Differential Algebra and Applications

MMath in Mathematics

Richard Tanburn

Hilary Term 2017

Contents

1 Introduction

Differential equations have been studied since 1675, following the invention of calculus [42]. Since then, much effort has been devoted to the analytic study of differential equations and many analytic methods exist. However, it is only relatively recently that algebraic methods have been developed, which can be used to study the geometry of a differential equation, non-dimensionalise equations or find identifiable parameter combinations in biological models algorithmically [35].

Much is known about polynomials and polynomial rings from abstract commutative algebra and there is a rich computational theory which can be used to study given systems of polynomials, in which Groebner bases play a central part [10, Chapter 2]. It is possible to adapt and generalise these methods to the study of differential equations [39].

In section 2 we present the theory of differential algebra and consider some of the implications for differential equations. In sections 3 and 4, we consider dynamical systems and show how to efficiently compute scaling symmetries and reduced dynamical systems. Finally, we give a novel extension of these algorithms to the scaling symmetries of arbitrary systems of partial differential equations (PDEs) in section 5.

1.0.1 Differential Algebra

Differential algebra was started by Joseph Ritt in 1932 and his later book "Differential Algebra" remains a key reference [39]. Ritt lays out his theory of differential polynomials, differential rings and differential ideals - with an emphasis on explicit constructions and algorithms. He goes on to define and examine the algebraic differential manifold, which begins the study of the geometry of differential equations.

One of Ritt's most influential students was Ellis Kolchin [38]. Indeed, in 1957 Irving Kaplansky, another differential algebraist, described differential algebra as being 99% the work of Ritt and Kolchin [16]. One of Kolchin's major contributions was the book "Differential Algebra and Algebraic Groups" [29]. In this work Kolchin also looks at differential field extensions, taking inspiration from Galois Theory. He goes on to define the differential analogue of the Zariski topology of affine space and describes differential algebraic geometry. Finally, differential algebraic groups are introduced in the final chapter of the book, which is used in the Galois theory of differential fields.

More recently, the constructions of Ritt and Kolchin have found applications in modern computer algebra, for instance in automated reasoning and proof [16]. However, for larger problems of interest to applied mathematicians the methods are too computationally expensive for practical use [17], hence the motivation for this work and that of Hubert and Labahn [26], which we extend. Other modern generalisations have included a schematic approach to differential algebraic geometry: see [31] for an introduction.

1.0.2 Dynamical Systems and Non-dimensionalisation

In applied mathematics, particularly mathematical biology, most modelling involves the analysis and solution of differential equations. The solution of such equations involves variables (for example, concentrations) and parameters (for example, rate constants for a reaction) which may need to be estimated or eliminated, depending on the analysis. Differential algebra has been successfully applied to the problem of structural identifiability, a pre-requisite for estimating the values of parameters from data [35].

Generally, each variable or parameter has an intrinsic unit. For example: the concentration of a chemical species might be measured in moles per litre; the birth rate of a species in a predator/prey model might be measured in number per year.

One of the very first steps in analysing these systems is to *non-dimensionalise*: the process of turning the variables and parameters into dimensionless quantities, reducing the complexity of the system. This removes the arbitrary choice of units, which introduces a scaling symmetry; we could scale time so that it is measured in seconds or years while not fundamentally changing the system.

Indeed, the Buckingham Pi theorem states that if we have n physical variables, each expressed in some combination of m fundamental dimensions, then there are m fundamental variables and the remaining $n - m$ variables can be expressed in non-dimensional quantities [46, Section 5].

There are many benefits to non-dimensionalisation [18]:

- It reduces the number of variables and parameters, making the system easier to analyse or solve analytically.
- Variables can be easier to interpret. For example, considering the ratio $\frac{\text{number of predators}}{\text{number of prey}}$ may be of more interest than the actual numbers themselves.
- Reducing the number of parameters vastly reduces the parameter space. Analysing the solutions to a model by varying $\frac{k_1}{k_2} \in \mathbb{R}$ is much cheaper than analysing it over $k_1 \times k_2 \in \mathbb{R}^2$.
- Non-dimensionalising can help with model comparison. Population dynamics of mammals and bacteria look very different over a series of days, but look very similar after adjusting for lifespans.

Usually non-dimensionalisation is done by hand and is a skill practised over many years. In addition to being cumbersome, there is no well defined method and superficially different non-dimensionalised systems can be found.

Non-dimensionalisation is an example of reducing the number of scaling symmetries of a system. More generally, variables of differential equations may exhibit extra symmetries, which are integral to the system being modelled. Analytic methods of exploiting the symmetry of an arbitrary Lie group action on the solutions to obtain reduced systems exist. However, many of these techniques are not widely known in the physics and applied mathematics communities [14, Chapter 16]. While we will not examine these, we will take some guidance from them in later sections. Further reading on this topic can be found in [13, 2, 36, 14, Chapter 16].

However these methods are computationally expensive for large systems, due to their reliance on the algorithms of differential algebra. Herein lies the motivation for restricting our attention to scaling actions - they are very efficient to find and exploit computationally. Evelyne Hubert and George Labahn have demonstrated that in the case of finding scaling symmetries it is possible to do so using only computationally efficient methods relying on linear algebra over the integers [26]. It is their methods we present and expand upon in sections 3 and 4.

First we show how non-dimensionalisation is currently done, by providing a motivational example, to highlight the value of an algorithmic approach.

Example 1.1. Consider question 2 from problem sheet 1 of the Part B course *Further Mathematical Biology* [4].

We consider the system of differential equations

$$\frac{ds}{dt} = -k_1 e_0 s + (k_1 s + k_{-1})c \tag{1}$$

$$\frac{dc}{dt} = k_1 e_0 s - (k_1 s + k_{-1} + k_2)c \tag{2}$$

where s, c are the concentrations of two chemical species, s_0, e_0 are the initial concentrations and k_1, k_{-1}, k_2 are rate constants.

The first part of the question is to non-dimensionalise with the non-dimensionalisation:

$$u = \frac{s}{s_0}, \quad v = \frac{c}{e_0}, \quad \lambda = \frac{k_{-1}}{k_1 s_0}, \quad K = \frac{k_{-1} + k_2}{k_1 s_0}, \quad \epsilon = \frac{e_0}{s_0}, \quad \sigma = k_1 s_0 t. \tag{3}$$

This gives a new system:

$$\frac{du}{d\sigma} = \frac{1}{k_1 s_0^2}\frac{ds}{dt} = -\epsilon u + \epsilon uv + \lambda \epsilon v, \tag{4}$$

$$\frac{dv}{d\sigma} = \frac{1}{k_1 e_0 s_0}\frac{dc}{dt} = u - uv - Kv. \tag{5}$$

This begs the question, how does one find a non-dimensionalisation of a system? Let us forget the non-dimensionalisation given in equation (3). One way is to choose non-dimensional quantities from the model. Clearly, s and s_0 will have the same units, so it is natural to consider $\frac{s}{s_0}$. This is how the first two non-dimensional variables of equation (3) were found.

The more general ad-hoc method is as follows:

1. **Identify the variables of your system.** That is: s, c, t.

2. **Rewrite them** as the product of a *scaling parameter* and a dimensionless variable:

$$s = \alpha u \qquad c = \beta v \qquad t = \gamma \tau \tag{6}$$

   where α, β, γ are the scaling parameters and u, v, τ are the new variables.

3. **Substitute them** back into the original equations:

$$\frac{du}{d\tau} = -k_1 e_0 \gamma u + (k_1 u + \frac{k_{-1}}{\alpha})\beta\gamma v \tag{7}$$

$$\frac{dv}{d\tau} = \frac{k_1 e_0 \alpha \gamma}{\beta} u - (k_1 \alpha u + k_{-1} + k_2)\gamma v \tag{8}$$

4. **Choose the scaling parameters** so that we remove as many constants as possible. There is no one way to do this but often some 'natural' choices will present themselves.

   The term $k_1 e_0 \gamma u$ suggests that we should choose $\gamma = \frac{1}{k_1 e_0}$. Equation (1) becomes

$$\frac{du}{d\tau} = -u + \frac{(k_1 u + \frac{k_{-1}}{\alpha})\beta v}{k_1 e_0} = -u + \frac{\beta uv}{e_0} + \frac{k_{-1}\beta v}{k_1 e_0 \alpha}$$

   which suggests the choice $\beta = e_0$.

We could then choose $\alpha = \frac{k_{-1}}{k_1}$, but instead we will look at equation (2) for a more natural choice of scaling parameter. By natural, we mean that our original quantity s represents a concentration and, while $\frac{k_1}{k_{-1}}$ has the same units after inspection, there may be a more canonical choice.

Equation (2) becomes:

$$\frac{dv}{d\tau} = \frac{\alpha}{e_0}u - \frac{k_1\alpha u + k_{-1} + k_2}{k_1 e_0}v$$

which suggests a more natural choice of $\alpha = e_0$.

After gathering constants and defining:

$$\lambda = \frac{k_{-1}}{k_1 e_0} \qquad \mu = \frac{k_2}{k_1 e_0}$$

our final system of equations becomes:

$$\frac{du}{d\tau} = -u + uv + \lambda v \tag{9}$$

$$\frac{dv}{d\tau} = u - uv + (\lambda + \mu)v \tag{10}$$

The new system given by equations (9) and (10) is much simpler, involving fewer parameters, than our original system. However, arbitrary choices were made and it is clear non-dimensionalisation is non-unique; rational invariant theory will provide an explanation in section 3.4.

2 Differential Algebra

Differential algebra is the study of differential rings, which are rings with extra structure: a set of derivations. Many of the classical results from constructive ideal theory, commutative algebra and algebraic geometry have useful analogues in differential algebra. In this section, we give the definitions and theorems that will be needed later. We will also give a brief description of some nice results which, while not needed for analysing scaling symmetries of dynamical systems, will be familiar to algebraists and have consequences for differential equations.

2.1 Definitions and First Results

For this section, proofs will generally be omitted when similar to their counterparts in commutative algebra but will be included when instructive. See Ritt [39], Kolchin [29] for the original constructions or [37, 28] for more modern treatments.

Definition 2.1. Let R be a ring. A map $\partial : R \to R$ is called a *derivation* if:

1. $\partial(a + b) = \partial(a) + \partial(b)$ and
2. $\partial(ab) = \partial(a)b + a\,\partial(b)$

for all $a, b \in R$.

Definition 2.2. A *differential ring* is a pair (R, Δ) with R a commutative unital ring and $\Delta = \{\partial_1, \dots, \partial_m\}$ a set of commuting derivatives $\partial_i : R \to R$.

By saying that Δ commutes, we mean that

$$\partial_i(\partial_j(r)) = \partial_j(\partial_i(r))$$

for all $1 \le i < j \le m$.

If R is also a field, then we call R a *differential field*.

We will often denote the differential ring (R, Δ) by just R.

Definition 2.3. If $\Delta = \{\partial\}$, then we call R an *ordinary differential ring*. In this case, for $x \in R$ we also write the image under ∂ as

$$x' := \partial(x).$$

We also define $x^{(n)} := \partial^n(x)$. By convention, $x^{(0)} = x$.

Theorem 2.4 ([28, Theorem 1.1])**.** *Given a differential ring (R, Δ), where R is also an integral domain, we can extend each $\partial_i \in \Delta$ to the field of fractions $F = \mathrm{Frac}(R)$ as follows:*

$$\partial_i\left(\frac{a}{b}\right) = \frac{(\partial_i a)b - a(\partial_i b)}{b^2}$$

Furthermore this extension is unique. We will refer to (F, Δ) as the field of fractions *of (R, Δ).*

Definition 2.5. The *constants* of a differential ring (R, Δ) is the set

$$R^\Delta = \{r \in R \mid \partial(r) = 0 \quad \forall\, \partial \in \Delta\}.$$

R^Δ is then a sub-ring of R, or a sub-field if R is a field.

Example 2.6. Let $\partial : R \to R$ be a derivation. Then:

$$\partial(1) = \partial(1 \cdot 1) = \partial(1) \cdot 1 + 1 \cdot \partial(1) = 2\,\partial(1)$$

which implies $\partial(1) = 0$. Furthermore, if $\mathbb{Q} \subset R$ then $\mathbb{Q} \subset R^\Delta$.

Example 2.7. Let $r \in R$, with R a differential ring. Since R is commutative, r and $\partial(r)$ commute and it can be easily shown by induction that:

$$\partial(r^n) = nr^{n-1}\,\partial(r)$$

Furthermore, we can extend ∂ to a derivation of $R[x]$, the polynomial ring over R, by choosing $\partial(x)$ arbitrarily and extending by linearity and the Leibniz product rule.

2.1.1 Differential Polynomials

In classical commutative algebra, one of the most important constructions is the polynomial ring $R[x_1, \dots, x_n]$. The corresponding construction in differential algebra is the ring of differential polynomials $R\{x_1, \dots, x_n\}$. This is the set of polynomials in the x_i and their derivatives $\partial_1^{e_1} \dots \partial_m^{e_m} x_i$, together with the natural derivations. We give a formal construction taking inspiration from [37, Section 2].

For the rest of this section, let $Y = \{y_1, \dots, y_n\}$ be a set of n indeterminates.

Definition 2.8. Let (R, Δ) be a differential ring.

1. Let Θ be the free monoid on Δ:

$$\Theta = \{\theta = \partial_1^{e_1} \dots \partial_m^{e_m} \mid e_i \in \mathbb{N}\}.$$

   We say that $\mathrm{Id}_\Theta \in \Theta$ is the empty word. We say that the *order* of $\theta = \partial_1^{e_1} \dots \partial_m^{e_m}$ is $e_1 + \dots + e_m$.

2. For $s \in \mathbb{N}$ we define

$$\Theta(s) := \left\{\theta = \partial_1^{e_1} \dots \partial_m^{e_m} \mid e_i \in \mathbb{N}, \sum e_i \leq s\right\}.$$

3. We define

$$\Theta Y := \{y_{i,\theta} \mid \theta \in \Theta, i = 1, \dots, n\}.$$

Definition 2.9. The *differential ring of polynomials* over (R, Δ) is the differential ring

$$(R\{Y\}, \Delta) := (R[\Theta Y], \Delta)$$

with derivations extended to $y_{i,\theta}$ by

$$\partial_i (y_{i,\theta}) := y_{i,\partial_i \theta}$$

where $\partial_i \theta$ is the concatenation in Θ.

The $y_i \in Y$ are known as *differential indeterminates* of $R\{Y\}$. The $\theta y_i \in \Theta$ are known as *differentials*.

For $\theta = \partial_1^{e_1} \dots \partial_m^{e_m}$, we will also write:

$$y_{i,(e_1,e_2,\dots,e_m)} := y_{i,\theta} = \theta y_i$$

where $y_{i,(0,\dots,0)} = y_i$ by convention.

Definition 2.10. Suppose that R is an integral domain. Then $R\{x\}$ is an integral domain and we can form its field of fractions. This is the set of *rational differential functions* and we denote it by

$$R\langle Y\rangle := \mathrm{Frac}\,(R\{x\}).$$

Definition 2.11. By the *degree* of a differential monomial, we mean the algebraic degree considered as an element of $R[\Theta Y]$.

The order (or degree) of a differential polynomial F is the maximal order (or degree) over the monomials appearing in F.

Example 2.12. For instance, working in $(\mathbb{C}\{y_1, y_2, y_3\}, \{\partial_1, \partial_2\})$:

- $y_1\, \partial_1\, y_2$ is a monomial of order 1, degree 2.
- $(\partial_2^2\, y_1)(\partial_1\, y_2) y_1^2$ is a monomial of order 2, degree 4.
- $\partial_1^3\, y_2 + \partial_2\, y_1$ is a *linear* differential polynomial (degree 1) of order 3.

2.1.2 Differential Ideals

Now we define differential ideals and quotients and note that the first isomorphism theorem holds.

Definition 2.13. An ideal $I \lhd R$ of a differential ring is a *differential ideal* if $a \in I \Rightarrow \partial a \in I$ for all $a \in R$, $\partial \in \Delta$. Equivalently, if $\partial I := \{\partial a \mid \forall a \in I\} \subset I$ for each $\partial \in \Delta$.

Given a subset $S \subset R$ we define $[S]$ to be the smallest differential ideal containing S.

Note ([29, Section 1.2]). Given a set of differential ideals $\{I_j\}_{j\in J}$ both $\bigcap_{j\in J} I_j$ and $\sum_{j\in J} I_j$ are differential ideals.

Definition 2.14. Let $I \lhd R$ be a differential ideal. Then for $\partial \in \Delta$ we can define a derivative on the quotient differential ring:

$$\partial_q(a + I) := \partial(a) + I$$

This definition is independent of the choice of representative and gives us the *quotient differential ring* $(R/I, \Delta_q)$, though we will drop the q in general.

Definition 2.15. Let $(R, \Delta_R), (S, \Delta_S)$ be differential rings. A *differential ring homomorphism* is a ring homomorphism $\varphi : R \to S$ and map of sets $\psi : \Delta_R \to \Delta_S$ that are compatible:

$$\varphi(\partial(r)) = \psi(\partial)(\varphi(r)), \quad \forall r \in R, \partial \in \Delta_R.$$

The definition for a *differential ring isomorphism* follows analogously.

Theorem 2.16 ([29, Section 1.2]). *Let $\varphi : R \to S$ be a map between differential rings. Let $I = \ker\varphi$. Then $I \lhd R$ is a differential ideal and*

$$R/I \cong \mathrm{im}(\varphi).$$

Corollary 2.17 ([37, Corollary 2.2]). *An ideal $I \lhd R$ is a differential ideal if and only if $(R/I, \partial)$ is a differential ring.*

Example 2.18. We briefly discuss how we describe differential rings and fields. If we have a differential ring R and want to add an element y such that $y' = x$ for some $x \in R$, we can formally construct our new ring as:

$$R\{y\}/[y' - x]$$

in the same way we adjoin roots of polynomials in commutative algebra. That is, $\sqrt{2}$ satisfies $t^2 - 2$, so if $\sqrt{2} \notin R$ we have an isomorphism

$$R[\sqrt{2}] \cong R[t]/(t^2 - 2).$$

In the differential case, e^x satisfies the differential polynomial $y' - y$ in a ring R, with $e^x \in R$ and $\partial = \frac{d}{dx}$. Hence we have

$$R[e^x] \cong R\{y\}/[y' - y].$$

Note that this is equivalent to considering the ring $R[y]$ and extending $\partial : R \to R$ by $\partial(y) = y$. When describing differential rings, we may use either convention.

2.1.3 Radical Differential Ideals

Now we come to one of the first real differences in the theory of commutative algebra and differential algebra: the importance of *radical differential ideals*. We will soon see that the differential ideals of 'well behaved' differential rings fail the ascending chain condition. This suggests that if we want generalisations of results from commutative algebra, like Hilbert's Nullstellensatz or Noether's Normalization Lemma, we will need to work with something else: radical differential ideals.

Even *a priori*, this will not be overly restrictive when we come to generalise algebraic geometry. Indeed, for k an algebraically closed field, we have a 1-1 correspondence

$$\{\text{algebraic subvarieties of } \mathrm{k}^n\} \leftrightarrow \{\text{radical ideals } I \lhd \mathrm{k}[x_1, \dots, x_n]\}$$

so there is reason to hope differential algebraic geometry will still behave in a similar way to regular algebraic geometry.

Definition 2.19. An ideal $I \lhd R$ is:

1. a *prime differential ideal* if it is a prime ideal and a differential ideal.
2. a *radical differential ideal* if it is a radical ideal and a differential ideal: $I = \operatorname{rad} I$.

Lemma 2.20. *The intersection of differential radical ideals is again a differential radical ideal.*

Proof. The intersection of a set of radical ideals is radical and the intersection of a set of differential ideals is a differential ideal. □

Definition 2.21. Given a subset $S \subset R$, there exists a unique minimal radical differential ideal containing S denoted by $\{S\}$.

Example 2.22 ([37, Example 1.11])**.** Constructing $\{S\}$ is not always trivial. Recall that $[S]$ is the differential ideal generated by S. In general: $\{S\} \neq \operatorname{rad}([S])$.

Consider the ordinary differential ring $R = \mathbb{Z}_2[x, y]$, with $x' = y$ and $y' = 0$, and the differential ideal $I = [x^2]$. Since $(x^2)' = 2xy = 0$, it is clear that $I = x^2 R$. If follows that $\operatorname{rad} I = x\, R$. However $\operatorname{rad} I = x\, R$ is not a differential ideal as $x' = y \notin \operatorname{rad} I$.

The following example illustrates the difference between differential ideals and radical differential ideals.

Example 2.23. Suppose $I \lhd R$ is a radical differential ideal and $ab \in I$ for $a, b \in R$. Then for each $\partial \in \Delta$:

$$\partial(a)b,\; a\,\partial(b) \in I.$$

Note $\partial(ab) = \partial(a)b + a\,\partial(b) \in I$ so in particular $a\,\partial(b)\,\partial(ab) = ab\,\partial(a)\,\partial(b) + (a\,\partial(b))^2 \in I$. As I is radical, $(a\,\partial(b))^2 \in I \Rightarrow a\,\partial(b) \in I$. This generalises to:

$$a_1 a_2 \dots a_n \in I \Rightarrow \partial(a_1) a_2 \dots a_n \in I$$

In contrast, if I is not radical, we only know that

$$\partial(ab) = \partial(a)b + a\,\partial(b) \in I.$$

From commutative algebra, we know that the radical of an ideal is the intersection of prime ideals containing it. There is an analogous statement in differential algebra.

Theorem 2.24 ([37, Theorem 1.2]). *Let $I \lhd R$ be a radical differential ideal. Then there exists a set of prime ideals $\{P_j\}_{j\in J}$ such that:*

$$I = \bigcap_{j\in J} P_j$$

In commutative algebra, the radical of an ideal plays an important role when considering affine varieties. While the radical of a differential ideal is not always a differential ideal, it is when we are working over a Ritt algebra [28, Chapter 1].

Definition 2.25. A *Ritt algebra* is a differential ring (R, Δ) such that $\mathbb{Q} \subset R$.

Lemma 2.26 ([28, Chapter 1]). *Let $I \lhd R$ be a differential ideal of a Ritt algebra with $a \in R$ such that $a^n \in I$ for some $n \in \mathbb{N}$. Then $(\partial(a))^{2n-1} \in I$ for each $\partial \in \Delta$.*

Lemma 2.27 ([28, Chapter 1]). *Let $I \lhd R$ be a differential ideal of a Ritt algebra. Then* rad I *is also a differential ideal.*

Proof. Let $a \in \operatorname{rad} I$, so $a^n \in I$ for some n. By the previous lemma, $(\partial(a))^{2n-1} \in I$ which implies $\partial(a) \in \operatorname{rad} I$. □

Corollary 2.28. *For $S \subset R$, R a Ritt algebra, we have*

$$\{S\} = \operatorname{rad}([S]).$$

Lemma 2.29 ([37, Corollary 1.1]). *Let $I \lhd R$ be maximal among proper differential ideals of a Ritt algebra. Then I is prime.*

Note ([37, Example 1.13]). Not all maximal differential ideals are prime in general. Consider $M = [x^2]$ in $R = \mathbb{Z}_2[x]$ with $\partial(x) = 1$. Then M is a maximal differential ideal but not prime

2.1.4 Ritt-Noetherian Rings

Definition 2.30. We say that a set Z of ideals in R satisfies the *ascending chain condition* if every chain (subset of Z totally ordered by inclusion) has a maximal element in the chain.

Example 2.31 ([7, Example 1.4]). The set of differential ideals of $\mathbb{Q}\{x\}$ does **not** satisfy the ascending chain condition (ACC). The chain of differential ideals

$$[xx'] \subset [xx', x'x''] \subset [xx', x'x'', x''x'''] \subset \dots$$

does not stabilise.

Definition 2.32. We say that a differential ring R is *Ritt-Noetherian* if the set of its radical differential ideals satisfies the ascending chain condition.

Theorem 2.33 (Ritt-Raudenbush, [37, Theorem 2.1]). *If R is a Ritt-Noetherian differential ring, then so is $R\{y\}$.*

Proof. See [37, Theorem 2.1] for a proof using characteristic sets, which are introduced in the next section. Alternatively, see [19, Theorem 3.23]. □

Corollary 2.34 ([37, Corollary 2.1]). *For all radical differential ideals J of $R = \mathcal{F}\{y_1, \dots y_n\}$ there exists a finite subset $S \subset R$ such that $J = \{S\}$.*

In particular, the solutions to any infinite set of differential polynomials in finitely many differential indeterminates over a differential field are perfectly described by a finite number of differential polynomials.

Lemma 2.35 ([37, Theorem 2.2])**.** *If R is a Ritt-Noetherian Ritt algebra, then for every radical differential ideal I there exists a finite set $\{P_1, \dots P_k\}$ of prime differential ideals such that:*

$$I = \bigcap_{i=1}^{k} P_i.$$

2.2 Ritt-Kolchin Theory

In this section, we focus our attention on differential polynomials and the Ritt problem.

The Generalised Ritt Problem: Given finite subset $W \subset \mathcal{F}\{y_1, \dots, y_n\}$ of differential polynomials, decompose the radical differential ideal $\{W\}$ as an irredundant intersection of prime differential ideals:

$$\{W\} = P_1 \cap P_2 \cap \dots \cap P_r$$

The Ritt problem can be formulated in many equivalent ways [15] and remains unsolved in general. However, it is possible to find a decomposition (not necessarily irredundant) algorithmically [44]. The core algorithm for this decomposition is the Rosenfeld-Groebner algorithm, introduced by Boulier *et al.*in 1995 [3], which decomposes a radical differential ideal into an intersection of prime differential ideals and gives their corresponding characteristic sets [44]. It is implemented in Maple as the ROSENFELDGROEBNER function [34, 37, Section 4].

The Rosenfeld-Groebner algorithm should not be confused with the *Ritt-Kolchin algorithm*. While the Ritt-Kolchin algorithm solves the same problem and precedes the Rosenfeld-Groebner algorithm, it relies on the solution of the so-called factorisation problem: determine whether a given ideal is prime and, if not, find two polynomials lying outside the ideal whose product lies in the ideal [15]. This is computationally expensive, so has not been implemented [15].

The Rosenfeld-Groebner and Ritt-Kolchin algorithms, which work with characteristic sets, should also not be confused with differential Groebner bases methods [33, 21, 1].

The Rosenfeld-Groebner algorithm has many uses [34]:

- Solving differential systems with an elimination ranking by putting the differential equations in triangular form.
- Solving differential systems with an orderly ranking; finding the lowest order differential equations vanishing the solutions of a system. In particular, it may be possible to find purely algebraic constraints.
- Constrained systems: mixed ranking. It is possible to obtain equivalent systems of differential equations that are unconstrained by eliminating any constraints.
- Solving PDEs: lexicographic ranking. It is possible to derive ODEs from PDEs - a process that has applications in determining Lie group symmetry.
- Simplifying systems of differential polynomials: it can remove any polynomials from a system that can be derived from the others.

For an excellent in-depth tutorial on these algorithms see [44, 29]. For a comparison of some of these algorithms see [1]. Furthermore, a modern summary with a focus on symbolic computation can be found in [45]. For an example application - automatically deriving Newton's gravitational laws from Kepler's laws - see [48].

2.2.1 Definitions and Basic Theory

For the rest of this section, let $(\mathcal{F}, \Delta = \{\partial_1, \ldots, \partial_m\})$ be a differential field, where $\mathcal{F}$ is of characteristic 0. Let

$$R = \mathcal{F}\{Y\}$$

be the polynomial ring in n differential indeterminates, where $Y = \{y_1, \ldots, y_n\}$.

Definitions in this subsection are taken from [37, Section 2.2].

Definition 2.36. A *differential ranking* $<_{\mathfrak{R}}$ on ΘY is a well-ordering (a total ordering where every non-empty subset has a least element) such that:

1. For all $u, v \in \Theta Y$ and $\partial \in \Delta$:
$$u <_{\mathfrak{R}} v \Rightarrow \partial u <_{\mathfrak{R}} \partial v.$$

2. For all $\theta \neq \mathrm{Id}_\Theta$:
$$u <_{\mathfrak{R}} \theta u.$$

For the rest of this subsection, fix a differential ranking $<_{\mathfrak{R}}$ on ΘY.

Definition 2.37. 1. The *leader* of a differential polynomial $F \in R$ is u_F, the largest differential θy_j that appears in F with respect to $<_{\mathfrak{R}}$.

2. The *initial* of F is I_F, the leading coefficient of F when written as a univariate polynomial in the leader u_F:
$$F = I_F u_F^p + a_{p-1} u_F^{p-1} + \ldots + a_0.$$

3. The *separant* of F is $S_F := \frac{\partial F}{\partial u_F}$.

Definition 2.38. For $F, G \in R$ we say that F is *partially reduced* with respect to G if none of the terms of F contain a proper derivative of u_G.

We say that F is *reduced* with respect to G if it is partially reduced with respect to G and if $u_F = u_G$ then $\deg_{u_F}(F) < \deg_{u_G}(G)$.

We say that a subset $A \subset R$ is *autoreduced* if for all $F, G \in A$ with $F \neq G$ then F is reduced with respect to G.

Lemma 2.39 ([29, Section 1.9])**.** *Every autoreduced set is finite.*

Now we introduce an ordering on autoreduced sets, so that we can introduce minimal autoreduced sets [37].

Let $A = \{A_1, \ldots, A_p\}, B = \{B_1, \ldots, B_q\}$ be autoreduced sets.

Given $<_{\mathfrak{R}}$, we define a partial order $<$ on the differential polynomials: $F < G$ if $u_F < u_G$, or if $u_F = u_G$ and $\deg_{u_F}(F) < \deg_{u_G}(G)$.

We can order the A_i by $<$, as it is a total order on autoreduced sets. We also order the B_j. We compare A and B by comparing A_i and B_i, starting from $i = 1$, declaring the largest (as a set) the smaller autoreduced set if we run out of pairs.

In particular, $A < B$ if:

1. there exists $i \leq \min(p, q)$ such that $A_i < B_i$ and $\neg(B_j < A_j)$ for $j < i$, or

2. $q < p$ and $\neg(B_j < A_j)$ for $i = 1, \ldots, \min(p, q)$.

Definition 2.40. Let $I \lhd R$ be a differential ideal. A minimal autoreduced subset of I is called a *characteristic set*.

Lemma 2.41 ([1, Theorem 5.10])**.** *Every set $S \subset R$ has a characteristic set.*

Definition 2.42 ([44, Definition 6.1])**.** A *partial remainder* of $F \in R$ with respect to a nonempty autoreduced set A is a differential polynomial $\tilde{F}$, partially reduced with respect to A, such that there exist $a \in A, m_a \in \mathbb{N}$ such that

$$\prod_{a \in A} S_a^{m_a} F - \tilde{F} \in [A],$$

where S_a is the separant of $a \in A$.

2.2.2 Algorithms

Algorithm 2.43 ([44, Procedure 6.3])**.** ***Ritt-Kolchin's Partial Remainder Theorem***
Input *A differential ranking, a non-empty autoreduced set $A \subset R$ and $F \in R$.*
Output *$\tilde{F}$, a partial remainder of F.*

Algorithm 2.44 ([44, Section 10])**.** ***Rosenfeld-Groebner***
Input *A differential polynomial ring $R = \mathcal{F}\{Y\}$, with* char $\mathcal{F} = 0$ *and $|Y| < \infty$, a ranking on ΘY and a non-empty finite set $W \subset R$.*
Output *A finite, possible empty, set $\mathcal{A}$ of autoreduced sets of R such that:*

1. *for each $A \in \mathcal{A}$, there exists a prime differential ideal P_A with characteristic set A;*

2. $\{W\} = \bigcap_{A \in \mathcal{A}} P_A$.

Using algorithm 2.43, it is also possible to test ideal membership, using the characteristic sets of each P_A [1, Section 5]. For details of the Rosenfeld-Groebner algorithm, see [3, 21, 44]. We leave out the details of these two algorithms and discuss their uses instead.

2.2.3 Examples

Example 2.45. Consider example 1.1, our non-dimensionalisation example, given by:

$$\frac{ds}{dt} = -k_1 e_0 s + (k_1 s + k_{-1})c,$$
$$\frac{dc}{dt} = k_1 e_0 s - (k_1 s + k_{-1} + k_2)c.$$

The output with respect to the elimination ranking $s <_{\mathfrak{R}} c$ is:

$$c = \frac{\frac{ds}{dt} + k_1 e_0 s}{k_1 s + k_{-1}},$$
$$\frac{d^2 s}{dt^2} = \frac{\left(\frac{ds}{dt}\right)^2 k_1 + \left(-k_1^2 s^2 - k_1\left(k_2 + 2k_{-1}\right)s - k_{-1}\left(k_1 e_0 + k_2 + k_{-1}\right)\right)\frac{ds}{dt} - e_0 k_1 k_2 s\left(k_1 s + k_{-1}\right)}{k_1 s + k_{-1}}.$$

This results in a differential equation for s and an exact solution for c as a function of s and $\frac{ds}{dt}$.

Example 2.46. Maple's ROSENFELDGROEBNER is very sensitive to the ranking. Consider the system described by:

$$\left(\frac{dx}{dt}\right)^2 - x\,(x+y) = \frac{dy}{dt} + y\,(x+y) = \left(\frac{d^2x}{dt^2}\right)^2 + \frac{d^2y}{dt^2} - 1 = 0.$$

With an elimination ranking $x <_{\mathfrak{R}} y$, ROSENFELDGROEBNER algorithm terminates within a couple of seconds, indicating the system is inconsistent and no solution exists. However the ranking $y <_{\mathfrak{R}} x$ terminates after 6 minutes producing an 'out of memory error'. [1]

2.2.4 Symmetries of Partial Differential Equations

We very briefly sketch out how the Rosenfeld-Groebner algorithm is used to find all Lie symmetries of a system of partial differential equations, using [36, Section 3]. See [6, 41, 27, 36] for more details or [32] for an exemplary symmetry analysis of Burgers' equation.

Let $u = (u_1, \ldots, u_n)$ be a set of dependent variables, $x = (x_1, \ldots x_m)$ be the set of independent variables and $F_l\left(x_i, u_j, \frac{d^k u_j}{dx_i^k}\right)$ for $k \leq N \in \mathbb{N}$ be functions defining the system of PDEs. We search for infinitesimal 1-parameter symmetries of the form:

$$\tilde{x}_i = x_i + \epsilon\xi_i(x,u),$$
$$\tilde{u}_j = u_j + \epsilon\eta_j(x,u)$$

where ϵ is the infinitesimal parameter. The *infinitesimal generator* associated with this transformation is

$$X = \sum_{i=1}^{m} \xi_i(x,u)\frac{\partial}{\partial x_i} + \sum_{j=1}^{n} \eta_j(x,u)\frac{\partial}{\partial u_j}.$$

Then the system of PDEs admits the scaling symmetry described by X if and only if

$$\left(X^{(N)}F_l\right)\left(x_i, u_j, \frac{d^k u_j}{dx_i^k}\right) = 0 \qquad \text{when} \qquad F_l\left(x_i, u_j, \frac{d^k u_j}{dx_i^k}\right) = 0$$

for each l, where $X^{(N)}$ is the *Nth prolongation* of X [36, Section 3].

This leads to an over-determined system of linear differential equations, called the *determining equations*, in terms of the ξ_i, η_j and their partial derivatives [36, Section 3]. Now the Rosenfeld-Groebner algorithm is used to simplify this system and put it in a triangular form, so that the Lie group describing the symmetries can be found. For large problems, this is computationally expensive.

2.3 Differential Algebraic Field Extensions

For the rest of this subsection, let $(\mathcal{F}, \Delta)$ be a differential field of characteristic zero.

Definition 2.47. Let $\mathcal{F} \subset \mathcal{U}$ be an extension of differential fields. Then $a \in \mathcal{U}$ is called *differentially algebraic* over $\mathcal{F}$ if there exists a non-zero $p \in \mathcal{F}\{y\}$ such that $p(a) = 0$. Equivalently, a is differential algebraic if $\{a, a', a'', \ldots\}$ is algebraically independent over $\mathcal{F}$.

[1] Experiments performed using Maple 2016.0 running on Ubuntu 16.04.2 LTS.

Example 2.48. Consider $\mathcal{F} = R(t)$, $\mathcal{U}$ the field of all infinitely differentiable functions in t.

Then $y = e^t \in \mathcal{U}$ is differentially algebraic over $\mathcal{F}$ but algebraically independent. Furthermore, any polynomial f in t is also differential algebraic over $\mathcal{F}$, since $\partial_t^{n+1} f = 0$ for n the degree of f [37, Example 3.2]. The gamma function $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$ is not differentially algebraic over $\mathcal{F}$ [19].

Definition 2.49. A differential field extension $\mathcal{F} \subset \mathcal{U}$ is *differentially finitely generated* if there exists $a_1, a_2, \ldots, a_n \in \mathcal{U}$ such that:

$$\mathcal{U} = \mathcal{F}\langle a_1, \ldots a_n\rangle := \text{Frac}\left(\mathcal{F}\{a_1, \ldots, a_n\}\right).$$

The notion of being differentially closed is similar to the notion of being algebraically closed. The construction is non-trivial and brings together model theory and differential algebra [7].

Definition 2.50.
1. Let $f_1, \ldots, f_r, g \in \mathcal{F}\{y_1, \ldots, y_n\}$ be polynomials of positive degree. The system

$$f_1 = \ldots = f_n = 0, \quad g \neq 0$$

is *consistent* if there exists a differential field extension $\mathcal{F} \subset \mathcal{U}$ and $a \in \mathcal{U}^n$ such that

$$f_1(a) = \ldots = f_n(a) = 0, \quad g(a) \neq 0$$

2. A differential field $\mathcal{F}$ is *differentially closed* if every consistent system of differential polynomials and inequations has a solution in $\mathcal{F}^n$.

Definition 2.51. A set Δ is *independent* over $\mathcal{F}$ if there exist $a_1, \ldots, a_m \in \mathcal{F}$ such that the $m \times m$ matrix

$$(\partial_i a_j)$$

is non-singular.

Theorem 2.52 ([37, Theorem 3.2])**.** *(Primitive Element Theorem for Differential Algebra) Let $\mathcal{F} \subset \mathcal{U}$ be a differentially finitely generated differential algebraic field extension. Suppose Δ is independent over $(\mathcal{F}, \Delta)$. Then there exists $a \in \mathcal{U}$ such that*

$$\mathcal{U} = \mathcal{F}\langle a\rangle$$

Example 2.53. Consider $\mathcal{F} = R \subset \mathbb{R}\langle t, e^t, \ln t\rangle = \mathcal{U}$ where the derivation is given by $\partial = \frac{d}{dt}$. We show $\mathcal{U}$ is differentially generated by $te^t \ln t$. Let $L = \mathbb{R}\langle te^t \ln t\rangle$ so clearly $L \subset \mathcal{U}$.

Now $\frac{d}{dt}(te^t \ln t) = e^t \ln t + te^t \ln t + e^t \in L$ so $e^t \ln t + e^t \in L$. Differentiating again we see $(e^t \ln t + e^t)' = e^t \ln t + \frac{e^t}{t} + e^t \in L$. Subtracting the previous two quantities gives $\frac{e^t}{t} \in L$. Differentiating and subtracting again gives $\frac{e^t}{t} - \frac{e^t}{t^2} \in L$ and $\frac{e^t}{t^2} \in L$. Hence $\frac{\frac{e^t}{t}}{\frac{e^t}{t^2}} = t \in L$. Finally, $t\frac{e^t}{t} = e^t \in L$ and $\ln t \in L$ so $\mathcal{U} \subset L$.

2.3.1 Differential Galois Theory

The study of differential algebraic fields extends naturally to a Galois theory of differential equations. Though we will not go into detail here, we will briefly describe some of the results of the area.

Differential Galois theory can be used to prove that some differential equations have no solutions that can be written in terms of 'elementary' functions, such as the Γ function discussed previously, or $u' = e^{-x^2}$ whose solution is the error function used in statistics [47]. This is done in an analogous way to algebraic Galois theory: a differential field extension $\mathcal{F} \subset \mathcal{U}$ is called a *Liouville* extension if $\mathcal{U}$ can be built up by adding a finite sequence of generators $\alpha \in \mathcal{U}$ such that α' or $\frac{\alpha'}{\alpha}$ is in the previous differential field. The differential Galois group of a Liouville extension is always solvable and certain converses also hold [11, Section 6].

Furthermore, it is also possible to derive closed form solutions in the same way the quadratic, cubic and quartic formulae exist for polynomials. In 1986 Jerald Kovacic [30] published a celebrated algorithm for finding the 'differential quadratic equation'. In particular, the algorithm found closed form solutions of

$$y'' + a(t)y' + b(t) = 0$$

for $a(t), b(t)$ rational functions of the complex independent variables t. For modern summaries, see [19, 37, 11, 7].

2.4 Differential Algebraic Geometry

While Ritt had already studied the "differential manifold" of a system of differential equations [39, Chapter 2] and used it to study its "components", it was his student Kolchin who introduced the differential version of the Zariski topology: the Kolchin topology [29, Chapter 4]. We give the details here, largely following Cassidy's more modern introduction [7].

For the rest of this subsection, let $\mathcal{F}$ be a differential field of characteristic 0, equipped with a set $\Delta = \{\partial_1, \ldots, \partial_m\}$ of differentials. Let $\mathcal{U}$ be the differential closure of $\mathcal{F}$ (or, more generally, any differential field extension). Finally, let $R = \mathcal{U}\{y_1, \ldots, y_n\}$.

Definition 2.54. For $V \subset \mathcal{U}^n$, define the set of all differential polynomials vanishing on V to be the ideal

$$\mathcal{I}(V) := \{f \in \mathcal{R} \mid f(v) = 0 \, \forall v \in V\} \lhd \mathcal{R}.$$

For $I \subset \mathcal{R}$ define the vanishing set of I to be

$$\mathcal{V}(I) := \{x \in \mathcal{U}^n \mid f(x) = 0 \, \forall f \in I\}.$$

A subset $V \subset \mathcal{U}^n$ is called *Kolchin closed* (K closed) if there exists $I \subset \mathcal{R}$ such that $V = \mathcal{V}(I)$.

Lemma 2.55. *We list some immediate properties of $\mathcal{V}$ and $\mathcal{I}$.*

1. *$\mathcal{V}$ and $\mathcal{I}$ are order reversing: $I \subset J \subset \mathcal{R} \Rightarrow \mathcal{V}(J) \subset \mathcal{V}(I)$ and $U \subset V \subset \mathcal{U}^n \Rightarrow \mathcal{I}(V) \subset \mathcal{I}(U)$.*

2. *For a subset $S \subset \mathcal{R}$, the vanishing sets of S and the differential variety generated by S are identical:*
$$\mathcal{V}(S) = \mathcal{V}([S]) = \mathcal{V}(\{S\})$$
Without loss of generality we work with ideals of R rather than arbitrary subsets.

3. *Let I, J, $\{I_j\}_{j \in K}$ be ideals of $\mathcal{R}$. Then:*
$$\mathcal{V}(\{0\}) = \mathcal{U}^n; \qquad \mathcal{V}(\{1\}) = \varnothing; \qquad \mathcal{V}(I + J) = \mathcal{V}(I) \cap \mathcal{V}(J);$$
$$\mathcal{V}(I \cap J) = \mathcal{V}(IJ) = \mathcal{V}(I) \cup \mathcal{V}(J); \qquad \mathcal{V}\left(\sum_{j \in K} I_j\right) = \bigcap_{j \in K} \mathcal{V}(I_j).$$

4. *The Kolchin closed sets $V \subset \mathcal{U}^n$ are the closed sets for a topology on $\mathcal{U}^n$.*

Proof. Follows similarly to the corresponding properties of the Zariski topology, see [10, Chapters 1, 4]. □

Definition 2.56. The topology on $\mathcal{U}^n$ whose closed sets are the Kolchin closed sets is called the *Kolchin topology*. By *differential affine space*, we mean $\mathcal{U}^n$ equipped with the Kolchin topology. We denote this topological space $\mathbb{A}^n_{\mathcal{U}}$.

While there is little difference in the definitions of the Kolchin and Zariski topologies, they behave very differently.

Example 2.57 ([7, Section 1.1])**.** We know from Part B commutative algebra that every Zariski closed subset in $\mathbb{A}^1$ is finite. However $\mathbb{Q} \subset \mathcal{V}([y'])$ is infinite.

Theorem 2.58 ([7, Theorem 1.7])**.** *(Ritt Nullstellensatz) If $\mathcal{U}$ is differentially closed, there exists an inclusion reversing bijection between the Kolchin-closed subsets of $\mathbb{A}^n$ and the set of radical differential ideals of $\mathcal{U}\{y_1, \ldots, y_n\}$, given by $\mathcal{V}$ and $\mathcal{I}$.*

Definition 2.59 ([19, Definition 6.1])**.** Let V be a Kolchin closed subset of $\mathbb{A}^n$. The *differential coordinate ring* of V is the differential ring:

$$\mathcal{U}\{V\} := \mathcal{U}\{y_1, \ldots, y_n\} / \mathcal{I}(V).$$

3 Torus Actions

As discussed in the introduction and section 2.2.4, arbitrary symmetries are computationally expensive to find and exploit. This is why we restrict our attention to non-dimensionalisation and scaling symmetries in general. In this section we discuss affine groups, representations and the algebraic theory needed later.

For the rest of this section, let k be a field of characteristic 0, $n \in \mathbb{N}$, $\mathbb{A}^n$ be the n-dimensional affine space k^n endowed with the Zariski topology.

3.1 Affine Algebraic Groups

Definition 3.1. An *affine algebraic group* is an affine algebraic variety $\mathcal{G}$ along with an identity element $e \in \mathcal{G}$, a composition map $\circ : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ and an inverse map $i : \mathcal{G} \to \mathcal{G}$ satisfying the group axioms. We will write $xy := x \circ y$ and $x^{-1} := i(x)$ for $x, y \in \mathcal{G}$. An *affine group action* on a set Z is simply a $\mathcal{G}$-action on Z.

Example 3.2 ([40, Section 8])**.**

1. Every finite group can be realised as a set of distinct points.
2. $(\mathrm{k}, +)$ is an affine group.
3. $SL(n, \mathrm{k})$, the $n \times n$ matrices of determinant 1, is an affine group. This is because the determinant can be written as a polynomial $D = \det(a_{ij}) \in \mathrm{k}[a_{ij}]$ in the n^2 coefficients of a matrix. Then $SL(n, \mathrm{k}) \cong \mathcal{V}(D - 1) \subset \mathbb{A}^{n^2}$.
4. $(\mathrm{k}^*, \cdot)$ is an affine group. We can construct an isomorphism $\mathrm{k}^* \to \mathcal{V}(xy - 1) \subset \mathbb{A}^2$ given by $x \mapsto (x, x^{-1})$.
5. $GL(n, \mathrm{k}) \cong \mathcal{V}(yD - 1) \subset \mathbb{A}^{n^2+1}$.

Definition 3.3. The r-dimensional algebraic torus is the affine algebraic group:

$$\mathbb{T} := \left(\mathrm{k}^*\right)^{\mathrm{r}} \cong \mathcal{V}(x_1 x_2 x_3 \ldots x_r y - 1) \subset \mathbb{A}^{r+1}.$$

The identity element is $(1, \ldots, 1) \in \mathbb{T}$ and multiplication is given by $*$.

The isomorphism is given by $(x_1, \ldots, x_n) \mapsto \left(x_1, \ldots, x_n, \frac{1}{x_1 \ldots x_n}\right)$.

Definition 3.4. Given an affine variety X, we can define the *coordinate ring* of X to be:

$$\mathrm{k}[X] := \frac{\mathrm{k}[z_1, \ldots, z_n]}{\mathcal{I}(X)}$$

where $\mathcal{I}(X)$ is the usual ideal of polynomials vanishing on X.

If X is an irreducible variety ($\Leftrightarrow \mathcal{I}(X)$ is prime [40, Section 2.6]) then we denote the fraction field of the coordinate ring by:

$$\mathrm{k}(X) := \mathrm{Frac}\,\mathrm{k}[X].$$

3.2 Matrix Notation for Scaling Actions

Now we consider the different actions of the algebraic torus $\mathbb{T} = \left(\mathrm{k}^*\right)^{\mathrm{r}}$ acting on $\mathbb{A}^n$, where $r, n \in \mathbb{Z}$. We represent the actions by full row rank matrices $A \in M_{r \times n}(\mathbb{N})$. In order to do so we introduce some notation from [26, Section 3].

Definition 3.5. For $x = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}$, $y = \begin{bmatrix} y_1 & \ldots & y_n \end{bmatrix}$ we define:

$$x * y = \begin{bmatrix} x_1 y_1 & x_2 y_2 & \ldots & x_n y_n \end{bmatrix}$$

to be component-wise multiplication.

Definition 3.6. Let $a = [a_1 \ldots a_r]^T \in \mathbb{Z}^r$ be a column vector.

For $\lambda = \begin{pmatrix} \lambda_1 & \ldots & \lambda_r \end{pmatrix} \in \mathbb{T}$ we define:

$$\lambda^a := \lambda_1^{a_1} \lambda_2^{a_2} \ldots \lambda_r^{a_r} \in \mathrm{k}^*$$

For $\lambda = \begin{pmatrix} \lambda_1 & \ldots & \lambda_r \end{pmatrix}$, a row vector of r indeterminates, we define:

$$\lambda^a := \lambda_1^{a_1} \lambda_2^{a_2} \ldots \lambda_r^{a_r} \in \mathrm{k}[\lambda_1, \lambda_1^{-1}, \ldots, \lambda_r, \lambda_r^{-1}]$$

Note. $\mathrm{k}[\lambda_1, \lambda_1^{-1}, \ldots, \lambda_r, \lambda_r^{-1}] \subsetneq \mathrm{k}(\lambda_1, \ldots, \lambda_r)$ is a proper subring of the field of rational functions.

Example 3.7.

1. For $a = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $\lambda = \begin{bmatrix} 6 & 3 \end{bmatrix}$ then $\lambda^a = 6^3 3^{-2} = 24$.

2. For $a = \begin{bmatrix} 2 \\ -5 \\ 4 \end{bmatrix}$, $\lambda = \begin{bmatrix} x & y & z \end{bmatrix}$ then $\lambda^a = x^2 y^{-5} z^4$ and $\lambda^{-a} = x^{-2} y^5 z^{-4}$.

3. For $a = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$, $\lambda = \begin{bmatrix} xy & x^2 y^{-1} & u + v \end{bmatrix}$ then $\lambda^a = (xy)^2 (x^2 y^{-1})^{-1} (u + v)^1 = y^3 (u + v)$.

We then extend this definition to matrices.

Definition 3.8. Let $A = (a_{ij}) \in M_{r\times n}(\mathbb{N})$, $\lambda = \begin{bmatrix}\lambda_1 & \dots & \lambda_r\end{bmatrix}$ where the λ_i could be indeterminates or elements of k^*. We define:

$$\begin{aligned}\lambda^A &:= \begin{bmatrix}\lambda^{A_{\cdot,1}} & \dots & \lambda^{A_{\cdot,n}}\end{bmatrix}\\ &= \begin{bmatrix}(\lambda_1^{a_{1,1}}\lambda_2^{a_{2,1}}\dots\lambda_r^{a_{r,1}}) & (\lambda_1^{a_{1,2}}\lambda_2^{a_{2,2}}\dots\lambda_r^{a_{r,2}}) & \dots & (\lambda_1^{a_{1,n}}\lambda_2^{a_{2,n}}\dots\lambda_r^{a_{r,n}})\end{bmatrix}.\end{aligned}$$

Lemma 3.9 ([26, Proposition 3.1])**.** *Let $A, B \in M_{r\times n}(\mathbb{N})$, $C \in M_{n\times n}(\mathbb{N})$ and λ, μ be row vectors with r components. Let $A = \begin{bmatrix}A_1 & A_2\end{bmatrix}$ be a partition of A into columns. Then:*

$$\lambda^A = \begin{bmatrix}\lambda^{A_1} & \lambda^{A_2}\end{bmatrix}, \qquad \lambda^{AC} = \left(\lambda^A\right)^C, \qquad (\lambda * \mu)^A = \lambda^A * \mu^A \quad \textit{and} \quad \lambda^{(A+B)} = \lambda^A + \lambda^B.$$

Definition 3.10. Let $A \in M_{r\times n}(\mathbb{Z})$. The *torus group* associated with A is $\mathbb{T}_A := \left(\mathrm{k}^*\right)^{\mathrm{r}}$. Then the $\mathbb{T}_A$-action on $\mathbb{A}^n$ by A is the linear group action $\cdot : \mathbb{T}_A \times \mathbb{A}^n \to \mathbb{A}^n$ defined by:

$$(\lambda, z) \mapsto \lambda \cdot z := \lambda^A * z.$$

This is a well defined group action by the previous lemma. We will often drop the subscript and write $\mathbb{T}$ when there is no ambiguity.

Definition 3.11. Since $\mathrm{k}^n \cong \mathcal{V}(0)$, we can define a $\mathbb{T}_A$-action by A on the coordinate ring $\mathrm{k}[\mathbb{A}^n] = \mathrm{k}[z_1, z_2, \dots, z_n]$. The action $\cdot : \mathbb{T}_A \times \mathrm{k}[z_1, \dots, z_n] \to \mathrm{k}[z_1, \dots, z_n]$ is defined by:

$$(\lambda \cdot f)(z_1, \dots, z_n) = f(\lambda^{-1} \cdot (z_1, \dots, z_n)).$$

Example 3.12. It is important to remember the distinction between the action on points in the topological space X and the action on the coordinate projection functions $z_i \in \mathrm{k}[X]$.

Take $A = [2]$ acting on $\mathbb{A}$ with coordinate ring $\mathrm{k}[z]$. Let $\lambda \in \mathbb{T}$. Then

$$\lambda \cdot z = \lambda^{-2} z$$

as for all points $x \in \mathbb{A}$:

$$(\lambda \cdot z)(x) = z(\lambda^{-1} \cdot x) = z(\lambda^{-2} x) = \lambda^{-2} x = \lambda^{-2} z(x).$$

3.3 Hermite Normal Form

We now introduce the Hermite normal form of a matrix. Over the integers this plays the role of the reduced row echelon form, obtained by Gaussian elimination, over a field. Hermite normal forms are very useful because information about the kernel of a map $A \in M_{m\times n}(\mathbb{Z})$ can be read off easily, such as its dimension and a basis.

The Hermite normal form is also useful because it is efficient to calculate: a key advantage of considering scaling symmetries. We will see in section 4 that all of the information we need to reduce a dynamical system can be read from two Hermite normal form decompositions.

In this section we will deal only with column Hermite normal forms, which are obtained by acting on the columns. However the *row* Hermite normal form can be formed completely analogously by transposing each definition and proof.

Definition 3.13 ([20, Section 6])**.** A matrix $A = (a_{ij}) \in M_{m\times n}(\mathbb{Z})$ is said to be in *column Hermite normal form* if:

1. the first r columns of A are nonzero;

2. for $1 \leq j \leq r$, if $a_{i_j j}$ is the first non-zero entry of column j, then $i_1 < i_2 < \ldots < i_r$;

3. $0 < a_{i_j j}$ for $1 < j < r$;

4. $0 \leq a_{i_j k} < a_{i_j j}$ for $1 \leq k < j \leq r$.

The *pivot* of each column is the first non-zero entry $a_{i_j j}$.

Remark ([43, Section 4.1])**.** An equivalent definition would be: A has block form $\begin{bmatrix} H & 0 \end{bmatrix}$ where H is lower triangular and each pivot row is non-negative with a maximal element at the end of the row .

If A has full row rank, then A is in column Hermite normal form if and only if H is a non-negative lower diagonal matrix where the maximum value of each row is found on the main diagonal.

Definition 3.14. The *elementary unimodular column operations* are given by: exchanging two columns; multiplying a column by -1; and adding an integral multiple of one column to another. We also refer to these as column operations.

Lemma 3.15. *1. A can be brought into Hermite normal form by a finite sequence of column operations on A [43, Theorem 4.1].*

2. If A has full row rank then the Hermite normal form is unique [43, Corollary 4.2a].

Corollary 3.16. *There exists a unimodular matrix U, called a* Hermite multiplier *of A, such that $H = AU$ is in Hermite normal form. This is because the elementary column operations are equivalent to multiplication on the right by a unimodular matrix.*

Note. Even if the Hermite normal form of A is unique, the Hermite multiplier might not be. This is discussed further after definition 3.17.

The column Hermite normal form can be computed in polynomial time [20]. We use the Diophantine Python package [8], which is an implementation of the method found in [20]. The reason it is easy to compute the Hermite normal form is that it is essentially calculating the highest common factors of integers, which can be done via Euclid's algorithm. However, this naive implementation results in coefficient explosion - the entries of the matrices become prohibitively large - and is not practical [9, p.69]. Hence the need for the LLL algorithm described in [20].

We introduce some non-standard terminology that will be used extensively in sections 4 and 5, taken from [26, Section 2.1].

Definition 3.17. Let $A \in M_{r\times n}(\mathbb{Z})$ have full row rank, with $r < n$, and V be a Hermite multiplier of A. Then $AV = \begin{bmatrix} H & 0 \end{bmatrix}$ where $H \in M_{r\times r}(\mathbb{Z})$ is a non-negative non-singular lower-triangular matrix. Then we can partition V and $W = V^{-1}$ as follows:

1. $V = \begin{bmatrix} \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix}$ with $\mathrm{V}_{\mathfrak{i}} \in M_{n\times r}(\mathbb{Z})$ and $\mathrm{V}_{\mathfrak{n}} \in M_{n\times(n-r)}(\mathbb{Z})$;

2. $W = \begin{bmatrix} \mathrm{W}_{\mathfrak{u}} \\ \mathrm{W}_{\mathfrak{d}} \end{bmatrix}$ with $\mathrm{W}_{\mathfrak{u}} \in M_{r\times n}(\mathbb{Z})$ and $\mathrm{W}_{\mathfrak{d}} \in M_{(n-r)\times n}(\mathbb{Z})$.

Since the Hermite multiplier is non-unique, different algorithms will give different Hermite multipliers. That is, we may perform any column operations using the columns of $\mathrm{V}_{\mathfrak{n}}$ and find another Hermite multiplier for A. In order to fix the Hermite multiplier in general, we introduce the normal Hermite multiplier [26, Proposition 2.3], which essentially puts $\mathrm{V}_{\mathfrak{n}}$ in Hermite normal form.

For the rest of this section, let $A \in M_{r\times n}(\mathbb{Z})$ be a full row rank matrix.

Definition 3.18 ([26, Section 2.2])**.** Let $V \in M_{n\times n}(\mathbb{Z})$ be a Hermite multiplier of A such that $AV = \begin{bmatrix} H & 0 \end{bmatrix}$ with $H \in M_{r\times r}(\mathbb{Z})$. V is said to be the *normal Hermite multiplier* of A if:

1. $V = \begin{bmatrix} \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix}$ with $\mathrm{V}_{\mathfrak{n}} \in M_{n\times(n-r)}(\mathbb{Z})$ in column Hermite normal form.

2. If $i_1 < \ldots < i_{(n-r)}$ are the pivot rows of $\mathrm{V}_{\mathfrak{n}}$ then

$$0 \leq (\mathrm{V}_{\mathfrak{i}})_{i_j k} < (\mathrm{V}_{\mathfrak{n}})_{i_j j}$$

for $1 \leq k \leq r$. In particular, $\mathrm{V}_{\mathfrak{i}}$ is reduced with respect to the pivot rows of $\mathrm{V}_{\mathfrak{n}}$.

Theorem 3.19 ([26, Proposition 2.3])**.** *1. A Hermite multiplier V of A is unique up to multiplication on the right by matrices of the form*

$$\begin{bmatrix} I_r & 0 \\ B & U \end{bmatrix}$$

where $U \in M_{(n-r)\times(n-r)}(\mathbb{Z})$ is unimodular and $B \in M_{(n-r)\times r}(\mathbb{Z})$.

2. The normal Hermite multiplier of A exists and is unique.

Lemma 3.20 ([9, Proposition 2.4.9])**.** *If V is a Hermite multiplier of A, then the last $n-r$ columns of V form a basis for the kernel of A.*

Lemma 3.21 ([26, Section 2.2])**.** *For a matrix $A \in M_{r\times n}(\mathbb{Z})$, the Hermite normal form and normal Hermite multiplier can be extracted from the column Hermite normal form decomposition of one $(n+r)\times n$ matrix.*

Proof. Let $V = \begin{bmatrix} \mathrm{V}_{\mathfrak{i}}\, \mathrm{V}_{\mathfrak{n}} \end{bmatrix}$ be a Hermite multiplier of A, so $\begin{bmatrix} H & 0 \end{bmatrix} = AV$ for $H \in M_{r\times r}(\mathbb{Z})$ full rank. Then

$$\begin{bmatrix} A \\ I_n \end{bmatrix} \begin{bmatrix} \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix} = \begin{bmatrix} A\,\mathrm{V}_{\mathfrak{i}} & A\,\mathrm{V}_{\mathfrak{n}} \\ \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix} = \begin{bmatrix} H & 0 \\ \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix}. \tag{11}$$

It follows that equation (11) is in Hermite normal form if and only if V is the normal Hermite multiplier.

Hence H and V can be found by a single Hermite normal form decomposition of $\begin{bmatrix} A \\ I_n \end{bmatrix}$. □

3.4 Rational Invariant Theory

In this section we show how, given a scaling matrix, we can calculate a generating set of rational invariants and re-write any invariant function in terms of them. In doing so, we draw on results from rational invariant theory [12, 25].

Definition 3.22 ([25, Definition 2.1])**.** Let G be an affine algebraic group acting on an affine variety X. Then $f \in \mathrm{k}[X]$ is said to be *invariant* if $\lambda \cdot f = f$ for all $\lambda \in G$.

The set of G-invariants is:

$$\mathrm{k}[X]^G := \{f \in \mathrm{k}[X] \mid \lambda \cdot f = f \ \ \forall \lambda \in G\}.$$

This set is a subalgebra of $\mathrm{k}[X]$ and is called the *invariant ring* of G.

If in addition X is irreducible, a *rational invariant* is an element $f \in \mathrm{k}(X)$ such that $\lambda \cdot f = f$ for all $\lambda \in G$. The set of rational invariants is:

$$\mathrm{k}(X)^G := \left\{ \frac{p}{q} \in \mathrm{k}(X) \mid p, q \in \mathrm{k}[X] \quad \text{and} \quad \frac{\lambda \cdot p}{\lambda \cdot q} = \frac{p}{q} \ \forall \lambda \in G \right\}.$$

This set is a subfield of $\mathrm{k}(X)$.

We now return our attention to the action of the torus $\mathbb{T}_A$.

Definition 3.23. For $f = \sum_{u \in \mathbb{Z}^n} a_u z^u \in \mathrm{k}[z]$ we define the *support* of f to be $Y_f = \{u \in \mathbb{Z}^n \mid a_u \neq 0\}$. Note that $|Y_f| < \infty$ by definition of the polynomial ring.

Lemma 3.24 ([26, Lemma 4.1])**.** *Let $f = \frac{p}{q} \in \mathrm{k}(z)^{\mathbb{T}_A}$ with $p, q \in \mathrm{k}[z]$ coprime. Then there exists $u \in \mathbb{Z}^n$ such that:*

$$p = \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v z^{u+v} \qquad \textit{and} \qquad q = \sum_{v \in \ker A \cap \mathbb{Z}^n} b_v z^{u+v}.$$

Proof. As $f \in \mathrm{k}(z)^{\mathbb{T}_A}$ it follows that

$$(\lambda^{-1} \cdot f)(z) = f(\lambda \cdot z) = \frac{p(\lambda \cdot z)}{q(\lambda \cdot z)} = \frac{p(z)}{q(z)} = f(z)$$

for $\lambda \in \mathbb{T}_A$. Hence $p(\lambda \cdot z) q(z) = p(z) q(\lambda \cdot z)$ in $\mathrm{k}(\lambda)[z]$. As p, q are coprime, we must have that $p(z)$ divides $p(\lambda \cdot z)$. Since these polynomials have the same degree in the z_i, we have that $p(\lambda \cdot z) = \chi(\lambda) p(z)$ for some $\chi(\lambda) \in \mathrm{k}(\lambda)$. It also follows that $\chi(\lambda) q(z) = q(\lambda \cdot z)$.

Now, for our specific group action:

$$p(\lambda \cdot z) = p(\lambda^A * z) = \sum_{w \in \mathbb{Z}^n} a_w \lambda^{Aw} z^w.$$

So if $p(\lambda \cdot z) = \chi(\lambda) p(z)$ then we must have that $Au = Av$ for any u, v in the support of p.

Fix u in the support of p, define $\chi(\lambda) = \lambda^{Au}$. Then for w in the support of p, we have that $A(w - u) = 0$ so $v = (w - u) \in \ker A$ for each w in the support of p as required.

Then

$$q(\lambda \cdot z) = \sum_{w \in \mathbb{Z}^n} b_v \lambda^{Av} z^v = \lambda^{Au} \sum_{v \in \mathbb{Z}^n} b_v z^v$$

so $Au = Av$ for v in the support of q. Then, as before, there exists $v = w - u \in \ker A$ such that $w = u + v$ for all w in the support of q. □

Example 3.25. Let $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$, so $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ generates $\ker A \cap \mathbb{Z}^2$ over $\mathbb{Z}$. Then $f = \frac{z_1 + z_2}{z_1}$ is a rational invariant. Picking $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ we see that $f = \frac{z_1 + z_2}{z_1} = \frac{z^u + z^{u-v}}{z^u}$.

Lemma 3.26 ([26, Theorem 4.2])**.** *Let $V = \begin{bmatrix} V_{\mathfrak{i}} & V_{\mathfrak{n}} \end{bmatrix}$ be a Hermite multiplier of $A \in M_{r \times n}(\mathbb{Z})$ and $W = V^{-1} = \begin{bmatrix} W_{\mathfrak{u}} \\ W_{\mathfrak{d}} \end{bmatrix}$.*

1. *The $n - r$ components of $g = \begin{bmatrix} z_1 & z_2 & \dots & z_n \end{bmatrix}^{V_{\mathfrak{n}}}$ give a generating set of $\mathrm{k}(z)^{\mathbb{T}_A}$, the set of rational invariants.*

2. *Any rational invariant $f \in \mathrm{k}(z)^{\mathbb{T}_A}$ can be written in terms of the components of g by substituting:*

$$z = \begin{bmatrix} z_1 & z_2 & \dots & z_n \end{bmatrix} \mapsto g^{\mathrm{W}_\mathfrak{d}}$$

In particular $f(z) = f\left((z^{\mathrm{V}_\mathfrak{n}})^{\mathrm{W}_\mathfrak{d}}\right)$.

Proof. First we show the components of g are rational invariants, then we show any invariant can be written in terms of them.

The columns of $\mathrm{V}_\mathfrak{n}$ span $\ker A$ by lemma 3.20, so $(\lambda^A \star z)^{\mathrm{V}_\mathfrak{n}} = \lambda^{A\,\mathrm{V}_\mathfrak{n}} \star z^{\mathrm{V}_\mathfrak{n}} = z^{\mathrm{V}_\mathfrak{n}}$. Hence the components of g are rational invariants.

Since $I_n = \mathrm{V}_\mathfrak{i}\,\mathrm{W}_\mathfrak{u} + \mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}$, for $v \in \mathbb{Z}^n$ we have $z^{\mathrm{V}_\mathfrak{i}\,\mathrm{W}_\mathfrak{u} + \mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}} = z$ where $z = \begin{bmatrix} z_1 & z_2 & \dots & z_n \end{bmatrix}$ is a vector of the indeterminates. Take $v \in \ker A$. By lemma 3.20 the columns of $\mathrm{V}_\mathfrak{n}$ form a basis for $\ker A$. Hence $v = \mathrm{V}_\mathfrak{n}\,u$ for some $u \in \mathbb{Z}^{n-r}$. As $I_n = WV = \begin{bmatrix} \mathrm{W}_\mathfrak{u}\,\mathrm{V}_\mathfrak{i} & \mathrm{W}_\mathfrak{u}\,\mathrm{V}_\mathfrak{n} \\ \mathrm{W}_\mathfrak{d}\,\mathrm{V}_\mathfrak{i} & \mathrm{W}_\mathfrak{d}\,\mathrm{V}_\mathfrak{n} \end{bmatrix}$ we have that $\mathrm{W}_\mathfrak{u}\,v = \mathrm{W}_\mathfrak{u}\,\mathrm{V}_\mathfrak{n}\,u = 0$. Hence it is also true that $z^v = z^{\mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}\,v} = g^{\mathrm{W}_\mathfrak{d}\,v}$ as $\ker A \subset \ker \mathrm{W}_\mathfrak{u}$.

Now take $f = \frac{p}{q} \in \mathrm{k}(z)^{\mathbb{T}_A}$ with $p, q \in \mathrm{k}[z]$ coprime. By lemma 3.24 there exists $u \in \mathbb{Z}^n$ such that

$$p = \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v z^{u+v} \qquad \text{and} \qquad q = \sum_{v \in \ker A \cap \mathbb{Z}^n} b_v z^{u+v}.$$

Working in $\mathrm{k}(z)$ we see that

$$p(z) = z^u \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v (z^{\mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}})^v \qquad \text{and} \qquad q(z) = z^u \sum_{v \in \ker A \cap \mathbb{Z}^n} b_v (z^{\mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}})^v$$

which implies

$$f(z) = \frac{p(z)}{q(z)} = \frac{p(z^{\mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}})}{q(z^{\mathrm{V}_\mathfrak{n}\,\mathrm{W}_\mathfrak{d}})} = \frac{p(g^{\mathrm{W}_\mathfrak{d}})}{q(g^{\mathrm{W}_\mathfrak{d}})} = f(g^{\mathrm{W}_\mathfrak{d}}).$$

□

Example 3.27. Consider $f \in \mathrm{k}(z_0, \dots, z_n)$ and $A = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}$. Then A has Hermite multiplier

$$V = \left[\begin{array}{c|c} \mathrm{V}_\mathfrak{i} & \mathrm{V}_\mathfrak{n} \end{array}\right] = \left[\begin{array}{c|cccc} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 1 & -1 & -1 & \cdots & -1 \end{array}\right].$$

so the rational functions

$$\begin{bmatrix} z_0, \dots, z_n \end{bmatrix}^{\mathrm{V}_\mathfrak{n}} = \left[\frac{z_0}{z_n}, \frac{z_1}{z_n}, \dots, \frac{z_{n-1}}{z_n}\right]$$

generate the set $\mathrm{k}(z)^{\mathbb{T}_A}$.

We have found is a generating set for the function field $\mathrm{k}\,(U_{z_n})$ of $U_{z_n} \subset \mathbb{P}^n$, the subset of projective space where $z_n \neq 0$ [40, Section 9.3]. This is because f is a well defined rational function on U_{z_n} if and only if $f(\lambda z_0, \dots, \lambda z_n) = f(z_0, \dots, z_n)$ for $\lambda \in \mathrm{k}^*$, hence our choice of A. We also have re-write rules for any function that is well defined on U_{z_n} in terms of these invariants.

3.5 Determining Maximal Scaling Actions

In this subsection, we consider the following problem: given a set of rational functions $\{f_1, f_2, \dots, f_m\} \subset \mathrm{k}\,(z_1, z_2, \dots, z_n)$, find a maximal $r \in \mathbb{N}$ and corresponding $A \in M_{r\times n}(\mathbb{Z})$ such that $\{f_i\}_{i=1}^m$ are invariant under the $\mathbb{T}_A$ action by A. We follow the methodology of [26, Section 5].

Example 3.28. Let $\mathrm{k} = \mathbb{R}$. Consider the rational function $f = \frac{z_1 + z_2^2}{z_3}$. Again, suppose k^* acts on $\mathrm{k}[z_1, z_2, z_3]$ by $A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$ such that f is invariant.

Then we must have:

$$\begin{aligned}(\lambda.f)(z_1, z_2, z_3) &= f(\lambda^{-1}.(z_1, z_2, z_3)) \\ &= f(\lambda^{-a_1} z_1, \lambda^{-a_2} z_2, \lambda^{-a_3} z_3) \\ &= \frac{\lambda^{-a_1} z_1 + \lambda^{-2a_2} z_2^2}{\lambda^{-a_3} z_3} \\ &= \frac{z_1 + z_2^2}{z_3} = f.\end{aligned}$$

Letting $\mu = \frac{1}{\lambda}$ and rearranging gives:

$$\mu^{a_3} z_1 z_3 + \mu^{a_3} z_2^2 z_3 - \mu^{a_1} z_1 z_3 - \mu^{2a_2} z_2^2 z_3 = 0$$

By viewing this equation as a polynomial in μ it is clear that f is invariant if and only if $a_1 = a_3$ and $a_3 = 2a_2$. These equations describe a one dimensional space spanned by $\begin{bmatrix} 2 & 1 & 2 \end{bmatrix}$, which is the maximal scaling matrix for this equation.

Definition 3.29. Let $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{Z}^n$, with $u \neq v$. Define a total order $<$ on $\mathbb{Z}^n$:

$$\begin{aligned} u < v \Leftrightarrow & |u_1| < |v_1|, \text{ or} \\ & |u_1| = |v_1|, |u_2| < |v_2|, \text{ or} \\ & \vdots \\ & |u_i| = |v_i| \ \ \forall i = 1, \ldots, n-1 \text{ and } |u_n| < |v_n|, \end{aligned}$$

where the $<$ above is the usual ordering on $\mathbb{Z}$. In other words, we compare the magnitude of entries, starting from the left.

Definition 3.30. Let $f = \frac{p}{q} \in \mathrm{k}(z)$ with $p = \sum_{u \in \mathbb{N}^n} a_u z^u, q = \sum_{v \in \mathbb{N}^n} b_v z^v \in \mathrm{k}[z]$ coprime. Note $2 \leq N = |Y_p \cup Y_q| < \infty$. Pick $w \in Y_q$ minimal with respect to $<$. Define the *exponent matrix* of f to be $K_f \in M_{(N-1) \times n}(\mathbb{Z})$, the matrix with columns made up of $v - w$ for $v \in (Y_p \cup Y_q) \setminus \{w\}$.

Note. If $f \in \mathrm{k}[z]$ then we take $q = 1$ and $w = (0, \ldots, 0)^T$.

Example 3.31.

1. $f = \frac{z_1 z_2}{1}$, $w = (0,0)^T$, $K_f = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

2. $f = \frac{z_1 z_2 + z_3}{z_2 z_3}$, then $q = z_2 z_3 = (z_1, z_2, z_3)^w$ where $w = (0,1,1)^T$. So $K_f = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -1 & 0 \end{bmatrix}$.

3. $f = \frac{z_1 z_2 + z_3 z_4}{z1 + z4}$, $Y_q = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$ so we pick $w = (0,0,0,1)^T$. $K_f = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$

Lemma 3.32 ([26, Section 5]). *Let $A \in M_{r \times n}(\mathbb{Z})$, $f \in \mathrm{k}(z)$. Then f is a rational invariant with respect to $\mathbb{T}_A$ if and only if:*

$$AK_f = 0.$$

Proof. Let $f = \frac{p}{q} \in \mathrm{k}(z)$ with $p = \sum_{u\in\mathbb{N}^n} a_u z^u, q = \sum_{v\in\mathbb{N}^n} b_v z^v \in \mathrm{k}[z]$ coprime.

⇒ By lemma 3.24, we have immediately have that $u - v \in \ker A$ for $u, v \in Y_p \cup Y_q$.

⇐ Then pick w in the support of q minimal with respect to $<$ defined in definition 3.29. Note $AK_f = 0$ if and only if $A(u-w) = 0$ for all $u \in Y_p \cup Y_q$. Hence

$$\begin{aligned}(\lambda^{-1}.f)(z) &= f(\lambda^A * z)\\ &= \frac{\sum_{u\in\mathbb{N}^n} a_u \lambda^{Au} z^u}{\sum_{v\in\mathbb{N}^n} b_v \lambda^{Av} z^v}\\ &= \frac{\sum_{u\in\mathbb{N}^n} \frac{a_u \lambda^{Au} z^u}{\lambda^{Aw}}}{\sum_{v\in\mathbb{N}^n} \frac{b_v \lambda^{Av} z^v}{\lambda^{Aw}}}\\ &= \frac{\sum_{u\in\mathbb{N}^n} a_u \lambda^{A(u-w)} z^u}{\sum_{v\in\mathbb{N}^n} b_v \lambda^{A(v-w)} z^v}\\ &= \frac{\sum_{u\in\mathbb{N}^n} a_u z^u}{\sum_{v\in\mathbb{N}^n} b_v z^v}\\ &= f(z).\end{aligned}$$

□

Now that we have found necessary and sufficient conditions for a rational function f to be invariant with respect to $\mathbb{T}_A$, we use the following lemma to find a maximal A.

Lemma 3.33 ([26, Proposition 5.1]). *Let $K \in M_{n\times m}(\mathbb{Z})$ and $U \in M_{n\times n}(\mathbb{Z})$ unimodular such that*

$$UK = \begin{bmatrix} K_0 \\ 0 \end{bmatrix}$$

is in row Hermite normal form, with $K_0 \in M_{(n-r)\times m}(\mathbb{Z})$ of full row rank, so that there are exactly r zero rows. Let A be the last r rows of U. Then

1. *$AK = 0$;*
2. *the column Hermite normal form of A is $\begin{bmatrix} I_r & 0 \end{bmatrix}$;*
3. *an integer matrix B satisfies $BK = 0$ if and only if there exists an integer matrix M such that $B = MA$.*

Proof. 1. $UK = \begin{bmatrix} * \\ A \end{bmatrix} K = \begin{bmatrix} K_0 \\ 0 \end{bmatrix}$.

2. Note $I_n = UU^{-1} = \begin{bmatrix} * \\ A \end{bmatrix} U^{-1}$. As U^{-1} is unimodular, acting on the right of A, we have that A is equivalent to $\begin{bmatrix} 0 & I_r \end{bmatrix}$ by elementary column operations. Permuting the columns shows that the column Hermite normal form of A is $\begin{bmatrix} I_r & 0 \end{bmatrix}$.

3. ⇒ A $\mathbb{Z}$-basis for the left kernel of K (which is defined as $\{x \in \mathbb{Z}^n \mid xK = 0\}$) is given by A, by taking the row version of lemma 3.20. Hence if $BK = 0$, each row of B is a $\mathbb{Z}$-linear combination of the rows of A. That is, there exists an integer matrix M such that $B = MA$.

   ⇐ $BK = MAK = 0$.

□

Corollary 3.34. *Let A be as in lemma 3.33.*

1. *A has full row rank.*

2. *r is maximal and A is a maximal scaling matrix. This is because any other matrix $B = MA$ such that $BK = 0$ will have row rank less than or equal to r.*

Example 3.35. We revisit example 3.28 and consider the rational function: $f = \frac{z_1+z_2^2}{z_3}$.

Then $K_f = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & -1 \end{bmatrix}$. The row Hermite normal form decomposition is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ -2 & -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Hence $A = \begin{bmatrix} -2 & -1 & -2 \end{bmatrix}$ describes a maximal scaling action on k^3 under which f is invariant. This is equivalent to our previous answer, showing it is maximal.

Corollary 3.36 ([26, Section 5])**.** *Let $F = (f_1, f_2, \ldots, f_m) \in (\mathrm{k}(z))^m$ be a vector of rational functions. Define $K = \begin{bmatrix} K_{f_1} & K_{f_2} & \ldots & K_{f_m} \end{bmatrix}$. Then*

1. *For any integer matrix B,*

$$BK = 0 \Leftrightarrow f_i \in \mathrm{k}(z)^{\mathbb{T}_B} \ \ \forall i = 1, \ldots, m.$$

2. *Suppose U is a unimodular matrix such that $UK = \begin{bmatrix} K_0 \\ 0 \end{bmatrix}$ is in row Hermite normal form with exactly r zero rows. Let A be the last r rows of U. Then A is a maximal scaling matrix for which F is invariant.*

Example 3.37. 1. Define

$$F = (F_1, F_2) = \left(\frac{z_1 z_2 + z_3 z_4}{z_1 z_3}, \frac{z_1 z_4 + z_3}{z_2 + z_5} \right).$$

Then

$$K = \left[\begin{array}{c|c} K_{F_1} & K_{F_2} \end{array} \right] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{array} \right].$$

The row Hermite normal form decomposition is given by:

$$UK = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 \\ -1 & -2 & -2 & -1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = H.$$

Hence our maximal scaling action is given by $A = \begin{bmatrix} 1 & 2 & 2 & 1 & 2 \end{bmatrix}$. Indeed for $\lambda \in \mathbb{T}_A = \mathrm{k}^*$:

$$\lambda^{-1} \cdot \left(\frac{z_1 z_2 + z_3 z_4}{z_1 z_3} \right) = \frac{(\lambda z_1)(\lambda^2 z_2) + (\lambda^2 z_3)(\lambda z_4)}{(\lambda z_1)(\lambda^2 z_3)} = \frac{\lambda^3 (z_1 z_2 + z_3 z_4)}{\lambda^3 z_1 z_3} = \frac{z_1 z_2 + z_3 z_4}{z_1 z_3}.$$

4 Scaling Symmetries of Dynamical Systems

In this section, we show how to find scaling symmetries of a dynamical system and produce an equivalent reduced system, as in [26].

We consider systems of ordinary differential equations of the form

$$\frac{dz}{dt} = G(t, z), \tag{12}$$

where $z = (z_1(t), \ldots, z_n(t))$ is a vector of functions of t and $G = (G_1(t, z), \ldots G_n(t, z)) \in \mathrm{k}(t, z)^n$ is a rational map $\mathrm{k} \times \mathrm{k}^n \to \mathrm{k}^n$. Since we are working with rational functions, we can write

$$\frac{dz}{dt} = G(t, z) = \frac{z * F(t, z)}{t}. \tag{13}$$

We will mostly work with equation (13) as it turns out that the solutions are invariant with respect to a $\mathbb{T}_A$-action if and only if F is a rational invariant of $\mathbb{T}_A$.

For the rest of this section, $\mathrm{k} = \mathbb{Q}$ and $\mathbb{T} = (\mathrm{k}^*)^{\mathrm{r}}$. Define

$$\frac{dz}{dt} := \left(\frac{dz_1}{dt}, \frac{dz_2}{dt}, \ldots, \frac{dz_n}{dt} \right), \quad z^{-1} := \left(z_1^{-1}, z_2^{-1}, \ldots, z_n^{-1} \right).$$

For $\lambda \in \mathbb{T}$ and vectors $x = (x_1, \ldots, x_n)$ for which $\lambda \cdot x_i$ is defined, define:

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

4.1 The Analytic Interpretation

As noted in [26, 12, Section 5.8], it is possible to find invariants and equivalent dynamical systems without knowing any differential algebra. Let $A = \begin{bmatrix} a_0 & a_1 & \ldots & a_n \end{bmatrix} \in M_{1 \times (n+1)}(\mathbb{Z})$ and $\mathbb{T}_A$ act on $\mathrm{k}(t, z)$.

Definition 4.1. 1. We say $(\tilde{t}, \tilde{z}_1, \ldots, \tilde{z}_n)$ is a solution to equation (13) if

$$\frac{d\tilde{z}}{d\tilde{t}} = G(\tilde{t}, \tilde{z}) = \frac{\tilde{z} * F(\tilde{t}, \tilde{z})}{\tilde{t}}. \tag{14}$$

2. We say that $\mathbb{T}_A$ defines a scaling symmetry for equation (13) if for all solutions (t, z) of equation (13) and $\lambda \in \mathbb{T}_A$, $\lambda^{-1} \cdot (t, z)$ is also a solution.

Theorem 4.2 ([26, Section 6.1]). *$\mathbb{T}_A$ defines a scaling symmetry for equation* (13) *if and only if $F_i \in \mathrm{k}(t, z)^{\mathbb{T}_A}$ is a rational invariant for each $i = 1, \ldots, n$.*

Proof. If (t, z) is a solution to equation (13) then $\lambda^{-1} \cdot (t, z)$ is a solution if:

$$\begin{aligned} \frac{d(\lambda^{a_i} z_i)}{d(\lambda^{a_0} t)} &= G_i\left(\lambda^{a_0} t, \lambda^{a_1} z_1, \ldots, \lambda^{a_n} z_n\right) \\ &= \frac{\lambda^{a_i}}{\lambda^{a_0}} \frac{z_i}{t} F_i\left(\lambda^{a_0} t, \lambda^{a_1} z_1, \ldots, \lambda^{a_n} z_n\right) \end{aligned}$$

for each $i = 1, \ldots, n$. Furthermore:

$$\frac{d(\lambda^{a_i} z_i)}{d(\lambda^{a_0} t)} = \lambda^{a_i} \frac{dz_i}{dt} \frac{dt}{d\left(\lambda^{a_0} t\right)} = \lambda^{a_i} \frac{dz_i}{dt} \frac{1}{\frac{d(\lambda^{a_0} t)}{dt}} = \frac{\lambda^{a_i}}{\lambda^{a_0}} \frac{dz_i}{dt}. = \frac{\lambda^{a_i}}{\lambda^{a_0}} \frac{z_i F_i(t, z)}{t}.$$

Hence $\lambda^{-1} \cdot (t, z)$ is a solution to equation (12) if and only if $F_i(t, z) = F_i\left(\lambda^{-1} \cdot (t, z)\right) \in \mathrm{k}\left(t, z\right)^{\mathbb{T}_A}$ is a rational invariant. □

4.2 The Differential Algebraic Geometry Interpretation

We consider how we might describe scaling symmetries using differential algebra. While we reach the same conclusion as the analytic methods of [26, Section 6.1] and section 4.1 we do so by a novel approach which is more general and satisfactory.

For the rest of this subsection, let $\mathcal{F}$ be an ordinary differential field of characteristic 0, $\mathcal{F} \subset \mathcal{U}$ be a differential field extension with $t \in \mathcal{U}$, $t' = 1$ and $R = \mathcal{F}\{y_1 \ldots, y_n\}$. Consider $\Sigma \subset R$ a finite set of differential polynomials and $I = [\Sigma]$. Definition 4.1 then inspires the following definition.

Definition 4.3. The torus $\mathbb{T}_A$ defines a scaling symmetry for Σ if the $\mathbb{T}_A$-action on $\mathbb{A}^n$ descends to a well defined $\mathbb{T}_A$-action on the differential affine variety $\mathcal{V}(I)$.

4.2.1 Scaling Symmetries of Dependent Variables

To begin with, we restrict ourselves to torus actions that only act on the dependent variables. Consider the system of differential equations

$$\frac{dz}{dt} = G(z) \tag{15}$$

given by rational $G_i = \frac{P_i}{Q_i} \in \mathcal{U}(z)$ for $P_i, Q_i \in \mathcal{U}[z_1, \ldots, z_n]$ as in equation (12). This corresponds to a differential ideal

$$I = [Q_1(y)y_1' - P_1(y), \ldots, Q_n(y)y_n' - P_n(y)] \lhd \mathcal{U}\{y_1, \ldots, y_n\} \tag{16}$$

where $y = (y_1, \ldots, y_n)$ are differential indeterminates. Furthermore, let

$$V = \mathcal{V}(I) \cap \{z \in \mathbb{A}^n_{\mathcal{U}} \mid Q_i(z) \neq 0 \, \forall 1 \leq i \leq n\} \subset \mathbb{A}^n_{\mathcal{U}}$$

be the solutions of the system in $\mathcal{U}^n$ and suppose $V \neq \varnothing$. Note that V is isomorphic to an affine differential variety by the same trick used repeatedly in example 3.2.

Lemma 4.4. *Let $\mathbb{T}_A = \mathrm{k}^*$ act non-trivially on $\mathbb{A}^n_{\mathcal{U}}$ by $A = \begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix}$. In particular, for $\lambda \in \mathbb{T}_A$ and $p \in \mathcal{V}(I)$ we have: $\lambda \cdot (p_1, \ldots, p_n) = (\lambda^{a_1} p_1, \ldots, \lambda^{a_n} p_n)$.*

Then $\mathbb{T}_A$ acts on V if and only if

$$Q_i(\lambda \cdot z) \neq 0 \quad \forall 1 \leq i \leq n \qquad \textit{and} \qquad G(\lambda \cdot z) = G(\lambda^A * z) = \lambda^A * G(z).$$

Proof. The $\mathbb{T}_A$-action descends to an action on V if and only if $\lambda \cdot (z_1, \ldots, z_n) \in V$ for all $\lambda \in \mathrm{k}^*$ and all $z \in V$.

$\Rightarrow$ If $\mathbb{T}_A$ acts on V, then for each $i = 1, \ldots, n$:

1. $Q_i(\lambda \cdot z) \neq 0$ as $\lambda \cdot z \in V$,
2. $Q_i(\lambda \cdot z)\,(\lambda^{a_i} z_i)' - P_i(\lambda \cdot z) = \lambda^{a_i} Q_i(\lambda \cdot z) z_i' - P_i(\lambda \cdot z) = 0$.

Then

$$G_i(\lambda \cdot z) = \frac{P_i(\lambda \cdot z)}{Q_i(\lambda \cdot z)} = \lambda^{a_i} z_i' = \lambda^{a_i} \frac{P_i(z)}{Q_i(z)} = \lambda^{a_i} G_i(z) \tag{17}$$

as required.

$\Leftarrow$ As $Q_i(\lambda \cdot z) \neq 0$ for each i, we have $\lambda \cdot z \in \{z \in \mathbb{A}^n_{\mathcal{U}} \mid Q_i(z) \neq 0 \,\forall 1 \leq i \leq n\}$.
Also equation (17) holds, so

$$Q_i(\lambda \cdot z)\,(\lambda^{a_i} z_i)' - P_i(\lambda \cdot z) = \lambda^{a_i} Q_i(\lambda \cdot z) z_i' - \lambda^{a_i} Q_i(\lambda \cdot z) \frac{P_i(z)}{Q_i(z)} = \lambda^{a_i} Q_i(\lambda \cdot z)\,(z_i' - G_i(z)) = 0$$

which implies $\lambda \cdot z \in \mathcal{V}(I)$. □

Corollary 4.5. *As V is infinite (it is non-empty, the action of $\mathbb{T}_A$ is faithful and $\mathbb{T}_A$ is an infinite group), we know from commutative algebra that $\mathbb{T}_A = \mathrm{k}^*$ acts on $\mathbb{A}^n_{\mathcal{U}}$ if and only if $\lambda \cdot G = \lambda^{-A} * G$ for all $\lambda \in \mathbb{T}_A$.*

Corollary 4.6. *$\mathbb{T}_A$ acts on V if and only if for each $i = 1, \ldots, n$:*

$$\lambda \cdot F_i = \lambda \cdot \left(\frac{tG_i}{z_i} \right) = \frac{\lambda^{-a_i} t G_i}{\lambda^{-a_i} z_i} = \frac{tG_i}{z_i} = F_i \in \mathcal{U}(y_i)^{\mathbb{T}_A}.$$

That is $\lambda \cdot F = F$ is invariant.

Example 4.7. Consider the system

$$\frac{d(z_1, z_2)}{dt} = \left(z_1 \left(1 - \frac{z_1}{z_2} \right), z_2 \left(1 + \frac{z_1}{z_2} \right) \right). \tag{18}$$

We know by corollary 4.6 and the methods of section 3.5 that $\mathbb{T} = \mathrm{k}^*$ acts on z_1, z_2 by $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$. Let $\mathcal{F} = \mathbb{C}$ and let $\mathcal{U} = \mathcal{F}(t)$. We have $P_1 = z_1(z_2 - z_1)$, $P_2 = z_2(z_2 + z_1)$, $Q_1 = Q_2 = z_2$, so the system corresponds to differential ideal:

$$I = [y_2 y_1' - y_1(y_2 - y_1), y_2 y_2' - y_2(y_2 + y_1), y_2 y_3 - 1] \subset \mathcal{U}\{y_1, y_2, y_3\}$$

where we are considering the system embedded in $\mathbb{A}^3$ so that we can ensure the z_2 coordinate does not vanish, making concrete the identification of V with an affine variety. This is an alternative way of acting on the solutions, avoiding the region where $Q_i = 0$, to the methods of lemma 4.4.

Then it is easy to check that $(z_1, z_2, z_2^{-1}) \in \mathcal{V}(I)$ implies $(\lambda z_1, \lambda z_2, \lambda^{-1} z_2^{-1}) \in \mathcal{V}(I)$ so the action descends to an action on V. Furthermore

$$\lambda \cdot G_1 = \frac{\lambda \cdot P_1}{\lambda \cdot Q_1} = \frac{\lambda^{-2} z_1(z_2 - z_1)}{\lambda^{-1} z_2} = \lambda^{-1} \frac{P_1}{Q_1} = \lambda^{-1} G_1, \quad \text{and} \quad \lambda \cdot G_2 = \lambda^{-1} G_2.$$

4.2.2 General Case

The differential algebra we have used so far does not allow us to act on t, since there is no way to act on the derivation in the manner required to make sense of definition 4.1. In order to generalise and consider actions on the independent variable, we employ a trick from [26, Section 6.2]. We introduce a dummy variable $z_0(t)$ which will behave like a scaled version of our time variable: $z_0(t) = \mu t$ for some constant $\mu \in \left(\mathcal{F}^{\Delta}\right)^*$. The z_0 we desire is characterised by the differential equation $\frac{d}{dt}\left(\frac{z_0}{t}\right) = \frac{tz_0' - z_0}{t^2} = 0$.

Since $\frac{dz_i}{dz_0} = \frac{dz_i}{dt}\frac{dt}{dz_0} = \frac{z_i'}{z_0'}$, instead of considering the ideal in equation (16) we consider the differential ideal:

$$J = [Q_1(y_0, y)y_1' - P_1(y)y_0', \dots, Q_n(y_0, y)y_n' - P_n(y_0, y)y_0', ty_0' - y_0] \lhd \mathcal{U}\{y_0, y_1, \dots, y_n\} \quad (19)$$

where we make the substitution $t \mapsto y_0$ in each of the G_i.

By repeating the calculations in the proof of lemma 4.4 for J we see that $\mathbb{T}_A$ acts on $\mathcal{V}(J)$ if and only if for $i = 1, \dots n$:

$$G_i(\lambda \cdot \bar{z}) = \frac{\lambda^{a_i}}{\lambda^{a_0}} G_i(\bar{z}) \Leftrightarrow F_i(\lambda \cdot \bar{z}) = F_i(\bar{z}). \quad (20)$$

This agrees with theorem 4.2 and provides a rigorous geometric interpretation of scaling symmetries.

Example 4.8. Consider the ODE:

$$\frac{dz}{dt} = \left(\frac{z}{t}\right)^2 \quad (21)$$

which has solutions $z(t) = \frac{t}{ct+1}$, where c is a constant. Let $\mathcal{F} = \mathbb{C} \subset \mathbb{C}(t) = \mathcal{U}$ with $t' = 1$.

The ideal J corresponding to equation (21) is given by $J = [F, G]$, where $s(t)$ is our dummy independent variable, $F(s, z) = s^2 z' - z^2 s'$ and $G(s, z) = s't - s$.

Then for $c \in \mathbb{C}$ we have:

$$F\left(t, \frac{t}{ct+1}\right) = \frac{t^2}{(ct+1)^2} - \left(\frac{t}{ct+1}\right)^2 = 0 \qquad \text{and} \qquad G\left(t, \frac{t}{ct+1}\right) = 0.$$

Furthermore, for $\lambda \in \mathbb{C}^*$:

$$F\left(\lambda t, \lambda \frac{t}{ct+1}\right) = \frac{\lambda^3 t^2}{(ct+1)^2} - \lambda\left(\frac{\lambda t}{ct+1}\right)^2 = 0,$$
$$G\left(\lambda t, \lambda \frac{t}{ct+1}\right) = \lambda t - \lambda t = 0$$

so $\left(\lambda t, \lambda \frac{t}{ct+1}\right) \in \mathcal{V}(\Sigma)\ \forall \lambda \in \mathbb{C}^*$.

Recall that $y(t) = \frac{t}{ct+1}$ is a function of t and that $s = \lambda t \Rightarrow t = \frac{s}{s'}$. It follows that given a solution $\left(s, y\left(\frac{s}{s'}\right)\right)$ of equation (21) and $\lambda \in \mathbb{C}^*$:

$$\left(\lambda s, \lambda y\left(\frac{\lambda s}{(\lambda s)'}\right)\right)$$

is also a solution to equation (21).

4.3 Construction of Reduced Systems

4.3.1 Torus Action on Dependent Variables

In the previous section we saw that finding the maximal torus action leaving the solutions to equation (13) invariant is equivalent to finding the maximal torus actions leaving $\{F_i\}_{i=1}^n$ invariant, where

$$\frac{dz}{dt} = G(t,z) = \frac{z * F(t,z)}{t}.$$

Now we use symmetries of the system to find a simpler system, involving fewer variables, such that there is a correspondence of solutions between the original and reduced systems. In particular, if a torus group $\mathbb{T}_A$ acts on the solutions of equation (13) there will be rational invariants of this map, which will have dynamics of their own. Rational invariant theory tells us we can re-write the $\{F_i\}_{i=1}^n$ in terms of these invariants.

Lemma 4.9 ([26, Lemma 6.1])**.** *Suppose* $z(t) = \begin{bmatrix} z_1(t) & z_2(t) & \dots & z_n(t) \end{bmatrix}$ *is a vector of functions in* t *and* $A \in M_{n\times r}(\mathbb{Z})$*. Then:*

$$\frac{d}{dt}\left(z^A\right) = z^A * \left(\left(z^{-1} * \frac{dz}{dt}\right).A\right)$$

Proof. Suppose that $a = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}^T$ is a column of A. Then by the product and chain rule:

$$\begin{aligned}
\frac{d(z^a)}{dt} &= \frac{d(z_1^{a_1} z_2^{a_2} \dots z_n^{a_n})}{dt} \\
&= a_1 \frac{z^a}{z_1}\frac{dz_1}{dt} + \dots + a_n \frac{z^a}{z_n}\frac{dz_n}{dt} \\
&= z^a \left(\left(z^{-1} * \frac{dz}{dt}\right).a\right).
\end{aligned}$$

Applying to each column of A gives the result. □

Theorem 4.10 ([26, Theorem 6.3])**.** *Consider a map* $F : \mathbb{A} \times \mathbb{A}^n \to \mathbb{A}^n$ *that is invariant under the group action of* $\mathbb{T}_A$ *on* $\mathbb{A}^n$ *given by* $A \in M_{r\times n}(\mathbb{Z})$ *(so* $\mathbb{T}_A$ *acts trivially on* t*). Let* $V = \begin{bmatrix} \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix}$ *be the Hermite multiplier of* A *with inverse* $W = \begin{bmatrix} \mathrm{W}_{\mathfrak{u}} \\ \mathrm{W}_{\mathfrak{d}} \end{bmatrix}$.

1. *If* $z(t)$ *is a solution of* $\frac{dz}{dt} = \frac{z*F(t,z)}{t}$ *where none of the components vanish, then*

   $$\begin{bmatrix} x(t) & y(t) \end{bmatrix} = \begin{bmatrix} z^{\mathrm{V}_{\mathfrak{i}}} & z^{\mathrm{V}_{\mathfrak{n}}} \end{bmatrix}$$

   is a solution to the dynamical system:

   $$\frac{dy}{dt} = \frac{y}{t} * \left(F\left(t, y^{\mathrm{W}_{\mathfrak{d}}}\right).\mathrm{V}_{\mathfrak{n}}\right), \tag{22}$$
   $$\frac{dx}{dt} = \frac{x}{t} * \left(F\left(t, y^{\mathrm{W}_{\mathfrak{d}}}\right).\mathrm{V}_{\mathfrak{i}}\right). \tag{23}$$

2. *If* $y(t), x(t)$ *are solutions to equations* (22) *and* (23)*, respectively, and none of their components vanish then*

   $$z(t) = \begin{bmatrix} x(t) & y(t) \end{bmatrix}^W$$

   is a solution to the dynamical system:

   $$\frac{dz}{dt} = \frac{z * F(t,z)}{t}.$$

Proof. 1. Since F is $\mathbb{T}_A$-invariant, lemma 3.26 implies that

$$F(t,z) = F\left(t, \left(z^{V_{\mathfrak{n}}}\right)^{W_{\mathfrak{d}}}\right) = F\left(t, y^{W_{\mathfrak{d}}}\right)$$

when considering $F \in \mathrm{k}(t)(z)$. Lemma 4.9 implies that

$$\begin{aligned}\frac{dy}{dt} &= \frac{d}{dt}\left(z^{V_{\mathfrak{n}}}\right)\\ &= z^{V_{\mathfrak{n}}} * \left(\left(z^{-1} * \frac{dz}{dt}\right).V_{\mathfrak{n}}\right)\\ &= y * \left(\frac{F(t,z)}{t}.V_{\mathfrak{n}}\right)\\ &= \frac{y}{t} * \left(F(t,y^{W_{\mathfrak{d}}}).V_{\mathfrak{n}}\right).\end{aligned}$$

Similarly $\frac{dx}{dt} = \frac{x}{t} * (F(t,y^{W_{\mathfrak{d}}}).V_{\mathfrak{i}})$.

2. From equations (22) and (23) we know that:

$$\frac{d}{dt}\left(\begin{bmatrix}x & y\end{bmatrix}\right) = \frac{1}{t}\begin{bmatrix}x & y\end{bmatrix} * \left(F\left(t,y^{W_{\mathfrak{d}}}\right)V\right).$$

5 Furthermore, $W\,V_{\mathfrak{n}} = \begin{bmatrix}0\\ I_{n-r}\end{bmatrix}$ so $z^{V_{\mathfrak{n}}} = y$. Finally, by lemma 3.26, $F(t,y^{W_{\mathfrak{d}}}) = F(t,z)$. Bringing these facts together with lemma 4.9 gives:

$$\begin{aligned}\frac{dz}{dt} &= \frac{d}{dt}\left(\begin{bmatrix}x & y\end{bmatrix}^W\right)\\ &= \begin{bmatrix}x & y\end{bmatrix}^W * \left(\left(\begin{bmatrix}x & y\end{bmatrix}^{-1} * \frac{d}{dt}\left(\begin{bmatrix}x & y\end{bmatrix}\right)\right)W\right)\\ &= z * \left(\frac{1}{t}\left(F\left(t,y^{W_{\mathfrak{d}}}\right)V\right)W\right)\\ &= \frac{z}{t} * F\left(t,z\right).\end{aligned}$$

□

Equation (22) describes the *reduced system* under $\mathbb{T}$; it expresses the dynamics of the $n-r$ rational invariants under the group action.

Note that given solutions to equation (22), it is possible to solve equations 23 by integration:

$$x^{-1} * \frac{dx}{dt} = \left(F\left(t,y^{W_{\mathfrak{d}}}\right)V_{\mathfrak{i}}\right) \Rightarrow x = \exp\left(\int F\left(t,y^{W_{\mathfrak{d}}}\right)V_{\mathfrak{i}}\,dt\right).$$

Example 4.11. Consider the system:

$$\frac{d(z_1,z_2)}{dt} = \left(z_1\left(1-\frac{z_1}{z_2}\right), z_2\left(1+\frac{z_1}{z_2}\right)\right) = \frac{(z_1,z_2)}{t} * \left(t\left(1-\frac{z_1}{z_2}\right), t\left(1+\frac{z_1}{z_2}\right)\right). \quad (24)$$

Note that we consider $t\left(1-\frac{z_1}{z_2}\right)$ as a rational polynomial in z_1, z_2 over the field $\mathrm{k}(t)$. Then we claim that the maximal $\mathbb{T}$-action on z_1, z_2 is given by $A = \begin{bmatrix}1 & 1\end{bmatrix}$.

The corresponding Hermite multiplier and inverse are

$$V = \left[\begin{array}{c|c}1 & 1\\ 0 & -1\end{array}\right] \qquad W = \left[\begin{array}{cc}1 & 1\\ \hline 0 & -1\end{array}\right]$$

which gives invariant and auxiliary variable:

$$y = \frac{z_1}{z_2}, \qquad x = z_1.$$

Then

$$\frac{dy}{dt} = yF\left(1, \frac{1}{y}\right)\mathrm{V}_{\mathfrak{n}} = y\begin{bmatrix}1-y & y+1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = -2y^2,$$
$$\frac{dx}{dt} = xF\left(1, \frac{1}{y}\right)\mathrm{V}_{\mathfrak{i}} = x\begin{bmatrix}1-y & y+1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = x(1-y).$$

Solving gives:

$$y(t) = \frac{1}{2t+c_0} \qquad x(t) = c_1 \exp\left(t - \frac{1}{2}\log(2t+c_0)\right)$$

for some constants c_0, c_1. This leads us to the final solutions:

$$z_1(t) = x(t) = c_1 \exp\left(t - \frac{1}{2}\log(2t+c_0)\right),$$
$$z_2(t) = \frac{x(t)}{y(t)} = c_1(2t+c_0)\exp\left(t - \frac{1}{2}\log(2t+c_0)\right).$$

4.3.2 General Case

In this subsection we show that the construction in section 4.3.1 can be extended to torus actions acting non-trivially on the independent variable, using the method of [26, Section 6.2].

Suppose we also have a torus $\mathbb{T}_{\bar{A}}$ acting k^{n+1} by

$$\bar{A} = \begin{bmatrix}A_0 & A\end{bmatrix} = \begin{bmatrix}A_0 & A_1 & \dots & A_N\end{bmatrix} \in M_{r\times(n+1)}(\mathbb{Z})$$

that is a symmetry of the system. We know from section 4.1 that this is equivalent to $F_i(t,z) \in \mathrm{k}(t,z)^{\mathbb{T}_{\bar{A}}}$ for each $i = 1, \dots, n$.

We introduce a new dependent variable $z_0(t)$ which will act like a scaled version of the independent variable t. Define

$$\bar{z} = (z_0, z_1, \dots, z_n), \quad \bar{F} = \begin{bmatrix}1 & F\end{bmatrix} \tag{25}$$

and consider the dynamical system

$$\frac{d\bar{z}}{dt} = \frac{\bar{z}}{t} * \bar{F}(\bar{z}) \tag{26}$$

where $\bar{F}(\bar{z}(t)) = \bar{F}(z_0(t), z(t))$. Note the first equation of equation (26) is $\frac{dz_0}{dt} = \frac{z_0}{t}$ which has solution $z_0(t) = ct$ for some constant c.

Lemma 4.12 ([26, Section 6.2]). *1. If $z(t) = (z_1(t), \dots, z_n(t))$ is a solution of equation* (13), *then $\bar{z}(t) = (t, z_1(t), \dots, z_n(t))$ is a solution of equation* (26).

2. Suppose $\bar{z}(t) = (\bar{z}_0(t), \bar{z}_1(t), \dots, \bar{z}_n(t))$ is a solution of equation (26) *and $\bar{z}_0(t) = ct$ for c nonzero. Then $z(t) = (\bar{z}_1(\frac{t}{c}), \dots, \bar{z}_n(\frac{t}{c}))$ is a solution of equation* (13).

Proof. 1. Immediate.

2. Let $z_i(t) = \bar{z}_i\left(\frac{t}{c}\right)$. Then for $i = 1, \ldots, n$:

$$\begin{aligned}\frac{dz_i(t)}{dt} &= \frac{d\bar{z}_i(\frac{t}{c})}{dt} \\ &= \frac{\bar{z}_i(\frac{t}{c})}{c\frac{t}{c}} F_i\left(\bar{z}_0\left(\frac{t}{c}\right), \bar{z}_1\left(\frac{t}{c}\right), \ldots, \bar{z}_n\left(\frac{t}{c}\right)\right) \\ &= \frac{z_i(t)}{t} F_i(t, z(t)).\end{aligned}$$

Hence $z(t)$ satisfies equation equation (13).

□

We obtain the general version of theorem 4.10 as a corollary.

Corollary 4.13 ([26, Theorem 6.5])**.** *Consider a map $F : \mathrm{k} \times \mathrm{k}^n \to \mathrm{k}^n$ that is invariant under the group action of $\mathbb{T}_{\bar{A}}$ given by $\bar{A} \in M_{r\times(n+1)}(\mathbb{Z})$. Let $V = \begin{bmatrix} \mathrm{V}_{\mathfrak{i}} & \mathrm{V}_{\mathfrak{n}} \end{bmatrix}$ be the Hermite multiplier of $\bar{A}$ with inverse $W = \begin{bmatrix} \mathrm{W}_{\mathfrak{u}} \\ \mathrm{W}_{\mathfrak{d}} \end{bmatrix}$. Define $\bar{F} = \begin{bmatrix} 1 & F \end{bmatrix}$ as above.*

1. *If $z(t) = (z_1(t), \ldots, z_n(t))$ is a solution of $\frac{dz}{dt} = \frac{z * F(t,z)}{t}$ where none of the components vanish and $\bar{z}(t) = (t, z_1(t), \ldots, z_n(t))$ then*

$$\begin{bmatrix} x(t) & y(t) \end{bmatrix} = \begin{bmatrix} \bar{z}^{\mathrm{V}_{\mathfrak{i}}} & \bar{z}^{\mathrm{V}_{\mathfrak{n}}} \end{bmatrix} \tag{27}$$

is a solution to the dynamical system:

$$\frac{dy}{dt} = \frac{y}{t} * \left(\bar{F}\left(y^{\mathrm{W}_{\mathfrak{d}}}\right) \cdot \mathrm{V}_{\mathfrak{n}}\right) \tag{28}$$

$$\frac{dx}{dt} = \frac{x}{t} * \left(\bar{F}\left(y^{\mathrm{W}_{\mathfrak{d}}}\right) \cdot \mathrm{V}_{\mathfrak{i}}\right), \tag{29}$$

where the reduced system is given by equation (28)*.*

2. *Suppose $y(t)$, $x(t)$ are solutions to equations* (28) *and* (29) *respectively and none of their components vanish. Let*

$$\begin{bmatrix} \bar{z}_0(t) & \bar{z}_1(t) & \ldots & \bar{z}_n(t) \end{bmatrix} = \begin{bmatrix} x(t) & y(t) \end{bmatrix}^W \tag{30}$$

where $\bar{z}_0(t) = ct$ for some nonzero constant c. Then

$$z(t) = \left(\bar{z}_1\left(\frac{t}{c}\right), \ldots, \bar{z}_n\left(\frac{t}{c}\right)\right) \tag{31}$$

is a solution to

$$\frac{dz}{dt} = \frac{z * F(z)}{t}.$$

4.4 Parameter Reduction by Scaling Symmetries

The previous section gives a very general framework for dealing with symmetries of specific systems of differential equations. In practice, mathematical models have many parameters which are constant, but unknown. In the majority of these cases there is a simpler algorithm for reduction than the method of section 4.3.2, which will only reduce the number of parameters.

Suppose we have a dynamical model of some state variables $z_1, \ldots, z_q$ that depend on the time t. Suppose their dynamics involve some constants $c_1, \ldots, c_p$. The parametrised dynamical system can be written:

$$\frac{dz}{dt} = G(t, z, c) \tag{32}$$

which can be trivially extended to the form of equation (12) by extending the system with the equations $\frac{dc}{dt} = 0$ [26, Section 7]. Then the matrix $A \in M_{r \times n}(\mathbb{Z})$, with $n = 1 + q + p$, defines a scaling symmetry if and only if the map $F(t, z, c) = t\, z^{-1} \star G(t, z, c)$ is an invariant of the action of $\mathbb{T}_A$.

The rest of this subsection assumes that the normal Hermite multiplier of A has the form

$$V = \left[\begin{array}{c|cc} 0 & I_{q+1} & 0 \\ \mathrm{V}_{\mathfrak{i}} & V_{\bar{v}} & V_c \end{array}\right] \quad \text{with inverse} \quad W = \left[\begin{array}{cc} \multicolumn{2}{c}{\mathrm{W}_{\mathfrak{u}}} \\ \hline I_{q+1} & 0 \\ \multicolumn{2}{c}{\mathrm{W}_{\mathfrak{d}}} \end{array}\right]. \tag{33}$$

This assumption may seem opaque but it has a natural interpretation. The I_{q+1} in the top row of V is exactly saying that our new non-constant invariants $y_0, \ldots, y_q$ are given by $y_0 = c^{u_0} t$, $y_i = c^{u_i} v_i$ for $i = 1, \ldots, n$, $u_j \in \mathbb{Z}^p$. The zeros on the top row ensure that our auxiliary variables and remaining invariants are only functions of the parameters. In particular, our original dependent and time variables each appear in exactly one invariant and we reduce only the number of parameters.

Let $V_{\bar{v}} = \begin{bmatrix} V_t & V_v \end{bmatrix}$ where V_t is the first column of $V_{\bar{v}}$. We can find the invariants explicitly

$$y = \begin{bmatrix} t & z & c \end{bmatrix}^{\mathrm{V}_{\mathfrak{n}}} = \begin{bmatrix} c^{V_t} t & c^{V_v} \star z & c^{V_c} \end{bmatrix},$$

giving us independent, dependent and constant invariants:

$$\mathfrak{t} = (c_1, c_2, \ldots, c_p)^{V_t} t, \tag{34}$$

$$\mathfrak{z} = (\mathfrak{z}_1, \mathfrak{z}_2, \ldots, \mathfrak{z}_q) = (c_1, \ldots, c_p)^{V_v} \star z, \tag{35}$$

$$\mathfrak{c} = (\mathfrak{c}_1, \mathfrak{c}_2, \ldots, \mathfrak{c}_{p-r}) = (c_1, \ldots, c_p)^{V_c}. \tag{36}$$

The auxiliary variables, which are constant under our assumptions, are given by:

$$x = c^{\mathrm{V}_{\mathfrak{i}}} \tag{37}$$

We also make the novel observation that if we only assume that the single invariant involving t is $c^a t$ for some scaling vector $a \in \mathbb{Z}^p$, then we can perform the t substitution outlines in this section, followed by the dependent variable reduction of section 4.3.1. This will avoid the introduction of a new variable by the general reduction of section 4.3.2.

Theorem 4.14 ([26, Section 7])**.** *Let W be as above and*

$$\mathrm{W}_{\mathfrak{d}} = \begin{bmatrix} W_t & W_v & W_c \end{bmatrix} \tag{38}$$

where $W_t \in M_{(p-r)\times 1}(\mathbb{Z})$, $W_v \in M_{(p-r)\times q}(\mathbb{Z})$, $W_c \in M_{(p-r)\times p}(\mathbb{Z})$.

Then the reduced system is given by making the substitution

$$t \mapsto \mathfrak{c}^{W_t}\mathfrak{t}, \qquad z \mapsto \mathfrak{c}^{W_v} * \mathfrak{z}, \qquad c \mapsto \mathfrak{c}^{W_c} \tag{39}$$

into equation (13).

Proof. Recall that by corollary 4.13 the reduced system is given by:

$$\frac{dy}{dt} = \frac{y}{t} * \left(\bar{F}\left(y^{\mathrm{W}_{\mathfrak{d}}}\right) \mathrm{V}_{\mathfrak{n}}\right).$$

Then

$$\begin{aligned} y^{\mathrm{W}_{\mathfrak{d}}} &= \begin{bmatrix} \mathfrak{t} & \mathfrak{z} & \mathfrak{c} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & I_q & 0 \\ W_t & W_v & W_c \end{bmatrix} \\ &= \begin{bmatrix} \mathfrak{c}^{W_t}\mathfrak{t} & \mathfrak{c}^{W_v} * \mathfrak{z} & \mathfrak{c}^{W_c} \end{bmatrix}. \end{aligned}$$

Furthermore, every component after the first $q+1$ components of $\bar{F}$ are equal to 0, so

$$\bar{F}\left(y^{\mathrm{W}_{\mathfrak{d}}}\right). \mathrm{V}_{\mathfrak{n}} = \begin{pmatrix} 1 & F_1\left(y^{\mathrm{W}_{\mathfrak{d}}}\right) & \dots & F_q\left(y^{\mathrm{W}_{\mathfrak{d}}}\right) & \underbrace{0 \dots 0}_{p-r \text{ times}} \end{pmatrix}$$

is equal to $\bar{F}\left(y^{\mathrm{W}_{\mathfrak{d}}}\right)$ with the last r zeros removed. The last $p-r$ variables of y are our new invariant constants $\mathfrak{c}$, which we require to have derivative 0 as above, so we can forget these equations. Hence we have the reduced system given by:

$$\frac{d\bar{y}}{dt} = \frac{\bar{y}}{t} * \begin{pmatrix} 1 & F_1\left(\mathfrak{c}^{W_t}\mathfrak{t}, \mathfrak{c}^{W_v} * \mathfrak{z}, \mathfrak{c}^{W_c}\right) & \dots & F_q\left(\mathfrak{c}^{W_t}\mathfrak{t}, \mathfrak{c}^{W_v} * \mathfrak{z}, \mathfrak{c}^{W_c}\right) \end{pmatrix}$$

where $\bar{y} = \begin{pmatrix} \mathfrak{t} & \mathfrak{z} \end{pmatrix}$.

The first equation is

$$\frac{d\mathfrak{t}}{dt} = \frac{\mathfrak{t}}{t}.$$

Dividing the other equations by $\frac{d\mathfrak{t}}{dt}$ gives:

$$\frac{d\mathfrak{z}}{d\mathfrak{t}} = \frac{\mathfrak{z}}{\mathfrak{t}} * F\left(\mathfrak{c}^{W_t}\mathfrak{t}, \mathfrak{c}^{W_v} * \mathfrak{z}, \mathfrak{c}^{W_c}\right). \tag{40}$$

Recalling that

$$\frac{d\left(\mathfrak{c}^{V_v} * \mathfrak{z}\right)}{d\left(\mathfrak{c}^{V_t}\mathfrak{t}\right)} = \frac{\mathfrak{c}^{V_v}}{\mathfrak{c}^{V_t}} * \frac{d\mathfrak{z}}{d\mathfrak{t}}, \quad \text{as } \mathfrak{c} \text{ is constant, and}$$

$$\frac{\mathfrak{c}^{V_v} * \mathfrak{z}}{\mathfrak{c}^{V_t}\mathfrak{t}} = \frac{\mathfrak{c}^{V_v}}{\mathfrak{c}^{V_t}} * \frac{\mathfrak{z}}{\mathfrak{t}},$$

suffices to show that the direct substitution gives the reduced system. □

Example 4.15. Recall example 1.1:

$$\frac{ds}{dt} = -k_1e_0s + (k_1s + k_{-1})c = \frac{s}{t}\left(\frac{-k_1e_0st + k_1sct + k_{-1}ct}{s}\right) = \frac{s}{t}F_1$$
$$\frac{dc}{dt} = k_1e_0s - (k_1s + k_{-1} + k_2)c = \frac{c}{t}\left(\frac{k_1e_0st - k_1sct + k_{-1}ct + k_2ct}{c}\right) = \frac{c}{t}F_2$$

We order our variables by: $(t, s, c, e_0, k_1, k_2, k_{-1})$. Then by section 3.5:

$$K_{F_1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad K = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The row Hermite normal form decomposition is:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 & 0 & 0 \end{bmatrix} \cdot K = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

giving us the maximal scaling matrix, given by the bottom two rows of the multiplier:

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

The next step is to perform the column Hermite normal form decomposition of A. The normal Hermite multiplier of A and its inverse is given by:

$$V = \left[\begin{array}{cc|c|cccc} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & 1 & -1 & -1 & -1 & -1 \end{array}\right] \quad \text{and} \quad W = \left[\begin{array}{ccccccc} -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

We can read off the invariants from the last 5 columns of V:

$$\mathfrak{t} = k_{-1}t \tag{41}$$
$$\mathfrak{s} = \frac{k_1}{k_{-1}}s \tag{42}$$
$$\mathfrak{c} = \frac{k_1}{k_{-1}}c \tag{43}$$
$$\mathfrak{k} = (\mathfrak{k}_1, \mathfrak{k}_2) = \left(\frac{e_0k_1}{k_{-1}}, \frac{k_2}{k_{-1}}\right) \tag{44}$$

We can also read off the substitutions from the bottom 5 rows of W that give the reduced system:

$$(t, s, c, e_0, k_1, k_2, k_{-1}) \mapsto (\mathfrak{t}, \mathfrak{s}, \mathfrak{c}, \mathfrak{k}_1, 1, \mathfrak{k}_2, 1). \tag{45}$$

Finally, the reduced system is given by:

$$\frac{d\mathfrak{s}}{d\mathfrak{t}} = -\mathfrak{k}_1\mathfrak{s} + (\mathfrak{s}+1)\mathfrak{c} \tag{46}$$

$$\frac{d\mathfrak{c}}{d\mathfrak{t}} = \mathfrak{k}_1\mathfrak{s} - (\mathfrak{s}+1+\mathfrak{k}_2)\mathfrak{c} \tag{47}$$

For more examples, see [26, Section 7].

4.4.1 Parameter Modification

We outline a novel way to generalise this method for cases when the normal Hermite multiplier has the form:

$$V = \left[\begin{array}{c|cc} 0 & D & 0 \\ \mathrm{V}_{\mathfrak{i}} & V_{\bar{v}} & V_c \end{array}\right], \quad \text{where} \quad D = \begin{bmatrix} d_0 & 0 & \dots & 0 \\ 0 & d_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_q \end{bmatrix} \quad \text{for } d_i \in \mathbb{N}_{>0}. \tag{48}$$

Suppose $d_i > 1$ for some $1 \le i \le q$ (the same method applies to d_0, except replace z_i with t). Then we have an invariant $\mathfrak{z}_i = c_1^{a_1} \dots c_p^{a_p} z_i^{d_i}$ for some $a_i \in \mathbb{Z}$. Since the c_j are parameters, almost always real, we suppose we can pick a d_ith root.

$$\text{Define} \quad \tilde{c}_i = \begin{cases} \sqrt[d_i]{c_i} \text{ if } a_i \ne 0 \\ c_i \text{ if } a_i = 0, \end{cases} \quad \text{and substitute} \quad c_i \mapsto \begin{cases} \tilde{c}_i^{d_i} \text{ if } a_i \ne 0 \\ \tilde{c}_i \text{ if } a_i = 0, \end{cases} \tag{49}$$

into our original system. This system will clearly be equivalent to our original system. Under this substitution our old invariant will become $(\tilde{c}_1^{a_1} \dots \tilde{c}_p^{a_p} z_i)^{d_i}$, so that the invariant of our new system will be: $\tilde{c}_1^{a_1} \dots \tilde{c}_p^{a_p} z_i$. The effect of this process is setting the entry d_i to 1, so that after repeated application the normal Hermite multiplier has the form given in equation (33).

Example 4.16. Consider the equation

$$\frac{dz}{dt} = \frac{cz^3}{t},$$

where c is a constant, and the scaling matrix $A = \begin{bmatrix} 0 & 1 & -2 \end{bmatrix}$ whose normal Hermite multiplier is $V = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

We make the substitution $c \mapsto \tilde{c}^2$ and consider the system

$$\frac{dz}{dt} = \frac{\tilde{c}^2 z^3}{t}$$

which has a corresponding scaling matrix $A = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$ and normal Hermite multiplier $V = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

By theorem 4.14, our original system is equivalent to the system:

$$\frac{d\mathfrak{z}}{d\mathfrak{t}} = \frac{\mathfrak{z}^3}{\mathfrak{t}}.$$

4.5 Scaling Symmetries and Non-Dimensionalisation

Recall that there are many different ways to non-dimensionalise a system of differential equations. This is explained by the non-uniqueness of the Hermite multiplier, which in turn is explained by the fact that there are many different choices generators for the field of invariants.

While the methods already discussed remove some ambiguity - through choice of the normal Hermite multiplier - we still have to choose the ordering of our variables at the outset. Given an order, the normal Hermite multiplier will try and normalise them using later variables. This is exactly why we choose an order:

$$(\text{independent variable}, \text{dependent variables}, \text{parameters})$$

since the dynamics of $f(c)z_i$ are much simpler than the dynamics of $f(z)$ for f an arbitrary rational function. This can be seen in example 4.15, where we act on the variables by k_2 and k_{-1} - the last variables possible. Hence, if there are parameters in a system that are natural to divide or multiply by, it is advisable to put them at the end.

Example 4.17. Continuing with example 4.15, the invariants $\mathfrak{s} = \frac{k_1}{k_{-1}}c, \mathfrak{c} = \frac{k_1}{k_{-1}}s$ are not particularly natural. If we choose to re-order the parameters, placing e_0 at the end, the method produces invariants

$$\hat{\mathfrak{s}} = \frac{s}{e_0} \quad \text{and} \quad \hat{\mathfrak{c}} = \frac{c}{e_0}.$$

5 Scaling Symmetries of PDEs

In this section, we extend the methods of Hubert and Labahn [26] to PDEs in an entirely novel way. While there is literature on the analytic reduction of PDEs and work by Hubert on its connection to differential algebra [22, 23, 24], it describes the differential algebra generated by invariant solutions and *invariant derivations* - which is different and beyond the scope of this work.

We will restrict ourselves to the reduction of parameters as in section 4.4 but only use the theoretical framework of differential algebra. Initially, this will restrict us further to the case where we can only act on the dependent variables, since differential algebra has no natural way of acting on a derivation ∂, though it could be generalised using differential ring isomorphisms (allowing a change in the derivations), or the trick from [26, Section 6.2] and 4.2.2, which we sketch at the end of this section.

We set up some notation of the rest of this section. We denote the independent variables by $x_1, x_2, \ldots, x_m$. The corresponding derivations are $\Delta = \{\partial_{x_1}, \partial_{x_2}, \ldots, \partial_{x_m}\}$ and we will also denote them by $\partial_i := \partial_{x_i}$. Let $\mathrm{k} = \mathbb{Q}$, $\mathcal{F}$ be a differential field with derivations Δ, $\mathrm{k} \subset \mathcal{F}$ and $x_i \in \mathcal{F}$. Let $R = \mathcal{F}\{y_1, \ldots, y_n\}$. Suppose we are given a system of differential equations described by:

$$\Sigma = \{F_1, \ldots, F_k\} \subset R\,. \tag{50}$$

Let $I = [\Sigma]$ be the differential ideal corresponding to our system and $X = \mathcal{V}(I) \subset \mathbb{A}^n$ be the differential variety of solutions. Finally let $\mathbb{T} = \mathrm{k}^*$ act on $\mathbb{A}^n$ by $A = \begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix}$.

Our aim will be to find torus actions on X and show how these actions can be used to reduce the number of parameters.

5.1 Semi-Invariants

Definition 5.1 ([12, Remark 2.1.8])**.** Let a group G act on the affine n-space $\mathbb{A}^n$ over a field k. Then $f \in \mathrm{k}[x_1, \dots x_n]$ is a *semi-invariant* if there exists $\chi : G \to \mathrm{k}^*$ such that $\sigma \cdot f = \chi(\sigma) f$ for all $\sigma \in G$.

Lemma 5.2. *Let $f = \sum_{w \in \mathbb{N}^n} a_w z^w \in \mathrm{k}[z]$ be a semi-invariant of a $\mathbb{T}_A$-action. Then there exists $u \in \mathbb{Z}^n$ such that:*

$$f = \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v z^{u+v}$$

where only finitely many a_v are non-zero.

Proof. Pick w in the support of f, so $a_w \neq 0$. If $f = a_w z^w$ then we are done, else f and $a_w z^w$ are coprime. Then $\frac{f}{a_w z^w}$ is a rational invariant so

$$\frac{f}{a_w z^w} = \frac{p}{q} \quad \text{for} \quad p = \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v z^{u+v},\ q = \sum_{v \in \ker A \cap \mathbb{Z}^n} b_v z^{u+v} \text{ coprime} \tag{51}$$

by lemma 3.24. Now $f = \frac{a_w z^w}{q} p \in \mathrm{k}[z]$ so q must divide $a_w z^w$. Hence $\frac{a_w z^w}{q} \in \mathrm{k}[z] \Rightarrow \frac{a_w z^w}{q} = b z^{w'}$ for some $w' \in \mathbb{N}^n, b \in \mathrm{k}$. Absorbing this factor into p gives the required result. □

5.2 Finding Torus Actions

We find torus actions leaving a finite subset of polynomials semi-invariant by modifying the methods of [26, Section 5] and section 3.5.

Definition 5.3. A $\mathbb{T} = (\mathrm{k}^*)^r$-action on $\mathbb{A}^n$ induces a map on the differential coordinate ring. For $\lambda \in \mathbb{T}$, we have

$$(\lambda \cdot \theta y_i)(z) = \theta y_i \circ \lambda^{-1}(z) = \theta y_i \left(\lambda^{-1} \cdot z\right) = \theta(\lambda_1^{-a_{1i}} \dots \lambda_r^{-a_{ri}} z_i) = \lambda_1^{-a_{1i}} \dots \lambda_r^{-a_{ri}} \theta(z_i) = \theta\left(\lambda \cdot y_i\right)(z) \tag{52}$$

for $1 \leq i \leq n$ as $\partial_j \lambda_k = 0$. Note the group action commutes with derivations.

In this way, every monomial $m = \prod_{i=1}^k \theta_i y_i$ for $\theta_i y_i \in \Theta Y$ has a *weight vector* a_m such that

$$\lambda^{-1} \cdot m = \lambda^{a_m} m.$$

Then $F \in R$ is a semi-invariant of the $\mathbb{T}_A$-action if each of its monomials has the same weight vector, so that the scaling action can be factored out. In particular, there exists $a \in \mathbb{Z}^n$ such that $\lambda^{-1} \cdot F = \lambda^a F$.

Lemma 5.4. *If $\{F_i\}_{i=1}^k$ are semi-invariant with respect to the $\mathbb{T}$-action, then the action descends to a well-defined action on X.*

Proof. Let $z = (z_1, \dots, z_n) \in X$, $\lambda \in \mathbb{T}$. Then

$$F_i\left(\lambda \cdot z\right) = (\lambda^{-1} \cdot F_i)(z) = \chi(\lambda^{-1}) F_i(z) = 0$$

for each i. Therefore $\lambda \cdot z \in X$ as required. □

The reverse implication does not hold, even in the algebraic case.

Example 5.5. Consider the polynomials $F_1 = x+y-z, F_2 = z \in \mathrm{k}[x,y,z]$. There exists a torus action given by $A = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$ on the solution set $\mathcal{V}(\langle F_1, F_2 \rangle) = \{(x,-x,0) \in \mathbb{A}^3 \mid x \in \mathrm{k}\} \subset \mathbb{A}^3$. However, the $\mathbb{T}$-action does not factor:

$$(\lambda^{-1} \cdot F_1)(x,y,z) = (\lambda x) + (\lambda y) - z \neq \chi(\lambda) F_1(x,y,z)$$

for any $\chi : \mathbb{T} \to \mathrm{k}^*$, so F_1 is not a semi-invariant.

Now we show how to find $\mathbb{T}$-actions leaving Σ semi-invariant. While this will not find all of the possible $\mathbb{T}$-actions on X, it will find some of them.

Definition 5.6. Consider $\mathbb{T} = (\mathrm{k}^*)^n$ acting on $\mathbb{A}^n$ by $A = I_n$. Let

$$F = \sum_{i=1}^{k} r_i m_i \in R,$$

where m_i is a product of differentials, $r_i \in \mathcal{F}$.

Let $M = \{a_{m_i} \in \mathbb{N}^n \mid m_i \text{ a monomial appearing in } F\}$ be the set of weight vectors and pick $w \in M$ minimal with respect to the ordering given in definition 3.29. We define the *exponent matrix* of F to be the matrix $\bar{K}_F$ with columns $v - w$ for $v \in M \setminus \{w\}$.

Example 5.7.

1. $F = y_1 y_2 - \partial_1^2 y_1$ has monomial weights $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$ so $\bar{K}_F = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

2. $F = y_1 \, \partial_2 \, y_2 - \partial_1^2 \, y_1 y_3^2 + y_2 y_3$ has monomial weights $\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$ so $\bar{K}_F = \begin{bmatrix} 1 & 1 \\ 0 & -1 \\ -1 & 1 \end{bmatrix}$.

Lemma 5.8. *Let $A \in M_{r \times n}(\mathbb{Z})$, $F \in R$. Then F is a semi-invariant with respect to the $\mathbb{T}$-action by A if and only if:*

$$A\bar{K}_F = 0.$$

Proof. Suppose that F has order p (order of the highest derivative in F). Consider F as a polynomial in the ring $S = \mathcal{F}[\Theta(p)\, z]$, where the action of $\mathbb{T}$ on $F \in S$ is inherited from the action on the differential coordinate ring. Let m be the lowest order monomial appearing in F, with respect to the ordering in definition 3.29.

Since $K_{\frac{F}{m}} = \begin{bmatrix} \bar{K}_F & 0 \end{bmatrix}$, it is clear that $A\bar{K}_F = 0 \Leftrightarrow \frac{F}{m}$ is a rational invariant, by lemma 3.32. It is also clear that this happens if and only if F is a semi-invariant, since m is a monomial. □

Corollary 5.9. *Let $U \in M_{n \times n}(\mathbb{Z})$ be a row Hermite normal multiplier, such that the last r rows of $U\bar{K}_F$ are 0. Let A be the last r rows of U. Then:*

1. *A has full row rank.*

2. *r is maximal and A is a maximal scaling matrix with F a semi-invariant.*

Definition 5.10. Given a set of differential polynomials Σ, we can form $\bar{K}_{F_i}$ for each $F_i \in \Sigma$ and define the *exponent matrix* of the system:

$$\bar{K}_\Sigma = \begin{bmatrix} \bar{K}_{F_1} & \dots & \bar{K}_{F_k} \end{bmatrix}. \tag{53}$$

We can then find the *maximal scaling matrix A* such that the F_i are semi-invariant by performing a Hermite normal form decomposition, exactly as we did for the rational invariant case in corollary 3.36.

Now that we can find the maximal action leaving $\{F_i\}_{i=1}^k$ semi-invariant, we briefly discuss the choice of $\{F_i\}_{i=1}^k$. Since we are really interested in acting on $\mathcal{V}(I)$, we can choose any set of generators $\{G_i\}_{i=1}^{k_1}$ for I to be semi-invariant. The optimum choice remains an open problem and is not discussed in-depth here.

One approach is to compute a Groebner basis (forgetting the differential structure) or compute characteristic sets using an elimination ranking. This would certainly find the torus action in example 5.5.

Given a system of differential equations where the variables and parameters have units, we can immediately write down a scaling matrix. This comes from the arbitrary choice of unit for fundamental quantities, like time or distance. Each fundamental unit induces a scaling action on the quantities that are written in terms of it. This is actually the beginning of the proof for the Buckingham Pi Theorem [46, Section 5].

Example 5.11. Consider question 1 from problem sheet 2 of the Part B course [5], modelling the population of fish near a fishing port. The system is given by

$$\partial_t U = rU\left(1 - \frac{U}{K}\right) - EU + D\,\partial_x^2 U$$

for constants r, K, E, D, where $U(t,x)$ the number of fish.

Let $F = K\,\partial_t U - rU(K-U) + KEU - KD\,\partial_x^2 U$ and pick a variable order: (U, r, K, E, D). We calculate the maximal scaling matrix by performing a row Hermite normal form decomposition of the exponent matrix $\bar{K}_F$:

$$U.\bar{K}_F = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have found a one dimensional scaling action given by $A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix}$.

5.3 Construction of Reduced Systems

We show that we can perform similar substitutions to theorem 4.14 to reduce the number of parameters.

Suppose we have a system with dependent variables $z = (z_1, \ldots, z_q)$ and parameters $c = (c_1, \ldots, c_p)$. In particular, $\Sigma = \{F_1, \ldots, F_m\} \subset \mathcal{F}\{z_1, \ldots, z_q, c_1, \ldots, c_p\} / [\partial_i c_j]$, as $\partial_i c_j = 0$ for $1 \le i \le m$ and $1 \le j \le p$. Let $n = p + q$.

5.3.1 Scaling Actions on Dependent Variables

Suppose that we have a scaling matrix $A \in M_{r \times n}(\mathbb{Z})$, acting only on the dependent variables of the solution set X. Assume that the normal Hermite multiplier of A has the form

$$V = \left[\begin{array}{c|cc} 0 & I_q & 0 \\ \mathrm{V_i} & V_v & V_c \end{array}\right] \quad \text{with inverse} \quad W = \left[\begin{array}{cc} \multicolumn{2}{c}{\mathrm{W_u}} \\ \hline I_q & 0 \\ W_v & W_c \end{array}\right]. \tag{54}$$

Equation (54) has a natural interpretation, as in section 4.4: our invariants are given by $c^{u_i} z_i$ and c^{v_j} for some $u_i, v_j \in \mathbb{Z}^p$. Under this assumption, we will only reduce the number of parameters.

By considering the algebraic ring $\mathcal{F}[z_1, \ldots, z_q, c_1, \ldots, c_p]$ and applying the methods of section 4.4, we can read off candidate invariants:

$$\mathfrak{z} = (\mathfrak{z}_1, \mathfrak{z}_2, \ldots, \mathfrak{z}_q) = (c_1, \ldots, c_p)^{V_v} * z \tag{55}$$
$$\mathfrak{c} = (\mathfrak{c}_1, \mathfrak{c}_2, \ldots, \mathfrak{c}_{p-r}) = (c_1, \ldots, c_p)^{V_c}. \tag{56}$$

The auxiliary constants are

$$\mathfrak{k} = c^{\mathrm{V}_{\mathfrak{i}}}. \tag{57}$$

Lemma 5.12. *Let $F \in R$ be a semi-invariant differential polynomial of the $\mathbb{T}_A$-action. Then we can rewrite F in terms of $\mathfrak{z}, \mathfrak{c}$ and their derivatives up to a multiple of constants. In particular*

$$c^a F(z,c) = F\left(\left((z,c)^{\mathrm{V}_{\mathfrak{n}}}\right)^{\mathrm{W}_{\mathfrak{d}}}\right) = F\left((\mathfrak{z},\mathfrak{c})^{\mathrm{W}_{\mathfrak{d}}}\right)$$

for some scaling vector $a \in \mathbb{Z}^p$.

Proof. The idea of the proof is to pick a Hermite multiplier such that the basis of invariants has a nice form and use lemma 3.26. We will prove the lemma for a concrete, illustrative example and note that it generalises easily. Recall $F(U,r,K,E,D) = K\,\partial_t U - rU(K-U) + KEU - KD\,\partial_x^2 U$, A and V as in example 5.11.

The highest order of F is 2, coming from the differential $\partial_x^2 U$. We consider the algebraic ring:

$$\mathcal{F}[\Theta(2)\{U,r,K,E,D\}] = \mathcal{F}[U,r,K,E,D,U_t,U_x,U_{tt},U_{tx},U_{xx}]$$

whose field of fractions will contain every rational differential polynomial of order 2.

The action of $\mathbb{T}$ on the original variables extends naturally to the matrix:

$$A^{(2)} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

as $\partial_i \lambda \cdot U = \lambda \cdot \partial_i U$ for $\lambda \in \mathbb{T}$. We first reduce the 1s in the columns corresponding to the θU, by subtracting the column of U, which gives our first sequence of column actions described by multiplication on the right by $V_1 = \begin{bmatrix} I_5 & -M \\ 0 & I_5 \end{bmatrix}$, where M is the 5×5 matrix with top row 1 and other rows 0.

Then $A^{(2)}V_1 = \begin{bmatrix} A & 0 \end{bmatrix}$ and we can multiply on the right by the matrix $V_2 = \begin{bmatrix} V & 0 \\ 0 & I_5 \end{bmatrix}$ to put it in column Hermite normal form.

Thus AV_1V_2 is in Hermite normal form, where

$$V_1V_2 = \begin{bmatrix} I_5 & -M \\ 0 & I_5 \end{bmatrix}\begin{bmatrix} V & 0 \\ 0 & I_5 \end{bmatrix} = \begin{bmatrix} V & -M \\ 0 & I_5 \end{bmatrix}.$$

Since V has the form assumed in equation (33) we can add the columns corresponding to the U-invariant to the columns of the θU-invariants again, eliminating the -1 top row while scaling θU by the same constants as U. Hence our final Hermite multiplier for $A^{(2)}$ is

$$V_1V_2\begin{bmatrix} I_5 & M \\ 0 & I_5 \end{bmatrix} = \begin{bmatrix} V & VM - M \\ 0 & I_5 \end{bmatrix}.$$

Recalling the torus action defined in equation (52), by lemma 3.26 we have that

$$\left\{\theta g \mid \theta \in \Theta(2), g \in (U,r,K,E,D)^{\mathrm{V}_{\mathfrak{n}}}\right\}$$

form a generating set for the rational invariants of order 2, as desired. In particular, it suffices to consider the action on the indeterminates (of 0th order), find a generating set of the rational invariants and evaluate the differential polynomial F at these.

Furthermore, since any semi-invariant f gives us a rational invariant by dividing by a monomial appearing f, $\frac{F}{KEU}$ has the form of 3.26 and we have the desired re-write rule. This substitution applied to semi-invariants is correct up to some constants, by lemma 5.2; the scaling vector of the constants will be exactly u appearing in lemma 5.2.

Returning to our example,

$$\begin{aligned} F\left(\left((U,r,K,E,D)^{\mathrm{V}_{\mathfrak{n}}}\right)^{\mathrm{W}_{\mathfrak{d}}}\right) =& F\left(\frac{U}{K},r,1,E,D\right) \\ =& \partial_t\left(\frac{U}{K}\right) - r\frac{U}{K}(1-\frac{U}{K}) + E\frac{U}{K} - D\,\partial_x^2\left(\frac{U}{K}\right) \\ =& K^{-2}\left(K\,\partial_t U - rU(K-U) + KEU - KD\,\partial_x^2 U\right) \\ =& K^{-2}F. \end{aligned}$$

This proof generalises easily enough by considering higher order rings $\mathcal{F}\left[\Theta(N)z,c\right]$ and performing more column operations for further dependent variables. □

Theorem 5.13. *Consider the substitution:*

$$z \mapsto \mathfrak{c}^{W_v} * \mathfrak{z}, \tag{58}$$

$$c \mapsto \mathfrak{c}^{W_c} \tag{59}$$

and let $G_i(\mathfrak{z},\mathfrak{c}) = F_i(\mathfrak{c}^{W_v} * \mathfrak{z}, \mathfrak{c}^{W_c})$.

1. *If* $(z,c) \in \mathcal{V}(F_i)$ *is a solution to the original system, then* $(\mathfrak{z},\mathfrak{c}) = (z,c)^{\mathrm{V}_{\mathfrak{n}}} \in \mathcal{V}(G_i)$ *is a solution of the reduced system.*

2. *If* $(\mathfrak{z},\mathfrak{c}) \in \mathcal{V}(G_i)$ *is a solution to the reduced equations, then* $(z,c) = (\mathfrak{k},\mathfrak{z},\mathfrak{c})^W \in \mathcal{V}(F_i)$ *is a solution to the original system, where* $\mathfrak{k}$ *is a vector of arbitrary constants.*

Proof. 1. By lemma 5.12, there exists $a \in \mathbb{Z}^p$ such that:

$$G_i(\mathfrak{z},\mathfrak{c}) = F_i\left(\mathfrak{c}^{W_v} * \mathfrak{z}, \mathfrak{c}^{W_c}\right) = F_i\left((\mathfrak{z},\mathfrak{c})^{\mathrm{W}_{\mathfrak{d}}}\right) = F_i\left(\left((z,c)^{\mathrm{V}_{\mathfrak{n}}}\right)^{\mathrm{W}_{\mathfrak{d}}}\right) = c^a F_i(z,c) = 0.$$

2. Consider F as a polynomial in $\mathcal{F}[\Theta(N)z]$ for sufficiently large N and denote the extended action of A by A', with corresponding V' and W'. There exists $u \in \mathbb{Z}^n$ such that:

$$F = \sum_{v \in \ker A' \cap \mathbb{Z}^n} a_v z^{u+v},$$

where only finitely many a_v are non-zero, by lemma 5.2. Recalling that $\ker A' \subset \ker \mathrm{W}_{\mathfrak{u}}'$ we see that

$$\begin{aligned} F\left(c^{\mathrm{W}_{\mathfrak{u}}} * z\right) &= \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v \left(c^{\mathrm{W}_{\mathfrak{u}}} z\right)^{u+v} \\ &= \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v c^{\mathrm{W}_{\mathfrak{u}}(u+v)} z^{u+v} \\ &= \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v c^{\mathrm{W}_{\mathfrak{u}}\, u} z^{u+v} \\ &= c^{\mathrm{W}_{\mathfrak{u}}\, u} \sum_{v \in \ker A \cap \mathbb{Z}^n} a_v z^{u+v} \\ &= c^{\mathrm{W}_{\mathfrak{u}}\, u} F(z). \end{aligned}$$

Therefore:

$$F_i(z,c) = F_i\left((\mathfrak{k},\mathfrak{z},\mathfrak{c})^W\right) = F_i\left(\mathfrak{k}^{W_u} * (\mathfrak{z},\mathfrak{c})^{W_\partial}\right) = \mathfrak{k}^{W_u u} F_i\left((\mathfrak{z},\mathfrak{c})^{W_\partial}\right) = \mathfrak{k}^{W_u u} G_i(\mathfrak{z},\mathfrak{c}) = 0.$$

□

5.3.2 General Case

If we wish to act on independent variables as well as dependent variables, we can introduce dummy variables that behave like scaled versions of our independent parameters, as in section 4.2.2. In particular, if we have independent variables x_i and derivations ∂_i, then we introduce new independent variables $\tilde{x}_i$, such that $\tilde{x}_i = \lambda x_i$ for some constant $\lambda \in \mathcal{F}$. We do this by adding $x_i \partial_i \tilde{x}_i - \tilde{x}_i$ and $\partial_j \tilde{x}_i$ to Σ for $i \neq j$.

We then substitute

$$x_i \mapsto \tilde{x}_i, \quad \partial_i z_j \mapsto \frac{\partial_i z_j}{\partial_i \tilde{x}_i}$$

into our original F_i and clear the fractions introduced by the $\partial_i \tilde{x}_i$ to create new differential polynomials $\tilde{F}_i$. Let this new system of equations be denoted

$$\tilde{\Sigma} = \left\{\tilde{F}_i \mid 1 \le i \le k\right\} \cup \{x_i \partial_i \tilde{x}_i - \tilde{x}_i \mid 1 \le i \le m\} \cup \{\partial_j \tilde{x}_i \mid i \neq j\}.$$

Theorem 5.14. *1. If $(z(x), c)$ is a solution to Σ then $(z(x), c, x)$ is a solution to $\bar{\Sigma}$.*

2. Suppose $(z(x), c, \tilde{x})$ is a solution to $\bar{\Sigma}$ where $\tilde{x} = \mu * x$ with μ a vector of non-zero constants. Then $\left(z\left(\frac{x}{\mu}\right), c\right)$ is a solution to Σ.

Proof. Very similar to that of lemma 4.12. □

Example 5.15. We return to example 5.11, except we also act on the independent variables. Let s, y be the dummy variables of t, x respectively. Then we substitute s, y into F:

$$F \mapsto \frac{K\,\partial_t U}{\partial_t s} - rU(K-U) + KEU - KD\frac{\partial_x^2 U}{(\partial_x y)^2}$$

Let

$$G = K(\partial_t U)(\partial_x y)^2 - rU(K-U)(\partial_t s)(\partial_x y)^2 + KEU(\partial_t s)(\partial_x y)^2 - KD(\partial_x^2 U)(\partial_t s). \quad (60)$$

We can then find $\mathbb{T}$ actions for which G is semi-invariant, where we choose the variable order (s, y, U, r, K, E, D). We find $A = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$ with normal Hermite multiplier

$$V = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

which does not satisfy equation (54). This is because the only invariant of the second transformation (corresponding to the second row) involving y and D is $\frac{Ey^2}{D}$. We employ the method of section 4.4.1, square E and D and consider the equivalent equation

$$G = K(\partial_t U)(\partial_x y)^2 - rU(K-U)(\partial_t s)(\partial_x y)^2 + KE^2U(\partial_t s)(\partial_x y)^2 - KD^2(\partial_x^2 U)(\partial_t s).$$

Now we have an action defined by $A = \begin{bmatrix} 2 & 0 & 0 & -2 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$ which has normal Hermite multiplier and inverse:

$$V = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & -1 & 0 & 2 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} 2 & 0 & 0 & -2 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{61}$$

From this we can read off the invariants:

$$\mathfrak{x}_1 = \frac{s}{E^2}, \qquad \mathfrak{x}_2 = \frac{Ey}{D}, \qquad \mathfrak{z} = \frac{U}{K}, \qquad \mathfrak{c} = \frac{r}{D^2} \tag{62}$$

and rewrite rules:

$$(s, y, U, r, K, E, D) \mapsto (\mathfrak{x}_1, \mathfrak{x}_2, \mathfrak{z}, \mathfrak{c}, 1, 1, 1). \tag{63}$$

Our reduced system is given by the PDE:

$$(\partial_t \mathfrak{z})(\partial_x \mathfrak{x}_2)^2 - \mathfrak{c}\,\mathfrak{z}(1-\mathfrak{z})(\partial_t \mathfrak{x}_1)(\partial_x \mathfrak{x}_2)^2 + \mathfrak{z}(\partial_t \mathfrak{x}_1)(\partial_x \mathfrak{x}_2)^2 - (\partial_x^2 \mathfrak{z})(\partial_t \mathfrak{x}_1) = 0. \tag{64}$$

Dividing through by $(\partial_t \mathfrak{x}_1)(\partial_x \mathfrak{x}_2)^2$ we see that

$$\frac{\partial_t \mathfrak{z}}{\partial_t \mathfrak{x}_1} - \mathfrak{c}\,\mathfrak{z}(1-\mathfrak{z}) + \mathfrak{z} - \frac{\partial_x^2 \mathfrak{z}}{(\partial_x \mathfrak{x}_2)^2} = 0, \tag{65}$$

hence the original system is equivalent to the reduced system

$$\frac{d\mathfrak{z}}{d\mathfrak{x}_1} - \mathfrak{c}\,\mathfrak{z}(1-\mathfrak{z}) + \mathfrak{z} - \frac{d^2\mathfrak{z}}{d\mathfrak{x}_2^2} = 0, \tag{66}$$

which has three parameters fewer than the original.

References

[1] Christian Aistleitner and Univ-Prof DI Franz Winkler. *Relations between Gröbner bases, differential Gröbner bases, and differential characteristic sets*. PhD thesis, Johannes Kepler Universitat, 2010.

[2] Ian M Anderson, Mark E Fels, I M Anderson, and M E Fels. Exterior Differential Systems with Symmetry. *Acta Applicandae Mathematicae*, 87:3–31, 2005.

[3] F. Boulier, D. Lazard, F. Ollivier, and M. Petitot. Representation for the radical of a finitely generated differential ideal. In *Proceedings of the 1995 international symposium on Symbolic and algebraic computation - ISSAC '95*, pages 158–166, New York, New York, USA, 1995. ACM Press.

[4] Helen Byrne. Oxford Mathematics Part B Further Mathematical Biology: Problem Sheet 1, 2016.

[5] Helen Byrne. Oxford Mathematics Part B Further Mathematical Biology: Problem Sheet 2, 2016.

[6] John Carminati and Khai Vu. Symbolic Computation and Differential Equations: Lie Symmetries. *Journal of Symbolic Computation*, 29(1):95–116, jan 2000.

[7] Phyllis J Cassidy. Introduction to Differential Algebraic Geometry and Differential Algebraic Groups. In *The Kolchin Seminar in Differential Algebra*, 2005.

[8] Thomas Close. Diophantine Python Package, URL: https://github.com/tclose/Diophantine, 2017.

[9] Henri. Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag, 1993.

[10] David A. Cox, John B. Little, and Donal. O'Shea. *Ideals, Varieties and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2007.

[11] Teresa Crespo, Zbigniew Hajto, and Juan J Morales-Ruiz. Introduction to Differential Galois Theory, URL: http://www4.ncsu.edu/~singer/papers/Crespo_Hajto.pdf, 2007.

[12] Harm Derksen and Gregor Kemper. *Computational Invariant Theory*. Springer, 2002.

[13] Mark Fels and Peter J Olver. Moving Coframes: II. Regularization and Theoretical Foundations. *Acta Applicandae Mathematicae*, 55:127–208, 1999.

[14] Robert Gilmore. *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers and Chemists*. Cambridge University Press, 2008.

[15] O. D. Golubitsky, M. V. Kondratieva, and A. I. Ovchinnikov. On the generalized Ritt problem as a computational problem. *Journal of Mathematical Sciences*, 163(5):515–522, dec 2009.

[16] Li Guo, Phyllis J Cassidy, William F Keigher, and William Y. Sit. Foreword. In *Differential Algebra and Related Topics*, pages i–xiii. World Scientific, may 2002.

[17] Richard Gustavson, Alexey Ovchinnikov, and Gleb Pogudin. Bounds for Orders of Derivatives in Differential Elimination Algorithms. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation - ISSAC '16*, pages 239–246, New York, New York, USA, 2016. ACM Press.

[18] Richard Hall. Nondimensionalization for Dummies, URL: http://wp.auburn.edu/radich/wp-content/uploads/2014/08/Nondimensionalization-Resource-2.pdf, 2012.

[19] Charlotte Hardouin. Galoisian approach to differential transcendence, URL: http://www.math.univ-toulouse.fr/~hardouin/notescolombiacharlotte.pdf, 2015.

[20] George Havas, Bohdan S. Majewski, and Keith R. Matthews. Extended GCD and Hermite Normal Form Algorithms via Lattice Basis Reduction. *Experimental Mathematics*, 7(2):125–136, jan 1998.

[21] Evelyne Hubert. Factorization-free Decomposition Algorithms in Differential Algebra. *J. Symbolic Computation*, 29:641–662, 2000.

[22] Evelyne Hubert. Differential algebra for derivations with nontrivial commutation rules. *Journal of Pure and Applied Algebra*, 200(1):163–190, 2005.

[23] Evelyne Hubert. Differential invariants of a Lie group action: Syzygies on a generating set. *Journal of Symbolic Computation*, 44(4):382–416, apr 2009.

[24] Evelyne Hubert. *Algebraic and Differential Invariants*, pages 165–187. London Mathematical Society Lecture Note Series. Cambridge University Press, 2012.

[25] Evelyne Hubert. Rational Invariants of a Group Action. In *Journées Nationales de Calcul Formel*, volume 3, page 10. CEDRAM - Center for Diffusion of Academic Mathematical Journals, 2013.

[26] Evelyne Hubert and George Labahn. Scaling Invariants and Symmetry Reduction of Dynamical Systems. *Foundations of Computational Mathematics*, 13(4):479–516, aug 2013.

[27] Peter E. Hydon. Introduction to symmetry methods in the solution of differential equations that occur in chemistry and chemical biology. *International Journal of Quantum Chemistry*, 106(1):266–277, 2006.

[28] Irving Kaplansky. *An Introduction to Differential Algebra*. Hermann, Paris, 1957.

[29] Ellis Robert Kolchin. *Differential algebra and algebraic groups*. Acad. Press, 1973.

[30] Jerald J. Kovacic. An algorithm for solving second order linear homogeneous differential equations. *Journal of Symbolic Computation*, 2(1):3–43, mar 1986.

[31] Jerald J. Kovacic. Differential Schemes. In *Differential Algebra and Related Topics*, pages 71–94. World Scientific, may 2002.

[32] Hanze Liu, Jibin Li, and Quanxin Zhang. Lie symmetry analysis and exact explicit solutions for general Burgers' equation. *Journal of Computational and Applied Mathematics*, 228(1):1–9, jun 2009.

[33] Elizabeth L Mansfield. Algorithms for Symmetric Differential Systems. *Found. Comput. Math*, 1:335–383, 2001.

[34] MapleSoft. DifferentialAlgebra - Maple Programming Help Webpage.

[35] Nicolette Meshkat, Marisa Eisenberg, and Joseph J. DiStefano. An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases. *Mathematical Biosciences*, 222(2):61–72, dec 2009.

[36] Francesco Oliveri and Francesco. Lie Symmetries of Differential Equations: Classical Results and Recent Contributions. *Symmetry*, 2(2):658–706, apr 2010.

[37] Alexey Ovchinnikov. Differential Algebra Lecture Notes, URL: http://qcpages.qc.cuny.edu/˜aovchinnikov/MATH86600/notes.pdf, 2014.

[38] Joseph Ritt. The Mathematics Genealogy Project, URL: https://www.genealogy.math.ndsu.nodak.edu/id.php?id=37354.

[39] Joseph Fels Ritt. Differential Algebra. *The Mathematical Gazette*, 33:137, 1950.

[40] Alexander Ritter. C3.4 Algebraic Geometry Lecture Notes, 2016.

[41] Tarcísio M. Rocha Filho and Annibal Figueiredo. [SADE] a Maple package for the symmetry analysis of differential equations. *Computer Physics Communications*, 182(2):467–476, feb 2011.

[42] John E Sasser. History of Ordinary Differential Equations: The First Hundred Years, URL: http://www2.fiu.edu/˜yuasun/ODE_History.pdf, 1992.

[43] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley, 1998.

[44] William Y. Sit. The Ritt-Kolchin Theory For Differential Polynomials. In *Differential Algebra and Related Topics*, pages 1–70. World Scientific, may 2002.

[45] William Y. Sit. Introduction to Computational Differential Algebra, I. In *Graduate Center Series For Kolchin Seminar in Differential Algebra*, 2005.

[46] Frank M White. *Fluid Mechanics*. McGraw-Hill Education, 2008.

[47] Wikipedia. Liouville's theorem (differential algebra) — Wikipedia, 2016.

[48] Wen-tsun Wu. Mechanical Derivation of Newton's Gravitational Laws from Kepler's Laws. In *Selected Works of Wen-Tsun Wu*, pages 295–301. WORLD SCIENTIFIC, may 2008.