

ALGORITHMS FOR NONCOMMUTATIVE
DIFFERENTIAL OPERATORS

By
Yang Zhang

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO
JULY 26, 2004

© Copyright by Yang Zhang, 2004

UNIVERSITY OF WESTERN ONTARIO
DEPARTMENT OF
APPLIED MATHEMATICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled **“Algorithms for Noncommutative Differential Operators”** by **Yang Zhang** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: July 26, 2004

External Examiner: _____

Research Supervisor: _____

Examining Committee: _____

Table of Contents

Table of Contents	iii
Abstract	v
The Co-Authorship Statement	vii
Acknowledgements	viii
1 Introduction	1
1.1 Skew polynomial rings	1
1.2 Poincaré-Birkhoff-Witt extensions	7
1.3 Non-commutative Riquier Bases	12
Bibliography	14
2 Factoring and Decomposing Ore Polynomials over $\mathbb{F}_q(t)$	20
2.1 Introduction	21
2.2 Canonical skew polynomial rings	24
2.2.1 Reducing to the pure automorphism and derivation cases	24
2.2.2 Automorphism classes and the shift and dilation cases	25
2.2.3 Derivation operators	27
2.2.4 Representation and basic operations with skew polynomials	27
2.3 The eigenring of Ore polynomials over finite fields	28
2.3.1 The centre in the pure automorphism case	29
2.3.2 The centre and minimal central multiples in $\mathbb{F}_q(t)[\mathcal{D}; \delta]$	31
2.3.3 Constructing the eigenring	32
2.3.4 Reducibility and the eigenring	34
2.3.5 Decomposability and the eigenring	36
2.4 Factoring modular Ore polynomials	38

2.5	Computing a complete LCLM decomposition	39
2.6	Conclusion	42
Bibliography		44
3	Non-Commutative Gröbner Bases in Poincaré-Birkhoff-Witt Extensions	48
3.1	Introduction	49
3.2	Poincaré-Birkhoff-Witt Extensions	51
3.3	Gröbner Bases in PBW Extensions	55
3.4	Application to Moving Frames	64
Bibliography		70
4	Non-commutative Riquier Theory in Moving Frames of Differential Operators	73
4.1	Introduction	74
4.2	An Example - the Nonlinear Diffusion Equation	80
4.3	Derivations	82
4.4	Rankings	85
4.5	Reduction	87
4.6	Parametric Derivations, Principal Derivations and Non-commutative Riquier Bases	89
4.7	The Formal Non-commutative Riquier Existence Theorem	91
4.8	Sufficient Finite Sets of Integrability Conditions	96
4.9	Relative Riquier Bases	103
4.10	Analyticity Issues	110
Bibliography		118
5	Conclusion and future work	123
A	Curriculum Vitae	125

Abstract

The aim of this work is to study some noncommutative differential operators. We take an algorithmic approach as well as further developing the mathematics, and design and analyse algorithms to solve fundamental problems. We also give applications to differential equations.

First we consider how to factor skew polynomials. These are polynomials in a differential or difference operator. Using the eigenring method, we present algorithms for computing factorizations and least common left multiple decompositions of skew polynomials over $\mathbb{F}_q(t)$, for a prime power $q = p^\mu$ (here \mathbb{F}_q is the finite field with q elements). Our algorithms are effective in the skew polynomial ring $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ (where $\mathcal{D}t = \sigma(t)\mathcal{D} + \delta(t)$), for any automorphism σ and any σ -derivation δ of $\mathbb{F}_q(t)$. Most importantly, these algorithms are the first to run in time polynomial in the degree of the input

In the second part of this thesis, we presents theory and algorithms for noncommutative Gröbner bases in Poincaré-Birkhoff-Witt extensions. These extension rings generalize the previous domains over which non-commutative Gröbner bases have been applied. Our approach to noncommutative Gröbner bases differs from previous work, which assumes that the coefficients are from a field or commutative ring. This is relevant for computations involving Cartan's theory of moving frames, and we explore these applications.

In the third part of this thesis, we further our study of computations with moving frames, and extend the Rust-Riquier existence and uniqueness theory to analytic

PDEs written in terms of moving frames of non-commuting partial differential operators. The main idea for the theoretical development is to use the commutation relations between the partial differential operators to place them in a standard order. This normalization is exploited to generalize the corresponding steps of the commuting Rust-Riquier Theory to the noncommutative case.

The Co-Authorship Statement

The paper version of Chapter 2 coauthored with Mark Giesbrecht. The paper version of Chapter 3 coauthored with Mark Giesbrecht and Greg Reid. The paper version of Chapter 4 coauthored with Francois Lemaire and Greg Reid.

Acknowledgements

I would like to thank Mark Giesbrecht, my supervisor, for his invaluable advice during all stages of my work and also for his financial support.

I also want to thank my co-supervisor Greg Reid. Our numerous discussions have given me the most interesting insights into analytic methods. Thank you for challenging me for many interesting projects and also financial support.

George Labahn, David Jeffrey and Stephen Watt who kindly agreed to be the official examiners of this thesis deserve my sincerest gratitude. Their critically constructive comments were appreciatively received and treasured

My thanks also goes to other members in ORCCA: Rob Corless, Keith Geddes, Francois Lemaire, Arne Storjohann, and all the other ORCCAians for their many interesting discussions.

London, Ontario

Yang Zhang

June 1, 2004

Chapter 1

Introduction

1.1 Skew polynomial rings

One of the most active and important research areas in noncommutative algebra is the investigation of *skew polynomial rings* (sometimes called Ore extensions or rings). Noether and Schmeidler [32] were the first to consider this kind of ring, and they were later systematically studied by Ore [37] in the 1930's both in the context of differential equations, and as operators on finite fields. Skew polynomial rings were one of the earliest examples in noncommutative algebra. Since then, these rings have been extensively studied, for example, for characterizing various kinds of radicals (Jacobson and Baer) and global and Krüll dimensions of such rings, for constructing finite-dimensional algebras, classifying all valuations of these algebras, etc. One can find a theoretical treatment of this material in many standard books on noncommutative algebras (for example, [10], [11], [21], [24], [25] and [28]).

Over the past ten years, skew polynomials have been successfully applied in many areas, including for example solving Ordinary Differential Equations (ODEs: see

Bronstein and Petkovšek[7], Chyzak and B. Salvy[9], van Hoeij[22] and Singer[44], etc.), control theory (see Chyzak-Quadrat-Robertz[8], Fliess-Mounier[16], Gluesing-Luerssen[20], etc.), and coding theory (see McEliece[29], Piret[39] etc.).

We will begin with an abstract definition of a skew polynomial. For a ring R , let σ be an endomorphism of R . We define a σ -derivation as a linear map $\delta : R \rightarrow R$ such that $\delta(ab) = \sigma(a)\delta(b) + \delta(a)b$. Intuitively, when σ is the identity, δ behaves like a differential operator.

Definition 1.1.1. Given an associative ring R , with σ an endomorphism of R and δ a derivation of R , a *skew polynomial ring* $R[x; \sigma, \delta]$ is a ring generated freely over R by an element x subject to the relation $ax = x\sigma + \delta(a)$ for all $a \in R$.

There are two archetypical cases of skew polynomials.

- (1) If σ is an identity mapping, then $R[x; \sigma, \delta] := R[x; \delta]$ is called a *skew polynomial with derivation type*;
- (2) if $\delta = 0$, then $R[x; \sigma, \delta] := R[x; \sigma]$ is called a *skew polynomial with endomorphism type*;

These cases correspond respectively to some commonly used rings in computer algebra applications:

Example 1.1.1. Let $R = \mathbb{Q}[t]$ be the usual rational polynomial ring over \mathbb{Q} .

- (1) *The differential polynomials:* $\mathbb{Q}[t][x; \sigma_1, \delta] := \mathbb{Q}[t][x; \delta]$, $\sigma(t) = t$, $\delta(t) = 1$.
 $xt = tx + 1$
- (2) *The Shift polynomials:* $\mathbb{Q}[t][x; \sigma, \delta_0] := \mathbb{Q}[t][x; \sigma]$, where $\sigma(t) = t + 1$, $\delta_0(t) = 0$.
 $xt = (t + 1)x$

The noncommutative relations $xt = tx + 1$ and $xt = (t + 1)x$ originally came from the requirements of the composition properties of differential operators and difference operators. For example, let ∂ and $t\partial^2 + 1$ be two differential operators over $\mathbb{Q}[t]$ with the action $\partial(t) = 1$ (this will make ∂ correspond to “usual” differentiation). Suppose we want to write a differential operator

$$\partial(t\partial^2 + 1)$$

in a standard form $\sum_{i=1}^n a_i(t)\partial^i$. In order to do this, we have to “swap” the position of t and ∂ , but since they do not commute, there are some restrictions. One is the composition rule, which requires that

$$(\partial(t\partial^2 + 1)) \circ F(t) = \partial \circ ((t\partial^2 + 1) \circ F(t))$$

where we use \circ for the ring action of the differential operator on the space of differentiable functions in one variable, and, for example, $F(t) = t^2$ is such a differentiable function. Note that ∂ and $t\partial^2 + 1$ correspond to x and $tx^2 + 1$ in the above skew polynomial ring $R[x; \delta]$ respectively, and the rule $xt = tx + 1$ corresponds to $\partial t = t\partial + 1$. We check the composition rule as follows:

$$\begin{aligned} (\partial(t\partial^2 + 1)) \circ F(t) &= \partial \circ ((t\partial^2 + 1) \circ F(t)) \\ &= \partial \circ (((t\partial^2) \circ F(t)) + F(t)) \\ &= \partial \circ (tF''(t) + F(t)) \\ &= F''(t) + tF'''(t) + F'(t). \end{aligned}$$

$$\begin{aligned}
(\partial(t\partial^2 + 1)) \circ F(t) &= (\partial t\partial^2 + \partial) \circ F(t) \\
&= ((t\partial + 1)\partial^2 + \partial) \circ F(t) \quad \text{where we use } \partial t = t\partial + 1 \\
&= (t\partial^3 + \partial^2 + \partial) \circ F(t) \\
&= tF'''(t) + F''(t) + F'(t)
\end{aligned}$$

That is, the noncommutative rule $\partial t = t\partial + 1$ is the appropriate one to make the set of differential operators $\sum_{i=1}^n a_i(t)\partial^i$ become a ring.

Similarly, for two difference operators σ and $t\sigma^2 + 1$ with $\sigma(t) = t + 1$, the rule $\sigma t = (t + 1)\sigma$ transforms $\sigma(t\sigma^2 + 1)$ into a difference operator in a standard form again, that is,

$$\begin{aligned}
(\sigma(t\sigma^2 + 1)) \circ F(t) &= \sigma \circ ((t\sigma^2 + 1) \circ F(t)) \\
&= \sigma \circ (t\sigma^2(F(t)) + F(t)) \\
&= \sigma \circ (tF(t+2) + F(t)) \\
&= \sigma \circ (tF(t+2)) + \sigma \circ F(t) \\
&= (t+1)F(t+3) + F(t+1)
\end{aligned}$$

$$\begin{aligned}
(\sigma(t\sigma^2 + 1)) \circ F(t) &= (\sigma t\sigma^2 + \sigma) \circ F(t) \\
&= ((t+1)\sigma^3 + \sigma) \circ F(t) \quad \text{where we use } \sigma t = (t+1)\sigma \\
&= (t+1)\sigma^3 \circ F(t) + \sigma \circ F(t) \\
&= (t+1)F(t+3) + F(t+1)
\end{aligned}$$

Skew polynomial rings have a number of important structural properties which make them mathematically rich, correspond to real applications, and also allow for effective and efficient algorithms.

One important property of skew polynomials is that $R[x; \sigma, \delta]$ is a principal right ideal domain if σ is injective and R is a division ring. Therefore, the Euclidian algorithms hold in skew polynomial rings.

It is generally much more convenient and computationally efficient to consider only the pure derivation case or pure automorphism case. The following theorem gives one method to transform $R[x; \sigma, \delta]$ into the two archetypical cases above:

Proposition 1.1.1. (*[10]*) *Let $R = K[x; \sigma, \delta]$ be a skew polynomial ring over a field K with endomorphism σ and σ -derivation δ . Then either*

- (1) *δ is inner, and by a suitable choice of x , δ may be taken to be 0;*
- (2) *σ is inner, and by a suitable choice of x , α may be taken to be the identity;*
- (3) *σ leaves the center C of K fixed and δ maps it to 0; in this case C is contained in the center of R .*

We will define this transformation explicitly, and it forms an important part of efficient algorithms, described in later sections of this thesis.

One of the significant differences between the usual polynomial rings and skew polynomial rings is that skew polynomial rings are not unique factorization domains. For example, let $k = \mathbb{C}(X)$ and $L = \partial^2$. Then there are two factorizations $\partial^2 = \partial \circ \partial = (\partial + \frac{f'}{f})(\partial - \frac{f'}{f})$ with f a monic polynomial in z of degree ≤ 1 . If we only consider the degrees, Ore[37] gave the following uniqueness theorem, which can be proven as a consequence of the Jordan-Hölder theorem, see Jacobson[24].

Theorem 1.1.2. (*Ore[37]*). *If $f \in R[x; \sigma, \delta]$ factors completely as*

$$\begin{aligned} f &= f_1 f_2 \cdots f_n \\ &= g_1 g_2 \cdots g_m \end{aligned}$$

where $f_1, \dots, f_n, g_1, \dots, g_m \in R[x; \sigma, \delta]$ are irreducible, then $n = m$ and there exists a permutation ϕ of $\{1, \dots, n\}$ such that for $1 \leq i \leq n$, $\deg(f_i) = \deg(g_{\phi(i)})$.

We refer to Singer and van der Put[45] for more details. Moreover, the usual Gauss lemma does not apply. A indicative example of this is as follows:

Example 1.1.2. Let $R = \bar{C}[t][x; \delta]$, with $\delta(t) = 1$, be a polynomial ring, where \bar{C} is the algebraic closure of a field C . It is easy to check that

$$tx^2 + t^2x - t = (x + t)(tx - 1)$$

Clearly the GCD of the leading coefficients of $x + t$ and $tx - 1$ is 1, but the coefficients of the left hand side can be divided by t . That is, Gauss lemma does not hold!

This unfortunate property makes the study of skew polynomials considerably more difficult than that of the usual polynomials. In particular, factoring algorithms are inherently much more complex.

Since the 1990's skew polynomials have attracted the interest of many computer algebraists. A primary reason is that one can use them to compute with ordinary differential equations. In fact, this was Ore's starting point in the 1930's, but development was not continued, possibly due to the lack of computers at that time. Work on differential factoring algorithms resumed recently, for example, Brostein and Petkovsek [7], Giesbrecht [18, 19], van Hoeij [22, 23] and Singer [44].

As noted earlier, algorithms for factoring and decomposing skew polynomials are very important in computer algebra, and are used for solving systems of differential and difference operators, for example in Maple. They also arise in cryptography, specifically in attacks on the Hidden Field Equation cryptosystem (see [12])

The earliest and most famous method for factoring differential operators goes back to Beke in 1894 [4]. Since then a number of authors have pursued different approaches, and developed a number of distinct algorithms. Some of these have been implemented in well-known mathematical software systems such as Mathematica and Maple. However, none of these previous algorithms run in time polynomial in the input size.

Our goal in Chapter 2 is to give the first polynomial-time factoring and decomposition algorithms, at least in the modular case. As mentioned above, skew polynomial rings are not unique factorization domains and the usual Gauss' lemma does not apply. As well, this ring is non-commutative. Therefore, one cannot apply the now well-refined methods of factoring commutative polynomials to skew polynomials. Our method is purely algebraic. The main idea of our algorithms is to use the properties of an associative algebra known as the *eigenring* related to the polynomial to be factored. Decompositions of this associative algebra will correspond to factorizations and decompositions of the skew polynomial. The eigenring can be used to factor and find least common left multiple decompositions efficiently. Up to now, we have successfully obtained such algorithms for skew polynomials over the function fields of a finite field. These results are included in Chapter 2, and have been published in Proc. ACM Symposium on Symbolic and Algebraic Computation ISSAC2003, pages 127-134. ACM, 2003.

1.2 Poincaré-Birkhoff-Witt extensions

Many different rings with derivation operators have been explored in the mathematical literature, such as (iterated) skew polynomial rings (see, e.g., Chyzak and B.

Salvy[9]), Weyl algebra (see, e.g., Saito-Sturmfels-Takayama[43]), rings of differential operators (see, e.g., Bronstein and Petkovšek[7]), smash products (see, e.g., Bergen-Montgomery[6]), etc. As well, other rings with derivation-like properties, such as enveloping Lie algebras, solvable algebras, and some quantum groups can be analysed with similar techniques. In [5], Bell and Goodearl defined the following so-called Poincaré-Birkhoff-Witt extensions, which provides a unified treatment of these rings.

Definition 1.2.1. Let R and T be two associative rings, with $R \subseteq T$. T is called a *(finite) Poincaré-Birkhoff-Witt extension* of R if there exist $x_1, x_2, \dots, x_n \in T$ such that

- (1) the monomials $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ form a basis for T as a free left R -module, where $i_1, i_2, \dots, i_n \in \mathbb{N}$;
- (2) $x_i r - r x_i := [x_i, r] \in R$ for each $i = 1, \dots, n$ and any $r \in R$;
- (3) $x_i x_j - x_j x_i := [x_i, x_j] \in R + R x_1 + \cdots + R x_n$ for all $i, j = 1, \dots, n$.

We write $T = R\langle x_1, \dots, x_n \rangle$.

Conditions (2) and (3) allow one to swap the positions of elements such that every element in T has a unique representation $\sum r_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ where $r_\alpha \in R$ and $\{\alpha_i\} \in \mathbb{N}$.

Example 1.2.1. *Some typical examples of PBW extensions are as follows:*

- (1) When $x_i r - r x_i = 0$ and $x_i x_j - x_j x_i = 0$ for all $r \in R$ and $i, j \in \{1, \dots, n\}$, the PBW extension $R\langle x_1, \dots, x_n \rangle$ is a usual polynomial ring $R[x_1, \dots, x_n]$ with n -variables $\{x_1, \dots, x_n\}$.
- (2) When $R = K[y_1, \dots, y_n]$, a polynomial ring over a field K , and $x_i k - k x_i = 0$, $x_i y_i - y_i x_i = 1$, $x_i y_j - y_j x_i = 0$ ($i \neq j$), $x_i x_j - x_j x_i = 0$ for all $k \in K, i, j \in \{1, \dots, n\}$.

$\{1, \dots, n\}$ the PBW extension $R\langle x_1, \dots, x_n \rangle$ is so-called n -Weyl algebra. These correspond to partial differential equations.

(3) When R is a field and $x_i r - r x_i = 0$, the PBW extension $R\langle x_1, \dots, x_n \rangle$ is so-called (universal) enveloping Lie algebra of a Lie algebra with dimension n .

Remark 1.2.1. We have to remind the reader of the differences between PBW extensions and other similar algebraic structures. For example, in [26] algebras of solvable type are defined as follows: let $A = k[x_1, \dots, x_n]$ be a finitely generated k -algebra with n generators x_1, \dots, x_n and $\mathcal{M} = \{x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n\}$ be the set of all standard monomials in A . Then A is called an *algebra of solvable type* if \mathcal{M} is a k -basis of A and $x^\alpha x^\beta = q_{\alpha, \beta} x^{\alpha+\beta} + f$, where $x^\alpha, x^\beta \in \mathcal{M}$, $q_{\alpha, \beta} \in k \setminus \{0\}$, $f \in A$ satisfies the leading monomial of f is smaller than that of $x^{\alpha+\beta}$ for some admissible order. We note a number of important differences. For example, algebras of solvable type require that coefficients and indeterminates commute, whereas PBW extensions make no such requirement. The smash product $R \# U(L)$ is not an algebra of solvable type in general, but is a PBW extension (see [21]).

One of the most important properties of PBW extensions is that the associated graded ring of $R\langle x_1, \dots, x_n \rangle$ is a usual polynomial ring $R[x_1, \dots, x_n]$, which gives a simple proof of the following noetherian property:

Theorem 1.2.1. *If R is noetherian, then so is $R\langle x_1, \dots, x_n \rangle$.*

This property, which says that all ascending chains of ideals are finite, is extremely useful for inductive proofs of mathematical properties, and for showing that our algorithms terminate.

It is well-known that commutative Gröbner Bases are a well established technique with many applications, including polynomial solving, constructive approaches to commutative algebra, algebraic geometry, coding theory, statistics, etc. (see, for example, Becker-Weispfenning[3], Eisenbud[13], Pistone-Riccomagno-Wynn[38]).

Recently, noncommutative Gröbner Bases have become a focus of research activity. To our knowledge, noncommutative Gröbner bases were studied as early as Galligo [17] in 1985, where he discussed Gröbner bases in Weyl algebras. A Weyl algebra of dimension n over a field k is the free associative k -algebra $\mathbb{A}_n = k[t_1, \dots, t_n; \partial_1, \dots, \partial_n]$ modulo the commutation rules:

$$t_i t_j = t_j t_i, \quad \partial_i \partial_j = \partial_j \partial_i, \quad \partial_i t_j = t_j \partial_i \quad \text{for } i \neq j, \quad \text{and } \partial_i t_i = t_i \partial_i + 1$$

Later Apel and Lassner [2] in 1988 considered Gröbner bases in enveloping Lie algebras, and Kandri-Rody and Weispfenning [26] considered them in solvable algebras. Since then many papers have been published in this direction. In particular, the algorithms of Chyzak and Salvy [9] have been implemented in Maple.

On the other hand, Mora [30, 31] considered noncommutative Gröbner bases over free algebras. Let $X = \{x_i\}_{i \in \Omega}$ be a set and S be the set of all words in the alphabet X plus the empty word ϕ , i.e., $S = \{x_{i_1} x_{i_2} \cdots x_{i_m} \mid x_{i_j} \in X, m \geq 1\} \cup \{\phi\}$ and the multiplication on S is defined as the concatenation of words. Then the k -space $k\langle X \rangle$ with basis S and above multiplication is called a *free k -algebra* on the set X . Because of the freeness condition on the algebra, the variables are non-commuting among themselves. However, the variables do commute with elements of k . A typical element f of $k\langle X \rangle$ is a polynomial over k in (finite) non-commuting variables of X .

Although these two main streams of research have some intersection, they have developed quite independent approaches. In some sense, the Weyl algebra case attracts

more interest since it can be applied to solving problems with differential equations, and has useful noetherian properties implying Gröbner bases always exist. For free algebras, Gröbner bases may not exist due to the lack of a noetherian property, but this approach is still useful for term rewriting and word problems [31].

Most recently, studying noncommutative differential operators has become an active research area. For example, combining invariant theory and elimination theory, or elimination in moving frames of partial differential operators invariant under an equivalence group (see [14], [15]), requires the use of noncommutative Gröbner bases in PBW extensions.

In Chapter 3, we present a theory and algorithms for noncommutative Gröbner bases in Poincaré-Birkhoff-Witt extensions. These extension rings generalize the previous domains over which non-commutative Gröbner bases have been applied. Our approach to noncommutative Gröbner bases differs from previous work, which assumes that the coefficients are from a field or commutative ring. In applications such as Cartan's method of moving frames, this is not the case. Note that in PBW extensions the coefficients and variables don't commute in general. These two non-commutative properties make PBW extensions more complicated than other rings with derivation types. For example, one of the four cornerstones in classic Gröbner bases theory, the S -pair, does not exist in PBW extensions. We have to calculate a generating set and syzygy set. The theory that we present are applied to moving frames by using a so-called Drach transformation, which transforms a system with n -dependent variables into a new system with only one dependent variable.

The main results in this chapter have been published in the proceedings of 5th Conference on Computer Algebra in Scientific Computation (2002).

1.3 Non-commutative Riquier Bases

Noncommutative Gröbner bases are a powerful tool for working with linear partial differential operators, but are not directly applicable to polynomially nonlinear or analytic systems of partial differential equations. For commutative differential operators, the Cauchy-Kovalevskaya Theorem gives an “Existence and Uniqueness Theorem” for analytic solutions to systems of analytic partial differential equations in a certain form. However, not all systems can be converted to this form. In particular, systems with more equations than dependent variables generally cannot be converted, motivating the need for the Riquier-Janet theory developed in the early part of the twentieth century. This theory can be applied to arbitrary linear (and some non-linear) systems of analytic partial differential equations to give existence and uniqueness theorems for formal and local analytic (real or complex) solutions and is restricted to systems written in terms of commutative differential operators.

Using ideas from Gröbner bases theory, Rust[41] and Rust et *al* [42] gives a Gröbner-style development of Riquier theory and generalized this to the nonlinear case for systems written in terms of commutative differential operators.

The method of moving frames given by Cartan is a powerful theoretical tool for studying the geometric properties of submanifolds and their invariants under the action of a transformation group. However, Cartan’s methods have not been widely applied and are restricted mainly to problems in classical differential geometry. Recently Fels and Olver formulated a new approach to the moving frame theory that can be systematically applied to general transformation groups. The key idea is to formulate a moving frame as an equivariant map to the transformation group. We

refer to Olver [33, 34, 35, 36] for details. Their new approach extends the applicability of Cartan's methods, and is much better suited to the development of symbolic computation algorithms.

Recently, the need to analyse systems of partial differential equations written in terms of non-commutative differential operators has become even more relevant because of potential applications to moving frames. We extend the Rust-Riquier existence and uniqueness theory to analytic PDEs written in terms of non-commuting partial differential operators. The main idea for the theoretical development is to use the commutation relations between the differential operators to place them in a standard order. This normalization is exploited to generalize the corresponding steps of the commuting Rust-Riquier theory to the noncommutative case. These results have been submitted to the journal *Foundations of Computational Mathematics*. They are applied to the problem of group classification of classes of nonlinear diffusion equations and make rigorous the methods developed by Lisle for classification problems.

Bibliography

- [1] W. W. Adams and P. Loustau, *An Introduction to Gröbner Bases*, Amer. Math. Soc. 1994.
- [2] J. Apel and W. Lassner, *An extension of Buchberger's Algorithm and calculations in enveloping fields of Lie algebras*, J. Symb. Comp., 6(1988), 361-370.
- [3] T. Becker and V. Weispfenning, *Gröbner bases*, Springer, 1993.
- [4] E. Beke. Die irreducibilität der homogenen linearen differentialgleichungen. *Math. Annalen*, 45:278–300, 1894.
- [5] A. D. Bell and K. R. Goodearl, *Uniform rank over differential operator rings and Poincaré-Birkhoff-Witt extensions*, Pacific Journal of Mathematics, vol.131(1)(1988), 13-37.
- [6] J. Bergen and S. Montgomery. Smash products and outer derivations. *Israel Journal of Mathematics*, 53(1):321–345, 1986.
- [7] M. Bronstein and M. Petkovšek. An introduction to pseudo-linear algebra. *Theoretical Computer Science*, 157:3–33, 1996.

- [8] F. Chyzak, A. Quadrat and D. Robertz, *Linear control systems over Ore algebras: Effective algorithms for computation of parametrizations*, preprint.
- [9] F. Chyzak and B. Salvy, *Non-commutative elimination in Ore algebras proves multivariate identities*, J. Symb. Comp. 26(1998), 187-227.
- [10] P. Cohn. *Free Rings and their Relations*. Academic Press, London, 1985.
- [11] P. Cohn. *Skew Fields: Theory of General Division Rings*. Cambridge, London, 1995.
- [12] R. Coulter, G. Havas and M. Henderson, *Giesbrecht's algorithm, the HFE cryptosystem and Ore's ps-polynomials*, Computer Mathematics, Proceedings of the Fifth Asian Symposium (ASCM 2001), Lecture Notes Series on Computing 8, World Scientific (2001) 36-45
- [13] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer, 1995.
- [14] M. Fels and P.J. Olver, *Moving Coframes. I. A practical algorithm*, Acta. Appl. Math. **51** (1998), 161–213.
- [15] M. Fels and P.J. Olver, *Moving Coframes. II. Regularization and theoretical foundations*, Acta. Appl. Math. **55** (1999), 127–208.
- [16] M. Fliess and H. Mounier, *Controllability and observability of linear delay systems: an algebraic approach*, ESAIM COCV, vol. 3(1998), pp.301-314.
- [17] A. Galligo, *Some algorithmic questions on ideals of differential operators*, Proc. EUROCAL'85, Springer LNCS 204, 413-421.

- [18] M. Giesbrecht. Factoring in skew-polynomial rings. In *Proceedings of Latin American Theoretical INformatics Conference (LATIN)*, pages 191–203, Sao Paulo, Brasil, 1992.
- [19] M. Giesbrecht. Factoring in skew-polynomial rings over finite fields. *J. of Symbolic Computation*, 24(5):463–486, 1998.
- [20] H. Gluesing-Luerssen, *Linear Delay-Differential Systems with Commensurate Delays: An Algebraic Approach*, Lecture Notes in Mathematics 1770, Springer, 2002.
- [21] K. R. Goodearl and R. B. Warfield, *An Introduction to Non-Commutative Noetherian Rings*, vol. 16, London Math. Soc. Student Texts, Cambridge University press, Cambridge, 1989.
- [22] M. van Hoeij. Rational solutions of the mixed differential equation and its application to factorization of differential operators. In *Proc. ISSAC'96*, pages 219–225, 1996.
- [23] M. van Hoeij. Factorization of differential operators with rational functions coefficients. *J. Symb. Comp.*, 24: 537–561, 1997.
- [24] N. Jacobson. *The Theory of Rings*. American Math. Soc., New York, 1943.
- [25] N. Jacobson. *Finite-Dimensional Division Algebras over Fields*. Springer-Verlag, 1996.
- [26] A. Kandri-Rody and V. Weispfenning, *Non-commutative Gröbner bases in algebras of solvable type*, *J. Symb. Comp.* vol. 9(1990), 1-26.

- [27] E. R. Kolchin, *Differential algebraic groups*, vol. 114 of Pure and Applied Mathematics. Academic Press Inc., 1985.
- [28] J. C. McConnell and J. C. Robson, *Non-commutative Noetherian Rings*, Wiley 1987.
- [29] R. J. McEliece, *The algebraic theory of convolutional codes*, In V. Pless and W. Huffman, edits, *Handbook of Coding Theory*, Vol 1, pages 1065-1138, Elsevier, Amsterdam, 1998.
- [30] T. Mora, *Gröbner basis for noncommutative polynomial rings*, Proc. AAEECC3, Lecture Notes in Computer Science 229(1986).
- [31] T. Mora, *An introduction to commutative and noncommutative Gröbner bases*, Theor. Comp. Sci., 134: 131-173, 1994.
- [32] E. Noether and W. Schmeidler, *Moduln in nichtkommutativen Bereichen*, insbesondere aus Differential-und Differenzanusdrücken. Mathematische Zeitschrift, 1-35, 1920.
- [33] P.J. Olver, *Applications of Lie Groups to Differential Equations*, Second Edition, Graduate Texts in Mathematics **107**, Springer-Verlag, New York, 1993.
- [34] P.J. Olver, *Equivalence, Invariants, and Symmetry*, Cambridge University Press, 1995.
- [35] P.J. Olver, *Geometric foundations of numerical algorithms and symmetry*, Appl. Alg. Engin. Comput. **11** (2001), 417–436.

- [36] P.J. Olver, *Moving Frames*, pre-print, University of Minnesota (see <http://www.math.umn.edu/~olver/xtra.html>).
- [37] O. Ore, *Theory of non-commutative polynomials*, Annals of Mathematics, 34:480-508, 1933.
- [38] Pistone, G., Riccomagno, E. and Wynn, H. P. *Algebraic Statistics: Computational Commutative Algebra in Statistics*. Chapman & Hall/CRC, 2001.
- [39] P. Piret, *Structure and constructions of cyclic convolutional codes*, IEEE Trans. Inform. Theory, 22:147-155, 1976.
- [40] J. F. Differential Algebra, AMS 1950.
- [41] C. J. Rust, *Rankings on derivatives for elimination algorithms and formal solvability of analytic partial differential equations*, Ph.D. thesis, University of Chicago, 1998.
- [42] C.J. Rust, G.J. Reid and A.D. Wittkopf, *Existence and uniqueness theorems for formal power series solutions of analytic differential systems*, Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation, 105–112 (electronic), ACM, New York, 1999.
- [43] M. Saito, B. Sturmfels and N. Takayama, *GröbnerDeformations of Hypergeometric Differential Equations*, Springer, 2000.
- [44] M. F. Singer. Testing reducibility of linear differential operators: A group theoretic perspective. *Applicable Algebra in Engineering, Communication and Computing*, 7 (2): 77–104, 1996.

- [45] M. Singer and van der Put, *Galois Theory of Linear Differential Equations*, Grundlehren der mathematischen Wissenschaften, Volume 328, Springer, 2003

Chapter 2

Factoring and Decomposing Ore Polynomials over $\mathbb{F}_q(t)$

We present algorithms for computing factorizations and least common left multiple (LCLM) decompositions of Ore polynomials over $\mathbb{F}_q(t)$, for a prime power $q = p^\mu$. Our algorithms are effective in $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, for any automorphism σ and σ -derivation δ of $\mathbb{F}_q(t)$. On input $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, the algorithms run in time polynomial in $\deg_{\mathcal{D}}(f)$, $\deg_t(f)$, p and μ .

The earliest and most famous method for factoring differential operators goes back to [2]. Many factorization algorithms are based on Beke's algorithm for computing first order factors and then computing higher order factors by using the exterior power method. In many cases these methods are quite expensive in practice.

Recently, a number of authors have pursued different approaches. By considering the exponential parts of linear differential operators, van Hoeij [13] gives a new efficient factorization algorithm. Since the existence of hyper-exponential solutions is equivalent to the existence of right factors of degree 1, Bronstein and Petkovšek [5] describes an algorithm that reduces the problem of factoring in Ore polynomial rings to finding all the irreducible right factors of degree 1. Singer [29] gives a method to

decide if a linear differential operator is reducible without having to find a factor. He uses the fact that each differential operator L is associated to a linear algebraic group G , its Galois group, and the reductive property of G decides if L is reductive. He shows that factorization can be reduced to solving a so-called mixed equation in many cases. Van Hoeij [12] provides an efficient method to compute the solutions of this equation.

Most closely related to our work here, van der Put [25, 26, 27] gives procedures for factoring differential operators over $\mathbb{Q}(t)$ and $\mathbb{F}_p(t)$ by considering the so-called p -curvature. In principle these techniques can be generalized to the case of difference operators: see Singer and van der Put [28], Sections 5.1 and 5.2. Very recently, Cluzeau [6] presents algorithms for factoring differential systems with coefficients in $\mathbb{F}_p(t)$.

2.1 Introduction

The Ore polynomials $F(t)[\mathcal{D}; \sigma, \delta]$, over a field of rational function $F(t)$, are the polynomials in $F(t)[\mathcal{D}]$ under the usual addition and (generally non-commutative) multiplication such that

$$\mathcal{D}a(t) = \sigma(a(t))\mathcal{D} + \delta(a(t)) \quad \text{for any } a(t) \in F(t).$$

Here σ is any automorphism of $F(t)$ and δ is a σ -derivation, that is, δ is an F -linear map such that for $a, b \in F(t)$, $\delta(ab) = \sigma(a)\delta(b) + \delta(a)b$. This ring has been well studied mathematically at least since [23]; we draw heavily on the excellent expositions in [7, 8], as well as [16, 17] in this paper. $F(t)[\mathcal{D}; \sigma, \delta]$ is a principal left (right) ideal domain, and hence admits unique monic least common left (right) multiples (LCLMs)

and greatest common right (left) divisors (GCRDs). Efficient algorithms for basic operations and an introduction to the computational theory are given in [4].

We present algorithms for the following two problems in Ore polynomial domains over $\mathbb{F}_q(t)$, the field of rational functions over the finite field \mathbb{F}_q with $q = p^\mu$ elements where p is prime. Given $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$:

- (1) **Factorization:** find $g, h \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta] \setminus \mathbb{F}_q(t)$ such that $f = gh$, or certify that f is irreducible in $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ (there is no such factorization).
- (2) **LCLM-decomposition:** find $g, h \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ of positive degree in \mathcal{D} such that $f = \text{lcm}(g, h)$ and $\text{gcd}(g, h) = 1$, or certify that f is indecomposable in $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ (there is no such decomposition).

The number of steps required by our algorithms is polynomial in $\deg_{\mathcal{D}}(f)$, $\deg_t(f)$, p and μ .

We consider a general class of Ore polynomials. Over $\mathbb{F}_q(t)$, we let σ be any automorphism fixing \mathbb{F}_q , so $\sigma(t) = (\sigma_1 t + \sigma_2)/(\sigma_3 t + \sigma_4)$, for $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \mathbb{F}_q$ such that $\sigma_1 \sigma_4 - \sigma_2 \sigma_3 \neq 0$. The σ -derivation δ is arbitrary, and can be specified completely by $\delta(t) \in \mathbb{F}_q(t)$. Standard canonicalization to the pure shift, pure dilation, and pure derivation cases will be presented in Section 2.2. As well, we summarize the costs and coefficient bounds on basic operations in Section 2.2.

The main idea of our algorithms is that the *eigenring* can be used to factor and LCLM-decompose any $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$. A similar idea is employed in [10, 11] to give efficient algorithms for factoring and LCLM-decomposing Ore polynomials over $\mathbb{F}_q[\mathcal{D}; \tau]$, where τ is a Frobenius automorphism of \mathbb{F}_q . Indeed, this is one of the original settings for Ore polynomials explored by [22, 24]. Properties of the eigenring, and an efficient algorithm to compute it, are presented in Section 2.3. For notational

convenience we will generally write $\mathfrak{S} := \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$. Following [21, 23, 24], but essentially using the language of [7], we define the *idealizer* of $\mathfrak{S}f$ to be $I(\mathfrak{S}f) = \{u \in \mathfrak{S} \mid fu \in \mathfrak{S}f\}$. $I(\mathfrak{S}f)$ is the largest subalgebra of \mathfrak{S} in which $\mathfrak{S}f$ is a two-sided ideal. The *eigenring* $E(\mathfrak{S}f)$ of $\mathfrak{S}f$ is defined as the quotient $E(\mathfrak{S}f) = I(\mathfrak{S}f)/\mathfrak{S}f$. The eigenring is an associative algebra over the field of constants $\mathbb{K} = \{a(t) \in \mathbb{F}_q(t) : a(t)\mathcal{D} = \mathcal{D}a(t)\}$.

In Section 2.3, we develop the central theory behind our algorithms. The first key point, which has been used in the differential case by [26], is that $\mathbb{F}_q(t)$ is a *finite algebraic extension* of \mathbb{K} . In particular, $\mathbb{K} \cong \mathbb{F}_q(T)$, for an indeterminate T . We prove here that this is the case for all rings of Ore polynomials over $\mathbb{F}_q(t)$. (This is in contrast to the case over $\mathbb{F}(t)$, for a field \mathbb{F} of characteristic zero, where the field of constants is \mathbb{F} .) Over $\mathbb{F}_q(t)$, the eigenring $E(\mathfrak{S}f)$ is isomorphic to a finite dimensional associative algebra (of relatively small dimension) over \mathbb{K} .

The second key point is that every non-trivial zero-divisor in $E(\mathfrak{S}f)$ yields a non-trivial factorization of f , and f is irreducible if and only if $E(\mathfrak{S}f)$ is a division algebra. For LCLM-decompositions, we show correspondingly that any pair of orthogonal idempotents in $E(\mathfrak{S}f)$ yield a non-trivial LCLM-decomposition of f , and f is LCLM-decomposable if and only if $E(\mathfrak{S}f)$ possesses no such orthogonal idempotents.

Finally, we note that there are efficient algorithms for the problem of finding zero divisors and idempotents in finite dimensional associative algebras over $\mathbb{K} = \mathbb{F}_q(T)$. [15] provides an efficient algorithm for computing the Jacobson radical and primitive orthogonal idempotents in the semi-simple part. We extend this in a straightforward way to produce orthogonal idempotents in the algebra itself (should they exist). That we can demonstrate our algorithms to require polynomial time relies on the fact that

the algorithm of [15] requires time polynomial in the dimension of the input associative algebra and the degree (in T) of the structure constants.

In Section 2.4, we employ this correspondence between zero divisors in the eigenring and factorizations to split reducible polynomials. Similarly, in Section 2.5, we use the correspondence between orthogonal idempotents in the eigenring and LCLM-decompositions to compute LCLM-decompositions.

While we do not give the explicit exponents, the dominant cost is the decomposition of the eigenring, and this requires time about $O((\deg_{\mathcal{D}}(f) + \deg_t(f) + p + \mu)^6)$. Note that the algorithm runs in time polynomial in p , not $\log p$, reflecting the fact the dimension of the eigenring is polynomial in p .

2.2 Canonical skew polynomial rings

While Ore's skew polynomials, with both an automorphism and a derivation, appear quite general, there are in fact only a small number of representative cases. In this section we briefly present this well-known reduction. Throughout this section we consider the ring $\mathfrak{S} := \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, where σ is an automorphism of $\mathbb{F}_q(t)$ fixing \mathbb{F}_q and δ is a σ -derivation of $\mathbb{F}_q(t)$. While we state the results over $\mathbb{F}_q(t)$, much of what we say will hold over $\mathbb{F}(t)$ for any perfect field \mathbb{F} .

2.2.1 Reducing to the pure automorphism and derivation cases

It is well known that if $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ has both a non-trivial automorphism σ and non-trivial σ -derivation δ , then, after a change of variables, \mathfrak{S} is isomorphic to a ring

$\mathfrak{S}' = \mathbb{F}_q(t)[\mathcal{D}'; \sigma']$ with only an automorphism (i.e., whose derivation is identically zero) by means of the substitution:

$$\mathcal{D} \rightarrow \frac{\mathcal{D} + \delta(t)}{t - \sigma(t)},$$

See [7], Proposition 3.1, page 498. This change of variables is computationally efficient. Thus, we need only consider the pure automorphism case $\mathbb{F}_q(t)[\mathcal{D}; \sigma]$ and the pure derivation case $\mathbb{F}_q(t)[\mathcal{D}; \delta]$.

2.2.2 Automorphism classes and the shift and dilation cases

In this subsection we assume that $\mathfrak{S} := \mathbb{F}_q(t)[\mathcal{D}; \sigma]$. We show that any such ring is isomorphic to a ring $\mathbb{F}_q(t)[\overline{\mathcal{D}}; \overline{\sigma}]$ of difference operators (i.e., where $\overline{\sigma}(t) = t + \gamma$ for $\gamma \in \mathbb{F}_q$), or a ring of dilation operators (i.e., where $\overline{\sigma}(t) = \xi t$ for some $\xi \in \mathbb{F}_{q^2}^*$) over a quadratic field extension. This classification is straightforward (and well known), and the transformation is computationally efficient.

Every automorphism of $\mathbb{F}_q(t)$ which fixes \mathbb{F}_q has the property that

$$\sigma(t) = \frac{\sigma_1 t + \sigma_2}{\sigma_3 t + \sigma_4} \quad \text{where} \quad \det \begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_3 & \sigma_4 \end{pmatrix} \neq 0.$$

The automorphisms form a group under composition, and it is easily proven that $\text{Aut}(\mathbb{F}_q(t)) \cong \text{PGL}(2, \mathbb{F}_q)$, the projective general linear group of invertible 2×2 matrices over \mathbb{F}_q modulo scalar multiples of the identity. In particular, there is an isomorphism

$$\text{Aut}(\mathbb{F}_q(t)) \rightarrow \text{PGL}(2, \mathbb{F}_q), \quad \frac{\sigma_1 t + \sigma_2}{\sigma_3 t + \sigma_4} \mapsto \begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_3 & \sigma_4 \end{pmatrix}.$$

Since every matrix is similar to a matrix in Jordan form, every $s \in \mathbb{F}_q^{2 \times 2}$ satisfies

either

$$\text{Case (1) } \exists u \in \mathbb{F}_{q^2}^{2 \times 2}, \text{ such that } usu^{-1} = \begin{pmatrix} \alpha^q & 0 \\ 0 & \alpha \end{pmatrix},$$

for $\alpha \in \mathbb{F}_{q^2}$;

$$\text{Case (2) } \exists u \in \mathbb{F}_q^{2 \times 2} \text{ such that } usu^{-1} = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix},$$

for $\alpha \in \mathbb{F}_q$.

Note that in Case (1), s generally has distinct eigenvalues and hence a generally irreducible minimal polynomial over \mathbb{F}_q . Thus the eigenvalues, and transformation matrices to the Jordan form, lie in a quadratic extension \mathbb{F}_{q^2} of \mathbb{F}_q . In Case (2) s has repeated eigenvalues, which must lie in \mathbb{F}_q .

This notion of normal form is easily extended to the skew polynomial ring itself. Suppose σ is represented in $\text{PGL}(2, \mathbb{F}_q)$ by $s \in \mathbb{F}_q^{2 \times 2}$ and usu^{-1} is in Jordan form as above. Let τ be the fractional linear transformation corresponding to u . Then

$$\bar{\sigma} := \tau \circ \sigma \circ \tau^{-1} = \begin{cases} \xi t, & \text{for } \xi = \alpha^{q-1} \in \mathbb{F}_{q^2}^*, \text{ a } \textit{dilation}, \text{ or} \\ t + \gamma, & \text{for } \gamma = \alpha^{-1} \in \mathbb{F}_q, \text{ a } \textit{shift}. \end{cases}$$

The map τ , which is itself an automorphism of $\mathbb{F}_q(t)$, is naturally extended to $\mathbb{F}_q(t)[\mathcal{D}; \sigma]$, by $\tau(\mathcal{D}) = \bar{\mathcal{D}}$, whence $\mathbb{F}_q(t)[\mathcal{D}; \sigma] \cong \mathbb{F}_q(t)[\bar{\mathcal{D}}; \bar{\sigma}]$.

To factor an $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma]$ we may thus factor the polynomial $\tau(f)$. This ring isomorphism is efficiently computable, and thus we assume from now on that σ is either a shift or dilation over $\mathbb{F}_q(t)$, where q is redefined as appropriate.

2.2.3 Derivation operators

It is easily observed that the derivation operator, in the pure derivation ring $\mathbb{F}_q(t)[\mathcal{D}; \delta]$, satisfies the standard algebraic properties of differentiation. In particular, δ is \mathbb{F}_q -linear and $\delta(t^n) = nt^{n-1}\delta(t)$. We can specify the derivation operator completely by specifying the value of $\delta(t) \in \mathbb{F}_q(t)$, and for any rational function $r(t) \in \mathbb{F}_q(t)$, we have $\delta(r(t)) = r'(t)\delta(t)$, where $r'(t) \in \mathbb{F}_q(t)$ is the usual first derivative of r with respect to t . For simplicity, we will assume that $\delta(t)$ is of constant degree, and do not consider its degree explicitly in our analysis.

2.2.4 Representation and basic operations with skew polynomials

To standardize our representation of $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, we write

$$f = \sum_{0 \leq i \leq n} a_i(t)\mathcal{D}^i,$$

where the $a_0(t), \dots, a_n(t) \in \mathbb{F}_q(t)$ are always written to the left of the power of \mathcal{D} . Let $c \in \mathbb{F}_q[t]$ be the LCM of the denominators of coefficients of f . It is obvious that $c \cdot f \in \mathbb{F}_q[t][\mathcal{D}; \sigma, \delta]$, and if $c \cdot f = f_1 f_2 \cdots f_k$, for $f_1, \dots, f_k \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, then $f = (c^{-1} \cdot f_1) f_2 \cdots f_k$. There is no necessity that polynomials in $\mathbb{F}_q[t][\mathcal{D}; \sigma, \delta]$ factor over $\mathbb{F}_q[t][\mathcal{D}; \sigma, \delta]$, and indeed there may be reducible polynomials in $\mathbb{F}_q[t][\mathcal{D}; \sigma, \delta]$ such that every complete factorization involves at least one factor in $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta] \setminus \mathbb{F}_q[t][\mathcal{D}; \sigma, \delta]$. We may, however, assume that our input comes from $\mathbb{F}_q[t][\mathcal{D}; \sigma]$ or $\mathbb{F}_q[t][\mathcal{D}; \delta]$ and our output is in $\mathbb{F}_q(t)[\mathcal{D}; \sigma]$ or $\mathbb{F}_q(t)[\mathcal{D}; \delta]$ respectively.

Basic operations with skew polynomials are performed with the polynomials in

standard representation. These basic operations are addition, subtraction, multiplication, division with remainder, greatest common right (left) divisor (GCRD), and least common left multiple (LCLM). Many good algorithms have been developed for these operations (see, e.g., [4], [18], [14]) We note the following crude bounds on the costs of these algorithms and on the size of the output (see [19]).

Theorem 2.2.1. *Let $f, g \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, and $h_1 = f + g$, $h_2 = fg$, $h_3 = \text{rem}(f, g)$, $h_4 = \text{quo}(f, g)$ (i.e., $f = \text{quo}(f, g)g + \text{rem}(f, g)$ for the unique $\text{quo}(f, g)$, and $\text{rem}(f, g) \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ such that $\deg_{\mathcal{D}}(\text{rem}(f, g)) < \deg_{\mathcal{D}}(g)$), $h_5 = \text{lclm}(f, g)$, and $h_6 = \text{gcd}(f, g)$). Then bounds on the degrees in \mathcal{D} for h_1, \dots, h_6 are as in the usual polynomial case, while*

$$\deg_t(h_i) = O((\deg_{\mathcal{D}}(f) + \deg_{\mathcal{D}}(g))(\deg_t(f) + \deg_t(g)))$$

for $1 \leq i \leq 6$. The cost of computing h_1, \dots, h_6 is bounded by

$$O((\deg_{\mathcal{D}}(f) + \deg_{\mathcal{D}}(g))^5(\deg_t(f) + \deg_t(g))^2)$$

operations in \mathbb{F}_q .

2.3 The eigenring of Ore polynomials over finite fields

In this section we describe the centre of the ring of Ore polynomials $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, where $\mathfrak{S} := \mathbb{F}_q(t)[\mathcal{D}; \sigma]$ for an automorphism σ , or $\mathfrak{S} := \mathbb{F}_q(t)[\mathcal{D}; \delta]$ for a derivation δ . In each case the centre turns out to be the usual polynomial ring $\mathbb{F}_q(T)[Y]$, for independent indeterminates T and Y . Hence the centre is a unique factorization

domain. We then observe that every polynomial in \mathfrak{S} has a non-trivial left multiple in the centre, the *minimal central multiple*. We then describe the eigenring of a polynomial and give its properties. In particular, how zero-divisors in the eigenring correspond to factors of the original polynomial. We also show how to construct the eigenring efficiently.

An element $f \in \mathfrak{S}$ is *invariant* if $\mathfrak{S}f = f\mathfrak{S}$. An invariant element f^* such that $\mathfrak{S}f \supseteq \mathfrak{S}f^*$ is called a *bound* for f , and f is said to be *bounded*, if there exists a non-trivial bound. Closely related to the invariant elements in \mathfrak{S} is the centre \mathfrak{C} of \mathfrak{S} , those elements which commute with every element of \mathfrak{S} . Every element $f \in \mathfrak{S}$ has a minimal central left (and right) multiple $\hat{f} \in \mathfrak{C} \setminus \{0\}$, the (monic) polynomial of lowest degree in the centre which is a left (right) multiple of f . [16], Chapter 3, Theorem 11 shows that the central left and right multiples are in fact equal. Thus, we refer to \hat{f} as simply the *minimal central multiple* of f .

2.3.1 The centre in the pure automorphism case

We first consider the shift and dilation cases. To describe the centre of \mathfrak{S} , we must first understand the constant field \mathbf{K} of σ , whose elements are fixed by the automorphism. Following [22], we say a $\varphi \in \mathbb{F}_q[t]$ is *additive* if $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$ for any $\alpha, \beta \in \overline{\mathbb{F}_q}$, where $\overline{\mathbb{F}_q}$ is the algebraic closure of \mathbb{F}_q . The additive polynomials in $\mathbb{F}_q[t]$, where $q = p^\mu$, are exactly those of the form $\varphi = \sum_{0 \leq i \leq \mu} \varphi_i t^{p^i} \in \mathbb{F}_q[t]$.

Theorem 2.3.1. *Let \mathbf{K} be the constant field of σ .*

- *If $\sigma(t) = t + \gamma$ for some $\gamma \in \mathbb{F}_q^*$ then $\mathbf{K} = \mathbb{F}_q(\varphi(t))$, where $\varphi \in \mathbb{F}_q[t]$ is the additive polynomial of smallest degree such that $\varphi(\gamma) = 0$.*

- If $\sigma(t) = \xi t$ for some $\xi \in \mathbb{F}_q^*$ then $\mathbf{K} = \mathbb{F}_q(t^\nu)$, where ν is the multiplicative order of ξ .

Proof. First observe that $\sigma(\varphi(t)) = \varphi(t+\gamma) = \varphi(t) + \varphi(\gamma) = \varphi(t)$, so $\varphi(t)$ is invariant under σ , as is $t^q - t$. Thus $\mathbb{F}_q(t^q - t) \subseteq \mathbf{K} \subseteq \mathbb{F}_q(t)$. By Luröth's theorem $\mathbf{K} = \mathbb{F}_q(v(t))$ for some $v \in \mathbb{F}_q[t]$, and there exists a $u \in \mathbb{F}_q[t]$ such that $t^q - t = u(v(t))$. Since $t^q - t$ is additive, by Theorem 3.3 of [9] we find u, v are also additive. Letting φ be the additive polynomial of minimal degree with root γ ensures that $\mathbf{K} = \mathbb{F}_q(\varphi(t))$.

In the dilation case, $h(t) = \sum_{0 \leq i \leq m} h_i t^i \in \mathbb{F}_q[t]$ is fixed by σ when $\sigma(h(t)) = h(\xi t)$. This is true only if for all i , $h_i = 0$ or $\xi^i = 1$, i.e., when $h \in \mathbb{F}_q[t^\nu]$. By [20], Page 71, the only rational functions which are fixed by σ are quotients of polynomials in \mathbf{K} , and hence $\mathbf{K} = \mathbb{F}_q(t^\nu)$ □

For consistency, we will let $\nu := \deg \varphi$ in the shift case. Thus, in both the shift and the dilation case we can write the ground field as an algebraic extension of degree ν over the constant field \mathbf{K} of σ . It should be noted that in all cases it is straightforward to compute the constant field.

The centre of \mathfrak{S} is characterized by the following:

Theorem 2.3.2. *The centre \mathfrak{C} of a pure automorphism Ore polynomial ring is characterized as follows.*

- *Usual:* When $\sigma(t) = t$, $\mathfrak{C} = \mathbb{F}_q(T)[Y]$ where $T = t$ and $Y = \mathcal{D}$;
- *Shift:* When $\sigma(t) = t + \gamma$ for $\gamma \in \mathbb{F}_q^*$, then $\mathfrak{C} = \mathbb{F}_q(T)[Y]$, where $T = \varphi(t)$ as in Theorem 2.3.1, and $Y = \mathcal{D}^\nu$;
- *Dilation:* When $\sigma(t) = \xi t$ is a dilation, then $\mathfrak{C} = \mathbb{F}_q(T)[Y]$ where $T = t^\nu$ and $Y = \mathcal{D}^\nu$, and σ has multiplicative order ν .

Thus in all cases \mathfrak{C} is a usual, commutative polynomial ring.

Proof. For $f = \sum_{0 \leq i \leq n} a_i(t) \mathcal{D}^i \in \mathfrak{S}$ to be in the centre, $t \cdot f = f \cdot t$ or

$$\sum_{0 \leq i \leq n} a_i(t) t \cdot \mathcal{D}^i = \sum_{0 \leq i \leq n} a_i(t) \mathcal{D}^i \cdot t = \sum_{0 \leq i \leq n} a_i(t) \cdot \sigma^i(t) \cdot \mathcal{D}^i,$$

whence either $a_i(t) = 0$ or $\sigma^i(t) = t$. In other words $\mathfrak{C} \subseteq \mathbb{F}_q(t)[\mathcal{D}^\nu; \sigma]$.

We note also that $\mathcal{D} \cdot f = f \cdot \mathcal{D}$ or

$$\mathcal{D} \cdot \sum_{0 \leq i \leq n} a_i(t) \mathcal{D}^i = \sum_{0 \leq i \leq n} \sigma(a_i(t)) \mathcal{D}^{i+1} = \sum_{0 \leq i \leq n} a_i(t) \cdot \mathcal{D}^{i+1}.$$

Thus, $\sigma(a_i(t)) = a_i(t)$, or in other words, $a_i(t) \in \mathbb{K}$ for $0 \leq i \leq n$.

Combining these two conditions yields the centre, as specified in the theorem. \square

There is a close relationship between the centre, the minimal central multiples and the invariant elements. In the automorphism case $\mathbb{F}_q(t)[\mathcal{D}; \sigma]$, [17], Theorem 1.1.22, shows every invariant element f^* has the form $f^* = a(t) \mathcal{D}^k \hat{f}$ for $\hat{f} \in \mathfrak{C}$, $a(t) \in \mathbb{F}_q(t)$ and $k \in \mathbb{Z}_{\geq 0}$.

2.3.2 The centre and minimal central multiples in $\mathbb{F}_q(t)[\mathcal{D}; \delta]$

In the derivation case, the constant subfield of $\mathbb{F}_q(t)$ in $\mathbb{F}_q(t)[\mathcal{D}; \delta]$ is $\mathbb{K} = \{a(t) \in \mathbb{F}_q(t) : \delta(a(t)) = 0\}$.

Lemma 2.3.3. *The constant subfield of $\mathbb{F}_q(t)[\mathcal{D}; \delta]$ is $\mathbb{F}_q(t^p)$.*

Proof. We first consider the kernel of δ in the polynomial ring $\mathbb{F}_q[t]$. Suppose $a(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_0 \in \mathbb{F}_q[t]$, and $\delta(a(t)) = 0$. Then

$$\begin{aligned} 0 &= \delta(a(t)) \\ &= n a_n t^{n-1} \delta(t) + (n-1) a_{n-1} t^{n-2} \delta(t) + \cdots + a_1 \delta(t) \\ &= (n a_n t^{n-1} + (n-1) a_{n-1} t^{n-2} + \cdots + a_1) \delta(t). \end{aligned} \tag{2.3.1}$$

Since $\delta(t) \neq 0$, it follows that

$$a(t) = a_{mp}(t^p)^m + a_{(m-1)p}(t^p)^{m-1} + \cdots + a_p t^p + a_0 \in \mathbb{F}_q[t^p],$$

where $m = n/p \in \mathbb{N}$, or in other words, $a(t) \in \mathbb{F}_q[t^p]$.

To identify the kernel of δ in $\mathbb{F}_q(t)$, Corollary 3.9, [3] implies that the only fractions in the kernel of δ are fractions of polynomials in the kernel of δ . That is, the kernel of δ is $\mathbb{F}_q(t^p)$. \square

We note that $\mathbb{F}_q(t)$ is an algebraic extension of $\mathbb{K} = \mathbb{F}_q(t^p)$ of degree p . For consistency with the automorphism case we let $\nu = p = [\mathbb{F}_q(t) : \mathbb{K}]$.

Lemma 2.3.4. *The centre of $\mathbb{F}_q(t)[\mathcal{D}; \delta]$ is $\mathbb{F}_q(t^p)[\mathcal{D}^p]$.*

Proof. This is proven in [1]. See [17], Theorem 1.1.32. \square

For consistency with the automorphism case, we let $T = t^p$ and $Y = D^p$, so the centre of $\mathbb{F}_q(t)[\mathcal{D}; \delta]$ is the commutative polynomial ring $\mathbb{F}_q(T)[Y]$.

Again there is a close relationship between the centre, the minimal central multiples and the invariant elements. In the pure derivation case $\mathbb{F}_q(t)[\mathcal{D}; \delta]$, [1] shows $f^* = a(t)\hat{f}$ for $\hat{f} \in \mathfrak{C}$ and $a(t) \in \mathbb{F}_q(t)$.

2.3.3 Constructing the eigenring

To completely factor an $f \in \mathfrak{S}$, we construct a finite dimensional associative algebra \mathfrak{A} over the constant subfield \mathbb{K} , with the property that each non-trivial zero divisor in \mathfrak{A} yields a non-trivial factorization of f . A candidate for \mathfrak{A} might be the quotient $\mathfrak{S}/\mathfrak{S}f$, but it is in general only an \mathfrak{S} -module, and not an algebra. While we could in principle decompose this module directly, the algorithmic machinery to do so has not

been completely developed. $\mathfrak{S}/\mathfrak{S}f$ is only an algebra when $\mathfrak{S}f$ is a two-sided ideal in \mathfrak{S} . To regain some of the desirable structure of finite algebras, we follow [7], Section 0.7, and introduce the eigenring. Define

$$I(\mathfrak{S}f) = \{u \in \mathfrak{S} \mid fu \in \mathfrak{S}f\}$$

the *idealizer* of $\mathfrak{S}f$. The set $I(\mathfrak{S}f)$ is the largest subalgebra of \mathfrak{S} in which $\mathfrak{S}f$ is a two-sided ideal. The *eigenring* $E(\mathfrak{S}f)$ of $\mathfrak{S}f$ is defined as the quotient

$$E(\mathfrak{S}f) = I(\mathfrak{S}f)/\mathfrak{S}f,$$

a finite dimensional \mathbb{K} -algebra since \mathfrak{S} is an \mathbb{K} -algebra and $\mathfrak{S}f$ a two-sided ideal in $I(\mathfrak{S}f)$.

The key facts about $E(\mathfrak{S}f)$, which we shall prove in the coming subsections, are that it is a division ring if and only if f is irreducible, and that non-trivial zero divisors in $E(\mathfrak{S}f)$ allow us to compute non-trivial factors of f efficiently. We shall also prove that pairs of orthogonal idempotents summing to the identity in $E(\mathfrak{S}f)$ correspond to LCLM-decompositions.

Computationally, we will represent the eigenring as a finite dimensional subalgebra of a matrix ring over \mathbb{K} , where \mathbb{K} is the field of constants in $\mathbb{F}_q(t)$. If $\deg f = n$, the eigenring $E(\mathfrak{S}f)$ is isomorphic to the \mathbb{K} -algebra

$$\begin{aligned} \mathfrak{A} &= \{u \in I(\mathfrak{S}f) \mid \deg u < n\} \\ &= \{u \in \mathfrak{S} \mid fu \in \mathfrak{S}f \text{ and } \deg u < n\} \cong E(\mathfrak{S}f) \end{aligned}$$

under addition in \mathfrak{S} and multiplication in \mathfrak{S} reduced modulo f (i.e., each element in $E(\mathfrak{S}f)$ is represented by its unique residue modulo f of degree less than n).

To compute a \mathbf{K} -basis for \mathfrak{A} , let $W \subseteq \mathfrak{S}$ be the set of all $g \in \mathfrak{S}$ with $\deg g < n$. As a \mathbf{K} -vector space W is isomorphic to $\mathfrak{S}/\mathfrak{S}f$, with \mathbf{K} -basis

$$\{t^i \mathcal{D}^j \mid 0 \leq i < \nu, 0 \leq j < n\},$$

and dimension $n\nu$. Multiplication on the left by f induces an \mathbf{K} -linear map $\Psi : W \rightarrow W$: if $u \in W$ then $\Psi(u) = v$, where $fu = wf + v$ for $w \in \mathfrak{S}$ and $v \in \mathfrak{S}$, the unique remainder, with degree less than n . Clearly $v \in W$. The elements of \mathfrak{A} are exactly those elements in the null space of Ψ , a basis which is found by constructing a matrix for Ψ (an $n\nu \times n\nu$ matrix over \mathbf{K}) and then using linear algebra over \mathbf{K} to compute a basis for the null space. This matrix is computed by evaluating Ψ at each of the basis elements of W , i.e., finding $\Psi(t^i \mathcal{D}^j)$ for $0 \leq i < \nu$ and $0 \leq j < n$.

Thus, \mathfrak{A} can be presented by means of a \mathbf{K} -basis $A_1, \dots, A_m \in W$ of polynomials under multiplication in \mathfrak{S} reduced modulo f , or as a \mathbf{K} -basis $\mathfrak{A}_1, \dots, \mathfrak{A}_m \in \mathbf{K}^{n\nu \times n\nu}$ of matrices, where \mathfrak{A}_i specifies the linear action of A_i on W . Finding such a basis for \mathfrak{A} involves only $n\nu$ divisions with remainder in \mathfrak{S} of polynomials of degree less than n in \mathcal{D} and less than $\max\{\deg_t f, q\}$ in t , followed by linear algebra to find the kernel of Ψ . We obtain the following.

Theorem 2.3.5. *A basis $A_1, \dots, A_m \in W$ for \mathfrak{A} as a reduced polynomial algebra, or a basis $\mathfrak{A}_1, \dots, \mathfrak{A}_m \in \mathbf{K}^{n\nu \times n\nu}$ for \mathfrak{A} as a matrix algebra, can be found in time polynomial in $\deg_{\mathcal{D}} f$, $\deg_t f$ and p .*

2.3.4 Reducibility and the eigenring

We show that non-trivial zero-divisors exists in $E(\mathfrak{S}f)$ if and only if f has a non-trivial factorization.

Theorem 2.3.6. *Let $f \in \mathfrak{S}$. Then f has a non-trivial factorization if and only if $E(\mathfrak{S}f)$ has non-trivial zero divisors.*

Proof. Any non-trivial zero divisor in $E(\mathfrak{S}f)$ yields a non-trivial factorization of f . For if $u, v \in I(\mathfrak{S}f)$, with $u, v \notin \mathfrak{S}f$ and $uv \in \mathfrak{S}f$, it follows that $\text{gcd}(f, u) \neq 1$. To see this, suppose conversely that $\text{gcd}(f, u) = 1$. There exist $s, t \in \mathfrak{S}$ such that $sf + tu = 1$ and $sfv + tuv = v$. But $fv \in \mathfrak{S}f$ and $uv \in \mathfrak{S}f$ so $v \in \mathfrak{S}f$, a contradiction. If u, v are represented in the basis W , then they have degree (in \mathcal{D}) less than that of f , and hence $\text{gcd}(f, u)$ is a non-trivial right factor of f .

To prove the converse, assume f is not irreducible, and $f = gh$ for $g, h \in \mathfrak{S}$ of positive degree. Let $\hat{f} \in \mathfrak{C} = \mathbb{K}[Y]$ be the minimal central multiple of f .

First suppose \hat{f} is reducible as a polynomial in $\mathbb{K}[Y]$, that is, $\hat{f} = \hat{g}\hat{h}$ for $\hat{g}, \hat{h} \in \mathbb{K}[Y] \setminus \mathbb{K}$. Since $\hat{g}, \hat{h} \in \mathfrak{C}$, both \hat{g} and $\hat{h} \in I(\mathfrak{S}f)$. By the minimality of $\deg \hat{f}$, we also know $\hat{g}, \hat{h} \notin \mathfrak{S}f$, i.e., $\hat{g} + \mathfrak{S}f, \hat{h} + \mathfrak{S}f \in E(\mathfrak{S}f) \setminus \{0\}$. Finally, since $\hat{g}\hat{h} \in \mathfrak{S}f$, we see $\hat{g} + \mathfrak{S}f, \hat{h} + \mathfrak{S}f$ are zero divisors in $E(\mathfrak{S}f)$.

Now assume that \hat{f} is irreducible as a polynomial in $\mathbb{K}[Y]$ (and $f = gh$ is reducible in \mathfrak{S}), we show that in fact $f = \text{lcm}(f_1, f_2)$ for some $f_1, f_2 \in \mathfrak{S}$ of positive degree such that $\text{gcd}(f_1, f_2) = 1$, i.e., f is *LCLM-decomposable*. In this case there exist $g_1, g_2 \in \mathfrak{S}$ such that $g_1f_1 + g_2f_2 = 1$. Let $h_1 = g_1f_1$ and $h_2 = g_2f_2$, neither of which are equivalent to zero modulo f . Then

$$fh_1 = f(1 - g_2f_2) \in \mathfrak{S}f_2 \quad \text{and} \quad fh_1 = fg_1f_1 \in \mathfrak{S}f_1,$$

so $fh_1 \in \mathfrak{S}f$. Similarly $fh_2 \in \mathfrak{S}f$, so $h_1, h_2 \in I(\mathfrak{S}f)$. Moreover, $h_1h_2 = h_1 - h_1^2 = h_2 - h_2^2$, which is right-equivalent to zero modulo both f_1 and f_2 , and hence modulo f . Thus $(h_1 + \mathfrak{S}f)(h_2 + \mathfrak{S}f) \in \mathfrak{S}f$ and $h_1 + \mathfrak{S}f$ and $h_2 + \mathfrak{S}f$ are non-trivial zero divisors in $E(\mathfrak{S}f)$.

Still assuming that \hat{f} is irreducible as a polynomial in $K[Y]$ yet f is reducible, we now show that f is in fact LCLM-decomposable. Since $\mathfrak{S}\hat{f}$ is a maximal two-sided ideal in \mathfrak{S} , $E(\mathfrak{S}f) = \mathfrak{S}/\mathfrak{S}\hat{f}$ is a (finite dimensional) simple algebra and

$$0 \subsetneq \underbrace{E(\mathfrak{S}f)(f + \mathfrak{S}\hat{f})}_{\mathfrak{F}} \subsetneq \underbrace{E(\mathfrak{S}f)(h + \mathfrak{S}\hat{f})}_{\mathfrak{H}} \subseteq E(\mathfrak{S}f)$$

is a tower of left ideals in $E(\mathfrak{S}f)$. Since $E(\mathfrak{S}f)$ is simple there exists a complementary ideal \mathfrak{G} such that $\mathfrak{H} \cap \mathfrak{G} = \mathfrak{F}$ and $\mathfrak{F} + \mathfrak{G} = 1$. $E(\mathfrak{S}f)$ inherits from \mathfrak{S} the property of being a left principal ideal ring. That is, there exists a unique monic $\bar{g} \in \mathfrak{S}$ of minimal degree such that $\mathfrak{G} = \bar{g} + \mathfrak{S}\hat{f}$, called the *minimal left modular generator* of \mathfrak{G} . This follows easily from the fact that if $\bar{g}_1 + \mathfrak{S}\hat{f}, \bar{g}_2 + \mathfrak{S}\hat{f} \in \mathfrak{G}$, then $\text{gcd}(\bar{g}_1, \bar{g}_2) + \mathfrak{S}\hat{f} \in \mathfrak{G}$. Thus $\mathfrak{G} = \bar{g} + \mathfrak{S}\hat{f}$, $f = \text{lcm}(h, \bar{g})$ and $\text{gcd}(h, \bar{g}) = 1$. Thus f is decomposable, and hence its eigenring has zero divisors. \square

Note that the above theorem does not hold in general, at least in the derivation case in characteristic 0. [29] exhibits reducible polynomials whose eigenrings are division algebras. Essentially, it is the fact that $\mathbb{F}_q(t)$ is an *algebraic* extension of the field of constants that allows for the more representative structure of the eigenring.

2.3.5 Decomposability and the eigenring

For a given polynomial $f \in \mathfrak{S}$, we now show a correspondence between the existence of a pair of non-trivial orthogonal idempotents summing to the identity in $E(\mathfrak{S}f)$, and the existence of a non-trivial LCLM-decomposition of f .

Recall that two idempotents e_1, e_2 in an algebra are orthogonal if $e_1e_2 = e_2e_1 = 0$.

Theorem 2.3.7. *Let $e_1, e_2 \in I(\mathfrak{S}f)$ be such that $\bar{e}_1 = e_1 + \mathfrak{S}f$ and $\bar{e}_2 = e_2 + \mathfrak{S}f \in E(\mathfrak{S}f)$ are non-trivial orthogonal idempotents such that $e_1 + e_2 \in 1 + \mathfrak{S}f$ (so $e_1e_2 \in \mathfrak{S}f$*

and $e_1^2 - e_1, e_2^2 - e_2 \in \mathfrak{S}f$). Let $f_i \in \mathfrak{S} \setminus \{0\}$ be the polynomial of minimal degree such that $f_i e_i \subseteq \mathfrak{S}f$, for $i = 1, 2$. Then $f = \text{lcm}(f_1, f_2)$ and $\text{gcd}(f_1, f_2) = 1$.

Proof. For $i = 1, 2$, the set $J_i = \{u \in \mathfrak{S} : ue_i \in \mathfrak{S}f\}$ is a left ideal in \mathfrak{S} . Since \mathfrak{S} is a principal left ideal ring, J_i is generated by a unique monic $f_i \in \mathfrak{S}$. Note that f_i is a right factor of f since $f \in J_i$ (because $e_i \in I(\mathfrak{S}f)$).

For any $h \in \mathfrak{S}$ such that $h \in \mathfrak{S}f_1$ and $h \in \mathfrak{S}f_2$, then $h + \mathfrak{S}f = h(e_1 + e_2) + \mathfrak{S}f$. But $he_i \in \mathfrak{S}f$, so $h \in \mathfrak{S}f$. Hence $f = \text{lcm}(f_1, f_2)$.

We now show $\text{gcd}(f_1, f_2) = 1$. Let $g_1 = \text{gcd}(f, e_1)$, $g_2 = \text{gcd}(f, e_2)$. Clearly $\text{gcd}(g_1, g_2) = \text{gcd}(f, e_1, e_2) = 1$ since $e_1 + e_2 \in 1 + \mathfrak{S}f$. There exist $v_i, w_i \in \mathfrak{S}$ such that $v_i f + w_i e_i = g_i$. Now $g_2 e_1 = (v_2 f + w_2 e_2) e_1 = v_2 f e_1 + w_2 e_2 e_1 \in \mathfrak{S}f$. Thus $f_1 \mid g_2$. Similarly $f_2 \mid g_1$. Since $\text{gcd}(g_1, g_2) = 1$, it follows that $\text{gcd}(f_1, f_2) = 1$. \square

Theorem 2.3.8. *Given $f \in \mathfrak{S}$, and f_1, f_2 of positive degree such that $\text{gcd}(f_1, f_2) = 1$ and $f = \text{lcm}(f_1, f_2)$, there exist non-trivial orthogonal idempotents $\bar{e}_1, \bar{e}_2 \in E(\mathfrak{S}f)$ whose sum is $1 \in E(\mathfrak{S}f)$.*

Proof. Since $\text{gcd}(f_1, f_2) = 1$, there exists $g_1, g_2 \in \mathfrak{S}$ such that $g_1 f_1 + g_2 f_2 = 1$. Let $h_1 = g_1 f_1$ and $h_2 = g_2 f_2$, neither of which lie in $\mathfrak{S}f$. Then $fh_1 = f(1 - g_2 f_2) = f - fg_2 f_2 \in \mathfrak{S}f_2$ and $fh_1 = fg_1 f_1 \in \mathfrak{S}f_1$, so $fh_1 \in \mathfrak{S}f$. Similarly $fh_2 \in \mathfrak{S}f$, so $h_1, h_2 \in I(\mathfrak{S}f)$. We assign $\bar{e}_1 = h_1 + \mathfrak{S}f$, and $\bar{e}_2 = h_2 + \mathfrak{S}f$ and note that both are in $E(\mathfrak{S}f) = I(\mathfrak{S}f)/\mathfrak{S}$. Moreover, $h_1 h_2 = h_1 - h_1^2 = h_2 - h_2^2 \in \mathfrak{S}f$, since it is clearly in $\mathfrak{S}f_1$ and $\mathfrak{S}f_2$, and hence in $\mathfrak{S}f$. Thus \bar{e}_1 and \bar{e}_2 are orthogonal idempotents in $E(\mathfrak{S}f)$. As well, $h_1 + h_2 = 1$, so $\bar{e}_1 + \bar{e}_2 = 1 \in E(\mathfrak{S}f)$. \square

2.4 Factoring modular Ore polynomials

Theorem 2.3.6 effectively reduces the problem of factoring in \mathfrak{S} to finding zero divisors in a finite dimensional associative algebra over the field of constants \mathbb{K} in both the difference and differential case. To find such zero divisors quickly we use the algorithm of [15]. For any finite-dimensional associative algebra over $\mathbb{F}_q(T)$, their algorithm finds the Jacobson radical if it exists, and, for a semi-simple algebra, finds a system of primitive orthogonal idempotents. Otherwise it reports that \mathfrak{A} is a division algebra. In either case, the output immediately yields zero divisors or states that none exists. The algebra is presented to the zero-divisor finding algorithm as a basis of generating matrices over $\mathbb{F}_q(T)$ (also known as the *structure constants* for the algebra). The algorithm of [15] runs in time polynomial in the dimension of the algebra and the height (degree in T) of the structure constants (entries in the basis). Computing a basis for the eigenring $\mathfrak{A} \cong E(\mathfrak{S}f)$ has been discussed in the previous section.

The cost of their algorithm is polynomial in the dimension of the algebra, the maximum degree of the entries of the generating matrices, and $\log q$.

Again, we let $\mathfrak{S} = \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, transformed to a pure differential or automorphism ring as in Section 2.2.

Algorithm: Factorization

Input: ▶ $f \in \mathfrak{S}$ of degree n ;

Output: ▶ $g, h \in \mathfrak{S}$, or a message that f is irreducible;

(1) Compute a basis $A_1, \dots, A_m \subseteq \mathbb{K}^{n\nu \times n\nu}$ for $\mathfrak{A} \cong E(\mathfrak{S}f)$ as in Subsection 2.3.3;

(2) If \mathfrak{A} is a division ring then report “ f is irreducible”

Else

- (3) Find a non-trivial left zero divisor $u \in \mathfrak{A}$;
 Compute $h := \text{gcd}(f, u)$ and $g \in \mathfrak{S}$ with $f = gh$;
- (5) Return g, h ;

Theorem 2.4.1. *The above algorithm works as specified. It runs in time polynomial in $\deg_{\mathcal{D}}(f)$, $\deg_t(f)$ and q .*

Proof. Correctness when f is irreducible follows from Theorem 2.3.6, as does the existence of u when f is irreducible. Suppose that $v \in \mathfrak{A} \setminus \{0\}$ is such that $uv = 0$. To see that $\text{gcd}(f, u)$ is a non-trivial factor of f , assume to the contrary that $\text{gcd}(f, u) = 1$. Then there exist $s, t \in \mathfrak{S}$ such that $sf + tu = 1$ and $sfv + tuv = v$. But $fv \in \mathfrak{S}f$ and $uv \in \mathfrak{S}f$, so $v \in \mathfrak{S}f$, a contradiction to v being non-trivial in \mathfrak{A} (which is defined modulo f).

The costs are dominated by finding a basis for the eigenring (see Theorem 2.3.5) and finding the zero-divisors in \mathfrak{A} (see [15]), all of which are polynomial in $\deg_{\mathcal{D}}(f)$, $\deg_t(f)$ and q . □

2.5 Computing a complete LCLM decomposition

In this section we present a method for the LCLM-decomposition of a polynomial $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$. Again, we let $\mathfrak{S} = \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, transformed to a pure differential or automorphism ring as in Section 2.2.

Using the relationship between the existence of an LCLM-decomposition and the existence of non-trivial, orthogonal idempotents in $E(\mathfrak{S}f)$ summing to 1 developed in Theorems 2.3.7 and 2.3.8 of the previous section, we reduce the problem to finding such idempotents.

We need to be able to find a pair of orthogonal idempotents in \mathfrak{A} which sum to the identity. We again employ the algorithm of [15], this time to compute $\text{Rad}(\mathfrak{A})$ and an idempotent in $\mathfrak{A} \bmod \text{Rad}(\mathfrak{A})$. We use this to construct an idempotent in \mathfrak{A} as follows.

Algorithm: Orthogonal-Idempotents

Input: ▶ a basis $\mathfrak{A}_1, \dots, \mathfrak{A}_m \in \mathbb{K}^{r \times r}$ for an associative algebra (containing the identity). Here $\mathbb{K} = \mathbb{F}_q(T)$ for an indeterminate T , and q some power of a prime q .

Output: ▶ non-trivial orthogonal idempotents $e_1, e_2 \in \mathfrak{A}$ such that $e_1 + e_2 = 1$.

- (1) Compute $\text{Rad}(\mathfrak{A})$, and assume $\text{Rad}(\mathfrak{A})^n = 0$, using the algorithm of [15].
- (2) If $\mathfrak{A}/\text{Rad}(\mathfrak{A})$ is a division algebra;
- (3) Then output “ \mathfrak{A} has no non-trivial idempotents”;
- (3) Else
- (4) Let $\bar{e} \in \mathfrak{A}$ be such that $\bar{e} \bmod \text{Rad}(\mathfrak{A})$ is a non-trivial idempotent in $\mathfrak{A}/\text{Rad}(\mathfrak{A})$, using the algorithm of [15].
- (5) Compute $u \in \mathbb{K}[x]: u \equiv 1 \pmod{x^n}, u \equiv 0 \pmod{(x-1)^n};$
 $v \in \mathbb{K}[x]: v \equiv 0 \pmod{x^n}, v \equiv 1 \pmod{(x-1)^n};$
- (6) Output $e_1 := u(\bar{e}), e_2 = v(\bar{e});$

Theorem 2.5.1. *Orthogonal-Idempotents works as specified. The time required is polynomial in r and the maximum degree of any entry in \mathfrak{A}_i for $1 \leq i \leq m$.*

Proof. We need to show that e_1, e_2 are orthogonal idempotents in \mathfrak{A} . First, note that $\bar{e}^2 - \bar{e} \in \text{Rad}(\mathfrak{A})$, so $(\bar{e}^2 - \bar{e})^n = 0$. Thus, the minimal polynomial in $\mathbb{K}[x]$ of \bar{e} divides

$(x^2 - x)^\eta = x^\eta(x - 1)^\eta$. With u, v as in step (5), we know

$$\begin{aligned} u(x) &= q_1(x)x^\eta + 1 & u(x) &= q_2(x)(x - 1)^\eta \\ v(x) &= r_1(x)x^\eta & v(x) &= r_2(x)(x - 1)^\eta + 1 \end{aligned}$$

for some $q_1, q_2, r_1, r_2 \in \mathbb{K}[x]$. Now

$$\begin{aligned} e_1^2 - e_1 &= e_1(e_1 - 1) = q_2(\bar{e})(\bar{e} - 1)^\eta q_1(\bar{e})\bar{e}^\eta \\ &= \bar{e}^\eta(\bar{e} - 1)^\eta q_1(\bar{e})q_2(\bar{e}) = 0, \\ e_2^2 - e_2 &= e_2(e_2 - 1) = r_2(\bar{e})(\bar{e} - 1)^\eta r_1(\bar{e})\bar{e}^\eta \\ &= \bar{e}^\eta(\bar{e} - 1)^\eta r_1(\bar{e})r_2(\bar{e}) = 0, \\ e_2e_1 &= e_1e_2 = q_1(\bar{e})(\bar{e} - 1)^\eta r_1(\bar{e})\bar{e}^\eta \\ &= \bar{e}^\eta(\bar{e} - 1)^\eta q_1(\bar{e})r_1(\bar{e}) = 0. \end{aligned}$$

Finally $e_1 + e_2 = u(\bar{e}) + v(\bar{e}) = (u + v)(\bar{e})$. But $u + v \equiv 1 \pmod{x^\eta(x - 1)^\eta}$ by construction, so $e_1 + e_2 = 1$.

That the algorithm runs in polynomial in the dimension r and the maximum degree of a structure constant follows immediately from [15]. \square

We can now present our algorithm to compute an LCLM-decomposition.

Algorithm: LCLM-Decomposition

Input: $\blacktriangleright f \in \mathfrak{S}$ of degree n ;

Output: $\blacktriangleright g, h \in \mathfrak{S}$ with $f = \text{lclm}(g, h)$ and $\text{gcd}(g, h) = 1$ for $1 \leq i \leq k$;

(1) Compute a basis $A_1, \dots, A_m \subseteq \mathbb{K}^{n\nu \times n\nu}$ for $\mathfrak{A} \cong E(\mathfrak{S}f)$ as in Section 2.3;

(2) Using **Orthogonal-Idempotents**, attempt to find a pair of orthogonal idempotents $e_1, e_2 \in \mathfrak{A}$ such that $e_1 + e_2 = 1 \in \mathfrak{A}$;

(3) If no such idempotents exist, report “ f is indecomposable”;

Else

(4) Let $g, h \in \mathfrak{S}$ each be of minimal degree such that $ge_1 \in \mathfrak{S}f$, and $he_2 \in \mathfrak{S}f$;

(5) Return g, h ;

Theorem 2.5.2. *The above algorithm works as specified. It runs in time polynomial in $\deg_{\mathcal{D}}(f)$, $\deg_t(f)$ and p .*

Proof. Correctness of Step (3) follows from Theorem 2.3.8, and that of Step (5) follows from Theorem 2.3.7.

The dominant cost in this algorithm is computing the orthogonal idempotents with `Orthogonal-Idempotents`, which can be done in time polynomial in the dimension and degree of entries in the basis for \mathfrak{A} , i.e., polynomial in n , $\deg_t(f)$ and p . As well, we must compute g and h in step (4), but this is simply linear algebra in W (see Subsection 2.3.3). \square

2.6 Conclusion

Given $f \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, we have considered the problems of computing factorizations $f = gh$ for $g, h \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$, and LCLM-decompositions $f = \text{lclm}(g, h)$ for relatively (right) prime g, h . The algorithms presented require time polynomial in $\deg_{\mathcal{D}} f$, $\deg_t f$ and q . Our main idea was to work by decomposing the eigenring, and use the correspondence between these decompositions and factorizations to split f .

Many important questions remain to be considered. We have shown here how to split and LCLM-decompose polynomials into two factors, and can iterate the algorithm on these factors until a complete factorization or complete LCLM-decomposition

is obtained. However, we must be careful to manage the growth in the degree of t in the factors. We can also ask if there are any factors of a specific given degree s . For this an analogue to the method of [11] should apply.

As noted earlier, one of the main goals was to apply this work to factoring in $\mathbb{Q}(t)[\mathcal{D}; \sigma, \delta]$. The problems of lifting and factor reconstruction are considerably more difficult in the non-commutative case. Interesting mathematical questions also arise in understanding the factorization pattern under modular reduction, which could greatly affect the complexity of algorithms for these problems.

Bibliography

- [1] S. A. Amitsur. Derivations in simple rings. *Proc. London Mathematical Society*, VII, 3rd series:87–112, 1957.
- [2] E. Beke. Die irreducibilität der homogenen linearen differentialgleichungen. *Math. Annalen*, 45:278–300, 1894.
- [3] J. Bergen and S. Montgomery. Smash products and outer derivations. *Israel Journal of Mathematics*, 53(1):321–345, 1986.
- [4] M. Bronstein and M. Petkovšek. On Ore rings, linear operators and factorisation. *Programmírovanie*, 20:27–45, 1994.
- [5] M. Bronstein and M. Petkovšek. An introduction to pseudo-linear algebra. *Theoretical Computer Science*, 157:3–33, 1996.
- [6] T. Cluzeau. Factorization of differential systems in characteristic p . These proceedings. 2003.
- [7] P. Cohn. *Free Rings and their Relations*. Academic Press, London, 1985.
- [8] P. Cohn. *Skew Fields: Theory of General Division Rings*. Cambridge, London, 1995.

- [9] M. Giesbrecht. Some results on the functional decomposition of polynomials. Master's thesis, University of Toronto, 1988. Also available as University of Toronto Technical Report 209/88.
- [10] M. Giesbrecht. Factoring in skew-polynomial rings. In *Proceedings of Latin American Theoretical INformatics Conference (LATIN)*, pages 191–203, Sao Paulo, Brasil, 1992.
- [11] M. Giesbrecht. Factoring in skew-polynomial rings over finite fields. *J. of Symbolic Computation*, 24(5):463–486, 1998.
- [12] M. van Hoeij. Rational solutions of the mixed differential equation and its application to factorization of differential operators. In *Proc. ISSAC'96*, pages 219–225, 1996.
- [13] M. van Hoeij. Factorization of differential operators with rational functions coefficients. *J. Symb. Comp.*, 24: 537–561, 1997.
- [14] J. van der Hoeven. FFT-like multiplication of linear differential operators. *J. Symb. Comp.*, 33 (1): 123–127, 2002.
- [15] G. Ivanyos, L. Rónyai, and A. Szántó. Decomposition of algebras over $\mathbb{F}_q(x_1, \dots, x_m)$. *Applicable Algebra in Engineering, Communication and Computing*, 5:71–90, 1994.
- [16] N. Jacobson. *The Theory of Rings*. American Math. Soc., New York, 1943.
- [17] N. Jacobson. *Finite-Dimensional Division Algebras over Fields*. Springer-Verlag, 1996.

- [18] Z. Li. A subresultant theory for ore polynomials with applications. In *Proc. ISSAC 1998*, pages 132–139, 1998.
- [19] Z. Li. Some bounds for skew polynomials, 2002. Preprint.
- [20] S. Montgomery. *Fixed rings of finite automorphism groups of associative rings*, volume 818 of *Lecture Notes in Mathematics*. Springer-Verlag, 1980.
- [21] O. Ore. Formale Theorie der linearen Differentialgleichungen. *J. reine angew. Math.*, 168: 233–252, 1932.
- [22] O. Ore. *On a special class of polynomials*. *Trans. Amer. Math. Soc.*, 35: 559–584.
- [23] O. Ore. Theory of non-commutative polynomials. *Annals of Mathematics*, 34 (22): 480–508.
- [24] O. Ore. Contributions to the theory of finite fields. *Trans. Amer. Math. Soc.*, 36: 243–274, 1934.
- [25] M. van der Put. Differential equations in characteristic p . *Compositio Mathematica*, 97: 227–251, 1995.
- [26] M. van der Put. Reduction modulo p of differential equations. *Indag. Mathem.*, 7 (3): 367–387, 1996.
- [27] M. van der Put. Modular methods for factoring differential operators. Unpublished manuscript, 34 pp., 1997.
- [28] M. van der Put and M.F. Singer. *Galois Theory of Difference Equations*. LNM 1666. Springer, 1997.

- [29] M. F. Singer. Testing reducibility of linear differential operators: A group theoretic perspective. *Applicable Algebra in Engineering, Communication and Computing*, 7 (2): 77–104, 1996.

Chapter 3

Non-Commutative Gröbner Bases in Poincaré-Birkhoff-Witt Extensions

Commutative Gröbner Bases are a well established technique with many applications, including polynomial solving and constructive approaches to commutative algebra and algebraic geometry. Noncommutative Gröbner Bases are a focus of much recent research activity. For example, combining invariant theory and elimination theory, or elimination in moving frames of partial differential operators invariant under an equivalence group, requires the use of noncommutative Gröbner bases. This paper presents theory and algorithms for noncommutative Gröbner bases in Poincaré-Birkhoff-Witt extensions. These extension rings generalize the previous domains over which noncommutative Gröbner bases have been applied. Our approach to noncommutative Gröbner bases differs from previous work which assumes that the coefficients are from a field or commutative ring. In applications such as Cartan's method of moving frames, this is not the case, and the theory that we present can be applied.

3.1 Introduction

Since the mid-1980's, noncommutative Gröbner bases have developed as an active research area in Computer Algebra, with many applications. See, for example, Chyzak and Salvy [6] for Ore algebras, Green [10] for path algebras, Kandri-Rody and Weispfenning [12] for algebras of solvable type, Mora [17] for free algebras over fields. Generally speaking, there are two streams in these studies. One stream is focused on free algebras, which preserve properties of semigroups. The other stream focuses on algebras of solvable type (including rings of differential operators) which admit Dickson's lemma allowing algorithms to terminate.

In most of the above papers, the authors assume that the coefficients are from a field or commutative ring, and that these commute with the indeterminates (although the indeterminates may not commute with each other).

There are many interesting and useful rings which the above papers do not address. Examples include some kinds of homogenous partial differential equations with non-constant coefficients (see Adams et al. [2]). The method of choosing coordinates which are invariant under a given symmetry group (e.g., polar coordinates), in its most general form requires the introduction of a moving frame of non-commuting partial differential operators (Cartan's famous equivalence method). Elimination theories for such systems, by necessity, require a non-commutative Gröbner Basis method of the type presented in this paper (see Lisle & Reid [14], and Mansfield [15]). Another example is the skew enveloping algebra $R\#U(L)$ (see McConnell and Robson [16]), which is important in associative ring theory. This motivates us to define Gröbner bases in Poincaré-Birkhoff-Witt (PBW) extensions. We prove the left division rule and many fundamental properties of such Gröbner bases, and give an algorithm to

construct them. As a special case, we consider the graded lexicographic ordering, and reduce computing Gröbner bases of PBW extensions to the commutative case. Finally we apply this theory to the moving frame approach in Section 3.4.

Differential elimination algorithms have been effective in pre-processing and simplifying systems for the subsequent application of the methods of scientific computation. Such methods include numerical integration techniques and symmetry techniques.

A popular new research area is the area of Geometric Integration [18]. The general philosophy is to include as many qualitative features of the physical system being studied in the tools that you use to study the system. For example numerical integrators, which are invariant under the symplectic group (*geometric integrators*) are used to numerically solve Hamilton's equations which are also symplectically invariant. This paper represents progress in this direction for differential elimination algorithms by, for example, enabling the differentiations and eliminations of such algorithms to be executed in a moving frame, invariant under a group admitted by the given problem. Our extension of such algorithms to coefficients which do not come from a field (i.e., are not invertible) is potentially relevant for matrix formulations arising in non-commutative field theories. In such instances it is helpful to be able to perform such calculations, as physicists would, in the non-commutative matrix formalism, instead of breaking it down to components, and using commutative differential algebra, as is the current practise. This is an area which we are investigating.

3.2 Poincaré-Birkhoff-Witt Extensions

We first introduce the definition of PBW extensions used in this paper, which is defined by Bell and Goodearl [5]. This definition leads to a unified treatment of many polynomial rings currently studied in associative ring theory and Computer Algebra.

Definition 3.2.1. Let R and T be two associative rings with $R \subseteq T$. T is called a (*finite*) *PBW extension* of R if there exist $x_1, x_2, \dots, x_n \in T$ such that

- (1) the monomials $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ form a basis for T as a free left R -module, where $i_1, i_2, \dots, i_n \in \mathbb{N}$;
- (2) $x_i r - r x_i := [x_i, r] \in R$ for each $i = 1, \dots, n$ and any $r \in R$;
- (3) $x_i x_j - x_j x_i := [x_i, x_j] \in R + R x_1 + \cdots + R x_n$ for all $i, j = 1, \dots, n$.

We write $T = R\langle x_1, \dots, x_n \rangle$.

Conditions (2) and (3) allow one to swap the positions of elements such that every element in T has a unique representation $\sum r_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ where $r_\alpha \in R$ and $\{\alpha_i\} \in \mathbb{N}$.

We call the monomials of form $x_1^{i_1} \cdots x_n^{i_n}$ *standard monomials*. Note also that any non-standard monomial $x_{j_1}^{i_1} \cdots x_{j_n}^{i_n}$ can be expressed as a finite sum of standard monomials.

Remark 3.2.1. We have to remind the reader of the differences between PBW extensions and other similar algebraic structures. For example, comparing PBW extensions with algebras of solvable type defined in [12], we note a number of important differences. Algebras of solvable type require coefficients and indeterminates commute, whereas PBW extensions make no such requirement. For example, $R\#U(L)$ is not an algebra of solvable type in general, but is a PBW extension. On the other hand,

algebras of solvable type define a “quantum” or “ordering” version of condition (3) above, which states roughly that the commutator is smaller than the product under a term ordering. In an upcoming paper we define skew-PBW extensions which include both algebras of solvable type and PBW extensions. A comparison of PBW extensions with free algebras can assist in understanding their differences.

Example 3.2.1. *There are several prototypical examples of PBW extensions:*

- (1) *The usual multivariate polynomial rings over R , Ore algebras and PBW algebras discussed in Computer algebra. The skew enveloping algebra (or smash product) $R\#U(L)$ is an example of PBW extensions. In particular, the universal enveloping algebra $U(L)$, the n -Weyl algebra $A_n(K)$ and skew polynomial ring (derivation type) $R[x, \delta]$ also are examples of PBW extensions (see McConnell and Robson [16]).*
- (2) *Some kinds of PDEs with non-constant coefficients. For example, let $R = K(x_1, \dots, x_n)$ be a function field over a field K . It is well-known that all partial derivations over R form a (possibly infinite dimensional) Lie algebra under the usual brace product. In the language of PDEs, one partial differential equation is written as:*

$$a_n(x_1, \dots, x_n) \frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_n}}{\partial x_n^{i_n}} + \cdots + a_0(x_1, \dots, x_n) = 0.$$

Corresponding to the skew enveloping algebra, the above equation is in the ring $K(x_1, \dots, x_n)\#U(L)$, and can be written as:

$$a_n(x_1, \dots, x_n) \bar{x}_1^{i_1} \cdots \bar{x}_n^{i_n} + \cdots + a_0(x_1, \dots, x_n) = 0.$$

- (3) *Moving frames (see Section 3.4) and some systems of linear homogeneous partial differential equations with non-constant coefficients (see [2], [14] and [15]).*

We denote by $\mathcal{M}(X)$ or $\mathcal{M}(x_1, \dots, x_n)$ the set $\{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \mid \alpha_i \in \mathbb{N}\}$ of all standard monomials of $\{x_1, \dots, x_n\}$. For simplicity, we write $X^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$. Given a total order \prec on the set of standard monomials, we define the *leading monomial* $\text{lm}(f)$ of $f \in R\langle x_1, \dots, x_n \rangle$ to be the largest standard monomial occurring in f with non-zero coefficient, the *leading coefficient* $\text{lc}(f)$ to be the coefficient of $\text{lm}(f)$ and *leading term* $\text{lt}(f) = \text{lc}(f) \cdot \text{lm}(f)$. For a subset $S \subseteq R\langle x_1, \dots, x_n \rangle$, $\text{lm}(S) = \{\text{lm}(s) \mid s \in S\}$ while $\text{lc}(S)$ and $\text{lt}(S)$ are similarly defined and $\ell(S)$ will be the left ideal generated by S in $R\langle x_1, \dots, x_n \rangle$. The *degree* of $X^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is $\text{deg}(X^\alpha) = |\alpha| := \alpha_1 + \cdots + \alpha_n$.

Definition 3.2.2. An *admissible order* \prec on $\mathcal{M}(X)$ is a total order of $\mathcal{M}(X)$ which satisfies:

- (1) *multiplicative*, i.e., $r \prec X^\alpha$ and $X^\alpha \prec X^\beta$ imply that $\text{lm}(X^\eta X^\alpha X^\gamma) \prec \text{lm}(X^\eta X^\beta X^\gamma)$, where $X^\eta, X^\alpha, X^\beta, X^\gamma \in \mathcal{M}(X)$ and $r \in R$.
- (2) *degree compatible*, i.e., $\text{deg}(X^\alpha) \prec \text{deg}(X^\beta)$ implies $X^\alpha \prec X^\beta$, where $X^\alpha, X^\beta \in \mathcal{M}(X)$.

Remark 3.2.2. From condition (2), we get the *descending chain condition*, i.e., there are no infinite strictly descending chains of standard monomials, which is vital in the proof of termination of algorithms.

Example 3.2.2. *The typical example of this order is the graded lexicographic ordering. That is, $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \prec x_1^{\beta_1} x_2^{\beta_2} \cdots x_n^{\beta_n}$ if and only if the first nonzero component of $(\sum_{k=1}^n (\alpha_k - \beta_k), \alpha_1 - \beta_1, \dots, \alpha_n - \beta_n)$ is negative.*

Before computing Gröbner bases in PBW extensions, we first need a notion of the coefficient ring being computable.

Definition 3.2.3. An associative (though not necessarily commutative) ring R is *(left) computable*, if in addition to the usual arithmetic operations being computable, the following two conditions hold:

- (1) *(left ideal membership)* Given $a, a_1, \dots, a_m \in R$, there is an algorithm which decides whether a is in the left ideal $R(a_1, \dots, a_m)$ and if so, finds $b_1, b_2, \dots, b_m \in R$ such that $a = \sum_{i=1}^m b_i a_i$.
- (2) *(left syzygies)* Given $a_1, \dots, a_m \in R$ there is an algorithm which finds a finite set of generators for the R -module

$$\text{Syz}(a_1, \dots, a_m) := \{(b_1, \dots, b_m) \in R^m \mid \sum_{i=1}^m b_i a_i = 0\}.$$

If R is a field, the condition “left ideal membership” is trivial. But if R is a ring, it is a useful and necessary condition. The condition “left syzygies” is needed to guarantee that the algorithm **GröbnerPBW** which follows is implementable, since from the noetherian condition we only know that there exist a finite number of generators. In fact these conditions have been used in many papers, for example, Gianni, Trager and Zacharias [9]. It is a common condition when one considers Gröbner theory on rings instead of fields. There are many rings satisfying this condition, for example, the usual polynomial ring over a field and the universal enveloping Lie algebra over a field (see Apel and Lassner [4]).

In the remainder of the paper, we assume that R is a left computable and noetherian (not necessarily commutative) ring with a finite PBW extension $R\langle x_1, \dots, x_n \rangle$, and \prec is an admissible order on $\mathcal{M}(X)$.

3.3 Gröbner Bases in PBW Extensions

In [8], Galligo first considered Gröbner bases in the ring of linear differential operators. Later, many authors extended Galligo's idea to various rings. For example, Chyzak and Salvy [6] consider Gröbner bases in Ore algebras, and Insa and Pauer [11] for Gröbner bases in the ring of differential operators, assuming the coefficients form a subring of a function field. In this paper we consider the more general case, PBW extensions.

Definition 3.3.1. For $f, g \in \mathcal{M}(X)$, f is a *factor* of g if there exist $p, q \in \mathcal{M}(X)$ such that $g = \text{lm}(pfq)$. Also g is said to be divisible by f , denoted by $f|g$.

It is easy to see that if $f = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and $g = x_1^{\beta_1} \cdots x_n^{\beta_n}$, then f is a factor of g if and only if $\alpha_k \leq \beta_k$, for all $1 \leq k \leq n$.

In the following definition, we only require that the leading terms are reduced. The conditions involving reduced polynomials (over rings) are no longer required, and so reducing non-leading terms is unnecessary (see [1], p.211).

Definition 3.3.2. Let G be a finite subset of $R\langle x_1, \dots, x_n \rangle$ and $f, h \in R\langle x_1, \dots, x_n \rangle$. We say h is a *one step reduction* of f modulo G , denoted $f \rightarrow^G h$, if for the leading term aX^α in f , there exist $g_1, \dots, g_t \in G$ and $r_1, \dots, r_t \in R$ such that

- (1) $\text{lm}(g_i)$'s are factors of X^α , say $\text{lm}(X^{\beta_i} g_i X^{\gamma_i}) = X^\alpha$, for all $i = 1, \dots, t$;
- (2) $a = \sum_{i=1}^t r_i \text{lc}(g_i)$ for some $r_i \in R, i = 1, \dots, t$;
- (3) $h = f - \sum_{i=1}^t r_i X^{\beta_i} g_i X^{\gamma_i}$.

Furthermore, we say f *reduces to h modulo G* if and only if there exist $h_1, \dots, h_s \in R\langle x_1, \dots, x_n \rangle$ such that $f \rightarrow^G h_1 \rightarrow^G h_2 \rightarrow^G \dots \rightarrow^G h_s \rightarrow^G h$.

Remark 3.3.1. (1) Note that the conditions of (1) and (3) are equivalent to:

(1)' $\text{lm}(g_i)$'s are factors of X^α , say $\text{lm}(X^{\beta_i}g_i) = X^\alpha$, for all $i = 1, \dots, t$;

(3)' $h = f - \sum_{i=1}^t r_i X^{\beta_i} g_i$.

(2) We remind the reader that Gauss reduction does not work in PBW extensions since the elements of R are not necessarily invertible. Therefore we have to use "sum" to cancel some terms. For example, consider the PBW extension $K[y_1, y_2, y_3]\langle x_1, x_2, x_3 \rangle$, where $K[y_1, y_2, y_3]$ is the usual polynomial ring over a field K and $\{x_1, x_2, x_3\}$ is the 3-dimensional Lie algebra with $[x_1, x_2] = x_1$, $[x_1, x_3] = -2x_1$ and $[x_2, x_3] = -2x_3$. Set $G = \{g_1 := y_1x_1, g_2 := y_2x_2\}$ and $f := (2y_1 + 3y_2)x_1x_2 + x_1 + 1$. Since $K[y_1, y_2, y_3]$ is just a ring, not a field, we do elimination as follows:

$$\begin{aligned} h &= f - (2g_1x_2 + 3g_2x_1) = (2y_1 + 3y_2)x_1x_2 + x_1 + 1 - (2(y_1x_1)x_2 + 3(y_2x_2)x_1) \\ &= (2y_1 + 3y_2)x_1x_2 + x_1 + 1 - (2(y_1x_1x_2) + 3y_2(x_1x_2 - x_3)) = x_1 - 3y_2x_3 + 1. \end{aligned}$$

Definition 3.3.3. An element $f \in R\langle x_1, \dots, x_n \rangle$ is said to be in *reduced form* with respect to G if f can not be reduced modulo G . A *reduced form* of f modulo G is an element $h \in R\langle x_1, \dots, x_n \rangle$ such that h is in reduced form with respect to G and $f \rightarrow^G h$.

As in the commutative case we have a division rule, but here it is one-sided.

Proposition 3.3.1. (Left division rule) Let $G := \{g_1, \dots, g_t\} \subseteq R\langle x_1, \dots, x_n \rangle$ and $f \in R\langle x_1, \dots, x_n \rangle$. Then there exist $h_1, \dots, h_t, \psi \in R\langle x_1, \dots, x_n \rangle$ such that $f = h_1g_1 + \dots + h_tg_t + \psi$, where ψ is reduced modulo G and

$$\text{lm}(f) = \max\{\max\{\text{lm}(\text{lm}(h_i) \text{lm}(g_i))\}_{i=1}^t, \text{lm}(\psi)\}$$

Proof. If f is reduced modulo G , then there is nothing to do. Assume then that there is a reduction chain $f \rightarrow \psi_1 \rightarrow \psi_2 \rightarrow \dots$. By the definition of reduction, the leading term is decreased, that is, $\text{lm}(f) \succ \text{lm}(\psi_1) \succ \text{lm}(\psi_2) \succ \dots$. Since \succ is a well-ordering, the reduction chain has to stop, say, $f \rightarrow \psi_1 \rightarrow \dots \rightarrow \psi_s \rightarrow \psi$, where ψ is reduced modulo G . By Remark 3.3.1, we have $f - \psi_1 = r_{11}X^{\alpha_{11}}g_1 + \dots + r_{1t}X^{\alpha_{1t}}g_t$, where $\{r_{1i}\} \in R$, $\text{lt}(f) = r_{11}\text{lt}(X^{\alpha_{11}}g_1) + \dots + r_{1t}\text{lt}(X^{\alpha_{1t}}g_t)$ (some r_{ji} and α_{ji} could be zero) and $\text{lm}(f) = \text{lm}(X^{\alpha_{1i}}g_i)$ for all i such that $r_{1i} \neq 0$ (since R could have zero-divisors). Similarly we have the representation for $\psi_1 - \psi_2 = r_{21}X^{\alpha_{21}}g_1 + \dots + r_{2t}X^{\alpha_{2t}}g_t$. Therefore

$$f - \psi_2 = (f - \psi_1) + (\psi_1 - \psi_2) = (r_{11}X^{\alpha_{11}} + r_{21}X^{\alpha_{21}})g_1 + \dots + (r_{1t}X^{\alpha_{1t}} + r_{2t}X^{\alpha_{2t}})g_t.$$

Note that the coefficients of $\{g_i\}$ are not all zero if $s > 2$. Continuing in this way, we get the representation for $f - \psi$ as required. □

The above proposition gives a method to calculate the reduced form:

Algorithm: ReducedForm

Input: $\blacktriangleright G = \{g_1, \dots, g_t\} \subseteq R\langle x_1, \dots, x_n \rangle$;

$\blacktriangleright f \in R\langle x_1, \dots, x_n \rangle$;

Output: \blacktriangleright a reduced form of f modulo G : $h_1, \dots, h_t, \psi \in R\langle x_1, \dots, x_n \rangle$ such that

$$f = h_1g_1 + \dots + h_tg_t + \psi, \text{ where } \psi \text{ is reduced modulo } G \text{ and } \text{lm}(f) = \max\{\max\{\text{lm}(\text{lm}(h_i)\text{lm}(g_i))\}_{i=1}^t, \text{lm}(\psi)\};$$

Set $\psi := f$ and $h_1, \dots, h_t := 0$;

While $\psi \neq 0$ and $\text{lc}(\psi) \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) \mid \text{lm}(\psi))$ Do

$$\text{Find } \{r_i\}_1^t \in R, \{X^{\alpha_i}\}_1^t \text{ so } \text{lc}(\psi) = \sum_{i=1}^t r_i \text{lc}(g_i), \text{lm}(X^{\alpha_i}g_i) = \text{lm}(\psi);$$

$$\psi := \psi - \sum_{i=1}^t r_i X^{\alpha_i} g_i.$$

For $i = 1$ to t Do $h_i := h_i + r_i X^{\alpha_i}$;

End

Definition 3.3.4. Let I be a left ideal of $R\langle x_1, \dots, x_n \rangle$ and G a subset of I . Then G is called a (left) Gröbner basis of I if for all $f \in I$, $f \rightarrow^G 0$.

While some of the definitions and theorems in the PBW extensions case are equivalent to those in the commutative case, others do not hold.

Theorem 3.3.2. Let I be a left ideal of $R\langle x_1, \dots, x_n \rangle$ and let G be a finite subset of I . The following assertions are equivalent:

- (1) G is a left Gröbner basis of I ;
- (2) For all $0 \neq f \in I$, f is reducible modulo G ;
- (3) For all $0 \neq f \in I$, there exist $g_1, \dots, g_t \in G$ such that $\text{lm}(g_i), i = 1, \dots, t$ are factors of $\text{lm}(f)$ and $\text{lc}(f) \in \ell(\text{lc}(g_1), \dots, \text{lc}(g_t))$;
- (4) For $\alpha \in \mathbb{N}^n$, let

$\ell(\alpha, I) := \ell(\text{lc}(f) : f \in I, \text{lm}(f) = X^\alpha)$. Then for all $\alpha \in \mathbb{N}^n$ the left ideal $\ell(\alpha, I)$ is generated by $\{\text{lc}(g) : g \in G, \text{lm}(g) | X^\alpha\}$.

Proof. (1) \Rightarrow (2) follows from the definition of Gröbner bases.

(2) \Rightarrow (1): Let $G := \{g_1, \dots, g_t\}$ and $0 \neq f \in I$. Since f is reducible, by induction and Proposition 3.3.1, there exist $h_1, \dots, h_t, \psi \in R\langle x_1, \dots, x_n \rangle$ such that $f = h_1 g_1 + \dots + h_t g_t + \psi$, where ψ is reduced modulo G . This implies that $\psi = f - (h_1 g_1 + \dots + h_t g_t) \in I$, and so ψ is reducible by (2), a contradiction. Therefore $\psi = 0$, that is, f can be reduced to 0 modulo G .

(2) \Leftrightarrow (3) follows from the definition of reduction.

(3) \Rightarrow (4): For fixed $\alpha \in \mathbb{N}^n$, for any $r \in \ell(\alpha, I)$, since $\ell(\alpha, I)$ is usually not a principal ideal, there exist $f_1, \dots, f_t \in I$ with $\text{lm}(f_i) = X^\alpha, i = 1, \dots, t$ and $a_1, \dots, a_t \in R$ such that $r = a_1 \text{lc}(f_1) + \dots + a_t \text{lc}(f_t)$. From (3), for each f_i , we have $\text{lc}(f_i) = b_1 \text{lc}(g_{i1}) + \dots + b_s \text{lc}(g_{is})$, where $g_{ij} \in G, \text{lm}(g_{ij}) | \text{lm}(f_i), b_j \in R, j = 1, \dots, s$. Thus $r \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) | X^\alpha)$.

Conversely, let $a \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) | X^\alpha)$. Then $a = r_1 \text{lc}(g_1) + \dots + r_t \text{lc}(g_t)$ for some $r_i \in R, g_i \in G, \text{lm}(g_i) | X^\alpha, i = 1, \dots, t$. Choose $\{X^{\alpha_i}\}_1^t \in \mathcal{M}(X)$ such that $\text{lm}(X^{\alpha_i} g_i) = X^\alpha$ for $i = 1, \dots, t$. Note that $X^{\alpha_i} g_i \in I$, we get

$$\begin{aligned} a &= r_1 \text{lc}(g_1) + \dots + r_t \text{lc}(g_t) = r_1 \text{lc}(X^{\alpha_1} g_1) + \dots + r_t \text{lc}(X^{\alpha_t} g_t) \\ &\in \ell(\text{lc}(f) : f \in I, \text{lm}(f) = X^\alpha). \end{aligned}$$

(4) \Rightarrow (3): For any $0 \neq f \in I$, let $\text{lm}(f) = X^\alpha$. Then $\text{lc}(f) \in \ell(\alpha, I)$. By (4), $\text{lc}(f) \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) | X^\alpha)$. \square

Corollary 3.3.3.

- (1) If G is a Gröbner basis for the left ideal I in $R\langle x_1, \dots, x_n \rangle$, then $I = \ell(G)$.
- (2) If G is a Gröbner basis and $f \in \ell(G)$ and $f \rightarrow^G h$, where h is reduced, then $h = 0$.
- (3) Let G be the Gröbner basis of left ideal I and $f \in R\langle x_1, \dots, x_n \rangle$. Then $f \in I$ if and only if $f = 0$ modulo G .

Next we give a method to construct Gröbner bases in PBW extensions. We remind the reader that in general rings the S -pair criteria are not necessarily available. A simple example can be found in ([4], p.248).

Theorem 3.3.4. *Let I be the left ideal of $R\langle x_1, \dots, x_n \rangle$ generated by a finite set G . For $F := \{g_1, \dots, g_s\} \subseteq G$, let $\text{LCM}(F)$ be the least common multiple of $\{\text{lm}(g_i)\}_{i=1}^s$.*

Let \mathcal{B}_F be a finite set of generators of $\text{Syz}(\text{lc}(g_1), \dots, \text{lc}(g_s))$. Then the following assertions are equivalent:

(1) G is a Gröbner basis of I ;

(2) For all $F := \{g_i\}_1^s \subseteq G$ and for all $\{b_1, \dots, b_s\} \in \mathcal{B}_F$, we have that

$$\sum_{i=1}^s b_i(X^{\text{LCM}(F)} / \text{lm}(g_i))g_i \text{ reduces to zero with respect to } G.$$

Proof. (1) \Rightarrow (2) follows from $\sum_{i=1}^s b_i(X^{\text{LCM}(F)} / \text{lm}(g_i))g_i \in \ell(G)$.

(2) \Rightarrow (1): Let $f \in I$. By Theorem 3.3.2 we need to show: $\text{lc}(f) \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) | \text{lm}(f))$. Let $G := \{g_1, \dots, g_t\}$. Then $f \in I$ implies that there exists a representation $f = \sum_{i=1}^t f_i g_i$. Furthermore we may choose $\{f_i\}_1^t$ such that $\max_{\prec} \{\text{lm}(\text{lm}(f_i) \text{lm}(g_i))\}_{i=1}^t$ is minimal, say X^{α_0} . Let $F := \{g_i \in G | \text{lm}(\text{lm}(f_i) \text{lm}(g_i)) = X^{\alpha_0}\}$. Without loss of generality, we may assume that $F = \{g_1, \dots, g_s\}, 1 \leq s \leq t$. The following argument considers cases based on the leading monomial of f .

If $\text{lm}(f) = X^{\alpha_0}$, then $\text{lt}(f) = \sum_{i=1}^s \text{lt}(f_i g_i)$ and $\text{lc}(f) = \sum_{i=1}^s \text{lc}(f_i) \text{lc}(g_i) \in \ell(\text{lc}(g) : g \in F)$. Note that $g_i \in F$ implies $\text{lm}(f) = \text{lm}(\text{lm}(f_i) \text{lm}(g_i))$. Thus $\text{lm}(g_i) | \text{lm}(f)$. Therefore $\text{lc}(f) \in \ell(\text{lc}(g) : g \in G, \text{lm}(g) | \text{lm}(f))$.

If $\text{lm}(f) \prec X^{\alpha_0}$, then

$$\sum_{i=1}^s \text{lc}(f_i) \text{lc}(g_i) = 0 \text{ and } (\text{lc}(f_1), \dots, \text{lc}(f_s)) \in \text{Syz}(\text{lc}(g_1), \dots, \text{lc}(g_s)).$$

Let \mathcal{B}_F be a basis of $\text{Syz}(\text{lc}(g_1), \dots, \text{lc}(g_s))$, say,

$$\mathcal{B}_F := \{\vec{b}_1, \dots, \vec{b}_k\} := \{(b_{11}, \dots, b_{1s}), \dots, (b_{k1}, \dots, b_{ks})\}.$$

There exist $r_1, \dots, r_k \in R$ such that $(\text{lc}(f_1), \dots, \text{lc}(f_s)) = r_1 \vec{b}_1 + \dots + r_k \vec{b}_k = (r_1 b_{11} + \dots + r_k b_{k1}, \dots, r_1 b_{1s} + \dots + r_k b_{ks})$. That is, for $1 \leq i \leq s$, we have $\text{lc}(f_i) = r_1 b_{1i} + \dots + r_k b_{ki}$. Now

$$\begin{aligned}
f &= \sum_{i=1}^t f_i g_i = \sum_{i=1}^s f_i g_i + \sum_{i=s+1}^t f_i g_i \\
&= \sum_{i=1}^s (f_i - \sum_{j=1}^k r_j b_{ji} \text{lm}(f_i)) g_i + \sum_{i=1}^s \sum_{j=1}^k r_j b_{ji} \text{lm}(f_i) g_i + \sum_{i=s+1}^t f_i g_i.
\end{aligned}$$

For the first sum of equation (3.2), from $\text{lm}(f_i - \sum_{j=1}^k r_j b_{ji} \text{lm}(f_i)) \prec \text{lm}(f_i)$, we have that $\text{lm}(\text{lm}(f_i - \sum_{j=1}^k r_j b_{ji} \text{lm}(f_i)) \text{lm}(g_i)) \prec X^{\alpha_0}, i = 1, \dots, s$.

For the third sum of equation (3.2), by the definition of F , we have

$$\max\{\text{lm}(\text{lm}(f_i) \text{lm}(g_i))\}_{i=s+1}^t \prec X^{\alpha_0}.$$

In order to get the contradiction to the minimality of α_0 , we have to rewrite the second sum of equation (3.2). Note that for all $g_i \in F$, $\text{lm}(\text{lm}(f_i) \text{lm}(g_i)) = X^{\alpha_0}$.

Thus there is an element $0 \neq \gamma \in \mathbb{N}^n$ such that $\alpha_0 = \text{LCM}(F) + \gamma$. Then

$$\begin{aligned}
\sum_{i=1}^s \sum_{j=1}^k r_j b_{ji} \text{lm}(f_i) g_i &= \sum_{j=1}^k r_j X^\gamma \left(\sum_{i=1}^s b_{ji} \frac{X^{\text{LCM}(F)}}{\text{lm}(g_i)} g_i \right) \\
&\quad + \sum_{j=1}^k r_j \left(\sum_{i=1}^s (b_{ji} \text{lm}(f_i) - X^\gamma b_{ji} \frac{X^{\text{LCM}(F)}}{\text{lm}(g_i)}) g_i \right).
\end{aligned}$$

Clearly, the leading monomials of every product in the second sum are smaller than X^{α_0} . Thus we only need to consider the first sum. By the assumption, $\sum_{i=1}^s b_{ji} \frac{X^{\text{LCM}(F)}}{\text{lm}(g_i)} g_i$ reduces to zero modulo G . Then by the division rule there exist $h_1, \dots, h_t \in R\langle x_1, \dots, x_n \rangle$ such that $\sum_{i=1}^s b_{ji} \frac{X^{\text{LCM}(F)}}{\text{lm}(g_i)} g_i - \sum_{i=1}^t h_i g_i = 0$ and $\text{lm}(\text{lm}(h_i) \text{lm}(g_i)) \prec X^{\alpha_0 - \gamma}$. Therefore

$$\sum_{j=1}^k r_j X^\gamma \left(\sum_{i=1}^s b_{ji} \frac{X^{\text{LCM}(F)}}{\text{lm}(g_i)} g_i \right) = \sum_{j=1}^k r_j X^\gamma \left(\sum_{i=1}^t h_i g_i \right) = \sum_{j=1}^k \sum_{i=1}^t X^\gamma h_i g_i.$$

It is easy to see that for all i , $\text{lm}(\text{lm}(X^\gamma h_i) \text{lm}(g_i)) \prec X^{\alpha_0}$, a contradiction as required. \square

The above theorem suggests an algorithm to construct Gröbner bases:

Algorithm: GröbnerPBW

Input: $\blacktriangleright F = \{f_1, \dots, f_s\} \subseteq R\langle x_1, \dots, x_n \rangle$.

Output: $\blacktriangleright G = \{g_1, \dots, g_t\}$, a Gröbner basis for $\ell(f_1, \dots, f_s)$.

Initialization: $G := F, G' :=$ all subsets of F ;

While $G' \neq \emptyset$ Do

Choose $\emptyset \neq S \in G'$, say, $S := \{f_{i_1}, \dots, f_{i_k}\}$;

$G' := G' \setminus S$;

Compute \mathcal{B}_S , a generating set for $Syz(\text{lc}(f_{i_1}), \dots, \text{lc}(f_{i_k}))$ and $\text{LCM}(S)$;

For each $b := (b_{i_1}, \dots, b_{i_k}) \in \mathcal{B}_S$ Do

$b_{i_1}(X^{\text{LCM}(S)} / \text{lm}(f_{i_1}))f_{i_1} + \dots + b_{i_k}(X^{\text{LCM}(S)} / \text{lm}(f_{i_k}))f_{i_k} \xrightarrow{G} \psi$, ψ is reduced modulo G .

If $\psi \neq 0$, then

$G := G \cup \{\psi\}$;

$G' := \{S' \cup \{\psi\}\}$; add ψ to every nonempty subset S' .

End Do

End Do

By the noetherian properties of the ring R , we know that $R\langle x_1, \dots, x_n \rangle$ is also noetherian and we obtain the following proposition.

Corollary 3.3.5. *Let I be a non-zero left ideal of $R\langle x_1, \dots, x_n \rangle$. Then I has a finite Gröbner basis.*

If R has some special properties, we *can* define a notion of S -pairs. For example, when R is a field, we can define reduction and S -pairs in a similar way to the way they are defined in the commutative case, and easily prove that:

Corollary 3.3.6. *Assume that R is a field. Let I be a left ideal of $R\langle x_1, \dots, x_n \rangle$ generated by a finite set G . Then G is a Gröbner basis if and only if all S -pairs reduce to zero.*

For noncommutative PIDs, we can define S -pairs in a similar fashion to commutative PIDs, though the situation becomes much more complex. For example, in noncommutative PIDs and UFDs, factors are only unique with respect to some invariant factors and left greatest common divisors are may not same as right greatest common divisors. We will discuss these issues in a forthcoming paper.

Discussing Gröbner bases from the graded point of view has a long history; for example, Robbiano [19], and more recently Apel [3], consider Gröbner bases on general graded rings. In PBW extensions the graded lexicographic ordering is the most popular ordering. Let $T = R\langle x_1, \dots, x_n \rangle$ be a PBW extension and \prec a term ordering compatible with the graded lexicographic ordering. The associated graded ring $\text{gr } T$ is defined as $R[\bar{x}_1, \dots, \bar{x}_n]$ with $r\bar{x}_i = \bar{x}_i r$ and $\bar{x}_i \bar{x}_j = \bar{x}_j \bar{x}_i$ for all $r \in R$, that is, usual polynomial ring over R .

We set up a relation between the Gröbner bases of $R\langle x_1, \dots, x_n \rangle$ and its associated graded ring $\text{gr } R\langle x_1, \dots, x_n \rangle$. For an element $f \in R\langle x_1, \dots, x_n \rangle$, \bar{f} will be the image of f in $\text{gr } R\langle x_1, \dots, x_n \rangle$. Recall that the standard filtration of $R\langle x_1, \dots, x_n \rangle$ is $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n$ where \mathcal{F}_i is the R -subspace generated by all X^α with $|\alpha| \leq i$ and $\text{gr } R\langle x_1, \dots, x_n \rangle = \bigoplus \mathcal{F}_n / \mathcal{F}_{n-1} = R[\bar{x}_1, \dots, \bar{x}_n]$. For any left ideal I of $R\langle x_1, \dots, x_n \rangle$, there is a graded ideal $\text{gr } I$ of $\text{gr } R\langle x_1, \dots, x_n \rangle$ which is defined by setting $(\text{gr } I)_i = (I + \mathcal{F}_{i-1}) \cap \mathcal{F}_i / \mathcal{F}_{i-1} \simeq I \cap \mathcal{F}_i / I \cap \mathcal{F}_{i-1} \subset \mathcal{F}_i / \mathcal{F}_{i-1}$, and $\text{gr } I = \bigoplus_i (\text{gr } I)_i$.

Lemma 3.3.7. (*[16]*) *Let $T = R\langle x_1, \dots, x_n \rangle$ be a PBW extension. If R is a noetherian ring, then T and $\text{gr } T$ are noetherian rings.*

Theorem 3.3.8. *Let I be a left ideal of $T = R\langle x_1, \dots, x_n \rangle$. If $G = \{f_1, \dots, f_m\}$ is a Gröbner basis of I , then $\{\bar{f}_1, \dots, \bar{f}_m\}$ is a Gröbner basis of $\text{gr } I$. Conversely if $\{g_1, \dots, g_t\}$ is a Gröbner basis of $\text{gr } I$, then choose $\{f_1, \dots, f_t\} \subseteq I$ such that $\bar{f}_1 = g_1, \dots, \bar{f}_t = g_t$ and $\{f_1, \dots, f_t\}$ is a Gröbner basis of I .*

Proof. Let $G = \{f_1, \dots, f_m\}$ be a Gröbner bases of I . For a homogeneous element $g \in (\text{gr } I)_i$, there exists $f \in (I + \mathcal{F}_{i-1}) \cap \mathcal{F}_i/\mathcal{F}_{i-1}$ such that $\bar{f} = g$. Since $f \in I$, we assume that $f = h_1 f_1 + \dots + h_t f_t$ with $\text{lm}(\text{lm}(h_j) \text{lm}(f_j)) \prec \text{lm}(f), 1 \leq j \leq t$. Thus $g = \bar{f} = f + \mathcal{F}_{i-1} = h_1 f_1 + \dots + h_t f_t + \mathcal{F}_{i-1} = (h_1 f_1 + \mathcal{F}_{i-1}) + \dots + (h_t f_t + \mathcal{F}_{i-1}) = \bar{h}_1 \bar{f}_1 + \dots + \bar{h}_t \bar{f}_t$. Therefore $\{\bar{f}_1, \dots, \bar{f}_m\}$ is a Gröbner basis of $\text{gr } I$. We leave the remainder of the proof to the reader (note that S -pair criteria are usually not available in $\text{gr } T$). \square

3.4 Application to Moving Frames

Standard differential elimination algorithms are based on commuting derivations. In many applications it is natural to choose instead a basis for these derivations which is adapted to the geometry of the application (e.g., a basis which is invariant under some given symmetry group of the application).

In these cases, the given commuting frame of commuting partial differential operators $\frac{\partial}{\partial x_i}$ may not be well suited to the application.

One class of moving frames is moving frames of differential operators of the form

$$\Delta_i = \sum_i A_i^j(z) \frac{\partial}{\partial z^j}, \quad (3.4.1)$$

and these satisfy frame commutation relations of the form:

$$[\Delta_i, \Delta_j] = \sum_k \gamma_{ij}^k(z) \Delta_k . \quad (3.4.2)$$

A significant abstraction and generalization of these ideas was initiated by Cartan, and in recent times developed and applied by Fels and Olver [7].

A simple example of such an approach is using polar coordinates for cylindrically invariant problems (where the operators $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial r}$ commute). Given an arbitrary Lie group \mathcal{G} the power of the general method of moving frames is that, on some sufficiently prolonged space, a \mathcal{G} -invariant frame always exists.

A study of differential-elimination methods in moving frames of differential operators was given by Lisle in his PhD thesis [13]. However Lisle did not give a rigorous Gröbner basis theory for his approach. He was able, however, to do very complex *classification problems* which were beyond the power of differential elimination packages based on commuting derivations.

In this section we show that the Gröbner theory developed in this paper can be applied to moving frames of differential operators for systems of linear homogeneous partial differential equations.

We treat an illustrative example which arises from the group classification problem for the class of nonlinear diffusion equations of the form

$$u_t = (D(u)u_x)_x . \quad (3.4.3)$$

This example was used by Lisle and Reid [14] and Lisle [13] to illustrate Lisle's moving frame method. The method of group classification attempts, for every form of $D(u)$, to describe the symmetry properties of the above partial differential equation. This is easy for the illustrative example, but in general leads to intractable overdetermined

systems of partial differential equations for the symmetries when commuting derivations are used. The idea of Lisle's method, was to exploit equivalence transformations which mapped members of the class of partial differential equations to another member of the class. In the above case $u \mapsto au + b, x \mapsto cx + d, t \mapsto ex + f$, are simple examples of such transformations. Then the method constructs a moving frame of differential operators invariant under such an equivalence group.

One branch of the calculation, for the nonlinear diffusion equation, leads to the following system of partial differential equations in the *frame standard form* of Lisle and Reid [14]:

$$\begin{aligned} \Delta_1 \Delta_1 \theta^1 &= 0 & \Delta_1 \theta^2 &= 0 & \Delta_1 \theta^3 &= 0 \\ \Delta_2 \theta^1 &= 0 & \Delta_2 \theta^2 &= 2\Delta_1 \theta^1 - \theta^3 & \Delta_2 \theta^3 &= 0 \\ \Delta_3 \theta^1 &= -\frac{1}{2}\theta^1 & \Delta_3 \theta^2 &= 0 & \Delta_3 \theta^3 &= 0. \end{aligned}$$

This is the system of equations just after equation (18) of Lisle and Reid [14]. In this case the frame derivations $\Delta_j, j = 1, 2, 3$, have vanishing commutators except for

$$[\Delta_1, \Delta_3] = -\frac{1}{2}\Delta_1. \quad (3.4.4)$$

In terms of the original physical variables x, t, u , and the commuting coordinate frame $\frac{\partial}{\partial x}, \frac{\partial}{\partial t}, \frac{\partial}{\partial u}$, the frame derivations are given by

$$\Delta_1 := D^{1/2} \frac{\partial}{\partial x}, \quad \Delta_2 := \frac{\partial}{\partial t}, \quad \Delta_3 := D/\dot{D} \frac{\partial}{\partial u}, \quad (3.4.5)$$

and the dependent variables $\theta^1, \theta^2, \theta^3$, yield the infinitesimal symmetries via the relation (14) given in Lisle and Reid [14].

Notice that the theory of this paper, which is directed to linear homogeneous systems is not directly applicable, since the above system has 3 dependent variables.

To transform it to an equivalent system, with one dependent variable, we use the *Drach Transformation* which, although written for the commutative case, easily generalizes to the non-commutative case. Consider systems with n independent variables x_1, \dots, x_n (here $n = 3$) and m dependent variables (here $m = 3$). The Drach transformation proceeds by introducing m new independent variables $x_{n+j}, j = 1, \dots, m$ and is defined by:

$$\theta^j := \Delta_{n+j}w, j = 1, \dots, m, \quad \Delta_{n+j} := \frac{\partial}{\partial x_{n+j}} \quad (3.4.6)$$

together with the additional relations

$$\Delta_{n+j}\Delta_{n+k}w = 0, \quad 1 \leq j, k \leq m. \quad (3.4.7)$$

The only non-vanishing commutators remain as

$$[\Delta_i, \Delta_j] = \sum_k \gamma_{ij}^k(z)\Delta_k, \quad 1 \leq i, j \leq n. \quad (3.4.8)$$

Under this transformation our system becomes

$$\begin{array}{lll} \Delta_1\Delta_1\Delta_4w = 0 & \Delta_1\Delta_5w = 0 & \Delta_1\Delta_6w = 0 \\ \Delta_2\Delta_4w = 0 & \Delta_2\Delta_5w = 2\Delta_1\Delta_4w - \Delta_6w & \Delta_2\Delta_6w = 0 \\ \Delta_3\Delta_4w = -\frac{1}{2}\Delta_4w & \Delta_3\Delta_5w = 0 & \Delta_3\Delta_6w = 0, \end{array}$$

together with the extra relations

$$\Delta_{3+j}\Delta_{3+k}w = 0, 1 \leq j, k \leq 3, \quad (3.4.9)$$

and the single non-vanishing commutator amongst the $\Delta_1, \dots, \Delta_6$ remains as $[\Delta_1, \Delta_3] = -\frac{1}{2}\Delta_1$. Now the system for w is a system to which our PBW Gröbner methods can be applied.

Next we outline the main idea. Set the independent variables $x = \{x_1, \dots, x_m\}$, dependent variables $u = \{u_1, \dots, u_n\}$, derivatives $\{\Delta_1, \dots, \Delta_m\}$ and $\Delta = \{\Delta^\alpha u_i | \alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m, i \in \{1, \dots, n\}\}$. For $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$, let $|\alpha| = \alpha_1 + \dots + \alpha_m$. Now we define the ranking on Δ as following:

Without loss of generality, we may assume that $x_1 \prec x_2 \prec \dots \prec x_m$ and $u_1 \prec u_2 \prec \dots \prec u_n$. The *total degree ordering* on Δ is given by:

$$\begin{aligned} \Delta^\alpha u_i \prec \Delta^\beta u_j &\iff |\alpha| < |\beta|, \text{ or } |\alpha| = |\beta|, \text{ and } i < j, \text{ or} \\ &|\alpha| = |\beta|, \text{ } i = j \text{ and } \alpha_1 < \beta_1, \text{ or} \\ &|\alpha| = |\beta|, \text{ } i = j, \alpha_1 = \beta_1, \dots, \alpha_{k-1} = \beta_{k-1}, \\ &\text{and } \alpha_k < \beta_k \text{ for some } 2 \leq k \leq m - 1. \end{aligned}$$

By the commutation rule, we also can write $\Delta_j \Delta^\alpha u_i$ as a polynomial of standard monomials and u_i . Thus we can define $\text{HD} \Delta_j \Delta^\alpha u_i$ as the highest derivative (highest with respect to \prec).

Definition 3.4.1. A *positive ranking* \prec of Δ is a total degree ordering of Δ which is compatible with differentiation and well-ordering:

$$\Delta^\alpha u_i \prec \Delta^\beta u_j \implies \text{HD} \Delta^\gamma \Delta^\alpha u_i \prec \text{HD} \Delta^\gamma \Delta^\beta u_j \quad (3.4.10)$$

$$\text{HD} \Delta^\alpha u_i \prec \text{HD} \Delta^\gamma \Delta^\alpha u_i \text{ for } |\gamma| \neq 0. \quad (3.4.11)$$

It is easy to see that positive ranking is compatible with Definition 3.2.2 and that the Drach transformation keeps the positive ranking invariant. Let I be the left ideal generated by w -system. Then $f(u) = 0$ for all $f \in I$. This point of view enables us to study w -systems through Gröbner bases for left ideals in PBW extensions. Given fw, gw in w -system, the S -pair is defined to be $S(fw, gw) =$

$\text{LCM}(\text{lm}(f), \text{lm}(g)) / \text{lm}(f) \cdot fw - \text{lc}(f) / \text{lc}(g) \cdot \text{LCM}(\text{lm}(f), \text{lm}(g)) / \text{lm}(g) \cdot gw$. In particular if we assume that all S -pairs are reduced to zero in the original untransformed system, then it is easy to show that all the S -pairs in the w -system are reduced to zero and we can use Corollary 3.3.6 to construct Gröbner bases. Thus we have a Buchberger-like algorithm for moving frames.

The above example from Lisle and Reid [14] is fairly simple, to do, even in the original non-invariant commuting coordinate frame. However Lisle and Reid [14] apply their moving frame method to some highly non-trivial systems (also given earlier in Lisle's thesis [13]). These include group classification of a class of potential convection diffusion equation and also a large class of linear wave equations admitting an infinite equivalence group.

Lisle and Reid conjectured [14], but did not prove, that if their linear homogeneous frame systems had their S -pairs reduce to zero, then they would obtain a Gröbner basis. This conjecture is rigorously proved in the current paper.

Some of the linear homogeneous systems in [14] involve functions of the dependent variables (the *class variables*) in their coefficients. The sometimes nonlinear auxiliary relations satisfied by the class functions, need a separate treatment, and had to be checked on a case by case basis. A fully algorithmic approach involving the class functions in addition to the linear homogeneous frame systems, is an important open problem. Finally, we note that we are currently working on generalizing the treatment of this paper to PBW extensions over modules, and this would allow the example to be treated directly without using the indirect Drach Transformation.

Bibliography

- [1] W. W. Adams and P. Loustau, *An Introduction to Gröbner Bases*, Amer. Math. Soc. 1994.
- [2] W. Adams, P. Loustau and D. Struppa, *Applications of commutative and computational algebra to partial differential equations*, Proc. Adv. in Sci. Comp. and Modeling, S. Dey and J. Ziebarth eds., 153-157 (1996).
- [3] J. Apel, *Effective Gröbner structures*, Informatik Report 12, Institut für Informatik, Universität Leipzig, 1997.
- [4] J. Apel and W. Lassner, *An extension of Buchberger's Algorithm and calculations in enveloping fields of Lie algebras*, J. Symb. Comp., 6(1988), 361-370.
- [5] A. D. Bell and K. R. Goodearl, *Uniform rank over differential operator rings and Poincaré-Birkhoff-Witt extensions*, Pacific Journal of Mathematics, vol.131(1)(1988), 13-37.
- [6] F. Chyzak and B. Salvy, *Non-commutative elimination in Ore algebras proves multivariate identities*, J. Symb. Comp. 26(1998), 187-227.
- [7] M. Fels and P.J. Olver, *Moving coframes. II. Regularization and theoretical foundations*, Acta Appl. Math. 55 (1999) 127-208.

- [8] A. Galligo, *Some algorithmic questions on ideals of differential operators*, Proc. EUROCAL'85, Springer LNCS 204, 413-421.
- [9] P. Gianni, B. Trager and G. Zacharias, *Gröbner bases and primary decomposition of polynomial ideals*, J. Symb. Comp. 6(1988), 149-167.
- [10] E. Green, *An introduction to noncommutative Gröbner bases*, In: Fisher K. G.(ed.), Computational Algebra, Dekker, New York.(Lecture Notes in Pure and Applied Mathematics 151): 167-190.
- [11] M. Insa and F. Pauer, *Gröbner bases in rings of differential operators*, In B. Buchberger and F. Winkler, editors, *Gröbner Bases and applications*, vol. 251, LMS Lec. Notes Series, 367-381, Cambridge University Press, 1998.
- [12] A. Kandri-Rody and V. Weispfenning, *Non-commutative Gröbner bases in algebras of solvable type*, J. Symb. Comp. vol. 9(1990), 1-26.
- [13] I. G. Lisle, *Equivalence Transformations for Classes of Differential Equations*, PhD thesis, Univ. of British Columbia, 1992.
- [14] I. G. Lisle and G. J. Reid, *Symmetry classification using invariant moving frame*, Ontario Research Centre for Computer Algebra, Technical Report TR-00-08, 2000, at <http://www.orcca.on.ca/TechReports>,
- [15] E. L. Mansfield, *Algorithms for symmetric differential systems*, Foundations of Computational Math. (2001) 1:335-383.
- [16] J. C. McConnell and J. C. Robson, *Non-commutative Noetherian Rings*, Wiley 1987.

- [17] T. Mora, *An introduction to commutative and noncommutative Gröbner bases* ,
Theor. Comp. Sci., 134: 131-173, 1994.
- [18] P.J. Olver, *Geometric foundations of numerical algorithms and symmetry*, Appl.
Alg. Engin. Comp. Commun. 11 (2001) 417-436.
- [19] L. Robbiano, *On the theory of graded structures*, J. Symb. Comp. 2(1986), 139-
170.

Chapter 4

Non-commutative Riquier Theory in Moving Frames of Differential Operators

Symbolic manipulation algorithms which apply a finite number of differentiations and eliminations to over-determined systems of partial differential equations (PDE) to yield them in forms suitable for the subsequent application of analytic and numeric solution techniques have recently attracted attention. Such algorithms usually should have an existence and uniqueness theorem for their output.

Such differential elimination algorithms, which can be regarded as a differential generalization of the method of Gröbner Bases (an algorithmic method for systems of polynomials), although enjoying success, can become intractable due to expression explosion on many problems. One symbolic approach based on extending the ancient

idea of choosing appropriate coordinates is to make the steps of the differential elimination methods share geometric features of the problems to which they are applied. Such features include the invariance under a known Lie group and often require the expression of the given PDE in terms of a set of non-commuting Partial Differential Operators (PDO) invariant under the given group.

Generalizing the commutative differential elimination algorithms to the case of non-commutative PDO, requires the construction of an existence and uniqueness theory for systems of PDE expressed in terms of such PDO.

In this article such an existence and uniqueness theory is given for systems of analytic PDE. It can be algorithmically applied to polynomially non-linear PDE. The main idea for the theoretical development is to use the commutation relations between the PDO to place them in a standard order. This normalization is exploited to generalize to the non-commutative case the corresponding steps of a previously developed commutative theory (by Rust, Reid and Wittkopf). This theory is applied to problems from Lisle's methods for examining equivalence and symmetry properties of families of PDE, and potentially can be applied to problems arising in the method of moving frames, as developed in recently by Olver, Fels and others.

4.1 Introduction

We consider systems of partial differential equations (PDE) with independent variables $x = (x_1, x_2, \dots, x_m) \in \mathbb{F}^m$ and dependent variables $u = (u^1, \dots, u^n) \in \mathbb{F}^n$ where $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . We restrict our attention to (mainly over-determined) systems which are \mathbb{F} -analytic functions of their independent, dependent variables and Partial Differential

Operators (PDO) applied to u where the PDO have the form:

$$\tilde{\partial}_i = \sum_{j=1}^m a_{ij}(x, u) \frac{\partial}{\partial x_j}, \quad i = 1, \dots, m, \quad (4.1.1)$$

Here the $a_{ij}(x, u)$ are \mathbb{F} -analytic functions and the matrix $a_{ij}(x, u)$ is invertible (i.e. $\det(a_{ij}(x, u))_{m \times m} \neq 0$) for (x, u) in some connected open subset V of $\mathbb{F}^m \times \mathbb{F}^n$. Such over-determined systems arise in applications motivating the need for the existence and uniqueness theory developed in this paper.

An easy computation shows that the $\tilde{\partial}_i$ satisfy commutation relations of the form:

$$[\tilde{\partial}_i, \tilde{\partial}_j] = \sum_k \gamma_{ij}^k \tilde{\partial}_k, \quad 1 \leq i \leq j \leq m, 1 \leq k \leq m, \quad (4.1.2)$$

where the γ_{ij}^k are so-called structure functions of x, u and first order derivatives of u . Thus the $\tilde{\partial}_i$ are generally non-commutative in comparison to the usual commuting partial derivatives $\frac{\partial}{\partial x_j}$ (which will be abbreviated as ∂_{x_j} or ∂_j).

By inversion of the relation (4.1.1) to solve for the $\frac{\partial}{\partial x_j}$ any analytic PDE system written in terms x, u and ∂_i can be written in terms of x, u and $\tilde{\partial}_i$.

It is not immediately clear why one would want to give up commutativity to express PDE in terms of non-commuting operators. The motivation is that such a non-commuting set may enjoy properties not shared by the standard commuting frame. For example such properties might be geometrical properties such as invariance under a certain Lie group \mathcal{G} . A special case is that of using polar coordinates for cylindrically invariant problems (where the operators are $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial r}$ and in fact commute).

This process of choosing appropriate coordinates to avoid unnecessarily complicated expressions, has a long history. Given a \mathcal{G} -invariant problem, however, it is not possible to always choose \mathcal{G} -invariant coordinates. A classic example of a set of invariant PDO which does not yield a global coordinate system is the existence of

such a set of PDO on the Torus. Cartan, with his method of moving frames, found a significant and far-reaching generalization of such ideas (see [1, Chapt 5] for historical remarks and also the foundational work of Tresse [34]). More recent works include those of Griffiths [11], and a new more general constructive approach was given to Moving Frame Theory by Fels and Olver [7, 8]. Given an arbitrary Lie group \mathcal{G} the power of the general method of moving frames is that under fairly weak conditions, on some sufficiently prolonged space, a \mathcal{G} -invariant frame exists. Invariant sets of PDO are just one aspect of this theory.

A major motivation for our work was provided by the work of Lisle [17] (also see [18]). That work concerned the computation and exploitation of Lie symmetries of classes of differential equations. For example, in modelling diffusion, one may be interested in classes of nonlinear diffusion equations of the form

$$u_t = (K(u)u_x)_x, \quad (4.1.3)$$

where the diffusion is assumed to be nonlinear ($K_u(u) = \dot{K}(u) \neq 0$). A common objective is to determine functional forms of the diffusion coefficient $K(u)$, capable of modelling physically important diffusion processes, for which exact solutions of the diffusion PDE can be found. Lie group classification methods can in theory determine the $K(u)$ for which such nonlinear diffusion PDE have large symmetry groups, and give procedures for identifying corresponding classes of exact solutions.

Algorithms [27, 28] based on commuting partial derivatives, exist for identifying the size and structure of the symmetry groups of classes of PDE such as (4.1.3). Computer implementations of the above algorithms using commuting partial derivatives rely on differential elimination packages such as the *RifSimp*, *Diffalg* and *Diffgrob*

packages in Maple. These packages manipulate the defining equations for infinitesimal Lie symmetries of the physical PDE of interest. These defining equations are overdetermined linear homogeneous PDE with coefficients depending on the so-called classification functions (e.g. the $K(u)$ in the PDE above). We direct the reader to the review article of Hereman on symbolic packages for differential equations [13]. These differential generalizations of Gröbner Bases [3], when applied to such systems, typically build up coefficients involving derivatives of the classification functions. These coefficients can become so large and complicated [17], that they can fail to terminate in the available time and memory. This problem persists today, despite considerable progress in both computer speed/memory and improvements in differential elimination algorithms based on commuting PDO.

The idea of Lisle's method [17, 18] to address the expression explosion often encountered in such classification problems, was to exploit easily determined transformations that mapped one member of such a class to another member. Such transformations form what is called an *equivalence group*. For example it is easily seen that the class of transformations:

$$x = \beta x', \quad t = t', \quad u = \gamma u' + \alpha, \quad \beta, \gamma \neq 0 \quad (4.1.4)$$

map the diffusion equation to $\gamma u'_t = \gamma \beta^{-2} (K(\gamma u' + \alpha) u'_{x'})_{x'}$. Hence the coefficient $K(u)$ is mapped to a new coefficient, given by $K'(u) = \beta^{-2} K(\gamma u' + \alpha)$. Symmetries are self-equivalences, which map a given member (e.g. here a nonlinear diffusion equation) to itself.

Lisle [17, 18] gives a method to easily determine a subclass of equivalence transformations (e.g. such as those above). His method then exploits the equivalence transformations to ease the more difficult problem of finding the full symmetry group

for each member of the class. This is done by recasting the equations for symmetries in a form which is invariant under the equivalence group. He was able to complete group classification problems, which could not be done by computer algebra methods based on commuting derivations. For these and other non-trivial examples, the reader is directed to [17, 18].

Another motivation for our work, is the revitalized interest in Cartan’s method of moving frames, its applications and generalizations (see [22] and the review paper [24]). Applications include: various forms of equivalence problem such as deciding when two objects are equivalent [4] under the projective group (a fundamental problem in computer vision), and deciding when two differential equations are equivalent by a change of variables. The design of group invariant numerical methods is also an important application which falls under the new area of geometric integration [12].

In his review Olver [24, page 2-3] states that “... any serious application ... will rely on computer algebra”, and further that “large scale applications ... will require the development of a suitable noncommutative Gröbner basis theory for such algebras, complicated by the non-commutativity of the invariant differential operators ...”.

In this paper we give existence and uniqueness theorems for systems of analytic PDE in a certain form with respect to moving frames of differential operators. This analytic non-commutative Gröbner-style theory is a partial answer to Olver’s open problem stated above. It allows nonlinearity which is not present in the linear differential-algebraic theory we presented in [9]. That linear theory did however allow the coefficient rings to be noncommutative which is relevant in non-commutative physical field theories having for example, non-commutative matrix coefficients.

We briefly discuss the dichotomy between such analytic and differential algebraic

approaches. Rust [31, 32] has given a Gröbner style development of Riquier Theory and generalized this to the nonlinear case. This work has helped bring analytic differential elimination methods (in the spirit of Riquier) and differential-algebraic approaches (as initiated by Ritt and Kolchin) closer together. Still, neither theory strictly contains the other. Specializing analytic functions to polynomials, does not yield all the results in differential algebra. Conversely the setting of Differential Algebra at this time, is too narrow to yield the full generality of the analytic approaches. Joint work with Hubert is ongoing to try to bring both approaches into a common theoretical setting. For the moment parallel developments seem necessary.

The main idea of our theory is to use the commutation relations to put the operators in a normal order modulo lower order terms. Then a non-commutative theory is built by mimicking the commutative theory across leading order derivatives (the commutative theory of Rust [32]). In particular we exploit a bijection between the commuting partial derivatives and the non-commutative differential operators $\tilde{\partial}_i$ (see Lisle and Reid (2000) [18, Appendix A] in which the treatment detailed in the current article is first sketched). We give a rigorous foundation and justification for the group classification-equivalence method of Lisle [17, 18]. Our results are not quite as general as those that would be required for a complete treatment of Olver's open problem, but are at the same time applicable to frames of operators enjoying geometric features other than \mathcal{G} -invariance. In particular a moving frame in Olver's approach is a \mathcal{G} -invariant map from a manifold to a group. The difficulty of establishing a rigorous noncommutative Gröbner basis theory for the moving frames case has become apparent since the seminal work of Mansfield [19], which produces interesting results, but similarly to the less ambitious work of Lisle, lacks an existence and uniqueness

theorem.

4.2 An Example - the Nonlinear Diffusion Equation

As a running example we treat the group classification problem for the nonlinear diffusion equation (4.1.3). That problem is to identify for all possible functional forms of the diffusion coefficient and the corresponding Lie symmetry algebras of vector fields, $\mathbf{X} = \xi(x, t, u)\partial_x + \tau(x, t, u)\partial_t + \eta(x, t, u)\partial_u$ leaving invariant (4.1.3).

The components ξ, τ, η of the symmetry vector field obey defining equations which can be automatically produced by many computer algebra packages

$$\tau_x = \tau_u = \xi_u = \eta_{uu} = 0 \quad (4.2.1a)$$

$$K(2\xi_x - \tau_t) - \dot{K}\eta = 0 \quad (4.2.1b)$$

$$K(2\eta_{xu} - \xi_{xx}) + 2\dot{K}\eta_x + \xi_t = 0 \quad (4.2.1c)$$

$$K\eta_{xx} - \eta_t = 0. \quad (4.2.1d)$$

For a given $K(u)$, this is an overdetermined linear homogeneous system. This system is simple enough to have all of its cases analyzed using differential elimination packages based on commuting PDO and is used for purposes of illustration (see [25] for the first time that this PDE was group classified).

We seek to write the defining equations (4.2.1) for its symmetries in a form invariant under the action of the equivalence group (4.1.4).

Following the method of Lisle [17] leads to the following frame:

$$\tilde{\partial}_1 := K^{1/2}\partial_x, \quad \tilde{\partial}_2 := \partial_t, \quad \tilde{\partial}_3 := K/\dot{K}\partial_u. \quad (4.2.2)$$

The reader can verify that this frame of PDO is invariant under the equivalence group. For example $\tilde{\partial}_1 = K^{1/2}\partial_x = \beta(K')^{1/2}\frac{1}{\beta}\partial_{x'} = (K')^{1/2}\partial_{x'}$. Lisle's method also requires introducing new infinitesimals defined by $\theta^1\tilde{\partial}_1 + \theta^2\tilde{\partial}_2 + \theta^3\tilde{\partial}_3 = \xi\partial_x + \tau\partial_t + \eta\partial_u$ yielding

$$\theta^1 := K^{-1/2}\xi, \quad \theta^2 := \tau, \quad \theta^3 := \dot{K}/K\eta \quad (4.2.3)$$

which the reader can verify are invariant. Lisle's method also yields the scalar equivalence group invariant

$$J := \frac{K\ddot{K}}{\dot{K}^2} - 1, \quad \tilde{\partial}_1 J = 0, \quad \tilde{\partial}_2 J = 0 \quad (4.2.4)$$

Computation of the structure relations for the frame by making the replacements (4.2.2) yields

$$[\tilde{\partial}_1, \tilde{\partial}_3] = -\frac{1}{2}\tilde{\partial}_1, \quad [\tilde{\partial}_1, \tilde{\partial}_2] = 0, \quad [\tilde{\partial}_2, \tilde{\partial}_3] = 0. \quad (4.2.5)$$

The defining system (4.2.1) becomes

$$\begin{aligned} \tilde{\partial}_3\theta^1 + \frac{1}{2}\theta^1 = 0 & \quad \tilde{\partial}_1\theta^2 = 0 & \quad \tilde{\partial}_1\tilde{\partial}_1\theta^3 - \tilde{\partial}_2\theta^3 = 0 \\ \tilde{\partial}_2\theta^2 - 2\tilde{\partial}_1\theta^1 + \theta^3 = 0 & \quad \tilde{\partial}_1\tilde{\partial}_3\theta^3 - \frac{1}{2}\tilde{\partial}_1\tilde{\partial}_1\theta^1 - (J-1)\tilde{\partial}_1\theta^3 + \frac{1}{2}\tilde{\partial}_2\theta^1 = 0 \\ \tilde{\partial}_3\theta^2 = 0 & \quad \tilde{\partial}_3\tilde{\partial}_3\theta^3 - J\tilde{\partial}_3\theta^3 - \tilde{\partial}_3J\theta^3 = 0 \end{aligned} \quad (4.2.6)$$

For example $\xi_u = 0$ implies that $\left(\frac{\dot{K}}{K}\tilde{\partial}_3\right)(K^{1/2}\theta^1) = 0$ and that $\tilde{\partial}_3\theta^1 + \frac{1}{2}\theta^1 = 0$ by using $\tilde{\partial}_3K = K$.

Our aim with the above system was not to give a detailed explanation of how Lisle's method works (which is described elsewhere [17, 18]), but instead to give the reader some insight, on the origin of such systems written in terms of non-commuting PDO.

The goal of the rest of the paper is to develop an existence and uniqueness theory for analytic systems such as (4.2.6) which are expressed in terms of non-commuting PDO.

4.3 Derivations

Let \mathbb{F} be a field (\mathbb{R} or \mathbb{C} in this paper) with characteristic zero, $x = (x_1, \dots, x_m)$ be the independent variables and $u = (u^1, \dots, u^n)$ be the dependant variables for a system of PDE.

In the usual commutative approaches to differential algebra and differential elimination theory [31, 2], a set of indeterminates corresponding to the partial derivatives is defined:

$$\Omega = \{v_\alpha^i \mid \alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m, i = 1, \dots, n\}.$$

Each indeterminate of Ω corresponds to a partial derivative by:

$$v_\alpha^i \leftrightarrow (\partial_m)^{\alpha_m} \dots (\partial_1)^{\alpha_1} u^i(x_1, \dots, x_m) := \partial^\alpha u^i(x_1, \dots, x_m).$$

As usual the commutative total derivative operators are then introduced to act on members of Ω by a unit increment of the i -th index of their vector subscript:

$$D_i v_\alpha^k := v_{\beta}^k.$$

where $\beta = (\alpha_1, \dots, \alpha_i + 1, \dots, \alpha_m)$. The usual (commutative) total derivative $D_{x_i} \equiv D_i$ action on functions of $\{x\} \cup \Omega$ is then given by:

$$D_i = \partial_i + \sum_{v \in \Omega} (D_i v) \frac{\partial}{\partial v}. \quad (4.3.1)$$

The corresponding construction for the non-commutative case is as follows.

We suppose that there are m derivations $\tilde{\partial}_1, \dots, \tilde{\partial}_m$ which act on formal power series in the x_i with coefficients in \mathbb{F} . The derivation operators do not necessarily commute, that is, $\tilde{\partial}_i \tilde{\partial}_j \neq \tilde{\partial}_j \tilde{\partial}_i$ (e.g. see (4.2.5)).

Theorem 4.3.1. *Since the $\tilde{\partial}_i$ are derivations, they are of the form:*

$$\tilde{\partial}_i = \sum_{j=1}^m a_{ij} \partial_j \text{ where } a_{ij} = \tilde{\partial}_i(x_j). \quad (4.3.2)$$

Proof. A derivation $\tilde{\partial}_i$ satisfies $\tilde{\partial}_i(f+g) = \tilde{\partial}_i(f) + \tilde{\partial}_i(g)$ and $\tilde{\partial}_i(fg) = \tilde{\partial}_i(f)g + f\tilde{\partial}_i(g)$ where f and g are any power series. Using those two properties, a derivation on the formal power series is uniquely defined by its action on the variables x_k . Since $\sum_{j=1}^m a_{ij} \partial_j$ is a derivation (a linear combination of derivations is a derivation) and satisfies $\sum_{j=1}^m a_{ij} \partial_j(x_k) = a_{ik} = \tilde{\partial}_i(x_k)$, equation (4.3.2) follows. \square

Consider the set of indeterminates

$$\tilde{\Omega} = \{\tilde{v}_\alpha^i \mid \alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m, i = 1, \dots, n\}.$$

Each indeterminate of this set corresponds to a derivation by:

$$\tilde{v}_\alpha^i \leftrightarrow (\tilde{\partial}_m)^{\alpha_m} \dots (\tilde{\partial}_1)^{\alpha_1} u^i(x_1, \dots, x_n) := \tilde{\partial}^\alpha u^i(x_1, \dots, x_n).$$

In contrast to the commutative case this correspondence only gives a subset of the set of all derivations. However the commutation relations will enable us to extend this correspondence to the whole set.

Note that the full set of derivations of dependant variables of order r contains nm^r members which is far greater than the corresponding number of r -order derivations of the form above (which is $n \binom{r+m-1}{r}$).

To be able to apply a reduction process (described in section 4.5), and prove uniqueness and existence in Theorem 4.7.2, we impose:

Blanket Hypothesis 4.3.2 (Analyticity-Invertibility Assumption). *Throughout this paper we assume that the $a_{ij}(x, u)$ in (4.3.2) (also see (4.1.1)) are \mathbb{F} -analytic functions and the matrix $a_{ij}(x, u)$ is invertible (i.e. $\det(a_{ij}(x, u))_{m \times m} \neq 0$) for (x, u) in some connected open subset V of $\mathbb{F}^m \times \mathbb{F}^n$.*

Proposition 4.3.3. *With Hypothesis 4.3.2, we have the following commutation rules:*

$$\tilde{\partial}_i \tilde{\partial}_j - \tilde{\partial}_j \tilde{\partial}_i = \sum_{k=1}^m b_{ij}^k \tilde{\partial}_k \quad (4.3.3)$$

where the b_{ij}^k are analytic functions of x, u and first derivatives of u with values in \mathbb{F} .

Proof. By replacing in $\tilde{\partial}_i \tilde{\partial}_j - \tilde{\partial}_j \tilde{\partial}_i$ the expressions $\tilde{\partial}_i$ and $\tilde{\partial}_j$ given by (4.3.2), we get a linear combination of the ∂_i (order 2 derivations are cancelled). By inverting the matrix $(a_{ij}(x, u))$, each ∂_k is itself a linear combination of the $\tilde{\partial}_k$'s. Thus $\tilde{\partial}_i \tilde{\partial}_j - \tilde{\partial}_j \tilde{\partial}_i$ is a linear combination of $\tilde{\partial}_k$'s. \square

Nontrivial examples of moving frames of PDO can be found in Lisle and Reid [18], Mansfield [19] and Spivak [33].

From (4.3.2) it is natural to define the (non-commuting) formal total derivation by:

$$\tilde{D}_i = \sum_{j=1}^m a_{ij}(x, u) D_j \quad (4.3.4)$$

By the commutation rule (4.3.3), any $\tilde{D}_j \tilde{v}$ can be rewritten (normalized) as a function of $\{x\} \cup \tilde{\Omega}$. Assuming this normalization gives as a consequence of (4.3.1), (4.3.4)

$$\tilde{D}_i = \tilde{\partial}_i + \sum_{\tilde{v} \in \tilde{\Omega}} (\tilde{D}_i \tilde{v}) \frac{\partial}{\partial \tilde{v}} \quad (4.3.5)$$

on functions of $\{x\} \cup \tilde{\Omega}$. As a consequence we can now extend our normalization process to functions of $\{x\} \cup \tilde{\Omega}$.

Blanket Hypothesis 4.3.4 (Normalization Assumption for Derivations). *In this article, each time a derivation is applied to a function of $\{x\} \cup \tilde{\Omega}$ we assume that the commutation rules are applied to get an expression only involving elements of $\tilde{\Omega}$.*

For example, in system (4.2.6) all of the derivations are in $\tilde{\Omega}$ except for $\tilde{\partial}_1 \tilde{\partial}_3 \theta^3$. So using the commutation relation $[\tilde{\partial}_1, \tilde{\partial}_3] = -\frac{1}{2} \tilde{\partial}_1$ in (4.2.5) we can replace $\tilde{\partial}_1 \tilde{\partial}_3 \theta^3$ in (4.2.6) by $\tilde{\partial}_3 \tilde{\partial}_1 \theta^3 - \frac{1}{2} \tilde{\partial}_1 \theta^3$ so that the fifth equation of (4.2.6) is replaced with

$$\tilde{\partial}_3 \tilde{\partial}_1 \theta^3 - \frac{1}{2} \tilde{\partial}_1 \tilde{\partial}_1 \theta^1 - (J - \frac{1}{2}) \tilde{\partial}_1 \theta^3 + \frac{1}{2} \tilde{\partial}_2 \theta^1 = 0 . \quad (4.3.6)$$

Denoting $\tilde{\partial}^\alpha = \tilde{\partial}_m^{\alpha_m} \cdots \tilde{\partial}_1^{\alpha_1}$ and $\partial^\alpha = \partial_m^{\alpha_m} \cdots \partial_1^{\alpha_1}$ where $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$, we have the following property:

Theorem 4.3.5 (Bijection between derivations and partial derivatives). *Each order q derivation operator (resp. partial differential operator) can be expressed as an (x, u) -linear-combination of partial differential operators (resp. derivation operators) of order q or less.*

Proof. Using relation (4.3.2), any derivation monomial $\tilde{\partial}^\alpha$ can be rewritten as an (x, u) -linear combination of ∂^α . Conversely, any derivation monomial ∂^α can be rewritten as an (x, u) -linear combination of $\tilde{\partial}^\alpha$'s using Hypotheses 4.3.2 and 4.3.4. \square

4.4 Rankings

As with any Gröbner style theory, rankings play a fundamental role.

Suppose \prec is a total order on the set of (normalized) derivations $\tilde{\Omega}$. For an analytic function f of $\{x\} \cup \tilde{\Omega} = \{x_1, \dots, x_m\} \cup \tilde{\Omega}$ let $\text{HD}f$ denote the greatest derivation with respect the occurring in f . For $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$, let $|\alpha| = \alpha_1 + \dots + \alpha_m$.

Definition 4.4.1. A *positive ranking* \prec of $\tilde{\Omega}$ is a total ordering on $\tilde{\Omega}$ which is compatible with differentiation and well-ordering:

$$\tilde{v}_\alpha^i \prec \tilde{v}_\beta^j \Rightarrow \text{HD}\tilde{D}^\gamma \tilde{v}_\alpha^i \prec \text{HD}\tilde{D}^\gamma \tilde{v}_\beta^j \quad (4.4.1)$$

$$\tilde{v}_\alpha^i \prec \text{HD}\tilde{D}^\gamma \tilde{v}_\alpha^i \text{ for } |\gamma| \neq 0. \quad (4.4.2)$$

Throughout this paper a positive ranking \prec is fixed. Moreover, we suppose that \prec is *compatible with the total degree ordering* that is:

$$|\alpha| < |\beta| \implies \tilde{v}_\alpha^i \prec \tilde{v}_\beta^j \text{ for any } 1 \leq i, j \leq n \quad (4.4.3)$$

Thanks to conditions (4.4.3) and (4.3.3), we have the following property:

$$\text{If } \tilde{v}_\alpha^i \text{ is the highest derivative of } f, \text{ then } \text{HD}(\tilde{D}^\beta f) = \tilde{v}_{\alpha+\beta}^i \quad (4.4.4)$$

There obviously exist positive rankings satisfying (4.4.3) such as:

$$\begin{aligned} \tilde{v}_\alpha^i \prec \tilde{v}_\beta^j \iff & |\alpha| < |\beta|, \text{ or} \\ & |\alpha| = |\beta|, \text{ and } i < j, \text{ or} \\ & |\alpha| = |\beta|, i = j \text{ and } \alpha_1 < \beta_1, \text{ or} \\ & |\alpha| = |\beta|, i = j, \alpha_1 = \beta_1, \dots, \alpha_{k-1} = \beta_{k-1}, \\ & \text{and } \alpha_k < \beta_k \text{ for some } 2 \leq k \leq m-1. \end{aligned}$$

As a consequence on our example this ranking implies:

$$\begin{aligned} \theta^1 \prec \theta^2 \prec \theta^3 \prec \tilde{\partial}_1 \theta^1 \prec \tilde{\partial}_2 \theta^1 \prec \tilde{\partial}_3 \theta^1 \prec \tilde{\partial}_1 \theta^2 \prec \tilde{\partial}_2 \theta^2 \prec \tilde{\partial}_3 \theta^2 \prec \tilde{\partial}_1 \theta^3 \prec \tilde{\partial}_2 \theta^3 \prec \tilde{\partial}_3 \theta^3 \\ \prec \tilde{\partial}_1 \tilde{\partial}_1 \theta^1 \prec \tilde{\partial}_2 \tilde{\partial}_1 \theta^1 \prec \tilde{\partial}_2 \tilde{\partial}_2 \theta^1 \prec \dots \end{aligned}$$

According to this ranking, the highest derivation in each equation of (4.2.6) with its 5-th equation replaced with (4.3.6) can be determined. Solving each equation for its

highest derivative with respect to the above ranking yields the system:

$$\begin{aligned}
\tilde{\partial}_3\theta^1 &= -\frac{1}{2}\theta^1 & \tilde{\partial}_1\theta^2 &= 0 & \tilde{\partial}_1\tilde{\partial}_1\theta^3 &= \tilde{\partial}_2\theta^3 \\
\tilde{\partial}_2\theta^2 &= 2\tilde{\partial}_1\theta^1 - \theta^3 & \tilde{\partial}_3\tilde{\partial}_1\theta^3 &= \frac{1}{2}\tilde{\partial}_1\tilde{\partial}_1\theta^1 + (J - \frac{1}{2})\tilde{\partial}_1\theta^3 - \frac{1}{2}\tilde{\partial}_2\theta^1 \\
\tilde{\partial}_3\theta^2 &= 0 & \tilde{\partial}_3\tilde{\partial}_3\theta^3 &= J\tilde{\partial}_3\theta^3 + (\tilde{\partial}_3J)\theta^3
\end{aligned} \tag{4.4.5}$$

To check the conditions for our existence and uniqueness theorem for such systems in solved form, we need to determine if certain integrability conditions are satisfied, or reduced to zero modulo the system. Hence in the next section we define and study a suitable reduction process.

4.5 Reduction

Let f be an analytic function of $\{x\} \cup \tilde{\Omega}$. We say that f is \prec -*monic* if f has the form $f = \text{HD}f + g$, with $\text{HD}g \prec \text{HD}f$. For example the system (4.4.5) above is \prec -monic.

In the remainder of the paper, let a finite set \mathcal{M} of \prec -monic analytic functions of $\{x\} \cup \tilde{\Omega}$ be fixed. (Other restrictions will be made on \mathcal{M} in section 4.6).

For g, h two analytic functions of $\{x\} \cup \tilde{\Omega}$, we say that h is a *one step reduction* of g if there exist $f \in \mathcal{M}$ and $\alpha \in \mathbb{N}^m$ such that, with $\tilde{v}^* := \text{HD}\tilde{D}^\alpha f$, h can be given by substituting $\tilde{v}^* - \tilde{D}^\alpha f$ for \tilde{v}^* in g :

$$h = g(x, (\tilde{v})_{\tilde{v} \neq \tilde{v}^*}, (\tilde{v}^* - \tilde{D}^\alpha f)_{\tilde{v} = \tilde{v}^*}).$$

This is denoted $g \mapsto^{(\alpha, f)} h$, or simply $g \mapsto h$.

We say that g *reduces to* h if h can be obtained from g by a finite chain of one step reductions. That is, g reduces to h if there exists a positive integer k and k functions

h_1, \dots, h_k of $\{x\} \cup \tilde{\Omega}$ such that

$$g = h_1 \mapsto h_2 \mapsto \dots \mapsto h_k = h.$$

We write $g \mapsto^\mu h$ or $g \mapsto h$, where μ is of the form

$$\mu = ((\alpha_1, f_1), \dots, (\alpha_{k-1}, f_{k-1}))$$

with $h_i \mapsto^{(\alpha_i, f_i)} h_{i+1}$. We also write $h = \text{red}(g, \mu)$.

We say that g *completely reduces* to h if g reduces to h and h reduces to h' implies that $h = h'$.

Remark 4.5.1. The complete reduction may not be unique since may exist two different functions h and \bar{h} such that g completely reduces to both h and \bar{h} .

Example 4.5.1. *As a consequence of the system (4.4.5) the following integrability condition is satisfied $\tilde{D}_2(\tilde{\partial}_3\theta^2) - \tilde{D}_3(\tilde{\partial}_2\theta^2) = \tilde{D}_2(0) - \tilde{D}_3(2\tilde{\partial}_1\theta^1 - \theta^3)$. Normalization of this equation using commutation relations implies that $-\tilde{D}_3(2\tilde{\partial}_1\theta^1 - \theta^3) = 0$. Reduction of this last equation with respect to $\tilde{\partial}_3\theta^1 = -\frac{1}{2}\theta^1$ and use of the normalization yields $\tilde{\partial}_3\theta^3 = 0$. Using this relation to reduce $\tilde{\partial}_3\tilde{\partial}_3\theta^3 = J\tilde{\partial}_3\theta^3 + (\tilde{\partial}_3J)\theta^3$ yields $(\tilde{\partial}_3J)\theta^3 = 0$ and in summary we have obtained the equations*

$$\tilde{\partial}_3\theta^3 = 0, \quad (\tilde{\partial}_3J)\theta^3 = 0. \quad (4.5.1)$$

The ad hoc simplification achieved here is only given as an illustration of how reduction can be used to uncover hidden relations from a system. Determination of all the hidden relations, awaits the full development of the theory in the next few sections.

4.6 Parametric Derivations, Principal Derivations and Non-commutative Riquier Bases

Recall that \mathcal{M} is a finite set of \prec -monic analytic functions of $\{x\} \cup \tilde{\Omega}$. As usual all derivations are assumed to be normalized.

The *principal derivations* of \mathcal{M} are defined as

$$\text{Prin}\mathcal{M} := \{\tilde{v} \in \tilde{\Omega} \mid \text{there exist } f \in \mathcal{M} \text{ and } \alpha \in \mathbb{N}^m \text{ with } \tilde{v} = \text{HD}\tilde{D}^\alpha f\}$$

The *parametric derivations* of \mathcal{M} , which we denote $\text{Par}\mathcal{M}$, are those derivations that are not principal.

All leading derivations of elements in \mathcal{M} are in $\text{Prin}\mathcal{M}$, and using Normalization Hypothesis 4.3.4 it is easily shown that $\text{Prin}\mathcal{M}$ are elements of $\tilde{\Omega}$ which contain some highest derivation as a factor. Therefore a reduction h of g is a complete reduction if and only if h depends on $\{x\} \cup \text{Par}\mathcal{M}$ only.

In this paper, fix a non-empty open subset U of $\mathbb{F}^{\{x\} \cup \tilde{\Omega}}$. Moreover, we now assume that \mathcal{M} is a set of \prec -monic analytic functions which are polynomials in $\text{Prin}\mathcal{M}$.

Lemma 4.6.1. *Let $f, f' \in \mathcal{M}$ and g be an analytic function on U that is a polynomial in $\text{Prin}\mathcal{M}$. If there exist non-empty one step reductions: $h = \text{red}(g, (\alpha, f))$, $k = \text{red}(g, (\beta, f'))$, then:*

$$(1) \text{ if } \text{HD}\tilde{D}^\alpha f \prec \text{HD}\tilde{D}^\beta f' \text{ then } \text{red}(h, ((\beta, f'), (\alpha, f))) = \text{red}(k, (\alpha, f)).$$

$$(2) \text{ if } \tilde{D}^\alpha f - \tilde{D}^\beta f' \rightarrow^\mu 0 \text{ and } \text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^\beta f', \text{ then } \text{red}(h, \mu) = \text{red}(k, \mu)$$

In both cases, there exists an analytic function l such that $h \rightarrow l$ and $k \rightarrow l$.

Proof. Let $\tilde{v}^* = \text{HD}\tilde{D}^\alpha f$ and $\tilde{v}^{**} = \text{HD}\tilde{D}^\beta f'$. If $\tilde{v}^* \prec \tilde{v}^{**}$, we have

$$\begin{aligned} \text{red}(k, (\alpha, f)) &= g(x, (\tilde{v})_{\tilde{v} \neq \tilde{v}^*, \tilde{v}^{**}}, (\tilde{v}^* - \tilde{D}^\alpha f)_{\tilde{v} = \tilde{v}^*}, (\tilde{v}^{**} - \tilde{D}^\beta f'(x, (\tilde{v})_{\tilde{v} \neq \tilde{v}^*}, \\ &\quad (\tilde{v}^* - \tilde{D}^\alpha f)_{\tilde{v} = \tilde{v}^*})_{\tilde{v} = \tilde{v}^{**}})) \\ &= \text{red}(h, ((\beta, f'), (\alpha, f))) \end{aligned}$$

If $\tilde{v}^* = \tilde{v}^{**}$, then

$$\begin{aligned} \text{red}(h, \mu) &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v} - \tilde{D}^\alpha f, \mu))_{\tilde{v} = \tilde{v}^*}) \\ &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v}, \mu) - \text{red}(\tilde{D}^\alpha f, \mu))_{\tilde{v} = \tilde{v}^*}) \\ &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v}, \mu) - \text{red}(\tilde{D}^\beta f', \mu))_{\tilde{v} = \tilde{v}^*}) \\ &= \text{red}(k, \mu). \end{aligned}$$

□

Lemma 4.6.2 (Diamond Lemma). Fix $\tilde{v} \in \tilde{\Omega}$. Suppose that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \leq \tilde{v}$, we have $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$. Let g be an analytic function on U that is polynomial in $\text{Prin}\mathcal{M}$ with $\text{HD}g \leq \tilde{v}$, and two non-empty reductions $g \rightarrow h, g \rightarrow k$. Then there exists l with $h \rightarrow l$ and $k \rightarrow l$. In particular, g has a unique complete reduction.

Proof. The proof is very similar to the proof of the uniqueness of the normal form of a polynomial modulo a Gröbner basis (for example, see [3]) by using Lemma 4.6.2. Also see [31, page 67] where it is given in the commutative case. The full proof for the more general case can be found in Lemma 4.9.2.

□

The use of the bound \tilde{v} on the highest derivative is needed later in the proofs of Lemma 4.8.1 and Theorem 4.8.4.

Definition 4.6.1. \mathcal{M} is called a *non-commutative Riquier Basis* if for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f'$, the integrability condition $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$.

Theorem 4.6.3. *Suppose that \mathcal{M} is a non-commutative Riquier Basis and g is an analytic function on U that is polynomial in $\text{Prin}\mathcal{M}$. Then g has an unique complete reduction.*

Proof. It follows directly from the definition of Riquier basis, Lemmas 4.6.1 and 4.6.2. □

We denote the complete reduction of g by $\text{red}(g, \mathcal{M})$.

4.7 The Formal Non-commutative Riquier Existence Theorem

Let f be an \mathbb{F} -analytic function of $\{x\} \cup \tilde{\Omega}$, $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m$, and x^0 be a point in \mathbb{F}^m and let $u(x) = (u^1(x), \dots, u^n(x))$ be a vector of formal power series in $\mathbb{F}[[x - x^0]]^n$.

If f is defined at the point $(x^0, (\tilde{D}^\alpha u^i(x^0))_{\tilde{v}_\alpha^i \in \tilde{\Omega}})$, let $f[u](x)$ denote the formal power series at x^0 given by

$$f[u](x) := f(x, ((\tilde{D}^\alpha u^i(x))_{\tilde{v}_\alpha^i \in \tilde{\Omega}}).$$

where the subscript “ $\tilde{v}_\alpha^i \in \tilde{\Omega}$ ” indicates that $\tilde{D}^\alpha u^i(x)$ is to be substituted in the argument of f corresponding to \tilde{v}_α^i for each $\tilde{v}_\alpha^i \in \tilde{\Omega}$.

We illustrate these concepts with a simple example.

Example 4.7.1. Let $m = n = 1$ and $x^0 = 1$. Here u^1 , x_1 , $\tilde{\partial}_1$ and ∂_1 are simply denoted u , x , $\tilde{\partial}$ and ∂ . The relation (4.3.2) is simply denoted $\tilde{\partial} = a(x, u)\partial$.

Let $u(x)$ be the formal power series

$$u(x) = 1 + (x - 1) + 2!(x - 1)^2 + \cdots = \sum_{k=0}^{\infty} k!(x - 1)^k$$

and let $f = \ln(\tilde{v}_{(1)}^1)$ (recall that $\tilde{v}_{(1)}^1[u](x) = \tilde{\partial}u(x)$).

For $k \geq 1$, we have $\tilde{\partial}((x - 1)^k) = a(x)k(x - 1)^{k-1}$. Differentiating $u(x)$ term by term (which is the definition of the derivative of a formal power series) we obtain

$$\tilde{v}_{(1)}^1[u](x) = \tilde{\partial}(u(x)) = a(x) \sum_{k=0}^{\infty} (k + 1)(k + 1)!(x - 1)^k$$

Note that the $\ln(y)$ function is analytic at the constant term of the series $u(x)$, i.e. at the point $y = 1$. Thus, the series $f[u](x)$ is well defined and equals:

$$\begin{aligned} f[u](x) &= \ln(\tilde{v}_{(1)}^1[u](x)) \\ &= -\sum_{j=1}^{\infty} \frac{(-1)^j}{j} (\tilde{v}_{(1)}^1[u](x) - 1)^j \\ &= -a(x) \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \left(\sum_{k=1}^{\infty} (k + 1)(k + 1)!(x - 1)^k \right)^j \end{aligned}$$

Definition 4.7.1. We say that $u(x) \in \mathbb{F}[[x - x^0]]^n$ (for some $x^0 \in \mathbb{F}^m$) is a *formal power series solution* to a system of analytic PDE if $f[u](x)$ is well-defined and $f[u](x) = 0$ for all f in the system.

Suppose that $u(x) \in \mathbb{F}[[x - x^0]]^n$ is a formal power series solution to \mathcal{M} . Clearly, $\tilde{D}^\alpha f[u](x) = 0$ for all $\alpha \in \mathbb{N}^m$ and $f \in \mathcal{M}$. Therefore for g, h analytic, if h is a one step reduction of g then $h[u](x)$ is well-defined if and only if $g[u](x)$ is well-defined and in this case $g[u](x) = h[u](x)$. Furthermore, u is a formal power series solution to $\mathcal{M} \cup \{g\}$ iff u is a formal power series solution to $\mathcal{M} \cup \{h\}$.

Definition 4.7.2. A *specification of initial data* for \mathcal{M} is a map

$$\phi : \{x\} \cup \text{Par } \mathcal{M} \rightarrow \mathbb{F}$$

For $x^0 \in \mathbb{F}^m$, we say that ϕ is a specification at x^0 if

$$\phi(x) := (\phi(x_1), \phi(x_2), \dots, \phi(x_m)) = x^0.$$

For g a function of $\{x\} \cup \tilde{\Omega}$, let $\phi(g)$ be the function of the principal derivations obtained from g by evaluating x and the parametric derivations using ϕ :

$$\phi(g) := g(\phi(x), (\phi(\tilde{v}))_{\tilde{v} \in \text{Par } \mathcal{M}}).$$

For an illustration see Example 4.10.1.

Lemma 4.7.1 (Uniform Reduction). *let $\mathcal{G} := \{g_1, g_2, \dots, g_k\}$ be a finite set of functions of $\{x\} \cup \tilde{\Omega}$. Then there exists μ such that $\text{red}(g, \mu)$ is a complete reduction of g for all $g \in \mathcal{G}$. Moreover, if \mathcal{M} is a non-commutative Riquier Basis, then there exists μ such that $\text{red}(g, \mu) = \text{red}(g, \mathcal{M})$ for all $g \in \mathcal{G}$.*

Proof. By Dickson's lemma complete reductions always exist and we can choose μ_1 such that $\text{red}(g_1, \mu_1)$ is a complete reduction of g_1 . Recursively we could construct $\text{red}(g_j, (\mu_1, \mu_2, \dots, \mu_j))$ which for $j = 2, \dots, k$, is a complete reduction of $\text{red}(g_j, (\mu_1, \dots, \mu_{j-1}))$ and hence a complete reduction of g_j . Set $\mu = (\mu_1, \dots, \mu_k)$. Thus for $j \in \{1, \dots, k\}$ we have

$$\text{red}(g_j, \mu) = \text{red}(\text{red}(g_j, (\mu_1, \dots, \mu_j)), (\mu_{j+1}, \dots, \mu_k)) = \text{red}(g_j, (\mu_1, \dots, \mu_j)),$$

which is a complete reduction of g_j by construction. □

Theorem 4.7.2 (Formal Non-commutative Riquier Existence Theorem).

Let \mathcal{M} be a non-commutative Riquier Basis such that each $f \in \mathcal{M}$ is polynomial in the principal derivations (e.g. \mathcal{M} is a reduced non-commutative Riquier basis). For $x^0 \in \mathbb{F}^m$, let ϕ be a specification of initial data for \mathcal{M} at x^0 such that $\phi(f)$ is well-defined for all $f \in \mathcal{M}$. Then there is an unique formal power series solution $u(x) \in \mathbb{F}[[x - x^0]]^n$ to \mathcal{M} at x^0 such that $\tilde{D}^\alpha u^i(x^0) = \phi(\tilde{v}_\alpha^i)$ for all $\tilde{v}_\alpha^i \in \text{Par}\mathcal{M}$. Furthermore, every formal power series solution to \mathcal{M} at x^0 may be obtained in this way for some ϕ .

Proof. By the bijective correspondence of Theorem 4.3.5, there exists a n -tuple formal power series $u(x) \in \mathbb{F}[[x - x^0]]^n$ satisfying

$$\tilde{D}^\alpha u^i(x^0) := \phi(\text{red}(\tilde{v}_\alpha^i, \mathcal{M})) \quad (4.7.1)$$

for $i \in \{1, \dots, n\}$.

Since $u(x)$ must satisfy equation (4.7.1) for all $i \in \{1, \dots, n\}$ and $\alpha \in \mathbb{N}^m$ and since by Theorem 4.3.5 we have a bijection between derivations and partial derivatives the formal power series solution (if it exists) is unique.

We now prove that $u(x)$ is a formal power series solution of the system, which will prove the existence part of the theorem.

(1) We have first to check that $\phi(\text{red}(\tilde{v}_\alpha^i, \mathcal{M}))$ is well-defined.

Note that $\phi(\text{red}(\tilde{v}_\alpha^i, \mathcal{M}))$ depends only on the parametric derivations and so it is an element of \mathbb{F} , so long as it is well-defined.

(2) Then we have to verify that $u(x)$ is a formal power series solution to \mathcal{M} .

Clearly, $u(x)$ satisfies $\tilde{D}^\alpha u^i(x^0) = \phi(\tilde{v}_\alpha^i)$ for all $\tilde{v}_\alpha^i \in \text{Par}\mathcal{M}$. Now it suffices to verify that $\tilde{D}^\beta f[u](x^0) = 0$ for all $f \in \mathcal{M}$ and $\beta \in \mathbb{N}^m$. Hypothesis 4.3.2 will imply

$D^\beta f[u](x^0) = 0$ for all $f \in \mathcal{M}$ and $\beta \in \mathbb{N}^m$ and consequently that $f[u]$ is the zero formal power series.

Fix f, β . We have

$$\begin{aligned} \tilde{D}^\beta f[u](x^0) &= (\tilde{D}^\beta f)(x^0, (\tilde{D}^\beta u^i(x^0))) \\ &= (\tilde{D}^\beta f)(\phi(x^0), (\phi(\text{red}(\tilde{v}_\alpha^i, \mathcal{M})))) \\ &= \phi(\tilde{D}^\beta f(x, (\text{red}(\tilde{v}_\alpha^i, \mathcal{M})))) \end{aligned}$$

Let $\tilde{\Omega}'$ be the finite subset of $\tilde{\Omega}$ on which $\tilde{D}^\beta f$ depends. By Lemma 4.7.1 there exists μ such that for all $\tilde{v} \in \tilde{\Omega}'$

$$\text{red}(\tilde{v}_\alpha^i, \mathcal{M}) = \text{red}(\tilde{v}_\alpha^i, \mu).$$

Therefore

$$\begin{aligned} \tilde{D}^\beta f[u](x^0) &= \phi(\tilde{D}^\beta f(x, (\text{red}(\tilde{v}_\alpha^i, \mu)))) \\ &= \phi(\text{red}(\tilde{D}^\beta f(x, \tilde{v}_\alpha^i), \mu)). \end{aligned}$$

Note that $\text{red}(\tilde{D}^\beta f(x, \tilde{v}_\alpha^i), \mu)$ depends only on the parametric derivations and x .

Hence it is a complete reduction of $\tilde{D}^\beta f(x, \tilde{v}_\alpha^i)$ and we have

$$\begin{aligned} \text{red}(\tilde{D}^\beta f(x, \tilde{v}_\alpha^i), \mu) &= \text{red}(\tilde{D}^\beta f(x, \tilde{v}_\alpha^i), \mathcal{M}) \\ &= \text{red}(\tilde{D}^\beta f(x, \tilde{v}_\alpha^i), (\beta, f)) \\ &= 0 \end{aligned}$$

Thus $(\tilde{D}^\beta f[u])(x^0) = 0$ as required.

This completes the proof of existence part of the theorem.

□

4.8 Sufficient Finite Sets of Integrability Conditions

Note that the Formal Non-commutative Riquier Existence Theorem 4.7.2 requires the checking of infinitely many integrability conditions. In this section we show that only finitely many integrability conditions need to be checked.

Lemma 4.8.1 (Reduction of a sum). *Suppose h, k are polynomials in $\text{Prin}\mathcal{M}$. Suppose $h \rightarrow^\mu 0$ and $k \rightarrow^\nu 0$. Suppose further that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \leq \text{HD}k$, we have $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$. Then we have $h+k \rightarrow 0$.*

Proof. There are two cases:

(1) If $\text{red}(k, \mu)$ is an empty reduction, then

$$\begin{aligned} \text{red}(h+k, (\mu, \nu)) &= \text{red}(h, (\mu, \nu)) + \text{red}(k, (\mu, \nu)) \\ &= \text{red}(0, \nu) + \text{red}(k, \nu) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

(2) If $\text{red}(k, \mu)$ is a non-empty reduction, say $l = \text{red}(k, \mu)$, then by Lemma 4.6.2 there exist j, l with $0 \rightarrow j$ and $l \rightarrow j$. Since $0 \rightarrow j$, we have $j = 0$ and hence $l \rightarrow 0$, say $0 = \text{red}(l, \rho)$. Then we have:

$$\begin{aligned} \text{red}(h+k, (\mu, \rho)) &= \text{red}(h, (\mu, \rho)) + \text{red}(k, (\mu, \rho)) \\ &= \text{red}(0, \rho) + \text{red}(l, \rho) \\ &= 0. \end{aligned}$$

Therefore, $h+k \rightarrow 0$, as required.

□

Lemma 4.8.2 (Reduction of a Derivation). *Take $\alpha \in \mathbb{N}^m$, $f \in \mathcal{M}$, $i \in \{1, \dots, n\}$ and g an analytic function on U that is polynomial in $\text{Prin}\mathcal{M}$. Let $\tilde{v}^{**} = \text{HD}\tilde{D}^\alpha f$ and let \tilde{v}^* be given by $\tilde{v}^{**} = \text{HD}\tilde{D}_i\tilde{v}^*$, if this is well-defined.*

Then

$$\begin{aligned} \text{red}(\tilde{D}_i g, (\alpha, f)) &= \tilde{D}_i \text{red}(g, (\alpha, f)) + \text{red}\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i \tilde{D}^\alpha f \\ &\quad - \text{red}\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) \tilde{D}^\alpha f. \end{aligned}$$

If \tilde{v}^* is not well-defined, then the last term is omitted in the above formula.

Proof. By the definition of reduction we have:

$$\text{red}(g, (\alpha, f)) = g(x, (\tilde{v})_{\tilde{v} \neq \tilde{v}^{**}}, (\tilde{v}^{**} - \tilde{D}^\alpha f)_{\tilde{v} = \tilde{v}^{**}}). \quad (4.8.1)$$

Therefore, using equation (4.3.5) and using the property that the operations red commute with any ∂_i or $\frac{\partial}{\partial \tilde{v}}$ yields

$$\begin{aligned} \tilde{D}_i \text{red}(g, (\alpha, f)) &= \text{red}(\tilde{\partial}_i g, (\alpha, f)) + \sum_{\tilde{v} \neq \tilde{v}^{**}} \text{red}\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} \\ &\quad + \text{red}\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i (\tilde{v}^{**} - \tilde{D}^\alpha f). \end{aligned} \quad (4.8.2)$$

Also by the definition of total derivation we have

$$\tilde{D}_i g = \tilde{\partial}_i g + \sum_{\tilde{v} \in \tilde{\Omega}} \frac{\partial g}{\partial \tilde{v}} \tilde{D}_i \tilde{v}.$$

Thus

$$\begin{aligned} \text{red}(\tilde{D}_i g, (\alpha, f)) &= \text{red}(\tilde{\partial}_i g, (\alpha, f)) + \sum_{\tilde{v} \neq \tilde{v}^*} \text{red}\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} \\ &\quad + \text{red}\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) \text{red}(\tilde{D}_i \tilde{v}^*, (\alpha, f)) \end{aligned} \quad (4.8.3)$$

Solving (4.8.2) for $red(\tilde{\partial}_i g, (\alpha, f))$, and then eliminating this from (4.8.3) yields

$$\begin{aligned}
red(\tilde{D}_i g, (\alpha, f)) &= \tilde{D}_i red(g, (\alpha, f)) - red\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i(\tilde{v}^{**} - \tilde{D}^\alpha f) \\
&\quad - \sum_{\tilde{v} \neq \tilde{v}^{**}} red\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} \\
&\quad + \sum_{\tilde{v} \neq \tilde{v}^*} red\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} + red\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) red(\tilde{D}_i \tilde{v}^*, (\alpha, f)) \\
&= \tilde{D}_i red(g, (\alpha, f)) - red\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i(\tilde{v}^{**} - \tilde{D}^\alpha f) \\
&\quad - \sum_{\tilde{v} \neq \tilde{v}^{**}, \tilde{v}^*} red\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} - red\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) \tilde{D}_i \tilde{v}^* \\
&\quad + \sum_{\tilde{v} \neq \tilde{v}^*, \tilde{v}^{**}} red\left(\frac{\partial g}{\partial \tilde{v}}, (\alpha, f)\right) \tilde{D}_i \tilde{v} + red\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i \tilde{v}^{**} \\
&\quad + red\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) red(\tilde{D}_i \tilde{v}^*, (\alpha, f)) \\
&= \tilde{D}_i red(g, (\alpha, f)) + red\left(\frac{\partial g}{\partial \tilde{v}^{**}}, (\alpha, f)\right) \tilde{D}_i \tilde{D}^\alpha f + \\
&\quad red\left(\frac{\partial g}{\partial \tilde{v}^*}, (\alpha, f)\right) [red(\tilde{D}_i \tilde{v}^*, (\alpha, f)) - \tilde{D}_i \tilde{v}^*]
\end{aligned} \tag{4.8.4}$$

Since the term $\tilde{D}_i \tilde{v}^*$ is not normalized, we have to be careful before applying the red operation. We can write $\tilde{D}_i \tilde{v}^* = \tilde{v}^{**} + \sum_\nu a_\nu \tilde{v}_\nu$ where the sum is finite, the a_ν 's analytic functions and the \tilde{v}_ν belong to $\tilde{\Omega}$ and are different from \tilde{v}^{**} . Thus we have

$$\begin{aligned}
red(\tilde{D}_i \tilde{v}^*, (\alpha, f)) - \tilde{D}_i \tilde{v}^* &= red(\tilde{v}^{**} + \sum_\nu a_\nu \tilde{v}_\nu, (\alpha, f)) - (\tilde{v}^{**} + \sum_\nu a_\nu \tilde{v}_\nu) \\
&= red(\tilde{v}^{**}, (\alpha, f)) - \tilde{v}^{**} \\
&= (\tilde{v}^{**} - \tilde{D}^\alpha f) - \tilde{v}^{**} \\
&= -\tilde{D}^\alpha f
\end{aligned}$$

Inserting this expression into (4.8.4) ends the proof of the lemma. \square

Lemma 4.8.3. *Let g be an analytic function on U such that $g \rightarrow 0$ with respect*

to \mathcal{M} . Fix $i \in \{1, \dots, m\}$. Suppose that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \prec \text{HD}\tilde{D}_i g$, $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$. Then $\tilde{D}_i g \rightarrow 0$.

Proof. By the induction on the length of the minimal chain required to reduce g to 0, we may assume that there exists an analytic function $h \neq g$ of $\{x\} \cup \tilde{\Omega}$ with $g \rightarrow h \rightarrow 0$ and $\tilde{D}_i h \rightarrow 0$, say $h = \text{red}(g, (\alpha_h, f_\alpha))$. By Lemma 4.8.2, we have an expression of the form

$$\text{red}(\tilde{D}_i g, (\alpha_h, f_h)) = \tilde{D}_i h + k \tilde{D}_i \tilde{D}^{\alpha_h} f_h + l \tilde{D}^{\alpha_k} f_h. \quad (4.8.5)$$

with k and l analytic functions of $\{x\} \cup \tilde{\Omega}$ satisfying $\text{HD}k \prec \text{HD}g$ and $\text{HD}l \prec \text{HD}g$. Furthermore, either $\text{HD}\tilde{D}_i h \prec \text{HD}\tilde{D}_i g$ or $\text{HD}\tilde{D}_i \tilde{D}^{\alpha_h} f_h \prec \text{HD}\tilde{D}_i g$. In any case, at least two of three summands in above equation have highest derivative strictly less than $\text{HD}\tilde{D}_i g$. Therefore by two applications of Lemma 4.8.1, we have $\text{red}(\tilde{D}_i g, (\alpha_h, f_h)) \rightarrow 0$ and so $\tilde{D}_i g \rightarrow 0$.

□

The *least common multiple* of $\alpha = (\alpha_1, \dots, \alpha_m)$ and $\alpha' = (\alpha'_1, \dots, \alpha'_m)$ is defined by $(\max(\alpha_1, \alpha'_1), \dots, \max(\alpha_m, \alpha'_m))$.

Definition 4.8.1. Let $f, f' \in \mathcal{M}$ with $\text{HD}f = \tilde{D}^{\alpha} u^i$ and $\text{HD}f' = \tilde{D}^{\alpha'} u^{i'}$, and β be the least common multiple of α and α' . Then if $i = i'$, define the *minimal integrability condition* of f and f' to be $\text{IC}(f, f') = \tilde{D}^{\beta-\alpha} f - \tilde{D}^{\beta-\alpha'} f'$. If $i \neq i'$, then $\text{IC}(f, f')$ is said to be undefined.

Theorem 4.8.4. Suppose that for each pair $f, f' \in \mathcal{M}$ with $\text{IC}(f, f')$ well-defined we have $\text{IC}(f, f') \rightarrow 0$. Then \mathcal{M} is a non-commutative Riquier Basis.

Proof. Take $f, f' \in \mathcal{M}$ and $\alpha, \alpha' \in \mathbb{N}^m$ such that $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f'$. We have to show that $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$. We proceed by induction on the highest derivation in $\tilde{D}^\alpha f$.

The basis for the induction is ensured by the assumption $\text{IC}(f, f') \rightarrow 0$.

Now assume that $\tilde{D}^{\alpha^*} f^* - \tilde{D}^{\alpha^{**}} f^{**} \rightarrow 0$ for $f^*, f^{**} \in \mathcal{M}$ and $\alpha^*, \alpha^{**} \in \mathbb{N}^m$ with $\text{HD}\tilde{D}^{\alpha^*} f^* = \text{HD}\tilde{D}^{\alpha^{**}} f^{**} \prec \text{HD}\tilde{D}^\alpha f$.

Suppose that $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f'$ is not equal to $\text{IC}(f, f')$. Thus there exist γ, β and β' in \mathbb{N}_n such that $\gamma \neq (0, \dots, 0)$, $\alpha = \gamma + \beta$, $\alpha' = \gamma + \beta'$ and $\tilde{D}^\beta f - \tilde{D}^{\beta'} f' = \text{IC}(f, f')$.

In contrast to the commutative case, we do not have $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' = \tilde{D}^\gamma(\text{IC}(f, f'))$ for some γ . However, using the commutation rules, we have the following relation:

$$\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' = \tilde{D}^\gamma(\text{IC}(f, f')) + \sum_{\nu} a_{\nu} \tilde{D}^{\nu} f + \sum_{\nu'} a_{\nu'} \tilde{D}^{\nu'} f'$$

where the two sums are finite, a_{ν} and $a_{\nu'}$ are analytic functions and $\text{HD}\tilde{D}^{\nu} f \prec \text{HD}\tilde{D}^\alpha f$ and $\text{HD}\tilde{D}^{\nu'} f' \prec \text{HD}\tilde{D}^\alpha f$.

Using the induction hypothesis and by (repeated) applications of Lemma 4.8.1, we have $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow 0$.

□

In fact, the proof of above lemma gives a more general criterion as following:

Theorem 4.8.5. *Suppose that for each pair $(f, f') \in \mathcal{M}^2$ with $\text{IC}(f, f')$ well-defined there exists an expansion of $\text{IC}(f, f')$ of the form:*

$$\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' = \sum_{(b, b') \in B_{f, f'}} \tilde{D}^{\alpha_{f, f', b, b'}}(\text{IC}(f, f')) + \sum_{\nu} a_{\nu} \tilde{D}^{\nu} f + \sum_{\nu'} a_{\nu'} \tilde{D}^{\nu'} f'$$

where $B_{f,f'}$ is a subset of \mathcal{M}^2 such that for each $(b,b') \in B_{f,f'}$, $\text{IC}(b,b') \rightarrow 0$ and $\text{HD}(\tilde{D}^{\alpha_{f,f',b,b'}} \text{LCM}(\text{HDb}, \text{HDb}')) \prec \text{LCM}(\text{HD}f, \text{HD}f')$, $\text{HD}\tilde{D}^{\nu}f \prec \text{HD}\tilde{D}^{\alpha}f$ and $\text{HD}\tilde{D}^{\nu'}f' \prec \text{HD}\tilde{D}^{\alpha}f$, the two sums are finite, a_{ν} and $a_{\nu'}$ are analytic functions. Then \mathcal{M} is a Riquier basis.

Similar to commutative case, the following corollary gives an efficient criterion for constricting Riquier bases.

Corollary 4.8.6. *Let B be a set of \mathcal{M}^2 such that $\text{IC}(b,b') \rightarrow 0$ for all $(b,b') \in B$ and for all $f, f' \in \mathcal{M}$ with $\text{IC}(f, f')$ well-defined, there exists $b \in \mathcal{M}$ such that*

$$(1) (f, b), (b, f') \in B \text{ and}$$

$$(2) \text{HDb divides the least common multiple of } \text{HD}f \text{ and } \text{HD}f'.$$

Then \mathcal{M} is a Riquier basis.

Proof. From HDb divides $\text{LCM}(\text{HD}f, \text{HD}f')$, we have $\text{LCM}(\text{HD}f, \text{HDb})$ and $\text{LCM}(\text{HDb}, \text{HD}f')$.

Using the commutation rule, there exist $\alpha, \beta \in \mathbb{N}^m$ such that

$$\text{LCM}(\text{HD}f, \text{HD}f') = \text{HD}(\tilde{D}^{\alpha} \text{LCM}(\text{HD}f, \text{HDb})) \text{ and}$$

$$\text{LCM}(\text{HD}f, \text{HD}f') = \text{HD}(\tilde{D}^{\beta} \text{LCM}(\text{HDb}, \text{HD}f')).$$

Therefore we have the expansion

$$\text{IC}(f, f') = \tilde{D}^{\alpha} \text{IC}(f, b) + \tilde{D}^{\beta} \text{IC}(b, f') + \sum_{\nu} a_{\nu} \tilde{D}^{\nu}f + \sum_{\nu'} a_{\nu'} \tilde{D}^{\nu'}f',$$

and the desired result follows from Theorem 4.8.5. \square

For the constant coefficient homogeneous linear systems, the Riquier basis corresponds to Gröbner basis and the more general results can be found in [10].

Example 4.8.1. We return to our running example, the frame treatment of the defining system (4.2.1) of PDE for infinitesimal symmetries of the nonlinear Heat equation (4.1.3). From (4.5.1) for our system (4.4.5) we have two cases.

Case 1: $\theta^3 = 0$, $\tilde{\partial}_3 J \neq 0$; **Case 2:** $\tilde{\partial}_3 J = 0$.

Case 1 ($\theta^3 = 0$, $\tilde{\partial}_3 J \neq 0$). Reducing the system with respect to $\theta^3 = 0$ yields $\frac{1}{2}\tilde{\partial}_1\tilde{\partial}_1\theta^1 - \frac{1}{2}\tilde{\partial}_2\theta^1 = 0$. Computing and reducing the integrability condition between this equation and $\tilde{\partial}_3\theta^1 = -\frac{1}{2}\theta^1$ gives $\tilde{\partial}_1\tilde{\partial}_1\theta^1 = 0$ and $\tilde{\partial}_2\theta^1 = 0$. In summary the system for this case becomes:

$$\begin{aligned} \tilde{\partial}_1\tilde{\partial}_1\theta^1 &= 0 & \tilde{\partial}_1\theta^2 &= 0 & \theta^3 &= 0 \\ \tilde{\partial}_2\theta^1 &= 0 & \tilde{\partial}_2\theta^2 &= 2\tilde{\partial}_1\theta^1 \\ \tilde{\partial}_3\theta^1 &= -\frac{1}{2}\theta^1 & \tilde{\partial}_3\theta^2 &= 0 \end{aligned} \quad (4.8.6)$$

It can be checked that all the integrability conditions for this system are satisfied and it satisfies all the conditions for a non-commutative Riquier Basis. There are three parametric derivations θ^1 , θ^2 , $\tilde{\partial}_1\theta^1$. Hence by the Non-commutative Riquier Existence and Uniqueness Theorem its symmetry algebra is of dimension three.

Case 2 ($\tilde{\partial}_3 J = 0$). Further compatibility conditions and reductions yield the condition $(3 - 4J)\tilde{\partial}_1\theta^3 = 0$. Thus there are two cases: **Case 2a:** $J \neq \frac{3}{4}$, $\tilde{\partial}_1\theta^3 = 0$ and **Case 2b:** $J = \frac{3}{4}$.

Case 2a: ($J \neq \frac{3}{4}$, $\tilde{\partial}_1\theta^3 = 0$). We obtain:

$$\begin{aligned} \tilde{\partial}_1\tilde{\partial}_1\theta^1 &= 2(1 - J)\tilde{\partial}_1\theta^3 & \tilde{\partial}_1\theta^2 &= 0 & \tilde{\partial}_1\tilde{\partial}_1\theta^3 &= 0 \\ \tilde{\partial}_2\theta^1 &= 0 & \tilde{\partial}_2\theta^2 &= 2\tilde{\partial}_1\theta^1 - \theta^3 & \tilde{\partial}_2\theta^3 &= 0 \\ \tilde{\partial}_3\theta^1 &= -\frac{1}{2}\theta^1 & \tilde{\partial}_3\theta^2 &= 0 & \tilde{\partial}_3\theta^3 &= 0 \end{aligned} \quad (4.8.7)$$

It can be checked that all the conditions for a non-commutative Riquier Basis are satisfied. There are four parametric derivations $\theta^1, \theta^2, \tilde{\partial}_1\theta^1, \theta^3$. Hence its symmetry algebra is of dimension four.

Case 2b: $J = \frac{3}{4}$. The system becomes the non-commutative Riquier Basis:

$$\begin{array}{lll} \tilde{\partial}_1\tilde{\partial}_1\theta^1 = 0 & \tilde{\partial}_1\theta^2 = 0 & \tilde{\partial}_1\theta^3 = 0 \\ \tilde{\partial}_2\theta^1 = 0 & \tilde{\partial}_2\theta^2 = 2\tilde{\partial}_1\theta^1 - \theta^3 & \tilde{\partial}_2\theta^3 = 0 \\ \tilde{\partial}_3\theta^1 = -\frac{1}{2}\theta^1 & \tilde{\partial}_3\theta^2 = 0 & \tilde{\partial}_3\theta^3 = 0. \end{array}$$

There are five parametric derivations $\theta^1, \theta^2, \theta^3, \tilde{\partial}_1\theta^1, \tilde{\partial}_1\theta^3$, yielding a five-dimensional symmetry algebra.

4.9 Relative Riquier Bases

For this section, let \mathcal{M} be a finite set of \prec -monic analytic functions of $\tilde{\Omega}$ that are polynomially nonlinear in $\text{Prin}\mathcal{M}$ and let \mathcal{N} be a finite set of analytic functions of $\{x\} \cup \text{Par}\mathcal{M}$.

Definition 4.9.1. Let U be a non-empty open subset of $\mathbb{F}^{\tilde{\Omega}}$ on which $\mathcal{M} \cup \mathcal{N}$ is analytic. For η an analytic function of $\mathbb{F}^{\tilde{\Omega}}$, we say that η is a *special consequence* of \mathcal{N} if η admits an expansion of the form

$$\eta = \sum_{i=1}^k h_i g_i \tag{4.9.1}$$

with $g_1, \dots, g_k \in \mathcal{N}$ and h_1, \dots, h_k analytic functions on U that depend on $\{x\} \cup \text{Par}\mathcal{M}$ only.

We say that \mathcal{M} is a *Riquier basis relative to \mathcal{N} on U* if for all $\alpha, \alpha' \in \mathbb{N}^m$ and

$f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}_\alpha f = \text{HD}\tilde{D}_{\alpha'} f'$, the integrability condition $\tilde{D}_\alpha f - \tilde{D}_{\alpha'} f'$ can be reduced to a special consequence of \mathcal{N} on U .

Lemma 4.9.1. *Let f, f', g be functions of $\tilde{\Omega}$ with $f, f' \in \mathcal{M}$, let α, β be elements of \mathbb{N}^m . Let h and k denote the one step reductions: $h = \text{red}(g, (\alpha, f))$ and $k = \text{red}(g, (\beta, f'))$. Then:*

(1) *If $\text{HD}\tilde{D}_\alpha f \prec \text{HD}\tilde{D}_\beta f'$ then $\text{red}(h, ((\beta, f'), (\alpha, f))) = \text{red}(k, (\alpha, f))$.*

(2) *Suppose that $\text{HD}\tilde{D}_\alpha f = \text{HD}\tilde{D}_\beta f'$ and g is analytic and polynomial in $\text{HD}\tilde{D}_\alpha f$. Suppose further that $\tilde{D}_\alpha f - \tilde{D}_\beta f' \rightarrow^\mu \eta$ for some special consequence η of \mathcal{N} . Then*

$$\text{red}(h, \mu) - \text{red}(k, \mu) = h'\eta$$

for some analytic function h' defined on U with $\text{HD}h' \prec \text{HD}g$.

Proof. Let $\tilde{v}^* = \text{HD}\tilde{D}^\alpha f$ and $\tilde{v}^{**} = \text{HD}\tilde{D}^\beta f'$. If $\tilde{v}^* \prec \tilde{v}^{**}$. The proof of (1) is similar to lemma 6.1(1). Next we give the proof of (2).

Let $\tilde{v}^* = \tilde{v}^{**}$ and p be the degree of g in \tilde{v}^* , then

$$\begin{aligned} \text{red}(h, \mu) &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v} - \tilde{D}^\alpha f, \mu))_{\tilde{v} = \tilde{v}^*}) \\ &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v}, \mu) - \text{red}(\tilde{D}^\alpha f, \mu))_{\tilde{v} = \tilde{v}^*}) \\ &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v}, \mu) - \text{red}(\tilde{D}^\beta f', \mu) - \eta)_{\tilde{v} = \tilde{v}^*}) \\ &= g(x, (\text{red}(\tilde{v}, \mu))_{\tilde{v} \neq \tilde{v}^*}, (\text{red}(\tilde{v}, \mu) - \text{red}(\tilde{D}^\beta f', \mu))_{\tilde{v} = \tilde{v}^*}) + h'\eta \\ &= \text{red}(k, \mu) + h'\eta. \end{aligned}$$

where h' is obtained by expanding g in term of η and using the finite degree of \tilde{v}^* .

□

Lemma 4.9.2 (Diamond Lemma for relative reduction). *Fix $\tilde{v} \in \tilde{\Omega}$. Suppose that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \leq \tilde{v}$, we have $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f'$ reduces to a special consequence of \mathcal{N} . Let g be an analytic function on U that is polynomial in $\text{Prin}\mathcal{M}$ with $\text{HD}g \leq \tilde{v}$, and two non-empty reductions $g \rightarrow h, g \rightarrow k$. Then there exists l, l' with $h \rightarrow l$ and $k \rightarrow l'$ such that $l - l'$ is a special consequence of \mathcal{N} .*

Proof. The proof is repeated from the proof of Theorem 6.3.2 in Rust's thesis.

Suppose that g, h, k satisfy the hypotheses of the lemma and that there so not exist l, η with $h \rightarrow l$ and $k \rightarrow l + \eta$, such that η is a special consequence of \mathcal{N} . Without loss of generality, we may assume that g is minimal in the sense that for g' with $g \rightarrow g'$ and $g' \neq g$, for any h', k' with $g' \rightarrow h'$ and $g' \rightarrow k'$, there do exist l', η' with $h' \rightarrow l'$ and $k' \rightarrow l' + \eta'$, such that η' is a special consequence of \mathcal{N} . We may assume that $h \neq g$ and $k \neq g$. Take h_1, k_1 with $h + 1 \neq g$ and $k_1 \neq g$ such that $g \rightarrow h_1 \rightarrow h$ and $g \rightarrow k_1 \rightarrow k$. By the previous lemma, there exist l_1, η_1 with $h_1 \rightarrow l_1$ and $k_1 \rightarrow l_1 + \eta_1$ such that η_1 is a special consequence of \mathcal{N} . By the minimality assumption, since h_1 is a proper reduction of g , there exist l', η' such that $h \rightarrow l'$ and $l_1 \rightarrow l' + \eta'$ with η' a special consequence of \mathcal{N} . Note that $k_1 \rightarrow k_1 \rightarrow l' + \eta' + \eta_1$ and $k_1 \rightarrow k$. Therefore, again by the minimality assumption, there exist l'', η'' such that $l' + \eta' + \eta_1 \rightarrow l''$ and $k \rightarrow l'' + \eta''$ with η'' a special consequence of \mathcal{N} . Thus there exists μ such that $\text{red}(l' + \eta' + \eta_1, \mu) = l''$, so $\text{red}(l', \mu) + \eta' + \eta_1 = l''$ and $\text{red}(l', \mu) = l'' - \eta' - \eta_1$. Thus $h \rightarrow l'' - \eta' - \eta_1$ and $k \rightarrow l'' + \eta''$, a contradiction with $l = l'' - \eta' - \eta_1$ and $\eta = \eta'' + \eta' + \eta_1$. \square

From above lemmas, it is easy to see that:

Theorem 4.9.3. *Suppose that \mathcal{M} is a non-commutative Riquier Basis relative to \mathcal{N} and g is an analytic function on U that is polynomial in $\text{Prin}\mathcal{M}$. Then the difference of any two complete reductions of g is a special consequence of \mathcal{N} .*

For g an analytic function on U polynomial in $\text{Prin}\mathcal{M}$, let $\text{redmod}(g, \mathcal{M})$ denote a particular choice of complete reduction of g with respect to \mathcal{M} . Let ϕ be a specification of initial data for \mathcal{M} that satisfies $\phi(\eta) = 0$ for all $\eta \in \mathcal{N}$. Theorem 4.9.3 implies that $\text{redmod}(g, \mathcal{M})$ is uniquely defined up to a special consequence of \mathcal{N} . In particular, $\phi(\text{redmod}(g, \mathcal{M}))$ is independent of the choice of complete reduction.

Theorem 4.9.4 (Formal Non-commutative relative Riquier Existence Theorem). *Let \mathcal{M} be a non-commutative relative Riquier Basis relative to \mathcal{N} such that each $f \in \mathcal{M}$ is polynomial in $\text{Prin}\mathcal{M}$. For $x^0 \in \mathbb{F}^m$, let ϕ be a specification of initial data for \mathcal{M} at x^0 such that $\phi(f)$ is well-defined for all $f \in \mathcal{M}$. Then there is a unique formal power series solution $u(x) \in \mathbb{F}[[x - x^0]]^n$ to \mathcal{M} at x^0 such that $\tilde{D}^\alpha u^i(x^0) = \phi(\text{redmod}(\tilde{v}_\alpha^i, \mathcal{M}))$ for all $\tilde{v}_\alpha^i \in \text{Par}\mathcal{M}$. Furthermore, every formal power series solution to \mathcal{M} at x^0 may be obtained in this way for some ϕ .*

Proof. The proof is same as the proof of Theorem 4.7.2. We only need to use $\text{redmod}(\tilde{v}_\alpha^i, \mathcal{M})$ instead of $\text{red}(\tilde{v}_\alpha^i, \mathcal{M})$ in some places and use above lemmas. □

Lemma 4.9.5 (Relative reduction of a sum). *Suppose h and k are polynomials in $\text{Prin}\mathcal{M}$. Suppose $h \rightarrow^\mu \eta_h$ and $k \rightarrow^\nu \eta_k$. Suppose further that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \leq \text{HD}k$, there exists a special consequence η of \mathcal{N} such that $\tilde{D}^\alpha f - \tilde{D}^{\alpha'} f' \rightarrow \eta$. Then we have $h + k \rightarrow \eta'$ for a special consequence of η' of \mathcal{N} .*

Proof. There are two cases:

(1) If $\text{red}(k, \mu)$ is an empty reduction, then

$$\begin{aligned} \text{red}(h + k, (\mu, \nu)) &= \text{red}(h, (\mu, \nu)) + \text{red}(k, (\mu, \nu)) \\ &= \text{red}(0, \nu) + \text{red}(k, \nu) \\ &= 0 + \eta_k \\ &= \eta_k \end{aligned}$$

(2) If $\text{red}(k, \mu)$ is a non-empty reduction, say $l = \text{red}(k, \mu)$, then by Lemma 4.9.2 there exist j with $\eta_k \rightarrow j$ and $l \rightarrow \eta_l + j$ for some special consequence η_l of \mathcal{N} . Since $\eta_k \rightarrow j$, we have $j = \eta_k$ and hence $l \rightarrow \eta_l + \eta_k$, say $\eta_l + \eta_k = \text{red}(l, \rho)$. Then we have:

$$\begin{aligned} \text{red}(h + k, (\mu, \rho)) &= \text{red}(h, (\mu, \rho)) + \text{red}(k, (\mu, \rho)) \\ &= \text{red}(0, \rho) + \text{red}(l, \rho) \\ &= \eta_h + \eta_l + \eta_k. \end{aligned}$$

Therefore, $\eta' := \eta_h + \eta_l + \eta_k$, as required. □

Lemma 4.9.6. *Let g be an analytic function on U such that $g \rightarrow \eta_g$ for some special consequence of \mathcal{N} . Fix $i \in \{1, \dots, m\}$. Suppose that for all $\alpha, \alpha' \in \mathbb{N}^m$ and $f, f' \in \mathcal{M}$ with $\text{HD}\tilde{D}^\alpha f = \text{HD}\tilde{D}^{\alpha'} f' \prec \text{HD}\tilde{D}_i g, \tilde{D}^\alpha f - \tilde{D}^{\alpha'} f'$ reduces to a special consequence of \mathcal{N} for all $\eta \in \mathcal{N}$. Then there exists a special consequence η' of \mathcal{N} such that $\tilde{D}_i g \rightarrow \eta'$.*

Proof. There are two cases:

Case 1. g is a special consequence of \mathcal{N} , say, $g := \eta_g := \sum_{j=1}^r g_j h_j$.

It is easy to see that the Leibniz rule is still true for \tilde{D}_i . Then $\tilde{D}_i g = \sum_{j=1}^r (\tilde{D}_i g_j) h_j + \sum_{j=1}^r g_j (\tilde{D}_i h_j)$.

Since for $j \in \{1, \dots, r\}$, $\text{HD}(\tilde{D}_i g_j) \prec \tilde{D}_i(\max_{\eta \in \mathcal{N}}(\text{HD}\eta))$, by $(r-1)$ -applications of lemma 4.9.5, $\sum_{j=1}^r (\tilde{D}_i g_j) h_j$ reduces to a special consequence of \mathcal{N} . Observe that an

arbitrary reduction μ of $\sum_{j=1}^r g_j(\tilde{D}_i h_j)$ by \mathcal{M} is of the form $\text{red}(\sum_{j=1}^r g_j(\tilde{D}_i h_j), \mu) = \sum_{j=1}^r g_j \text{red}(\tilde{D}_i h_j, \mu)$. Therefore any complete reduction of $\sum_{j=1}^r g_j \tilde{D}_i h_j$ is a special consequence of \mathcal{N} . By lemma 4.9.5, $\tilde{D}_i g$ can be reduced to a special consequence of \mathcal{N} .

Case 2. $g \rightarrow \eta_g$ with $g \neq \eta_g$.

By the induction on the length of the minimal chain required to reduce g to a special consequence of \mathcal{N} , we may assume that there exists an analytic function $h \neq g$ of $\{x\} \cup \tilde{\Omega}$ with $g \rightarrow h \rightarrow \eta_g$ and $\tilde{D}_i h \rightarrow \eta_h$, say $h = \text{red}(g, (\alpha_h, f_\alpha))$. By Lemma 4.8.2, we have an expression of the form

$$\text{red}(\tilde{D}_i g, (\alpha_h, f_h)) = \tilde{D}_i h + k \tilde{D}_i \tilde{D}^{\alpha_h} f_h + l \tilde{D}^{\alpha_k} f_h. \quad (4.9.2)$$

with k and l analytic functions of $\{x\} \cup \tilde{\Omega}$ satisfying $\text{HD}k \prec \text{HD}g$ and $\text{HD}l \prec \text{HD}g$. Furthermore, either $\text{HD}\tilde{D}_i h \prec \text{HD}\tilde{D}_i g$ or $\text{HD}\tilde{D}_i \tilde{D}^{\alpha_h} f_h \prec \text{HD}\tilde{D}_i g$. In any case, at least two of three summands in above equation have the highest derivative strictly less than $\text{HD}\tilde{D}_i g$. Therefore by two applications of Lemma 4.9.5, there exists a special consequence η' of \mathcal{N} such that $\text{red}(\tilde{D}_i g, (\alpha_h, f_h)) \rightarrow \eta'$ and so $\tilde{D}_i g \rightarrow \eta'$.

□

Definition 4.9.2. We say that $(\mathcal{M}, \mathcal{N})$ is in *rif'* form on U if \mathcal{M} is a Riquier basis relative to \mathcal{N} on U and for all $i \in \{1, \dots, m\}$ and $g \in \mathcal{N}$, $\tilde{D}_i g$ can be reduced to a special consequence of \mathcal{N} on U .

Theorem 4.9.7. *Suppose that for each pair $f, f' \in \mathcal{M}$ with $\text{IC}(f, f')$ well-defined we have $\text{IC}(f, f')$ reduces to a special consequence of \mathcal{N} and $\tilde{D}_i \eta$ reduces to a special consequence of \mathcal{N} for all $\eta \in \mathcal{N}$. Then $(\mathcal{M}, \mathcal{N})$ is in *rif'* form.*

Proof. The proof is similar to the proof of Theorem 4.8.4 by using Lemmas 4.9.5 and 4.9.6. □

For \mathcal{S} a finite set of analytic functions of $\tilde{\Omega}$, let $\mathcal{L} := L_{\rightarrow}\mathcal{S}$ (resp. $\mathcal{N} := N_{\rightarrow}\mathcal{S}$) denote the set of leading linear (resp. nonlinear) elements of \mathcal{S} . Clearly $N_{\rightarrow}\mathcal{S} = \mathcal{S} \setminus L_{\rightarrow}\mathcal{S}$.

Next let \mathcal{S} be a finite set of analytic functions of $\tilde{\Omega}$ in rif' form on an open subset U of $\{x\} \cup \text{Par}(\mathcal{L})$ and ϕ be a specification of initial data for \mathcal{L} with $\phi \in U$. Clearly, for ϕ to correspond to a formal solution $u(x)$ of \mathcal{S} it is necessary that $\phi(g) = 0$ for all $g \in \mathcal{N}$.

Theorem 4.9.8. *Let $u(x) \in \mathbb{F}[[x - x^0]]^n$ be the solution to \mathcal{L} that corresponds to ϕ via Theorem 4.9.4. Then $u(x)$ is a solution to \mathcal{S} .*

Proof. By induction we first show that for all $g \in \mathcal{N}$ and $\alpha \in \{1, \dots, m\}$, there exists an expansion of $\text{redmod}(\tilde{D}_{\alpha}g, \mathcal{L})$ of the form

$$\text{redmod}(\tilde{D}_{\alpha}g, \mathcal{L}) = \sum_{g_i \in \mathcal{N}} h_i g_i \tag{4.9.3}$$

with $\phi(h_i)$ is well-defined.

For $\alpha = 0$, $\text{redmod}(\tilde{D}_{\alpha}g, \mathcal{L}) = g$ since g is independent of the principal derivatives. By induction on $|\alpha|$, we may assume that $\tilde{D}_{\alpha}g$ satisfies an equation of the form (4.9.3). For any j , we have to show that $\tilde{D}_j \tilde{D}_{\alpha}g$ is also satisfies an equation of this form.

Note that

$$\tilde{D}_{\alpha}g - \text{redmod}(\tilde{D}_{\alpha}g, \mathcal{L}) \rightarrow_{\mathcal{L}} 0.$$

From the commutative rule, $\tilde{D}_j(\tilde{D}_\alpha g - \text{redmod}(\tilde{D}_\alpha g, \mathcal{L}))$ is linear in $\text{Par}\mathcal{L}$. Thus by lemma 4.9.5,

$$\tilde{D}_j(\tilde{D}_\alpha g - \text{redmod}(\tilde{D}_\alpha g, \mathcal{L})) \rightarrow_{\mathcal{L}} \eta$$

for some η of the form $\eta = \sum_{i=1}^k h_i g_i$ with each $g_i \in \mathcal{N}$ and so

$$\begin{aligned} \text{redmod}(\tilde{D}_j \tilde{D}_\alpha g, \mathcal{L}) - \eta &= \text{redmod}(\tilde{D}_j \text{redmod}(\tilde{D}_\alpha g, \mathcal{L}) - \mathcal{L}) \\ &= \text{redmod}(\tilde{D}_j \sum_{i=1}^k h_i g_i, \mathcal{L}) \\ &= \sum_{i=1}^k \text{redmod}((\tilde{D}_j h_j) g_j + h_j \tilde{D}_j g_j, \mathcal{L}) \\ &= \sum_{g_i, g' \in \mathcal{N}} \text{redmod}(\tilde{D}_j h_j, \mathcal{L}) g_j + h_j h(j, g_i, g') g'. \end{aligned}$$

Since h_i depends only on the parametric derivatives, $\tilde{D}_j h_i$ is a polynomial (in fact linear) in the principal derivatives. Therefore, since $\phi(h_i)$ is well-defined, it follows that $\phi(\text{redmod}(\tilde{D}_j h_i, \mathcal{L}))$ is well-defined and we have the equation 4.9.3.

Now since $u(x)$ is a solution to \mathcal{L} , we have that $\tilde{D}_\alpha g[u](x) = \text{redmod}(\tilde{D}_\alpha g, \mathcal{L})[u](x)$. Therefore,

$$\begin{aligned} \tilde{D}_\alpha g[u](x^0) &= \text{redmod}(\tilde{D}_\alpha g, \mathcal{L})[u](x^0) \\ &= \phi(\text{redmod}(\tilde{D}_\alpha g, \mathcal{L})) \\ &= \sum_{g_i \in \mathcal{N}} \phi(h_i) \phi(g_i) \\ &= 0. \end{aligned}$$

□

4.10 Analyticity Issues

Theorem 4.7.2 gives existence and uniqueness conditions for a formal power series solution for an associated specification of initial data. The Riquier-Janet Existence and

Uniqueness Theorem for the commutative case states that, under certain assumptions, a specification of *analytic* initial data yields an *analytic* power series solution.

In this section, we investigate the generalization of this analyticity theorem to the non-commutative case by seeking conditions on the initial data specification ensuring that the formal power series solution is analytic.

Riquier [30] and Janet [14] consider systems of PDE with commuting derivations. They consider orthonomic and passive systems to express the analyticity theorem. Instead of defining the orthonomic and passive systems, we will use the non-commutative Riquier Basis described in this paper to state the Riquier analyticity theorem. We need the following definitions.

Definition 4.10.1. A Riquier ranking \prec is a positive ranking satisfying $\tilde{\partial}^\alpha u^i \prec \tilde{\partial}^\beta u^i \iff \tilde{\partial}^\alpha u^j \prec \tilde{\partial}^\beta u^j$ for any i and j .

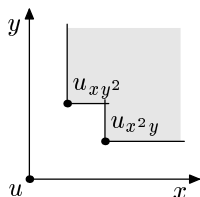
Definition 4.10.2. A *specification of initial data* ϕ for a system \mathcal{M} is *analytic* if there exist two positive real numbers M and r such that $|\phi(\tilde{\partial}^\alpha u^i)| \leq Mr^{|\alpha|}\alpha!$ for all $\tilde{\partial}^\alpha u^i$ in $\text{Par}\mathcal{M}$.

With the above definitions we can state the Riquier analyticity theorem in the following alternative form.

Theorem 4.10.1. *Let \prec be a Riquier ranking compatible with the total degree ordering. Suppose that the $\tilde{\partial}_i$ commute. Consider a non-commutative Riquier Basis \mathcal{M} and analytic initial data specification ϕ . Then the unique formal power series solution thus defined is also analytic.*

In the Riquier-Janet Theory, specifying the initial conditions is equivalent to fixing the values of the dependant variables and their derivatives at a point x^0 .

Example 4.10.1.
$$\begin{cases} u_{xxy} = f(u) \\ u_{xyy} = g(u) \end{cases}$$



Choosing a specification of initial data around the origin $x = 0, y = 0$ in the Riquier-Janet approach amounts to fixing the value of u_{xy} at $x = y = 0$, and fixing the value of u on $\{x = 0\} \cup \{y = 0\}$. A precise construction of the initial data based on multiplicative variables [14] would consist in fixing the value of u_{xy} on $M_0 = \{(0, 0)\}$; fixing the value u on $M_1 = \{(x, 0) : x \in \mathbb{R}\}$; and fixing the value u_y on $M_2 = \{(0, y) : y \in \mathbb{R}\}$.

More generally, a specification of initial data can be expressed in the following geometric manner, which is more suitable for our investigation of the non-commutative case. Choosing a specification of initial data around x^0 is equivalent to assigning functions to some parametric derivatives along specific sub-manifolds M_i . The choice of the parametric derivatives and the M_i is described in [14]. Moreover, choosing an analytic initial data specification is equivalent to assigning analytic functions to the dependant variables on the sub-manifolds M_i .

Extending these results to the non-commutative case leads to the following questions:

- What is the geometric meaning of fixing an initial data specification?
- What criteria must the initial data satisfy to ensure the analyticity of the associated formal power series solution?

We first consider the case where there are a finite number of parameters in the formal power series solution. As is well-known the integration of such systems can be reduced

to integration of ODE (and existence of smooth solutions established via the Frobenius Theorem). We include the following remarks to assist in algorithmic implementations related to this case.

Remark 4.10.1 (Analyticity in the finite parameter case). Let \prec be a Riquier ranking compatible with the total degree ordering. Consider a non-commutative Riquier Basis \mathcal{M} with a finite set of parametric derivations $\text{Par}\mathcal{M} = \{w^1, \dots, w^k\}$. Then the formal power series solution about x^0 with initial data $w^1(x^0) = w_0^1, \dots, w^k(x^0) = w_0^k$ is analytic at x^0 .

Proof. For $i = 1, \dots, m$ any $\tilde{D}_i w^l \in \text{Prin}\mathcal{M}$ can be completely reduced by \mathcal{M} to an analytic function f_i^l of $\{x\} \cup \text{Par}\mathcal{M}$ such that

$$\tilde{D}_i w^l = f_i^l. \quad (4.10.1)$$

Now from (4.3.4) it follows that (4.10.1) is equivalent to

$$D_i w^l = \sum_j b_{ij}(x, u) f_j^l, \quad (4.10.2)$$

where $b(x, u)$ is the inverse matrix of $a(x, u)$.

The easily computed integrability conditions of (4.10.2) are analytic functions of $\{x\} \cup \text{Par}\mathcal{M}$. If one integrability condition was not satisfied, there would a set of initial conditions for (4.10.2) such that (4.10.2) does not admit a solution. Thus, for the initial conditions, (4.10.1) and \mathcal{M} would not admit a solution, which contradicts the existence theorem of a solution for the non-commutative Riquier Basis \mathcal{M} . Thus the system (4.10.2) is a commutative Riquier Basis, and by the standard commutative theory, must have a formal power series solution with the given data, which is analytic at x^0 . \square

Remark 4.10.2. Under the hypotheses of Remark 4.10.1, the integration of \mathcal{M} is equivalent to integrating a system of ODE along an analytic curve.

Proof. Consider an analytic curve $x(\tau) = x_i(\tau)$, with $x(0) = x^0$. Then $\frac{dw^l}{d\tau} = \sum_i \frac{dx_i}{d\tau} \frac{\partial w^l}{\partial x_i}$ which from (4.10.2) yields the system of ODE:

$$\frac{dw^l}{d\tau} = \sum_i \frac{dx_i}{d\tau} \sum_j b_{ij}(x, u) f_j^l. \quad (4.10.3)$$

□

Traditional commutative differential elimination packages often use elimination rankings to decouple ODE which can then be sometimes exactly integrated by ODE solvers. Remark 4.10.2 gives an alternative method for exposing ODE systems, and although well-known in classical differential geometry, does not appear to have been implemented in the common computer algebra solvers for over-determined PDE systems (even in the commutative case). It is interesting to explore to what extent geometric ODE integrators (numerical integrators invariant under the admitted Lie group), could be fruitfully applied to such systems using Theorem 4.10.2.

with referee comments We now consider the case of a non-commutative Riquier Basis \mathcal{M} where $\text{Par}\mathcal{M}$ is infinite. In the commutative case, the geometric theory of PDE [26], the geometric prolongation of the system to an order r is obtained by applying D_i to the equations of the system until no undifferentiated equations of order r or less remain. Equivalently one may obtain the geometric prolongation by similarly using \tilde{D}_i . Since the system is a non-commutative Riquier Basis any prolongation of the system is formally integrable (as defined in the geometric theory). It is also a consequence of the geometric theory that some finite order prolongation of the system has involutive symbol, and hence the system is also involutive. Indeed in our case the

Mansfield Prolongation Theorem [20] can be used to determine a bound for that order. Once the system is involutive, then an analytic existence and uniqueness theorem can be given.

We sketch below some partial results obtained in the infinite non-commutative case.

We can generalize the Riquier analyticity theorem to the non-commuting infinite case by assuming that the $\tilde{\partial}_i$ are lined up with an analytical system of coordinates. In particular we make the hypothesis **(H)** that there exist m analytical functions X_i and a neighborhood $N(x^0)$ of the expansion point x^0 satisfying:

- $\tilde{\partial}_i X_j = 0$ in $N(x^0)$ if $i \neq j$
- The Jacobian of (X_1, \dots, X_m) does not vanish in $N(x^0)$

This is less stringent than assuming that the $\tilde{\partial}_i$ are associated to a set of coordinates since in that case we would have $\tilde{\partial}_i X_j = 1$, if $i = j$ and 0 otherwise.

Theorem 4.10.2. *Let \prec be a Riquier ranking compatible with the total degree ordering. Suppose that the frame $\tilde{\partial}_i$ satisfies **(H)**. Consider a non-commutative Riquier Basis \mathcal{M} and analytic initial data specification ϕ . Then the unique formal power series solution thus defined is also analytic.*

The proof consists in transforming our problem into a commuting derivative problem in the set of coordinates X_i where we can apply the Riquier theorem.

A sketch of the straightforward proof follows.

Proof. Sketch of proof

- Without loss of generality set $x^0 = 0$.

- Introduce new commuting derivations $\widehat{\partial}_i$ which are based on the system of coordinates X .
- Prove that there exist m scalar functions $A_i(x)$ analytic at x^0 such that $\widetilde{\partial}_i = A_i \widehat{\partial}_i$ with $A_i(0) \neq 0$.
- Any derivation $\widetilde{\partial}^\alpha$ can be rewritten in terms of $A_1^{\alpha_1} \cdots A_m^{\alpha_m} \widehat{\partial}^\alpha$ plus a finite sum of terms $f_\beta \widehat{\partial}^\beta$ where f_β is an analytic function and β strictly “divides” α (i.e. $\beta_k \leq \alpha_k$ for $1 \leq k \leq m$).
- Replace the $\widetilde{\partial}_i$ by $\widehat{\partial}_i$ in \mathcal{M} and obtain an orthonomic system $\widehat{\mathcal{M}}$. The set of leaders of $\widehat{\mathcal{M}}$ coincide with the ones in \mathcal{M} by replacing the $\widetilde{\partial}_i$ with $\widehat{\partial}_i$.
- The analytic initial condition specification ϕ of \mathcal{M} defines an analytic initial condition specification $\widehat{\phi}$ of $\widehat{\mathcal{M}}$.
- Since the problem has been reduced to the commutative case, $\widehat{\phi}$ has a geometrical meaning. In particular this demonstrates that the dependant variables and some of their derivatives have been fixed to analytic functions on unions of sub-manifolds of the form $X_i = 0$.
- Since the ranking is Riquier and compatible with total order, the commutative Riquier analyticity theorem applies.

□

It is interesting to note that the condition **(H)** is always satisfied in the case $m = 2$ (the case $m = 1$ is obvious).

In general the condition **(H)** of the previous subsection is not satisfied as the following example shows:

Example 4.10.2.

$$\begin{cases} \tilde{\partial}_1 = \partial_x & + \partial_z \\ \tilde{\partial}_2 = & \partial_y + z\partial_z \\ \tilde{\partial}_3 = & \partial_z \end{cases}$$

Since $\tilde{\partial}_1\tilde{\partial}_2 - \tilde{\partial}_2\tilde{\partial}_1 = \tilde{\partial}_3$, the relations $\tilde{\partial}_1X_3 = 0$ and $\tilde{\partial}_2X_3 = 0$ imply $\tilde{\partial}_3X_3 = 0$.

Thus, X_3 cannot generate a system of coordinates.

Thus in the general case the above geometric interpretation is lost and it is not straightforward to adapt the proof of the analyticity theorem ([30], [14]) to the non-commuting case. Indeed, the use of majorizing functions to prove convergence of the formal series rely on commuting derivatives. The analyticity results obtained from prolonging the system to involution indicates there is a reasonable chance of proving a suitable non-commutative Riquier analyticity theorem in the non-commutative case.

Bibliography

- [1] M.A. Akivis and B.A. Rosenfeld, *Élie Cartan (1869-1951)*, Translations Math. Monographs, vol. 123, American Math. Soc., Providence, R.I., 1993.
- [2] François Boulier, Daniel Lazard, François Ollivier, and Michel Petitot, *Representation for the radical of a finitely generated differential ideal*, proceedings of ISSAC'95 (Montréal, Canada) (A.H.M. Levelt, ed.), ACM Press, 1995, pp. 158–166.
- [3] T. Becker, V. Weispfenning, and H. Kredel, *Gröbner Bases, a Computational Approach to Commutative Algebra*, corrected second printing ed., vol. 141 of Graduate Texts in Mathematics. Springer, New York, 1998.
- [4] M. Boutin, *Numerically invariant signature curves*, Int. J. Computer Vision **40** (2000), 235–248.
- [5] C.J. Budd and A. Iserles, *Geometric Integration: numerical integration of differential equations on manifolds*, Phil. Trans. Roy. Soc: London A **357** (1999), 945–956.
- [6] É. Cartan, *La Méthode du Repère Mobile, la Théorie des Groupes Continus, et les Espaces Généralisés*, Exposés de Géométrie No. 5, Hermann, Paris 1935.

- [7] M. Fels and P.J. Olver, *Moving Coframes. I. A practical algorithm*, Acta. Appl. Math. **51** (1998), 161–213.
- [8] M. Fels and P.J. Olver, *Moving Coframes. II. Regularization and theoretical foundations*, Acta. Appl. Math. **55** (1999), 127–208.
- [9] M. Giesbrecht, G. Reid and Y. Zhang, *Non-commutative Gröbner Bases in Poincaré-Birkhoff-Witt Extensions*, Proc. of the Fifth International Workshop on Computer Algebra in Scientific Computing (CASC 2002, Yalta, Ukraine), edited by V.G. Ganzha, E.W. Mayr & E.V. Vorozhtsov (Pub: Technical University of Munich) 97–106.
- [10] M. Giesbrecht, G. Reid and Y. Zhang, *Gröbner Bases in modules over Poincaré-Birkhoff-Witt Extensions*, preprint.
- [11] P.A. Griffiths, *On Cartan's method of Lie Groups and moving frames as applied to uniqueness and existence questions in differential geometry*, Duke Math. J. **41** (1974), 775–814.
- [12] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, New York, 2002.
- [13] W. Hereman, *Review of symbolic software for Lie symmetry analysis*, Math. and Comp. Modelling **25** (1997) 115–132.
- [14] M. Janet, *Sur les systèmes d'équations aux dérivées partielles*, J. de Math, 3(1920), 65-151.
- [15] I.A. Kogan, *Inductive construction of moving frames*, Contemp. Math. **285** (2001), 157–170.

- [16] E. R. Kolchin, *Differential Algebra and Algebraic Groups*, Academic Press, New York 1973.
- [17] I.G. Lisle, *Equivalence Transformations for Classes of Differential equations*, PhD Thesis (University of British Columbia, 1992). Available at www.ise.canberra.edu.au/MathStat/StaffPages/LisleDissertation.pdf.
- [18] I.G. Lisle and G.J. Reid (2000), *Symmetry classification using invariant moving frame*, ORCCA Technical Report TR-00-08. Available at <http://www.orcca.on.ca/TechReports/2000/TR-00-08.html>
- [19] E.L. Mansfield, *Algorithms for symmetric differential systems*, Foundations of Computational Math.(2001) 1:335-383.
- [20] E.L. Mansfield, *A simple criterion for involutivity*, J. London Math. Soc. (2) **54**, (1996) 2:323–345.
- [21] P.J. Olver, *Applications of Lie Groups to Differential Equations*, Second Edition, Graduate Texts in Mathematics **107**, Springer-Verlag, New York, 1993.
- [22] P.J. Olver, *Equivalence, Invariants, and Symmetry*, Cambridge University Press, 1995.
- [23] P.J. Olver, *Geometric foundations of numerical algorithms and symmetry*, Appl. Alg. Engin. Comput. **11** (2001), 417–436.
- [24] P.J. Olver, *Moving Frames*, pre-print, University of Minnesota (see <http://www.math.umn.edu/~olver/xtra.html>).

- [25] L.V. Ovsiannikov, *Group Analysis of Differential Equations*, Academic Press, New York, 1982.
- [26] Jean-François Pommaret, *Systems of Partial Differential Equations and Lie pseudogroups*, Gordon and Breach science publishers Inc., 1978.
- [27] G. Reid, *Algorithms for reducing a system of PDEs to standard form, determining the dimension of its solution space and calculating its Taylor series solution*, Euro. J. Appl. Maths., 2(1991), 293-318.
- [28] G.J. Reid, I.G. Lisle, A. Boulton, and A.D. Wittkopf, *Algorithmic determination of commutation relations for Lie symmetry algebras of PDEs*, proceedings of ISSAC'92 (Berkeley, CA, USA) (Paul S. Wang, ed.), ACM Press, 1992, pp. 63–68.
- [29] G. Reid, A. Wittkopf and A. Boulton, *Reduction of systems of nonlinear partial differential equations to simplified involutive forms*, Euro. J. Appl. Maths., 7 1996, 604–635.
- [30] Charles Riquier. *Les systèmes d'équations aux dérivées partielles*, Gauthier–Villars, Paris, 1910.
- [31] C. J. Rust, *Rankings on derivatives for elimination algorithms and formal solvability of analytic partial differential equations*, Ph.D. thesis, University of Chicago, 1998.
- [32] C.J. Rust, G.J. Reid and A.D. Wittkopf, *Existence and uniqueness theorems for formal power series solutions of analytic differential systems*, Proceedings

of the 1999 International Symposium on Symbolic and Algebraic Computation, 105–112 (electronic), ACM, New York, 1999.

- [33] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Houston, Texas, 2ed, 1979.
- [34] A. Tresse, *Sur les invariants différentiels des groupes continus de transformations*, Acta. Math. **18**, (1894) 1–88.

Chapter 5

Conclusion and future work

In Chapter 2 eigenring methods were used to develop algorithms to compute factorizations $f = gh$ for $g, h \in \mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ and LCLM-decompositions $f = \text{lcm}(g, h)$ for relatively (right) prime g, h . These algorithms run in time polynomial in $\deg_{\mathcal{D}} f$, $\deg_t f$ and q .

There are many interesting questions left for future work. For example, how to lift the factors from $\mathbb{F}_q(t)[\mathcal{D}; \sigma, \delta]$ to $\mathbb{Q}(t)[\mathcal{D}; \sigma, \delta]$. As we mentioned before $\mathbb{Q}(t)[\mathcal{D}; \sigma, \delta]$ is not a unique factorization domain. Therefore usual reconstruction methods in commutative polynomial rings do not work for $\mathbb{Q}(t)[\mathcal{D}; \sigma, \delta]$. Another interesting question is how to extend our methods to the multivariate case. As an application, in Weyl algebras, such methods for factoring elements could greatly assist in solving the corresponding partial differential operator equations.

Noncommutative Gröbner bases in PBW extensions were defined and discussed in Chapter 3. We give algorithms to construct such Gröbner bases. We extend the method to systems with more than one dependent variable by using the Drach Transformation, and use this extension in an application to the method of moving frames.

In the future we will consider Gröbner bases in q -PBW extensions, a generalization which includes most quantum groups. We plan to implement these algorithms in the symbolic language Maple.

In chapter 4 we defined non-commutative Riquier bases and extended the Riquier existence and uniqueness theory to non-linear analytic PDE systems written in terms of moving frames of non-commuting Partial Differential Operators. Many interesting questions remain. Such questions include the practical and efficient implementation of non-commutative Riquier Bases (and their non-linear generalizations) given in this thesis. Other interesting questions include the investigation of the relations between non-commutative Gröbner bases and non-commutative Riquier bases.

Appendix A

Curriculum Vitae

Yang Zhang was born in Kaifeng, one of the six oldest capitals in China, where he attended the attached elementary school, high school of Henan University and Mathematics Department at Henan University, and received his Bachelor degree. In 2000, he entered the University of Western Ontario to study Computer Algebra under the direction of Mark Giesbrecht and Greg Reid, and received a Doctor of Philosophy degree in Applied Mathematics in July 2004. He is currently an assistant professor in the Department of Mathematics and Computer Science at Brandon University, Manitoba.