2,\ldots,v_N)$ be the p-generator sequence given by (3.6). Then $(v_1,v_2,\ldots,v_N)$ is called a **minimal Gröbner p-basis** for M.

Let us denote

$$\mathrm{LC}(v) := \mathrm{lc}(v)\,\mathrm{e}_{\mathrm{lpos}(v)} \in \mathbb{Z}_{p^r}^q$$

for a nonzero $v \in \mathcal{D}^q$.

Remark 3.2.5 Let $(v_1,v_2,\ldots,v_N)$ be a minimal Gröbner p-basis of M. Then

$$(\mathrm{LC}(v_1),\ldots,\mathrm{LC}(v_N))$$

is a p-generator sequence in $\mathbb{Z}_{p^r}^q$.

Proof: By definition, the sequence $(\mathrm{LC}(v_1),\ldots,\mathrm{LC}(v_N))$ coincides to

$$(\mathrm{LC}(g_1),\mathrm{LC}(pg_1),\ldots,\mathrm{LC}(p^{\beta_1-1}g_1),\ldots,\mathrm{LC}(g_m),\ldots,\mathrm{LC}(p^{\beta_m-1}g_m)).$$

Note that $\mathrm{LC}(p^\alpha g_i) = p^\alpha\,\mathrm{LC}(g_i)$ for all $1 \leq \alpha \leq \beta_i - 1$. Therefore to prove the claim, it is sufficient to show that $p\,\mathrm{LC}(p^{\beta_j-1}g_j)$ is contained in the p-span of the sequence $(\mathrm{LC}(g_{j+1}),\ldots,\mathrm{LC}(p^{\beta_m-1}g_m))$. Note that $\mathrm{LC}(g_j)$ can be written as $u_j p^{r-\mathrm{ord}(g_j)}\,\mathrm{e}_{\mathrm{lpos}(g_j)}$ for a unit u_j, which yields

$$p\,\mathrm{LC}(p^{\beta_j-1}g_j) = p^{r-\mathrm{ord}(g_j)+\beta_j}\,\mathrm{e}_{\mathrm{lpos}(g_j)}\,.$$

Case I: $\beta_j = \mathrm{ord}(g_j)$

Then $r - \mathrm{ord}(g_j) + \beta_j = r$ and thus $p\,\mathrm{LC}(p^{\beta_j-1}g_j) = 0$.

Case II: $\beta_j \neq \mathrm{ord}(g_j)$

Then there exists a smallest integer $i > j$ such that $\mathrm{lpos}(g_i) = \mathrm{lpos}(g_j)$ and $\beta_j = \mathrm{ord}(g_j) - \mathrm{ord}(g_i)$. Further note that $\mathrm{LC}(g_i) = u_i p^{r-\mathrm{ord}(g_i)}\,\mathrm{e}_{\mathrm{lpos}(g_j)}$ for a unit u_i. Since

$$r - \mathrm{ord}(g_j) + \beta_j = r - \mathrm{ord}(g_j) + \mathrm{ord}(g_j) - \mathrm{ord}(g_i) = r - \mathrm{ord}(g_i),$$

we obtain

$$p\,\mathrm{LC}(p^{\beta_j-1}g_j) = u_j p^{r-\mathrm{ord}(g_i)}\,\mathrm{e}_{\mathrm{lpos}(g_j)} = u_j u_i^{-1}\,\mathrm{LC}(g_i).$$

Writing $u_j u_i^{-1} = \theta_0 + \theta_1 p + \cdots + \theta_{r-1}p^{r-1}$ yields

$$\begin{aligned} u_j u_i^{-1}\,\mathrm{LC}(g_i) = \overbrace{\theta_0\,\mathrm{LC}(g_i) + \theta_1 p\,\mathrm{LC}(g_i) + \cdots + \theta_{\beta_i-1}p^{\beta_i-1}\,\mathrm{LC}(g_i)}^{\in p\text{-span}(\mathrm{LC}(g_1),\ldots,\mathrm{LC}(p^{\beta_i-1}g_i))} \\ + p^{\beta_i}(\theta_{\beta_i} + \cdots + \theta_{r-1}p^{r-1-\beta_i})\,\mathrm{LC}(g_i). \end{aligned}$$

Let u denote the unit $\theta_{\beta_i} + \cdots + \theta_{r-1}p^{r-1-\beta_1}$. It is still left to show that $up^{\beta_i}\,\mathrm{LC}(g_i)$ is contained in the p-span of $(\mathrm{LC}(g_i),\ldots,\mathrm{LC}(p^{\beta_m-1}g_m))$. In case $\beta_i = \mathrm{ord}(g_i)$, it follows that $up^{\beta_i}\,\mathrm{LC}(g_i) = 0$. Else there exists a smallest integer $k > i$ such that $\mathrm{lpos}(g_k) = \mathrm{lpos}(g_i)$ and $\beta_i = \mathrm{ord}(g_i) - \mathrm{ord}(g_k)$. And furthermore $\mathrm{LC}(g_k) = u_k p^{r-\mathrm{ord}(g_k)} e_{\mathrm{lpos}\,g_j}$ for a unit u_k. Then

$$up^{\beta_i}\,\mathrm{LC}(g_i) = uu_i p^{\beta_i+r-\mathrm{ord}(g_i)}\,\mathrm{e}_{\mathrm{lpos}\,g_j} = uu_i u_k^{-1}\,\mathrm{LC}(g_k).$$

The claim follows by induction. □

Remark 3.2.6 Let $(v_1, v_2, \ldots, v_N)$ be a minimal Gröbner p-basis of M. Then

$$(\mathrm{LC}(v_1),\ldots,\mathrm{LC}(v_N))$$

is p-linearly independent in $\mathbb{Z}_{p^r}^q$.

Proof: Suppose the claim does not hold, that is, there exist $a_i \in \mathrm{A}_p$, not all equal to zero, such that

$$\sum_{i=1}^{N} a_i\,\mathrm{LC}(v_i) = 0. \tag{3.14}$$

Since $\mathrm{LC}(v_i)$ is of the form $d\,\mathrm{e}_j$ for $d \in \mathbb{Z}_{p^r}$ and $1 \le j \le q$, we can without loss of generality suppose that q equals 1. Let k be the smallest integer such that a_k in (3.14) is nonzero. Then $a_k\,\mathrm{lc}(v_k) = -\sum_{i=k+1}^{N} a_i\,\mathrm{lc}(v_i)$. Recall that $\mathrm{ord}(v_k) > \cdots > \mathrm{ord}(v_N)$ by definition, and obviously $\mathrm{ord}(v_i) = \mathrm{ord}(-v_i)$. Then

$$a_k\,\mathrm{lc}(v_k) = p^{r-\mathrm{ord}(v_{k+1})}(-\sum_{i=k+1}^{N} a_i u_i p^{(-\,\mathrm{ord}\,v_i + \mathrm{ord}\,v_{k+1})}),$$

where $\mathrm{lc}(v_i) = u_i p^{r-\mathrm{ord}\,v_i}$ for a unit u_i. But this leads to a contradiction, since

$$\mathrm{ord}(a_k\,\mathrm{lc}(v_k)) = \mathrm{ord}(\mathrm{lc}(v_k)) = \mathrm{ord}(v_k) > \mathrm{ord}(v_{k+1}) = \mathrm{ord}(p^{r-\mathrm{ord}(v_{k+1})})$$

and

$$\mathrm{ord}(p^{r-\mathrm{ord}(v_{k+1})}) > \mathrm{ord}(p^{r-\mathrm{ord}(v_{k+1})}(-\sum_{i=k+1}^{N} a_i u_i p^{(-\,\mathrm{ord}\,v_i + \mathrm{ord}\,v_{k+1})}))$$

hold. □

Theorem 3.2.7 *Let M be a submodule of $\mathcal{D}^q$ with a minimal Gröbner basis corresponding to the* TOP *ordering. Then the associated minimal Gröbner p-basis $(v_1, \ldots, v_N)$ is a reduced p-basis of M.*

Proof: To prove the claim, we need to show that $(v_1^{\mathrm{lrc}}, \ldots, v_N^{\mathrm{lrc}})$ is p-linearly independent. For this purpose note:

1. Due to Remark 3.2.6, we already know that $(\mathrm{LC}(v_1), \ldots, \mathrm{LC}(v_N))$ is p-linearly independent.

2. By the definition of TOP, see Definition 2.1.6, each v_i^{lrc} can be written as

$$v_i^{\mathrm{lrc}} = \mathrm{LC}(v_i) + \sum_{j=\mathrm{lpos}(v_i)+1}^{q} u_j^{(i)}\, \mathrm{e}_j$$

   for suitable $u_j^{(i)}$ and all $1 \leq i \leq N$.

These two items yield the claim. □

One can easily see that the requirement in the previous remark on the minimal Gröbner basis to be calculated with respect to the TOP ordering is crucial for the result. Suppose we would work with respect to POT and consider for instance the module $\langle (\underline{1}\,,\, x), (0\,,\, \underline{x}) \rangle = M \subseteq \mathcal{D}^2$. Then the generators of M already give a minimal Gröbner basis of M corresponding to POT. But the leading row coefficient vectors both coincide to $(0\,,\, 1)$. Algorithm 1 shows how to obtain a reduced p-basis from a minimal Gröbner p-basis in general. Suppose the input of the Algorithm to be M generated by a minimal Gröbner p-basis $(v_1, \ldots, v_N)$ and suppose the output to be $(w_1, \ldots, w_k)$. Then

$$p\text{-span}(v_1, \ldots, v_N) = p\text{-span}(w_1, \ldots, w_k). \tag{3.15}$$

This is easy to see. Since $(w_1, \ldots, w_k)$ is a p-basis of M, v_i is contained in the p-span of $(w_1, \ldots, w_k)$ for all $1 \leq i \leq N$. Conversely let us recall the operations acting on the input vectors during the algorithm: First, the sequence $(v_1, \ldots, v_N)$ is expanded by power of p multiples to

$$W := (v_1, pv_1, \ldots, p^{r-1}v_1, \ldots, v_N, \ldots, p^{r-1}v_N).$$

Since $(v_1, \ldots, v_N)$ is a p-generator sequence, the sequence W is contained in the p-span of $(v_1, \ldots, v_N)$ due to Theorem 3.1.4. Apart from re-ordering and the deletion of zero vectors, the algorithm changes W in the following way. Elements w_ℓ of W are suitably replaced by

$$\tilde{w}_\ell := w_\ell + \alpha_1 x^{\deg(w_\ell)-\deg(w_{\ell+1})} w_{\ell+1} + \cdots + \alpha_{k-\ell} x^{\deg(w_\ell)-\deg(w_k)} w_k,$$

where $\alpha_i \in \mathbb{Z}_{p^r}$ and $w_i \in W$. Since each w_i is contained in the span of $(v_1, \ldots, v_N)$ and $(v_1, \ldots, v_N)$ is a p-basis, the element $\tilde{w}_\ell$ is contained in the p-span of $(v_1, \ldots, v_N)$ as well. Altogether we have shown that equality (3.15) holds.
We can further conclude that $k = N$, since both sequences are p-linearly independent. Thus the number of elements in a minimal Gröbner p-basis equals the p-dimension of the associated module M. This allows to give the following characterization.

Lemma 3.2.8 *Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner p-basis $(v_1, \ldots, v_N)$. Then the p-dimension of M is given by*

$$p\text{-dim}(M) = N.$$

Remark 3.2.9 Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\text{lm}(g_1) > \cdots > \text{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G as per Definition 2.1.26. Then

$$p\text{-dim}(M) = \sum_{i=1}^{m} \beta_i.$$

Proof: The claim follows by the definition of a minimal Gröbner p-basis. □

Corollary 3.2.10 *Let M be a submodule of $\mathcal{D}^q$. Then the p-dimension of M is bounded by rq, that is,*

$$p\text{-dim}(M) \leq rq.$$

Proof: By the definition of the sequence of ordered differences, each β_i takes exclusively vectors of equal leading positions into account. In order to prove the claim, we can therefore assume $q = 1$ without loss of generality and show that $p\text{-dim}(M) \leq r$. Let $\{g_1, \ldots, g_m\}$ be a minimal Gröbner basis of M, ordered so that $\text{lm}(g_1) > \cdots > \text{lm}(g_m)$. Then by the previous remark,

$$p\text{-dim}(M) = \sum_{i=1}^{m} \beta_i = (\sum_{i=1}^{m-1} \text{ord}(g_i) - \text{ord}(g_{i+1})) + \text{ord}(g_m) = \text{ord}(g_1)$$

holds. Since additionally $\text{ord}(m) \leq r$ for all $m \in \mathcal{D}$, the claim follows. □

Corollary 3.2.10 leads to the following remark.

Remark 3.2.11 Let M be a submodule of $\mathcal{D}^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\text{lm}(g_1) > \cdots > \text{lm}(g_m)$. Let $(\beta_1, \ldots, \beta_m)$ be the sequence of order differences of G. Further let $g^{(j)}$ denote the element g_i of smallest index in G such that $\text{lpos}(g_i) = j$. Then

$$\sum_{i=1;\text{lpos}(g_i)=j}^{m} \beta_i = \text{ord}(g^{(j)}) \leq r.$$

Example 3.2.12 Recall Example 3.0.21.

- Using the TOP ordering:
  By Theorem 3.2.3, the sequence (g_1, g_2, g_3, g_4) is a minimal Gröbner p-basis for M. Further due to Remark 3.2.7, the p-generator sequence (g_1, g_2, g_3, g_4) is a reduced p-basis; it has the p-PLM property. Furthermore, $p\text{-dim}(M) = \beta_1 + \beta_2 + \beta_3 + \beta_4 = 4$.

- Using the POT ordering:
  According to Theorem 3.2.3, the sequence $(R_1, 3R_1, R_2, 3R_2)$ is a minimal Gröbner p-basis for M; it has the p-PLM property. Note that $\beta_1 + \beta_2$ indeed equals $4 = p\text{-dim}(M)$.

3.3 Application to signals and systems

In the publications [KP08b, KP08a], one can find applications for minimal Gröbner p-bases. In [KSb], these applications are studied in detail.

Parametrization of all shortest linear recurrence relations

In Example 2.1.21, we have already discussed the field case. The results that will be presented are given in [KP08b] without the use of minimal Gröbner p-bases. Suppose given a sequence $S_0, \ldots, S_{n-1}$ over $\mathbb{Z}_{p^r}$. We call a polynomial $f \in \mathcal{D}$, written as $f(x) = f_L x^L + f_{L-1}x^{L-1} + \cdots + f_1 x + f_0$, a **linear recurrence relation** of length L for $S_0, \ldots, S_{n-1}$ if f_L is a unit and

$$f_L S_{L+j} + \sum_{i=1}^{L} f_{L-i} S_{L+j-i} = 0 \qquad \text{for } j = 0, \ldots, n-L-1. \tag{3.16}$$

As usual, we call the polynomial f **monic** if $f_L = 1$. As in Example 2.1.21, defining the **partial impulse response trajectory** $\boldsymbol{b}$ on the time-axis $\mathbb{N}$ as

$$\boldsymbol{b} = \left(\begin{bmatrix} S_0 \\ 0 \end{bmatrix}, \begin{bmatrix} S_1 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} S_{n-1} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \cdots \right), \tag{3.17}$$

we can reformulate (3.16) as $[d(\mathbf{s}) \quad -h(\mathbf{s})]\,\boldsymbol{b} = 0$, where $h(x)$ is a polynomial of degree $\leq L$ and $\mathbf{s}$ is the backward shift operator, acting on trajectories $\boldsymbol{w}$ on $\mathbb{N}$ as $(\mathbf{s}\boldsymbol{w})(k) = \boldsymbol{w}(k+1)$. A linear recurrence relation for $S_0, \ldots, S_{n-1}$ thus corresponds to a kernel representation

$$[d(\mathbf{s}) \quad -h(\mathbf{s})]\,\boldsymbol{w} = 0$$

whose behavior includes the so-called **partial impulse response behavior**

$$\mathcal{B} := \operatorname{span}\{\boldsymbol{b}, \mathbf{s}\boldsymbol{b}, \mathbf{s}^2\boldsymbol{b}, , \ldots, \mathbf{s}^n\boldsymbol{b}\}, \tag{3.18}$$

where $\boldsymbol{b}$ is defined by (3.17). The search for shortest linear recurrence relations now translates into a search for an annihilator $[d(\mathbf{s}) \quad -h(\mathbf{s})]\,\boldsymbol{w} = 0$ for $\mathcal{B}$ that has minimal row degree and satisfies $\deg(h) \leq \deg(d)$. As in Example 2.1.21, define the polynomial $S(x)$ as

$$S(x) := S_0 x^n + S_1 x^{n-1} + \cdots + S_{n-1}x \tag{3.19}$$

and consider the module M spanned by the vectors $\begin{bmatrix} 1 & -S(x) \end{bmatrix}$ and $\begin{bmatrix} 0 & x^{n+1} \end{bmatrix}$. It is easily verified that M consists of all annihilators of $\mathcal{B}$.
The vectors $\begin{bmatrix} 1 & -S(x) \end{bmatrix}$ and $\begin{bmatrix} 0 & x^{n+1} \end{bmatrix}$ are a minimal Gröbner basis of M with respect to the POT ordering and the corresponding sequence of order differences is (r, r). Thus it is clear that

$$p\text{-}\dim(M) = 2r.$$

Let $\{g_1, \ldots, g_m\}$ be a minimal Gröbner basis of M corresponding the TOP ordering, so that $\operatorname{lm}(g_1) > \cdots > \operatorname{lm}(g_m)$. Further let $(v_1, v_2, \ldots, v_{2r})$ be the associated minimal Gröbner p-basis, with v_i written as $v_i = [d_i \quad -h_i] \in \mathcal{D}^2$ for $i = 1, \ldots, 2r$. Then due to Remark 3.2.11, there exists an index $h \in \{1, \ldots, m\}$ such that $\operatorname{lpos}(g_h) = 1$ and $\operatorname{ord}(g_h) = r$. Let $\ell \in \{1, \ldots, 2r\}$ be such that $v_\ell = g_h$. By the definition of TOP, we obtain that $\deg(h_\ell) \leq \deg(d_\ell)$. Thus it follows that d_ℓ is a linear recurrence relation for the sequence $S_0, \ldots, S_{n-1}$.

Theorem 3.3.1 *Using the previous notation, we claim that:*

1. *The polynomial d_ℓ is a shortest linear recurrence relation and furthermore, a parametrization of all shortest linear recurrence relations for $S_0, \ldots, S_{n-1}$ is given by*
$$q_\ell d_\ell + \sum_{i>\ell} q_i d_i, \tag{3.20}$$
with $0 \neq q_\ell \in \mathrm{A}_p$ and $q_i \in \mathrm{A}_p[x]$ with $\deg(q_i) \leq \deg(v_\ell) - \deg(v_i)$ for $i = \ell+1, \ldots, 2r$.

2. *The shortest linear recurrence relation d_ℓ is unique up to units if and only if $v_\ell = g_m$, that is, if and only if $\mathrm{lpos}(g_m) = 1$ and $\mathrm{ord}(g_m) = r$.*

Proof:

1. Suppose that a polynomial $d^\star \in \mathcal{D}$ is a shortest linear recurrence relation for $S_0, \ldots, S_{n-1}$. Then there exists a polynomial $h^\star \in \mathcal{D}$ of smaller or equal degree such that $[d^\star \quad -h^\star] \in M$. Since $(v_1, v_2, \ldots, v_{2r})$ is a minimal Gröbner p-basis of M, we can write $[d^\star \quad -h^\star]$ as a p-linear combination of v_1, $v_2 \ldots, v_{2r}$. Since v_ℓ is the unique vector in this Gröbner p-basis of leading position 1 and order r, this p-linear combination must use v_ℓ. Because of the p-PLM property of $\{v_1, \ldots, v_N\}$, it follows that $\deg(d^\star) \geq \deg(v_\ell)$. This implies that v_ℓ is a shortest linear recurrence relation for $S_0, \ldots, S_{n-1}$. Moreover, it also follows from the p-PLM property of $\{v_1, \ldots, v_N\}$ that the above p-linear combination can not use v_i for $i < \ell$. This proves the parametrization (3.20).

2. If v_ℓ equals g_m, then d_ℓ is obviously unique up to units. Now let d_ℓ be unique up to units and suppose v_ℓ is not equal to the element of smallest leading monomial of G. Then by definition of a minimal Gröbner p-basis and the choice of v_ℓ, there exists an element g_h with $1 \leq h < m$ such that $v_\ell = g_h$. But then $g_m = v_i \in (v_1, \ldots, v_{2r})$ for a suitable $i > \ell$ and since $\mathrm{lm}(g_h) > \mathrm{lm}(g_m)$, it follows that $\deg(d_i) < \deg(d_\ell)$. This would yield the shortest linear recurrence relation $d_\ell + d_i$ and introduce a contradiction. □

Example 3.3.2 Consider the sequence $S_0, S_1, S_2, S_3, S_4 = 1, 4, 4, 7, 7$ over the ring $\mathbb{Z}_9$. Let M be the submodule of $\mathbb{Z}_9[x]^2$, defined by $M = \mathrm{span}\{s_1, s_2\}$, where $s_1(x) = [1 \quad 8x^5 + 5x^4 + 5x^3 + 2x^2 + 2x]$ and $s_2(x) = [0 \quad x^6]$. Using the TOP ordering, as shown in Example 3.0.21, the SINGULAR computer algebra system [GPS05] computes the minimal Gröbner basis $G = \{g_1, \ldots, g_4\}$ of M corresponding to the rows of
$$\begin{bmatrix} 8 & \underline{x^5} + 4x^4 + 4x^3 + 7x^2 + 7x \\ x + 5 & \underline{3x^4} + 3x^2 + x \\ \underline{x^2} + 3x + 2 & x^2 + 4x \\ \underline{3x} + 6 & 3x \end{bmatrix}.$$

According to Theorem 3.3.1, g_3 gives a shortest linear recurrence relation $x^2 + 3x + 2$ that is not unique. A parametrization of all shortest linear recurrence relations is given by $\Theta_1(x^2 + 3x + 2) + (\Theta_2 x + \Theta_3)(3x + 6)$, where $\Theta_i \in \{0, 1, 2\}$ for $i = 1, 2, 3$; $\Theta_1 \neq 0$.

It is easily seen that a parametrization of all monic shortest linear recurrence relations is given by $(x^2+3x+2)+\Theta(3x+6)$, where $\Theta \in \{0,1,2\}$. That is, we obtain

$$S_{2+j} + (3+3\Theta)S_{1+j} + (2+6\Theta)S_j = 0, \quad \text{where } \; j = 0 \ldots 2.$$

Minimal state realization

In Example 2.1.22, we have already discussed minimal state realization for the field case. Consider a finite support convolutional code $\mathcal{C}$ of length n over $\mathbb{Z}_{p^r}$, that is, a submodule of $\mathbb{Z}_{p^r}^n[x]$. Denote the p-dimension of $\mathcal{C}$ by κ. Let $(v_1, v_2, \cdots, v_\kappa)$ be a minimal Gröbner p-basis for $\mathcal{C}$ (under TOP ordering). Denote the $\kappa \times n$ polynomial matrix $(v_1^T, v_2^T, \cdots, v_\kappa^T)^T$ by V. We call V a **minimal Gröbner p-encoder** for $\mathcal{C}$. The matrix V is realized in controller canonical form as

$$V(x) = B(x^{-1}I - A)^{-1}C + D,$$

see [Kai80] and [KP08a]. Here $(A,B,C,D) \in \mathbb{Z}_{p^r}^{\gamma\times\gamma} \times \mathbb{Z}_{p^r}^{\kappa\times\gamma} \times \mathbb{Z}_{p^r}^{\gamma\times n} \times \mathbb{Z}_{p^r}^{\kappa\times n}$, where $\gamma = \sum_{i=1}^{\kappa} \deg(v_i)$. In the terminology of [KP08a], the corresponding controller canonical trellis representation is denoted as $\mathcal{X}_V := \{X_t\}_{t\in\mathbb{Z}_+}$, where $X_t = (\mathbb{Z}_{p^r}^n, S_t, S_t', K_t)$ with

$$S_0 = \{0\} \text{ and } S_t' = \{sA + uB : s \in S_t \text{ and } u \in \mathrm{A}_p^\kappa\}, \; t \in \mathbb{Z}_+ \quad \text{and}$$

$$K_t = \{(s(t), s(t)C + u(t)D, s(t)A + u(t)B \mid s(t) \in S_t \text{ and } u(t) \in \mathrm{A}_p^\kappa\}.$$

The next theorem follows immediately from [KP08a, Thm. 2].

Theorem 3.3.3 *Let $\mathcal{C}$ be a finite support convolutional code over $\mathbb{Z}_{p^r}$ of length n and p-dimension κ, and let $V \in \mathbb{Z}_{p^r}^{\kappa\times n}[x]$ be a minimal Gröbner p-encoder of $\mathcal{C}$; write $V = (v_1^T, v_2^T, \cdots, v_\kappa^T)^T$ and denote $\gamma := \sum_{i=1}^{\kappa} \deg(v_i)$ and $\gamma_{max} := \max_{1\le i\le\kappa}\{\deg(v_i)\}$. Then the controller canonical trellis representation $\mathcal{X}_V$ is a minimal trellis representation for $\mathcal{C}$. In particular, the number of trellis states of $\mathcal{X}_V$ equals p^γ, for $t \ge \gamma_{max}$.*

Example 3.3.4 Consider the finite support convolutional code

$$\mathcal{C} = \mathrm{span}\{\begin{bmatrix} x^2+1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 2x & 1 & 2 \end{bmatrix}\}$$

of length 3 over $\mathbb{Z}_4$. Using the TOP ordering, SINGULAR computes the minimal Gröbner basis $G = \{g_1, g_2, g_3, g_4\}$ of M, where $g_1 = \begin{bmatrix} x^2+1 & 1 & 0 \end{bmatrix}$, $g_2 = \begin{bmatrix} 2x & 1 & 2 \end{bmatrix}$, $g_3 = \begin{bmatrix} 2 & x & 2x \end{bmatrix}$ and $g_4 = \begin{bmatrix} 0 & 2 & 0 \end{bmatrix}$. A minimal Gröbner p-encoder is then given by

$$V(x) = \begin{bmatrix} x^2+1 & 1 & 0 \\ 2x & 1 & 2 \\ 2 & x & 2x \\ 0 & 2 & 0 \end{bmatrix}.$$

Its controller canonical trellis $\mathcal{X}_V$ is given by

$$A = \begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}, \; B = \begin{bmatrix} 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \\ 0&0&0&0 \end{bmatrix}, \; C = \begin{bmatrix} 0&0&0 \\ 1&0&0 \\ 0&1&2 \\ 2&0&0 \end{bmatrix}, \; D = \begin{bmatrix} 1&1&0 \\ 0&1&2 \\ 2&0&0 \\ 0&2&0 \end{bmatrix}.$$

According to the above theorem, this trellis is minimal with $2^4 = 16$ trellis states for $t \ge 2$.

Chapter 4

One-dimensional time-varying systems

As already outlined in Section 1.4, the system and control theoretical interpretations of a diagonal form and a normal form are of great interest. The proof for the existence of a normal form is usually constructive and therefore, it offers an algorithm directly. However, such a direct algorithm is not very efficient, in general. Naturally, the proposed approaches getting a grip on the occurring difficulties benefit from the structure of the underlying ring, see for instance [Lüb02].

Here, we choose the framework of skew polynomial rings, see Section 2.2. Skew polynomial rings, among others, offer the possibility to describe time-varying systems. A system and control theoretical motivation to consider these rings is established in [Zer06a, Zer07c, IM05, INS84]. In [CQ05], applications to systems of partial differential equations are shown and several concrete examples are studied.
Beyond that, many known operator algebras can be realized as skew polynomial rings or solvable polynomial rings [Kre93], some of them can be even realized as Ore algebras [CS98, CQR04]. However, general solvable polynomial rings are hard to tackle constructively (say, in a computer algebra system), while the class of Ore algebras is restrictive. Based on PBW algebras [BGTV03] or G-algebras [Lev05a, GLH05], we have proposed a new class of univariate skew polynomial rings, which are obtained as Ore localizations of G-algebras, in Section 2.2. Thus altogether, the chosen framework permits a large variety of relevant algebras, i.e., one improvement given here is the universality of the proposed method.
Certainly, we need to fix some requirements on the ring to ensure the existence of a decoupled form or rather, a normal form. More precisely, we assume the skew polynomial ring to be an Euclidean domain. The corresponding assumptions on the rings are given in Theorem 2.2.6. For more theoretical background on normal forms for matrices over those rings, we refer to [Coh71].
Apart of the goal to introduce an algorithm permitting a variety of interesting rings, we want to get a grip on the swell of the entries in the transformation matrices. This problem, known from the Smith form, becomes even worse with the non-commutative analog, the Jacobson form, see Remark 4.1.12.
We present an approach to calculate a decoupled form or rather, a normal form, which is based on Gröbner bases. To the best of our knowledge, this idea was originally

introduced for the commutative ring $K[x]$ in [Her05]. There, an algorithm for the computation of the Smith normal form was presented. We generalize this idea to our framework and develop a polynomial strategy. We stress that this approach is completely constructive. It is important that it can be realized in any computer algebra system which can handle G-algebras or polynomial Ore algebras.
Note that the crucial improvement of the proposed method is introduced in Subsection 4.1.1, where we show in detail how to handle the problem in a completely fraction-free polynomial framework. We point out advantages of the polynomial strategy and illustrate some of them with interesting examples in Subsection 4.2.1. At the same time, we compare our results with the output of the implementation of an algorithm which uses fractions directly. In many examples, our approach delivers much more compact results with small coefficients. We want to stress that these examples have not been specially selected for this purpose; instead, we picked a couple from a bigger family of examples. In our opinion, this phenomenon is quite ubiquitous.

The implementation is realized as a library called `jacobson.lib` for the computer algebra system SINGULAR::PLURAL [GPS05, GLH05], and it has been incorporated into the official distribution of SINGULAR (version 3-1-0).
There are other implementations, which are realized in MAPLE. The implementation [CQ05] works for Ore algebras, while the implementation [Mid08] is done for the first Weyl algebra with coefficients in a differential field. Middeke [Mid08] has reported that the classical algorithm for computing the Jacobson form of a matrix over the Weyl algebra over a differential field is polynomial-time. The complexity of our implementation is left open. But it seems to us (due to the polynomial strategy approach), that the subalgebra of invertible elements must be involved in the complexity analysis. Perhaps one should consider different models for studying complexity, since experience with practical applications suggests that the important role played by the coefficient arithmetics (which is not the arithmetics over a number field anymore) must be appropriately reflected in the overall complexity. Otherwise, the complexity of operations over the skew field of invertible elements remains hidden.
The results of this chapter are accepted for publication in the Journal of Symbolic Computation [SL].

Within this chapter, let A be a division ring and suppose $\mathcal{D} = A[\partial; \sigma, \delta]$ to be a skew polynomial ring that is Euclidean, see Theorem 2.2.6. Recall that the degree of $s \in \mathcal{D}$ is defined as the highest exponent of ∂ appearing in s.

4.1 Decoupling systems over Ore extensions

The proposed algorithm will require to work with left and right modules over $\mathcal{D}$. Operations from the left corresponds to manipulations on the rows, and operations from the right to manipulations on the columns. Thus one has to use both $\mathcal{D}$ and its opposite algebra $\mathcal{D}^{op}$. Recall that $\mathcal{D}^{op}$ is the same set as $\mathcal{D}$, endowed with the opposite multiplication, that is, for all $a, b \in \mathcal{D}^{op}$, the equality $a \star_{\mathcal{D}^{op}} b = b \star_{\mathcal{D}} a$ holds, where $\star_{\mathcal{D}}$ stands for the multiplication in the algebra $\mathcal{D}$. A natural map turns a right (resp. left) $\mathcal{D}$-module into a left (resp. right) $\mathcal{D}^{op}$-module. There is an algorithmic procedure

to set up an opposite algebra for a given G-algebra, see [Lev05a]. Alternatively, for "swapping sides" one can employ an involution on $\mathcal{D}^{g\times q}$. For this purpose, we define an involution on a ring first.

Definition 4.1.1 An automorphism θ of the additive group of $\mathcal{D}$ which satisfies

$$\theta(ab) = \theta(b)\theta(a) \quad \text{for all } \ a, b \in \mathcal{D}$$

is called an anti-automorphism of $\mathcal{D}$. If moreover, θ satisfies $\theta \circ \theta = \mathrm{id}$, then θ is called an **involution** of $\mathcal{D}$.

Note that since $\mathcal{D}$ is a domain an anti-automorphism maps 1 to 1.

Remark 4.1.2 Note that for every left ideal ${}_{\mathcal{D}}\langle f_1, \ldots, f_k\rangle = I \subseteq \mathcal{D}$, its image $\theta(I)$ under the involution becomes a right ideal. This follows since the equality

$$\sum_{j=1}^{k} \theta(f_j)a_j = \theta^2(\sum_{j=1}^{k} \theta(f_j)a_j) = \theta(\sum_{j=1}^{k} \underbrace{\theta(a_j)}_{\in\mathcal{D}} f_j) \in \theta(I)$$

holds for all $a_j \in \mathcal{D}$. One can show that involutions preserve the degree.

In classical operator algebras, particularly simple involutions are known [CQR04]. A constructive advantage of using involutions versus using opposite algebras lies in the fact that one does not need to create the opposite algebra and associate to any object its opposite. Instead, we apply an involution to an object and remain in the same ring.
Now we still need to extend the introduced map to $\mathcal{D}^{g\times g}$. An involution can be defined on matrices as follows. Let $\theta : \mathcal{D} \to \mathcal{D}$ be an involution as above. We define

$$\widetilde{\theta} : \mathcal{D}^{g\times g} \to \mathcal{D}^{g\times g}, \quad R \mapsto (\theta(R))^T,$$

where R^T is the transposed matrix of R, and $\theta(R) = [\theta(R_{ij})], 1 \le i, j \le g$. Then one can easily show that

$$\widetilde{\theta}(B \cdot C) = \widetilde{\theta}(C) \cdot \widetilde{\theta}(B) \quad \text{for all } B, C \in \mathcal{D}^{g\times g}$$

and, moreover, $(\widetilde{\theta})^2 = \mathrm{id}_{g\times g}$.
Let ${}_{\mathcal{D}}R$ denote the left $\mathcal{D}$-module generated by the rows of a matrix R, and $R_{\mathcal{D}}$ the right $\mathcal{D}$-module generated by the columns of a matrix R.

Remark 4.1.3 Extending Remark 4.1.2 to the module case, we obtain

$${}_{\mathcal{D}}R = {}_{\mathcal{D}}\widetilde{\theta}(\widetilde{\theta}({}_{\mathcal{D}}R)_{\mathcal{D}}) \qquad \text{and} \qquad R_{\mathcal{D}} = \widetilde{\theta}({}_{\mathcal{D}}\widetilde{\theta}(R_{\mathcal{D}}))_{\mathcal{D}}.$$

Thus all feasible left operations executed in ${}_{\mathcal{D}}R$, that is, on the rows of R, correspond to suitable ones in $\widetilde{\theta}({}_{\mathcal{D}}R)_{\mathcal{D}}$, that is, right operations on the columns of R under $\widetilde{\theta}$. Analogously, this holds for $R_{\mathcal{D}}$.

Note that in case of $\mathcal{D}$ being commutative, θ can be chosen to be the identity and $\widetilde{\theta}$ becomes simply the transposition.

In the sequel, let θ denote an involution on $\mathcal{D}$ and $\widetilde{\theta}$ an involution on $\mathcal{D}^{g\times g}$, respectively. Indeed, the map $\widetilde{\theta}$ can easily be extended to the set of non-square matrices by setting $\widetilde{\theta}(A\cdot C) = \widetilde{\theta}(C)\cdot\widetilde{\theta}(A)$ for $A \in \mathcal{D}^{g\times q}$, $C \in \mathcal{D}^{q\times k}$. Applied twice, we get back $A\cdot C$.

Let $R \in \mathcal{D}^{g\times q}$. In the spirit of system-theoretical applications, we assume, without loss of generality, that $g \leq q$. Then there exist unimodular matrices $U \in \mathcal{D}^{g\times g}$ and $V \in \mathcal{D}^{q\times q}$ such that

$$URV = [D, 0], \quad \text{where} \quad D = \operatorname{diag}(r_1, \ldots, r_g).$$

Recall that a matrix $U \in \mathcal{D}^{g\times g}$ is called **unimodular** if and only if there exists $U^{-1} \in \mathcal{D}^{g\times g}$ such that $UU^{-1} = U^{-1}U = \mathrm{I}_g$. There are several ways to prove the claim, all based on the Euclidean (and thus PID) property of the underlying ring. We present an algorithm to obtain the unimodular matrices via Gröbner bases. The proof that the algorithm terminates with the desired result will give an additional way to verify the claim. The main idea about the computation is an alternation between the computation of a reduced Gröbner basis of the submodule generated by, say, the rows of a matrix, and the application of the involution $\widetilde{\theta}$. As already mentioned, this idea was applied to $K[x]$ in order to compute the Smith normal form in the PhD thesis [Her05]. However, in the commutative case, θ becomes the identity map and thus $\widetilde{\theta}$ yields the matrix transposition.

Recall that by ${}_{\mathcal{D}}R$ we denote the left $\mathcal{D}$-module generated by the rows of a matrix R. Further on, by $\mathcal{G}({}_{\mathcal{D}}R)$ we denote the reduced left Gröbner basis of the submodule ${}_{\mathcal{D}}R$ with respect to a module ordering $<$ giving priority to the last component. More precisely, for $r, s \in \operatorname{Mon}(\mathcal{D})$ and $\mathrm{e}_i := (0, \cdots, 0, \underbrace{1}_{i}, 0, \cdots, 0)$

$$r\,\mathrm{e}_i < s\,\mathrm{e}_j \quad :\Leftrightarrow \quad i < j \text{ or if } i = j \text{ then } \deg(r) < \deg(s).$$

Note that this is different from the choice in Definition 2.1.6, where the priority was given to the first component. For the i-th row of a matrix R, we write R_i and R_{ij} stands, as usual, for the entry in the i-th row and j-th column. Recall that $\operatorname{lpos}(g)$ denotes the leading position for all $g \in \mathcal{D}^{1\times q}$. Define the **degree** of an element $0 \neq a \in \mathcal{D}^{1\times q}$ to be the degree of the corresponding leading monomial, that is, $\deg(a) := \deg(\operatorname{lm}(a))$. This notion differs from the notion of "row degree" used in previous chapters.
Note that the elements of $\mathcal{G}({}_{\mathcal{D}}R)$ have pairwise distinct leading monomials, since they form a reduced Gröbner basis.

Due to the definition of a reduced Gröbner basis $\operatorname{lm}(\mathcal{G}({}_{\mathcal{D}}R)_i)$ divides $\operatorname{lm}(\mathcal{G}({}_{\mathcal{D}}R)_j)$ for $\mathcal{G}({}_{\mathcal{D}}R)_i, \mathcal{G}({}_{\mathcal{D}}R)_j \in \mathcal{G}({}_{\mathcal{D}}R)$ if and only if $\mathcal{G}({}_{\mathcal{D}}R)_i = \mathcal{G}({}_{\mathcal{D}}R)_j$. Thus we can suppose without loss of generality that

$$\operatorname{lm}(\mathcal{G}({}_{\mathcal{D}}R)_1) < \cdots < \operatorname{lm}(\mathcal{G}({}_{\mathcal{D}}R)_m).$$

Moreover, we can formulate an even more precise lemma.

From now on, let us assume that $g = q$ and R is of full rank, that is, row and column ranks of R are equal to g first.

Lemma 4.1.4 *As before, assume that* $\mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_1) < \cdots < \mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_m)$. *Then*

$$\begin{bmatrix} \mathcal{G}({}_{\mathcal{D}}R)_1 \\ \vdots \\ \mathcal{G}({}_{\mathcal{D}}R)_m \end{bmatrix}$$

is a lower triangular matrix.

Proof: Suppose the claim does not hold. Then there exists $\mathcal{G}({}_{\mathcal{D}}R)_i$ and $\mathcal{G}({}_{\mathcal{D}}R)_j$ with $\mathrm{lpos}(\mathcal{G}({}_{\mathcal{D}}R)_i) = \mathrm{lpos}(\mathcal{G}({}_{\mathcal{D}}R)_j)$ for $i < j$. Thus $\mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_i) = \partial^{\alpha_i} \mathrm{e}_k$ and $\mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_j) = \partial^{\alpha_j} \mathrm{e}_k$ such that $\alpha_i < \alpha_j$. But then evidently $\mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_i)$ divides $\mathrm{lm}(\mathcal{G}({}_{\mathcal{D}}R)_j)$, which is a contradiction to $\mathcal{G}({}_{\mathcal{D}}R)$ being reduced. □

Due to the previous lemma, we may assume the matrix induced by $\mathcal{G}({}_{\mathcal{D}}R)$ to be lower triangular without loss of generality. We need some further preparation to introduce the algorithm and to prove its correctness.

Lemma 4.1.5 *Let* $\mathcal{I}$ *denote the left ideal generated by the elements in the last column of* $\widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R))$, *that is, by* $\theta(\mathcal{G}({}_{\mathcal{D}}R)_{g1}), \ldots, \theta(\mathcal{G}({}_{\mathcal{D}}R)_{gg})$. *Then*

$$\mathcal{I} = {}_{\mathcal{D}}\langle\, \mathcal{G}({}_{\mathcal{D}}\widetilde{\theta}(\, \mathcal{G}({}_{\mathcal{D}}R)\,)\,)_{gg}\,\rangle.$$

Proof: Note, that due to Lemma 4.1.4

$$\underbrace{\begin{bmatrix} * & & \\ \vdots & \ddots & \\ \mathcal{G}({}_{\mathcal{D}}R)_{g1} & \cdots & \mathcal{G}({}_{\mathcal{D}}R)_{gg} \end{bmatrix}}_{\mathcal{G}({}_{\mathcal{D}}R)} \overset{\widetilde{\theta}}{\rightsquigarrow} \begin{bmatrix} & & \theta(\mathcal{G}({}_{\mathcal{D}}R)_{g1}) \\ & \iddots & \vdots \\ * & \cdots & \theta(\mathcal{G}({}_{\mathcal{D}}R)_{gg}) \end{bmatrix} \overset{\mathcal{G}}{\rightsquigarrow} \begin{bmatrix} * & & \\ \vdots & \ddots & \\ * & \cdots & \mathcal{G}({}_{\mathcal{D}}\widetilde{\theta}(\, \mathcal{G}({}_{\mathcal{D}}R)\,)\,)_{gg} \end{bmatrix}.$$

According to the definition of $\mathcal{G}$ the left ideal generated by $\mathcal{G}({}_{\mathcal{D}}\widetilde{\theta}(\, \mathcal{G}({}_{\mathcal{D}}R)\,)\,)_{gg}$ coincides with ${}_{\mathcal{D}}\langle \theta(\mathcal{G}({}_{\mathcal{D}}R)_{g1}), \ldots, \theta(\mathcal{G}({}_{\mathcal{D}}R)_{gg})\rangle$. □

Now we can formulate the algorithm that yields the desired diagonal form.

Input : $R \in \mathcal{D}^{g\times g}$ of full rank, $\widetilde{\theta}$ involution as above
Output: Matrices $U, V, D \in \mathcal{D}^{g\times g}$, such that
$U \cdot R \cdot V = \mathrm{diag}(r_1, \ldots, r_g) = D$, where U, V are unimodular

$R^{(0)} \leftarrow R$, $U \leftarrow \mathrm{I}_g$, $V \leftarrow \mathrm{I}_g$
$i \leftarrow 0$
while *($R^{(i)}$ is not a diagonal matrix* **or** $i \equiv_2 1$*)* **do**
  $i \leftarrow i + 1$
  Compute $U^{(i)}$ such that $U^{(i)} \cdot R^{(i-1)} = \mathcal{G}({}_{\mathcal{D}}R^{(i-1)})$
  $R^{(i)} \leftarrow \widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R^{(i-1)}))$
  if *($i \equiv_2 0$)* **then**
    $V \leftarrow V \cdot \widetilde{\theta}(U^{(i)})$
  end
  else
    $U \leftarrow U^{(i)} \cdot U$
  end
end
return $(U, V, R^{(i)})$

Algorithm 2: Diagonalization with Gröbner Bases

Theorem 4.1.6 *The Algorithm 2 terminates and it is correct.*
That is, for $R \in \mathcal{D}^{g\times g}$, let $R^{(i)}$ denote the matrix we get after the i-th execution of the ***while*** *loop. Then there exists an element $k \in \mathbb{N}$ such that $R^{(k)}$ is a diagonal matrix. If k is odd, then the* ***while*** *loop is repeated just one more time (define $l := k+(k \mod 2)$ in this case). In both cases, the* ***while*** *loop is terminated by its condition. The matrices U, V obtained in the last loop are unimodular and satisfy $URV = \mathrm{diag}(r_1, \ldots, r_g)$.*

Proof: We prove the claim by induction on g, the size of the square matrix R. For $g = 1$, there is nothing to show.
Using Lemma 4.1.5, the equality ${}_{\mathcal{D}}\langle\theta((R^{(i+1)})_{gg})\rangle = {}_{\mathcal{D}}\langle(R^{(i)})_{1g}, (R^{(i)})_{2g}, \ldots, (R^{(i)})_{gg}\rangle$ holds. Hence we get

$$ {}_{\mathcal{D}}\left\langle(R^{(i)})_{gg}\right\rangle \subseteq {}_{\mathcal{D}}\left\langle\theta((R^{(i+1)})_{gg})\right\rangle \qquad \text{for all } i.$$

By degree arguments, this implies that ${}_{\mathcal{D}}\langle(R^{(r)})_{gg}\rangle = {}_{\mathcal{D}}\langle\theta((R^{(r+1)})_{gg})\rangle$ for some r. Using Lemma 4.1.5 and $(R^{(r)})_{gg} \neq 0$ (since R is of full rank), we obtain that $(R^{(r)})_{gg}$ is a right divisor of $(R^{(r)})_{ig}$ for each $1 \leq i \leq g-1$. Then the definition of $\mathcal{G}$ yields that

$$R^{(r+1)} = \begin{pmatrix} & & & 0 \\ & R' & & \vdots \\ & & & 0 \\ 0 & \ldots & 0 & (R^{(r+1)})_{gg} \end{pmatrix}.$$

The $(g-1)\times(g-1)$ matrix R' can be transformed to a diagonal matrix via unimodular operations by induction. It remains to consider the transformation matrices U and V.

For each $i \in \mathbb{N}$, after executing the **while** loop i times, we obtain

$$\begin{cases} R^{(i)} = U^{(i-1)} \cdot U^{(i-3)} \cdots U^{(1)} \cdot \ R \ \cdot \widetilde{\theta}(U^{(2)}) \cdot \widetilde{\theta}(U^{(4)}) \cdots \widetilde{\theta}(U^{(i)}), & \text{if } i \text{ is even} \\ R^{(i)} = U^{(i-1)} \cdot U^{(i-3)} \cdots U^{(1)} \cdot \ \widetilde{\theta}(R) \ \cdot \widetilde{\theta}(U^{(2)}) \cdot \widetilde{\theta}(U^{(4)}) \cdots \widetilde{\theta}(U^{(i)}), & \text{if } i \text{ is odd,} \end{cases}$$

which completes the proof. □

Remark 4.1.7 Note that the algorithm requires only Gröbner basis computations for row modules, and not for column modules. This is an advantage in practice, since most non-commutative computer algebra systems provide either implementations, not both.

Let us illustrate the algorithm with the following example. At first, we consider a matrix over a commutative ring. Non-commutative examples appear in Example 4.1.22 and in Subsection 4.2.1.

Example 4.1.8 Let $\mathcal{D} = K[x]$ and let $\widetilde{\theta}$ be the transposition on $\mathcal{D}^{2\times 2}$. Consider

$$R = \begin{bmatrix} x^2-1 & x+1 \\ x^3+x^2+1 & -x \end{bmatrix} \in \mathcal{D}^{2\times 2}.$$

Then $R^{(0)} := R$, $U = V = \mathrm{I}_2$ and $i = 0$.

1: Since $R^{(0)}$ is not diagonal, we enter the while loop

- $i \leftarrow 1$
- Since $\begin{bmatrix} x & x+1 \\ 1 & 1 \end{bmatrix} R^{(0)} = \mathcal{G}({}_{\mathcal{D}}R^{(0)})$ and $i \equiv_2 1$

  $R^{(1)} \leftarrow \begin{bmatrix} x^4+3x^3+x^2+1 & x^3+2x^2 \\ 0 & 1 \end{bmatrix}$

  $U \leftarrow \begin{bmatrix} x & x+1 \\ 1 & 1 \end{bmatrix}$

2: Since $R^{(1)}$ is not diagonal, we enter the while loop

- $i \leftarrow 2$
- Since $\begin{bmatrix} 1 & -x^3-2x^2 \\ 0 & 1 \end{bmatrix} R^{(1)} = \mathcal{G}({}_{\mathcal{D}}R^{(1)})$ and $i \equiv_2 0$

  $R^{(2)} \leftarrow \begin{bmatrix} x^4+3x^3+x^2+1 & 0 \\ 0 & 1 \end{bmatrix}$

  $V \leftarrow \begin{bmatrix} 1 & 0 \\ -x^3-2x^2 & 1 \end{bmatrix}$

3: Since i is even and $R^{(2)}$ is diagonal, the **while** loop terminates. The algorithm returns U, V and $R^{(2)}$.

A routine check ensures that $URV = R^{(2)}$.

Remark 4.1.9 In order to extend Theorem 4.1.6 and Algorithm 2 to non-square and non-full rank matrices, we need to add suitable syzygies to U and V, and zero rows and columns to the diagonal matrix, in order to maintain the initial size of R. For a computational solution, it is sufficient to extend Algorithm 2 in the following way. Let $R^{(i)} \in \mathcal{D}^{s\times t}$, where either $s = g, t = q$ or $s = q, t = g$ in the i-th while loop. Instead of computing $U^{(i)}$, satisfying $U^{(i)} \cdot R^{(i-1)} = \mathcal{G}({}_{\mathcal{D}}R^{(i-1)})$, we compute $\mathcal{G}({}_{\mathcal{D}}\tilde{R})$ for the extended matrix $\tilde{R} := [\mathrm{I}_s, R^{(i-1)}]$. Then $\tilde{R}$ is obviously a full row rank matrix. Defining

$$U^{(i)} := [\mathcal{G}({}_{\mathcal{D}}\tilde{R})_1^T, \ldots, \mathcal{G}({}_{\mathcal{D}}\tilde{R})_s^T]^T, \text{ and } R^{(i)} := [\mathcal{G}({}_{\mathcal{D}}\tilde{R})_{s+1}^T, \ldots, \mathcal{G}({}_{\mathcal{D}}\tilde{R})_t^T]^T$$

it is easy to see that

$$U^{(i)}R^{(i-1)} = R^{(i)}.$$

The matrix $R^{(i)}$ consists of the rows of $\mathcal{G}({}_{\mathcal{D}}R^{(i-1)})$ and additional zero rows, such that $R^{(i)} \in \mathcal{D}^{s\times t}$.

4.1.1 Polynomial decoupling

We are given a matrix R over a non-commutative Euclidean domain $\mathcal{D}$. In this section, we show our main result of this chapter. We introduce a method that allows to execute Algorithm 2 in a completely polynomial (that is, fraction-free) framework. The idea comes from the commutative case and was elaborated e.g. in [GTZ98].
To specify this, let A_* denote a G-algebra and $A = \mathrm{Quot}(A_*)$. Moreover, let $\mathcal{D} = A[\partial; \sigma, \delta]$ and let $\mathcal{D}_*$ denote $A_*[\partial; \sigma, \delta]$, which is a G-algebra. For a detailed introduction to G-algebras and the connection to Ore localizations, we refer to Subsection 2.2. Evidently $\mathcal{D}_* \subseteq \mathcal{D}$, since $A_* \subseteq A$. Without loss of generality, we suppose that R does not contain a zero row.
We define the **degree** of an element in $\mathcal{D}_*$ or $\mathcal{D}_*^{1\times q}$ analogously to $\mathcal{D}$ or $\mathcal{D}^{1\times q}$, that is, by assigning degree 0 to the nonzero elements of A_* and degree 1 to ∂. In particular, the degree is invariant under the multiplication by elements in A_*.

Lemma 4.1.10 *Let $R \in \mathcal{D}^{g\times q}$. Then there exists a $\mathcal{D}$-unimodular matrix $T \in \mathcal{D}_*^{g\times g}$ such that $TR \in {\mathcal{D}_*}^{g\times q}$.*

Proof: If $R \in {\mathcal{D}_*}^{g\times q}$, there is nothing to do, so assume R contains elements with fractions. At first, we show how to bring two fractional elements $a^{-1}b, c^{-1}d$ for $a, c \in A_*$, $b, d \in \mathcal{D}_*$ to a common left denominator, cf. [Ape88]. For any $h_1, h_2 \in A_*$ such that $h_1a = h_2c$, it is easy to see that

$$(h_1a)^{-1}(h_1b) = a^{-1}h_1^{-1}h_1b = a^{-1}b \text{ and } (h_1a)^{-1}(h_2d) = (h_2c)^{-1}(h_2d) = c^{-1}d,$$

hence $h_1a = (a^{-1}h_1^{-1})^{-1} = h_2c$ is a common left denominator. Analogously, we can compute a common left denominator for any finite set of fractions, hence we can do it for any element in $\mathcal{D}$ as well as for any row vector in $\mathcal{D}^{1\times q}$. Let T_{ii} be the common left denominator of non-zero elements from the i-th row of R, then TR contains no fractions. Moreover, T is a diagonal matrix with non-zero polynomial entries, so it is $\mathcal{D}$-unimodular. □

Remark 4.1.11 Note that the computation of compatible factors h_i for $a_1, a_2 \in A_*$ can be achieved by computing syzygies, since $\{(h_1, h_2) \in A_*^2 \mid h_1 a_1 = h_2 a_2\}$ is exactly the module $\mathrm{Syz}(a_1, -a_2) \subset A_*^2$. The factors h_i for more than two a_i's can be obtained by iterating this procedure.

Define $R_* := TR \in \mathcal{D}_*^{g\times q}$ using the notation of Lemma 4.1.10. Then the relations ${}_{\mathcal{D}_*}R_* \subseteq {}_{\mathcal{D}}R$ and ${}_{\mathcal{D}}R_* = {}_{\mathcal{D}}R$ hold obviously. Thus whenever we speak about a finitely generated submodule ${}_{\mathcal{D}}R \subset \mathcal{D}^{1\times q}$, then ${}_{\mathcal{D}}R_*$ will denote a presentation of ${}_{\mathcal{D}}R$ with generators contained in $\mathcal{D}_*$. In what follows, we will show how to find $\mathcal{D}$-unimodular matrices $U \in \mathcal{D}_*^{g\times g}$ and $V \in \mathcal{D}_*^{q\times q}$ such that

$$U(TR)V = \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_q \\ & 0 & \end{bmatrix} \in \mathcal{D}_*^{g\times q}.$$

Since the equality $U(TR)V = (UT)RV$ holds and UT is a $\mathcal{D}$-unimodular matrix, our initial aim follows.
As in the previous subsection, by $\mathcal{G}({}_{\mathcal{D}_*}R_*)$ we denote the reduced left Gröbner basis of the submodule ${}_{\mathcal{D}_*}R_*$ with respect to the module ordering $<_*$, giving priority to the last component and a lexicographical ordering on $\mathcal{D}_*$, that is, satisfying

$$\partial > x_n > \cdots > x_1.$$

For $r, s \in \mathrm{Mon}(\mathcal{D}_*)$, this extends to

$$r\,\mathrm{e}_i < s\,\mathrm{e}_j \quad :\Leftrightarrow \quad i < j \text{ or if } i = j \text{ then } r < s.$$

Unlike the rational case, the leading monomials of elements in $\mathcal{D}_*^{1\times g}$ are of the form $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \partial^\beta$ for $\alpha_i, \beta \in \mathbb{N}$.

Remark 4.1.12 Using the polynomial strategy, two improvements can be observed. On the one hand, the quotient field has not to be used at all, once we have mapped the matrix we work with from $\mathcal{D}^{g\times q}$ to $\mathcal{D}_*^{g\times q}$. The other improvement lies in the nature of generation of normal forms for matrices and the corresponding transformation matrices. The naive approach would be to apply elementary operations including division by invertibles on the rows and columns, that is, operations from the left and from the right. Indeed, there are methods using different techniques like for instance p-adic arguments to calculate the invariant factors of the Smith form over the integers [Lüb02]. However, this method does not help with the generation of transformation matrices. Obviously, the swap from left to right has no influence in the commutative framework. But suppose for instance $\mathcal{D}$ to be the rational Weyl algebra B_1, see Example 2.2.5. Then $\frac{1}{x}$ is an unit in B_1 and

$$\partial \frac{1}{x} = \frac{1}{x}\partial - \frac{1}{x^2}.$$

Comparing the multiplication with the inverse element, that is, with x, we see that $\partial x = x\partial + 1$ holds. Thus a multiplication of any polynomial containing ∂ with the element $\frac{1}{x}$ in the field of fractions causes an immediate coefficient swell. Since a normal

form of a matrix is given modulo unimodular operations, the previous example illustrates the variations of possible representations. In Section 4.2.1, we will give examples where the polynomial method dams up the coefficient increase in a very impressive way.

But note that on the other hand, the changeover to the polynomial framework brings a problem. The underlying ring $\mathcal{D}_*$ is not a PID anymore, as we can already see from the form of the leading monomials. However, this is the essential property for the existence of a diagonal form. In the sequel, we show how this problem can be resolved by fixing the chosen module ordering and a suitable sorting condition. Referring to the argumentation of Remark 4.1.4 yields the block-diagonal form

$$\mathcal{G}({}_{\mathcal{D}_*}R_*) = \begin{bmatrix} 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \\ \boxed{*} & & & \\ \vdots & & 0 & \\ * & & & \\ & \boxed{*} & & \\ & \vdots & & 0 \\ & * & & \\ & & \ddots & \\ & & & \boxed{*} \\ & * & & \vdots \\ & & & * \end{bmatrix}, \tag{4.1}$$

where the rows with the boxed element have the smallest leading term with respect to the chosen ordering in the corresponding block. A block denotes all elements of the same leading position in $\mathcal{G}({}_{\mathcal{D}_*}R_*)$. In Theorem 4.1.17, we show that exactly these elements generate ${}_{\mathcal{D}}R$. Further in Lemma 4.1.15, we will see that these elements provide us with additional information. However, this result demands some more preparation.

Lemma 4.1.13 *Let P be $\mathcal{D}$ or $\mathcal{D}_*$. For $R \in P^{g\times q}$ of full row rank and every $1 \leq i \leq g$, define $\alpha_i := \min\{\deg(a) \mid a \in {}_PR\backslash\{0\}$ and $\operatorname{lpos}(a) = i\}$. Then for all $1 \leq i \leq g$, there exists $h_i \in \mathcal{G}({}_PR)$ of degree α_i with $\operatorname{lpos}(h_i) = i$.*

Proof: Recall that $\partial > x_j$ for all j. Let $f \in {}_PR$ with $\operatorname{lpos}(f) = i$ and $\deg(f) = \alpha_i$. Suppose that for all $g \in \mathcal{G}({}_PR)$ with leading position i, $\deg(g) > \alpha_i$ holds. Since $\mathcal{G}({}_PR)$ is a Gröbner basis, there exists $g \in \mathcal{G}({}_PR)$ such that $\operatorname{lm}(g)$ divides $\operatorname{lm}(f)$. This happens if and only if $\deg(g) \leq \deg(f)$, because $\mathcal{D}_*$ is a G-algebra and $\mathcal{D}$ is an Ore PID. This yields a contradiction. □

The full rank assumption in the previous lemma guaranties the existence of α_i for each component $1 \leq i \leq g$. Note that over $\mathcal{D}_*$, there can exist more than one element in $\{\deg(a) \mid a \in {}_PR\backslash\{0\}$ and $\operatorname{lpos}(a) = i\}$. We propose the selection strategy described in Theorem 4.1.17.

Corollary 4.1.14 *Lemma 4.1.13 and Lemma 4.1.4 yield*

$$\deg(\mathcal{G}({}_{\mathcal{D}}R)_i) = \min\{\deg(a) \mid a \in {}_{\mathcal{D}}R \text{ and } \operatorname{lpos}(a) = i\}.$$

Lemma 4.1.15 *Let α_i be the degree of the boxed entry with leading position in the i-th column, that is*

$$\alpha_i := \deg(\min_{<_*}\{b \mid b \in \mathcal{G}({}_{\mathcal{D}_*}R_*) \text{ and } \operatorname{lpos}(b) = i\}).$$

Then $\deg(\operatorname{lm}(h)) \geq \alpha_i$ *for all* $h \in {}_{\mathcal{D}}R$ *with* $\operatorname{lpos}(h) = i$.

Proof: Now suppose the claim does not hold and there is $h \in {}_{\mathcal{D}}R$ with $\operatorname{lpos}(h) = i$ of degree smaller than α_i. Using Lemma 4.1.10, there exists $a \in A_*$ such that $ah \in {}_{\mathcal{D}_*}R_*$. Then $\deg(ah) = \deg(h)$ and $\operatorname{lpos}(ah) = i$. Due to Lemma 4.1.13, $\deg(f) \geq \alpha_i$ for all $f \in {}_{\mathcal{D}_*}R_*$ with leading position i, hence we obtain a contradiction. □

Corollary 4.1.16 *Lemma 4.1.15 and Corollary 4.1.14 provide for all* $1 \leq i \leq g$ *the equality*

$$\min\{\deg(a) \mid a \in {}_{\mathcal{D}}R \text{ and } \operatorname{lpos}(a) = i\} = \min\{\deg(a) \mid a \in {}_{\mathcal{D}_*}R_* \text{ and } \operatorname{lpos}(a) = i\}.$$

Theorem 4.1.17 *Let $R \in \mathcal{D}^{g\times g}$ be of full rank. For each $1 \leq i \leq g$, let us define*

$$b_i := \min_{<_*}\{b \mid b \in \mathcal{G}({}_{\mathcal{D}_*}R_*) \text{ and } \operatorname{lpos}(b) = i\}.$$

Note that the set $\{b_1, \ldots, b_g\}$ corresponds to the subset of all rows with a boxed entry in the block triangular form (4.1). Moreover ${}_{\mathcal{D}}\langle b_1, \ldots, b_g\rangle = {}_{\mathcal{D}}R$.

Remark 4.1.18 Note that the minimum b_i, defined in the previous theorem, exists for each $1 \leq i \leq g$, since R is of full rank.

Proof: Let $f \in {}_{\mathcal{D}}R$. Due to Corollary 4.1.16, there exists $1 \leq k \leq g$ such that $\operatorname{lpos}(b_k) = \operatorname{lpos}(f)$ and $\deg(b_k) \leq \deg(f)$. Thus there exists an element $s_k \in \mathcal{D}$ providing that $\deg(f - s_k b_k) < \deg(b_k)$. Since $f - s_k b_k \in {}_{\mathcal{D}}R$, Corollary 4.1.16 implies that we have $\operatorname{lpos}(f - s_k b_k) < \operatorname{lpos}(f)$. Iterating this reduction leads to the remainder zero and thus $f = \sum_{i=1}^{k} s_i b_i$. □

Using the notation of the previous theorem, let

$$G^*({}_{\mathcal{D}}R) := \begin{bmatrix} b_1 \\ \vdots \\ b_g \end{bmatrix},$$

which is by definition a lower triangular matrix. In the sequel, let $R \in \mathcal{D}^{g\times g}$ be of full row rank. Note that then obviously $G^*({}_{\mathcal{D}}R)$ is a square matrix.

Proposition 4.1.19 Suppose $R \in \mathcal{D}^{g\times g}$ is a full row rank matrix and there is $U_* \in \mathcal{D}_*^{l\times g}$ such that $U_*R_* = \mathcal{G}({}_{\mathcal{D}_*}R_*)$. Let us select the indices

$$\{t_1,\dots,t_g\} \subseteq \{1,\dots,l\} \text{ such that } \{(U_*R_*)_{t_1},\dots,(U_*R_*)_{t_g}\} = G^*({}_{\mathcal{D}}R) \tag{4.2}$$

(see notation of Theorem 4.1.17). Then $U := [(U_*)_{t_1},\dots,(U_*)_{t_g}]^T$ is $\mathcal{D}$-unimodular in $\mathcal{D}^{g\times g}$ and $UR_* = G^*({}_{\mathcal{D}}R)$.

Proof: The equality $UR_* = G^*({}_{\mathcal{D}}R)$ follows by the definition of U. Still left to show is that U is $\mathcal{D}$-unimodular. Note that ${}_{\mathcal{D}}(UR_*) = {}_{\mathcal{D}}G^*({}_{\mathcal{D}}R) = {}_{\mathcal{D}}R = {}_{\mathcal{D}}R_*$ holds and $UR_* \in \mathcal{D}^{g\times g} \ni R_*$. Thus there exists $V \in \mathcal{D}^{g\times g}$ such that $R_* = V(UR_*)$. Then $VU = \mathrm{I}_g$ and so one can easily show that $UV = \mathrm{I}_g$, since R has full row rank. □

Lemma 4.1.20 *The equality of the following left ideals holds:*

$${}_{\mathcal{D}}\langle\theta(G^*({}_{\mathcal{D}}R)_{g1}),\dots,\theta(G^*({}_{\mathcal{D}}R)_{gg})\rangle = {}_{\mathcal{D}}\langle G^*(\widetilde{\theta}(G^*({}_{\mathcal{D}}R))_{gg}\rangle.$$

Proof: Using the argumentation given in the proof of Lemma 4.1.5 we obtain

$${}_{\mathcal{D}}\langle\theta(G^*({}_{\mathcal{D}}R)_{g1}),\dots,\theta(G^*({}_{\mathcal{D}}R)_{gg})\rangle = {}_{\mathcal{D}}\langle\mathcal{G}(\widetilde{\theta}(G^*({}_{\mathcal{D}}R))_{gg}\rangle.$$

Note the module identities ${}_{\mathcal{D}}G^*({}_{\mathcal{D}}R) = {}_{\mathcal{D}}\mathcal{G}({}_{\mathcal{D}}R)$

$$\Rightarrow \quad \widetilde{\theta}(G^*({}_{\mathcal{D}}R))_{\mathcal{D}} = \widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R))_{\mathcal{D}} \quad \Rightarrow \quad {}_{\mathcal{D}}G^*(\widetilde{\theta}(G^*({}_{\mathcal{D}}R))) = {}_{\mathcal{D}}\mathcal{G}(\widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R))).$$

According to the latter identity and since $\mathcal{G}(\widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R))$ and $G^*(\widetilde{\theta}(G^*({}_{\mathcal{D}}R)))$ are lower triangular matrices, we obtain ${}_{\mathcal{D}}\langle\mathcal{G}(\widetilde{\theta}(\mathcal{G}({}_{\mathcal{D}}R))_{gg}\rangle = {}_{\mathcal{D}}\langle G^*(\widetilde{\theta}(G^*({}_{\mathcal{D}}R))_{gg}\rangle$. □

Now we are ready to formulate the polynomial version of Algorithm 2.

Input : $R \in \mathcal{D}^{g\times g}$ of full rank, θ an involution on $\mathcal{D}_*$ and $\widetilde{\theta}$ the extension of θ to $\mathcal{D}_*^{g\times g}$
Output: $\mathcal{D}$-unimodular matrices $U, V \in \mathcal{D}_*^{g\times g}$ such that $U \cdot R \cdot V = \mathrm{diag}(r_1,\dots,r_g)$

Find $T \in \mathcal{D}^{g\times g}$ unimodular such that $TR \in \mathcal{D}_*^{g\times g}$
$R^{(0)} \leftarrow TR, \quad U \leftarrow T, \quad V \leftarrow \mathrm{I}_g$
$i \leftarrow 0$
while $R^{(i)}$ *is not a diagonal matrix* **or** $i \equiv_2 1$ **do**
  $i \leftarrow i+1$
  Compute $U^{(i)}$ so that $U^{(i)} \cdot R^{(i-1)} = \mathcal{G}({}_{\mathcal{D}_*}R^{(i-1)}) \in \mathcal{D}_*^{l\times g}$
  Select $\{t_1,\dots,t_g\} \subseteq \{1,\dots,l\}$ as in (4.2)
  $U^{(i)} \leftarrow [(U^{(i)})_{t_1},\dots,(U^{(i)})_{t_g}]^T$
  $R^{(i)} \leftarrow \widetilde{\theta}(G^*({}_{\mathcal{D}}R))$
  if $i \equiv_2 0$ **then**
    $V \leftarrow V \cdot \widetilde{\theta}(U^{(i)})$
  end
  else
    $U \leftarrow U^{(i)} \cdot U$
  end
end
return $(U, V, R^{(i)})$

Algorithm 3: Polynomial diagonalization with Gröbner Bases

Theorem 4.1.21 *Algorithm 3 terminates with the claimed result.*

Proof: Using Proposition 4.1.19 and replacing Lemma 4.1.5 by Lemma 4.1.20 in the proof of Theorem 4.1.6 provides the claim. □

Let us again illustrate the algorithm via an example.

Example 4.1.22 Suppose $\mathcal{D} = K(x)[\partial; \mathrm{id}, \frac{d}{dx}]$ and $\mathcal{D}_* = K[x][\partial; \mathrm{id}, \frac{d}{dx}]$. Let us define a K-linear map

$$\theta : \mathcal{D}_* \to \mathcal{D}_*, \left\{ \begin{array}{c} \partial \mapsto -\partial \\ x \mapsto x. \end{array} \right.$$

Indeed, θ gives rise to an involution (anti-automorphism), which obeys $\theta(ab) = \theta(b)\theta(a)$ for all $a, b \in \mathcal{D}_*$. Let

$$R = \left[\begin{array}{cc} \partial^2 - 1 & \partial + 1 \\ \partial^2 + 1 & \partial - x \end{array} \right] \in \mathcal{D}^{2\times 2}.$$

Evidently $T = \mathrm{I}_2$ and thus $R^{(0)} := R$, $U = V = \mathrm{I}_2$ and $i = 0$.

1: Since $R^{(0)}$ is not diagonal, go into the while loop

- $i \leftarrow 1$
- Since $\left[\begin{array}{cc} -x\partial - \partial + x^2 + x + 1 & x\partial + \partial + x \\ -\partial^2 + x\partial - \partial + x + 2 & \partial^2 + 2\partial + 1 \\ \partial - x & -\partial - 1 \end{array} \right] R^{(0)} = \mathcal{G}({}_{\mathcal{D}_*}R^{(0)})$

  where $\mathcal{G}({}_{\mathcal{D}_*}R^{(0)}) = \left[\begin{array}{cc} x^2\partial^2 + 2x\partial^2 + \partial^2 + 2x\partial + 2\partial - x^2 - 1 & 0 \\ x\partial^3 + \partial^3 + x\partial^2 + 5\partial^2 - x\partial + 3\partial - x - 1 & 0 \\ -x\partial^2 - \partial^2 - 2\partial + x - 1 & 1 \end{array} \right]$

  and $i \equiv_2 1$

  $R^{(1)} \leftarrow \left[\begin{array}{cc} x^2\partial^2 + 2x\partial^2 + \partial^2 + 2x\partial + 2\partial - x^2 - 1 & -x\partial^2 - \partial^2 + x - 1 \\ 0 & 1 \end{array} \right]$

  $U \leftarrow \left[\begin{array}{cc} -x\partial - \partial + x^2 + x + 1 & x\partial + \partial + x \\ \partial - x & -\partial - 1 \end{array} \right]$

2: Since $R^{(1)}$ is not diagonal, go into the while loop

- $i \leftarrow 2$
- Since $\left[\begin{array}{cc} 1 & x\partial^2 + \partial^2 - x + 1 \\ 0 & 1 \end{array} \right] R^{(1)} = \mathcal{G}({}_{\mathcal{D}_*}R^{(1)})$ and $i \equiv_2 0$

  $R^{(2)} \leftarrow \left[\begin{array}{cc} x^2\partial^2 + 2x\partial^2 + \partial^2 + 2x\partial + 2\partial - x^2 - 1 & 0 \\ 0 & 1 \end{array} \right]$

  $V \leftarrow \left[\begin{array}{cc} 1 & 0 \\ t\partial^2 + \partial^2 + 2\partial - x + 1 & 1 \end{array} \right]$

3: Since i is even and $R^{(2)}$ is diagonal, the algorithm returns U and V.

And indeed, the algorithm outputs the claimed result, since

$$URV = \left[\begin{array}{cc} x^2\partial^2 + 2x\partial^2 + \partial^2 + 2x\partial + 2\partial - x^2 - 1 & 0 \\ 0 & 1 \end{array} \right].$$

Algorithm 3 can be extended to $R \in \mathcal{D}^{g\times q}$ along the lines already presented in Remark 4.1.9.

4.2 Normal forms for time-varying systems

As in the previous section, suppose $\mathcal{D}$ to be a left and right Euclidean domain. Inspired by the Smith form, we will focus on how to sharpen the result of the already discussed diagonal form. Following [Coh71, Jac43], we obtain the following theorem.

Theorem 4.2.1 *Every matrix $R \in \mathcal{D}^{g\times q}$ is equivalent to a certain diagonal matrix, namely* $\operatorname{diag}(r_1, \ldots, r_\ell, 0, \ldots, 0)$*, such that additionally*

$$\mathcal{D}r_{i+1}\mathcal{D} \subseteq r_i\mathcal{D} \cap \mathcal{D}r_i \tag{4.3}$$

holds for all $i = 1, \ldots, \min\{g, q\} - 1$.

Remark 4.2.2 [Jac43, Theorem 31] The elements r_i are unique up to an equivalence relation called **similarity**, which will be defined in Definition 4.2.7.

Definition 4.2.3 If the relation (4.3) is satisfied, then r_i is called a **total divisor** of r_{i+1}, short $r_i \mid\mid r_{i+1}$.

Note that if r_i is a total divisor of r_{i+1}, then r_i is a left and a right divisor of r_{i+1}, but not necesserily vice versa. A counterexample is given by every non-invertible non-zero element of a simple ring $\mathcal{D}$.

Remark 4.2.4

1. If $\mathcal{D}$ is simple, then $r_i \mid\mid r_{i+1}$ implies that r_i is a unit or r_{i+1} equals zero.
2. If $\mathcal{D}$ is commutative, then $r_i \mid\mid r_{i+1}$ if and only if $r_{i+1}\mathcal{D} \subseteq r_i\mathcal{D}$, that is, $r_i \mid r_{i+1}$.
3. If $\mathcal{D}r_{i+1}\mathcal{D} \subseteq r_i\mathcal{D}$, then $r_i \mid\mid r_{i+1}$.

Proof: The first and second item are easy to see. To prove the third item, we need to show that $\mathcal{D}r_{i+1}\mathcal{D} \subseteq \mathcal{D}r_i$. Without loss of generality, let $\mathcal{D}r_{i+1}\mathcal{D} \neq \{0\}$. We show first that there exists an element $\ell \in \mathcal{D}$ such that

$$\mathcal{D}r_{i+1}\mathcal{D} = \ell\mathcal{D} = \mathcal{D}\ell. \tag{4.4}$$

Since $\mathcal{D}$ is a left and right Euclidean domain, there exist $\ell, r \in \mathcal{D}$ satisfying

$$\mathcal{D}r_{i+1}\mathcal{D} = \mathcal{D}r = \mathcal{D}\ell.$$

We claim that ℓ is associated to r: There exist $a, b \in \mathcal{D}$ such that $\ell = ar$ and $r = \ell b$. Thus $\ell = a\ell b$, where $a\ell \in \mathcal{D}r_{i+1}\mathcal{D}$. This implies that $a\ell = \ell b'$ for $b' \in \mathcal{D}$, and thus $\ell(1 - b'b) = 0$. Since $\mathcal{D}r_{i+1}\mathcal{D} \neq \{0\}$, we have that $\ell \neq 0$. This implies that b is a unit and the claim follows.
Finally, we claim that $\ell\mathcal{D} \subseteq r_i\mathcal{D}$ implies $\mathcal{D}\ell \subseteq \mathcal{D}r_i$, which implies that

$$\mathcal{D}r_{i+1}\mathcal{D} \subseteq \mathcal{D}r_i,$$

using (4.4). Due to the structure of $\mathcal{D}$, there exists an element $f \in \mathcal{D}$ such that $\mathcal{D}r_i + \mathcal{D}\ell = \mathcal{D}f$. Thus there exist suitable $a, b \in \mathcal{D}$ such that $f = ar_i + b\ell$. Since $\ell = r_i h$ for a non-zero element $h \in \mathcal{D}$, the equality

$$fh = ar_ih + b\ell h =_{(4.4)} ar_ih + bh'\ell = ar_ih + bh'r_ih = (a + bh')r_ih$$

holds. This yields $\mathcal{D}f \subseteq \mathcal{D}r_i$ and hence $\mathcal{D}\ell \subseteq \mathcal{D}r_i$. □

Note that (4.3) is hard to tackle constructively in general, since there is no grip on the intersection of a left and a right ideal. This difficulty is overcome if $\mathcal{D}$ is simple. Theorem 4.2.1 and Remark 4.2.4 then yields the existence of unimodular matrices $U \in \mathcal{D}^{g\times g}$ and $V \in \mathcal{D}^{q\times q}$ such that

$$URV = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} =: J,$$

where $D = \operatorname{diag}(1, \ldots, 1, m_R) \in \mathcal{D}^{r\times r}$ such that $r = \operatorname{rank}(R)$. Then J is called the **Jacobson form** of R, see Theorem 1.4.1.

In order derive a normal form, the uniqueness of the element m_R needs to be discussed. For this purpose, suppose $R, S \in \mathcal{D}^{g\times q}$ to be equivalent, that is,

$$URV = S$$

for unimodular matrices U, V. Then $R_{\mathcal{D}} \cong S_{\mathcal{D}}$ and furthermore, we have the following isomorphism between the corresponding modules

$$\mathcal{D}/m_R\mathcal{D} \oplus \mathcal{D}^{q-\operatorname{rank}(R)} \cong \mathcal{D}^g/R\mathcal{D}^q \cong \mathcal{M}_R \cong \mathcal{M}_S \cong \mathcal{D}^g/S\mathcal{D}^q \cong \mathcal{D}/m_S\mathcal{D} \oplus \mathcal{D}^{q-\operatorname{rank}(S)}.$$

Since S and R are of equal rank and $\mathcal{D}$ is a Euclidean domain, the isomorphism yields

$$\mathcal{D}/m_R\mathcal{D} \cong \mathcal{D}/m_S\mathcal{D}.$$

Thus there exists a right module isomorphism $\phi : \mathcal{D}/m_R\mathcal{D} \to \mathcal{D}/m_S\mathcal{D}$. Since ϕ is linear, we have

$$[x] \mapsto a[x] = [ax]$$

for a suitable $a \in \mathcal{D}$. The map ϕ is well-defined, which implies that $am_R\mathcal{D} \subseteq m_S\mathcal{D}$, that is,

$$m_R\mathcal{D} \subseteq \{x \in \mathcal{D} \mid ax \in m_S\mathcal{D}\}. \tag{4.5}$$

The injectivity of ϕ implies conversely that

$$m_R\mathcal{D} \supseteq \{x \in \mathcal{D} \mid ax \in m_S\mathcal{D}\}. \tag{4.6}$$

Furthermore since ϕ is surjective, we obtain that

$$\forall\, c \in \mathcal{D}\ \exists r \in \mathcal{D} : c - ar \in m_S\mathcal{D},$$

which yields

$$\mathcal{D} = a\mathcal{D} + m_S\mathcal{D}. \tag{4.7}$$

Consideration of (4.5), (4.6) and (4.7) motivates the following claim:

Theorem 4.2.5 *Let $m_R, m_S \in \mathcal{D}$. Then*

$$\mathcal{D}/m_R\mathcal{D} \cong \mathcal{D}/m_S\mathcal{D}$$

if and only if there exists an element $a \in \mathcal{D}$ such that

(i) $\mathcal{D} = a\mathcal{D} + m_S\mathcal{D}$,

(ii) $m_R\mathcal{D} = \{x \in \mathcal{D} \mid ax \in m_S\mathcal{D}\}$.

Proof: If $\mathcal{D}/m_R\mathcal{D}$ and $\mathcal{D}/m_S\mathcal{D}$ are isomorphic, then (4.5) and (4.6) imply (ii) and (4.7) yields (i). Conversely, (i) leads to

$$\mathcal{D}/m_S\mathcal{D} \cong (a\mathcal{D} + m_S\mathcal{D}) \,/\, m_S\mathcal{D} \cong a\mathcal{D} \,/\, (a\mathcal{D} \cap m_S\mathcal{D}).$$

Using the epimorphism

$$\mathcal{D} \to a\mathcal{D} \,/\, (a\mathcal{D} \cap m_S\mathcal{D}), \quad x \mapsto [ax]$$

and (ii), we obtain

$$\mathcal{D}/m_R\mathcal{D} \cong a\mathcal{D} \,/\, (a\mathcal{D} \cap m_S\mathcal{D}),$$

which completes the proof. □

The characterization introduced in the previous theorem can be specified and leads to the next corollary.

Corollary 4.2.6 *Let $a, m_S, m_R \in \mathcal{D}$ and $\mathcal{D} = a\mathcal{D} + m_S\mathcal{D}$. Then*

$$m_R\mathcal{D} = \{x \in \mathcal{D} \mid ax \in m_S\mathcal{D}\} \tag{4.8}$$

holds if and only if there exists an element $b \in \mathcal{D}$ such that

$$am_R = m_S b \quad \text{and} \quad \mathcal{D} = \mathcal{D}b + \mathcal{D}m_R \tag{4.9}$$

hold.

Proof: First suppose m_S to be zero. Then the assumption $\mathcal{D} = a\mathcal{D} + m_S\mathcal{D}$ implies that a is a unit. Since $\mathcal{D}$ is a domain, (4.8) coincides with $m_R = 0$ and (4.9) coincides with $m_R = 0$ too. Now suppose $m_S \neq 0$.
It is easy to see that $m_R\mathcal{D} \subseteq \{x \in \mathcal{D} \mid ax \in m_S\mathcal{D}\}$ holds if and only if there exists an element $b \in \mathcal{D}$ such that am_R and $m_S b$ coincide.
Now suppose (4.8) holds, that is, $am_R = m_S b$ and for all $x \in \mathcal{D}$, the relation $ax \in m_S\mathcal{D}$ implies that $x \in m_R\mathcal{D}$. Further, we have the ideal identity

$$\mathcal{D}b + \mathcal{D}m_R = \mathcal{D}y, \tag{4.10}$$

where $0 \neq y \in \mathcal{D}$. Else $m_R = 0$, which would contradict Theorem 4.2.5. Let $x, c, z \in \mathcal{D}$ be suitable such that $ax = m_S c$ and $x = m_R z$. Then $m_S c = ax = am_R z = m_S bz$ and therefore $bz = c$, since $m_S \neq 0$. Thus

$$\ker([a, -m_S]\cdot) \subseteq \operatorname{im}\left(\begin{bmatrix} m_R \\ b \end{bmatrix} \cdot \right). \tag{4.11}$$

Due to (4.10), there exist $m', b' \in \mathcal{D}$ such that $m_R = m'y$ and $b = b'y$. Since $0 = am_R - m_S b = am'y - m_S b'y$, we obtain that

$$\mathrm{im}(\begin{bmatrix} m' \\ b' \end{bmatrix} \cdot) \subseteq \ker([a, -m_S]\cdot). \tag{4.12}$$

Combining the two inclusions (4.11) and (4.12) leads to

$$\ker([a, -m_S]\cdot) \subseteq \mathrm{im}(\begin{bmatrix} m_R \\ b \end{bmatrix} \cdot) = \mathrm{im}(\begin{bmatrix} m' \\ b' \end{bmatrix} y \cdot) \subseteq \mathrm{im}(\begin{bmatrix} m' \\ b' \end{bmatrix} \cdot) \subseteq \ker([a, -m_S]\cdot).$$

Thus we obtain the equality

$$\mathrm{im}(\begin{bmatrix} m_R \\ b \end{bmatrix} \cdot) = \mathrm{im}(\begin{bmatrix} m' \\ b' \end{bmatrix} \cdot).$$

Therefore y is a unit and thus (4.9) follows.
Conversely suppose that (4.9) holds. Then there exist $m'_R, m'_S, a', b' \in \mathcal{D}$ such that

$$b'b + m'_R m_R = 1, \quad aa' + m_S m'_S = 1 \quad \text{and} \quad am_R = m_S b.$$

Using this, one can easily verify that the matrix $\begin{bmatrix} m'_R & b' \\ a & -m_S \end{bmatrix}$ is invertible. To prove the theorem we still need to show that $ax = m_S c$ implies $x \in m_R\mathcal{D}$. The equivalence

$$\begin{aligned}
& \begin{bmatrix} x \\ c \end{bmatrix} \in \ker([a, -m_S]\cdot) \\
\Leftrightarrow \quad & \begin{bmatrix} z \\ 0 \end{bmatrix} = \begin{bmatrix} m'_R & b' \\ a & -m_S \end{bmatrix} \begin{bmatrix} x \\ c \end{bmatrix} \\
\Leftrightarrow \quad & \begin{bmatrix} m'_R & b' \\ a & -m_S \end{bmatrix} \begin{bmatrix} m_R \\ b \end{bmatrix} z = \begin{bmatrix} m'_R & b' \\ a & -m_S \end{bmatrix} \begin{bmatrix} x \\ c \end{bmatrix} \\
\Leftrightarrow \quad & \begin{bmatrix} m_R \\ b \end{bmatrix} z = \begin{bmatrix} x \\ c \end{bmatrix} \\
\Leftrightarrow \quad & \begin{bmatrix} x \\ c \end{bmatrix} \in \mathrm{im}(\begin{bmatrix} m_R \\ b \end{bmatrix} \cdot)
\end{aligned}$$

completes the proof. □

Definition 4.2.7 In the situation of Corollary 4.2.6, we call m_R and m_S **similar**.

Remark 4.2.8 Up to similarity, the Jacobson form is a normal form.

Now let $\mathcal{D} = A[\partial; \sigma, \delta]$ for the fraction algebra A of a G-algebra. Recall that the degree of $s \in \mathcal{D}$ is defined as the highest exponent of ∂ appearing in s.

Remark 4.2.9 According to [Ore33], the degrees of two similar polynomials are equal, that is, the degree of m_R is an invariant.

However, it is not necessary to compute the Jacobson form to determine the degree of m_R. As the next remark will show, it is sufficient to have a diagonal form.

Remark 4.2.10 Let U, V be unimodular and $a, b, c \in \mathcal{D} \setminus \{0\}$ such that

$$U \operatorname{diag}(a, b) V = \operatorname{diag}(1, c). \tag{4.13}$$

Then $\deg(a) + \deg(b) = \deg(c)$.

Proof: Due to (4.13) there exists a $\mathcal{D}$-module isomorphism

$$\phi : \mathcal{D}/a\mathcal{D} \oplus \mathcal{D}/b\mathcal{D} \to \mathcal{D}/c\mathcal{D}.$$

Recall that A denotes the fraction algebra $\operatorname{Quot}(A_*)$ of a G-algebra A_*. Then A is a skew field and ϕ is a A-vector space isomorphism. Thus the A-dimensions of $\mathcal{D}/a\mathcal{D} \oplus \mathcal{D}/b\mathcal{D}$ and $\mathcal{D}/c\mathcal{D}$ coincide. But since the A-dimension of $\mathcal{D}/a\mathcal{D} \oplus \mathcal{D}/b\mathcal{D}$ is $\deg(a) + \deg(b)$ and the A-dimension of $\mathcal{D}/c\mathcal{D}$ is $\deg(c)$, the claim follows. □

In what follows, we focus on the rational Weyl algebra over a field of characteristic zero, see Remark 2.2.7. Thus let $\mathcal{D}$ denote the rational Weyl algebra $K(x)[\partial; \mathrm{id}, \frac{\partial}{\partial x}]$ for a field of characteristic zero for the remaining section. Note that the restriction of the field to be of characteristic zero is indeed essential as already pointed out in Remark 2.2.7.

Lemma 4.2.11 *Consider $a, b \in \mathcal{D}$ with $\deg(a) > 0$, $b \neq 0$ and $\deg(b) \geq \deg(a)$. Then there exists $i \in \{0, \ldots, \deg(b) - \deg(a) + 1\}$ such that a is not a right divisor of bx^i.*

Proof: Suppose that for every $i \in \{0, \ldots, \deg(b) - \deg(a) + 1\}$ there exists a $q_i \in \mathcal{D}$ such that $bx^i = q_i a$. Let $b = b_n(x)\partial^n + \cdots + b_1(x)\partial + b_0(x)$, where $b_n \neq 0$. Note that for any $k \in \mathbb{N}$ the equality $\partial^k x = x\partial^k + k\partial^{k-1}$ holds. Thus

$$\begin{aligned} bx &= b_n(x)\partial^n x + \cdots + b_1(x)\partial x + b_0(x)x \\ &= b_n(x)(x\partial^n + n\partial^{n-1}) + \cdots + b_1(x)(x\partial + 1) + xb_0(x) \\ &= xb + \underbrace{\sum_{i=1}^{n} b_i(x) i \partial^{i-1}}_{=:r_1} \end{aligned}$$

with $\deg(r_1) = n - 1 < \deg(b)$ and $r_1 \neq 0$ since $\deg(b) \geq 1$. Since $b = q_0 a$ and $bx = q_1 a$, it follows that $r_1 = bx - xb = (q_1 - xq_0)a$, that is, a is a right divisor of r_1. By proceeding with bx^2 and so on, we obtain a sequence of non-zero polynomials r_i, such that $\deg(b) > \deg(r_1) > \ldots$ and $a \mid r_i$. Since the degree of r_i decreases exactly by 1 at each step, after at most $\deg(b) - \deg(a) + 1$ iterations, we obtain a polynomial of degree $\deg(a) - 1$, which is non-zero. Such a polynomial must contain a right factor of degree $\deg(a)$, which is a contradiction. □

Due to Lemma 4.2.11, we can describe constructively how to compute the Jacobson form from a diagonal matrix. Suppose $R \in R^{g\times q}$, where $g = q = 2$. The extension to $g, q \in \mathbb{N}$ is evident. Algorithm 3 returns unimodular matrices U, V such that $URV = \mathrm{diag}(r_1, r_2)$. Without loss of generality, let us assume that $\deg(r_2) \leq \deg(r_1)$.

1) If r_2 is a unit, we get the Jacobson form just by replacing U by $\begin{bmatrix} 0 & r_2^{-1} \\ 1 & 0 \end{bmatrix} U$ and V by $V \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

   Otherwise, choose according to Lemma 4.2.11 an exponent $i \in \mathbb{N}$ such that $r_1 x^i = a r_2 + b$ with $\deg(b) < \deg(r_2)$ and $b \neq 0$. Then

$$\begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & x^i \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_1 & b \\ 0 & r_2 \end{bmatrix}.$$

   Replace U by $\begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix} U$ and V by $V \begin{bmatrix} 1 & x^i \\ 0 & 1 \end{bmatrix}$.

2) Now we apply Algorithm 3 to the matrix $\begin{bmatrix} r_1 & b \\ 0 & r_2 \end{bmatrix}$. The result is then the diagonal matrix $\mathrm{diag}(r_1', r_2')$, where $\deg(r_2') < \deg(r_2)$.

Thus, by iterating 1) and 2), we compute U and V such that $URV = \mathrm{diag}(m_R, 1)$.

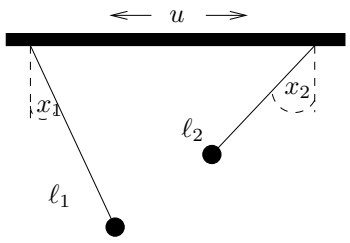
Remark 4.2.12 We claim (without proof) that the Jacobson form can be calculated once having a diagonal form in the following way:
Due to Remark 4.2.10, the sum $\Sigma_{\deg}$ of the degrees of all diagonal entries is an invariant of the module. Following the ideas of [Mid08], we can use this information to compute the cyclic generator of $\mathcal{D}^{1\times q}/_{\mathcal{D}}R$. The algorithmic idea is to compute random annihilators a_i of each diagonal entry r_i. Let $\langle c\rangle = \bigcap \langle a_i\rangle$. Then $[c, 1, \ldots, 1]$ is a cyclic generator of $\mathcal{D}^{1\times q}/_{\mathcal{D}}R$ if c is of degree $\Sigma_{\deg}$. In that case $\mathrm{diag}(1, \ldots, 1, c, 0, \ldots, 0)$ is a Jacobson form of R.

4.2.1 Examples, Applications and Comparison

To the best of our knowledge, the Jacobson normal form algorithm has been implemented in Maple by Culianez and Quadrat [CQ05], by Robertz et.al. [BCG+03, CQR04] and by Middeke [Mid08]. However, we could not locate the download version of the implementation of [CQ05]. The implementation of Middeke [Mid08] was, according to its author, merely a check of ideas and was not supposed to become a freely distributed package for Maple. Robertz informed us that his implementation [BCG+03] directly follows the classical algorithm and it has not been specially optimized. Nevertheless, in what follows, we compare our implementation with the one in the Maple package Janet [BCG+03] on some nontrivial examples.

As we already pointed out in the introduction, behind normal forms there are various application-driven motivations. See e. g. [CQ05] for several interesting examples.

Example 4.2.13 Consider a bar on which two pendula of length ℓ_1 and ℓ_2 are fixed, a so-called bipendulum. The bar can be moved horizontally and u denotes its position. Furthermore let x_1 denote the angle between the left pendulum and the vertical position and x_2 analogously the angle between the rigth pendulum and the vertical position.



The linearization of this problem leads to the system of linear ordinary differential equations

$$\ell_1 \frac{d^2}{dt^2} x_1(t) + g x_1(t) - g u(t) = 0$$
$$\ell_2 \frac{d^2}{dt^2} x_2(t) + g x_2(t) - g u(t) = 0,$$

where g is the gravitational constant. Using our notation this coincides with the linear system $R \bullet [x_1, x_2, u]^T = 0$, where

$$R = \begin{bmatrix} \ell_1\partial^2 + g & 0 & -g \\ 0 & \ell_2\partial^2 + g & -g \end{bmatrix}.$$

Since the variable t does not appear in R, the ground ring can be thought of as $\mathbb{R}(\ell_1, \ell_2, g)[\partial]$. Thus, indeed one can compute the Smith normal form.
We run our implementation of the Jacobson form of R on this example and obtain

$$U = \begin{bmatrix} -1/g & 0 \\ -1/g & 1/g \end{bmatrix} \text{ and } V = \begin{bmatrix} 0 & g\ell_2 & -g\ell_2\partial^2 - g^2 \\ 0 & g\ell_1 & -g\ell_1\partial^2 - g^2 \\ 1 & \ell_1\ell_2\partial^2 + g\ell_2 & -\ell_1\ell_2\partial^4 - g\ell_1 - g\ell_2\partial^2 - g^2 \end{bmatrix}$$

such that

$$U\,R\,V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & g\ell_1 - g\ell_2 & 0 \end{bmatrix}.$$

This result agrees with results obtained in [CQ05]. Note that a purely fractional method would return 1 instead of $g(\ell_1 - \ell_2)$. With our polynomial approach, we obtain a polynomial matrix, which is useful for further investigations. In particular, in the current example we see, that setting $\ell_1 = \ell_2$ implies a drop of the rank of the Smith form from 2 to one, thus the properties of the corresponding system will change. More precisely, in case $\ell_1 \neq \ell_2$ the underlying system module is isomorphic to the free module $\mathbb{R}(\ell_1, \ell_2, g)[\partial]$, that is, the linear abstract system given by R is controllable due to Theorem 1.2.9. In case $\ell_1 = \ell_2$ we obtain

$$\underbrace{\begin{bmatrix} -1/g & 0 \\ 1 & -1 \end{bmatrix}}_{=:U} R \underbrace{\begin{bmatrix} 0 & 0 & -g \\ 0 & 1 & -g \\ 1 & 0 & -\ell_1\partial^2 - g \end{bmatrix}}_{=:V} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\ell_1\partial^2 - g & 0 \end{bmatrix}.$$

Since

$$V^{-1} = \begin{bmatrix} -\frac{\ell_1}{g}\partial^2 - 1 & 0 & 1 \\ -1 & 1 & 0 \\ -\frac{1}{g} & 0 & 0 \end{bmatrix},$$

the decomposition introduced in Section 1.4 implies that

$$\mathcal{B} = \{ \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \in \mathcal{A}^3 \mid R \bullet \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} = 0 \},$$

where $\mathcal{A} = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$, can be written as $\mathcal{B}_a \oplus \mathcal{B}_c$ with the autonomous subsystem

$$\begin{aligned} \mathcal{B}_a &= \{\omega \in \mathcal{A}^3 \mid \begin{bmatrix} -\frac{\ell_1}{g}\partial^2 - 1 & 0 & 1 \\ \ell_1\partial^2 + g & -\ell_1\partial^2 - g & 0 \end{bmatrix} \bullet \omega = 0 \text{ and } \begin{bmatrix} -\frac{1}{g} & 0 & 0 \end{bmatrix} \bullet \omega = 0\} \\ &= \{ \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \in \mathcal{A}^3 \mid x_1 = u = 0 \text{ and } (-\ell_1\partial^2 - g) \bullet x_2 = 0 \} \end{aligned}$$

and the controllable subsystem

$$\begin{aligned} \mathcal{B}_c &= \{\omega \in \mathcal{A}^3 \mid \exists v \in \mathcal{A} \colon \omega = \begin{bmatrix} -g \\ -g \\ -\ell_1\partial^2 - g \end{bmatrix} \bullet v \} \\ &= \{ \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \in \mathcal{A}^3 \mid x_1 = x_2 \text{ and } u = -\frac{\ell_1}{g}\partial^2 \bullet x_1 - x_1 \}. \end{aligned}$$

Remark 4.2.14 In [LZ07], an algorithm for finding so-called "obstructions to genericity" was derived and discussed. A lesson learned from that paper can be applied for an implementation of the Jacobson form as follows. It is recommended to split the algorithm (resp. the implementation) into two parts. In the first part, one computes a diagonal matrix, where cancellation of invertible elements of the ground field is artificially avoided. The second part applies the normalization on the invertible elements; this part is trivial to achieve. Note that our polynomial algorithm allows one to keep track of these invertible elements due to this scheme.

Let us start with a non-commutative example of a small matrix with entries of low degree. The results of both implementations (namely, the JANET one and the SINGULAR one) are quite similar.

Example 4.2.15 Consider the Example 3.1.2. of [CQ05], where there is a module presented by the matrix

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

over the first Weyl algebra in x and $\partial = \frac{\partial}{\partial x}$ over $\mathbb{Q}$. Our implementation computes

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

while JANET returns

$$\begin{bmatrix} 1 & -x \\ x^2\partial & -x^3\partial \end{bmatrix} R \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence, the difference is only in the degree of the x terms in the left transformation matrix.

Example 4.2.16 Over the first rational Weyl algebra in x, ∂, let us consider the 3×3 matrix

$$R = \begin{bmatrix} \partial^2 & \partial + 1 & 0 \\ \partial + 1 & 0 & \partial^3 - x^2\partial \\ 2\partial + 1 & \partial^3 + \partial^2 & \partial^2 \end{bmatrix}.$$

The implementation of Algorithm 3 in SINGULAR returns the matrix $\mathrm{diag}(g, 1, 1)$ together with transformation matrices $U, V \in \mathbb{Q}[x][\partial; \mathrm{id}, \frac{d}{dx}]^{3\times 3}$, that is,

$$URV = \mathrm{diag} =: J.$$

Below, we write down just the leading term of each matrix entry and moreover, we write "l.o.t." for "lower order terms" corresponding the degree lexicographical ordering on $\mathbb{Q}[x][\partial; \mathrm{id}, \frac{d}{dx}]$, that is: Recall that the total degree is given by $\deg(x^\alpha \partial^\beta) = \alpha + \beta$, then for unequal monomials we have

$$x^{\alpha_1}\partial^{\beta_1} < x^{\alpha_2}\partial^{\beta_2} \quad \Leftrightarrow \quad \deg(x^{\alpha_1}\partial^{\beta_1}) < \deg(x^{\alpha_2}\partial^{\beta_2})$$

or if

$$\deg(x^{\alpha_1}\partial^{\beta_1}) = \deg(x^{\alpha_2}\partial^{\beta_2}) \quad \text{then} \quad \beta_1 < \beta_2.$$

Then the diagonal matrix can be written as

$$J = \begin{bmatrix} 2x^2\partial^8 + 33 \text{ l.o.t.} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the transformation matrices are

$$U = \begin{bmatrix} \frac{1}{2}x\partial^{13} + 24 \text{ l.o.t.} & \frac{1}{2}x\partial^{10} + 19 \text{ l.o.t.} & \frac{1}{2}x\partial^{11} + 44 \text{ l.o.t.} \\ \frac{1}{2} & 0 & 0 \\ -\frac{1}{4}\partial^5 + 2 \text{ l.o.t.} & -\frac{1}{4}\partial^2 & \frac{1}{4} + 2 \text{ l.o.t.} \end{bmatrix}$$

and

$$V = \begin{bmatrix} 2x\partial^2 + 3 \text{ l.o.t.} & 2\partial^2 & 2\partial^2 + 1 \text{ l.o.t.} \\ -2x\partial^3 + 2 \text{ l.o.t.} & -2\partial^3 + 3 \text{ l.o.t.} & -2\partial^3 \\ x\partial^8 + 28 \text{ l.o.t.} & \partial^8 + 11 \text{ l.o.t.} & \partial^8 + 16 \text{ l.o.t.} \end{bmatrix}.$$

With the help of JANET, we obtain the diagonal matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (279936x^{14} + 14 \text{ l.o.t. })^{-1}(279936x^{14}\partial^8 + 145 \text{ l.o.t. }) \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 & 0 \\ (6x^2 + 2 \text{ l.o.t.})^{-1}(\partial^2 + 1 \text{ l.o.t. }) & (6x^2 + 2 \text{ l.o.t.})^{-1}(\partial^3 + 3 \text{ l.o.t. }) & (6x^2 + 2 \text{ l.o.t.})^{-1} \\ u_{31} & u_{32} & u_{33} \end{bmatrix},$$

where

$$\begin{aligned} u_{31} &= (559872x^{14} + 14 \text{ l.o.t.})^{-1}(-279936x^{14}\partial^9 + 158 \text{ l.o.t.}) \\ u_{32} &= (559872x^{14} + 14 \text{ l.o.t.})^{-1}(279936x^{14}\partial^{10} + 182 \text{ l.o.t.}) \\ u_{33} &= (559872x^{14} + 14 \text{ l.o.t.})^{-1}(279936x^{14}\partial^7 + 127 \text{ l.o.t.}) \end{aligned}$$

The right transformation matrix is

$$V = \begin{bmatrix} 1 & \frac{1}{2}\partial^6 + 15 \text{ l.o.t.} & (279936x^{14} + 14 \text{ l.o.t. })^{-1}(46656x^{12}\partial^7 + 110 \text{ l.o.t. }) \\ \partial + 1\text{l.o.t.} & -\frac{1}{2}\partial^7 + 15 \text{ l.o.t.} & (-1679614x^{16} + 16 \text{ l.o.t. })^{-1}(279936x^{14}\partial^8 + 138 \text{ l.o.t. }) \\ 0 & 1 & (6x^2 + 2 \text{ l.o.t. })^{-1}(2\partial^2 + 1 \text{ l.o.t} \end{bmatrix},$$

What we observe in this example is quite a typical behavior of any implementation which directly uses arithmetics over a skew field. Namely, one gets big polynomials with long coefficients, which, as we can see, are of approximately the same size. This stays in distinct contrast with the output of our algorithm, where polynomials are of moderate size and the coefficients are rather small.

Example 4.2.17 Suppose $\mathcal{D} = \mathbb{Q}(y, x)[\partial; \mathrm{id}, \frac{d}{dx}]$, $\mathcal{D}_* = \mathbb{Q}[y, x][\partial; \mathrm{id}, \frac{d}{dx}]$ and consider the following matrix coming from the system of differential equations

$$R = \begin{bmatrix} y^2 & 0 \\ 0 & x^2 \end{bmatrix} \partial^2 + \begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix} \partial + \begin{bmatrix} 1 & 1 \\ 0 & y \end{bmatrix}.$$

Our implementation returns the matrix $\mathrm{diag}(g, 1)$, where

$$g = -y^2x^2\partial^4 - x^2\partial^3 - x^2\partial^2 - y^2\partial^3 + x\partial + (-y^3 - 1)\partial^2 + (-y - 1)\partial - y$$

and the corresponding transformation matrices, such that

$$\begin{bmatrix} -x^2\partial^2 - \partial - y & 1 \\ 1 & 0 \end{bmatrix} M \begin{bmatrix} 1 & 0 \\ -y^2\partial^2 - \partial - 1 & 1 \end{bmatrix} = \begin{bmatrix} g & 0 \\ 0 & 1 \end{bmatrix}.$$

The diagonal element g can be easily represented in the rational form as

$$\tilde{g} = \partial^4 + \frac{x^2 + y^2}{x^2y^2}\partial^3 + \frac{x^2 + y^3 + 1}{x^2y^2}\partial^2 + \frac{-x + y + 1}{x^2y^2}\partial + \frac{1}{x^2y}.$$

If we suppose $R \in \mathbb{Z}_2(y, x)[\partial; \mathrm{id}, \frac{d}{dx}]^{2\times 2}$, we obtain the single example from [Mid08]. Then the rational form of our result is exactly the result obtained in [Mid08], namely $\mathrm{diag}(1, g')$, where

$$g' = \partial^4 + \frac{x^2 + y^2}{x^2y^2}\partial^3 + \frac{x^2 + y^3 + 1}{x^2y^2}\partial^2 + \frac{x + y + 1}{x^2y^2}\partial + \frac{1}{x^2y}.$$

Note that in our method, no computations with 4×4 matrices as in [Mid08] are needed. As demonstrated, selecting the characteristic 2 is not crucial for this example.

Chapter 5

Multi-dimensional time-varying systems

Section 1.6 motivates the impact of exact linear modeling for time-invariant systems. The most powerful unfalsified model was introduced and its properties were discussed for the model classes

- $\mathcal{D} = \mathbb{C}[\partial_1, \ldots, \partial_n]$ and polynomial-exponential signals in $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$, that is, trajectories of the form $p(x)\exp(\lambda_1 x_1 + \cdots + \lambda_n x_n)$, where $p \in \mathbb{C}\left[x_1, \cdots, x_n\right]^q$ and $\lambda \in \mathbb{C}^n$
- $\mathcal{D} = \mathbb{C}[\mathbf{s}_1, \ldots, \mathbf{s}_n]$ and polynomial-exponential signals in $\mathbb{C}^{\mathbb{N}^n}$, that is, trajectories of the form $p(x)\lambda_1^{x_1} \cdots \lambda_n^{x_n}$, where $p \in \mathbb{C}\left[x_1, \cdots, x_n\right]^q$ and $\lambda \in \mathbb{C}^n$.

In this chapter, we extend the discussed model classes. The results of this chapter are accepted for publication in the Journal of Symbolic Computation [SLZ].
In the sequel let K be a field and $\mathcal{D}$ be an operator algebra over K. Further let $\mathcal{A}_\mathcal{D}$ be a function space over K possessing a $\mathcal{D}$-module structure. In most cases of interest, we have $K \subseteq \mathcal{D}$ and $ok = ko$ for all $k \in K$, $o \in \mathcal{D}$. Then the corresponding linear abstract system is a K-vector space, and thus the introduced model class is linear. Within such a model class, we want to perform modeling now.
Let us recall the idea of linear exact modeling. Suppose to observe a set of signals $\Omega \subseteq \mathcal{A}_\mathcal{D}^m$. The aim is to find a model $\mathcal{B}_\Omega$ in the model class such that

1. $\mathcal{B}_\Omega$ is unfalsified by Ω, i.e. $\Omega \subseteq \mathcal{B}_\Omega$.
2. $\mathcal{B}_\Omega$ is most powerful, i.e. for every model $\mathcal{B}$ with $\Omega \subseteq \mathcal{B}$, it follows that $\mathcal{B}_\Omega \subseteq \mathcal{B}$.

If $\mathcal{B}_\Omega$ is invariant under the action of $\mathcal{D}$, that is, if we have for all $o \in \mathcal{D}$:

$$\omega \in \mathcal{B}_\Omega \quad \Rightarrow \quad o \bullet \omega \in \mathcal{B}_\Omega,$$

it is called **most powerful unfalsified model**, short MPUM of Ω. Else, if $\mathcal{B}_\Omega$ varies under $\mathcal{D}$, it is called **variant most powerful unfalsified model**, short VMPUM of Ω. We denote the VMPUM of Ω by $\mathcal{B}_\Omega^V$.
The following example shows how the choice of the model class affects the model.

Example 5.0.18 Consider the signal set consisting of a single polynomial signal

$$\Omega = \{\omega\}, \quad \text{where } \omega(x) = x \text{ for all } x \in \mathbb{R}.$$

1. Let $\mathcal{D} = \mathbb{C}[\partial]$ and $\mathcal{A}_\mathcal{D} = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, where $\partial \bullet f := \frac{df}{dx}$. Using the commutative structure of the operator ring, the underlying system is invariant under differentiation:

$$R \bullet w = 0 \quad \Rightarrow \quad R(\partial \bullet w) = (R\partial) \bullet w = (\partial R) \bullet w = \partial(R \bullet w) = 0.$$

Since we are searching for a differentiation-invariant model, we obtain that besides ω, also its derivative, the constant function 1, belongs to $\mathcal{B}_\Omega$. Using that the model is $\mathbb{C}$-linear, we get that

$$\mathcal{B}_\Omega = \{w \mid \exists a, b \in \mathbb{C} : \forall x \in \mathbb{R} : w(x) = ax + b\}.$$

An element $w \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ is contained in $\mathcal{B}_\Omega$ if and only if

$$\partial^2 \bullet w = 0,$$

i.e. the MPUM is specified by a single ordinary differential equation with constant coefficients.

2. Now let $\mathcal{D} = \mathbb{C}[x]\langle\partial\rangle$, where $\partial \bullet f := \frac{df}{dx}$ and $\mathcal{A}_\mathcal{D}$ is defined as above. We want to describe ω as a solution of homogeneous ordinary differential equations with polynomial coefficients. The equations

$$\partial^2 \bullet w = 0 \quad \text{and} \quad x\partial \bullet w - w = 0$$

are satisfied by ω. We will see later that these two generate a kernel representation of the VMPUM of Ω. The corresponding solution space equals

$$\mathcal{B}_\Omega^V = \{w \mid \exists a \in \mathbb{C} : \forall x \in \mathbb{R} : w(x) = ax\}.$$

Notice that this example demonstrates the variance under ∂, since we have $\partial \bullet \omega \notin \mathcal{B}_\Omega^V$. Another property that should be pointed out is that the VMPUM yields a more precise description of Ω than the MPUM.

5.1 Preliminaries

Example 5.0.18 deals with continuous signals. But in applications, there are also discrete phenomena or combinations of discrete and continuous signals that are of great interest too. As already motivated in Example 2.2.5, many of the relevant operator algebras have the structure of an Ore algebra, see Section 2.2.
Let A be a ring and $\mathcal{D} := A[\partial_1; \sigma_1, \delta_1] \cdots [\partial_m; \sigma_m, \delta_m]$ be an Ore algebra. With the action

$$\partial_i \bullet p := \delta_i(p) \quad \text{and} \quad a \bullet p := a \cdot p \quad \text{for all } p \in A \text{ and } a \in A, \tag{5.1}$$

the ring A becomes a $\mathcal{D}$-module. For this, we have to show that

1. $(o_1 \cdot o_2) \bullet p = o_1 \bullet (o_2 \bullet p)$ for all $o_1, o_2 \in \mathcal{D}$ and $p \in A$
2. $(o_1 + o_2) \bullet p = o_1 \bullet p + o_2 \bullet p$ for all $o_1, o_2 \in \mathcal{D}$ and $p \in A$
3. $o \bullet (p + q) = o \bullet p + o \bullet q$ for all $o \in \mathcal{D}$ and $p, q \in A$.

To show 1. it suffices to consider $o_1 = a\partial_i$ and $o_2 = b\partial_j$ with $a, b \in A$. Then

$$\begin{aligned}(o_1 \cdot o_2) \bullet p &= (\, a(\, \sigma_i(b)\partial_i + \delta_i(b)\,)\partial_j\,) \bullet p \\ &= (\, a\sigma_i(b)\partial_i\partial_j + a\delta_i(b)\partial_j\,) \bullet p \\ &= a\sigma_i(b)\delta_i(\delta_j(p)) + a\delta_i(b)\delta_j(p) \\ &= a\,\delta_i(b\delta_j(p)) \\ &= a\partial_i \bullet (\, b\partial_j \bullet p\,) \\ &= o_1 \bullet (\, o_2 \bullet p\,).\end{aligned}$$

The equality in 2. and 3. holds by similar arguments.

Using this action, we can define the kernel of a linear operator f from the Ore algebra $\mathcal{D}$ over the ring A to be $\ker_A f := \{a \in A \mid f \bullet a = 0\}$. The proposed module structure (5.1) is not interesting in case ∂_i operates trivially on A for all i, because then $\ker_A f$ would equal A for all $\sum_{\alpha \in \mathbb{N}^n} f_\alpha \partial^\alpha = f \in \mathcal{D}$ with $f_0 = 0$. Our approach would then model exclusively inner relations of A, or rather A^q. Therefore we will consider operator algebras which act nontrivially on the function space. But this requirement is not a restriction, as the following results will show.

Lemma 5.1.1 *Let A be a ring, and $A[\partial; \sigma, \delta]$ be an Ore extension of A. For any $\alpha \in A$, there exists an Ore extension $A[\Delta_\alpha; \sigma, \delta']$ with $\delta'(a) = \sigma(a)\alpha - \alpha a + \delta(a)$, such that $A[\partial; \sigma, \delta] \cong A[\Delta_\alpha; \sigma, \delta']$ as rings.*

Proof: For all $a \in A$, the equality $\partial a = \sigma(a)\partial + \delta(a)$ holds. For $\alpha \in A$, define $\Delta_\alpha := \partial - \alpha$. Then it obeys the relation $\Delta_\alpha a = \sigma(a)\Delta_\alpha + \sigma(a)\alpha - \alpha a + \delta(a)$. Let us define $\delta'(a) := \sigma(a)\alpha - \alpha a + \delta(a)$, which is clearly linear. It is easy to see that δ' is a σ-derivation. Namely, by expanding $\sigma(a)\delta'(b) + \delta'(a)b$, we get

$$\sigma(a)\delta'(b) + \delta'(a)b = \sigma(a)\sigma(b)\alpha - \sigma(a)\alpha b + \sigma(a)\delta(b) + \sigma(a)\alpha b - \alpha a b + \delta(a)b =$$

$$\sigma(a)\sigma(b)\alpha - \alpha a b + \sigma(a)\delta(b) + \delta(a)b = \delta'(ab).$$

This shows that $A[\Delta_\alpha; \sigma, \delta']$ is a subring of $A[\partial; \sigma, \delta]$ by construction. Hence the ring homomorphism $\varphi_\alpha : A[\partial; \sigma, \delta] \to A[\Delta_\alpha; \sigma, \delta']$, being the identity on A, which sends ∂ to Δ_α is surjective. Since Δ_α was defined to be $\partial - \alpha$, the injectivity follows. □

Lemma 5.1.2 *Let K be a field, A be a K-algebra, δ be a K-linear operator, acting on A and $B = A[\partial; \sigma, \delta]$ be the corresponding operator algebra (that is, for all $a \in A$ we have $\partial a = \sigma(a)\partial + \delta(a)$). Then the following holds:*

(*i*) $\ker_A \partial = A \Leftrightarrow \delta = 0 \Leftrightarrow B = A[\partial; \sigma, 0]$.

(*ii*) *If* $\ker_A \partial = A$, *then we have for* $\Delta := \partial - 1$*:* $A[\partial; \sigma, 0]$ *is isomorphic as a* K*-algebra to the operator algebra* $A[\Delta; \sigma, \delta']$ *with* $\delta' := \sigma - 1$.

Proof:

(*i*) We have $\ker_A \partial = \{a \in A \mid \partial \bullet a = 0\} = \{a \in A \mid \delta(a) = 0\}$, thus the claim follows by definition.

(*ii*) Referring to Lemma 5.1.1, we define $\alpha := 1$. Then $\Delta_\alpha = \partial - 1 = \Delta$. Since further due to the previous item $\delta = 0$, we obtain the equality

$$\sigma(a)\alpha - \alpha a + \delta(a) = \sigma(a) - a = \delta'(a).$$

Thus the claim follows by Lemma 5.1.1. □

Remark 5.1.3 Using Lemmas 5.1.1 and 5.1.2, we pass to the new setting of operators ∂_i which act nontrivially on A. From now on, we will work only with such operators.

The following example shows how to pass to the feasible setting in case δ equals the zero-map.

Example 5.1.4

1. The first **forward shift algebra** is defined by $K[x]\,[\mathbf{s}; \sigma, 0]$ with $(\sigma f)(x) = f(x+1)$ for all $f \in K[x]$. The commutation rule is $\mathbf{s}x = x\mathbf{s}+\mathbf{s}$. There is a natural operator associated to $\mathbf{s}$, namely the difference operator $\Delta = \mathbf{s} - 1$, already defined in Example 2.2.5, obeying the relation $\Delta x = x\Delta + \Delta + 1$. Applying Lemma 5.1.2, we see by degree arguments that $\ker \Delta = K$, and the two algebras are isomorphic both as Ore extensions and as K-algebras.

2. Let q be transcendental over K. Then the first **q-commutative algebra** (or Manin's quantum plane) is defined as $K_q[x, y] := K(q)[x][\partial; \sigma, 0]$ with $(\sigma f)(x) = f(qx)$ for $f \in K[x]$. Again, there is a natural q-difference operator $\Delta_q := \partial - 1$ and the corresponding operator algebra. It has been already described in Example 2.2.5 as the first continuous q-difference algebra. Its commutation rule reads as $\partial x = qx\partial + (q-1)x$.

For $o_1, \ldots, o_k \in \mathcal{D}^n$, we denote by ${}_{\mathcal{D}}\langle o_1, \ldots, o_k\rangle$ the left submodule of $\mathcal{D}^n$ generated by $o_1, \ldots, o_k$.

Theorem 5.1.5 *Let* $\mathcal{D}$ *be an Ore* A*-algebra built from operators* $\partial_1, \ldots, \partial_s$*. Then there is an isomorphism of left* $\mathcal{D}$*-modules*

$$\mathcal{D}\,/_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s\rangle \cong A.$$

Proof: There is a left $\mathcal{D}$-module homomorphism

$$\varphi : \mathcal{D} \to A, \quad a = \sum_{\alpha \in \mathbb{N}^s} a_\alpha \partial^\alpha \mapsto a \bullet 1,$$

since $\varphi(b \cdot a) = (b \cdot a) \bullet 1 = b \bullet \varphi(a)$. Due to Definition 2.2.4, we have $\delta(1) = 0$ and thus $a \bullet 1 = a_0$. The kernel of φ is given by the left ideal ${}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle$. Further, φ is clearly surjective. So the claim follows from the homomorphism theorem. □

Following Theorem 5.1.5, every $p \in A$ can be viewed as an element of the left $\mathcal{D}$-module $\mathcal{D}/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle$ by identifying p with $p + {}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle =: [p]$. Then the action of ∂_i is exactly the σ_i-derivation δ_i, since

$$\partial_i[p] = [\partial_i p] = [\sigma_i(p)\partial_i + \delta_i(p)] = [\delta_i(p)] = [\partial_i \bullet p].$$

Remark 5.1.6 Let $p \in A$ and $o \in \mathcal{D}$. Then the equivalence

$$o \bullet p = 0 \quad \text{if and only if} \quad o \cdot p \in {}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle$$

holds.

Proof: By Theorem 5.1.5, we have a $\mathcal{D}$-module isomorphism $A \cong \mathcal{D}/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle$ given by

$$A \xrightarrow{\cong} \mathcal{D}/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle, \qquad p \mapsto [p].$$

Since the $\mathcal{D}$-module structure is respected, $o \bullet p$ maps to $[o \cdot p]$ and hence the claim follows. □

Remark 5.1.6 gives the possibility to describe and to compute the annihilator of an element $p \in A$. Consider the map

$$\kappa_p : \mathcal{D} \to \mathcal{D}/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle, \qquad o \mapsto o \cdot [p], \tag{5.2}$$

which is clearly a left $\mathcal{D}$-module homomorphism with the kernel

$$\ker(\kappa_p) = \operatorname{ann}_{\mathcal{D}}(p) := \{o \in \mathcal{D} \mid o \bullet p = 0\},$$

which is a left ideal in $\mathcal{D}$. See Corollary 2.2.15 for its algorithmic computation. This construction lifts to the non-scalar case. Suppose $p = [p_1, \ldots, p_m]^T \in A^m$. An element of $o \in \mathcal{D}^{1 \times m}$ acts on p by

$$o \bullet p := \sum_{i=1}^{m} o_i \bullet p_i.$$

A subset $B \subseteq A^m$ is called **invariant** under $G \subseteq \mathcal{D}^{1 \times m}$ if and only if $o \bullet p = 0$ for all $o \in G$ and $p \in B$. The set of elements under which p is invariant has a $\mathcal{D}$-module structure and equals the kernel of

$$\kappa_p : \mathcal{D}^{1 \times m} \to \mathcal{D}/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s \rangle, \qquad o = [o_1, \ldots, o_m] \mapsto \sum_{i=1}^{m} o_i \cdot [p_i].$$

Moreover the following isomorphism holds

$$\mathcal{D}^{1\times m} / \ker(\kappa_p) \cong {}_{\mathcal{D}}\langle p_1, \ldots, p_m\rangle / {}_{\mathcal{D}}\langle p_1, \ldots, p_m\rangle \cap {}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s\rangle.$$

The image of κ_p equals $({}_{\mathcal{D}}\langle p_1, \ldots, p_m\rangle + {}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s\rangle)/{}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s\rangle$. This is isomorphic to ${}_{\mathcal{D}}\langle p_1, \ldots, p_m\rangle / {}_{\mathcal{D}}\langle p_1, \ldots, p_m\rangle \cap {}_{\mathcal{D}}\langle \partial_1, \ldots, \partial_s\rangle$. So the claim follows, since κ_p is a homomorphism.

Remark 5.1.7 If $\mathcal{D}$ is Noetherian (see [MR01]), then the left submodule $\ker(\kappa_p) \subseteq \mathcal{D}^{1\times m}$ is finitely generated.

For an m-tuple $p \in A^m$, we consider

$$\operatorname{ann}_{\mathcal{D}}(p) = \{o \in \mathcal{D} \mid o \bullet p = 0\} = \{o \in \mathcal{D} \mid o \bullet p_i = 0 \ \forall\ i\} = \bigcap \operatorname{ann}_{\mathcal{D}}(p_i),$$

which is a left ideal in $\mathcal{D}$. As we see immediately, $\operatorname{ann}_{\mathcal{D}}(p)^{1\times m}$ is a (usually strict) submodule of $\ker(\kappa_p)$ and hence, the latter typically has a richer and more interesting structure.

5.2 Application to linear exact modeling

We will now use the results from above to define an unfalsified and most powerful model over an Ore algebra.

Assumptions and notations: Suppose $\mathcal{D} = A[\partial_1; \sigma_1, \delta_1] \ldots [\partial_n; \sigma_n, \delta_n]$ to be a Noetherian Ore algebra (see Remark 2.2.13). Recall that $\mathcal{A}_{\mathcal{D}}$ denotes a function space over K possessing a $\mathcal{D}$-module structure. Suppose further that $A \subseteq \mathcal{A}_{\mathcal{D}}$.

Starting with a single signal $p \in A^m$, we want to find the VMPUM of p, that is a behavior, invariant under some finitely generated submodule of $\mathcal{D}^{1\times m}$.

Theorem 5.2.1 *Let $p \in A^m$ be given. Consider the map κ_p from (5.2). Let* $\ker(\kappa_p) = {}_{\mathcal{D}}\langle k_1, \ldots, k_r\rangle$ *and let $R \in \mathcal{D}^{r\times m}$ be a matrix whose i-th row equals k_i. Then the* VMPUM *of $\{p\}$ is given by*

$$\mathcal{B}^V_{\{p\}} = \{g \in \mathcal{A}^m_{\mathcal{D}} \mid R \bullet g = 0\}.$$

Proof: By the definition of R and Remark 5.1.6, it is clear that $\{p\} \subseteq \mathcal{B}^V_{\{p\}}$.
It remains to show that $\mathcal{B}^V_{\{p\}}$ is most powerful. Suppose there exists another behavior $\mathcal{B}'$ unfalsified by p. The behavior $\mathcal{B}'$ possesses a kernel representation $R' \in \mathcal{D}^{r'\times m}$. By the definition of R, there exists a matrix $X \in \mathcal{D}^{r'\times r}$ such that $R' = XR$. But since $(X \cdot R) \bullet p = X \bullet (R \bullet p)$, it follows that $\mathcal{B}^V_{\{p\}} \subseteq \mathcal{B}'$. □

Example 5.2.2 Let us consider a more interesting example than Example 5.0.18 with respect to the algebras from Example 2.2.5. Let $\Omega = \{\omega\}$ consist of the cusp

$$\omega(x_1, x_2) = x_1^3 - x_2^2.$$

Let us denote by $\mathcal{A}_\mathcal{D} = \mathbb{C}[[x_1, x_2]]$ the ring of formal power series and consider the VMPUM $\mathcal{B}^V_{\{\omega\}} = \{f \in \mathcal{A}_\mathcal{D} \mid R_{\text{VMPUM}} \bullet f = 0\}$ of Ω with respect to the following operator algebras $\mathcal{D}$.

1. Suppose $\mathcal{D}$ to be the second **Weyl algebra**. Then by using SINGULAR we obtain:

$$R_{\text{VMPUM}} = \begin{bmatrix} \partial_2^3 \\ \partial_1\partial_2 \\ \partial_1^3 + 3\partial_2^2 \\ x_2\partial_2^2 - \partial_2 \\ x_2\partial_1^2 + 3x_1\partial_2 \\ 2x_1\partial_1 + 3x_2\partial_2 - 6 \end{bmatrix}.$$

   Now let us determine $\mathcal{B}^V_{\{\omega\}}$ to see how precise the description given by the VMPUM is. Let $f \in \mathcal{A}_\mathcal{D}$.

   (a) $\partial_2^3 \bullet f = 0 \;\Rightarrow\; f = c_0 + c_1x_2 + c_2x_2^2$, where $c_i \in \mathbb{C}[[x_1]]$.

   (b) $\partial_1\partial_2 \bullet f = 0 \;\Rightarrow\; \partial_1 \bullet c_1 + 2x_2\partial_1 \bullet c_2 = 0 \;\Rightarrow\; \partial_1 \bullet c_1 = 0 \;\wedge\; \partial_1 \bullet c_2 = 0 \;\Rightarrow\; c_1, c_2 \in \mathbb{C}$.

   (c) $(\partial_1^3 + 3\partial_2^2) \bullet f = 0 \;\Rightarrow\; \partial_1^3 \bullet c_0 + 6c_2 = 0 \;\Rightarrow\; c_0 = -c_2x_1^3 + d_2x_1^2 + d_1x_1 + d_0$, where $d_i \in \mathbb{C}$.

   (d) $(x_2\partial_2^2 - \partial_2) \bullet f = 0 \;\Rightarrow\; c_1 = 0$.

   (e) $(3x_1\partial_2 + x_2\partial_1^2) \bullet f = 0 \;\Rightarrow\; d_2 = 0$.

   (f) $(2x_1\partial_1 + 3x_2\partial_2 - 6) \bullet f = 0 \;\Rightarrow\; -4d_1x_1 - 6d_0 = 0 \;\Rightarrow\; d_1 = 0 = d_0$.

   Hence, we obtain that $f = -c_2(x_1^3 - x_2^2)$, thus

$$\mathcal{B}^V_{\{\omega\}} = \{c(x_1^3 - x_2^2) \mid c \in \mathbb{C}\}.$$

   With respect to the requirement of being most powerful and linear, the VMPUM is as significant as possible. We observe that the VMPUM of a single non-zero signal has $\mathbb{C}$-dimension one. Actually, this holds in general, as will be shown in Theorem 5.2.7.

2. Suppose $\mathcal{D}$ to be the second **difference algebra**. Then by using SINGULAR we obtain:

$$R_{\text{VMPUM}} = \begin{bmatrix} \Delta_2^3 \\ \Delta_1\Delta_2 \\ \Delta_1^3 + 3\Delta_2^2 \\ 2x_2\Delta_2^2 + \Delta_2^2 - 2\Delta_2 \\ 2x_2\Delta_1^2 + \Delta_1^2 + 6x_1\Delta_2 + 6\Delta_2 \\ 8\Delta_1^2 + 21\Delta_2^2 + 24x_1\Delta_1 + 36x_2\Delta_2 - 24\Delta_1 - 18\Delta_2 - 72 \end{bmatrix}.$$

   Similar arguments as above lead to

$$\mathcal{B}^V_{\{\omega\}} = \{c(x_1^3 - x_2^2) \mid c \in \mathbb{C}\}.$$

3. Suppose $\mathcal{D}$ to be the second $\mathcal{SW}$ **algebra**. Then by using SINGULAR we obtain:

$$R_{\text{VMPUM}} = \begin{bmatrix} \Delta_2^3 \\ \Delta_1\Delta_2 \\ \Delta_1^3 + 3\Delta_2^2 \\ 2\partial_2 + \Delta_2^2 - 2\Delta_2 \\ 2\partial_1 + \Delta_1^2 - 2\Delta_1 + 2\Delta_2^2 \\ 2x_2\Delta_2^2 + \Delta_2^2 - 2\Delta_2 \\ 2x_2\Delta_1^2 + \Delta_1^2 + 6x_1\Delta_2 + 6\Delta_2 \\ 8\Delta_1^2 + 21\Delta_2^2 + 24x_1\Delta_1 + 36x_2\Delta_2 - 24\Delta_1 - 18\Delta_2 - 72 \end{bmatrix}.$$

Note that the output depends on the monomial ordering of the operators. In this example $\Delta_{1,2}$ were chosen to be greater than $\partial_{1,2}$. Taking a reverse ordering produces different (but equivalent) answers.

Comparing this matrix with the matrix above, we see that the rows of the matrix belonging to the difference case appear also here. We conclude that

$$\mathcal{B}^V_{\{\omega\}} = \{c(x_1^3 - x_2^2) \mid c \in \mathbb{C}\}.$$

Thus, taking $\mathcal{SW}$ as the operator algebra, we have got more equations than with the difference algebra. However, we have obtained interesting mixed differential-difference equations, which show the interplay of two different operator settings.

4. The second q**-difference algebra**:

$$R_{\text{VMPUM}} = \begin{bmatrix} \partial_2^2 + (-q^2+1)\partial_2 \\ (-q-1)\partial_1 + (-q^2-q-1)\partial_2 + (q^4+q^3-q-1) \\ x_1^3\partial_2 - x_2^2\partial_2 + (q^2-1)x_2^2 \end{bmatrix}$$

Let $f \in \mathcal{A}_{\mathcal{D}}$. Then $f = \sum_{i,j} c_{i,j} x_1^i x_2^j$.

(a) The first equation yields:

$$\begin{aligned} & \sum_{i,j} c_{i,j}(q^j-1)^2 x_1^i x_2^j + (-q^2+1)\sum_{i,j} c_{i,j}(q^j-1)x_1^i x_2^j = 0 \\ \Leftrightarrow\ & (q^j-1)^2 + (-q^2+1)(q^j-1) = 0 \\ \Leftrightarrow\ & j = 0 \vee j = 2. \end{aligned}$$

(b) Now consider the second equation.

i. Suppose $j = 2$, then

$$\begin{aligned} & (-q-1)\sum_{i,j} c_{ij}(q^i-1)x_1^i x_2^2 + (-q^2-q-1)\sum_{i,j} c_{ij}(q^2-1)x_1^i x_2^2 \\ & + (q^4+q^3-q-1)\sum_{i,j} c_{ij} x_1^i x_2^2 = 0 \\ \Leftrightarrow\ & (-q-1)(q^i-1) + (-q^2-q-1)(q^2-1) + (q^4+q^3-q-1) = 0 \\ \Leftrightarrow\ & i = 0. \end{aligned}$$

ii. Suppose $j = 0$, then

$$(-q-1)\sum_{i,j} c_{i0}(q^i-1)x_1^i + (q^4+q^3-q-1)\sum_i c_{i0}x_1^i = 0$$
$$\Leftrightarrow\ i = 3.$$

Thus $\quad f = c_{30}x_1^3 + c_{02}x_2^2.$

(c) Applying the last equation, we get

$$x_1^3 c_{02}(q^2-1)x_2^2 - x_2^2 c_{02}(q^1-1)x_2^2 + (q^2-1)x_2^2(c_{30}x_1^3 + c_{02}x_2^2) = 0$$
$$\Leftrightarrow\ \ x_1^3 t_2^2 (q^2-1)(c_{30}+c_{02}) = 0 \ \ \Leftrightarrow\ \ c_{30} = -c_{02}.$$

Thus we obtain once more

$$\mathcal{B}^V_{\{\omega\}} = \{c(x_1^3 - x_2^2) \mid c \in \mathbb{C}\}.$$

Remark 5.2.3 As we have seen in the previous example, the number of equations giving the VMPUM depends strongly on the underlying Ore algebra. In all cases, with Gröbner bases we get more equations than might be actually necessary. However, it is often possible to find a smaller generating set, which is usually not a Gröbner basis. Namely, one computes a left syzygy module of a given system and almost directly deduces a smaller generating set from it. For instance, only 3 of the 6 equations from the first example of 5.2.2 generate the whole ideal, namely $\partial_1\partial_2, \partial_1^3 + 3\partial_2^2, 2x_1\partial_1 + 3x_2\partial_2 - 6$. Analogous smaller generating sets can be obtained for the other examples.

Theorem 5.2.1 can be generalized to a set of several signals directly. A kernel representation of the VMPUM of $\Omega = \{\omega_1, \ldots, \omega_N\}$ is determined by stacking a set of generators of

$$\bigcap_{i=1}^{N} \ker(\kappa_{\omega_i})$$

row-wise into a matrix R.

Theorem 5.2.4 *Using the notation from above, the* VMPUM *of* Ω *equals*

$$\mathcal{B}^V_\Omega = \{g \in \mathcal{A}^m_\mathcal{D} \mid R \bullet g = 0\}.$$

Proof: By the definition of R, it is clear that $\Omega \subseteq \mathcal{B}^V_\Omega$. Also the property of being most powerful follows by the same arguments as used in the proof of Theorem 5.2.1. □

Example 5.2.5 Suppose $\mathcal{D} = \mathrm{W}_1$ to be the first Weyl algebra and $\mathcal{A}_\mathcal{D} = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$. Consider the signal set $\Omega = \{x, v_0x - v_1x^2\}$, where $v_0, v_1 \in \mathbb{C} \setminus \{0\}$. The second trajectory will appear in Example 5.2.9 again. Since

$$\ker(\kappa_x) \cap \ker(\kappa_{v_0x - v_1x^2})$$

$$
\begin{aligned}
&= {}_{\mathrm{W}_1}\langle x\partial - 1,\ \partial^2\rangle \cap {}_{\mathrm{W}_1}\langle -v_0^2\partial^2 + (4v_1^2 x - 2v_0v_1)\partial - 8v_1^2,\ \partial^3\rangle \\
&= {}_{\mathrm{W}_1}\langle x^2\partial^2 - 2x\partial + 2, \partial^3\rangle,
\end{aligned}
$$

the VMPUM of Ω is given by

$$\mathcal{B}_\Omega^V = \{c_1 x + c_2 x^2 \mid c_1, c_2 \in \mathbb{C}\}.$$

The intersection of submodules of a free module over a Noetherian Ore algebra can be computed as in Corollary 2.2.15, for instance with the system SINGULAR::PLURAL [GLH05].

5.2.1 VMPUM using the Weyl algebra

In this section, we suppose $\mathcal{D}$ to be the n-th Weyl algebra

$$\mathcal{D} = \mathrm{W}_n := \mathbb{C}[x_1, \ldots, x_n][\partial_1; \mathrm{id}_{\mathrm{W}_n}, \tfrac{\partial}{\partial x_1}] \cdots [\partial_n; \mathrm{id}_{\mathrm{W}_n}, \tfrac{\partial}{\partial x_n}].$$

Thus for $p \in \mathbb{C}[x_1, \ldots, x_n]$, we obtain

$$\partial_i \bullet p := \frac{\partial p}{\partial x_i}.$$

Further suppose $\mathcal{A}_\mathcal{D}$ to be $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$, the space of smooth functions. Identifying a polynomial with the corresponding polynomial function, we obtain $\mathbb{C}[x_1, \ldots, x_n] = A \subseteq \mathcal{A}_\mathcal{D}$.
In this context, the VMPUM was already introduced in [SLZ08]. Here, we will recall some results and additionally point out a new interesting property.

$\mathbb{C}$-dimension

Since all partial differential equations with constant coefficients are contained in the set of partial differential equations with polynomial coefficients, the VMPUM of a certain signal set is contained in the corresponding MPUM. Thus due to Theorem 1.6.1, the VMPUM is a finitely generated $\mathbb{C}$-vector space. In some cases, we can determine the dimension more precisely. We claim that the VMPUM of a single non-zero signal has $\mathbb{C}$-dimension one.
Suppose $p \in A^m$. Every polynomial p_i can be written as $\sum_{k=1}^{h_i} c_{ik} x^{\beta_{ik}}$, where $c_{ik} \in \mathbb{C}$ for all i, k. Let $\mathcal{E}_i := \{\beta_{i1}, \ldots, \beta_{ih_i}\} \subset \mathbb{N}^n$ denote the set of all exponent multi-indices occurring in p_i and let

$$d_{ij} := \max_{1 \le k \le h_i} \{(\beta_{ik})_j \mid \beta_{ik} \in \mathcal{E}_i\} \tag{5.3}$$

be the highest degree in x_j of p_i. Recall that by e_i we denote the i-th canonical generator of the free module A^m. The set

$$\mathcal{E}_{p_i} = \{\alpha \in \mathbb{N}^n \mid \alpha_j \le d_{ij} + 1 \ \text{ for } \ 1 \le j \le n\}$$

is finite, that is, $\mathcal{E}_{p_i} = \{\alpha_{i1}, \ldots, \alpha_{il_i}\}$. Define for $p \in A^m$:

$$\mathrm{Der}_p = \left(p_1, \frac{\partial^{|\alpha_{11}|} p_1}{\partial^{\alpha_{11}}}, \ldots, \frac{\partial^{|\alpha_{1l_1}|} p_1}{\partial^{\alpha_{1l_1}}}, \ldots, p_m, \frac{\partial^{|\alpha_{m1}|} p_m}{\partial^{\alpha_{m1}}}, \ldots, \frac{\partial^{|\alpha_{ml_m}|} p_m}{\partial^{\alpha_{ml_m}}}\right).$$

Let $\mathrm{Syz}(\mathrm{Der}_p)$ denote the module of polynomial syzygies. Define the A-module homomorphism

$$\Phi_p : \mathrm{Syz}(\mathrm{Der}_p) \to \ker(\kappa_p), \quad (q_1, \ldots, q_l) \mapsto (q_1, \ldots, q_l) \cdot \begin{bmatrix} 1 & & \\ \partial^{\alpha_{11}} & & \\ \vdots & & \\ \partial^{\alpha_{1l_1}} & & \\ & \ddots & \\ & & 1 \\ & & \partial^{\alpha_{m1}} \\ & & \vdots \\ & & \partial^{\alpha_{ml_m}} \end{bmatrix},$$

which is clearly injective.

Lemma 5.2.6 *The equality* ${}_{\mathrm{W}_n}\langle \mathrm{im}(\Phi_p) \rangle = \ker(\kappa_p)$ *holds.*

Proof: Evidently ${}_{\mathrm{W}_n}\langle \mathrm{im}(\Phi_p) \rangle \subseteq \ker(\kappa_p)$. Now suppose that $a \in \ker(\kappa_p)$. Since every element in W_n can be written in normal form, we obtain

$$\begin{aligned} a \bullet p &= \sum_k a_k \bullet p_k = \sum_k (\sum_j c_{kj} x^{\beta_{kj}} \partial^{\gamma_{kj}}) \bullet p_k \\ &= \sum_k (\sum_j c_{kj} x^{\beta_{kj}}) (\partial^{\gamma_{kj}} \bullet p_k). \end{aligned}$$

Let us split the element a in a_z and a_{nz} such that $a = a_z + a_{nz}$ and $(a_z)_k$ consists of the parts of a_k where $\partial^{\gamma_{kj}} \bullet p_k$ is zero.
By the choice of d_{ij}, the set $\{\partial_j^{(d_{ij}+1)} \mid 1 \leq j \leq n,\ 1 \leq i \leq m\}$ generates the set of ∂^γ with the property that there exists $1 \leq i \leq m$ such that $\partial^\gamma \bullet p_i = 0$. Then a_z is contained ${}_{W_n}\langle \partial^{(d_{ij}+1)\,\mathrm{e}_j} \mid 1 \leq j \leq n,\ 1 \leq i \leq m \rangle$. But by the choice of Der_p, the element a_z is in the image of Φ_p. Suppose $\partial^{\gamma_{kj}} \bullet p_k \neq 0$, then γ_{kj} is equal or smaller than $(d_{k1}, \ldots, d_{kn})$ in each component and again by the choice of Der_p, the element a_{nz} is contained in the image of Φ_p. Thus it follows that $a \in \mathrm{im}(\Phi_p)$. □

Theorem 5.2.7 *The* VMPUM *of* $p \neq 0$ *is a one-dimensional vector space over* $\mathbb{C}$.

Proof: We use the notation of the Lemma 5.2.6, which reduces the problem to commutative calculations. It is easy to see that the equivalence

$$s \bullet (f_1, \partial^{\alpha_{11}} \bullet f_1, \ldots, f_m, \ldots, \partial^{\alpha_{ml_m}} \bullet f_m)^T = 0 \quad \Leftrightarrow \quad \Phi_p(s) \bullet f = 0 \tag{5.4}$$

holds for every $s \in \mathrm{Syz}(\mathrm{Der}_p)$ and each $f \in A^m$. Now let us discuss the left hand side. Consider the solution space $\mathrm{Sol}(\mathrm{Syz}(\mathrm{Der}_p)) := \{\omega \in \mathcal{A}_{\mathcal{D}}^l \mid S \bullet \omega = 0\}$, where the rows of S generate the module $\mathrm{Syz}(\mathrm{Der}_p)$. Since Der_p contains all non-zero derivatives of

p_i for all i, there exists a component of Der_p which is a non-zero constant $k \in \mathbb{C}$. We can suppose $k = 1$ and without loss of generality let $\partial^{\alpha_{ml_m}} \bullet p_m = 1$. Then

$$(-1, 0, \ldots, 0, p_1),\ \ldots,\ (0, \ldots, 0, -1, \frac{\partial^{|\alpha_{m(l_m-1)}|} p_m}{\partial^{\alpha_{m(l_m-1)}}}) \in \mathrm{Syz}(\mathrm{Der}_p)$$

and thus

$$\mathrm{Sol}(\mathrm{Syz}(\mathrm{Der}_p)) = \{c \cdot (p_1, \frac{\partial^{|\alpha_{11}|} p_1}{\partial^{\alpha_{11}}}, \ldots, \frac{\partial^{|\alpha_{m(l_m-1)}|} p_m}{\partial^{\alpha_{m(l_m-1)}}}) \mid c \in \mathbb{C}\}. \tag{5.5}$$

Now suppose that f is contained in the VMPUM of p. From Lemma 5.2.6 together with (5.4) and (5.5), we deduce the claim. □

Remark 5.2.8 In the case of a single non-zero signal, the VMPUM gives the most precise description one can get with a linear system.

Example 5.2.9 Consider the trajectory $\omega(x) = v_0 x - v_1 x^2$, where $v_0, v_1 \in \mathbb{C} \setminus \{0\}$. Then the MPUM of ω is given by

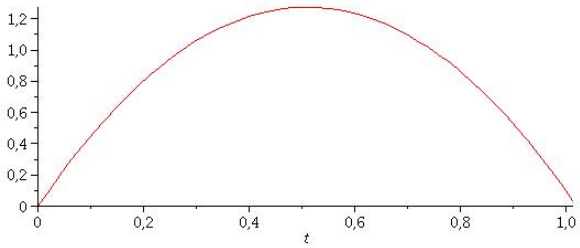


Figure 5.1: $\omega(x) = 5x - 0.5\,\mathrm{g}\,x^2$

$$\begin{aligned}\mathcal{B}_{\{\omega\}} &= \{\alpha(v_0 x - v_1 x^2) + \beta(v_0 - 2v_1 x) + \gamma(-2v_1) \mid \alpha, \beta, \gamma \in \mathbb{C}\} \\ &= \{ax^2 + bx + c \mid a, b, c \in \mathbb{C}\} \\ &= \{w \in \mathcal{A}_{\mathcal{D}} \mid \partial^3 \bullet w = 0\}.\end{aligned}$$

Thus there are three free parameters to choose. The VMPUM of ω is given by

$$\begin{aligned}\mathcal{B}^V_{\{\omega\}} &= \{w \in \mathcal{A}_{\mathcal{D}} \mid \begin{bmatrix} -v_0^2\partial^2 + (4v_1^2 x - 2v_0 v_1)\partial - 8v_1^2 \\ \partial^3 \end{bmatrix} \bullet w = 0\} \\ &= \{c\,(v_0 x - v_1 x^2) \mid c \in \mathbb{C}\},\end{aligned}$$

that is, two degrees of freedom vanish when we consider the time-variant model.

Structural properties

Let us discuss some structural properties of the W_n-module $\ker(\kappa_p)$. Since every element of W_n can be transformed into normal form, the degree of a non-zero element $a = \sum_{\alpha,\beta\in\mathbb{N}^n} a_{\alpha,\beta} x^\alpha \partial^\beta \in \mathrm{W}_n$, where $a_{\alpha,\beta} \in \mathbb{C}$, can be introduced as

$$\deg(a) := \max\{|(\alpha,\beta)| \mid a_{\alpha,\beta} \neq 0\}.$$

Then $\mathcal{F}^i(\mathrm{W}_n) := \{a \in \mathrm{W}_n \mid \deg(a) \leq i\}$ induces a filtration on W_n. The corresponding associated graded ring $\mathrm{Gr}(\mathrm{W}_n)$ is isomorphic to $\mathbb{C}[x_1,\ldots,x_n,\partial_1,\ldots,\partial_n]$ as a $\mathbb{C}$-algebra. For every finitely generated W_n-module M, we define the Hilbert polynomial $\mathrm{HP}_M^{\mathrm{W}_n} := \mathrm{HP}_{\mathrm{Gr}(M)}^{\mathrm{Gr}(\mathrm{W}_n)}$. The dimension of M is defined as $\dim_{\mathrm{W}_n}(M) := \deg(\mathrm{HP}_M) + 1$. Furthermore M is called **holonomic** if it has dimension n. A holonomic module is of minimal dimension, since the dimension of W_n-modules is bounded below by n and bounded above by $2n$. Holonomic W_n-modules are additionally cyclic and torsion modules. For details see [Cou95].

We write $A = \mathbb{C}[x_1,\ldots,x_n] \subset \mathrm{W}_n$.

Theorem 5.2.10 *Let p be unequal to zero. There is an isomorphism of W_n-modules*

$$\mathrm{W}_n^{1\times m} / \ker(\kappa_p) \cong A.$$

Proof: Since κ_p is a homomorphism of W_n-modules, we get

$$\mathrm{W}_n^{1\times m} / \ker(\kappa_p) \cong \mathrm{im}(\kappa_p) \subseteq \mathrm{W}_n /_{\mathrm{W}_n}\langle \partial_1,\ldots,\partial_n\rangle \cong A.$$

Thus $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ is isomorphic to a submodule of A. Due to the fact that A is a simple W_n-module and $\mathrm{W}_n^{1\times m} / \ker(\kappa_p) \neq 0$, the claim follows. □

Corollary 5.2.11 *Let p be unequal to zero. The W_n-module $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ is holonomic.*

Proof: This follows from Theorem 5.2.10 and the property of A being a simple W_n-module (see [Cou95], Chapter 5, Proposition 1.2). □

Corollary 5.2.12 *Since $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ is holonomic, there exists a left ideal L_p, depending on p, such that $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ is isomorphic to the cyclic left W_n-module W_n / L_p.*

An algorithm, using Gröbner bases, to compute a generator of $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ is given in [Ley04]. On the other hand, [Cou95] shows that a generic element of $\mathrm{W}_n^{1\times m} / \ker(\kappa_p)$ can be taken as a generator for the cyclic module.

Example 5.2.13 Suppose $\omega = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ for $c_1, c_2, c_3 \in \mathbb{C} \setminus \{0\}$. Then

$$\ker(\kappa_\omega) =_{\mathrm{W}_1} \langle [0, c_3, -c_2], [c_3, 0, -c_1], [0, 0, \partial] \rangle.$$

Since

$$\begin{bmatrix} 0 & c_3 & -c_2 \\ c_3 & 0 & -c_1 \\ 0 & 0 & \partial \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 0 & 1/c_3 & c_1/c_3 \\ 1/c_3 & 0 & c_2/c_3 \\ 0 & 0 & 1 \end{bmatrix}}_{=:C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \partial \end{bmatrix},$$

we obtain

$$\mathrm{W}_1^{1\times 3} / \ker(\kappa_\omega) \cong \mathrm{W}_1^{1\times 3} / \ker(\kappa_\omega) C \cong \mathrm{W}_1 /_{\mathrm{W}_1} \langle \partial \rangle \cong \mathbb{C}[x].$$

VMPUM of polynomial-exponential signals

In this section, we extend the signal space that should be modeled. The goal is to compute the VMPUM of

$$p = \begin{bmatrix} p_1 \exp_{\lambda^1} \\ \vdots \\ p_m \exp_{\lambda^m} \end{bmatrix}, \tag{5.6}$$

where for all $1 \le i \le m$, we have $p_i \in A$, $\lambda^i \in \mathbb{C}^n$ and

$$\exp_\lambda(x) := \exp(\lambda_1 x_1 + \cdots + \lambda_n x_n) \quad \text{for } \lambda \in \mathbb{C}^n.$$

By the action $\partial_j \bullet \exp_\lambda = \lambda_j \exp_\lambda$ for all $1 \le j \le n$, the space of polynomial-exponential functions becomes a W_n-module.
Consider the scalar setting first, that is, $m = 1$. Define for $\lambda \in \mathbb{C}^n$ the W_n-homomorphism

$$\sigma_\lambda : \mathrm{W}_n \to \mathrm{W}_n, \quad \partial_i \mapsto (\partial_i - \lambda_i), \quad x_i \mapsto x_i.$$

It is easy to see that σ_λ is a W_n-automorphism. We claim that for $a \in \mathrm{W}_n$ and $g \in A$

$$a \bullet p = 0 \quad \text{if and only if} \quad \sigma_\lambda(a) \bullet (p \exp_\lambda) = 0. \tag{5.7}$$

For the proof suppose $a = \sum_i c_i x^{\alpha_i} \partial^{\beta_i}$.
Using the identity $(\partial_i - \lambda_i) \bullet (p \exp_\lambda) = (\partial_i \bullet p) \exp_\lambda$, the claim follows by

$$\begin{aligned}
\sigma_\lambda(a) \bullet (p \exp_\lambda) &= (\sum_i c_i x^{\alpha_i} \sigma_\lambda(\partial^{\beta_i})) \bullet (p \exp_\lambda) \\
&= \sum_i c_i x^{\alpha_i} ((\partial_1 - \lambda_1)^{\beta_{i1}} \cdots (\partial_n - \lambda_n)^{\beta_{in}}) \bullet (p \exp_\lambda) \\
&= \sum_i c_i x^{\alpha_i} ((\partial_1^{\beta_{i1}} \cdots \partial_n^{\beta_{in}}) \bullet p) \exp_\lambda \\
&= \left(\sum_i c_i x^{\alpha_i} ((\partial_1^{\beta_{i1}} \cdots \partial_n^{\beta_{in}}) \bullet p) \right) \exp_\lambda \quad = \quad (a \bullet p) \exp_\lambda .
\end{aligned}$$

Extending the dimension, there are two special cases requiring attention. First suppose $\lambda^1, \ldots, \lambda^m$ to be equal, that is, $p = [p_1, \ldots, p_m]^T \exp_\lambda$, where $\lambda := \lambda^1$. Then claim (5.7) can be generalized directly and it follows that

$$\sum_{i=1}^{m} a_i \bullet (p_i \exp_\lambda) = 0 \text{ if and only if } [a_1, \ldots, a_m] \in \sigma_\lambda(\ker(\kappa_p)). \tag{5.8}$$

Assume now that $\lambda^1, \ldots, \lambda^m$ are pairwise different. Then

$$\sum_{j=1}^{m} a_j \bullet (p_j \exp_{\lambda^j}) = 0 \quad \text{if and only if} \quad [a_1, \ldots, a_m] \in \bigoplus_{j=1}^{m} \sigma_{\lambda^j}(\ker(\kappa_{p_j})). \tag{5.9}$$

Since $\exp_{\lambda^1}, \ldots, \exp_{\lambda^m}$ are algebraically independent over A, the claim follows from

$$\begin{aligned}
& \sum_{j=1}^{m} a_j \bullet (p_j \exp_{\lambda^j}) = 0 \\
\Leftrightarrow \quad & \sum_{j=1}^{m} \left(\sum_{i=1}^{h_j} c_{ji} x^{\alpha_{ji}} \partial^{\beta_{ji}} \bullet (p_j \exp_{\lambda^j}) \right) = 0 \\
\Leftrightarrow \quad & \sum_{j=1}^{m} \left(\sum_{i=1}^{h_j} c_{ji} x^{\alpha_{ji}} (\partial_1 + \lambda_1^j)^{(\beta_{ji})_1} \ldots (\partial_n + \lambda_n^j)^{(\beta_{ji})_n} \bullet p_j \right) \exp_{\lambda^j} = 0 \\
\Leftrightarrow \quad & \sum_{j=1}^{m} (\sigma_{\lambda^j}^{-1}(a_j) \bullet p_j) \exp_{\lambda^j} = 0 \\
\Leftrightarrow \quad & \sigma_{\lambda^j}^{-1}(a_j) \in \ker(\kappa_{p_j}) \quad \text{for all } 1 \leq j \leq m.
\end{aligned}$$

Recapitulating we get:

Theorem 5.2.14 *Let f be of the form (5.6). Further let*

$$K_i := \left\{ j \mid \lambda^j = \lambda^i \right\} = \{k_{i1}, \ldots, k_{il_i}\}$$

and let l be chosen minimal such that we have a disjoint union

$$K_1 \dot\cup \ldots \dot\cup K_l = \{k_{11}, \ldots, k_{1h_1}, \ldots, k_{l1}, \ldots, k_{lh_l}\} = \{1, \ldots, m\}.$$

Further define the vector $h_i := [f_{k_{i1}}, \ldots, f_{k_{il_i}}]^T$ and $H_i := \sigma_{\lambda^i}(\ker(\kappa_{h_i}))$. Let $\mathrm{e}_{k_{ij}}$ denote the k_{ij}-th canonical generator of $\mathrm{W}_n^{1\times m}$ for $1 \leq i \leq l$ and $1 \leq l \leq h_i$. Defining for $1 \leq i \leq l$,

$$\phi_i : H_i \to \mathrm{W}_n, \quad [a_1, \ldots, a_{h_i}] \mapsto \sum_{j=1}^{h_i} a_j \, \mathrm{e}_{k_{ij}},$$

the VMPUM *of f is given by*

$$\bigoplus_{i=1}^{l} \phi_i(H_i).$$

Proof: After choosing a suitable projection, the claim follows by (5.7) and (5.8). □

5.2.2 VMPUM using the difference algebra

Suppose K to be of characteristic zero. Recall the definition of the n-th difference algebra:

$$\mathcal{S}_n := K[x_1, \ldots, x_n][\Delta_1; \sigma_1, \delta_1] \cdots [\Delta_n; \sigma_n, \delta_n].$$

For $p \in K[x_1, \ldots, x_n]$, we have

$$\Delta_i \bullet p = \delta_i(p) = \sigma_i(p) - p \quad \text{and thus} \quad (\Delta_i \bullet p)(x) = p(x + \mathrm{e}_i) - p(x).$$

Further suppose that $\mathcal{A}_\mathcal{D} = K^{\mathbb{N}^n}$. Identifying a polynomial with the corresponding polynomial function, we obtain $A = K[x_1, \ldots, x_n] \subseteq \mathcal{A}_\mathcal{D}$.
Similarly to the continuous case, the kernel of κ_p can be computed in a completely commutative framework. For this purpose, we assume p to be represented as in (1.14), namely

$$p = \sum_{\nu \in \mathbb{N}^n, \nu \leq_{cw} \varrho} c_\nu p_\nu,$$

where c_ν are suitable chosen and

$$p_\nu : \mathbb{N}^n \to K, \quad x \mapsto \binom{x_1}{\nu_1} \cdots \binom{x_n}{\nu_n}.$$

Remark 5.2.15 Connecting Remark 1.6.2 and (1.16), we get that $\delta^\mu p = 0$ for all μ with $\mu_i > \varrho_i$ for at least one $1 \leq i \leq n$. Now consider the sequence

$$\mathrm{Diff}_p = (\delta^{\mu_1} p, \delta^{\mu_2} p, \ldots, \delta^{\mu_\ell} p)$$

for pairwise different μ_i satisfying $\mu_i \leq_{cw} \varrho$ for all i. Since A is a Noetherian ring, the corresponding syzygy module $\mathrm{Syz}(\mathrm{Diff}_p)$ is finitely generated by $s_1, \ldots, s_d$. Analogously to the continuous case, we can give an A-module homomorphism from $\mathrm{Syz}(\mathrm{Diff}_p)$ to $\ker(\kappa_p)$, such that the image of $s_1, \ldots, s_d$ under this map generates $\ker(\kappa_p)$, that is, $\ker(\kappa_p)$ is finitely generated as an A-module. This implies that $\ker(\kappa_p)$ is finitely generated as an $\mathcal{S}_n$-module.

Example 5.2.16 Let $p = [x^3, x]^T$. Then the continuous VMPUM is the same as the discrete VMPUM, that is, equal to $\{c[x^3, x]^T \mid c \in K\}$. Direct computation over $\mathcal{S}_1 = K[x][\Delta; \sigma, \delta]$ yields

$$\ker_{\mathcal{S}_1}(\kappa_p) = {}_{\mathcal{S}_1}\langle [0, \Delta^2],\ [0, x\Delta - 1],\ [1, -x^2] \rangle$$

and this means that

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

is a kernel representation of the VMPUM of p. Note that over the first polynomial Weyl algebra W_1, we have

$$\ker_{\mathrm{W}_1}(\kappa_p) = {}_{\mathrm{W}_1}\langle [0, \partial^2],\ [0, x\partial - 1],\ [1, -x^2] \rangle.$$

Alternatively, we can compute $\ker_{\mathcal{S}_1}(\kappa_p)$ in the commutative framework, using the analogon of the "difference algebra" approach. At first, we observe that

$$\mathrm{Diff}_{[x^3,x]^T} = {}_{K[x]}\langle x^3, 3x^2+3x+1, 6x+6, 6, x, 1\rangle$$

so

$$\mathrm{Syz}(\mathrm{Diff}_{[x^3,x]^T}) = {}_{K[x]}\langle \begin{bmatrix} 0\\0\\0\\0\\1\\-x \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\1\\0\\-6 \end{bmatrix}, \begin{bmatrix} 0\\0\\1\\0\\0\\-6x-6 \end{bmatrix}, \begin{bmatrix} 0\\1\\0\\0\\0\\-3x^2-3x-1 \end{bmatrix}, \begin{bmatrix} 1\\0\\0\\0\\0\\-x^3 \end{bmatrix}\rangle.$$

Finally we get that

$$\begin{aligned} &\ker(\kappa_p)\\ =&_{\mathcal{S}_1}\langle [0,-x\Delta+1],\ [\Delta^3,-6\Delta],\ [\Delta^2,(-6x-6)\Delta],\\ &\quad [\Delta,(-3x^2-3x-1)\Delta],\ [1,-x^3\Delta],\ [\Delta^4,0],\ [0,\Delta^2]\rangle\\ =&_{\mathcal{S}_1}\langle [0,\Delta^2],\ [0,x\Delta-1],\ [1,-x^2]\rangle. \end{aligned}$$

VMPUM of polynomial-exponential signals

For $\lambda=(\lambda_1,\ldots,\lambda_n)\in K^n$, the discrete exponential function is given by

$$\exp_\lambda : \mathbb{N}^n \to K,\ \ x\mapsto \lambda^x=\lambda_1^{x_1}\cdots\lambda_n^{x_n}.$$

First suppose that $m=1$, that is, we want to construct the VMPUM of a scalar polynomial exponential trajectory of the form $p\exp_\lambda$, where $p\in A$. Without loss of generality, we can assume $\lambda_i\neq 0$ for all $1\leq i\leq n$, because in case $\lambda_i=0$ for one $1\leq i\leq n$, we have

$$p(x)\exp_\lambda(x)=\begin{cases} 0 & \text{if } x_i\neq 0\\ g(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) & \text{if } x_i=0,\end{cases}$$

where g is a polynomial-exponential function on $\mathbb{N}^{n-1}$. Consider the $\mathcal{S}_n$-algebra-automorphism

$$\chi_\lambda : \mathcal{S}_n\to\mathcal{S}_n,\ \ \begin{cases} x_i\mapsto x_i\\ \Delta_i\mapsto \frac{1}{\lambda_i}(\Delta_i-\lambda_i+1).\end{cases}$$

Since the equality

$$\begin{aligned} \chi_\lambda(\Delta_i)\bullet(p\exp_\lambda) &= \frac{1}{\lambda_i}(\Delta_i-\lambda_i+1)\bullet(p\exp_\lambda)\\ &= \frac{1}{\lambda_i}\left(\Delta_i\bullet(p\exp_\lambda)-\lambda_i p\exp_\lambda+p\exp_\lambda\right)\\ &= \frac{1}{\lambda_i}\left(\lambda_i\exp_\lambda\sigma_i(p)-p\exp_\lambda-\lambda_i p\exp_\lambda+p\exp_\lambda\right)\\ &= \frac{1}{\lambda_i}\left(\lambda_i\exp_\lambda(\sigma_i(p)-p)\right)\\ &= \exp_\lambda\Delta_i\bullet p \end{aligned}$$

holds, we obtain the identity

$$\chi_\lambda(\Delta_i^k) \bullet (p \exp_\lambda) = \exp_\lambda \Delta_i^k \bullet p$$

that finally extends to

$$\chi_\lambda(\Delta^\mu) \bullet (p \exp_\lambda) = \chi_\lambda^\mu(\Delta) \bullet (p \exp_\lambda) = \exp_\lambda \Delta^\mu \bullet p. \tag{5.10}$$

Now using (5.10), we can deduce for $a = \sum_{i=1}^h a_i \Delta^{\alpha_i} \in \mathcal{S}_n$ the equivalence

$$a \bullet p = 0 \quad \Leftrightarrow \quad \chi_\lambda(a) \bullet (\exp_\lambda p) = 0, \tag{5.11}$$

since

$$\begin{aligned}
\chi_\lambda(a) \bullet (\exp_\lambda p) &= \sum_{i=1}^h a_i \chi_\lambda(\Delta^{\alpha_i}) \bullet (\exp_\lambda p) \\
&= \sum_{i=1}^h a_i (\Delta^{\alpha_i} \bullet p) \exp_\lambda \\
&= \exp_\lambda \sum_{i=1}^h a_i (\Delta^{\alpha_i} \bullet p) \\
&= \exp_\lambda a \bullet p.
\end{aligned}$$

For $m = 1$, that is, $p \in A$, we obtain the following result.

Theorem 5.2.17 *Let us denote by $R \in \mathcal{S}_n^{l\times 1}$ a kernel representation matrix of the* VMPUM *of p. Then the kernel representation matrix of $p \exp_\lambda$ is given by $(\chi_\lambda(R_i))_i$.*

Proof: The proof is given by (5.11). □

Now consider $p \in A^m$, where

$$p = \begin{bmatrix} p_1 \exp_{\lambda^{(1)}} \\ \vdots \\ p_m \exp_{\lambda^{(m)}} \end{bmatrix}, \qquad \lambda^{(i)} \in (K\backslash\{0\})^n, \;\; p_i \in A \tag{5.12}$$

and $\lambda^{(1)}, \ldots, \lambda^{(m)}$ are pairwise different. Then

$$\sum_{j=1}^m a_j \bullet (p_j \exp_{\lambda^{(j)}}) = 0 \quad \text{if and only if} \quad [a_1, \ldots, a_m] \in \bigoplus_{j=1}^m \chi_{\lambda^{(j)}}(\ker(\kappa_{p_j})), \tag{5.13}$$

which follows from

$$\begin{aligned}
&\sum_{j=1}^m a_j \bullet (p_j \exp_{\lambda^{(j)}}) = 0 \\
\Leftrightarrow &\sum_{j=1}^m \left(\sum_{i=1}^{h_j} c_{ji} x^{\alpha_{ji}} \Delta^{\beta_{ji}} \bullet (p_j \exp_{\lambda^{(j)}}) \right) = 0
\end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \textstyle\sum_{j=1}^{m}\left(\sum_{i=1}^{h_j} c_{ji}x^{\alpha_{ji}}(\lambda_1^{(j)}\Delta_1+\lambda_1^{(j)}-1)^{(\beta_{ji})_1}\ldots(\lambda_n^{(j)}\Delta_n+\lambda_n^{(j)}-1)^{(\beta_{ji})_n}\bullet p_j\right)\exp_{\lambda^{(j)}}=0\\
&\Leftrightarrow \sum_{j=1}^{m}(\chi_{\lambda^{(j)}}^{-1}(a_j)\bullet p_j)\exp_{\lambda^{(j)}}=0\\
&\Leftrightarrow \chi_{\lambda^{(j)}}^{-1}(a_j)\in\ker(\kappa_{p_j})\quad\text{for all } 1\leq j\leq m.
\end{aligned}$$

Additionally we get

Theorem 5.2.18 *Let p be of the form (5.12). Further let $K_i := \{j \mid \lambda^j = \lambda^i\} = \{k_{i1},\ldots,k_{il_i}\}$ and l chosen minimal such that the disjoint union*

$$K_1\dot\cup\ldots\dot\cup K_l = \{k_{11},\ldots,k_{1h_1},\ldots,k_{l1},\ldots,k_{lh_l}\} = \{1,\ldots,m\}.$$

Further define the vector $h_i := [f_{k_{i1}},\ldots,f_{k_{il_i}}]^T$ and $H_i := \chi_{\lambda^{(i)}}(\ker(\kappa_{h_i}))$. Let $\mathrm{e}_{k_{ij}}$ denote the k_{ij}-th standard generator of $\mathcal{S}_n^{1\times m}$ for $1\leq i\leq l$ and $1\leq l\leq h_i$. Defining for $1\leq i\leq l$

$$\phi_i : H_i\to\mathcal{S}_n,\quad [a_1,\ldots,a_{h_i}]\mapsto\sum_{j=1}^{h_i}a_j\,\mathrm{e}_{k_{ij}}$$

the VMPUM *of p is given by*

$$\bigoplus_{i=1}^{l}\phi_i(H_i).$$

Proof: Choosing a suitable projection, the claim follows by (5.11) and (5.13). □

Conclusion and future work

The main result of Chapter 3 is Theorem 3.2.3, which leads to the novel concept of "Gröbner p-basis" (Definition 3.2.4). In fact, Theorem 3.2.3 shows that a Gröbner p-basis has a particular type of PLM property which can be applied to yield straightforward solutions to several problems involving systems over $\mathbb{Z}_{p^r}$. A topic of future research is to investigate the use of the POT ordering to derive novel results on a Smith-McMillan like form for polynomial matrices over $\mathbb{Z}_{p^r}$. Further it remains open to elaborate connections with Janet bases [GY05, PR05] for multivariate Gröbner bases over fields, where restrictions on coefficients are also used.

In Chapter 4, novel methods to obtain a diagonalization, that is, a decoupled form, are investigated. Algorithm 2 and Algorithm 3 can be applied to a very general setting which is one of the major achievements of this chapter. Another issue that should be stressed is the purely fraction-free setting in which Algorithm 3 works. Topic of future studies is an implementation of the ideas proposed in Remark 4.2.12, namely the application of a random vector to generate a Jacobson form from a decoupled form.

Generalizing ideas from systems theory, we have defined a "varying most powerful unfalsified model" (VMPUM) over polynomial Ore algebras such as the Weyl algebra or the difference algebra in Chapter 5. Structural properties of the resulting models were presented, and we have seen, in terms of examples, that models with polynomial coefficients provide a more precise description of the data than models with constant coefficients. It remains to be investigated how rational coefficients perform in that respect. Another topic of future concerns may be a characterization of the vector space dimension of the VMPUM of several trajectories, thus generalizing Theorem 5.2.7. Moreover, it seems possible to develop VMPUMs with polynomial coefficients for data represented by rational and by rational-exponential functions.

Bibliography

[AL94] W. W. Adams and P. Loustaunau. *An Introduction to Gröbner Bases*. Graduate Studies in Mathematics. American Mathematical Society, 1994.

[Ape88] J. Apel. *Gröbnerbasen in nichtkommutativen Algebren und ihre Anwendung*. PhD thesis, Universität Leipzig, 1988.

[AW93] A.C. Antoulas and J.C. Willems. A behavioral approach to linear exact modeling. *IEEE Transactions on Automatic Control*, 38:1776–1802, 1993.

[BCG+03] Y. A. Blinkov, C. F. Cid, V. P. Gerdt, W. Plesken, and D. Robertz. The maple package "janet": II. linear partial differential equations. In *Proceedings of the 6th International Workshop on Computer Algebra in Scientific Computing*, Passau, Germany, 2003.

[BF01] E. Byrne and P. Fitzpatrick. Gröbner bases over Galois rings with an application to decoding alternant codes. *Journal of Symbolic Computation*, 31:565–584, 2001.

[BGTV03] J. Bueso, J. Gómez-Torrecillas, and A. Verschoren. *Algorithmic Methods in Non-commutative Algebra. Applications to Quantum Groups*. Kluwer Academic Publishers, 2003.

[BHK92] S. Boztaş, R. Hammons, and P.V. Kumar. 4-Phase sequences with near-opimum correlation properties. *IEEE Transactions on Information Theory*, 38:1101–1113, 1992.

[Bou05] H. Bourlès. Structural properties of discrete and continuous linear time-varying systems. In *Advanced Topics in Control Systems Theory*, 2005.

[Buc01] B. Buchberger. Gröbner bases: A short introduction for systems theorists. In *Computer Aided Systems Theory, EUROCAST 2001*. Springer Berlin / Heidelberg, 2001.

[CBSW02] K. Camlibel, M. N. Belur, A. J. Sasane, and J. C. Willems. On a class of time-varying behaviors. In *Proceedings of the 15th International Symposium on Mathematical Theory of Networks and Systems*, South Bend, USA, 2002.

[Chy98] F. Chyzak. Gröbner bases, symbolic summation and symbolic integration. In B. Buchberger and F. Winkler, editors, *Gröbner Bases and Applications*. London Mathematical Society Lecture Notes Series, 1998.

[Coh71] C. Cohn. *Free Rings and their Relations*. Academic Press, 1971.

[Cou95] S. C. Coutinho. *A Primer of Algebraic D-modules*. Cambridge University Press, 1995.

[CQ05] G. Culianez and A. Quadrat. Formes de Hermite et de Jacobson: implementations et applications. Technical report, INRIA Sophia Antipolis, 2005.

[CQR04] F. Chyzak, A. Quadrat, and D. Robertz. OreModules: A symbolic package for the study of multidimensional linear systems. In *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*, Leuven, Belgium, 2004.

[CQR05] F. Chyzak, A. Quadrat, and D. Robertz. Effective algorithms for parametrizing linear control systems over Ore algebras. Technical report, INRIA Sophia Antipolis, 2005.

[CQR07] F. Chyzak, A. Quadrat, and D. Robertz. OreModules: A symbolic package for the study of multidimensional linear systems. In J. Chiasson and J. J. Loiseau, editors, *Applications of Time-Delay Systems*. Springer, 2007.

[CS98] F. Chyzak and B. Salvy. Non-commutative elimination in Ore algebras proves multivariate identities. *Journal of Symbolic Computation*, 26:187–227, 1998.

[DR95] E. Delaleau and J. Rudolph. An intrinsic characterization of properness for linear time-varying systems. *Journal of Mathematical Systems, Estimation, and Control*, 5:1–18, 1995.

[Fit95] P. Fitzpatrick. On the key equation. *IEEE Transactions on Information Theory*, 41:1290–1302, 1995.

[Fli90] M. Fliess. Some basic structural properties of generalized linear systems. *Systems & Control Letters*, 15:391–396, 1990.

[FO98] S. Fröhler and U. Oberst. Continuous time-varying linear systems. *Systems & Control Letters*, 35:97–110, 1998.

[For75] G. D. Forney, Jr. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal of Control*, 13:493–520, 1975.

[FZ97] F. Fagnani and S. Zampieri. Canonical kernel representations for behaviors over finite abelian groups. *Systems & Control Letters*, 32:271–282, 1997.

[GL00] H. Gluesing-Luerssen. *Linear Delay-Differential Systems with Commensurate Delays: An Algebraic Approach*. Springer, 2000.

[GLH05] G. M. Greuel, V. Levandovskyy, and H.Schönemann. SINGULAR::PLURAL 3.0., A Computer Algebra System for Noncommutative Polynomial Algebras. Centre for Computer Algebra, University of Kaiserslautern, 2005.

[GLS07] H. Gluesing-Luerssen and G. Schneider. State space realizations and monomial equivalence for convolutional codes. *Linear Algebra and its Applications*, 425:518–533, 2007.

[GML] J. I. G. Garcia, J. G. Miranda, and F. J. Lobillo. Elimination orderings and localization in PBW algebras. *Linear Algebra and its Applications*, in print, 2009.

[GPS05] G. M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3.0.4, A Computer Algebra System for Polynomial Computations. Centre for Computer Algebra, University of Kaiserslautern, 2005.

[GTZ98] P. Gianni, B. Trager, and G. Zacharias. Gröbner bases and primary decomposition of polynomial ideals. *Journal of Symbolic Computation*, 6:149–167, 1998.

[GY05] V. P. Gerdt and D. A. Yanovich. Parallel computation of Janet and Gröbner bases over rational numbers. *Programming and Computer Software*, 31, 2005.

[Her05] M. A. I. Hermo. *Varias perspectives sobre las bases de Gröbner: forma normal de Smith, algorithme de Berlekamp y algebras de Leibniz*. PhD thesis, Universidade de Santiago de Compostela, 2005.

[HKC+94] A.R. Hammons, P.V. Kumar, A.R. Calderbank, N.J.A. Sloane, and P. Solé. The Z_4-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Transactions on Information Theory*, 40:301–319, 1994.

[IM05] A. Ilchmann and V. Mehrmann. A behavioral to time-varying linear system. *SIAM Journal of Control Opimization*, 44:1725–1747, 2005.

[INS84] A. Ilchmann, I. Nürnberger, and W. Schmale. Time-varying polynomial matrix systems. *International Journal of Control*, 40:329–362, 1984.

[Jac43] N. Jacobson. *The Theory of Rings*. American Mathematical Society, 1943.

[JW93] R. Johannesson and Z-X. Wan. A linear algebra approach to minimal convolutional encoders. *IEEE Transactions on Information Theory*, 39:1219–1233, 1993.

[JZ99] R. Johannesson and K.S. Zigangirov. *Fundamentals of Convolutional Coding*. IEEE Press Series in Digital and Mobile Communication, 1999.

[Kai80] T. Kailath. *Linear Systems*. Prentice Hall, Englewood Cliffs, New Jersey, 1980.

[KP02] M. Kuijper and J.W. Polderman. Behavioral models for list decoding. *Journal of Mathematical and Computer Modeling of Dynamical Systems*, 8:429–443, 2002.

[KP04] M. Kuijper and J.W. Polderman. Reed-Solomon list decoding from a system theoretic perspective. *IEEE Transactions on Information Theory*, 50:259–271, 2004.

[KP08a] M. Kuijper and R. Pinto. Minimal trellis construction for finite support convolutional ring codes. In A. Barbero, editor, *Coding Theory and Applications (ICMCTA)*. Springer, 2008.

[KP08b] M. Kuijper and R. Pinto. Parametrization of linear recurrence relations by row reduction for sequences over a finite ring. In *Proceedings of the 18th International Symposium on Mathematical Theory of Networks and Systems*, Blacksburg, USA, 2008.

[KP09] M. Kuijper and R. Pinto. On minimality of convolutional ring encoders. *IEEE Transactions on Information Theory*, in print, 2009.

[KPP07] M. Kuijper, R. Pinto, and J. W. Polderman. The predictable degree property and row reducedness for systems over a finite ring. *Linear Algebra and its Applications*, 425:776–796, 2007.

[KPPR06] M. Kuijper, R. Pinto, J. W. Polderman, and P. Rocha. Autonomicity and the absence of free variables for behaviors over finite rings. In *Proceedings of the 7th Portuguese Conference on Automatic Control*, Lisbon, Portugal, 2006.

[Kre93] H. Kredel. *Solvable Polynomial Rings*. Shaker Aachen, 1993.

[KSa] M. Kuijper and K. Schindelar. Gröbner bases and behaviors over finite rings. In Proceedings of the 48th IEEE Conference Decision and Control, Shanghai, China, 2009.

[KSb] M. Kuijper and K. Schindelar. Minimal Gröbner bases and the predictable leading monomial property. Submitted (2009) to Linear Algebra and its Applications.

[Kui01] M. Kuijper. Algorithms for decoding and interpolation. In B. Marcus and J. Rosenthal, editors, *Codes, Systems, and Graphical Models*. Springer-Verlag, 2001.

[KvDHO01] M. Kuijper, M. van Dijk, H. Hollmann, and J. Oostveen. A unifying system-theoretic framework for errors-and-erasures Reed-Solomon decoding. In S. Boztas and I.E. Shparlinski, editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*. Springer, 2001.

[KW97] M. Kuijper and J.C. Willems. On constructing a shortest linear recurrence relation. *IEEE Transactions on Automatic Control*, 42:1554–1558, 1997.

[KWP05] M. Kuijper, X. Wu, and U. Parampalli. Behavioral models over rings — minimal representations and applications to coding and sequences. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, 2005.

[Lev05a] V. Levandovskyy. *Non-commutative Computer Algebra for Polynomial Algebra: Gröbner Bases, Applications and Implementation.* PhD thesis, Technische Universität Kaiserslautern, 2005.

[Lev05b] V. Levandovskyy. On preimages of ideals in certain non–commutative algebras. In G. Pfister, S. Cojocaru, and V. Ufnarovski, editors, *Computational Commutative and Non-Commutative Algebraic Geometry.* IOS Press, 2005.

[Ley04] A. Leykin. Algorithmic proofs of two theorems of Stafford. *Journal of Symbolic Computation*, 38(6):1535–1550, 2004.

[LLO04] P. Lu, M. Liu, and U. Oberst. Linear recurring arrays, linear systems and multidimensional cyclic codes over Quasi-Frobenius rings. *Acta Applicandae Mathematicae*, 80:175–198, 2004.

[LO08] K. Lee and M.E. O'Sullivan. List decoding of Reed-Solomon codes from a Gröbner basis perspective. *Journal of Symbolic Computation*, 43:645–658, 2008.

[LS03] V. Levandovskyy and H. Schönemann. Plural — a computer algebra system for noncommutative polynomial algebras. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, Philadelphia, USA, 2003.

[Lüb02] F. Lübeck. On the computation of elementary divisors of integer matrices. *Journal of Symbolic Computation*, 33(1):57–65, 2002.

[LXB08] Z. Lin, L. Xu, and N. K. Bose. A tutorial on Gröbner bases with applications in signals and systems. *IEEE Transactions on circuits and systems*, 55:445–461, 2008.

[LZ07] V. Levandovskyy and E. Zerz. Obstructions to genericity in study of parametric problems in control theory. *Radon Series on Computational and Applied Mathematics*, 3:191–214, 2007.

[Mal64] B. Malgrange. Systèmes différentiels à coefficients constants. *Séminaire Bourbaki 15*, 264, 1964.

[Mid08] J. Middeke. A polynomial-time algorithm for the Jacobson form for matrices of differential operators. Technical report, RISC Report Series, 2008.

[Mor03] K. Mori. A new parametrization method for all stabilizing controllers of nD systems without coprime factorizabiliy. In *Proceedings of the International Symposium on Circuits and Systems*, Bangkok, Thailand, 2003.

[MR01] J. C. McConnell and J. C. Robson. *Noncommutative Noetherian Rings*. American Mathematical Society, 2001.

[Obe90] U. Oberst. Multidimensional constant linear systems. *Acta Applicandae Mathematicae*, 20:1–175, 1990.

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

[PQ98] J. F. Pommaret and A. Quadrat. Generalized Bézout identity. *Applicable Algebra in Engineering, Communication and Computing*, 9:91–116, 1998.

[PQ99] J. F. Pommaret and A. Quadrat. Algebraic analysis of linear multidimensional control systems. *IMA Journal of Mathematical Control and Information*, 16:275–297, 1999.

[PR05] W. Plesken and D. Robertz. Janet's approach to presentations and resolutions for polynomials and linear PDE's. *Archiv der Mathematik*, 84, 2005.

[PR07] H. Park and G. Regensburger, editors. *Gröbner Bases in Control Theory and Signal Processing*. Walter de Gruyter, 2007.

[Rob06] Daniel Robertz. *Formal Computational Methods for Control Theory*. PhD thesis, RWTH Aachen University, 2006.

[Rot79] J. J. Rotman. *An Introduction to Homological Algebra*. Academic Press, New York, 1979.

[RS99] J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Applicable Algebra in Engineering, Communication and Computing*, 10(1):15–32, 1999.

[RSY96] J. Rosenthal, J. M. Schumacher, and E. V. York. On behaviors and convolutional codes. *IEEE Transactions on Information Theory*, 42:1881–1891, 1996.

[RW01] P. Rocha and J. Wood. Trajectory control and intersection of nd systems. *SIAM Journal of Control Opimization*, 40:107–134, 2001.

[SL] K. Schindelar and V. Levandovskyy. Computing normal forms using Gröbner bases. To appear in Journal of Symbolic Computation.

[SLZ] K. Schindelar, V. Levandovskyy, and E. Zerz. Excat linear modeling using Ore algebras. To appear in Journal of Symbolic Computation.

[SLZ08] K. Schindelar, V. Levandovskyy, and E. Zerz. Linear exact modeling with variable coefficients. In *Proceedings of the 18th International Symposium on Mathematical Theory of Networks and Systems, Blacksburg, USA*, 2008.

[US00] P. Udaya and M. U. Siddiqi. Generalized GMW quadriphase sequences satisfying the Welch bound with equality. *Applicable Algebra in Engineering, Communication and Computing*, 10:203–225, 2000.

[VSR96] V. V. Vazirani, H. Saran, and B. S. Rajan. An efficient algorithm for constructing minimal trellises for codes over finite abelian groups. *IEEE Transactions on Information Theory*, 42:1839–1854, 1996.

[Wed34] J. H. M. Wedderburn. *Lectures on Matrices*. Dover Phoenix, 1934.

[Wil88] J. C. Willems. Models for dynamics. *Dynamics Report*, 2:171–282, 1988.

[Woo00] J. Wood. Modules and behaviours in nd systems theory. *Multidimensional Systems and Signal Processing*, 11:11–48, 2000.

[Woo02] J. Wood. Key problems in the extension of module-behaviour duality. *Linear Algebra and its Applications*, 351-352:761–798, 2002.

[WRO98] J. Wood, E. Rogers, and D. Owens. Minimum lag descriptions and minimal Gröbner bases. *Systems & Control Letters*, 34:289–293, 1998.

[WRO00] J. Wood, E. Rogers, and D. Owens. Controllable and autonomous nd linear systems. *Multidimensional Systems and Signal Processing*, 10:226–241, 2000.

[WW89] K. R. Woodearl and R. B. Warfield. *An Introduction to Noncommutative Noetherian Rings*. Cambridge University Press, 1989.

[WZ99] J. Wood and E. Zerz. Notes on the definition of behavioral controllability. *Systems & Control Letters*, 37:31–37, 1999.

[Zer00] E. Zerz. *Topics in Multidimensional Linear Systems Theory*. Springer, London, 2000.

[Zer05] E. Zerz. Characteristic frequencies, polynomial-exponential trajectories, and linear exact modeling with multidimensional behaviors. *SIAM Journal of Control Opimization*, 44(3):1148–1163, 2005.

[Zer06a] E. Zerz. An algebraic analysis approach to linear time-varying systems. *IMA Journal of Mathematical Control and Information*, 23:113–126, 2006.

[Zer06b] E. Zerz. Recursive computation of the multidimensional MPUM. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 1893–1902, Kyoto, Japan, 2006.

[Zer06c] Eva Zerz. Algebraic System Theory. Lecture at the RWTH Aachen, 2006.

[Zer07a] E. Zerz. Autonomy properties of mutidimensional linear systems over rings. In *Proceedings of the 5th International Workshop on Multidimensional Systems*, Aveiro, Portugal, 2007.

[Zer07b] E. Zerz. Discrete multidimensional systems over $\mathbb{Z}_n$. *Systems & Control Letters*, 56:702–708, 2007.

[Zer07c] E. Zerz. State representations of time-varying linear systems. *Radon Series on Computational and Applied Mathematics*, 3:235–251, 2007.

[Zer08] E. Zerz. The discrete multidimensional MPUM. *Multidimensional Systems and Signal Processing*, 19:307–321, 2008.

[ZL01] E. Zerz and V. Lomadze. A constructive solution to interconnection and decomposition problems with mulidimensional behaviors. *SIAM Journal of Control Opimization*, 40:1072–1086, 2001.

Ich erkläre eidesstattlich, dass ich die Dissertation selbstständig verfasst und alle in Anspruch genommenen Hilfen in der Dissertation angegeben habe.

Kristina Schindelar

Lebenslauf

Kristina Schindelar

03.09.1981	geboren in Bratislava Staatsangehörigkeit: deutsch
1988 - 1982	Katholische Grundschule Düsseldorf
1992 - 1994	Geschwister Scholl Gymnasium Düsseldorf
1994 - 2001	Nelly Sachs Gymnasium Neuss
Okt. 2001 - Sep. 2003	Grundstudium der Mathematik an der Heinrich-Heine-Universität Düsseldorf
Okt. 2003 - Aug. 2006	Hauptstudium der Mathematik an der RWTH Aachen
Aug. 2006	Diplom in Mathematik an der RWTH Aachen
Aug. 2006 - Mai 2010	Doktorandin an der RWTH Aachen