Invariant Theory, Tensors and Computational Complexity

by

Visu Makam

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in the University of Michigan
2018

Doctoral Committee:

Professor Harm Derksen, Co-Chair
Associate Professor Andrew Snowden, Co-Chair
Professor Sergey Fomin
Associate Professor Kayvan Najarian
Professor Martin Strauss

visu@umich.edu
ORCID iD: 0000-0001-8470-7860

2018

Dedicated to the memory of my father.

ACKNOWLEDGEMENTS

I must thank first and foremost my mother as none of this would have happened without her love and support. Next, I must thank my two sisters who will always be my favourite source of happiness.

I will be forever grateful to my advisor, Harm Derksen, for showing me a truly unique brand of mathematics. I would also like to thank K. V. Subrahmanyam, Jerzy Weyman, Joseph Landsberg and Avi Wigderson for their support and encouragement. I look forward to more mathematical interactions with them in the future.

There are several faculty that I would like to thank at the university of Michigan, in particular Andrew Snowden, Thomas Lam, Sergey Fomin, Bill Fulton, John Stembridge, Karen Smith and Paul Kessenich.

PREFACE

There is an exciting new field of interdisciplinary research at the interface of invariant theory, tensors and computational complexity that is being established. This area represents problems that are of interest to mathematicians, physicists and computer scientists alike. It is only natural then that there are several perspectives, which allow for progress on many fronts, often simultaneously, with a variety of ideas and techniques. In this dissertation, we will focus on the problem of giving upper bounds for the degree of generators for invariant rings, especially for the cases of matrix invariants and matrix semi-invariants. We will provide polynomial bounds for these cases. The problem of degree bounds for matrix semi-invariants is very closely related to the problem of rational identity testing in computational complexity. The polynomial bounds are instrumental in giving a polynomial time algorithm for rational identity testing. Tensors are the link between the two, and serve as an intermediary language to translate problems between invariant theory and computational complexity.

TABLE OF CONTENTS

Dedication . . . . . **ii**

Acknowledgments . . . . . **iii**

Preface . . . . . **iv**

List of Figures . . . . . **viii**

Abstract . . . . . **ix**

Chapter

1 Introduction **1**

2 Invariant theory **4**

2.1 Introduction 4
2.2 Method of Popov and Derksen for degree bounds 5
2.3 Weyl's polarization theorem 7
2.4 Separating invariants 9
2.5 Matrix invariants and matrix semi-invariants: Main Results 11
2.5.1 Matrix invariants 11
2.5.2 Matrix semi-invariants 13

3 Quivers **16**

3.1 Quiver representations 16
3.2 Invariant theory of quivers 19
3.2.1 Invariants 19
3.2.2 Semi-invariants 20

4 Linear matrices and non-commutative rank **22**

4.1 Introduction 22
4.1.1 Linear subspaces of matrices 23
4.2 Regularity Lemma 25
4.2.1 The ring of generic matrices 26
4.2.2 Universal division algebras 27
4.3 Failure of the weakly increasing property in blow-ups 29
4.3.1 Rational identities 31
4.3.2 Non-commutative arithmetic circuits with division 31

4.4 Combinatorics of ranks of blow-ups . . . 32
4.5 Ratio of non-commutative rank to commutative rank . . . 35
4.5.1 Preliminaries from Linear Algebra . . . 37
4.5.2 Effects of scaling basis vectors on the matrices of L_i's . . . 39
4.5.3 Examples . . . 41
4.5.4 The general case . . . 42
4.5.5 More examples . . . 45

5 Degree bounds for matrix invariants and matrix semi-invariants . . . 47

5.1 Degree bounds for matrix semi-invariants . . . 47
5.1.1 Null cone . . . 47
5.1.2 Degree bounds in characteristic 0 . . . 48
5.1.3 Adaptation to positive characteristic: Good filtrations . . . 49
5.2 Matrix invariants . . . 50
5.3 Semi-invariants of quivers . . . 52
5.3.1 Stability conditions and the null cone . . . 53
5.3.2 Bounds for the null cone . . . 54
5.3.3 Bounds for generating semi-invariants . . . 57
5.3.4 Removing dependence on $\dim \mathrm{SI}(Q, \alpha)$ . . . 58
5.3.5 Exponential lower bound . . . 59
5.4 Quadratic lower bounds for matrix semi-invariants . . . 61
5.4.1 Combinatorial description of $R(n, m)$ . . . 61
5.4.2 Lower bounds for $\beta(R(n, m))$ . . . 63
5.5 Matrix semi-invariants for 3×3 matrices . . . 65
5.5.1 Krull dimension of $R(n, m)$ . . . 65
5.5.2 Invariants defining the null cone . . . 66
5.5.3 A hsop for $R(3, m), m \geq 3$ . . . 67
5.5.4 Upper bounds for $\beta(R(3, m))$ . . . 67
5.6 Hilbert series computations . . . 68

6 Orbit closure problem and separating invariants . . . 71

6.1 Introduction . . . 71
6.1.1 Known algorithms for matrix invariants . . . 71
6.2 Time complexity equivalence of orbit closure problems for matrix invariants and matrix semi-invariants . . . 72
6.2.1 Reduction from matrix invariants to matrix semi-invariants . . . 72
6.2.2 Reduction from matrix semi-invariants to matrix invariants . . . 73
6.3 A polynomial time algorithm for finding a basis for a subalgebra of $\mathrm{Mat}_{n,n}$ 77
6.4 Orbit closure problem for matrix invariants . . . 79
6.4.1 The positive characteristic case . . . 80
6.5 Bounds for separating invariants . . . 84
6.5.1 Matrix invariants . . . 84
6.5.2 Matrix semi-invariants . . . 85

7 Computational complexity . . . 89

7.1 Non-commutative circuits 89
7.2 Rational identity testing 90

8 Tensor rank 94

8.1 Introduction 94
8.2 Strassen's equations 95
8.2.1 Application to 3×3 determinant and permanent tensors 96
8.3 Flattenings 98
8.4 Border rank of tensors in $K^d \otimes K^d \otimes K^d$ 99
8.4.1 The case d is odd 99
8.4.2 The case d is even 102

BIBLIOGRAPHY 103

LIST OF FIGURES

3.1 m-loop quiver 16
3.2 Kronecker quiver 17

5.1 Quiver for exponential lower bounds 53
5.2 2-Kronecker quiver 59
5.3 Indecomposable representation 1 59
5.4 Indecomposable representation 2 59

7.1 Example of a circuit 89

ABSTRACT

The main problem addressed in this dissertation is the problem of giving strong upper bounds on the degree of generators for invariant rings. In the cases of matrix invariants and matrix semi-invariants, we give polynomial upper bounds. An exciting consequence of these bounds is a polynomial time algorithm for rational identity testing. We use an approach inspired by ideas from Popov and Derksen to reduce the problem to finding invariants that define the null cone. The theory of blow-ups of matrix spaces and non-commutative rank is crucial in finding invariants that define the null cone. We also give a polynomial time algorithm for deciding if the orbit closures of two points intersect for matrix invariants and semi-invariants. In addition, we give some applications for proving lower bounds on the border rank of tensors.

CHAPTER 1

Introduction

As early as in high school mathematics, we learn about algebraic identities. Here are some simple ones:

- $(a+b)(a+b) = a^2 + 2ab + b^2$.
- $(a+b+c)(a^2+b^2+c^2-ab-bc-ac) = a^3+b^3+c^3-3abc$.

Most of us would be able to verify these identities without much effort. Here are some more complicated ones:

- $(a+b+c)^7 - a^7 - b^7 - c^7 =$
  $7(a+b)(b+c)(a+c)((a^2+b^2+c^2+ab+bc+ac)^2 + abc(a+b+c))$
- $(a_1^2+a_2^2+a_3^2+a_4^2)(b_1^2+b_2^2+b_3^2+b_4^2) =$
  $(a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 +$
  $(a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 +$
  $(a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 +$
  $(a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2$

These identities would of course take much longer for us to verify. The first happens to be an identity used by Lamé in his proof of Fermat's last theorem for the case $n = 7$, and the second is Euler's four square identity. Nevertheless, this brings us to a fundamental question – How quickly can you verify whether a given identity is correct? This problem is often called polynomial identity testing (PIT) in theoretical computer science.

There is of course a naive algorithm to do this: expand out all multiplications, and then check whether the coefficient of each monomial occurring on both sides match. This is of course a very tedious process, and one might wonder if there is a quicker way to do this. One strategy is to assign in a choice of numbers for each variable, and then test whether the two sides evaluate to the same number. If we start with something that is not an algebraic

identity, then it is likely that for a random choice, both sides will not match. For example, if we are trying to verify whether $(x-1)^2 + y^2 = (x+y-1)^2$, then this will not be true unless we end up assigning in 1 for x or 0 for y. This is unlikely to happen when we choose these numbers at random, but nevertheless possible. Hence, if both sides agree for a particular assignment, we cannot be fully convinced that it is an identity. If we take numerous assignments, and both sides keep agreeing, then we become more and more convinced that it is an identity. In fact, to be convinced with a very high degree of confidence, we do not need too many assignments, and this analysis usually goes by Schwarz-Zippel Lemma. In the language of computer science, this is an efficient randomized algorithm. However, to be convinced beyond any doubt, the number of assignments we would need is likely too large. One of the major open questions in this subject is whether we can find a small number of assignments that will suffice to verify the identity beyond doubt.

One can formalize this more rigorously by constructing polynomials with arithmetic circuits, and asking whether there is an algorithm to decide whether a given circuit computes an identity or not. Ideally one would want an algorithm whose run time is polynomial in the size of the circuit. We will not make this rigorous formulation, but will be content to mention that a polynomial time algorithm is not known, and remains one of the big open questions in computational complexity.

If we live in a world where multiplication is not commutative (and everyone who deals with matrices lives in such a world), then we might be interested in "non-commutative" polynomial identities. This is a world in which $(a+b)^2 = a^2 + ab + ba + b^2$, as opposed to $a^2 + 2ab + b^2$. Let us increase the difficulty a bit more by allowing inverses as well! The problem of verifying whether two given non-commutative rational expressions are equal is called non-commutative rational identity testing (RIT). In the commutative world, any rational function can be written as a quotient of two polynomial functions. Rational functions in the non-commutative world can be far more complicated. For example the expression $(a + ab^{-1}a)^{-1}$ illuminates the possibility of nested inverses – an inverse inside an inverse. It is in fact not possible to simplify the expression into one of the form PQ^{-1} as in the commutative situation.

Consider the equation

$$(a + ab^{-1}a)^{-1} + (a+b)^{-1} = a^{-1}.$$

When we were considering PIT, there was a naive algorithm to verify any given identity – simplify both sides by expanding all multiplications and additions and check the coefficients of monomials on both sides. It is unclear how to simplify a non-commutative

rational expression, primarily because nested inverses cannot always be removed. This raises an alarming question – how then would one verify an identity like the one above? It so happens that the above identity is indeed correct, and is known as Hua's identity. It is not very difficult to prove Hua's identity by hand and I invite the reader to try it so as to witness first hand the challenges that a nested inverse presents.

At this juncture, it seems like this problem (RIT) is far more complicated than the simple one (PIT) we began with. If we cannot find an efficient algorithm for PIT, then is there any hope of finding one for RIT? How about even an efficient randomized algorithm? Recall, for PIT, a randomized algorithm is to evaluate the identity by assigning numbers to the variables. In this non-commutative world, assigning numbers will clearly not be sufficient. We would need to specialize the variables to something that is not commutative. The canonical example of things that do not multiply commutatively are matrices! From the work of Amitsur, Cohn and others, it can be deduced that if an identity is not true, then indeed there is a assignment of matrices to the variables for which both sides do not agree. The question then becomes whether we can find such an assignment? And if so, how quickly can we find one? These questions throw us into an enchanting world at the interface of invariant theory, tensors and computational complexity which we will explore in this dissertation.

A lot of the work in this dissertation is joint with Harm Derksen. Most of the results have already been published or in the form of a preprint, see [11, 12, 13, 14, 15, 16, 62].

CHAPTER 2

Invariant theory

In this chapter, we will first introduce invariant theory in Section 2.1, and then in Section 2.2 describe the method of Popov and Derksen for finding bounds for the degree of generators. In Section 2.3, we recall a fantastic result of Weyl on polarization, and in Section 2.4 we treat the notion of separating invariants. The central objects in this dissertation – matrix invariants and matrix semi-invariants – are introduced in Section 2.5 along with the main problems that we treat in this dissertation.

2.1 Introduction

Fix an infinite field K. Let G be an algebraic group acting on an (affine) algebraic variety X. We denote the ring of regular functions on X by $K[X]$. A function $f \in K[X]$ is called invariant if it is constant along orbits, i.e., $f(g \cdot x) = f(x)$ for all $g \in G$ and $x \in X$. The set of all invariant functions forms a subring $K[X]^G \subseteq K[X]$ called the *ring of invariants* or the *invariant ring*.

A vector space V is an example of an algebraic variety. If a group G acts on V, then V is called a representation of G. A representation V of G can also be seen as a homomorphism $G \to \mathrm{GL}(V)$. If the map $G \to \mathrm{GL}(V)$ is given by regular functions, then we call V a rational representation of G. The following result due to Hilbert is a landmark result in the history of invariant theory.

Theorem 2.1.1 (Hilbert)**.** *Let V be a rational representation of a linearly reductive group G, then the ring of invariants $K[V]^G$ is finitely generated.*

In fact, Hilbert's papers on invariant theory (see [43, 44]) also propelled the rise of algebraic geometry and commutative algebra, both of which have had a profound influence on invariant theory. In characteristic 0, reductive groups and linearly reductive groups coincide. In positive characteristic, this is not the case, and even the classical groups –

$\mathrm{GL}_n, \mathrm{SL}_n$ etc are not linearly reductive. Nevertheless, Nagata generalized Hilbert's result by weakening the hypothesis from linearly reductive to geometrically reductive, see [67]. Haboush proved that all reductive groups are geometrically reductive, see [41], and as a result, we now know that invariant rings for rational representations of reductive groups are finitely generated.

Theorem 2.1.2 (Nagata-Haboush)**.** *Let V be a rational representation of a reductive group G, then the ring of invariants $K[V]^G$ is finitely generated.*

These theorems are unsatisfactory on one count – they are not constructive. In other words, while we know that these invariant rings are finitely generated, the proofs do not produce a set of generators. A fundamental question in classical invariant theory is the following:

Problem 2.1.3. *Construct a minimal set of generators for $K[V]^G$.*

We know minimal sets of generators in very few cases, and it is an extremely difficult question to answer in general. Note that $K[V]$ is a polynomial ring and hence has a natural grading. This grading descends to the invariant ring $K[V]^G$. A more tractable problem is to find a bound on the degree of generators. To make this precise, we make a definition.

Definition 2.1.4. *The number $\beta(K[V]^G)$ is defined to be the smallest nonnegative integer d such that the invariants of degree $\leq d$ generate $K[V]^G$, i.e.,*

$$\beta(K[V]^G) = \min\{d \in \mathbb{Z}_{\geq 0} |\ K[V]^G_{\leq d} \textit{ generates } K[V]^G\},$$

where $K[V]^G_{\leq d}$ denotes the invariants of degree $\leq d$.

Problem 2.1.5. *Can we give an upper bound for $\beta(K[V]^G)$?*

2.2 Method of Popov and Derksen for degree bounds

We assume K is an algebraically closed field of characteristic 0 for this section. Several decades after Hilbert, Popov came up with a strategy to find upper bounds for the degree of generators. The core of this method is the Hochster-Roberts theorem and Kempf's results on the Hilbert series of an invariant ring. Let us recall the Hochster-Roberts theorem, see [45].

Theorem 2.2.1 (Hochster-Roberts)**.** *Let V be a rational representation of a linearly reductive group G. Then the invariant ring $K[V]^G$ is Cohen-Macaulay.*

We will explain the significance of this theorem, without bothering to define what a Cohen-Macaulay ring is. A set of homogeneous invariants $\{f_1, \ldots, f_r\}$ is called a homogeneous system of parameters (hsop) if $f_1, \ldots, f_r$ are algebraically independent and $K[V]^G$ is a finite module over $K[f_1, \ldots, f_r]$. In particular r must be the Krull dimension of $K[V]^G$. That $K[V]^G$ is Cohen-Macaulay implies that $K[V]^G$ is in fact a finite free module over $K[f_1, \ldots, f_r]$ for any hsop $\{f_1, \ldots, f_r\}$.

Now, suppose we have a hsop $\{f_1, \ldots, f_r\}$, and let $g_1, \ldots, g_s$ be a set of free module generators (these can be chosen to be homogeneous). The f_i are called primary invariants and the g_i are called secondary invariants. Observe that $\{f_1, \ldots, f_r, g_1, \ldots, g_s\}$ is a generating set, and hence $\beta(K[V]^G)$ is bounded above by the largest degree among the primary and secondary invariants. Popov's method was to utilize this.

In fact, a result of Kempf (see [54]) tells us that $\deg(g_j) \leq \sum_{i=1}^{r} \deg(f_j)$. Kempf's result is often seen in the following form – the Hilbert series of the invariant ring is a proper rational function. Let us define the Hilbert series.

Definition 2.2.2. *For a graded K-algebra $R = \bigoplus_{d \in \mathbb{Z}} R_d$, we define its Hilbert series*

$$H(R, t) = \sum_{d \in \mathbb{Z}} \dim(R_d) t^d.$$

The two different formulations of Kempf's results are the same, and this follows from the fact that

$$H(K[V]^G, t) = \frac{\sum_j t^{\deg(g_j)}}{\prod_i (1 - t^{\deg(f_i)})}.$$

In any case, let us observe that Kempf's result reduces the problem to finding a hsop, and let us record this.

Proposition 2.2.3. *Let $f_1, \ldots, f_r$ be a hsop. Then we have*

$$\beta(K[V]^G) \leq \sum_{i=1}^{r} \deg(f_i).$$

Improvements to Kempf's results were made by Knop in [57, 58], but we will not recall them here. Having reduced the problem to finding a hsop, we give an alternate characterization of a hsop in terms of the null cone. For a subset $S \subset K[V]$, define $\mathbb{V}(S) = \{v \in V \mid f(v) = 0 \; \forall f \in S\}$.

Definition 2.2.4. *The null cone $\mathcal{N}(G, V)$ is defined as the set of points in V that vanish on*

all invariant polynomials with no constant terms, i.e,

$$\mathcal{N}(G,V) = \mathbb{V}(K[V]^G_+),$$

where $K[V]^G_+ = \bigoplus_{d\in\mathbb{Z}_{>0}} K[V]^G_d$.

It turns out that a set of homogeneous invariants $\{f_1,\dots,f_r\}$ is a hsop if and only if the f_i's are algebraically independent and $\mathbb{V}(f_1,\dots,f_r) = \mathcal{N}(G,V)$.

Definition 2.2.5. *The number $\gamma(K[V]^G)$ is defined as the smallest non negative integer d such that non-constant homogeneous invariants of degree $\leq d$ cut out the null cone, i.e.,*

$$\gamma(K[V]^G) = \max\{d\in\mathbb{Z}_{\geq 0} \mid \mathbb{V}(\bigoplus_{i=1}^{d} K[V]^G_i) = \mathcal{N}(G,V)\}.$$

Popov showed the existence of a hsop in terms of $\gamma(K[V]^G)$. More precisely, he observed that if $D = lcm(1,2,\dots,\gamma(K[V]^G))$, then there exists a hsop $f_1,\dots,f_r$ with $\deg(f_i) = D$. So, if $r = \dim(K[V]^G)$, we have

$$\beta(K[V]^G) \leq r(lcm(1,2\dots,\gamma(K[V]^G)).$$

In effect, Popov reduced the problem to finding a bound for $\gamma(K[V]^G)$. Derksen sharpened this result considerably in [8].

Theorem 2.2.6 (Derksen)**.** *We have*

$$\beta(K[V]^G) \leq \frac{3}{8} r\gamma(K[V]^G)^2.$$

Derksen's result is much stronger than Popov's in the sense that Derksen's bound is polynomial in $\gamma(K[V]^G)$, whereas Popov's is not. However, this still requires us to find a bound for $\gamma(K[V]^G)$. We consider the representation $G \to \mathrm{GL}(V) \subset \mathrm{End}(V)$, we look at the image of G as a subvariety of the affine space $\mathrm{End}(V)$. Popov showed that the degree of this image happens to be an upper bound for $\gamma(K[V]^G)$. The best bounds for the degree are also due to Derksen, but we omit the details, referring the interested reader to [8].

2.3 Weyl's polarization theorem

Let V be a representation of G. We are interested in considering the diagonal action of G on several copies of V. More precisely, for $g\in G$ and $(v_1,\dots,v_m)\in V^m := V^{\oplus m}$, we

have

$$g \cdot (v_1, \ldots, v_m) = (g \cdot v_1, \ldots, g \cdot v_m).$$

It is a theorem of Weyl (see [56]) in characteristic 0 that invariants in $K[V^m]^G$ can be obtained by polarization and restitution from invariants in $K[V^n]^G$, where $n = \dim V$. As far as degree bounds are concerned, this translates to the following:

Theorem 2.3.1 (Weyl)**.** *Assume* $\mathrm{char}(K) = 0$*. Let* V *be an* n*-dimensional representation of* G*. Then we have the inequality*

$$\beta(K[V^m]^G) \leq \beta(K[V^n]^G) \; \forall m \in \mathbb{Z}_{>0}.$$

We outline a proof of the above theorem not in terms of polarization and restitution, but in terms of Schur modules. For a vector space V, let $S_\lambda(V)$ denote the Schur module associated to a partition λ. If λ has at most $n = \dim V$ parts, then this is the irreducible representation of $\mathrm{GL}(V)$ with highest weight λ. If λ has more than n parts, then $S_\lambda(V) = 0$. The assignment $V \to S_\lambda(V)$ is a functor, which is called the Schur functor. We recall some basic facts on Schur functors.

Proposition 2.3.2. *Let* V, W *be vector spaces and* λ *a partition. Then, we have*

$$S_\lambda(V \otimes W) = \bigoplus_{\mu,\nu} (S_\mu(V) \otimes S_\nu(W))^{a_{\lambda,\mu,\nu}},$$

where $a_{\lambda,\mu,\nu}$ *are known as the Kronecker coefficients.*

When we take $\lambda = (m)$, the Schur module is a symmetric power Sym^m. We write $\lambda \vdash d$ to denote that λ is a partition of d. Specializing the above proposition to this case, we get

Corollary 2.3.3. *Let* V, W *be vector spaces. Then we have*

$$\mathrm{Sym}^d(V \otimes W) = \bigoplus_{\lambda \vdash d} S_\lambda(V) \otimes S_\lambda(W).$$

Now, we can identify the representation V^m with $V \otimes W$, where W is an m-dimensional vector space on which G acts trivially. Consider the degree d homogeneous polynomials

$$K[V^m]_d := \mathrm{Sym}^d(V^* \otimes W^*) = \bigoplus_{\lambda \vdash d} S_\lambda(V^*) \otimes S_\lambda(W^*).$$

We make two observations about the above formula. The first is that we can restrict to partitions having at most n parts, since otherwise $S_\lambda(V^*) = 0$. The second is that in this

decomposition, G acts on only on $S_\lambda(V^*)$ and the action on $S_\lambda(W^*)$ is trivial. Hence, we have

$$K[V^m]_d^G = \bigoplus_{\lambda \vdash d,\ l(\lambda) \leq n} S_\lambda(V^*)^G \otimes S_\lambda(W^*), \tag{2.1}$$

where $l(\lambda)$ denotes the number of parts in the partition, also called the length of the partition. The group $\mathrm{GL}(W)$ acts on $V \otimes W$ by acting on the second tensor factor. This action commutes with the action of G, and hence we have an action of $\mathrm{GL}(W)$ on $K[V \otimes W]^G = K[V^m]^G$. Further the multiplication in the ring $K[V^m]^G$ is $GL(W)$ equivariant.

Proof of Theorem 2.3.1. We identify V^m with $V \otimes W$, where $\dim W = m$. Suppose $\beta(K[V^m]^G) = D$. This means that there is some invariant of degree D that cannot be formed as linear combinations of products of invariants of smaller degree. In other words, the multiplication map

$$\bigoplus_{i=1}^{D-1} K[V^m]_i^G \otimes K[V^m]_{D-i}^G \to K[V^m]_D^G$$

is not surjective. By equivariance of the multiplication map, and the semisimplicity of the representation theory, the image is a direct summand of $K[V^m]_D^G$. That the map is not surjective implies that there is at least one irreducible $\mathrm{GL}(W)$–subrepresentation $P \subseteq K[V^m]_D^G$ that is not in the image. We must have $P \cong S_\lambda(W^*)$ for some $\lambda \vdash D$. For P to be nonempty, we must have $l(\lambda) \leq \dim W$. From Equation 2.1, we also know that $l(\lambda) \leq n$.

We also observe that this happens on a purely functorial level, and whether a summand is in the image or not doesn't really depend on the vector space W. As long as the summand is non-empty, the invariants in that summand cannot be generated from invariants of smaller degree. Since $l(\lambda) \leq n$, this summand is non-empty if we take $\dim W = n$. This shows the existence of invariants of degree D in $K[V^n]^G$ that cannot be generated from invariants of smaller degree. This argument is perhaps best seen in the framework of twisted commutative algebras, but we will refrain from doing that here. □

2.4 Separating invariants

Assume K is algebraically closed. Let V be a representation of G. For a point $v \in V$, its orbit $G \cdot v = \{g \cdot v \mid g \in G\} \subseteq V$ is not necessarily closed. It follows from continuity that any invariant polynomial must take the same value on all points of the closure of an orbit.

Hence invariant polynomials cannot distinguish two points whose orbit closures intersect.

We can ask the converse question – If $v, w \in V$ such that $\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset$, then is there an invariant polynomial $f \in K[V]^G$ such that $f(v) \neq f(w)$? The answer to this question is in general negative. Indeed, consider the additive group $\mathbb{G}_a$ acting on the affine plane K^2 by $t \cdot (x, y) \mapsto (x, tx + y)$. The invariant ring is $K[x] \subset k[x, y]$, and hence every invariant polynomial takes the same value on every point on the y axis. On the other hand, every point on the y-axis is a closed orbit! Hence invariant functions do not suffice to separate points whose orbit closures do not intersect. However, if we enforce additional hypothesis, we get a positive answer as the theorem below shows.

Theorem 2.4.1. *Let V be a rational representation of a reductive group G. Then for $v, w \in V$, there exists $f \in K[V]^G$ such that $f(v) \neq f(w)$ if and only if $\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset$.*

While separating invariants can be defined for more general group actions, we will assume from here on that V is a rational representation of a reductive group G as all the cases we are interested in are of this form. For a more general treatment, we refer to [10].

Definition 2.4.2. *A subset of invariants $S \subset K[V]^G$ is called a separating set of invariants if for every pair $v, w \in V$ such that $\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset$, there exists $f \in S$ such that $f(v) \neq f(w)$.*

Any generating set of invariants is a separating set, but we may be able to find simpler separating sets that do not generate. We make another definition.

Definition 2.4.3. *We define $\beta_{\text{sep}}(K[V]^G)$ to be the smallest non-negative integer d such that the invariants of degree $\leq d$ form a separating set of invariants, i.e.,*

$$\beta_{\text{sep}}(K[V]^G) = \min\{d \in \mathbb{Z}_{\geq 0} \mid K[V]^G_{\leq d} \text{ is a separating set of invariants}\}.$$

Clearly $\beta_{\text{sep}}(K[V]^G) \leq \beta(K[V]^G)$. Weyl's polarization theorem is unfortunately not true in positive characteristic, see [22, 24]. Nevertheless, Draisma, Kemper and Wehlau proved a version of this result for separating invariants in [32].

Theorem 2.4.4 (Draisma, Kemper, Wehlau)**.** *Let V be an n-dimensional representation of G. Then for all $m \in \mathbb{Z}_{>0}$, we have*

$$\beta_{\text{sep}}(K[V^m]^G) \leq \beta_{\text{sep}}(K[V^n]^G).$$

2.5 Matrix invariants and matrix semi-invariants: Main Results

A lot of extremely interesting connections between invariant theory and computational complexity have been found. Perhaps the most interesting and well known is Mulmuley and Sohoni's reformulation of Valiant's algebraic version of P vs NP into a question about orbits of algebraic groups. This is known as the Geometric Complexity Theory (GCT) program, see [65, 66]. The most relevant to us is however the connection between non-commutative circuits and identity testing with the ring of matrix semi-invariants formulated by Hrubes and Wigderson in [46]. We will explain the connections in more detail later. Here we will introduce matrix invariants and matrix semi-invariants, the primary objects of study in this dissertation. We will also point out some of the main results.

2.5.1 Matrix invariants

Let $\mathrm{Mat}_{p,q}$ denote the set of $p \times q$ matrices. Consider the group $G = \mathrm{GL}_n$ acting by simultaneous conjugation on $\mathrm{Mat}_{n,n}^m$, the space of m-tuples of $n \times n$ matrices. More precisely, for $g \in \mathrm{GL}_n$ and $(X_1, \ldots, X_m) \in \mathrm{Mat}_{n,n}^m$, we have

$$g \cdot (X_1, \ldots, X_m) = (gX_1g^{-1}, \ldots, gX_mg^{-1}).$$

We write $S(n, m) = K[\mathrm{Mat}_{n,n}^m]^{\mathrm{GL}_n}$ for the ring of invariants for this action. Let us consider the simplest case when $m = 1$. For an $n \times n$ matrix X, consider it's characteristic polynomial $\det(I + tX)$. This is a degree n polynomial in t. Let $\sigma_j(X)$ denote the coefficient of t^j. Then it is easy to see that $\sigma_j(X)$ is a polynomial in the entries of X, and moreover it is invariant under the action of GL_n by conjugation, i.e., $\sigma_j \in K[\mathrm{Mat}_{n,n}]^{\mathrm{GL}_n} = S(n, 1)$. In fact, these form an algebraically independent set of generators for $S(n, 1)$.

Proposition 2.5.1. *We have* $S(n, 1) = K[\sigma_j \mid 1 \leq j \leq n]$.

Proof. Let Σ_n denote the symmetric group on n letters. Let $W = K^n$ and consider the natural action of Σ_n on W by permuting the coordinates. Let $x_1, \ldots, x_n$ denote the coordinate functions, then the invariant ring $K[W]^{\Sigma_n} = K[x_1, \ldots, x_n]^{\Sigma_n}$ is known as the ring of symmetric functions. Let

$$e_j = \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq n} x_{i_1} x_{i_2} \ldots x_{i_j},$$

denote the j^{th} elementary symmetric polynomial. The elementary symmetric functions $e_1, \ldots, e_n$ are algebraically independent and generate $K[W]^{\Sigma_n}$, a result that can be traced back to Newton.

We have an inclusion $\phi : W \hookrightarrow \mathrm{Mat}_{n,n}$ sending $\lambda = (\lambda_1, \ldots, \lambda_n)$ to the diagonal matrix D_λ whose diagonal entries are $\lambda_1, \ldots, \lambda_n$. We get a map on the coordinate rings $\phi^* : K[\mathrm{Mat}_{n,n}] \to K[W]$. We will show that this map descends to an isomorphism on the invariant rings $\phi^* : K[\mathrm{Mat}_{n,n}]^{\mathrm{GL}_n} = S(n,1) \to K[W]^{\Sigma_n}$ sending $\sigma_j \mapsto e_j$.

For $f \in S(n,1)$, we first show that $\phi^*(f) \in K[W]^{\Sigma_n}$. Indeed, if we take $\lambda, \mu \in W$ such that λ is a permutation of μ, then D_λ and D_μ are in the same GL_n orbit. Thus $\phi^*(f)(\lambda) = f(D_\lambda) = f(D_\mu) = \phi^*(f)(\mu)$.

Next, we show that the map is injective. Indeed if $\phi^*(f) = 0$, then f vanishes on all diagonal matrices. Since f must be constant on GL_n–orbits, it vanishes on all diagonalizable matrices, which are dense in $\mathrm{Mat}_{n,n}$. Since f vanishes on a dense subset of $\mathrm{Mat}_{n,n}$, it vanishes on all of $\mathrm{Mat}_{n,n}$, i.e., $f = 0$.

To show that it is surjective, observe that the image of σ_j is e_j, and e_j for $1 \leq j \leq n$ form a generating set. This gives the required conclusion. □

Remark 2.5.2. *In characteristic* 0*, we also have* $S(n,1) = K[\mathrm{Tr}_j \mid 1 \leq j \leq n]$*, where* Tr_j *is the polynomial that maps a matrix* $X \mapsto \mathrm{Tr}(X^j)$*. This is easy to see since* $\mathrm{Tr}(X^j)$ *is the power sum symmetric function in the eigenvalues, whereas* $\sigma_j(X)$ *is the elementary symmetric function in the eigenvalues. In characteristic* 0*, the power sum symmetric functions are also an algebraically independent set of generators for the ring of symmetric functions, and this is seen by Newton's identites. However, these identities involve denominators, and hence this is no longer true in positive characteristic.*

Procesi generalized the description in terms of traces to get a description of generators for $S(n,m)$ in [70]. A word w in an alphabet set S is a string of elements $w = w_1 w_2 \ldots w_k$ with $w_i \in S$. We denote the set of all words in an alphabet set S by $\mathrm{words}(S)$. For a word $w = w_1 w_2 \ldots w_k$ in the alphabet set $[m] := \{1, 2, \ldots, m\}$, we define $\mathrm{Tr}_w \in S(n,m)$ as the invariant polynomial that sends $(X_1, \ldots, X_m) \mapsto \mathrm{Tr}(X_{w_1} X_{w_2} \ldots X_{w_k})$

Theorem 2.5.3 (Procesi)**.** *Assume* $\mathrm{char}(K) = 0$*. Then the invariants* $\{\mathrm{Tr}_w \mid w \in \mathrm{words}([m])\}$ *generate the invariant ring* $S(n,m)$*.*

Procesi's theorem gives an infinite set of generators. A couple of years prior to Procesi's theorem, Razmyslov had studied trace identities in [72], and in particular showed that the trace of any word of length $> n^2$ can be written in terms of traces of smaller words. In our notation, this can be reformulated as follows:

Theorem 2.5.4 (Razmyslov). *Assume* $\mathrm{char}(K) = 0$. *Then we have* $\beta(S(n,m)) \leq n^2$.

In the case of positive characteristic, Donkin gave a description of generators in terms of coefficients of the characteristic polynomials rather than traces, see [29]. For a word $w = w_1 w_2 \dots w_k \in \mathrm{words}([m])$, define $\sigma_j(w) \in S(n,m)$ by $\sigma_j(w)(X_1, \dots, X_m) = \sigma_j(X_{w_1} \dots X_{w_k})$.

Theorem 2.5.5 (Donkin). *The invariants* $\{\sigma_j(w) \mid 1 \leq j \leq n, w \in \mathrm{words}([m])\}$ *generate the invariant ring* $S(n,m)$.

Bounds on the degree of generators were given by Domokos, see [22].

Theorem 2.5.6 (Domokos). *We have* $\beta(S(n,m)) = O(n^7 m^n)$.

It is unfortunate that the bound on the degree of generators is far worse in positive characteristic, and it is only natural to ask if whether there exists a polynomial bound (in n and m)?

Problem 2.5.7. *Is there a polynomial bound for* $\beta(S(n,m))$ *in positive characteristic?*

Generating sets give separating sets, but can we find better bounds for separating invariants?

Problem 2.5.8. *Give good bounds for* $\beta_{\mathrm{sep}}(S(n,m))$.

We will resolve both these problems as part of this dissertation, see Corollary 5.2.3 and Theorem 6.5.1.

2.5.2 Matrix semi-invariants

Consider the left-right action of $G = \mathrm{SL}_n \times \mathrm{SL}_n$ on $V = \mathrm{Mat}_{n,n}^m$. For $(A,B) \in \mathrm{SL}_n \times \mathrm{SL}_n$ and $(X_1, \dots, X_m) \in \mathrm{Mat}_{n,n}^m$, this action is given by

$$(A,B) \cdot (X_1, \dots, X_m) = (AX_1B^{-1}, \dots, AX_mB^{-1}).$$

We set $R(n,m) = K[V]^G$, the ring of invariants for this action. Once again, let us consider the simplest case $m = 1$. We define $\det$ to be the polynomial that sends a matrix $X \mapsto \det(X)$, the determinant of X. The following can be proved in a fashion similar to Proposition 2.5.1.

Proposition 2.5.9. *The ring* $R(n,1) = K[\det]$.

The case of $m = 2$ is already non-trivial and goes back to Happel. Consider the expression $\det(X_1 + tX_2)$ as a polynomial in t, and denote by f_i, the coefficient of t^i. Then it is easy to see that f_i is a polynomial in the entries of X_1 and X_2. In fact, it is an invariant polynomial, i.e., $f_i \in R(n, 2)$.

Proposition 2.5.10 (Happel)**.** *The invariants $f_0, f_1, \ldots, f_n$ are algebraically independent and generate $R(n, 2)$.*

The description for the ring of invariants for general m is deduced from the description for semi-invariants of quivers. Matrix semi-invariants are a special case – $R(n, m)$ is the ring of semi-invariants for the m-Kronecker quiver for the dimension vector (n, n) and we will discuss this later. For now, we are content to give the description in this special case.

Given two matrices $A = (a_{ij})$ of size $m \times n$, and $B = (b_{ij})$ of size $p \times q$, we define their tensor (or Kronecker) product to be

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1}B & \cdots & \cdots & a_{mn}B \end{bmatrix} \in \mathrm{Mat}_{mp,nq}\,.$$

For $T = (T_1, T_2, \ldots, T_m) \in \mathrm{Mat}_{d,d}^m$, we define a homogeneous invariant $f_T \in R(n, m)$ of degree dn by

$$f_T(X_1, X_2, \ldots, X_m) = \det(X_1 \otimes T_1 + X_2 \otimes T_2 + \cdots + X_m \otimes T_m). \tag{2.2}$$

The following result can be found in [17, 25, 73].

Theorem 2.5.11. *The invariant ring $R(n, m)$ is spanned by all f_T with $T \in \mathrm{Mat}_{d,d}^m$ and $d \geq 1$.*

Remark 2.5.12. *There are no homogeneous invariants of degree d in $R(n, m)$ unless d is a multiple of n.*

In fact, there are two ways of viewing the homogeneous invariant f_T, simply because the definition of the Kronecker product of matrices is not very symmetric. We could just as well have defined f_T with the tensor factors switched because we have

$$\det(T_1 \otimes X_1 + T_2 \otimes X_2 + \cdots + T_m \otimes X_m) = \det(X_1 \otimes T_1 + \cdots + X_m \otimes T_m),$$

even though on the level of matrices, $\sum_{i=1}^m T_i \otimes X_i$ and $\sum_{i=1}^m X_i \otimes T_i$ will usually be different. Both descriptions have their advantages, and some results are more transparent in one description. However, to avoid confusion, we will always use $\sum_{i=1}^m X_i \otimes T_i$.

The following bounds were known if K has characteristic 0:

1. $\beta(R(n,1)) = \beta(n,2) = n$;
2. $\beta(R(1,m)) = 1$;
3. $\beta(R(2,m)) \leq 4$;
4. $\beta(R(3,3)) = 9$;
5. $\beta(R(n,m)) = O(n^4((n+1)!)^2)$.

The bounds in (1) follow from the descriptions of $R(n,1)$ and $R(n,2)$ above and (2) is trivial. The bound (3) can be found in [20] (see also [50]). This bound also follows from the First Fundamental Theorem of Invariant Theory for SO_4, because $\mathrm{SL}_2 \times \mathrm{SL}_2$ is a finite central extension of SO_4 and the representation $\mathrm{Mat}_{2,2}$ of $\mathrm{SL}_2 \times \mathrm{SL}_2$ corresponds to the standard 4-dimensional representation of SO_4. The bound (4) was given in [21]. The general bound on the degree of generating invariants due to Derksen mentioned in Section 2.2 gives a bound of $O(n^8 16^{n^2})$ and Ivanyos, Qiao and Subrahmanyam showed in [49, 50] that this bound can be improved to the factorial one (5).

Problem 2.5.13. *Are there polynomial bounds for $\beta(R(n,m))$ and $\beta_{\mathrm{sep}}(R(n,m))$?*

One of the main results of this dissertation is a positive answer to this question, see Corollary 5.1.12 and Corollary 6.5.8.

CHAPTER 3

Quivers

Matrix invariants and matrix semi-invariants are special cases of invariant rings associated to quivers. Hence, it is only natural to generalize our results to this generality. In Section 3.1, we introduce quivers and their representations. In Section 3.2, we discuss various invariant rings associated to quivers, connecting them to matrix invariants and matrix semi-invariants.

3.1 Quiver representations

A quiver is a (finite) directed graph. More formally, a quiver Q is a 4-tuple (Q_0, Q_1, h, t) where Q_0 is the set of vertices, Q_1 is the set of arrows, and $h, t : Q_1 \to Q_0$ denote the head and tail of the arrow respectively.

Example 3.1.1 (m-loop quiver)**.** *The m-loop quiver $\Upsilon(m)$ is a quiver with one vertex $Q_0 = \{v\}$, and m arrows $Q_1 = \{a_1, \ldots, a_m\}$ whose head and tail are the only vertex v.*

Example 3.1.2 (m-Kronecker quiver)**.** *The m-Kronecker quiver $\Theta(m)$ is the quiver with two vertices $Q_0 = \{x, y\}$ and m arrows $Q_1 = \{a_1, \ldots, a_m\}$, with $h(a_i) = x$ and $t(a_i) = y$ for all $1 \leq i \leq m$.*

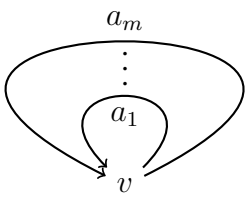

Figure 3.1: m-loop quiver

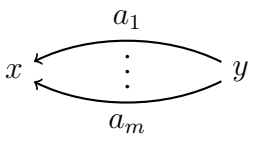

Figure 3.2: Kronecker quiver

A representation V of a quiver Q is a vector space V_x for each vertex $v \in Q_0$ and a linear map from $V(a) : V(ta) \to V(ha)$ for each arrow $a \in Q_1$. For two representations V, W, a morphism $f : V \to W$ is given by a collection of linear maps $f_x : V(x) \to W(x)$ such that for each arrow a, the following diagram commutes.

$$\begin{array}{ccc} V(ta) & \xrightarrow{V(a)} & V(ha) \\ \downarrow{\scriptstyle f_{ta}} & & \downarrow{\scriptstyle f_{ha}} \\ W(ta) & \xrightarrow{W(a)} & W(ha) \end{array}$$

The notions of subrepresentation, kernel, quotient, direct sum, summand etc are defined in the obvious way. A representation V is called finite dimensional if $V(x)$ is finite dimensional for each $x \in Q_0$. We denote by $\mathrm{Rep}(Q)$, the category of finite dimensional representations of a quiver Q.

A path of length k is a sequence $p = a_k a_{k-1} \cdots a_1$ where $a_1, \ldots, a_k$ are arrows such that $ha_{i-1} = ta_i$ for $i = 2, 3, \ldots k$. The head and tail of the path are defined by $hp = ha_k$ and $tp = ta_1$ respectively. For every vertex $x \in Q_0$ we also have a trivial path ε_x of length 0 such that $h\varepsilon_x = t\varepsilon_x = x$. For a path $p = a_k a_{k-1} \cdots a_1$ is a path, then we define

$$V(p) = V(a_k)V(a_{k-1}) \cdots V(a_1) : V(tp) \to V(hp).$$

We define $V(\varepsilon_x)$ is the identity map from $V(x)$ to itself.

We define the path algebra KQ first as a vector space with basis all paths of Q. For two paths p and q, we define $p \cdot q$ to be the concatenation of p and q if it forms a path, and 0 otherwise. This gives KQ the structure of an algebra. To any representation V of Q, we can associate a module M_V over KQ in the following way. Let $M_V = \bigoplus_{x \in Q_0} V(x)$ as a vector space. To describe an action of KQ, it suffices to describe what each path p acts by. The path p acts on $V(tp)$ as the map $V(p) : V(tp) \to V(hp)$, and on all other summands of M_V by 0. This is in fact an equivalence of categories, i.e., the category of finitely generated KQ-modules is equivalent to the category $\mathrm{Rep}(Q)$.

For a finite dimensional representation V of Q, we define its dimension vector $\underline{\dim}(V) : Q_0 \to \mathbb{Z}_{\geq 0}$ sending $x \mapsto \dim(V(x))$. For a dimension vector α, we define the representation space

$$\mathrm{Rep}(Q, \alpha) = \bigoplus_{a \in Q_1} \mathrm{Mat}_{\alpha(ha), \alpha(ta)} .$$

If V is a representation with dimension vector α and we identify $V(x) \cong K^{\alpha(x)}$ for all $x \in Q_0$, then V can be viewed as an element of $\mathrm{Rep}(Q, \alpha)$. Consider the group $\mathrm{GL}(\alpha) = \prod_{x \in Q_0} \mathrm{GL}_{\alpha(x)}$. The group $\mathrm{GL}(\alpha)$ acts on $\mathrm{Rep}(Q, \alpha)$ by:

$$(A(x) \mid x \in Q_0) \cdot (V(a) \mid a \in Q_1) = (A(ha)V(a)A(ta)^{-1} \mid a \in Q_1).$$

For $V \in \mathrm{Rep}(Q, \alpha)$, choosing a different basis means acting by the group $\mathrm{GL}(\alpha)$. The $\mathrm{GL}(\alpha)$-orbits in $\mathrm{Rep}(Q, \alpha)$ correspond to isomorphism classes of representations of dimension α.

Example 3.1.3 (1-loop quiver)**.** *For a dimension vector n, the representation space $\mathrm{Rep}(\Upsilon_1, n) = \mathrm{Mat}_{n,n}$. The group $\mathrm{GL}(n)$ acts by base change with the formula $g \cdot X = gXg^{-1}$ for $g \in \mathrm{GL}(n)$ and $X \in \mathrm{Rep}(\Upsilon_1, n)$. A classic linear algebra result translates to the following – the orbits are in $1-1$ correspondence with Jordan canonical forms (if K is algebraically closed).*

Example 3.1.4 (1-Kronecker quiver)**.** *For a dimension vector (p, q), the representation space $\mathrm{Rep}(\Theta_1, (p, q)) = \mathrm{Mat}_{p,q}$. The base change group $\mathrm{GL}(p) \times \mathrm{GL}(q)$ acts by left-right multiplication, i.e., $(A, B) \cdot X = AXB^{-1}$ for $(A, B) \in \mathrm{GL}(p) \times \mathrm{GL}(q)$ and $X \in \mathrm{Rep}(\Theta_1, (p, q))$. It is easy to see that two matrices are in the same orbit if and only if they have the same rank.*

Of course, these examples are very basic. If we take more complicated quivers, then trying to understand the isomorphism classes becomes exceptionally challenging. A quiver is called finite type if the number of indecomposable representations (of any dimension) is finite. The set of finite type quivers are precisely the simply laced Dynkin diagrams, and in particular, cannot have any directed cycles. There is tripartite classification of quivers – finite type, tame and wild. Tame quivers have the property that the number of indecomposables up to isomorphism in any particular dimension vector are parametrized by a finite set and finitely many 1-parameter families (in a way that Kac made precise). Any quiver that is not tame is called wild, as no one has a clue how to classify their indecomposables.

3.2 Invariant theory of quivers

3.2.1 Invariants

Fix a quiver Q, and a dimension vector α. The group $\mathrm{GL}(\alpha)$ acts (on the left) on the ring $K[\mathrm{Rep}(Q,\alpha)]$ of polynomial functions on $\mathrm{Rep}(Q,\alpha)$ by

$$A \cdot f(V) = f(A^{-1} \cdot V)$$

where $f \in K[\mathrm{Rep}(Q,\alpha)]$, $V \in \mathrm{Rep}(Q,\alpha)$ and $A \in \mathrm{GL}(\alpha)$.

The ring of invariants for this action is denoted

$$\mathrm{I}(Q,\alpha) := K[\mathrm{Rep}(Q,\alpha)]^{\mathrm{GL}(\alpha)}.$$

Example 3.2.1. *For the m-loop quiver Υ_m and the dimension vector n, the representation space is $\mathrm{Mat}_{n,n}^m$ and the base change group $\mathrm{GL}(n)$ acts on the representation space by simultaneous conjugation, i.e., we have*

$$\mathrm{I}(\Upsilon_m, n) = S(n,m).$$

In fact, for any quiver Q and any dimension vector α, we can get a description of the invariant ring in terms of the descriptions we have for $S(n,m)$. A cyclic path is a path p of positive length such that $hp = tp$. For a cyclic path p, we define an invariant $T_p : \mathrm{Rep}(Q,\alpha) \to K$ by $T_p(V) = \mathrm{Tr}(V(p))$, the trace of the endomorphism $V(p)$. It is easy to see that this function is invariant under the action of $\mathrm{GL}(\alpha)$. LeBruyn and Procesi showed that such invariants generate $\mathrm{I}(Q,\alpha)$ in characteristic zero.

Although it was only proved a couple of decades after Procesi's result, the description is only a little more complicated in positive characteristic. For a cyclic path p, we define $\sigma_j(p) \in \mathrm{I}(Q,\alpha)$ by $\sigma_j(p)(V) = \sigma_j(V(p))$. That such invariants generate $\mathrm{I}(Q,\alpha)$ is due to Donkin. We will present this result in a slightly different format.

Let $N = \sum_{x \in Q_0} \alpha(x)$. For each $a \in Q_1$, consider the natural inclusion map

$$\mathrm{Mat}_{\alpha(ha),\alpha(ta)} = \mathrm{Hom}(K^{\alpha(ta)}, K^{\alpha(ha)}) \hookrightarrow \mathrm{End}(\bigoplus_{x \in Q_0} K^{\alpha(x)}) = \mathrm{Mat}_{N,N}\,.$$

Putting these inclusions together, we get an inclusion $\iota : \mathrm{Rep}(Q,\alpha) \to \mathrm{Mat}_{N,N}^M$, where $M = |Q_1|$. This gives a map $\iota^* : K[\mathrm{Mat}_{N,N}^M] \to K[\mathrm{Rep}(Q,\alpha)]$. Then LeBruyn–Procesi and Donkin's result can be succinctly presented as the statement that the map ι^* descends

to a surjection on invariant rings.

Theorem 3.2.2 (LeBruyn–Procesi, Donkin)**.** *Let Q, α, N and M be as in the discussion above. Then the map $\iota^* : S(N, M) \to \mathrm{I}(Q, \alpha)$ is surjective.*

Remark 3.2.3. *As a consequence of the above theorem, degree bounds for $S(n, m)$ immediately translate into degree bounds for $\mathrm{I}(Q, \alpha)$.*

Remark 3.2.4. *There are no non-trivial invariants if the quiver Q has no oriented cycles.*

3.2.2 Semi-invariants

If a quiver Q has no oriented cycles, then there are no invariants. Yet, it may have an interesting ring of semi-invariants. For a quiver Q with no oriented cycles and a dimension vector α, consider the subgroup

$$\mathrm{SL}(\alpha) = \prod_{x \in Q_0} \mathrm{SL}(\alpha(x)) \subseteq \mathrm{GL}(\alpha).$$

The invariant ring $\mathrm{SI}(Q, \alpha) = K[\mathrm{Rep}(Q, \alpha)]^{\mathrm{SL}(\alpha)}$ is called the ring of semi-invariants. A multiplicative character of the group GL_α is of the form

$$\chi_\sigma : (A(x) \mid x \in Q_0) \in \mathrm{GL}_\alpha \mapsto \prod_{x \in Q_0} \det(A(x))^{\sigma(x)} \in K^\star,$$

where $\sigma : Q_0 \to \mathbb{Z}$ is called the weight of the character χ_σ. Define

$$\mathrm{SI}(Q, \alpha)_\sigma = \{f \in K[\mathrm{Rep}(Q, \alpha)] \mid \forall A \in \mathrm{GL}(\alpha)\ A \cdot f = \chi_\sigma(A) f\}.$$

Then we have a weight space decomposition

$$\mathrm{SI}(Q, \alpha) = \bigoplus_\sigma \mathrm{SI}(Q, \alpha)_\sigma.$$

If $\sigma \cdot \alpha = \sum_{x \in Q_0} \sigma(x)\alpha(x) \neq 0$, then $\mathrm{SI}(Q, \alpha)_\sigma = 0$. Assume that $\sigma \cdot \alpha = 0$. We can write $\sigma = \sigma_+ - \sigma_-$ where $\sigma_+(x) = \max\{\sigma(x), 0\}$ and $\sigma_-(x) = \max\{-\sigma(x), 0\}$. Define $n = \sigma_+ \cdot \alpha = \sigma_- \cdot \alpha$.

Now we define an $n \times n$ linear matrix

$$A : \bigoplus_{x \in Q_0} V(x)^{\sigma_+(x)} \to \bigoplus_{x \in Q_0} V(x)^{\sigma_-(x)}$$

where each block $\mathrm{Hom}(V(x), V(y))$ is of the form $t_1V(p_1)+\cdots+t_rV(p_r)$ where $t_1, t_2, \ldots, t_r$ are indeterminates and $p_1, p_2, \ldots, p_r$ are all paths from x to y. We use different indeterminates for the different blocks, so the linear matrix has $m = \sum_{x\in Q_0}\sum_{y\in Q_0}\sigma_+(x)b_{x,y}\sigma_-(y)$ indeterminates where $b_{x,y}$ is the number of paths from x to y. We can write $A = t_1X_1 + \cdots + t_mX_m$ with $X_1, \ldots, X_m \in \mathrm{Mat}_{n,n}$. We have the following result (see [17, Corollary 3], [25] and [73]).

Theorem 3.2.5. *The space* $\mathrm{SI}(Q,\alpha)_\sigma$ *is spanned by* $\det(t_1X_1+\cdots+t_mX_m)$ *with* $t_1, \ldots, t_m \in K$.

Corollary 3.2.6. *For any positive integer* d*, the space* $\mathrm{SI}(Q,\alpha)_{d\sigma}$ *is spanned by* $\det(X_1 \otimes T_1 + \cdots + X_m \otimes T_m)$ *with* $T_1, \ldots, T_m \in \mathrm{Mat}_{d,d}$.

Proof. This follows from the construction for $d\sigma$ instead of σ. □

Let us define a subring $\mathrm{SI}(Q,\alpha,\sigma) = \bigoplus_{d\in\mathbb{Z}_{\geq 0}}\mathrm{SI}(Q,\alpha)_\sigma \subseteq \mathrm{SI}(Q,\alpha)$. The projective variety $\mathrm{Proj}(\mathrm{SI}(Q,\alpha,\sigma)$, if nonempty, is a moduli space for the α-dimensional representations of Q, see [55]. The above discussion can be formulated as the following:

Corollary 3.2.7. *Let* Q, α, σ, n *and* m *be as in the above discusiion. Then we have a surjective ring homomorphism* $\psi : R(n,m) \to \mathrm{SI}(Q,\alpha,\sigma)$ *which sends homogeneous elements of degree* dn *into* $\mathrm{SI}(Q,\alpha)_{d\sigma}$.

Remark 3.2.8. *Once again, degree bounds for* $R(n,m)$ *will immediately give degree bounds for the subrings* $\mathrm{SI}(Q,\alpha,\sigma)$*. However, to give degree bounds for the entire ring of semi-invariants* $\mathrm{SI}(Q,\alpha)$*, we will need additional work.*

CHAPTER 4

Linear matrices and non-commutative rank

We introduce linear matrices and the notions of commutative and non-commutative rank in Section 4.1. In Section 4.2, we present an alternate proof of the regularity lemma. We present some strange behaviour in the ranks of blow-ups in Section 4.3. A crucial result on the rank of blow-ups is proved in Section 4.4, and in Section 4.5, we provide examples where the ratio of non-commutative rank to commutative rank is large.

4.1 Introduction

We fix an infinite field K. A linear matrix (or matrix pencil) A over K is a matrix whose coefficients are linear expressions in variables $t_1, t_2, \ldots, t_m$, i.e.,

$$A = A_0 + t_1 A_1 + t_2 A_2 + \cdots + t_m A_m, \text{ with } A_i \in \mathrm{Mat}_{n,n}(K).$$

There are several interesting ranks one can define on a linear matrix. We start with the most obvious one.

Definition 4.1.1. *The commutative rank* $\mathrm{crk}(A)$ *is defined as the rank over the commutative function field* $K(t_1, t_2, \ldots, t_m)$

We can also take $t_1, t_2, \ldots, t_m$ to be independent non-commuting variables and compute the rank over the free skew field $K\langle\!\langle t_1, t_2, \ldots, t_m \rangle\!\rangle$.

Definition 4.1.2. *The non-commutative rank* $\mathrm{ncrk}(A)$ *is defined as the rank over the free skew field* $K\langle\!\langle t_1, t_2, \ldots, t_m \rangle\!\rangle$

We will not treat the skew field in detail as we will give a different characterization of non-commutative rank. We refer to [46, 38, 37] for more details on the skew field. Over any skew field (a.k.a division algebra), the rank of a matrix is defined as the (left) row rank,

which is equal to the (right) column rank. In particular, adding left multiplied rows to other rows and right multiplied columns to other columns does not affect the rank. Note also that a square matrix is invertible over a skew field if and only if it is of full rank.

The following well-known example shows that the commutative and non-commutative rank of a linear matrix may differ.

Example 4.1.3. *This example is based on skew symmetric matrices. Let*

$$A = \begin{pmatrix} 0 & 1 & t_1 \\ -1 & 0 & t_2 \\ -t_1 & -t_2 & 0 \end{pmatrix}.$$

It is easy to see that $\mathrm{crk}(A) = 2$*. However, over the free skew field* $K\langle\langle t_1, t_2 \rangle\rangle$*, we can do row and column transformations to transform* A *to*

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & [t_2, t_1] \end{pmatrix},$$

which is clearly of full rank since $[t_2, t_1] = t_2 t_1 - t_1 t_2$ *is non-zero over the skew field* $K\langle\langle t_1, t_2 \rangle\rangle$*.*

Example 4.1.4. *In [33], several low rank examples can be found. For example, they show that*

$$\begin{pmatrix} c & d & 0 & 0 \\ 0 & 0 & c & d \\ -a & 0 & -b & 0 \\ 0 & -a & 0 & -b \end{pmatrix} \text{ and } \begin{pmatrix} -b & -d & 0 & 0 \\ 0 & 0 & -c & -d \\ -d & 0 & b & 0 \\ c & a & 0 & b \end{pmatrix}$$

are linear matrices that have commutative rank 3*, and non-commutative rank* 4*. Here* a, b, c, d *are variables.*

4.1.1 Linear subspaces of matrices

Linear matrices can also be studied from the point of view of linear subspaces and their tensor blow-ups.

Definition 4.1.5. *We define the rank of a linear subspace* $\mathcal{X} \subseteq \mathrm{Mat}_{p,q}$ *to be the maximal rank among its members, and denote it by* $\mathrm{rk}(\mathcal{X})$*.*

The set of matrices in $\mathcal{X}$ having this maximal rank is Zariski open. Since the underlying field K is infinite, we can relate the commutative rank of a linear matrix to the rank of a linear subspace (see [37, Lemma 3.1]) as follows:

Lemma 4.1.6. *Let $A = X_0 + t_1X_1 + t_2X_2 + \cdots + t_mX_m$ be a linear matrix and let $\mathcal{X} = \operatorname{span}(X_0, X_1, X_2, \ldots, X_m)$. Then*

$$\operatorname{crk}(A) = \operatorname{rk}(\mathcal{X}).$$

Proof. Suppose $\operatorname{crk}(A) = r$. This means that there is an $r \times r$ minor M of A which is non-zero. This minor is a polynomial in the variables $t_1, \ldots, t_m$, i.e., $0 \neq M = p(t_1, \ldots, t_m) \in K[t_1, \ldots, t_m]$. Since the field is infinite, we can find $a_1, \ldots, a_m$ such that $p(a_1, \ldots, a_m) \neq 0$. Hence $A_0 + a_1A_1 + \cdots + a_mA_m \in \mathcal{X}$ with rank at least r. This shows that $\operatorname{rk}(\mathcal{X}) \geq \operatorname{crk}(A)$.

Now, suppose there is a matrix $a_0A_0 + \cdots + a_mA_m \in \mathcal{X}$, with rank $s > r$. Then there is some $s \times s$ minor that is non-zero. Let t be a variable and consider the same $s \times s$ minor in $tA_0 + a_1A_1 + \cdots + a_mA_m$. This is a polynomial in t, say $p(t)$, and we know that $p(t) \neq 0$ since $p(a_0) \neq 0$. Hence, there exists $a_0' \neq 0$ such that $p(a_0') \neq 0$. Thus w.l.o.g., replacing a_0 by a_0' if necessary, we can assume $a_0 \neq 0$. Then $A_0 + \frac{a_1}{a_0}A_1 + \cdots + \frac{a_m}{a_0}A_m \in \mathcal{X}$ has rank $s > r$. But this means that some $s \times s$ minor with $s > r$ is non-zero. This same minor in $A_0 + t_1A_1 + \ldots t_mA_m$ is therefore a non-zero polynomial. Hence $\operatorname{crk}(A) \geq k > r = \operatorname{crk}(A)$, which is a contradiction.

□

It turns out that non-commutative rank can also be understood from the perspective of linear subspaces. In order to do this, we require the notion of tensor blow-ups for linear subspaces.

Definition 4.1.7. *Let $\mathcal{X}$ be a linear subspace of $\operatorname{Mat}_{k,n}$. We define its (p, q) tensor blow-up $\mathcal{X}^{\{p,q\}}$ to be*

$$\mathcal{X} \otimes \operatorname{Mat}_{p,q} = \left\{ \sum_i X_i \otimes T_i \;\middle|\; X_i \in \mathcal{X}, T_i \in \operatorname{Mat}_{p,q} \right\}$$

viewed as a linear subspace of $\operatorname{Mat}_{kp,nq}$. We will write $\mathcal{X}^{\{d\}} = \mathcal{X}^{\{d,d\}}$.

The following characterization of non-commutative rank in terms of ranks of tensor blow-ups appears in [49].

Lemma 4.1.8 ([49]). *Let $A = X_0 + t_1X_1 + t_2X_2 + \cdots + t_mX_m$ be a linear matrix and let $\mathcal{X} = \mathrm{span}(X_0, X_1, X_2, \ldots, X_m)$. Then:*

$$\mathrm{ncrk}(A) = \max_d \frac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d} = \lim_{d\to\infty} \frac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}.$$

It is not obvious that $\max \dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ is an integer, or even that $\lim\limits_{d\to\infty} \dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ exists. These will be justified in the following sections. But first, we observe that we can define non-commutative ranks for linear subspaces of matrices, and do so.

Definition 4.1.9. *For a linear subspace of matrices $\mathcal{X}$, define*

$$\mathrm{ncrk}(\mathcal{X}) = \max \frac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d} = \lim_{d\to\infty} \frac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}.$$

4.2 Regularity Lemma

Given $X \in \mathcal{X}$ having rank $r = \mathrm{rk}(\mathcal{X})$, observe that $X \otimes I \in \mathcal{X}^{\{d\}}$ has rank rd. Hence $\mathrm{rk}(\mathcal{X}^{\{d\}})$ is at least $d \cdot \mathrm{rk}(\mathcal{X})$.

Example 4.2.1. *Let $\mathcal{X}$ denote the linear subspace of skew symmetric 3×3 matrices. The rank of this subspace is 2. Let X_1, X_2, X_3 be any basis of $\mathcal{X}$. Domokos showed in [21] that $X_1 \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + X_2 \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + X_3 \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{X}^{\{2\}}$ has full rank, i.e., 6, which is larger than $2 \cdot 2 = 4$.*

The above example shows $\mathrm{rk}(\mathcal{X}^{\{d\}})$ could very well be larger than $d \cdot \mathrm{rk}(\mathcal{X})$. However, Ivanyos, Qiao and Subrahmanyam showed that there is a very strong restriction on the possible ranks of tensor blow-ups. They proved the following regularity lemma ([49, Lemma 11 and Remark 10]).

Proposition 4.2.2 (Regularity Lemma). *If $\mathcal{X}$ is a linear subspace of matrices, then $\mathrm{rk}(\mathcal{X}^{\{d\}})$ is a multiple of d.*

In [49], this is proved by giving an algorithm that takes a matrix of rank $\geq rd+1$ in $\mathcal{X}^{\{d\}}$ and produces another matrix in $\mathcal{X}^{\{d\}}$ of rank $\geq (r+1)d$. Analyzing their algorithm (see [49, 50]), they show that it runs in polynomial time. We give another proof of the regularity lemma using Amitsur's universal division algebras. While our proof is less constructive than the original proof, it is conceptually more satisfying.

Let $A = X_0 + t_1X_1 + t_2X_2 + \cdots + t_mX_m$ be an $p \times q$ linear matrix. The $(i,j)^{th}$ entry of A is a linear function in the indeterminates t_k's with coefficients in K. In fact if $c_k \in K$ is

the $(i,j)^{th}$ entry of X_k, then the $(i,j)^{th}$ entry of A is given by $A_{i,j} = c_0 + c_1 t_1 + \cdots + c_m t_m$. Suppose $S_1, \ldots, S_m$ are $d \times d$ matrices, then $X_0 \otimes I + X_1 \otimes S_1 + \cdots + X_m \otimes S_m$ is a $p \times q$ block matrix and the size of each block is $d \times d$. Moreover, the $(i,j)^{th}$ block is $c_0 I + c_1 S_1 + \cdots + c_m S_m$.

In effect $X_0 \otimes I + X_1 \otimes S_1 + \cdots + X_m \otimes S_m$ is simply the block matrix obtained by substituting S_k for t_k in the linear matrix A. Hence, we make the following definition.

Definition 4.2.3. *Let $A = X_0 + t_1 X_1 + \cdots + t_m X_m$ be a linear matrix. For any m-tuple of matrices $S = (S_1, S_2, \ldots, S_m)$, we define*

$$A(S) = X_0 \otimes I + X_1 \otimes S_1 + \cdots + X_m \otimes S_m.$$

Example 4.2.4. *Let* $A = \begin{pmatrix} 0 & 1 & t_1 \\ -1 & 0 & t_2 \\ -t_1 & -t_2 & 0 \end{pmatrix}$. *Then* $A(S_1, S_2) = \begin{pmatrix} 0 & I & S_1 \\ -I & 0 & S_2 \\ -S_1 & -S_2 & 0 \end{pmatrix}$.

4.2.1 The ring of generic matrices

Let $\{t^i_{j,k} | 1 \leq j,k \leq d, i \in \mathbb{Z}_{>0}\}$ be a collection of independent commuting indeterminates. For each $i \in \mathbb{Z}_{>0}$, define the $d \times d$ matrix $T_i = [t^i_{j,k}]$. By a generic matrix, we will refer to a matrix of indeterminates. Let $K[\{t^i_{j,k}\}]$ denote the polynomial ring in the variables $t^i_{j,k}$ for $1 \leq j,k \leq d$, $i \in \mathbb{Z}_{>0}$. Observe that for each i, the generic matrix T_i lies in $\mathrm{Mat}_{d,d}(K[\{t^i_{j,k}\}])$.

Definition 4.2.5. *The ring of generic matrices $R_d \subseteq \mathrm{Mat}_{d,d}(K[\{t^i_{j,k}\}])$ is defined as the subalgebra generated by $\{T_i | i \in \mathbb{Z}_{>0}\}$.*

Lemma 4.2.6. *Let $A = X_0 + t_1 X_1 + \cdots + t_m X_m$ be a linear matrix, and let $\mathcal{X} = \mathrm{span}(X_1, X_2, \ldots, X_m)$. Then, we have*

$$\mathrm{rk}(\mathcal{X}^{\{d\}}) = \mathrm{rk}(X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m),$$

where T_i is a generic matrix for $i = 1, 2, \ldots, m$.

Proof. We first show $\mathrm{rk}(\mathcal{X}^{\{d\}}) \leq \mathrm{rk}(X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m)$. For $S = (S_0, S_1, \ldots, S_m)$ in a non-empty Zariski open subset of $\mathrm{Mat}^{m+1}_{d,d}$, we have $\mathrm{rk}(X_0 \otimes S_0 + X_1 \otimes S_1 + \cdots + X_m \otimes S_m) = \mathrm{rk}(\mathcal{X}^{\{d\}}) = r$, since K is infinite. There is an S in this Zariski open subset for which S_0 is invertible. For such an S, observe that $\mathrm{rk}(X_0 \otimes I + X_1 \otimes S_0^{-1} S_1 + \cdots + X_m \otimes S_0^{-1} S_m) = r$. The corresponding $r \times r$ minor in $X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m$ must be a non-zero polynomial.

The other inequality, i.e., $\mathrm{rk}(\mathcal{X}^{\{d\}}) \geq \mathrm{rk}(X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m)$ is straightforward, and follows the same argument as Lemma 4.1.6. □

Example 4.2.7. *Let* $A = X_0 + t_1X_1 + t_2X_2 = \begin{bmatrix} 0 & 1 & t_1 \\ -1 & 0 & t_2 \\ -t_1 & -t_2 & 0 \end{bmatrix}$. *Then for generic matrices* T_1, T_2*, we have:*

$$A(T_1, T_2) = X_0 \otimes I + X_1 \otimes T_1 + X_2 \otimes T_2 = \begin{bmatrix} 0 & I & T_1 \\ -I & 0 & T_2 \\ -T_1 & -T_2 & 0 \end{bmatrix}.$$

As observed in the introduction, we can do row and column transformations to transform

$$\begin{bmatrix} 0 & I & T_1 \\ -I & 0 & T_2 \\ -T_1 & -T_2 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & [T_2, T_1] \end{bmatrix}.$$

Hence $\mathrm{rk}\, A(T_1, T_2) = 2d + \mathrm{rk}([T_1, T_2])$*. If the* T_i *are generic matrices of size* 1×1*, then* $[T_1, T_2] = 0$*, and if* T_1, T_2 *are generic matrices of size* $d \times d$ *for* $d \geq 2$*, then* $[T_1, T_2]$ *is invertible, and hence of full rank. Thus for* T_1, T_2 *generic matrices of size* $d \times d$*, we have*

$$\mathrm{rk}\, A(T_1, T_2) = \begin{cases} 2 & if\ d = 1, \\ 3d & if\ d \geq 2. \end{cases}$$

In particular, observe that $\mathrm{rk}\, A(T_1, T_2)$ *is always a multiple of* d*. Using Lemma 4.2.6, one sees that the regularity lemma is satisfied for the linear subspace of* 3×3 *skew symmetric matrices.*

4.2.2 Universal division algebras

Observe, as in Example 4.2.7, that for generic $d \times d$ matrices, the expression $[T_1, T_2]$ was either identically zero, or invertible depending upon the value of d. This is a special case of a surprising general phenomenon, namely that any non-zero non-commutative rational expression in some $d \times d$ generic matrices must in fact be invertible! This follows from the fact that Amitsur's universal division algebras are division algebras. We describe these universal division algebras.

Recall the ring of generic matrices $R_d \subseteq \mathrm{Mat}_{d,d}(K[\{t^i_{j,k}\}])$. Let Z_d denote the center of R_d, and let the field of fractions of Z_d be Q_d. The following result is due to Amitsur (see

[2, 3, 4]). One can also find it in standard texts (for example [7, Section 7.2]).

Theorem 4.2.8 (Amitsur)**.** $\mathrm{UD}(d) := Q_d \otimes_{Z_d} R_d$ *is a division algebra and is called a universal division algebra of degree* d*.*

Proof. Posner proved that the central quotient of a prime PI-ring is a simple algebra (see [69]). The ring R_d satisfies a polynomial identity, namely the Amitsur-Levitzki polynomial. Amitsur showed (see [2, Theorem 4]) that R_d is in fact a (non commutative) integral domain, and in particular a prime ring. Hence its central quotient $\mathrm{UD}(d)$ is a simple algebra. By the Wedderburn-Artin theorem (see [52, Section 3.13]), it must be a matrix algebra over a division algebra, i.e., $\mathrm{UD}(d) \cong \mathrm{Mat}_{r,r}(D)$ for some division algebra D and $r \in \mathbb{Z}_{>0}$. Further, since R_d is an integral domain, $\mathrm{UD}(d)$ has no nilpotents. Hence $\mathrm{UD}(d) \cong \mathrm{Mat}_{1,1}(D) \cong D$. □

Note that $\mathrm{UD}(d) \subseteq \mathrm{Mat}_{d,d}(K(\{t^i_{j,k}\}))$. We now give another proof of the regularity lemma, as we mentioned in the introduction.

Proof of Theorem 4.2.2. Let $X_0, X_1, X_2, \ldots, X_m$ span the linear subspace $\mathcal{X} \subseteq \mathrm{Mat}_{p,q}$, and set $A = X_0 + t_1X_1 + \cdots + t_mX_m$. Then by Lemma 4.2.6, we have

$$\mathrm{rk}(\mathcal{X}^{\{d\}}) = \mathrm{rk}(X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m).$$

$A(T_1, T_2, \ldots, T_m) = X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m$ can be viewed as a $p \times q$ block matrix whose blocks are linear expressions in the generic matrices T_i, and in particular elements of $\mathrm{UD}(d)$, a division algebra. By row and column operations in $\mathrm{UD}(d) \subseteq \mathrm{Mat}_{d,d}(K(\{t^i_{j,k}\}))$, we can make the transformation:

$$(X_0 \otimes I + X_1 \otimes T_1 + \cdots + X_m \otimes T_m) \longrightarrow \left[\begin{array}{ccc|c} I & & & \\ & \ddots & & 0 \\ & & I & \\ \hline & 0 & & 0 \\ & & & \end{array}\right]$$

Since each I denotes a $d \times d$ identity matrix, it contributes d to the rank. Hence, it is clear that $\mathrm{rk}(A(T_1, \ldots, T_m)) = \mathrm{rk}(\mathcal{X}^{\{d\}})$ is a multiple of d. □

4.3 Failure of the weakly increasing property in blow-ups

Observe that $\mathrm{ncrk}(\mathcal{X}) \geq \mathrm{rk}(\mathcal{X})$, since $\mathrm{rk}(\mathcal{X}^{\{d\}})$ is at least $d \cdot \mathrm{rk}(\mathcal{X})$. On the other hand, it is shown in [37] using an argument of Flanders that $\mathrm{ncrk}(\mathcal{X}) \leq 2\,\mathrm{rk}(\mathcal{X})$. Modifying their argument slightly, one can show that the ratio must be < 2.

Proposition 4.3.1. *For any linear subspace $\mathcal{X}$, we have* $\mathrm{ncrk}(\mathcal{X}) < 2\,\mathrm{rk}(\mathcal{X})$.

Proof. Let r be the smallest non-negative integer such that we have a linear subspace $\mathcal{X} \subseteq \mathrm{Mat}_{p,q}$ of rank r for some p, q, such that $\mathrm{ncrk}(\mathcal{X}) = 2r$. We have $r > 1$ since $\mathrm{rk}(\mathcal{X}) = 1$ implies $\mathrm{ncrk}(\mathcal{X}) = 1$ (see [37, Remark 1]). This also follows from Lemma 4.4.6 that we will prove in the following section.

We use a result of Flanders (see [34, Lemma 1]) to see that $\mathcal{X}$ is equivalent to a subspace of the form $\left\{\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}\right\}$ with C of size $r \times r$ (see also [37, Corollary 2]). Since $\mathrm{ncrk}(\mathcal{X}) = 2r$, we must have $\mathrm{ncrk}(A) \geq r$, since we must have at least $2r$ linearly independent rows. But A has only r columns, and hence $\mathrm{ncrk}(A) = r$. A similar argument considering columns shows that $\mathrm{ncrk}(B) = r$.

We have $\mathrm{rk}(A), \mathrm{rk}(B) \geq r/2$ because the ratio is at most 2. We cannot have $\mathrm{rk}(A) = r/2$ or $\mathrm{rk}(B) = r/2$ as that would violate the minimality of r. Thus $\mathrm{rk}(A), \mathrm{rk}(B) > r/2$. However, this means that $\mathrm{rk}(\mathcal{X}) \geq \mathrm{rk}(A) + \mathrm{rk}(B) > r$. □

The existence of $\max_d \dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ and $\lim_{d\to\infty} \dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ follows from the following partial increasing property of ranks of blow-ups (see [49, Corollary 12]), along with the regularity lemma.

Lemma 4.3.2. *Let $\mathcal{X}$ be a linear subspace of matrices, Then for $d \geq n$, $\dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ is weakly increasing.*

The authors of [49] comment that the statement of Lemma 4.3.2 is perhaps true for $d < n$ as well, but are unable to prove it. We are able to strengthen the result:

Proposition 4.3.3. *Let $\mathcal{X}$ be a linear subspace of matrices, Then for $d \geq \frac{n}{2} - 1$, $\dfrac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$ is weakly increasing.*

Proof. Proof of Proposition 4.3.3. Suppose $\mathrm{rk}(\mathcal{X}^{\{d\}})/d = r$. Choose a basis $X_1, \ldots, X_m$ of $\mathcal{X}$. There exist $T_1, \ldots, T_m \in \mathrm{Mat}_{d,d}$ such that $\sum_i X_i \otimes T_i \in \mathcal{X}^{\{d\}}$ has rank rd. Choose $a_1, \ldots, a_m \in K$ such that $\sum_i a_i X_i \in \mathcal{X}$ has rank equal to $\mathrm{rk}(\mathcal{X})$.

Then let $\widetilde{T_i} \in \mathrm{Mat}_{d+1,d+1}$ be given by

$$\widetilde{T_i} = \left[\begin{array}{ccc|c} & & & 0 \\ & T_i & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & a_i \end{array}\right].$$

Then it is easy to see that

$$\mathrm{rk}(\textstyle\sum_{i=1}^m X_i \otimes \widetilde{T_i}) \geq \mathrm{rk}(\sum_{i=1}^m X_i \otimes T_i) + \mathrm{rk}(\sum_{i=1}^m a_i X_i) = rd + \mathrm{rk}(\mathcal{X})$$

Furthermore, we have

$$\mathrm{rk}(\mathcal{X}) > \tfrac{1}{2}\,\mathrm{ncrk}(\mathcal{X}) \geq \tfrac{1}{2}r.$$

In the above, the first inequality follows from Proposition 4.3.1, and the second follows from the Definition 4.1.9. Hence, we have

$$\mathrm{rk}(\mathcal{X}^{\{d+1\}}) \geq \mathrm{rk}(\textstyle\sum_i X_i \otimes \widetilde{T_i}) > rd + \frac{1}{2}r.$$

Since $d \geq \frac{n}{2} - 1 \geq \frac{r}{2} - 1$, we have $\mathrm{rk}(\mathcal{X}^{\{d+1\}}) > rd + \frac{1}{2}r \geq (r-1)(d+1)$. Now, by the regularity lemma (Proposition 4.2.2) we must have $\mathrm{rk}(\mathcal{X}^{\{d+1\}})/(d+1) \geq r = \mathrm{rk}(\mathcal{X}^{\{d\}})/d$. □

□

More importantly, we show that the increasing property need not hold for small values of d. We combine a surprising construction of Bergman in [5] of an explicit rational identity satisfied by 3×3 matrices but not by 2×2 matrices, with a construction of Hrubeš and Wigderson in [46] to give a counterexample.

Proposition 4.3.4. *There exists a linear subspace $\mathcal{X}$ such that*

$$\frac{\mathrm{rk}(\mathcal{X}^{\{2\}})}{2} > \frac{\mathrm{rk}(\mathcal{X}^{\{3\}})}{3}.$$

In fact, using an existential result in [5], we can show:

Theorem 4.3.5. *For any $n, m \in \mathbb{Z}_{>0}$ such that $n \nmid m$, there is a linear subspace $\mathcal{X}$ such that*

$$\frac{\mathrm{rk}(\mathcal{X}^{\{n\}})}{n} > \frac{\mathrm{rk}(\mathcal{X}^{\{m\}})}{m}.$$

In order to prove the above claims, we will need to recall some results on rational identities, and a construction from [46].

4.3.1 Rational identities

In [5, 6], Bergman proved a number of remarkable results on rational relations and rational identities in division rings. In particular, he came up with an explicit construction of a rational expression which is an identity on 3×3 matrices, but invertible on general 2×2 matrices. We introduce some notation. Let Y' denote the commutator $[X, Y]$, and let $\delta(Y)$ denote $(Y^2)'[(Y^{-1})']^{-1}$. In [5], Bergman proves the following result (see also [7, Theorem 7.4.3]).

Theorem 4.3.6 ([5]). *Let $n = 2$ or 3. For X, Y generic $n \times n$ matrices, we have:*

$$\psi = \delta(Y')\delta(Y'')[(\delta(Y'')^{-1})'][(\delta(Y''')^{-1})'] = \begin{cases} 1 & if\ n = 3, \\ 0 & if\ n = 2. \end{cases}$$

Note that in the above theorem, 1 denotes the identity matrix, i.e., the identity element in the ring of $n \times n$ matrices, and 0 denotes the zero matrix.

Corollary 4.3.7. *The rational expression $\psi - 1$ is an identity for 3×3 matrices, but is invertible for general choices of 2×2 matrices.*

Bergman also showed the existence of such rational functions more generally. Let $\mathcal{E}(d)$ be the set of rational expressions that can be evaluated on generic $d \times d$ matrices.

Theorem 4.3.8 ([5]). *Assume $n, m \in \mathbb{Z}_{>0}$. Then $\mathcal{E}(n) \subseteq \mathcal{E}(m)$ if and only if $n \mid m$.*

4.3.2 Non-commutative arithmetic circuits with division

A non-commutative arithmetic circuit is a directed acyclic graph, whose vertices are called gates. Gates of in-degree 0 are elements of K or variables t_i. The other allowed gates are inverse, addition and multiplication gates of in-degrees 1, 2 and 2 respectively. The edges going into an multiplication gate are labelled left and right to indicate the order of multiplication. A formula is a circuit, where every node has out-degree at most 1. The number of gates in a circuit is called its size. Let Φ be a circuit in m variables. It is easy to observe that each output gate of a circuit Φ computes a rational expression. We denote by $\widehat{\Phi}(T)$ the evaluation of Φ at $T = (T_1, T_2, \ldots, T_m) \in \mathrm{Mat}_{p,p}^m$. In the process of evaluation, if the input of an inverse gate is not invertible, then $\widehat{\Phi}(T)$ is undefined. Φ is called a correct circuit if $\widehat{\Phi}(T)$ is defined for some T. For further details, we refer to [46].

Given a non-commutative formula of size n, Hrubeš and Wigderson construct a family of linear matrices A_u for each gate u of the formula with certain properties outlined in the

proposition below. We refer to [46, Theorem 2.5] for details. We are content to remark that these matrices can be constructed explicitly in time which is polynomial in n. We recall [46, Propostion 7.1].

Proposition 4.3.9 ([46])**.** *Let R be a ring which contains K in its center. For a formula Φ, and $a_1, a_2, \ldots, a_m \in R$, the following are equivalent:*

1. *$\widehat{\Phi}(a_1, a_2, \ldots, a_m)$ is defined.*
2. *For every gate u, the $A_u(a_1, a_2, \ldots, a_m)$ is invertible.*

Now, we can put Bergman's results together with Hrubeš and Wigderson's results to give a proof of Proposition 4.3.4.

Proof of Proposition 4.3.4. Let Φ be the non-commutative formula that computes the rational expression $(\psi - 1)^{-1}$, where ψ is as in Theorem 4.3.6. By the construction of Hrubeš and Wigderson mentioned above, we have linear matrices A_u for each gate u. Observe that $\widehat{\Phi}(T)$ is defined for $T = (T_1, T_2, \ldots, T_m)$ where the T_i are generic 2×2 matrices by Theorem 4.3.6. Thus, the $A_u(T)$ is invertible for all u.

On the other hand, if the T_i are generic 3×3 matrices, then once again by Theorem 4.3.6, $\widehat{\Phi}(T)$ is not defined. Thus, for some u, A_u is not invertible. For this u, write $A_u = X_0 + t_1 X_1 + \cdots + t_m X_m$ and let $\mathcal{X} = \operatorname{span}(X_0, X_1, \ldots, X_m)$. Then, using Lemma 4.2.6, we conclude

$$\frac{\operatorname{rk}(\mathcal{X}^{\{2\}})}{2} > \frac{\operatorname{rk}(\mathcal{X}^{\{3\}})}{3}.$$

□

For the general case, we use Theorem 4.3.8.

Proof of Theorem 4.3.5. If $n \nmid m$, then there exists $r \in \mathcal{E}(n)$ such that $r \notin \mathcal{E}(m)$. Let Φ be the non-commutative formula that computes r. The argument in the proof of Proposition 4.3.4 applied to Φ gives the required conclusion. □

4.4 Combinatorics of ranks of blow-ups

In this section, we prove an extremely useful stabilization result on the ranks of blow-ups. Recall that

$$\operatorname{ncrk}(\mathcal{X}) = \max \frac{\operatorname{rk}(\mathcal{X}^{\{d\}})}{d} = \lim_{d \to \infty} \frac{\operatorname{rk}(\mathcal{X}^{\{d\}})}{d}.$$

However, to compute the non-commutative rank, one potentially has to go to a very large blow-up. This raises an important problem.

Problem 4.4.1. *Given a linear subspace $\mathcal{X}$ what is the smallest d such that* $\mathrm{ncrk}(\mathcal{X}) = \frac{\mathrm{rk}(\mathcal{X}^{\{d\}})}{d}$*?*

In this section, we will in fact show that it suffices to consider small blow-ups – If $\mathcal{X} \subseteq \mathrm{Mat}_{n,n}$, then blow-ups of size $n-1$ suffice! Let us start with a definition.

Definition 4.4.2. *We define the function $r : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by*

$$r(p,q) = \mathrm{rk}(\mathcal{X}^{\{p,q\}}).$$

Remark 4.4.3. *Note that the set of all $T = (T_1, \ldots, T_m) \in \mathrm{Mat}_{p,q}^m$ for which $\sum_{i=1}^m X_i \otimes T_i$ has maximal rank $r(p,q)$ is Zariski dense in $\mathrm{Mat}_{p,q}^m$.*

Lemma 4.4.4. *The function r has the following properties:*

1. $r(p,q+1) \geq r(p,q)$*;*
2. $r(p+1,q) \geq r(p,q)$*;*
3. $r(p,q+1) \geq \frac{1}{2}(r(p,q) + r(p,q+2))$*;*
4. $r(p+1,q) \geq \frac{1}{2}(r(p,q) + r(p+2,q))$*;*
5. $r(p,q)$ *is divisible by* $\gcd(p,q)$.

Proof.
(1) follows from viewing $\mathcal{X}^{\{p,q\}}$ as a subspace of $\mathcal{X}^{\{p,q+1\}}$.

Now we will prove (3). Let $T = (T_1, \ldots, T_m) \in \mathrm{Mat}_{p,q+2}^m$. For a subset $J \subseteq \{1, 2, \ldots, q+2\}$, let T_i^J be the submatrix where all the columns with index in J are omitted, and let $\mathcal{Y}_J$ be the column span of $\sum_i X_i \otimes T_i^J$. If we choose T general enough, then $\sum_i X_i \otimes T_i^J$ will have rank $r(p, q+2-|J|)$ for all $J \subseteq \{1, 2, \ldots, q+2\}$. We have $\mathcal{Y}_1 + \mathcal{Y}_2 = \mathcal{Y}_\emptyset$ and $\mathcal{Y}_{1,2} \subseteq \mathcal{Y}_1 \cap \mathcal{Y}_2$. It follows that

$$r(p,q) = \dim \mathcal{Y}_{1,2} \leq \dim \mathcal{Y}_1 \cap \mathcal{Y}_2 = \dim \mathcal{Y}_1 + \dim \mathcal{Y}_2 - \dim(\mathcal{Y}_1 + \mathcal{Y}_2) = 2r(p,q+1) - r(p,q+2).$$

Parts (2) and (4) follow from (1) and (3) respectively by symmetry.

To see (5), write $p = dp'$ and $q = dq'$. Then we have $\mathcal{X}^{\{p,q\}} = (\mathcal{X}^{\{p',q'\}})^{\{d\}}$ and the result follows from the regularity lemma $\square$

In the above lemma, parts (1) and (3) give us that $r(p,q)$ is weakly increasing and weakly concave in the second variable, and parts (2) and (4) give the same conclusion for the first variable.

Corollary 4.4.5. *The function $r(p, q)$ is weakly increasing and weakly concave in either variable.*

Lemma 4.4.6. *If $r(1, 1) = 1$, then we have $r(d, d) = d$ for all d.*

Proof. Choose a nonzero matrix $A \in \mathcal{X}$ of rank 1. Using left and right multiplication with matrices in $\mathrm{GL}_n(K)$ we may assume without loss of generality that

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

It is clear that $r(d, d) \geq d$. If $i > 1$, $j > 1$ and $B \in \mathcal{X}$ then $B_{i,j}$ has to be zero, otherwise $tA + B$ will have rank at least 2 for some t. So $\mathcal{X}$ is contained in

$$\begin{bmatrix} * & * & \cdots & * \\ * & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & 0 \end{bmatrix}.$$

Because all matrices of $\mathcal{X}$ have rank at most 1, $\mathcal{B}$ must be contained in the union $W_1 \cup W_2$, where

$$W_1 = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & 0 \end{bmatrix} \text{ and } W_2 \begin{bmatrix} * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Because $\mathcal{X}$ is a subspace, it is entirely contained in W_1 or in W_2. Now it is clear that the matrices in $\mathcal{X}^{\{d\}}$ have at most d nonzero columns, or at most d nonzero rows, so $r(d, d) \leq d$.

□

Proposition 4.4.7. *Let $n \geq 2$, and let $d + 1 \geq n$. If $r(d + 1, d + 1) = n(d + 1)$, then $r(d, d) = nd$ as well.*

Proof. Suppose that $r(d+1, d+1) = n(d+1)$. If $1 \leq a \leq d$, then weak concavity implies that

$$r(d + 1, a) \geq \frac{(d + 1 - a)r(d + 1, 0) + ar(d + 1, d + 1)}{d + 1} = \frac{an(d + 1)}{d + 1} = an.$$

The inequality $r(d+1,a) \leq an$ is clear, so $r(d+1,a) = an$. Similarly, we have $r(a,d+1) = an$. If $r(1,1) = 1$ then we get $r(d+1,d+1) = d+1$ by Lemma 4.4.6 which contradicts $r(d+1,d+1) = n(d+1)$. So we have $r(1,1) \geq 2$. Since $r(p,q)$ is weakly concave in the second variable, we have

$$r(1,d) \geq \frac{(d-1)\cdot r(1,d+1) + 1 \cdot r(1,1)}{d} \geq \frac{(d-1)n+2}{d} = n - \frac{n-2}{d} > n-1,$$

where the last inequality follows as $d \geq n-1$. Since $r(1,d)$ must be an integer, we have $r(1,d) \geq n$. Now, by the weak concavity in the first variable, we have

$$r(d,d) \geq \frac{(d-1)\cdot r(d+1,d) + 1 \cdot r(1,d)}{d} \geq \frac{(d-1)nd+n}{d} = nd - n + \frac{n}{d}.$$

Note that since $d \geq n-1$, we have $d + \frac{n}{d} > n$ or equivalently that $-n + \frac{n}{d} > -d$. Thus, we have

$$r(d,d) \geq nd - n + \frac{n}{d} > d(n-1).$$

Recall that $r(d,d)$ must be a multiple of d by the regularity lemma. Thus $r(d,d) = nd$.

□

From the above proposition and Proposition 4.3.3, we deduce:

Corollary 4.4.8. *We have* $\mathrm{ncrk}(\mathcal{X}) = n$ *if and only if* $r(d,d) = nd$ *for all* $d \geq n-1$.

4.5 Ratio of non-commutative rank to commutative rank

We know that for any linear subspace of matrices $\frac{\mathrm{ncrk}(\mathcal{X})}{\mathrm{crk}(\mathcal{X})} < 2$ by Proposition 4.3.1. On the other hand, Example 4.1.3 is an explicit linear matrix for which the ratio of non-commutative rank to commutative rank is $3/2$. Fortin and Reutenauer suspected that this example was extremal, suggesting that the actual bound for the ratio is $3/2$. We show that this is not the case.

Given an element in $v \in K^n$, we can define a map $L_v : \bigwedge^p K^n \to \bigwedge^{p+1} K^n$ given by $x \mapsto v \wedge x$. This gives a linear map $L : K^n \to \mathrm{Hom}(\bigwedge^p K^n, \bigwedge^{p+1} K^n)$ sending $v \mapsto L_v$. The image is a linear subspace.

Theorem 4.5.1. *Let* $\mathcal{X}(p, 2p+1)$ *denote the image of*

$$L : K^{2p+1} \to \mathrm{Hom}(\bigwedge^{p} K^{2p+1}, \bigwedge^{p+1} K^{2p+1}).$$

We have

$$\frac{\mathrm{ncrk}(\mathcal{X}(p,2p+1))}{\mathrm{rk}(\mathcal{X}(p,2p+1))} = \frac{2p+1}{p+1}.$$

First, let us compute $\mathrm{rk}(\mathcal{X}(p,n))$. Note that a basis for $\bigwedge^p(K^n)$ is given by

$$\{e_{i_1} \wedge e_{i_2} \wedge \ \ldots \ \wedge e_{i_p} | 1 \leq i_1 < i_2 < \cdots < i_p \leq n\}.$$

Let $A(p,n)$ denote the linear matrix given by $t_1 L_{e_1} + t_2 L_{e_2} + \cdots + t_n L_{e_n}$.

Lemma 4.5.2. *For a particular choice of basis, the linear matrix $A(p,n)$ has the form*

$$\left[\begin{array}{c|c} t_n I & A(p-1,n-1) \\ \hline A(p,n-1) & 0 \end{array}\right]$$

Proof. Let

$$A = \{(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_{p-1}}) \wedge e_n \mid 1 \leq i_1 < \cdots < i_{p-1} \leq n-1\}, \text{ and}$$
$$B = \{e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p} \mid 1 \leq i_1 < \cdots < i_p \leq n-1\}.$$

Then clearly $A \cup B$ is a basis for $\bigwedge^p(K^n)$. Similarly, let

$$C = \{(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}) \wedge e_n \mid 1 \leq i_1 < \cdots < i_p \leq n-1\}, \text{ and}$$
$$D = \{e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_{p+1}} \mid 1 \leq i_1 < \cdots < i_{p+1} \leq n-1\}.$$

Then $C \cup D$ is a basis for $\bigwedge^{p+1}(K^n)$. It is easy to see that there $L_{e_n} : B \to C$ is a bijection. Now, order the basis elements for $\bigwedge^p(K^n)$ by taking the basis vectors from B first, and then from A. For $\bigwedge^{p+1}(K^n)$, order the basis vectors by taking the basis vectors from C first, and then from D. Within the basis vectors of C, we order them in the same order as the vectors from B via the aformentioned bijection given by L_{e_1}. □

Remark 4.5.3. *The description in [60, Section 4] is the same as the one above.*

Corollary 4.5.4. *If $A(p,n)$ has full column rank, then so does $A(p-1,n-1)$. Similarly, if $A(p,n)$ has full row rank, then so does $A(p,n-1)$.*

Corollary 4.5.5. *For any non-zero $v \in K^n$, $\mathrm{rk}(L_v) = \binom{n-1}{p}$.*

Proof. Assume without loss of generality that $v = e_n$. By the above choice of basis $L_{e_n} = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right]$. Hence

$$\mathrm{rk}(L_v) = |B| = |C| = \binom{n-1}{p}.$$

□

Let $\mathcal{X}(p,n)$ denote the image of the linear map $K^n \to \mathrm{Hom}(\bigwedge^p(K^n), \bigwedge^{p+1}(K^n))$ given by $v \mapsto L_v$.

Corollary 4.5.6. *We have* $\mathrm{crk}(A(p,n)) = \mathrm{rk}(\mathcal{X}(p,n)) = \binom{n-1}{p}$.

Proof. This follows from Lemma 4.1.6. □

Corollary 4.5.7. *We have* $\mathrm{rk}(\mathcal{X}(p,2p+1)) = \binom{2p}{p}$.

Proposition 4.5.8. *Let* $e_1, \ldots e_{2p+1}$ *denote the standard basis for* K^m. *Let* L_i *denote* L_{e_i}. *For* $1 \leq r \leq 2p+1$, *let* S_r *be the* $(p+1) \times (p+1)$ *matrix such that*

$$S_r(j,k) = \begin{cases} 1 & \text{if } k - j = p + 1 - r \\ 0 & \text{otherwise} \end{cases}$$

Then $L := L_1 \otimes S_1 + L_2 \otimes S_2 + \cdots + L_{2p+1} \otimes S_{2p+1}$ *is invertible.*

The S_i are the most obvious basis of the space of $(p+1) \times (p+1)$ Toeplitz matrices. Since $L \in \mathcal{X}(p, 2p+1)^{2p+1}$ is invertible, we have:

Corollary 4.5.9. *We have* $\mathrm{ncrk}(\mathcal{X}(p,2p+1))$ *is full, i.e.,* $\mathrm{ncrk}(\mathcal{X}(p,2p+1)) = \binom{2p+1}{p}$.

Proof of Theorem 4.5.1. This follows from Corollary 4.5.7 and Corollary 4.5.9. □

Now, it remains to prove Proposition 4.5.8. In characteristic 0, this can be found in [61]. However, the argument does not extend to positive characteristic. Nevertheless, we give a different argument using only elementary linear algebra to extend the result to arbitrary characteristic. The rest of this section is devoted to this.

4.5.1 Preliminaries from Linear Algebra

Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ denote an ordered basis for an n-dimensional vector space V. Consider the alternating power $\bigwedge^r V$. For a subset $I = \{i_1, \ldots, i_r\} \subseteq [n]$ of size r, with $i_1 < i_2 < \cdots < i_r$, we define $v_I = v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_r}$. Here $[n]$ denotes the set $\{1, 2, \ldots, n\}$. The following lemma is a well known fact.

Lemma 4.5.10. *For a given ordered basis $\mathscr{B} = (v_1, \ldots, v_n)$ for K^n, define $\mathscr{B}(r)$ as the set $\{v_I \mid I \subseteq \{1, 2, \ldots, n\},$ with $|I| = r\}$ ordered lexicographically. Then $\mathscr{B}(r)$ is an ordered basis for $\bigwedge^r V$.*

Example 4.5.11. *Let $n = 3$, and $r = 2$, then $\mathscr{B}(r)$ is the ordered basis $(v_{1,2}, v_{1,3}, v_{2,3})$.*

Definition 4.5.12. *Given an ordered basis $\mathscr{B} = (v_1, \ldots, v_n)$ of V and an ordered basis $\mathscr{C} = \{w_1, \ldots, w_m\}$ of W, we define $x_{i,j} = v_i \otimes w_j$. By $\mathscr{B} \otimes \mathscr{C}$, we mean the set $\{x_{i,j} \mid i \in [n], j \in [m]\}$ ordered lexicographically. This is a basis of $V \otimes W$.*

Example 4.5.13. *Let $n = 2, m = 2$, then $\mathscr{B} \otimes \mathscr{C} = (v_1 \otimes w_1, v_1 \otimes w_2, v_2 \otimes w_1, v_2 \otimes w_2) = (x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2})$.*

Suppose that $\mathscr{B}$ is a basis of V and $\mathscr{C}$ is a basis of W and $L : V \to W$ is a linear map. Then $L_{\mathscr{C},\mathscr{B}}$ denotes the matrix of the transformation L with respect to the bases $\mathscr{B}$ and $\mathscr{C}$. If $M : W \to Z$ is a linear map and $\mathscr{D}$ is a basis of Z, then we have $(ML)_{\mathscr{D},\mathscr{B}} = M_{\mathscr{D},\mathscr{C}} L_{\mathscr{C},\mathscr{B}}$.

Let $\mathscr{B} = (b_1, b_2, \ldots, b_n)$ and $\mathscr{B}' = (b'_1, b'_2, \ldots, b'_n)$ be two ordered bases for V. Then denote by $X_{\mathscr{B},\mathscr{B}'} = (\mathrm{id}_V)_{\mathscr{B},\mathscr{B}'}$ be the matrix of the identity with respect to $\mathscr{B}$ and $\mathscr{B}'$. This is the base change matrix and its colums are the vectors $b'_1, b'_2, \ldots, b'_n$ expressed in the basis $\mathscr{B}$. Note that $X_{\mathscr{B}',\mathscr{B}} = X^{-1}_{\mathscr{B},\mathscr{B}'}$. We recall the base change formula for linear transformations.

Lemma 4.5.14 (Base change formula)**.** *We have*

$$L_{\mathscr{C}',\mathscr{B}'} = X_{\mathscr{C}',\mathscr{C}} L_{\mathscr{C},\mathscr{B}} X_{\mathscr{B},\mathscr{B}'} = X^{-1}_{\mathscr{C},\mathscr{C}'} L_{\mathscr{C},\mathscr{B}} X_{\mathscr{B},\mathscr{B}'}.$$

Let $\mathscr{B} = (b_1, b_2, \ldots, b_n)$ be an ordered basis of V and we multiply the i^{th} basis vector by some scalar $\lambda \neq 0$ to obtain the basis $\mathscr{B}' = (b_1, \ldots, b_{i-1}, \lambda b_i, b_{i+1}, \ldots, b_n)$. Then $X_{\mathscr{B},\mathscr{B}'}$ is a diagonal matrix. The i^{th} diagonal entry of $X_{\mathscr{B},\mathscr{B}'}$ is λ and all other diagonal entries are 1. In particular, we have $\det(X_{\mathscr{B},\mathscr{B}'}) = \lambda$. For our purposes we need to understand a more interesting base change matrix.

Proposition 4.5.15. *With $\mathscr{B}$ and $\mathscr{B}'$ as above, we have we have $\det(X_{\mathscr{B}(r),\mathscr{B}'(r)}) = \lambda^{\binom{n-1}{r-1}}$.*

Proof. It is easy to see that the basis $\mathscr{B}'(r)$ is gotten from $\mathscr{B}(r)$ by scaling some of its basis vectors. More precisely, if a subset I contains i, then the basis vector b_I is scaled by λ. All other basis vectors remain unchanged. The number of subsets containing i is given by $\binom{n-1}{r-1}$. Hence $X_{\mathscr{B}(r),\mathscr{B}'(r)}$ is a diagonal matrix in which $\binom{n-1}{r-1}$ diagonal entries are λ and all other diagonal entries are 1. The proposition follows since the determinant of a diagonal matrix is the product of the diagonal entries. □

We also need to understand what happens to a linear map $L \in \mathrm{Hom}(\bigwedge^r V, \bigwedge^{r+1} V)$ when we change basis. For a basis $\mathscr{B}$ of V, let $L_{\mathscr{B}} = L_{\mathscr{B}(r+1),\mathscr{B}(r)}$ denote the matrix of L in the basis $\mathscr{B}(r)$ and $\mathscr{B}(r+1)$ for the domain and codomain respectively.

Corollary 4.5.16. *Let $\mathscr{B}$ and $\mathscr{B}'$ be as in Proposition 4.5.15. Then for $L \in \mathrm{Hom}(\bigwedge^r V, \bigwedge^{r+1} V)$, we have* $\det(L_{\mathscr{B}'}) = \lambda^{\binom{n-1}{r-1}-\binom{n-1}{r}} \det(L_{\mathscr{B}})$.

Proof. This follows from applying Proposition 4.5.15 to the base change formula

$$L_{\mathscr{B}'} = X^{-1}_{\mathscr{B}(r+1),\mathscr{B}'(r+1)} L_{\mathscr{B}} X_{\mathscr{B}(r),\mathscr{B}'(r)}.$$

□

In fact, we need slightly more general results. An argument along the lines of the proof of Proposition 4.5.15 gives the following lemma.

Lemma 4.5.17. *Let $\mathscr{B}$ and $\mathscr{B}'$ be as in Proposition 4.5.15. Let W be a c-dimensonal vector space with ordered basis $\mathscr{C}$. Then we have* $\det(X_{\mathscr{B}(r)\otimes\mathscr{C},\mathscr{B}'(r)\otimes\mathscr{C}}) = \lambda^{c\binom{n-1}{r-1}}$.

For a linear transformation $L \in \mathrm{Hom}((\bigwedge^r V) \otimes W, (\bigwedge^{r+1} V) \otimes W)$, let $L_{\mathscr{B}\otimes\mathscr{C}}$ denote the matrix for the linear transformation of L in the bases $\mathscr{B}(r) \otimes \mathscr{C}$ and $\mathscr{B}(r+1) \otimes \mathscr{C}$ for the domain and codomain respectively. Following the same idea as Corollary 5.3.9, we get the following:

Corollary 4.5.18. *Let $\mathscr{B}$ and $\mathscr{B}'$ be as in Propositon 4.5.15. Then for a linear transformation $L \in \mathrm{Hom}((\bigwedge^r V) \otimes W, (\bigwedge^{r+1} V) \otimes W)$, we have*

$$\det(L_{\mathscr{B}'\otimes\mathscr{C}}) = \lambda^{c(\binom{n-1}{r-1}-\binom{n-1}{r})} \det(L_{\mathscr{B}\otimes\mathscr{C}}).$$

4.5.2 Effects of scaling basis vectors on the matrices of L_i's

Let $m = 2p+1$ be a positive integer. Let $\mathscr{E} = (e_1, \ldots, e_m)$ denote the standard ordered basis of K^m. Recall that for a $v \in K^m$, $L_v \in \mathrm{Hom}(\bigwedge^p K^m, \bigwedge^{p+1} K^m)$ is the linear map that sends w to $v \wedge w$. Let $\mathscr{E}'$ be the ordered basis obtained from $\mathscr{E}$ by scaling the i^{th} basis vector by λ, i.e., $\mathscr{E}' = (e_1, \ldots, e_{i-1}, \lambda e_i, e_{i+1} \ldots, e_m)$. It is easy to understand the effect of this base change on the matrices of L_i.

Lemma 4.5.19. *We have* $(L_j)_{\mathscr{E}'} = \begin{cases} (L_j)_{\mathscr{E}} & \text{if } j \neq i, \\ \lambda^{-1}(L_i)_{\mathscr{E}} & \text{if } j = i. \end{cases}$

Proof. It is easy to see that for any basis $\mathscr{B} = (b_1, \ldots, b_m)$ of K^m, the matrix of L_{b_i} written in the basis $\mathscr{B}(r)$ and $\mathscr{B}(r+1)$ is the same, i.e., $(L_{b_i})_{\mathscr{B}} = (L_{c_i})_{\mathscr{C}}$ for any other basis $\mathscr{C} = (c_1, \ldots, c_m)$. For $j \neq i$, we have $e_j = e'_j$, and hence

$$(L_j)_{\mathscr{E}'} := (L_{e_j})_{\mathscr{E}'} = (L_{e'_j})_{\mathscr{E}'} = (L_{e_j})_{\mathscr{E}} =: (L_j)_{\mathscr{E}}.$$

For $j = i$, we have $e_i = \lambda^{-1} e'_i$, and so

$$(L_i)_{\mathscr{E}'} := (L_{e_i})_{\mathscr{E}'} = (L_{\lambda^{-1}e'_i})_{\mathscr{E}'} = \lambda^{-1}(L_{e'_i})_{\mathscr{E}'} = \lambda^{-1}(L_{e_i})_{\mathscr{E}} =: \lambda^{-1}(L_i)_{\mathscr{E}}.$$

□

Let

$$L = L_1 \otimes S_1 + L_2 \otimes S_2 + \cdots + L_{2p+1} \otimes S_{2p+1} \in \operatorname{Hom}\left((\textstyle\bigwedge^p K^m) \otimes K^{p+1}, (\bigwedge^{p+1} K^m) \otimes K^{p+1}\right),$$

where S_i is defined as in Proposition 4.5.8. Let $\mathscr{F}$ denote the standard basis of K^{p+1}. Hence we have the bases $\mathscr{E}(p) \otimes \mathscr{F}$ and $\mathscr{E}'(p) \otimes \mathscr{F}$ for the domain and the bases $\mathscr{E}(p+1) \otimes \mathscr{F}$ and $\mathscr{E}'(p+1) \otimes \mathscr{F}$ for the codomain. Recall that for a linear transformation $L \in \operatorname{Hom}\left((\bigwedge^r V) \otimes W, (\bigwedge^{r+1} V) \otimes W\right)$, $L_{\mathscr{B}\otimes\mathscr{C}}$ denotes the matrix for the linear transformation of L in the bases $\mathscr{B}(r) \otimes \mathscr{C}$ and $\mathscr{B}(r+1) \otimes \mathscr{C}$ for the domain and codomain respectively, where $\mathscr{B}$ is a basis for V and $\mathscr{C}$ is a basis for W.

Lemma 4.5.20. *We have* $\det(L_{\mathscr{E}'\otimes\mathscr{C}}) = \lambda^{-\binom{2p}{p}} \det(L_{\mathscr{E}\otimes\mathscr{C}})$

Proof. This follows from Corollary 4.5.18, since $(p+1)(\binom{2p}{p-1} - \binom{2p}{p}) = -\binom{2p}{p}$. □

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in K^m$ such that $\lambda_i \neq 0$ for $1 \leq i \leq m$. Given an ordered basis $\mathscr{E} = (e_1, \ldots, e_m)$, we define another ordered basis $\lambda \cdot \mathscr{E} = (\lambda_1 e_1, \lambda_2 e_2, \ldots, \lambda_m e_m)$. Applying the above lemma several times, we get:

Corollary 4.5.21. *We have* $\det(L_{(\lambda\cdot\mathscr{E})\otimes\mathscr{C}}) = \left(\prod_{i=1}^m \lambda_i\right)^{-\binom{2p}{p}} \det(L_{\mathscr{E}\otimes\mathscr{C}})$.

Definition 4.5.22. *Let* M_i *denote the matrix* $(L_i)_{\mathscr{E}}$. *We define*

$$M(t_1, \ldots, t_{2p+1}) := t_1 M_1 \otimes S_1 + t_2 M_2 \otimes S_2 + \cdots + t_{2p+1} M_{2p+1} \otimes S_{2p+1}.$$

Define $p(t_1, \ldots, t_{2p+1}) := \det(M(t_1, \ldots, t_{2p+1}))$.

Corollary 4.5.23. *We have* $p(t_1, \ldots, t_m) = \left(\prod_{i=1}^m t_i\right)^{\binom{2p}{p}} p(1, 1, \ldots, 1)$.

Proof. Apply Lemma 4.5.19 to Corollary 4.5.21, where $\lambda = (t_1^{-1}, t_2^{-1}, \ldots, t_m^{-1})$. □

4.5.3 Examples

Let us first recall that for an $m \times n$ matrix $A = (a_{i,j})$ and a $B = (b_{k,l})$, we define the Kronecker product $A \otimes B$ by

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \dots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \dots & a_{m,n}B \end{pmatrix}$$

If $A = (a_{i,j})$ is a square $n \times n$ matrix, then its determinant is equal to $\sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) r_\sigma$, where σ runs over all elements of the symmetric group Σ_n, $\operatorname{sgn}(\sigma)$ is the sign of the permutation σ and $r_\sigma = \prod_{i=1}^n a_{i,\sigma(i)}$. To proceed further, we believe it is necessary to acquaint the reader with small examples.

Example 4.5.24 ($p = 1$)**.** *Suppose that $p = 1$ and $m = 3$. Let $\mathscr{E} = (e_1, e_2, e_3)$ be the standard basis of K^3. Then the basis $\mathscr{E}(1)$ is $\mathscr{E}$ itself, and the basis $\mathscr{E}(2) = (e_{1,2}, e_{1,3}, e_{2,3})$. In this basis $t_1 L_1 \otimes S_1 + t_2 L_2 \otimes S_2 + t_3 L_3 \otimes S_3$ is given by the block matrix*

$$A := \begin{pmatrix} -t_2 S_2 & t_1 S_1 & 0 \\ -t_3 S_3 & 0 & t_1 S_1 \\ 0 & -t_3 S_3 & t_2 S_2 \end{pmatrix}$$

In other words $A = M(t_1, t_2, t_3)$. We also write out S_i. We have

$$S_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, S_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Observe that the matrix A is a 6×6 matrix with entries in $\mathbb{Z}[t_1, t_2, t_3]$. We will try to compute $\det A$ as an element of this ring. We will analyze the situation thoroughly as it will be useful in handling the general case. We know $\det A = k(t_1 t_2 t_3)^2$ by Corollary 4.5.23 and we want to establish that $k = \pm 1$.

Recall that $\det A = \sum_{\sigma \in \Sigma_6} \operatorname{sgn}(\sigma) r_\sigma$, with $r_\sigma = \prod_{i=1}^6 a_{i,\sigma(i)}$. Now, observe that each entry of A is either 0 or $\pm t_i$. Hence each r_σ is either 0 or $\pm$ monomial (in the t_i's). We know that the final answer must be a multiple of the monomial $(t_1 t_2 t_3)^2$. So, it suffices to focus on the permutations σ such that $r_\sigma = \pm t_1^2 t_2^2 t_3^2$.

We claim that there is at most one permutation σ such that $r_\sigma = \pm t_1^2 t_2^2 t_3^2$. In other words, there is at most one choice of 6 entries, satisfying the condition that no two entries are in the same row and no two entries are in the same column such that the product of their entries is $\pm t_1^2 t_2^2 t_3^2$.

To see this, observe first there are only two entries of the form $\pm t_1$, since $t_1 S_1 = \begin{pmatrix} 0 & t_1 \\ 0 & 0 \end{pmatrix}$ and there are exactly two blocks which are $\pm t_1 S_1$. So, in order to get t_1^2, we have no choice but to pick both entries.

Now, there are four entries of the form $\pm t_2$, two in each block of the form $\pm t_2 S_2$. Consider the northwest $-t_2 S_2$ block. This block occurs in the same block row as a $t_1 S_1$. We focus on these two blocks in the top block row.

$$\left(-t_2 S_2 \quad | \quad t_1 S_1\right) = \left(\begin{array}{cc|cc} -t_2 & 0 & 0 & t_1 \\ 0 & -t_2 & 0 & 0 \end{array}\right)$$

We have already argued that we must pick the blue t_1 in the $t_1 S_1$, since all $\pm t_1$'s must be picked. Hence we cannot pick any other entry from that row. This rules out the $-t_2$ that we have colored red. So only the $-t_2$ from the bottom row is available, which we have colored blue. A similar argument shows that you can only pick the t_2 in the left column of the southeast most block of the form $t_2 S_2$. Since there are only two $\pm t_2$'s available, we have no choice but to pick both of them.

Remark 4.5.25. *We want to think of this in the following way. While considering the northwest block entry $-t_2 S_2$, we observe that there is exactly 1 block entry of the form $\pm t_i S_i$ in the same row with $i < 2$. This is the condition that rules out the top 1 rows. Similarly, there are 0 block entries of the form $\pm t_i S_i$ in the same column with $i < 2$. This is the condition that rules out the right 0 columns. This leaves precisely one non-zero entry in the northwest $t_2 S_2$ to choose from. A generalization of such an argument (see Proposition 4.5.36) will be the key to unlocking the general case.*

Continuing with the example, observe that there are only two $\pm t_3$'s, and hence we must pick both of them. These $\pm t_3$'s could potentially be in the same row or column as the choices of t_1's and t_2's, which would be disastrous. However, this doesn't happen. In this case, one can check explicitly. In the general case, however, instead of an explicit check we will use the generalization of the argument mentioned in the above remark. Hence, there is exactly one permutation σ for which $r_\sigma = \pm(t_1 t_2 t_3)^2$. Thus we have that $\det A = \pm(t_1 t_2 t_3)^2$.

4.5.4 The general case

We will prove Proposition 4.5.8 in this section. Let $m = 2p + 1$ be a positive integer, and let $A := M(t_1, \ldots, t_m)$. We will begin with some structural results on the matrix A. Let $M = t_1 M_1 + \cdots + t_{2p+1} M_{2p+1}$.

Example 4.5.26. *For $p = 1$, we have*

$$M = \begin{pmatrix} -t_2 & t_1 & 0 \\ -t_3 & 0 & t_1 \\ 0 & -t_3 & t_2 \end{pmatrix},$$

and

$$A = \begin{pmatrix} -t_2 S_2 & t_1 S_1 & 0 \\ -t_3 S_3 & 0 & t_1 S_1 \\ 0 & -t_3 S_3 & t_2 S_2 \end{pmatrix}.$$

Lemma 4.5.27. *The matrix M is a $\binom{2p+1}{p} \times \binom{2p+1}{p}$ matrix, whose block entries are either 0 or $\pm t_i$.*

Proof. The positions of the nonzero entries of M_i's are clearly distinct. □

Lemma 4.5.28. *For each $i \in [2p+1]$, there are $\binom{2p}{p}$ entries of the form $\pm t_i$ in M, and all other entries are 0.*

Proof. There are $\binom{2p}{p}$ subsets I of size p that do not contain i. For each such subset I, we have $L_i(e_I) = \pm e_{I \cup i}$. The corresponding entry in the matrix is ± 1, and all other entries are 0. Thus $t_i M_i$ is a matrix with $\binom{2p}{p}$ entries of the form $\pm t_i$, and all other entries 0. Since the positions of the nonzero entries of the $t_i M_i$ are distinct from the positions of nonzero entries of $t_j M_j$ for $i \neq j$, we have the required conclusion. □

Lemma 4.5.29. *Fix an entry $\pm t_i$ in M. Then for each $j \neq i$, then the number of entries of the form $\pm t_j$ in the same row or column is exactly 1.*

Proof. The fixed entry $\pm t_i$ in M corresponds to the fact that $L_i(e_I) = \pm e_{I \cup \{i\}}$ for some I that does not contain i. Now, if $j \in I$, then let $J = I \cup \{i\} \setminus \{j\}$. Then we have $L_j(e_J) = \pm e_{J \cup \{j\}} = \pm e_{I \cup \{i\}}$. This corresponds to a $\pm t_j$ in the same row. On the other hand if $j \notin I$, then $L_j(e_I) = \pm e_{I \cup \{j\}}$ which corresponds to a $\pm t_j$ in the same column. □

Remark 4.5.30. *It follows from the definition of the tensor product of matrices that by replacing each t_i in M with the block matrix $t_i S_i$, we get the block matrix A. See Example 4.5.26.*

The above remark applied to the above lemmas yield:

Corollary 4.5.31. *The matrix A is a $\binom{2p+1}{p}$-block matrix, whose block entries are either 0 or $\pm t_i S_i$.*

Corollary 4.5.32. *For each $i \in [2p+1]$, there are $\binom{2p}{p}$ block entries of the form $\pm t_i S_i$ in A, and all other block entries are 0.*

Corollary 4.5.33. *Fix a block entry $\pm t_i S_i$ in A. Then for each $j \neq i$, the number of block entries of the form $\pm t_j S_j$ in the same block row or same block column is exactly 1.*

Definition 4.5.34. *Let $P = \pm t_i S_i$ be a block entry of A. Suppose there are x entries of the form $\pm t_j S_j$ with $j < i$ in the same block row and y entries of the form $\pm t_j S_j$ in the same block column. Then we call the $(x+1, p-y)^{th}$ entry of P, the* elusive entry *of P.*

Lemma 4.5.35. *The elusive entry of any block $P = \pm t_i S_i$ is a $\pm t_i$. Further, with x and y as defined in the previous definition, all other nonzero entries of P are in the top x rows or the right y columns.*

Proof. The equality $x + y = i - 1$ follows from Corollary 4.5.33. Indeed, we have $S_i(x+1, p-y) = 1$ as $p - y = x + 1 - i + p + 1$ follows from $x + y = i - 1$. Thus there is a t_i in position $(x+1, p-y)$ in the block P. The second statement is obvious since the only nonzero entries are along the diagonal containing $(x+1, p-y)$. □

Let us recall that a permutation $\sigma \in \Sigma_n$ is a choice of n entries subject to the condition that there are no two entries in the same row and no two entries in the same column. In order for $r_\sigma = \pm(t_1 t_2 \dots t_{2p+1})^{\binom{2p}{p}}$, we must make such a choice, where each entry chosen is of the form $\pm t_i$ and for each i, there are $\binom{2p}{p}$ entries chosen of the form $\pm t_i$.

Proposition 4.5.36. *In order for $r_\sigma = \pm(t_1 t_2 \dots t_{2p+1})^{\binom{2p}{p}}$, we must choose the elusive entry from each nonzero block entry.*

Proof. Let $P = \pm t_i S_i$ be a nonzero block entry of A. We proceed by induction on i.

- **Base Case:** $i = 1$**.**

  In this case, observe that there is exactly one nonzero entry, which is $\pm t_1$, and that is precisely the elusive entry. There are $\binom{2p}{p}$ such block entries. In order for the power of t_1 in r_σ to be $\binom{2p}{p}$, we have no choice but to choose the elusive entries from each block entry of the form $\pm t_1 S_1$.

- **Induction Step:**

  Suppose the claim is true for all $j < i$. Let the block entries in the same row of the form $\pm t_k S_k$ with $k < i$ be $Q_1 = \pm t_{j_1} S_{j_1}, Q_2 = \pm t_2 S_{j_2}, \dots, Q_x = \pm t_{j_x} S_{j_x}$ with $1 \leq j_1 < j_2 < \dots < j_x < i$. Then clearly the block entry Q_k satisfies the hypothesis

of the claim for $k-1$. Hence, by induction we would have picked the $\pm t_{j_k}$ from the k^{th} row. Hence, we cannot pick the $\pm t_i$'s in the first x rows of P.

By a similar argument, we cannot pick the t_i's in the right y columns, where y is the number of the block entries of the form $\pm t_k S_k$ with $k < i$ in the same column. This leaves precisely one non-zero entry in P, which is the elusive entry. Now, once again we have precisely $\binom{2p}{p}$ blocks of the form $\pm t_i S_i$, and we can pick at most one $\pm t_i$ from each one. Since we want the power of t_i in r_σ to be $\binom{2p}{p}$, we have no choice but to pick all of them.

□

Corollary 4.5.37. *There is at most one permutation σ such that $r_\sigma = \pm(t_1 t_2 \dots t_{2p+1})^{\binom{2p}{p}}$.*

Proof of Proposition 4.5.8. We know that

$$p(t_1, \dots, t_m) = \det(M(t_1, \dots, t_m)) = k(t_1 t_2 \dots t_{2p+1})^{\binom{2p}{p}},$$

where $k = p(1, \dots, 1) \in K$ by Corollary 4.5.23. We also know that each r_σ is $\pm$ monomial. Further, by the above Proposition, there is exactly one r_σ which gives us $\pm(t_1 t_2 \dots t_{2p+1})^{\binom{2p}{p}}$, and hence we must have $k = \pm 1 \neq 0$. But $k = p(1, \dots, 1)$, and hence L is invertible, since $p(1, \dots, 1) = \det M(1, \dots, 1)$ and $M(1, \dots, 1)$ is the matrix for L in some coordinates.

□

4.5.5 More examples

The linear subspaces in Theorem 4.5.1 provide a family of examples for which the ratio approaches 2. In fact, these linear subspaces give rise to more examples which have a discrepancy between the commutative and non-commutative rank.

Corollary 4.5.38. *Let $\mathcal{X}(i,n)$ denote the image of $L : K^n \to (\bigwedge^i K^n, \bigwedge^{i+1} K^n)$. Then*

1. $\operatorname{ncrk}(\mathcal{X}(i,n))$ *is full;*
2. *If $i \neq 0, n-1$, then $\operatorname{rk}(\mathcal{X}(i,n))$ is not full.*

Proof of Corollary 4.5.38. To prove (1), consider $A(i,n)$. If $i < n/2$, then let $k = n - 2i - 1$. The linear matrix $A(i+k, n+k)$ has full column rank by Proposition 4.5.8, since $2(i+k)+1 = n+k$. By repeated application of Corollary 4.5.4, we conclude that $A(i,n)$

has full column rank. Since $i < n/2$, the matrix $A(i, n)$ has more rows than columns, and hence has full non-commutative rank.

If $i \geq n/2$, then we observe that $A(i, 2i + 1)$ has full non-commutative rank. Once again by repeated application of Corollary 4.5.4, we conclude that $A(i, n)$ has full row rank. Since $A(i, n)$ has more columns than rows, it has full non-commutative rank. Finally, observe that the linear subspace defined by $A(i, n)$ is the linear subspace $\mathcal{X}(i, n)$.

To prove (2), use Corollary 4.5.6. $\square$

CHAPTER 5

Degree bounds for matrix invariants and matrix semi-invariants

We prove polynomial bounds for generators for the ring of matrix semi-invariants in Section 5.1. In Section 5.2, we deduce bounds for matrix invariants and invariants of quivers. We analyze the case of the ring of semi-invariants for a quiver in Section 5.3. In Section 5.4, we show an invariant of degree n^2 in $R(n,m)$ ($m \geq n^2$) that cannot be generated by invariants of smaller degree. We treat the case of $R(3,m)$ in more detail in Section 5.5, and in Section 5.6, we compute the Hilbert series in several cases.

For this chapter, we will assume K to be an algebraically closed field. Most of the statements remain true for an infinite field, but the proofs are cleaner for algebraically closed fields.

5.1 Degree bounds for matrix semi-invariants

5.1.1 Null cone

Recall that the method of Derksen and Popov relied on giving degree bounds for invariant rings by giving bounds for invariants defining the null cone. Let us recall the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$ on $\mathrm{Mat}_{n,n}^m$, i.e., for $(A,B) \in \mathrm{SL}_n \times \mathrm{SL}_n$ and $X = (X_1, \ldots, X_m) \in \mathrm{Mat}_{n,n}^m$, we have

$$(A,B) \cdot (X_1, \ldots, X_m) = (AX_1B^{-1}, \ldots, AX_mB^{-1}).$$

The ring $R(n,m) = K[\mathrm{Mat}_{n,n}^m]^{\mathrm{SL}_n \times \mathrm{SL}_n}$ is the ring of matrix semi-invariants. Recall that for each $T \in \mathrm{Mat}_{d,d}^m$, we defined the homogeneous invariant f_T of degree dn by $f_T(X) = \det(\sum_i X_i \otimes T_i)$, and that these invariants spanned $R(n,m)$.

We will denote the null cone for the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$ on $\mathrm{Mat}_{n,n}^m$ by $\mathcal{N}(n,m)$ in this section for notational convenience.

Proposition 5.1.1. *Let $X = (X_1, \ldots, X_m) \in \mathrm{Mat}_{n,n}^m$. Denote by $\mathcal{X} = \mathrm{span}(X_1, \ldots, X_m)$ be the linear subspace spanned by $X_1, \ldots, X_m$, and let $r(p,q) = \mathrm{rk}(\mathcal{X}^{\{p,q\}})$. Then the following are equivalent.*

1. *$X \notin \mathcal{N}(n,m)$;*
2. *$\exists$ non-constant homogeneous invariant $f \in R(n,m)$ such that $f(X) \neq 0$;*
3. *$\exists d such that \exists\, T \in \mathrm{Mat}_{d,d}^m$ such that $f_T(X) \neq 0$;*
4. *$r(d,d) = nd$ for some d;*
5. *$\mathrm{ncrk}(\mathcal{X}) = n$;*
6. *$r(d,d) = nd$ for any $d \geq n-1$;*
7. *$\forall d \geq n-1 \exists\, T \in \mathrm{Mat}_{d,d}^m$ such that $f_T(X) \neq 0$.*

Proof. The equivalence of (1) and (2) follows from the definition of null cone. Since invariants of the form f_T are homogeneous and span the invariant ring $R(n,m)$, we have $(2) \Leftrightarrow (3)$. The equivalence of (3) and (4) follows from the definition of tensor blow up. $(4) \Leftrightarrow (5)$ follows from the characterization of non-commutative rank in terms of ranks of blow-ups. The equivalence $(5) \Leftrightarrow (6)$ is the most crucial equivalence, and is the statement of Corollary 4.4.8. Finally $(6) \Leftrightarrow (7)$ follows from the definition of tensor blow-ups. □

Corollary 5.1.2. *Let $r = \dim R(n,m)$. Then there exists $f_1, \ldots, f_r \in R(n,m)_{n(n-1)}$ which forms a hsop.*

Proof. From the above Proposition, we know that the degree $n(n-1)$ invariants cut out the null cone. Hence, by Noether normalization lemma, there exist $f_1, \ldots, f_r$ of degree $n(n-1)$ form a hsop. □

5.1.2 Degree bounds in characteristic 0

We can now get degree bounds by using Hochster-Roberts theorem and Kempf's results directly.

Theorem 5.1.3. *If K has characteristic 0, then we have*

$$\beta(R(n,m)) \leq mn^4.$$

Proof. Note that we have $r = \dim R(n,m) \leq \dim \mathrm{Mat}_{n,n}^m = mn^2$. For $n \geq 2$, we use the hsop from Corollary 5.5.3 in Proposition 2.2.3 to get

$$\beta(R(n,m)) \leq r(n^2-n) \leq mn^2(n^2-n) < mn^4.$$

It is clear that $\beta(R(1,m)) = 1$, so we have $\beta(R(n,m)) \leq mn^4$ for all n and m. □

Corollary 5.1.4. *If K has characteristic 0, then we have $\beta(R(n,m)) \leq n^6$.*

Proof. This follows from Weyl's polarization theorem, i.e., Theorem 2.3.1. □

Remark 5.1.5. *The above Theorem and Corollary continue to be true as long as the field K is infinite. This is because we have $R_K(n,m) \otimes \overline{K} = R_{\overline{K}}(n,m)$ where $\overline{K}$ denotes the algebraic closure of K.*

5.1.3 Adaptation to positive characteristic: Good filtrations

Notice that the result on the null cone is independent of characteristic. To adapt the method to positive characteristic, we need the theory of good filtrations. The theory of good filtrations is very powerful in positive characteristic. A comprehensive introduction to this theory can be found in [26] (see also [27, 28, 30, 42, 63]). We also refer the reader to [23, 75] for an exposition with a view of using them for invariant rings coming from quivers.

Let G be a reductive group. Fix a maximal torus T and fix a Borel subgroup B containing the torus. Let Λ denote the set of dominant weights. Given $\lambda \in \Lambda$, we have $\lambda : T \to K^*$, and we can extend it to a map $\lambda : B \to K^*$, by composing with the natural surjection $B \twoheadrightarrow T$.

Definition 5.1.6. *For $\lambda \in \Lambda$, the dual Weyl module $\nabla(\lambda)$ is defined as*

$$\nabla(\lambda) := \{f \in K[G] \mid f(bg) = \lambda(b)f(g) \ \forall \ (b,g) \in B \times G\}.$$

When $G = GL_n$, the dual Weyl modules coincide with Schur modules.

Definition 5.1.7. *A G-module V is called a good G-module if V has a filtration of the form $0 \subseteq V_0 \subseteq V_1 \subseteq \ \ldots$ such that $\bigcup_i V_i = V$ and each quotient V_i/V_{i-1} is a dual Weyl module. Such a filtration is called a good filtration.*

It is well known that the coordinate ring of the representation space of a quiver for a fixed dimension vector has a good filtration. We refer to [74] and [25] for details.

Proposition 5.1.8. *$K[\mathrm{Rep}(Q,\alpha)]$ is a good $\mathrm{GL}(\alpha)$-module and hence a good $\mathrm{SL}(\alpha)$-module.*

We recall the main result from [42], which will be crucial to our purposes.

Theorem 5.1.9 (Hashimoto)**.** *Assume $\mathrm{char}\, K > 0$. Let V be a rational representation of a connected reductive group G, and assume its coordinate ring $K[V]$ is a good G-module. Then $K[V]^G$ is strongly F-regular and hence Cohen-Macaulay.*

Corollary 5.1.10. *The invariant rings $\mathrm{SI}(Q,\alpha), \mathrm{I}(Q,\alpha), S(n,m)$ and $R(n,m)$ are Cohen-Macaulay.*

It is a well known fact that if a G-module V has a good filtration, as a consequence, the Hilbert series of the invariant ring is independent of the underlying field. In [26, page 399], Donkin proves this, albeit for a special case. However, it is easy to see that the argument holds in much greater generality.

Proposition 5.1.11. *The Hilbert series $H(\mathrm{SI}(Q,\alpha),t), H(\mathrm{I}(Q,\alpha),t), H(R(n,m),t)$ and $H(S(n,m),t)$ are independent of the underlying field K.*

We can replace the Hochster-Roberts theorem with Corollary 5.1.10 and Kempf's results with Corollary 5.1.11 to get the bounds we require in positive characteristic.

Corollary 5.1.12. *Let K be an infinite field of any characteristic. Then we have*

$$\beta(R(n,m)) \leq mn^4.$$

5.2 Matrix invariants

We consider the map

$$\begin{aligned} \phi : \mathrm{Mat}_{n,n}^m &\longrightarrow \mathrm{Mat}_{n,n}^{m+1} \\ (X_1,\ldots,X_m) &\longmapsto (I, X_1,\ldots,X_m) \end{aligned}$$

This gives a surjection on the coordinate rings $\phi^* : K[\mathrm{Mat}_{n,n}^{m+1}] \to K[\mathrm{Mat}_{n,n}^m]$, which descends to a surjective map on invariant rings (see [21, 13]).

Proposition 5.2.1. *The map $\phi^* : R(n,m+1) \twoheadrightarrow S(n,m)$ is surjective.*

Proof. We want to first show that the image $\phi^*(R(n,m+1)) \subseteq S(n,m)$. Since $\{f'_T \mid T \in \mathrm{Mat}_{d,d}^{m+1}\}$ is a spanning set for the ring $R(n,m+1)$, it suffices to show that $\phi^*(f_T) \in S(n,m)$ for all $T \in \mathrm{Mat}_{d,d}^{m+1}$. By definition, we have

$$\phi^*(f_T)(X_1, X_2, \ldots, X_m) = f_T(I, X_1, X_2, \ldots, X_m).$$

For $g \in \mathrm{GL}_n$, we have

$$\begin{aligned}
\phi^*(f_T)(g \cdot (X_1, X_2, \ldots, X_m)) &= \phi^*(f_T)(gX_1g^{-1}, gX_2g^{-1}, \ldots, gX_mg^{-1}) \\
&= f_T(I, gX_1g^{-1}, \ldots, gX_mg^{-1}) \\
&= \det(T_1 \otimes I + T_2 \otimes gX_1g^{-1} + \ldots T_{m+1} \otimes gX_mg^{-1}) \\
&= \det((I \otimes g)(T_1 \otimes I + \cdots + T_{m+1}X_m)(I \otimes g)^{-1}) \\
&= \det(T_1 \otimes I + T_2 \otimes X_1 + \cdots + T_{m+1}X_m) \\
&= f_T(I, X_1, \ldots, X_m) \\
&= \phi^*(f_T)(X_1, \ldots, X_m).
\end{aligned}$$

Now, we show that the image of ϕ^* surjects onto $S(n,m)$. For each $f \in S(n,m)$, we need to find an $\widetilde{f}$ such that $\phi^*(\widetilde{f}) = f$. Define $\widetilde{f}$ by

$$\widetilde{f}(X_1, \ldots, X_{m+1}) = f(Ad(X_1)X_2, Ad(X_1)X_3, \ldots, Ad(X_1)X_{m+1}),$$

where $Ad(X)$ denotes the adjoint of a matrix. It is easy to see that $\widetilde{f}$ is invariant for the action of $\mathrm{SL}_n \times \mathrm{SL}_n$. Further, we have

$$\begin{aligned}
(\phi^*(\widetilde{f}))(X_1, \ldots, X_m) &= \widetilde{f}(I, X_1, \ldots, X_m) \\
&= f(Ad(I)X_1, \ldots, Ad(I)X_m) \\
&= f(X_1, \ldots, X_m)
\end{aligned}$$

□

In fact, from the above proof, we can see that for $f \in S(n,m)$, we can construct a pre-image easily. We record this as a corollary.

Corollary 5.2.2. *For $f \in S(n,m)$, the invariant polynomial $\widetilde{f} \in R(n,m+1)$ defined by*

$$\widetilde{f}(X_1, \ldots, X_{m+1}) = f(Ad(X_1)X_2, Ad(X_1)X_3, \ldots, Ad(X_1)X_{m+1})$$

is a pre-image of f under ϕ^*, i.e., $\phi^*(\widetilde{f}) = f$.

Corollary 5.2.3. *We have* $\beta(S(n,m)) \leq \beta(R(n,m+1)) \leq (m+1)n^4$.

Corollary 5.2.4. *For a quiver* Q *and a dimension vector* α*, the ring of invariants* $\mathrm{I}(Q,\alpha)$ *is generated by invariants of degree* $(M+1)N^4$*, where* $M = |Q_1|$ *and* $N = \sum_{i \in Q_0} \alpha_i$.

Proof. Use the surjection from Theorem 3.2.2. □

5.3 Semi-invariants of quivers

Let Q be a quiver with no oriented cycles. A representation $V \in \mathrm{Rep}(Q,\alpha)$ is called σ-semistable if there exists an semi-invariant $f \in \mathrm{SI}(Q,\alpha)_{d\sigma}$ with $f(V) \neq 0$ (see [55]). From the degree bounds for $R(n,m)$ and the surjection in Corollary 3.2.7, we deduce

Proposition 5.3.1. *Let* $Q = (Q_0, Q_1)$ *be a quiver with no oriented cycles. Let* $\sigma \in \mathbb{Z}^{Q_0}$ *be a weight such that* $\sigma \cdot \alpha = 0$*, and let* $|\sigma|_\alpha := \sigma_+ \cdot \alpha - \sigma_- \cdot \alpha$*. Then*

1. *If* V *is* σ*-semistable, and* $d \geq |\sigma|_\alpha - 1$*, then there exists a semi-invariant* $f \in \mathrm{SI}(Q,\alpha)_{d\sigma}$ *with* $f(V) \neq 0$*;*
2. *If* $\mathrm{char}\, K = 0$*, then the ring* $\mathrm{SI}(Q,\alpha,\sigma)$ *is generated in degree* $\leq |\sigma|_\alpha^5$.

While we have given bounds for these subrings of $\mathrm{SI}(Q,\alpha)$, we require additional work to give bounds for the entire ring of semi-invariants. We will prove the following results in this section. Recall that $||\alpha||_1 = \sum_{i \in Q_0} |\alpha_i|$ and $||\alpha||_2 = (\sum_{i \in Q_0} \alpha_i^2)^{1/2}$. We prove:

Theorem 5.3.2. *Let* $Q = (Q_0, Q_1)$ *be a quiver with no oriented cycles, and let* $|Q_0| = n$*. Then the null cone for the action of* SL_α *on* $\mathrm{Rep}(Q,\alpha)$ *is defined by semi-invariants for nonzero weights* σ *such that* $|\sigma|_\alpha \leq \dfrac{||\alpha||_1^{2n}}{4(n-1)^{2n-2}}$.

The bounds on the degree of the invariants defining the null cone can be translated into bounds for the degree of generating invariants.

Theorem 5.3.3. *Let* $Q = (Q_0, Q_1)$ *be a quiver with no oriented cycles, and let* $|Q_0| = n$*. Assume* $\mathrm{char}\, K = 0$ *and let* r *be the Krull dimension of* $\mathrm{SI}(Q,\alpha)$*. The ring* $\mathrm{SI}(Q,\alpha)$ *is generated by semi-invariants of weights* σ *with*

$$|\sigma|_\alpha \leq \frac{3rn^2||\alpha||_1^{4n}}{128(n-1)^{4n-4}}.$$

Note that $\dim(\mathrm{SI}(Q,\alpha)) \leq \dim \mathrm{Rep}(Q,\alpha)$, which depends on Q_0 and Q_1. We show that using a version of Weyl's polarization theorem, we can give a bound that depends only on $n = |Q_0|$ and α.

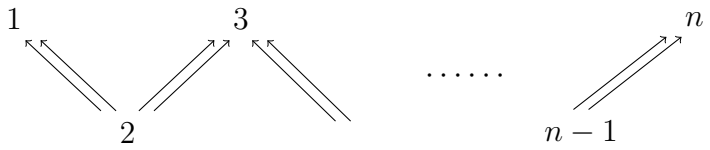

Figure 5.1: Quiver for exponential lower bounds

Corollary 5.3.4. *Let $Q = (Q_0, Q_1)$ be a quiver with no oriented cycles, and let $|Q_0| = n$ Assume* $\operatorname{char} K = 0$*. The ring* $\mathrm{SI}(Q, \alpha)$ *is generated by semi-invariants of weights σ with*

$$|\sigma|_\alpha \leq \frac{3}{256}\left(||\alpha||_1^2 - ||\alpha||_2^2\right)\frac{n^2||\alpha||_1^{4n}}{(n-1)^{4n-4}}.$$

Even though our bounds are not polynomial in $n = |Q_0|$, we give an example to show that it is not possible to obtain general bounds that are polynomial. Indeed consider the quiver Q_n shown in the above figure.

Proposition 5.3.5. *For the quiver Q_n, and dimension vector $\alpha = (2, 3, \ldots, 3, 1)$, the semi-invariants of weights σ with $|\sigma|_\alpha < 2^n - 2$ do not define the null cone, and hence do not generate* $\mathrm{SI}(Q_n, \alpha)$.

The rest of this section will be devoted to proving all the aforementioned results.

5.3.1 Stability conditions and the null cone

Fix a quiver $Q = (Q_0, Q_1)$ with no oriented cycles and let $|Q_0| = n$. There is a criterion for deciding σ-semistability of a representation in terms of the dimension vectors of subrepresentations due to King (see [55]). We use the conventions in [19]. Given a weight σ and a dimension vector β, we define $\sigma \cdot \beta = \sum_{i \in Q_0} \sigma_i \beta_i$.

Theorem 5.3.6 ([55])**.** *Let $Q = (Q_0, Q_1)$ be a quiver with no oriented cycles, α be a sincere dimension vector, and σ be a weight. Then we have:*

1. *A representation $V \in \mathrm{Rep}(Q, \alpha)$ is σ-semistable if and only if $\sigma \cdot \underline{\dim} V = 0$ and $\sigma \cdot \underline{\dim} W \leq 0$ for all subrepresentations $W \subset V$;*
2. *A representation $V \in \mathrm{Rep}(Q, \alpha)$ is σ-stable if and only if $\sigma \cdot \underline{\dim} V = 0$ and $\sigma \cdot \underline{\dim} W < 0$ for all proper subrepresentations $0 \neq W \subsetneq V$.*

The set of σ-semistable representations form an abelian subcategory of the category of finite dimensional representations of a quiver Q. The simple objects in the category are precisely the σ-stable representations.

Corollary 5.3.7. *If V is σ-semistable and $\sigma \cdot \underline{\dim} W = 0$ for some non-zero proper subrepresentation W of V, then W and V/W are also σ-semistable.*

In fact, we have a Jordan-Hölder filtration $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_m = V$. The composition factors V_i/V_{i-1} are unique up to rearrangement and isomorphism. Further these composition factors are σ-stable representations. We can define $\mathrm{gr}_\sigma(V) = \bigoplus_i V_i/V_{i-1}$.

Remark 5.3.8. *Let $d \in \mathbb{Z}_{>0}$. From Theorem 5.3.6, it follows that a representation V is σ-semistable (resp. stable) if and only if V is $d\sigma$-semistable (resp. stable). Hence, in particular, we have $\mathrm{gr}_\sigma(V) = \mathrm{gr}_{d\sigma}(V)$.*

Lemma 5.3.9. *A representation $V \in \mathrm{Rep}(Q, \alpha)$ is not in the null cone if and only if there exists a nonzero weight σ such that V is σ-semistable.*

Proof. We have already remarked that the null cone is the zero set of the semi-invariants of nonzero weights. Thus if a representation V is not in the null cone, then there is a semi-invariant $f \in \mathrm{SI}(Q, \alpha)_\sigma$ for some nonzero weight σ such that $f(V) \neq 0$. Consequently for this σ, f is σ-semistable. Conversely, if $V \in \mathrm{Rep}(Q, \alpha)$ is σ-semistable for some nonzero weight σ, then there is an invariant $f \in \mathrm{SI}(Q, \alpha)_{d\sigma}$ such that $f(V) \neq 0$ for some $d \in \mathbb{Z}_{>0}$. Hence V is not in the null cone. □

5.3.2 Bounds for the null cone

We first discuss some linear algebra that we require. For any vector $w = (w_1, w_2, \ldots, w_n) \in \mathbb{Q}^n$, recall that

$$||w||_1 = |w_1| + |w_2| + \cdots + |w_n|,$$

and

$$||w||_2 = (w_1^2 + w_2^2 + \cdots + w_n^2)^{1/2}.$$

We have the inequalities

$$||w||_2 \leq ||w||_1 \leq \sqrt{n}||w||_2.$$

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1} \in \mathbb{Z}^n_{\geq 0}$ be linearly independent over $\mathbb{Q}$, with $\vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_{n-1} = \vec{v} \in \mathbb{Z}^n_{\geq 0}$. Considering each $\vec{v}_i$ as a row vector, we can write a $(n-1) \times n$ matrix M whose i^{th} row is $\vec{v}_i$. Since the $\vec{v}_i$ are linearly independent over $\mathbb{Q}$, the rank of this matrix is $n-1$. Hence it has a 1-dimensional kernel. The following proposition bounds the smallest nonzero integral vector on this 1-dimensional kernel:

Proposition 5.3.10. *Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}, \vec{v}$ and M be as above. Then there is a nonzero integral vector $\vec{u} = (u_1, u_2, \ldots, u_n) \in \mathrm{Ker}(M)$ such that $|u_i| \leq \left(\frac{||\vec{v}||_1}{n-1}\right)^{n-1}$*

Proof. Let $\widehat{M}(i)$ denote the $(n-1) \times (n-1)$ minor of M obtained by removing the i^{th} column. Then define $u_i = (-1)^i \widehat{M}(i)$. It is clear that $\vec{u} = (u_1, u_2, \ldots, u_n)$ is an integral vector and that it is in the kernel of M. Further, we have

$$\begin{aligned}
|u_i| &\leq ||\vec{v}_1||_2 \cdot ||\vec{v}_2||_2 \cdots ||\vec{v}_{n-1}||_2 \\
&\leq ||\vec{v}_1||_1 \cdot ||\vec{v}_2||_1 \cdots ||\vec{v}_{n-1}||_1 \\
&\leq \left(\frac{||\vec{v}_1||_1 + ||\vec{v}_2||_1 + \cdots + ||\vec{v}_{n-1}||_1}{n-1}\right)^{n-1} \\
&= \left(\frac{||\vec{v}||_1}{n-1}\right)^{n-1}
\end{aligned}$$

□

Without loss of generality, we can assume α is sincere, i.e., $\alpha(x) \neq 0 \ \forall x \in Q_0$. If not, one can work with the subquiver $\widetilde{Q}$ defined by $\mathrm{supp}(\alpha)$, to get better bounds. Given a representation $V \in \mathrm{Rep}(Q, \alpha)$ which is not in the null cone, we denote by $\mathrm{C}(V)$, the set of weights σ for which V is σ-semistable, i.e,

$$\mathrm{C}(V) = \{\sigma | V \text{ is } \sigma\text{-semistable}\}.$$

Notice that $\mathrm{C}(V) \subset \mathbb{Z}^{Q_0}$ is cut out by a linear equation $\sigma(\alpha) = 0$ and by linear inequalities $\sigma(\underline{\dim} W) \leq 0$ for proper subrepresentations W of V. Let L be an extremal ray of $\mathrm{C}(V)$. It is clear that this extremal ray is defined by degenerating a subset of the linear inequalities to equalities. Hence, there exist subrepresentations $W_i, i \in I$ such that the equalities $\sigma(\underline{\dim} W_i) = 0$, and $\sigma(\alpha) = 0$ define $\mathbb{Q}L$.

Lemma 5.3.11. *There exist dimension vectors $\beta(i)$, $1 \leq i \leq n-1$ with $||\sum_{i=1}^{n-1} \beta(i)||_1 \leq ||\alpha||_1$ such that the line $\mathbb{Q}L$ is defined by the linear equalities $\sigma(\beta(i)) = 0$, $1 \leq i \leq n-1$.*

Before we prove the lemma we remark that V is σ-stable for a weight σ precisely when σ is in the interior of $C(V)$.

Proof. Let $\widetilde{\sigma} \in L$. Consider any Jordan-Hölder series (in the abelian subcategory of $\widetilde{\sigma}$-semistable representations) $0 = V_0 \subseteq V_1 \subset V_2 \subset \cdots \subset V_k = V$, and let $Z_i = V_i/V_{i-1}$,

$1 \leq i \leq k$ be the composition factors. Let $\alpha(i) = \dim Z_i$. We have $\widetilde{\sigma}(\alpha(i)) = 0$ since Z_i are $\widetilde{\sigma}$-semistable.

Let us look at the set of subrepresentations W_i such that the linear equalities $\sigma(\underline{\dim} W_i) = 0$ and $\sigma(\alpha) = 0$ define $\mathbb{Q}L$. Each W_j is a $\widetilde{\sigma}$-semistable subrepresentation of V, by Corollary 5.3.7. Therefore, the composition factors for W_j must be a subset of $\{Z_i \mid 1 \leq i \leq k\}$, and hence we have $\dim W_j = \alpha(i_1) + \alpha(i_2) + \cdots + \alpha(i_l)$, for some subset $\{i_1, i_2, \ldots, i_l\} \subseteq \{1, 2, \ldots, k\}$. Hence the condition $\sigma(\underline{\dim} W_i) = 0$ is a consequence of the conditions $\sigma(\alpha(i)) = 0$. In particular, we get that $\mathbb{Q}L$ is defined by the linear equalities $\sigma(\alpha(i)) = 0$ for $1 \leq i \leq k$.

However, some of these may be redundant. Since these equalities define a 1-dimensional subspace, there is a subset of the $\alpha(i)$'s of size $n-1$, say $\{\beta(1), \beta(2), \ldots, \beta(n-1)\}$, such that $\sigma(\beta(i)) = 0$ for $i = 1, 2, \ldots, n-1$ define $\mathbb{Q}L$.

□

Proposition 5.3.12. *Given a representation $V \in \mathrm{Rep}(Q, \alpha)$ such that $\mathrm{C}(V)$ is non-empty, V is σ-semistable for some weight σ such that each coordinate of σ has size $\leq \left(\frac{||\alpha||_1}{n-1}\right)^{n-1}$*

Proof. Pick an extremal ray L in $C(V)$. By Lemma 5.3.11, we have that $\mathbb{Q}L$ is the kernel of the $(n-1) \times n$ matrix whose rows are the dimension vectors $\beta(i)$. Let σ be the smallest integral vector in L. Then apply Proposition 5.3.10. □

We can now translate this into a bound for $|\sigma|_\alpha$.

Corollary 5.3.13. *Given a representation $V \in \mathrm{Rep}(Q, \alpha)$ such that $\mathrm{C}(V)$ is non-empty, V is σ-semistable for some weight σ such that*

$$|\sigma|_\alpha \leq \left(\frac{||\alpha||_1}{n-1}\right)^{n-1} \left(\frac{||\alpha||_1}{2}\right) = \frac{||\alpha||_1^n}{2(n-1)^{n-1}}.$$

Proof. If every coordinate $|\sigma_i| \leq M$ for some M, then we have $|\sigma_i|\alpha_i \leq M\alpha_i$. Note further that since $\sigma(\alpha) = 0$, we have $\sum_{i=1}^{n} |\sigma_i|\alpha_i = \sigma_+ \cdot \alpha + \sigma_- \cdot \alpha = 2|\sigma|_\alpha$. Thus $|\sigma|_\alpha \leq \frac{1}{2}M||\alpha||_1$. □

Proof of Theorem 5.3.2. Given $V \in \mathrm{Rep}(Q, \alpha)$ which is not in the null cone, we have that $C(V)$ is nonzero by Lemma 5.3.9. Hence there exists some σ with $|\sigma|_\alpha \leq \frac{||\alpha||_1^n}{2(n-1)^{n-1}}$, such that V is σ-semistable. Then by the first part of Proposition 5.3.1, there is a semi-invariant $f \in \mathrm{SI}(Q, \alpha)_{d\sigma}$ that does not vanish on V for each $d \geq |\sigma|_\alpha - 1$. Observe that $|d\sigma|_\alpha = d|\sigma|_\alpha$. Taking $d = |\sigma|_\alpha$ gives the required conclusion. □

Remark 5.3.14. *It might seem very wasteful to find bounds using an extremal ray L, as it is very likely that smaller weights lie in the interior of $\mathrm{C}(V)$. However, observe that if σ is an integral weight on an extremal ray L of $\mathrm{C}(V)$, then for $\mathrm{gr}_\sigma(V)$ we have $\mathrm{C}(\mathrm{gr}_\sigma(V)) = L$. Hence these extremal rays cannot be avoided.*

5.3.3 Bounds for generating semi-invariants

The ring $\mathrm{SI}(Q,\alpha)$ has two natural gradings. We have the weight space decomposition $\mathrm{SI}(Q,\beta) = \bigoplus_\sigma \mathrm{SI}(Q,\alpha)_\sigma$. We also have the natural grading inherited from viewing $K[\mathrm{Rep}(Q,\alpha)]$ as a polynomial ring. While the weight space decomposition is the more interesting one, all the results from Computational Invariant Theory hold for the latter grading. In the previous section, we found bounds for invariants defining the null cone in terms of the weight space decomposition. In order to use get degree bounds for generating invariants, we must switch to the latter grading.

Lemma 5.3.15. *Let $f \in \mathrm{SI}(Q,\alpha)_\sigma$, then its homogeneous components are non-trivial only for degrees between $|\sigma|_\alpha$ and $n|\sigma|_\alpha$.*

Proof. A set of semi-invariants spanning $f \in \mathrm{SI}(Q,\alpha)_\sigma$ was given in Theorem 3.2.5. A semi-invariant in this set is given by the determinant of a matrix, whose size is $|\sigma|_\alpha$. The matrix is described in block form, where each block defines a linear combinations of paths between two different vertices. Such paths have length at least 1 and at most n. Hence the entries of this matrix are polynomials whose homogeneous components are non-trivial only for degrees between 1 and n. □

The above lemma can then be used to convert the bounds given in Theorem 5.3.2 with respect to weight spaces to one in the total degree grading.

Corollary 5.3.16. *The null cone for the action of SL_α on $\mathrm{Rep}(Q,\alpha)$ is defined by homogeneous invariants of degree $\leq \dfrac{n||\alpha||_1^{2n}}{4(n-1)^{2n-2}}$, i.e.,*

$$\gamma(\mathrm{SI}(Q,\alpha)) \leq \frac{n||\alpha||_1^{2n}}{4(n-1)^{2n-2}}.$$

Corollary 5.3.17. *Assume $\mathrm{char}\, K = 0$ and let $r = \dim(\mathrm{SI}(Q,\alpha))$. The ring of semi-invariants $\mathrm{SI}(Q,\alpha)$ is generated by invariants of degree $\leq \dfrac{3}{8} r \left(\dfrac{n||\alpha||_1^{2n}}{4(n-1)^{2n-2}}\right)^2$.*

Proof. When K is algebraically closed, we can apply Theorem 2.2.6 to get the required bound. As a result of the discussion on good filtrations, the bound continues to hold in

positive characteristic as well. If K is only an infinite field, then working over the algebraic closure $\overline{K}$, observe that the ring of semi-invariants is simply $\mathrm{SI}(Q,\alpha) \otimes_K \overline{K}$. Hence the same bound holds over K as well. □

Proof of Theorem 5.3.3. This follows from Lemma 5.3.15 and Corollary 5.3.17. □

Remark 5.3.18. *The bounds given in Theorem 5.3.3 depend on* $\dim(\mathrm{SI}(Q,\alpha))$. *Kac gave a formula (see [53]) for* $\dim(\mathrm{SI}(Q,\alpha))$ *in terms of the canonical decomposition. There is in fact an efficient algorithm to compute the canonical decomposition, see [18]. More importantly, as remarked in the introduction,* $\dim(\mathrm{SI}(Q,\alpha))$ *is bounded by* $\dim(\mathrm{Rep}(Q,\alpha)) = \sum_{a\in Q_1} \alpha(ha)\alpha(ta)$.

5.3.4 Removing dependence on $\dim \mathrm{SI}(Q,\alpha)$

The bounds in Theorem 5.3.3 depend on $|Q_0| = n$, α and $\dim(\mathrm{SI}(Q,\alpha))$. Note that $\dim(\mathrm{SI}(Q,\alpha))$ depends on Q_1. We now show how one can use Weyl's theorem on polarization of invariants to remove the dependence on $\dim(\mathrm{SI}(Q,\alpha))$, and get a bound which is purely in terms of $|Q_0| = n$ and α.

Given a quiver $Q = (Q_0, Q_1)$ with no oriented cycles, we can label the vertices $1, 2, \ldots, n$ so that for every arrow, $ta < ha$. Let $n(i,j)$ denote the number of arrows with tail i and head j. Fix a dimension vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. Now, observe that

$$\mathrm{Rep}(Q,\alpha) = \bigoplus_{i<j} \mathrm{Mat}_{\alpha_j,\alpha_i}^{n(i,j)}.$$

Observe further that each $\mathrm{Mat}_{\alpha_j,\alpha_i}$ is a representation of GL_α as well as SL_α. Observe that $\dim \mathrm{Mat}_{\alpha_j,\alpha_i} = \alpha_i\alpha_j$. Hence, as a consequence of Weyl's theorem on polarization of invariants Theorem 2.3.1, we can obtain the semi-invariant ring $\mathrm{SI}(Q,\alpha)$ by the process of polarization from $K[\bigoplus_{i<j} \mathrm{Mat}_{\alpha_j,\alpha_i}^{\alpha_i\alpha_j}]^{\mathrm{SL}_\alpha}$. See also [32, Theorem 0.1] for a version that is better suited to our situation. In other words, for the purposes of finding a bound on the generating invariants, we can assume $n(i,j) = \alpha_i\alpha_j$.

Define a quiver $\widetilde{Q}$ whose vertex set is $1, 2, \ldots, n$, and has $\alpha_i\alpha_j$ arrows from i to j. The above discussion can be summarized as follows:

Proposition 5.3.19. *Assume* $\mathrm{char}\, K = 0$, *then we have*

$$\beta(\mathrm{SI}(Q,\alpha)) \leq \beta(\mathrm{SI}(\widetilde{Q},\alpha)).$$

$$x \underset{b}{\overset{a}{\rightrightarrows}} y$$

Figure 5.2: 2-Kronecker quiver

$$V = K \underset{\binom{0}{1}}{\overset{\binom{1}{0}}{\rightrightarrows}} K^2$$

Figure 5.3: Indecomposable representation 1

Proof of Corollary 5.3.4. For $\widetilde{Q}$, we have:

$$\begin{aligned}\dim(\mathrm{SI}(\widetilde{Q},\alpha)) &\leq \dim \mathrm{Rep}(\widetilde{Q},\alpha)\\ &= \sum_{i<j} \alpha_i\alpha_j\\ &= \frac{||\alpha||_1^2 - ||\alpha||_2^2}{2}.\end{aligned}$$

Now, use this bound for r in Corollary 5.3.17, and apply Lemma 5.3.15. □

5.3.5 Exponential lower bound

We first recall some results on the 2-Kronecker quiver, the quiver with 2 vertices x and y and two arrows a, b from x to y.

We look at two particular indecomposable representations. The indecomposable representation V has dimension vector $(1, 2)$ (see above), and the indecomposable representation W has dimension vector $(2, 1)$ (see below).

It is easy to check that V is σ-semistable precisely when $\sigma \in \mathbb{Z}_{>0}(2, -1)$ and that W is σ-semistable precisely when $\sigma \in \mathbb{Z}_{>0}(1, -2)$.

$$W = K^2 \underset{\left(\begin{smallmatrix}0&1\end{smallmatrix}\right)}{\overset{\left(\begin{smallmatrix}1&0\end{smallmatrix}\right)}{\rightrightarrows}} K$$

Figure 5.4: Indecomposable representation 2

Proof of Proposition 5.3.5. Consider the quiver Q_n, and observe that the odd vertices are sources and the even vertices are sinks. For any $i \in \{1, 2 \dots, n-1\}$, one of i and $i+1$ is a source and the other is a sink. Let ψ_i be the embedding of the 2-Kronecker quiver, that maps the vertices to i and $i+1$, with source begin mapped to source and sink to sink. Under this embedding, we see that $\psi_i(V)$ and $\psi_i(W)$ are indecomposable representations of the quiver Q_n. We consider the representation

$$\begin{aligned} R &= \psi_1(V) \oplus \psi_2(W) \oplus \psi_3(V) \cdots \\ &= \bigoplus_{i \text{ odd}} \psi_i(V) \oplus \bigoplus_{i \text{ even}} \psi_i(W). \end{aligned}$$

We have $\dim(R) = (2, 3, 3, \dots, 3, 1)$. Moreover, R is σ-semistable for the indivisible integral weight $\sigma = (-1, 2, -4, 8, \dots)$. Since R is a direct sum of indecomposables, it suffices to check σ-semistability of these indecomposables. That each of these indecomposables is σ-semistable follows from the above discussion above on 2-Kronecker quivers. Thus, in particular, $\mathrm{C}(R)$ is non-empty, and R is not in the null cone.

Moreover, we have that R is a direct sum of $n-1$ indecomposables, and their dimension vectors are linearly independent vectors, and hence it follows from King's stability conditions that $\mathrm{C}(R)$ is at most 1-dimensional. Since $\mathrm{C}(R)$ is non-empty, and $(-1, 2, -4, 8, \dots)$ is indivisible, we have that $\mathrm{C}(R) = \mathbb{Z}_{\geq 0}(-1, 2, -4, 8, \dots)$. More concretely, we have the condition that $\sigma \in \mathrm{C}(R)$, then σ is in the kernel of

$$\begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & \ddots & \ddots & \\ & & & 2 & 1 \end{pmatrix}.$$

The kernel of the above matrix is precisely the 1-dimensional subspace spanned by the vector $(-1, 2, -4, 8, \dots)$, and the smallest integral vector in this 1-dimensional subspace is $(-1, 2, -4, 8, \dots)$ by virtue of being indivisible. For the weight $\sigma = (-1, 2, -4, 8, \dots)$, we get $|\sigma|_\alpha = 2^n - 2$ by computation. Thus in this case, the semi-invariants of weights σ with $|\sigma|_\alpha < 2^n - 2$ do not define the null cone.

□

Remark 5.3.20. *For any given quiver, one might be able to generate stronger bounds by improving the estimates we make at various stages of obtaining our bounds.*

5.4 Quadratic lower bounds for matrix semi-invariants

5.4.1 Combinatorial description of $R(n, m)$

In this section, we introduce a combinatorial description of the invariants. We write $\lambda \vdash d$ to denote that λ is a partition of d. We denote by S_λ, the Schur functor corresponding to the partition λ. We identify $\mathrm{Mat}_{n,n}$ with $\mathbb{C}^n \otimes \mathbb{C}^n$, and consequently identify $\mathrm{Mat}_{n,n}^m$ with $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^m$. Thus

$$\mathbb{C}[\mathrm{Mat}_{n,n}^m] = \mathbb{C}[\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^m] = \bigoplus_{d=0}^{\infty} \mathrm{Sym}^d(\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^m).$$

Let $\lambda \vdash d$. We have the decomposition

$$S_\lambda(V \otimes W) = \bigoplus_{\mu,\nu \vdash d} (S_\mu(V) \otimes S_\nu(W))^{a_{\lambda,\mu,\nu}},$$

where $a_{\lambda,\mu,\nu}$ are known as the Kronecker coefficients. A particular case is the Cauchy formula,

$$\mathrm{Sym}^d(V \otimes W) = \bigoplus_{\lambda \vdash d} S_\lambda(V) \otimes S_\lambda(W).$$

Applying the above two decompositions, we get

$$\begin{aligned}
\mathrm{Sym}^d(V \otimes W \otimes Z) &= \bigoplus_{\lambda \vdash d} S_\lambda(V \otimes W) \otimes S_\lambda(Z) \\
&= \bigoplus_{\lambda \vdash d} \left(\bigoplus_{\mu,\nu \vdash d} (S_\mu(V) \otimes S_\nu(W))^{a_{\lambda,\mu,\nu}} \right) \otimes S_\lambda(Z) \\
&= \bigoplus_{\lambda,\mu,\nu \vdash d} (S_\mu(V) \otimes S_\nu(W) \otimes S_\lambda(Z))^{a_{\lambda,\mu,\nu}}.
\end{aligned}$$

This shows in particular that the Kronecker coefficients are invariant under permuting λ, μ and ν. The above is essentially a decomposition of $\mathrm{Sym}^d(V \otimes W \otimes Z)$ as a direct sum of irreducible representations of $\mathrm{GL}(V) \times \mathrm{GL}(W) \times \mathrm{GL}(Z)$.

Proposition 5.4.1. *The invariants of $R(n, m)$ have the following description.*

1. $R(n, m)_d = 0$ *if* $n \nmid d$.
2. $R(n, m)_{kn} = S_{k^n}(\mathbb{C}^n) \otimes S_{k^n}(\mathbb{C}^n) \otimes \left(\bigoplus_{\lambda \vdash kn} S_\lambda(\mathbb{C}^m)^{a_{k^n,k^n,\lambda}} \right)$.

Proof. We want the polynomials which are invariant under the action of SL. SL_n acts trivially on the irreducible representations of GL_n corresponding to the rectangular partitions of length n (i.e the powers of the determinant representation). On all other irreducible representations, SL_n acts with no invariants.

Thus the SL invariants of degree d are the summands $(S_\mu(\mathbb{C}^n)\otimes S_\nu(\mathbb{C}^n)\otimes S_\lambda(\mathbb{C}^m))^{a_{\lambda,\mu,\nu}}$ in the decomposition of $\mathrm{Sym}^d(\mathbb{C}^n\otimes\mathbb{C}^n\otimes\mathbb{C}^m)$ where μ,ν are rectangular partitions of length n, i.e, $\mu=\nu=k^n$ for some k. So, in particular, unless d is a multiple of n, we cannot have any invariants. This proves (1). For (2),

$$\begin{aligned} R(n,m)_{kn} &= \bigoplus_{\lambda\vdash kn}(S_{k^n}(\mathbb{C}^n)\otimes S_{k^n}(\mathbb{C}^n)\otimes S_\lambda(\mathbb{C}^m))^{a_{k^n,k^n,\lambda}} \\ &= S_{k^n}(\mathbb{C}^n)\otimes S_{k^n}(\mathbb{C}^n)\otimes\left(\bigoplus_{\lambda\vdash kn} S_\lambda(\mathbb{C}^m)^{a_{k^n,k^n,\lambda}}\right). \end{aligned}$$

□

Lemma 5.4.2. *The dimenson of $R(n,m)_{kn}$ is given by the computable formula*

$$\dim(R(n,m)_{kn}) = \sum_{\lambda\vdash kn} a_{k^n,k^n,\lambda}(\dim\ S_\lambda(\mathbb{C}^m)).$$

Proof. Since $S_{k^n}(\mathbb{C}^n)$ is 1-dimensional, as GL_m representations, we have

$$R(n,m)_{kn} = \bigoplus_{\lambda\vdash kn} S_\lambda(\mathbb{C}^m)^{a_{k^n,k^n,\lambda}}.$$

Hence we get the formula

$$\dim(R(n,m)_{kn}) = \sum_{\lambda\vdash kn} a_{k^n,k^n,\lambda}(\dim\ S_\lambda(\mathbb{C}^m)).$$

□

Remark 5.4.3. *Let $\lambda,\mu\vdash d$. If T_λ $(resp.\ T_\mu)$ denotes the irreducible representation of the symmetric group on d letters corresponding to the partition λ $(resp.\ \mu)$, then*

$$T_\lambda\otimes T_\mu = \bigoplus_\nu T_\nu^{a_{\lambda,\mu,\nu}}.$$

Example 5.4.4. $T_{1^n} \otimes T_{1^n} = T_n$. *Therefore by Lemma 5.4.2,*

$$\dim(R(n,m)_n) = \dim(S_n(\mathbb{C}^m)) = \binom{m+n-1}{n}.$$

Example 5.4.5. $T_{2^3} \otimes T_{2^3} = T_6 + T_{(4,2)} + T_{2^3} + T_{(3,1,1,1)}$. *Therefore by Lemma 5.4.2,*

$$\dim(R(3,m)_6) = \dim S_6(\mathbb{C}^m) + \dim S_{(4,2)}(\mathbb{C}^m) + \dim S_{2^3}(\mathbb{C}^m) + \dim S_{(3,1,1,1)}(\mathbb{C}^m).$$

5.4.2 Lower bounds for $\beta(R(n,m))$

In this section, we prove the following theorem.

Theorem 5.4.6. *Assume $K = \mathbb{C}$, and $m \geq n^2$. Then we have $\beta(R(n,m)) \geq n^2$.*

Let $R \subset \mathbb{C}[W \otimes V]$ be a $\mathrm{GL}(V)$ stable graded subring. Then each R_d is a finite dimensional $\mathrm{GL}(V)$ representation, and we can decompose it as a direct sum of irreducibles, i.e,

$$R_d = \bigoplus_{\lambda \vdash d} (S_\lambda(V))^{n_\lambda}, n_\lambda \in \mathbb{N}.$$

Note here that as $\mathrm{GL}(V)$ representations, the k^{th} exterior power $\bigwedge^k(V)$ is $S_{1^k}(V)$ for all positive integers k.

Proposition 5.4.7. *Let $R \subset \mathbb{C}[W \otimes V]$ be a $\mathrm{GL}(V)$ stable graded subring. Assume*

1. *$\bigwedge^i(V)$ does not occur in the decomposition of R_i, for $i = 1, 2, ..., t-1$;*
2. *$\bigwedge^t(V)$ occurs in the decomposition of R_t at least once ;*
3. *$\dim V \geq t$.*

Then R is not generated by invariants of degree $< t$.

Proof. We have a $\mathrm{GL}(V)$ equivariant map $R_i \otimes R_{t-i} \to R_t$ given by multiplication. We can collect the maps for various i to get a map

$$\varphi : \bigoplus_{i=1}^{\lfloor t/2 \rfloor} R_i \otimes R_{t-i} \to R_t.$$

It is clear that if R is generated by invariants of degree $< t$, then φ is surjective.

Let $\lambda \vdash a$ and $\mu \vdash b$. Recall the well known identity

$$S_\lambda(V) \otimes S_\mu(V) \;=\; \oplus_\nu \, (S_\nu(V))^{c^\nu_{\lambda,\mu}},$$

where $c^{\nu}_{\lambda,\mu}$ are the Littlewood-Richardson coefficients. By the Littlewood-Richardson rule, if $\bigwedge^{a+b}(V) = S_{1^{a+b}}(V)$ appears in the decomposition for $S_\lambda(V) \otimes S_\mu(V)$, then $\lambda = 1^a$ and $\mu = 1^b$.

We assume that for $1 \leq i \leq t-1$, $\bigwedge^i(V) = S_{1^i}(V)$ does not occur in the decomposition for R_i. Hence, $\bigwedge^t(V)$ does not occur in the decomposition for $R_i \otimes R_{t-i}$, and hence does not occur in the decomposition for $\bigoplus_{i=1}^{\lfloor t/2 \rfloor} R_i \otimes R_{t-i}$, and thus not in the decomposition of its image under φ. But in the decomposition of R_t, there is at least one copy of $\bigwedge^t(V)$. Since $\dim V \geq t$, we are guaranteed that $\bigwedge^t(V)$ is non-empty. So φ cannot be surjective.

Thus R cannot be generated in degree $< t$ as the invariants corresponding to the isotypic component for $\bigwedge^t(V)$ cannot be generated by smaller degree invariants. □

Proof of Theorem 5.4.6. We want to apply Proposition 5.4.7 to the ring $R(n, m)$, via the combinatorial description in Section 5.4.1. Take $W = \mathbb{C}^n \otimes \mathbb{C}^n$ and $V = \mathbb{C}^m$. Then by the results in Section 5.4.1, we have that

$$R_i = \begin{cases} \bigoplus_{\lambda \vdash kn} (S_\lambda(V))^{a_{k^n,k^n,\lambda}} & \text{if } i = kn, \\ 0 & \text{otherwise.} \end{cases}$$

From the representation theory of the symmetric group, we know that $T_{k^n} \otimes T_{1^{kn}} = T_{n^k}$. Moreover, since the Kronecker coefficients are invariant under permutations, we have

$$a_{k^n,k^n,1^{kn}} = \begin{cases} 0 & \text{if } k \neq n, \\ 1 & \text{if } k = n. \end{cases}$$

This gives the first two conditions required for Proposition 5.4.7 (for $t = n^2$) . Since we assume $\dim V = m \geq n^2$, the third condition holds as well. Hence we have $\beta(R(n, m)) \geq n^2$.

□

In fact, we can describe explicitly these invariants in degree n^2. For a matrix M, denote by $\overline{M}$, a column matrix obtained by stacking the columns of M.

Example 5.4.8. *If* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *then* $\overline{M} = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$.

Define a function f on n^2-tuples of $n \times n$ matrices by

$$f(X_1, X_2, \ldots, X_{n^2}) = \det \left(\begin{array}{c|c|c|c} \overline{X_1} & \overline{X_2} & \ldots & \overline{X_{n^2}} \end{array} \right).$$

Then $f \in R(n, n^2)_{n^2}$ is the unique invariant (upto scalars) in the isotypic component corresponding to $\bigwedge^{n^2}(\mathbb{C}^{n^2})$. For $n = 2$, this is the invariant of degree 4 constructed in [20].

5.5 Matrix semi-invariants for 3×3 matrices

In this section, we treat $R(3, m)$ in more detail. We exhibit invariants that define the null cone for $R(3, m)$. We have

Proposition 5.5.1. *The null cone for $R(3, m)$ is defined by a finite set of invariants of degree ≤ 6, namely*

- *A set of $\leq 9m - 16$ degree 3 invariants that define the same subvariety as the vanishing of all degree 3 invariants.*
- *The degree 6 invariants $g_{i,j,k} := det \left(\begin{array}{c|c} X_j & X_i \\ \hline X_i & X_k \end{array} \right)$ for $1 \leq i < j < k \leq m$.*

We can deduce from the proof of Proposition 5.5.1 that the invariants of degree 6 are necessary to define the null cone if $m \geq 3$.

Corollary 5.5.2. *For $m \geq 3$, we have $\gamma(R(3, m)) = 6$.*

This in turn gives us a hsop.

Proposition 5.5.3. *For $m \geq 3$, there exists a set of $9m - 16$ invariants of degree 6 that form a hsop for $R(3, m)$.*

Proposition 5.5.4. *The ring $R(3, m)$ is generated by invariants of degree ≤ 309 for all m.*

5.5.1 Krull dimension of $R(n, m)$

There is a formula for the dimension of the ring of semi-invariants of a quiver for a given dimension vector in terms of the canonical decomposition of the dimension vector due to Kac. In the case of the m-Kronecker quiver, the canonical decomposition of the dimension vector (n, n) is the following :

- $m = 1, 2$: The canonical decomposition is $(n, n) = (1, 1)^{\oplus n}$;
- $m \geq 3$: (n, n) is an imaginary Schur root and its canonical decomposition is trivial, i.e, $(n, n) = (n, n)$.

The cases for $m = 1, 2$ have already been dealt with, so we only apply Kac's formula for $m \geq 3$ to get the following lemma.

Lemma 5.5.5. *If $m \geq 3$, then we have* $\dim R(n, m) = mn^2 - 2(n^2 - 1)$.

5.5.2 Invariants defining the null cone

Proposition 5.5.1 gives a finite set of invariants that define the null cone. We rely heavily on the results in [21] for proving it.

Proof of Proposition 5.5.1. Let Z denote the vanishing set of all the degree 3 invariants. Note that the dimension of $R(3, m)$ is $9m - 16$, by Lemma 5.5.5. Hence, there is a set of $\leq 9m - 16$ degree 3 invariants that defines Z, by the Noether normalization lemma (see [10, Lemma 2.4.7]).

In [21], Domokos analyses the maximal singular matrix spaces in order to compute a hsop for $R(3, 3)$. We quickly summarize the results which we'll use.

A singular matrix space is a linear subspace of the space of matrices which does not contain an invertible matrix. The m-tuples in Z are precisely the m-tuples whose span is a singular matrix space, by the determinantal description of invariants.

In [21], Domokos classifies the maximal singular 3×3 spaces as being equivalent to one of 4 types, which are denoted $\mathcal{H}_i, i = 1, 2, 3, 4$ (see [21, Proposition 2.2]). A triple of matrices belongs to the null cone if and only if it belongs to a maximal singular space of type $\mathcal{H}_1, \mathcal{H}_2$, or $\mathcal{H}_3$, by [21, Proposition 3.2]. The same proof goes through for an m-tuple of matrices for any $m \geq 3$. Domokos remarks after [21, Proposition 2.2] that any 2-dimensional singular space is contained in $\mathcal{H}_1, \mathcal{H}_2$, or $\mathcal{H}_3$. $\mathcal{H}_4$ is the space of skew-symmetric matrices, and in particular is a 3-dimensional space. In [21], Domokos shows that the invariant $\det \left(\begin{array}{c|c} X_2 & X_1 \\ \hline X_1 & X_3 \end{array} \right)$ does not vanish on a triple of matrices (X_1, X_2, X_3) if the span of the triple is equivalent to $\mathcal{H}_4$.

Suppose an m-tuple $(X_1, X_2, \ldots, X_m)$ is in Z, but not in the null cone, then the span of the m-tuple is equivalent to $\mathcal{H}_4$, and hence 3-dimensional. Hence, there exist 3 matrices X_i, X_j, X_k which span the space. Hence $g_{i,j,k}$ is an invariant that does not vanish on the given m-tuple. □

Proof of Corollary 5.5.2. By Proposition 5.5.1, the invariants of degree ≤ 6 define the null cone. We observe from the proof of Proposition 5.5.1, that the degree 3 invariants are not sufficient to define the null cone if $m \geq 3$. □

Remark 5.5.6. *The set of invariants in Proposition 5.5.1 forms a hsop for $m = 3$, but not for $m \geq 4$ as the number of invariants is larger than the dimension of the ring.*

5.5.3 A hsop for $R(3, m), m \geq 3$

Proof of Proposition 5.5.3. Recall that $\dim(R(3, m)) = 9m - 16$. Since invariants of degree 3 and degree 6 define the null cone, it is clear that just the set of invariants of degree 6 define the null cone. By the Noether normalization lemma (see [10, Lemma 2.4.7]), we conclude that there exists $9m - 16$ degree 6 invariants that form a hsop. □

5.5.4 Upper bounds for $\beta(R(3, m))$

We want to bound the degrees of primary and secondary generators, in order to obtain upper bounds on $\beta(R(3, m))$. The following result of Knop is very useful in that regard.

Theorem 5.5.7 ([57])**.** *Let V be a rational representation of a semisimple connected group G. Let r be the Krull dimension of $\mathbb{C}[V]^G$, then*

$$\deg(H(\mathbb{C}[V]^G, t)) \leq -r.$$

In [10], this is used to get the following result.

Proposition 5.5.8 ([10])**.** *Let V and G be as in the Theorem 5.5.7. Suppose $f_1, f_2, \ldots, f_l \in \mathbb{C}[V]^G$ are homogeneous invariants that define the null cone. Let $d_i = \deg f_i$. Then*

$$\beta(\mathbb{C}[V]^G) \leq \max\{d_1, d_2, \ldots, d_l, d_1 + d_2 + \cdots + d_l - l\}.$$

There is a stronger result by Knop on the degree of the Hilbert series.

Theorem 5.5.9 ([58])**.** *Let V be a rational representation of a semisimple connected group G. Let $Z = \{v \in V | \dim G_v > 0\}$. Then*

$$\deg(H(\mathbb{C}[V]^G, t)) = -\dim V \iff \operatorname{codim}(Z) \geq 2.$$

In [20], the codimension condition was proved for $R(n, m)$ for $m \geq 3$, and $n \geq 2$. Since this stronger result on the degree of the Hilbert series holds, we can get a stronger

result by repeating the proof of Proposition 5.5.8 (see the proof of [10, Corollary 4.7.7]). Lemma 5.5.5 implies that for $m \geq 3$, the difference between $\dim R(n,m)$ and $\dim \mathrm{Mat}_{n,n}^m$ is $2n^2 - 2$. So, we get

Proposition 5.5.10. *Let $m \geq 3$ and $n \geq 2$. Suppose $f_1, f_2, \ldots, f_l \in R(n,m)$ are homogeneous invariants that define the null cone. Let $d_i = \deg f_i$. Then*

$$\beta(R(n,m)) \leq \max\{d_1, d_2, \ldots, d_l, d_1 + d_2 + \cdots + d_l - l - 2n^2 + 2)\}.$$

For computing upper bounds for $\beta(R(3,m))$, we can apply Proposition 5.5.10 to the set of invariants defining the null cone given by either Proposition 5.5.1 or Proposition 5.5.3. For $m \leq 3$, tight upper bounds have already been computed. For $4 \leq m \leq 6$, Proposition 5.5.1 gives better bounds, whereas for $m \geq 7$, Proposition 5.5.3 gives better bounds. So we get the following table.

m	Upper bound for $\beta(R(3,m))$
1	3
2	3
3	9
4	44
5	92
6	160
7	219
8	264
9	309

Proof of Proposition 5.5.4. As remarked in the introduction, a theorem of Weyl (see [56, Section 7.1,Theorem A]) gives us $\beta(R(3,m)) \leq \beta(R(3,9)) \leq 309$. □

5.6 Hilbert series computations

We know that for $m = 1, 2$, $R(n,m)$ is a polynomial ring. It is also clear that $R(1,m)$ is a polynomial ring since SL_1 is trivial. For $R(2,m)$, the Hilbert series has already been

computed by Domokos in [20]. So, we restrict to the cases $m \geq 3$, $n \geq 3$. Notice that for these cases, we have $\deg H(R(n,m),t) = -\dim \mathrm{Mat}_{n,n}^m$, as discussed in Section 5.5.4.

Remark 5.6.1. *If we can compute a denominator for the Hilbert series of $R(n,m)$, then we can compute the polynomial in the numerator once we know the dimensions of the graded pieces of $R(n,m)$ upto the degree of the numerator, which is given by*

$$\deg(Numerator) = \deg(Denominator) - n^2 m.$$

Remark 5.6.2. *Computing $\dim R(n,m)_{kn}$ is a hard task even with a computer, and is the bottleneck for these computations. So, it is desirable to minimize the degree of the numerator as much as possible, and hence it is desirable to minimize the degree of the denominator.*

Fortunately, the theory of universal denominators (see [19]) gives us strong results in our case. We first renormalize our grading to agree with the grading in [19].

Definition 5.6.3. *The renormalized Hilbert series is defined as*

$$\widetilde{H}(R(n,m),t) = \sum_{k=0}^{\infty} \dim R(n,m)_{kn} t^k.$$

Remark 5.6.4. *The usual Hilbert series and the renormalized Hilbert series are related by*

$$\widetilde{H}(R(n,m),t^n) = H(R(n,m),t).$$

The most relevant result is [19, Corollary 1]. We restate it for our situation.

Proposition 5.6.5 ([19])**.** *Let r be the Krull dimension of $R(n,m)$. Then*

$$\widetilde{H}(R(n,m),t) = \frac{P(t)}{(1-t)^r},$$

where $P(t)$ is a polynomial with integer coefficients.

This gives us denominators of the lowest degree possible, making several computations accessible. Domokos proved a functional equation for the Hilbert series of $R(n,m)$ for $m \geq 3$, $n \geq 2$ in [20]. This implies that when we use the universal denominator, the coefficients of the polynomial in the numerator are palindromic, so we need to compute only half the coefficients. In view of Remarks 5.6.1-5.6.2, this makes a few more computations feasible.

We give a few explicit computations that we are able to compute.

1. $\widetilde{H}(R(3,3),t) = \dfrac{1-t+t^2}{(1-t)^{11}}$. This was already computed by Domokos in [21]. We remark that even though $\beta(R(3,3)) = 9$, we only needed the $\dim(R(3,3)_3)$ to compute the Hilbert series.

2. $\widetilde{H}(R(3,4),t) = \dfrac{1+20t^2+20t^3+55t^4+20t^5+20t^6+t^8}{(1-t)^{20}}$.

3. $\widetilde{H}(R(3,5),t) = \dfrac{P(t)}{(1-t)^{29}}$, where the coefficients of $P(t)$ are 1, 6, 141, 931, 4816, 13916, 27531, 33391, 27531, 13916, 4816, 931, 141, 6, 1.

4. $\widetilde{H}(R(3,6),t) = \dfrac{P(t)}{(1-t)^{38}}$, where the coefficients of $P(t)$ are 1, 18, 626, 10246, 114901, 830484, 4081260, 13763184, 32507115, 54176230, 64224060, 54176230, 32507115, 13763184, 4081260, 830484, 114901, 10246, 626, 18, 1.

5. $\widetilde{H}(R(3,7),t) = \dfrac{P(t)}{(1-t)^{47}}$, where the coefficients of $P(t)$ are 1, 37, 2033, 62780, 1301634, 18067706, 173883458, 1186198090, 5851715254, 21192401230, 57013957462, 114926408114, 174616665986, 200665719450, 174616665986, 114926408114, 57013957462, 21192401230, 5851715254, 1186198090, 173883458, 18067706, 1301634, 62780, 2033, 37, 1.

6. $\widetilde{H}(R(4,3),t) = \dfrac{1-3t+9t^2+8t^3+9t^4-3t^5+t^6}{(1-t)^{18}}$.

7. $\widetilde{H}(R(4,4),t) = \dfrac{P(t)}{(1-t)^{34}}$, where the coefficients of $P(t)$ are 1, 1, 141, 981, 8534, 39193, 139348, 325823, 556368, 652716, 556368, 325823, 139348, 39193, 8534, 981, 141, 1, 1.

8. $\widetilde{H}(R(5,3),t) = \dfrac{P(t)}{(1-t)^{27}}$, where the coefficients of $P(t)$ are 1, -6, 36, -70, 231, -189, 419, -189, 231, -70, 36, -6, 1.

CHAPTER 6

Orbit closure problem and separating invariants

We introduce the orbit closure problem in Section 6.1 and describe the known algorithms. In Section 6.2, we show polynomial reductions (in both directions) between the orbit closure problems for matrix invariants and matrix semi-invariants. We give a polynomial time algorithm for finding the basis of a subalgebra of $\mathrm{Mat}_{n,n}$ in Section 6.3, and use this to give an algorithm for matrix invariants in Section 6.4. Finally in Section 6.5, we give bounds on separating invariants.

6.1 Introduction

Deciding whether the orbit closures of two points is an important problem in computational invariant theory. For example, Mulmuley and Sohoni reformulated Valiant's algebraic version of P vs NP to a problem of deciding whether the orbit closures of two points intersect.

Problem 6.1.1. *The orbit closure problem for the action of a group G on V is the problem of deciding whether the orbit closures of two given points $v, w \in V$ intersect.*

We make a definition for ease of notation.

Definition 6.1.2. *Two points $v, w \in V$ are said to be closure equivalent if $\overline{G \cdot v} \cap \overline{G \cdot w} \neq \emptyset$. We write $v \sim_G w$ if v and w are closure equivalent, and we write $v \not\sim_G w$ if they are not closure equivalent. We drop the subscript G if there is no ambiguity in the group being considered.*

6.1.1 Known algorithms for matrix invariants

In characteristic 0, Forbes and Shpilka show the existence of a quasi-polynomial sized set of separating invariants for the simultaneous conjugation action of GL_n on $\mathrm{Mat}_{n,n}^m$ (see [35]), but this alone is not sufficient to get a polynomial time algorithm.

Nevertheless, Forbes and Shpilka give a polynomial time algorithm for the orbit closure problem. Given an input $X \in \mathrm{Mat}_{n,n}^m$, they construct in polynomial time a noncommutative polynomial P_X with the feature that the coefficients of the monomials in P_X are the evaluations of a generating set of invariants on X. Hence, to check if the orbit closures of two points $X, Y \in \mathrm{Mat}_{n,n}^m$ intersect, one needs to determine whether the noncommutative polynomial $P_X - P_Y$ is the zero polynomial. There is an efficient algorithm to test whether $P_X - P_Y$ is the zero polynomial (see [71]).

This algorithm has two shortcomings. The first is that it is not easily adapted to positive characteristic. The second is that when the orbit closures of X and Y do not intersect, this algorithm does not provide a specific invariant $f \in S(n, m)$ such that $f(X) \neq f(Y)$.

In this chapter, we will rectify both shortcomings with a different algorithm. We will also solve the corresponding problem for matrix semi-invariants.

6.2 Time complexity equivalence of orbit closure problems for matrix invariants and matrix semi-invariants

In this section, we will show polynomial reductions between the orbit closure problem for matrix invariants and the orbit closure problem for matrix semi-invariants. We will in fact show a more robust reduction.

Let G be a group acting on V.

Definition 6.2.1. *An algorithm for the orbit closure problem with witness is an algorithm that decides if $v \sim w$ for any two points $v, w \in V$, and if $v \not\sim w$, provides a witness $f \in K[V]^G$ such that $f(v) \neq f(w)$.*

We use $\sim_{LR}$ to denote closure equivalence for the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$ on $\mathrm{Mat}_{n,n}^m$. We continue to use $\sim_{\mathrm{GL}_n}$ to denote the closure equivalence for the simultaneous conjugation action of GL_n on $\mathrm{Mat}_{n,n}^m$.

6.2.1 Reduction from matrix invariants to matrix semi-invariants

Let $A, B \in \mathrm{Mat}_{n,n}^m$. We can consider $\phi(A), \phi(B) \in \mathrm{Mat}_{n,n}^{m+1}$, where $\phi : \mathrm{Mat}_{n,n}^m \to \mathrm{Mat}_{n,n}^{m+1}$ is the map described in Section 5.2. Recall that this gives a surjection $\phi^* : R(n, m+1) \twoheadrightarrow S(n, m)$.

Proposition 6.2.2. *The following are equivalent:*

1. *There exists $f \in S(n, m)$ such that $f(A) \neq f(B)$*

2. *There exists* $g \in R(n, m+1)$ *such that* $g(\phi(A)) \neq g(\phi(B))$.

Proof. Let's first prove $(1) \implies (2)$. Given $f \in S(n, m)$ such that $f(A) \neq f(B)$, take g to be a preimage of f, i.e., $\phi^*(g) = f$. Now,

$$g(\phi(A)) = \phi^*(g)(A) = f(A) \neq f(B) = \phi^*(g)(B) = g(\phi(B)).$$

To prove $(2) \implies (1)$, simply take $f = \phi^*(g)$. □

Corollary 6.2.3. *Let* $A, B \in \mathrm{Mat}_{n,n}^m$. *Then we have*

$$A \sim_{\mathrm{GL}_n} B \text{ if and only if } \phi(A) \sim_{LR} \phi(B).$$

Corollary 6.2.4. *There is a polynomial reduction that reduces the orbit closure problem with witness for matrix invariants to the orbit closure problem with witness for matrix semi-invariants*

Proof. Given $A, B \in \mathrm{Mat}_{n,n}^m$, we construct $\phi(A)$ and $\phi(B)$. Appeal to the orbit closure problem with witness for matrix semi-invariants with input $\phi(A)$ and $\phi(B)$. There are two possible outcomes. If $\phi(A) \sim_{LR} \phi(B)$, then we conclude that $A \sim_{\mathrm{GL}_n} B$. If $\phi(A) \not\sim_{LR} \phi(B)$ and $f \in R(n, m+1)$ separates $\phi(A)$ and $\phi(B)$, then $\phi^*(f)$ is an invariant that separates A and B. The reduction is clearly a polynomial one. □

6.2.2 Reduction from matrix semi-invariants to matrix invariants

We will show that the orbit closure problem for matrix semi-invariants can be reduced to the orbit closure problem for matrix invariants. We will need some preparatory lemmas before we give the algorithm.

Lemma 6.2.5. *Assume* $A = (I, A_2, \ldots, A_m)$ *and* $B = (I, B_2, \ldots, B_m)$. *If we denote* $\widetilde{A} = (A_2, \ldots, A_m)$ *and* $\widetilde{B} = (B_2, \ldots, B_m)$. *Then we have*

$$A \sim_{LR} B \iff \widetilde{A} \sim_{\mathrm{GL}_n} \widetilde{B}.$$

Proof. This follows from Corollary 6.2.3 applied to $\widetilde{A}$ and $\widetilde{B}$. □

The above lemma paves the way for a slightly more general result.

Proposition 6.2.6. *Assume* $A, B \in \mathrm{Mat}_{n,n}^m$ *such that* $\det(A_1) = \det(B_1) \neq 0$. *If we denote* $\widetilde{A} = (A_1^{-1}A_2, \ldots, A_1^{-1}A_m)$ *and* $\widetilde{B} = (B_1^{-1}B_2, \ldots, B_1^{-1}B_m)$, *then we have*

$$A \sim_{LR} B \iff \widetilde{A} \sim_{\mathrm{GL}_n} \widetilde{B}.$$

Proof. Let us first suppose that $\det(A_1) = \det(B_1) = 1$. Then for $g = (A_1^{-1}, \mathrm{id}) \in \mathrm{SL}_n \times \mathrm{SL}_n$, we have $g \cdot A = (I, A_1^{-1}A_2, \ldots, A_1^{-1}A_m) = \phi(\widetilde{A})$. Similarly for $h = (B_1^{-1}, \mathrm{id}) \in \mathrm{SL}_n \times \mathrm{SL}_n$, we have $h \cdot B = \phi(\widetilde{B})$. Now, we have

$$A \sim_{LR} B \iff \phi(\widetilde{A}) \sim_{LR} \phi(\widetilde{B}) \iff \widetilde{A} \sim_{\mathrm{GL}_n} \widetilde{B}.$$

The general case for $\det(A_1) \neq 0$ follows because the orbit closures of A and B intersect if and only if the orbit closures of $\lambda \cdot A = (\lambda A_1, \ldots, \lambda A_m)$ and $\lambda \cdot B = (\lambda B_1, \ldots, \lambda B_m)$ intersect for any $\lambda \in K^*$. □

Lemma 6.2.7. *For any non-zero row vector* $\mathbf{v} = (v_1, \ldots, v_m)$*, we can construct efficiently a matrix* $P \in \mathrm{GL}_m$ *such that the top row of the matrix* P *is* $\mathbf{v}$.

Proof. This is straightforward and left to the reader. □

For $1 \leq j, k \leq d$, we define $E_{j,k} \in \mathrm{Mat}_{d,d}$ to be the $d \times d$ matrix which has a 1 in the $(j,k)^{th}$ entry, and 0 everywhere else.

Definition 6.2.8. *If* $X = (X_1, \ldots, X_m) \in \mathrm{Mat}_{n,n}^m$*, we define* $X^{\otimes d} = (X_i \otimes E_{j,k})_{i,j,k} \in \mathrm{Mat}_{nd,nd}^{md^2}$*, where the tuples* $(i,j,k) \in [m] \times [d] \times [d]$ *are ordered lexicographically.*

Proposition 6.2.9. *The following are equivalent*

1. *There exists* $f \in R(n,m)$ *such that* $f(A) \neq f(B)$*;*
2. *There exists* $g \in R(nd, md^2)$ *such that* $g(A^{\otimes d}) \neq g(B^{\otimes d})$ *for either* $d = n-1$ *or* $d = n$*.*

Proof. We first show $(1) \implies (2)$. We can assume $f = f_T$ for some $T \in \mathrm{Mat}_{e,e}^m$ for some $e \geq 1$. Without loss of generality, assume $f(A) \neq 0$. Then we have $\mu = f(B)/f(A) \neq 1$. For any $\mu \neq 1$, both μ^{n-1} and μ^n cannot be 1. Hence for at least one of $d \in \{n-1, n\}$, we have $\mu^d = f(B)^d/f(A)^d \neq 1$, and hence $f(A)^d \neq f(B)^d$. Now, it suffices to show the existence of $g \in R(nd, md^2)$ such that $g(A^{\otimes d}) = f(A)^d$ for all $A \in \mathrm{Mat}_{n,n}^m$.

But now, consider

$$\begin{aligned} f_T(A)^d &= \det(\sum_{i=1}^m T_i \otimes A_i)^d \\ &= \det(\sum_{i=1}^m T_i^{\oplus d} \otimes A_i)) \\ &= \det(\sum_{i=1}^m (\sum_{k=1}^d T_i \otimes E_{k,k} \otimes A_i)) \\ &= \det(\sum_{i,k} T_i \otimes (A_i \otimes E_{k,k})) \end{aligned}$$

Let $S \in \mathrm{Mat}_{e,e}^{md^2}$ given by $S_{i,j,k} = \delta_{j,k} T_i$. We can take $g = f_S$.

We now show (2) $\implies$ (1). Indeed, we can choose $g = f_S$ for some $S \in \mathrm{Mat}_{d',d'}^{md^2}$, with $d' \geq 1$. We have

$$\begin{aligned} f_S(A^{\otimes d}) &= \det\left(\sum_{i,j,k} S_{i,j,k} \otimes (A^{\otimes d})_{i,j,k}\right) \\ &= \det\left(\sum_{i,j,k} S_{i,j,k} \otimes A_i \otimes E_{j,k}\right) \\ &= \det\left(\sum_i \left(\sum_{j,k} S_{i,j,k} \otimes E_{j,k}\right) \otimes A_i\right) \\ &= \det\left(\sum_i \widetilde{S}_i \otimes A_i\right), \end{aligned}$$

where $\widetilde{S}_i = \sum_{j,k} S_{i,j,k} \otimes E_{j,k}$. Let $\widetilde{S} = (\widetilde{S}_1, \ldots, \widetilde{S}_m) \in \mathrm{Mat}_{dd',dd'}^m$. Then the above calculation tells us that $f_{\widetilde{S}}(A) = f_S(A^{\otimes d}) = g(A^{\otimes d})$. Hence we have

$$f_{\widetilde{S}}(A) = g(A^{\otimes d}) \neq g(B^{\otimes d}) = f_{\widetilde{S}}(B).$$

We can take $f = f_{\widetilde{S}}$.

□

Corollary 6.2.10. *The orbit closures of A and B do not intersect if and only if the orbit closures of $A^{\otimes d}$ and $B^{\otimes d}$ do not intersect for at least one choice of $d \in \{n-1, n\}$.*

6.2.2.1 Commuting action of another group

Let G be a group acting on V. Suppose we have another group H acting on V, and the actions of G and H commute. The orbit closure problem for the action of G on V also commutes with the action of H. More precisely, we have the following:

Lemma 6.2.11. *Let $v, w \in V$ and $h \in H$. Then $v \sim_G w$ if and only if $h \cdot v \sim_G h \cdot w$.*

We have a natural identification of $V = \mathrm{Mat}_{n,n}^m$ with $\mathrm{Mat}_{n,n} \otimes K^m$. The latter viewpoint illuminates an action of GL_m on V that commutes with the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$, as well as the simultaneous conjugation action of GL_n. In explicit terms, for $P = (p_{i,j}) \in \mathrm{GL}_m$ and $X = (X_1, \ldots, X_m) \in \mathrm{Mat}_{n,n}^m$, we have

$$P \cdot (X_1, \ldots, X_m) = (\sum_j p_{1,j} X_j, \sum_j p_{2,j} X_j, \ldots, \sum_j p_{m,j} X_j).$$

Corollary 6.2.12. *The orbit closure problem for both the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$ and the simultaneous conjugation action of GL_n on $\mathrm{Mat}_{n,n}^m$ commutes with the action of GL_m.*

6.2.2.2 The reduction

Now, we can give the algorithm to reduce the orbit closure problem with witness for matrix semi-invariants to the orbit closure problem with witness for matrix invariants. Let the input be $A, B \in \mathrm{Mat}_{n,n}^m$.

Step 1: Check if A or B are in the null cone. This can be done in polynomial time by the algorithm in [50]. We will henceforth refer to this as the IQS algorithm. If both of them are in the null cone, then $A \sim_{LR} B$. If precisely one of them is in the null cone, then $A \not\sim_{LR} B$ and the IQS algorithm gives an invariant that separates A and B. If neither are in the null cone, then we proceed to Step 2.

Step 2: A and B are both not in the null cone. Now, for $d \in \{n-1, n\}$, the IQS algorithm constructs $T(d) \in \mathrm{Mat}_{d,d}^m$ such that $f_{T(d)}(A) \neq 0$ in polynomial time. We denote $f_{T(d)} = f_d$. If $f_d(A) \neq f_d(B)$, then $A \not\sim_{\mathrm{GL}_n} B$ and f_d is the separating invariant. Else $f_d(A) = f_d(B)$ for both choices of $d \in \{n-1, n\}$, and we proceed to Step 3.

Step 3: For $d \in \{n-1, n\}$, we have $f_d(A) = \det(\sum_{i,j,k} (T(d)_i)_{j,k} (A_i \otimes E_{j,k}))$. We can construct efficiently a matrix $P \in \mathrm{Mat}_{md^2, md^2}$ such that the first row is $(T(d)_i)_{j,k})_{i,j,k}$ by Lemma 6.2.7. Consider $U = P \cdot A^{\otimes d}, V = P \cdot B^{\otimes d} \in \mathrm{Mat}_{nd,nd}^{md^2}$. This has the

property that $\det(U_1) \neq 0$. Since we did not terminate in Step 2, we know that $\det(U_1) = \det(V_1)$. Let us recall that by Corollary 6.2.10, $A \sim_{LR} B$ if and only $A^{\otimes d} \sim B^{\otimes d}$ for both $d = n-1$ and $d = n$. By Lemma 6.2.11, $A^{\otimes d} \sim_{LR} B^{\otimes d}$ if and only if $U \sim_{LR} V$.

To decide whether $U \sim_{LR} V$, we do the following. Let $\widetilde{U} = (U_1^{-1}U_2, \ldots, U_1^{-1}U_{md^2})$ and $\widetilde{V} = (V_1^{-1}V_2, \ldots, V_1^{-1}V_{md^2})$. By the above Proposition, we have $U \sim_{LR} V$ if and only if $\widetilde{U} \sim_{GL_{nd}} \widetilde{V}$. But this can be seen as an instance of an orbit closure problem with witness for matrix invariants. Also note the fact if we get an invariant separating $\widetilde{U}$ and $\widetilde{V}$, the steps can be traced back to get an invariant separating A and B.

Corollary 6.2.13. *There is a polynomial time reduction from the orbit closure problem with witness for matrix semi-invariants to the orbit closure problem with witness for matrix invariants.*

6.3 A polynomial time algorithm for finding a basis for a subalgebra of $\mathrm{Mat}_{n,n}$

Let $\{C_1, \ldots, C_m\} \subseteq \mathrm{Mat}_{n,n}$ be a finite subset of $\mathrm{Mat}_{n,n}$. Consider the (unital) subalgebra $\mathcal{C} \subseteq \mathrm{Mat}_{n,n}$ generated by $C_1, \ldots, C_m$. In other words, $\mathcal{C}$ is the smallest subspace of $\mathrm{Mat}_{n,n}$ containing $\mathrm{Id}, C_1, \ldots, C_m$ closed under addition and multiplication. We will describe a polynomial time algorithm for finding a basis for $\mathcal{C}$. First observe that $\mathcal{C}$ is spanned by $\{C(w) \mid w \in \mathrm{words}([m])\}$. While this is an infinite spanning set, we will extract a basis from this, in polynomial time. We define a total order on $\mathrm{words}([m])$.

Definition 6.3.1. *Given two words $w_1 = i_1 i_2 \ldots i_b$ and $w_2 = j_1 j_2 \ldots j_c$, we write $w_1 \prec w_2$ if either*

1. *$l(w_1) < l(w_2)$ or*
2. *$l(w_1) = l(w_2)$ and for the smallest integer m for which $i_m \neq j_m$, we have $i_m < j_m$.*

Remark 6.3.2. *If $w \prec w'$, we will say w is smaller than w'.*

We call a word a non-pivot if C_w is a finite linear combination of words w_k with $w_k \prec w$, i.e., $C_w = \sum_k a_k C_{w_k}$, with $a_k \in K$ and $w_k \prec w$. We call a word pivot if it is not a non-pivot. For a pivot (resp. non-pivot) word w, we will also refer to C_w as pivot (resp. non-pivot). The following lemma is straightforward.

Lemma 6.3.3. *Let $P = \{w \mid w \text{ is pivot}\}$. Then $\{C_w \mid w \in P\}$ is a basis for $\mathcal{C}$. We will call this the pivot basis.*

Definition 6.3.4. *For two words $w = i_1 i_2 \dots i_b$ and $w' = j_1 j_2 \dots j_c$, we define the concatenation*

$$w \circ w' = i_1 i_2 \dots i_b j_1 j_2 \dots j_c.$$

Lemma 6.3.5. *If w is a non-pivot, then $w \circ w'$ is a non-pivot for all words w'.*

Proof. If w is non-pivot, then $C_w = \sum_k a_k C_{w_k}$ for $w_k \prec w$ and $a_k \in K$. Then we have $C_{w \circ w'} = \sum_k a_k C_{w_k \circ w'}$. Hence, $w \circ w'$ is non-pivot as well. □

Corollary 6.3.6. *If there are no pivot words of length t, then there are no pivot words of length $\geq t$.*

Corollary 6.3.7. *There are no pivot words of length $> n^2$.*

Proof. Let N be the largest length of a pivot word. Then there must be at least one pivot word of length d for each $1 \leq d \leq N$, by Corollary 6.3.6. Hence,

$$n^2 = \dim(\mathrm{Mat}_{n,n}) \geq \dim(\mathcal{C}) = \text{number of pivots} \geq N.$$

□

Lemma 6.3.8. *If w is a non-pivot, then C_w is a finite linear combination of pivots x_k with $x_k \prec w$.*

Proof. If w is non-pivot, then C_w is a finite linear combination of words w_k with $w_k \prec w$. For each one of these words w_k, if it is not already a pivot, we can write w_k as a finite linear combination of smaller pivots, by induction. □

Corollary 6.3.9. *The set of pivots form a basis.*

Proof. The above lemma shows that every non-pivot is a finite linear combination of pivots, and hence the pivots form a spanning set. Now, suppose there was a non-trivial linear combination of pivots $\sum_k a_k C_{w_k} = 0$. Let w_k be the largest pivot such that $a_k \neq 0$. This means w_k is a linear combination of smaller pivots, which contradicts the fact that w_k is a pivot. Hence, the pivots are linearly independent. □

Now, we describe an efficient algoirithm to construct the set of pivots.

Step 0: Set $t = 1$. Set $P = P_0 = \{(e, \mathrm{Id})\}$, where e represents the empty word.

Step 1: Create a list P_t consisting of tuples (w, C_w), where w is a word of length t of the form $w' \circ i$, where $w' \in P_{t-1}$ is a pivot and $i \in [m]$. Create this list in order.

Step 2: Proceeding through the list P_t, check if an entry (w, C_w) is a pivot. This can be done in polynomial time, as we have to simply check if C_w is a linear combination of smaller pivots. If it is a pivot, add it to P. If it is not a pivot, then remove it from P_t. Upon completing this step, the list P_t contains all pivots of length t.

Step 3: If P_t is nonempty, $t = t + 1$, and go back to Step 1. Else, proceed to Step 4.

Step 4: Return P.

Corollary 6.3.10. *There is a polynomial time algorithm to construct the set of pivots. Further, this algorithm also records the word associated to each pivot.*

Proof. To show that the above algorithm runs in polynomial time, it suffices to show that the number of words we consider is at most polynomial. Indeed, if there are k pivots of length d, then we only consider $k \times m$ words of length $d + 1$. Since $k \leq n^2$, the number of words we consider in each degree is at most $n^2 m$. We only consider words of length up to n^2, since there will be no pivots larger than n^2. Hence, the number of words considered is polynomial (in n and m). □

6.4 Orbit closure problem for matrix invariants

We will first discuss the orbit closure problem for matrix invariants in characteristic 0. We will need some additional results to adapt the algorithm to positive characteristic, and we do that subsequently. Suppose $A, B \in \mathrm{Mat}_{n,n}^m$ and write $A = (A_1, \ldots, A_m)$, $B = (B_1, \ldots, B_m)$. For every $i \in [m]$, define a block matrix

$$C_i = \begin{pmatrix} A_i & 0 \\ 0 & B_i \end{pmatrix}.$$

Let $\mathcal{C} \subseteq \mathrm{Mat}_{2n,2n}$ be the algebra generated by $C_1, \ldots, C_m$.

Proposition 6.4.1. *We have $A \sim_{\mathrm{GL}_n} B$ if and only if for every matrix*

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \in \mathcal{C}$$

we have $\operatorname{Tr}(X) = \operatorname{Tr}(Y)$.

Proof. Two orbit closures do not intersect if and only if there is an invariant that separates them. We know that invariants of the form $X \mapsto \operatorname{Tr} X_w$ for some word w in the alphabet $\{1, 2, \ldots, m\}$ are a set of generators. Note that $\mathcal{C}$ is the span of all

$$C_w = \begin{pmatrix} A_w & 0 \\ 0 & B_w \end{pmatrix},$$

where w is a word. Now the proposition follows by linearity of trace. □

Theorem 6.4.2. *Assume* $\operatorname{char}(K) = 0$. *There is a polynomial time algorithm for the orbit closure problem with witness for matrix invariants.*

Proof. We can construct the set of pivots for $\mathcal{C} \subseteq \operatorname{Mat}_{2n,2n}$ in polynomial time by Corollary 6.3.10. Since the pivots form a basis for $\mathcal{C}$, it suffices to check whether $\operatorname{Tr} A_w = \operatorname{Tr} B_w$ for each pivot $C_w = \begin{pmatrix} A_w & 0 \\ 0 & B_w \end{pmatrix}$. In fact, we only need to check for pivots w with $l(w) \leq n^2$ by the degree bound for generating invariants.

If $\operatorname{Tr} A_w \neq \operatorname{Tr} B_w$ for some pivot w, then T_w is an invariant that separates A and B. □

6.4.1 The positive characteristic case

If $\operatorname{char}(K) = 0$, then characteristic coefficients of a matrix can be written in terms of traces of powers of a matrix. If $\operatorname{char}(K) = p > 0$, this is no longer true, and hence we are forced to consider characteristic coefficients in the description of the invariant ring $S(n, m)$. Higher characteristic coefficients lack the linearity of trace, and we must make a more involved argument.

Suppose that K is an algebraically closed field, R is a finite dimensional associative K-algebra and $N : R \to K$ is a norm, meaning that it has the following properties:

1. N is a polynomial;
2. $N(1) = 1$;
3. $N(ab) = N(a)N(b)$ for all $a, b \in R$.

We define a trace function as follows:

$$\operatorname{Tr}(a) = \frac{d}{dt} N(1 + at)\Big|_{t=0}.$$

The map Tr is K-linear.

Lemma 6.4.3. *We have*

$$\frac{d}{dt}N(a+bt) = N(a+bt)\,\mathrm{Tr}((a+bt)^{-1}b)$$

for all $t \in K$ for which $a+bt$ is invertible.

Proof.

$$\begin{aligned}\frac{d}{dt}N(a+bt) &= \frac{d}{ds}N(a+bt+bs)\Big|_{s=0}\\ &= N(a+bt)\frac{d}{ds}N(1+(a+bt)^{-1}bs)\Big|_{s=0}\\ &= N(a+bt)\,\mathrm{Tr}((a+bt)^{-1}b).\end{aligned}$$

□

Lemma 6.4.4. *Suppose that $N = N_1^{k_1}N_2^{k_2}\cdots N_r^{k_r}$ where $N_1, N_2, \ldots, N_r$ are distinct irreducible polynomials, $k_1, k_2, \ldots, k_r$ are positive and $N_1(1) = N_2(1) = \cdots = N_r(1) = 1$. Then $N_1, N_2, \ldots, N_r$ are norms as well.*

Proof. We have

$$\begin{aligned}N_1^{k_1}(ab)N_2^{k_2}(ab)\cdots N_r^{k_r}(ab) &= N(ab)\\ &= N(a)N(b)\\ &= N_1^{k_1}(a)N_2^{k_2}(a)\cdots N_r^{k_r}(a)N_1^{k_1}(b)N_2^{k_2}(b)\cdots N_r^{k_r}(b)\end{aligned}$$

If the irreducible polynomial $N_i(a)$ divides $N_j(ab)$, then it also divides $N_j(a)$ by setting $b = 1$. It follows that $N_i(a)$ and $N_j(a)$ have to be the same up to a scalar. Since $N_i(1) = N_j(1)$ we have $N_i = N_j$ and $i = j$. So $N_i(a)$ and $N_i(b)$ divide $N_i(ab)$ and $N_i(ab) = \lambda N_i(a)N_i(b)$ with $\lambda \in K$. For $a = b = 1$ we get $\lambda = 1$. □

Theorem 6.4.5. *Suppose that $N_1, N_2 : R \to K$ are two norms, $a_1, a_2, \ldots, a_k \in R$ span R as a K-vector space and $N_1(1+ta_i) = N_2(1+ta_i)$ for all i and all t. Then we have $N_1 = N_2$.*

Proof. Without loss of generality we may assume that $N_1(a)$ and $N_2(a)$ do not have a common factor as a polynomial. If the characteristic of F is $p > 0$ then we can also assume that N_1 and N_2 are not p-th powers (because otherwise we replace N_1 and N_2 by their (unique) p-th roots). For $a, b \in R$ and $t \in F$ with $a + tb \in R^\times$ we have

$$\frac{d}{dt}\frac{N_1(a+tb)}{N_2(a+tb)} = \frac{N_1(a+tb)}{N_2(a+tb)}\Big(\mathrm{Tr}_1((a+tb)^{-1}b) - \mathrm{Tr}_2((a+tb)^{-1}b)\Big)$$

For $a = 1$, $b = a_i$ and $t = 0$ we get $\mathrm{Tr}_1(a_i) = \mathrm{Tr}_2(a_i)$ for all i. Because Tr_1 and Tr_2 are linear and $a_1, \ldots, a_k$ spans R, we get $\mathrm{Tr}_1 = \mathrm{Tr}_2$. It follows that $\frac{d}{dt} N_1(a+tb)/N_2(a+tb) = 0$. This implies that $N_1(a)/N_2(a)$ is constant, or $N_1(a)/N_2(a)$ is a p-th power of a rational function, where p is the characteristic. In the first case, $N_1(a)/N_2(a) = N_1(1)/N_2(1) = 1$ and $N_1 = N_2$ and we are done. In the second case, $N_1(a)$ and $N_2(a)$ both must be p-th powers because they do not have a common factor, but this contradicts our assumptions.

□

Remark 6.4.6. *For the previous theorem, it is not necessary that K is algebraically closed. Any norm $N : R \to K$ can be extended to a norm $\overline{N} : R \otimes_K \overline{K} \to \overline{K}$ where $\overline{K}$ is the algebraic closure of K.*

Remark 6.4.7. *For $A, B \in \mathrm{Mat}_{n,n}^m(K)$, we have $A \sim_{\mathrm{GL}_n} B$ if and only if $A \sim_{\mathrm{GL}_n(\overline{K})} B$. This is because $S(n,m) \otimes_K \overline{K}$ is the invariant ring for the action of $\mathrm{GL}_n(\overline{K})$ on $\mathrm{Mat}_{n,n}^m(\overline{K})$.*

Theorem 6.4.8. *Let $A, B \in \mathrm{Mat}_{n,n}^m$ with $A = (A_1, \ldots, A_m)$ and $B = (B_1, \ldots, B_m)$. Define*

$$C_i = \begin{pmatrix} A_i & 0 \\ 0 & B_i \end{pmatrix}$$

for all i. Let $\mathcal{C}$ be the algebra generated by $C_1, C_2, \ldots, C_m$. Let $Z_1, Z_2, \ldots, Z_s$ be the pivot basis of $\mathcal{C}$ and write

$$Z_j = \begin{pmatrix} X_j & 0 \\ 0 & Y_j \end{pmatrix}$$

for all j. Then $A \sim_{\mathrm{GL}_n} B$ if and only if $\det(I + tX_j) = \det(I + tY_j)$ as a polynomial in t for all j.

Proof. By the above two remarks, we can assume K is algebraically closed. We define two norms on $\mathcal{C}$, by

$$N_1\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \det(A), \quad N_2\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \det(B).$$

Suppose for every j, we have

$$N_1(I + tZ_j) = \det(I + tX_j) = \det(I + tY_j) = N_2(I + tZ_j).$$

Since $Z_1, \ldots, Z_s$ spans $\mathcal{C}$, we have $N_1 = N_2$. This means that $\det(I + tA) = N_1(I + tZ) =$

$N_2(1+tZ) = \det(I+tB)$ for every

$$Z = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \in \mathcal{C}.$$

In particular, if w is a word, we have

$$C_w = \begin{pmatrix} A_w & 0 \\ 0 & B_w \end{pmatrix} \in \mathcal{C},$$

so

$$\det(I + tA_w) = \det(1 + tB_w)$$

for all words $w \in \text{words}([m])$. Hence $\sigma_j(w)(A) = \sigma_j(w)(B)$ for all $1 \leq j \leq n$ and $w \in \text{words}([m])$. By Theorem 2.5.5, these are a set of generating invariants, and hence $A \sim_{\text{GL}_n} B$.

For the other direction, suppose $\det(I + tX_j) \neq \det(I + Y_j)$ for some j. Since $Z_1, \ldots, Z_s$ is a pivot basis, $Z_j = C_w$ for some $w \in \text{words}([m])$. So, we have $X_j = A_w$ and $Y_j = B_w$, and $\det(I + tA_w) \neq \det(I + tB_w)$. In particular, $\sigma_j(w)(A) \neq \sigma_j(w)(B)$ for some j. For this j, $\sigma_j(w)$ is an invariant that separates A and B, and hence $A \not\sim_{\text{GL}_n} B$. □

Theorem 6.4.9. *There is a polynomial time algorithm for the orbit closure problem with witness for matrix invariants.*

Proof. Given $A, B \in \text{Mat}_{n,n}^m$, let $C_i = \begin{pmatrix} A_i & 0 \\ 0 & B_i \end{pmatrix}$. Let $\mathcal{C}$ be the subalgebra generated by $C_1, \ldots, C_m$. Construct the pivot basis $Z_1, \ldots, Z_s$ of $\mathcal{C}$. For each $Z_i = \begin{pmatrix} A_i & 0 \\ 0 & B_i \end{pmatrix}$, we need to check if $\det(I+tA_i) = \det(I+tB_i)$ as a polynomial in t for each i. But this can be done efficiently, as one only needs to check whether $\det(I + kA_i) = \det(I + kB_i)$ for $k \in S$, where S is a predetermined finite subset of size $n+1$ in K.

When $A \not\sim_{\text{GL}_n} B$, the algorithm finds $k \in S$ and $w \in \text{words}([m])$ such that $\det(I + kA_w) \neq \det(I + kB_w)$. This means $\det(I + kX_w) \in S(n, m)$ is an invariant that separates A and B.

□

6.5 Bounds for separating invariants

6.5.1 Matrix invariants

We will prove the following:

Theorem 6.5.1. *We have* $\beta_{\mathrm{sep}}(S(n,m)) \leq 2n^2\sqrt{n}$. *If we assume* $\mathrm{char}(K) = 0$, *then we have* $\beta_{\mathrm{sep}}(S(n,m)) \leq 2n\sqrt{n}$.

The bound in characteristic 0 is especially interesting because there are quadratic lower bounds for the degree of generating invariants in this case, see [36]. This also improves the bound in [13] for the degree of invariants defining the null cone.

Proposition 6.5.2. *Suppose that* $w = w_1 w_2 \cdots w_d \in \mathrm{words}([m])$ *is a word of length* d, *and* w *has no subword of the form* u^n *(i.e., a subword that is repeated* $\geq n$ *times). Then the total number of subwords of* w *is at least* $\binom{d+1}{2}/n$.

Proof. Suppose that $w_{i+1}w_{i+2}\cdots w_k = w_{j+1}w_{j+2}\cdots w_l$ with $1 \leq j < i$. Then we have $w_{j+1}w_{j+2}\cdots w_i = w_{i+1}w_{i+2}\cdots w_{2i-j}$, $w_{i+1}w_{i+2}\cdots w_{2i-j} = w_{2i-j+1}w_{2i-j+2}\cdots w_{3i-2j}$ etc. The subword $w_{j+1}w_{j+2}\cdots w_i$ appears at least $\lfloor (k-j)/(i-j) \rfloor$ inside $w_{j+1}w_{j+1}\cdots w_k$. If $i \leq k/n$, then a subword appears at least $\lfloor (k-j)/(i-j) \rfloor = \lfloor (ni-j)/(i-j) \rfloor \geq \lfloor ki/i \rfloor = n$ consecutive times. This contradicts the assumption.

Let S be the set of all subwords of the form $w_{i+1}w_{i+2}\cdots w_k$ with $k \leq d$ and $i \leq k/n$. These words are all distinct. For every k with $0 \leq k \leq d$ there are $\lfloor k/n \rfloor + 1$ choices for i. So the total number of subwords is:

$$\sum_{k=0}^{d} \left\lfloor \frac{k}{n} \right\rfloor + 1 > \sum_{k=0}^{d} \frac{k}{n} = \frac{\binom{d+1}{2}}{n}.$$

□

Let $A, B \in \mathrm{Mat}_{n,n}^m$ with $A \not\sim_{\mathrm{GL}_n} B$ and write $C_i = \begin{pmatrix} A_i & 0 \\ 0 & B_i \end{pmatrix}$. Let $\mathcal{C} \subseteq \mathrm{Mat}_{2n,2n}$ be the subalgebra generated by $C_1, \ldots, C_m$. An argument similar to the proof of Lemma 6.3.5 gives us the following result.

Lemma 6.5.3. *Every subword of a pivot word is again a pivot word.*

Lemma 6.5.4. *For any word* $u \in \mathrm{words}([m])$, u^n *cannot be a pivot.*

Proof. If u^n is a pivot, then so is u^i for all $i < n$. However, the Cayley-Hamilton theorem tells us that $I, C_u, C_{u^2} = C_u^2, \ldots C_{u^n} = C_u^n$ are linearly dependent, which is a contradiction.

□

Corollary 6.5.5. *Any word w containing a subword of the form u^n cannot be a pivot.*

Since the upper right and lower left quadrants are always zero for every matrix in $\mathcal{C}$, the algebra $\mathcal{C}$ which is at most $2n^2$-dimensional,. Hence the number of pivots is at most n^2. .

Lemma 6.5.6. *Suppose the length of any pivot w is d, then we must have*

$$d \leq 2n\sqrt{n}.$$

Proof. For w to be a pivot, it cannot contain any subword of the form u^n by the above corollary. Hence by Proposition 6.5.2, there are at least $\binom{d+1}{2}$ subwords. All these subwords must be pivots by Lemma 6.5.3. But the number of pivots is at most $2n^2$, and so we must have $2n^2 \geq \binom{d+1}{2}/n \geq \frac{d^2}{2n}$. Hence, we have $d^2 \leq 4n^3$, and so we get $d \leq 2n\sqrt{n}$. □

Proof of Theorem 6.5.1. Given $A, B \in \mathrm{Mat}_{n,n}$ with $A \not\sim_{\mathrm{GL}_n} B$, we construct the set of pivots of $\mathcal{C}$. In characteristic 0, we must have $\mathrm{Tr}(A_w) \neq Tr(B_w)$ for some pivot w. This means there is an invariant of degree $= \deg(T_w) = l(w) \leq 2n\sqrt{n}$ that separates them.

In characteristic $p > 0$, we must have $\det(I + tA_w) \neq \det(I + tB_w)$ for some pivot w, and hence for some $1 \leq j \leq n$, $\sigma_j(w)(A) \neq \sigma_j(w)(B)$. This gives an invariant of degree $\leq 2n^2\sqrt{n}$ that separates them. □

6.5.2 Matrix semi-invariants

The reduction given in Section 6.2.2 is good enough for showing that the orbit closure problems for matrix invariants and matrix semi-invariants are in the same complexity class. In this section we give a stronger reduction with the aim of finding better bounds for the degree of separating invariants for matrix semi-invariants. This reduction can also be made algorithmic, and can replace the reduction in Section 6.2.2. However, we will only focus on obtaining bounds for separating invariants.

Theorem 6.5.7. *We have $\beta_{\mathrm{sep}}(R(n,m)) \leq n^2\beta_{\mathrm{sep}}(S(n,mn^2))$.*

Using the bounds on matrix invariants in Theorem 6.5.1, we get the bounds for matrix semi-invariants.

Corollary 6.5.8. *We have $\beta_{\mathrm{sep}}(R(n,m)) \leq 2n^4\sqrt{n}$. If we assume $\mathrm{char}(K) = 0$, then we have $\beta_{\mathrm{sep}}(R(n,m)) \leq 2n^3\sqrt{n}$.*

Let $T \in \mathrm{Mat}_{d,d}^m$ such that $f_T \neq 0$. For $X \in \mathrm{Mat}_{n,n}^m$, consider

$$L_T(X) = \sum_{k=1}^{m} T_k \otimes X_k = \begin{pmatrix} L_{1,1}(X) & \dots & L_{1,d}(X) \\ \vdots & \ddots & \vdots \\ L_{d,1}(X) & \dots & L_{d,d}(X) \end{pmatrix},$$

where $L_{i,j}(X)$ represents an $n \times n$ block. From the definition of Kronecker product of matrices, one can check that $L_{i,j}(X) = \sum_{k=1}^{m} (T_k)_{i,j} X_k$, i.e., a linear combination of the X_i. By definition $f_T(X) = \det(L_T(X))$. Let

$$M_T(X) = Ad(L_T(X)) = \begin{pmatrix} M_{1,1}(X) & \dots & M_{1,d}(X) \\ \vdots & \ddots & \vdots \\ M_{d,1}(X) & \dots & M_{d,d}(X) \end{pmatrix},$$

where $M_{i,j}(X)$ represents an $n \times n$ block. The entries of $M_T(X)$ are not linear in the entries of the matrices X_k. Instead the entries are $\deg(dn-1)$ polynomials in the entries $(X_k)_{i,j}$. For $X \in Mat_{n,n}^m$, let us define

$$X_{i,j,k} = X_k M_{i,j}(X),$$

for $1 \leq k \leq m$ and $1 \leq i, j \leq d$.

Consider the map $\zeta : \mathrm{Mat}_{n,n}^m \to \mathrm{Mat}_{n,n}^{md^2}$ given by $X \mapsto (X_{i,j,k})_{i,j,k}$. This gives a map on the coordinate rings $\zeta^* : K[\mathrm{Mat}_{n,n}^{md^2}] \to K[\mathrm{Mat}_{n,n}^m]$.

Lemma 6.5.9. *The map ζ^* descends to a map on invariant rings $\zeta^* : S(n, md^2) \to R(n, m)$.*

Proof. Observe that $L_{i,j}(X)$ is a linear expression in the matrices X_i. Hence for $g = (P, Q^{-1}) \in \mathrm{SL}_n \times \mathrm{SL}_n$, we have

$$L_{i,j}(g \cdot X) = L_{i,j}(PX_1Q, PX_2Q, \dots, PX_mQ) = PL_{i,j}(X)Q.$$

In particular, we see that $L_T(g \cdot X) = (P \otimes \mathrm{Id}) L_T(X) (Q \otimes \mathrm{Id})$. For any two square matrices

A, B of the same size, we have $Ad(AB) = Ad(B)Ad(A)$. Hence, we have

$$\begin{aligned} M_T(g \cdot X) &= Ad(L_T(g \cdot X)) \\ &= Ad((P \otimes \mathrm{Id})L_T(X)(Q \otimes \mathrm{Id})) \\ &= Ad(Q \otimes \mathrm{id})M_T(X)Ad(P \otimes \mathrm{Id}) \\ &= (Q^{-1} \otimes \mathrm{id})M_T(X)(P^{-1} \otimes \mathrm{Id}). \end{aligned}$$

The last equality follows from the fact that for a matrix whose determinant is equal to 1, the inverse and adjoint are the same. We deduce that $M_{i.j}(g \cdot X) = Q^{-1}M_{i,j}(X)P^{-1}$.

Hence, we have $(g \cdot X)_{i,j,k} = (PX_kQ)(Q^{-1}M_{i,j}(X)P^{-1}) = PX_{i,j,k}P^{-1}$. Now, it is clear that $\zeta^*(g) \in R(n,m)$ since $\zeta^*(g \cdot X) = g(\zeta(g \cdot X)) = g(P\zeta(X)P^{-1}) = g(\zeta(X)) = \zeta^*(g)(X)$. □

Corollary 6.5.10. *Suppose we have $g \in S(n, md^2)$ such that $g(\zeta(A)) \neq g(\zeta(B))$, then $A \not\sim_{LR} B$.*

Assume $A, B \in \mathrm{Mat}_{n,n}$ and assume $A \not\sim_{LR} B$. We want to show the existence of an invariant $f \in R(n,m)$ such that $f(A) \neq f(B)$ such that $\deg(f)$ is as small as possible. Indeed, since $A \not\sim_{LR} B$, there is a choice of $S \in \mathrm{Mat}_{k,k}^m$, for some $k \geq 1$, such that $f_S(A) \neq f_S(B)$. Without loss of generality, assume $f_S(B) \neq 0$. Hence $f_S(A)/f_S(B) \neq 1$. Once again we must have $f_S(A)^d/f_S(B)^d \neq 1$ for at least one choice of $d \in \{n-1, n\}$. In particular, for such a d, $(f_S)^d$ is an invariant of degree dkn that separates A and B.

Lemma 6.5.11. *Let $A, B \in \mathrm{Mat}_{n,n}^m$ and let $T \in \mathrm{Mat}_{d,d}^m$ such that $f_T(A) = f_T(B) \neq 0$. Then there exists $g \in S(n, md^2)$ such that $\zeta^*(g)(A) \neq \zeta^*(g)(B)$.*

Proof. We have a degree dkn invariant f such that $f(A) \neq f(B)$ by the above discussion. We can find $U \in \mathrm{Mat}_{dk,dk}^m$ such that $f_U(A) \neq f_U(B)$ since such invariants are a spanning set for invariants of degree dkn. Now for $X \in \mathrm{Mat}_{n,n}^m$, consider

$$N(X) = \left(\sum_{k=1}^{m} U_k \otimes X_k\right)(M_T(X) \otimes \mathrm{Id_k}).$$

We observe that $N(X)$ is a $dk \times dk$ block matrix, where each block has size $n \times n$ and is a linear combination of $X_{i,j,k}$. In other words, the $(p,q)^{th}$ block $N(X)_{p,q}$ is a linear combination $\sum_{i,j,k} \lambda_{p,q}^{i,j,k} X_{i,j,k}$ for some $\lambda_{p,q}^{i,j,k} \in K$. Now we can define an invariant $g \in S(n, md^2)$. For $Z = (Z_{i,j,k})_{i,j,k} \in \mathrm{Mat}_{n,n}^{md^2}$, we define N_Z to be the $dk \times dk$ block matrix, where the $(p,q)^{th}$ block is given by $\sum_{i,j,k} \lambda_{p,q}^{i,j,k} Z_{i,j,k}$. Let $g(Z) = \det(N_Z)$. It is easy to

check that $g \in S(n, md^2)$. Moreover, we have

$$\zeta^*(g)(X) = g(\zeta(X)) = \det(N_{\zeta(X)}) = \det(N(X)) = f_U(X)\det(M_T(X))^k.$$

Since $f_T(A) = f_T(B) \neq 0$, we have that $\det(M_T(A)) = \det(M_T(B)) \neq 0$. In particular, since $f_U(A) \neq f_U(B)$, we have $\zeta^*(g)(A) \neq \zeta^*(g)(B)$ as required.

□

Now, we can finally prove Theorem 6.5.7.

Proof of Theorem 6.5.7. Suppose $A, B \in \mathrm{Mat}_{n,n}^m$ with $A \not\sim_{LR} B$. Both A and B cannot be in the null cone. If A is in the null cone, then we have an invariant f with $\deg(f) = n(n-1)$ such that $f(B) \neq 0 = f(A)$. Similarly if B is in the null cone.

Now, let us assume both A and B are not in the null cone. By the above discussion, for at least one choice of $d \in \{n-1, n\}$, we either have $T \in \mathrm{Mat}_{d,d}^m$ such that $f_T(A) \neq f_T(B)$ or we have an invariant of the form $f = \zeta^*(g)$ such that $f(A) \neq f(B)$. In the former case, we have an invariant of degree $nd \leq n^2$ that separates A and B. The latter implies that $\zeta(A) \not\sim_{\mathrm{GL}_n} \zeta(B)$. Hence, we have an invariant $g \in S(n, md^2)$ of degree $\leq \beta_{\mathrm{sep}}(S(n, md^2))$ such that $g(\zeta(A)) \neq g(\zeta(B))$.

Now, since ζ is a map of degree dn, we have $\zeta^*(g) \in R(n, m)$ is a polynomial of degree $= \deg(g)dn \leq n^2\beta_{\mathrm{sep}}(S(n, md^2)) \leq n^2\beta_{\mathrm{sep}}(S(n, mn^2))$ that separates A and B. □

CHAPTER 7

Computational complexity

In Section 7.1, we introduce non-commutative circuits, and in Section 7.2, we discuss applications of our results to rational identity testing and other problems on non-commutative circuits.

7.1 Non-commutative circuits

Hrubes and Wigderson in their paper titled "Non-commutative arithmetic circuits with division" ([46]) raised four open problems. As a consequence of the results in this dissertation, we settle two of the open problems, and make partial progress on the other two. Let us first describe (non-commutative) arithmetic circuits, and we start with an example.

The circuit shown below is a simple circuit, which computes the expression $ab^{-1} + c$. There are three input gates a, b and c, a muliplication gate (labelled $\times$), an addition gate (labelled $+$), and an inverse gate (labelled $\square^{-1}$). Further, for the multiplication gate, the two incoming arrows are labelled L and R to tell us the order in which to multiply the inputs. The size of the circuit is taken as the number of gates, which in the above example is 6.

We now give a formal definition. A non-commutative arithmetic circuit Φ over a field K is a finite directed acyclic graph as follows. Nodes (or gates) of in–degree zero are

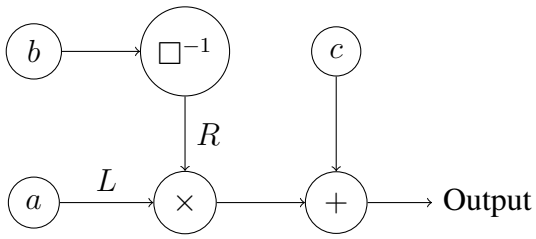

Figure 7.1: Example of a circuit

labelled by either a variable or a field element in K. All other nodes have in–degree one or two. The gates of in–degree one are labelled by $\square^{-1}$, and the gates of degree two by either $+$ or $\times$. The edges going into a gate labelled by $\times$ are labelled L and R, to determine the order of multiplication. The gates are called input, inverse, addition and multiplication gates. The nodes of out–degree zero are called output gates.

The size of a circuit Φ is the number of gates. A formula is circuit where every node has out–degree at most one. A gate in a circuit Φ in variables $x_1, \ldots, x_n$ will compute (in the obvious way) a non-commutative rational expression, i.e., an element of the free skew-field $K\langle\!\langle x_1, \ldots, x_n \rangle\!\rangle$. For a gate $u \in \Phi$, we will denote the non-commutative rational expression it computes by $\widehat{u}$. However, it may happen that the gate u is an inverse gate, and the input to this gate is zero. In this case, we say that $\widehat{u}$ is undefined.

A circuit Φ is called a correct circuit if $\widehat{u}$ is defined for every gate $u \in \Phi$. A correct circuit computes a set of non-commutative rational expressions, one for each output gate.

One can try to evaluate a non-commutative arithmetic circuit (with single output gate) by specializing the variables to elements in any (non-commutative) K-algebra R. Let Φ be a circuit in variables $x_1, \ldots, x_n$. We can define a partial map

$$\Phi^R : R^n \to R.$$

Indeed, for $(a_1, \ldots, a_n) \in R^n$, we try to evaluate the circuit. We may come across an inverse gate whose input is undefined, in which case we say $\Phi^R(a_1, \ldots, a_n)$ is undefined.

7.2 Rational identity testing

We will consider rational expressions as given by formulas. We want to address the fundamental question: How can one decide if two rational expressions define the same rational function? This is equivalent to deciding whether a single rational expression defines the zero function. To decide whether a rational expression defines the zero function is the same as deciding whether the inverse of the rational expression exists. This is the same as adding an inverse gate to the output, and asking if the formula is correct.

Problem 7.2.1 (Rational Identity Testing)**.** *How do we decide if a given formula is correct?*

As we mentioned in the introduction, we can test this on matrices.

Proposition 7.2.2. *A formula Φ in variables $x_1, \ldots, x_m$ is correct if and only if the evaluation $\Phi(A_1, \ldots, A_m)$ is defined for some $A_1, \ldots, A_m \in \mathrm{Mat}_{d,d}^m$ for some $d \in \mathbb{Z}_{\geq 0}$.*

The above Proposition is really just a consequence of the construction of the free skew-field. However, this raises an important question.

Problem 7.2.3. *Given a formula Φ in variables $x_1, \ldots, x_m$ let $d(\Phi)$ be the smallest integer such that $\Phi(A)$ is defined for some $A = (A_1, \ldots, A_m) \in \mathrm{Mat}_{d,d}^m$. How do we compute $d(\Phi)$?*

It is of course too ambitious in general to try and find the smallest such integer, and so instead we might ask for a bound. Another thing we might ask for, is whether there is a bound that depends only upon the size of the circuit. It is conceivable that even if we consider the formulas of size n, for each integer D, there might be a correct formula Φ_D such that $d(\Phi) > D$.

Problem 7.2.4. *Is $d(n) := \max\{d(\Phi) \mid \mathrm{size}(\Phi) = n\} < \infty$.*

While both these problems have very good answers now, let us first connect these problems to the null cone for matrix semi-invariants. For this we need a construction of Hrubeš and Wigderson that we mentioned previously in Section 4.3.2. We will demonstrate this with an example.

Consider the rational expression $\Phi = (ba - ab)^{-1}$. We will look at the linear matrix $L_\Phi = \begin{pmatrix} 0 & 1 & a \\ -1 & 0 & b \\ -a & -b & 0 \end{pmatrix}$

Consider the sequence of row and column transformations

1. $R_3 \to R_3 + bR_1$;
2. $R_2 \to R_2 - aR_1$;
3. $C_3 \to C_3 + bC_1$;
4. $C_3 \to C_3 - aC_2$.

This transforms the linear matrix L_Φ to $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & ba - ab \end{pmatrix}$. Row and column transformations do not change the rank. Hence, $\Phi(A, B)$ is correct if and only if $L_\Phi(A, B)$ has full rank. But now we see that the linear matrix $L_\Phi = X_0 + aX_1 + bX_2$. There exist matrices $d \times d$ matrices A, B such that $L_\Phi(A, B)$ is invertible if and only if there is an invariant $f \in R(3, 3)_{3d}$ such that $f(X) \neq 0$ for the 3-tuple $X = (X_0, X_1, X_2) \in \mathrm{Mat}_{3,3}^3$. This follows from Proposition 5.1.1 and Lemma 4.2.6.

We have not explained how one gets the matrix L_Φ from a formula Φ. There is a procedure to do this in [46], where they spell out in detail how to deal with each kind of gate. The most important part of the construction is of course that size of L_Φ is at most twice the size of Φ. We can assume that size of L_Φ is exactly twice the size of Φ by padding with some diagonal 1's if necessary.

This connection alone tells us that there is a positive answer to Problem 7.2.4 since we know by general results on invariant theory that bounds for invariants defining the null cone exist. However, given the powerful results we have on the null cone for matrix semi-invariants, we get much stronger results.

Theorem 7.2.5. *We have* $d(n) \leq 2n - 1$.

We will not develop the connections more here, referring the interested reader instead to [46], in particular their appendix on invariant theory. We will simply list the consequences that our results on linear matrices and matrix semi-invariants have.

Rational identity testing

Deciding whether a non-commutative formula computes the zero function is called the rational identity testing problem. Hrubeš and Wigderson give a randomized algorithm for rational identity testing whose run time is polynomial in n and $d(n)$. See [46, Section 7] for the details. Thus the above bound on $d(n)$ gives a polynomial time randomized algorithm for rational identity testing for infinite fields in arbitrary characteristic.

If $K = \mathbb{Q}$, Gurvits' algorithm can decide the invertibility of a linear matrix if it satisfies a certain property. In [49], they showed that a polynomial bound for $d(n)$ (equivalently $\gamma(\mathrm{SL}_n \times \mathrm{SL}_n, \mathrm{Mat}_{n,n}^m)$) would suffice to extend Gurvits' algorithm to all linear matrices. Hence we have a deterministic polynomial time algorithm if $K = \mathbb{Q}$. Although, in [38], they showed that Gurvits' algorithm suffices independent of our bounds.

Remark 7.2.6. *Using the bounds on $d(n)$, Ivanyos, Qiao and Subrahmanyam give a different deterministic polynomial time algorithm that works for any sufficiently large field, see [50].*

Eliminating inverse gates

Let f be a non-commutative polynomial in $K\langle t_1, t_2, \ldots, t_m\rangle$ of degree k, which can be computed by a formula of size n. Then f can be computed by a formula of size $n^{O(\log^2(k)\log(n))}$ without inverse gates. (see [46, Corollary 8.4]).

Lower bounds on formula size

Problem 1 in [46] asks for an explicit family of non-commutative polynomials which cannot be computed by a polynomial size formula with divisions. We give an answer to this problem. In [68], it was proved that any formula without divisions computing the non-commutative determinant (or permanent) of degree k must have size $2^{\Omega(k)}$. To find the size of a formula that allows divisions, we use our bound for eliminating inverse gates, and solve $2^{\Omega(k)} = n^{O(\log^2(k)\log(n))}$ for n. This shows that any formula with divisions computing the non-commutative determinant (or permanent) of degree k has size $2^{\Omega(\sqrt{k}/\log(k))}$.

CHAPTER 8

Tensor rank

We introduce the notions of tensor and border rank in Section 8.1. In Section 8.2, we recall Strassen's equations for showing lower bounds on border rank, and strengthen the result by reformulating it in terms of linear matrices. We also give an application of this method to compute the 3×3 determinant and permanent tensors. In Section 8.3, we describe the general method of finding lower bounds by using flattenings arising from linear matrices. Finally in Section 8.4, we provide explicit tensors in $K^d \otimes K^d \otimes K^d$ of border rank at least $2d - 3$ when d is odd, and $2d - 2$ when d is even.

8.1 Introduction

Over the last decade, tensors have received a lot of attention as a consequence of its wide ranging applications in mathematics as well as other scientific disciplines. We refer to [59] for several open conjectures in the subject, as well as a detailed introduction to the subject. The subject begins with the concept of tensor rank which is a generalization of matrix rank.

Definition 8.1.1. *For a tensor $T \in K^{a_1} \otimes K^{a_2} \otimes \cdots \otimes K^{a_l}$, we define its tensor rank $\mathrm{trk}(T)$ to be the smallest integer m such that T can be written as a sum of m pure tensors.*

Let Z_m denote the set of tensors of rank $\leq m$. The set Z_m need not be Zariski closed, and we consider its Zariski closure $\overline{Z}_m$. This gives rise to the definition of border rank.

Definition 8.1.2. *For a tensor T, we define its border rank $\mathrm{brk}(T)$ to be the smallest integer m such that $T \in \overline{Z}_m$.*

It is only natural to try and understand the polynomials that define the closed subset $\overline{Z}_m$. If f is a polynomial that vanishes on $\overline{Z}_m$ (or even Z_m), then if $f(T) \neq 0$ for some tensor, we immediately know that $\mathrm{brk(T)} > \mathrm{m}$. In other words, f can be used a test to prove that a tensor has border rank $> m$.

8.2 Strassen's equations

We recall the well known result of Strassen for showing lower bounds on the border rank.

Theorem 8.2.1 (Strassen)**.** *Let $T \in K^3 \otimes K^m \otimes K^m$. We write $T = e_1 \otimes A + e_2 \otimes B + e_3 \otimes C$. Identifying $K^m \otimes K^m$ with $\mathrm{Mat}_{m,m}$, we think of A, B and C as $m \times m$ matrices. Suppose A is invertible, then we have*

$$\mathrm{brk}(T) \geq m + \frac{1}{2}\,\mathrm{rk}(BA^{-1}C - CA^{-1}B).$$

We can view Strassen's equations from the perspective of linear matrices and linear subspaces of matrices. Let

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \text{ and } X_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix},$$

Let $L : K^3 \to \mathrm{Mat}_{3,3}$ be the map given by $e_i \mapsto X_i$, and consider the corresponding linear matrix $A = t_1X_1 + t_2X_2 + t_3X_3$. Identifying $K^m \otimes K^m$ with $\mathrm{Mat}_{m,m}$, we consider the map

$$\phi : K^3 \otimes K^m \otimes K^m = K^3 \otimes \mathrm{Mat}_{m,m} \to \mathrm{Mat}_{3,3} \otimes \mathrm{Mat}_{m,m} \to \mathrm{Mat}_{3m,3m},$$

where the first map is given by $L \otimes \mathrm{id}$ and the second map is given by taking the Kronecker product of matrices. Now, observe that the image of any rank 1 tensor under ϕ has rank at most 2, since $\mathrm{crk}(A) = 2$. Since the map is additive, the image of any rank r tensor will be at most $2r$. Thus the $2r + 1 \times 2r + 1$ minors of the image are polynomials that vanish on $\mathcal{Z}_r$, and hence on $\overline{\mathcal{Z}}_r$. In other words, we have the inequality

$$\mathrm{brk}(T) \geq \frac{\mathrm{rk}(\phi(T))}{2}. \tag{8.1}$$

Let us look at the map more carefully. We have

$$\phi(e_1 \otimes A + e_2 \otimes B + e_3 \otimes C) = \begin{pmatrix} 0 & A & B \\ -A & 0 & C \\ -B & -C & 0 \end{pmatrix}.$$

But now, consider the sequence of (block) row and column transformations.

$$\begin{pmatrix} 0 & A & B \\ -A & 0 & C \\ -B & -C & 0 \end{pmatrix} \xrightarrow[R_1 \mapsto A^{-1}R_1]{R_2 \mapsto A^{-1}R_2} \begin{pmatrix} 0 & I & A^{-1}B \\ -I & 0 & A^{-1}C \\ -B & -C & 0 \end{pmatrix}$$
$$\xrightarrow[R_3 \mapsto R_3 + CR_1 - BR_2]{C_3 \mapsto C_3 - C_2(A^{-1}B) + C_1(A^{-1}C)} \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & CA^{-1}B - BA^{-1}C \end{pmatrix}.$$

Hence, Equation 8.1 turns into

$$\operatorname{brk}(T) \geq \frac{\operatorname{rk}(\phi(T))}{2} = \frac{2m + \frac{1}{2}\operatorname{rk}(BA^{-1}C - CA^{-1}B)}{2} = m + \frac{1}{2}\operatorname{rk}(BA^{-1}C - CA^{-1}B).$$

This point of view yields a generalization of Strassen's result to the case even when A is not invertible.

Proposition 8.2.2. *Let $T \in K^3 \otimes K^m \otimes K^m$. We write $T = e_1 \otimes A + e_2 \otimes B + e_3 \otimes C$. Identifying $K^m \otimes K^m$ with $\operatorname{Mat}_{m,m}$, we think of A, B and C as $m \times m$ matrices. Then we have*

$$\operatorname{brk}(T) \geq \frac{1}{2}\operatorname{rk}\begin{pmatrix} 0 & A & B \\ -A & 0 & C \\ -B & -C & 0 \end{pmatrix}.$$

8.2.1 Application to 3×3 determinant and permanent tensors

We illustrate the method described above to compute the border rank and tensor rank for the 3×3 determinant and permanent tensors. The 3×3 determinant tensor is

$$\det_3 = \sum_{\sigma \in \Sigma_3} \operatorname{sgn}(\sigma) \mathrm{e}_{\sigma(1)} \otimes \mathrm{e}_{\sigma(2)} \otimes \mathrm{e}_{\sigma(3)},$$

where Σ_3 denotes the symmetric group in 3 letters. The 3×3 permanent tensor is

$$\operatorname{perm}_3 = \sum_{\sigma \in \Sigma_3} \mathrm{e}_{\sigma(1)} \otimes \mathrm{e}_{\sigma(2)} \otimes \mathrm{e}_{\sigma(3)}.$$

Lemma 8.2.3. *We have $\operatorname{brk}(\det_3) \geq 5$ if $\operatorname{char} K \neq 2$.*

Proof. The matrix $\phi(\det_3)$ is an explicit 9×9 matrix, which can be checked to be invertible

if $\mathrm{char}\, K \neq 2$. We write the matrix explicitly.

$$\phi(\det_3) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{-1} \\ 0 & 0 & 0 & 0 & 0 & \textcircled{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{1} & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & \textcircled{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \textcircled{1} & 0 & 0 & 0 & 0 & 0 \\ \textcircled{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

This matrix contains only 12 nonzero entries of the form ± 1. Six of these entries (circled) are in a column or a row with no other nonzero entry, reducing our computation to a 3×3 minor. It is easy to see that this minor is of full rank if $\mathrm{char}\, K \neq 2$ (and drops rank by 1 if $\mathrm{char}\, K = 2$).

Hence, by Proposition 8.2.2, we have $\mathrm{brk}(\det_3) \geq \frac{9}{2} = 4.5$. Since the border rank must be an integer, it must be at least 5.

□

On the other hand, there is an explicit decomposition of $\det_3$ as a sum of 5 simple tensors if $\mathrm{char}\, K \neq 2$, see [9].

Corollary 8.2.4. *Assume* $\mathrm{char}\, K \neq 2$. *Then we have* $\mathrm{trk}(\det_3) = \mathrm{brk}(\det_3) = 5$.

Lemma 8.2.5. *We have* $\mathrm{brk}(\mathrm{perm}_3) \geq 4$.

Proof. A similar computation shows that rank of $\phi(\mathrm{perm}_3)$ is 8. Hence we have

$$\mathrm{brk}(\mathrm{perm}_3) \geq \frac{8}{2} = 4.$$

□

Once again, if $\mathrm{char}\, K \neq 2$, there is an explicit decomposition of perm_3 as a sum of 4 simple tensors due to Glynn, see [40].

Corollary 8.2.6. *Assume* $\mathrm{char}\, K \neq 2$. *Then we have* $\mathrm{brk}(\mathrm{perm}_3) = \mathrm{trk}(\mathrm{perm}_3) = 4$.

In characteristic 0, the tensor rank of $\det_3$ and perm_3 were shown to be 5 and 4 respectively in [47]. While the arguments for bounding the tensor rank from above are still the

same (i.e., explicit decompositions), the arguments for bounding the tensor rank from below are more complicated. Their approach is to analyze certain Fano schemes parametrizing linear subspaces contained in the hypersurfaces $\det_3 = 0$ and $\mathrm{perm}_3 = 0$, and even involves a computation done with the help of a computer. The aforementioned method is far more elementary, and works as long as characteristic is not two.

8.3 Flattenings

We consider tensor product spaces with three tensor factors. Given a tensor in $T \in K^a \otimes K^b \otimes K^c$, we can write $T = \sum_i s_i \otimes X_i$, with $s_i \in K^a$ and $X_i \in K^b \otimes K^c$. Let $L : K^a \to \mathrm{Mat}_{p,q}$ be a linear map, and denote the image by $\mathcal{X}_L$. We identify $K^b \otimes K^c$ with $\mathrm{Mat}_{b,c}$ and identify $\mathrm{Mat}_{p,q} \otimes \mathrm{Mat}_{b,c}$ with $\mathrm{Mat}_{pb,qc}$. This gives the following map.

$$\begin{aligned} \psi_L : K^a \otimes K^b \otimes K^c &\longrightarrow \mathrm{Mat}_{pb,qc} \\ \sum_i s_i \otimes X_i &\longmapsto \sum_i L(s_i) \otimes X_i. \end{aligned}$$

Lemma 8.3.1. *For a tensor $T \in K^a \otimes K^b \otimes K^c$ we have $\mathrm{rk}(\psi_L(T)) \leq \mathrm{brk}(T)\,\mathrm{rk}(\mathcal{X}_L)$.*

Proof. Let $T = s \otimes b \otimes c$ be a tensor of rank 1. Then $\psi_L(T) = L(a) \otimes (b \otimes c)$, and hence $\mathrm{rk}(\psi_L(T)) \leq \mathrm{rk}(L(a)) \leq \mathrm{rk}(\mathcal{X}_L)$. Therefore, if we take a tensor $T \in K^a \otimes K^b \otimes K^c$ of rank r, then $\mathrm{rk}(\psi_L(T)) \leq r\,\mathrm{rk}(\mathcal{X}_L) =: D$. Observe that the $(D+1) \times (D+1)$ minors of $\psi_L(T)$ are polynomial equations that vanish all tensors of rank $\leq r$, i.e they vanish on Z_r. Hence these equations vanish on $\overline{Z_r}$ as well.

Hence if $\mathrm{brk}(T) = r$, we must have $\mathrm{rk}(\psi_L(T)) \leq D = r\,\mathrm{rk}(\mathcal{X}_L) = \mathrm{brk}(T)\,\mathrm{rk}(\mathcal{X}_L)$. □

Remark 8.3.2. *In particular, $\dfrac{\mathrm{rk}(\psi_L(T))}{\mathrm{rk}(\mathcal{X}_L)}$ is a lower bound for $\mathrm{brk}(T)$. Further, observe that $\psi_L(T) \in \mathcal{X}_L^{\{p,q\}}$, and hence $\mathrm{rk}(\psi_L(T)) \leq \mathrm{rk}(\mathcal{X}_L^{\{b,c\}})$. Hence in order to get a good lower bound, it would be useful for the blow-up to have large rank, which in turn is only possible if $\mathcal{X}_L$ has a large ratio of non-commutative rank to commutative rank.*

Corollary 8.3.3. *Let $D = r\,\mathrm{rk}(\mathcal{X}_L)$. Then the $(D+1) \times (D+1)$ minors of $\psi_L(T)$ give equations that are satisfied by all tensors of border rank $\leq r$.*

Landsberg's technique (see [61]) for obtaining lower bounds for border rank is the same as the one we describe above. For any r, the above corollary gives polynomials that are

satisfied by all tensors having border rank $\leq r$. It follows that if these polynomials do not vanish on a tensor T, then we must have $\underline{\mathrm{rk}}(T) > r$, providing a possible method for showing lower bounds for border rank. However, this method is only useful if these polynomials are non-trivial, i.e., not identically zero. The non-triviality of these equations essentially depends on the rank of the blow-up $\mathcal{X}_L^{\{b,c\}}$.

Lemma 8.3.4. *One of the $d \times d$ minors of $\psi_L(T)$ is a non-trivial polynomial if and only if $\mathrm{rk}(\mathcal{X}_L^{\{b,c\}}) \geq d$.*

Proof. Suppose $\mathrm{rk}(\mathcal{X}_L^{\{b,c\}}) \geq d$. Since $\mathrm{im}(\psi_L) = \mathcal{X}_L^{\{b,c\}}$, there exists $T_1 \in K^a \otimes K^b \otimes K^c$ such that $\mathrm{rk}(\psi_L(T_1)) = \mathrm{rk}(\mathcal{X}_L^{\{b,c\}}) \geq d$. Hence there is a $d \times d$ minor in $\psi_L(T_1)$ that is non-zero, and hence that $d \times d$ minor is a non-trivial polynomial.

The converse follows immediately since the underlying field K is infinite. □

8.4 Border rank of tensors in $K^d \otimes K^d \otimes K^d$

8.4.1 The case d is odd

Let $d = m = 2p + 1$ be a positive odd integer. Let $L : K^m \to \mathrm{Hom}(\bigwedge^p K^m, \bigwedge^{p+1} K^m)$ be the linear map defined in Theorem 4.5.1. For, this L, we define ψ_L as in the previous section, i.e.,

$$\begin{aligned} \psi_L : K^m \otimes K^m \otimes K^m &\longrightarrow \mathrm{Mat}_{\binom{2p+1}{p}m, \binom{2p+1}{p}m} \\ \sum_i s_i \otimes X_i &\mapsto \sum_i L(s_i) \otimes X_i. \end{aligned}$$

Theorem 8.4.1. *Let ψ_L be as above, and let $D = \binom{2p}{p}(2m-4)$. Then at least one of the $(D+1) \times (D+1)$ minors of ψ_L gives a non-trivial equation for tensors in $K^m \otimes K^m \otimes K^m$ of border rank $\leq 2m - 4$.*

Proof. Observe that $\mathrm{rk}(\mathcal{X}_L) = \binom{2p}{p}$ by Corollary 4.5.6. Hence, by Corollary 8.3.3 and Lemma 8.3.4, it suffices to show that $\mathrm{rk}(\mathcal{X}_L^{\{m\}}) \geq D + 1$.

By Proposition 4.5.8, we know that $\mathcal{X}_L^{\{p+1\}}$ has full rank and so we have $\mathcal{X}_L^{\{2p+2\}}$ has full rank as well. To find lower bounds on $\mathrm{rk}(\mathcal{X}_L^{\{2p+1\}})$, we use the properties from Corollary 4.4.5.

Let $M = \dim \bigwedge^p K^m = \dim \bigwedge^{p+1} K^m = \binom{2p+1}{p}$, and let $r(p,q) = \mathrm{rk}(\mathcal{X}^{\{p,q\}})$. Then we have $r(p+1, p+1) = (p+1)M$, and $r(2p+2, 2p+2) = (2p+2)M$ by the above

discussion. We have

$$r(p+1, 2p+1) \geq r(p+1, p+1) \geq (p+1)M.$$

Further, by concavity in the second variable, we have

$$\begin{aligned} r(2p+2, 2p+1) &\geq \frac{(2p+1)r(2p+2, 2p+2) + r(2p+2, 0)}{2p+2} \\ &\geq \frac{(2p+1)(2p+2)M}{2p+2} \\ &= (2p+1)M. \end{aligned}$$

Now, by concavity in the first variable, we have

$$\begin{aligned} r(2p+1, 2p+1) &\geq \frac{pr(2p+2, 2p+1) + r(p+1, 2p+1)}{p+1} \\ &\geq \frac{p(2p+1)M + (p+1)M}{p+1} \\ &= \frac{2p^2+2p+1}{p+1}M. \end{aligned}$$

Hence, we have

$$\begin{aligned} \frac{\mathrm{rk}(\mathcal{X}_L^{\{2p+1\}})}{\binom{2p}{p}} &\geq \frac{(2p^2+2p+1)\binom{2p+1}{p}}{(p+1)\binom{2p}{p}} \\ &= \frac{(2p^2+2p+1)(2p+1)}{(p+1)(p+1)} \\ &> 4p-2 \\ &= 2m-4 \end{aligned}$$

Thus $\mathrm{rk}(\mathcal{X}_L^{\{m\}}) > \binom{2p}{p}(2m-4)$ as required.

□

Recall that the non-commutative rank is at most twice the commutative rank. Hence

$$\frac{\mathrm{rk}(\mathcal{X}_L^{\{m\}})}{\mathrm{crk}(\mathcal{X}_L)} \leq \frac{m \cdot \mathrm{ncrk}(\mathcal{X}_L)}{\mathrm{crk}(\mathcal{X}_L)} < 2m.$$

This alone shows that there is very little room for improvement for the lower bounds we obtain using this method.

Remark 8.4.2. *For $m = 5$ i.e., $p = 2$, Landsberg shows that in fact $\mathcal{X}_L^{\{m\}}$ has full rank, thus giving non-trivial equations for tensors of border rank 8. Experimental evidence shows that in fact this is true for $p = 3$ and 4 as well, suggesting that it is perhaps true for all p, which would give non-trivial equations for tensors of border rank $2m - 2$.*

In $K^m \otimes K^m \otimes K^m$, Landsberg gives explicit tensors having border rank $\geq 2m - 2$ (resp. $2m - 4$) when m is even (resp. odd) (see [61]). For m odd, we can give explicit tensors whose border rank is $\geq 2m - 3$.

Let $m = 2p + 1$ be odd, and let $S_i \in \mathrm{Mat}_{p+1,p+1}$ be as in Proposition 4.5.8. For each r, consider $Q_r = S_r \oplus S_r \in \mathrm{Mat}_{2p+2,2p+2}$, and let $\widetilde{Q}_r \in \mathrm{Mat}_{2p+1,2p+1}$ be the matrix obtained from Q_r by removing the last column and last row of Q_r. Identifying $\mathrm{Mat}_{2p+1,2p+1}$ with $K^m \otimes K^m$, we can consider the tensor $T = \sum_{i=1}^{m} e_i \otimes \widetilde{Q}_i \in K^m \otimes K^m \otimes K^m$, where $e_1, e_2, \ldots, e_m$ is the standard basis for K^m.

Proposition 8.4.3. *The tensor $T = \sum_{i=1}^{m} e_i \otimes \widetilde{Q}_i \in K^m \otimes K^m \otimes K^m$ has border rank $\geq 2m - 3$.*

Proof. Let $L : K^m \to \mathrm{Hom}(\bigwedge^p K^m, \bigwedge^{p+1} K^m)$ be the linear map defined in Theorem 4.5.1. We have $\psi_L(T) = \sum_{i=1}^{m} L(e_i) \otimes \widetilde{Q}_i \in \mathrm{Mat}_{\binom{2p+1}{p}m,\binom{2p+1}{p}m}$. Observe that $A = \sum_{i=1}^{m} L(e_i) \otimes Q_i \in \mathrm{Mat}_{\binom{2p+1}{p}(m+1),\binom{2p+1}{p}(m+1)}$ has full rank. Observe that $\psi_L(T)$ is obtained by removing $\binom{2p+1}{p}$ columns and $\binom{2p+1}{p}$ rows from A. Hence, we have

$$\begin{aligned}
\mathrm{rk}(\psi_L(T)) &\geq \mathrm{rk}(A) - 2\binom{2p+1}{p} \\
&= (m+1)\binom{2p+1}{p} - 2\binom{2p+1}{p} \\
&= \binom{2p+1}{p}(2p).
\end{aligned}$$

□

Thus, we have

$$\begin{aligned}
\mathrm{brk}(T) &\geq \frac{\binom{2p+1}{p}(2p)}{\mathrm{rk}(\mathcal{X}_L)} \\
&= \frac{\binom{2p+1}{p}(2p)}{\binom{2p}{p}} \\
&> 2m - 4.
\end{aligned}$$

Hence $\mathrm{brk}(T) \geq 2m - 3$ as required.

8.4.2 The case d is even

In this case, we set $m = 2p + 1 = d - 1$. Let ψ_L be as in the previous section.

$\det(\psi_L)$ is a polynomial on $K^{d-1} \otimes K^d \otimes K^d$. Take any projection $\pi : K^d \to K^{d-1}$, and let $\phi = \pi \otimes \mathrm{id} \otimes \mathrm{id} : K^d \otimes K^d \otimes K^d \to K^{d-1} \otimes K^d \otimes K^d$. Let $f = \phi^*(\det \psi_L)$ be the pull back of the polynomial $\det(\psi_L)$ under ϕ.

Corollary 8.4.4. *The polynomial f is a non-trivial polynomial that vanishes on tensors of border rank $\leq 2d - 3$.*

Proof. The fact that f vanishes on tensors of border rank $\leq 2d-3$ is by a similar calculation as in the previous section. The polynomial f does not vanish on the tensor $T = \sum_{i=1}^{m} e_i \otimes (S_i \oplus S_i)$, and hence is non-trivial. □

Remark 8.4.5. *The tensor $T = \sum_{i=1}^{m} e_i \otimes (S_i \oplus S_i)$has border rank $\geq 2d - 2$.*

BIBLIOGRAPHY

[1] K. Akin, D. Buchsbaum and J. Weyman, *Schur functors and Schur complexes*, Adv. in Math. **44** (1982), 207–278.

[2] S. A. Amitsur, *The T-ideals of the free ring*, J. London Math. Soc. **30** (1955), 470–475.

[3] S. A. Amitsur, *On central division algebras*, Israel J. Math. **12** (1972), 408–420.

[4] S. A. Amitsur, *The generic division rings*, Israel J. Math. **17** (1974), 241–247.

[5] G. Bergman, *Rational relations and rational identities in division rings. I*, Journal of Algebra **43**, (1976), 252–266.

[6] G. Bergman, *Rational relations and rational identities in division rings. II*, Journal of Algebra **43**, (1976), 267–297.

[7] P. M. Cohn *Skew fields. Theory of general division rings*, Encyclopedia of Mathematics and its Applications **57**, Cambridge University Press, Cambridge, 1995.

[8] H. Derksen, *Polynomial bounds for rings of invariants*, Proc. Amer. Math. Soc. **129** (2001), no. 4, 955–963.

[9] H. Derksen, *On the nuclear norm and singular value decomposition of tensors*, Found. Comput. Math. **16** (2016), no. 3, 779–811.

[10] H. Derksen and G. Kemper, *Computational Invariant Theory.* Invariant Theory and Algebraic Transformation Groups. I. Encyclopaedia of Mathematical Sciences **130**, Springer-Verlag, 2002.

[11] H. Derksen and V. Makam, *On non-commutative rank and tensor rank*, Linear and Multilinear Algebra, published online (2017).

[12] H. Derksen and V. Makam, *Polynomial degree bounds for matrix semi-invariants*, Adv. Math. **310** (2017), 44–63.

[13] H. Derksen and V. Makam, *Generating invariant rings of quivers in arbitrary characteristic*, J. Algebra **489** (2017), 435–445.

[14] H. Derksen and V. Makam, Degree bounds for semi-invariant rings of quivers, J. Pure and Applied Algebra, published online (2017).

[15] H. Derksen and V. Makam, Explicit tensors of border rank at least $2d-2$ in $K^d \otimes K^d \otimes K^d$ in arbitrary characteristic, `arXiv:1709.06131` [math.RA], 2017.

[16] H. Derksen and V. Makam, Algorithms for orbit closure separation for invariants and semi-invariants of matrices, `arXiv: 1801.02043`, [math.RA], 2018.

[17] H. Derksen and J. Weyman, *Semi-invariants of quivers and saturation of Littlewood-Richardson co-efficients*, Journal of the American Math. Soc. **13** (2000), 467–479.

[18] H. Derksen and J. Weyman, *On the canonical decomposition of quiver representations*, Compositio Mathematica **133** (2002), 245–265.

[19] H. Derksen and J. Weyman, *On Littlewood-Richardson polynomials*, Journal of Algebra **255** (2002), 247–257.

[20] M. Domokos, *Poincaré series of semi-invariants of* 2×2 *matrices*, Linear Algebra and its Applications **310** (2000), 183–194.

[21] M. Domokos, *Relative invariants of* 3×3 *matrix triples*, Linear and Multilinear Algebra **47** (2000), 175–190.

[22] M. Domokos, *Finite generating system of matrix invariants*, Math. Pannon **13** (2002), 175–181.

[23] M. Domokos, *Matrix invariants and the failure of Weyl's theorem*, Polynomial identities and combinatorial methods (Pantelleria, 2001), Lecture Notes in Pure and Appl. Math. **235**, Dekker, New York (2003), 215–236.

[24] M. Domokos, S. G. Kuzmin and A. N. Zubkov, *Rings of matrix invariants in positive characteristic*, J. of Pure and Applied Algebra **176** (2002), 61–80.

[25] M. Domokos and A. N. Zubkov, *Semi-invariants of quivers as determinants*, Transformation groups **6** (2001), 9–24.

[26] S. Donkin, *Rational representations of algebraic groups: tensor products and filtrations*, Lecture Notes in Math., **1140**, Springer, 1985.

[27] S. Donkin, *Skew modules for reductive groups*, J. Algebra **113** (1988), 465–479.

[28] S. Donkin, *The normality of closures of conjugacy classes of matrices*, Inv. Math. **101** (1990), 717–736.

[29] S. Donkin, *Invariants of several matrices*, Inv. Math. **110** (1992), 389–401.

[30] S. Donkin, *On tilting modules for algebraic groups*, Math. Z. **212** (1993), 39–60.

[31] S. Donkin, *Polynomial invariants of representations of quivers*, Comment. Math. Helvetici. **69** (1994), 137–141.

[32] J. Draisma, G. Kemper and D. Wehlau, *Polarization of separating invariants*, Canad. J. Math. **60** (2008), 556–571.

[33] D. Eisenbud and J. Harris, *Vector spaces of matrices of low rank*, Adv. in Math. **70**, (1988), 135-155.

[34] H. Flanders, *On spaces of linear transformations with bounded rank*, J. London Math. Soc. **37** (1962), 10–16.

[35] M. A. Forbes and A. Shpilka, *Explicit Noether normalization for simultaneous conjugation*, Approximation, randomization, and combinatorial optimization, Lecture Notes in Comput. Sci., vol **8096**, 527–542.

[36] E. Formanek, *Generating the ring of matrix invariants*, in: F. M. J. van Oystaeyen, editor, *Ring Theory*, Lecture Notes in mathematics **1197**, Springer Berlin Heidelberg, 1986, 73–82.

[37] M. Fortin and C. Reutenauer, *Commutative/non-commutative rank of linear matrices and subspaces of matrices of low rank*, Sém. Lothar. Combin. 52:B52f, 2004.

[38] A. Garg, L. Gurvits, R. Oliveira and A. Widgerson, *A deterministic polynomial time algorithm for non-commutative rational identity testing*, FOCS 2016, 109–117, IEEE Computer Soc., Los Alamitos, CA.

[39] L. Gurvits, *Classical complexity and quantum entanglement*, Journal of Computer and System Sciences **69** (2004), 448–484.

[40] D. G. Glynn, *The permanent of a square matrix*, European J. Combin. **31** (2010), no. 7, 1887-1891.

[41] W. Haboush, *Reductive groups are geometrically reductive*, Ann. of Math. **102** (1975), 67–85.

[42] M. Hashimoto, *Good filtrations of symmetric algebras and strong F-regularity of invariant subrings*, Math. Z. **236** (2001), 605-623.

[43] D. Hilbert, *Über die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), 473–534.

[44] D. Hilbert, *Über die vollen Invariantensysteme*, Math. Ann. **42** (1893), 313–370.

[45] M. Hochster and J. L. Roberts, *Rings of invariants of reductive groups acting on regular rings are Cohen-Macaulay*,Adv. in Math. **13** (1974), 115–175.

[46] P. Hrubeš and A. Wigderson, *Non-commutative arithmetic circuits with division*, ITCS'14, Princeton, NJ, USA, 2014.

[47] N. Ilten and Z. Teitler, *Product ranks of* 3×3 *determinant and permanent*, Canad. Math. Bull. **59** (2016), no. 2, 311-319.

[48] G. Ivanyos, M. Karpinski, Y. Qiao and M. Santha, *Generalized Wong sequences and their applications to Edmonds' problems*, J. Comput. System Sci. **81** (2015), 1373–1386.

[49] G. Ivanyos, Y. Qiao and K. V. Subrahmanyam, *Non-commutative Edmonds' problem and matrix semi-invariants* Computational Complexity **26**, (2017), no. 3, 717–763.

[50] G. Ivanyos, Y. Qiao and K. V. Subrahmanyam, *Constructive non-commutative rank computation in deterministic polynomial time over fields of arbitrary characteristics*, `arXiv:1512.03531` [cs.CC], 2016.

[51] N. Jacobson, PI-*algebras*, Lecture Notes in Mathematics, Vol. 441, Springer-Verlag, Berlin-New York, 1975.

[52] N. Jacobson, *Basic algebra. II*, Second Edition, W. H. Freeman and Company, New York, 1989.

[53] V. Kac, *Infinite root systems, representations of graphs and invariant theory. II*, Journal of Algebra **78** (1982), 141-162.

[54] G. Kempf, *The Hochster-Roberts theorem of invariant theory*, Michigan Math. Journal **26**, issue 1 (1979), 19–32.

[55] A. D. King, *Moduli of representations of finite-dimensional algebras*, Quart. J. Math. Oxford Ser. **45** (1994), no. 180, 515–530.

[56] H. Kraft and C. Procesi, *Classical Invariant Theory : A primer.* `http://www.unibas.math.ch`.

[57] F. Knop. *Uber die Glattheit von Quotientenabbildungen.* Manuscripta Math. 56 (4): 419–427, 1986.

[58] F. Knop. *Der Kanonische Modul eines Invariantenringes.* J. Algebra 127(1): 40–54, 1989.

[59] J. M. Landsberg, *Tensors: Geometry and Applications*, Graduate Studies in Mathematics **128**, American Math. Soc., Providence, RI, 2012.

[60] J. M. Landsberg, *New lower bounds for the rank of matrix multiplication*, SIAM Journal on Computing **43** (2014), 144–149.

[61] J. M. Landsberg, *Nontriviality of equations and explicit tensors in* $\mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m$ *of border rank at least* $2m - 2$, J. Pure Appl. Algebra **219** (2015), 3677–3684.

[62] V. Makam, *Hilbert series and degree bounds for matrix (semi)-invariants*, J. Algebra **454** (2016), 14–28.

[63] O. Mathieu, *Filtrations of G-modules*, Ann. Scient. Ec. Norm. Sup (2) **23** (1990), 625–644.

[64] K. Mulmuley, *Geometric Complexity Theory V: Equivalence between blackbox derandomization of polynomial identity testing and derandomization of Noether's normalization lemma*, `arXiv:1209.5993`.

[65] K Mulmuley and M. Sohoni, *Geometric Complexity Theory I: an approach to the P vs NP and related problems*, SIAM J. Comput. **31** (2001), 496–526.

[66] K. Mulmuley and M. Sohoni, *Geometric Complexity Theory II: towards explicit obstructions for embeddings among class varieties*, SIAM. J. Comput. **38** (2008), 1175–1206.

[67] M. Nagata, *Invariants of a group in an affine ring*, J. Math. Kyoto Univ. **3** (1963/1964), 369–377.

[68] N. Nisan, *Lower bounds for non-commutative computation*, In *Proceedings of the 23rd STOC* (1991), 410–418.

[69] E. C. Posner, *Prime rings satisfying a polynomial identity*, Proc. Amer. Math. Soc. **11** (1960), 180–184.

[70] C. Procesi, *The invariant theory of $n \times n$ matrices*, Adv. in Math. **19** (1976), 306–381.

[71] R. Raz and A. Shpilka, *Deterministic polynomial identity testing in non-commutative models*, Comput. Complexity **14** (2005), 1–19.

[72] Y. Razmyslov, *Trace identities of full matrix algebras over a field of characteristic zero*, Comm. in Alg. **8** (1980), Math. USSR Izv. **8** (1974), 727–760.

[73] A. Schofield and M. van den Bergh, *Semi-invariants of quivers for arbitrary dimension vectors*, Indag. Mathem., N.S **12** (2001), 125–138.

[74] A. Skowroński and J. Weyman, *The algebras of semi- invariants of quivers*, Transformation Groups **5** (2000), No. 4., 361– 402.

[75] A. N. Zubkov, *Matrix invariants of an infinite field of arbitrary characteristic*, Siberian Math. J. **34** (1993), 68–74.