Geometry and Complexity Theory

J.M. Landsberg
## Contents

Preface ix

Chapter 1. Introduction 1
   §1.1. Matrix multiplication 1
   §1.2. Separation of algebraic complexity classes 7
   §1.3. Supplementary material 11

Chapter 2. The complexity of Matrix multiplication I: lower bounds 15
   §2.1. Matrix multiplication and multi-linear algebra 15
   §2.2. Complexity of bilinear maps 17
   §2.3. Definitions and examples from representation theory 20
   §2.4. Definitions and examples from algebraic geometry 24
   §2.5. Strassen’s equations and generalizations 34
   §2.6. Lower bounds for the rank of matrix multiplication 43
   §2.7. Additional equations for $\sigma_r(Seg(PA \times PB \times PC))$ 46

Chapter 3. The complexity of matrix multiplication II: asymptotic upper bounds 57
   §3.1. Overview 57
   §3.2. The upper bounds of Bini, Capovani, Lotti, and Romani 59
   §3.3. Schönhage’s upper bounds 62
   §3.4. Strassen’s laser method 65
   §3.5. The Coppersmith-Winograd method 70
   §3.6. The Cohn-Umans program 72
Chapter 4. The complexity of Matrix multiplication III: algorithms and “practical” upper bounds 79

§4.1. Geometry of decompositions 80

§4.2. Strassen’s algorithm for $M_{2}$ 81

§4.3. Border rank algorithms for $M_{(2,2,n)}$ 86

§4.4. Alternating least squares approach to algorithms 89

§4.5. Algorithms for $3 \times 3$ matrix multiplication with 23 multiplications 89

§4.6. Smirnov’s algorithm showing the border rank of $M_{(3)}$ is at most 20 89

Chapter 5. Representation theory for complexity theory 91

§5.1. Double-Commutant Theorem and first consequences 91

§5.2. Representations of $S_d$ and $GL(V)$ 98

§5.3. Methods for determining equations of secant varieties of Segre varieties using representation theory 109

Chapter 6. Geometric Complexity Theory 113

§6.1. Circuits and the flagship conjecture of GCT 114

§6.2. A program to find modules in $I[Det_n]$ via representation theory 118

§6.3. Necessary conditions for modules of polynomials to be useful for GCT 124

§6.4. GCTV: Blackbox derandomization and the search for explicit linear spaces 128

§6.5. GCTVI: The needle exists 128

§6.6. Dual varieties and GCT 128

§6.7. Remarks on Kronecker coefficients and plethysm coefficients 138

§6.8. Symmetries of polynomials and coordinate rings of orbits 139

Chapter 7. Shallow circuits and the Chow variety 149

§7.1. Shallow circuits and geometry 150

§7.2. Hilbert functions and the method of shifted partial derivatives 154

§7.3. The Chow variety 154

Chapter 8. Advanced Topics 167

§8.1. Asymptotic surjectivity of the Hadamard-Howe map 167

§8.2. Non-normality of $Det_n$ 172

Hints and Answers to Selected Exercises 175
Contents

Bibliography  179
Index            187
Preface

The purpose of this book is to describe recent applications of algebraic geometry and representation theory to complexity theory. I focus on two central problems: the complexity of matrix multiplication and Geometric Complexity Theory (GCT).

I have attempted to make this book accessible to both computer scientists and geometers, and the exposition as self-contained as possible. The two main goals of this book are to convince computer scientists of the utility of techniques from algebraic geometry and representation theory, and to show geometers beautiful, interesting, and important questions arising in complexity theory.

Computer scientists have made extensive use of tools from mathematics such as combinatorics, graph theory, asymptotic estimates, and especially linear algebra. I hope to show that even elementary techniques from algebraic geometry and representation theory can substantially advance the search for lower, and even upper bounds in complexity theory. For questions such as lower bounds for the complexity of matrix multiplication and Valiant’s algebraic variants of \( \mathbf{P} \) v. \( \mathbf{NP} \), I believe this additional mathematics will be necessary for further advances, as there is strong evidence that these problems cannot be solved by linear algebra alone. I have attempted to make these techniques accessible, introducing them as needed to deal with concrete problems.

For geometers, I expect that complexity theory will be as good a source for questions in algebraic geometry as modern physics has been. Recent work has indicated that tools such as Fulton-McPherson intersection theory, Castelnuovo-Mumford regularity, and the Kempf-Weyman method for
computing minimal free resolutions all have something to add to complexity theory. In addition, complexity theory has a way of rejuvenating old questions that had been nearly forgotten but remain beautiful and intriguing: questions of Hadamard, Darboux, Luroth, and the classical Italian school. At the same time, complexity theory has brought different areas of mathematics together in new ways—combinatorics, representation theory, and algebraic geometry all play a role in understanding the coordinate ring of the orbit closure of the determinant.

This book evolved from several classes I have given on the subject: a spring 2013 semester course at Texas A&M, summer courses at: Scuola Matematica Inter-universitaria, Cortona (July 2012), CIRM, Trento (June 2014), and an IMA summer school at U. Chicago (July 2014), and most recently, and importantly, a fall 2014 semester course at UC Berkeley as part of the semester long program, Algorithms and Complexity in Algebraic Geometry, at the Simons Institute for the Theory of Computing.

**Overview.** Chapters 2,3,4 deal with the complexity of matrix multiplication. Chapter 2, in addition to introducing background material in algebraic geometry and representation theory, focuses on lower bounds. This is the first exposition that includes recent advances in the subject. Chapter 3 deals with asymptotic upper bounds. There are many excellent expositions of this material. The exposition in this book emphasizes geometric aspects of the advances and includes a detailed overview of the Cohn-Umans program, neither of which is available elsewhere. Chapter 4 provides a new geometric perspective on, and an exposition of, the existing algorithms for fast matrix multiplication.

Chapter 5 covers further topics in representation theory such as the algebraic Peter-Weyl theorem (which plays a central role in GCT) and includes additional details on the use of representation theory in the study of matrix multiplication.

Chapters 6,7,8 deal with the Geometric Complexity Theory (GCT) initiated in [MS01], as well as complexity results (such as reductions to shallow circuits) that may be approached from a GCT perspective. Chapter 6 is on the flagship permanent v. determinant conjecture of [MS01]. In addition to content from [MS08, Mula, Mulb], it covers direct geometric methods. Chapter 7 approaches the doors opened by [GKKS13b] towards paths to separate VP from VNP via depth three or restricted depth five circuits from a geometric perspective. These translate to questions about familiar varieties in algebraic geometry (secant varieties of the Chow variety, and of Veronese re-embeddings of secant varieties of Veronese varieties). More advanced tools from algebraic geometry are unavoidable when approaching GCT, and in Chapter 8, I present two (*possibly more*) results that require
a more advanced background; the asymptotic surjectivity of the Hadamard-
Howe map due to M. Brion, and the non-normality of the orbit closures of
the determinant and permanent, due to S. Kumar.

Prerequisites. I have attempted to limit prerequisites to a solid back-
ground in linear algebra, although such a reader would have to accept several
basic results in algebraic geometry without proof (e.g. Noether normaliza-
tion). In chapter 6 some further, but still elementary algebraic geometry is
needed, but nothing beyond [Sha94] is used. The exception is chapter 8,
where more advanced results are utilized.

Attendees of the 2014 UC Berkeley course included first year graduate
students in mathematics and more advanced graduate students in computer
science.

Acknowledgments. To be written later

Layout. To be written later

Dependency of chapters. To be written later
Chapter 1

Introduction

A dramatic leap in signal processing occurred in the 1960’s with the implementation of the fast Fourier transform, an algorithm that surprised the engineering community with its efficiency. How could one predict when such fast, perhaps non-intuitive, algorithms exist? Can we prove when they do not? Complexity theory addresses these questions.

This book is concerned with the use of geometry in attaining these goals. I focus on two central questions: the complexity of matrix multiplication, and algebraic variants of the famous \( P \) versus \( NP \) problem. In the first case, a surprising algorithm exists and it is conjectured that even more amazing algorithms exist. In the second case it is conjectured that no surprising algorithm exists.

1.1. Matrix multiplication

Much of scientific computation is linear algebra, and the basic operation of linear algebra is matrix multiplication. All operations of linear algebra, solving systems of linear equations, computing determinants etc., use matrix multiplication.

1.1.1. The standard algorithm. The standard algorithm for multiplying matrices is row-column multiplication: Let \( A, B \) be \( 2 \times 2 \) matrices

\[
A = \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1^1 & b_1^2 \\ b_2^1 & b_2^2 \end{pmatrix}.
\]

1To this day, it is not known if there is an even more efficient algorithm than the FFT. See [Val77, KLPSMN09, GHIL].
The usual algorithm to calculate the matrix product \( C = AB \) is

\[
\begin{align*}
    c_1^1 &= a_1^1 b_1^1 + a_2^1 b_2^1, \\
    c_2^1 &= a_1^1 b_2^1 + a_2^1 b_2^2, \\
    c_1^2 &= a_1^2 b_1^1 + a_2^2 b_1^2, \\
    c_2^2 &= a_1^2 b_2^1 + a_2^2 b_2^2.
\end{align*}
\]

It requires 8 multiplications and 4 additions to execute, and applied to \( n \times n \) matrices, it uses \( n^3 \) multiplications and \( n^3 - n^2 \) additions.

This algorithm has been around for a long time.

In 1968, V. Strassen set out to prove the standard algorithm was optimal in the sense that no algorithm using fewer multiplications exists – at least for two by two matrices – at least over \( \mathbb{Z}_2 \). His spectacular failure opened up a whole new area of research:

1.1.2. Strassen’s algorithm for multiplying \( 2 \times 2 \) matrices using seven scalar multiplications [Str69]. Set

\[
\begin{align*}
    I &= (a_1^1 + a_2^1)(b_1^1 + b_2^1), \\
    II &= (a_1^2 + a_2^2)b_1^1, \\
    III &= a_1^1(b_1^1 - b_2^2) \\
    IV &= a_2^2(-b_1^1 + b_1^2) \\
    V &= (a_1^1 + a_2^1)b_2^2 \\
    VI &= (-a_1^1 + a_2^1)(b_1^1 + b_1^2), \\
    VII &= (a_2^1 - a_2^2)(b_1^2 + b_2^2),
\end{align*}
\]

Exercise 1.1.2.1: Show that if \( C = AB \), then

\[
\begin{align*}
    c_1^1 &= I + IV - V + VII, \\
    c_2^1 &= II + IV, \\
    c_1^2 &= III + V, \\
    c_2^2 &= I + III - II + VI.
\end{align*}
\]

This raises questions:

1. Can one find an algorithm that uses just six multiplications?
2. Could Strassen’s algorithm have been predicted in advance?
3. Since it uses more additions, is it actually better in practice?
4. What about algorithms for \( n \times n \) matrices?

I address the last question first:
1.1.3. Fast multiplication of $n \times n$ matrices. In Strassen’s algorithm, the entries of the matrices need not be scalars - they could themselves be matrices. Let $A, B$ be $4 \times 4$ matrices, and write

$$A = \begin{pmatrix} a^1_1 & a^1_2 \\ a^2_1 & a^2_2 \end{pmatrix}, \quad B = \begin{pmatrix} b^1_1 & b^1_2 \\ b^2_1 & b^2_2 \end{pmatrix},$$

where $a^i_j, b^i_j$ are $2 \times 2$ matrices. One may apply Strassen’s algorithm to get the blocks of $C = AB$ in terms of the blocks of $A, B$ performing 7 multiplications of $2 \times 2$ matrices. Since one can apply Strassen’s algorithm to each block, one can multiply $4 \times 4$ matrices using $7^2 = 49$ multiplications instead of the usual $4^3 = 64$.

If $A, B$ are $2^k \times 2^k$ matrices, one may multiply them using $7^k$ multiplications instead of the usual $8^k$. If $n$ is not a power of two, enlarge the matrices with blocks of zeros to obtain matrices whose size is a power of two. Asymptotically, by recursion and block multiplication one can multiply $n \times n$ matrices using approximately $n^{\log_2(7)} \approx n^{2.81}$ arithmetic operations. To see this, let $n = 2^k$ and write $7^k = (2^k)^a$ so $k \log_2 7 = a k \log_2 2$ so $a = \log_2 7$.

1.1.4. Regarding the number of additions. The number of additions in Strassen’s algorithm also grows like $n^{2.81}$, so this algorithm is more efficient in practice when the matrices are large. For any efficient algorithm for matrix multiplication, the total complexity is governed by the number of multiplications, see [BCS97, Prop. 15.1]. This is fortuitous because there is a geometric object, tensor rank, that counts the number of multiplications in an optimal algorithm (within a factor of two), and thus provides us with a geometric measure of the complexity of matrix multiplication.

Just how large matrices one needs to obtain a substantial savings with Strassen’s algorithm (one needs matrices of size thousands by thousands) and other practical matters are addressed in [BB].

1.1.5. An even better algorithm? Regarding question (1) above, one cannot improve upon Strassen’s algorithm for $2 \times 2$ matrices. This was first shown in [Win71]. I will give a proof, using geometry and representation theory, of a stronger statement in §5.3.5. However for $n > 2$ very little is known, as is discussed below and in Chapters 2-4. It is known that better algorithms than Strassen’s exist for $n \times n$ matrices, even if they are not known explicitly.

1.1.6. How to predict in advance? The answer to question (2) is yes! In fact it could have been predicted 100 years ago.

Had someone asked Terracini in 1913, he would have been able to predict the existence of something like Strassen’s algorithm from geometric considerations alone. Matrix multiplication is a bilinear map (see §1.1.8).
Terracini would have been able to tell you (and we will see why in §2.4.14) that even a generic bilinear map $\mathbb{C}^4 \times \mathbb{C}^4 \rightarrow \mathbb{C}^4$ can be executed using seven multiplications and thus, fixing any $\epsilon > 0$, one can perform any bilinear map $\mathbb{C}^4 \times \mathbb{C}^4 \rightarrow \mathbb{C}^4$ within an error of $\epsilon$ using seven multiplications.

1.1.7. An astonishing conjecture. The following quantity is the standard measure of the complexity of matrix multiplication:

**Definition 1.1.7.1.** The exponent $\omega$ of matrix multiplication is

$$\omega := \inf \{ h \in \mathbb{R} \mid n \times n \text{ matrices may be multiplied using } O(n^h) \text{ arithmetic operations} \}$$

where inf denotes the infimum.

By Theorem 1.1.9.2 below, Strassen’s algorithm shows $\omega \leq \log_2(7) < 2.81$. Determining $\omega$ is a central open problem in complexity theory. The current “world record” is $\omega < 2.373$ [Wil, Gal, Sto]. This result and the substantial work leading up to it is the topic of Chapter 4.

This work has led to the following astounding conjecture:

**Conjecture 1.1.7.2.** $\omega = 2$.

That is, it is conjectured that asymptotically, it is nearly just as easy to multiply matrices as it is to add them!

Although I am unaware of anyone taking responsibility for the conjecture, all computer scientists I have discussed it with expect it to be true.

Since I have no opinion on whether the conjecture should be true or false, I discuss both upper and lower bounds for $\omega$, focusing on the role of geometry. I start with lower bounds:

1.1.8. Matrix multiplication as a bilinear map. I will use the notation $M_{(n,m,l)} : \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times l} \rightarrow \mathbb{C}^{n \times l}$ for matrix multiplication, writing $M_{(n)} = M_{(n,n,n)}$.

Matrix multiplication is a bilinear map, that is, for all $X, Y \in \mathbb{C}^{n \times m}$, $a_j, b_j \in \mathbb{C}$,

$$M_{(n,m,l)}(a_1 X_1 + a_2 X_2, Y) = a_1 M_{(n,m,l)}(X_1, Y) + a_2 M_{(n,m,l)}(X_2, Y), \quad \text{and}$$

$$M_{(n,m,l)}(X, b_1 Y_1 + b_2 Y_2) = b_1 M_{(n,m,l)}(X, Y_1) + b_2 M_{(n,m,l)}(X, Y_2).$$

The set of all bilinear maps $\mathbb{C}^a \times \mathbb{C}^b \rightarrow \mathbb{C}^c$ is a vector space. (In our case $a = nm, b = ml$, and $c = ln$.) Write $a_1, \ldots, a_a$ for a basis of $\mathbb{C}^a$ and similarly for $\mathbb{C}^b, \mathbb{C}^c$. Then $T : \mathbb{C}^a \times \mathbb{C}^b \rightarrow \mathbb{C}^c$ is uniquely determined by its action on basis vectors, $T(a_i, b_j) = \sum_{k=1}^c t_{ijk} c_k$. That is, the vector space of bilinear maps $\mathbb{C}^a \times \mathbb{C}^b \rightarrow \mathbb{C}^c$, which I will denote by $\mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^c$, has
1.1. Matrix multiplication

dimension abc and when I discuss polynomials on this space, I will mean polynomials on the coefficients $t^{ijk}$.

1.1.9. Tensor rank. Our measure of complexity for a bilinear map will be its tensor rank. Given an $a \times b$ matrix $X$, one can always change bases, i.e., multiply $X$ on the left by an invertible $a \times a$ matrix and on the right by an invertible $b \times b$ matrix, to obtain a matrix with some number of 1’s along the diagonal and zeros elsewhere. The number of 1’s appearing is called the rank of the matrix and is an invariant of the linear map $X$ determines. In other words, the only property of a linear map $\mathbb{C}^a \to \mathbb{C}^b$ that is invariant under changes of bases is its rank, and for each rank we have a normal form. This is not surprising because the dimension of the space of such linear maps is $ab$ and we have $a^2 + b^2$ worth of parameters of changes of bases we can make in a matrix representing the map.

For bilinear maps $\mathbb{C}^a \times \mathbb{C}^b \to \mathbb{C}^c$ we are not so lucky, as usually $abc > a^2 + b^2 + c^2$, i.e., there are fewer free parameters of changes of bases than the number of parameters needed to describe the map.

Nonetheless, there are properties of a bilinear map that will not change under a change of basis. The main property we will use is tensor rank. It is a generalization of the rank of a linear map. A tensor $T$ has tensor rank $r$ if it can be written as the sum of $r$ rank one tensors but no fewer, and a tensor has tensor rank one if in some coordinate system the multidimensional matrix representing it has exactly one nonzero entry. The set of bilinear maps in $\mathbb{C}^{a*} \otimes \mathbb{C}^{b*} \otimes \mathbb{C}^{c*}$ of tensor rank at most $r$ will be denoted $\hat{\sigma}^0_r$. For a bilinear map $T : \mathbb{C}^a \times \mathbb{C}^b \to \mathbb{C}^c$, write $R(T) \leq r$ if $T \in \hat{\sigma}^0_r$.

Remark 1.1.9.1. The peculiar notation $\hat{\sigma}^0_r$ will be explained in §2.4.7. To have an idea where it comes from for now: $\sigma_r = \sigma_r(\text{Seg}(\mathbb{P}^{a-1} \times \mathbb{P}^{b-1} \times \mathbb{P}^{c-1}))$ is standard notation in algebraic geometry for the $r$-th secant variety of the Segre variety, which is the object we will study. The hat denotes its cousin in affine space and the 0 indicates that we are now only considering an open subset of this set.

The following theorem shows that tensor rank is a legitimate measure of complexity:

**Theorem 1.1.9.2** (Strassen [Str69], see [BCS97] §15.1). $R(M(n)) = O(n^\omega)$.

Our goal is thus to determine, for a given $r$, whether or not matrix multiplication lies in $\hat{\sigma}^0_r$.

1.1.10. How to use algebraic geometry to prove lower bounds for the complexity of matrix multiplication? Algebraic geometry deals with the study of zero sets of polynomials. It may be used to prove both upper and lower complexity bounds. For lower bounds:
Plan to show $M_{\langle n,m,l \rangle} \notin \hat{\sigma}_r^0$ via algebraic geometry.

- Find a polynomial $P$ on the space of bilinear maps $\mathbb{C}^{nm} \times \mathbb{C}^{ml} \rightarrow \mathbb{C}^{nl}$ such that $P(T) = 0$ for all $T \in \hat{\sigma}_r^0$.
- Show that $P(M_{\langle n,m,l \rangle}) \neq 0$.

Chapters 2 and 5 discuss techniques for finding such polynomials, using algebraic geometry and representation theory, the study of symmetry in linear algebra.

### 1.1.11. Representation theory and matrix multiplication.

The study of polynomials on any vector space is facilitated by sorting the polynomials by degree. In our situation, our vector space of bilinear maps has additional structure (the groups of changes of bases in each of $\mathbb{C}^a$, $\mathbb{C}^b$, and $\mathbb{C}^c$ act) and we will exploit this to obtain a finer sorting of polynomials.

**1.1.12. How to use algebraic geometry to prove upper bounds for the complexity of matrix multiplication?** Based on the above discussion, one could try:

**Plan to show $M_{\langle n,m,l \rangle} \in \hat{\sigma}_r^0$ with algebraic geometry.**

- Find a set of polynomials $\{P_j\}$ on the space of bilinear maps $\mathbb{C}^{nm} \times \mathbb{C}^{ml} \rightarrow \mathbb{C}^{nl}$ such that $T \in \hat{\sigma}_r^0$ if and only if $P_j(T) = 0$ for all $j$.
- Show that $P_j(M_{\langle n,m,l \rangle}) = 0$ for all $j$.

*This plan has a problem:* Consider the set $S = \{(w,z) \in \mathbb{C}^2 \mid z = 0, w \neq 0\}$, whose real picture looks like the $z$-axis with the origin removed.

Any polynomial vanishing on $S$ will also vanish at the origin.

Just as in this example, the zero set of the polynomials vanishing on $\hat{\sigma}_r^0$ is larger than $\hat{\sigma}_r^0$ when $r > 1$ (see §2.2.2) so one cannot determine membership in $\hat{\sigma}_r^0$ via polynomials (although one can exclude membership with polynomials).

**Definition 1.1.12.1.** Given a subset $Z$ of a vector space $U$, define the *ideal* of $Z$, denoted $I(Z)$, to be the set of polynomials vanishing at all points of $Z$. Define the Zariski closure of $Z$, denoted $\overline{Z}$, to be the set of $u \in U$ such that $P(u) = 0$ for all $P \in I(Z)$. The common zero set of a collection of polynomials (such as $\overline{Z}$) is called an *algebraic variety*. 
In the example above, \( \overline{S} = \{(w, z) \in \mathbb{C}^2 | z = 0\} \).

When \( U = \mathbb{C}^{a*} \otimes \mathbb{C}^{b*} \otimes \mathbb{C}^c \), let \( \tilde{\sigma}_r := \sigma_r^0 \) denote the Zariski closure of the set of bilinear maps of tensor rank at most \( r \).

We will say \( T \) has border rank at most \( r \) if \( T \in \tilde{\sigma}_r \) and write \( R(T) \leq r \).

It is expected that \( R(M_{(n)}) < R(M_{(n)}) \), although for \( n = 2 \) they both equal seven. For \( n = 3 \) we only know \( 14 \leq R(M_{(3)}) \leq 20 \) and \( 19 \leq R(M_{(3)}) \leq 23 \). See Chapter 2 for the lower bounds and Chapter 4 for the upper bounds.

For the study of the exponent of matrix multiplication, we have good luck: the problem goes away.

**Theorem 1.1.12.2** (Bini [Bin80], see §3.2). \( R(M_{(n)}) = O(n^\omega) \).

That is, although we may have \( R(M_{(n)}) < R(M_{(n)}) \), they are not different enough to effect the exponent.

### 1.1.13. How to use geometry to find algorithms?

At first glance, geometry, which studies properties of objects that are independent of coordinates, seems ill-suited for developing explicit algorithms, which are given in coordinates. In fact, one uses geometry not to write down a single algorithm, but rather families of algorithms all sharing common geometric features. A highlight of this is a purely geometric derivation of Strassen’s algorithm for \( 2 \times 2 \) matrices in §4.2.3. Other algorithms and principles for finding them are discussed in Chapter 4.

### 1.2. Separation of algebraic complexity classes

In a 1956 letter to von Neumann (see [Sip92, Appendix]) Gödel tried to quantify the apparent difference between intuition and systematic problem solving. Around the same time, researchers in the Soviet Union were trying to determine if “brute force search” was avoidable in solving problems such as the famous traveling salesman problem where there seems to be no fast way to find a solution, but a proposed solution can be easily checked. (The problem is to determine if there exists a way to visit, say twenty cities traveling less than a thousand miles. If I claim to have an algorithm to do so, you just need to look at my plan and check the distances.) These discussions eventually gave rise to the complexity classes \( P \), which models problems admitting a fast algorithm to produce a solution, and \( NP \) which models problems admitting a fast algorithm to verify a proposed solution, and the famous conjecture of Cook, Karp and Levin that these two classes are distinct. See [Sip92] for a history of the problem.

The transition of this problem to geometry goes via algebra:
1.2.1. From complexity to algebra. The P v. NP conjecture is generally believed to be out of reach at the moment, so there have been weaker conjectures proposed that might be more tractable. One such comes from a standard counting problem discussed in §1.3.1. This variant has the advantage that it admits a clean algebraic formulation that I now discuss.

L. Valiant [Val79a] conjectured that a sequence of polynomials that is “easy” to write down should not necessarily admit a fast evaluation. He defined algebraic complexity classes that are now called VP and VNP (see §6.1.2 for their definitions), and conjectured:

**Conjecture 1.2.1.1 (Valiant [Val79a]).** $\text{VP} \neq \text{VNP}$.

For the precise relationship between this conjecture and the $\text{P} \neq \text{NP}$ conjecture see [BCS97, Chap. 21].

Valiant also showed that a particular polynomial sequence, the permanent $(\text{perm}_n)$, is complete for the class VNP, in the sense that $\text{VP} \neq \text{VNP}$ if and only if $(\text{perm}_n) \not\in \text{VP}$. As explained in §1.3.1, the permanent is natural for computer science. Although it is not immediately clear, the permanent is also natural to geometry, see §5.2.6. The formula for the permanent of an $n \times n$ matrix $x = (x_{ij})$ is:

\[
\text{perm}_n(x) := \sum_{\sigma \in \mathfrak{S}_n} x_{\sigma(1)}^1 \cdots x_{\sigma(n)}^n.
\]

Here $\mathfrak{S}_n$ denotes the group of permutations of $\{1, \ldots, n\}$.

How would one show there is no fast algorithm for the permanent? In §6.1.2 we will define algebraic circuits, which are a class of algorithms for computing a polynomial, and their size, which is a measure of the complexity of the algorithm. Let circuit-size$(\text{perm}_n)$ denote the size of the smallest algebraic circuit computing $\text{perm}_n$.

**Conjecture 1.2.1.2 (Valiant [Val79a]).** circuit-size$(\text{perm}_n)$ grows faster than any polynomial in $n$.

Since this conjecture is expected to be extremely difficult, one could simply try to find lower bounds for circuit-size$(\text{perm}_n)$.

1.2.2. From algebra to algebraic geometry. As with our earlier discussion, one could work as follows:

Let $S^n\mathbb{C}^{n^2}$ denote the vector space of all homogeneous polynomials of degree $n$ in $n^2$ variables, so $\text{perm}_n$ is a point of the vector space $S^n\mathbb{C}^{n^2}$. (Here $S^n$ is short for the $n$-th symmetric tensor power.) If we write an element of $S^n\mathbb{C}^{n^2}$ as $p = \sum_{1 \leq i_1 \leq \cdots \leq i_n \leq n^2} c_{i_1 \cdots i_n} y_{i_1} \cdots y_{i_n}$, then we may view
1.2. Separation of algebraic complexity classes

the coefficients $c_{i_1\cdots i_n}$ as coordinates on this vector space. We will look for polynomials on our space of polynomials, that is, polynomials on the $c_{i_1\cdots i_n}$.

Plan to show $(\text{perm}_n) \not\in \text{VP}$, or at least bound its circuit size by $r$ with algebraic geometry.

- Find a polynomial $P$ on $S^n \mathbb{C}^{n^2}$ such that $P(p) = 0$ for all $p \in S^n \mathbb{C}^{n^2}$ with circuit-size($p$) $\leq r$.
- Show that $P(\text{perm}_n) \neq 0$.

By the discussion above on Zariski closure, this may be a more difficult problem: we are not just trying to exclude $\text{perm}_n$ from having a circuit, but we are also requiring it not be “near” to having a small circuit. I return to this issue below.

1.2.3. Shallow circuits and algebraic geometry. As with matrix multiplication, one would like to use symmetry, more precisely, representation theory, to attack Valiant’s conjectures. Unfortunately, representation theory gives little insight regarding polynomials for small circuits.

Fortunately, there are reduction theorems that allow one to restrict the classes of circuits one works with. Chapter 7 deals with two reductions that are amenable to algebraic geometry and representation theory. On one hand, they have very appealing algebraic varieties associated to them (e.g., secant varieties of Chow varieties). On the other, one pays a heavy price: instead of needing to prove non-polynomial growth, one needs to prove non-nearly-exponential growth.

1.2.4. Another path to algebraic geometry. The permanent resembles one of the most, perhaps the most, studied polynomial, the determinant of an $n \times n$ matrix $x = (x_{ij})$:

\begin{equation}
\det_n(x) := \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{\sigma(1)}^1 \cdots x_{\sigma(n)}^n.
\end{equation}

Here $\text{sgn}(\sigma)$ denotes the sign of the permutation $\sigma$. The determinant, despite its enormous formula of $n!$ terms, can be computed very quickly, e.g., by Gaussian elimination. (See §6.1.2 for an explicit division free algorithm.) In particular $(\det_n) \in \text{VP}$. It is not known if $\det_n$ is complete for $\text{VP}$, that is, whether or not a sequence of polynomials is in $\text{VP}$ if and only if it can be reduced to the determinant in the sense made precise below.

Although

\[ \text{perm} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & -b \\ c & d \end{pmatrix}, \]
Marcus and Minc [MM61], building on work of Pólya and Szegő (see [Gat87]), proved that one could not express $\text{perm}_m(y)$ as a size $m$ determinant of a matrix whose entries are affine linear functions of the $x^i_j$ when $m > 2$. This raised the question that perhaps the permanent of an $m \times m$ matrix could be expressed as a slightly larger determinant. More precisely, we say $p(y^1, \ldots, y^M)$ is an affine linear projection of $q(x^1, \ldots, x^N)$, if there exist affine linear functions $x^i_\alpha(y) = x^i_\alpha(y^1, \ldots, y^M)$ such that $p(y) = q(x(y))$. For example

\[
\text{perm}_3(y) = \det_7 \begin{pmatrix}
0 & y_1^1 & y_1^2 & y_1^3 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & y_3^3 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & y_3^1 & y_3^3 \\
0 & 0 & 0 & 1 & y_3^1 & 0 & y_3^1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
y_2^2 & 0 & 0 & 0 & 0 & 1 & 0 \\
y_2^3 & 0 & 0 & 0 & 0 & 0 & 1 \\
y_2^1 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

This formula is due to B. Grenet [Gre11], who also generalized it to express $\text{perm}_m$ as a determinant of size $2^m - 1$.

Valiant conjectured that one cannot do much better than this. Let $p$ be a polynomial. Let $dc(p)$ denote the smallest $n$ such that $p$ is an affine linear projection of the determinant. ($\det_n$ would be $\text{VP}$ complete if $dc(p_m)$ grew no faster than a polynomial for all sequences $(p_m) \in \text{VP}$.)

**Conjecture 1.2.4.1** (Valiant [Val79a]). $dc(\text{perm}_m)$ grows faster than any polynomial in $m$.

### 1.2.5. Geometric Complexity Theory

The “Zariski closed” version of Conjecture 1.2.4.1 is the flagship conjecture of Geometric Complexity Theory and the topic of Chapter 6. To state it in a useful form, rephrase Valiant’s conjecture as follows:

Let $\text{End}(\mathbb{C}^{n^2})$ denote the space of all linear maps $\mathbb{C}^{n^2} \to \mathbb{C}^{n^2}$, which acts on $S^n \mathbb{C}^{n^2}$ under the action $L \cdot p(x) := p(L^T(x))$, where $x$ is viewed as a column vector of size $n^2$, $L$ is an $n^2 \times n^2$ matrix, and $T$ denotes transpose. (The transpose is used so that $L_1 \cdot (L_2 \cdot p) = (L_1 L_2) \cdot p$.) Define an auxiliary variable $\ell \in \mathbb{C}^1$ so $\ell^{n-m} \text{perm}_m \in S^n \mathbb{C}^{m^2+1}$. Consider any linear inclusion $\mathbb{C}^{(m^2+1)} \to \mathbb{C}^{n^2}$, so we may consider $\ell^{n-m} \text{perm}_m \in S^n \mathbb{C}^{n^2}$. Then

\[
dc(\text{perm}_m) \leq n \iff \ell^{n-m} \text{perm}_m \in \text{End}(\mathbb{C}^{n^2}) \cdot \det_n.
\]

For $p \in S^d \mathbb{C}^M$ with $d \leq n$ and $M \leq n^2$, let $\overline{dc}(p)$ denote the smallest $n$ such that $\ell^{n-d}p \in \text{End}(\mathbb{C}^{n^2}) \cdot \det_n$.

**Conjecture 1.2.5.1.** [MS01] $\overline{dc}(\text{perm}_m)$ grows faster than any polynomial in $m$. 
Representation theory indicates a path towards solving Conjecture 1.2.5.1. To explain the path, introduce the following terminology: a polynomial \( p \in S^n \mathbb{C}^N \) is characterized by its symmetries if, letting \( G_p := \{ g \in GL_N \mid g \cdot p = p \} \), for any \( q \in S^n \mathbb{C}^N \) with \( G_q \supseteq G_p \), one has \( p = \lambda q \) for some \( \lambda \in \mathbb{C} \).

There are two essential observations:

- \( \text{End}(\mathbb{C}^n^2) \cdot \det_n = GL_n^2 \cdot \det_n \), that is \( \text{End}(\mathbb{C}^n^2) \cdot \det_n \) is an orbit closure.
- \( \det_n \) and \( \text{perm}_n \) are characterized by their symmetries.

Representation theory (more precisely, the Peter-Weyl Theorem, see §5.1), in principle gives a description of the polynomials vanishing on an orbit closure modulo the effect of the boundary. (More precisely, it describes the ring of regular functions on the orbit.) In the problem at hand this leads to questions about much-studied, but little understood, quantities in representation theory and combinatorics (Kronecker and plethysm coefficients).

A potential problem is that there might be different orbits with the same coordinate ring. The property of being characterized by symmetries avoids this problem.

Unlike matrix multiplication, this problem is in its infancy and I do not expect it to be fully resolved in the near future. But there are several paths to making progress with this conjecture that I discuss in Chapter 6. To gain insight as to what techniques might work, it will be useful to examine “toy” versions of the problem - these questions are of mathematical significance in their own right, and lead to interesting connections between combinatorics, representation theory and geometry. See §7.3 for one such, called the Hadamard-Howe problem.

### 1.3. Supplementary material

This section consists of three topics: §1.3.1 discusses why the permanent is natural for computer science, §1.3.2 is intended for readers who have no experience with algebraic geometry, and §1.3.3 is intended for those not used to working with big numbers.

#### 1.3.1. Remarks on the permanent and complexity theory

A standard problem in graph theory, for which the only known algorithms are exponential in the size of the graph, is to count the number of perfect matchings of a bipartite graph, that is, a graph with two sets of vertices and edges only joining vertices from one set to the other.

A perfect matching is a subset of the edges such that each vertex shares an edge from the subset with exactly one other vertex.
To a bipartite graph one associates an incidence matrix $x_{ij}$, where $x_{ij} = 1$ if an edge joins the vertex $i$ above to the vertex $j$ below and is zero otherwise. For example the graph above has incidence matrix

$$
\begin{pmatrix}
1 & 1 & 0 \\
0 & 1 & 1 \\
0 & 1 & 1
\end{pmatrix}.
$$

A perfect matching corresponds to a matrix constructed from the incidence matrix by setting some of the entries to zero so that the resulting matrix has exactly one 1 in each row and column, i.e., is a matrix obtained by applying a permutation to the columns of the identity matrix.

For example, $\text{perm}_3 \begin{pmatrix}
1 & 1 & 0 \\
0 & 1 & 1 \\
0 & 1 & 1
\end{pmatrix} = 2$.

**Exercise 1.3.1.1:** Show that if $x$ is the incidence matrix of a bipartite graph, then $\text{perm}_n(x)$ indeed equals the number of perfect matchings.

Thus a classical problem: determine the complexity of counting the number of perfect matchings of a bipartite graph (which is complete for the complexity class $\#P$), can be studied via algebra - determine the complexity of evaluating the permanent.

1.3.2. **Exercises for those not familiar with algebraic geometry.** In §1.1.11 we defined the rank of a matrix in a way that made it clear it was a property of the linear map the matrix represented.
Theorem 1.3.2.1 (Fundamental theorem of linear algebra). Let \( V, W \) be finite dimensional vector spaces, let \( f : V \to W \) be a linear map and let \( A_f \) be a matrix representing \( f \). Then

1. \[
\text{rank}(f) = \dim f(V) = \dim(\text{span}\{\text{columns of } A_f\}) = \dim(\text{span}\{\text{rows of } A_f\}) = \dim V - \dim \ker f.
\]

In particular \( \text{rank}(f) \leq \min\{\dim V, \dim W\} \).

2. For generic \( f \), \( \text{rank}(f) = \min\{\dim V, \dim W\} \).

3. If a sequence of linear maps \( f_t \) of rank \( r \) has a limit \( f_0 \), then \( \text{rank}(f_0) \leq r \).

4. \( \text{rank}(f) \leq r \) if and only if, in any choice of bases, the determinants of all size \( r + 1 \) submatrices of the matrix representing \( f \) are all zero.

Exercise 1.3.2.2: Prove the theorem. ☐

Note that (4) says the Zariski closure of the space of linear maps of rank \( r \) is the space of linear maps of rank \( \leq r \) and (3) says that the Euclidean closure (i.e., closure under taking limits) of the linear maps of rank \( r \) is contained in the set of linear maps of rank \( \leq r \) (of course equality holds).

Now consider polynomials on spaces of polynomials. Let \( p \in S^n\mathbb{C}^N \). How can one test if \( p \) is an \( n \)-th power of a linear form, \( p = \ell^n \) for some \( \ell \in \mathbb{C}^N \)?

Exercise 1.3.2.3: Show that \( p = \ell^n \) for some \( \ell \in \mathbb{C}^N \) if and only if \( \dim \{ \frac{\partial p}{\partial x^1}, \ldots, \frac{\partial p}{\partial x^N} \} = 1 \), where \( x^1, \ldots, x^N \) are coordinates on \( \mathbb{C}^N \).

How would one test if \( p \) is the sum of two \( n \)-th powers, \( p = \ell_1^n + \ell_2^n \) for some \( \ell_1, \ell_2 \in \mathbb{C}^N \)?

Exercise 1.3.2.4: Show that \( p = \ell_1^n + \ell_2^n \) for some \( \ell_j \in \mathbb{C}^N \) implies \( \dim \{ \frac{\partial^2 p}{\partial x^i \partial x^j} \} \leq 2 \).

Exercise 1.3.2.5: Show that the above condition is not sufficient by considering \( p = \ell_1^{n-1} \ell_2 \).

Exercise 1.3.2.6: Show that \( p = \ell_1^n + \ell_2^n \) for some \( \ell_j \in \mathbb{C}^N \) implies \( \dim \{ \frac{\partial^2 p}{\partial x^i \partial x^j} \} \leq 2 \). Is this condition sufficient?

Exercise 1.3.2.7: Show that any polynomial vanishing on all polynomials of the form \( p = \ell_1^n + \ell_2^n \) for some \( \ell_j \in \mathbb{C}^N \) also vanishes on \( x^{n-1}y \). ☐
**Exercise 1.3.2.8:** More generally, show that the Euclidean closure (i.e., closure under taking limits) of a set is always contained in its Zariski closure.

A question to think about: how would one detect if \( p \in S^n \mathbb{C}^N \) was a product of \( n \) linear forms? (See §7.3.3 for the answer.)

### 1.3.3. Big numbers.

\[
\begin{align*}
(1.3.1) & \quad n! \gtrsim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \\
(1.3.2) & \quad \ln(n!) = n \ln(n) - O(\ln(n)) \\
(1.3.3) & \quad \binom{2n}{n} \gtrsim \frac{4^n}{\sqrt{\pi n}} \\
(1.3.4) & \quad \ln \left( \frac{\alpha n}{\beta n} \right) = \alpha H_e \left( \frac{\beta}{\alpha} \right) n - O(\ln n)
\end{align*}
\]

where \( H_e(x) := -x \ln x - (1 - x) \ln(1 - x) \) is the **Shannon entropy**. All these identities follow from (1.3.1). Equation (1.3.1) is proved via complex analysis (analysis of a contour integral) - it is not a simple matter of linear algebra.

**Exercise 1.3.3.1:** Show \( a^{\log(b)} = b^{\log(a)} \).

**Exercise 1.3.3.2:** Consider the following sequences of \( n \):

\[
\log_2(n), \ n, \ 100n, \ n^2, \ n^3, \ n^{\log_2(n)}, \ 2^{\log^2(n)}, \ n^{\sqrt{\log(n)}}, \ 2^n, \ \binom{2n}{n}, \ n!, \ n^n.
\]

In each case, determine for which \( n \), the sequence surpasses the number of atoms in the known universe. (It is estimated that there are between \( 10^{78} \) and \( 10^{82} \) atoms in the known universe.)

**Exercise 1.3.3.3:** Compare the sizes of \( s^{\sqrt{d}} \) and \( 2^{\sqrt{\log ds}} \).
The complexity of Matrix multiplication
I: lower bounds

I begin in §2.1 with a brief tour of multi-linear algebra and present matrix multiplication from an invariant point of view. In §2.2 the basic complexity measures tensor rank and border rank are defined. In §2.3 (resp. §2.4) basic definitions and examples from representation theory (resp. algebraic geometry) are presented. To prove lower bounds, one needs to find equations on spaces of tensors, and an introduction to this subject, as well as the state of the art regarding lower bounds for the border rank of matrix multiplication, are given in in §2.5. At first sight, algebraic geometry does not appear useful for proving lower bounds on tensor rank beyond border rank bounds, but thanks to the symmetries of the matrix multiplication tensor, it can be used to prove rank lower bounds for $M_{(n)}$, which is explained in §2.6. The chapter concludes with §2.7, where I present the state of the art regarding equations for border rank. As explained in Chapter 3, these equations may also be useful for proving upper bounds for the complexity of matrix multiplication.

2.1. Matrix multiplication and multi-linear algebra

To better understand matrix multiplication as a bilinear map, I first review basic facts from linear and multi-linear algebra. For more details on this topic, see [Lan12, Chap. 2].

2.1.1. Linear maps. In what follows it will be essential to work without bases, so instead of writing $\mathbb{C}^V$ I will work with a complex vector space $V$ of
dimension $v$. Given $V$, one can form a second vector space, called the dual space to $V$, whose elements are linear maps from $V$ to $\mathbb{C}$:

$$V^* := \{ \alpha : V \to \mathbb{C} \mid \alpha \text{ is linear} \}$$

If one is working in bases and represents elements of $V$ by column vectors, then elements of $V^*$ are naturally represented by row vectors and the map $v \mapsto \alpha(v)$ is just row-column matrix multiplication. Given a basis $v_1, \ldots, v_v$ of $V$, it determines a basis $\alpha^1, \ldots, \alpha^v$ of $V^*$ by $\alpha^i(v_j) = \delta_{ij}$.

**Exercise 2.1.1.1:** Assuming $V$ is finite dimensional, write down a canonical isomorphism $V \to (V^*)^*$. ⊗

Let $V^* \otimes W$ denote the vector space of all linear maps $V \to W$. Given $\alpha \in V^*$ and $w \in W$ define a linear map $\alpha \otimes w : V \to W$ by $\alpha \otimes w(v) := \alpha(v)w$. (In bases, if $\alpha$ is represented by a row vector and $w$ by a column vector, $\alpha \otimes w$ will be represented by the matrix $wa$.) Such a linear map is said to have rank one.

Define the rank of an element $f \in V^* \otimes W$ is the smallest $r$ such $f$ may be expressed as a sum of $r$ rank one linear maps.

A linear map $f : V \to W$ determines a linear map $f^T : W^* \to V^*$ defined by $f^T(\beta)(v) := \beta(f(v))$ for all $v \in V$ and $\beta \in W^*$. Note that this is consistent with the notation $V^* \otimes W \cong W \otimes V^*$, being interpreted as the space of all linear maps $(W^*)^* \to V^*$, that is, the order we write the factors does not matter. If we work in bases and insist that all vectors are column vectors, the matrix of $f^T$ is just the transpose of the matrix of $f$.

**Exercise 2.1.1.2:** Show that we may also consider an element $f \in V^* \otimes W$ as a bilinear map $b_f : V \times W^* \to \mathbb{C}$ defined by $b_f(\beta, v) := \beta(f(v))$.

**Exercise 2.1.1.3:** If $v_1, \ldots, v_v$ is a basis of $V$ and $v_1^*, \ldots, v_v^*$ the dual basis of $V^*$, defined by $v^j(v_i) = \delta^j_i$, show that the identity map on $V$ is $Id_V = \sum_j v^j \otimes v_j$.

**Exercise 2.1.1.4:** Show that there is a canonical isomorphism $(V^* \otimes W)^* \to V \otimes W^*$ where $\alpha \otimes w(\omega \otimes \beta) := \alpha(w)\beta(v)$. Now let $V = W$ and let $Id_V \in V^* \otimes V \cong (V^* \otimes V)^*$ denote the identity map. What is $Id_V(f)$ for $f \in V^* \otimes V$? ⊗

### 2.1.2. Multi-linear maps and tensors

We say $V \otimes W$ is the tensor product of $V$ with $W$. More generally, for vector spaces $A_1, \ldots, A_n$ define their tensor product $A_1 \otimes \cdots \otimes A_n$ to be the space of $n$-linear maps $A_1^* \times \cdots \times A_n^* \to \mathbb{C}$, equivalently the space of $(n-1)$-linear maps $A_1^* \times \cdots \times A_{n-1}^* \to A_n$ etc.

Let $a_j \in A_j$ and define an element $a_1 \otimes \cdots \otimes a_n \in A_1 \otimes \cdots \otimes A_n$ by $a_1 \otimes \cdots \otimes a_n(a_1^*, \ldots, a_n^*) := a_1^1(a_1) \cdots a_n^1(a_n)$.
2.2. Complexity of bilinear maps

**Exercise 2.1.2.1:** Show that if \( \{ a_j^{s_j} \} \), \( 1 \leq s_j \leq a_j \), is a basis of \( A_j \), then \( a_1^{s_1} \otimes \cdots \otimes a_n^{s_n} \) is a basis of \( A_1 \otimes \cdots \otimes A_n \). In particular \( \dim(A_1 \otimes \cdots \otimes A_n) = a_1 \cdots a_n. \)

**Remark 2.1.2.2.** One may identify \( A_1 \otimes \cdots \otimes A_n \) with any re-ordering of the factors. When I need to be explicit about this, I will call this identification the re-ordering isomorphism.

### 2.1.3. Matrix multiplication.

We may view matrix multiplication

\[
M_{(U,V,W)} : (U^* \otimes V)^* \times (V^* \otimes W)^* \to W^* \otimes U
\]

as a trilinear map which I will still denote by \( M_{(U,V,W)} \):

\[
M_{(U,V,W)} : (U^* \otimes V)^* \times (V^* \otimes W)^* \times (U \otimes W^*)^* \to \mathbb{C}.
\]

**Exercise 2.1.3.1:** Show that in bases this map is \((X,Y,Z) \mapsto \text{trace}(XYZ)\).

We have \( M_{(U,V,W)} \in (U^* \otimes V) \otimes (V^* \otimes W) \otimes (W^* \otimes U) \simeq (U \otimes U^*) \otimes (V \otimes V^*) \otimes (W \otimes W^*) \).

**Exercise 2.1.3.2:** Show that as a tensor \( M_{(U,V,W)} = \text{Id}_U \otimes \text{Id}_V \otimes \text{Id}_W \).

**Exercise 2.1.3.3:** Show that \( \text{Id}_U \otimes \text{Id}_V \otimes \text{Id}_V \), when viewed as a bilinear map \((U^* \otimes V) \times (V^* \otimes W) \to U^* \otimes W\), takes a linear map \( f : U \to V \), and a second linear map \( g : V \to W \) and sends it to their composition \( g \circ f : U \to W \).

**Exercise 2.1.3.4:** Show that \( \text{Id}_V \otimes \text{Id}_W \in V \otimes V^* \otimes W^* = (V \otimes W) \otimes (V \otimes W)^* \) equals \( \text{Id}_V \otimes \text{Id}_W \).

**Exercise 2.1.3.5:** Show that \( M_{(n,m,1)} \otimes M_{(n',m',1')} = M_{(nn',mm',ll')} \).

### 2.2. Complexity of bilinear maps

**2.2.1. Tensor rank.** An element \( T \in A_1 \otimes \cdots \otimes A_n \) is said to have rank **one** if there exist \( a_j \in A_j \) such that \( T = a_1 \otimes \cdots \otimes a_n \).

We will use the following measure of complexity:

**Definition 2.2.1.1.** Let \( T \in A_1 \otimes \cdots \otimes A_n \). Define the rank (or tensor rank) of \( T \) to be the smallest \( r \) such that \( T \) may be written as the sum of \( r \) rank one tensors. We write \( R(T) = r \). Let \( \delta_r^{0} = \delta_{r,A_1 \otimes \cdots \otimes A_n} \subseteq A_1 \otimes \cdots \otimes A_n \) denote the set of tensors of rank at most \( r \).

The rank of \( T \in A \otimes B \otimes C \) is comparable to all other standard measures of complexity on the space of bilinear maps, see, e.g., [BCS97, §14.1].

For example, letting \( x_i^j, y_j^a, z_i^u \) respectively be bases of \( A = \mathbb{C}^{nn}, B = \mathbb{C}^{ml}, C = \mathbb{C}^{ln} \), then the standard expression of matrix multiplication is

\[
M_{(l,m,n)} = \sum_{i=1}^n \sum_{a=1}^m \sum_{u=1}^l x_i^j \otimes y_j^a \otimes z_i^u
\]
so we conclude $R(M_{(n,m,l)}) \leq nml$.

**Exercise 2.2.1.2:** Write Strassen’s algorithm out as a tensor.

Strassen’s algorithm shows $R(M_{(2,2,2)}) \leq 7$.

Recall our plan from §1.1.10 and that it was actually a plan for bounding border rank. I now elaborate on the difference between rank and border rank by comparing Zariski and Euclidean closure.

### 2.2.2. The Fundamental theorem of linear algebra is false for tensors

Recall the fundamental theorem of linear algebra from §1.3.2.

**Theorem 2.2.2.1.** Let $n \geq 3$.

1. If $T \in A_1 \otimes \cdots \otimes A_n$ is outside the zero set of a certain finite collection of polynomials (in particular outside a certain set of measure zero), then $R(T) \geq \frac{a_1 \cdots a_n}{a_1 \cdots + a_n + n + 1} \gg a_i$.

2. Rank can jump up (or down) under limits.

The first assertion is proved in Exercise 2.4.14.3. To see that (2) holds, at least when $r = 2$, consider

$$T(t) := \frac{1}{t} \left[ a_1 \otimes b_1 \otimes c_1 - (a_1 + ta_2) \otimes (b_1 + tb_2) \otimes (c_1 + tc_2) \right]$$

and note that

$$\lim_{t \to 0} T(t) = a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_1 + a_2 \otimes b_1 \otimes c_1$$

which does not have rank two (exercise).

To visualize why rank can jump up while taking limits, consider the following picture, where the curve represents the points of $\hat{\sigma}_r^0$. Points of $\hat{\sigma}_2^0$ (e.g., the dots limiting to the dot labelled $T$) are those on a secant line to $\hat{\sigma}_1^0$, and the points where the rank jumps up, such at the dot labelled $T$, are those that lie on a tangent line to $\hat{\sigma}_1^0$. (This phenomena fails to occur for matrices because for matrices, every point on a tangent line is also on an honest secant line.)

We appear to have a problem:

- The set $\hat{\sigma}_r^0$ is not closed under taking limits. (I will say a set that is closed under taking limits is closed in the Euclidean topology.)

- It is also not closed in the Zariski topology, i.e., the zero set of all polynomials vanishing on $\hat{\sigma}_r^0$ includes tensors that are of rank greater than $r$.

The tensors that are honestly “close” to tensors of rank $r$ would be the Euclidean closure, but to deal with polynomials, we needed to work with the Zariski closure.
Often the Zariski closure is much larger than the Euclidean closure. For example, the Zariski closure of $\mathbb{Z} \subset \mathbb{C}$ is $\mathbb{C}$, while $\mathbb{Z}$ is already closed in the Euclidean topology.

However, in this case we have good luck: the Zariski and Euclidean closures of $\hat{\sigma}_r^0$ coincide. I outline the proof in §2.4.3.

2.2.3. Border rank. Recall that $\hat{\sigma}_r$ denotes the Zariski (and by the above discussion Euclidean) closure of $\hat{\sigma}_0^r$, and the border rank of $T \in A_1 \otimes \cdots \otimes A_n$, denoted $R(T)$, is the smallest $r$ such that $T \in \hat{\sigma}_r$. By our above discussion, border rank is semi-continuous.

Border rank is easier to work with than rank for several reasons. For example, the maximal rank of a tensor in $C^m \otimes C^m \otimes C^m$ is not known in general. In contrast, the maximal border rank is known to be $\left\lceil \frac{m^3 - 1}{3m^2 - 3} \right\rceil$ for all $m \neq 3$, and is 5 when $m = 3$ [Lic85]. The method of proof is a differential-geometric calculation that dates back to Terracini [Ter11], see §2.4.14.

Exercise 2.2.3.1: Prove that if $T \in A \otimes B \otimes C$ and $T' := T|_{A' \times B' \times C'}$ for some $A' \subseteq A^*$, $B' \subseteq B^*$, $C' \subseteq C^*$, then $R(T) \geq R(T')$ and $R(T) \geq R(T')$. ⊙

Exercise 2.2.3.2: Let $T_j \in A_j \otimes B_j \otimes C_j$, $1 \leq j, k, l \leq s$. Consider $\oplus_j T_j \in (\oplus_j A_j) \otimes (\oplus_k B_k) \otimes (\oplus_l C_l)$ and, letting $A = \otimes_j A_j$, $B = \otimes_k B_k$, and $C = \otimes_l C_l$, $T_1 \otimes \cdots \otimes T_s \in A \otimes B \otimes C$. Show that $R(\oplus_j T_j) \leq \sum_{i=1}^s R(T_i)$ and $R(\otimes_{i=1}^s T_i) \leq \Pi_{i=1}^s R(T_i)$, and that the statements also hold for border rank.

2.2.4. Our first lower bound. Given $T \in A \otimes B \otimes C$, write $T \in A \otimes (B \otimes C)$ and think of $T$ as a linear map $T_A : A^* \to B \otimes C$.

Proposition 2.2.4.1. $R(T) \geq \text{rank}(T_A)$.

Exercise 2.2.4.2: Prove Proposition 2.2.4.1. ⊙
Exercise 2.2.4.3: Find a choice of bases such that

\[ M_{(n)}(A^*) = \begin{pmatrix} x & \cdots & x \end{pmatrix} \]

where \( x = (x_i^j) \) is \( n \times n \), i.e., the image in the space of \( n^2 \times n^2 \) matrices is block diagonal with all blocks the same.

Exercise 2.2.4.4: Show that \( R(M_{(n)}) \geq n^2 \).

A fancier proof that \( R(M_{(n)}) \geq n^2 \), which will be useful for proving further lower bounds, is as follows: Write \( A = U^* \otimes V \), \( B = V^* \otimes W \), \( C = W^* \otimes U \), so \( (M_{(n)})_A : A^* \rightarrow B \otimes C \) is a map \( U \otimes V^* \rightarrow V^* \otimes W \otimes W^* \otimes U \). This map is, for \( f \in A^* \), \( f \mapsto f \otimes \text{Id}_W \), and thus is clearly injective.

Exercise 2.2.4.5: Show \( R(M_{(m,n,1)}) = mn \) and \( R(M_{(m,1,1)}) = m \).

Exercise 2.2.4.6: Let \( b = c \) and assume \( T_A \) is injective. Show that if \( T(A^*) \) is diagonalizable under the action of \( GL(B) \times GL(C) \) then \( R(T) \leq b \), and therefore if \( T(A^*) \) is the limit of diagonalizable subspaces then \( R(T) \leq b \).

2.3. Definitions and examples from representation theory

2.3.1. First definitions. Let \( GL(V) \) denote the group of invertible linear maps \( V \rightarrow V \). If we have chosen a basis, this is the set of invertible \( v \times v \) matrices. One says \( GL(V) \) acts on \( V \) and that \( GL(V) \) is a representation of \( G \). The action is irreducible if there does not exist a proper subspace \( U \subset V \) such that \( \mu(g) \cdot u \in U \) for all \( u \in U \), \( g \in G \) (such a subspace \( U \) is called a submodule), and otherwise the action is decomposable. The word reducible is used when there exists both a subspace \( U \) and a complementary subspace \( U^c \) both of which are preserved by \( G \). A module \( W \) is trivial if for all \( g \in G \), \( \mu(g) = \text{Id}_W \).

2.3.2. First examples.

Example 2.3.2.1. \( GL(V) \) acts irreducibly on \( V \) and \( GL(V) \times GL(W) \) acts irreducibly on \( V^* \otimes W \) by \( (g,h) \cdot \alpha \otimes w = (g \cdot \alpha) \otimes (h \cdot w) \) and extending linearly. To see the first is irreducible, any nonzero vector can be mapped to any other
nonzero vector by $GL(V)$. Exercise 2.3.3.5 below will show the second is irreducible.

**Example 2.3.2.2.** Let
\[
G = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \bigm| t \in \mathbb{C} \right\} \subset GL_2
\]
Then $G$ acts on $\mathbb{C}^2$, the line $\mathbb{C}\{e_1\}$ is a submodule, but there is no complementary subspace preserved, so this action is decomposable, but not reducible.

**Example 2.3.2.3.** The permutation group $S_d$ acts on $\mathbb{C}^d$ equipped with its standard basis $e_1, \ldots, e_d$ by $\sigma(e_j) = e_{\sigma(j)}$ and extending linearly. This action is reducible, it decomposes into the direct sum of the trivial representation, given by the span of $e_1 + \cdots + e_d$, and a second irreducible representation which has basis $e_1 - e_2, \ldots, e_1 - e_d$. The first is usually denoted $[d]$, and the second is denoted $[d-1,1]$. (These notations are explained in §5.2.2.)

There is a second one-dimensional representation of $S_d$, denoted $[1^d] := [1, \ldots, 1]$, namely, for $v \in \mathbb{C}^1$, define $\sigma \cdot v = \text{sgn}(\sigma)v$. This is called the sign representation.

**Proposition 2.3.2.4.** The action of $GL(V)$ on $V^* \otimes V$ given by $g \cdot (\alpha \otimes v) = (g \cdot \alpha) \otimes (g \cdot v)$ is reducible. It decomposes as $V^* \otimes V = \mathbb{C}\{Id_V\} \oplus \mathfrak{sl}(V)$, where $Id_V$ is the identity map and $\mathfrak{sl}(V)$ are the linear maps $f : V \to V$ such that, $Id_{V^*}(f) = 0$, where we consider $Id_{V^*} \in V \otimes V^* \simeq (V^* \otimes V)^*$. In bases, for a matrix $X$, the action is $g \cdot X = gXg^{-1}$.

**Exercise 2.3.2.5:** Prove Proposition 2.3.2.4. Show that moreover $\mathbb{C}\{Id\}$ is a trivial $GL(V)$-module.

**Remark 2.3.2.6.** In any choice of basis, $\mathfrak{sl}(V)$ will be identified with the traceless matrices, and by Exercise 2.1.1.4, the map $f \mapsto Id_{V^*}(f)$ is just $f \mapsto \text{trace}(f)$, i.e., $f$ maps to the sum of its eigenvalues. For more details see [Lan12, §2.3.2.5].

**Proposition 2.3.2.7.** The action of $GL(V)$ on $V \otimes V$ given by $g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w)$ decomposes into two irreducible submodules. In bases this decomposition corresponds to the splitting of $v \times v$ matrices into the direct sum of symmetric and skew-symmetric matrices. If we represent the bilinear map $v \otimes w$ by a matrix $X$, where, if $\alpha, \beta \in V^*$ are represented by row vectors, so $(v \otimes w)(\alpha, \beta) = \alpha X \beta^T$, then the matrix $g \in GL(V)$ acts by $g \cdot X = gXg^T$.

**Exercise 2.3.2.8:** Prove Proposition 2.3.2.7.

Propositions 2.3.2.4 and 2.3.2.7 illustrate the importance of keeping track of the difference between $V$ and $V^*$. In bases, both $V \otimes V$ and $V \otimes V^*$
are just $v \times v$ matrices, but the action of $GL(V)$ on the two spaces is very different.

### 2.3.3. Schur’s lemma and consequences.

**Definition 2.3.3.1.** Let $\rho_j : G \to GL(W_j)$, $j = 1, 2$ be representations. A $G$-module homomorphism, or $G$-module map, is a linear map $f : W_1 \to W_2$ such that $f(\rho_1(g) \cdot v) = \rho_2(g) \cdot f(v)$ for all $v \in W_1$ and $g \in G$.

One says $W_1$ and $W_2$ are isomorphic $G$-modules if there exists a $G$-module homomorphism $W_1 \to W_2$ that is a linear isomorphism.

**Exercise 2.3.3.2:** Show that the image and kernel of a $G$-module homomorphism are $G$-modules.

The following easy lemma is central to representation theory:

**Lemma 2.3.3.3** (Schur’s Lemma). Let $G$ be a group, let $V$ and $W$ be irreducible $G$-modules and let $f : V \to W$ be a $G$-module homomorphism. Then either $f = 0$ or $f$ is an isomorphism. If further $V = W$, then $f = \lambda \text{Id}$ for some constant $\lambda$.

**Exercise 2.3.3.4:** Prove Schur’s Lemma.

**Exercise 2.3.3.5:** Let $V$ be an irreducible $G$-module and let $W$ be an irreducible $H$-module. Show that $V \otimes W$ is an irreducible $G \times H$-module.

### 2.3.4. Spaces of tensors.

Give $V \otimes d$ the structure of a $GL(V)$-module by $g \cdot (v_1 \otimes \cdots \otimes v_d) := (g \cdot v_1) \otimes \cdots \otimes (g \cdot v_d)$ and extending linearly.

**Exercise 2.3.4.1:** Give $V \otimes d$ the structure of a $S_d$-module by defining $\sigma \cdot (v_1 \otimes \cdots \otimes v_d) := v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(d)}$.

Show that the actions of $S_d$ and $GL(V)$ on $V \otimes d$ commute with each other.

**Definition 2.3.4.2.** A tensor $T \in V \otimes d$ is said to be symmetric if $T(\alpha_1, \ldots, \alpha_d) = T(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(d)})$ for all $\sigma \in S_d$, and skew-symmetric if $T(\alpha_1, \ldots, \alpha_d) = \text{sgn}(\sigma)T(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(d)})$ for all $\sigma \in S_d$. Let $S^d V \subset V \otimes d$ (resp. $\Lambda^d V \subset V \otimes d$) denote the space of symmetric (resp. skew-symmetric) tensors.

Note that $\Lambda^d V$ and $S^d V$ are $GL(V)$-submodules of $V \otimes d$.

Introduce the notations:

$$x_1 x_2 \cdots x_k := \frac{1}{k!} \sum_{\sigma \in S_k} x_{\sigma(1)} \otimes x_{\sigma(2)} \otimes \cdots \otimes x_{\sigma(k)} \in S^k V.$$
and
\[ x_1 \land x_2 \land \cdots \land x_k := \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \text{sgn}(\sigma) x_{\sigma(1)} \otimes x_{\sigma(2)} \otimes \cdots \otimes x_{\sigma(k)} \in \Lambda^k V. \]

If \( v_1, \ldots, v_v \) is a basis of \( V \), then \( v_i \otimes \cdots \otimes v_i \) with \( i_j \in [v] \) is a basis of \( V^\otimes d \), \( v_i \otimes v_i \) with \( 1 \leq i_1 \leq \cdots \leq i_d \leq v \) is a basis of \( S^d V \) and \( v_i \land \cdots \land v_i \) with \( 1 \leq i_1 < \cdots < i_d \leq v \) is a basis of \( S^d V \). With respect to this basis, if \( x_j = (x_j^1, \ldots, x_j^v)^T \), then the coefficient of \( x_1 \land \cdots \land x_k \) on the basis vector \( v_i \land \cdots \land v_i \) is
\[ \det \begin{pmatrix} x_1^{i_1} & \cdots & x_k^{i_k} \\ \vdots & & \vdots \\ x_1^{i_1} & \cdots & x_k^{i_k} \end{pmatrix}. \]

**Definition 2.3.4.3.** If \( l > k \), define a contraction map \( V^\otimes k \times V^* \otimes l \to V^* \otimes l-k \) by
\[(X, \phi) = (x_1 \otimes \cdots \otimes x_k, \phi_1 \otimes \cdots \otimes \phi_l) \mapsto X \phi := \phi_1(x_1) \cdots \phi_k(x_k) \phi_{k+1} \otimes \cdots \otimes \phi_l.\]

**2.3.5. Exercises.**

1. Show that the \( \mathfrak{S}_2 \)-module \( V^\otimes 2 \) is reducible. Show that in matrices, this decomposition corresponds to writing an \( v \times v \) matrix as the direct sum of a symmetric matrix and a skew-symmetric matrix.

2. Show that as a \( GL(V) \times \mathfrak{S}_2 \)-module, \( V \otimes V = (S^2 V \otimes [2]) \oplus (\Lambda^2 V \otimes [1, 1]). \)

3. Given \( T \in S^d V \), show that \( T \) defines a homogeneous polynomial of degree \( d \) on \( V^* \), which, for the purposes of this exercise, denote it by \( P_T \), via \( P_T(\alpha) = T(\alpha, \ldots, \alpha) \). Similarly, given a homogeneous polynomial \( P \) of degree \( d \) on \( V^* \), define a tensor \( \overline{P} \) by \( \overline{P}(\alpha_1, \ldots, \alpha_d) \) is \( \frac{1}{d!} \) times the coefficient of \( t_1 \cdots t_d \) in the polynomial \( P(t_1 \alpha_1 + \cdots + t_d \alpha_d) \). The tensor \( \overline{P} \) is called the polarization of \( P \). I will generally use the the same notation for symmetric tensors and homogenous polynomials.

4. Let \( P \in S^2 \mathbb{C}^2 \). If \( P \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = a\alpha^2 + b\alpha\beta + c\beta^2 \), what is \( \overline{P} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} \)?

5. One can also define partial polarizations. Consider the surjective map \( S^q V \otimes S^p V \to S^{p+q} V \) given by multiplication of polynomials. Note that it is a \( GL(V) \)-module map. Consider the transpose map \( S^{p+q} V^* \to S^q V^* \otimes S^p V^* \). The image of a polynomial \( P \in S^{p+q} V^* \) will be denoted \( P_{q,p} \). Consider \( P_{q,p} : S^q V \to S^p V^* \) and show that in bases, its image is the span of the partial derivatives of \( P \) of order \( q \).

6. Given a linear map \( f : V \to W \), one obtains linear maps \( f^\otimes k : V^\otimes k \to W^\otimes k \) defined by \( f(v_1 \otimes \cdots \otimes v_k) = f(v_1) \otimes \cdots \otimes f(v_k) \). Show
that \( f^\otimes k \) descends to give linear maps \( f^\otimes k : S^k V \rightarrow S^k W \) and \( f^\wedge k : \Lambda^k V \rightarrow \Lambda^k W \). Show that if \( \dim V = \dim W = v \), the map \( f^\wedge v : \Lambda^v V \rightarrow \Lambda^v W \) is multiplication by a scalar. If \( V = W \), show that the scalar is the determinant of the matrix of \( f \) with respect to any choice of basis. Show that more generally, in bases, the matrix of \( f^\wedge k \) has entries \( \pm \) the size \( k \) minors of the matrix of \( f \), in particular the matrix of \( f^\wedge v^{-1} \) is (up to transpose) the cofactor matrix of the matrix of \( f \).

(7) Show that as a \( GL(V) \times GL(W) \)-module, \( S^2(V \otimes W) = (S^2 V \otimes S^2 W) \oplus (\Lambda^2 V \otimes \Lambda^2 W) \) and that \( \Lambda^2(V \otimes W) = (S^2 V \otimes \Lambda^2 W) \oplus (\Lambda^2 V \otimes S^2 W) \).

For \( T \in A \otimes B \otimes C \), define
\[
G_T := \{ g = (g_1, g_2, g_3) \in GL(A) \times GL(B) \times GL(C) \mid g \cdot T = T \} \subset GL(A) \times GL(B) \times GL(C),
\]
called the stabilizer of \( T \).

Thanks to Exercises 2.1.3.2 and 2.3.2.5, we conclude

**Proposition 2.3.5.1.** Matrix multiplication \( M_{(U,V,W)} \) is invariant under \( GL(U) \times GL(V) \times GL(W) \). In other words,
\[
GL(U) \times GL(V) \times GL(W) \subseteq G_{M_{(U,V,W)}} \subset GL(U^* \otimes V) \times GL(V^* \otimes W) \times GL(W^* \otimes U).
\]

**Exercise 2.3.5.2:** Let \( U = V = W \) so we are dealing with multiplication of square matrices. Show that \( M_{(n)} \) is also invariant under the group \( \mathbb{Z}_3 \) of cyclic permutations of the three matrices. Since \( \text{trace } X^T = \text{trace } X \) show that there is also a \( \mathbb{Z}_2 \)-invariance, but note that this \( \mathbb{Z}_2 \) is not contained in the \( \mathfrak{S}_3 \) permuting the factors.

We conclude there is a map \( \rho : GL_n^3 \times \mathbb{Z}_3 \times \mathbb{Z}_2 \rightarrow G_{M_{(n)}} \), where, recalling \( a = b = c = n^2 \), \( G_{M_{(n)}} \subset GL(A) \times GL(B) \times GL(C) \) denotes the subgroup preserving the tensor \( M_{(n)} \).

### 2.4. Definitions and examples from algebraic geometry

Standard references for this material are [Har95, Mum95, Sha94]. The first is very good for examples, the second and third have clean proofs, with the proofs in the second more concise and those in the third more elementary.

**2.4.1. Varieties.** For our purposes, an algebraic variety \( \hat{X} \subset V \) is defined to be the common zero set of a collection of homogeneous polynomials on the vector space \( V \). Since we only deal with homogeneous polynomials, the zero set will be invariant under re-scaling. For this, and other reasons, it will be convenient to work in projective space \( \mathbb{P}V := (V \setminus 0)/\sim \) where \( v \sim w \) if and only if \( v = \lambda w \) for some \( \lambda \in \mathbb{C} \setminus 0 \). Write \( \pi : V \setminus 0 \rightarrow \mathbb{P}V \) for the projection.
map. For \( X \subset \mathbb{P}V \), write \( \pi^{-1}(X) \cup \{0\} =: \hat{X} \subset V \), and \( \pi(y) = [y] \). If \( \hat{X} \subset V \) is a variety, I will also refer to \( X \subset \mathbb{P}V \) as a variety. For subsets \( Z \subset V \), I’ll write \( \mathbb{P}Z \subset \mathbb{P}V \) for its image under \( \pi \).

Let \( \text{Sym}(V) := \oplus_{d=0}^{\infty} S^dV \) and note that \( \text{Sym}(V) \) is an algebra under multiplication of polynomials. The \emph{ideal} of \( X \) is defined to be

\[
I(X) := \{ P \in \text{Sym}(V^*) \mid P(x) = 0 \ \forall [x] \in X \}.
\]

Note that \( I(X) \) is indeed an ideal in the algebra \( \text{Sym}(V^*) \). Write \( \mathbb{C}[\hat{X}] := \text{Sym}(V^*)/I(X) \) and note that \( \mathbb{C}[\hat{X}] \) is a ring, called the \emph{ring of regular functions} on \( \hat{X} \).

A variety \( X \) is said to be \emph{irreducible} if it is not possible to nontrivially write \( X = Y \cup Z \) with \( Y, Z \) varieties. If \( P \in S^dV^* \) is an irreducible polynomial, then \( Z(P) \subset \mathbb{P}V \) is an irreducible variety, called a \emph{hypersurface of degree} \( d \).

We will be mostly concerned with varieties in spaces of tensors (for the study of matrix multiplication) and spaces of polynomials (for geometric complexity theory).

\[2.4. \text{ First examples of varieties.}\]

(1) Projective space \( \mathbb{P}V \subset \mathbb{P}V \).

(2) The \emph{Segre variety} of rank one tensors

\[
\sigma_1 = \text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n) := \mathbb{P}\{ T \in A_1 \otimes \cdots \otimes A_n \mid \exists a_j \in A_j \text{ such that } T = a_1 \otimes \cdots \otimes a_n \} \subset \mathbb{P}(A_1 \otimes \cdots \otimes A_n).
\]

(3) The \emph{Veronese variety}

\[
v_d(\mathbb{P}V) = \mathbb{P}\{ P \in S^dV \mid P = x^d \text{ for some } x \in V \} \subset \mathbb{P}S^dV.
\]

(4) The \emph{Grassmannian}

\[
G(k, V) := \mathbb{P}\{ T \in \Lambda^k V \mid \exists v_1, \ldots, v_k \in V \text{ such that } T = v_1 \wedge \cdots \wedge v_k \} \subset \mathbb{P}\Lambda^k V.
\]

(5) The \emph{Chow variety}

\[
Ch_d(V) := \mathbb{P}\{ P \in S^dV \mid \exists v_1, \ldots, v_d \in V \text{ such that } P = v_1 \cdots v_d \} \subset \mathbb{P}S^dV.
\]

The Grassmannian admits the geometric interpretation as the space parametrizing the \( k \)-planes through the origin in \( V \) via the correspondence \( [v_1 \wedge \cdots \wedge v_k] \leftrightarrow \text{span}\{v_1, \ldots, v_k\} \).

By definition, projective space is a variety (the zero set of no equations).

Here is a proof that the Segre is a variety: Recall the two factor Segre \( \text{Seg}(\mathbb{P}A \times \mathbb{P}B) \) is a variety defined by \( 2 \times 2 \) minors, for \( T \in A \otimes B \), of the linear map \( T_A : A^* \rightarrow B \). Now given \( T \in A_1 \otimes \cdots \otimes A_n \), consider the linear maps \( T_j : A_j^* \rightarrow A_1 \otimes \cdots \otimes A_{j-1} \otimes A_{j+1} \otimes \cdots \otimes A_n \). If \( T = a_1 \otimes \cdots \otimes a_n \), then \( T_j(a_j) = a_j(a_j)a_1 \otimes \cdots \otimes a_{j-1} \otimes a_{j+1} \otimes \cdots \otimes a_n \) so the size two minors are zero. It remains to show that if all these maps have rank one, then
Exercise 2.4.2.1: Show that the Veronese variety and Grassmanian are varieties. Explicitly:

1. Show that the Veronese is the zero set of the $2 \times 2$ minors of the linear map, for $P \in S^dV$, $P_{1,d-1} : V^* \to S^d-1V$, where the map is contraction. ($V^*$ may also be thought of as the space of first order homogeneous linear differential operators on $S^dV$.)

(2) The Grassmanian is the zero set of equations parametrized by $\Lambda^{k-2j}V^* \otimes \Lambda^{k+2j}V^*$ for $1 \leq j \leq \min\left\{\left\lfloor\frac{k-2}{2}\right\rfloor, \left\lfloor\frac{k}{2}\right\rfloor\right\}$ as follows: for $Y \in \Lambda^{k-2j}V^*$ and $Z \in \Lambda^{k-2j}V^*$, recalling Definition 2.3.4.3, define $P_Y \otimes Z(T) := (T \cdot Z)(Y \cdot T)$, the evaluation of an element of $\Lambda^{2j}V^*$ on an element of $\Lambda^{2j}V$. Note that these are quadratic equations in the coefficients of $T$. Show that these equations indeed vanish on the Grassmanian. We will prove later, using representation theory, that they span the ideal of the Grassmanian in degree two (in fact they generate the ideal), and moreover that there are redundancies in each space, but the equations from different spaces are independent of each other. 

The Chow variety is shown to be a variety in Exercise 2.4.3.1.

Exercise 2.4.2.2: Show that $v_d(\mathbb{P}V) = \text{Seg}(\mathbb{P}V \times \cdots \times \mathbb{P}V) \cap \mathbb{P}S^dV$.

Exercise 2.4.2.3: Let $X_j \subset \mathbb{P}A_j$ be varieties. Show that

$\text{Seg}(X_1 \times \cdots \times X_n) := \{[a_1 \otimes \cdots \otimes a_n] \mid [a_j] \in X_j \} \subset \text{Seg}(\mathbb{P}A_1 \otimes \cdots \otimes \mathbb{P}A_n) \subset \mathbb{P}(A_1 \otimes \cdots \otimes A_n)$

is a variety. When $A_j = A$ and $X_j = X$, I will write $X^{x_n} = \text{Seg}(X \times \cdots \times X)$.

Example 2.4.2.4. Define the subspace variety

$\hat{\text{Sub}}_{a',b',c'}(A \otimes B \otimes C) := \{T \in A \otimes B \otimes C \mid \exists A' \subset A, \dim A' = a', B' \subset B, \dim B' = b', C' \subset C, \dim C' = c', \text{ such that } T \in A' \otimes B' \otimes C'\}$

To see that $\text{Sub}_{a',b',c'}(A \otimes B \otimes C)$ is a variety, note that the existence of such an $A'$ is exactly that the map $T_A : A^* \to B \otimes C$ has rank at most $a'$ and similarly for the other factors.

We may rephrase the statement that $\overline{R}(M_{(n)}) \geq n^2$ as combining the observations $[M_{(n)}] \notin \text{Sub}_{n^2-1,n^2,n^2}(A \otimes B \otimes C)$ and $\sigma_{n^2-1} \subset \text{Sub}_{n^2-1,n^2,n^2}(A \otimes B \otimes C)$
2.4. Definitions and examples from algebraic geometry

(in fact $\sigma_{n^2-1} \subset Sub_{n^2-1,n^2-1,n^2-1}(A \otimes B \otimes C)$). This perspective illustrates a general strategy:

*Find varieties containing $\sigma_r$ whose equations are easy to describe, and then show $M_{[n]}$ (or whatever tensor we are trying to prove lower bounds for) is not a point of the intersection of the varieties.*

When trying to prove upper bounds, the goal is to find enough such varieties that their intersection is exactly $\sigma_r$.

**Exercise 2.4.2.5:** Define

$$\tilde{Sub}_r(S^dV) := \{ P \in S^dV \mid \exists V' \subset V, \dim V' = r, \text{ such that } P \in S^dV' \}. $$

Show $Sub_r(S^dV)$ is a variety.

In the cases of the Veronese, Segre and Grassmannian, the equations are the minors of some matrix and are naturally expressed as such. Most varieties do not admit such natural determinental equations.

2.4.3. Basic facts about varieties. If $X \subset \mathbb{P}W$ is a variety, $L \subset W$ is a subspace with $\mathbb{P}L \cap X = \emptyset$, and one considers the projection map $p : W \to W/L$, then $\mathbb{P}p(X) \subset \mathbb{P}W/L$ is also a variety. This is part of the Noether normalization theorem (see, e.g., [Sha94, §5.4] or [Mum95, §2C]). It is proved via elimination theory. This projection property does not hold when one works over $\mathbb{R}$, or in affine space: The surface in $\mathbb{C}^3$ given by $xz - y = 0$, when projected to the $x - y$ plane is the union of the open set $\{(x,y) \mid x \neq 0\}$ and the point $(0,0)$. Similarly, the surface in $\mathbb{R}^3$ given by $x^2 + z^2 - y^2 = 0$ when projected from $[1,0,0]$ is not a real algebraic variety.

**Exercise 2.4.3.1:** Show that if $W = V^\otimes d$ and $L$ is the $GL(V)$-complement to $S^dV$ in $V^\otimes d$, taking $p : V^\otimes d \to V^\otimes d/L \simeq S^dV$, then $p(\text{Seg}(\mathbb{P}V \times \cdots \times \mathbb{P}V)) = \text{Ch}_d(V)$. Conclude that the Chow variety is indeed a variety.

Equations for the Chow are known, see §7.3.3. However generators of the ideal of the Chow variety are not known explicitly.

2.4.4. Dimension. Informally, the dimension of a variety is the number of parameters needed to describe it locally. For example, the dimension of $\mathbb{P}V$ is $v - 1$ because in coordinates on the open neighborhood where $x_1 \neq 0$, points of $\mathbb{P}V$ have a unique expression as $[1, x_2, \ldots, x_v]$, where $x_2, \ldots, x_v$ are free parameters.

Two definitions of dimension are as follows:

First, the codimension of $X \subset \mathbb{P}V$ is the unique non-negative integer $c$ such that there exists some $\mathbb{P}^{c-1} \subset \mathbb{P}V$ with $X \cap \mathbb{P}^{c-1} = \emptyset$, but for all $\mathbb{P}^c \subset \mathbb{P}V$ one has $X \cap \mathbb{P}^c \neq \emptyset$ (see, e.g. [Har95, §11]), and the dimension
is \( \dim(X) = \dim \mathbb{P}V - c \). In this case \( \text{proj} : X \to \mathbb{P}(V/\mathbb{C}^c) \) can be defined, is surjective, and finite to one. This definition illustrates a remarkable property of projective space: that two projective varieties of complementary dimension must intersect. One defines the degree of \( X \) to be the number of points of intersection of \( X \) with a general \( \mathbb{P}^c \), equivalently the number of points in the inverse image of a general point of \( \text{proj}(X) \).

Second, the dimension of an irreducible variety \( \hat{X} \subset V \) is the dimension of the tangent space at a smooth point of \( \hat{X} \): define \( \hat{T}_x \hat{X} = \hat{T}_{[x]}X \subset V \) to be the span of the tangent vectors \( x'(0) \) to curves \( x(t) \) on \( \hat{X} \) with \( x(0) = x \). A point \( x \in \hat{X} \) is defined to be a smooth point if \( \dim \hat{T}_y \hat{X} \) is constant for all \( y \) in some neighborhood of \( x \). If \( x \) is a smooth point, \( \dim X = \dim \hat{X} - 1 = \dim \hat{T}_x \hat{X} - 1 \).

If a Zariski open subset of a variety is given parametrically, then one can calculate the tangent space to the variety via the parameter space. For example \( \hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times BPC) \) may be thought of as the image of the map

\[
A \times B \times C \to A \otimes B \otimes C
\]

\[
(a, b, c) \mapsto a \otimes b \otimes c,
\]

so to compute \( \hat{T}_{[a \otimes b \otimes c]} \hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \), take curves \( a(t) \subset A \) with \( a(0) = a \) and similarly for \( B, C \), then

\[
\frac{d}{dt}\big|_{t=0} a(t) \otimes b(t) \otimes c(t) = a' \otimes b \otimes c + a \otimes b' \otimes c' + a \otimes b \otimes c'
\]

by the Leibniz rule. Since \( a' \) can be any vector in \( A \) and similarly for \( b', c' \) we conclude

\[
\hat{T}_{[a \otimes b \otimes c]} \hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) = a \otimes b \otimes c + a \otimes B \otimes c + a \otimes b \otimes C.
\]

The right hand side spans a space of dimension \( a + b + c - 2 \) so \( \dim(\hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) = a + b + c - 3 \).

### 2.4.5. Zariski and Euclidean closure

Recall from § Zarcldef that the Zariski closure of a set can be larger than the Euclidean closure. Nevertheless, the following theorem, proved using Noether normalization, shows that in our situation, the competing definitions of closure agree:

**Theorem 2.4.5.1.** Let \( Z \subset \mathbb{P}V \) be a subset. Then the Euclidean closure of \( Z \) is contained in the Zariski closure of \( Z \). If \( Z \) is a Zariski open subset of its Zariski closure and the closure is irreducible, then the two closures coincide. The same assertions hold for subsets \( Z \subset V \).

For the proof, see, e.g. [Mum95, Thm. 2.33]. Here is a sketch: The first assertion is easy, as if \( x_j \) is a sequence of points converging to \( x_0 \) such that \( P(x_j) = 0 \), then \( P(x_0) = 0 \) because polynomials are continuous functions.

For the second, let \( X^n \subset \mathbb{P}V \) be a variety of dimension \( n \) and let \( X_0 \subset X \) be a Zariski open subset whose Zariski closure is \( X \). Let \( p \in X \setminus X_0 \), we need to construct a sequence of points in \( X_0 \) limiting to \( p \). Project \( X \) onto a \( \mathbb{P}^m \),
2.4. Definitions and examples from algebraic geometry

call the projection $\pi$. The image of $X \setminus X_0$ will be a proper subvariety of $\mathbb{P}^n$. Let $f$ be a nonzero equation in its ideal. Take a point $a \in \mathbb{P}^n$ such that $f(a) \neq 0$. Now restrict $f$ to the $\mathbb{P}^1$ spanned by $a$ and $\pi(p)$. It will have a finite number of zeros on this line, so one can take a sequence of points $a_j$ on this line, starting at $a$, limiting to $\pi(p)$ that avoids the zeros of $f$. Finally there exists a choice of points $x_j$ on $X_0$ that project to the $a_j$, which limit to $p$.

**picture to be added here**

**Definition 2.4.5.2.** A general point of a variety $Z$ is a point not lying on some explicit Zariski closed subset of $Z$. This subset is often understood from the context and so not mentioned.

**2.4.6. A geometric description of the ideal of a variety.** For a linear subspace $U \subset V$, define its annihilator

$$U^\perp := \{ \alpha \in V^* \mid \alpha(u) = 0 \ \forall u \in U \} \subset V^*.$$

**Exercise 2.4.6.1:** Show that $(U^\perp)^\perp = U$.

**Proposition 2.4.6.2.** Let $X \subset \mathbb{P}V$ be a variety. Then $I_d(X) = (\text{span}(\hat{v}_d(X)))^\perp$.

**Proof.** We have the chain of equalities:

$$P \in I_d(X) \iff P(x) = 0 \ \forall x \in \hat{X}$$

$$\iff \overline{P}(x, \ldots, x) = 0 \ \forall x \in \hat{X}$$

$$\iff \langle \overline{P}, x^d \rangle = 0 \ \forall x \in \hat{X}$$

$$\iff P \in (\hat{x}^d)^\perp \ \forall x \in \hat{X}$$

where the brackets in the third line represent the pairing of a vector in a vector space and a vector in the dual space. The intersection of all the hyperplanes $(\hat{x}^d)^\perp$ is $(\text{span}(\hat{v}_d(X)))^\perp$. 

**2.4.7. Secant Varieties.** In order to better study $\sigma_r$, which governs the complexity of $M_{(n)}$, it will be useful to place the study in the broader context of secant varieties, an extensively studied class of varieties.

Given a variety $X \subset \mathbb{P}V$, define the $X$-rank of $p \in V$, $R_X(p)$, to be the smallest $r$ such that there exist $x_1, \ldots, x_r \in \hat{X}$ such that $p$ is in the span of of $x_1, \ldots, x_r$, and the $X$-border rank $R_X(p)$ is defined to be the smallest $r$ such that there exist curves $x_1(t), \ldots, x_r(t) \in \hat{X}$ such that $p$ is in the span of the limiting plane $\lim_{t \to 0} \langle x_1(t), \ldots, x_r(t) \rangle$, where $\langle x_1(t), \ldots, x_r(t) \rangle$ is viewed as a curve the Grassmannian. I will use the same terms and notation for points in projective space.
Let $\sigma_r(X) \subset \mathbb{P}V$ denote the set of points of $X$-border rank at most $r$, called the $r$-th secant variety of $X$. (Theorem 2.4.5.1 assures us that $\sigma_r(X)$ is indeed a variety.) When $X = \sigma_1 = 	ext{Seg} (\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n)$ is the set of rank one tensors, $\sigma_r(X) = \sigma_r$.

2.4.8. Glynn’s algorithms for the permanent. **Glynn’s theorem to go here***

2.4.9. Homogeneous varieties, orbit closures, and $G$-varieties. The Segre, Veronese and Grassmannian are examples of homogeneous varieties:

**Definition 2.4.9.1.** A subvariety $X \subset \mathbb{P}V$, is homogeneous if it is a closed orbit of some point $x \in \mathbb{P}V$ under the action of some group $G \subset GL(V)$. If $G_x \subset G$ is the subgroup fixing $x$, we write $X = \mathbb{G}(x, G_x)$.

Most orbits are not varieties, so one takes their orbit closures to obtain a variety. When $d \leq v$ the Chow variety is an orbit closure, namely, if $x_1, \ldots, x_v$ is a basis of $V$,

$$Ch_d(V) = \overline{GL(V) \cdot \langle x_1^1 \cdots x_d \rangle} \subset \mathbb{P}S^d V.$$ 

**Exercise 2.4.9.2:** Show that $Ch_{d+1}(\mathbb{C}^d)$ is also an orbit closure.

When $r \leq a_i$ for $1 \leq i \leq n$, $\sigma_r(\text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n))$ is an orbit closure. Namely, letting $a_{j_1}^{\alpha_1}, \ldots, a_{j_n}^{\alpha_n}$, be a basis of $A_j$,

$$\sigma_r(\text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n)) = \overline{GL(A_1) \times \cdots \times GL(A_n) \cdot [a_1^1 \otimes \cdots a_n^1 + \cdots + a_1^r \otimes \cdots \otimes a_n^r]} \subset \mathbb{P}(A_1 \otimes \cdots \otimes A_n)$$

In particular,

$$\sigma_r(\text{Seg}(\mathbb{P}^{r-1} \times \mathbb{P}^{r-1} \times \mathbb{P}^{r-1})) = \overline{GL_r \times GL_r \times GL_r \cdot [M_{r+1}^{r+1}]}.$$ 

2.4.10. Lie algebras. In much of mathematics it is convenient to work infinitesimally to linearize problems. The same is true when studying group actions. This linearization of the group action is via its Lie algebra.

Let $\mathfrak{gl}(V) := (\text{End}(V), [, ,])$ be the endomorphisms of $V$ endowed with the product $[X, Y] := XY - YX$. With this structure, $\mathfrak{gl}(V)$ is called the Lie algebra of $GL(V)$. Define an action of $\mathfrak{gl}(V)$ on $V \otimes d$ by

$$X.(v_1 \otimes \cdots \otimes v_d) := (Xv_1) \otimes v_2 \otimes \cdots \otimes v_d + v_1 \otimes (Xv_2) \otimes \cdots \otimes v_d + \cdots + v_1 \otimes \cdots \otimes v_d \otimes (Xv_d).$$

This is an infinitesimal version of the action of $GL(V)$ on $V \otimes d$: let $g(t) \subset GL(V)$ be a curve with $g(0) = Id$ and $g'(0) = X$. Then $\frac{d}{dt} \big|_{t=0} g(t) \cdot v_1 \otimes \cdots \otimes v_d = X.(v_1 \otimes \cdots \otimes v_d)$. If $G \subset GL(V)$ is a subgroup, and we restrict to curves in $G$, the resulting subset of $\mathfrak{gl}(V)$ obtained is a sub-algebra, called the Lie algebra of $G$, and is denoted $\mathfrak{g}$. There is a $1-1$ correspondence between $GL(V)$-modules in $V \otimes d$ and $\mathfrak{gl}(V)$-modules.
2.4. Definitions and examples from algebraic geometry

In particular, the action of \( \oplus_j \mathfrak{gl}(A_j) \) on \( A_1 \otimes \cdots \otimes A_n \) is
\[
(X_1 \oplus \cdots \oplus X_n).a_1 \otimes \cdots \otimes a_n = (X_1 a_1) \otimes \cdots \otimes a_n + a_1 \otimes (X_2 a_2) \otimes \cdots \otimes a_n + \cdots + a_1 \otimes \cdots \otimes a_{n-1} \otimes (X_n a_n).
\]

Exercise 2.4.10.1: Show that \( \mathfrak{sl}(U) \) defined in Proposition 2.3.2.4 is the Lie algebra of \( SL(U) := \{ g \in GL(U) \mid \det(g) = 1 \} \).

2.4.11. Tangent spaces to orbit closures. If \( X \subseteq G \cdot [v] \subseteq \mathbb{P} V \) is an orbit closure, we can compute the tangent space \( \hat{T}_v \hat{X} \subseteq V \) using the Lie algebra:
\[
\hat{T}_v \hat{X} = \mathfrak{g}.v,
\]
as these are the points \( \frac{d}{dt}|_{t=0} g(t) \cdot v \) for curves \( g(t) \subseteq G \) with \( g(0) = Id \).

Exercise 2.4.11.1: Compute \( \hat{T}_{e_1 \otimes f_1} \hat{Seg}(\mathbb{P} V \times \mathbb{P} W), \hat{T}_{e_1 \otimes f_1} \hat{Id}(\mathbb{P} V) \), and \( \hat{T}_{e_1 \otimes \cdots \otimes e_k} \hat{G}(k, V) \) by this method.

If \( X \subseteq G \cdot [v] \subseteq \mathbb{P} V \) is an orbit closure, let \( \mathfrak{g}_v := \{ Z \in \mathfrak{g} \mid Z.v = 0 \} \subseteq \mathfrak{g} \) and \( \mathfrak{g}_v \) as the Lie algebras of the stabilizers of \( v \) and \( [v] \). (Here \( x \equiv 0 \mod v \) means \( x = \lambda v \) for some \( \lambda \in \mathbb{C} \).)

Exercise 2.4.11.2: Show that \( \mathfrak{g}_v \) and \( \mathfrak{g}_v \) are indeed the Lie algebras of \( G_v \) and \( G_{[v]} \).

Exercise 2.4.11.3: Show that \( \dim X = \dim \mathfrak{g} - \dim \mathfrak{g}_v \).

2.4.12. Border rank and orbit closures. This section will be important for our study of geometric complexity theory.

Proposition 2.4.12.1. If \( T' \subseteq GL(A) \times GL(B) \times GL(C) \cdot \hat{T} \subseteq A \otimes B \otimes C \), then \( \mathcal{R}(T') \leq \mathcal{R}(T) \).

Exercise 2.4.12.2: Prove Proposition 2.4.12.1. \( \oplus \)

Consider the orbit closure of the matrix multiplication tensor
\[
GL(A) \times GL(B) \times GL(C) \cdot [M_{(U,V,W)}] \subseteq \mathbb{P}(A \otimes B \otimes C).
\]

Exercise 2.4.12.3: Let \( V \) be a \( G \)-module and let \( u, w \in V \). Show that \( w \in G \cdot v \) if and only if \( G \cdot w \subseteq G \cdot v \).

By Exercise 2.4.12.3, we may rephrase our characterization of border rank as, taking inclusions \( A, B, C \subseteq \mathbb{C}^r \),
\[
\mathcal{R}(M_{(u)}) > r \Leftrightarrow [M_{(u)}] \not\in \sigma_r(\text{Seg}(\mathbb{P} A \times \mathbb{P} B \times \mathbb{P} C))
\]
\[
\Leftrightarrow GL_r \times GL_r \times GL_r \cdot [M_{(u)}] \not\in \sigma_r(\text{Seg}(\mathbb{P}^{r-1} \times \mathbb{P}^{r-1} \times \mathbb{P}^{r-1}))
\]
\[
\Leftrightarrow GL_r \times GL_r \times GL_r \cdot [M_{(u)}] \not\in GL_r \times GL_r \times GL_r \cdot [M_{(1)}^{(r)}].
\]
The last equality indicates that this problem may be viewed as a “toy” case of GCT.

Although it is not immediately clear why the following containment is useful, we will see in §3.4 that it is an important step in implementing Strassen’s “laser” method to prove upper bounds on the exponent of matrix multiplication:

**Theorem 2.4.12.4 (Strassen [Str87]).** Set \( r = \lceil \frac{3}{4}n^2 \rceil \) and choose a linear embedding \( \mathbb{C}^r \subset \mathbb{C}^{n^2} \). Then

\[
\sigma_r(\text{Seg}(\mathbb{P}^{r-1} \times \mathbb{P}^{r-1} \times \mathbb{P}^{r-1})) \subset GL_{n^2} \times GL_{n^2} \times GL_{n^2} \cdot [M(n)],
\]

i.e.,

\[
GL_r \times GL_r \times GL_r \cdot [M^{\oplus r}_{(1)}] \subset GL_{n^2} \times GL_{n^2} \times GL_{n^2} \cdot [M(n)].
\]

**Proof.** Write \( x_{ij} \) for a basis of \( A \) etc. Then, setting \( h = \lceil \frac{3n^2}{2} \rceil + 1 \), we have

\[
\sum_{i+j+k=h} x_{ij} \otimes y_{jk} \otimes z_{ki} = \lim_{t \to 0} t^{3r} \sum_{i,j,k=1}^n (t^{\frac{1}{2}(i^2+j^2)+2ij+(\frac{h}{3} - i - j)h+r} x_{ij})
\]

\[
\otimes (t^{\frac{1}{2}(k^2+j^2)+2kj+(\frac{h}{3} - k - j)h+r} y_{jk}) \otimes (t^{\frac{1}{2}(i^2+k^2)+2ik+(\frac{h}{3} - i - k)h+r} z_{ki})
\]

because the exponent of \( t \) in each term in the summation is

\[
i^2 + j^2 + k^2 + 2(ij + jk + ik) + h^2 - 2(i + j + k)h + 3r = (i + j + k - h)^2 + 3r.
\]

and dividing by \( t^{3r} \) we obtain an exponent of \( (i + j + k - h)^2 \) which is zero exactly when \( i + j + k = h \) and is positive otherwise. Since any two indices in a zero-th order term uniquely determine the third, each variable appears in at most one term, and there are \( \lceil \frac{3}{4}n^2 \rceil \) such terms. Now just relabel. \( \square \)

**Exercise 2.4.12.5:** Verify the assertions that the exponents of \( t \) are all non-negative and that there are \( \lceil \frac{3}{4}n^2 \rceil \) choices of triples.

**Remark 2.4.12.6.** Theorem 2.4.12.4 may be interpreted as saying that one can degenerate \( M(n) \) to a tensor that computes \( \lceil \frac{3}{4}n^2 \rceil \) independent scalar multiplications. If we have any tensor realized as \( M(n) \otimes T \), the same degeneration procedure works to degenerate it to \( M^{\oplus r}_{(1)} \otimes T \), which is the key to Strassen’s “laser method”.

### 2.4.13. G-varieties

Secant varieties of the Segre variety are not orbit closures in general, but they are invariant under the action of \( G := GL(A_1) \times \cdots \times GL(A_n) \subset GL(A_1 \otimes \cdots \otimes A_n) \), in the sense that if \( x \in \sigma_r(\text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n)) \) and \( g \in G \), then \( g \cdot x \in \sigma_r(\text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n)) \).

**Definition 2.4.13.1.** A variety invariant under the action of a group is called a **G-variety**.
An elementary, but important observation is:

If $X \subset \mathbb{P}V$ is a $G$-variety, then $I(X)$ is a $G$-submodule of $\text{Sym}(V^*)$.

**Exercise 2.4.13.2:** Prove the observation. ⊠

Thus one can use representation theory to find equations for a $G$-variety $X \subset \mathbb{P}V$ by decomposing $\text{Sym}(V^*)$ as a $G$-module. Consider the example of $X = \text{Seg}(\mathbb{P}A^* \times \mathbb{P}B^*) \subset \mathbb{P}(A \otimes B)^*$, the variety of rank one linear maps. By Exercise 2.3.5.7, we have $S^2(A \otimes B) = (S^2A \otimes S^2B) \oplus (\Lambda^2A \otimes \Lambda^2B)$. Given $[a \otimes b] \in \text{Seg}(\mathbb{P}A \times \mathbb{P}B)$, we have $v_2([a \otimes b]) = [a^2 \otimes b^2]$. Using Exercise 2.4.6.2, we recover that as a $\text{GL}(A) \times \text{GL}(B)$-module, $I_2(\text{Seg}(\mathbb{P}A^* \times \mathbb{P}B^*)) = \Lambda^2A \otimes \Lambda^2B$.

**Exercise 2.4.13.3:** Determine $I_2(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ as a $\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)$-module. ⊠

In order to make use of representation theory in the study of matrix multiplication, it will be useful to have information about the decomposition of $S^d(A \otimes B \otimes C)$ as a $\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)$-module. I postpone discussion of this decomposition until §5.2.

### 2.4.14. Dimensions of secant varieties

Let $X \subset \mathbb{P}V$ be a variety. A naïve parameter count gives that if $\dim X = n$, one expects $\dim \sigma_r(X) = \min\{rn + r - 1, v - 1\}$, because to locate a point on $\sigma_r(X)$, one gets to pick $r$ points on $x$ and then a point on the $\mathbb{P}^{r-1}$ that they span. Call this number the **expected dimension** of $\sigma_r(X)$. An efficient way to compute the actual dimension is to use local differential geometry: Let $([x_1], \ldots, [x_r]) \in X^{\times r}$ be a general point in the sense that the next statement fails for a Zariski closed subset of $X^{\times r}$. If $p = x_1 + \cdots + x_r$, then the dimension of $\sigma_r(X)$ is its expected dimension minus the “number of ways” one can move the $x_i$ within $X$ such that their span still contains $p$. Terracini’s lemma is an infinitesimal version of this remark.

Consider the map

$$
\text{add} : V \times \cdots \times V \rightarrow V \\
(v_1, \ldots, v_r) \mapsto v_1 + \cdots + v_r.
$$

When restricted to $\hat{X}^{\times r}$, the image is $\hat{\sigma}_0^r(X)$. In particular any curve $w(t)$ in $\hat{\sigma}_0^r(X)$ may be written in the form $w(t) = v_1(t) + \cdots + v_r(t)$ for some curves $v_j(t) \subset \hat{X}$. Differentiating at $t = 0$, we conclude

**Lemma 2.4.14.1** (Terracini’s lemma).

$$
(2.4.2) \quad \hat{T}_{[v_1 + \cdots + v_r]} \sigma_r(X) \subseteq \hat{T}_{[v_1]} X + \cdots + \hat{T}_{[v_r]} X.
$$

Equality holds for a Zariski open subset of points $([v_1], \ldots, [v_r]) \in X^{\times r}$. 

If the ambient space is large enough, usually the tangent spaces will only intersect at the origin and we recover the computation of the expected dimension.

**Corollary 2.4.14.2.** If \((x_1, \ldots, x_r) \in (X^r)^{\text{general}}\), then \(\dim \sigma_r(X) = \dim(\hat{T}_{x_1}X + \cdots + \hat{T}_{x_r}X) - 1\).

**Exercise 2.4.14.3:** Show that the rank of a general element in \(A_1 \otimes \cdots \otimes A_n\) is at least \(\sum_j a_j - n + 1\).

**Example 2.4.14.4.** Let \(X = \text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)\). Write \(p = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2\) for a general point of \(\hat{\sigma}_2(X)\), where \(a_1, a_2\) etc. are linearly independent. Then

\[
\hat{T}_{[p]} \sigma_2(X) = \hat{T}_{[a_1 \otimes b_1 \otimes c_1]}X + \hat{T}_{[a_2 \otimes b_2 \otimes c_2]}X
\]

\[
= (a_1 \otimes b_1 \otimes C + a_1 \otimes B \otimes c_1 + A \otimes b_1 \otimes c_1)
+ (a_2 \otimes b_2 \otimes C + a_2 \otimes B \otimes c_2 + A \otimes b_2 \otimes c_2)
\]

and it is easy to see that \(\hat{T}_{[a_1 \otimes b_1 \otimes c_1]}X\) and \(\hat{T}_{[a_2 \otimes b_2 \otimes c_2]}X\) intersect only at the origin. Thus \(\dim \sigma_2(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) = 2 \dim \text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) + 1\).

With a little more work, one can show the following, which would have anticipated Strassen’s algorithm:

**Proposition 2.4.14.5.** \(\sigma_7(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) = \mathbb{P}^{63}\), so any tensor \(T \in \mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4\) is either the sum of at most seven rank one tensors, or a limit of sums of seven rank one tensors.

### 2.5. Strassen’s equations and generalizations

This section gives a history of, and presents the state of the art for, lower bounds for \(\text{R}(M_{(n)})\).

#### 2.5.1. Beyond the classical equations

Recall the equations for \(\sigma_{n-1}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))\) from §2.2.4, which were, for \(T \in A \otimes B \otimes C\), the maximal minors of the linear map \(T_A : A^* \to B \otimes C\) (I suppress reference to the obvious exchanges of \(A, B, C\) in what follows). To extract more information, let’s examine the image of this map. Assume \(b = c\) so the image is a space of linear maps \(\mathbb{C}^b \to \mathbb{C}^b\). (If \(b < c\), just restrict to some \(\mathbb{C}^b \subset C\).) If \(\text{R}(T) = b\), then \(T_A(A^*) = T(A^*)\) will be spanned by \(b\) rank one linear maps.

**Exercise 2.5.1.1:** Show, assuming \(a = b = c\) and \(T_A\) is injective, that \(\text{R}(T) = b\) if and only if \(T(A^*)\) is spanned by \(b\) rank one linear maps.
2.5. Strassen’s equations and generalizations

How can we test if the image is spanned by $b$ rank one linear maps? If $T = a_1 \otimes b_1 \otimes c_1 + \cdots + a_n \otimes b_n \otimes c_n$ with each set of vectors a basis, then

$$T(A^*) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_j \in \mathbb{C} \right\},$$

and this is the case for a general rank $a$ tensor in $\mathbb{C}^a \otimes \mathbb{C}^a \otimes \mathbb{C}^a$. That is, $T_A(A^*) \subset B \otimes C$, when $T$ has border rank $a$ lies in the Zariski closure of the subspaces that, under the action of $GL(B) \times GL(C)$ are diagonalizable. So we have the problem: determine polynomials on $B \otimes C$ that vanish of the set of diagonalizable subspaces.

A set of equations whose zero set is exactly the Zariski closure of the diagonalizable matrices is not known! What follows are some equations. Recall that $B \otimes C = \text{Hom}(B^*, C)$. If instead we had $\text{Hom}(C, C)$, a necessary condition for endomorphisms to be simultaneously diagonalizable is that they must commute, and the algebraic test for a subspace $U \subset \text{Hom}(C, C)$ to be abelian is simple: the commutators $[X_i, X_j] := X_i X_j - X_j X_i$ must vanish on a basis $X_1, \ldots, X_u$ of $U$. (Note that commutators only make sense for maps from a vector space to itself.) These degree two equations exactly characterize abelian subspaces. We do not have maps from a vector space to itself, but we can fix the situation if there exists $\alpha \in A^*$ such that $T_A(\alpha) : B^* \rightarrow C$ is invertible, as then we could test if the commutators $[T_A(\alpha_1) T_A(\alpha)^{-1}, T_A(\alpha_2) T_A(\alpha)^{-1}]$ are zero. So we now have a test, but it is not expressed in terms of polynomials on $A \otimes B \otimes C$, and we cannot apply it to all tensors.

To fix these problems, we need a substitute for the inverse of a linear map. Recall that the inverse of a $b \times b$ invertible matrix $X$ is $X^{-1} = \frac{1}{\det(X)} \Delta^T_X$, where $\Delta_X$ is the cofactor matrix of $X$, whose entries are the minors of size $b - 1$ of $X$. In bases,

$$[T_A(\alpha_1) T_A(\alpha)^{-1}, T_A(\alpha_2) T_A(\alpha)^{-1}] = [T_A(\alpha_1) \frac{1}{\det(T_A(\alpha))} \Delta^T_{T_A(\alpha)}, T_A(\alpha_2) \frac{1}{\det(T_A(\alpha))} \Delta^T_{T_A(\alpha)}]$$

and, as long as $\det(T_A(\alpha)) \neq 0$ (i.e., as long as the expression makes sense), we can multiply the expression by $\det(T_A(\alpha))^2$ to obtain the equations

$$[T(\alpha_1) \Delta^T_{T(\alpha)}, T(\alpha_2) \Delta^T_{T(\alpha)}] = 0.$$

The equations will not be robust if $T(\alpha)$ is not of full rank.

Here is a coordinate free description of the cofactor matrix which will facilitate generalizations, as well as determining redundancies in equations. Recall from Exercise 2.3.5(6) that a linear map $f : V \rightarrow W$ induces linear
maps \( f^k : \Lambda^k V \to \Lambda^k W \). Assume \( v = w \). Fixing a volume form on \( V \), i.e., a nonzero element of \( \Omega \in \Lambda^V W^* \), we may identify \( \Lambda^V W \simeq V^* \) by the map
\[
\Lambda^V W \ni \alpha \mapsto \alpha \cdot \Omega.
\]

Fixing volume elements and the corresponding identifications, consider \( f^\wedge_{V^{-1}} : V^* \to W^* \), and taking transpose, \( (f^\wedge_{V^{-1}})^T : W \to V \).

**Exercise 2.5.1.2:** What is the relationship between \( f^\wedge_{V^{-1}} \) and \( f^{-1} \) when \( f \) is invertible? What happens when \( \det(f) = 0 \)?

### 2.5.2. Strassen’s test as polynomials.

Combining the discussions of the previous two subsections, we obtain:

**Proposition 2.5.2.1.** Let \( T \in A \otimes B \otimes C \) and assume \( b = c \). Then \( R(T) \leq b \) implies that for all \( \alpha, \alpha_1, \alpha_2 \in A^* \), the linear map \([(T(\alpha)^\wedge b^{-1})^T T(\alpha_1), (T(\alpha)^\wedge b^{-1})^T T(\alpha_2)]\) is zero.

To see the degree, these equations are of degree one in the entries of the commutator, the entries of \( T(\alpha_j) \) are linear in the entries of \( T \), and the entries of \((T(\alpha)^\wedge b^{-1})^T \) are of degree \( b - 1 \) in the entries of \( T \), so the polynomials are of degree \( 2b \). We can lower the degree by observing that
\[
[(T(\alpha)^\wedge b^{-1})^T T(\alpha_1), (T(\alpha)^\wedge b^{-1})^T T(\alpha_2)]
= (T(\alpha)^\wedge b^{-1})^T T(\alpha_1)T(\alpha_2)(T(\alpha)^\wedge b^{-1})^T T(\alpha_2) - (T(\alpha)^\wedge b^{-1})^T T(\alpha_2)(T(\alpha)^\wedge b^{-1})^T T(\alpha_1)
\]
so we can work instead with
\[
T(\alpha_1)(T(\alpha)^\wedge b^{-1})^T T(\alpha_2) - T(\alpha_2)(T(\alpha)^\wedge b^{-1})^T T(\alpha_1),
\]
dropping the degree to \( b + 1 \).

### 2.5.3. Strassen’s equations.

If \( T \in A \otimes B \otimes C \) is “close to” having rank \( a = b = c \), one expects, using \( \alpha \) with \( T(\alpha) \) invertible, that \( T(A^*) T(\alpha)^{-1} \) will be “close to” being abelian. The following theorem makes this precise:

**Theorem 2.5.3.1 (Strassen).** [Str83] Let \( T \in A \otimes B \otimes C \) and assume \( b = c \).

Then \( R(T) \leq r \) implies that for all \( \alpha, \alpha_1, \alpha_2 \in A^* \), the size \( 2(r - b) + 1 \) minors of the linear map
\[
(2.5.1)
\]
\[
STR_{\alpha, \alpha_1, \alpha_2}(T) := T(\alpha_1)(T(\alpha)^\wedge b^{-1})^T T(\alpha_2) - T(\alpha_2)(T(\alpha)^\wedge b^{-1})^T T(\alpha_1)
\]
are all zero. In particular
\[
R(T) \geq \frac{1}{2} \text{rank}(STR_{\alpha, \alpha_1, \alpha_2}(T)) + b.
\]

I prove Theorem 2.5.3.1 for the case of the maximal minors in 2.5.4 below and in general in 2.7.2.
We now have potential tests for border rank for tensors in $\mathbb{C}^N \otimes \mathbb{C}^N \otimes \mathbb{C}^N$ up to $r = \frac{3}{2}N$.

Strassen uses Theorem 2.5.3.1 to show that $R(M(\langle n \rangle)) \geq \frac{3}{2}n^2$.

**Exercise 2.5.3.2:** Prove $R(M(\langle n \rangle)) \geq \frac{3}{2}n^2$. ⊙

A proof that $R(M(\langle n \rangle)) \geq \frac{3}{2}n^2$ using representation theory is given in §2.5.5.

A natural question arises: we actually have three sets of such equations - are the three sets of equations the same or different? We should have already asked this question for the three types of usual flattenings: are the equations coming from the minors of $T_A, T_B, T_C$ the same or different? Representation theory will enable us to answer these questions.

**Remark 2.5.3.3.** One can generalize Strassen’s equations by taking higher order commutators, see [LM08]. These generalizations do give new equations, but they do not give equations for border rank beyond the $\frac{3}{2}b$ of Strassen’s equations.

2.5.4. Reformulation and proof of Strassen’s equations. Let $\dim A = 3$ and $\dim B = \dim C = b$. We augment the linear map $T_B: B^* \to A \otimes C$ by tensoring it with $\text{Id}_A$, to get a linear map

$$T_B \otimes \text{Id}_A: B^* \otimes A \to A \otimes A \otimes C.$$  

So far we have done nothing interesting, but the target of this map decomposes as $(\Lambda^2 A \otimes C) \oplus (S^2 A \otimes C)$, and we may project onto these factors. Write the projections as:

$$T_{BA}^\wedge: A \otimes B^* \to \Lambda^2 A \otimes C \quad \text{and} \quad T_{BA}^\circ: A \otimes B^* \to S^2 A \otimes C.$$  

**Exercise 2.5.4.1:** Show that if $T = a \otimes b \otimes c$ is a rank one tensor, then $\text{rank}(T_{BA}^\wedge) = 2$ and $\text{rank}(T_{BA}^\circ) = 3$.

Exercise 2.5.4.1 implies:

**Proposition 2.5.4.2.** If $T$ has rank $r$, $\text{rank}(T_{BA}^\wedge) \leq 2r$ and $\text{rank}(T_{BA}^\circ) \leq 3r$.

Since the dimension of the sources are $3b$, by this method one respectively gets potential equations for $\sigma_r(\text{Seg}(\mathbb{P}^2 \times \mathbb{P}^{b-1} \times \mathbb{P}^{b-1}))$ up to $r = \frac{3}{2}b - 1$ and $r = b - 1$. Thus only the first gives interesting bounds. The first set is Strassen’s equations, as I now show.

**Remark 2.5.4.3.** As presented, this derivation of the equations seems a bit “rabbit out of the hat”. However it comes from a long tradition of finding determinantal equations for algebraic varieties that is out of the scope of this book. For the experts, given a variety $X$ and a subvariety $Y \subset X$, one way to find defining equations for $Y$ is to find vector bundles $E,F$
over $X$ and a vector bundle map $\phi : E \to F$ such that $Y$ is realized as the degeneracy locus of $\phi$, that is, the set of points $x \in X$ such that $\phi_x$ drops rank. Strassen’s equations in the partially symmetric case had been discovered by Barth in this context. Variants of Strassen’s equations date back to Frahm-Toeplitz and Aronhold. See [Lan12, §3.8.5] for a discussion. We will also see in §5.3.3 and §5.3.2 two different ways of deriving Strassen’s equations via representation theory.

Let $a_1, a_2, a_3$ be a basis of $A$, with dual basis $\alpha^1, \alpha^2, \alpha^3$ of $A^*$ so $T \in A \otimes B \otimes C$ may be written as $T = a_1 \otimes T(\alpha^1) + a_2 \otimes T(\alpha^2) + a_3 \otimes T(\alpha^3)$. Then $T_A^\Lambda$ will be expressed by a $3b \times 3b$ matrix. Ordering the basis of $A \otimes B^*$ by $a_3 \otimes \beta^1, \ldots, a_3 \otimes \beta^b, a_2 \otimes \beta^1, \ldots, a_2 \otimes \beta^b, a_1 \otimes \beta^1, \ldots, a_1 \otimes \beta^b$, and that of $\Lambda^2 A \otimes C$ by $(a_1 \wedge a_2) \otimes c_1, \ldots, (a_1 \wedge a_2) \otimes c_b, (a_1 \wedge a_3) \otimes c_1, \ldots, (a_1 \wedge a_3) \otimes c_b, (a_2 \wedge a_3) \otimes c_1, \ldots, (a_2 \wedge a_3) \otimes c_b$, we obtain the block matrix

$$\begin{equation}
T_A^\Lambda := T_B^\Lambda = \begin{pmatrix}
0 & T(\alpha^1) & -T(\alpha^2) \\
T(\alpha^2) & T(\alpha^3) & 0 \\
T(\alpha^1) & 0 & T(\alpha^3)
\end{pmatrix}
\end{equation}
$$

Recall the following basic identity about determinants, assuming $W$ is invertible:

$$\begin{equation}
\det \begin{pmatrix}
X & Y \\
Z & W
\end{pmatrix} = \det(W) \det(X - YW^{-1}Z).
\end{equation}
$$

Assume $T(\alpha^3)$ is invertible and use the $(b, 2b) \times (b, 2b)$ blocking (so $X = 0$ in (2.5.4)) to obtain

$$\begin{equation}
\det \text{Mat}(T_A^\Lambda) = \det(T(\alpha^1)T(\alpha^3)^{-1}T(\alpha^2) - T(\alpha^2)T(\alpha^3)^{-1}T(\alpha^1))
\end{equation}
$$

Equation (2.5.5) shows the new formulation is equivalent to the old, at least in the case of maximal rank, and combined with Proposition 2.5.4.2, proves Theorem 2.5.3.1.

### 2.5.5. Proof that $R(M_{(n)}) \geq \frac{3}{2}n^2$ using representation theory.

**Lemma 2.5.5.1.** Let $T_A^\Lambda : B^* \otimes A \to \Lambda^2 A \otimes C$ be injective, and let $A' \subset A^*$ be a general subspace. Consider the tensor $T' := T|_{A' \otimes B^* \otimes C^*}$. Then $(T')^\Lambda_A : B^* \otimes (A')^* \to \Lambda^2 (A')^* \otimes C$ is injective.

**Proof.** Without loss of generality, assume $A'$ is a hyperplane, choose a complementary hyperplane $\tilde{A}$ to $(A')^\perp$ in $A$ and identify it with $(A')^*$. Consider $(T')^\Lambda_A$ as the composition of the injection $B^* \otimes \tilde{A} \to \Lambda^2 \tilde{A} \otimes C$ with the projection to $\Lambda^2 \tilde{A} \otimes C$. The image of the first map is contained in $(\Lambda^2 \tilde{A} \otimes C) \oplus ((A')^\perp \otimes \tilde{A} \otimes C)$. The kernel of the projection map is the intersection of the image with $(A')^\perp \otimes \tilde{A} \otimes C$. If we modify $\tilde{A}$ by a small multiple
2.5. Strassen’s equations and generalizations

...nontrivially.

Recall that $A_M$.

Recalling that $U_\Lambda \text{ vector in each irreducible component is not in the kernel. By Exercise 2.3.5.7,}$

2.5.6. Koszul flattenings. The reformulation of Strassen’s equations sug-

A (2.5.6) T obtained by first taking $T$ summation convention, then projecting to $\Lambda U$.

Proposition 2.5.6.1.

This is a $GL(U) \times GL(V)$-module map, so we can proceed by decomposing the source into irreducible representations, and just show that one vector in each irreducible component is not in the kernel. By Exercise 2.3.5.7, $\Lambda^2(U^* \otimes V) = (S^2U^* \otimes \Lambda^2 V) \oplus (\Lambda^2U^* \otimes S^2V)$, and $V \otimes V = \Lambda^2V \oplus S^2V$. The source is thus $(U^* \otimes S^2V) \oplus (U^* \otimes \Lambda^2 V)$. Now just check that the vectors $v \otimes (\alpha \otimes v)$ and $v_1 \otimes (\alpha \otimes v_2) - v_2 \otimes (\alpha \otimes v_1)$ do not map to zero.

2.5.6. Koszul flattenings. The reformulation of Strassen’s equations suggests the following generalization: let $\dim A = 2p+1$, let $b = c$ and consider

(2.5.6) $T_A^{\Lambda^p} : B^* \otimes \Lambda^p A \to \Lambda^{p+1} A \otimes C$

obtained by first taking $T_B \otimes Id_{\Lambda^p A} : B^* \otimes \Lambda^p A \to \Lambda^p A \otimes A \otimes C$, and then projecting to $\Lambda^{p+1} A \otimes C$.

If $a_i, b_j, c_k$ are bases of $A, B, C$ and $T = t^{ijk} a_i \otimes b_j \otimes c_k$, where I use the summation convention, then

(2.5.7) $T_A^{\Lambda^p}(\beta \otimes f_1 \wedge \cdots \wedge f_p) = t^{ijk} \beta(b_j) a_i \wedge f_1 \wedge \cdots \wedge f_p \otimes c_k$.

The map $T_A^{\Lambda^p}$ is called a Koszul flattening. Note that if $T = a \otimes b \otimes c$ has rank one, then rank($T_A^{\Lambda^p}$) = $\binom{2p}{p}$ as the image is $a \wedge \Lambda^p A \otimes c$. By linearity of the map $T \mapsto T_A^{\Lambda^p}$ we conclude:

Proposition 2.5.6.1. [LO] Let $T \in A \otimes B \otimes C$. If rank($T_A^{\Lambda^p}$) $\geq r$, then

$$R(T) \geq \frac{r}{\binom{2p}{p}}.$$

Since the source (resp. target) has dimension $\binom{2p+1}{p} b$ (resp. $\binom{2p+1}{p} c$), assuming $b \leq c$, we potentially obtain equations for $\sigma_c(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$
up to
\[ r = \left( \frac{2p+1}{p} \right) b - 1 = \frac{2p + 1}{p + 1} b - 1. \]

Just as with Strassen’s equations (case \( p = 1 \)), if \( \dim A > 2p + 1 \), one obtains the best bound for these equations by restricting to subspaces of \( A^* \) of dimension \( 2p + 1 \).

The skew-symmetrization map \( A \otimes \Lambda^p A \to \Lambda^{p+1} A \) is a surjective \( GL(A) \)-module map, so its kernel is a \( GL(A) \)-module generalizing \( S_{21} A \) that we saw in \( \S 4.2.2 \). Its isomorphism class as a \( GL(A) \)-module is denoted \( S_{2p-1}^{p-1} \).

In \( \S 5.2 \) we will see that it is irreducible.

Thus the projection map, \( A \otimes \Lambda^p A \otimes C \to \Lambda^{p+1} A \otimes C \), since it is a \( GL(A) \times GL(C) \)-module map, has kernel isomorphic to \( S_{2p-1}^{p-1} \otimes C \), so one way to determine the rank of \( T_A^p \) would be to compute the dimension of \( S_{2p-1} A \otimes C \cap (T_B(B^*) \otimes \Lambda^p A) \).

2.5.7. Exercises.

(1) Prove an analogue of Lemma 2.5.5.1 for Koszul flattenings.

(2) Consider the symmetric cousin of Strassen’s equations: for \( P \in S^3 V \subset V \otimes V \otimes V \), first take \( P_{1,2} \otimes \text{Id}_V : V^* \otimes V \to S^2 V \otimes V \), then considering \( S^2 V \subset V \otimes V \), skew symmetrize on the second and third factors, to obtain a map \( P^\wedge : V^* \otimes V \to V \otimes \Lambda^2 V \). Now assume further that \( \dim V = 3 \), so, after the choice of a volume element, \( \Lambda^2 V = V^* \). Show that the resulting matrix in properly ordered bases is skew-symmetric, and thus its determinant is zero. Moreover, one can show that all the Pfaffians of size 8 centered about the diagonal are equivalent and give rise to a polynomial called the Aronhold invariant. What bound does this equation give on symmetric border rank? See [Lan12, \S 3.10.1] for more details.

(3) More generally, consider the symmetric cousins of Koszul flattenings: for \( P \in S^d V \), first take the map \( P_{d-k,k} : S^{d-k} V^* \to S^k V \), tensor it with \( \text{Id}_{\Lambda^p V} \), and project to get a map
\[ P_{d-k,k}^\wedge : S^{d-k} V^* \otimes \Lambda^p V \to S^{k-1} V \otimes \Lambda^{p+1} V. \]

If \( P = x^d \), what is the rank of \( P_{d-k,k}^\wedge \)? (For those familiar with the rudiments of differential geometry, note that the map \( S^k V \otimes \Lambda^p V \to S^{k-1} V \otimes \Lambda^{p+1} V \) is just the exterior derivative.) We will see later that \( S^k V \otimes \Lambda^p V \) decomposes into the direct sum of two irreducible modules, called \( S_{k,1} V \) and \( S_{k+1,1} V \), so \( P_{d-k,k}^\wedge \) can at best surject onto \( S_{k,1} V \).
2.5. Strassen’s equations and generalizations

(4) For those handy with computers: determine the rank of \((M_{(3)})^\wedge A\).
What lower bound does this give on the border rank?

Answer for those not handy: 14, the same as Strassen.

(5) For those handy with computers: determine the rank of \((M_{(4)})^\wedge A\).
What lower bound does this give on the border rank?

This beats Strassen by one. Lickteig’s bound [Lic84]: \(\mathbf{R}(M_{(n)}) \geq \frac{3n^2}{2} + \frac{n}{2} - 1\) gives the same answer when \(n = 4\).

(6) For those handy with computers: determine the rank of \((M_{(5)})^\wedge A\).
What lower bound does this give on the border rank?

This also ties Lickteig’s bound.

(7) For those handy with computers: determine the rank of \((M_{(6)})^\wedge A\).
What lower bound does this give on the border rank?

Now it looks like something interesting, as \(58 > 56 = \frac{3(6^2)}{2} + \frac{6}{2} - 1\).

How would one possibly determine the rank for arbitrary \(n\)? If we knew how to decompose the source and target into irreducible representations, we could then see what modules must be in the kernel of \((M_{(n)})^\wedge \left[\frac{n^2}{2}\right]^{-1}\), and then for the remaining modules, we could check on one vector. However, by restricting to a well-chosen subspace, one can obtain a good lower bound, which I now present.

2.5.8. Koszul flattenings and matrix multiplication. Our map is

\((M_{(U,V,W)})^\wedge A : V \otimes W^* \otimes \Lambda^p(U^* \otimes V) \to \Lambda^{p+1}(U^* \otimes V) \otimes (W^* \otimes U)\).

The presence of \(Id_W = Id_{W^*}\) means the map factors as \((M_{(U,V,W)})^\wedge A = (M_{(u,v,1)})^\wedge A \otimes Id_{W^*}\), where

\[(2.5.8)\]

\[(M_{(u,v,1)})^\wedge A : V \otimes \Lambda^p(U^* \otimes V) \to \Lambda^{p+1}(U^* \otimes V) \otimes U.\]

\[v \otimes (\xi^1 \otimes e_1) \wedge \cdots \wedge (\xi^p \otimes e_p) \mapsto \sum_{s=1}^n u_s \otimes (\gamma^s \otimes v) \wedge (\xi^1 \otimes e_1) \wedge \cdots \wedge (\xi^p \otimes e_p).\]

where \(u_1, \ldots, u_n\) is a basis of \(U\) with dual basis \(\gamma^1, \ldots, \gamma^u\) of \(U^*\), so \(Id_U = \sum_{s=1}^n \gamma^s \otimes u_s\).

**Theorem 2.5.8.1.** [LO] Let \(n \leq m\). Then

\[\mathbf{R}(M_{(m,n,l)}) \geq \frac{nl(n + m - 1)}{m}.\]

In particular \(\mathbf{R}(M_{(n)}) \geq 2n^2 - n\).
Proof. Set \( \dim U = n \), \( \dim V = m \), \( \dim W = l \). The idea is to choose a subspace \( A' \subset A^* \) of dimension \( 2p + 1 \) on which \( (M_{(1,m,n)}|_{A' \otimes U \otimes V^*})^p_A \) becomes injective for \( p = n - 1 \), where \( \hat{A} \subset A \) is a choice of inclusion \( (A')^* \subset A \). We need a choice that on one hand is generic enough to force injectivity, but on the other is easy to compute with. We will use representation theory to help us by giving \( A \) extra structure as a \( G \)-module for some \( G \subset GL(A) \).

In fact we’ll just use \( G = SL_2 \) as follows: Take a vector space \( E \) of dimension 2, and fix isomorphisms \( U \simeq S^{n-1}E, V \simeq S^{m-1}E^* \). Let \( A' \) be the \( SL(E) \)-submodule \( S^{m+n-2}E^* \subset S^{n-1}E^* \otimes S^{m-1}E^* = U^* \otimes V \). The \( SL(E) \)-module structure of \( A \) determines an inclusion \( (A')^* \subset A \).

Remark 2.5.8.2. The Koszul flattenings actually give equations for border limits, as the image of a linear map is closed.

Consider the transposed map
\[
((M_{(1,m,n)}|_{A' \otimes U \otimes V^*})^p_A)^T : S^{m-1}E^* \otimes \Lambda^n S^{m+n-2}E \rightarrow S^{n-1}E \otimes \Lambda^{n-1} S^{m+n-2}E
\]
\[
g \circ (f_1 \wedge \cdots \wedge f_n) \mapsto \sum_{i=1}^n (-1)^{i-1} (g \circ f_i) \otimes f_1 \wedge \cdots \hat{f}_i \cdots \wedge f_n.
\]

The map \( ((M_{(1,m,n)}|_{A' \otimes U \otimes V^*})^p_A)^T \) is surjective: Let \( \sum_{i=1}^l (m+n-2 \wedge \cdots \wedge m+n-2) \in S^{m+n-2}E \otimes \Lambda^{n-1} S^{m+n-2}E \) with \( l, l_i \in E \). Such elements span the target so it will be sufficient to show any such element is in the image. Assume first that \( l \) is distinct from the \( l_i \). Since \( n \leq m \), there is a polynomial \( g \in S^{m+n-2}E^* \) which vanishes on \( l_1, \ldots, l_{n-1} \) and is nonzero on \( l \). Then, up to a nonzero scalar, \( g \otimes (l_1^{m+n-2} \wedge \cdots \wedge l_{n-1}^{m+n-2} \wedge l^{m+n-2}) \) maps to our element.

The condition that \( l \) is distinct from the \( l_i \) may be removed by taking limits, as the image of a linear map is closed.

We conclude:
\[
\mathcal{R}(M_{(n,m,l)}) \geq 1 + \frac{\text{rank}(M_{(1,m,n)}|_{A' \otimes U \otimes V^*})^p_A}{\binom{n+m-2}{n-1}} = nl \frac{\binom{n+m-1}{n-1}}{\binom{n+m-2}{n-1}} = \frac{nl(n+m-1)}{m}.
\]

\[\square\]

Remark 2.5.8.2. The Koszul flattenings actually give equations for border rank in \( \mathbb{C}^N \otimes \mathbb{C}^N \otimes \mathbb{C}^N \) up to \( 2N - 3 \), see [Lan]. On the other hand, the map \( (M_{(n)}|_{A}^\Lambda \rightleftharpoons \Lambda^{n^2/2} - 1) A \otimes B^* \rightarrow \Lambda^{n^2/2} A \otimes C \) is far from being injective,
2.6. Lower bounds for the rank of matrix multiplication

2.6.1. The results. Any tensor $T$ with a positive dimensional symmetry group is a good candidate to have lower border rank than rank. This intuition comes from symmetry generally implying non-uniqueness of an expression of $T$ as a sum of $R(T)$ rank one tensors, as discussed in §4.2.

Lower bounds for the rank of matrix multiplication beyond the border rank arise by exploiting both a generic property and a pathological property see §5.2.10. This shows that there are nontrivial equations for border rank $2n^2 - 3$ that are satisfied by $M_{(n)}$.

2.5.9. Koszul flattenings in coordinates. To prove lower bounds on the rank of matrix multiplication, and to facilitate a comparison with Griesser’s equations discussed in §2.7.2, it will be useful to view $T^\wedge_p$ in coordinates. Let $\dim A = 2p + 1$. Write $T = a_0 \otimes X_0 + \cdots + a_{2p} \otimes X_{2p}$ where $a_j$ is a basis of $A$ with dual basis $\alpha^j$ and $X_j = T(\alpha^j)$. An expression of $T^\wedge_p$ in bases is as follows: write $a_I := a_{i_1} \wedge \cdots \wedge a_{i_p}$ for $\Lambda^p A$, require that the first $(2p-1)\binom{2p}{p}$ basis vectors of $\Lambda^p A$ have $i_1 = 0$, that the second $(2p)\binom{2p}{p}$ do not, and call these multi-indices $0J$ and $K$. Order the bases of $\Lambda^{p+1} A$ such that the first $(2p+1)\binom{2p}{p}$ multi-indices do not have 0, and the second $(2p)\binom{2p}{p}$ do, and furthermore that the second set of indices is ordered the same way as $K$, only we write $0K$ since a zero index is included. The resulting matrix is of the form

\begin{equation}
(2.5.9) \begin{pmatrix}
0 & Q \\
\bar{Q} & R
\end{pmatrix}
\end{equation}

where this matrix is blocked $(\binom{2p}{p+1}b, \binom{2p}{p}b) \times (\binom{2p}{p+1}b, \binom{2p}{p}b)$,

$$R = \begin{pmatrix}
X_0 \\
\vdots \\
X_0
\end{pmatrix},$$

and $Q, \bar{Q}$ have entries in blocks consisting of $X_1, \ldots, X_{2p}$ and zero. Thus if $X_0$ is of full rank and we change coordinates such that it is the identity matrix, so is $R$ and the determinant equals the determinant of $Q\bar{Q}$ by (2.5.4). When $p = 1$, $Q\bar{Q} = [X_1, X_2]$, and when $p = 2$

\begin{equation}
(2.5.10) \begin{pmatrix}
0 & [X_1, X_2] & [X_1, X_3] & [X_1, X_4] \\
[X_2, X_1] & 0 & [X_2, X_3] & [X_2, X_4] \\
[X_3, X_1] & [X_3, X_2] & 0 & [X_3, X_4] \\
[X_4, X_1] & [X_4, X_2] & [X_4, X_3] & 0
\end{pmatrix}
\end{equation}

In general $Q\bar{Q}$ is a block $(\binom{2p}{p-1}b \times \binom{2p}{p-1}b)$ matrix whose block entries are either zero or commutators $[X_i, X_j]$. 

2.6. Lower bounds for the rank of matrix multiplication
of $M_{(n)}$. Of these two, the pathological property is much more important, the generic property only modifies the error term.

The pathological property is that $M_{(l,m,n)}$ satisfies equations coming from the Koszul flattenings if and only if $M_{(1,m,n)}$ satisfies analogous equations (which are of lower degree).

Using the pathological property, one obtains:

**Theorem 2.6.1.1.** [Lan14b] Let $p < n - 1$. Then

$$\text{rank}(M_{(n,m,n)}) \geq \frac{2p+1}{p+1}nm + n^2 - \binom{2p+1}{p}n.$$

This gives a bound of the form $\text{rank}(M_{(n)}) \geq 3n^2 - o(n^2)$ by taking, e.g., $p = \log(\log(n))$.

The generic property is as follows: Recall that when deriving Strassen’s equations, there was a complication because we could not assume there existed an invertible linear map in $T(A^*) \subset B \otimes C$. However, if $T \in A \otimes B \otimes C$ is a tensor such that there does exist an invertible linear map in $T(A^*)$, we can simplify the equations by choosing it. The error term may be improved by examining the equations in coordinates, using the genericity property and exploiting identities about determinants. The result is:

**Theorem 2.6.1.2.** [MR13] Let $p \leq n$ be a natural number. Then

\begin{equation}
\text{rank}(M_{(n,n,n)}) \geq (1 + \frac{p}{p+1})nm + n^2 - 2\binom{2p}{p+1} - \binom{2p-2}{p-1} + 2)n.
\end{equation}

When $n = m$,

\begin{equation}
\text{rank}(M_{(n)}) \geq (3 - \frac{1}{p+1})n^2 - 2\binom{2p}{p+1} - \binom{2p-2}{p-1} + 2)n.
\end{equation}

For example, when $p = 1$ one recovers Bläser’s bound of $\frac{5}{2}n^2 - 3n$. When $p = 3$, the bound (2.6.2) becomes $\frac{11}{4}n^2 - 26n$, which improves Bläser’s for $n \geq 132$. A modification of the method also yields $\text{rank}(M_{(n)}) \geq \frac{5}{2}n^2 - 7n$.

See [MR13, Lan14b] for proofs of the modifications of the error terms.

### 2.6.2. Proof of Theorem 2.6.1.1.

The following standard Lemma, also used in [Blä03], appears in this form in [Lan12, Lemma 11.5.0.2]:

**Lemma 2.6.2.1.** Let $\mathbb{C}^a$ be given a basis $e_1, \ldots, e_a$. Given a polynomial $P$ of degree $d$ on $\mathbb{C}^a$, there exists a subset $\{e_{i_1}, \ldots, e_{i_d}\}$ such that $P |_{\langle e_{i_1}, \ldots, e_{i_d}\rangle}$ is not identically zero.

The lemma follows by simply choosing the basis vectors from a monomial that appears in $P$. For example, Lemma 2.6.2.1 implies that a quadric surface in $\mathbb{P}^3$ cannot contain six lines whose pairwise intersections span $\mathbb{P}^3$. 
Recall the Grassmannian $G(k,V) \subset \mathbb{P} \Lambda^k V$ from §2.4.2. The Plücker coordinates $(x^\mu_\alpha)$, $k + 1 \leq \mu \leq \dim V = v$, $1 \leq \alpha \leq k$ are obtained by choosing a basis $e_1, \ldots, e_v$ of $V$, centering the coordinates at $[e_1 \wedge \cdots \wedge e_k]$, and writing a nearby $k$-plane as $[(e_1 + \sum x^\mu_i e_\mu) \wedge \cdots \wedge (e_k + \sum x^\mu_i e_\mu)]$. One says a function on $G(k,V)$ is a polynomial of degree $d$ if, as a function in the Plücker coordinates, it is a degree $d$ polynomial. If the polynomial is also homogeneous in the $x^\mu_\alpha$, it is the restriction of a homogeneous degree $d$ polynomial on $\Lambda^k V$ to $G(k,V)$.

**Lemma 2.6.2.2.** Let $A$ be given a basis. Given a homogeneous polynomial of degree $d$ on the Grassmannian $G(k,A)$, there exists at most $dk$ basis vectors such that, denoting their (at most) $dk$-dimensional span by $A'$, $P$ restricted to $G(k,A')$ is not identically zero.

**Proof.** Consider the map $f : \Lambda^k A \to G(k,A)$ given by $(a_1, \ldots, a_k) \mapsto [a_1 \wedge \cdots \wedge a_k]$. Then $f$ is surjective. Take the polynomial $P$ and pull it back by $f$. Here the pullback $f^*(P)$ is defined by $f^*(P)(a_1, \ldots, a_k) := P(f(a_1, \ldots, a_k))$. The pullback is of degree $d$ in each copy of $A$. (I.e., fixing $k-1$ of the $a_j$, it becomes a degree $d$ polynomial in the $k$-th.) Now simply apply Lemma 2.6.2.1 $k$ times to see that the pulled back polynomial is not identically zero restricted to $A'$, and thus $P$ restricted to $G(k,A')$ is not identically zero. □

**Remark 2.6.2.3.** The bound in Lemma 2.6.2.2 is sharp, as give $A$ a basis $a_1, \ldots, a_n$ and consider the polynomial on $\Lambda^k A$ with coordinates $x^I = x^{i_1} \wedge \cdots \wedge x^{i_k}$ corresponding to the vector $\sum_I x^I a_{i_1} \wedge \cdots \wedge a_{i_k}$:

$$P = x^{1} \wedge x^{k+1} \wedge x^{2k} \wedge \cdots \wedge x^{(d-1)k+1} \wedge \cdots \wedge x^{dk}.$$ 

Then $P$ restricted to $G(k,\langle a_1, \ldots, a_{dk}\rangle)$ is non-vanishing but there is no smaller subspace spanned by basis vectors on which it is non-vanishing.

**Proof of Theorem 2.6.1.1.** Say $R(M_{(n,n,m)}) = r$ and write

$$(2.6.3) \quad M_{[n,n,m]} = \sum_{j=1}^{r} a_j \otimes b_j \otimes c_j.$$ 

We will show that the Koszul-flattening equation is already non-zero restricted to a subset of this expression for a judicious choice of $\tilde{A} \subset A$ of dimension $2p+1$ with $p < n - 1$. Then the rank will be at least the border rank bound plus the number of terms not in the subset. Here are the details:

Define

$$P : G(2p+1,A) \to \mathbb{C}$$

$$\tilde{A} \mapsto \det\left(\langle M_{[n,n,m]} \rangle_{A' \times B' \times C'}\right)_{\tilde{A}}^{\otimes P} : \Lambda^P \tilde{A} \otimes B^* \to \Lambda^{p+1} \tilde{A} \otimes C).$$ 

Since $p \leq n$ and we proved $P$ is not identically zero when $p = n$, $P$ will not be identically zero when $p < n$ either.
Now $P$ is a polynomial of degree $(2^{p+1})\frac{nm}{p} > nm$, so at first sight, e.g., when $m \sim n$, Lemma 2.6.2.2 will be of no help because $dk > \dim A = n^2$, but since
\[(M_{(n,n,m)}|_{A' \times B' \times C'})^p_{\tilde{A}} = (M_{(n,n,1)}|_{A' \times U' \times V'})^p_{\tilde{A}} \otimes \text{Id}_{W'},\]
we actually have $P = \tilde{P}_{nm}$, where
\[
\tilde{P} : G(2p+1, A) \to \mathbb{C}
\]
\[
\tilde{A} \mapsto \det((M_{(n,n,1)}|_{A' \times U' \times V'})^p_{\tilde{A}} : \Lambda^p \tilde{A} \otimes U' \to \Lambda^{p+1} \tilde{A} \otimes V).
\]
Hence we may work with $\tilde{P}$ which is of degree $(2^{p+1})\frac{n}{p}$ which will be less than $n^2$ if $p$ is sufficiently small. Since $(M_{(n,n,m)}|_A) : A' \to B \otimes C$ is injective, some subset of the $a_j$ forms a basis of $A$. Lemma 2.6.2.2. implies that there exists a subset of those basis vectors of size $dk = (2^{p+1})\frac{n}{p}(2p+1)$, such that if we restrict to terms of the expression (2.6.3) that use only $a_j$ whose expansion in the fixed basis has nonzero terms from that subset of $dk$ basis vectors, calling the sum of these terms $M'$, we have $\mathbf{R}(M') \geq \frac{2p+1}{p+1} nm$. Let $M''$ be the sum of the remaining terms in the expression. There are at least $n^2 - (2^{p+1})\frac{n}{p}(2p+1)$ of the $a_j$ appearing in $M''$ (the terms corresponding to the complementary basis vectors). Since we assumed we had an optimal expression for $M_{(n,n,m)}$, we have
\[
\mathbf{R}(M_{(n,n,m)}) = \mathbf{R}(M') + \mathbf{R}(M'') \geq \frac{2p+1}{p+1} nm + \left[n^2 - (2^{p+1})\frac{n}{p}(2p+1)\right].
\]

The further lower bounds are obtained by lowering the degree of the polynomial by localizing the equations. An easy such localization is to set $X_0 = \text{Id}$ which reduces the determinant of (2.5.9) to that of (2.5.10) when $p = 2$ and yields a similar reduction of degree in general. Further localizations both reduce the degree and the size of the Grassmannian, both of which improve the error term.

2.7. Additional equations for $\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$

There are numerous motivations for studying equations for $\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ from algebraic geometry, signal processing, quantum information theory etc. Two motivations coming from matrix multiplication that we have discussed are: First, to prove lower bounds, if we have $P \in I(\sigma_r)$ such that $P(M_{(n)}) \neq 0$, we know $\mathbf{R}(M_{(n)}) > r$. Second, to prove upper bounds, it would suffice to find set-theoretic equations for $\sigma_R$, that is, polynomials $P_1, \ldots, P_q$ such that there common zero set is exactly $\sigma_R$, and then show
2.7. Additional equations for \( \sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \)

\[ P_j(M_{(u)}) = 0 \text{ for all } j = 1, \ldots, q \] to conclude \( \mathbf{R}(M_{(u)}) \leq R \). As I describe below, for the second problem, something weaker will suffice (because \( M_{(u)} \) is 1-generic), but even the easier problem seems currently out of reach in the range of interest. Despite this, even small results in this direction, that I believe are within reach, will be useful for proving upper bounds on the exponent as I will explain in our discussion of upper bounds.

We will see further techniques for finding equations after we have covered Schur-Weyl duality in \( \S 5.2.1 \). In this section I present equations of Griesser and Friedland, each of which illustrates possible further paths towards finding new equations. I first discuss some generalities regarding reducing the problem of finding equations for \( \sigma_r \) to questions in linear algebra.

Recall the classical flattenings: \( T_A : A^* \rightarrow B \otimes C \). If \( T_A \) is injective, to gain further information, we examined the nature of the image (or equivalently, the kernel of \( T_A^* : B^* \otimes C^* \rightarrow A \)). All equations I am aware of are of this nature, in analogy with the study of secondary characteristic classes. Friedland’s equations discussed below examine the kernel of \( T_A^* \) to obtain further information.

2.7.1. Multi-linear algebra via linear algebra. The following result shows that when studying the rank and border rank of a tensor \( T \in A \otimes B \otimes C \), there is no loss of information in restricting attention to \( T(A^*) \): 

\[ \text{Theorem 2.7.1.1. Let } T \in A \otimes B \otimes C, \text{ Then } R(T) \text{ equals the number of rank one matrices needed to span (a space containing) } T(A^*) \subset B \otimes C \text{ (and similarly for the permuted statements).} \]

Let \( Z_r \subset G(a, B \otimes C) \) denote the set of \( a \)-planes in \( B \otimes C \) that can be spanned by \( r \) rank one elements (so \( R(T) \leq r \) if and only if \( T \in Z_r \)). Then \( R(T) \leq r \) if and only if \( T(A^*) \in Z_r \).

\[ \text{Proof. Let } T \text{ have rank } r \text{ so there is an expression } T = \sum_{i=1}^r a_i \otimes b_i \otimes c_i. \text{ (I remind the reader that the vectors } a_i \text{ need not be linearly independent, and similarly for the } b_i \text{ and } c_i. \) \] Then \( T(A^*) \subseteq \langle b_1 \otimes c_1, \ldots, b_r \otimes c_r \rangle \) shows that the number of rank one matrices needed to span \( T(A^*) \subset B \otimes C \) is at most \( R(T) \).

On the other hand, say \( T(A^*) \) is spanned by rank one elements \( b_1 \otimes c_1, \ldots, b_r \otimes c_r \). Let \( a^1, \ldots, a^a \) be a basis of \( A^* \), with dual basis \( a_1, \ldots, a_a \) of \( A \). Then \( T(a^i) = \sum_{s=1}^r x^i_s b_s \otimes c_s \) for some constants \( x^i_s \). But then \( T = \sum_{s=1}^r a_i \otimes (x^i_s b_s \otimes c_s) = \sum_{s=1}^r (\sum_i x^i_s a_i) \otimes b_s \otimes c_s \) proving \( R(T) \) is at most the number of rank one matrices needed to span \( T(A^*) \subset B \otimes C \).

\[ \text{Exercise 2.7.1.2: Prove the border rank assertion.} \]
Call $T \in A \otimes B \otimes C$ concise if it does not lie in any non-trivial subspace variety, i.e., if the maps $T_A, T_B, T_C$ are injective. For a concise tensor $T$, we will say $T$ is $1_A$-generic if $T(A^*)$ contains an element of maximal rank, and that $T$ is $1$-generic if it is $1_A, 1_B$ and $1_C$-generic.

**Exercise 2.7.1.3:** Given $T \in \mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m = A \otimes B \otimes C$ that is concise, show that $\mathbb{P}(T^*) \cap \text{Seg}(\mathbb{P}B \times \mathbb{P}C) = \emptyset$ implies $R(T) > m$. ⊠

All the equations we have seen so far arise as Koszul flattenings. Koszul flattenings provide robust equations only if $T$ is $1_A, 1_B$ or $1_C$-generic, because otherwise the presence of $T(A)^* T(\alpha)^{-1}$ in the expressions make them likely to vanish. When $T$ is $1_A$-generic, the Koszul flattenings $T_A^p : \Lambda^p A \otimes B^* \rightarrow \Lambda^{p+1} A \otimes C$ provide measures of the failure of $T(A^*) T(\alpha)^{-1} \subset \text{End}(B)$ to be an abelian subspace.

Fortunately $M(n)$ is $1$-generic. In order to test for upper bounds for a $1$-generic tensor, we only need equations for $\sigma_r$ such that the intersection of the zero set of the equations with the set of $1$-generic tensors is the intersection of $\sigma_r$ with the set of $1$-generic tensors.

How far are our equations from having this property? Let’s consider the first case where Strassen’s commutator is identically zero, i.e., $T(A^*) T(\alpha)^{-1} \subset \text{End}(B)$ is abelian.

**Proposition 2.7.1.4.** [LM] There exist $1_A$-generic tensors in $\mathbb{C}^a \otimes \mathbb{C}^a \otimes \mathbb{C}^a$ satisfying Strassen’s equations of border rank $\frac{a^2}{8}$.

**Proof.** Assume $a = b = c$. Consider $T$ such that

\[
T(A^*) \subset \begin{pmatrix}
a_1 & * & \cdots & * \\
* & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
* & \cdots & * & a_1
\end{pmatrix}
\]

and set $a_1 = 0$. We obtain a generic tensor in $\mathbb{C}^{a-1} \otimes \mathbb{C}^\frac{a}{2} \otimes \mathbb{C}^\frac{a}{2}$, which will have border greater than $\frac{a^2}{8}$. Conclude by applying Exercise 2.2.3.1. ⊠

If $T(A^*) T(\alpha)^{-1}$ is diagonalizable, then $R(T) = b$, and if $T(A^*) T(\alpha)^{-1} \subset G(b, \text{End}(B))$ is in the Zariski closure of the set of $E \in G(b, \text{End}(B))$ that are diagonalizable, then $\overline{R}(T) = b$.

A classical result in linear algebra says a subspace $U \subset \text{End}(B)$ is simultaneously diagonalizable if and only if $U$ is abelian and every $x \in U$ (or equivalently for each $x_j$ in a basis of $U$), $x$ is diagonalizable, answering the question for rank.
Additional equations for $\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$

***This paragraph to be updated*** If $b \leq 4$, then the Zariski closure of the diagonalizable spaces in $G(b, \text{End}(B))$ equals the abelian spaces $[IM05]$, so we have a complete answer. For $b \geq 7$ it just gives a component of the variety of abelian spaces in $G(b, \text{End}(B))$, and for $b = 6, 7$ it is not known whether or not there are additional components. Proposition 2.7.1.7 below gives a sufficient condition.

For $x \in \text{End}(B)$, define the centralizer of $x$, denoted $C(x)$, by

$$C(x) := \{y \in \text{End}(B) \mid [y, x] = 0\}.$$  

**Exercise 2.7.1.5:** Show that $\dim C(x) \geq b$. ⊙

**Exercise 2.7.1.6:** Show that $\dim C(x) = b$ if and only if the minimal polynomial of $x$ equals the characteristic polynomial.

We will say $x \in \text{End}(B)$ is regular if $\dim C(x) = b$. We say $x$ is regular semi-simple if $x$ is diagonalizable with distinct eigenvalues. Note that $x$ is regular semi-simple if and only if $C(x) \subset \text{End}(B)$ is diagonalizable.

**Proposition 2.7.1.7.** Let $U \subset \text{End}(B)$ be an abelian subspace of dimension $b$, and such that there exists $x \in U$ that is regular. Then $U$ lies in the Zariski closure of the diagonalizable $b$-planes in $G(b, \text{End}(B))$.

**Proof.** Since the Zariski closure of the regular semi-simple elements is all of $\text{End}(B)$, for any $x \in \text{End}(B)$, there exists a curve $x_t$ of regular semi-simple elements with $\lim_{t \to 0} x_t = x$. Consider the induced curve in the Grassmannian $C(x_t) \subset G(b, \text{End}(B))$. Then $C_0 := \lim_{t \to 0} C(x_t)$ exists and is contained in $C(x) \subset \text{End}(B)$ and since $U$ is abelian, we also have $U \subset C(x)$. But if $x$ is regular, then $\dim C_0 = \dim(U) = b$, so $\lim_{t \to 0} C(x_t)$, $C_0$ and $U$ must all be equal and thus $U$ is a limit of diagonalizable subspaces. □

**Problem 2.7.1.8.** Determine a criterion for $U \in G(b, \text{End}(B))$ to be in the closure of the diagonalizable $b$-planes, when $U$ does not contain a regular element.

For a $1_A$-generic tensor $T \in A \otimes B \otimes C$, define $T$ to be $2_A$-generic if there exist $\alpha \in A^*$ such that $T(\alpha) : C^* \to B$ is of maximal rank and $\alpha' \in A^*$ such that $T(\alpha')T(\alpha)^{-1} : B \to B$ is regular semi-simple. Unfortunately $M_{(h)}$ is not $2_A$-generic. The equations coming from Koszul flattenings, and even more so Griesser’s equations, are less robust for tensors that fail to be $2_A$-generic. This partially explains why it satisfies some of the Koszul flattening equations and Griesser’s equations below. Thus an important problem is to identify modules of equations for $\sigma_r$ that are robust for non-$2$-generic tensors.
2.7.2. Griesser’s equations. The following theorem dates back to 1985. It posits potential equations for \( \sigma_r \) in the range \( b < r \leq 2b - 1 \).

**Theorem 2.7.2.1.** [Gri86] Let \( b = c \). Given a 1\(_A\)-generic tensor \( T \in A \otimes B \otimes C \) with \( R(T) \leq r \), let \( \alpha \in A^* \) be such that \( T(\alpha) \) is invertible. For \( \alpha' \in A^* \), let \( X(\alpha') = T(\alpha')T(\alpha_0)^{-1} \in \text{End}(B) \). Fix \( \alpha_1 \in A^* \). Consider the space of endomorphisms \( U := \{ [X(\alpha_1), X(\alpha')] : B \to B \mid \alpha' \in A^* \} \subseteq \mathfrak{sl}(B) \). Then there exists \( E \in G(2b - r, B) \) such that \( \dim(U.E) \leq r - b \).

**Remark 2.7.2.2.** As stated, Theorem 2.7.2.1 potentially could provide equations for \( \sigma_r \) up to \( r = 2b - 1 \), but it was not determined in [Gri86] whether or not these equations are trivial beyond \( r = \frac{3}{2}b \).

**Remark 2.7.2.3.** Because of the way the equations are described, they are difficult to write down.

**Remark 2.7.2.4.** Compared with the minors of \( T_A^{\wedge p} \), here one is just examining the first block column of the matrix appearing in the expression \( Q\overline{Q} \), but one is apparently extracting more refined information from it.

**Proof.** For the moment assume \( R(T) = r \) and \( T = \sum_{j=1}^r a_j \otimes b_j \otimes c_j \). Let \( \hat{B} = \mathbb{C}^r \) equipped with basis \( e_1, \ldots, e_r \). Define \( \pi : \hat{B} \to B \) by \( \pi(e_j) = b_j \). Let \( i : B \to \hat{B} \) be such that \( \pi \circ i = 1_{\hat{B}} \). Choose \( B' \subset \hat{B} \) of dimension \( r - b \) such that \( \hat{B} = i(B) \oplus B' \), and denote the inclusion and projection respectively \( i' : B' \to B \) and \( \pi' : \hat{B} \to B' \). Pictorially:

\[
\begin{array}{c}
\hat{B} \\
\begin{array}{c}
i \nearrow \\
\searrow \\
\pi' \searrow \\
\swarrow \end{array}
\end{array}
\begin{array}{c}
B \\
B' \end{array}
\]

Let \( \alpha_0, \alpha_1, \ldots, \alpha_{a-1} \) be a basis of \( A^* \). Let \( \hat{T} = \sum_{j=1}^r a_j \otimes e_j \otimes e_j^* \in A \otimes \hat{B} \otimes \hat{B}^* \) and let \( \hat{X}_j := \hat{T}(\alpha_j)\hat{T}(\alpha_0)^{-1} \). Now in \( \text{End}(\hat{B}) \) all the commutators \([\hat{X}_1, \hat{X}_j] \) are zero because \( R(\hat{T}) = r \). We have, for all \( 2 \leq s \leq a - 1 \), \([\hat{X}_1, \hat{X}_s] = 0 \) implies

\[
0 = \pi[\hat{X}_1, \hat{X}_s]i
= [X_1, X_s] + (\pi\hat{X}_1 i')(\pi'\hat{X}_s i') - (\pi'\hat{X}_s i') (\pi\hat{X}_1 i')
\]

(2.7.1)

Now take \( E \subseteq \ker \pi'\hat{X}_1 i \subset B \) of dimension \( 2b - r \). Then for all \( s \), \([X_1, X_s] \cdot E \subset \text{Image } \pi\hat{X}_1 i' \), which has dimension at most \( r - b \) because \( \pi\hat{X}_1 i' : B' \to B \) and \( \dim B' = r - b \). The general case follows by taking limits. \( \square \)

**Proof of Theorem 2.5.3.1.** Here there is just one commutator \([X_1, X_2] \) and its rank is at most the sum of the ranks of the other two terms in (2.7.1). But each of the other two terms is a composition of linear maps
including \( i' \) which can have rank at most \( r - b \), so their sum can have rank at most \( 2(r - b) \).

\[ \square \]

**Proposition 2.7.2.5.** Let \( \dim A = a, \dim B = \dim C = b \). Then Griesser’s equations for \( \hat{\sigma}_r \) have the following properties:

1. They are trivial for \( r = 2b - 1 \) and all \( a \).
2. They are trivial for \( r = 2b - 2, \) and \( a \leq b + 2 \), in particular \( a = b = 4 \).
3. Setting \( b = n^2 \), matrix multiplication \( M(n) \) fails to satisfy the equations for \( r \leq \frac{3}{2}n^2 - 1 \) when \( n \) is even and \( r \leq \frac{3}{2}n^2 + \frac{9}{2} - 2 \) when \( n \) is odd, and satisfies the equations for all larger \( r \).

I do not know whether or not the equations are trivial for \( r = 2b - 2, \) \( a = b \) and \( b > 4 \). If they are nontrivial for \( r = 2b - 2 \) and even \( b \), they would give equations beyond the maximal minors of \( T_A^{\wedge p} \).

**Proof.** Assuming \( T \) is sufficiently generic, we may choose \( X_1 \) to be diagonal with distinct entries on the diagonal, and this is a generic choice of \( X_1 \). Let \( \mathfrak{s}\mathfrak{l}(B)_R \) denote the matrices with zero on the diagonal (the span of the root spaces). Then

\[
\text{ad}(X_1) : \mathfrak{s}\mathfrak{l}(B)_R \to \mathfrak{s}\mathfrak{l}(B)_R,
Y \mapsto [X,Y],
\]

is a linear isomorphism, and \( \text{ad}(X_1) \) kills the diagonal matrices. Write \( U_s = [X_1,X_s], 2 \leq s \leq a-1 \), so the \( U_s \) will be matrices with zero on the diagonal, and by picking \( T \) generically we can have any such matrices, and this is the most general choice of \( T \) possible, so if the equations vanish for a generic choice of \( U_s \subset \mathfrak{s}\mathfrak{l}(B)_R \), they vanish identically.

Proof of (1): In the case \( r = 2b - 1 \), so \( r - b = b - 1 \) and \( a \leq b + 1 \) the equations are trivial as we only have \( a - 2 \leq b - 1 \) linear maps. When \( a \geq b + 2 \) a naïve dimension count makes it possible for the equations to be non-trivial, the equations are that there exists \( v \in B \) such that \( \dim \langle U_2v,\ldots,U_{a-1}v \rangle \leq b - 1 \). Taking \( v = (1,0,\ldots,0)^T \) (the superscript \( T \) denotes transpose), the \( U_jv \) will be contained in the hyperplane of vectors with their first entry zero. Since we only made genericity assumptions, we conclude.

Proof of (2): In the case \( r = 2b - 2 \), the equations will be nontrivial if and only if there exist \( U_2,\ldots,U_{a-1} \in \mathfrak{s}\mathfrak{l}(B)_R \) such that for all linearly independent \( v, w \)

\[
\dim \langle U_2v,\ldots,U_{a-1}v,U_2w,\ldots,U_{a-1}w \rangle \geq b - 1.
\]
The map $U_2^{-1}U_3$ will in general have $b$ linearly independent eigenvectors. Take $v, w$ to be two such, then $\langle U_2v \rangle = \langle U_3v \rangle$ and $\langle U_2w \rangle = \langle U_3w \rangle$, and the span will have dimension at most $2(a - 2) - 2 \leq b - 2$, and we conclude.

Proof of (3): Consider matrix multiplication $M_{(n)} \in \mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2} \otimes \mathbb{C}^{n^2} = A \otimes B \otimes C$. Recall from Exercise 2.2.4.3 that with a judicious ordering of bases, $M_{(n)}(A^*)$ is block diagonal

\[
(2.7.2) \quad \begin{pmatrix} x & \cdots & x \end{pmatrix}
\]

where $x = (x_j^i)$ is $n \times n$. In particular, the image is closed under brackets. Choose $X_0$ so it is the identity. It is not possible to have $X_1$ diagonal with distinct entries on the diagonal, the most generic choice for $X_1$ is to be block diagonal with each block having the same $n$ distinct entries. For a subspace $E$ of dimension $2b - r = dn + e$ (recall $b = n^2$) with $0 \leq e \leq n - 1$, the image of a generic choice of $[X_1, X_2], \ldots, [X_1, X_{n^2-1}]$ applied to $E$ is of dimension at least $(d+1)n$ if $e \geq 2$, at least $(d+1)n - 1$ if $e = 1$ and $dn$ if $e = 0$, and equality will hold if we choose $E$ to be, e.g., the span of the first $2b - r$ basis vectors of $B$. (This is because the $[X_1, X_s]$ will span the entries of type (2.7.2) with zeros on the diagonal.) If $n$ is even, taking $2b - r = \frac{n^2}{2} + 1$, so $r = \frac{3n^2}{2} - 1$, the image occupies a space of dimension $\frac{n^2}{2} + n - 1 > \frac{n^2}{2} - 1 = r - b$. If one takes $2b - r = \frac{n^2}{2}$, so $r = \frac{3n^2}{2}$, the image occupies a space of dimension $\frac{n^2}{2} = r - b$, showing Griesser’s equations cannot do better for $n$ even. If $n$ is odd, taking $2b - r = \frac{n^2}{2} - \frac{n}{2} + 2$, so $r = \frac{3n^2}{2} + \frac{n}{2} - 2$, the image will have dimension $\frac{n^2}{2} + \frac{n}{2} > r - b = \frac{n^2}{2} + \frac{n}{2} - 1$, and taking $2b - r = \frac{n^2}{2} - \frac{n}{2} + 1$ the image can have dimension $\frac{n^2}{2} - \frac{n}{2} + (n - 1) = r - b$, so the equations vanish for this and all larger $r$. Thus Griesser’s equations for $n$ odd give Lickteig’s bound $\text{R}(M_{(n)}) \geq \frac{3n^2}{2} + \frac{n}{2} - 1$. \hfill \qed

2.7.3. Friedland’s degree 16 equations. Secant varieties of Segre varieties appear in the study of algebraic statistical models corresponding to bifurcating phylogenetic trees, see [AR08], in particular $\sigma_4(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3))$ plays a central role. In 2008, equations for this variety were just beyond the state of the art, so E. Allmann, who resides in Alaska, offered to hand-catch, smoke and send an Alaskan salmon to anyone who could find the generators of the ideal of this variety. In [LM08] the problem was reduced to finding generators of the ideal for $\sigma_4(\text{Seg}(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3))$. The first major breakthrough to this conjecture was by S. Friedland [Fri13]. The breakthrough had two essential steps: finding new equations and proving the known equations plus the new ones were sufficient to cut out the variety set-theoretically.
I explain the new equations in this section. For a proof of the second step, see [Fri13] or [Lan12, §7.7.3].

First I describe a genericity condition for tensors $T \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4$ satisfying Strassen’s equations that assures $R(T) \leq 4$. I then describe additional equations that distinguish which tensors among those that fail to satisfy the genericity conditions have border rank at most four.

**Proposition 2.7.3.1.** Let $a = b$, let $T \in A \otimes B \otimes C$ and assume $T^A_1$ has a nontrivial kernel. Let $\psi_{AB} \in \ker(T^A_1 : A \otimes B^* \to \Lambda^2 A \otimes C)$.

If $\psi_{AB} \in \ker(T^A_1)_{AB}$ is of maximal rank, then $T$ is equivalent to a tensor in $S^2 A \otimes C$.

If furthermore $c = a$ and there also exists $\psi_{AC} \in \ker(T^A_1_{AC})$ of maximal rank, then $T$ is equivalent to a tensor in $S^3 A$.

**Proof.** Consider $(Id_A \otimes \psi_{AB} \otimes Id_C)(T) \in A \otimes A \otimes C$. The condition $\psi_{AB} \in \ker(T^A_1)$ implies $(Id_A \otimes \psi_{AB} \otimes Id_C)(T) \in S^2 A \otimes C$, and since $\psi_{AB}$ is injective, it is $GL(A) \times GL(B) \times GL(C)$-equivalent to $T$. The second assertion is similar. \hfill \Box

**Proposition 2.7.3.2.** If $T \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4$ satisfies Strassen’s equations for border rank 4, and $T^A_1_{AB}$ contains an element of maximal rank, then $R(T) \leq 4$.

**Exercise 2.7.3.3:** Show that $\sigma_4(Seg(v_2(\mathbb{P}^2) \times \mathbb{P}^3)) = \mathbb{P}(S^2 \mathbb{C}^3 \otimes \mathbb{C}^4)$.

**Proof.** Under the hypotheses, $T$ is equivalent to an element of $S^2 \mathbb{C}^3 \otimes \mathbb{C}^4$, so we conclude by Exercise 2.7.3.3. \hfill \Box

Consider the tensor

$T_F = (a_1 \otimes b_1 + a_2 \otimes b_2) \otimes c_1 + (a_1 \otimes b_1 + a_2 \otimes b_3) \otimes c_2 + (a_1 \otimes b_1 + a_3 \otimes b_2) \otimes c_3 + (a_1 \otimes b_1 + a_3 \otimes b_3) \otimes c_4$.

**Exercise 2.7.3.4:** Show $T_F$ satisfies Strassen’s degree nine equations and $\ker((T_F)_{A})^\perp$ does not contain an element of maximal rank.

We will prove that $R(T_F) \geq 5$ by finding additional equations for $\sigma_4(Seg(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3))$ that it does not satisfy. How to find such equations?

Consider a general rank four element of $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4$; $T = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3 + (a_1 + a_2 + a_3) \otimes (b_1 + b_2 + b_3) \otimes c_4$. (This is indeed general, see [Lan12].)

**Exercise 2.7.3.5:** Show that $\ker(T^A_{AB})$ is spanned by $\psi_{AB} := a_1 \otimes \beta^1 + a_2 \otimes \beta^2 + a_3 \otimes \beta^3$, where $\beta^j$ is the dual basis to $b_j$ and similarly $\ker(T^A_{AB})$ is spanned by $\psi_{BA} := b_1 \otimes \alpha^1 + b_2 \otimes \alpha^2 + b_3 \otimes \alpha^3$. 
Note that for any choices of basis of the kernel, \( \psi_{AB} \psi_{BA} = \lambda \text{Id}_A \) and \( \psi_{BA} \psi_{AB} = \mu \text{Id}_B \) for some scalars \( \lambda, \mu \). To obtain equations from this observation, recall that \( A \otimes A^* = \mathfrak{sl}(A) \oplus \mathbb{C}\{\text{Id}_A\} \). The closed property to extract from this information is:

**Proposition 2.7.3.6.**[Fri13] If \( T \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^4 \), \( T_{AB}^{\wedge 1}, T_{BA}^{\wedge 1} \) have one dimensional kernels with bases \( \psi_{AB}, \psi_{BA} \), and \( T \in \hat{\sigma}_4 \), then

\[
(2.7.3) \quad \text{proj}_{\mathfrak{sl}(A)}(\psi_{AB} \psi_{BA}) = 0, \quad \text{proj}_{\mathfrak{sl}(B)}(\psi_{BA} \psi_{AB}) = 0.
\]

**Exercise 2.7.3.7:** Show that \( R(T_F) \geq 5 \).

These are equations of degree 16 for \( \sigma_4(\text{Seg}(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3)) \). I do not know what they are as a module.

**Remark 2.7.3.8.** These equations are now obsolete, in the sense that lower degree equations found in [LM04] via representation theory were shown to be sufficient to cut out \( \sigma_4(\text{Seg}(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3)) \) in [BO11, FG10]. I include them because they illustrate a technique that can be generalized. A similar technique to Proposition 2.7.3.1 was used in [Qi] for tensors in \( A_1 \otimes \cdots \otimes A_n \).

2.7.4. **Summary.** Our first equations for \( \sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \) came from the flattenings, e.g. \( T_A : A^* \to B \otimes C \). To go beyond that, Strassen examined the nature of the image. Assuming \( b = c \) (or restricting to a subspace), if \( T \) is \( 1_A \)-generic, we can use a general \( \alpha \in A \) to identify \( B \otimes C \simeq \text{End}(B) \) and then measure the failure of \( T(A^*) \) to be abelian. That is, we exploited the structure of the image of the linear map to obtain further equations. If these next equations are satisfied, i.e., if it is abelian, a second genericity condition (that \( T(A^*) \) contains a regular element) assures us \( R(T) = b \). If \( T(A^*) \) fails to be abelian but there are no commutators of full rank, in some special situations, one can prove, again under a further genericity condition on \( \ker T_A^{\wedge 1} \) (e.g. Proposition 2.7.3.2), that the border rank equals the value predicted by Strassen’s equations. To obtain further equations when this genericity condition fails, one exploits the structure of \( \ker T_A^{\wedge 1} \) as in Proposition 2.7.3.6.

We have the additional equations obtained from minors of \( T_A^{\wedge p} \), and matrix multiplication satisfies some of them. An important problem is to determine if these equations are enough to conclude an upper bound on \( M_{(n)} \). The discussion above suggests we should study the nature of the kernel (or image) of \( T_A^{\wedge p} \):

**Problem 2.7.4.1.** Find equations for tensors where \( T_A^{\wedge p} : \Lambda^p A \otimes B^* \to \Lambda^{p+1} A \otimes C \) is not of full rank. In particular find equations for the case \( T_{C^9}^{\wedge 4} : \Lambda^4 C^9 \otimes C^9^* \to \Lambda^5 C^9 \otimes C^9 \).
Assume $b = c$ and $a = 2p + 1$. Say $\text{rank}(T_A^{\wedge p})$ is “small” (so that there are many equations satisfied). Are there conditions on $\ker(T_A^{\wedge p})$ that assure $R(T) = \frac{\text{rank}(T_A^{\wedge p})}{\binom{2p}{p}}$?
This chapter discusses progress towards the astounding conjecture that asymptotically, the complexity of multiplying two $n \times n$ matrices is nearly the same as the complexity of adding them. I cover the main advances in upper bounds for the exponent of matrix multiplication, emphasizing a geometric perspective.

3.1. Overview

The exponent $\omega$ of matrix multiplication is naturally defined in terms of tensor rank:

$$\omega := \inf \{ \tau \in \mathbb{R} \mid R(M(n)) = O(n^\tau) \}.$$ 

See [BCS97, §15.1] why tensor rank yields the same exponent as other complexity measures.

The above-mentioned conjecture is that $\omega = 2$. I do not know of any general tools for studying tensor rank. However, we have seen tools for studying border rank. Fortunately, Bini et. al. showed that one may also define the exponent in terms of border rank, namely (see Proposition 3.2.1.10)

$$\omega = \inf \{ \tau \in \mathbb{R} \mid R(M(n)) = O(n^\tau) \},$$
which implies:

\[ \omega \leq \frac{\log R(M(n))}{\log n}. \]

A systematic path to proving the astonishing conjecture would be to find a function \( r(n) = O(n^2) \) and set-theoretic equations for \( \sigma_r(n)(Seg(\mathbb{P}^{m-1} \times \mathbb{P}^{m-1} \times \mathbb{P}^{m-1})) \), and show that \([M(n)]\) is in their zero set. As discussed in §2.7.1, an important simplification is that it would be sufficient to find set-theoretic equations for the open subset of 1-generic elements of \( \sigma_r(n)(Seg(\mathbb{P}^{m-1} \times \mathbb{P}^{m-1} \times \mathbb{P}^{m-1})) \).

At this writing, I know of no evidence that \( R(M(n)) > 2n^2 - 1 \) for any \( n \).

While we have equations for \( r(n) \leq 2n^2 - 4 \) these equations are not satisfied by matrix multiplication. This is why Problem 2.7.4.1 is important for the study of upper bounds.

To disprove the conjecture, it would suffice to find a function \( r(n) \) growing faster than \( Cn^2 \) for any constant \( C \), and one equation (for each \( n \)) \( P_n \) in the ideal of \( \sigma_r(n)(Seg(\mathbb{P}^{m-1} \times \mathbb{P}^{m-1} \times \mathbb{P}^{m-1})) \) such that \( P_n(M(n)) \neq 0 \). All proposed equations for \( \sigma_r(Seg(\mathbb{P}^{m-1} \times \mathbb{P}^{m-1} \times \mathbb{P}^{m-1})) \) when \( 2r \geq m \) that I am aware of have matrix multiplication in their zero set.

Fortunately, there are results that enable one to prove upper bounds on \( \omega \) while avoiding border rank beyond the range we understand. The first is easy and part of Proposition 3.2.1.10, namely we can also consider rectangular matrix multiplication:

\[ \omega \leq \frac{3\log R(M(m,n,l))}{\log (mnl)}, \text{ i.e., } (lmn)^{\frac{3}{2}} \leq R(M(m,n,l)). \]

The second result, due to Schönhage (Theorem 3.3.3.1) and described in §3.3, is more involved: it says it is sufficient to prove upper bounds on sums of disjoint matrix multiplications:

\[ \sum_{i=1}^{s} (m_i,n_i,l_i)^{\frac{3}{2}} \leq R(\bigoplus_{i=1}^{s} M(m_i,n_i,l_i)). \]

Schönhage also proves that the assertion is non-vacuous, that is there exist disjoint tensors where the border rank of their sum is less than the sum of their border ranks, see §3.3.2. Sometimes the border rank of the sum is so low, it becomes in the range where we may determine its border rank. (However researchers proving lower bounds by this method have usually done so simply by writing out an explicit limit.)

As mentioned in §3.3 the inequalities regarding \( \omega \) above are strict, e.g., there does not exist \( n \) with \( R(M(n)) \) equal to \( n^\omega \). (This does not rule out
3.2. The upper bounds of Bini, Capovani, Lotti, and Romani

3.2.1. Rank, border rank, and the exponent of matrix multiplication.

Proposition 3.2.1.1. [Bin80] For all $n$, $\omega \leq R(M_{(n)})$, i.e., $\omega \leq \frac{\log R(M_{(n)})}{\log(n)}$.

Proof. By Exercise 2.2.3.1, $R(M_{(n)}) \leq r$ implies $R(M_{(n^k)}) \leq r^k$. Since $R(M_{(v)})$ is a non-decreasing function of $v$ and $n^{k-1} \leq v < n^k$ for some $k$, we conclude that the $\tau$ in the definition of the exponent satisfies $\tau \leq \frac{\log r}{\log n}$. \qed

Proposition 3.2.1.2. For all $l,m,n$, $(lmn)^{\frac{2}{3}} \leq R(M_{(m,n,l)})$, i.e., $\omega \leq \frac{3\log R(M_{(m,n,l)})}{\log(mnl)}$.

Exercise 3.2.1.3: Prove Proposition 3.2.1.2. ⊙

To show that $\omega$ may also be defined in terms of border rank, introduce a sequence of ranks that interpolate between rank and border rank.

$R(M_{(n)})$ equal to $2n^\omega$ for all $n$.) Thus one cannot exactly determine $\omega$ by the above methods. One way to extend the above methods is to find sequences of sums $\otimes_{i=1}^{\rho} M_{l_i(N),m_i(N),n_i(N)}$ with the border rank of the sums giving upper bounds on $\omega$ that converge to the actual value. This is one aspect of Strassen’s “laser method” described in §3.4. A second new ingredient of his method is that instead of dealing with the sum of a collection of disjoint rectangular matrix multiplications, one looks for a larger tensor $T \in A \otimes B \otimes C$, that has special structure rendering it easy to study, that can be degenerated into a collection of disjoint matrix multiplications. More precisely, to obtain a sequence as in the previous paragraph, one degenerates the tensor powers $T^\otimes N$. Strassen’s degeneration is in the sense of the $GL(A) \times GL(B) \times GL(C)$-orbit closure of $T^\otimes N$.

After Strassen, all other subsequent upper bounds on $\omega$ use what I will call combinatorial restrictions of $T^\otimes N$, where entries of a coordinate presentation of $T^\otimes N$ are simply set equal to zero. The choice of entries to zero out is subtle. I describe these developments in §3.6.
We say \( R_h(T) \leq r \) and \( T \in \hat{\sigma}_r^{0,h} \) if there exists an expression

\[
T = \lim_{\epsilon \to 0} \frac{1}{\epsilon^h} (a_1(\epsilon) \otimes b_1(\epsilon) \otimes c_1(\epsilon) + \cdots + a_r(\epsilon) \otimes b_r(\epsilon) \otimes c_r(\epsilon))
\]

where \( a_j(\epsilon), b_j(\epsilon), c_j(\epsilon) \) are analytic functions of \( \epsilon \).

**Proposition 3.2.1.4.** \( R(T) \leq r \) if and only if there exists an \( h \) such that \( R_h(T) \leq r \).

To prove Proposition 3.2.1.4, we need the following lemma:

**Lemma 3.2.1.5.** Let \( Z \subset \mathbb{P}V \) be an irreducible variety and let \( Z^0 \subset Z \) be a Zariski open subset. Let \( p \in Z \setminus Z^0 \). Then there exists an analytic curve \( C(t) \) such that \( C(t) \in Z^0 \) for all \( t \neq 0 \) and \( \lim_{t \to 0} C(t) = p \).

For the proof of the lemma, we will need the following basic fact: every irreducible algebraic curve \( C \subset \mathbb{P}V \) (i.e., every one dimensional algebraic variety) is such that there exists a smooth algebraic curve \( \hat{C} \) and a surjective algebraic map \( \pi : \hat{C} \to C \) that is one-to-one over the smooth points of \( C \).

See, e.g., [Sha94, §II.5], Thms. 3 and 6 for a proof. The curve \( \hat{C} \) is called the *normalization* of \( C \).

**Proof.** Let \( c \) be the codimension of \( Z \) and take a general \( L \subset \mathbb{P}V \) of dimension \( c + 1 \) that contains \( p \). Then \( L \cap Z \) will be a possibly reducible algebraic curve containing \( p \). Take a component \( C \) of the curve that contains \( p \). If \( p \) is a smooth point of the curve we are done, as we can expand a Taylor series about \( p \). Otherwise take the the normalization \( \pi : \hat{C} \to C \) and a point of \( \pi^{-1}(p) \), expand a Taylor series about that point and compose with \( \pi \) to obtain the desired curve. \( \square \)

**Proof of Proposition 3.2.1.4.** Recall that every curve in \( Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \) is of the form \( [a(t) \otimes b(t) \otimes c(t)] \) for some curves \( a(t) \subset A \) etc., and if the curve is analytic, the functions \( a(t), b(t), c(t) \) can be taken to be analytic as well. Thus every analytic curve in \( \sigma_r^0(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \) may be written as \( [\sum_{j=1}^r a(t) \otimes b(t) \otimes c(t)] \) for some analytic curves \( a_j(t) \subset A \) etc. We conclude that if \( T \in \hat{\sigma}_r \), then \( T \in \hat{\sigma}_r^{0,h} \) for some \( h \). \( \square \)

**Remark 3.2.1.6.** In the matrix multiplication literature, e.g. [BCS97], \( \hat{\sigma}_r \) is often defined to be the the set of \( T \) that are in \( \sigma_r^{0,h} \) for some \( h \). One then must show that this set is algebraically closed.

**Proposition 3.2.1.7.** If \( T \in \hat{\sigma}_r^{0,h} \), then \( R(T) \leq r(\frac{h+2}{2}) < rh^2 \).

**Proof.** Write \( T \) as in (3.2.1). Then \( T \) is the coefficient of the \( \epsilon^h \) term of the expression in parentheses. For each monomial, there is a contribution of \( \left( \frac{h+2}{2} \right) \) terms of degree \( h \) in \( \epsilon \). \( \square \)

**Remark 3.2.1.8.** In fact \( R(T) \leq rh \), see Proposition 3.6.1.6.
Exercise 3.2.1.9: Show that if \( T \in \hat{\sigma}_r^{0,h} \) then \( T \otimes N \in \hat{\sigma}_{r,N}^{0,Nh} \).

Proposition 3.2.1.10. For all \( l, m, n, \omega \leq 3\log R(M(l,m,n)) \frac{\log N}{\log(k)} \).

Proof. Write \( r = R(M(l,m,n)) \). As above, set \( N = mnl \). We have \( M(l,m,n) \in \hat{\sigma}_{r,l}^{0,h} \) for some \( h \) and thus \( R(M(l,m,n)) \leq r^{3k}(hk)^2 \), which implies

\[
\omega \leq \frac{3\log r \log N}{\log k} + \frac{2\log(hk) \log N}{k\log N}
\]

and the second term goes to zero as \( k \to \infty \). \( \square \)

3.2.2. Bini et. al’s algorithm. Bini, Capovani, Lotti, and Romani had been looking for a rank five expression for the tensor corresponding to \( M_{(2)} \) when the first matrix has an entry equal to zero. Say the entry is \( x_{22} = 0 \) and denote the tensor by

\[
T_{BCLR} := x_1^1 \otimes (y_1^1 \otimes z_1^1 + y_2^1 \otimes z_2^1) + x_2^1 \otimes (y_1^2 \otimes z_1^1 + y_2^2 \otimes z_2^1) + x_1^1 \otimes (y_1^1 \otimes z_2^1 + y_1^2 \otimes z_2^1) + x_2^1 \otimes (y_2^1 \otimes z_2^1 + y_2^2 \otimes z_2^1)
\]

Their method was to minimize the norm of \( T_{BCLR} \) minus a rank five tensor that varied, and their computer kept on producing rank five tensors with larger and larger coefficients. At first they thought there was a problem in their computer code, and only later realized it was a manifestation of the phenomenon Bini named border rank \([Bin80]\). (This information from Bini, personal communication.)

The expression for the tensor \( T_{BCLR} \) that their computer search found is

\[
T_{BCLR} = \lim_{t \to 0} \frac{1}{t} [((x_2^2 + tx_2^1) \otimes (y_2^1 + ty_2^2) \otimes z_2^1 - x_1^2 \otimes (z_1^1 + z_2^2) - x_2^1 \otimes (y_2^1 + y_2^2) \otimes z_1^1 + (x_2^1 + x_2^2) \otimes (y_2^2 + ty_1^2) \otimes (z_1^1 + z_2^1))]
\]

and thus \( R(T_{BCLR}) \leq 5 \). I discuss the geometry of this algorithm in §4.3.1.

Exercise 3.2.2.1: Use (3.2.2) to show \( R(M_{(2,2,3)}) \leq 10 \). ⊗

Using Proposition 3.2.1.10 we conclude:

Theorem 3.2.2.2. \([BCRL79]\) (1979) \( \omega < 2.78 \).
3.3. Schöhage’s upper bounds

The next contribution to upper bounds for the exponent of matrix multiplication was Schöhage’s realization that a border rank version of Strassen’s additivity conjecture described below is false, and that this failure could be exploited to prove further upper bounds on the exponent. From a geometric perspective, this result is useful because it allows one to prove upper bounds with tensors that are easier to analyze because of their low border rank.

3.3.1. Strassen’s additivity conjecture. Given \( T_1 \in A_1 \otimes B_1 \otimes C_1 \) and \( T_2 \in A_2 \otimes B_2 \otimes C_2 \), if one considers \( T_1 + T_2 \in (A_1 \oplus A_2) \otimes (B_1 \oplus B_2) \otimes (C_1 \oplus C_2) \), we saw that \( R(T_1 + T_2) \leq R(T_1) + R(T_2) \).

Conjecture 3.3.1.1. [Str73] With the above notation, \( R(T_1 + T_2) = R(T_1) + R(T_2) \).

Let’s consider a potential border rank version of Conjecture 3.3.1.1. Say \( T_1 + T_2 \) is as above and consider \((T_1 + T_2)(C_1^* \oplus C_2^*)\). The image looks like

\[
\begin{pmatrix}
T(C_1^*) & 0 \\
0 & T(C_2^*)
\end{pmatrix}.
\]

One might attempt to construct a curve leading to a counter-example via geometry as follows: say that \( \dim C_2 = 1 \), e.g., that \( T_2 = M_{(1,1,N)} \). If one takes a curve with zero-th order terms in \( A_1 \otimes B_1 \otimes C_2 \), and takes one derivative, one can have terms in \( A_1 \otimes B_1 \otimes C_1 \) and after two derivatives, one can have terms in both \( A_1 \otimes B_1 \otimes C_1 \) and \( A_2 \otimes B_2 \otimes C_2 \). The zero-th order terms must be arranged to all cancel. Schöhage accomplishes this in the simplest possible way: he takes dimensions sufficiently unbalanced that there are more terms than the dimension of \( A_1 \otimes B_1 \otimes C_2 \), so they are linearly dependent and easily arranged to cancel. What is more subtle is the cancellation of the first order terms, whose geometry I leave to the reader to explore.

3.3.2. Schöhage’s example. Recall from Exercise 2.2.4.5 that \( R(M_{(1,m,n)}) = mn \) and \( R(M_{(1,1,N)}) = N \).

Theorem 3.3.2.1 (Schöhage [Sch81]). Set \( N = (n-1)(m-1) \). Then

\[
R(M_{(1,m,n)} \oplus M_{(1,1,N)}) = mn + 1.
\]
3.3. Schönhage’s upper bounds

Proof.

\[
T = \lim_{t \to 0} \frac{1}{t^2} \left( \sum_{u=1}^{m-1} \sum_{v=1}^{n-1} \left( x_u + tx_{uv} \right) \otimes (b_v + ty_{uv}) \otimes (z + t^2 z_{uv}) 
+ \sum_{u=1}^{m-1} x_u \otimes (y_n + t(-\sum_{v} y_{uv})) \otimes (z + t^2 z_{un}) 
+ \sum_{v=1}^{n-1} (x_m + t(-\sum_{u} x_{uv})) \otimes b_v \otimes (z + t^2 z_{mv}) 
+ x_m \otimes y_n \otimes (z + t^2 z_{mn}) - (\sum_{i} x_i) \otimes (\sum_{s} y_s) \otimes z \right)
\]

\[\square\]

3.3.3. Schönhage’s asymptotic sum inequality. To develop intuition how an upper bound on a sum of matrix multiplications could give an upper bound on a single matrix multiplication, say we knew \( R(M^+_{(n)}) \leq r \) with \( s \leq n^3 \). Then to compute \( M_{(n^2)} \) we could write \( M_{(n^2)} = M_{(n)} \otimes M_{(n)} \). At worst this is evaluating \( n^3 \) disjoint copies of \( M_{(n)} \). Now group these disjoint copies in groups of \( s \) and apply the bound to obtain a savings.

Here is the precise statement:

**Theorem 3.3.3.1.** [Sch81] [Schönhage’s asymptotic sum inequality] For all \( l_i, m_i, n_i \), with \( 1 \leq i \leq s \):

\[
\sum_{i=1}^{s} (m_i, n_i, l_i) \leq R(\bigoplus_{i=1}^{s} M_{(m_i, n_i, l_i)}).
\]

**Remark 3.3.3.2.** A similar result (also proven in [Sch81]) holds for the border rank of the multiplication of matrices with some entries equal to zero, where the product \( m_i, n_i, l_i \) is replaced by the number of multiplications in the naïve algorithm for the matrices with zeros.

Following notes of M. Bläser [Blä13], we first analyze a special case that isolates the new ingredient:

**Lemma 3.3.3.3.** If \( R(M_{(n)}^\oplus s) \leq r \), then \( \omega \leq \frac{\log(s)}{\log(n)} \). In particular, \( sn^\omega \leq R(M_{(n)}^\oplus s) \).

**Proof.** It is sufficient to show that for all \( N \),

\[
(3.3.1) \quad R(M_{(n^N)}^\oplus s) \leq \left[ \frac{r}{s} \right] N s
\]
because (3.3.1) would imply \( R(M_{\langle nN \rangle}) \leq \lceil \frac{r}{s} \rceil N^s \) and thus one obtains

\[
\omega \leq \lim_{N \to \infty} \frac{\log(\lceil \frac{r}{s} \rceil N^s)}{\log(nN)} = \log \left( \frac{\lceil \frac{r}{s} \rceil}{\log n} \right).
\]

We prove (3.3.1) by induction. The hypothesis is the case \( N = 1 \). Assume (3.3.1) holds up to \( N \) and compute

\[
M_{\langle nN+1 \rangle}^{\oplus s} = M_{\langle n \rangle}^{\oplus s} \otimes M_{\langle n^N \rangle}^{\oplus s}
\]

Now \( R(M_{\langle n \rangle}^{\oplus s}) \leq r \) implies \( M_{\langle n \rangle}^{\oplus s} \in GL_{r \times 3}^{\otimes s} \cdot M_{\langle \langle 1 \rangle \rangle}^{\otimes s} \). Thus \( R(M_{\langle n^N+1 \rangle}^{\oplus s}) \leq R(M_{\langle \langle 1 \rangle \rangle}^{\otimes s} \otimes M_{\langle n^N \rangle}^{\oplus s}) \). Recall that \( M_{\langle \langle 1 \rangle \rangle}^{\otimes s} \otimes M_{\langle n^N \rangle}^{\oplus s} = M_{\langle \langle n^N \rangle \rangle}^{\otimes s} \). Now

\[
R(M_{\langle n^N \rangle}^{\otimes s}) \leq R(M_{\langle n^N \rangle}^{\otimes s})
\]

\[
\leq R(M_{\langle \langle 1 \rangle \rangle}^{\otimes \frac{r}{s}} \otimes M_{\langle n^N \rangle}^{\oplus s})
\]

\[
\leq R(M_{\langle \langle 1 \rangle \rangle}^{\otimes \frac{r}{s}}) R(M_{\langle n^N \rangle}^{\oplus s})
\]

\[
\leq \lceil \frac{r}{s} \rceil \left( \left\lceil \frac{r}{s} \right\rceil N^s \right)
\]

where the last line follows from the induction hypothesis. \( \square \)

The general case essentially follows from the above lemma and repeating arguments used previously: one first takes a high tensor power of the sum, then switches to rank at the price of introducing an \( h \) that washes out in the end. The new tensor is a sum of products of matrix multiplications that one converts to a sum of matrix multiplications. One then takes the worst term in the summation and estimates with respect to it (multiplying by the number of terms in the summation), and applies the lemma to conclude.

**Corollary 3.3.3.4.** [Sch81] (1981) \( \omega < 2.55 \).

**Proof.** Applying Theorem 3.3.3.1 to \( R(M_{\langle 1,m,n \rangle}^{\otimes s} \oplus M_{\langle (m-1)(n-1),1,1 \rangle}^{\langle 1 \rangle}) = mn + 1 \) gives

\[
(mn)^{\frac{r}{s}} + ((m - 1)(n - 1))^{\frac{r}{s}} \leq mn + 1
\]

and taking \( m = n = 4 \) gives the result. \( \square \)

In [CW82] they prove that for any tensor \( T \) that is a direct sum of disjoint matrix multiplications, if \( R(T) \leq r \), then there exists \( N \) and \( k \) such that \( R(T^{\otimes k} \oplus M_{\langle 1,1,N \rangle}) \leq r^k + 1 \). In particular, setting \( N = R(M_{\langle 1,m,n \rangle}) - 1(m + n - 1) \),

\[
R(M_{\langle 1,m,n \rangle} \oplus M_{\langle 1,1,N \rangle}) \leq R(M_{\langle 1,m,n \rangle}) + 1.
\]

This implies
3.4. Strassen’s laser method

3.4.1. Introduction. Recall our situation: we don’t understand rank or even border rank in the range we would need to prove upper bounds on \( \omega \) via \( M(n) \), so we showed upper bounds on \( \omega \) could be proved first with rectangular matrix multiplication, then with sums of disjoint matrix multiplications which had the property that the border rank of the sum was less than the sum of the border ranks, and the border ranks were sufficiently low that they were in the range of \( R \) that we can determine with equations (although Schönhage determined the border rank via an explicit algorithm).

We also saw that to determine the exponent, one would need to deal with sequences of tensors. Strassen’s laser method is based on taking high tensor powers of a fixed tensor, but then degenerating it in a different way for each power. Because the degenerations are different, there is no theoretical obstruction to determining \( \omega \) exactly via Strassen’s method.

Starting with Strassen’s method, all attempts to determine \( \omega \) aim at best for a Pyrrhic victory in the sense that even if \( \omega \) were determined by these methods, they would not give any indication as to what would be optimally fast matrix multiplication for any given size matrix.

Definition 3.4.1.1. Given \( T \in A \otimes B \otimes C \), if \( T' \in GL(A) \times GL(B) \times GL(C) \cdot T \), one says \( T \) degenerates to \( T' \).

Recall from Proposition 2.4.12.1 that if \( T \) degenerates to \( T' \), then \( R(T') \leq R(T) \).

In Schönhage’s work, raising tensors to a power was used to prove results about particular starting tensors. In Strassen’s work one utilizes the tensor powers directly to obtain upper bounds on \( \omega \) that could potentially determine \( \omega \) exactly. Note that if \( T \) degenerates to \( T' \) then \( T^{\otimes N} \) degenerates to \( T'^{\otimes N} \).

What tensors \( T \) are good candidates to have both low border rank and admit a degeneration to a “good” sum of matrix multiplication tensors? To formalize this question, we make a definition:

Definition 3.4.1.2. For \( T \in A \otimes B \otimes C \), \( N \geq 1 \) and \( \rho \in [2, 3] \) define \( V^d_{\rho,N}(T) \) to be the maximum of \( \sum_{i=1}^{s} (l_i m_i n_i)^{\frac{\rho}{N}} \) over all degenerations of \( T^{\otimes N} \) to
The normal form:
\[ T_{i}^{\sigma} \]

is the general tangent line to the Segre, in which case their rank is three, and then their border rank.

\[ \sigma_{3}. \]

4. Tensors of low border rank.

Not much is known about the pos-

s, l, m, n and define the degeneracy value of T to be
\[ V_{\rho}^{d}(T) := \sup N V_{\rho,N}^{d}(T) \frac{1}{N}. \]

\[ \omega \]

The supremum in the definition can be replaced by a limit since the sequence \( V_{\rho,N}^{d}(T) \) is super-additive, so Feketes lemma implies \( \log \frac{1}{N} V_{\rho,N}^{d}(T) \) tends to a limit. See [AFL14] for details.

The utility of value is that an analogue of the asymptotic sum inequality holds:

**Theorem 3.4.1.3.**

\[ \sum_{i=1}^{s} V_{\omega}^{d}(T_{i}) \leq R(\oplus_{i=1}^{s} T_{i}). \]

The proof is similar to the proof of the asymptotic sum inequality. It is clear that \( V_{\omega}^{d}(T_{1} \otimes T_{2}) \geq V_{\omega}^{d}(T_{1}) \otimes V_{\omega}^{d}(T_{2}) \). To show \( V_{\omega}^{d}(T_{1} \oplus T_{2}) \geq V_{\omega}^{d}(T_{1}) + V_{\omega}^{d}(T_{2}) \) one expands out \( V_{\omega,N}^{d}(T_{1} \oplus T_{2}) \), the result is a sum of products with coefficients, but as with the asymptotic sum inequality, one can essentially just look at the largest term, and as N tends to infinity, the coefficient becomes irrelevant after taking N-th roots.

Since degeneracy value is a non-decreasing function of \( \rho \), if we could show \( V_{\rho}^{d}(T) \geq R(T) \) for some \( \rho \), we would obtain the upper bound \( \omega \leq \rho \). Thus tensors of low border rank with high degeneracy value give upper bounds on \( \omega \). The problem is that we have no systematic way of estimating degeneracy value. In subsequent work, researchers restrict to a special type of value that is possible to estimate.

I expect that value is related to tensor rank, so that to obtain upper bounds on \( \omega \), one should look for tensors whose rank is much larger than their border rank.

### 3.4.2. Tensors of low border rank.

Not much is known about the possible ranks of tensors of low border rank. It is classical that elements of \( \sigma_{2}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \setminus \text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \) are either of rank two, or lie on a general tangent line to the Segre, in which case their rank is three, and then there is the normal form: \( T = a_{1} \otimes b_{1} \otimes c_{1} + a_{1} \otimes b_{1} \otimes c_{2} + a_{1} \otimes b_{2} \otimes c_{1} + a_{2} \otimes b_{1} \otimes c_{1} \).

Normal forms for the third secant variety were determined in [BL14]:

**Theorem 3.4.2.1.** [BL14] Normal forms points of \( \sigma_{3}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \setminus \sigma_{2}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \) are (up to permuting the roles of A, B, C in the last case) as follows:

1. \( a_{1} \otimes b_{1} \otimes c_{1} + a_{2} \otimes b_{2} \otimes c_{2} + a_{3} \otimes b_{3} \otimes c_{3} \)
2. \( a_{1} \otimes b_{1} \otimes c_{1} + a_{2} \otimes b_{2} \otimes c_{3} + a_{2} \otimes b_{3} \otimes c_{2} + a_{3} \otimes b_{2} \otimes c_{2} \)
3. \( a_{1} \otimes b_{1} \otimes c_{2} + a_{1} \otimes b_{2} \otimes c_{1} + a_{2} \otimes b_{1} \otimes c_{1} + a_{1} \otimes b_{3} \otimes c_{3} + a_{3} \otimes b_{1} \otimes c_{3} + a_{3} \otimes b_{3} \otimes c_{1} \)
(4) \( a_3 \otimes b_1 \otimes c_1 + a_1 \otimes b_2 \otimes c_2 + a_1 \otimes b_1 \otimes c_2 + a_2 \otimes b_3 \otimes c_1 + a_2 \otimes b_1 \otimes c_3 \).

The points of type (1) form a Zariski open subset, those of type (2) are of the form \( x + y + y' \) where \( x, y \in \hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \) and \( y' \in \hat{T}_y \hat{\text{Seg}}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \) and have rank four, those of type (3) are of the form \( x + x' + x'' \) where \( x(t) \) is a curve in the Segre, and have rank five, and those of type (4) are of the form \( x' + y' \) where \([x], [y]\) are two points lying on a line in the Segre and have rank five.

For higher secant varieties, all these types of points generalize, and many kinds of new phenomena arise as well.

Recall how Schönhage’s example began: one took \( q + 1 \) points on a \( \mathbb{P}^q \subset \text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \). Then any point in the linear span of the tangent spaces is a point on \( \sigma_{q+1}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \).

**Exercise 3.4.2.2:** Show that for a linear space on the Segre, e.g., a \( \mathbb{P}^q \), there are points of \( \hat{\sigma}_{q+2}(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) \) of the form

\[
x_1'' + \cdots + x_q''
\]

where \( x_1(t), \ldots, x_q(t) \) are curves in the Segre with \( x_j(0) \) all lying on at most a \( \mathbb{P}^{q-2} \) and summing to zero, and \( x_j(0)' \) all summing to zero as well. An extreme case of this is when all the \( x_j(0) \) coincide.

It may be useful to consider the corresponding curve in the Grassmanian. For example, consider the extreme case and the curve

\[
[(a_0 \otimes b_0 \otimes c_0) \land (t(a_0 + ta_1) \otimes (b_0 + tb_1) \otimes (c_0 + tc_1)) \land \\
\cdots \land (t(a_0 + ta_q) \otimes (b_0 + tb_q) \otimes (c_0 + tc_q)) \land (a_0 - t^2 \sum_{j=1}^{q} a_j) \otimes (b_0 - t^2 \sum_{j=1}^{q} b_j) \otimes (c_0 - t^2 \sum_{j=1}^{q} c_j)]
\]

The resulting tensor is

\[
(3.4.2) \quad T_{CW} := \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0 \in \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1}
\]
Exercise 3.4.2.3: Show the tensor

\[ \tilde{T}_{CW} := \sum_{j=1}^{q} (a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j + a_j \otimes b_j \otimes c_0) \]

\[ + a_0 \otimes b_0 \otimes c_{q+1} + a_0 \otimes b_{q+1} \otimes c_0 + a_{q+1} \otimes b_0 \otimes c_0 \in \mathbb{C}^{q+2} \otimes \mathbb{C}^{q+2} \otimes \mathbb{C}^{q+2} \]

also has border rank \( q + 2 \) by modifying the curves used to obtain \( T_{CW} \).

Exercise 3.4.2.4: Show \( R(T_{CW}) \leq 2q \).

For future reference, we record our observation:

Proposition 3.4.2.5. \( R(T_{CW}) = R(\tilde{T}_{CW}) = q + 2 \).

**need to add proof of lower bound**

Here the matrices are

\[
T_{CW}(C^*) = \begin{pmatrix}
0 & c_1 & \cdots & c_q \\
c_1 & c_0 & 0 & \cdots \\
c_2 & 0 & c_0 & \cdots \\
\vdots & \vdots & \ddots & \cdots \\
c_q & 0 & \cdots & 0 & c_0
\end{pmatrix}
\]

and

\[
\tilde{T}_{CW}(C^*) = \begin{pmatrix}
c_{q+1} & c_1 & \cdots & c_q & c_0 \\
c_1 & c_0 & 0 & \cdots & 0 \\
c_2 & 0 & c_0 & \cdots & \cdots \\
\vdots & \vdots & \ddots & \cdots & \cdots \\
c_q & 0 & \cdots & 0 & c_0 \\
c_0 & 0 & \cdots & 0 & 0
\end{pmatrix}.
\]

3.4.3. Strassen’s tensor. Strassen uses the following tensor,

\[ T_{STR} = \sum_{j=1}^{q} a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j \in \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1} \otimes \mathbb{C}^{q} \]

which has border rank \( q + 1 \), as these are tangent vectors to \( q \) points of the \( \mathbb{P}^{q-1} = \{ a_0 \otimes b_0 \otimes (c_1, \ldots, c_q) \} \) that lies on the Segre as discussed in §4.3.1. The vector \( a_0 \otimes b_j \otimes c_j + a_j \otimes b_0 \otimes c_j \) is in the tangent space to \( a_0 \otimes b_0 \otimes c_j \), showing the border rank is at most \( q + 1 \), but since the tensor is concise, we obtain equality.
The corresponding linear spaces are:

$$T_{STR}(C^*) = \begin{pmatrix}
0 & c_1 & \cdots & c_q \\
c_1 & 0 & \cdots & 0 \\
c_2 & \vdots & \ddots & \vdots \\
c_q & 0 & \cdots & 0 
\end{pmatrix},$$

and

$$T_{STR}(A^*) = \begin{pmatrix}
a_1 & a_2 & \cdots & a_q \\
a_0 & 0 & \cdots & 0 \\
0 & a_0 & 0 & \vdots \\
0 & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & a_0 
\end{pmatrix}.$$

Consider $\tilde{T} := T_{STR} \otimes \sigma(T_{STR}) \otimes \sigma^2(T_{STR})$ where $\sigma$ is a cyclic permutation of $T_{STR}$. Thus $\tilde{T} \in \mathbb{C}^{q(q+1)^2} \otimes \mathbb{C}^{q(q+1)^2} \otimes \mathbb{C}^{q(q+1)^2}$ and $R(\tilde{T}) \leq (q + 1)^3$.

Write $a_{\alpha\beta\gamma} := a_{\alpha} \otimes a_{\beta} \otimes a_{\gamma}$ and similarly for $b$’s and $c$’s. Then, omitting the $\otimes$’s:

$$\tilde{T} = \sum_{i,j,k=1}^{q} (a_{ij0}b_{0jk}c_{i0k} + a_{ijk}b_{0jk}c_{i00} + a_{ij0}b_{00k}c_{ijk} + a_{ij0}b_{00k}c_{ij0}$$

$$+ a_{0jk}b_{ij0}c_{i0k} + a_{0jk}b_{ij0}c_{i00} + a_{0jk}b_{i0k}c_{ijk} + a_{0jk}b_{i0k}c_{ij0})$$

There are four kinds of $A$-indices, $ij0$, $ijk$, $0j0$ and $0jk$. To emphasize this, and to suggest what will come next, label these with superscripts $11$, $21$, $12$ and $22$. Label each of the $B$ and $C$ indices (which come in four types as well) similarly. We obtain:

$$\tilde{T} = \sum_{i,j,k=1}^{q} (a_{11ij0}b_{01jk}c_{i10k} + a_{21ij0}b_{01jk}c_{i100} + a_{11ij0}b_{120k}c_{i2j0} + a_{21ij0}b_{120k}c_{i2j0}$$

$$+ a_{120j0}b_{21ij}c_{j0} + a_{220j0}b_{21ij}c_{j00} + a_{02j0}b_{21ij}c_{i00} + a_{02j0}b_{21ij}c_{i00} + a_{220j0}b_{22ij}c_{i2j0} + a_{220j0}b_{22ij}c_{i2j0}).$$

This expression has the structure of block $2 \times 2$ matrix multiplication, where inside the blocks are rectangular matrix multiplications. E.g., the first block is $M_{(q^2,q^2)}$, the second is $M_{(q^2,q^2)}$ etc.

Were the matrix multiplications disjoint, we could apply the asymptotic sum inequality. To obtain disjoint matrix multiplications, we could set some coefficients to zero. For example we could set all terms with $a_{121}, b_{121}, c_{121}, a_{212}, b_{212}, c_{212}$ to zero to be left with two independent matrix multiplications. But it makes far more sense to use Theorem 2.4.12.4 to keep three ($= \lfloor \frac{3}{4}(2^2) \rfloor$) independent matrix multiplications. Applying the asymptotic
3. Asymptotic upper bounds for matrix multiplication

Asymptotic upper bounds for matrix multiplication, we obtain $3^q \omega \leq (q + 1)^3$ and taking $q = 7$ gives $\omega < 2.642$, which is not as good as Schönhage’s bound.

We can do better by taking a high tensor power of $\tilde{T}$, as $\tilde{T} \otimes N$ contains $(2^N)^2$ matrix multiplications $M_{(l,m,n)}$, all with $lmn = q^3N$, and again by Theorem 2.4.12.4 we may keep $\frac{3}{4}2^N$ of them. The asymptotic sum inequality applied to the degenerated tensor gives

$$\frac{3}{4}2^N q^N \omega \leq (q + 1)^3 N.$$

Taking $N$-th roots and letting $N$ tend to infinity, the $\frac{3}{4}$ goes away and we obtain

$$2^N q^\omega \leq (q + 1)^3.$$

Finally, the case $q = 5$ implies:

**Theorem 3.4.3.1.** [Str87] (1987) $\omega < 2.48$.

### 3.5. The Coppersmith-Winograd method

Degeneracy value had the defect that we have no way of computing it in general. I now define two weaker notions of value, the second has the advantage that it is computable, but the disadvantage that it depends on a choice of coordinates. (However, Strassen only applied his method to coordinate degenerations as well.)

#### 3.5.1. The value of a tensor

Let $\text{End}(A) \times \text{End}(B) \times \text{End}(C)$ act on $A \otimes B \otimes C$ by the action inherited from the $\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)$ action (not the Lie algebra action). Then for all $X \in \text{End}(A) \times \text{End}(B) \times \text{End}(C)$ and $T \in A \otimes B \otimes C$, we have $R(X \cdot T) \leq R(T)$ and $R(X \cdot T) \leq R(T)$ by Exercise 2.2.3.1.

**Definition 3.5.1.1.** One says $T$ restricts to $T'$ if there exists $X \in \text{End}(A) \times \text{End}(B) \times \text{End}(C)$ such that $T' = X \cdot T$.

**Definition 3.5.1.2.** For $T \in A \otimes B \otimes C$, $N \geq 1$ and $\rho \in [2, 3]$ define $V_{\rho,N}^r(T)$ to be the maximum of $\sum_{i=1}^s (l_i m_i n_i)^{\rho \frac{2}{3}}$ over all restrictions of $T \otimes N$ to $\otimes_{i=1}^s M_{(q_i, m_i, n_i)}$ and define the restriction value of $T$ to be $V_{\rho}^r(T) := \sup_N V_{\rho,N}^r(T)^{\frac{3}{N}}$.

All the results for degeneracy value hold for restriction value, in particular the analog of the asymptotic sum inequality. Coppersmith-Winograd and all subsequent work, use only the following type of restriction:

**Definition 3.5.1.3.** Let $A, B, C$ be given bases $b_A, b_B, b_C$. We say $T \in A \otimes B \otimes C$ combinatorially restricts to $T'$ with respect to $b_A, b_B, b_C$, if $T$ restricts to $T'$ by coordinate projections in these bases.
For $T'$ the matrix multiplication tensor, following [CU03], this may be phrased as follows: write $a_\alpha$, $b_\beta$, $c_\gamma$ for the given bases of $A, B, C$ and 

$$T = \sum_{\alpha=1}^{a} \sum_{\beta=1}^{b} \sum_{\gamma=1}^{c} t^{\alpha,\beta,\gamma} a_\alpha \otimes b_\beta \otimes c_\gamma$$

$T \in \mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^c$ combinatorially restricts to $M_{l|m|n}$ means that there exist functions

$$\alpha : [l] \times [m] \to [a]$$

$$\beta : [m] \times [n] \to [b]$$

$$\gamma : [n] \times [l] \to [c]$$

such that

$$(3.5.1) \quad t^{\alpha(i,u'),\beta(u,s'),\gamma(s,i')} = \begin{cases} 1 & \text{if } i = i', u = u', s = s' \\ 0 & \text{otherwise} \end{cases}.$$

**Definition 3.5.1.4.** For $T \in A \otimes B \otimes C$, $N \geq 1$ and $\rho \in [2, 3]$ define $V_{\rho,N}(T)$ to be the maximum of $\sum_{i=1}^{a} (l|m|n)^{\frac{\rho}{3}}$ over all combinatorial degenerations of $T^{\otimes N}$ to $\oplus_{i=1}^{a} M_{l|m|n}$ and define the **combinatorial value** (or value for short, since it is the value used in the literature) of $T$ to be $V_{\rho}(T) := \lim_{n \to \infty} V_{\rho,N}(T)^\frac{1}{N}$. (The limit is shown to exist in [DS13].)

Note that the values satisfy $V_{\rho}^d \geq V_{\rho}^r \geq V_{\rho}$. As with all the values we have

- $V_{\rho}(T)$ is a non-decreasing function of $\rho$,
- $V_{\omega}(T) \leq R(T)$.

Thus if $V_{\rho}(T) \geq R(T)$, then $\omega \leq \rho$.

Combinatorial value can be estimated in principle, as for each $N$, there are only a finite number of combinatorial restrictions. In practice, the tensor is presented in such a way that there are “obvious” combinatorial degenerations to disjoint matrix multiplication tensors and at first, one optimizes just among these obvious combinatorial degenerations. However, it may be that there are matrix multiplication tensors of the form $\sum_j a_0 \otimes b_j \otimes c_j$ as well as tensors of the form $a_0 \otimes b_k \otimes c_k$ where $k$ is not in the range of $j$. Then one can merge these tensors to $a_0 \otimes (\sum_j b_j \otimes c_j + b_k \otimes c_k)$ because although formally speaking they were not disjoint, they do not interfere with each other. So the actual procedure is to optimize among combinatorial restrictions with merged tensors.

Coppersmith and Winograd work with the tensors $T_{CW}$ and $\tilde{T}_{CW}$ of (3.4.2) and (3.4.3). They get their best result of $\omega < 2.3755$ by merging $\tilde{T}_{CW}^{\otimes 2}$ and then optimizing. In subsequent work Stothers [Sto], resp. V. Williams [Wil], resp. LeGall [Gal] used merging with $\tilde{T}_{CW}^{\otimes 4}$ resp. $\tilde{T}_{CW}^{\otimes 8}$, resp. $\tilde{T}_{CW}^{\otimes 16}$ and $\tilde{T}_{CW}^{\otimes 32}$ leading to the current “world record”:

**Theorem 3.5.1.5.** [Gal][2014] $\omega < 2.3728639$. 

---

**3.5. The Coppersmith-Winograd method**
Ambainis, Filmus and LeGall [AFL14] showed that taking higher powers of $\tilde{T}_{CW}$ when $q \geq 5$ cannot be used to prove $\omega < 2.30$ by this method alone. Thus one either needs to develop new methods, or find better base tensors.

To get an idea of how the optimization procedure works, start with some base tensor $T$ that contains a collection of matrix multiplication tensors $M_{l_i, m_i, n_i}, 1 \leq i \leq x$ that are not disjoint. Then $T^{\otimes N}$ will contain matrix multiplication tensors of the form $M_{l_\mu, m_\mu, n_\mu}$ where $l_\mu = l_{\mu_1} \cdots l_{\mu_N}$ and similarly for $m_\mu, n_\mu$, where $\mu_j \in [x]$.

Each matrix multiplication tensor will occur with a certain multiplicity and certain variables. The problem becomes to zero out variables in a way that maximizes the value of what remains. More precisely, for large $N$, one wants to maximize the sum $\sum_j M_j(l_\mu, m_\mu, n_\mu)\tilde{\rho}$ where the surviving matrix multiplication tensors are $M_{l_\mu, m_\mu, n_\mu}$ and disjoint. One then takes the smallest $\rho$ such that $\sum_j M_j(l_\mu, m_\mu, n_\mu)\tilde{\rho} \geq R(T)$ and concludes $\omega \leq \rho$.

There are many excellent expositions of this method, which involves exploiting the Salem-Spencer theorem on arithmetic progressions to be assured one can get away with only zero-ing out a relatively small number of terms and in the general case one assigns probability distributions and performs and optimization to determine what percentage of each type gets zero-ed out. I suggest [CW82] for the basic idea and [AFL14] for the state of the art.

3.6. The Cohn-Umans program

A conceptually appealing approach to proving upper bounds on $\omega$ was initiated by H. Cohn and C. Umans.

One works with two geometrically defined bases. In one (the “matrix coefficient basis”), one gets an upper bound on the rank of the tensor, and in the other (the “standard basis”) there are many potential combinatorial degenerations and one gets a lower bound on the value.

Instead of finding arbitrary tensors that combinatorially restrict to matrix multiplication, they look for more structured ones that do: the multiplication tensors in algebras (especially group algebras of finite groups) and more generally tensors constructed from such (“coherent configurations”).

I state the needed representation theory now, and defer proofs of the statements to §5.1. I then present their method.

3.6.1. Preliminaries.

Structure tensor of an algebra. Let $\mathcal{A}$ be an algebra with basis $a_1, \ldots, a_\alpha$ and dual basis $\alpha^1, \ldots, \alpha^\alpha$. Write $a_ia_j = \sum A^k_{ij}a_k$ for the multiplication in
3.6. The Cohn-Umans program

A. Define the structure tensor of $A$

\[
M_A := \sum_{i,j,k} A^k_{ij} \alpha^i \otimes \alpha^j \otimes a_k \in A^* \otimes A^* \otimes A.
\]

The group algebra of a finite group. Let $G$ be a finite group and let $\mathbb{C}[G]$ denote the vector space of complex-valued functions on $G$, called the group algebra of $G$. The following exercise justifies the name:

**Exercise 3.6.1.1:** Show that if the elements of $G$ are $g_1, \ldots, g_r$, then $\mathbb{C}[G]$ has a basis indexed $\delta_{g_1}, \ldots, \delta_{g_r}$, where $\delta_{g_i}(g_j) = \delta_{ij}$. Show that $\mathbb{C}[G]$ may be given the structure of an algebra by defining $\delta_{g_i} \delta_{g_j} := \delta_{g_i g_j}$ and extending linearly.

Thus if $G$ is a finite group with elements $g_1, \ldots, g_q$, then $M_{\mathbb{C}[G]} = \sum_{g,h \in G} \delta_g^* \otimes \delta_h^* \otimes \delta_{gh}$.

**Example 3.6.1.2.**

\[
M_{\mathbb{C}[\mathbb{Z}_m]} = \sum_{0 \leq i,j < m} \delta_i^* \otimes \delta_j^* \otimes \delta_{i+j \text{ mod } m}.
\]

Notice that, labeling the elements of $\mathbb{Z}_m \times 0, \ldots, x_{m-1}$ for $0, \ldots, m - 1$, one obtains a circulant matrix for $M_{\mathbb{C}[\mathbb{Z}_m]}(A^*)$:

\[
M_{\mathbb{C}[\mathbb{Z}_m]}(A^*) = \begin{pmatrix}
  x_0 & x_1 & \cdots & x_{m-1} \\
  x_{m-1} & x_0 & x_1 & \cdots \\
  \vdots & \ddots & \ddots & \ddots \\
  x_1 & x_2 & \cdots & x_0
\end{pmatrix}
\]

What are $R(M_{\mathbb{C}[\mathbb{Z}_m]})$ and $R(M_{\mathbb{C}[\mathbb{Z}_m]})$? The space of circulant matrices forms an abelian subspace, which indicates the rank and border rank might be minimal or nearly minimal among concise tensors.

**Exercise 3.6.1.3:** Show that the centralizer of $M_{\mathbb{C}[\mathbb{Z}_m]}(x_1)$ is $M_{\mathbb{C}[\mathbb{Z}_m]}(A^*)$ and thus $R(M_{\mathbb{C}[\mathbb{Z}_m]}) = m$.

We will determine the rank momentarily via the discrete Fourier transform.

**Burnside’s theorem.** In §5.1, we will see a proof of Burnside’s theorem and the structure theorem of $\mathbb{C}[G]$ when $G$ is finite:

**Theorem 3.6.1.4.** Let $A$ be a finite dimensional simple algebra over $\mathbb{C}$ acting irreducibly on a finite-dimensional vector space $V$. Then $A = \text{End}(V)$. More generally, a finite dimensional semi-simple algebra $A$ over $\mathbb{C}$ is isomorphic to a direct sum of matrix algebras:

\[
A \simeq \text{Mat}_{d_1 \times d_1}(\mathbb{C}) \oplus \cdots \oplus \text{Mat}_{d_q \times d_q}(\mathbb{C})
\]

for some $d_1, \ldots, d_q$. 

Theorem 3.6.1.5. Let $G$ be a finite group, then as a $G \times G$-module and as an algebra,

$$\mathbb{C}[G] = \oplus_i V_i^* \otimes V_i$$

where the sum is over all the distinct irreducible representations of $G$. In particular, if $\dim V_i = d_i$, then

$$\mathbb{C}[G] \simeq \oplus_i \text{Mat}_{d_i \times d_i} (\mathbb{C}).$$

The (generalized) discrete Fourier transform. We have two natural expressions for $M_{\mathbb{C}[G]}$, the original presentation in terms of the algebra multiplication and the matrix coefficient basis in terms of Theorem 3.6.1.5. The change of basis matrix from the standard expression to the matrix coefficient basis is called the (generalized) Discrete Fourier Transform (DFT).

The classical DFT is the case $G = \mathbb{Z}_m$ and the matrix

$$(e^{2\pi i (j+k)/m})_{0 \leq j,k \leq m-1}.$$

Proposition 3.6.1.6. $R(M_{\mathbb{C}[\mathbb{Z}_m]}) = R(M_{\mathbb{C}[\mathbb{Z}_m]}) = m$.

Proof. In the matrix coefficient basis:

$$M_{\mathbb{C}[\mathbb{Z}_m]}(A^*) = \begin{pmatrix} y_0 & y_1 & \cdots & \cdots & y_m-1 \\ y_1 & y_0 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & y_0 & y_1 \\ y_m-1 & y_m-2 & \cdots & \cdots & y_0 \end{pmatrix},$$

where all but the diagonal entries are zero. This is because the irreducible representations of $\mathbb{Z}_m$, where $\sigma \in \mathbb{Z}_m$ is a generator, are all of the form $\sigma \cdot v = e^{2\pi i k/m} v$ for $0 \leq k \leq m$. \qed

In other words, as a tensor $M_{\mathbb{C}[\mathbb{Z}_m]} = M_{\mathbb{Z}_m}^{\otimes m}$. 

Exercise 3.6.1.7: Show that if $T \in \hat{S}_r \circ \hat{S}_h$, then $R(T) \leq r(h+1)$. \(\bigcirc\)

Exercise 3.6.1.8: Obtain a fast algorithm for multiplying two polynomials in one variable by the method you used to solve the previous exercise. \(\bigcirc\)

Consider the example of $S_3$. In the standard basis,

$$M_{\mathbb{C}[S_3]}(A^*) = \begin{pmatrix} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 \\ x_1 & x_0 & x_4 & x_5 & x_2 & x_3 \\ x_2 & x_5 & x_0 & x_4 & x_3 & x_1 \\ x_3 & x_4 & x_5 & x_0 & x_1 & x_2 \\ x_4 & x_3 & x_1 & x_2 & x_5 & x_0 \\ x_5 & x_2 & x_3 & x_1 & x_0 & x_4 \end{pmatrix}$$

Here $x_0 = Id$, $x_1 = (12)$, $x_2 = (13)$, $x_3 = (23)$, $x_4 = (123)$ and $x_5 = (132)$. The irreducible representations of $S_3$ are the trivial [3], the sign [1,1,1] and
the two-dimensional standard representation \([2, 1]\). As a left \(S_3\)-module (rather than as a \(S_3 \times S_3\) module as in the statement of the theorem), \([2, 1] \otimes [2, 1] \simeq [2, 1] \otimes \mathbb{C}^2 \simeq [2, 1]^{\oplus 2}\), so \(M_{[\mathbb{C}[S_3]]} = M_{\{1\}}^{\oplus 2} \oplus M_{\{2\}}\), and in the matrix coefficient basis:

\[
M_{\mathbb{C}[S_3]}(A^*) = \begin{pmatrix}
y_0 \\
y_1 \\
y_2 & y_3 \\
y_4 & y_5 \\
y_2 & y_3 \\
y_4 & y_5
\end{pmatrix}
\]

where the blank entries are zero.

We conclude \(R(M_{\mathbb{C}[S_3]}) \leq 1 + 1 + 7 = 9\).

### 3.6.2. Upper bounds via finite groups

Here is the main idea:

*Use the standard basis to get a lower bound on the value of \(M_{\mathbb{C}[G]}\) and the matrix coefficient basis to get an upper bound on its cost (i.e., its border rank).*

Say \(M_{\mathbb{C}[G]}\) expressed in its standard basis degenerates (in particular, if in standard coordinates it combinatorially restricts) to a sum of matrix multiplications, say \(\oplus_{j=1}^s M_{\{1, m_j, n_j\}}\). The standard basis is particularly well suited to combinatorial restrictions because all the coefficients of the tensor in this basis are zero or one. Using the matrix coefficient basis, we see \(M_{\mathbb{C}[G]} = \oplus_{u=1}^q M_{\{d_u\}}\). Thus \(R(\oplus_{j=1}^s M_{\{1, m_j, n_j\}}) \leq R(\oplus_{u=1}^q M_{\{d_u\}})\) and \(R(\oplus_{j=1}^s M_{\{1, m_j, n_j\}}) \leq R(\oplus_{u=1}^q M_{\{d_u\}})\).

A variant of the proof of asymptotic sum inequality implies:

**Proposition 3.6.2.1.** [CU03, CU] If \(M_{\mathbb{C}[G]}\) degenerates to \(\oplus_{j=1}^s M_{\{1, m_j, n_j\}}\) and \(d_u\) are the degrees of the irreducible representations of \(G\), then \(\sum_{j=1}^s (l_j m_j n_j)^{\frac{3}{2}} \leq \sum d_u^2\).

In this section I will denote the canonical basis for \(\mathbb{C}[G]\) given by the group elements (which I have been denoting \(g_s\)) simply by \(g_s\).

Say \(\mathbb{C}[G]\) in the standard basis admits a combinatorial restriction to \(M_{\{1, m, n\}}\). The images of the functions \(\alpha, \beta, \gamma\) in the combinatorial restriction (3.5.1) may be identified with subsets of \(G\) of the form \(S_1^{-1}S_2, S_2^{-1}S_3\), and \(S_3^{-1}S_1\) with \(|S_1| = 1\), \(|S_2| = m\), \(|S_3| = n\) and such that for any \(s_j, s_j' \in S_j\), if \(s_1 s_1^{-1} s_2 s_2^{-1} s_3 s_3^{-1} = Id\), then \(s_1' = s_1, s_2' = s_2, s_3' = s_3\). This is called the *triple product property* in [CU03]. There is a corresponding simultaneous triple product property when there is a combinatorial restriction to a collection of disjoint matrix multiplication tensors.
Example 3.6.2.2. [CKSU05] Let $G = (\mathbb{Z}_N^3 \times \mathbb{Z}_N^3) \rtimes \mathbb{Z}_2$ where $\mathbb{Z}_2$ acts by switching the two factors. Write elements of $G$ as $[(\omega^a, \omega^b, \omega^c), (\omega^d, \omega^e, \omega^f)]$ where $0 \leq i, j, k, l, m, n \leq N - 1$, $\omega$ is a primitive $N$-th root of unity, and $u \in \{0, 1\}$. Set $1 = m = n = 2N(N - 1)$. Label the elements of $[m] = [2N(N - 1)]$ by a triple $(a, b, \epsilon)$ where $1 \leq a, a' \leq N - 1, 0 \leq b, b' \leq N - 1$ and $\epsilon, \epsilon' \in \{0, 1\}$, and define

$$\alpha : [l] \times [m] \to [a]$$

$$((a, b, \epsilon), (a', b', \epsilon')) \mapsto [(\omega^a, 1, 1)(1, \omega^b, 1)^{-1}[1, \omega^{a'}, 1)(1, 1, \omega^{b'})^{\epsilon'}]$$

$$\beta : [m] \times [n] \to [b]$$

$$((a, b, \epsilon), (a', b', \epsilon')) \mapsto [(1, \omega^a, 1)(1, 1, \omega^b)^{-1}[1, 1, \omega^{a'}, (1, 1, \omega^{b'})^{\epsilon'}]$$

$$\gamma : [n] \times [l] \to [c]$$

$$((a, b, \epsilon), (a', b', \epsilon')) \mapsto [(1, 1, \omega^a)(\omega^b, 1, 1)^{-1}[\omega^{a'}, 1, 1)(1, \omega^{b'}, 1)^{-1}]$$

In other words, $S_1 = [(\omega^a, 1, 1)(1, \omega^b, 1)^{-1}], S_2 = [(1, \omega^a, 1)(1, 1, \omega^b)^{-1}]$ and $S_3 = [(1, 1, \omega^a)(\omega^b, 1, 1)^{-1}], (1, 1, \omega^b)^{-1}]$, where $1 \leq a \leq N - 1, 0 \leq b \leq N - 1$ and $\epsilon \in \{0, 1\}$.

Then one verifies the triple product property (there are several cases).

Here $|G| = 2N^6$, there are $2N^3$ conjugacy classes with 1 element in it and $\binom{N^3}{2}$ conjugacy classes with 2 elements. Thus $R(M_{(m)}) \leq 2N^3 + 8\binom{N^3}{2}$, which is less than $n^3$ for all admissible $n \geq 40$. Asymptotically this is about $\frac{7}{18}n^3$. If one applies Proposition 3.6.2.1 with $N = 17$ (which is optimal), one obtains $\omega < 2.9088$. Note that this does not even exploit Strassen’s algorithm, so one actually has $R(M_{(m)}) \leq 2N^3 + 7\binom{N^3}{2}$, however this does not effect the asymptotics.

While this is worse than what one would obtain just using Strassen’s algorithm (writing $40 = 32 + 8$ and using Strassen in blocks), the algorithm is different. In [CKSU05] they obtain a bound of $\omega < 2.41$ by similar techniques as Coopersmith-Winograd.

3.6.3. Further ideas towards upper bounds. The structure tensor of $\mathbb{C}[G]$ had the convenient property that all the coefficients are zero or one. In [CU] they propose looking at combinatorial restrictions of more general structure tensors, where the coefficients can be more general. They make the following definition, which is very particular to matrix multiplication:

Definition 3.6.3.1. We say $T \in A \otimes B \otimes C$, given in bases $a_\alpha, b_\beta, c_\gamma$ of $A, B, C$, combinatorially supports $M_{(l,m,n)}$, if such that, writing $T = \sum t^{\alpha,\beta,\gamma}a_\alpha \otimes b_\beta \otimes c_\gamma,$
there exist functions

\[ \alpha : [l] \times [m] \to [a] \]
\[ \beta : [m] \times [n] \to [b] \]
\[ \gamma : [n] \times [l] \to [c] \]

such that \( t^{\alpha(i,u'),\beta(u,s')\gamma(s,u')} \neq 0 \) if and only if \( i = i', u = u' \) and \( s = s' \). (Recall that \( T \) combinatorially restricts to \( M_{\langle l,m,n \rangle} \) if moreover \( t^{\alpha(i,u),\beta(u,s)\gamma(s,u)} = 1 \) for all \( i, s, u \).)

\( T \) combinatorially supports \( M_{\langle m,n,l \rangle} \) if there exists a coordinate expression of \( T \) such that, upon setting some of the coefficients in the multidimensional matrix representing \( T \) to zero, one obtains \( mnl \) nonzero entries such that in that coordinate system, matrix multiplication is supported on exactly those \( mnl \) entries. They then proceed to define the \( s \)-rank of a tensor \( T' \), which is the lowest rank of a tensor \( T \) that combinatorially supports it. This is a strange concept because the \( s \)-rank of a generic tensor is one, because a generic tensor is combinatorially supported by \( T = (\sum_j a_j) \otimes (\sum_k b_k) \otimes (\sum_l c_l) \) where \( a_j \) are a basis of \( A \) etc..

Despite this, they show that \( \omega \leq \frac{3}{2} \omega_s - 1 \) where \( \omega_s \) is the analog of the exponent of matrix multiplication for \( s \)-rank. In particular, \( \omega_s = 2 \) would imply \( \omega = 2 \). The idea of the proof is that if \( T \) combinatorially supports \( M_{\langle n \rangle} \), then \( T^\otimes 3 \) combinatorially degenerates to \( M_{\langle n \rangle}^\boxtimes t \) with \( t = O(n^{2-o(1)}) \). (Compare with the situation when \( T \) combinatorially restricts to \( M_{\langle n \rangle} \), then \( T^\otimes 3 \) combinatorially restricts to \( M_{\langle n \rangle} \otimes M_{\langle n^2 \rangle} \) and thus (combinatorially) degenerates to \( M_{\langle n \rangle}^{\boxtimes \left\lceil \frac{3}{2} n^2 \right\rceil} \) by Theorem 2.4.12.4.)
Chapter 4

The complexity of Matrix multiplication III: algorithms and “practical” upper bounds

One might argue that the exponent of matrix multiplication is unimportant for the world we live in, since $\omega$ might not be relevant until the sizes of the matrices are on the order of number of atoms in the known universe. For implementation, it is more important to develop explicit algorithms that provide a savings for matrices of sizes that need to be multiplied in practice.

In this chapter I discuss the known algorithms. I begin, in §4.1 with general remarks on the geometry of decompositions. Then, in §4.2 I present a purely geometric derivation for Strassen’s algorithm for $M(2)$, and in fact all rank six expressions of $M(2)$. There are very few cases where we know the border rank of $M(l,m,n)$. One case where we have reasonable understanding is for $M(2,2,n)$. In §4.3 I describe the known border rank algorithms in these cases and their geometry. There have been numerous algorithms found by numerical methods. In §4.4 I describe one of the main methods for finding such, the alternating least squares method. For $M(3)$ we do not know the rank or border rank, so one cannot hope for too much geometry in any algorithm found. In §4.5 I present several of the rank 23 expressions for $M(3)$ and
remark on their geometry. I conclude in §4.6 with Smirnov’s border rank 20 algorithm for $M_{(3)}$ and a discussion of its geometry.

**Warning:** this chapter is in very rough form.

### 4.1. Geometry of decompositions

#### 4.1.1. The abstract secant variety

I now construct a variety that will facilitate the study of decompositions of a tensor. I make the construction in the more general context of secant varieties.

Let $X \subset \mathbb{P}V$ be a variety. Consider the set

$$
S_r(X)^0 := \{(x_1, \ldots, x_r, z) \in X^{\times r} \times \mathbb{P}V \mid z \in \text{span}\{x_1, \ldots, x_r\}\} \subset \text{Seg}(X^{\times r} \times \mathbb{P}V) \subset \mathbb{P}V^{\otimes r+1}
$$

and let $S_r(X) := \overline{S_r(X)^0}$ denote its Zariski closure. (For those familiar with quotients, it would be more convenient to deal with $X^{(\times r)} := X^{\times r}/\mathfrak{G}_r$.) We have a map $\pi^0 : S_r(X)^0 \to \mathbb{P}V$, extending to a map $\pi : S_r(X) \to \mathbb{P}V$, given by projection onto the last factor and the image is $\sigma_r^0(X)$ (resp. $\sigma_r(X)$).

We will call $S_r(X)$ the *abstract $r$-th secant variety* of $X$. As long as $r < v$ and $X$ is not contained in a linear subspace of $\mathbb{P}V$, $\dim S_r(X) = rn + r - 1$ because $\dim X^{\times r} = rn$ and a general set of $r$ points on $X$ will span a $\mathbb{P}^{r-1}$.

If $\sigma_r(X)$ is of the expected dimension, so its dimension equals that of $S_r(X)$, then for general points $z \in \sigma_r(X)^0$, $(\pi^0)^{-1}(z)$ will consist of a finite number of points and each point will correspond to a decomposition $\pi = \pi_1 + \cdots + \pi_r$ for $\pi_j \in \bar{x}_j$, $\pi \in \bar{z}$. In summary:

**Proposition 4.1.1.1.** If $X^n \subset \mathbb{P}N$ and $\sigma_r(X)$ is of (the expected) dimension $rn + r - 1 < N$, then a Zariski open subset of points on $\sigma_r(X)$ have a finite number of decompositions into a sum of $r$ elements of $X$.

If the fiber of $\pi^0$ over $z \in \sigma_r^0(X)$ is $k$-dimensional, then there is a $k$-parameter family of decompositions of $z$ as a sum of $r$ rank one tensors. This occurs, for example if $z \in \sigma_{r-1}(X)$, but it can also occur for points in $\sigma_r(X) \setminus \sigma_{r-1}(X)$. We will see this is indeed the case for $M_{(2,2,2)} \in \sigma_7(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3))$.

If $X$ is a $G$-variety, then $\sigma_r(X)$ is also a $G$-variety, and if $z \in \sigma_r^0(X)$ is fixed by $G_z \subset G$, then $G_z$ will act (possibly trivially) on $(\pi^0)^{-1}(z)$, and every distinct (up to trivialities) point in its orbit will correspond to a distinct decomposition of $z$. Let $q \in (\pi^0)^{-1}(x)$. If $\dim(G_z \cdot q) = d_z$, then there is at least a $d_z$ parameter family of decompositions of $z$ as a sum of $r$ elements of $X$.

#### 4.1.2. Decompositions of $M_{(n)}$

Let $A = \mathbb{C}^n$. The variety $\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}A \times \mathbb{P}A)) \subset \mathbb{P}(A \otimes A \otimes A)$ is a $G = \text{GL}(A) \times \text{GL}(A) \times \text{GL}(A) \times \mathfrak{G}_3$-variety. Recall from §2.1.3 that $M_{(n)} \in A \otimes A \otimes A$ is invariant under $(\text{SL}_n \times \text{SL}_n \times \text{SL}_n \times \text{SL}_n \otimes \text{SL}_n \otimes \text{SL}_n)$. 

4.2. Strassen’s algorithm for $M_{(2)}$

In this section I give the promised geometric derivation of Strassen’s equations. I begin with a discussion of discrete symmetries of tensors in general.

### 4.2.1. The generalized Comon conjecture

In 2008 there was an AIM workshop, *Geometry and Representation theory of tensors for computer science, statistics and other areas*, that brought together a very diverse group of researchers. Among them was Pierre Comon, an engineer working in signal processing. In signal processing (at least practiced by Comon), one wants to decompose tensors presumed to be of rank $r$ explicitly into a sum of $r$ rank one tensors. Sometimes the relevant tensors are symmetric. At the workshop Comon presented the conjecture that if a tensor happens to be symmetric and of rank $r$ (as a tensor), then it will admit a decomposition as a sum of $r$ rank one symmetric tensors. The algebraic geometers in the audience

...
reacted to the conjecture with, one might say, skepticism. The conjecture is still open, see [BGL13] for a discussion. Here is a generalization:

**Conjecture 4.2.1.1** (Generalized Comon Conjecture). [CHI+] Let $T \in A^{\otimes d}$ and say $R(T) = r$ and $T$ is invariant under some $\Gamma \subseteq S_d$. Then $T$ admits a $\Gamma$-invariant decomposition as a sum of $r$ rank one tensors.

In particular, I expect there to be an optimal decomposition of matrix multiplication that is $\mathbb{Z}_3$-invariant.

**Remark 4.2.1.2.** This conjecture is similar to Strassen’s additivity conjecture straddconj, in that in both cases a tensor that lives in a linear subspace of the ambient space of tensors is conjectured to have an optimal decomposition only from elements of that subspace. See §?? for a discussion.

### 4.2.2. $\mathbb{Z}_3$ invariant tensors in $V^{\otimes 3}$.

To understand $\mathbb{Z}_3$-invariant decompositions, it will be helpful to decompose $V^{\otimes 3}$ under the action of $S_3 \times GL(V)$. First let’s review the decomposition of $V^{\otimes 3}$ under $S_3 \times GL(V)$ from Exercise 2.3.5(2): $V^{\otimes 3} = [2] \otimes S^2 V \oplus [1,1] \otimes \Lambda^2 V$ where $[2] \otimes S^2 V$ and $[1,1] \otimes \Lambda^2 V$ are irreducible $S_3 \times GL(V)$-modules. As an $S_3$-module, the first is the direct sum of $(\binom{\gamma+1}{2})$ copies of the trivial representation and the second is the direct sum of $(\binom{\gamma}{2})$ copies of the sign representation. As a $GL(V)$-module, both modules are irreducible. (To see this, note that the $GL(V)$-orbit of any non-zero vector does not lie in a proper subspace.)

There are three irreducible representations of $S_3$, the trivial $[3]$, the sign $[1,1,1]$, and the complement of the trivial for the action of $S_3$ on $\mathbb{C}^3$, $[2,1]$ (see Example 2.3.2.3), which is two dimensional.

It is straightforward to see that $S^3 V, \Lambda^3 V$ are both irreducible for $GL(V)$, and are the isotypic components respectively of the trivial representation and the sign representation. By counting dimensions, we see we have not accounted for everything. Thus there must be a subspace corresponding to the isotypic component of $[2,1]$, of dimension $\mathbf{v}^3 - \binom{\mathbf{v}+2}{3} - \binom{\mathbf{v}}{2} = \frac{2}{3} (\mathbf{v}^3 - \mathbf{v})$. To get something complementary to $S^3 V$ and $\Lambda^3 V$, consider $V^{\otimes 3} = V^{\otimes 2} \otimes V = (S^2 V \oplus \Lambda^2 V) \otimes V$. We have the $GL(V)$-module symmetrization map $S^2 V \otimes V \to S^3 V$. Its kernel must be a $GL(V)$-module. Since the map is surjective, the dimension of the kernel is $\mathbf{v} (\binom{\mathbf{v}+1}{2} - \binom{\mathbf{v}+2}{3}) = \frac{1}{2} (\mathbf{v}^3 - \mathbf{v})$, exactly half of what we are missing. The kernel of the skew-symmetrization map $\Lambda^2 V \otimes V \to \Lambda^3 V$ provides the rest. Both of these kernels are irreducible $GL(V)$-modules and they are isomorphic, which will be proved in §5.2. Label the irreducible $GL(V)$-module $S_2 V$, so

\[
V^{\otimes 3} = (S^3 V \oplus \mathbf{[3]}) \oplus (S_2 V \otimes S^2) \oplus (\Lambda^3 V \otimes [1,1,1])
\]

as an irreducible $GL(V) \times S_3$-module. In particular, as a $GL(V)$-module,

\[
V^{\otimes 3} = S^3 V \oplus (S_2 V)^{\otimes 2} \oplus \Lambda^3 V.
\]
4.2. Strassen’s algorithm for $M_{(2)}$

Exercise 4.2.2.1: Show that the image of the map $V \otimes \Lambda^2 V \to S^2 V \otimes V$ obtained by inclusion into $V \otimes V \otimes V$ combined with symmetrization on the first two factors also has dimension $\frac{1}{3}(v^3 - v)$ and is isomorphic to $S_{21}V$ as a $GL(V)$-module.

We now consider how $\mathbb{Z}_3 \subset \mathfrak{S}_3$ acts on each factor. It is generated by a cyclic permutation.

Exercise 4.2.2.2: Show that the cyclic permutation $\sigma = (1,2,3)$ acts trivially on $S^3V \oplus \Lambda^3 V$ (i.e., these are the $+1$ eigenspaces for $(1,2,3)$, and $S_{21}V \otimes [2,1]$ splits into a direct sum of eigenspaces for $\omega$ and $\omega^2$ where $\omega = e^{2\pi i/3}$.

To see the splitting in bases, the $\mathbb{Z}_3$ invariant tensors in $V \otimes^3$ have naïve building blocks of three types: $v \otimes v \otimes v$, which lies in $S^3V$, $v \otimes w \otimes v + w \otimes w \otimes v$, which also lies in $S^3V$, and

\[
u_3 \otimes \nu_3 \otimes \nu_3 \otimes \nu_3 \otimes \nu_3 + v \otimes w \otimes v + \omega \otimes w \otimes v + w \otimes v \otimes w + u \otimes w \otimes v \in S^3V \oplus \Lambda^3V.
\]

Exercise 4.2.2.3: Decompose $V \otimes^4$ as a $GL(V)$-module, keeping in mind that $V \otimes^4 = V \otimes^3 \otimes V = V \otimes^2 \otimes V \otimes^2$.

4.2.3. Decompositions of $M_{(2)}$. We are looking for a nine-parameter family of decompositions of $M_{(2)}$, parametrized by $SL_2 \times SL_2 \times SL_2$, among which there should be $\mathbb{Z}_3$-invariant algorithms. Recall the invariant description of $M_{(2)}$: let $U, V, W = \mathbb{C}^2$, and write $M_{(2)} = Id_U \otimes Id_V \otimes Id_W \in (U^* \otimes V^* \otimes W^*)(V^* \otimes U) = A \otimes B \otimes C$. To obtain an $\mathbb{Z}_3$-invariant algorithm, we will need to identify $A, B, C$, which we may do by identifying $U, V, W$. Explicitly, choose isomorphisms $a_0 : U \to V$ and $b_0 : V \to W$, which determines $c_0 = a_0^{-1}b_0^{-1} : W \to U$. These elements, in bases for $V, W$ induced from those of $U$, will correspond to identity matrices. The choices have already killed off two of our $SL_2$’s and we are only left with one, the diagonal $SL_2$ inside $SL(U) \times SL(V) \times SL(W)$, which for convenience I will identify with $SL(U)$.

I have been a little sloppy above, since the group we are really interested in is the action of $GL(U) \times GL(V) \times GL(W)$ on our space, and when the general linear group acts on projective space, the image group is denoted $PGL_m := GL_m / \mathbb{C}^*$.

It remains to exactly use up a $PGL(U)$’s worth of freedom. Now $PGL_2$ acts three-transitively on $\mathbb{CP}^1$, that is, given any three distinct points
4. Algorithms for matrix multiplication

in \( \mathbb{CP}^1 \simeq S^2 \), there exists exactly one element of \( PGL_2 \) that will move them to e.g., \( 0, 1, \infty \), i.e., \([1, 0], [1, 1], [0, 1]\). So our algorithm should be obtained from the choice of \( a_0, b_0 \) and three distinct points \([u_1], [u_2], [u_3] \in \mathbb{P}U\). Moreover, the algorithm should be unique if it is also \( Z_3 \)-invariant.

We need to be careful here about what we mean by \( Z_3 \) and \( Z_2 \)-invariance. Write \( Z_3^{std} \) for the \( Z_3 \subset \mathfrak{S}_3 \) that is cyclic permutation and \( Z_2^T \) for the \( Z_2 \) corresponding to transpose. The algorithm should be \( Z_3^{std} \)-invariant if we represent \( a_0, b_0 \) (and hence \( c_0 \)) by identity matrices, and then the additional \( Z_2 \)-action will be \( Z_2^T \).

Our choices also determine three points \([u_1^\perp], [u_2^\perp], [u_3^\perp] \in \mathbb{P}U^*\), as the annihilator of a line in \( \mathbb{C}^2 \) is a line in \( \mathbb{C}^{2*} \). Applying \( a_0 \) etc., gives us triples of points in each of \( \mathbb{P}V, \mathbb{P}V^*, \mathbb{P}W, \mathbb{P}W^* \) as well.

By the above principles, we should be able to obtain matrix multiplication as a sum of seven rank one tensors in a \( Z_3 \)-invariant manner only from this data. Since the \( Z_3 \)-orbit of a rank one element of \( A \otimes B \otimes C \) is either a rank one element or rank three, the possible decompositions are \( 7 = 1+3+3 \), \( 7 = 1+1+1+1+3 \), or \( 7 = 1+1+1+1+1+1+1 \). The last case is ruled out because it would result in a symmetric tensor. The first case is most promising, as we have the singleton \( a_0 \otimes b_0 \otimes c_0 \) which is \( Z_3^{std} \) invariant when these are identity matrices, and to obtain two triplicities, we may use the six \( u_i^\perp \otimes v_j \) with \( i \neq j \) terms. Note that as linear maps, the terms \( u_i^\perp \otimes v_j \) all have rank one.

Indeed in Strassen’s algorithm, all terms in it except the first are the tensor product of three of rank one matrices, moreover, the first term, as matrices, is \( Id \otimes Id \otimes Id \), which corresponds to \( a_0 \otimes b_0 \otimes c_0 \).

Consider the term \((u_1^\perp \otimes v_2) \otimes (v_3^\perp \otimes w_1) \otimes (w_2^\perp \otimes u_3)\) The sum of its images under the action of \( Z_3 \) is

\[
\begin{align*}
((u_1^\perp \otimes v_3) \otimes (v_2^\perp \otimes w_1) \otimes (w_3^\perp \otimes u_2))_{Z_3} := \\
(u_1^\perp \otimes v_3) \otimes (v_2^\perp \otimes w_1) \otimes (w_3^\perp \otimes u_2) + (a_0^T(v_2^\perp) \otimes b_0^{-1}(w_1)) \otimes (b_0^T(w_3^\perp) \otimes c_0^{-1}(u_2)) \otimes (c_0(a_0^\perp) \otimes a_0^{-1}(v_3)) \\
+ ((c_0^{-1})^T(w_3^\perp) \otimes a_0(u_2)) \otimes ((a_0^{-1})^T(u_1^\perp) \otimes b_0(v_3)) \otimes ((b_0^{-1})^T(v_2^\perp) \otimes c_0(w_1)) \\
= \\
(u_1^\perp \otimes v_3) \otimes (v_2^\perp \otimes w_1) \otimes (w_3^\perp \otimes u_2) \\
+ (u_2^\perp \otimes v_1) \otimes (v_3^\perp \otimes w_2) \otimes (w_1^\perp \otimes u_3) \\
+ (u_3^\perp \otimes v_2) \otimes (v_1^\perp \otimes w_3) \otimes (w_2^\perp \otimes u_1).
\end{align*}
\]

We have slightly violated our principles, as \( u_i^\perp, v_j \) etc., are only defined up to scale, so we must adjust our expression to be invariant under our choices of scale. Set \( \lambda = u_1^\perp(v_2)v_3^\perp(w_1)w_2^\perp(u_3) \) and \( \mu = u_1^\perp(v_3)v_2^\perp(w_1)w_3^\perp(u_2) \).
When we divide the first triple by $\lambda$ and the second by $\mu$, we obtain elements that do not depend on our choices of scales.

**Theorem 4.2.3.1.** [CHI⁺] Notations as above. The following is a 9 parameter family, parametrized by $\text{SL}(U) \times \text{SL}(V) \times \text{SL}(W)$, of rank seven expressions for $M_{(2)}$:

\[
M_{(2)} = a_0 \otimes b_0 \otimes c_0 + \frac{1}{\mu} \langle (u_1^\perp \otimes v_3) \otimes (v_2^\perp \otimes w_1) \otimes (w_3^\perp \otimes u_2) \rangle_{\mathbb{Z}_3} \\
+ \frac{1}{\lambda} \langle (u_1^\perp \otimes v_2) \otimes (v_3^\perp \otimes w_1) \otimes (w_2^\perp \otimes u_3) \rangle_{\mathbb{Z}_3}
\]

These algorithms, are all possible expressions of $M_{(2)}$ as a sum of seven rank one tensors (up to trivialities).

If we fix bases such that $a_0 = b_0 = \text{Id}$, then the algorithms are also $\mathbb{Z}_3^{std}$-invariant and we obtain a three parameter family of $\mathbb{Z}_3^{std}$-invariant algorithms.

To prove (4.2.2) holds, of course one could check it directly, e.g., taking $u_1 = (1,0)$, $u_2 = (0,1)$, $u_3 = (1,1)$, $u_1^\perp = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $u_2^\perp = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $u_3^\perp = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, and $a_0 = b_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. However instead I will give a geometric proof.

**Proof.** Since $\mathbb{C}\{M_{(U,V,W)}\}$ is the unique instance of the trivial representation of $GL(U) \times GL(V) \times GL(W)$ in $A \otimes B \otimes C = (U^* \otimes V) \otimes (V^* \otimes W) \otimes (W^* \otimes U)$, the matrix multiplication tensor is characterized by its symmetry group, in the sense that no other tensor in $A \otimes B \otimes C$ is preserved by a group $G' \subseteq GL(U) \times GL(V) \times GL(W)$. Thus to prove (4.2.2) it is sufficient to show the right hand side is preserved by $SL(U) \times SL(V) \times SL(W)$, or equivalently annihilated by $\mathfrak{sl}(U) + \mathfrak{sl}(V) + \mathfrak{sl}(W)$.

Assume $a_0 = b_0 = \text{Id}$. By $\mathbb{Z}_3$-invariance it is sufficient to show the right hand side is annihilated by $\mathfrak{sl}(U)$, which has basis $u_1^\perp \otimes u_1$, $u_2^\perp \otimes u_2$, $u_3^\perp \otimes u_3$.

By the symmetry of the algorithm it is sufficient to show it is annihilated by $u_1^\perp \otimes u_1$, which is easy to check.

The penultimate assertion is Theorem 4.1.2.1, combined with the observation that the $\mathbb{Z}_3$ action takes an element of the family to another element of the family. The $\mathbb{Z}_3$-invariance in the final assertion was already shown. \(\square\)

Show that no expression is fixed by $\mathbb{Z}_2$ and that $\mathbb{Z}_2$ indeed takes an expression in the family to another expression in the family.

**Exercise 4.2.3.2:** Verify that $u_1^\perp \otimes u_1$, $u_2^\perp \otimes u_2$, $u_3^\perp \otimes u_3$ is a basis of $\mathfrak{sl}(U)$. 
4.3.1. Geometry of the expression (3.2.2) Recall
\[ T_{BCLR} = \lim_{t \to 0} \frac{1}{t} [(x^1_2 + tx^1_1) \otimes (y^1_1) \otimes (z^1_1 + tx^1_2) \\
+ (x^1_2 + tx^1_1) \otimes (y^1_1) \otimes (z^1_1 + tx^1_2) \\
- x^1_2 \otimes y^1_1 \otimes ((z^1_1 + z^2_1) + tx^2_2) \\
- x^1_1 \otimes (y^1_1 + y^2_1) \otimes (z^1_1 + z^2_1) \\
+ (x^1_2 + x^2_1) \otimes (y^1_2 + ty^2_1) \otimes (z^1_1 + z^2_2)] \]

***This subsection to be expanded***

If we want to find expressions such as (3.2.2) for larger matrix multiplication tensors, we will need to develop geometric insight for the examples we already have. What follows are remarks in this direction.

First, from our lower bounds for the rank and border rank of matrix multiplication, it should not be surprising that \( R(T_{BCLR}) < R(T_{BCLR}) \).

The group preserving \( T_{BCLR} \) includes \( SL_2 = SL(W) \) and \( \mathbb{Z}_2 \) which acts by \( (X, Y, Z) \mapsto (X^T, Z^T, Y^T) \).

The key geometric feature of the expression (3.2.2) is based on the following general fact: If \( X \subset \mathbb{P}V \) is a variety, and \( x_1, \ldots, x_r \in X \) are linearly dependent elements of \( V \), then any point on the sum of their tangent spaces is in \( \sigma_r(X) \). The idea of the proof is as follows: consider instead the Grassmannian \( G(r, V) \). For any curves \( x_j(t) \) with \( x_j(0) = x_j \) such that \( x_1(t) \wedge \cdots \wedge x_r(t) \neq 0 \) for \( t \neq 0 \), i.e., the points span an \( r \)-plane for all \( t \neq 0 \), consider the limiting plane in \( G(r, V) \):

\[
\lim_{t \to 0} [x_1(t) \wedge \cdots x_r(t)] = \lim_{t \to 0} [(x_1 + tx'_1 + O(t^2)) \wedge \cdots \wedge (x_r + tx'_r + O(t^2))] \\
= \lim_{t \to 0} \frac{1}{t} [(x_1 + tx'_1 + O(t^2)) \wedge \cdots \wedge (x_r + tx'_r + O(t^2))].
\]

The zero-th order term in the expansion is zero. The first order term in general will be nonzero, and one can arrange for it to be \( x_1 \wedge \cdots x_{r-1} \wedge v \) where \( v \) is an arbitrary linear combination of elements of the \( T_{x_j}X \). For details, see [Lan12, §11.2].

The five points in (3.2.2) are:

\[(4.3.1) x^1_2 \otimes y^1_2 \otimes z^1_2, x^2_2 \otimes y^1_2 \otimes z^1_1, -x^1_2 \otimes y^1_2 \otimes (z^1_1 + z^2_1), -x^2_1 \otimes y^1_1 \otimes z^1_1, (x^1_2 + x^2_1) \otimes y^1_2 \otimes z^1_1.\]
and indeed they sum to zero. Notice all but $x_1^2 + x_1^2$ of the six elements appearing in the expression, considered as matrices, have rank one.

Now consider the five tangent vectors:

\[
\begin{align*}
&x_2^1 y_2^2 \otimes z_2^2 + x_1^1 y_2^1 \otimes z_2^1, \\
&x_1^2 y_1^1 \otimes z_1^2 + x_1^1 y_1^1 \otimes z_1^1, \\
&-x_2^1 y_2^1 \otimes z_2^2, \\
&-x_1^2 y_1^1 \otimes z_1^1, \\
&(x_1^2 + x_2^2) y_1^2 \otimes z_1^1 + (x_1^1 + x_2^1) y_2^1 \otimes z_2^2.
\end{align*}
\]

The first two and the last contribute two terms that appear in the standard expression, the last term also contributes two “bad” terms that do not appear, which are canceled by the single terms appearing in the third and fourth terms. Note that none of these are generic tangent vectors, and that it was the point containing the rank two matrix that generated two bad terms along with the two good ones that had to be canceled.

In general, given $r - 1$ points on a variety $X \subset \mathbb{P}V$ of codimension at least $r - 1$, it is rare that one can find an $r$-th point on $X$ contained in their span. Thus it is worthwhile to see “why” one can find such points in our case, so I take short detour.

Consider the case $r = 3$: An easy way for there to exist a third point on $X$ on the line spanned by two points of $X$ is if the two points lie on a line contained in $X$.

The Segre variety indeed has many linear spaces on it, but the points used in (3.2.2) do not all lie on a single linear space contained in the Segre. However the structure of lines on the Segre does arise in their construction. So I begin with a discussion of lines on the smallest Segre varieties:

Consider first lines (i.e. linear $\mathbb{P}^1$’s) on $Seg(\mathbb{P}^1 \times \mathbb{P}^1) = Seg(\mathbb{P}A \times \mathbb{P}B)$. They come in two types, call them $\alpha$-lines, those of the form $[a \otimes B]$ for some $[a] \in \mathbb{P}A$, and $\beta$-lines, which are of the form $[A \otimes b]$ for some $[b] \in \mathbb{P}B$.

**picture of hyperbola with double ruling to appear here**

**Exercise 4.3.1.1:** Show that no two lines of the same type intersect, and that each $\alpha$-line intersects each $\beta$-line in one point.

Now consider $Seg(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) = Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$, and call the analogous lines of type $\alpha, \beta, \gamma$.

**Exercise 4.3.1.2:** Show that given a $\beta$-line and a $\gamma$-line, either the two lines do not intersect or they intersect at one point. Show that in either case there is a unique $\alpha$-line intersecting both lines.
Now for the description of the points (4.3.1): Since $B, C$ are play the same role in $T_{BCLR}$, one expects the point selection to be symmetric in $B, C$. Now $B, C$ have internal structure as $V^* \otimes W$ and $W^* \otimes U$. We first pick a $\beta$-line and a $\gamma$-line. To pick the $\beta$ line we need a line in $\mathbb{P}B$ and points in $\mathbb{P}A, \mathbb{P}C$: take a line on $\text{Seg}(\mathbb{P}V^* \times \mathbb{P}W) \subset \mathbb{P}B$, a point on $\mathbb{P}A \cap \text{Seg}(\mathbb{P}U^* \times \mathbb{P}V)$, and a point on $\text{Seg}(\mathbb{P}W^* \times \mathbb{P}U) \subset \mathbb{P}C$. Pick the $\gamma$-line by taking a line in $\text{Seg}(\mathbb{P}V^* \times \mathbb{P}W)$ containing the point picked for the $\beta$-line, a point in $\mathbb{P}C$ lying on the line in $\text{Seg}(\mathbb{P}W^* \times \mathbb{P}U)$ that was chosen for the $\beta$-line, and a point of $\mathbb{P}A \cap \text{Seg}(\mathbb{P}U^* \times \mathbb{P}V)$ that is not collinear to the point chosen for the $\beta$ line. The points chosen all lie in some $\mathbb{P}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2) \subset \mathbb{P}(A \otimes B \otimes C)$ and such that there will be a unique $\alpha$-line in this space intersecting the other two lines.

**picture of the configuration to go here**

The points (4.3.1) are taken by taking two points on the $\beta$-line, and two points on the $\gamma$-line and then the fifth point on the $\alpha$-line that passes through the other two lines.

Both the $\beta$ and $\gamma$-lines contain exclusively rank one matrices, whereas the $\alpha$ line consists of $B$ and $C$ matrices that are rank one, and all $A$-points on the line except the two points of intersection with the $\beta$ and $\gamma$-lines have rank two.

**Problem 4.3.1.3.** Can one show this construction, plus perhaps some additional geometric constraints, uniquely determines the algorithm up to the $G$-action?

**Problem 4.3.1.4.** Formulate general principles for algorithms for $T_{BCLR}$ and other restricted matrix multiplication operators based on the above remarks.

### 4.3.2. Smirnov’s algorithms for $M_{(2,2,n)}$.

The algorithm that shows $R(M_{(2,2,3)}) \leq 10$, due to Bini et. al., is just Bini’s algorithm for $2 \times 2$ matrices with one entry zero applied twice. Here is an algorithm for the next case:

By the principles described in §4.2.1, the optimal algorithm should have $\mathbb{Z}_2$ symmetry coming from the equality $\text{trace}(XYZ) = \text{trace}(Z^TY^TX^T)$ (although strictly speaking, this is more general than the generalized Comon conjecture) and come in families parametrized by $\text{SL}_2 \times \text{SL}_2 \times \text{SL}_4$. That is, they should correspond to a choice of 3 points in each of $U, V \simeq \mathbb{C}^2$ and five points in $W = \mathbb{C}^4$. 
4.6. Smirnov’s algorithm showing the border rank of $M_{(3)}$ is at most 20

The algorithm in [AS13] is $\frac{1}{7}$ times
\[
(x_2^2 + x_4^2 - tx_1^2)(-y_1^4 + y_2^2)(-z_1^1 + z_3^1 - tz_1^1 + t^2z_2^2)
+ (-x_2^2 + tx_1^2)(y_2^2 - ty_1^1)(-z_1^1 + z_3^1 + t(-z_1^1 + z_1^1))
+ (-x_2^2 + x_4^2) + t(x_2^2 - x_1^2) + t^2x_1^1)y_1^1(z_1^1 - t^2z_2^2)
+ (x_2^2 + x_4^2 + t(-x_1^4 - x_2^1))(-y_1^1 + ty_1^1)(-z_3^1 + tz_1^1)
+ x_2^2y_1^2(z_1^1 - z_3^1 + tz_1^1 - t^2z_2^2)
+ x_2^2(y_1^2 + t(-y_1^1 + y_1^1))(-z_1^1 + tz_1^1 + t^2z_2^2)
+ (x_2^2 + tx_1^3 + t^2x_1^3)(y_1^2 - ty_1^2)(-z_1^3 + z_3^3)
+ (x_1^4 + tx_1^3 + t^2x_1^3)(y_1^2 + ty_1^1 - t^2y_1^1)(z_3^1 + t^2z_3^1)
+ (x_1^4 + tx_1^3)(-y_1^2 + t(y_1^1 + y_1^2))z_3^2
+ (x_2^2 + t^2x_1^1)(y_1^2 + ty_1^1 - t^2y_1^1)z_2^2
+ (x_1^2 + x_2^2)(y_1^2 + ty_1^2)(z_1^1 + tz_1^1)
+ (x_2^4 - x_1^4 + t(x_2^3 - x_2^1))y_1^1z_3^1
+ x_2^2y_1^2(-z_1^1 + z_3^1 - t^2z_2^2)
\]

This algorithm is not far from being $\mathbb{Z}_2$ invariant. Comparing the $x$’s and the $z$’s, there are 3 terms each with $t$ appearing to orders 0, 1, 2, 4 terms each with $t^0$ only and then for orders 0, 2 $x$ has 1 term and $z$ two, while for orders 0, 1, $x$ has 5 terms and $z$ 4. This gives the first hint how to get a $\mathbb{Z}_2$-invariance.

4.4. Alternating least squares approach to algorithms

This section to be written.

4.5. Algorithms for $3 \times 3$ matrix multiplication with 23 multiplications

Will discuss how there are many algorithms for $3 \times 3$, what they all have in common, uniqueness issues. under construction

4.6. Smirnov’s algorithm showing the border rank of $M_{(3)}$ is at most 20

To be added.
Chapter 5

Representation theory for complexity theory

This chapter contains proofs of several results regarding matrix multiplication and establishes the basic representation theory needed for Geometric Complexity Theory. It begins, in §5.1 with the algebraic Peter-Weyl theorem, which is central to GCT. This section also contains a proof of Burnside’s theorem owed from Chapter 3. Next, in §5.2 the representation theory of \( \mathfrak{S}_d \) and \( GL(V) \) is developed with an emphasis on practical implementation and addressing issues that arise in the study of matrix multiplication. In particular, I explain the origin of the Koszul flattenings and the more general Young flattenings. It concludes, in §5.3, with a discussion of various methods for finding modules of equations for secant varieties of Segre varieties via representation theory.

For supplementary material to this chapter, I suggest Lectures 1-4 and 6 in [FH91] or [Lan12, Chap. 6].

5.1. Double-Commutant Theorem and first consequences

5.1.1. Algebras and their modules. It will sometimes be advantageous to work with the Lie algebra of \( GL(V) \) and the group algebra \( \mathbb{C}[G] \) of a finite group \( G \) because of the linear structure of the algebras. Any \( G \)-module is naturally a \( \mathbb{C}[G] \)-module and vice versa.

For an algebra \( \mathcal{A} \), and \( a \in \mathcal{A} \) the ideal \( \mathcal{A}a \) is a (left) \( \mathcal{A} \)-module.

Define a representation \( L : G \to GL(\mathbb{C}[G]) \) by \( L(g)\delta_y = \delta_{gy} \) and extending the action linearly. Define a second representation \( R : G \to GL(\mathbb{C}[G]) \)
by $R(g)\delta_{g_i} = \delta_{g_{g^{-1}}}$. Thus $\mathbb{C}[G]$ is a $G \times G$-module under the representation $(L, R)$, and for all $c \in \mathbb{C}[G]$, $\mathbb{C}[G]c$ is a $G$-module under the action $L$.

A representation $\rho : G \to GL(V)$ induces an algebra homomorphism $\mathbb{C}[G] \to \text{End}(V)$, and it is equivalent that $V$ is a $G$-module or a left $\mathbb{C}[G]$-module.

Terminology for algebras is slightly different than for groups. A module $M$ (for a group, ring, or algebra) is simple if it has no proper submodules. The module $M$ is semi-simple if it may be written as the direct sum of simple modules. An algebra is completely reducible if all its modules are semi-simple. For groups alone I will continue to use the terminology irreducible for a simple module, completely reducible for a semi-simple module, and reductive for a group such that all its modules can be decomposed into a direct sum of irreducible modules.

Exercise 5.1.1.1: Show that if $A$ is completely reducible, $V$ is an $A$-module with an $A$-submodule $U \subset V$, then there exists an $A$-invariant complement to $U$ in $V$ and a projection map $\pi : V \to U$ that is an $A$-module map. ⊙

5.1.2. The double-commutant theorem. Our sought-after decomposition of $V \otimes^d$ as a $GL(V)$-module will be obtained by exploiting the fact that the actions of $GL(V)$ and $S_d$ on $V \otimes^d$ commute. In this subsection we study commuting actions in general, as this is also the basis of the generalized DFT used in the Cohn-Umans method and the key to much of GCT. References for this section are [Pro07, Chap. 6], [GW09, §4.1.5] and [Lan02, Chap. XVII]. Let $S \subset \text{End}(V)$ be any subset. Define the centralizer or commutator of $S$ to be

$$S' := \{X \in \text{End}(V) \mid Xs = sX \forall s \in S\}$$

Proposition 5.1.2.1.

(1) $S'$ is an algebra.

(2) $S \subset (S')'$.

Exercise 5.1.2.2: Prove Proposition 5.1.2.1.

Theorem 5.1.2.3. [Double-Commutant Theorem] Let $A \subset \text{End}(V)$ be a completely reducible associative algebra. Then $A'' = A$.

Proof. By Proposition 5.1.2.1, $A \subseteq A''$. To show the reverse inclusion, say $T \in A''$. Fix a basis $v_1, \ldots, v_v$ of $V$. Since the action of $T$ is determined by its action on a basis, we need to find $a \in A$ such that $av_j = Tv_j$ for $j = 1, \ldots, v$. Let $w := v_1 \oplus \cdots \oplus v_v \in V^{\oplus v}$ and consider the submodule $Aw \subseteq V^{\oplus v}$. By Exercise 5.1.1.1, there exists an $A$-invariant complement to this submodule and an $A$-equivariant projection $\pi : V^{\oplus v} \to Aw \subset V^{\oplus v}$, that is, a projection $\pi$ that commutes with the action of $A$, i.e., $\pi \in A'$. Since
5.1. Double-Commutant Theorem and first consequences

Let \( T \in \mathcal{A}'' \) we have \( \pi(Tw) = T(\pi(w)) \) but \( T(\pi(w)) = T(w) = Tv_1 \oplus \cdots \oplus Tv_v \). But since \( \pi(Tw) \in \mathcal{A}w \), there must be some \( a \in \mathcal{A} \) such that \( aw = T(w) \), i.e., \( av_1 \oplus \cdots \oplus av_v = Tv_1 \oplus \cdots \oplus Tv_v \), i.e., \( av_j = Tv_j \) for \( j = 1, \ldots, v \). □

Adopt the notation \( \text{Hom}_S(V) := S' \).

Burnside’s theorem, stated in §3.6, has a similar proof:

**Theorem 5.1.2.4.** Let \( \mathcal{A} \) be a finite dimensional simple algebra over \( \mathbb{C} \) acting irreducibly on a finite-dimensional vector space \( V \). Then \( \mathcal{A} = \text{End}(V) \). More generally, a finite dimensional semi-simple algebra \( \mathcal{A} \) over \( \mathbb{C} \) is isomorphic to a direct sum of matrix algebras:

\[
\mathcal{A} \cong \text{Mat}_{d_1 \times d_1}(\mathbb{C}) \oplus \cdots \oplus \text{Mat}_{d_q \times d_q}(\mathbb{C})
\]

for some \( d_1, \ldots, d_q \).

**Proof.** For the first assertion, we need to show that given \( X \in \text{End}(V) \), there exists \( a \in \mathcal{A} \) such that \( av_j = Xv_j \) for \( v_1, \ldots, v_v \) a basis of \( V \). Now just imitate the proof of Theorem 5.1.2.3. For the second assertion, note that \( \mathcal{A} \) is a direct sum of simple algebras. □

**Remark 5.1.2.5.** A pessimist could look at this theorem as a disappointment: all kinds of interesting looking algebras over \( \mathbb{C} \), such as the group algebra of a finite group, are actually just plain old matrix algebras in disguise. An optimist could view this theorem as stating there is a rich structure hidden in matrix algebras. We will determine the matrix algebra structure explicitly for the group algebra of a finite group.

5.1.3. Consequences for reductive groups. Let \( S \) be a group or algebra and let \( V, W \) be \( S \)-modules, adopt the notation \( \text{Hom}_S(V, W) \) for the space of \( S \)-module maps \( V \to W \), i.e.,

\[
\text{Hom}_S(V, W) := \{ f \in \text{Hom}(V, W) \mid s(f(v)) = f(s(v)) \forall s \in S, \ v \in V \} = (V^* \otimes W)^S.
\]

**Theorem 5.1.3.1.** Let \( G \) be a reductive group and let \( V \) be a \( G \)-module. Then

1. The commutator \( \text{End}_G(V) \) is a semi-simple algebra.
2. The isotypic components of \( G \) and \( \text{End}_G(V) \) in \( V \) coincide.
3. Let \( U \) be one such isotypic component, say for irreducible representations \( A \) of \( G \) and \( B \) of \( \text{End}_G(V) \). Then, as a \( G \times \text{End}_G(V) \)-module,

\[
U = A \otimes B,
\]

as an \( \text{End}_G(V) \)-module

\[
B = \text{Hom}_G(A, U),
\]
and as a $G$-module

$$A = \text{Hom}_{\text{End}_G(V)}(B, U).$$

In particular, $\text{mult}(A, V) = \dim B$ and $\text{mult}(B, V) = \dim A$.

**Example 5.1.3.2.** Below we will see that $\text{End}_{GL(V)}(V \otimes^d) = \mathbb{C}[S_d]$. Recall from equation (4.2.1) the $S_3 \times GL(V)$-module decomposition $V \otimes^3 = ([3] \otimes S^3 V) \oplus ([2, 1] \otimes S_{21} V) \oplus ([1, 1, 1] \otimes \Lambda^3 V)$ which illustrates Theorem 5.1.3.1.

To prove the theorem, we will need the following lemma:

**Lemma 5.1.3.3.** For $U \subset V$ a $G$-submodule and $f \in \text{Hom}_G(U, V)$, there exists $a \in \text{End}_G(V)$ such that $a|_U = f$.

**Proof.** Consider the diagram

$$\begin{array}{ccc}
\text{End}(V) & \rightarrow & \text{Hom}(U, V) \\
\downarrow & & \downarrow \\
\text{End}_G(V) & \rightarrow & \text{Hom}_G(U, V)
\end{array}$$

The vertical arrows are $G$-equivariant projections, and the horizontal arrows are restriction of domain of a linear map. The diagram is commutative. Since the vertical arrows and upper horizontal arrow are surjective, we conclude the lower horizontal arrow is surjective as well. □

**Proof of Theorem.** I first prove (3): The space $\text{Hom}_G(A, V)$ is an $\text{End}_G(V)$-module because for $s \in \text{Hom}_G(A, V)$ and $a \in \text{End}_G(V)$, the composition $as : A \rightarrow V$ is still a $G$-module map. We need to show (i) that $\text{Hom}_G(A, V)$ is irreducible and (ii) that the isotypic component of $A$ in $V$ is $A \otimes \text{Hom}_G(A, V)$.

To show (i), it is sufficient to show that for all $s, t \in \text{Hom}_G(A, V)$, there exists $a \in \text{End}_G(V)$ such that $at = s$. By Lemma 5.1.3.3, there exists $a \in \text{End}_G(V)$ such that $a(tA) = sA$, i.e., $at : A \rightarrow sA$ is an isomorphism. Consider the isomorphism $S : A \rightarrow sA$, given by $a \mapsto sa$, so $S^{-1}at$ is a nonzero scalar $c$ times the identity. Then $\tilde{a} := \frac{1}{c}a$ has the property that $\tilde{a}t = s$.

To see (ii), let $U$ be the isotypic component of $A$, so $U = A \otimes B$ for some vector space $B$. Let $b \in B$ and define a map $\tilde{b} : A \rightarrow V$ by $a \mapsto a \otimes b$, which is a $G$-module map where the action of $G$ on the target is just the action on the first factor. Thus $B \subseteq \text{Hom}_G(A, V)$. Any $G$-module map $A \rightarrow V$ by definition has image in $U$, so equality holds.

(3) implies (2).

To see (1), note that $\text{End}_G(V)$ is semi-simple because if the irreducible components of $V$ are $U_i$, then $\text{End}_G(V) = \bigoplus_i \text{End}_G(U_i) = \bigoplus_i \text{End}_G(A; \otimes B_i) = \bigoplus_i \text{End}(B_i)$. □
5.1.4. Matrix coefficients. For a large class of groups, one can obtain all their irreducible representations from the ring of regular functions on \( G \), denoted \( \mathbb{C}[G] \). (Here \( G \) is not usually a projective variety, it is a quasi-projective variety, see, e.g., [Sha94, §4] and one can still define regular functions on \( G \). For a finite group, this just means any function.) Let \( G \) be any group. Let \( \rho : G \to GL(V) \) be a finite dimensional representation of \( G \). Define a map \( i_V : V^* \otimes V \to \mathbb{C}[G] \) by \( i_V(\alpha \otimes v)(g) := \alpha(\rho(g)v) \). The space of functions \( i_V(V^* \otimes V) \) is called the space of matrix coefficients of \( V \).

**Exercises.**

1. Show \( i_V \) is a \( G \times G \)-module map.
2. Show that if \( V \) is irreducible, \( i_V \) is injective. \( \circ \)
3. If we choose a basis \( v_1, \ldots, v_v \) of \( V \) with dual basis \( \alpha^1, \ldots, \alpha^v \), then \( i_V(\alpha^i \otimes v_j)(g) \) is the \((i,j)\)-th entry of the matrix representing \( \rho(g) \) in this basis (which explains the name “matrix coefficients”).
4. Compute the matrix coefficient basis of the three irreducible representations of \( S_3 \) in terms of the standard basis \( \delta_\sigma, \sigma \in S_3 \).
5. Let \( G = GL_2 \mathbb{C} \), write \( g = \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \in G \), and compute the matrix coefficient basis as functions of \( a, b, c, d \) when \( V = S_2^2 \mathbb{C}^2, S_3^3 \mathbb{C}^2 \) and \( \Lambda^2 \mathbb{C}^2 \).

**Theorem 5.1.4.1.** Let \( G \) be any group and \( V \) an irreducible \( G \)-module. Then \( i_V(V^* \otimes V) \) equals the isotypic component of type \( V \) in \( \mathbb{C}[G] \) under the action \( R \) and the isotypic component of \( V^* \) in \( \mathbb{C}[G] \) under the action \( L \).

**Proof.** It suffices to prove one of the assertions, consider the action \( R \). Let \( j : V \to \mathbb{C}[G] \) be a \( G \)-module map under the action \( R \). We need to show \( j(V) \subseteq i_V(V^* \otimes V) \). Define \( \alpha \in V^* \) by \( \alpha(v) := j(v)(Id) \). Then \( j(v) = i_V(\alpha \otimes v) \), as \( j(v)g = j(v)(Id \cdot g) = j(gv)(Id) = \alpha(gv) = i_V(\alpha \otimes v)g \). \( \square \)

The matrix coefficients provide many functions on a group. For reductive algebraic groups, we will see they provide all the algebraic functions \( \mathbb{C}[G] \). To make sense of this we need to define \( \mathbb{C}[G] \). Since we know what the functions on an affine variety are, we just need to realize \( G \) as an affine variety. Consider \( GL(W) \) as the affine subvariety of \( \mathbb{C}^{n^2+1} \), with coordinates \( (x_j, t) \) given by the equation \( t \det(x) = 1 \). Then \( \mathbb{C}[GL(W)] \) may be defined to be the restriction of polynomial functions on \( \mathbb{C}^{n^2+1} \) to the subvariety isomorphic to \( GL(W) \). If \( G \subset GL(W) \) is defined by algebraic equations, this also enables us to define \( \mathbb{C}[G] \) because \( G \subset GL(W) \) is a subvariety.

5.1.5. Characters.
Exercise 5.1.5.1: Let \( \rho : G \rightarrow GL(V) \) be a representation. Define a function \( \chi_\rho : G \rightarrow \mathbb{C} \) by \( \chi_\rho(g) = \text{trace}(\rho(g)) \). The function \( \chi_\rho \) is called the character of \( \rho \). Show that \( \chi_\rho \) is constant on conjugacy classes of \( G \). (In general, a function \( f : G \rightarrow \mathbb{C} \) such that \( f(hgh^{-1}) = f(g) \) for all \( g, h \in G \) is called a class function.) Show that for representations \( \rho_j : G \rightarrow GL(V_j) \), that \( \chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2} \).

Exercise 5.1.5.2: Given \( \rho_j : G \rightarrow GL(V_j) \) for \( j = 1, 2 \), define \( \rho_1 \otimes \rho_2 : G \rightarrow GL(V_1 \otimes V_2) \) by \( \rho_1 \otimes \rho_2(g)(v_1 \otimes v_2) = \rho_1(g)v_1 \otimes \rho_2(g)v_2 \). Show that \( \chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1} \chi_{\rho_2} \).

5.1.6. Application to representations of finite groups. Theorem 5.1.4.1 implies:

Theorem 5.1.6.1. Let \( G \) be a finite group, then as a \( G \times G \)-module under the action \( (L, R) \) and as an algebra,

\[
\mathbb{C}[G] := \bigoplus_i V_i^* \otimes V_i
\]

where the sum is over all the distinct irreducible representations of \( G \).

Exercise 5.1.6.2: Let \( G \) be a finite group and \( H \) a subgroup. Show that \( \mathbb{C}[G/H] = \oplus_i V_i^* \otimes (V_i)^H \) as a \( G \)-module under the action \( L \).

Theorem 5.1.6.1 has many generalizations. Its generalization to reductive algebraic groups was a major inspiration for the Geometric Complexity Theory program.

Remark 5.1.6.3. The classical Heisenberg uncertainty principle from physics, in the language of mathematics, is that it is not possible to localize both a function and its Fourier transform. A discrete analog of this uncertainty principle holds, in that the transforms of the delta functions have large support in terms of matrix coefficients and vice versa.

Theorem 5.1.6.1 is not yet useful, as we do not yet know what the \( V_i \) are. Let \( \mu_i : G \rightarrow GL(V_i) \) denote the representation. Recall from Exercise 5.1.5.1 the class function \( \chi_{\mu_i}(g) := \text{trace}(\mu_i(g)) \). It is not difficult to show that the functions \( \chi_{\mu_i} \) are linearly independent in \( \mathbb{C}[G] \). (One uses a \( G \)-invariant inner-product and shows that they are orthogonal with respect to the inner-product, see, e.g., [FH91, §2.2].) On the other hand, we have a natural basis of the class functions, namely the \( \delta \)-functions on each conjugacy class. Let \( C_j \) be a conjugacy class of \( G \) and define \( \delta_{C_j} := \sum_{g \in C_j} \delta_g \). It is straightforward to see, via the DFT, that the span of the \( \delta_{C_j} \)'s equals the span of the \( \chi_{\mu_i} \)'s, that is the number of distinct irreducible representations of \( G \) equals the number of conjugacy classes (see, e.g., [FH91, §2.2] for the
standard proof and [GW09, §4.4] for a DFT proof). Unfortunately, in another manifestation of the uncertainty principle, the relation between these two bases can be complicated. However, an apparent miracle occurs when $G = \mathfrak{S}_d$: we will associate irreducible representations directly to conjugacy classes.

A partition $\pi = (p_1, \ldots, p_r)$ of $d$ is a set of integers $p_1 \geq p_2 \geq \cdots \geq p_r$, $p_i \in \mathbb{N}$, such that $p_1 + \cdots + p_r = d$. Write $|\pi| = d$ and $\ell(\pi) = r$.

The conjugacy class of a permutation is determined by its decomposition into a product of disjoint cycles. The conjugacy classes of $\mathfrak{S}_d$ are in 1-1 correspondence with the set of partitions of $d$: to a partition $\pi = (p_1, \ldots, p_r)$ one associates the conjugacy class of an element with disjoint cycles of lengths $p_1, \ldots, p_r$. Let $[\pi]$ denote the isomorphism class of the irreducible $\mathfrak{S}_d$-module associated to $\pi$.

Thus in the case of $\mathfrak{S}_d$:

\begin{equation}
\mathbb{C}[\mathfrak{S}_d] = \bigoplus_{|\pi| = d} [\pi]_L \otimes [\pi]_R
\end{equation}

where I have included the $L, R$ as subscripts to remind the reader that different copies of $\mathfrak{S}_d$ are acting on the different factors.

We can say even more, as $\mathfrak{S}_d$ modules, $[\pi]$ is isomorphic to $[\pi]^*$. This is usually proved by first noting that for any finite group $G$, and any irreducible representation $\mu$, $\chi_{\mu^*} = \overline{\chi_\mu}$ where the overline denotes complex conjugate and then observing that the characters of $\mathfrak{S}_d$ are all real-valued functions. Thus we may rewrite (5.1.2) as

\begin{equation}
\mathbb{C}[\mathfrak{S}_d] = \bigoplus_{|\pi| = d} [\pi]_L \otimes [\pi]_R.
\end{equation}

5.1.7. The algebraic Peter-Weyl theorem. All reductive algebraic groups are affine varieties, so one can make sense of $\mathbb{C}[G]$ . Theorem 5.1.6.1 is still valid for reductive algebraic groups with the same proof, except that one has an infinite sum.

**Theorem 5.1.7.1.** Let $G$ be a reductive algebraic group. Then there are only countably many non-isomorphic irreducible $G$-modules. Let $\Lambda_G^+$ denote a set indexing the irreducible $G$-modules, and let $V_\lambda$ denote the irreducible module associated to $\lambda$. Then, as a $G \times G$-module

\[ \mathbb{C}[G] = \bigoplus_{\lambda \in \Lambda_G^+} V_\lambda \otimes V_\lambda^*. \]
5.2. Representations of $\mathfrak{S}_d$ and $GL(V)$

In this section we finally obtain our goal of the decomposition of $V^\otimes d$ as a $GL(V)$-module. Representations of $GL(V)$ occurring in $V^\otimes$ may also be indexed by partitions, which is explained in §5.2.1 where Schur-Weyl duality is stated and proved. I then implement the theory in §5.2.2.

5.2.1. Schur-Weyl duality. We have already seen that the actions of $GL(V)$ and $\mathfrak{S}_d$ on $V^\otimes d$ commute.

**Proposition 5.2.1.1.** $\text{End}_{GL(V)}(V^\otimes d) = \mathbb{C}[\mathfrak{S}_d]$.

**Proof.** We will show that $\text{End} \mathbb{C}[\mathfrak{S}_d]$ is the algebra generated by $GL(V)$ and conclude by the double commutant theorem. Since $\text{End}(V^\otimes d) = V^\otimes d \otimes (V^\otimes d)^*$ under the re-ordering isomorphism, $\text{End}(V^\otimes d)$ is spanned by elements of the form $X_1 \otimes \cdots \otimes X_d$ with $X_j \in \text{End}(V)$, i.e., elements of $\check{\text{Seg}}(\mathbb{P}(\text{End}(V)) \times \cdots \times \mathbb{P}(\text{End}(V)))$. The action of $X_1 \otimes \cdots \otimes X_d$ on $v_1 \otimes \cdots \otimes v_d$ induced from the $GL(V)$-action is $v_1 \otimes \cdots \otimes v_d \mapsto (X_1 v_1) \otimes \cdots \otimes (X_d v_d)$. Since $g \in GL(V)$ acts by $g \cdot (v_1 \otimes \cdots \otimes v_d) = gv_1 \otimes \cdots \otimes gv_d$, the image of $GL(V)$ lies in $S^d(V \otimes V^*)$, in fact it is a Zariski open subset of $\check{\mathbb{V}}_d(\mathbb{P}(V \otimes V^*))$ which spans $S^d(V \otimes V^*)$. In other words, the algebra generated by $GL(V)$ is $S^d(V \otimes V^*) \subset \text{End}(V^\otimes d)$. But by definition $S^d(V \otimes V^*) = (V \otimes V^*)^d$ and we conclude. □

Applying Theorem 5.1.3.1 we obtain:

**Theorem 5.2.1.2.** [Schur-Weyl duality] The irreducible decomposition of $V^\otimes d$ as a $GL(V) \times \mathbb{C}[\mathfrak{S}_d]$-module (equivalently, as a $GL(V) \times \mathfrak{S}_d$-module) is

$$V^\otimes d = \bigoplus_{|\pi|=d} S_\pi V \otimes [\pi],$$

where $S_\pi V := \text{Hom}_{\mathfrak{S}_d}([\pi], V^\otimes d)$ is an irreducible $GL(V)$-module.

Note that as far as we know, $S_\pi V$ could be zero. (It will be zero whenever $\ell(\pi) \geq \dim V$.)

5.2.2. Explicit realizations of representations of $\mathfrak{S}_d$ and $GL(V)$. By Theorem 5.1.6.1 we may explicitly realize each irreducible $\mathfrak{S}_d$-module via some projection from $\mathbb{C}[\mathfrak{S}_d]$. The question is, which projections?

To visualize $\pi$, define a Young diagram associated to a partition $\pi$ to be a collection of left-aligned boxes with $p_j$ boxes in the $j$-th row, as in Figure 5.2.1.
5.2. Representations of $\mathfrak{S}_d$ and $GL(V)$

Define the conjugate partition $\pi'$ to $\pi$ to be the partition whose Young diagram is the reflection of the Young diagram of $\pi$ in the north-west to south-east diagonal.

Figure 5.2.1. Young diagram for $\pi = (4, 2, 1)$

Figure 5.2.2. Young diagram for $\pi' = (3, 2, 1, 1)$, the conjugate partition to $\pi = (4, 2, 1)$.

Given $\pi$ we would like to find elements $c_\pi \in \mathbb{C}[\mathfrak{S}_d]$ such that $\mathbb{C}[\mathfrak{S}_d]c_\pi$ is isomorphic to $[\pi]$. I write $\overline{\pi}$ instead of just $\pi$ because the elements are far from unique; there is a vector space of dimension $\dim[\pi]$ of such projection operators by Theorem 5.1.6.1, and the overline signifies a specific realization. In other words, the $\mathfrak{S}_d$-module map $RM_{c_\pi} : \mathbb{C}[\mathfrak{S}_d] \to \mathbb{C}[\mathfrak{S}_d]$, $f \mapsto fc_\pi$ should kill all $\mathfrak{S}_d^R$-modules not isomorphic to $[\pi]_R$, and the image should be $[\pi]_L \otimes v$ for some $v \in [\pi]_R$. If this works, as a bonus, the map $c_\pi : V^\otimes d \to V^\otimes d$ induced from the $\mathfrak{S}_d$-action will have image a copy of $S_\pi V$ for the same reason.

Here are projection operators for the two representations we understand well:

When $\pi = (d)$, there is a unique up to scale $c_{(d)}$ and it is easy to see it must be $c_{(d)} := \sum_{\sigma \in \mathfrak{S}_d} \delta_\sigma$, as the image of $RM_{c_{(d)}}$ is clearly the line through $c_{(d)}$ on which $\mathfrak{S}_d$ acts trivially. Note further that $c_{(d)}(V^\otimes d) = S^d V$ as desired.

When $\pi = (1^d)$, again we have a unique up to scale projection, and its clear we should take $c_{(1^d)} = \sum_{\sigma \in \mathfrak{S}_d} \sgn(\sigma) \delta_\sigma$ as the image of any $\delta_\tau$ will be $\sgn(\tau)c_{(1^d)}$, and $c_{(1^d)}(V^\otimes d) = \Lambda^d V$.

The only other representation we have a reasonable understanding of is the standard representation $\pi = (d-1, 1)$ which corresponds to the complement of the trivial representation in the permutation action on $\mathbb{C}^d$. A basis of this space could be given by $e_1 - e_d, e_2 - e_d, \ldots, e_{d-1} - e_d$. Note that
the roles of $1, \ldots, d-1$ in this basis are the “same” in that if one permutes them, one gets the same basis, and that the role of $d$ with respect to any of the other $e_j$ is “skew” in some sense. To capture this behavior, consider

$$c_{(d-1,1)} := (\delta_{Id} - \delta_{(1,d)}) \left( \sum_{\sigma \in S_{d-1}[d-1]} \delta_{\sigma} \right)$$

where $S_{d-1}[d-1] \subset S_d$ is the subgroup permuting the elements $\{1, \ldots, d-1\}$. Note that $c_{(d-1,1)} \delta_{\tau} = c_{(d-1,1)}$ for any $\tau \in S_{d-1}[d-1]$ so the image is of dimension at most $d = \dim(\mathbb{C}[S_d]/\mathbb{C}[S_{d-1}])$.

**Exercise 5.2.2.1:** Show that the image is in fact $d-1$ dimensional.

Now consider $RM_{(d-1,1)}(V^\otimes d)$: after re-orderings, it is the image of the composition of the maps

$$V^\otimes d \to V^\otimes d-2 \otimes \Lambda^2 V \to S^{d-1}V \otimes V.$$

In particular, in the case $d = 3$, it is the image of

$$V \otimes \Lambda^2 V \to S^2 V \otimes V,$$

which we saw was isomorphic to $S_{21} V$ in §4.2.

Here is the general recipe to construct an $S_d$-module isomorphic to $[\pi]$:

fill the Young diagram of a partition $\pi$ of $d$ with integers $1, \ldots, d$ from top to bottom and left to right. For example let $\pi = (4, 2, 1)$ and write:

```
1 4 6 7
2 5
3
```

(5.2.2)

Define $S_{\pi} \simeq S_{q_1} \times \cdots \times S_{q_{p_1}} \subset S_d$ to be the subgroup that preserves the subsets of elements in the columns and $S_{\pi}$ is the subgroup of $S_d$ permuting the elements in the rows.

Explicitly, writing $\pi = (p_1, \ldots, p_{q_1})$ and $\pi' = (q_1, \ldots, q_{p_1})$, $S_{q_1}$ permutes the elements of $\{1, \ldots, q_1\}$, $S_{q_2}$ permutes the elements of $\{q_1 + 1, \ldots, q_1 + q_2\}$ etc. Similarly, $S_{\pi} \simeq S_{p_1} \times \cdots \times S_{p_{q_1}} \subset S_d$ is the subgroup where $S_{p_1}$ permutes the elements $\{1, q_1 + 1, q_1 + q_2 +1, \ldots, q_1 + \cdots + q_{p_1-1} +1\}$, $S_{p_2}$ permutes the elements $\{2, q_1 + 2, q_1 + q_2 +2, \ldots, q_1 + \cdots + q_{p_1-1} +2\}$ etc.

Define two elements of $\mathbb{C}[S_d]$:

$$s_{\pi} := \sum_{\sigma \in S_{\pi}} \delta_{\sigma}$$

and

$$a_{\pi} := \sum_{\sigma \in S_{\pi}} \text{sgn}(\sigma) \delta_{\sigma}.$$ 

Then $[\pi]$ is defined to be the isomorphism class of the $S_d$-module $\mathbb{C}[S_d]s_{\pi}a_{\pi}$. (It is also the isomorphism class of $\mathbb{C}[S_d]s_{\pi}\delta_{\pi}$, although these two realizations are generally distinct.)

**Exercise 5.2.2.2:** Show that $[\pi'] = [\pi] \otimes [1^d]$ as $S_d$-modules. ⊗
The space $V^\otimes d c_{\pi}$ will be a copy of the module $S_{\pi} V$ because $c_{\pi}$ kills all modules not isomorphic to $\pi$ and maps $[\pi]$ to a one dimensional vector space. The action on $V^\otimes d$ is first to map it to $\Lambda p_1 V \otimes \cdots \otimes \Lambda p_n V$, and then the module $S_{\pi} V$ is realized as the image of the map from this space to $S^{p_1} V \otimes \cdots \otimes S^{p_n} V$. So despite their original indirect definition, we may realize the modules $S_{\pi} V$ explicitly simply be skew-symmetrizations and symmetrizations.

Example 5.2.2.3. Consider $c_{(2,2)}$, associated to

$$(5.2.3) \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

which realizes a copy of $S_{(2,2)} V \subset V^\otimes 4$. It first maps $V^\otimes 4$ to $\Lambda^2 V \otimes \Lambda^2 V$ and then maps that to $S^2 V \otimes S^2 V$. Explicitly, the maps are

$$a \otimes b \otimes c \otimes c \mapsto (a \otimes b - b \otimes a) \otimes (c \otimes d - d \otimes c) = a \otimes b \otimes c \otimes d - a \otimes b \otimes d \otimes c - b \otimes a \otimes c \otimes d + b \otimes a \otimes d \otimes c$$

$$\mapsto (a \otimes b \otimes c \otimes d + c \otimes b \otimes a \otimes d + a \otimes d \otimes c \otimes b + c \otimes d \otimes a \otimes b)$$

$$- (a \otimes b \otimes d \otimes c + d \otimes b \otimes a \otimes c + a \otimes c \otimes d \otimes b + d \otimes c \otimes a \otimes b)$$

$$- (b \otimes a \otimes c \otimes d + c \otimes a \otimes b \otimes d + b \otimes d \otimes c \otimes a + c \otimes d \otimes b \otimes a)$$

$$+ (b \otimes a \otimes d \otimes c + d \otimes a \otimes b \otimes c + b \otimes c \otimes d \otimes a + d \otimes c \otimes b \otimes a)$$

5.2.3. Weights. Fix a basis $e_1, \ldots, e_V$ of $V$, let $T \subset GL(V)$ denote the subgroup of diagonal matrices, let $B \subset GL(V)$ be the subgroup of upper triangular matrices, and let $N \subset B$ be the upper triangular matrices with 1’s along the diagonal. (The notation is chosen because $T$ is usually called a torus, $B$ a Borel subgroup and the Lie algebra $\mathfrak{n}$ of $N$ consists of nilpotent matrices.) Call $z \in V^\otimes d$ a weight vector if $T[z] = [z]$. If

$$\begin{pmatrix} x_1 \\ \vdots \\ x_V \end{pmatrix}$$

we say $z$ has weight $\pi = (p_1, \ldots, p_V) \in \mathbb{Z}^V$.

Call $z$ a highest weight vector if $B[z] = [z]$, i.e., if $N z = z$. If $M$ is an irreducible $GL(V)$-module and $z \in M$ is a highest weight vector, call the weight of $z$ the highest weight of $M$.

If $G = GL(A_1) \times \cdots \times GL(A_n)$ and $V = V_1 \otimes \cdots \otimes V_n$ with $V_j$ a module for $GL(A_j)$, we define the maximal torus in $G$ to be the product of the maximal tori in the $GL(A_j)$, and similarly for the Borel. A weight is then defined to be an $n$-tuple of weights etc...

Exercise 5.2.3.1: Show that if $z \in V^\otimes d$ is a highest weight vector of weight $\pi$, then $\pi$ is a partition of $d$. 
Because of the relation with weights, it will often be convenient to pad a partition with a string of zeros to make it a string of \( v \) integers.

**Exercise 5.2.3.2:** Find highest weight vectors in \( V, S^2V, \Lambda^2V, S^3V, \Lambda^3V \) and the two realizations of \( S_{21}V \) given by the kernels of the symmetrization and skew-symmetrization maps \( V \otimes S^2V \to S^3V \) and \( V \otimes \Lambda^2V \to \Lambda^3V \).

**Exercise 5.2.3.3:** More generally, find a highest weight vector for the kernel of the symmetrization map \( V \otimes S^{d-1}V \to S^dV \) and of the kernel of the “exterior derivative” (or “Koszul”) map

\[
S^kV \otimes \Lambda^tV \to S^{k-1}V \otimes \Lambda^{t+1}V
\]

\[
x_1 \cdots x_k \otimes y_1 \wedge \cdots \wedge y_t \mapsto \sum_{j=1}^k x_1 \cdots \hat{x}_j \cdots x_k \otimes x_j \wedge y_1 \wedge \cdots \wedge y_t.
\]

A necessary condition for two irreducible \( GL(V) \)-modules to be isomorphic is that they have the same highest weight (because they must also be isomorphic \( T \)-modules). The condition is also sufficient, see Exercise 5.2.5.(2).

**5.2.4. Dimension formulas.** Given a Young diagram, and a square \( x \) in the diagram, define \( h(x) \), the hook length of \( x \) to be the number of boxes to the right of \( x \) in the same row, plus the number of boxes below \( x \) in the same column, plus one. Then for \( \pi \) a partition of \( d \),

\[
\dim[\pi] = \frac{d!}{\prod_{x \in \pi} h(x)}.
\]

For a box \( x \) occurring in a Young diagram, define the content of \( x \), denoted \( c(x) \), to be 0 if \( x \) is on the diagonal, \( s \) if \( x \) is \( s \) steps above the diagonal, and \( -s \) if \( x \) is \( s \) steps below the diagonal. Then

\[
\dim S_\pi C^n = \prod_{x \in \pi} \frac{n + c(x)}{h(x)}.
\]

See any of, e.g., [FH91, Pro07, Mac95] for proofs of these formulas.

For example \( \dim[(2,1)] = 2 = \frac{3!}{3+1+1} \) because the hook lengths are \( \begin{bmatrix} 3 & 1 \\ 1 \end{bmatrix} \) and \( \dim S_{21}V = \frac{n(n+1)(n-1)}{3+1+1} \) because the contents are \( \begin{bmatrix} 0 & 1 \\ -3 \end{bmatrix} \).

**5.2.5. Exercises:** Explicit submodules of \( V \otimes d \) isomorphic to \( S_\pi V \) and \( [\pi] \).
5.2. Representations of $\mathfrak{S}_d$ and $GL(V)$

(1) Let $\pi = (p_1, \ldots, p_\ell)$ be a partition with at most $v$ parts and let $\pi' = (q_1, \ldots, q_{p_1})$ denote the conjugate partition. Show that

\[ z_{\pi} := (e_1 \wedge \cdots \wedge e_{q_1}) \otimes (e_1 \wedge \cdots \wedge e_{q_2}) \otimes \cdots \otimes (e_1 \wedge \cdots \wedge e_{q_{p_1}}) \]

is a highest weight vector of weight $\pi$.

(2) Show that $z_{\pi} = [(e_1 \otimes \cdots \otimes e_{q_1}) \otimes (e_1 \otimes \cdots \otimes e_{q_2}) \otimes \cdots \otimes (e_1 \otimes \cdots \otimes e_{q_{p_1}})]s_{\pi}a_{\pi}$ and thus $\mathbb{C}[\mathfrak{S}_d]z_{\pi}$ is isomorphic to $[\pi]$ and the highest weight of $S_{\pi}V$ is (as the notation suggests) $\pi$. This also shows that the irreducible $GL(V)$-modules in $V^\otimes d$ are determined up to equivalence by their highest weight.

(3) Show that for any non-negative integers $(p_1, \ldots, p_\nu)$ with $p_1 + \cdots + p_\nu = d$, the set of vectors of weight $(p_1, \ldots, p_\nu)$ in $V^\otimes d$ is a $\mathfrak{S}_d$-submodule.

(4) Show that if $z$ is a highest weight vector, then so is $\sigma \cdot z$ for all $\sigma \in \mathfrak{S}_d$.

(5) Show that $S_{\pi}V \otimes [\pi]$ is the span of the $GL(V) \times \mathfrak{S}_d$-orbit of $z_{\pi}$.

(6) Verify directly that $S_{(d)}V, S_{(1d)}V$, and $S_{(d-1,1)}V$ respectively have multiplicities $1, 1, d-1$ in $V^\otimes d$ by computing the $\mathfrak{S}_d$-orbits of the corresponding $z_{\pi}$'s. ⊗

(7) Let $t \subset \text{End}(V)$ denote the diagonal matrices and $b$ the upper triangular matrices. Write $b = t \oplus n$ where $n$ is the space of strictly upper-triangular matrices. Show that each of these spaces are Lie subalgebras of $\text{gl}(V)$, i.e., they are closed under the Lie bracket.

(8) Show that $z \in V^\otimes d$ is a highest weight vector if and only if $n \cdot z = 0$.

(9) Show that $S_{\pi+\mu}V \subset S_{\pi}V \otimes S_{\mu}V$. Here, if $\pi = (p_1, \ldots, p_\nu)$ and $\mu = (m_1, \ldots, m_\nu)$, then $\pi + \mu = (p_1 + m_1, \ldots, p_\nu + m_\nu)$. In particular $S^{a+b}V \subset S^aV \otimes S^bV$, and the map $S^aV \times S^bV \rightarrow S^{a+b}V$ is just multiplication of polynomials. ⊗

By the exercises above, to determine a copy of $S_{\pi}V$ in $V^\otimes d$, it is sufficient to obtain its highest weight vector. Two ways to obtain a basis of the space of highest weight vectors of the isotypic component of $S_{\pi}V$ in $V^\otimes d$ are as follows: First, exercise 5.2.5(5) provides an algorithm to obtain spanning set from which we can extract a basis. Second, define a standard Young tableau to be a Young diagram filled with integers $1, \ldots, d$ where entries are increasing both along the rows and down the columns. Then for each standard Young tableau, form the corresponding projection operator $c_\pi$, and the vectors $(e_1^{\otimes p_1} \otimes \cdots \otimes e_\nu^{\otimes p_\nu})c_\pi$ will be a basis of the space of highest weight vectors. See e.g. [FH91] for details. I will use the notation $S_{\pi}V$ for a specific copy of $S_{\pi}V$ realized in $V^\otimes d$. 
5.2.6. Polynomials on spaces of matrices and the Cauchy formula.

To study polynomials on spaces of matrices (respectively tensors), we need to decompose \( S^d(A \otimes B) \) (resp. \( S^d(A \otimes B \otimes C) \)) as \( GL(A) \times GL(B) \) (resp. \( GL(A) \times GL(B) \times GL(C) \)) module. To obtain the first decomposition, write
\[
S^d(A \otimes B) = (A \otimes B)^{\otimes d} \mathcal{S}_d
\]
\[
= \bigoplus_{\pi, \mu} ([\pi] \otimes S_\pi A \otimes [\mu] \otimes S_\mu B)^{\otimes d}
\]
\[
= \bigoplus_{\pi, \mu} ([\pi] \otimes [\mu])^{\otimes d} \otimes S_\pi A \otimes S_\mu B.
\]

Since representations of \( \mathcal{S}_d \) are self-dual, \( ([\pi] \otimes [\mu])^{\otimes d} = \text{Hom}_{\mathcal{S}_d}([\pi], [\mu]) \). By Schur’s lemma we conclude:

**Proposition 5.2.6.1.** As a \( GL(A) \times GL(B) \)-module,
\[
S^d(A \otimes B) = \bigoplus_{\ell(\pi) \leq \min\{a, b\}} S_{\ell(\pi)} A \otimes S_{\pi} B.
\]

Formula (5.2.8) is called the *Cauchy formula*.

**Exercise 5.2.6.2:** Show that
\[
I_d(\text{Seg}(\mathbb{P} A \times \mathbb{P} B)) = \bigoplus_{\ell(\pi) \leq \min\{a, b\}} S_{\pi} A \otimes S_{\pi} B.
\]

What is \( I_d(\sigma_r(\text{Seg}(\mathbb{P} A \times \mathbb{P} B))) \)?

**Exercise 5.2.6.3:** Show that as a \( GL(A) \times GL(B) \)-module,
\[
\Lambda^d(A \otimes B) = \bigoplus_{\ell(\pi) \leq a; \ell(\mu) \leq b} S_{\pi} A \otimes S_{\mu} B.
\]

**Exercise 5.2.6.4:** Show that
\[
\text{mult}(S_{\pi_1} A_1 \otimes \cdots \otimes S_{\pi_n} A_n, S^d(A_1 \otimes \cdots \otimes A_n)) = \text{mult}([d], [\pi_1] \otimes \cdots \otimes [\pi_n]).
\]

**Exercise 5.2.6.5:** Show that the polynomials in \( S^p(A \otimes B \otimes C) \) arising from the usual flattenings \( \Lambda^p(A \otimes B) \otimes \Lambda^p C \), \( \Lambda^p(A \otimes C) \otimes \Lambda^p B \), \( \Lambda^p(B \otimes C) \otimes \Lambda^p A \) are such that when \( p = 2 \), any third is in the span of the other two, but any two are not completely redundant, and that for \( p = 3 \) all three are needed to span the space of equations.

Recall that \( SL(V) \subset GL(V) \) is the subgroup that acts trivially on \( \Lambda^V \), i.e., the maps \( V \to V \) with determinant equal to one.
Example 5.2.6.6. Let $A, B = \mathbb{C}^n$. By the Cauchy formula, $S^n(A \otimes B)$ contains a unique vector up to scale that is invariant under $SL(A) \times SL(B)$. This vector is $\det_n \in \Lambda^n A \otimes \Lambda^n B$. To see this explicitly, a highest weight vector of $\Lambda^n A \otimes \Lambda^n B$ is

$$(e_1 \wedge \cdots \wedge e_n) \otimes (f_1 \wedge \cdots \wedge f_n) = \frac{1}{(n!)^2} \left( \sum_{\sigma \in S_n} \text{sgn}(\sigma) e_{\sigma(1)} \otimes \cdots \otimes e_{\sigma(n)} \otimes \left( \sum_{\tau \in S_n} \text{sgn}(\tau) f_{\tau(1)} \otimes \cdots \otimes f_{\tau(n)} \right) \right).$$

Re-order to express this as an element of $(A \otimes B)^n$ and write $x_j^i = e_i \otimes f_j$ to obtain

$$\frac{1}{(n!)^2} \sum_{\sigma, \tau \in S_n} \text{sgn}(\sigma) \text{sgn}(\tau) x_{\tau(1)}^{\sigma(1)} \otimes \cdots \otimes x_{\tau(n)}^{\sigma(n)}.$$

The symmetrization map can be accomplished efficiently notationally by simply erasing the $\otimes$ symbols and multiplying by $n!$. Thus the image of this vector in $S^n(E \otimes F)$ under the projection given by symmetrization is

$$\frac{1}{n!} \sum_{\sigma \in S_n} \left( \sum_{\tau \in S_n} \text{sgn}(\sigma) \text{sgn}(\tau) x_{\tau(1)}^{\sigma(1)} \otimes \cdots \otimes x_{\tau(n)}^{\sigma(n)} \right).$$

We think of this as first summing over $\tau$, so for each fixed $\sigma$ in the sum over $\tau$, we can re-order the elements by $\sigma^{-1}$ to obtain

$$\frac{1}{n!} \sum_{\sigma \in S_n} \left( \sum_{\tau \in S_n} \text{sgn}(\sigma) \text{sgn}(\tau) x_{\tau(1)}^{\sigma^{-1}(1)} \cdots x_{\tau(n)}^{\sigma^{-1}(n)} \right) = \sum_{\sigma^{-1} \tau \in S_n} \text{sgn}(\sigma^{-1} \tau) x_{\tau(1)}^{\sigma^{-1}(1)} \cdots x_{\tau(n)}^{\sigma^{-1}(n)}$$

which is the familiar formula for the determinant if one renames $\sigma^{-1} \tau$ as $\sigma$.

Fixing a torus $T$, we may talk about permutation matrices. Let $\mathcal{W}_A \subset GL(A)$ denote the subgroup of permutation matrices. (The notation is chosen because $\mathcal{W}_A$ is the Weyl group of $GL(A)$.) The normalizer of $T$ in $GL(A)$ is $T \rtimes \mathcal{W}_A$, where $\mathcal{W}_A$ acts by conjugation.

Example 5.2.6.7. Let $A, B = \mathbb{C}^n$. Then $S^n(A \otimes B)$ contains a unique vector up to scale invariant under $(T_A \times \mathcal{W}_A) \times (T_B \times \mathcal{W}_B)$. This vector is $\text{perm}_n$. This vector lives in the unique weight $(1, \ldots, 1)_A \times (1, \ldots, 1)_B$ line inside $S^n A \otimes S^n B$. To recognize this as the permanent, write the weight vector in $S^n A \otimes S^n B$ as $(e_1 \cdots e_n) \otimes (f_1 \cdots f_n)$ re-order as above and write $x_j^i = e_i \otimes f_j$ to get $\sum_{\sigma, \tau \in S_n} x_{\tau(1)}^{\sigma(1)} \otimes \cdots \otimes x_{\tau(n)}^{\sigma(n)}$ and continue the calculation as before to obtain the familiar formula for the permanent when one projects to $S^n(E \otimes F)$.

Remark 5.2.6.8. From Example 5.2.6.7, one can begin to appreciate the beauty of the permanent. Since $\det_n$ is the only polynomial invariant under $SL(E) \times SL(F)$, to find other interesting polynomials on spaces of matrices, one has to be content with subgroups of this group. But what could be a more natural subgroup than the product of the normalizer of the tori?
In fact, say we begin by asking simply for a polynomial invariant under the action of $T_E \times T_F$. We need to look at $S^n(E \otimes F)_0$, where the 0 denotes the sl-weight zero subspace. This decomposes as $\bigoplus \pi (S_\pi E)_0 \otimes (S_\pi F)_0$. We will see (Corollary ??(i)) that these spaces are the $\mathfrak{S}_n^E \times \mathfrak{S}_n^F$-modules $[\pi] \otimes [\pi]$. Only one of these is trivial, namely $[n] \otimes [n] \subset S^n E \otimes S^n F$ and that corresponds to the permanent! More generally, if we consider the diagonal $\mathfrak{S}_n \subset \mathfrak{S}_n^E \times \mathfrak{S}_n^F$, then both $[\pi]$’s are modules for the same group, and since $|\pi| \simeq |\pi|^*$, there is then an element corresponding to the identity map. These vectors are Littlewood’s \textit{immanants}, of which the determinant and permanent are special cases.

5.2.7. The Pieri rule. Let’s decompose $S_\pi V \otimes V$ as a $GL(V)$-module. Write $\pi' = (q_1, \ldots, q_{p_1})$ and recall $z_\pi$ from (5.2.7). Consider the vectors:

\[
(e_1 \wedge \cdots \wedge e_{q_1} \wedge e_{q_1+1}) \otimes (e_1 \wedge \cdots \wedge e_{q_2}) \otimes \cdots \otimes (e_1 \wedge \cdots \wedge e_{q_{p_1}}) \\
\vdots \\
(e_1 \wedge \cdots \wedge e_{q_1}) \otimes (e_1 \wedge \cdots \wedge e_{q_2}) \otimes \cdots \otimes (e_1 \wedge \cdots \wedge e_{q_{p_1}+1}) \\
(e_1 \wedge \cdots \wedge e_{q_1}) \otimes (e_1 \wedge \cdots \wedge e_{q_2}) \otimes \cdots \otimes (e_1 \wedge \cdots \wedge e_{q_{p_1}}) \otimes e_1.
\]

These are all highest weight vectors obtained from $z_\pi$ tensored with a vector in $V$ and skew-symmetrizing appropriately, so the associated modules are contained in $S_\pi V \otimes V$. With a little more work, one can show these are highest weight vectors of all the modules that occur in $S_\pi V \otimes V$. If $q_j = q_j + 1$ one gets the same module if one inserts $e_{q_1+1}$ into either slot, but its multiplicity in $S_\pi V \otimes V$ is one. More generally one obtains:

\textbf{Theorem 5.2.7.1} (The Pieri formula). The decomposition of $S_\pi V \otimes S_d V$ is multiplicity free. The partitions corresponding to modules $S_\mu V$ that occur are those obtained from the Young diagram of $\pi$ by adding $d$ boxes to the diagram of $\pi$, with no two boxes added to the same column.

\textbf{Definition 5.2.7.2.} Let $\pi, \mu$ be partitions with $\ell(\mu) < \ell(\pi)$ One says $\mu$ interlaces $\pi$ if $p_1 \geq m_1 \geq p_2 \geq m_2 \geq \cdots \geq m_{\ell(\pi)} \geq p_{\ell(\pi)}$.

\textbf{Exercise 5.2.7.3:} Show that $S_\pi V \otimes S_d V$ consists of all the $S_\mu V$ such that $|\mu| = |\pi| + d$ and $\mu$ interlaces $\pi$.

Although a pictorial proof is possible, the standard proof of the Pieri formula uses a character calculation, computing $\chi_\pi \chi_{(d)}$ as a sum of $\chi_\mu$’s. See, e.g., [Mac95, §I.9]. A different proof, using Schur-Weyl duality is in [GW09, §9.2]. There is an algorithm to compute arbitrary tensor product decompositions called the \textit{Littlewood Richardson Rule}. See, e.g., [Mac95, §I.9] for details.

Similar considerations give:
Theorem 5.2.7.4. [The skew-Pieri formula] The decomposition of $S_\pi V \otimes \Lambda^k V$ is multiplicity free. The partitions corresponding to modules $S_\mu V$ that occur are those obtained from the Young diagram of $\pi$ by adding $k$ boxes to the diagram of $\pi$, with no two boxes added to the same row.

5.2.8. The $GL(V)$-modules not appearing in $V^\otimes$. The $GL(V)$-module $V^*$ does not appear in the tensor algebra of $V$. Nor do the one-dimensional representations $\det^{-k} : GL(V) \to GL(\mathbb{C}^1)$ given by, for $v \in \mathbb{C}^1$, $\det^{-k}(g)v := \det(g)^{-k}v$.

Exercise 5.2.8.1: Show that if $\pi = (p_1, \ldots, p_v)$ with $p_v > 0$, then $\det^{-1} \otimes S_\pi V = S_{(p_1-1, \ldots, p_v-1)}V$.

Every irreducible $GL(V)$-module is of the form $S_\pi V \otimes \det^{-k}$ for some $k \geq 0$. Thus they may be indexed by non-increasing sequences of integers $(p_1, \ldots, p_v)$ where $p_1 \geq p_2 \geq \cdots \geq p_v$. Such a module is isomorphic to $S_{(p_1-p_v, \ldots, p_v-1)} V \otimes \det^{p_v}$.

Exercise 5.2.8.2: Show that as a $GL(V)$-module, $V^* = \Lambda^{v-1} V \otimes \det^{-1} = S_{1^{v-1}} V \otimes \det^{-1}$.

Using $S_\pi V \otimes V^* = S_\pi V \otimes \Lambda^{v-1} V \otimes \det^{-1}$, we may compute the decomposition of $S_\pi V \otimes V^*$ using the skew-symmetric version of the Pieri rule.

Example 5.2.8.3. Let $w = 3$, then

$$S_{(32)} W \otimes W^* = S_{(43)} W \otimes \det^{-1} \oplus S_{(331)} W \otimes \det^{-1} \oplus S_{(421)} W \otimes \det^{-1} = S_{(43)} W \otimes \det^{-1} \oplus S_{(22)} W \oplus S_{(31)} W$$

so the first module does not occur in the tensor algebra but the rest do.

5.2.9. $SL(V)$-modules in $V^d$. Fact: Every $SL(V)$-module is the restriction to $SL(V)$ of some $GL(V)$-module. However distinct $GL(V)$-modules, when restricted to $SL(V)$ can become isomorphic, such as the trivial representation and $\Lambda^v V$.

Proposition 5.2.9.1. Let $\pi = (p_1, \ldots, p_v)$ be a partition. The $SL(V)$-modules in the tensor algebra $V^\otimes$ that are isomorphic to $S_\pi V$ are $S_\mu V$ with $\mu = (p_1 + j, p_2 + j, \ldots, p_v + j)$ for $-p_v \leq j < \infty$.

Exercise 5.2.9.2: Prove Proposition 5.2.9.1. ⊙

For example, for $SL_2$-modules, $S_{p_1,p_2} \mathbb{C}^2 \simeq S^{p_1-p_2} \mathbb{C}^2$. We conclude:

Corollary 5.2.9.3. A complete set of the finite dimensional irreducible representations of $SL_2$ are the $S^d \mathbb{C}^2$ with $d \geq 0$. 

The \( GL(V) \)-modules that are \( SL(V) \)-equivalent to \( S_\pi V \) may be visualized as being obtained by erasing or adding columns of size \( v \) from the Young diagram of \( \pi \).

![Figure 5.2.3. Young diagrams for \( SL_3 \)-modules equivalent to \( S_{421} \)](image)

### 5.2.10. The kernel of \( (M_{(U,V,C_1)})^A_p \)

Consider the case \( u = v = 3 \), take \( p = 4 \), so we need to study

\[
M_{(U,V,C_1)}^{A^4} : \Lambda^4(U^* \otimes V) \otimes V \to \Lambda^5(U^* \otimes V) \otimes V^*.
\]

Since \( v = u = 3 \), we have \( \Lambda^4(U^* \otimes V) = S_{211}U^* \otimes S_{31}V \oplus S_{22}U^* \otimes S_{22}V \oplus S_{31}U^* \otimes S_{21}V \). The Pieri rule implies

\[
\Lambda^4(U^* \otimes V) \otimes V = S_{31}U^* \otimes S_{311}V \oplus S_{41}U^* \otimes S_{221}V \oplus S_{22}U^* \otimes S_{32}V \oplus S_{22}U^* \otimes S_{221}V \oplus S_{211}U^* \otimes S_{31}V \oplus S_{211}U^* \otimes S_{32}V \oplus S_{211}U^* \otimes S_{311}V.
\]

Pictorially,

\[
\begin{array}{ccc}
0 & \otimes & 0 = & 0 & \oplus & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & \otimes & 0 = & 0 & \oplus & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & \otimes & 0 = & 0 & \oplus & 0 \\
\end{array}
\]

The kernel must contain all modules that do not appear in

\[
\Lambda^5(U^* \otimes V) \otimes V^* = S_{31}U^* \otimes S_{221}V \oplus S_{22}U^* \otimes S_{221}V \otimes V
\]

\[
= S_{31}U^* \otimes S_{311}V \oplus S_{211}U^* \otimes S_{311}V \oplus S_{22}U^* \otimes S_{32}V \oplus S_{22}U^* \otimes S_{221}V \oplus S_{211}U^* \otimes S_{31}V
\]

\[
\oplus (S_{43}U^* \otimes \det U^{-1}) \otimes S_{221}V.
\]

So the module \( S_{41}U^* \otimes S_{221}V \) coming from the the third picture is in the kernel and \( (S_{43}U^* \otimes \det U^{-1}) \otimes S_{221}V \) is not in the image. The map is an isomorphism on the remaining factors.
For example, consider $S_{31}U^* \otimes S_{311}V \subset \Lambda^4(U^* \otimes V) \otimes V$. We may write a highest weight vector of $S_{31}U^* \otimes S_{311}V \subset U^* \otimes V \otimes V$ as

$$e^1 \wedge e^2 \otimes e^1 \otimes f_1 \wedge f_2 \otimes f_3 \otimes f_1 \otimes f_1 \in S_{31}U^* \otimes S_{311}V.$$ 

This embeds into $\Lambda^4(U^* \otimes V) \otimes V$ as follows:

$$(e^1 \otimes f_1) \wedge (e^1 \otimes f_1) \wedge (e^2 \otimes f_1) \wedge (e^1 \otimes f_2) \wedge (e^1 \otimes f_3) \wedge f_3 \otimes f_1 \otimes f_1.$$ 

Observe that this expression is still skew in the first $e^1, e^2$ as well as the first $f^1, f^2, f^3$. When we map this to $\Lambda^8(U^* \otimes V) \otimes U$ we get

$$\sum_{j=1}^3 (e^1 \otimes f_1) \wedge (e^2 \otimes f_1) \wedge (e^1 \otimes f_2) \wedge (e^1 \otimes f_3) \wedge (w^j \otimes f_1) \otimes u_j$$

$$= (e^1 \otimes f_1) \wedge (e^2 \otimes f_1) \wedge (e^1 \otimes f_2) \wedge (e^1 \otimes f_3) \wedge (e^3 \otimes f_1) \otimes e_3$$

$$\neq 0.$$

Thus the rank of $M_{\Lambda^4}^{\Lambda^4}$ is $(3 \cdot \binom{9}{4} - 72)l = 306l$ and $R(M_{(3,3,1)}) \geq \lceil \frac{306l}{76} \rceil$ which is 14 when $l = 3$.

More generally:

**Proposition 5.2.10.1.** [LO] $\ker(M(u,v,b))_A^p \supset \bigoplus_\pi S_{\pi^*} U^* \otimes S_{\pi^+(1)} V \otimes \mathbb{C}^1$ where the summation is over partitions $\pi = (u, v_1, \ldots, v_{-1})$ where $\nu = (v_1, \ldots, v_{-1})$ is a partition of $p - u$, $v_1 \leq u$ and $\pi + (1) = (u + 1, v_1, \ldots, v_{-1})$.

**Exercise 5.2.10.2:** Show that the kernel contains the asserted modules.

Thus one expects that if one restricts to a sufficiently generic subspace $A'$ of $A^*$ of dimension $2u + 1$, the reduced map $\Lambda^{u-1}A \otimes V \to \Lambda^u A \otimes U^*$, will be injective. This is what was done in the proof of Theorem 2.5.8.1.

5.3. Methods for determining equations of secant varieties of Segre varieties using representation theory

***This section will be expanded****

5.3.1. Inheritance. Inheritance is a general technique for studying equations of $G$-varieties that come in series. It is discussed extensively in [Lan12, §7.4,16.4], so I will just give an overview here.

If $V \subset W$ then $S_\pi W$ induces a module $S_\pi V$ by, e.g., choosing a basis of $W$ whose first $v$ vectors are a basis of $V$. Then the two modules have the same highest weight vector and one obtains the $GL(W)$-module by the action of $GL(W)$ on the highest weight vector.

Because the realizations of $S_\pi V$ in $V^{\otimes d}$ do not depend on the dimension of $V$, one can reduce the study of $\sigma_\pi(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ to that of
There is an inclusion of modules $5.3.3. \text{Young flattenings.}$

Proposition 5.3.1.1. [LM04, Prop. 4.4] For all vector spaces $B_j$ with $\dim B_j = b_j \geq \dim A_j = a_j \geq r$, a module $S_{A_1} \otimes \cdots \otimes S_{A_n} A_n$ such that $\ell(\mu_j) \leq a_j$ for all $j$, is in $I_d(\sigma_r(\text{Seg}(\mathbb{P}B^*_1 \times \cdots \times \mathbb{P}B^*_n)))$ if and only if $S_{A_1} \otimes \cdots \otimes S_{A_n} A_n$ is in $I_d(\sigma_r(\text{Seg}(\mathbb{P}A_1^* \times \cdots \times \mathbb{P}A_n^*)))$.

See [Lan12, §7.4] for the proof.

Thus if $\dim A_j = r$ and $\ell(\mu_j) \leq r$,

$$S_{A_1} \otimes \cdots \otimes S_{A_n} A_n \in I(\sigma_r(\text{Seg}(\mathbb{P}A_1^* \times \cdots \times \mathbb{P}A_n^*)))$$

if and only if

$$S_{A_1} \otimes \cdots \otimes S_{A_n} C^{\ell(\mu_1)} \cdots C^{\ell(\mu_n)} \in I(\sigma_r(\text{Seg}(\mathbb{P}A_1^* \times \cdots \times \mathbb{P}A_n^*))).$$

In summary:

Corollary 5.3.1.2. [LM04, AR03] Let $\dim A_j \geq r, 1 \leq j \leq n$. The ideal of $\sigma_r(\text{Seg}(\mathbb{P}A_1 \times \cdots \times \mathbb{P}A_n^*))$ is generated by the modules inherited from the ideal of $\sigma_r(\text{Seg}(\mathbb{P}A_1^* \times \cdots \times \mathbb{P}A_n^-))$ and the modules generating the ideal of $\text{Sub}_{r-1,n-r}$. The analogous scheme and set-theoretic results hold as well.

5.3.2. Brute force Strassen. to type in

5.3.3. Young flattenings. There is an inclusion of modules $A \otimes B \otimes C \subset (S_{A_1} A \otimes S_{A_1} B \otimes S_{A_1} C)^* \otimes (S_{A_2} A \otimes S_{A_2} B \otimes S_{A_2} C)$ if and only if the Young diagram of $\pi_2$ is obtained from that of $\pi_1$ by adding a box, and similarly for the other two factors. In each such situation we obtain, given $T \in A \otimes B \otimes C$, a linear map:

$$T_{(\pi_1,\mu_1,\nu_1),(\pi_2,\mu_2,\nu_2)} : S_{A_1} A \otimes S_{A_1} B \otimes S_{A_1} C \to S_{A_2} A \otimes S_{A_2} B \otimes S_{A_2} C.$$

Say we are trying to find equations for $\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$. Let $[a \otimes b \otimes c] \in \text{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$. Then

$$(5.3.2) \quad \mathbb{R}(T) \geq \frac{\text{rank} T_{(\pi_1,\mu_1,\nu_1),(\pi_2,\mu_2,\nu_2)}}{\text{rank}(a \otimes b \otimes c)_{(\pi_1,\mu_1,\nu_1),(\pi_2,\mu_2,\nu_2)}}$$

The Koszul flattenings of §2.5.6 are the case $\pi_1 = (1^p), \pi_2 = (1^{p+1}), \nu_1 = \emptyset, \nu_2 = (1), \mu_1 = (-1), \mu_2 = \emptyset. (B^* \text{ does not occur in the tensor algebra of } B, \text{ a simple way to account for it is to represent it by a Young diagram with negative one boxes.})$

Young flattenings were originally used for symmetric border rank, i.e., to find equations for secant varieties of Veronese varieties in [LO13]. There the problem is to study inclusions $S^d V \subset S_\mu V \otimes S_\mu V$. I'll return to this in our study of geometric complexity theory.
5.3.4. Algorithms for determining the ideals of secant varieties of Segre varieties and proving lower bounds for $R(M(n))$. Write

\[(5.3.3)\quad S^d(A \otimes B \otimes C) = \bigoplus_{\pi, \mu, \nu} (S_\pi A \otimes S_\mu B \otimes S_\nu C) \oplus k_{\pi, \mu, \nu} .\]

The multiplicities $k_{\pi, \mu, \nu}$ are called Kronecker coefficients. By Exercise 5.2.6.4, they equal $\dim([\pi] \otimes [\mu] \otimes [\nu])^\mathfrak{S}_d$, so they can be calculated via characters of the symmetric group:

\[(5.3.4)\quad k_{\pi, \mu, \nu} = \frac{1}{d!} \sum_{\sigma \in \mathfrak{S}_d} \chi_\pi(\sigma) \chi_\mu(\sigma) \chi_\nu(\sigma)\]

where recall $\chi_\pi(\sigma) = \text{trace}(\rho_\pi(\sigma))$ where $\rho_\pi : \mathfrak{S}_d \to GL([\pi])$ is the representation $[\pi]$ (see, e.g., [Lan12, §6.6.2]).

**Exercise 5.3.4.1:** Show that $k_{\pi, \mu, \nu} \dim \text{Hom}_{\mathfrak{S}_d}([\pi], [\mu] \otimes [\nu])$.

Using the decomposition, one has systematic ways to determine equations for $\sigma_r$:

**Naïve algorithm using representation theory.** For each $d$, decompose $S^d(A \otimes B \otimes C)$. For each isotypic component, $S_\pi A \otimes S_\mu B \otimes S_\nu C \otimes ([\pi] \otimes [\mu] \otimes [\nu])^\mathfrak{S}_d \subset S^d(A \otimes B \otimes C)$ write down a basis of highest weight vectors, $v_1, \ldots, v_{k_{\pi, \mu, \nu}}$. Then pick $k_{\pi, \mu, \nu}$ “random” points on $\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$ and test $c_1v_1 + \cdots + c_{k_{\pi, \mu, \nu}}v_{k_{\pi, \mu, \nu}}$ on these points, where the $c_j$ are arbitrary. If for some choice of $c_j$ one gets zero, then with high probability the corresponding module is in $I_d(\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$.

If $a = b = c = n^2$, on can then test this module on $M(n)$. Say $P \in S_\pi A \otimes S_\mu B \otimes S_\nu C$ is such that $P(M(n)) \neq 0$, then .....}

5.3.5. Enhanced search using numerical methods. While we don’t have equations for $\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$, we do have a simple parametric description of a large open subset of it via the addition map $A^{\otimes r} \times B^{\otimes r} \times C^{\otimes r} \to A \otimes B \otimes C$, $(a_1, \ldots, a_r, b_1, \ldots, b_r, c_1, \ldots, c_r) \mapsto \sum_{j=1}^r a_j \otimes b_j \otimes c_j$. One picks a “random” point of $\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$, then takes a general linear space of complementary dimension intersecting it. The intersection will be a finite collection of points. The parametric description enables one to find other points of $\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$. Say we find $q$ of them. One can then systematically search for equations vanishing on these points. Once one finds the equations say of degree $\delta$ vanishing on it, one looks for additional points on $\sigma_r(Seg(\mathbb{P}A^* \times \mathbb{P}B^* \times \mathbb{P}C^*))$ to see if the polynomial also vanishes on it. If not, add the new point and check if any polynomials vanish on it. If not, go to degree $\delta + 1$, if so, add another point. One continues in this fashion, and there are methods to tell
one when to stop adding points and when one has found the ideal of the
point set. If everything has been done properly one now knows the dimen-
sion of the space of generators in the ideal of \( \sigma_r(\text{Seg}(\mathbb{P}^A \times \mathbb{P}^B \times \mathbb{P}^C^*) \) in
each degree. Now we use that the ideal of \( \sigma_r(\text{Seg}(\mathbb{P}^A \times \mathbb{P}^B \times \mathbb{P}^C^*) \) is a
\( G = GL(A) \times GL(B) \times GL(C) \)-module, we decompose \( S^d V \) as a \( G \)-module.
If one is lucky, there are only a few possible irreducible modules whose di-
mensions add to the correct dimension. One is then reduced to just checking
these candidates. These methods were used to prove the following theorem:

**Theorem 5.3.5.1. [HIL13]** With extremely high probability, the ideal of
\( \sigma_6(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) \) is generated in degree 19 by the module \( S_{5554} \mathbb{C} \otimes S_{5554} \mathbb{C} \otimes S_{5554} \mathbb{C} \). This module does not vanish on \( M_{(2)} \).

In the same paper, the trivial degree twenty module \( S_{5555} \mathbb{C} \otimes S_{5555} \mathbb{C} \otimes S_{5555} \mathbb{C} \)
is shown to be in the ideal of \( \sigma_6(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) \) by symbolic methods,
giving a proof that \( R(M_{(2)}) = 7 \). This was originally shown in [Lan06] by
a study of the components of \( \sigma_6(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) \setminus \sigma_6(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) \).

The same methods have shown \( I_{45}(\sigma_{15}(\text{Seg}(\mathbb{P}^3 \times \mathbb{P}^7 \times \mathbb{P}^0)) = 0 \) and that
\( I_{186,999}(\sigma_{18}(\text{Seg}(\mathbb{P}^6 \times \mathbb{P}^9 \times \mathbb{P}^0)) = 0 \) (this variety is a hypersurface), both of
which are relevant for determining the border rank of \( M_{(3)} \), see [HIL13].

5.3.6. Geometric Complexity Theory approach to equations. By
inheritance discussed below, in determining the ideal of \( \sigma_r(\text{Seg}(\mathbb{P}^A \times \mathbb{P}^B \times \mathbb{P}^C)) \) we may restrict attention to \( \sigma_r(\text{Seg}(\mathbb{P}^{r-1} \times \mathbb{P}^{r-1} \times \mathbb{P}^{r-1})) \) which has the
advantage that it is an orbit closure, the orbit closure of \( M_{(3)}^{(r)}. \) In \( 6.2 \) we will
see how to use the Peter-Weyl theorem to in principle obtain a description
of the module structure of the coordinate ring of the orbit. From that, as
explained in \( 6.2 \), if one understands the module structure of the space of
polynomials, one can sometimes find modules that must be in the ideal of
\( \sigma_r(\text{Seg}(\mathbb{P}^A \times \mathbb{P}^B \times \mathbb{P}^C)). \)
This chapter is devoted to the flagship conjecture of Geometric Complexity Theory 1.2.5.1 that $\text{perm}_m$ cannot be realized as a polynomial sized determinant, even approximately. It begins in §6.1 with a discussion of circuits and brief history and context. Next, §6.2 discusses the approach to the flagship conjecture proposed in [MS01, MS08]. This approach requires one to search for $GL(W)$-modules with certain properties. In §6.3 I describe two ways to narrow the search. The preprint [Mula] is discussed in §6.4: this relates blackbox derandomization with the search for an explicit large linear space that fails to intersect the orbit closure of the determinant. The preprint [Mulb] is the subject of §6.5, where a case is made for the existence of representation-theoretic obstructions as defined in §6.2.2. The best lower bound for the flagship conjecture was obtained via classical algebraic geometry - the study of hypersurfaces with degenerate dual varieties. An exposition of this result is presented in §6.6. The discussion in §6.2 leads to questions regarding Kronecker and plethysm coefficients. A summary of our knowledge of these coefficients relevant for the flagship conjecture is presented in §6.7. To better understand the flagship conjecture it will be useful to make similar studies of other polynomials. Several such are discussed in §6.8.

Unlike matrix multiplication, this is a subject in its infancy. There are almost no mathematical results, so the exposition will be focused not so much on the accomplishments, but on mathematics that I expect to be useful in the future for the subject.
6.1. Circuits and the flagship conjecture of GCT

To be precise about the conjectural differences between the complexity of the permanent and the determinant, one needs a model of computation. I will use the model of arithmetic circuits.


**Definition 6.1.1.1.** An arithmetic circuit $C$ is a finite, acyclic, directed graph with vertices of in-degree 0 or 2 and exactly one vertex of out-degree 0. The vertices of in-degree 0 are labeled by elements of $\mathbb{C} \cup \{x_1, \ldots, x_n\}$, and called inputs. Those of in-degree 2 are labeled with $+$ or $\ast$ and are called gates. If the out-degree of $v$ is 0, then $v$ is called an output gate. The size of $C$ is the number of edges.

![Circuit for $(x + y)^3$](image)

**Figure 6.1.1.** Circuit for $(x + y)^3$

**Exercise 6.1.1.2:** Show that if one instead uses the number of gates to define the size, the asymptotic size estimates are the same. (Size is sometimes defined as the number of gates.)

To each vertex $v$ of a circuit $C$, associate the polynomial that is computed at $v$, which will be denoted $C_v$. In particular the polynomial associated with the output gate is called the polynomial computed by $C$.

At first glance, circuits do not look geometrical, as they depend on a choice of coordinates. While computer scientists always view polynomials as being given in some coordinate expression, in geometric complexity theory one erases the choice of coordinates by allowing the group of all changes of bases to act. The reason circuits will still be reasonable is that one is not concerned with the precise size of a circuit, but its size up to, e.g., a polynomial factor. At a small price of an additional layer of computation,
one can think of the inputs to our circuits as arbitrary affine linear or linear
functions on a vector space. We will soon homogenize in which case the
inputs will become linear functions on a vector space.

6.1.2. Arithmetic circuits and complexity classes. Let \( \mathbb{C}[x_1, \ldots, x_N] \)
denote the space of all polynomials in \( x_1, \ldots, x_N \) and let \( \mathbb{C}[x_1, \ldots, x_N]_{\leq d} \)
denote the space of polynomials of degree at most \( d \).

**Definition 6.1.2.1.** Let \((f_n)\) be a sequence of polynomials, \( f_n \in \mathbb{C}[x_1, \ldots, x_{N(n)}]_{\leq d(n)} \). We say \((f_n) \in \text{VP}\) if \( \deg(f_n) \) and \( N(n) \) are bounded by a polynomial in \( n \) and there exists a sequence of circuits \( C_n \) of size polynomial in \( n \) computing \( f_n \).

Often the phrase “there exists a sequence of circuits \( C_n \) of size polynomial in \( n \) computing \( f_n \)” is abbreviated “there exists a polynomial sized circuit computing \((f_n)\)”.

The class \( \text{VNP} \), which consists of sequences of polynomials whose coef-
ficients are “easily” described, has a more complicated definition:

**Definition 6.1.2.2.** A sequence \((f_n)\) is defined to be in \( \text{VNP} \) if there exists a polynomial \( p \) and a sequence \((g_n) \in \text{VP} \) such that

\[
f_n(x) = \sum_{\epsilon \in \{0,1\}^{p(|x|)}} g_n(x, \epsilon).
\]

One may think of the class \( \text{VP} \) as a bundle over \( \text{VNP} \) where elements of \( \text{VP} \) are sequences of maps, say \( g_n : \mathbb{C}^{N(n)} \to \mathbb{C} \), and elements of \( \text{VNP} \) are projections of these sequences obtained by eliminating some of the variables by averaging or “integration over the fiber”. In algebraic geometry, it is well known that projections of varieties can be far more complicated than the original varieties. See [Bas14] for more on this perspective.

**Conjecture 6.1.2.3.** [Valiant [Val79b]] \( \text{VP} \neq \text{VNP} \).

**Definition 6.1.2.4.** One says that a sequence \((g_m(y_1, \ldots, y_{M(m)}))\) can be reduced to \((f_n(x_1, \ldots, x_{N(n)}))\) if there exists a polynomial \( n(m) \) and affine linear functions \( X_1(y_1, \ldots, y_M), \ldots, X_N(y_1, \ldots, y_M) \) such that \( g_m(y_1, \ldots, y_{M(m)}) = f_n(X_1(y), \ldots, X_N(y)) \). A sequence \((p_n)\) is hard for a complexity class \( C \) if every \((f_m) \in C \) can be reduced to \((p_n)\), and it is complete for \( C \) if furthermore \((p_n) \in C \).

**Theorem 6.1.2.5.** [Valiant [Val79b]] \( \text{(perm}_m) \text{ is complete for VNP}.\)

See the original [Val79b] or [BCS97, §21.4] for the proof.

Thus Conjecture 6.1.2.3 is equivalent to:

**Conjecture 6.1.2.6.** [Valiant][Val79b] There does not exist a polynomial size circuit computing the permanent.
If one accepts that counting the number of perfect matchings of a bipartite graph is a reasonable “difficult” problem, then by the discussion in §1.2.1, one can ignore all definitions regarding circuits and still grasp the essence of Valiant’s conjectures.

Now for the determinant:

**Proposition 6.1.2.7.** \((\det_n) \in \text{VP}\).

One can compute the determinant quickly via Gaussian elimination: one uses the group to put a matrix in a form where the determinant is almost effortless to compute (the determinant of an upper triangular matrix is just the product of its diagonal entries). However this algorithm as presented is not a circuit (there are divisions and one needs to check if pivots are zero). Here is a construction of a small circuit for the determinant:

**Proof of Proposition 6.1.2.7.** The determinant of a linear map is the product of its eigenvalues, \(e_n(\lambda) = \lambda_1 \cdots \lambda_n\). This is an elementary symmetric function, fitting into the family

\[
e_k(\lambda) := \sum_{J \subseteq [n] \mid |J|=k} \lambda_{j_1} \cdots \lambda_{j_k}.
\]

This family has a generating function

\[
E_n(t) := \sum_{k \geq 0} e_k(\lambda) t^k = \prod_{i=1}^n (1 + \lambda_i t).
\]

On the other hand, recall that \(\text{trace}(f)\) is the sum of the eigenvalues of \(f\), and more generally, letting \(f^k\) denote the composition of \(f\) with itself \(k\) times,

\[
\text{trace}(f^k) = p_k(\lambda) := \lambda_1^k + \cdots + \lambda_n^k.
\]

The quantities \(\text{trace}(f^k)\) can be computed with small circuits. The power sum symmetric functions have the generating function

\[
P_n(t) = \sum_{k \geq 1} p_k t^k = \frac{d}{dt} \log \Pi_{j=1}^n (1 - \lambda_j t)^{-1}.
\]

In what follows, I suppress \(n\) from the notation.

**Exercise 6.1.2.8:** Show that

\[
P(-t) = \frac{E'(t)}{E(t)}.
\]
Exercise 6.1.2.8, together with a little more work (see, e.g. [Mac95, p. 28]) shows that

\[ e_n(\lambda) = \frac{1}{n!} \det \begin{pmatrix} p_1 & 1 & 0 & \cdots & 0 \\ p_2 & p_1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n-1} & p_{n-2} & \cdots & n-1 \\ p_n & p_{n-1} & \cdots & p_1 \end{pmatrix}. \]

While this is still a determinant, it is almost lower triangular and its naïve computation, e.g., with Laplace expansion, can be done with an \( O(n^3) \)-size circuit and the full algorithm for computing \( \det_n \) can be executed with an \( O(n^4) \) size circuit. \( \square \)

**Remark 6.1.2.9.** A more restrictive class of circuits are formulas which are circuits that are trees. The circuit in the proof above is not a formula because results from computations are used more than once. It is known that the determinant admits a quasi-polynomial size formula, that is, a formula of size \( n^{O((\log n)^c)} \) for some constant \( c \), and it is complete for the complexity class \( \text{VQP} = \text{VP}_s \) consisting of sequences of polynomials admitting a quasi-polynomial size formula (or equivalently, by ***, a polynomial sized “weakly skew” circuit). See, e.g., [BCS97, §21.5]. It is not known whether or not the determinant is complete for \( \text{VP} \).

**6.1.3. Permanent v. determinant.** Despite the above remark, since the determinant is so beautiful, researchers often work with it. Valiant made the following possibly weaker conjecture:

**Conjecture 6.1.3.1 (Valiant).** [Val79b] Set \( n = mc \), for any \( c > 0 \). Then for all sufficiently large \( m \),

\[ e^{n-m} \text{perm}_m \not\in \text{End}(\mathbb{C}^{n^2}) \cdot \det_n. \]

Since Conjecture 6.1.3.1 is expected to be difficult, define the following measure of progress towards proving it: let \( dc(\text{perm}_m) \) denote the smallest \( n \) such that \( e^{n-m} \text{perm}_m \in \text{End}(\mathbb{C}^{n^2}) \cdot \det_n \). Lower bounds on \( dc(\text{perm}_m) \) give a benchmark for progress. So far we know \( dc(\text{perm}_m) \geq \frac{m^2}{2} \) from [MR04]. This bound was obtained with local differential geometry.

**this paragraph needs cleaning*** Conjecture 6.1.3.1 takes us to a question of algebra. It involves our favorite polynomial, the determinant, which recall from Example 5.2.6.6 is characterized by its symmetries because it is the unique instance of the trivial representation of \( SL(E) \times SL(F) \) in \( S^n(E \otimes F) \). Here \( E, F = \mathbb{C}^n \).
As we have seen in Example 5.2.6.7, the permanent is not so bad either: it is also characterized by its symmetries, as it is the unique instance of the trivial representation of $T^* \Sigma(E) \times V_F \times T^* \Sigma(F) \times V_F \in S^n(E \otimes F)$.

In order to use more tools from algebraic geometry and representation theory to separate complexity classes, the following conjecture appeared in [MS01]:

**Conjecture 6.1.3.2.** [MS01] Let $\ell$ be a linear coordinate on $\mathbb{C}^1$ and consider any linear inclusion $\mathbb{C}^1 \oplus \mathbb{C}^m \to W = \mathbb{C}^n$, so in particular $\ell^{n-m} \text{perm}_m \in S^n W$. Let $n(m)$ be a polynomial. Then for all sufficiently large $m$,

$$\ell^{n-m} \text{perm}_m \notin GL(W) \cdot [\det_n(m)].$$

Note that $GL(W) \cdot [\det_n] = \text{End}(W) \cdot [\det_n]$ so this is a strengthening of Conjecture 6.1.3.1. It will be useful to rephrase the conjecture slightly, to highlight that it is a question about determining whether one orbit closure is contained in another. Let

$$\mathcal{D}et_n := GL(W) \cdot [\det_n],$$

and let

$$\mathcal{P}erm^m_n := GL(W) \cdot [\ell^{n-m} \text{perm}_m].$$

**Conjecture 6.1.3.3.** [MS01] Let $n(m)$ be a polynomial. Then for all sufficiently large $m$,

$$\mathcal{P}erm^m_n(m) \nsubseteq \mathcal{D}et_n(m).$$

The equivalence follows as $\ell^{n-m} \text{perm}_m \nsubseteq \mathcal{D}et_n$ implies $GL(W) \cdot \ell^{n-m} \text{perm}_m \nsubseteq \mathcal{D}et_n$, and since $\mathcal{D}et_n$ is closed and both sides are irreducible, there is no harm in taking closure on the left hand side.

Now the goal is clear: both varieties are invariant under $GL(W)$ so their ideals will be $GL(W)$-modules. We look for a $GL(W)$-module $M$ such that $M \subset I[\mathcal{D}et_n]$ and $M \nsubseteq I[\mathcal{P}erm^m_n]$.

In the next section I explain a plan to use the algebraic Peter-Weyl theorem to find modules in $I[\mathcal{D}et_n]$.

### 6.2. A program to find modules in $I[\mathcal{D}et_n]$ via representation theory

In this section I present the program in [MS08] to find modules in the ideal of $\mathcal{D}et_n$.

**Remark 6.2.0.4.** In this section I deal with the $GL(W)$-orbit closure. In [MS08], the idea was to use the $SL(W)$-orbit closure of $\det_n$ in affine space. This has the advantage that the orbit is already closed, so one can apply the
6.2. A program to find modules in $I[\text{Det}_n]$ via representation theory

algebraic Peter-Weyl theorem directly, as discussed below. It has the disadvantage that one cannot distinguish isomorphic $SL(W)$-modules appearing in different degrees in the coordinate ring.

6.2.1. Preliminaries. Define $C[\hat{\text{Det}}_n] := \text{Sym}(S^nW^*)/I(\text{Det}_n)$, the homogeneous coordinate ring of the (cone over) $\text{Det}_n$. This is the space of polynomial functions on $\hat{\text{Det}}_n$ inherited from polynomials on the ambient space.

Since $I(\text{Det}_n) \subset \text{Sym}(S^nW^*)$ is a $GL(W)$-submodule, and since $GL(W)$ is reductive, we obtain the following splitting as a $GL(W)$-module:

$$\text{Sym}(S^nW^*) = I(\text{Det}_n) \oplus C[\hat{\text{Det}}_n].$$

In particular, if a module $S_\pi W$ appears in $\text{Sym}(S^nW^*)$ and it does not appear in $C[\hat{\text{Det}}_n]$, it must appear in $I(\text{Det}_n)$.

Define $C[GL(W) \cdot \text{det}_n] = C[GL(W)/G_{\text{det}_n}]$ to be $C[GL(W)]^{G_{\text{det}_n}}$. ***Need to add explanation as to why this ring is the ring of regular functions of the orbit - inclusion is clear *****. There is an injective map

$$C[\hat{\text{Det}}_n] \to C[GL(W) \cdot \text{det}_n]$$

given by restriction of functions. The map is an injection because any function identically zero on a Zariski open subset of an irreducible variety is identically zero on the variety. The algebraic Peter-Weyl theorem gives a description of the $G$-module structure of $C[G/H]$ when $G$ is a reductive algebraic group and $H$ is a subgroup.

**Plan:** Find a module $S_\pi W^*$ not appearing in $C[GL(W)/G_{\text{det}_n}]$ that does appear in $\text{Sym}(S^nW^*)$.

By the above discussion such a module must appear in $I(\text{Det}_n)$.

**Definition 6.2.1.1.** An irreducible $GL(W)$-module $S_\pi W^*$ appearing in $\text{Sym}(S^nW^*)$ and not appearing in $C[GL(W)/G_{\text{det}_n}]$ is called an orbit occurrence obstruction.

The precise condition a module must satisfy to be an orbit occurrence obstruction is explained in Proposition 6.2.5.2. Whether such a module will be useful or not for separating the determinant from the padded permanent will be taken up in §6.3.

We will see that orbit occurrence obstructions may be difficult to find, so several relaxations of the definition are given in the next subsection.

One might object that the coordinate rings of different orbits could coincide, or at least be very close. Indeed this is the case for generic polynomials, but in GCT one generally restricts to polynomials whose symmetry groups *characterize* the orbit as follows:
Definition 6.2.1.2. Let $V$ be a $G$-module. A point $P \in V$ is characterized by its stabilizer $G_P$ if any $Q \in V$ with $G_Q \supseteq G_P$ is of the form $Q = cP$ for some constant $c$.

We have seen in §5.2.6 that both the determinant and permanent polynomials are characterized by their stabilizers.

One can think of polynomial sequences that are complete for their complexity classes and are characterized by their stabilizers as “best” representatives of their class. We will see that if $P \in S^dV$ is characterized by its stabilizer, the coordinate ring of its $G$-orbit is unique as a module among orbits of points in $V$.

Recall that $Ch_n(\mathbb{C}^n)$ is the variety of polynomials that are products of linear forms and observe that $Ch_n(\mathbb{C}^n)$ is contained in $\text{Det}_n$ and $\text{Perm}_n^n$. The following conjecture, if true, would reduce the search space for occurrence obstructions:

**Conjecture 6.2.1.3 (Kumar [Kum]).** For all partitions $\pi$ with $\ell(\pi) \leq n$, the module $S_{n\pi} \mathbb{C}^n$ occurs in $\mathbb{C}[Ch_n(\mathbb{C}^n)]$. In particular, $S_{n\pi} \mathbb{C}^{n^2}$ occurs in $\mathbb{C}[\text{Det}_n]$ and $\mathbb{C}[\text{Perm}_n^n]$.

In §7.3.5 I discuss evidence for this conjecture and how it has interesting connections to a longstanding conjecture in combinatorics as well as the evaluation of certain integrals over the special unitary group.

### 6.2.2. Three relaxations of orbit occurrence obstructions.

First, if the multiplicity of $S_\pi W^*$ in $\mathbb{C}[GL(W)/G_{\text{det}_n}]$ is smaller than its multiplicity in $\text{Sym}(S^nW^*)$, then some copy of it must be in $I(\text{Det}_n)$.

**Definition 6.2.2.1.** An irreducible $GL(W)$-module $S_\pi W^*$ satisfying $\text{mult}(S_\pi W^*, \text{Sym}(S^nW^*)) > \text{mult}(S_\pi W^*, \mathbb{C}[GL(W)/G_{\text{det}_n}])$ is called an orbit representation-theoretic obstruction.

While orbit representation-theoretic obstructions may be easier to find than orbit occurrence obstructions, they have the disadvantage that to utilize them, one must either determine an explicit realization of the module in $I(\text{Det}_n)$ to test it on the padded permanent or have the multiplicity in $\mathbb{C}[\text{Perm}_m^n]$ be sufficiently large to guarantee at least one copy in $I(\text{Det}_n)$ is in $\mathbb{C}[\text{Perm}_m^n]$.

A different relaxation is that even if a module appears in $\mathbb{C}[GL(W)/G_{\text{det}_n}]$, it may not appear in $\mathbb{C}[\text{Det}_n]$.

**Definition 6.2.2.2.** An irreducible $GL(W)$-module $S_\pi W^*$ appearing in $\text{Sym}(S^nW^*)$ and not appearing in $\mathbb{C}[\text{Det}_n]$ is called an occurrence obstruction.
Occurrence obstructions have the disadvantage in that the extension problem of determining which modules in $\mathbb{C}[GL(W)/G_{\text{det}_n}]$ extend to be defined on all of $\mathbb{C}[\text{det}_n]$ is notoriously difficult, even in simple cases where of orbit closures where the boundary is well understood. See §8.1.1 for an example.

Finally one can combine the two relaxations to define a representation-theoretic obstruction, a module $S_\pi W^*$ satisfying $\text{mult}(S_\pi W^*, \text{Sym}(S^0 W^*)) > \text{mult}(S_\pi W^*, \mathbb{C}[\text{det}_n])$, which should be the easiest to find, but the hardest to utilize.

6.2.3. The coordinate ring of an orbit. Recall the algebraic Peter-Weyl Theorem from §5.1.7.

Corollary 6.2.3.1. Let $H \subset G$ be a closed subgroup. Then, as a $G$-module,

$$(6.2.1) \quad \mathbb{C}[G/H] = \mathbb{C}[G]^H = \bigoplus_{\lambda \in \Lambda_G^+} V_\lambda \otimes (V_\lambda^*)^H = \bigoplus_{\lambda \in \Lambda_G^+} V_\lambda \otimes \dim(V_\lambda^*)^H.$$

Here $G$ acts on the $V_\lambda$ and $(V_\lambda^*)^H$ is just a vector space whose dimension records the multiplicity of $V_\lambda$ in $\mathbb{C}[G/H]$.

Corollary 6.2.3.1 motivates the study of polynomials characterized by their stabilizers: if $P \in V$ is characterized by its stabilizer, then $G \cdot P$ is the unique orbit in $V$ with coordinate ring isomorphic to $\mathbb{C}[G \cdot P]$ as a $G$-module. Moreover, for any $Q \in V$ that is not a multiple of $P$, $\mathbb{C}[G \cdot Q] \not\subset \mathbb{C}[G \cdot P]$.

We next determine $G_{\text{det}_n} \subset GL(W)$ in order to get as much information as possible about the modules in $\mathbb{C}[GL(W)/G_{\text{det}_n}]$.

Remark 6.2.3.2. All $GL(W)$-modules $S_{(p_1, \ldots, p_w)} W$ may be graded using $p_1 + \cdots + p_w$ as the grading. One does not have such a grading for $SL(W)$-modules, which makes their use in GCT more difficult, as mentioned in Remark 6.2.0.4.

6.2.4. The stabilizer of the determinant. I follow [Die49] in this section. Write $\mathbb{C}^{n^2} = W = E \otimes F = \text{Hom}(E^*, F)$ with $E, F = \mathbb{C}^n$, and $\rho : GL_{n^2} \to GL(S^n \mathbb{C}^{n^2})$ for the representation.

Theorem 6.2.4.1 (Frobenius [Fro97]). Let $\phi \in GL_{n^2}$ be such that $\rho(\phi)(\text{det}_n) = \text{det}_n$. Then, identifying $\mathbb{C}^{n^2}$ with the space of $n \times n$ matrices,

$$\phi(z) = \begin{cases} gzh, & \text{or} \\ gz^T h \end{cases}$$

for some $g, h \in GL_n$, with $\text{det}_n(g) \text{det}_n(h) = 1$. Here $z^T$ denotes the transpose of $z$. 
Let $\mu_{n,r}$ denote the kernel of the product map $(\mathbb{Z}_n)^{\times r} \to \mathbb{Z}_n$ and think of $\mathbb{Z}_n \subset GL_n$ as the $n$-th roots of unity times the identity matrix. Write $\mu_n = \mu_{n,2}$.

**Corollary 6.2.4.2.** $G_{\text{det}_n} = (SL(E) \times SL(F))/\mu_n \times \mathbb{Z}_2$

To prove the Corollary, just note that the $\mathbb{C}^*$ corresponding to $\text{det}(g)$ above and $\mu_n$ are the kernel of the map $\mathbb{C}^* \times SL(E) \times SL(F) \to GL(E \otimes F)$.

**Exercise 6.2.4.3:** Prove the $n = 2$ case of the theorem. ⊗

**Lemma 6.2.4.4.** Let $U \subset W$ be a linear subspace such that $U \subset \{\text{det}_n = 0\}$. Then $\dim U \leq n^2 - n$. The subvariety of the Grassmannian $G(n^2 - n,W)$ consisting of maximal linear spaces on $\{\text{det}_n = 0\}$ has two irreducible components, call them $\Sigma_\alpha$ and $\Sigma_\beta$, where

\[(6.2.2) \quad \Sigma_\alpha = \{X \in G(n^2 - n,W) \mid \ker(X) = \hat{L} \text{ for some } L \in \mathbb{P}E\}, \text{ and} \]
\[(6.2.3) \quad \Sigma_\beta = \{X \in G(n^2 - n,W) \mid \text{Image}(X) = \hat{H} \text{ for some } H \in \mathbb{P}F^*\}.\]

Here $\ker(X)$ means the intersections of the kernels of all $f \in X$ and $\text{Image}(X)$ is the span of all the images.

Moreover, for any two distinct $X_j \in \Sigma_\alpha$, $j = 1,2$, and $Y_j \in \Sigma_\beta$ we have

\[(6.2.4) \quad \dim((X_1 \cap X_2)) = \dim((Y_1 \cap Y_2)) = n^2 - 2n, \text{ and} \]
\[(6.2.5) \quad \dim((X_1 \cap Y_j)) = n^2 - 2n + 1.\]

**Exercise 6.2.4.5:** Prove Lemma 6.2.4.4.

**Proof of theorem 6.2.4.1.** **ask Fulvio why confused*** Let $\Sigma = \Sigma_\alpha \cup \Sigma_\beta$. Then the map on $G(n^2 - n,W)$ induced by $\phi$ must preserve $\Sigma$. By the conditions (6.2.4),(6.2.5) of Lemma 6.2.4.4, in order to preserve dimensions of intersections, either every $X \in \Sigma_\alpha$ must map to a point of $\Sigma_\alpha$, in which case every $Y \in \Sigma_\beta$ must map to a point of $\Sigma_\beta$, or, every $X \in \Sigma_\alpha$ must map to a point of $\Sigma_\beta$, and every $Y \in \Sigma_\beta$ must map to a point of $\Sigma_\alpha$. If we are in the second case, replace $\phi$ by $\phi \circ T$, where $T(z) = z^T$, so we may now assume $\phi$ preserves both $\Sigma_\alpha$ and $\Sigma_\beta$.

Now $\Sigma_\alpha \simeq \mathbb{P}E$, so $\phi$ induces an algebraic map $\phi_E : \mathbb{P}E \to \mathbb{P}E$. If $L_1, L_2 \in \mathbb{P}E$ lie on a $\mathbb{P}^1$, in order for $\phi$ to preserve $\dim(X_{L_1} \cap X_{L_2})$, the images of the $L_j$ under $\phi_E$ must also lie on a $\mathbb{P}^1$, and thus $\phi_E$ must take lines to lines (and similarly hyperplanes to hyperplanes). But then, (see, e.g., [Har95, §18, p. 229]) $\phi_E \in PGL(E)$, and similarly, $\phi_F \in PGL(F)$, where $\phi_F : \mathbb{P}F^* \to \mathbb{P}F^*$ is the corresponding map. Here $PGL(E)$ denotes $GL(E)/\mathbb{C}^*$, the image of $GL(E)$ in its action on projective space. Write $\hat{\phi}_E \in GL(E)$ for any choice of lift and similarly for $F$.

Consider the map $\hat{\phi} \in \rho(GL(W))$ given by $\hat{\phi}(X) = \hat{\phi}_E^{-1}(\phi(X)) \hat{\phi}_F^{-1}$. The map $\hat{\phi}$ sends each $X_j \in \Sigma_\alpha$ to itself as well as each $Y_j \in \Sigma_\beta$, in particular
it does the same for all intersections. Hence it preserves \( \text{Seg}(\mathbb{P}E \times \mathbb{P}F) \subset \mathbb{P}(E \otimes F) \) point-wise, so it is up to scale the identity map. \( \square \)

**Remark 6.2.4.6.** For those familiar with Picard groups, M. Brion points out that there is a shorter proof of Theorem 6.2.4.1 as follows: In general, if a polynomial \( P \) is reduced and irreducible, then \( G_{\mathbb{Z}(P)} = G_{\mathbb{P}(P)} \) (This follows as the dual variety \( Z(P)^\vee \), see \S 6.6, satisfies \( (Z(P)^\vee)^\vee = Z(P) \)). The dual of \( Z(\mathrm{det}_n) \) is the Segre \( \text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}) \). Now the automorphism group of \( \mathbb{P}^{n-1} \times \mathbb{P}^{n-1} = \mathbb{P}E \times \mathbb{P}F \) acts on the Picard group which is \( \mathbb{Z} \times \mathbb{Z} \) and preserves the two generators \( \mathbb{O}_{\mathbb{P}E \times \mathbb{P}F}(1,0) \) and \( \mathbb{O}_{\mathbb{P}E \times \mathbb{P}F}(0,1) \) coming from the generators on \( \mathbb{P}E, \mathbb{P}F \). Thus, possibly composing with \( \mathbb{Z}_2 \) swapping the generators (corresponding to transpose in the ambient space), we may assume each generator is preserved. But then we must have an element of \( \text{Aut}(\mathbb{P}E) \times \text{Aut}(\mathbb{P}F) = \text{PGL}(E) \times \text{PGL}(F) \). Passing back to the ambient space, we obtain the result.

### 6.2.5. The coordinate ring of \( \text{GL}(W) \cdot \det_n \).**

We first compute the \( \text{SL}(E) \times \text{SL}(F) \)-invariants in \( S_\pi W \) where \( |\pi| = d \). Recall from \S 5.2.1 that by definition, \( S_\pi W = \text{Hom}_{\mathcal{S}_d}([\pi], W \otimes^d) \). Thus

\[
S_\pi W = \text{Hom}_{\mathcal{S}_d}([\pi], E \otimes^d F \otimes^d) = \text{Hom}_{\mathcal{S}_d}([\pi], (\bigoplus_{|\mu|=d} [\mu] \otimes S_\mu E \otimes (\bigoplus_{|\nu|=d} [\nu] \otimes S_\nu F))
\]

\[
= \bigoplus_{|\mu|=|\nu|=d} \text{Hom}_{\mathcal{S}_d}([\pi], [\mu] \otimes [\nu]) \otimes S_\mu E \otimes S_\nu F
\]

The vector space \( \text{Hom}_{\mathcal{S}_d}([\pi], [\mu] \otimes [\nu]) \) simply records the multiplicity of \( S_\mu E \otimes S_\nu F \) in \( S_\pi (E \otimes F) \). Recall from \S 5.3.4 that the numbers \( k_{\pi,\mu,\nu} = \dim \text{Hom}_{\mathcal{S}_d}([\pi], [\mu] \otimes [\nu]) \) are called **Kronecker coefficients**.

Recall from \S 5.2.9 that \( S_\mu E \) is a trivial \( \text{SL}(E) \) module if and only if \( \mu = (\delta^n) \) for some \( \delta \in \mathbb{Z} \). Thus so far, we are reduced to studying the Kronecker coefficients \( k_{\pi,\delta^n,\delta^n} \). Now take the \( \mathbb{Z}_2 \) action given by exchanging \( E \) and \( F \) into account. Write \( [\mu] \otimes [\mu] = S^2[\mu] \otimes \Lambda^2[\mu] \). The first module will be invariant under \( \mathbb{Z}_2 = \mathcal{S}_2 \), and the second will transform its sign under the transposition. So define the **symmetric Kronecker coefficients**

\[
ksk^{\pi}_{\mu,\mu} := \dim(\text{Hom}_{\mathcal{S}_d}([\pi], S^2[\mu])).
\]

We conclude:

**Proposition 6.2.5.1. [BLMW11]** Let \( W = \mathbb{C}^n \). The coordinate ring of the \( \text{GL}(W) \)-orbit of \( \det_n \) is

\[
\mathbb{C}[\text{GL}(W) \cdot \det_n] = \bigoplus_{d \in \mathbb{Z}} \bigoplus_{|\pi|=n\delta} (S_\pi W^*) \otimes \ksk^{\pi}_{\mu,\mu}.
\]

To phrase in the language of obstructions:
Proposition 6.2.5.2. Partitions $\pi$ of $dn$ such that $sk^\pi_{d^n,d^n} < \text{mult}(S_\pi W, S^d(S^n W))$ are orbit representation-theoretic obstructions, and if moreover $sk^\pi_{d^n,d^n} = 0$, $S_\pi W$ is an orbit occurrence obstruction.

I discuss the computation of Kronecker and symmetric Kronecker coefficients in §6.7.

C. Ikenmeyer [Ike12] has examined the situation for $\text{Det}_3$. He found on the order of $3,000$ orbit representation-theoretic obstructions, of which on the order of $100$ are orbit occurrence obstructions in degrees up to $d = 15$. There are two such partitions with seven parts, $(13^2, 2^5)$ and $(15^2, 5^6)$. The rest consist of partitions with at least $8$ parts (and many with $9$). Also of interest is that for approximately $2/3$ of the partitions $sk^\pi_{5^3,5^3} < k_{\pi 5^3 5^3}$. The lowest degree of an occurrence obstruction is $d = 10$, where $\pi = (9^2, 2^6)$ has $sk_{10^3 10^3}^\pi = k_{10^3 10^3} = 0$ but $\text{mult}(S_\pi W, S^{10}(S^3 W)) = 1$. In degree $11$, $\pi = (11^2, 2^5, 1)$ is an occurrence obstruction where $\text{mult}(S_\pi W, S^{11}(S^3 W)) = k_{11^3 11^3} = 1 > 0 = sk_{11^3 11^3}^\pi$.

In §6.3, I describe restrictions on which modules one should search for, taking into account just that one is comparing the determinant with a polynomial that is padded and only involves few variables.

6.3. Necessary conditions for modules of polynomials to be useful for GCT

The polynomial $\ell^{n-m} \text{perm}_m \in S^n \mathbb{C}^{n^2}$ has two properties that can be studied individually: it is padded, that is divisible by a large power of a linear form, and its zero set is a cone with a $n^2 - m^2 - 1$ dimensional vertex, that is, it only uses $m^2 + 1$ of the $n^2$ variables in an expression in good coordinates. I begin with the study of cones, as this is a classical topic.

6.3.1. Cones.

Let $\text{Sub}_k(S^d V) \subset \mathbb{P}S^d V$ be the variety of polynomials whose zero sets in projective space are cones with a $v - k$ dimensional vertex, called the subspace variety. If $[P] \in \text{Sub}_k(S^d V)$, then there exist linear coordinates on $V$ such that the expression of $P$ only involves at most $k$ of the coordinates.
6.3. Necessary conditions for modules of polynomials to be useful for $G\mathbb{C}^2$

Equations for $\text{Sub}_k(S^dV)$ are easy to describe: its ideal is generated by the size $k+1$ minors of

$$P_{1,d-1} : V^* \to S^{d-1}V.$$  

**Proposition 6.3.1.1.** $I_\delta(\text{Sub}_k(S^dV))$ consists of the isotypic components of the modules $S_\pi V^*$ appearing in $S^d(S^dV^*)$ such that $\ell(\pi) > k$.

**Exercise 6.3.1.2:** Prove Proposition 6.3.1.1.

With just a little more effort, one can prove our first equations generate the ideal:

**Theorem 6.3.1.3.** [Wey03, Cor. 7.2.3] The ideal of $\text{Sub}_k(S^dV)$ is generated by the image of $\Lambda^{k+1}V^* \otimes \Lambda^{k+1}S^{d-1}V^* \subset S^{k+1}(V^* \otimes S^{d-1}V^*)$ in $S^{k+1}(S^dV^*)$.

**Aside 6.3.1.4.** Here is further information about the variety $\text{Sub}_k(S^dV)$: First, it is a good example of a variety admitting a Kempf-Weyman desingularization, a type of desingularization that $G$-varieties often admit. Rather than discuss the general theory here (see [Wey03] for a full exposition or [Lan12, Chap. 17] for an elementary introduction), I just explain this example, which gives a proof of Theorem 6.3.1.3, although more elementary proofs are possible. Consider the Grassmannian $G(k, V)$ it has a tautological vector bundle $\pi : S \to G(k, V)$, where the fiber over a $k$-plane $E$ is just the $k$-plane itself, so the whole bundle is a sub-bundle of the trivial bundle $V$ with fiber $V$. Consider the bundle $S^dS \subset S^dV$. We have a projection map $p : S^dV \to S^dV$. The image of $S^dS$ under $p$ is $\hat{\text{Sub}}_k(S^dV)$. Moreover, the map is a desingularization, that is $S^dS$ is smooth, and the map to $\hat{\text{Sub}}_k(S^dV)$ is generically one to one. In particular, this implies $\dim \hat{\text{Sub}}_k(S^dV) = \dim (S^dS) = \binom{k+d-1}{d} + d(v-k)$. From these calculations, one obtains the entire minimal free resolution.

6.3.2. The variety of padded polynomials. Define the variety of padded polynomials

$$\text{Pad}_{n-m}(S^nW) := \mathbb{P}\{P \in S^nW \mid P = \ell^{n-m}h, \text{for some } \ell \in W, h \in S^mW^*\} \subset \mathbb{P}S^nW.$$  

**Theorem 6.3.2.1.** [KL14] Let $\pi = (p_1, \ldots, p_w)$ be a partition of $dn$.

1. If $p_1 < d(n-m)$, then the isotypic component of $S_\pi W$ in $S^d(S^nW)$ is contained in $I_d(\text{Pad}_{n-m}(S^nW^*))$.

2. If $p_1 \geq \min\{d(n-1), dn-m\}$, then $I_d(\text{Pad}_{n-m}(S^nW^*))$ does not contain a copy of $S_\pi W$.

**Proof.** An analog of inheritance for secant varieties of Segre varieties (see §5.3.1) applies in the current situation: if $k < w$ and we have an inclusion $\mathbb{C}^k \subset W$, then the modules $S_\pi \mathbb{C}^k$ each induce a unique module $S_\pi W$ by
6. Geometric Complexity Theory

considering the highest weight vector in \((\mathbb{C}^k)^{\otimes |\pi|}\) as a highest weight vector in \(W^{\otimes |\pi|}\).

Since all partitions in \(S^d(S^nW)\) have length at most \(d\), we may assume by Proposition 6.3.2.4 that \(w = d\). Fix a (weight) basis \(e_1, \ldots, e_d\) of \(W\) with dual basis \(x^1, \ldots, x^d\) of \(W^*\). Note any element \(e^{n-m}h \in \text{Pad}_{n-m}(S^nW)\) is in the \(GL(W)\)-orbit of \((e_1)^{n-m}h\) for some \(h\), so it will be sufficient to show that the ideal in degree \(d\) contains the modules vanishing on the orbits of elements of the form \((e_1)^{n-m}h\). The highest weight vector of any copy of \(S_{(p_1, \ldots, p_d)}W\) in \(S^d(S^nW)\) will be a linear combination of vectors of the form \(m_I := ((x^1)^{i_1} \cdots (x^d)^{i_d}) \cdots ((x^1)^{i_1} \cdots (x^d)^{i_d})\), where \(i_1 + \cdots + i_d = p_j\) and \(i_1^k + \cdots + i_d^k = n\) for all \(k\). Each \(m_I\) vanishes on any \((e_1)^{n-m}h\) unless \(p_1 \geq d(n-m)\), proving (1). For a coordinate-free proof, see [KL14].

To prove (2), define a map on basis elements by

\[
m_{x^1} : S^d(S^nW) \to S^d(S^nW)
((x^1)^{i_1} (x^2)^{i_2} \cdots (x^d)^{i_d}) \cdots ((x^1)^{i_1} \cdots (x^d)^{i_d}) \mapsto ((x^1)^{i_1} + (n-m)(x^2)^{i_2} \cdots (x^d)^{i_d}) \cdots ((x^1)^{i_1} + (n-m) \cdots (x^d)^{i_d})
\]

and extend linearly. A vector of weight \(\mu = (q_1, q_2, \ldots, q_d)\) is mapped under \(m_{x^1}\) to a vector of weight \(\pi = (p_1, \ldots, p_d) := \mu + (d(n-m)) = (q_1 + d(n-m), q_2, \ldots, q_d)\) in \(S^d(S^nW^*)\). Moreover, \(m_{x^1}\) takes highest weight vectors to highest weight vectors, as raising operators annihilate \(x_1\). The image of the space of highest weight vectors for the isotypic component of \(S_{\pi}W^*\) in \(S^d(S^nW^*)\) under \(m_{x^1}\) will be in \(\mathbb{C}[\text{Pad}_{n-m}(S^nW^*)]\) because, for example, such a polynomial will not vanish on \((e_1)^{n-m}[(e_1)^{i_1} \cdots (e^d)^{i_d} + \cdots + (e^1)^{i_1} \cdots (e^d)^{i_d}]\).

In §?? we will show

\[
\text{mult}(S_{\pi}W, S^d(S^nW)) = \text{mult}(S_{\pi + d(n-m)}W, S^d(S^nW))
\]
as soon as \(d \geq q_2 + \cdots + q_d\). This condition implies \(q_1 \geq dn - d\), so \(p_1 = q_1 + dn - dm \geq dn - d = d(n-1)\) as required. For the sufficiency of \(p_1 \geq dn - m\), note that if \(p_1 \geq (d-1)n + (n-m) = dn - m\), then in an element of weight \(\pi\), each of the exponents \(i_1, \ldots, i_d\) of \(x_1\) must be at least \(n-m\). So there again exists an element of \(\text{Pad}_{n-m}(S^nW)\) such that a vector of weight \(\pi\) does not vanish on it.

\[\square\]

**Remark 6.3.2.2.** For the subspace variety \(\text{Sub}_{k}(S^nW)\), a module appears in its ideal if and only if its entire isotypic component in \(\text{Sym}(S^nW^*)\) appears. This property fails for \(\text{Pad}_{n-m}(S^nW)\). For example, consider \(I_3(\text{Pad}_{n-2}(S^nC^2))\). For \(n \geq 6\), the module \(S_{(3(n-2),6)}W\) appears with multiplicity two in \(S^3(S^nW^*)\), by e.g., [Man98, Thm 4.2.2]. Using [Wey93, Cor. 4a], one computes that the module \(S_{(3(n-2),6)}W^*\) has multiplicity one.
6.3. Necessary conditions for modules of polynomials to be useful for GCT

in \( \mathbb{C}[\text{Pad}_{n-2}(S^nC^2)] \) and hence multiplicity one in \( I_3(\text{Pad}_{n-2}(S^nC^2)) \) as well. Thus as far as a GCT guide for “where to look” for good (abstract) modules is concerned, Theorem 6.3.3.3 appears to be sharp.

Regarding set-theoretic equations, we have the following result:

**Proposition 6.3.2.3. [KL14]** If \( 4m < n \), then the variety of padded polynomials \( \text{Pad}_{n-m}(S^nW^*) \) is cut out set-theoretically by equations of degree two. For any \( n > m \),

\[
I_2(\text{Pad}_{n-m}(S^nW^*)) = \bigoplus_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor + m} S_{2(n-m-k).2(m+k)}W.
\]

**Proof.** Write \( x = x_1, y = x_2 \), and note that the highest weight vector of \( S_{2n-2s,2s}W \subset S^2(S^nW) \) will be a linear combination of vectors of the form

\[
(x^n)(x^{n-2s}y^{2s}), (x^{n-1}y)(x^{n-2s+1}y^{2s-1}), \ldots, (x^{n-s}y^s)(x^{n-s}y^s)
\]

with a nonzero coefficient on the last factor. These vectors vanish on \( e_1^{n-m}h \) for any \( h \) if and only if \( s \geq m + 1 \). \( \square \)

What we really need to study is the variety \( \text{Pad}_{n-m}(\text{Sub}_k(S^dW)) \) of padded cones.

**Proposition 6.3.2.4. [KL14]** (Inheritance) Notations as above. \( I_d(\text{Pad}_{n-m}(\text{Sub}_k(S^nW^*))) \) consists of all modules \( S_\pi W \) such that \( S_\pi C^k \) is in the ideal of \( \text{Pad}_{n-m}(S^nC^{k^*}) \) and all modules whose associated partition has length at least \( k + 1 \).

**Exercise 6.3.2.5:** Prove Proposition 6.3.2.4.

6.3.3. GCT useful polynomials. The results of the previous two sections indicate that for the purposes of separating the determinant from any sequence of padded polynomials, we should restrict our search for modules in the ideal of \( \text{Det}_n \) as follows:

**Definition 6.3.3.1.** Let \( W = \mathbb{C}^{n^2} \). A module \( S_\pi W \subset I_d(\text{Det}_n) \) is \( (n,m) \)-GCT useful if \( S_\pi W \not\subset I_d(\text{Pad}_{n-m}(\text{Sub}_{m^2+1}(S^nW))) \).

**Remark 6.3.3.2.** For different complexity problems one makes a analogous definitions, see [KL14].

In summary:

**Corollary 6.3.3.3. [KL14]** A necessary condition for a sequence of modules \( S_\pi W_n \subset I_d(\text{Det}_n) \) to be \( (n,m) \)-GCT useful is

1. \( \ell(\pi_{n(m)}) \leq m^2 + 1 \),
2. If \( \pi_{n(m)} = (p_1, \ldots, p_t) \), then \( p_1 \geq d(n - m) \).
Moreover, if \( p_1 \geq \min\{d(n-1), dn-m\} \), then the necessary conditions are also sufficient. In particular, for \( p_1 \) sufficiently large, GCT usefulness depends only on the partition \( \pi \).

**Proof.** By Proposition 6.3.2.4, the ideal of \( \text{Pad}_{n-m}(\text{Sub}_k(S^nW)) \) may be deduced from that of \( \text{Pad}_{n-m}(S^nW) \), which explains condition (1), so we only consider the case \( k = w \). Condition (2) and the “moreover” assertion are a consequence of Theorem 6.3.2.1. \( \square \)

***mention state of the art here, also conjectures about ease of computation of Kronecker coefficients and transition from \( n = \text{poly}(m) \) to \( n = 2^m \).***

6.4. GCTV: Blackbox derandomization and the search for explicit linear spaces

This section to be written. One path to separating \( \text{Perm}_n^m \) from \( \text{Det}_n \) would be to find a large linear space that fails to intersect \( \text{Det}_n \). One then would have a reasonably tractable problem working with the projected varieties because the codimension becomes a polynomial in \( n \). This is the subject of [Mula] which will be discussed here. This is a problem of finding hay in a haystack, since a “random” linear space would work with probability one.

6.5. GCTVI: The needle exists

This section to be written. The search for representation-theoretic obstructions is a problem of finding a needle in a haystack. This section will discuss [Mulb], where arguments are made that at least the needle exists.

6.6. Dual varieties and GCT

In this section I describe modules in the ideal of \( \text{Det}_n \) that were found by solving a problem in classical algebraic geometry. I begin with a detour into the relevant classical geometry.

In mathematics, one often makes transforms to reorganize information, such as the Fourier transform. There are geometric transforms to “reorganize” the information in an algebraic variety. Taking the dual variety of a variety is one such, as I now describe.

6.6.1. Singular hypersurfaces and the dual of the Veronese. Perhaps the two most natural subvarieties in a space of polynomials are the Veronese variety \( X = v_d(\mathbb{P}V) \subset \mathbb{P}(S^dV) \), consisting of polynomials that are \( d \)-th powers, and the variety of polynomials \( P \in S^dV \) whose zero set \( Z(P) \subset \mathbb{P}V^* \) is not smooth. These two varieties are closely related:
6.6. Dual varieties and GCT

Let $P \in S^dV^*$, and let $Z_P \subset \mathbb{P}V$ be the corresponding hypersurface. By definition, $[x] \in Z_P$ if and only if $P(x) = 0$, i.e., $\mathcal{P}(x, \ldots, x) = \mathcal{P}(x^d) = 0$ considering $\mathcal{P}$ as a multilinear form. Similarly, $[x] \in (Z_P)_{\text{sing}}$ if and only if $P(x) = 0$, and $dP_x = 0$, i.e., for all $y \in W$, $0 = dP_x(y) = \mathcal{P}(x, \ldots, x, y) = \mathcal{P}(x^{d-1}y)$. Recall from Exercise 2.4.11.1 that $\hat{T}_x v_d(\mathbb{P}V) = \{x^{d-1}y \mid y \in V\}$. Introduce the notation $v_d(\mathbb{P}V)^{\vee} \subset \mathbb{P}S^dV^*$ for the variety of polynomials whose zero set is not smooth. We have

$$v_d(\mathbb{P}V)^{\vee} := \{[P] \in \mathbb{P}S^dV^* \mid \exists [x] \in v_d(\mathbb{P}V), \hat{T}_x v_d(\mathbb{P}V) \subset P_{\perp}\}$$

Here $P_{\perp} \subset S^dV$ is the hyperplane annihilating the vector $P$, when $P$ is considered as a linear form on $S^dV$.

The variety $v_d(\mathbb{P}V)^{\vee}$ is sometimes called the discriminant hypersurface.

**Exercise 6.6.1.1:** Show that $v_d(\mathbb{P}V) = \{[q] \in \mathbb{P}S^dV \mid \exists z \in v_d(\mathbb{P}V)^{\vee}_{\text{smooth}} \text{ such that } \hat{T}_z v_d(\mathbb{P}V)^{\vee} \subset q_{\perp}\}$.

**6.6.2. Dual varieties in general.** For a smooth irreducible variety $X \subset \mathbb{P}V$, define $X^{\vee} \subset \mathbb{P}V^*$, the dual variety of $X$, by

$$X^{\vee} := \{H \in \mathbb{P}V^* \mid \exists x \in X, \hat{T}_x X \subseteq H_{\perp}\}.$$  

**Exercise 6.6.2.1:** Show that $X^{\vee} = \{H \in \mathbb{P}V^* \mid X \cap H \text{ is not smooth}\}$.

Here $H$ refers both to a point in $\mathbb{P}V^*$ and the hyperplane $\mathbb{P}(H_{\perp}) \subset \mathbb{P}V$. The dual variety can be defined even when $X$ is singular, as long as it is irreducible, by

$$X^{\vee} := \{H \in \mathbb{P}V^* \mid \exists x \in X_{\text{smooth}}, \mathbb{P}\hat{T}_x X \subseteq H\}$$

$$= \{H \in \mathbb{P}V^* \mid \exists x \in X_{\text{smooth}} \text{ such that } X \cap H \text{ is not smooth at } x\}.$$  

For our purposes, the most important property of dual varieties is that for a smooth hypersurface other than a hyperplane, its dual variety is also a hypersurface. This will be a consequence of the B. Segre dimension formula 6.6.4.1 below. If the dual of $X \subset \mathbb{P}V$ is not a hypersurface, one says that $X^{\vee}$ is degenerate. It is a classical problem to study the varieties with degenerate dual varieties.

**Remark 6.6.2.2.** If $X$ is irreducible, then $X^{\vee}$ is irreducible, see, e.g., [Har95, p. 197].

Recall that we are looking for pathologies of the hypersurface $\{\det_n = 0\}$ to exploit. It turns out its dual variety is very degenerate:

**Proposition 6.6.2.3.** $\{\det_n = 0\}^{\vee} = Seg(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}) \subset \mathbb{P}^{n^2-1}$, in particular, $\dim\{\det_n = 0\}^{\vee} = 2n - 2 \leq n^2 - 2$.  

Proof. Let \( z \in \{ \det_n = 0 \}_{\text{smooth}} \), then without loss of generality

\[
z = \begin{pmatrix}
1 & & \\
& \ddots & \\
& & 1 \\
& & & 0
\end{pmatrix}
\]

where all off-diagonal entries are zero, so \( \hat{T}_z \{ \det_n = 0 \} \) consists of all matrices whose \((n, n)\)-entry is zero, and thus \((\hat{T}_z \{ \det_n = 0 \})^\perp\) is the line of matrices in the dual space where all entries are zero except the \((n, n)\)-slot. In particular, \((\hat{T}_z \{ \det_n = 0 \})^\perp\), is a rank one matrix, and thus \((\hat{T}_y \{ \det_n = 0 \})^\perp\) is a rank one matrix for all \( y \in \{ \det_n = 0 \}_{\text{smooth}} \). \( \square \)

6.6.3. What do varieties with degenerate dual varieties “look like”? Consider a curve \( C \subset \mathbb{P}^3 \), and a point \( p \in \mathbb{P}^3 \). Take \( J(C, p) \), the cone over \( C \) with vertex \( p \). Let \([x] \in J(C, p)\), write \( x = y + \bar{p} \).

Exercise 6.6.3.1: Show that if \( p \neq y \), \( \hat{T}_x J(C, p) = \text{span}\{\hat{T}_y C, \bar{p}\} \).

Thus the tangent space to the cone is constant along the rulings, and the surface only has a curves worth of tangent (hyper)-planes, so its dual variety is degenerate.

Consider again a curve \( C \subset \mathbb{P}^3 \), and this time let \( \tau(C) \subset \mathbb{P}^3 \) denote the Zariski closure of the union of all points on \( \mathbb{P}\hat{T}_x C \) as \( x \) ranges over the smooth points of \( C \).
Exercise 6.6.3.2: Show that if $y_1, y_2 \in \tau(C)$ are both on a tangent line to $x \in C$, then $\hat{T}_{y_1}\tau(C) = \hat{T}_{y_2}\tau(C)$, and thus $\tau(C)^{\vee}$ is degenerate.

A theorem of C. Segre from 1910 states the above two examples are the only surfaces with degenerate dual varieties:

**Theorem 6.6.3.3.** [Seg10, p. 105] Let $X^2 \subset \mathbb{P}^3$ be a surface with degenerate dual variety. Then $X$ is one of the following:

1. A linearly embedded $\mathbb{P}^2$,
2. A cone over a curve $C$, denoted $J(p, C)$,
3. A tangential variety to a curve $C$, that is the Zariski closure of the union of the points on tangent lines to $C$, denoted $\tau(C)$.

(1) is a special case of both (2) and (3) and is the only intersection of the two.

The proof is differential-geometric, see [IL03, §3.4]. Griffiths and Harris [GH79] vaguely conjectured a higher dimensional generalization of C. Segre’s theorem, namely that a variety with a degenerate dual is “built out of” cones and tangent developables. For example, \{det$_n = 0$\} may be thought of as the union of tangent lines to tangent lines to ... to the Segre variety $Seg(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})$.

Segre’s theorem already indicates that if we take the Zariski closure in $\mathbb{P}S^dV^{\ast}$ of the set of irreducible hypersurfaces of degree $d$ with degenerate dual varieties, we will obtain a reducible variety.

Define the (naïve) conormal space $N^\ast_x X := \hat{T}_x X^\perp \subset V^\ast$. (For the experts, this is a twist of the usual conormal space.) We obtain the geometric interpretation

$$\mathbb{P}N^\ast_x X = \{ H \in \mathbb{P}V^\ast \mid \mathbb{P}\hat{T}_x X \subseteq H \}.$$
For more on dual varieties see [?, §8.2].

6.6.4. B. Segre’s dimension formula.

Proposition 6.6.4.1 (B. Segre, see e.g., [GKZ94]). Let $P \in S^d W^*$ be irreducible and let $x \in \hat{Z}(P)$ be a general point. Then
\[
\dim Z(P)^\vee = \text{rank}(P_{d-2,2}(x^{d-2})) - 2.
\]

Here $(P_{d-2,2}(x^{d-2})) \in S^2 W^*$. In coordinates, $P_{d-2,2}$ may be written as a symmetric matrix whose entries are polynomials of degree $d - 2$ in the coordinates of $x$, and is called the Hessian.

Proof. Let $x \in \hat{Z}(P) \subset W$ be a smooth point, so $P(x) = \overline{P}(x, \ldots, x) = 0$ and $dP_x = \overline{P}(x, \ldots, x, \cdot) \neq 0$ and take $h = dP_x \in W^*$, so $[h] \in Z(P)^\vee$. Now consider a curve $h_t \subset \hat{Z}(P)^\vee$ with $h_0 = h$. There must be a corresponding (possibly stationary) curve $x_t \in \hat{Z}(P)$ such that $h_t = \overline{P}(x_t, \ldots, x_t, \cdot)$ and thus $h'_0 = \overline{P}(x^{d-2}, x'_0, \cdot)$. Thus the dimension of $\hat{T}_h Z(P)^\vee$ is the rank of $P_{d-2,2}(x^{d-2})$ minus one (we subtract one because $x'_0 = x$ yields the zero vector). Now just recall that $\dim X = \dim \hat{T}_x X - 1$. \hfill \Box

Exercise 6.6.4.2: Show that every $P \in \text{Sub}_k(S^d W)$ has $\dim Z(P)^\vee \leq k - 2$.

Exercise 6.6.4.3: Show that $\sigma_3(CH_n(\mathbb{C}^{n^2})) \not\subset \text{Det}_n$.

Exercise 6.6.4.4: Show that $\sigma_{2n+1}(v_n(\mathbb{P}^{n^2-1})) \not\subset \text{Det}_n$.

Exercise 6.6.4.5: Show that $\{x_1 \cdots x_n + y_1 \cdots y_n = 0\} \subset \mathbb{P}^{2n-1}$ is self dual, in the sense that it is isomorphic to its own dual variety.

***fix $V$ v. $W$ and duals***

6.6.5. First steps towards equations. Segre’s formula implies, for $P \in S^d W^*$, that $\dim Z(P)^\vee \leq k$ if and only if, for all $w \in W$,
\[
(6.6.1) \quad P(w) = 0 \ \Rightarrow \ \det_{k+3}(P_{d-2,2}(w^{d-2})|_F) = 0 \ \forall F \in G(k+3, W).
\]

Equivalently (assuming $P$ is irreducible), for any $F \in G(k + 3, W)$, the polynomial $P$ must divide $\det_{k+3}(P_{d-2,2}|_F) \in S^{(k+3)(d-2)} W^*$, where $\det_{k+3}$ is evaluated on the $S^2 W^*$ factor in $S^2 W^* \otimes S^{d-2} W^*$.

Thus to find polynomials on $S^d W^*$ characterizing hypersurfaces with degenerate duals, we need polynomials that detect if a polynomial $P$ divides a polynomial $Q$. Now $P$ divides $Q$ if and only if $Q \in P \cdot S^{e-d} V$, i.e.
\[
x^{I_1} P \land \cdots \land x^{I_d} P \land Q = 0
\]
where $x^{I_j}$ is a basis of $S^{e-d} V$ (and $D = \binom{n+e-d-1}{e-d}$), which potentially give a $\binom{n+e-1}{D+1}$ dimensional space of equations. Let $D_{k,d,N} \subset \mathbb{P} S^d \mathbb{C}^N$ denote the zero set of these equations.
Define $\text{Dual}_{k,d,N} \subset \mathbb{P}(S^dW^*)$ as the Zariski closure of the set of irreducible hypersurfaces of degree $d$ in $\mathbb{P}W \simeq \mathbb{P}^{N-1}$, whose dual variety has dimension at most $k$. Our discussion above implies $\text{Dual}_{k,d,N} \subseteq \mathcal{D}_{k,d,N}$.

Note that $[\det_n] \in \text{Dual}_{2n-2,n,n^2} \subseteq \mathcal{D}_{2n-2,n,n^2}$.

### 6.6. Dual varieties and GCT

**Exercise 6.6.6.1:** Compute $\text{perm}^{k,d,N}_{m-2,2}(x^{m-2})$ when

$$x = \begin{pmatrix} 1 - m & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

and show it is of maximal rank.

**Proposition 6.6.6.2.** Let $U = \mathbb{C}^M$, let $R \in S^mU^*$ be irreducible, let $\ell$ be a coordinate on $\mathbb{C}$ be nonzero, let $U^* \oplus \mathbb{C} \subset W^*$ be a linear inclusion.

If $[R] \in \mathcal{D}_{\kappa,m,M}$ and $[R] \not\in \mathcal{D}_{\kappa-1,m,M}$, then $[\ell^{d-m}R] \in \mathcal{D}_{\kappa,d,N}$ and $[\ell^{d-m}R] \not\in \mathcal{D}_{\kappa-1,d,N}$.

**Proof.** Choose a basis $u_1, \ldots, u_M, v, w_{M+2}, \ldots, w_N$ of $W$ so $(U^*)^\perp = \langle w_{M+2}, \ldots, w_N \rangle$ and $(L^*)^\perp = \langle u_1, \ldots, u_M, w_{M+2}, \ldots, w_N \rangle$. Let $c = (d-m)(d-m-1)$. In these coordinates, the matrix of $2n-2, n, n^2_{d-2,2}$ in $(M,1,N-M-1) \times (M,1,N-M-1)$-block form:

$$2n-2, n, n^2_{d-2,2} = \begin{pmatrix} \ell^{d-m}R_{m-2,2} & \ell^{d-m}R_{m-1,1} & 0 \\ \ell^{d-m}R_{m-1,1} & c\ell^{d-m}R_{m-2,2} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$  

First note that $\det_{M+1}(2n-2, n, n^2_{d-2,2} | F)$ for any $F \in G(M+1,W)$ is either zero or a multiple of $2n-2, n, n^2$. If $\dim Z(R)^\vee = M-2$ (the expected dimension), then for a general $F \in G(M+1,W)$, $\det_{M+1}(2n-2, n, n^2_{d-2,2} | F)$ will not be a multiple of $2n-2, n, n^2$, and more generally if $\dim Z(R)^\vee = \kappa$, then for a general $F \in G(\kappa+2,W)$, $\det_{\kappa+2}(\ell^{d-m}R_{d-2,2} | F)$ will not be a multiple of $\ell^{d-m}R$ but for any $F \in G(\kappa+3,W)$, $\det_{\kappa+3}(\ell^{d-m}R_{d-2,2} | F)$ will be a multiple of $\ell^{d-m}R$. This shows $[R] \not\in \mathcal{D}_{\kappa-1,m,M}$, implies $[\ell^{d-m}R] \not\in \mathcal{D}_{\kappa-1,d,N}$.

**Exercise 6.6.6.3:** Show that $[R] \in \mathcal{D}_{\kappa,m,M}$, implies $[\ell^{d-m}R] \in \mathcal{D}_{\kappa,d,N}$. ⊗

Combining Exercise 6.6.6.1 and Proposition 6.6.6.2, we conclude:

**Theorem 6.6.6.4.** [LMR13] $\text{Perm}_{m}^{n} \not\in \mathcal{D}_{2n-2,n,n^2}$ when $m < \frac{n^2}{2}$. In particular, $\overline{\text{de}}(\text{perm}_m) \geq \frac{m^2}{2}$.  

\[\square\]
On the other hand, by Exercise 6.6.4.2 cones have degenerate duals, so \( \ell^{n-m} \operatorname{perm}_m \in \mathcal{D}_{2n-2,n,n^2} \) whenever \( m \geq \frac{n^2}{2} \).

### 6.6.7. A better module of equations.

The equations above are of degree \( (v+e-d-1)+e \), where \( e = (k+e)(d-2) \). I now derive equations of much lower degree. Since \( P \in S^d W \) divides \( Q \in S^e W \) if and only if for each \( L \in G(2, W) \), \( P|_L \) divides \( Q|_L \), it will be sufficient to solve this problem for polynomials on \( \mathbb{C}^2 \). This will have the advantage of producing polynomials of much lower degree.

Let \( V = \mathbb{C}^2 \), let \( d \leq e \), let \( P \in S^d V \) and \( Q \in S^e V \). If \( P \) divides \( Q \) then \( S^{e-d}V \cdot P \) will contain \( Q \). That is, the vectors \( x^{e-d}P, x^{e-d-1}yP, \ldots, y^{e-d}P, Q \) in \( S^e V \) will fail to be linearly independent, i.e.,

\[
x^{e-d}P \wedge x^{e-d-1}yP \wedge \cdots \wedge y^{e-d}P \wedge Q = 0.
\]

Since \( \dim S^e V = e + 1 \), these potentially give a \( \left( \frac{e+1}{e-d+2} \right) \)-dimensional vector space of equations, of degree \( e - d + 1 \) in the coefficients of \( P \) and linear in the coefficients of \( Q \).

Let’s compute the highest weight of these equations as a \( GL(V) \)-module. Using the summation convention, write \( P = P_{i,j} x^i y^j \), \( Q = Q_{s,t} x^s y^t \). Then the highest weight of a vector appearing in

\[
x^{e-d}P \wedge x^{e-d-1}yP \wedge \cdots \wedge y^{e-d}P \wedge Q = 0.
\]

will be the coefficient of \( x^e \wedge x^{e-1}y \wedge \cdots \wedge x^{d-1}y^{e-d+1} \). The weight will be the sum of the subscript indices appearing in a monomial. All monomials appearing with this coefficient will have the same weight, so consider

\[
x^e P_{d,0} \wedge x^{e-1}yP_{d,0} \wedge \cdots \wedge x^{d-1}y^{e-d+1}P_{d,0} \wedge Q_{d-1,e-d+1},
\]

which has highest weight \( (ed - d^2 + 2d - 1, e - d + 1) \). The irreducible \( GL_2 \)-module with this highest weight has dimension \( (d-1)(2+e-d) + 1 \), which is much less than the number of equations, which is \( \left( \frac{e+1}{e-d+2} \right) \), so either there are other irreducible modules in the decomposition, or the equations are far from being linearly independent.

By taking our polynomials to be \( P = P|_L \) and \( Q = \det_{k+3}(P_{n-2,2}|_F)|_L \) for \( F \in G(k+3, V) \) and \( L \in G(2, F) \) (or, for those familiar with flag varieties, better to say \( (L, F) \in \text{Flag}_{2,k+3}(V) \)) we now have equations parametrized by the pairs \( (L, F) \). Note that \( \deg(Q) = e = (k + 3)(d-2) \).

**Remark 6.6.7.1.** More generally, with \( V = \mathbb{C}^2 \), given \( P \in S^d V \), \( Q \in S^e V \), one can ask if \( P, Q \) have at least \( r \) roots in common (counting multiplicity). Then \( P, Q \) having \( r \) points in common says the spaces \( S^{e-r}V \cdot P \) and \( S^{d-r}V \cdot Q \) intersect. That is,

\[
x^{e-r}P \wedge x^{e-r-1}yP \wedge \cdots \wedge y^{e-r}P \wedge x^{d-r}Q \wedge x^{d-r-1}yQ \wedge \cdots \wedge y^{d-r}Q = 0.
\]
6.6. Dual varieties and GCT

In the case \( r = 1 \), we get a single polynomial, called the resultant, which is of central importance. In particular, the proof that the projection of a projective variety \( X \subset \mathbb{P}W \) from a point \( y \in \mathbb{P}W \) with \( y \not\in X \), to \( \mathbb{P}(W/y) \) is still a projective variety relies on the resultant to produce equations for the projection. This in turn enables the definition of dimension and degree often used in projective algebraic geometry. See, e.g., [Har95, Lect. 3].

6.6.8. Module structure of the better set of equations. Write \( P = \sum_j \tilde{P}_j x^J \) with the sum over \( |J| = d \). The weight of a monomial \( \tilde{P}_j x^J \) is \( J_0 = (j_1, \ldots, j_n) \). Adopt the notation \([i] = (0, \ldots, 0, 1, 0, \ldots, 0)\) where the 1 is in the \( i \)-th slot and similarly for \([i, j]\) where there are two 1’s. The entries of \( P_{d-2,2} \) are, for \( i \neq j \), \( (P_{d-2,2})_{i,j} = \tilde{P}_{I+[i,j]} x^I \), and for \( i = j \), \( P_{I+2|j} x^I \), where \(|I| = d - 2 \), and \( P_J \) is \( \tilde{P}_J \) with the coefficient adjusted, e.g., \( P_{(d,0,\ldots,0)} = d(d-1)\tilde{P}_{(d,0,\ldots,0)} \) etc. (This won’t matter because we are only concerned with the weights of the coefficients, not their values.) To determine the highest weight vector, take \( L = \text{span}\{e_1, e_2\} \), \( F = \text{span}\{e_1, \ldots, e_{k+3}\} \). The highest weight term of

\[
(x_1^{e-d}P|_L) \wedge (x_1^{e-d-1}x_2P|_L) \wedge \cdots \wedge (x_2^{e-d}P|_L) \wedge (\det_{k+3}(P_{d-2,2}|_F))|_L
\]

is the coefficient of \( x_1^e \times x_2^{e-d} \). It will not matter how we distribute these for the weight, so take the coefficient of \( x_1^e \) in \( (\det_{k+3}(P_{d-2,2}|_F))|_L \). It has leading term \( P_{(d,0,\ldots,0)}P_{(d-2,2,0,\ldots,0)}P_{(d-2,0,2,0,\ldots,0)} \cdots P_{(d-2,0,\ldots,0,2,0,\ldots,0)} \) which is of weight \( (d+(k+2)(d-2),2k+2) \). For each \( (x_1^{e-d-s}x_2^sP|_L) \) take the coefficient of \( x_1^{e-s-1}x_2^{s+1} \) which has the coefficient of \( P_{(d-1,1,0,\ldots,0)} \) each time, to get a total weight contribution of \( ((e-d+1)(d-1),(e-d+1),0,\ldots,0) \) from these terms. Adding the weights together, and recalling that \( e = (k+3)(d-2) \) the highest weight is

\[
(d^2k + 2d^2 - 2dk - 4d + 1, dk + 2d - 2k - 3, 2k+1),
\]

which may be written as

\[
((k+2)(d^2 - 2d) + 1, (k+2)(d-2) + 1, 2k+1).
\]

In summary:

**Theorem 6.6.8.1.** [LMR13] The ideal of the variety \( D_{k,d,N} \subset \mathbb{P}(S^d\mathbb{C}^N^*) \) contains a copy of the \( GL_N \)-module \( S_{\pi(k,d)}\mathbb{C}^N \), where

\[
\pi(k,d) = ((k+2)(d^2 - 2d) + 1, d(k+2) - 2k - 3, 2k+1).
\]

Since \(|\pi| = d(k+2)(d-1)\), these equations have degree \((k+2)(d-1)\).

Observe that the module \( \pi(2n-2,n) \) indeed satisfies the requirements to be \((m,\frac{m^2}{2})\)-GCT useful, as \( p_1 = 2n^3 - 2n^2 + 1 > n(n-m) \) and \( \ell(\pi(2n-2,n)) = 2n+1 \).
Remark 6.6.8.2. I do not know whether or not the module of equations obtained by the span of the pairs \((L, P)\) is irreducible or not. The dimension count indicates either there is substantial redundancy among the equations or the module is not irreducible.

Proposition 6.6.8.3. When restricted to the open subset of irreducible hypersurfaces in \(S^d \mathbb{C}^N\), \(\text{Dual}_{k,d,N} = D_{k,d,N}\) as sets.

Proof. Let \(P \in D_{k,d,N}\) be irreducible. For each \((L, F)\), one obtains set-theoretic equations for the condition that \(P|_L\) divides \(Q|_L\), where \(Q = \det(P_{d-2,2}|_F)\). But \(P\) divides \(Q\) if and only if restricted to each plane \(P\) divides \(Q\), so these conditions imply that the dual variety of the irreducible hypersurface \(Z(P)\) has dimension at most \(k\). \(\square\)

Remark 6.6.8.4. The module with highest weight \(\pi(2n-2, n)\) in \(I_{2n(n-1)}(D_{et}) \subset S^{2n(n-1)}(\mathbb{C}^{n^2})\) is unlikely to be even a representation-theoretic obstruction. For example, when \(n = 3\), the module with highest weight \((19, 7, 2^5)\) occurs with multiplicity six in \(S^{12}(\mathbb{C}^3)\), but only one copy of it appears to be in the ideal.

6.6.9. \(D_{et}\) is an irreducible component of \(D_{2n-2,n,n^2}\). This section is for those familiar with Zariski tangent spaces to schemes. It is not used elsewhere. Given two schemes, \(X, Y\) with \(X\) irreducible and \(X \subseteq Y\), when one has an equality of Zariski tangent spaces, \(T_X X = T_X Y\) for some \(x \in X_{\text{smooth}}\), this implies that \(X\) is an irreducible component of \(Y\) (and in particular, if \(Y\) is irreducible, that \(X = Y\)). The following theorem is the main result of [LMR13]:

Theorem 6.6.9.1. [LMR13] The scheme \(D_{2n-2,n,n^2}\) is smooth at \([\text{det}_n]\), and \(D_{et}\) is an irreducible component of \(D_{2n-2,n,n^2}\).

**picture**

The idea of the proof is as follows: We need to show \(T_{[\text{det}_n]} D_{n,2n-2,n^2} = T_{[\text{det}_n]} D_{et}\). We already know \(T_{[\text{det}_n]} D_{n,2n-2,n^2} \subseteq T_{[\text{det}_n]} D_{et}\). Both of these vector spaces are \(G_{\text{det}_n}\)-submodules of \(S^n(E \otimes F) = \bigoplus_{|\pi| = n} S_\pi E \otimes S_\pi F\). It is easy to see that, as a \(GL(E) \times GL(F)\)-module, \(T_{[\text{det}_n]} D_{et} = S_{2,1^{n-2}} E \otimes S_{2,1^{n-2}} F\), so the problem becomes to show that none of the other modules are in \(T_{[\text{det}_n]} D_{n,2n-2,n^2}\). To do this, it suffices to check a single point in each module. A first guess would be to check highest weight vectors, but these are not so easy to write down in any uniform manner. Fortunately in this case there is another choice, namely the immanants (see Remark 5.2.6.8), and the proof in [LMR13] proceeds by checking that none of these other than \(IM_{2,1^{n-2}}\) are contained in \(T_{[\text{det}_n]} D_{n,2n-2,n^2}\).

Exercise 6.6.9.2: Show that \(T_{\text{det}_n} D_{et} = S_{1^n} E \otimes S_{1^n} F \oplus S_{2,1^{n-1}} E \otimes S_{2,1^{n-1}} F\).
Theorem 6.6.9.1 implies that the $GL(W)$-module of highest weight $\pi(2n-2,n)$ given by Theorem 6.6.8.1 gives local equations at $[\det_n]$ of $\text{Det}_n$, of degree $2n(n-1)$. Since $\text{Sub}_k(S^n\mathbb{C}^N) \subset \mathcal{D}_{k,n,N}$, the zero set of the equations is strictly larger than $\text{Det}_n$. Recall that $\dim \text{Sub}_k(S^n\mathbb{C}^n^2) = \binom{k+n+1}{n} + \binom{k}{2}(N-k-2) - 1$. For $k = 2n-2$, $N = n^2$, this is larger than the dimension of the orbit of $[\det_n]$, and therefore $\mathcal{D}_{2n-2,n,n^2}$ is not irreducible.

### 6.6.10. On the boundary of the orbit of the determinant

**begin with general discussion of finding components on the boundary**

Recall that the transposition $\tau \in G_{\det_n}$ allows us to write $\mathbb{C}^{n^2} = E \otimes E = S^2E \oplus \Lambda^2E$, where the decomposition is into the $\pm 1$ eigenspaces for $\tau$. For $M \in E \otimes E$, write $M = M_S + M_A$ reflecting this decomposition.

Define a polynomial $P_\lambda \in S^n(\mathbb{C}^{n^2})^*$ by

$$P_\lambda(M) = \overline{\det}_n(M_A, \ldots, M_A, M_S).$$

Let $P_f(\lambda_A)$ denote the Pfaffian of the skew-symmetric matrix, obtained from $\lambda_A$ by suppressing its $i$-th row and column. Write $M_S = (s_{ij})$.

**Exercise 6.6.10.1:** Show that

$$P_\lambda(M) = \sum_{i,j} s_{ij} P_f(\lambda_A) P_{f_j}(\lambda_A).$$

In particular, $P_\lambda = 0$ if $n$ is even but is not identically zero when $n$ is odd.

**Proposition 6.6.10.2. [LMR13]** $P_{\lambda,n} \in \mathcal{D}_{n,n}$. Moreover, $GL(W) \cdot P_\lambda$ is an irreducible codimension one component of the boundary of $\mathcal{D}_{n,n}$, not contained in $\text{End}(W) \cdot [\det_n]$. In particular $\text{dc}(\lambda_{\lambda,m}) = m < \text{dc}(P_{\lambda,m})$.

**Proof.** For the first assertion, note that

$$P_\lambda(M) = \lim_{t \to 0} \frac{1}{t} \det(M_A + tM_S).$$

To prove the second assertion, we compute the stabilizer of $P_\lambda$ inside $GL(M_n(\mathbb{C}))$.

The action of $GL(E)$ on $E \otimes E$ by $M \mapsto gMg^t$ preserves $P_\lambda$ up to scale, and the Lie algebra of the stabilizer of $[P_\lambda]$ is a $GL(E)$ submodule of $\text{End}(E \otimes E)$. Decompose $\text{End}(E \otimes E)$ as a $GL(E)$-module:

$$\text{End}(E \otimes E) = \text{End}(\Lambda^2E) \oplus \text{End}(S^2E) \oplus \text{Hom}(\Lambda^2E, S^2E) \oplus \text{Hom}(S^2E, \Lambda^2E)$$

$$= \Lambda^2E \otimes \Lambda^2E^* \oplus S^2E \otimes S^2E^* \oplus \Lambda^2E^* \otimes S^2E \oplus S^2E^* \otimes \Lambda^2E$$

$$= (\mathfrak{gl}(E) \oplus S_{2,1n-2}E) \oplus (\mathfrak{gl}(E) \oplus S_{4,2n-1}E) \oplus (\mathfrak{sl}(E) \oplus S_{3,1n-2}E) \oplus (\mathfrak{sl}(E) \oplus S_{32,2n-2}E)$$

By testing highest weight vectors, one concludes the Lie algebra of $G_{P_\lambda}$ is isomorphic to $\mathfrak{gl}(E) \oplus \mathfrak{gl}(E)$, which has dimension $2n^2 = \dim G_{\det_n} + 1$, implying $GL(W) \cdot P_\lambda$ has codimension one in $GL(W) \cdot [\det_n]$. Since it is

---

not contained in the orbit of the determinant, it must be an irreducible component of its boundary. Since the zero set is not a cone, $P_\Lambda$ cannot be in $\text{End}(W) \cdot \det_n$ which consists of $GL(W) \cdot \det_n$ plus cones, as any element of $\text{End}(W)$ either has a kernel or is invertible. □

Exercise 6.6.10.3: Verify by testing on highest weight vectors that the only ones in (6.6.2) annihilating $P_\Lambda$ are those in $\mathfrak{gl}(E) \oplus \mathfrak{gl}(E)$. Note that as a $\mathfrak{gl}(E)$-module, $\mathfrak{gl}(E) = \mathfrak{sl}(E) \oplus \mathbb{C}$ so one must test the highest weight vector of $\mathfrak{sl}(E)$ and $\mathbb{C}$.

The hypersurface defined by $P_\Lambda$ has interesting properties.

Proposition 6.6.10.4. [LMR13]

\[ Z(P_\Lambda)^\vee = \overline{\{v^2 \oplus v \wedge w \in S^2 \mathbb{C}^n \oplus \Lambda^2 \mathbb{C}^n, v, w \in \mathbb{C}^n\}} \subset \mathbb{P}^{n^2-1}. \]

As expected, $Z(P_\Lambda)^\vee$ resembles $\text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})$.

Remark 6.6.10.5. For those familiar with the notation, $Z(P_\Lambda)$ can be defined as the image of the projective bundle $\pi : \mathbb{P}(E) \to \mathbb{P}^{n-1}$, where $E = \mathcal{O}(-1) \oplus Q$ is the sum of the tautological and quotient bundles on $\mathbb{P}^{n-1}$, by a sub-linear system of $\mathcal{O}_E(1) \otimes \pi^* \mathcal{O}(1)$. This sub-linear system contracts the divisor $\mathbb{P}(Q) \subset \mathbb{P}(E)$ to the Grassmannian $G(2,n) \subset \mathbb{P} \Lambda^2 \mathbb{C}^n$.

Remark 6.6.10.6. A second way to realize the polynomial $P = \ast \ast \ast$ from Example ?? is via $P_\Lambda$: take

\[ M_\Lambda = \begin{pmatrix} 0 & x_3 & x_2 \\ -x_3 & 0 & x_1 \\ -x_2 & -x_1 & 0 \end{pmatrix}, \quad M_S = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_4 & 0 \\ 0 & 0 & x_2 \end{pmatrix}. \]

6.7. Remarks on Kronecker coefficients and plethysm coefficients

To pursue the path outlined in [MS01, MS08], the first task would be to find a useful module in $\mathcal{I}(\text{Det}_n)$. That is, by the results described in §6.3, to find a sequence (depending on $n$, which I suppress from the notation of $\pi = \pi(n)$ and $d = d(n)$) of partitions $\pi = (p_1, \ldots, p_\ell)$ of $dn$ such that

- $sk^{p_\ell}_{d^n, d^n} < \text{mult}(S_\pi W, S^d(S^n W))$
- $\ell(\pi) \leq m$
- $p_1 > d(n - m)$.

To advance the state of the art, by the results discussed in §6.6 it will necessary to take $m \leq \sqrt{n}$. Write $p_\pi(d[n]) := \text{mult}(S_\pi W, S^d(S^n W))$ for this plethysm coefficient.
What follows is information about Kronecker coefficients and plethysm coefficients that may be relevant for this problem.

6.7.1. **Remarks on Kronecker coefficients.** **This section under construction***

6.7.2. **Remarks on plethysm coefficients.** First observe that

\[
S^d(S^n V) = (V \otimes^{dn}) \mathbb{S}_n \wr \mathbb{S}_d
\]

where \( \mathbb{S}_n \wr \mathbb{S}_d = \mathbb{S}_n \times \mathbb{S}_d \subset \mathbb{S}_{dn} \) is the *wreath product*, which consists of permutations acting on \( V \otimes^{dn} \) considered as \( d \) blocks of \( n \) copies of \( V \), permuting within each of the \( n \) blocks and permuting the blocks. We have,

\[
S^d(S^n V) = \bigoplus_{|\pi| = dn} [\pi] \otimes S_{\pi} V \wr S_d
\]

and we conclude, assuming \( \dim V \geq dn \), that

\[
p_\pi(d[n]) := \text{mult}(S_\pi V, S^d(S^n V)) = \dim[\pi] \mathbb{S}_n \wr \mathbb{S}_d.
\]

Now let \( \dim V = dn \). Fix a basis of \( V \) so \( \mathcal{W}_V \simeq \mathbb{S}_{dn} \subset GL(V) \) is realized as the group of permutation matrices. (I use \( \mathcal{W}_V \) instead of \( \mathbb{S}_{dn} \) to avoid confusion with the \( \mathbb{S}_{dn} \) that acts on \( V \otimes^{dn} \) by permuting the factors: \( \mathcal{W}_V \) acts on \( V \otimes^{dn} \) by acting on each factor.) Let \( (S_\pi V)_0 \subset S_\pi V \) denote the \( (1^{dn}) \)-weight space (the \( \mathfrak{sl}(V) \)-weight zero subspace). Then \( \mathcal{W}_V \) acts on \( (S_\pi V)_0 \).

**Proposition 6.7.2.1.** Notation as above, let \( |\pi| = dn \), then \( (S_\pi V)_0 = [\pi] \) as a \( \mathcal{W}_V \)-module.

More generally:

**Theorem 6.7.2.2.** [Gay76] For \( |\pi| = d = vs \) and \( |\mu| = v \),

\[
\text{mult}_{\mathfrak{sl}_V}([\mu], (S_\pi V)_0) = \text{mult}_{GL(V)}(S_\pi V, S_\mu(S^s V)).
\]

In particular:

\[
p_\pi(d[n]) = \text{mult}_{\mathfrak{sl}_V}([d], (S_\pi V)_0).
\]

***Proofs to be added***

6.8. **Symmetries of polynomials and coordinate rings of orbits**

***Needs an introduction***

The calculations of the coordinate rings of the orbits follow [BLMW11]. Examples 6.8.2, 6.8.3, 6.8.5, and 6.8.6 follow [CKW10].
Throughout this section $G = GL(V)$, $\dim V = n$, and I use index ranges $1 \leq i, j, k \leq n$.

I begin with the cases of: a generic polynomial, and the most degenerate polynomial, namely $x^d$.

Other than §6.8.3, this section contains auxiliary results and may be skipped on a first reading.


Let $P \in S^d V$ be generic. If $d, n > 3$, then $G_P = \{ \lambda I_d : \lambda^d = 1 \} \simeq \mathbb{Z}_d$, hence $GL(V) \cdot P \simeq GL(V)/\mathbb{Z}_d$, where $\mathbb{Z}_d$ acts as multiplication by the $d$-th roots of unity, see [Pop75]. (If $P \in S^d V$ is any element, then $\mathbb{Z}_d \subset G_P$.)

We need to determine the $\mathbb{Z}_d$-invariants in $GL(V)$-modules. Since $S_\pi V$ is a submodule of $V \otimes |\pi|$, $\omega \in \mathbb{Z}_d$ acts on $S_\pi V \otimes (\text{det } V)^{-s}$ by the scalar $\omega |\pi|^{-ns}$.

By Theorem 5.1.7.1, we conclude the following equality of $GL(V)$-modules:

$$C[GL(V) \cdot P] = \bigoplus_{(\pi, s) \mid d | |\pi|−ns} (S_\pi V^*)^{\oplus \dim S_\pi V \otimes (\text{det } V^*)^{-s}}.$$  

When we pass to $C[GL(V) \cdot P] = \bigoplus_{\delta} S^\delta (S^d V^*)/I_\delta(GL(V) \cdot P)$ we loose all terms with $s > 0$. Since degree is respected, we may write:

$$(6.8.1) \quad C[GL(V) \cdot P]_\delta \subseteq \bigoplus_{\pi \mid |\pi| = \delta d} (S_\pi V^*)^{\oplus \dim S_\pi V}.$$

Note that $C[GL(V) \cdot P]_\delta \subset S^\delta (S^d V)$, but there are far fewer modules and multiplicities in $S^\delta (S^d V)$ than on the right hand side of (6.8.1).

### 6.8.2. A point of $v_d(\mathbb{P}V)$. Let $P = x_1^d \in S^d V$. Let $g = (g_1^j) \in GL(V)$.

Then $g \cdot (x_1^d) = (g_1^j x_j^d)$ so if $g \cdot (x_1^d) = x_1^d$, then $g_1^j = 0$ for $j > 1$ and $g_1^1$ must be a $d$-th root of unity. There are no other restrictions, thus

$$G_P = \left\{ g \in GL(V) \mid g = \begin{pmatrix} \omega & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix}, \omega^d = 1 \right\}, \quad G_{[P]} = \left\{ g \in GL(V) \mid g = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix} \right\}.$$  

The $GL(V)$ orbit of $[x_1^d]$ is closed and equal to the Veronese variety $v_d(\mathbb{P}V)$.

**Exercise 6.8.2.1:** Use Corollary 6.2.3.1 to determine $C[v_d(\mathbb{P}V)]$ (even if you already know it by a different method).
6.8.3. The Chow polynomial. Let \( P = \text{chow}_n := x_1 \cdots x_n \in S^n V \), which I will call the “Chow polynomial”. It is clear \( \Gamma_n := T_n^{SL} \ltimes \mathfrak{S}_n \subset G_{\text{chow}} \), we need to determine if the stabilizer is larger. We can work by brute force: \( g \cdot \text{chow}_n = (g^j x_j)(g^j x_j) \cdots (g^j x_j) \). In order that this be equal to \( x_1 \cdots x_n \), by unique factorization of polynomials, there must be a permutation \( \sigma \in \mathfrak{S}_n \) such that for each \( k \), we have \( g^j x_j = \lambda_k x_{\sigma(k)} \) for some \( \lambda_k \in \mathbb{C}^\ast \).

Composing with the inverse of this permutation we have \( g^j x_j = \delta^j_k \lambda_j \), and finally we see that we must further have \( \lambda_1 \cdots \lambda_n = 1 \), which means it is an element of \( T_n^{SL} \), so the original \( g \) is an element of \( \Gamma_n \). Thus \( G_{\text{chow}} = \Gamma_n \).

The orbit closure of \( \text{chow}_n \) is the Chow variety \( Ch_n(V) \subset \mathbb{P} S^n V \).

When \( v = n \) (which we may restrict to by inheritance), the space of \( T_n^{SL} \) invariants in a \( GL(V) \)-module \( S_\pi V \) is the \( (d^n) \)-weight space, which I will denote by \( (S_\pi V)_0 \) because it is the \( \mathfrak{sl}(V) \)-weight zero space. The algebraic Peter-Weyl theorem 5.1.7.1, implies

\[
\mathbb{C}[GL(V) \cdot (x_1 \cdots x_n)] = \bigoplus_{\ell(\pi) \leq n} (S_\pi V^*)^\oplus \dim(S_\pi V)_0^{\mathfrak{S}_n},
\]

Now specialize to the case of modules appearing in \( Sym(S^n V) \). Corollary ??(ii) says \( \dim(S_\pi V)_0^{\mathfrak{S}_n} = \text{mult}(S_\pi V, S^n(S^4 V)) \). If we consider all the \( \pi \)'s together, we conclude

\[
\mathbb{C}[GL(V) \cdot (x_1 \cdots x_n)]_{\text{poly}} = \bigoplus_s S^n(S^4 V^*).
\]

In particular, \( \bigoplus_s S^n(S^4 V^*) \) inherits a ring structure. We’ll return to this in §8.1.1.

6.8.4. Techniques. For other polynomials, the determination of \( G_P \) needs further techniques. Here is a brief overview of some methods.

One technique is to form auxiliary objects from \( P \) which have a symmetry group \( H \) that one can compute, and by construction \( H \) contains \( G_P \). Usually it is easy to find a group \( H' \) that clearly is contained in \( G_P \), so if \( H = H' \), we are done.

One can determine the connected component of the stabilizer by a Lie algebra calculation: If we are concerned with \( p \in S^dV \), the connected component of the identity of the stabilizer of \( p \) in \( GL(V) \) is the connected Lie group associated to the Lie subalgebra of \( \mathfrak{gl}(V) \) that annihilates \( p \). (The analogous statement holds for tensors.) To see this, let \( \mathfrak{h} \subset \mathfrak{gl}(V) \) denote the annihilator of \( p \) and let \( H = \exp(\mathfrak{h}) \subset GL(V) \) the corresponding Lie group. Then it is clear that \( H \) is contained in the stabilizer as \( h \cdot p = \exp(X) \cdot p = (Id + X + \frac{1}{2}XX + ...)p \) the first term preserves \( p \).
and the remaining terms annihilate it. Similarly, if $H$ is the group preserving $p$, taking the derivative of any curve in $H$ through $Id$ at $t = 0$ give $\frac{d}{dt}|_{t=0}h(t) \cdot p = 0$.

To recover the full stabilizer from knowledge of the connected component of the identity, we have the following observation, the first part was exploited in [BGL14]:

**Proposition 6.8.4.1.** Let $V$ be an irreducible $GL(W)$-module. Let $G^0$ be the identity component of the stabilizer $G_v$ of some $v \in V$ in $GL(W)$. Then $G_v$ is contained in the normalizer of $G^0$ in $GL(W)$. If $G^0$ is semi-simple and $[v]$ is determined by $G^0$, then equality holds up to scalar multiples of the identity in $GL(W)$.

**Proof.** First note that for any group $H$, the full group $H$ normalizes $H^0$. (If $h \in H^0$, take a curve $h_t$ with $h_0 = Id$ and $h_1 = h$, then take any $g \in H$, the curve $gh_tg^{-1}$ connects $gh_1g^{-1}$ to the identity.) So $G_v$ is contained in the normalizer of $G_0$ in $GL(W)$.

For the second assertion, let $h \in N(G^0)$ be in the normalizer. We have $h^{-1}ghv = g'v = v$ for some $g' \in G^0$, and thus $g(hv) = (hv)$. But since $[v]$ is the unique line preserved by $G^0$ we conclude $hv = \lambda v$ for some $\lambda \in \mathbb{C}^*$. □

The following lemma requires more representation theory than we have used so far, but since it only appears here, readers not familiar with root systems can skip it without harm:

**Lemma 6.8.4.2.** [BGL14, Prop. 2.2] Let $G^0$ be semi-simple and act irreducibly on $V$. Then its normalizer is generated by itself, the scalar matrices, and a finite group constructed as follows: Assume we have chosen a Borel for $G^0$, and thus have distinguished a set of simple roots $\Delta$ and a group homomorphism $Aut(\Delta) \rightarrow GL(V)$. Assume $V = V_\lambda$ is the irreducible representation with highest weight $\lambda$ of $G^0$ and consider the subgroup $Aut(\Delta, \lambda) \subset Aut(\Delta)$ that fixes $\lambda$. Then $N(G^0) = ((\mathbb{C}^* \times G^0)/Z) \rtimes Aut(\Delta, \lambda)$.

For the proof, see [BGL14].

Further techniques come from geometry. Consider the hypersurface $Z(P) := \{[v] \in \mathbb{P}V^* \mid P(v) = 0\} \subset \mathbb{P}V^*$. If all the irreducible components of $P$ are reduced, then $G_{Z(P)} = G_{[P]}$, as a reduced polynomial may be recovered up to scale from its zero set, and in general $G_{Z(P)} \supseteq G_{[P]}$. Moreover, we can consider its singular set $Z(P)_{sing}$, which may be described as the zero set of the image of $P_i d - 1$ (which is essentially the exterior derivative $dP$). If $P = a_1 x^1$, where $a_1 \ldots i_d$ is symmetric in its lower indices, then $Z(P)_{sing} = \{[v] \in \mathbb{P}V^* \mid a_{\iota_1 \iota_2 \ldots \iota_d} x^{\iota_1} (v) \cdots x^{\iota_d} (v) = 0 \ \forall \iota\}$. While we could consider the singular locus of the singular locus etc., it turns out
6.8. Symmetries of polynomials and coordinate rings of orbits

To be easier to work with what I will call the Jacobian loci. For an arbitrary variety $X \subset \mathbb{P}V$, define $X_{\text{Jac},1} := \{x \in \mathbb{P}V \mid dP_x = 0 \forall P \in I(X)\}$. If $X$ is a hypersurface, then $X_{\text{Jac},1} = X_{\text{sing}}$ but in general they can be different. Define $X_{\text{Jac},k} := (X_{\text{Jac},k-1})_{\text{Jac},1}$. Algebraically, if $X = Z(P)$ for some $P \in S^dV$, then the ideal of $Z(P)_{\text{Jac},k}$ is generated by the image of $P_{k,d-k} : S^kV^* \to S^{d-k}V$. The symmetry groups of these varieties all contain $G_P$.

6.8.5. The Fermat. Let $\text{fermat}^d_n := x_1^d + \cdots + x_n^d$. The $GL(V)$-orbit closure of $[\text{fermat}^d_n]$ is the $n$-th secant variety of the Veronese variety $\sigma_n(v_d(\mathbb{P}V)) \subset \mathbb{P}S^nV$. It is clear $\mathcal{G}_n \subset G_{\text{fermat}}$, as well as the diagonal matrices whose entries are $d$-th roots of unity. We need to see if there is anything else. The first idea, to look at the singular locus, does not work, as the zero set is smooth, so we consider $\text{fermat}_{2,d-2} = x_1^2 \otimes x^{d-2} + \cdots + x_n^2 \otimes x^{d-2}$. Write the further polarization $P_{1,d-2}$ as a symmetric matrix whose entries are homogeneous polynomials of degree $d-2$ (the Hessian matrix). We get

$$
\begin{pmatrix}
x_1^{d-2} \\
\vdots \\
x_n^{d-2}
\end{pmatrix}
$$

Were the determinant of this matrix $GL(V)$-invariant, we could proceed as we did with chow$_n$, using unique factorization. Although it is not, it is close enough as follows: Recall that for a linear map $f : W \to V$, where $\dim W = \dim V = n$, we have $f^\Lambda_n \in \Lambda^n W^* \otimes \Lambda^n V$ and an element $(h, g) \in GL(W) \times GL(V)$ acts on $f^\Lambda_n$ by $(h, g) \cdot f^\Lambda_n = (\det(h))^{-1}(\det(g)) f^\Lambda_n$. In our case $W = V^*$ so $P_{2,d-2}^\Lambda n(x) = \det(g)^2 P_{2,d-2}^\Lambda n(g \cdot x)$, and the polynomial obtained by the determinant of the Hessian matrix is invariant up to scale.

Arguing as above, $(g_1^1 x_1)^{d-2} \cdots (g_n^1 x_n)^{d-2} = x_1^{d-2} \cdots x_n^{d-2}$ and we conclude again by unique factorization that $g$ is in $\Gamma_n$. Composing with a permutation matrix to make $g \in T$, we see that, by acting on the Fermat itself, that the entries on the diagonal are $d$-th roots of unity.

**Exercise 6.8.5.1:** Show that the Fermat is characterized by its symmetries.

6.8.6. The sum-product polynomial. The following polynomial, called the sum-product polynomial, will be important when studying depth-3 circuits:

$$
SP_n^m := \sum_{i=1}^{m} \Pi_{j=1}^{n} x_{ij} \in S^n(C^{nm}).
$$

Its $GL(mn)$-orbit closure is the $m$-th secant variety of the Chow variety $\sigma_m(CH_n(C^{nm}))$. 
**Exercise 6.8.6.1:** Determine $G_{SP_n^m}$ and show that $SP_n^m$ is characterized by its symmetries.

**6.8.7. Iterated matrix multiplication.** Let $IMM_k^n \in S_n^k(C^{k^2n})$ denote the iterated matrix multiplication operator for $k \times k$ matrices, $(X_1, \ldots, X_n) \mapsto \text{trace}(X_1 \cdots X_n)$. Letting $V_j = C^k$, invariantly

$IMM_k^n = \text{Id}_{V_1} \otimes \cdots \otimes \text{Id}_{V_n} \in (V_1 \otimes V_2^*) \otimes \cdots \otimes (V_{n-1} \otimes V_n^*) \otimes (V_n \otimes V_1^*)$,

and the connected component of the identity of $G_{IMM_k^n} \subset GL(C^{k^2n})$ is clear.

The case of $IMM_3^n$ is important as this sequence is complete for the complexity class $VP_c$, of sequences of polynomials admitting small formulas, see [BOC92]. Moreover $IMM_n^n$ is complete for the same complexity class as the determinant, namely $VQP$, see [Blä01].

**Problem 6.8.7.1.** Find equations in the ideal of $GL_9^n \cdot IMM_3^n$. Determine lower bounds for the inclusions $\text{Perm}^n_m \subset GL_9^n \cdot IMM_3^n$ and study common geometric properties (and differences) of $\text{Det}^n_n$ and $GL_9^n \cdot IMM_3^n$.

**Problem 6.8.7.2.** Determine $G_{IMM_3^n}$ and $G_{IMM_3^n}$. Are $IMM_3^n, IMM_n^n$ determined by their stabilizers?

**6.8.8. The permanent.** Write $C^{n^2} = E \otimes F$. Then it is easy to see $(\Gamma^E_n \times \Gamma^F_n) \rtimes \mathbb{Z}_2 \subseteq G_{\text{perm}_n}$, where the nontrivial element of $\mathbb{Z}_2$ acts by sending a matrix to its transpose and recall $\Gamma^E_n = T^{SL}_{E} \rtimes \mathcal{S}_n$. We would like to show this is the entire symmetry group. However, it is not when $n = 2$.

**Exercise 6.8.8.1:** What is $G_{\text{perm}_2}$?

**Theorem 6.8.8.2.** [MM62] For $n \geq 3$, $G_{\text{perm}_n} = (\Gamma^E_n \times \Gamma^F_n)/\mu_n \rtimes \mathbb{Z}_2$.

Consider $Z(\text{perm}_n)_{\text{sing}} \subset \mathcal{P}(E \otimes F)^*$. It consists of the matrices all of whose size $n-1$ submatrices have zero permanent. (To see this, note the permanent has Laplace type expansions.) This seems even more complicated than the hypersurface $Z(\text{perm}_n)$ itself. Continuing, $Z(\text{perm}_n)_{\text{jac},k}$ consists of the matrices all of whose sub-matrices of size $n-k$ have zero permanent. In particular $Z(\text{perm}_n)_{\text{jac},n-2}$ is defined by quadratic equations. Its zero set has many components, but each component is easy to describe:

**Lemma 6.8.8.3.** Let $A$ be an $n \times n$ matrix all of whose size 2 submatrices have zero permanent. Then one of the following hold:

1. all the entries of $A$ are zero except those in a single size 2 submatrix, and that submatrix has zero permanent.
2. all the entries of $A$ are zero except those in the $j$-th row for some $j$. Call the associated component $C^j$. 


(3) all the entries of \( A \) are zero except those in the \( j \)-th column for some \( j \). Call the associated component \( C_j \).

The proof is straight-forward. Take a matrix with entries that don’t fit that pattern, e.g., one that begins
\[
\begin{array}{ccc}
  a & b & e \\
  * & d & *
\end{array}
\]
and note that it is not possible to fill in the two unknown entries and have all size two sub-permanents, even in this corner, zero. There are just a few such cases since we are free to act by \( S_n \times S_n \).

**Proof of theorem 6.8.8.2.** I follow \([Ye11]\). Any linear transformation preserving the permanent must send a component of \( Z(\text{perm}_n)_{\text{Jac,}n-2} \) of type (1) to another of type (1). It must send a component \( C^j \) either to some \( C^k \) or some \( C_i \). But if \( i \neq j \), \( C^j \cap C^i = 0 \) and for all \( i, j \), \( \dim(C^i \cap C^j) = 1 \).

Since intersections must be mapped to intersections, either all components \( C^i \) are sent to components \( C_k \) or all are permuted among themselves. By composing with an element of \( \mathbb{Z}_2 \), we may assume all the \( C^i \)’s are sent to \( C^i \)’s and the \( C^j \)’s are sent to \( C^j \)’s. Similarly, by composing with an element of \( S_n \times S_n \) we may assume each \( C_i \) and \( C^j \) is sent to itself. But then their intersections are sent to themselves. So we have, for all \( i, j \),
\[
(6.8.2) \quad (x^i_j) \mapsto (\lambda^i_j x^i_j)
\]
for some \( \lambda^i_j \) and there is no summation in the expression. Consider the image of a size 2 submatrix, e.g.,
\[
(6.8.3) \quad 
\begin{array}{ccc}
  x^1_1 & x^1_2 \\
  x^2_1 & x^2_2
\end{array} \mapsto 
\begin{array}{ccc}
  \lambda^1_1 x^1_1 & \lambda^1_2 x^1_2 \\
  \lambda^2_1 x^2_1 & \lambda^2_2 x^2_2
\end{array}
\]
In order that the map (6.8.2) be in \( G_{\text{perm}_n} \), when \( (x^i_j) \in Z(\text{perm}_n)_{\text{Jac,}n-2} \), the permanent of the matrix on the right hand side of (6.8.3) must be zero. The permanent of the right hand side of (6.8.3) when \( (x^i_j) \in Z(\text{perm}_n)_{\text{Jac,}n-2} \) is
\[
\lambda^1_1 \lambda^2_2 x^1_1 x^2_2 + \lambda^1_2 \lambda^2_1 x^2_1 x^1_2 = x^1_1 x^2_2 (\lambda^1_1 \lambda^2_2 - \lambda^2_1 \lambda^1_2)
\]
which implies \( \lambda^1_1 \lambda^2_2 - \lambda^2_1 \lambda^1_2 = 0 \), thus all the \( 2 \times 2 \) minors of the matrix \( (\lambda^i_j) \) are zero, so it has rank one and is the product of a column vector and a row vector, but then it is an element of \( T_E \times T_F \).

Because the stabilizer of the permanent is so small, the coordinate ring of its orbit will be much larger than that of the determinant. An analysis of the coordinate ring of the orbits of the permanent and padded permanent is given in \([BLMW11]\).
6.8.9. The Pascal determinant. Let $k$ be even, and let $A_j = \mathbb{C}^n$. Define the $k$-factor Pascal determinant $PD_{k,n}$ to be the unique up to scale element of $\Lambda^n A_1 \otimes \cdots \otimes \Lambda^n A_k \subset S^n(A_1 \otimes \cdots \otimes A_k)$. Choose the scale such that if $X = \sum x_{i_1 \ldots i_k} a_{1,i_1} \otimes \cdots \otimes a_{k,i_k}$ with $a_{\alpha,j}$ a basis of $A_\alpha$, then

$$PD_{k,n}(X) = \sum_{\sigma_2, \ldots, \sigma_k \in S_n} \text{sgn}(\sigma_2 \cdots \sigma_k)x_{1,\sigma_2(1),\ldots,\sigma_2(n)} \cdots x_{n,\sigma_2(n),\ldots,\sigma_k(n)}$$

By this expression we see, fixing $k$, that $(PD_{k,n}) \in \text{VNP}$.

Proposition 6.8.9.1 (Gurvits). The sequence $(PD_{4,n})$ is $\text{VNP}$ complete.

Proof. Set $x_{ijkl} = 0$ unless $i = j$ and $k = l$. Then $x_{i,\sigma_2(i),\sigma_3(i),\sigma_4(i)} = 0$ unless $\sigma_2(i) = i$ and $\sigma_3(i) = \sigma_4(i)$ so the only nonzero monomials are those where $\sigma_2 = \text{Id}$ and $\sigma_3 = \sigma_4$, since the sign of $\sigma_3$ is squared, the result is the permanent. Thus we could just as well work with the sequence $PD_{4,n}$ as the permanent. Initially this is appealing, as its stabilizer resembles that of the determinant, and, after all, $PD_{2,n} = \text{det}_n$.

It is clear the identity component of the stabilizer includes $SL_n \times \cdots \times SL_k / \mu_{n,k}$ where $\mu_n$ is as in §6.2.4, and a straight-forward Lie algebra calculation confirms this is the entire identity component. (Alternatively, one can use Dynkin’s classification [Dyn52] of maximal subalgebras.) It is also clear that $S_k$ preserves $PD_{n,k}$ by permuting the factors.

Theorem 6.8.9.2 (Garibaldi, personal communication). For all $k$ even

$$G_{PD_{k,n}} = SL_n \times \cdots \times SL_k / \mu_{n,k} \rtimes S_k$$

Note that this includes the case of the determinant, and gives a new proof.

It follows from

Lemma 6.8.9.3. [Garibaldi, personal communication] Let $V = A_1 \otimes \cdots \otimes A_k$. The normalizer of $SL_n \times \cdots \times SL_k / \mu_n$ in $GL(V)$ is $GL_n \times / Z \rtimes S_k$, where $Z$ denotes the kernel of the product map $(\mathbb{C}^*)^k \rightarrow \mathbb{C}^*$.

Proof of Lemma 6.8.9.3. We use Lemma 6.8.4.2. In our case, the Dynkin diagram for $(\Delta, \lambda)$ is

and $Aut(\Delta, \lambda)$ is clearly $S_k$. The theorem follows.
Figure 6.8.1. Marked Dynkin diagram for $V$
This chapter discusses approaches to GCT that do not involve the determinant.

In §7.1, I explain results from [GKKS13b] and their consequences for geometry. The result is that to show $\text{VP} \neq \text{VNP}$, one can show the padded permanent is not in the orbit closure of either of two relatively nice varieties. The price is that instead of a polynomial separation, one needs a much more severe separation.

In [GKKS13a], they utilize the method of shifted partial derivatives, which is a special case of Young-flattenings. I discuss this method and its relationship to Hilbert functions of Jacobian ideals in §7.2.

One of the two varieties mentioned above is $r$-th secant variety of the Chow variety of products of linear forms. The Chow variety itself has been studied for over 100 years, and in §7.3 I give a history and bring the reader up to the state of the art.

I believe the Chow variety is important for complexity theory for several reasons: the study of depth three circuits, as a “toy” orbit closure problem, and since it is contained in the orbit closures central to GCT, information about its coordinate ring will be useful for the study of $\text{Det}_n$, the orbit closure of the determinant. It also has a fascinating history, and deep connections to problems in combinatorics, representation theory and algebraic geometry.
7. Shallow circuits and the Chow variety

7.1. Shallow circuits and geometry

7.1.1. Introduction. The depth of a circuit $C$ is the length of (i.e., the number of edges in) the longest path in $C$ from an input to its output. If a circuit has small depth, it is called a shallow circuit, and the polynomial it computes can be computed quickly in parallel. When one studies circuits of bounded depth, one must allow gates to have an arbitrary number of edges coming in to them ("unbounded fanin"). For such circuits, multiplication by constants is considered "free."

**picture here**

The main theoretical interest in shallow circuits for separating complexity classes is that there are depth reduction theorems described in §7.1.2 that enable one substitute the problem of e.g., showing that there does not exist a small circuit computing the permanent to the problem of showing that there does not exist a "slightly less small" shallow circuit computing the permanent. Two classes of shallow circuits have very nice algebraic varieties associated to them: the depth three or $\Sigma\Pi\Sigma$ circuits, which consist of depth three formulas where the first layer of gates consist of additions, the second of multiplications, and the last gate is an addition gate, and the $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuits, which are special depth five circuits, where the first layer of gates are additions, the second layer consists of "powering gates", where a powering gate takes $f$ to $f^\delta$ for some natural number $\delta$, the third layer addition gates, the fourth layer again powering gates, and the fifth layer is an addition gate. I describe the associated varieties to these classes of circuits in §7.1.3 and §7.1.4.

One can restrict one’s class of circuits further by requiring that they are homogeneous in the sense that each gate computes a homogeneous polynomial. It turns out that for $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuits, this is not restrictive for the questions of interest, but for $\Sigma\Pi\Sigma$ circuits, there is a tremendous loss of computing power described in §7.1.4. Fortunately this loss of computing power can be overcome by something we are already familiar with: computing padded polynomials.

7.1.2. Depth reduction theorems. A major result in the study of shallow circuits was [VSBR83], where it was shown that if a polynomial of degree $d$ can be computed by a circuit of size $s$, then it can be computed by a circuit of depth $O(\log \log s)$ and size polynomial in $s$.

Here are the relevant results relevant for our discussion. They combine results of [Bre74, GKK13b, Tav, Koi, AV08]:

**Theorem 7.1.2.1.** Let $d$ be a polynomial in $n$ and let $P_n \in S^d \mathbb{C}^n$ be a sequence of polynomials that can be computed by a circuit of size $s = s(n)$. 

Then:

1. $P$ is computable by a $\Sigma\Pi\Sigma$ circuit of size roughly $s\sqrt{d}$, more precisely of size $2^{O(\sqrt{\log(n)\log(ds)})}$.

2. $P$ is computable, by a homogeneous $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuit of size roughly $s\sqrt{d}$, more precisely of size $2^{O(\sqrt{d}\log(ds)\log(n))}$, and both powering gates of size of roughly $\sqrt{d}$.

Here are ideas towards the proof: In [GKKS13b] they prove upper bounds for the size of an inhomogeneous depth three circuit computing a polynomial, in terms of the size of an arbitrary circuit computing the polynomial. They first apply the work of [Koi, AV08], which allows one to reduce an arbitrary circuit of size $s$ computing a polynomial of degree $d$ in $n$ variables to a formula of size $2^{O(\log s \log d)}$ and depth $d$. Next they reduce to a depth four circuit of size $s' = 2^{O(\sqrt{d} \log d \log s \log n)}$. This second passage is via iterated matrix multiplication. From the depth four circuit, they use Equation (7.1.1) to convert all multiplication gates to $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuits to have a depth five circuit of size $O(s')$ and of the form $\Sigma\Lambda\Sigma\Lambda\Sigma$. Finally, they convert the power sums to elementary symmetric functions which keeps the size at $O(s')$ and drops the depth to three.

7.1.3. Geometric reformulation of homogeneous $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuits.

Recall that computer scientists always work in bases and the inputs to the circuits are constants and variables. For homogeneous circuits, the inputs are simply the variables. The first layer of such a circuit is just to obtain arbitrary linear forms from these variables, so it plays no role in the geometry. The second layer sends a linear form $\ell$ to $\ell^\delta$, i.e., we are forming points of $v^\delta(\mathbb{P}V)$. The next layer consists of addition gates, which means we obtain sums of $d$-th powers, i.e., points of $\sigma_\tau(v^\delta(\mathbb{P}V))$. Then at the next layer, we take Veronese re-embeddings of these secant varieties to obtain points of $v^\delta\sigma_\tau(v^\delta(\mathbb{P}V)))$, and in the final addition gate we obtain a point of $\sigma_\tau(v^\delta\sigma_\tau(v^\delta(\mathbb{P}V)))$. Thus we may rephrase Theorem 7.1.2.1(2) as:

**Proposition 7.1.3.1.** [Lan14a] Let $d = n^{O(1)}$ and let $P \in S^d \mathbb{C}^n$ be a polynomial sequence that can be computed by a circuit of size $s$. Then $[P] \in \sigma_{r_1}(v^\delta_{\sigma_2}(v^\delta_{\mathbb{P}(\mathbb{P}^{m-1})})))$ with roughly $\delta \sim \sqrt{d}$ and $r_1 r_2 \sim s\sqrt{d}$, more precisely $r_1 r_2 \delta = 2^{O(\sqrt{\log(ds)\log(n)})}$.

**Corollary 7.1.3.2.** [GKKS13b] If for all but finitely many $m$, $\delta \simeq \sqrt{m}$, and all $r_1, r_2$ such that $r_1 r_2 = 2^{\sqrt{\log(m)\omega(1)}}$, one has $[\text{perm}_m] \notin \sigma_{r_1}(v^\delta_{\sigma_2}(v^\delta_{\mathbb{P}^{m-1}})))$, then there is no circuit of polynomial size computing the permanent, i.e., $\text{VP} \neq \text{VNP}$.

**Problem 7.1.3.3.** Find equations for $\sigma_{r_1}(v^\delta_{\sigma_2}(v^\delta_{\mathbb{P}^{m-1}})))$. 
7. Shallow circuits and the Chow variety

7.1.4. Geometric reformulation of \( \Sigma \Pi \Sigma \) circuits. The homogeneous \( \Sigma \Pi \Sigma \) circuits are easily seen to be modeled by \( \sigma_r(Ch_n(V)) \), but Theorem 7.1.2.1(1) requires inhomogeneous \( \Sigma \Pi \Sigma \) circuits. One might think that for a small price, one could restrict to homogeneous \( \Sigma \Pi \Sigma \) circuits. E.g., does \( \det_n \) admit a homogeneous \( \Sigma \Pi \Sigma \) circuit of size \( \sim (n^4)^{\sqrt{\pi}} \sim 2^{\pi \log n} \), i.e., is \( [\det_n] \in \sigma_{2,\pi \log n}(Ch_n(\mathbb{C}^{n^2})) \)? The answer is no:

**Proposition 7.1.4.1.** \( \det_n \notin \sigma_{\frac{2n}{\pi}}(Ch_n(\mathbb{C}^{n^2})) \).

**Proof.** Consider the flattenings:

\[
(det_n)^{\lceil \frac{n}{2} \rceil, \lceil \frac{n}{2} \rceil} : S^{\lceil \frac{n}{2} \rceil} W^* \to S^{\lceil \frac{n}{2} \rceil} W
\]

and

\[
(x_1 \cdots x_n)^{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor} : S^{\lfloor \frac{n}{2} \rfloor} W^* \to S^{\lfloor \frac{n}{2} \rfloor} W.
\]

In the first case the image is spanned by the minors of size \( \lfloor \frac{n}{2} \rfloor \), so it is of dimension \( (\binom{n}{\lfloor \frac{n}{2} \rfloor})^2 \), whereas in the second case the image is spanned by the monomials \( x_{i_1} \cdots x_{i_{\lfloor \frac{n}{2} \rfloor}} \) for \( I \subset \{1, \ldots, n\} \), so it has dimension \( (\binom{n}{\lfloor \frac{n}{2} \rfloor}) \). Thus

\[
\det_n \notin \sigma_{\binom{n}{\lceil \frac{n}{2} \rceil}}(Ch_n(\mathbb{C}^{n^2})).
\]

Recalling that \( (\frac{2m}{m}) \sim \frac{4^m}{\sqrt{m}} \), we conclude. \( \square \)

**Remark 7.1.4.2.** The flattening above gives a lower bound on the symmetric tensor border rank of points of the Chow variety. The symmetric rank is known: On one hand,

\[
(7.1.1) \quad x_1 \cdots x_n = \frac{1}{2(n-1)!} \sum_{\epsilon \in \{-1,1\}^{n-1}} (x_1 + \epsilon_1 x_2 + \cdots + \epsilon_{n-1} x_n)^n \epsilon_1 \cdots \epsilon_{n-1},
\]

so \( R_S(x_1 \cdots x_n) \leq 2^{n-1} \). (This expression dates at least back to \([Fis94]\).) On the other hand, in \([RS11]\) they showed

\[
(7.1.2) \quad R_S(x_1 \cdots x_n) = 2^{n-1}.
\]

**Exercise 7.1.4.3:** Show that \( [\text{perm}_n] \notin \sigma_{O(\frac{4^n}{\pi}) v_n(\mathbb{P} W)} \).

In summary:

**Proposition 7.1.4.4.** \([NW97]\) The polynomial sequences \( \det_n \) and \( \text{perm}_n \) do not admit depth three circuits of size \( 2^n \).

**Remark 7.1.4.5.** In \([NW97]\) they consider all partial derivatives of all orders simultaneously, but the bulk of the dimension is concentrated in the middle order flattening.
Thus homogeneous depth three circuits at first sight do not seem that powerful because a polynomial sized homogeneous depth 3 circuit cannot compute the determinant.

To make matters worse, consider the polynomial corresponding to iterated matrix multiplication of three by three matrices $IMM^k_3 \in S^k(\mathbb{C}^{3^k})$. It is in $VP$, i.e., it admits a polynomial size formula.

**Exercise 7.1.4.6:** Show that $IMM^k_3 \not\in \sigma_{poly}(Ch_k(W))$.

By Exercise 7.1.4.6, homogeneous depth three circuits (naïvely applied) cannot even capture sequences of polynomials admitting small formulas.

Fortunately this problem is easy to fix. If at the first level, we add a homogenizing variable $\ell$, so that the affine linear outputs become linear in our original variables plus $\ell$, the product gates will each produce a homogeneous polynomial. While the different product gates may produce polynomials of different degrees, if we were trying to produce a homogeneous polynomial, when we add them up what remains must be a sum of homogeneous polynomials, such that when we set $\ell = 1$, we obtain the desired homogeneous polynomial. Say the largest power of $\ell$ appearing in this sum is $q_L$. For each other term there is some other power of $\ell$ appearing, say $q_i$ for the $i$-th term. Then to the original circuit, add $q_L - q_i$ inputs to the $i$-th product gate, where each input is $\ell$. This will not change the size of the circuit by more than $q_L r$, so it remains of the same order. Our new homogeneous depth three circuit will output $\ell^{q_L} P$.

We conclude:

**Proposition 7.1.4.7.** [Lan14a] Let $d = n^{O(1)}$ and let $P \in S^d \mathbb{C}^n$ be a polynomial that can be computed by a circuit of size $s$.

Then $[\ell^{N-d}P] \in \sigma_r(Ch_N(\mathbb{C}^{n+1}))$ with roughly $r N \sim s \sqrt{d}$, more precisely, $r N = 2^{O(\sqrt{d \log(n) \log(ds)})}$.

**Corollary 7.1.4.8.** [GKKS13b] $[\ell^{n-m} \det m] \in \sigma_r(Ch_n(\mathbb{C}^{m^2+1}))$ where $r n = 2^{O(\sqrt{m \log m})}$.

**Proof.** The determinant admits a circuit of size $m^4$, so it admits a $\Sigma \Pi \Sigma$ circuit of size $2^{O(\sqrt{m \log(m) \log(m^2+1)})} \sim 2^{O(\sqrt{m \log m})}$, so its padded version lies in $\sigma_r(Ch_n(\mathbb{C}^{m^2+1}))$ where $r n = 2^{O(\sqrt{m \log m})}$.

**Corollary 7.1.4.9.** [GKKS13b] If for all but finitely many $m$ and all $r, n$ with $r n = 2^{\sqrt{m \log(m) \omega(1)}}$, one has $[\ell^{n-m} \text{perm}_m] \not\in \sigma_r(Ch_n(\mathbb{C}^{n^2+1}))$, then there is no circuit of polynomial size computing the permanent, i.e., $VP \neq VNP$. 


Proof. One just needs to observe that the number of edges in the first layer (which are invisible from the geometric perspective) is dominated by the number of edges in the other layers. 

□

Problem 7.1.4.10. [CKW10, Open problem 11.1] Find an explicit sequence of polynomials $P_m \in S^m \mathbb{C}^{w-1}$ such that for $m$ sufficiently large $\ell^{n-m} P_m \notin \sigma_r(Ch_n(W))$, whenever $r, w, n$ are polynomials in $m$.

Remark 7.1.4.11. The expected dimension of $\sigma_r(Ch_d(W))$ is $rdw + r - 1$. If we take $d' = d2^m$ and work instead with padded polynomials $\ell2^m P$, the expected dimension of $\sigma_r(Ch_d(W))$ is $2^r rdw + r - 1$. In contrast, the expected dimension of $\sigma_r(v_{d-a}(\sigma_p(v_{d}(FW))))$ does not change when one increases the degree, which gives some insight as to why padding is so useful for homogeneous depth three circuits but not for $\Sigma\Lambda\Sigma\Lambda\Sigma$ circuits.

7.1.5. Elementary symmetric functions and depth 3 circuits. **This section under construction***

7.2. Hilbert functions and the method of shifted partial derivatives

**This section to be written***

7.3. The Chow variety

Motivation for studying the ideal of the Chow variety was given above. This ideal has been studied for over 100 years, dating back at least to Brill, Gordan and Hadamard. The history is rife with rediscoveries and errors that only make the subject more intriguing.

7.3.1. The Hermite-Hadamard-Howe map and the ideal of the Chow variety. The following linear map was first defined when dim $W = 2$ by Hermite (1854), and in general independently by Hadamard (1897) and Howe (1988).

Definition 7.3.1.1. The Hermite-Hadamard-Howe map $h_{d,n} : S^d(S^nW) \rightarrow S^n(S^dW)$ is defined as follows: First include $S^d(S^nW) \subset W^\otimes nd$. Next, reorder the copies of $W$ from $d$ blocks of $n$ to $n$ blocks of $d$ and symmetrize the blocks of $d$ to obtain an element of $(S^dW)^\otimes n$. Finally, thinking of $S^dW$ as a single vector space, symmetrize the $n$ blocks.
For example, putting subscripts on \( W \) to indicate position:

\[
S^2(S^3W) \subset W^\otimes 6 = W_1 \otimes W_2 \otimes W_3 \otimes W_4 \otimes W_5 \otimes W_6
\]
\[
\rightarrow (W_1 \otimes W_4) \otimes (W_2 \otimes W_5) \otimes (W_3 \otimes W_6)
\]
\[
\rightarrow S^2W \otimes S^2W \otimes S^2W
\]
\[
\rightarrow S^3(S^2W)
\]

Note that \( h_{d,n} \) is a \( GL(W) \)-module map.

**Example 7.3.1.2.** Here is \( h_{2,2}((xy)^2) \):

\[
(xy)^2 \mapsto \frac{1}{4} [(x \otimes y + y \otimes x) \otimes (x \otimes y + y \otimes x)]
\]
\[
= \frac{1}{4} [x \otimes y \otimes x \otimes y + x \otimes y \otimes y \otimes x + y \otimes x \otimes x \otimes y + y \otimes x \otimes y \otimes x]
\]
\[
\mapsto \frac{1}{4} [(x \otimes x \otimes y \otimes y + x \otimes y \otimes y \otimes x + y \otimes x \otimes x \otimes y + y \otimes y \otimes x \otimes x]
\]
\[
\mapsto \frac{1}{4} [2(x^2 \otimes (y^2) + 2(xy) \otimes (xy)]
\]
\[
\mapsto \frac{1}{2} [(x^2)(y^2) + (xy)(xy)]).
\]

**Exercise 7.3.1.3:** Show that \( h_{d,n}(x_1^n \cdots x_d^n) = (x_1 \cdots x_d)^n \).

**Exercise 7.3.1.4:** Show that \( h_{d,n} : S^d(S^n V) \to S^{n}(S^dV) \) is “self-dual” in the sense that \( h_{d,n}^T = h_{n,d} : S^n(S^d V^*) \to S^d(S^n V^*) \). Conclude that \( h_{d,n} \) surjective if and only if \( h_{n,d} \) is injective.

**Theorem 7.3.1.5 (Hadamard [Had97]).** \( \ker h_{d,n} = I_d(Ch_n(W^*)) \).

**Proof.** Let \( P \in S^d(S^n W) \). Write \( P = \sum x_{1j}^n \cdots x_{dj}^n \) for some \( x_{\alpha,j} \in W \). Let \( \ell^1, \ldots, \ell^n \in W^* \).

\[
P(\ell^1 \cdots \ell^n) = \langle P, (\ell^1 \cdots \ell^n)^d \rangle
\]
\[
= \sum_j \langle x_{1j}^n \cdots x_{dj}^n, (\ell^1 \cdots \ell^n)^d \rangle
\]
\[
= \sum_j \langle x_{1j}^n, (\ell^1 \cdots \ell^n) \rangle \cdots \langle x_{dj}^n, (\ell^1 \cdots \ell^n) \rangle
\]
\[
= \sum_j \Pi_{s=1}^n \Pi_{i=1}^d x_{ij}(\ell_s)
\]
\[
= \sum_j \langle x_{1j} \cdots x_{dj}, (\ell^1)^d \rangle \cdots \langle x_{1j} \cdots x_{dj}, (\ell^n)^d \rangle
\]
\[
= \langle h_{d,n}(P), (\ell^1)^d \cdots (\ell^n)^d \rangle
\]
If $h_{d,n}(P)$ is nonzero, there will be some monomial it will pair with to be nonzero. On the other hand, if $h_{d,n}(P) = 0$, then $P$ annihilates all points of $Ch_n(W^*)$. □

**Exercise 7.3.1.6:** Show that if $h_{d,n} : S^d(S^n \mathbb{C}^m) \to S^n(S^d \mathbb{C}^m)$ is not surjective, then $h_{d,n} : S^d(S^n \mathbb{C}^k) \to S^n(S^d \mathbb{C}^k)$ is not surjective for all $k > m$, and that the partitions describing the kernel are the same in both cases if $d \leq m$.

**Exercise 7.3.1.7:** Show that if $h_{d,n} : S^d(S^n \mathbb{C}^m) \to S^n(S^d \mathbb{C}^m)$ is surjective, then $h_{d,n} : S^d(S^n \mathbb{C}^k) \to S^n(S^d \mathbb{C}^k)$ is surjective for all $k < m$.

**Example 7.3.1.8** (The case $\dim W = 2$). When $\dim W = 2$, every polynomial decomposes as a product of linear factors, so the ideal of $Ch_n(\mathbb{C}^2)$ is zero. We recover the following theorem of Hermite:

**Theorem 7.3.1.9** (Hermite reciprocity). The map $h_{d,n} : S^d(S^n \mathbb{C}^2) \to S^n(S^d \mathbb{C}^2)$ is an isomorphism for all $d, n$. In particular $S^d(S^n \mathbb{C}^2)$ and $S^n(S^d \mathbb{C}^2)$ are isomorphic $GL_2$-modules.

Often in modern textbooks only the “In particular” is stated.

**Theorem 7.3.1.10** ([Hadamard [Had99]]). The map $h_{3,3} : S^3(S^3 \mathbb{C}^3) \to S^3(S^3 \mathbb{C}^3)$ is an isomorphism.

**Proof.** Without loss of generality, assume $w = 3$ and $x_1, x_2, x_3 \in W^*$ are a basis. Say we had $P \in I_3(Ch_3(W^*))$. Consider $P(\mu(x_1^3 + x_2^3 + x_3^3) - \lambda x_1 x_2 x_3)$ as a cubic polynomial on $\mathbb{P}^1$ with coordinates $[\mu, \lambda]$. Note that it vanishes at the four points $[0, 1], [1, 3], [1, 3\omega], [1, 3\omega^2]$ where $\omega$ is a primitive third root of unity. A cubic polynomial on $\mathbb{P}^1$ vanishing at four points is identically zero. In particular, $P(1, 0) = 0$, i.e., $P$ vanishes on $P(x_1^3 + x_2^3 + x_3^3)$. Hence it must vanish identically on $\sigma_3(v_3(\mathbb{P}^2))$. But $I_3(\sigma_3(v_3(\mathbb{P}^2))) = 0$, see, e.g., Corollary ?? (in fact $\sigma_3(v_3(\mathbb{P}^2)) \subset \mathbb{P}S^3 \mathbb{C}^3$ is a hypersurface of degree four). □

**Remark 7.3.1.11.** The above proof is due to A. Abdesselam (personal communication). It is a variant of Hadamard’s original proof, where instead of $x_1^3 + x_2^3 + x_3^3$ one uses an arbitrary cubic $f$, and generalizing $x_1 x_2 x_3$ one uses the Hessian $H(f)$. Then the curves $f = 0$ and $H(f) = 0$ intersect in 9 points (the nine flexes of $f = 0$) and there are four groups of three lines going through these points, i.e. four places where the polynomial becomes a product of linear forms.

**Theorem 7.3.1.12.** ([BL89]) If $h_{d,n}$ is surjective, then $h_{d,n}$ is surjective for all $d' > d$. In other words, if $h_{n,d}$ is injective, then $h_{n,d'}$ is injective for all $d' > d$.

**Proof.** **proof to be inserted*** □
Remark 7.3.1.13. The statements and proofs in [BL89, McK08] were regarding the map $h_{d,n}$ defined in §7.3.2 below.

Originally Hadamard mistakenly thought the maps $h_{d,n}$ were always of maximal rank, but in [Hadamard99] he proved the map is an isomorphism when $d = n = 3$ and posed determining if injectivity holds in general when $d \leq n$ as an open problem. (Injectivity for $d \leq n$ is equivalent to surjectivity when $d \geq n$ by Exercise 7.3.1.4.) Howe [Howe87] also investigated the map $h_{d,n}$ and wrote “it is reasonable to expect” that $h_{d,n}$ is always of maximal rank.

Theorem 7.3.1.14. [MN05] The map $h_{5,5}$ is not surjective.

Remark 7.3.1.15. In [MN05] they showed the map $h_{5,5,0}$ defined in §7.3.2 below was not injective. A. Abdessalem realized their computation showed the map $h_{5,5}$ is not injective and pointed this out to them. Evidently there was some miscommunication because in [MN05] they mistakenly say the result comes from [Bri02] rather than their own paper.

The $GL(V)$-module structure of the kernel of $h_{5,5}$ was determined by C. Ikenmeyer and S. Mrktchyan as part of a 2012 AMS MRC program:

**Proposition 7.3.1.16** (Ikenmeyer and Mkrtchyan, unpublished). The kernel of $h_{5,5} : S^5(S^5\mathbb{C}^5) \rightarrow S^5(S^5\mathbb{C}^5)$ consists of irreducible modules corresponding to the following partitions:

$$
\{(14, 7, 2, 2), (13, 7, 2, 2, 1), (12, 7, 3, 2, 1), (12, 6, 3, 2, 2), (12, 5, 4, 3, 1), (11, 5, 4, 4, 1), (10, 8, 4, 2, 1), (9, 7, 6, 3)\}.
$$

All these occur with multiplicity one in the kernel, but not all occur with multiplicity one in $S^5(S^5\mathbb{C}^5)$. In particular, the kernel is not an isotypic component.

The Young diagrams of the kernel of $h_{5,5}$ are: 

![Young diagrams](image-url)
The phrase “with high probability” means the result was obtained numerically, not symbolically.

While the Hermite-Hadamard-Howe map is not always of maximal rank, it is “eventually” of maximal rank:

**Theorem 7.3.1.17.** [Bri93] The Hermite-Hadamard-Howe map

\[ h_{d,n} : S^d(S^nW^*) \to S^n(S^dW^*) \]

is surjective for \( d \) sufficiently large.

I present the proof of Theorem 7.3.1.17 in §8.1.1.

Moreover, in [Bri97] Brion gives an explicit exponential bound on the size of \( d \) required with respect to \( n \) that assures surjectivity, see Equation (8.1.3).

**Problem 7.3.1.18** (The Hadamard-Howe Problem). Determine the function \( d(n) \) such that \( h_{d,n} \) is surjective for all \( d \geq d(n) \).

A more ambitious problem would be:

**Problem 7.3.1.19.** Determine the kernel of \( h_{d,n} \).

The following special case of the problem would prove Conjecture 6.2.1.3.

**Conjecture 7.3.1.20** (Kumar [Kum]). Let \( n \) be even, then for all \( d \leq n \), \( S_{n^d}\mathbb{C}^n \not\subset \ker h_{d,n} \), i.e., \( S_{n^d}\mathbb{C}^n \subset \mathbb{C}[Ch_n(\mathbb{C}^n)] \).

From Exercise ??, the trivial \( SL_n \)-module \( S_n^a\mathbb{C}^n \) occurs in \( S^n(S^n\mathbb{C}^n) \) with multiplicity one when \( n \) is even and zero when \( n \) is odd.

It is not hard to see that the \( d = n \) case implies the others. Adopt the notation that if \( \pi = (p_1, \ldots, p_k) \), then \( m\pi = (mp_1, \ldots, mp_k) \). By taking Cartan products in the coordinate ring, Conjecture 7.3.1.20 would imply Conjecture 6.2.1.3.

**7.3.2. **\( S_{dn} \)-formulation of the Hadamard-Howe problem.** The dimension of \( V \), as long as it is at least \( d \), is irrelevant for the \( GL(V) \)-module structure of the kernel of \( h_{d,n} \). In this section assume \( \dim V = dn \). Fix a basis of \( V \). Then the Weyl group \( W_V \) of \( GL(V) \) is the subgroup of \( GL(V) \) consisting of the permutation matrices (in particular, it is isomorphic to \( S_{dn} \)). It acts on \( V^{\otimes dn} \) by acting on each factor. (I write \( W_V \-
7.3. The Chow variety

To distinguish this from the $\mathfrak{S}_{dn}$-action permuting the factors.) An element $x \in V^{\otimes dn}$ has $\mathfrak{sl}(V)$-weight zero if, in the standard basis $\{e_i\}_{1 \leq i \leq dn}$ of $V$ induced from the identification $V \cong \mathbb{C}^{nd}$, $x$ is a sum of monomials $x = \sum_{I=(i_1, \ldots, i_{nd})} I^{e_i_1 \otimes \cdots \otimes e_{i_{nd}}}$, where $I$ runs over the orderings of $[nd]$. If one restricts $h_{d,n}$ to the $\mathfrak{sl}(V)$-weight zero subspace, one obtains a $\mathcal{W}_V$-module map

\begin{equation}
(7.3.1) \quad h_{d,n,0} : S^d(S^n V)_0 \to S^n(S^d V)_0.
\end{equation}

These $\mathcal{W}_V$-modules are as follows: Let $\mathfrak{S}_n \wr \mathfrak{S}_d \subset \mathfrak{S}_{dn}$ denote the wreath product, which, by definition, is the normalizer of $\mathfrak{S}_d \times \mathfrak{S}_{dn}$ in $\mathfrak{S}_{dn}$. It is the semi-direct product of $\mathfrak{S}_d \times \mathfrak{S}_{dn}$ with $\mathfrak{S}_d$, where $\mathfrak{S}_d$ acts by permuting the factors of $\mathfrak{S}_d \times \mathfrak{S}_{dn}$, see e.g., [Mac95, p 158]. Since $\dim V = dn$, $S^d(S^n V)_0 = \text{Ind}_{\mathfrak{S}_n \wr \mathfrak{S}_d}^{\mathfrak{S}_{dn}} \text{triv}$, where triv denotes the trivial $\mathfrak{S}_n \wr \mathfrak{S}_d$-module as explained in §6.7.2.

In other words, as a $\mathcal{W}_V = \mathfrak{S}_{dn}$-module map, (7.3.1) is

\begin{equation}
(7.3.2) \quad h_{d,n,0} : \text{Ind}_{\mathfrak{S}_n \wr \mathfrak{S}_d}^{\mathfrak{S}_{dn}} \text{triv} \to \text{Ind}_{\mathfrak{S}_d \mathfrak{S}_n}^{\mathfrak{S}_{dn}} \text{triv}.
\end{equation}

Moreover, since every irreducible module appearing in $S^d(S^n V)$ has a non-zero $SL(V)$-weight zero subspace, $h_{d,n}$ is the unique $SL(V)$-module extension of $h_{d,n,0}$.

The map $h_{d,n,0}$ was defined purely in terms of combinatorics in [BL89] as a path to try to prove the following conjecture of Foulkes:

**Conjecture 7.3.2.1.** [Fou50] Let $d > n$, let $\pi$ be a partition of $dn$ and let $[\pi]$ denote the corresponding $\mathfrak{S}_{dn}$-module. Then,

\[
\text{mult}( [\pi], \text{Ind}_{\mathfrak{S}_n \wr \mathfrak{S}_d}^{\mathfrak{S}_{dn}} \text{triv} ) \geq \text{mult}( [\pi], \text{Ind}_{\mathfrak{S}_d \wr \mathfrak{S}_n}^{\mathfrak{S}_{dn}} \text{triv} ) .
\]

Equivalently,

\[
\text{mult}( S_{\pi} V, S^d(S^n V) ) \geq \text{mult}( S_{\pi} V, S^n(S^d V) ) .
\]

Conjecture 7.3.2.1, and in fact equality was shown to hold asymptotically by L. Manivel in [Man98], in the sense that for any partition $\mu$, the multiplicity of the partition $(dn - |\mu|, \mu)$ is the same in $S^d(S^n V)$ and $S^n(S^d V)$ as soon as $d$ and $n$ are at least $|\mu|$. (The conjecture also holds when $d$ is sufficiently large or small with respect to $n$ by Brion’s theorem.) Conjecture 7.3.2.1 is still open in general.

***Include survey of results in the combinatorics literature here***

7.3.3. Brill’s equations. Set theoretic equations of $Ch_d(V)$ have been known since 1894. Here is a modern presentation elaborating the presentation in [Lan12, §8.6], which was suggested by E. Briand.

Our goal is a polynomial test to see if $f \in S^d V$ is a product of linear factors. We can first try to see if $P$ is divisible by a power of a linear
The discussion in §6.3.2 will not be helpful as the conditions there are vacuous when \( n - m = 1 \). We could proceed as in §6.6.5 and check if \( \ell x^1 \wedge \cdots \wedge \ell x^D \wedge f = 0 \) where the \( x^f \) are a basis of \( S^{d-1}V \), but in this case there is a simpler test to see if a given linear form \( \ell \) divides \( f \):

Consider the map \( \pi_{d,d} : S^dV \otimes S^dV \to S_{d,d}V \) obtained by projection. (By the Pieri rule 5.2.7.1, \( S_{d,d}V \subset S^dV \otimes S^dV \) with multiplicity one.)

**Lemma 7.3.3.1.** Let \( \ell \in V, f \in S^dV \). Then \( f = \ell h \) for some \( h \in S^{d-1}V \) if and only if \( \pi_{d,d}(f \otimes \ell^d) = 0 \).

**Proof.** Since \( \pi_{d,d} \) is linear, it suffices to prove the lemma when \( f = \ell_1 \cdots \ell_d \).

In that case \( \pi_{d,d}(f \otimes \ell^d) \), up to a constant, is \( (\ell_1 \wedge \ell) \cdots (\ell_d \wedge \ell) \). Since the expression of a polynomial as a monomial is essentially unique, the result follows. \( \square \)

We would like a map that sends \( \ell_1 \cdots \ell_d \) to \( \sum_j \ell_j^d \otimes stuff_j \), as then we could apply \( \pi_{d,d} \otimes Id_{stuff} \) to \( f \) tensored with the result of our desired map to obtain our equations.

While it is not obvious how to obtain such a map for powers, there is an easy way to get elementary symmetric functions, namely the maps \( f \mapsto f_{j,d,j} \) because \( (\ell_1 \cdots \ell_d)_{j,d,j} = \sum_{|K|=j} \ell_K \otimes \ell_{K^c} \) where \( \ell_K = \ell_{k_1} \cdots \ell_{k_j} \) and \( K^c \) denotes the complementary index set in \([d]\). We can try to convert this to power sums by the conversion formula obtained from the relation between generating functions (6.1.2):

\[
(7.3.3) \quad p_d = \mathcal{P}_d(e_1, \ldots, e_d) := \text{det} \begin{pmatrix} e_1 & 1 & 0 & \cdots & 0 \\ 2e_2 & e_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ de_d & e_{d-1} & e_{d-2} & \cdots & e_1 \end{pmatrix}.
\]

The desired term comes from the diagonal \( e_1^d \) and the rest of the terms kill off the unwanted terms of \( e_1^d \). This idea almost works - the only problem is that our naïve correction terms have the wrong degree on the right hand side. For example, when \( d = 3 \), naïvely using \( p_3 = e_1^3 - 3e_1e_2 + 3e_3 \) would give, for the first term, degree \( 6 = 2 + 2 + 2 \) on the right hand side of the tensor product, the second degree \( 3 = 2 + 1 \) and the third degree zero. In general, the right hand side of the \( e_1^d \) term would have degree \( (d-1)^d \), whereas the \( de_d \) term would have degree zero. In addition to fixing the degree mismatch, we need to formalize how we will treat the right hand sides.

To these ends, recall that for any two algebras \( \mathcal{A}, \mathcal{B} \), one can give \( \mathcal{A} \otimes \mathcal{B} \) the structure of an algebra by defining \( (\alpha \otimes \beta) \cdot (\alpha' \otimes \beta') := \alpha \alpha' \otimes \beta \beta' \) and extending linearly. Give \( \text{Sym}(V) \otimes \text{Sym}(V) \) this algebra structure. Define
7.3. The Chow variety

maps

\begin{equation}
E_j : S^\delta V \to S^j V \otimes S^{(\delta-1)} V
f \mapsto f_{j,\delta-j} \cdot (1 \otimes f^{j-1}).
\end{equation}

The \((1 \otimes f^{j-1})\) fixes our degree problem. If \(j > \delta\) define \(E_j(f) = 0\).

Our desired map is

\begin{equation}
Q_d : S^d V \to S^d V \otimes S^{d(d-1)} V
f \mapsto \mathcal{P}_d(E_1(f), \ldots, E_d(f)).
\end{equation}

**Theorem 7.3.3.2** [Brill \[Br93\], Gordan \[Gor94\], Gelfand-Kapranov-Zelevinsky \[GKZ94\], Briand \[Br10\]]. Consider the map

\begin{align}
B : S^d V &\to S_{d,d} V \otimes S^{d-d} V \\
f &\mapsto (\pi_{d,d} \otimes \text{Id}_{S^{d-d} V})[f \otimes Q_d(f)].
\end{align}

Then \([f] \in Ch_d(V)\) if and only if \(B(f) = 0\).

For the proof, we’ll argue by induction, and thus need to generalize the definition of \(Q_d\) a little. Define

\begin{equation}
Q_{d,\delta} : S^\delta V \to S^d V \otimes S^{d(d-1)} V
f \mapsto \mathcal{P}_d(E_1(f), \ldots, E_d(f))
\end{equation}

**Lemma 7.3.3.3.** If \(f_1 \in S^\delta V\) and \(f_2 \in S^{d-d} V\), then

\[Q_{d,\delta}(f_1 f_2) = (1 \otimes f_1^\delta) \cdot Q_{d,d-\delta}(f_2) + (1 \otimes f_2^d) \cdot Q_{d,\delta}(f_1).\]

Assume Lemma 7.3.3.3 for the moment:

**Proof of Theorem 7.3.3.2.** Say \(f = \ell_1 \cdots \ell_d\). First note that for \(\ell \in V\), \(E_j(\ell) = \ell^j \otimes \ell^{j-1}\) and \(Q_{d,1}(\ell) = \ell^d \otimes 1\). Next, compute \(E_1(\ell_1 \ell_2) = \ell_1 \otimes \ell_2 + \ell_2 \otimes \ell_1\) and \(E_2(\ell_1 \ell_2) = \ell_1 \ell_2 \otimes \ell_1 \ell_2\), so \(Q_{2,2}(\ell_1 \ell_2) = \ell_1^2 \otimes \ell_2^2 + \ell_2^2 \otimes \ell_1^2\). By induction and Lemma 7.3.3.3,

\[Q_{d,\delta}(\ell_1 \cdots \ell_\delta) = \sum_j \ell_j^d \otimes (\ell_1^{r_1} \cdots \ell_{j-1}^{r_{j-1}} \ell_{j+1}^{r_{j+1}} \cdots \ell_\delta^{r_\delta}).\]

We conclude \(Q_d(f) = \sum_j \ell_j^d \otimes (\ell_1^{r_1} \cdots \ell_{j-1}^{r_{j-1}} \ell_{j+1}^{r_{j+1}} \cdots \ell_\delta^{r_\delta})\) and \(\pi_{d,d}(\ell_1 \cdots \ell_d, \ell_j^d) = 0\) for each \(j\) by Lemma 7.3.3.1.

For the other direction, first assume \(f\) is reduced, i.e., has no repeated factors. Let \(z \in \text{Zeros}(f)_{\text{smooth}}\), then \(Q_d(f) = (E_1(f))^d + \sum \mu_j \otimes \psi_j\) where \(\psi_j \in S^{d-d} V\), \(\mu_j \in S^d V\) and \(f\) divides \(\psi_j\) for each \(j\) because \(E_1(f)^d\) occurs as a monomial in the determinant (7.3.3) and all the other terms contain an \(E_j(f)\) with \(j > 1\), and so are divisible by \(f\).

Thus \(B(f)(\cdot, z) = \pi_{d,d}(f \otimes (dz)^d)\) because \(E_1(f)^d = (f_1,d-1)^d\) and \(f_1,d-1(\cdot, z) = df_z\), and all the \(\psi_j(z)\) are zero. By Lemma 7.3.3.1, \(df_z\) divides \(f\) for all
Finally, say \( f = g^k h \) where \( g \) is irreducible of degree \( q \) and \( h \) is of degree \( d - qk \) and is relatively prime to \( g \). Apply Lemma 7.3.3.3:

\[
Q_d(g^{k-1}h) = (1 \otimes g^d) \cdot Q_{d,d-q}(g^{k-1}h) + (1 \otimes (g^{k-1}h)^d) \cdot Q_{d,q}(g).
\]

A second application gives

\[
Q_d(g^k h) = (1 \otimes g^d) \cdot [(1 \otimes g^d) \cdot Q_{d,d-2q}(g^{k-2}h) + (1 \otimes (g^{k-2}h)^d) \cdot Q_{d,q}(g)] + (1 \otimes (g^{k-2}h)^d) \cdot Q_{d,q}(g).
\]

After \( k - 1 \) applications one obtains:

\[
Q_d(g^k h) = (1 \otimes g^{d(k-1)}) \cdot [k(1 \otimes h^d) \cdot Q_{d,q}(g) + (1 \otimes g^d) \cdot Q_{d,d-qk}(h)]
\]

and \((1 \otimes g^{d(k-1)})\) will also factor out of \( \mathcal{B}(f) \). Since \( \mathcal{B}(f) \) is identically zero but \( g^{d(k-1)} \) is not, we conclude

\[
0 = \pi_{d,d} \otimes I_{d,k} \cdot f \otimes k(1 \otimes h^d) \cdot Q_{d,q}(g) + (1 \otimes g^d) \cdot Q_{d,d-qk}(h)
\]

Let \( w \in \text{Zeros}(g) \) be a general point, so in particular \( h(w) \neq 0 \). Evaluating at \((z,w)\) with \( z \) arbitrary gives zero on the second term and the first implies

\[
\pi_{d,d} \otimes I_{d,k} (f \otimes Q_{d,q}(g)) = 0
\]

which implies \( dg_w \) divides \( g \), so \( g \) is a linear form. \( \square \)

**Proof of Lemma 7.3.3.3.** Define, for \( u \in \text{Sym}(V) \otimes \text{Sym}(V) \),

\[
\Delta_u : \text{Sym}(V) \to \text{Sym}(V) \otimes \text{Sym}(V)
\]

\[
f \mapsto \sum_j u^j \cdot f_{j,\deg(f)-j}.
\]

**Exercise 7.3.3.4:** Show that \( \Delta_u(fg) = \Delta_u(f) \cdot \Delta_u(g) \), and that the generating series for the \( E_j(f) \) may be written as

\[
\mathcal{E}_f(t) = \frac{1}{1 \otimes f} \cdot \Delta_{t(1 \otimes f)} f.
\]

Note that \((1 \otimes f)^* = 1 \otimes f^* \) and \((1 \otimes fg) = (1 \otimes f) \cdot (1 \otimes g)\). Thus

\[
\mathcal{E}_{fg}(t) = \left[ \frac{1}{1 \otimes f} \cdot \Delta_{t(1 \otimes g)(1 \otimes f)}(f) \right] \cdot \left[ \frac{1}{1 \otimes g} \cdot \Delta_{t(1 \otimes f)(1 \otimes g)}(g) \right],
\]

and taking the logarithmic derivative (recalling Equation (6.1.2)) we conclude. \( \square \)

**Remark 7.3.3.5.** There was a gap in the argument in [Gor94], repeated in [GKZ94], when proving the “only if” part of the argument. They assumed that the zero set of \( f \) contains a smooth point, i.e., that the differential of \( f \) is not identically zero. This gap was fixed in [Bri10]. In [GKZ94] they use \( G_0(d, \dim V) \) to denote \( \text{Ch}_d(V) \).
7.3.4. Brill’s equations as modules. Brill’s equations are of degree $d+1$ on $S^d V^*$. (The total degree of $S_{d,d} V \otimes S^{d^2-d} V$ is $d(d+1)$ which is the total degree of $S^{d+1}(S^d V)$.) Consider the $GL(V)$-module map
\[ S_{d,d} V \otimes S^{d^2-d} V \to S^{d+1}(S^d V) \]
given by Brill’s equations. The components of the target are not known in general and the set of modules present grows extremely fast. One can use the Pieri formula 5.2.7.1 to get the components of the first. Using the Pieri formula, we conclude:

**Proposition 7.3.4.1.** As a $GL(V)$-module, Brill’s equations are multiplicity free.

**Exercise 7.3.4.2:** Write out the decomposition and show that only partitions with three parts appear as modules in Brill’s equations. 

**Remark 7.3.4.3.** If $d < v = \dim V$, then $Ch_d(V) \subset \text{Sub}_d(S^d V)$ so $I(Ch_d(V)) \supset \Lambda^{d+1} V^* \otimes \Lambda^{d+1}(S^{d-1} V^*)$. J. Weyman (in unpublished notes from 1994) observed that these equations are not in the ideal generated by Brill’s equations. More precisely, the ideal generated by Brill’s equations does not include modules $S^\pi V^*$ with $\ell(\pi) > 3$, so it does not cut out $Ch_d(V)$ scheme theoretically when $d < v$. By Theorem 7.3.1.14 the same holds for $Ch_5(C^5)$ and almost certainly holds for all $Ch_n(C^n)$ with $n \geq 5$.

**Problem 7.3.4.4.** What is the kernel of $Brill : S_{n,n} W \otimes S^{n^2-n} W \to S^{n+1}(S^n W)$?

7.3.5. Conjecture 7.3.1.20 and a conjecture in combinatorics. Let $P \in S_{n^d}(C^d) \subset S^d(S^n C^d)$ be non-zero. Conjecture 7.3.1.20 may be stated as $P(x_1 \cdots x_n) \neq 0$. Our first task is to obtain an expression for $P$.

Let $V = C^d$. For any even $n$, the one-dimensional module $S^d V$ occurs with multiplicity one in $S^d(S^n V)$ (cf. [How87, Proposition 4.3]). Fix a volume form on $V$ so that $\det d \in S^d V$ is well defined.

**Proposition 7.3.5.1.** [KL] Let $n$ be even. The unique (up to scale) polynomial $P \in S_{n^d}(C^d) \subset S^d(S^n V)$ evaluates on
\[ x = (v_1^1 \cdots v_n^1)(v_1^2 \cdots v_n^2) \cdots (v_1^d \cdots v_n^d) \in S^d(S^n V^*), \text{ for any } v_j^i \in V^*, \]
to give
\[ \langle P, x \rangle = \sum_{\sigma_1, \ldots, \sigma_d \in S_n} \det_d(v_{\sigma_1(1)}^1, \ldots, v_{\sigma_d(1)}^d) \cdots \det_d(v_{\sigma_1(n)}^1, \ldots, v_{\sigma_d(n)}^d). \]

**Proof.** Let $\bar{P} \in (V) \otimes^{nd}$ be defined by the identity (7.3.9) (with $P$ replaced by $\bar{P}$). It suffices to check that
\begin{enumerate}
  \item $\bar{P} \in S^d(S^n V)$,
\end{enumerate}
(ii) $\bar{P}$ is $SL(V)$ invariant, and

(iii) $\bar{P}$ is not identically zero.

Observe that (iii) follows from the identity (7.3.9) by taking $v_j^i = x_i$ where $x_1, \ldots, x_d$ is a basis of $V^*$, and (ii) follows because $SL(V)$ acts trivially on $\det_d$.

To see (i), we show (ia) $\bar{P} \in S^d((V) \otimes^n)$ and (ib) $\bar{P} \in (S^nV) \otimes^d$ to conclude. To see (ia), it is sufficient to show that exchanging two adjacent factors in parentheses in the expression of $x$ will not change (7.3.9). Exchange $v_1^i$ with $v_2^j$ in the expression for $j = 1, \ldots, n$. Then, each individual determinant will change sign, but there are an even number of determinants, so the right hand side of (7.3.9) is unchanged. To see (ib), it is sufficient to show the expression is unchanged if we swap $v_1^1$ with $v_2^1$ in (7.3.9). If we multiply by $n!$, we may assume $\sigma_1 = \text{Id}$, i.e.,

$$\langle \bar{P}, x \rangle = \left( \begin{array}{c}
\det_d(v_1^1, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d) \cdot \det_d(v_1^2, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d) \cdots \det_d(v_1^n, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d) \\
\end{array} \right).$$

With the two elements $v_1^1$ and $v_2^1$ swapped, we get

(7.3.10)

$$n! \sum_{\sigma_2, \ldots, \sigma_d \in S_n} \det_d(v_1^2, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d) \cdot \det_d(v_1^1, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d) \cdots \det_d(v_1^n, v_{\sigma_2}^2, \ldots, v_{\sigma_d}^d).$$

Now right compose each $\sigma_s$ in (7.3.10) by the transposition $(1, 2)$. The expressions become the same. □

Now specialize to the case $d = n$ (this is the critical case) and evaluate on $(x_1 \cdots x_n)^n$, where $x_1, \ldots, x_n$ is a uni-modular basis of $V^*$.

(7.3.11)

$$\langle P, (x_1 \cdots x_n)^n \rangle = \sum_{\sigma_1, \ldots, \sigma_n \in S_n} \det_d(x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_n}).$$

For a fixed $(\sigma_1, \ldots, \sigma_n)$ the contribution will either be 0, 1 or $-1$. The contribution is zero unless for each $j$, the indices $\sigma_1(j), \ldots, \sigma_n(j)$ are distinct. Arrange these numbers in an array:

$$\begin{pmatrix}
\sigma_1(1) & \cdots & \sigma_n(1) \\
\vdots \\
\sigma_1(n) & \cdots & \sigma_n(n)
\end{pmatrix}$$

The contribution is zero unless the array is a Latin square, i.e., an $n \times n$ matrix such that each row and column consists of the integers $\{1, \ldots, n\}$. If it is a Latin square, the rows correspond to permutations, and the contribution of the term is the product of the signs of these permutations. Call this the row sign of the Latin square. There is a famous conjecture in combinatorics
regarding the products of both the signs of the row permutations and the column permutations, called the sign of the Latin square:

**Conjecture 7.3.5.2** (Alon-Tarsi [AT92]). Let \( n \) be even. The number of sign \(-1\) Latin squares of size \( n \) is not equal to the number of sign \(+1\) Latin squares of size \( n \).

Conjecture 7.3.5.2 is known to be true when \( n = p \pm 1 \), where \( p \) is an odd prime; in particular, it is known to be true up to \( n = 24 \) [Gly10, Dri97].

In [HR94], Huang and Rota showed:

**Theorem 7.3.5.3.** [HR94, Identities 8,9] The difference between the number of column even Latin squares of size \( n \) and the number of column odd Latin squares of size \( n \) equals the difference between the number of even Latin squares of size \( n \) and the number of odd Latin squares of size \( n \), up to sign. In particular, the Alon-Tarsi conjecture holds for \( n \) if and only if the column-sign Latin square conjecture holds for \( n \).

Thus

**Theorem 7.3.5.4.** [KL] The Alon-Tarsi conjecture holds for \( n \) if and only if \( S_n^\kappa(\mathbb{C}^n) \in \mathbb{C}\[Ch_n(\mathbb{C}^n)\] \).

In [KL] several additional statements equivalent to the conjecture were given. In particular, for those familiar with integration over compact Lie groups, the conjecture holds for \( n \) if and only if

\[
\int_{(g_j^i) \in SU(n)} \prod_{1 \leq i, j \leq n} g_j^i d\mu \neq 0
\]

where \( d\mu \) is Haar measure.
Chapter 8

Advanced Topics

In this chapter I present two (possibly more) results that require a more advanced background in algebraic geometry. In §8.1 I present M. Brion’s proof of the asymptotic surjectivity of the Hermite-Hadamard-Howe map. In §8.2 I present S. Kumar’s proof of the non-normality of the determinant orbit closure.

8.1. Asymptotic surjectivity of the Hadamard-Howe map

*This section is still in rough form*****

8.1.1. Coordinate ring of the normalization of the Chow variety.

*** introduction about normalization, and normal varieties to be added***

In this section I follow [Bri93]. There is another variety whose coordinate ring is as computable as the coordinate ring of the orbit, the normalization of the Chow variety. We work in affine space.

An affine variety $Z$ is normal if $\mathbb{C}[Z]$ is integrally closed, that is if every element of $\mathbb{C}(Z)$, the field of fractions of $\mathbb{C}[Z]$, that is integral over $\mathbb{C}[Z]$ (i.e., that satisfies a monic polynomial with coefficients in $\mathbb{C}[Z]$) is in $\mathbb{C}[Z]$. To every affine variety $Z$ one may associate a unique normal affine variety $\text{Nor}(Z)$, called the normalization of $Z$, such that there is a finite map $\pi : \text{Nor}(Z) \to Z$ (i.e. $\mathbb{C}[\text{Nor}(Z)]$ is integral over $\mathbb{C}[Z]$), in particular it is generically one to one and one to one over the smooth points of $Z$. For details see [Sha94, Chap II.5].

In particular, there is an inclusion $\mathbb{C}[Z] \to \mathbb{C}[\text{Nor}(Z)]$ given by pullback of functions, e.g., given $f \in \mathbb{C}[Z]$, define $\tilde{f} \in \mathbb{C}[\text{Nor}(Z)]$ by $\tilde{f}(z) = f(\pi(z))$. 

167
If the non-normal points of $Z$ form a finite set, then the cokernel is finite dimensional. If $Z$ is a $G$-variety, then $\text{Nor}(Z)$ will be too.

Recall that $\text{Ch}_n(W)$ is the projection of the Segre variety, but since we want to deal with affine varieties, we will deal with the cone over it. Consider the product map

$$\phi_n : W^\times n \to S^n W$$

$$(u_1, \ldots, u_n) \mapsto u_1 \cdots u_n$$

Note that i) the image of $\phi_n$ is $\hat{\text{Ch}}_n(W)$, ii) $\phi_n$ is $\Gamma_n = T_W \ltimes G_n$ equivariant.

For any affine algebraic group $\Gamma$ and any $\Gamma$-variety $Z$, define the GIT quotient $Z//\Gamma$ to be the affine algebraic variety whose coordinate ring is $\mathbb{C}[Z]^\Gamma$. (When $\Gamma$ is finite, this is just the usual set-theoretic quotient. In the general case, $\Gamma$-orbits will be identified in the quotient when there is no $\Gamma$-invariant regular function that can distinguish them.) If $Z$ is normal, then so is $Z//\Gamma$ (see, e.g. [Dol03, Prop 3.1]). In our case $W^\times n$ is an affine $\Gamma_n$-variety and $\phi_n$ factors through the GIT quotient because it is $\Gamma_n$-equivariant, so we obtain a map

$$\psi_n : W^\times n /\Gamma_n \to S^n W$$

whose image is $\hat{\text{Ch}}_n(W)$. By unique factorization, $\psi_n$ is generically one to one. Elements of $W^\times n$ of the form $(0, u_2, \ldots, u_n)$ cannot be distinguished from $(0, \ldots, 0)$ by $\Gamma_n$ invariant functions, so they are identified with $(0, \ldots, 0)$ in the quotient, which is consistent with the fact that $\phi_n(0, u_2, \ldots, u_n) = 0$. Observe that $\phi_n$ and $\psi_n$ are $GL(W) = SL(W) \times \mathbb{C}^*$ equivariant.

Consider the induced map on coordinate rings:

$$\psi^*_n : \mathbb{C}[S^n W] \to \mathbb{C}[W^\times n /\Gamma_n] = \mathbb{C}[W^\times n]^{\Gamma_n}.$$ 

For affine varieties, $\mathbb{C}[Y \times Z] = \mathbb{C}[Y] \otimes \mathbb{C}[Z]$ (see, e.g. [Sha94, §2.2]), so

$$\mathbb{C}[W^\times n] = \mathbb{C}[W]^{\otimes n} = Sym(W^*)^\otimes i_1 \cdots Sym(W^*)^\otimes i_n.$$

Taking torus invariants gives

$$\mathbb{C}[W^\times n]^{TSL} = \bigoplus_i S^i W^* \otimes \cdots \otimes S^i W^*,$$

and finally

$$(\mathbb{C}[W^\times n]^{TSL})^G_n = \bigoplus_i S^n (S^i W^*).$$

In summary,

$$\psi_n^*: \text{Sym}(S^n W^*) \to \bigoplus_i (S^n (S^i W^*)).$$
and this map respects $GL$-degree, so it gives rise to maps $\tilde{h}_{d,n} : S^d(S^nW^*) \to S^n(S^dW^*)$.

**Proposition 8.1.1.1.** $\tilde{h}_{d,n} = h_{d,n}$.

**Proof.** Since elements of the form $x_1^n \cdots x_d^n$ span $S^d(S^nW)$ it will be sufficient to prove the maps agree on such elements. By Exercise 7.3.1.3, $h_{d,n}(x_1^n \cdots x_d^n) = (x_1 \cdots x_d)^n$. On the other hand, in the algebra $\mathbb{C}[W]^\otimes_n$, the multiplication is $(f_1 \otimes \cdots \otimes f_n) \circ (g_1 \otimes \cdots \otimes g_n) = f_1 g_1 \otimes \cdots \otimes f_n g_n$ and this descends to the algebra $(\mathbb{C}[W]^\otimes_n)^\Gamma_n$ which is the target of the algebra map $\psi_n^*$, i.e.,

$$\tilde{h}_{d,n}(x_1^n \cdots x_d^n) = \psi_n^*(x_1^n \cdots x_d^n) = \psi_n^*(x_1^n) \otimes \cdots \otimes \psi_n^*(x_d^n) = x_1^n \otimes \cdots \otimes x_d^n = (x_1 \cdots x_d)^n.$$  

□

**Proposition 8.1.1.2.** $\psi_n : W^x \Gamma_n \to \hat{Ch}_n(W)$ is the normalization of $\hat{Ch}_n(W)$.

### 8.1.2. Brion’s asymptotic surjectivity result.

A regular (see, e.g., [Sha94, p.27] for the definition of regular) map between affine varieties $f : X \to Y$ such that $f(X)$ is dense in $Y$ is defined to be finite if $\mathbb{C}[X]$ is integral over $\mathbb{C}[Y]$ (see, e.g. [Sha94, p. 61]). To prove the proposition, we will need a lemma:

**Lemma 8.1.2.1.** Let $X, Y$ be affine varieties equipped with polynomial $\mathbb{C}^*$-actions with unique fixed points $0_X \in X$, $0_Y \in Y$, and let $f : X \to Y$ be a $\mathbb{C}^*$-equivariant morphism such that as sets, $f^{-1}(0_Y) = \{0_X\}$. Then $f$ is finite.

**Proof of Proposition 8.1.1.2.** Since $W^x \Gamma_n$ is normal and $\psi_n$ is regular and generically one to one, it just remains to show $\psi_n$ is finite.

Write $[0] = [0, \ldots, 0]$. To show finiteness, by Lemma 8.1.2.1, it is sufficient to show $\psi_n^{-1}(0) = [0]$ as a set, as $[0]$ is the unique $\mathbb{C}^*$ fixed point in $W^x \Gamma_n$, and every $\mathbb{C}^*$ orbit closure contains $[0]$. Now $u_1 \cdots u_n = 0$ if and only if some $u_j = 0$, say $u_1 = 0$. The $T$-orbit closure of $(0, u_2, \ldots, u_n)$ contains the origin so $[0, u_2, \ldots, u_n] = [0]$. □

**Proof of Lemma 8.1.2.1.** $\mathbb{C}[X], \mathbb{C}[Y]$ are $\mathbb{Z}_{\geq 0}$-graded, and the hypothesis $f^{-1}(0_Y) = \{0_X\}$ states that $\mathbb{C}[X]/f^*(\mathbb{C}[Y]_{>0})\mathbb{C}[X]$ is a finite dimensional vector space. We want to show that $\mathbb{C}[X]$ is integral over $\mathbb{C}[Y]$. This is a graded version of Nakayama’s Lemma (the algebraic implicit function theorem). □
In more detail (see, e.g. [Kum13, Lemmas 3.1,3.2], or [Eis95, p136, Ex. 4.6a]):

**Lemma 8.1.2.2.** Let \( R, S \) be \( \mathbb{Z}_{\geq 0} \)-graded, finitely generated domains over \( \mathbb{C} \) such that \( R_0 = S_0 = \mathbb{C} \), and let \( f^*: R \to S \) be an injective graded algebra homomorphism. If \( f^{-1}(R_{>0}) = \{ S_{>0} \} \) as sets, where \( f: \text{Spec}(S) \to \text{Spec}(R) \) is the induced map on the associated schemes, then \( S \) is a finitely generated \( R \)-module. In particular, it is integral over \( R \).

**Proof.** The hypotheses on the sets says that \( S_{>0} \) is the only maximal ideal of \( S \) containing the ideal \( \mathfrak{m} \) generated by \( f^*(R_{>0}) \), so the radical of \( \mathfrak{m} \) must equal \( S_{>0} \), and in particular \( S_{d>0} \) must be contained in it for all \( d > d_0 \), for some \( d_0 \). So \( S/\mathfrak{m} \) is a finite dimensional vector space, and by the next lemma, \( S \) is a finitely generated \( R \)-module. \( \square \)

**Lemma 8.1.2.3.** Let \( S \) be as above, and let \( M \) be a \( \mathbb{Z}_{\geq 0} \)-graded \( S \)-module. Assume \( M/(S_{>0} \cdot M) \) is a finite dimensional vector space over \( S/S_{>0} \simeq \mathbb{C} \). Then \( M \) is a finitely generated \( S \)-module.

**Proof.** Choose a set of homogeneous generators \( \{ \overline{x}_1, \ldots, \overline{x}_n \} \subset M/(S_{>0} \cdot M) \) and let \( x_j \in M \) be a homogeneous lift of \( \overline{x}_j \). Let \( N \subset M \) be the graded \( S \)-submodule \( Sx_1 + \cdots + Sx_n \). Then \( M = S_{>0}M + N \), as let \( a \in M \), consider \( \overline{a} \in M/(S_{>0}M) \) and lift it to some \( b \in N \), so \( a - b \in S_{>0}M \), and \( a = (a - b) + b \). Now quotient by \( N \) to obtain

\[
S_{>0} \cdot (M/N) = M/N.
\]

If \( M/N \neq 0 \), let \( d_0 \) be the smallest degree such that \( (M/N)^{d_0} \neq 0 \). But \( S_{>0} \cdot (M/N)^{\geq d_0} \subset (M/N)^{d_0+1} \) so there is no way to obtain \( (M/N)^{d_0} \) on the right hand side. Contradiction. \( \square \)

**Theorem 8.1.2.4.** [Bri93] For all \( n \geq 1 \), \( \psi_n \) restricts to a map

\[
\psi_n^* : (W^n/\Gamma_n)\langle 0 \rangle \to S^nW\langle 0 \rangle
\]

such that \( \psi_n^*: \mathbb{C}[S^nW\langle 0 \rangle] \to \mathbb{C}[(W^n/\Gamma_n)\langle 0 \rangle] \) is surjective.

**Corollary 8.1.2.5.** [Bri93] The Hermite-Hadamard-Howe map

\[
h_{d,n} : S^d(S^nW^*) \to S^n(S^dW^*)
\]

is surjective for \( d \) sufficiently large.

**Proof of Corollary.** Theorem 8.1.2.4 implies \( (\psi_n^*)_d \) is surjective for \( d \) sufficiently large, because the cokernel of \( \psi_n^* \) is supported at a point and thus must vanish in large degree. \( \square \)

The proof of Theorem 8.1.2.4 will give a second proof that the kernel of \( \psi_n^* \) equals the ideal of \( Ch_n(W) \).
8.1. Asymptotic surjectivity of the Hadamard-Howe map

Proof of Theorem. Since $\psi_n$ is $\mathbb{C}^*$-equivariant, we can consider the quotient to projective space

$$\tilde{\psi}_n : ((W^{\times n}/\Gamma_n)\setminus[0])/\mathbb{C}^* \to (S^n W\setminus[0])/\mathbb{C}^* = \mathbb{P}S^n W$$

and show that $\tilde{\psi}_n^*$ is surjective. Note that $((W^{\times n}/\Gamma_n)\setminus[0])/\mathbb{C}^*$ is $GL(V)$-isomorphic to $(\mathbb{P}W)^{\times n}/\mathcal{S}_n$, as

$$(W^{\times n}/\Gamma_n)\setminus[0] = (W\setminus0)^{\times n}/\Gamma_n$$

and $\Gamma_n \times \mathbb{C}^* = (\mathbb{C}^*)^{\times n} \times \mathcal{S}_n$. So

$$\tilde{\psi}_n : (\mathbb{P}W)^{\times n}/\mathcal{S}_n \to \mathbb{P}S^n W.$$ 

It will be sufficient to show $\tilde{\psi}_n^*$ is surjective on affine open subsets that cover the spaces. Let $w_1,\ldots,w_n$ be a basis of $W$ and consider the affine open subset of $\mathbb{P}W$ given by elements where the coordinate on $w_1$ is nonzero, and the corresponding induced affine open subsets of $(\mathbb{P}W)^{\times n}$ and $(\mathbb{P}S^n W)$1. We will show that the algebra of $\mathcal{S}_n$-invariant functions on $(\mathbb{P}W)^{\times n}$ is in the image of $(\mathbb{P}S^n W)$1. The restriction of the quotient by $\mathcal{S}_n$ of $(\mathbb{P}W)^{\times n}$ composed with $\tilde{\psi}_n$ to these open subsets in coordinates is

$$(w_1 + \sum_{s=2}^n x_s^w w_s),\ldots,(w_1 + \sum_{s=2}^n x_s^w w_s) \mapsto \Pi_{i=1}^n (w_1 + \sum_{s=2}^n x_s^w w_s).$$

Finally, by e.g., [Wey97, §II.3], the coordinates on the right hand side generate the algebra of $\mathcal{S}_n$-invariant functions in the $n$ sets of variables $(x_s^i)_{i=1,...,n}$. □

With more work, in [Bri97, Thm 3.3], Brion obtains an explicit (but enormous) function $d_0(n,w)$ which is

$$(8.1.3) \quad d_0(n,w) = (n - 1)(w - 1)((n - 1) \left\lfloor \frac{n + w - 1}{w - 1} \right\rfloor - n)$$

for which the $h_{d,n}$ is surjective for all $d > d_0$ where $\dim W = w$.

Problem 8.1.2.6. Improve Brion’s bound to say, a polynomial bound in $n$ when $n = w$.

Problem 8.1.2.7. Note that $\mathbb{C}[\text{Nor}(Ch_n(W))] = \mathbb{C}[GL(W) \cdot (x_1 \cdots x_n)]_{\geq 0}$ and that the the boundary of the orbit closure is irreducible. Under what conditions will a $GL(W)$-orbit closure with reductive stabilizer that has an irreducible boundary will be such that the coordinate ring of the normalization of the orbit closure equals the positive part of the coordinate ring of the orbit?
8.2. Non-normality of $\text{Det}_n$

**give context** be sure to include how $\text{SL}$-orbits are closed*** I follow [Kum13] in this section. Throughout this section I make the following assumptions and adopt the following notation:

Set up:
- $V$ is a $GL(W)$-module,
- Let $\mathcal{P}^0 := GL(W) \cdot P$ and $\mathcal{P} := GL(W) \cdot P$ denote its orbit and orbit closure, and let $\partial \mathcal{P} = \mathcal{P} \setminus \mathcal{P}^0$ denote its boundary, which we assume to be more than zero (otherwise $[\mathcal{P}]$ is homogeneous).

(8.2.1) Assumptions:
1. $P \in V$ is such that the $\text{SL}(W)$-orbit of $P$ is closed.
2. The stabilizer $G_P \subset GL(W)$ is reductive, which is equivalent (by a theorem of Matsushima [Mat60]) to requiring that $\mathcal{P}^0$ is an affine variety.

This situation holds when $V = S^n W$, $\dim W = n^2$ and $P = \det_n$ or $\text{perm}_n$ as well as when $\dim W = rn$ and $P = S^r_n := \sum_{j=1}^{r} x_1^j \cdots x_n^j$, the sum-product polynomial, in which case $\mathcal{P} = \hat{\sigma}_r(\text{Ch}_n(W))$.

Lemma 8.2.0.8. [Kum13] Assumptions as in (8.2.1). Let $M \subset C[\mathcal{P}]$ be a nonzero $GL(W)$-module, and let $Z(M) = \{y \in \mathcal{P} | f(y) = 0 \ \forall f \in M\}$ denote its zero set. Then $0 \subseteq Z(M) \subseteq \partial \mathcal{P}$.

If moreover $M \subset I(\partial \mathcal{P})$, then as sets, $Z(M) = \partial \mathcal{P}$.

Proof. Since $Z(M)$ is a $GL(W)$-stable subset, if it contains a point of $\mathcal{P}^0$ it must contain all of $\mathcal{P}^0$ and thus $M$ vanishes identically on $\mathcal{P}$, which cannot happen as $M$ is nonzero. Thus $Z(M) \subseteq \partial \mathcal{P}$. For the second assertion, since $M \subset I(\partial \mathcal{P})$, we also have $Z(M) \supseteq \partial \mathcal{P}$.

Proposition 8.2.0.9. [Kum13] Assumptions as in (8.2.1). The space of $\text{SL}(W)$-invariants of positive degree in the coordinate ring of $\mathcal{P}$, $C[\mathcal{P}]^{\text{SL}(W)}$, is non-empty and contained in $I(\partial \mathcal{P})$. Moreover,

1. any element of $C[\mathcal{P}]^{\text{SL}(W)}$ cuts out $\partial \mathcal{P}$ set-theoretically, and
2. the components of $\partial \mathcal{P}$ all have codimension one in $\mathcal{P}$.

Proof. To study $C[\mathcal{P}]^{\text{SL}(W)}$, consider the GIT quotient $\mathcal{P}//\text{SL}(W)$ whose coordinate ring, by definition, is $C[\mathcal{P}]^{\text{SL}(W)}$. It parametrizes the closed $\text{SL}(W)$-orbits in $\mathcal{P}$, so it is non-empty. Thus $C[\mathcal{P}]^{\text{SL}(W)}$ is nontrivial.
8.2. Non-normality of \( \text{Det}_n \)

Claim: every \( SL(W) \)-orbit in \( \partial P \) contains \( \{0\} \) in its closure, i.e., \( \partial P \) maps to zero in the GIT quotient. This will imply any \( SL(W) \)-invariant of positive degree is in \( I(\partial P) \) because any non-constant function on the GIT quotient vanishes on the inverse image of \([0]\). Thus \( 1 \) follows from Lemma 8.2.0.8. The zero set of a single polynomial, if it is not empty, has codimension one, which implies the components of \( \partial P \) are all of codimension one, proving \( 2 \).

It remains to show \( \partial P \) maps to zero in \( P/\!/SL(W) \), where \( \rho : GL(W) \to GL(V) \) is the representation. This GIT quotient inherits a \( \mathbb{C}^* \) action via \( \rho(\lambda Id) \), for \( \lambda \in \mathbb{C}^* \). Its normalization is just the affine line \( \mathbb{A}^1 = \mathbb{C} \). To see this, consider the \( \mathbb{C}^* \)-equivariant map \( \sigma : \mathbb{C} \to P \) given by \( z \mapsto \rho(zId) \cdot P \), which descends to a map \( \overline{\sigma} : \mathbb{C} \to P/\!/SL(W) \). Since the \( SL(W) \)-orbit of \( P \) is closed, for any \( \lambda \in \mathbb{C}^* \), \( \rho(\lambda Id)P \) does not map to zero in the GIT quotient, so we have \( \overline{\sigma}^{-1}(\{0\}) = \{0\} \) as a set. Lemma 8.1.2.1 applies so \( \overline{\sigma} \) is finite and gives the normalization. Finally, were there a closed nonzero orbit in \( \partial P \), it would have to equal \( SL(W) \cdot \sigma(\lambda) \) for some \( \lambda \in \mathbb{C}^* \) since \( \overline{\sigma} \) is surjective. But \( SL(W) \cdot \sigma(\lambda) \subset P^0 \).

Remark 8.2.0.10. That each irreducible component of \( \partial P \) is of codimension one in \( P \) is due to Matsushima [Mat60]. It is a consequence of his result mentioned above.

The key to proving non-normality of \( \hat{\text{Det}}_n \) and \( \hat{\text{perm}}_n^\text{et} \) is to find an \( SL(W) \)-invariant in the coordinate ring of the normalization (which has a \( GL(W) \)-grading), which does not occur in the corresponding graded component of the coordinate ring of \( S^n W \), so it cannot occur in the coordinate ring of any \( GL(W) \)-subvariety.

Lemma 8.2.0.11. Assumptions as in (8.2.1). Let \( P \in S^n W \) be such that \( SL(W) \cdot P \) is closed and \( G_P \) is reductive. Let \( d \) be the smallest positive \( GL(W) \)-degree such that \( \mathbb{C}[P^0]^{SL(W)} \neq 0 \). If \( n \) is even and \( d < nw \) (resp. \( n \) is odd and \( d < 2nw \)) then \( P \) is not normal.

Proof. Since \( P^0 \subset P \) is a Zariski open subset, we have the equality of \( GL(W) \)-modules \( \mathbb{C}(P) = \mathbb{C}(P^0) \). By restriction of functions \( \mathbb{C}[P] \subset \mathbb{C}[P^0] \) and thus \( \mathbb{C}[P]^{SL(W)} \subset \mathbb{C}[P^0]^{SL(W)} \). Now \( P^0/\!/SL(W) = P^0/SL(W) \simeq \mathbb{C}_w \), so \( \mathbb{C}[P^0]^{SL(W)} \simeq \bigoplus_{k \in \mathbb{Z}} \mathbb{C}\{z^k\} \). Under this identification, \( z \) has \( GL(W) \)-degree \( d \). By Proposition 8.2.0.9, \( \mathbb{C}[P]^{SL(W)} \neq 0 \). Let \( h \in \mathbb{C}[P]^{SL(W)} \) be the smallest element in positive degree. Then \( h = z^k \) for some \( k \). Were \( P \) normal, we would have \( k = 1 \).

But now we also have a surjection \( \mathbb{C}[S^n W] \to \mathbb{C}[P] \), and by Exercise ??? the smallest possible \( GL(W) \)-degree of an \( SL(W) \)-invariant in \( \mathbb{C}[S^n W] \) when \( n \) is even (resp. odd) is \( wn \) (resp. \( 2wn \)) which would occur in \( S^w(S^n W) \) (resp. \( S^{2w}(S^n W) \)). We obtain a contradiction. ☐
Theorem 8.2.0.12 (Kumar [Kum13]). For all $n \geq 3$, $\det_n$ and $\perm_n^m$ are not normal. For all $n \geq 2m$ (the range of interest), $\perm_n^m$ is not normal.

I give the proof for $\det_n$, the case of $\perm_n^m$ is an easy exercise. Despite the variety being much more singular, the proof for $\perm_n^m$ is more difficult, see [Kum13].

Proof. We will show that when $n$ is congruent to 0 or 1 mod 4, $\mathbb{C}[\det_n]^{SL(W)} \not \cong \mathbb{C}^* = 0$ and when $n$ is congruent to 2 or 3 mod 4, $\mathbb{C}[\det_n]^{SL(W)} \not \cong \mathbb{C}^*$. Since $n, 2n < (n^2)n$ Lemma 8.2.0.11 applies.

The $SL(W)$-trivial modules are $(\Lambda^n W)^{\otimes s} = S_{sn} W$. Write $W = E \otimes F$. We want to determine the lowest degree trivial $SL(W)$-module that has a $G_{\det_n} = (SL(E) \times SL(F)/\mu_n) \rtimes \mathbb{Z}_2$ invariant. We have the decomposition $(\Lambda^n W)^{\otimes s} = (\oplus_{\pi' = n^2} S_{\pi'} E \otimes S_{\pi'} F)^{\otimes s}$, where $\pi'$ is the conjugate partition to $\pi$. Thus $(\Lambda^n W)^{\otimes s}$ contains the trivial $SL(E) \times SL(F)$ module $(\Lambda^n E)^{\otimes ns} \otimes (\Lambda^n F)^{\otimes ns}$ with multiplicity one. (In the language of §6.2.5, $k_{sn^2,(sn)^n} = 1$.) Now we consider the effect of the $\mathbb{Z}_2 \subset G_{\det_n}$ with generator $\tau \in GL(W)$. It sends $e_i \otimes f_j$ to $e_j \otimes f_i$, so acting on $W$ it has +1 eigenspace $e_i \otimes f_j + e_j \otimes f_i$ for $i \leq j$ and -1 eigenspace $e_i \otimes f_j - e_j \otimes f_i$ for $1 \leq i < j \leq n$. Thus it acts on the one-dimensional vector space $(\Lambda^n W)^{\otimes s}$ by $(-1)^{\binom{s}{2}}$, i.e., by -1 if $n \equiv 2, 3 \mod 4$ and $s$ is odd and by 1 otherwise. We conclude that there is an invariant as asserted above. (In the language of §6.2.5, $sk_{sn^2,(sn)^n} = 1$ for all $s$ when $\binom{n}{2}$ is even, and $sk_{sn^2,(sn)^n} = 1$ for even $s$ when $\binom{n}{2}$ is odd and is zero for odd $s$.) \hfill \Box

Exercise 8.2.0.13: Write out the proof of the non-normality of $\perm_n^m$.

Exercise 8.2.0.14: Show the same method gives another proof that $Ch_n(W)$ is not normal, but that it fails (with good reason) to show $\sigma_n(v_d(\mathbb{P}^{n-1}))$ is not normal.

Exercise 8.2.0.15: Show a variant of the above holds for any reductive group with a nontrivial center (one gets a $\mathbb{Z}_k$-grading of modules if the center is $k$-dimensional), in particular it holds for $G = GL(A) \times GL(B) \times GL(C)$. Use this to show that $\sigma_n(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ is not normal when $\dim A = \dim B = \dim C = r > 2$. 

Hints and Answers to
Selected Exercises

1.3.2.2 For the second assertion, a generic matrix will have nonzero determinant. In general, the complement to the zero set of any polynomial over the complex numbers has full measure. For the last assertion, first say \( \text{rank}(f) = r' \leq r \) and let \( v_1, \ldots, v_\nu \) be a basis of \( V \) such that the kernel is spanned by the last \( \nu - r' \) vectors. Then the matrix representing \( f \) will be nonzero only in the upper \( r' \times r' \) block and thus all minors of size greater than \( r' \) will be zero. Next say \( \text{rank}(f) = s > r \). Taking basis in the same manner, we see the upper right size \( s \) submatrix will have a nonzero determinant. Taking a Laplace expansion, we see at least one size \( r + 1 \) minor of it is nonzero. In any other choice of basis minors expressed in the new basis are linear combinations of minors expressed in the old, so we conclude. If you need help with the third assertion, use Proposition 2.4.5.1.

1.3.2.7 Consider \( \lim_{\epsilon \to 0} \frac{1}{\epsilon} ((x + \epsilon y)^n - x^n) \).

?? answer coming soon

2.1.1.1 \( v \in V \) goes to the map \( \beta \mapsto \beta(v) \).

2.1.1.4 \( \text{trace}(f) \).

2.1.2.1 A multi-linear map is determined by its action on bases of \( A_1^*, \ldots, A_n^* \).

2.1.3.5 Use Exercise 2.3.5.2.1.3.4.

2.3.2.5 Show that a basis of \( \mathfrak{s}(V) \) may be obtained from elements of \( GL(V) \) acting on e.g. a matrix with a single nonzero entry off the diagonal, and the matrix whose entries are all zero except the \((1, 1)\), which is 1, and the \((2, 2)\) which is \(-1\). In fact it is sufficient to use permutation matrices.
2.3.2.8 We need $\alpha X \beta^T = (g \cdot \alpha)^T (g \cdot X) (g \cdot \beta)$. If $\alpha$ is a row vector then $g \cdot \alpha$ is the row vector $a g^{-1}$.

2.3.3.5 First note that $V \otimes W$ is an $H$ module, which decomposes as $\dim V$ copies of $W$, i.e., the irreducible $H$-submodules are of the form $v \otimes W$ for some $v \in V$. Thus if $U \subset V \otimes W$ is a $G \times H$-submodule, it must contain a vector of the form $v \otimes W$. But then it must contain $\text{span}(G \cdot v) \otimes W$ because it is a $G$-module as well. But the span of $G \cdot v$ is $V$ because $V$ is irreducible.

2.3.5.7 In each case show that both factors are included and conclude by counting dimensions.

2.2.1.2 See [Lan12, §2.4.4]

2.2.3.1 If $T = \sum_i a_i \otimes b_i \otimes c_i$, then, letting $\pi_A : A \to A/(A')^\perp$ be the projection, and similarly for $B$, then $T'_A \otimes T'_B \otimes T'_C = \sum_i \pi_A(a_i) \otimes \pi_B(b_i) \otimes \pi_C(c_i)$.

2.2.4.2 First assume $R(T) = R(T)$ and write $T = a_1 \otimes b_1 \otimes c_1 + \cdots + a_r \otimes b_r \otimes c_r$. 2.4.6.2 $P(x) = 0$ if and only if $P(x, \ldots, x) = P(x^d) = 0$.

2.4.12.2 Since the border rank of points in $GL(A) \times GL(B) \times GL(C) \cdot T$ equals the border rank of $T$, the border rank of points in the closure cannot increase.

2.4.13.2 Since $X$ is a $G$-variety, $P(x) = 0$ for all $x \in X$ implies $P(gx) = 0$ for all $g \in G$.

2.4.13.3 Use that

$$S^2(A \otimes B \otimes C) = S^2A \otimes S^2(B \otimes C) \oplus \Lambda^2A \otimes \Lambda^2(B \otimes C)$$

$$= S^2A \otimes S^2B \otimes S^2C \oplus S^2A \otimes \Lambda^2B \otimes \Lambda^2C \oplus \Lambda^2A \otimes S^2B \otimes \Lambda^2C \oplus \Lambda^2A \otimes \Lambda^2B \otimes S^2C$$

2.5.3.2 Use Exercise 2.2.4.3.

2.7.1.3 If $R(T) = m$, then $T$ is a limit of points $T_i$ with $\mathbb{P}T_i(A^*) \cap \text{Seg}(\mathbb{P}B \times \mathbb{P}C) \neq \emptyset$.

2.7.1.5 First note that if $x$ is generic, it is diagonalizable with distinct eigenvalues so if $x$ is generic, then $\dim C(x) = b$. Then observe that $\dim(C(x))$ is semi-continuous as the set $\{y \mid \dim C(y) \leq \perp\}$ is an algebraic variety.

3.2.1.3 Recall from Exercise 2.1.3.5 that $\otimes_j M_{ij,m_j,n_j} = M_{ij,m_j,n_j} \otimes M_{ij,m_j,n_j}$.

Set $N = m,m$ and consider $M_{(\cdot)} = M_{(m,n,l)} \otimes M_{(n,1,m)} \otimes M_{(l,m,n)}$.

3.2.1. Consider

\[
\begin{pmatrix}
\heartsuit & \heartsuit \\
\spadesuit & \spadesuit
\end{pmatrix}
\]
3.4.2.3 Instead of the curve $a_0 + ta_1$ use $a_0 + ta_1 + t^2a_{q+1}$ and similarly for $b, c$.

3.4.2.4 Consider the coefficient of $t^2$ in an expansion.

3.6.1.7 When writing $T = \lim_{t \to 0} T(t)$ we may take $t \in \mathbb{Z}_{h+1}$. 3.6.1.8 See [BCS97, §2.1]

4.2.2.3 Note that while for $V^{\otimes 3}$, the kernels of $S^2V \otimes V \to S^3V$ and $\Lambda^2V \otimes V \to \Lambda^3V$ were isomorphic $GL(V)$-modules, the kernels of $S^3V \otimes V \to S^4V$ and $\Lambda^3V \otimes V \to \Lambda^4V$ are not. One can avoid dealing with spaces like $S_{21}V \otimes V$ by using maps like, e.g. the kernel of $S^2V \otimes S^2V \to S^4V$ and keeping track of dimensions of spaces uncovered. The answer is given by Theorem 5.2.1.2.

4.2.3 If $a_0 = b_0 = c_0 = 1d$ then $(u_1 \oplus v_3) \oplus (v_2 \oplus w_1) \oplus (w_3 \oplus u_2)$ is mapped to $(w_T^T \oplus (w_3^T \oplus (w_3^T \oplus u_1^T)) \oplus (v_3^T \oplus (u_1^T))^T)$. The general case is just notationally more cumbersome.

4.2.3.3 See [CHI+]. The calculation is a little more involved than indicated in the section.

5.1.1.1 Prove an algebra version of Schur’s lemma.

5.1.4.2 If $V$ is an irreducible $G$-module, then $V^* \otimes V$ is an irreducible $G \times G$-module.)

5.2.5.6 The highest weight vectors of $S_{d-1,1}V$ are all permutations of $x_1^{d-1} \otimes x_2 - x_1^{d-2}x_2 \otimes x_1$, so the only choice one has is the position of $x_2$. But now if one sums over all possible positions one gets zero, and this is the only linear relation.

5.2.2.2 $c_{\pi'} = \sum_{\sigma \in \mathfrak{S}_d} \delta_{\sigma} \sum_{\sigma \in \mathfrak{S}_d} \text{sgn}(\sigma)\delta_{\sigma}$. Now show $c_{(1d)}c_{\pi} = c_{\pi'}$.

5.2.5.9 $x_{\pi} \otimes x_{\mu} \subset V^{\otimes (|\pi|+|\mu|)}$ and is a highest weight vector.

5.2.6.4 Recall that $S^d(A_1 \otimes \cdots \otimes A_n) = (A_1^{\otimes d} \otimes \cdots \otimes A_n^{\otimes d}) \mathfrak{S}_d$.

5.2.6.3 Use Exercise 5.2.2.2.

5.2.9.2 $g \cdot e_1 \wedge \cdots \wedge e_V = \det(g)e_1 \wedge \cdots \wedge e_V$

?? This is clear for the default Young tableau, now let $\mathfrak{S}_d$ act.

?? Consider $M^{\otimes N}_{(12,12,12)}$.

6.2.4.3 In this case the determinant is a smooth quadric.

?? Use the tangent space as described in §2.4.4.

?? $\mathbb{P}N^*_xX \subset X^V$.

?? The hypersurface $\{x_1 \cdots x_n + y_1 \cdots y_n = 0\}$ is self-dual.

6.6.6.3 Note that $\frac{\partial R}{\partial x_i} = \sum_j \frac{\partial^2 R}{\partial x_i \partial x_j}$ and now consider the last nonzero column.

6.8.8.1 $\{\text{perm}_2 = 0\}$ is a smooth quadric.
7.3.4.2 $Ch_d(C^2) = \mathbb{P}S^dC^2$. 
Bibliography


[Kum] Shrawan Kumar, A study of the representations supported by the orbit closure of the determinant, arXiv:1109.5996.


[MR13] Alex Massarenti and Emanuele Raviolo, *The rank of an n × n matrix multiplication is at least 3n^2 − 2√2n^{3/2} − 3n*, Linear Algebra Appl. 438 (2013), no. 11, 4500–4509. MR 3034546


[Pop75] A. M. Popov, Irreducible simple linear Lie groups with finite standard subgroups in general position, Funkcional. Anal. i Priložen. 9 (1975), no. 4, 81–82. MR MR0396847 (53 #707)


[Sha94] Igor R. Shafarevich, Basic algebraic geometry. 1, second ed., Springer-Verlag, Berlin, 1994, Varieties in projective space, Translated from the 1988 Russian edition and with notes by Miles Reid. MR 1328833 (95m:14001)


[Ter11] A. Terracini, Sulla varietà $V_h$ per cui la varietà degli $s_{h}(h+1)$-seganti ha dimensione minore dell'ordinario, Rend. Circ. Mat. Palermo 31 (1911), 392–396.


Index

2A-generic, 49
G-module, 20
G-module map, 22
\(Sub_{k}(S^{d}V)\)
equations of, 123
\(X^{\vee}\), dual variety of \(X\), 127
Pad, 123
\(\pi^{t}\), 96
chow, 139
\(VNP\), 113
\(VP\), 113
s-rank, 75
\(v_{d}(\mathbb{P}V)^{\vee}\), dual of Veronese, 127
affine linear projection, 10
algebraic variety, 6, 24
arithmetic circuit, 112
bilinear map, 4
border rank, 19, 59
Brill’s equations, 157
Burnside’s theorem, 91
Cauchy formulas, 101
centralizer, 49, 90
character of representation, 94
Chow variety
Brill’s equations for, 157
equations of, 152
class function, 94
codimension of variety, 27
combinatorial restriction, 57
combinatorial value, 69
commutator, 90
complete problem, 113
completely reducible module, 90
complexity class
complete problem for, 113
hard problem for, 113
concise, 47
cone, 122
conjugate partition, 96
content, 100
contraction map, 23
decomposable representation, 20
degeneracy value, 64
degree of variety, 28
determinant, 9
DFT, 72
Discrete Fourier Transform, 72
discriminant hypersurface, 127
dual variety, 126, 127
dual vector space, 16
elementary symmetric function, 114
exponent of matrix multiplication, 4
formula, 115
quasi-polynomial, 115
general point, 29
generating function
for elementary symmetric polynomials, 114
generic
1A, 47
group algebra, 71
hard problem, 113
highest weight, 99
highest weight vector, 99
hook length, 100
ideal, 6
immanant, 104
inheritance, 107
interlace, 104
irreducible action, 20
isomorphic G-modules, 22
Jacobian loci, 141
Kempf-Weyman desingularization, 123
Koszul flattening, 39
Kronecker coefficients, 109, 121
symmetric, 121
Lie algebra, 30
linear map
rank, 16
Littlewood Richardson Rule, 104
matrix coefficient, 72
matrix coefficients, 93
module, 20
completely reducible, 90
semi-simple, 90
simple, 90
module homomorphism, 22
module map, 22
normalization
of a curve, 58
obstruction
occurrence, 118
orbit occurrence, 117
orbit representation-theoretic, 118
representation-theoretic, 119
occurrence obstruction, 118
orbit occurrence obstruction, 117
orbit representation-theoretic obstruction, 118
partition, 95
partpolar, 23
Pascal determinant, 144
Pieri formula, 104
Plücker coordinates, 44
polarization, 23
pullback, 45
quasi-polynomial size formula, 115
quasi-projective variety, 93
rank of linear map, 16
re-ordering isomorphism, 17
reducible representation, 20
reduction of sequence, 113
regular endomorphism, 49
regular semi-simple, 49
representation
decomposable, 20
irreducible, 20
reducible, 20
representation-theoretic obstruction, 119
restriction value, 68
resultant, 133
Schur’s lemma, 22
semi-simple module, 90
Shannon entropy, 14
sign representation, 21
simple module, 90
size of
arithmetic circuit, 112
size of circuit, 112
skew-symmetric tensor, 22
smooth point, 28
stabilizer
characterizes point, 117
of tensor, 24
standard Young tableau, 101
structure tensor of an algebra, 71
submodule, 20
subspace variety, 26, 122
sum-product polynomial, 141
symmetric Kronecker coefficients, 121
symmetric subspace variety
equations of, 123
symmetric tensor, 22
tensor
combinatorial restriction, 68
degenerates to, 63
restricts to, 68
skew-symmetric, 22
symmetric, 22
tensor product, 16
tensor rank, 5
Terracini’s Lemma, 33
triple product property, 73
trivial representation, 20
value
combinatorial, 69
degeneracy, 64
restriction, 68
variety, 24
algebraic, 6
codimension of, 27
degree of, 28
of padded polynomials, 123
quasi-projective, 93
Index

weight vector, 99
weight, highest, 99
Weyl group, 103
wreath product, 137

Young diagram, 96

Zariski closure, 6