Sums of Squares and Symmetric Polynomials

by

Isabelle Shankar

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Serkan Hoşten, Co-chair
Professor Jon Wilkening, Co-chair
Professor Alper Atamturk
Professor Mark Haiman

Spring 2021

Sums of Squares and Symmetric Polynomials

Abstract

Sums of Squares and Symmetric Polynomials

by

Isabelle Shankar

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Serkan Hoşten, Co-chair

Professor Jon Wilkening, Co-chair

Real algebraic geometry has a long and beautiful history going back to the 1800s. It is the study of real polynomials using algebraic techniques. Convex optimization plays a key role in applied mathematics and engineering, studying the geometric structure of convex sets that arise in optimization problems. Representation theory allows us to understand, study, and exploit the symmetries that naturally arise in many mathematical problems. The work in this thesis lies in the intersection of these fields, studying the symmetries of geometric objects coming from convex algebraic geometry and optimization. In particular, we study the spectrahedra that arise in the theory of symmetric polynomials and sums of squares (SOS) polynomials.

Much of this thesis is motivated by polynomial optimization. It is through this lens that we arrive at the study of semidefinite programming, the feasible region of which is called a spectrahedron. A common method for solving a semidefinite program (SDP) is via interior-point methods. Interior-point algorithms cut out a path towards the optimal solution and taking the Zariski closure of this defines the central curve. In this thesis, we discuss the central curve in semidefinite programming, linear programming and quadratic programming. As our first contribution, we prove the degree of the central curve for a generic SDP is equal to the maximum likelihood (ML) degree of a statistical model formulated and discussed in [75]. As such the degree of the central curve of a generic SDP can be computed using complete quadrics [50, 53].

The study of sums of squares and invariant polynomials inevitably leads to the study of invariant semidefinite programs [32]. Such an SDP can be greatly simplified using symmetry reduction techniques. We explore the geometry of the spectrahedra that arise from the invariant SDPs when considering sums of squares and symmetric polynomials. The first spectrahedron we study is the *symmetry-adapted PSD cone*, denoted $PSD_N^{S_n}$, a spectrahedral

cone that gives representations of symmetric SOS polynomials. We compute its dimension, characterize its extremal rays, and determine that this convex cone is Terracini convex. Furthermore, using tools from representation theory, one can enforce a block-diagonalized structure on the set of matrices in $PSD_N^{S_n}$. We study the structure of these blocks, each associated to an irreducible representation of the symmetric group. For the trivial irreducible representation, we provide an explicit description of these matrix blocks.

The second spectrahedron is the symmetry-adapted Gram spectrahedron. For a given symmetric polynomial f of degree $2d$ in n variables, the *symmetry-adapted Gram spectrahedron* of f is the intersection of the affine subspace defined by the coefficients of f and the symmetry-adapted PSD cone. For particular n and d, we describe the facial structure of these spectrahedra, including for binary ($n = 2$), quadratic ($d = 2$), ternary quartic ($n = 3, d = 4$), and ternary sextic ($n = 3, d = 6$) polynomials.

Finally, as an application of the above theory, we find several counterexamples to a 2011 conjecture presented by Cuttler, Greene, and Skandera in [24] pertaining to symmetric mean inequalities.

To my family

who supported and encouraged me through every step of this journey.

Contents

List of Figures

Acknowledgments

I would first like to thank my advisor Serkan Hoşten for the incredible support and mentorship he provided me throughout graduate school. His belief in me was unwavering and his enthusiasm for math was infectious. It was an honor and a pleasure to work with him.

I also wish to thank the many mathematical mentors who helped me grow academically and professionally, including Bernd Sturmfels and the many brilliant scholars at the Max Planck Institute for Mathematics in the Sciences, Mario Sanchez, Cynthia Vinzant, Greg Blekherman, Curtis Greene, and Federico Ardila. Thank you to my collaborators Alexander Heaton and Angélica Torres who remind me why I love math. Thank you to my Berkeley advisor Jon Wilkening who stood by me when others wouldn't and helped keep me in the game. Thank you to my NRing comrades, including Katrin, Mariel, Maddie, Maddie, Max, Jess, and Catherine. You are my heros.

Last, I wish to thank my partner, Nathan Feigion, who kept me sane through grad school.

Chapter 1

Introduction

1.1 Motivation

Let $\mathbb{R}[x_1, \ldots, x_n]_{\leq d} = \mathbb{R}[x]_{\leq d}$ be the space of real polynomials in n variables of degree at most d. A polynomial f in $\mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ is said to be a *sums of squares* (SOS) polynomial if $f = q_1^2 + \cdots + q_r^2$ where $q_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}$, $i = 1, \ldots, r$. We will often simply say that f is SOS for short.

Why is this interesting to study? As motivation we consider the unconstrained polynomial optimization problem

$$f^* = \inf_{x \in \mathbb{R}^n} f(x) \tag{1.1}$$

where f is real a polynomial of degree $2d$. Polynomial optimization problems approximate an incredibly wide range of problems in theoretical and applied mathematics, and as such have a rich history of research while still being an ongoing area of study. However, even this seemingly simple unconstrained problem is NP-hard (in the number of variables n and fixed degree). Thus the optimization community continues to search for methods that can help relax this problem while still effectively and efficiently finding a good solution. This is where sums of squares come in.

We first reformulate the problem as follows

$$\begin{aligned} f^* &= \inf\{f(x) : x \in \mathbb{R}^n\} \\ &= \sup\{\gamma \in \mathbb{R} : f(x) \geq \gamma\} \\ &= \sup\{\gamma \in \mathbb{R} : f(x) - \gamma \geq 0\} \end{aligned}$$

so that we are now interested in certifying nonnegativity of the polynomial $f - \gamma$. This can be relaxed to a sums of squares problem,

$$f^{sos} = \sup\{\gamma : f(x) - \gamma \text{ is SOS}\}.$$

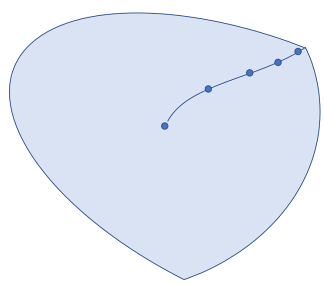
It is clear that if a polynomial is SOS, it is nonnegative. However the converse is not necessarily true except in the three cases proven by Hilbert: univariate polynomials ($n = 1$), quadratic polynomials ($2d = 2$), and bivariate quartics ($n = 2$ and $2d = 4$). Still, there is a great deal of practical evidence that sums of squares are highly useful even beyond these cases.

The next natural question to answer is why this is easier to solve than the original polynomial optimization problem. The short answer is because an SOS problem reduces to a semidefinite program (SDP) and there are efficient algorithms to solve it. Indeed a polynomial f is SOS if and only if there is a positive semidefinite matrix Q such that

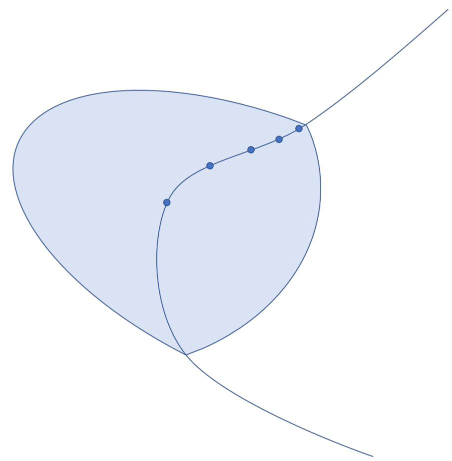
$$f(x) = [x]^t Q [x] \tag{1.2}$$

where $[x]$ is a vector of polynomials that form a basis of $\mathbb{R}[x]_{\leq d}$. If we take Q to be a matrix of unknowns, then by equating coefficients we get a system of linear equations in the matrix entries of Q. Thus to certify that a polynomial is SOS, we must search over the cone of positive semidefinite matrices (the PSD cone) for a matrix which satisfies the affine conditions imposed by Equation 1.2 which is precisely a semidefinite program. The set of all positive semidefinite matrices Q that satisfy the equation $f = [x]^t Q[x]$ is called the *Gram spectrahedron* of f. More generally, the feasible region of an SDP is called a spectrahedron.

One common method of solving an SDP is via an interior-point method. In this algorithm, one solves a series of optimization problems related to the original SDP. Each unique solution is a point inside the feasible region of the SDP, and if one takes the set of all such solutions, we create a path towards the boundary of the spectrahedron. This is called the *central path* and it leads to an optimal solution of the original optimization problem.



Suppose that we now take the Zariski closure of this path. This gives us the *central curve* of our optimization problem.



Studying the central curve gives insight into how complex the corresponding optimization problem could be. Indeed, in the linear programming case, this was studied in [25] and the degree of the central curve of a generic linear program was calculated to be $\binom{m-1}{k}$, where m is the number variables and k is the number of linear constraints. Chapter 2 discusses this in the semidefinite programming case as well as for linear programming and quadratic programming.

Now, the SDP corresponding to a given SOS program can grow quite quickly. As such, it is important to develop and study tools which reduce or simplify the problem size. The primary goal of this thesis is to study symmetry reduction techniques which do precisely that for the case when a given polynomial is symmetric. Symmetric polynomials are of vital importance in representation theory and combinatorics. With respect to sums of squares, the structure of the space of symmetric polynomials can be exploited, and it is for this reason that we are interested in studying symmetric polynomials.

The symmetry reduction techniques studied in this thesis were introduced in [32] and can be broken into two fundamental steps. In the first step, we reduce the number of indeterminants in the decision variable of our SDP. The second step block-diagonalizes the decision variable. We briefly summarize this process here, but we will go into more details in Chapter 3.

Let $D : S_n \to GL(N)$ be an orthogonal representation of S_n acting on $\mathbb{R}[x]_d$, the space of homogeneous polynomials of degree d in n variables, where $N = \binom{n+d-1}{d}$. By restricting the SOS problem to the *fixed point subspace*

$$\mathcal{F} = \{X : XD(\sigma) = D(\sigma)X, \forall \sigma \in S_n\},$$

the number of variables in the SDP is significantly reduced. This brings us to the first spectrahedron of interest. For fixed n and d, the *symmetry adapted PSD cone* is

$$PSD_N^{S_n} = \{Q \in PSD_N : QD(\sigma) = D(\sigma)Q, \text{ for all } \sigma \in S_n\},$$

a spectrahedral cone that gives representations of symmetric SOS polynomials. Next, by employing a change of basis matrix which block-diagonalizes the representation D, one can enforce a similar block-diagonalized structure to the set of matrices in $PSD_N^{S_n}$ [28]. This results in a cone that is isomorphic to a direct sum of smaller PSD cones. The smaller and more structured symmetry adapted PSD cone is much easier to optimize over compared to the much larger PSD cone. We use this to define the second spectrahedron of interest in this thesis. For a given symmetric polynomial f of degree $2d$, the *symmetry adapted Gram spectrahedron* of f is the set

$$K_f^{S_n} = L_f \cap PSD_N^{S_n},$$

that is, the intersection of the affine subspace defined by $f = [x]^T Q[x]$ (as before) and the symmetry adapted PSD cone. It is clear that the symmetry adapted Gram spectrahedron is a subset of the Gram spectrahedron of f.

1.2 Contributions

Chapter 2 discusses the degree of the central curve for semidefinite programming, linear programming, and quadratic programming, based on work in [42]. Our main contribution is Theorem 2.2.3 where we prove that the degree of the central curve for an SDP, when the cost function and the right-hand side vector are generic, is equal to the maximum likelihood degree (ML degree) of the linear concentration model generated by the constraint matrices and cost matrix of the given SDP. When the constraints are also generic, this degree is equal to the degree of the reciprocal variety associated to the linear subspace defined again by the constraint matrices and cost matrix. We further conclude in Corollary 2.2.1 and Corollary 2.2.2 that the degree of the central curve of a generic SDP is symmetric in the number of constraints, and it is polynomial in m (size of the matrices) of degree k (number of constraints).

In Section 2.2 we will revisit the degree of the central curve of a generic linear program. We relate this degree to the ML degree of linear concentration models generated by diagonal matrices and further provide a new proof that this degree is equal to $\binom{m-1}{k}$. The end of Section 2.2 extends this result and its proof technique to convex quadratic programs with linear constraints. Theorem 2.2.6 bounds the degree of the central curve of such programs when the objective function and the constraints are generic.

Based on work from [37], in Chapter 3 we compute the dimension of PSD_N^G, characterize its extremal rays, show it is Terracini convex, and in the case of $G = S_n$, we present the block in any symmetric matrix $Q \in \mathrm{PSD}_N^{S_n}$ corresponding to the trivial representation. Section 3.5 collects our results on binary and quadratic symmetric polynomials that are SOS. In the binary case, we compute the symmetry adapted matrix representations of all symmetric polynomials, and in the quadratic case, we do the same, and prove that, as the number of indeterminates tends to infinity, the ratio of SOS symmetric quadratic forms to all symmetric quadratic forms is $\frac{1}{8}$. Another interesting consequence obtained is that symmetric quadratic SOS polynomials in n variables can only be sums of 1, $n-1$ or n squares. In Section 3.6, we start with the classic case of ternary quartics, describing the associated symmetry adapted PSD cone. We then completely describe the geometric structure of the symmetry adapted

Gram spectrahedron for a generic, smooth, positive, symmetric ternary quartic including the rank of the matrices on its boundary. Further, we provide necessary conditions on the coefficients for a symmetric ternary quartic to be SOS. We continue the section by going up in degree and considering degree six symmetric polynomials in three variables. Here we show that the rank of a matrix in the symmetry adapted Gram spectrahedron of a generic symmetric ternary sextic will be at least 4.

In Chapter 4, using the machinery of Chapter 3, we study symmetric mean inequalities introduced in [24]. This chapter is based on the work done in [38]. Let h_λ be the homogeneous symmetric polynomial with respect to a partition λ. The term-normalized homogeneous symmetric polynomial is

$$H_\lambda(x) = \frac{h_\lambda(x)}{h_\lambda(1,\dots,1)}.$$

It was conjectured that

$$H_\lambda(x) \le H_\mu(x),\ x \ge 0 \Leftrightarrow \lambda \preceq \mu$$

where $x \ge 0$ means x is component-wise nonnegative and the partition order $\preceq$ is the usual dominance ordering. However, we provide counterexamples showing that $H_\lambda(x) \le H_\mu(x)$ can hold for incomparable pairs of partitions, μ and λ. Indeed we do this by proving that $H_\mu(x_1^2,\dots,x_n^2) - H_\lambda(x_1^2,\dots,x_n^2)$ is SOS for some incomparable pairs (λ,μ).

Chapter 2

The Degree of the Central Curve

The Zariski closure of the central path which interior point algorithms track in convex optimization problems such as linear, quadratic, and semidefinite programs is an algebraic curve. The degree of this curve has been studied in relation to the complexity of these interior point algorithms, and for linear programs it was computed by De Loera, Sturmfels, and Vinzant in 2012 [25], with a specific formula in the generic case. Semidefinite programming is a generalization of linear programming and thus a natural question to ask is what the degree of the central curve is for a generic SDP.

After some background on semidefinite programming, we show that the degree of the central curve for generic semidefinite programs is equal to the maximum likelihood degree of an associated linear concentration model. New results from the intersection theory of the space of complete quadrics imply that this is a polynomial in the size of semidefinite matrices with degree equal to the number of constraints. Besides its degree we explore the arithmetic genus of the same curve. For completeness, we also compute the degree of the central curve for generic linear programs with different techniques which extend to bounding the same degree for generic quadratic programs.

2.1 Preliminaries

Semidefinite Programming

In this section we go over key notions in semidefinite programming that will be required for the topics of this thesis. A semidefinite program (SDP) is a type of conic optimization problem that generalizes linear programming. It has a wide array of applications including control theory and dynamical systems, combinatorial optimization, and statistics. Thus semidefinite programming has been a topic of interest for several decades. In particular, it is necessary for understanding sums of squares of polynomials as we will see in Section 3.1. We begin with several definitions.

Let $\mathcal{S}^m_{\mathbb{R}}$ and $\mathcal{S}^m_{\mathbb{C}}$ be the vector spaces of $m \times m$ symmetric matrices with real and complex entries, respectively. If the underlying field is clear, we will at times denote the space of symmetric matrices simply as $\mathcal{S}^m$.

Definition 2.1.1. *A real symmetric matrix $A \in \mathcal{S}^m_{\mathbb{R}}$ is* positive semidefinite *if*

$$x^t A x \geq 0, \forall x \in \mathbb{R}^m$$

and we denote this by $A \succeq 0$. It is positive definite *if*

$$x^t A x > 0, \forall x \neq 0 \in \mathbb{R}^m$$

and we denote this by $A \succ 0$.

Definition 2.1.1 is not always the most useful characterization of positive semidefinite matrices. The following two propositions provide additional ways to determine if a matrix is positive semidefinite or positive definite.

Proposition 2.1.1. *Let $A \in \mathcal{S}^m_{\mathbb{R}}$ be a real symmetric matrix. The following are equivalent:*

1. *$A \succeq 0$.*
2. *$x^t A x \geq 0, \forall x \in \mathbb{R}^m$.*
3. *All the eigenvalues of A are nonnegative.*
4. *All the principle minors of A are nonnegative.*
5. *There exists a factorization $A = U^t U$ where U is a real $m \times m$ matrix.*

Proposition 2.1.2. *Let $A \in \mathcal{S}^m_{\mathbb{R}}$ be a real symmetric matrix. The following are equivalent:*

1. *$A \succ 0$.*
2. *$x^t A x > 0, \forall x \neq 0 \in \mathbb{R}^m$.*
3. *All the eigenvalues of A are positive.*
4. *All the principle minors of A are positive.*
5. *There exists a factorization $A = U^t U$ where U is a real invertible $m \times m$ matrix.*

The space of positive semidefinite matrices is in fact a cone called the positive semidefinite cone or PSD cone, denoted PSD_m. As a subset of $m \times m$ real symmetric matrices (isomorphic to $\mathbb{R}^{\binom{m+1}{2}}$), PSD_m is a full-dimensional closed convex cone in this vector space. It is a semi-algebraic set defined by $2^m - 1$ polynomial inequalities given by forcing the $2^m - 1$ principal minors of an $m \times m$ symmetric matrix to be nonnegative.

In semidefinite programming, we intersect the PSD cone with an affine subspace of symmetric matrices and optimize over the resulting convex body, called a spectrahedron. Recall that the standard Euclidean inner product on real symmetric matrices is $\langle Y, Z\rangle := \mathrm{Tr}(YZ) = \sum_{i,j} Y_{ij}Z_{ij}$. We may now present the primal SDP:

$$\begin{aligned} &\text{minimize } \langle C, X\rangle \\ &\text{subject to } \langle A_i, X\rangle = b_i, \quad i = 1, \dots, k \\ &\qquad X \succeq 0 \end{aligned} \tag{2.1}$$

where C and $A_i, i = 1, \dots, k$, are in $\mathcal{S}^m_{\mathbb{R}}$, and $b_i \in \mathbb{R}$ for $i = 1, \dots, k$.

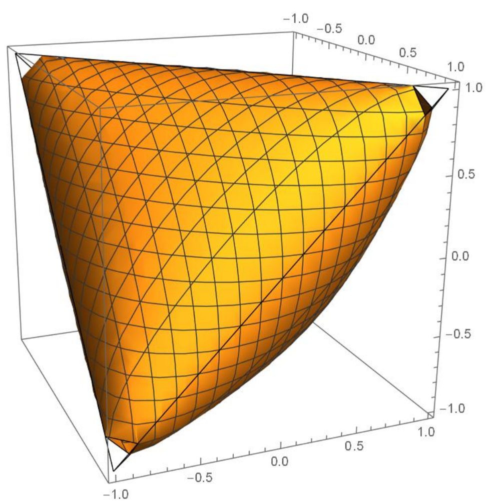
Example 2.1.1. *Consider the SDP with*

$$C = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and $b_i = 1$ *for* $i = 1, 2, 3$. *The feasible region of this SDP is*

$$\{X \in \mathcal{S}^3_{\mathbb{R}} : X = \begin{bmatrix} 1 & x_{12} & x_{13} \\ x_{12} & 1 & x_{23} \\ x_{13} & x_{23} & 1 \end{bmatrix} \succeq 0\}$$

and is carved out by the inequalities we get when we force the principle minors of X *to be nonnegative as in Proposition 2.1.1 (4). That is, we require that* $1 - x_{ij} \geq 0$ *and* $\det(X) \geq 0$. *The resulting convex body is affectionately called the samosa:*



The cost function $\langle C, X\rangle = -2x_{12} - 2x_{13} - 2x_{23}$ *is minimized by maximizing* $x_{12} + x_{13} + x_{23}$ *subject to the constraints. Note that* $1 - x_{ij} \geq 0$ *implies that* $x_{ij} \leq 1$. *Moreover, the determinant of*

$$X^* = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

is zero, thus X^* *is feasible and minimizes the cost function with an optimal value of -6.*

Remark 1. *In general, the set of positive semidefinite matrices with diagonal entries equal to 1 is called the elliptope. This object is of interest in some combinatorial problems and has been well studied (see for example [14, 45, 76]).*

In the particular case where $C = 0$ the SDP problem (2.1) reduces to determining if the feasible region is nonempty. This is called a *feasibility problem*.

Duality Theory

As with all conic programming, one can study semidefinite programming from a duality perspective. This section will briefly cover this duality theory for SDPs. First we define the dual problem to the primal SDP (2.1):

$$\begin{aligned} &\text{maximize } b^t y \\ &\text{subject to } \sum_{i=1}^{k} A_i y_i \preceq C \end{aligned} \tag{2.2}$$

where $b = (b_1, \dots, b_k)^t$ and $y = (y_1, \dots, y_k)$ is the decision vector. As with linear programming, the dual problem 2.2 is itself a semidefinite programming problem.

We always have weak duality, which means that the optimal solution value to the primal problem bounds the solution value to the dual problem and vice versa. To see this, let X^* be a feasible solution to the primal problem and y^* be a feasible solution to the dual problem. Then,

$$\langle C, X^* \rangle - b^t y^* = \langle C, X^* \rangle - \sum_{i=1}^{k} \langle A_i, X^* \rangle y_i^* = \langle C - \sum_{i=1}^{k} A_i y_i^*, X^* \rangle.$$

Now since X^* and y^* are feasible, $C - \sum_{i=1}^{k} A_i y_i^*$ and X^* are positive semidefinite. Thus their inner product must be nonnegative and we get weak duality, i.e. $\langle C, X^* \rangle \geq b^t y^*$. However, strong duality, where the optimal values of the primal and dual problems are equal, is not always guaranteed.

Example 2.1.2. *Consider the primal and dual problems defined by*

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$b_1 = 0$, *and* $b_2 = -1$. *The first constraint of the primal SDP,* $\langle A_1, X \rangle = b_1$, *implies that* $x_{11} = 0$ *which forces* $x_{12} = x_{13} = 0$. *Then the second constraint determines* $x_{22} = 1$, *thus the optimal value is 1.*

On the other hand, the dual problem is formulated as

$$\begin{aligned} &\textit{maximize } -y_2 \\ &\textit{subject to } \begin{bmatrix} 1 - y_1 & 0 & -y_2 \\ 0 & 1 + y_2 & 0 \\ -y_2 & 0 & 0 \end{bmatrix} \succeq 0 \end{aligned} \tag{2.3}$$

and a solution is only feasible if $y_2 = 0$. Thus the difference between the primal and dual optimal values, called the duality gap, is 1. Positive duality gap is an active area of research and the interested reader is encouraged to read [58].

Proposition 2.1.3. *Let X^* be a feasible solution to the primal SDP problem and y^* be a feasible solution to the dual SDP problem. Suppose further that*

$$(C - \sum_{i=1}^{k} A_i y_i^*) X^* = 0. \tag{2.4}$$

Then the cost functions of the primal and dual problem at X^* and y^* are equal (i.e. $\langle C, X^* \rangle = b^t y^*$) and they are optimal solutions.

Equation (2.4) is called complementary slackness. This together with the feasibility constrains of the dual and primal problems make what are called the Karush–Kuhn–Tucker conditions or KKT conditions for short:

- $(C - \sum_{i=1}^{k} A_i y_i) X = 0$
- $\langle A_i, X \rangle = b_i, \quad i = 1, \ldots, k$
- $X \succeq 0$
- $\sum_{i=1}^{k} A_i y_i \preceq C$

The KKT conditions are necessary for strong duality. Indeed one can check that complementary slackness fails in Example 2.1.2 above.

There are a few ways to ensure that strong duality holds. We include here the most commonly used condition known as Slater's condition.

Theorem 2.1.1. *Suppose that both the primal SDP (2.1) and dual SDP (2.2) are strictly feasible. Then both problems have optimal solutions whose optimal values are equal.*

Interior Point Method

Interior point algorithms are commonly used to solve semidefinite programming problems by solving a series of associated optimization problems. Each solution is a point within the spectrahedron (the feasible region of the SDP) and the set of solutions creates a path towards an extreme point of the spectrahedron. This path is called the central path. In this section we recall one interior point algorithm called the barrier method.

Given an SDP problem as (2.1), we consider the problem

$$\begin{aligned} &\text{minimize } \langle C, X \rangle - \lambda \log(\det(X)) \\ &\text{subject to } \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, k \end{aligned} \tag{2.5}$$

where $\lambda > 0$ is a real parameter. We first solve (2.5) for very large λ. Note that the original objective function $\langle C, X\rangle$ becomes negligible for sufficiently large λ. We then choose λ to be smaller by some appropriate step size and again solve (2.5). Continuing on as such and recording these solutions as λ goes to zero creates the central path.

As with the standard SDP, there are KKT conditions for (2.5):

$$\begin{aligned} & C - \lambda X^{-1} - \sum_{i=1}^{k} y_i A_i = 0, \\ & \langle A_i, X\rangle = b_i, \quad i = 1, \dots, k, \\ & X \succeq 0 \end{aligned} \tag{2.6}$$

where $y_1, \dots, y_k$ are the dual variables to the dual semidefinite program (2.2).

With the preliminary ingredients for semidefinite programming defined, we are now ready to consider the main contribution of this section of the thesis, the degree of the central curve of a generic SDP.

2.2 Central Curve

Throughout this section we will assume that the cost matrix C, the constraint matrices $A_1, \dots, A_k$, and $b = (b_1, \dots, b_k)^t$ are generic, unless otherwise stated. This assures, among other things, that if (2.1) is feasible, it is strictly feasible. The central curve of the primal SDP is obtained from the Karush-Kuhn-Tucker (KKT) conditions (2.6) to the auxiliary optimization problem (2.5). Hence we formally define the central curve as follows.

Definition 2.2.1. *Let $(X^*(\lambda), y^*(\lambda))$ be the unique solution of the system (2.6) for a fixed $\lambda > 0$. The (primal)* central curve *$\mathcal{C}_{SDP}(C, \{A_i\}, b)$ is the projection onto $\mathcal{S}^m_{\mathbb{C}}$ of the Zariski closure in $\mathcal{S}^m_{\mathbb{C}} \times \mathbb{C}^k$ of $\{(X^*(\lambda), y^*(\lambda)) : \lambda > 0\}$.*

The central curve contains the *central path* $\{X^*(\lambda) : \lambda > 0\}$. Interior point algorithms follow a piecewise linear approximation to the central path to obtain an optimal solution to (2.1) as λ approaches zero [13, 29, 30, 57, 55]. The degree of $\mathcal{C}_{SDP}(C, \{A_i\}, b)$ can be used to give an upper bound on the total curvature of the central path which is a heuristic measure on the number of steps interior point algorithms will take to find an optimal solution.

Interior point methods were first developed for linear programming problems, and the study of the central curve for linear programming from the perspective of algebraic geometry was initiated by Bayer and Lagarias in [4] and [3]. Dedieu, Malajovich, and Shub [27] studied the total curvature of the central path for linear programs in relation to bounding the number of iterations interior point algorithms take. By now we know that the total curvature can be exponential in the dimension of the ambient space [1]. Most relevant to our work, De Loera, Sturmfels, and Vinzant [25] obtained a breakthrough by computing the degree of the linear programming central curve. Given the linear program

$$\begin{aligned} \text{minimize } & cx \\ \text{subject to } & Ax = b \\ & x \geq 0, \end{aligned} \tag{2.7}$$

where $c \in \mathbb{R}^m$ is a row vector, A is $k \times m$ matrix of rank k, and $b \in \mathbb{R}^k$ is a column vector, they have related this degree to the degree of a *reciprocal variety* and a matroid invariant.

Theorem 2.2.1. *[25, Lemma 11] For generic b and c, the degree of the central curve of the linear program (2.7) is equal to the degree of the reciprocal variety*

$$\mathcal{L}_{A,c}^{-1} := \overline{\left\{(u_1, \dots, u_m) \in \mathbb{C}^m \,:\, \left(\frac{1}{u_1}, \dots, \frac{1}{u_m}\right) \in \operatorname{rowspan}\begin{pmatrix} A \\ c \end{pmatrix} \text{ and } u_i \neq 0, i = 1, \dots, m\right\}}$$

as well as the Möbius number $|\mu(A,c)|$ of the rank $k+1$ matroid associated to the row span of $\binom{A}{c}$. When A is also generic, the degree of the central curve is equal to $\binom{m-1}{k}$.

Our main contribution is Theorem 2.2.3 where we prove that the degree of the central curve for the SDP (2.1) when C and b are generic is equal to the maximum likelihood degree (ML degree) of the linear concentration model generated by $\{A_i\}$ and C. When $\{A_i\}$ are also generic, this degree is equal to the degree of the reciprocal variety associated to the linear subspace $\mathcal{L}_{\{A_i\},C} = \operatorname{span}\{A_1, \dots, A_k, C\}$:

$$\mathcal{L}_{\{A_i\},C}^{-1} := \overline{\left\{X \in \mathcal{S}_{\mathbb{C}}^m \,:\, X^{-1} \in \mathcal{L}_{\{A_i\},C}\right\}}.$$

We further show in Corollary 2.2.1 that, when $\{A_i\}, C$, and b are generic, the degree of $\mathcal{C}_{SDP}(C, \{A_i\}, b)$ is symmetric in the number of the linear equations defining (2.1). Corollary 2.2.2 concludes that in this case the degree of the central curve is a polynomial in m of degree k. This theorem and the two corollaries complete the work started in [65], proving Conjectures 4.3 and 4.4 in the same work.

Later in this section we report our observations on the arithmetic genus of $\mathcal{C}_{SDP}(C, \{A_i\}, b)$. We will also discuss semidefinite programs and the degree of their central curves associated to sum of squares (SOS) polynomials.

Finally at the end of the chapter we will revisit the degree of the central curve of the linear program (2.7) when A, c, and b are generic. Besides relating this degree to the ML degree of linear concentration models generated by diagonal matrices, in Theorem 2.2.5 we provide a different proof that this degree is equal to $\binom{m-1}{k}$. This result and its proof technique extend to convex quadratic programs with linear constraints. Theorem 2.2.6 bounds the degree of the central curve of such programs when the objective function and the constraints are generic.

Semidefinite Programs and Linear Concentration Models

In this section we consider the central curve $\mathcal{C}_{SDP}(C, \{A_i\}, b)$ when C and b are generic. In what follows, we describe the degree of this curve as the ML degree of a linear concentration model. When $\{A_i\}$ are also generic, we denote $\deg(\mathcal{C}_{SDP}(C, \{A_i\}, b))$ by $\psi_{SDP}(m, k)$.

Let $\mathcal{L}$ be a linear subspace of $\mathcal{S}^m_{\mathbb{R}}$ spanned by k linearly independent symmetric matrices $\{K_1, \ldots, K_k\}$. A linear concentration model is the set

$$\mathcal{L}^{-1}_{\succeq 0} := \{\Sigma \in \mathcal{S}^m_{\succeq 0} \, : \, \Sigma^{-1} \in \mathcal{L}\}$$

where $\mathcal{S}^m_{\succeq 0}$ is the cone of positive semidefinite matrices. Every matrix Σ in $\mathcal{L}^{-1}_{\succeq 0}$ is the covariance matrix of a multivariate normal distribution on $\mathbb{R}^m$, and the elements of $\mathcal{L}$ are concentration matrices.

Given a sample covariance matrix S, the maximum likelihood estimate $\hat{K}$ of S with respect to the linear concentration model defined by $\mathcal{L}$ is the unique positive semidefinite solution to the zero-dimensional polynomial equations

$$\Sigma K = Id_m, \qquad K \in \mathcal{L}, \qquad \Sigma - S \in \mathcal{L}^{\perp}. \tag{2.8}$$

The ML degree of this linear concentration model is defined as the number of solutions to (2.8) in $\mathcal{S}^m_{\mathbb{C}}$.

In [75] it was proven that when the matrices $K_1, \ldots, K_k$ are generic, the ML degree of the linear concentration model is precisely the degree of the reciprocal variety $\mathcal{L}^{-1}$.

Theorem 2.2.2. *[75, Theorem 2.3] The ML degree $\phi(m, k)$ of a linear concentration model defined by a generic linear subspace $\mathcal{L}$ of dimension d in $\mathcal{S}^m$ equals the degree of the projective variety $\mathcal{L}^{-1}$. This degree further satisfies*

$$\phi(m, k) = \phi\left(m, \binom{m+1}{2} + 1 - k\right).$$

Now we are ready to prove our main theorem.

Theorem 2.2.3. *Given an SDP as in (2.1) with C and b generic, $\deg(\mathcal{C}_{SDP}(C, \{A_i\}, b))$ is equal to the ML degree of the linear concentration model generated by $\mathcal{L} = \operatorname{span}\{C, A_1, \ldots, A_k\}$. If in addition $A_1, \ldots, A_k$ are generic, $\psi_{SDP}(m, k)$ is equal to the degree of $\mathcal{L}^{-1}$, and hence $\psi_{SDP}(m, k) = \phi(m, k + 1)$.*

Proof. By definition

$$\deg(\mathcal{C}_{SDP}(C, \{A_i\}, b)) = |\mathcal{C}_{SDP}(C, \{A_i\}, b) \cap \mathcal{H}|$$

where $\mathcal{H}$ is a generic hyperplane in $\mathcal{S}^m_{\mathbb{C}}$. Using the KKT conditions (2.6), the equations defining $\mathcal{C}_{SDP}(C, \{A_i\}, b) \cap \mathcal{H}$ are

$$\begin{aligned} X^{-1} &= \frac{1}{\lambda} C - \frac{1}{\lambda} \Sigma_{i=1}^k y_i A_i \\ \langle A_i, X\rangle - b_i &= 0, \quad i = 1, \ldots, k \\ \langle B, X\rangle - b_{k+1} &= 0, \end{aligned} \tag{2.9}$$

for some generic $B \in \mathcal{S}^m_{\mathbb{C}}$ and $b_{k+1} \in \mathbb{C}$.

The first equation in (2.9) means that $X^{-1} \in \mathcal{L}$, where $\mathcal{L} = \text{span}\{C, A_1, \ldots, A_k\}$. Since C is generic, in the last equation of (2.9) we can take $B = C$. Additionally, if we define S as a matrix such that $\langle A_i, S\rangle = b_i$, for $i = 1, \ldots, k$, and $\langle C, S\rangle = b_{k+1}$, the last $k+1$ equations in (2.9) mean that $X - S \in \mathcal{L}^{\perp}$. Note that these are precisely the likelihood equations of the linear concentration model determined by $\mathcal{L}$. This proves that $\deg(\mathcal{C}_{SDP}(C, \{A_i\}, b))$ is equal to the ML degree of the linear concentration model defined by $\mathcal{L}$. Additionally, if $A_1, \ldots, A_k$ are generic, Theorem 2.2.2 guarantees that $\phi(m, k+1)$ coincides with the degree of $\mathcal{L}^{-1}$, which means that $\psi(m, k)$ is equal to the degree of $\mathcal{L}^{-1}$ as well.

□

Corollary 2.2.1. *The degree of the central curve for a generic SDP satisfies*

$$\psi_{SDP}(m, k) = \psi_{SDP}\left(m, \binom{m+1}{2} - k - 1\right).$$

Proof.

$$\begin{aligned}\psi_{SDP}(m, k) &= \phi(m, k+1)\\ &= \phi(m, \binom{m+1}{2} + 1 - (k+1))\\ &= \phi(m, \binom{m+1}{2} - k))\\ &= \psi_{SDP}\left(m, \binom{m+1}{2} - k - 1\right).\end{aligned}$$

□

Corollary 2.2.2. *$\psi_{SDP}(m, k)$ is a polynomial in m of degree k.*

Proof. This result follows from the work of Michałek, Monin, Wiśniewski, Manivel, Seynaeve, and Vodička who employed the space of complete quadrics and intersection theory to prove the polynomiality of $\phi(m, k)$ ([53] and [50, Theorem 1.3]) and from the seperate work of Cid-Ruiz [20, Corollary C]. □

Arithmetic Genus

The ideal of polynomials $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ in $\mathbb{C}[x_{ij} : 1 \le i \le j \le m]$ vanishing on the reciprocal variety $\mathcal{L}^{-1}_{\{A_i\},C}$ is a prime ideal since this variety is irreducible. The proof of Theorem 2.2.2 (see [75, Theorem 2.3]) relies on the fact that $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ is Cohen-Macaulay when $\{A_i\}$ and C are generic [41, 44]. Since the central curve $\mathcal{C}_{SDP}(C, \{A_i\}, b)$ is obtained from intersecting the reciprocal variety with d generic linear equations in (2.6), the numerator of the Hilbert series of $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ and that of the defining ideal of the the central curve are identical. The Hilbert series for the central curve will be of the form

$$\frac{h_0 + h_1 t + h_2 t^2 + \cdots + h_l t^l}{(1-t)^2}$$

where the coefficients h_j are nonnegative integers with $h_0 = 1$ and $h_l \neq 0$. The arithmetic genus of the central curve can be calculated as

$$\text{genus}(m,k) := \text{genus}(\mathcal{C}_{SDP}(C, \{A_i\}, b)) = 1 - \sum_{j=0}^{l} (1-j)h_j.$$

The following table shows genus(m,k) for all values we can compute with Macaulay2 [36] and/or using the two propositions that follow.

$m\backslash k$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	0	0												
3	0	0	1	0	0									
4	0	1	10	20	22	20	10	1	0					
5	0	3										33	3	0

Proposition 2.2.1. *For $m \geq 2$,*

$$\text{genus}(m,1) = \text{genus}\left(m, \binom{m+1}{2} - 1\right) = 0.$$

In these cases, the central curve is a rational curve. Furthermore, when $k = 1$ the numerator of the Hilbert series is $1 + (m-2)t$, and when $k = \binom{m+1}{2} - 1$ it is 1.

Proof. In the case $k = \binom{m+1}{2} - 1$, the reciprocal variety is equal to $\mathbb{P}^k$, and therefore the central curve is $\mathbb{P}^1$. In the case $k = 1$, the reciprocal variety is the image of span$\{C, A_1\} \simeq \mathbb{P}^1$ under the rational map given by the $(m-1)$-minors of a generic $m \times m$ symmetric matrix. Hence it is a rational curve of degree $m-1$. This implies that the numerator of the Hilbert series of the ideal defining the reciprocal variety, and therefore that of the central curve, is $1 + (m-2)t$. This means that the central curve is also a rational curve, i.e., its genus is equal to zero. □

Proposition 2.2.2. *For $m \geq 2$,*

$$\text{genus}\left(m, \binom{m+1}{2} - 2\right) = \binom{m-2}{2}$$

and

$$\text{genus}\left(m, \binom{m+1}{2} - 3\right) = 1 + (m-1)^2(m-3).$$

Proof. In the first case, the reciprocal variety is a hypersurface defined by a single polynomial of degree $m-1$. Therefore the numerator of the Hilbert series is equal to $1+t+\cdots+t^{m-2}$. Therefore the arithmetic genus of the central curve is

$$1-\sum_{j=0}^{m-2}(1-j)=\sum_{j=1}^{m-3}j=\binom{m-2}{2}.$$

In the second case, the reciprocal variety is of codimension two, and it is a complete intersection generated by two degree $m-1$ generators; see [75, p. 611] and Lemma 2.2.1 below. Therefore the numerator of the Hilbert series is equal to $(1+t+\cdots+t^{m-2})^2 = 1+2t+\cdots+(m-2)t^{m-3}+(m-1)t^{m-2}+(m-2)t^{m-1}+\cdots+2t^{2m-5}+t^{2m-4}$. Using the formula for the arithmetic genus first yields $1+(2m-6)\binom{m-1}{2}+(m-3)(m-1)$. This in turn is equal to $1+(m-3)(m-1)^2$. □

Lemma 2.2.1. *When $k=\binom{m+1}{2}-3$, the reciprocal variety $\mathcal{L}^{-1}_{\{A_i\},C}$ associated to a generic linear subspace $\mathcal{L}_{\{A_i\},C}$ is a complete intersection of codimension two generated by two polynomials of degree $m-1$.*

Proof. Let V be the variety of codimension 3 in $\mathbb{P}^{\binom{m+1}{2}-1}$ defined by the $(m-1)$-minors of a generic $m\times m$ symmetric matrix, and let X be the quasiprojective variety $\mathbb{P}^{\binom{m+1}{2}-1}\setminus V$. Consider the regular map $F : X \longmapsto \mathbb{P}^{\binom{m+1}{2}-1}$ given by the $(m-1)$-minors of a generic $m\times m$ symmetric matrix. Given the generic codimension two subspace $\mathcal{L}_{\{A_i\},C}$, the inverse image $F^{-1}(\mathcal{L}_{\{A_i\},C})$ is an irreducible subvariety of X by Bertini's theorem [46, Theorem 3.3.1]. This subvariety is defined by two generic linear combinations of $(m-1)$-minors, f_1 and f_2, which are of degree $m-1$. The variety in $\mathbb{P}^{\binom{m+1}{2}-1}$ defined by the same two polynomials is a complete intersection of codimension two. This variety contains the reciprocal variety which is irreducible and has also codimension two. Therefore if the ideal $\langle f_1, f_2\rangle$ is prime it has to be the defining ideal of the reciprocal variety. But this is the case, since it is a complete intersection and hence all its components have the same codimension. Any component other than the one coming from $F^{-1}(\mathcal{L}_{\{A_i\},C})$ is associated to V, but V has codimension three. □

We note that in the above table the entry for $m=5$ and $k=12$ is computed using Proposition 2.2.2. However, the entry for $m=5$ and $k=3$, which is conjecturally equal to 33 is missing. Nevertheless, we venture to state the following conjecture.

Conjecture 1. $\text{genus}(m,k)=\text{genus}\left(m,\binom{m+1}{2}-k\right)$.

Although we cannot prove this conjecture, we can prove the analogous statement for the central curve of linear programs (2.7) when A, c and b are generic. The central curve for linear programs is defined as in Definition 2.2.1 but using the KKT conditions for linear programs; see (2.13) below.

Theorem 2.2.4. *Let A_k and A_{m-k} be generic matrices of size $k \times m$ and $(m-k) \times m$ and of rank k and $m-k$, respectively. Let b_k and b_{m-k} be two generic vectors in $\mathbb{R}^k$ and $\mathbb{R}^{m-k}$. The central curve of the linear program defined by A_k, b_k, and a generic vector c has the same arithmetic genus as the central curve of the linear program defined by A_{m-k}, b_{m-k} and c.*

Proof. Let $\mathcal{C}_{LP}(k)$ and $\mathcal{C}_{LP}(m-k)$ denote the central curve of the generic linear programs as in the statement. In this generic case, from [25] we have

$$\text{genus}(\mathcal{C}_{LP}(k)) = 1 - \sum_{j=0}^{k} (1-j) \binom{m-k+j-2}{j}, \tag{2.10}$$

$$\text{genus}(\mathcal{C}_{LP}(m-k)) = 1 - \sum_{j=0}^{m-k} (1-j) \binom{k+j-2}{j}, \tag{2.11}$$

where the binomial coefficients in each equation come from the coefficients of the Hilbert series computed in [25]. To check that both computations have the same value, we need the identities

$$\sum_{j=0}^{n} \binom{r+j}{j} = \binom{r+n+1}{n} \quad \text{ank} \quad \sum_{j=0}^{k} j \binom{m-k+j-2}{j} = (m-k-1)\binom{m-1}{k-1}.$$

First we get

$$\begin{aligned}
\text{genus}(\mathcal{C}_{LP}(k)) =& 1 - \sum_{j=0}^{k} \binom{m-k+j-2}{j} + \sum_{j=0}^{k} j \binom{m-k+j-2}{j} \\
=& 1 - \binom{m-k-2+k+1}{k} + (m-k-1)\binom{m-1}{k-1} \\
=& 1 - \binom{m-1}{k} + (m-k-1)\binom{m-1}{k-1} \\
=& 1 - \frac{(m-1)!}{(m-1-k)!k!} + (m-k-1)\frac{(m-1)!}{(m-1-k+1)!(k-1)!} \\
=& 1 - \frac{(m-1)!(m-mk+k^2)}{(m-k)!k!}
\end{aligned}$$

where the second line comes from the identities mentioned above with $r = m-k-2$. Doing a similar computation for genus($\mathcal{C}_{LP}(m-k)$) we get

$$\begin{aligned}\text{genus}(\mathcal{C}_{LP}(m-k)) =& 1 - \sum_{j=0}^{m-k} \binom{m-(m-k)+j-2}{j} + \sum_{j=0}^{m-k} j \binom{m-(m-k)+j-2}{j} \\ =& 1 - \binom{k-2+m-k+1}{m-k} + (k-1)\binom{m-1}{m-1-k} \\ =& 1 - \binom{m-1}{m-k} + (k-1)\binom{m-1}{m-1-k} \\ =& 1 - \frac{(m-1)!}{(k-1)!(m-k)!} + (k-1)\frac{(m-1)!}{k!(m-1-k)!} \\ =& 1 - \frac{(m-1)!(m-mk+k^2)}{(m-k)!k!}.\end{aligned}$$

□

Sum of Squares Polynomials

Before considering other types of convex optimization problems, we consider the semidefinite programs for sums of squares problems. More details on sums of squares can be found in Section 3.1, but we briefly review this here. Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a homogeneous polynomial of degree $2d$ and let L_p be the affine subspace of symmetric matrices Q satisfying the identity

$$p = [x]^T Q[x] \tag{2.12}$$

where $[x]$ is a vector of all monomials of degree d in n variables. The intersection of L_p with the cone of positive semidefinite matrices is the Gram spectrahedron of p, and it is nonempty if and only if p is a sum of squares (SOS) polynomial. That is, certifying that a polynomial is SOS reduces to checking the feasibility of an SDP. This can be achieved by solving an SDP using a random (generic) cost matrix C.

Example 2.2.1. *Suppose we wish to show that a generic ternary quartic is an SOS. The A_i's and b_i's come from equating coefficients in (2.12). For example, if we let*

$$[x] = [x^2, xy, xz, y^2, yz, z^2]$$

the linear equation for the x^2y^2 term will be

$$p_{(2,2,0)} = \langle A_{(2,2,0)}, Q\rangle$$

where $p_{(2,2,0)}$ is the x^2y^2 coefficient of the random ternary quartic p,

$$A_{(2,2,0)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and Q is the decision variable of the SDP which will have $k = \binom{3+4-1}{4} = 15$ constraints with matrices of size $m = \binom{3+2-1}{2} = 6$.

In general, the matrices $\{A_i\}$ for the linear constraints will be sparse and far from generic. However, if the polynomial p that we want to certify to be an SOS polynomial is generic, then the b_i's in the corresponding SDP will be also generic. Using a generic cost matrix C in this SDP allows us to consider the degree of the central curve for a generic SOS polynomial.

We wish to report our computations in three instances: binary sextics ($n = 2, 2d = 6$), binary octics ($n = 2, 2d = 8$), and ternary quartics ($n = 3, 2d = 4$). The corresponding SDPs are given by input data with $m = 4, k = 7$ for binary sextics, $m = 5, k = 9$ for binary octics, and $m = 6, k = 15$ for ternary quartics. We note that for the same size SDPs with generic $\{A_i\}$ we will obtain $\psi_{SDP}(4, 7) = 9$, $\psi_{SDP}(5, 9) = 137$, and $\psi_{SDP}(6, 15) = 528$. We believe that studying this invariant for various families of SOS polynomials is an interesting future project.

Proposition 2.2.3. *The degrees of the central curves for SDPs associated to generic binary sextics, binary octics, and ternary quartics, where generic cost matrices are used, are* 7*,* 45*, and* 66*, respectively.*

As a last remark about the SDP arising from sums of squares, we would like to mention that since C and b are generic, the computations in the previous proposition also correspond to the ML degree of a linear concentration model. Namely, the concentration model defined by catalecticants and an additional generic matrix corresponding to the cost matrix. Exploring this relation is also an interesting future project.

Linear Programs

By choosing C and $\{A_i\}$ in (2.1) to be diagonal matrices we recover linear programs (2.7). The central curve $\mathcal{C}_{LP}(c, A, b)$ for such a linear program can be defined as in the case of the central curve for a semidefinite program using the corresponding KKT conditions:

$$\begin{aligned} & c - \lambda \left(\frac{1}{x_1}, \ldots, \frac{1}{x_m} \right) - y^t A = 0 \\ & Ax = b, \\ & x \geq 0. \end{aligned} \tag{2.13}$$

When in the data defining (2.7), c and b are generic the degree of the central curve $\mathcal{C}_{LP}(c, A, b)$ is equal to the degree of the reciprocal variety $\mathcal{L}_{A,c}^{-1}$ [25, Lemma 11]. Further, if A is also generic, this degree is equal to $\binom{m-1}{k}$. For the case when all the data is generic, we will denote the degree of the linear programming central curve by $\psi_{LP}(m, k)$.

The observations that connect the ML degree of generic linear concentration models to the degree of the central curve of generic semidefinite programs have their counterpart here as well. One can consider the ML degree of linear concentration models generated by diagonal matrices as in [75, Section 3]. For generic models we denote the ML degree by $\phi_{\text{diag}}(m, k)$. A consequence of Corollary 3 in [75] is the following.

Corollary 2.2.3.

$$\phi_{\text{diag}}(m, k) \quad = \quad \binom{m-1}{k-1}.$$

An argument parallel to the one used in the proof of Theorem 2.2.3 gives

Corollary 2.2.4. $\psi_{LP}(m, k) = \phi_{\text{diag}}(m, k+1)$.

In the rest of this section we will develop another method to prove that $\psi_{LP}(m, k) = \binom{m-1}{k}$. This method will be extended for bounding the degree of the central curve for generic convex quadratic programs with linear constraints in the next section. We note that our techniques which are based on counting solutions to polynomial systems were employed for a similar purpose in [27].

First we consider the polynomial system obtained by clearing denominators and dropping the $x \geq 0$ condition in (2.13):

$$\begin{aligned} &c_i x_i - \lambda - (y^t a_i) x_i = 0, \quad i = 1, \ldots, m \\ &Ax = b \end{aligned} \tag{2.14}$$

where a_i is the ith column of the matrix A. For generic data, the central curve is obtained as the Zariski closure in $\mathbb{C}^m$ of the projection of the solution set in $(\mathbb{C}^*)^{m+k+1}$ to the equations (2.14). Further, the degree of this central curve would be equal to the number of points in $(\mathbb{C}^*)^m$ obtained as the intersection of the central curve with a generic hyperplane defined by $ex = f$.

Lemma 2.2.2. *The degree of $\mathcal{C}_{LP}(c, A, b)$ for generic c, A, and b is equal to the number of solutions in $(\mathbb{C}^*)^{m+k+1}$ to the system (2.14) together with an extra equation of the form $ex = f$ where the coefficients of this equation are generic.*

Proof. Clearly, every solution to (2.14) plus $ex = f$ in $(\mathbb{C}^*)^{m+k+1}$ projects to a point in $\mathcal{C}_{LP}(c, A, b) \cap \{x : ex = f\}$. Conversely, the genericity of $ex = f$ implies that the points in $\mathcal{C}_{LP}(c, A, b) \cap \{x : ex = f\}$ come from points in $(\mathbb{C}^*)^{m+k+1}$ that satisfy (2.14) and $ex = f$. We show that for each point x^* "downstairs" there is a unique point "upstairs". Suppose there are at least two points (x^*, y^*, λ^*) and (x^*, z^*, μ^*) with these properties. Then it is easy to check that $(x^*, ty^* + (1-t)z^*, t\lambda^* + (1-t)\mu^*)$ is also a solution with the same properties for any t. But this is a contradiction since we have only finitely many preimages by the genericity of the data. □

This lemma implies that in order to compute the degree of $\mathcal{C}_{LP}(c, A, b)$ for generic A, c, and b we need to count the solutions in $(\mathbb{C}^*)^{m+k+1}$ to

$$\begin{aligned} &c_i x_i - \lambda - (y^t a_i) x_i = 0, \quad i = 1, \dots, m \\ &Ax = b \\ &ex = f \end{aligned} \tag{2.15}$$

where $ex = f$ is also generic. Note that the rank of the matrix $\begin{pmatrix} A \\ e \end{pmatrix}$ is $k+1$ and the solutions to the last $k+1$ equations in (2.15) can be parametrized by

$$x = v_0 + t_1 v_1 + \cdots + t_{m-k-1} v_{m-k-1}$$

where $v_0, v_1, \dots, v_{m-k-1}$ are generic vectors. Substituting this into the first m equations in (2.15) we obtain m equations in m variables $\lambda, y_1, \dots, y_k, t_1, \dots, t_{m-k-1}$. Furthermore, the genericity assumptions guarantee that each equation will have support equal to

$$\lambda, 1, t_1, \dots, t_{m-k-1}, y_1, y_1 t_1, \dots, y_1 t_{m-k-1}, \dots, y_k, y_k t_1, \dots, y_k t_{m-k-1}.$$

The Newton polytope of a polynomial with this support is a pyramid of height one with base equal to the product of simplices $\Delta_{m-k-1} \times \Delta_k$.

Theorem 2.2.5. *$\psi_{LP}(m, k)$ is equal to the volume of $\Delta_{m-k-1} \times \Delta_k$:*

$$\binom{m-1}{k} = \sum_{j=0}^{m-k-1} \binom{m-j-2}{k-1}$$

Proof. The above lemma and the previous discussion imply that $\psi_{LP}(m, k)$ is equal to the number solutions in $(\mathbb{C}^*)^m$ to m equations in m variables, where each equation has support equal to the set of monomials listed above. Bernstein's Theorem implies that this number is bounded above by the normalized volume of the Newton polytope of these monomials. Since this polytope is a pyramid of height one over $\Delta_{m-k-1} \times \Delta_k$, we just need to compute the normalized volume of the product of simplices. Further, because every triangulation of $\Delta_{m-k-1} \times \Delta_k$ is unimodular we just need to count the number of simplices in any triangulation. One such triangulation is the staircase triangulation. The maximal simplices in this triangulation are described as follows. Consider a $(m-k) \times (k+1)$ rectangular grid. The simplices in the staircase triangulation of $\Delta_{m-k-1} \times \Delta_k$ are in bijection with paths from the northwest corner of this grid to the southeast corner where a path consists of steps in the east or south direction. The total number of steps in each path is $m-1$, and out of these steps k have to be south steps. Therefore there are a total of $\binom{m-1}{k}$ such paths. These paths can be partitioned into those which reach the south edge of the grid j steps before the southeast corner where $j = 0, \ldots, m-k-1$. The number of these kinds of paths for each j is $\binom{m-j-2}{k-1}$. Finally, the proof of Lemma 11 in [25] implies that $\psi_{LP}(m, d) \geq \binom{m-1}{d}$, and this concludes the proof. □

Quadratic Programs

To complete our study of central curves in optimization problems we will now consider convex quadratic programs with linear constraints.

$$\begin{aligned} \text{minimize } & \frac{1}{2}x^t Q x + cx \\ \text{subject to } & Ax = b \\ & x \geq 0, \end{aligned} \tag{2.16}$$

where Q is an $m \times m$ positive definite matrix, $c \in \mathbb{R}^m$, A is $k \times m$ matrix of rank k, and $b \in \mathbb{R}^k$. The KKT conditions that lead to the definition of the central curve are

$$\begin{aligned} & x^t Q + c - \lambda \left(\frac{1}{x_1}, \ldots, \frac{1}{x_m} \right) - y^t A = 0 \\ & Ax = b, \\ & x \geq 0. \end{aligned} \tag{2.17}$$

When Q, c, A, and b are generic, we denote by $\psi_{QP}(m, k)$ the degree of the central curve for generic quadratic programs. One can show by a homotopy continuation argument that it is sufficient to assume Q to be a generic diagonal matrix. For precise details of this result, we refer the reader to [71, Section 3.2]. With $Q = \text{diag}(q_1, \ldots, q_m)$, after clearing denominators and ignoring the nonnegativity constraints $x \geq 0$ in (2.17), we arrive to the following system of polynomial equations:

$$\begin{aligned} & q_i x_i^2 + c_i x_i - \lambda - (y^t a_i) x_i = 0 \quad i = 1, \ldots, m \\ & Ax = b, \end{aligned} \tag{2.18}$$

where a_i is the ith column of the matrix A. As in the linear programming case we have the following lemma.

Lemma 2.2.3. *$\psi_{QP}(m, k)$, the degree of the central curve of a generic quadratic program is equal to the number of solutions in $(\mathbb{C}^*)^{m+k+1}$ to the system (2.18) together with an extra equation of the form $ex = f$ where the coefficients of this equation are also generic.*

Proof. The proof of this lemma is identical to the proof of Lemma 2.2.2. □

Theorem 2.2.6.

$$\psi_{QP}(m,k) \leq \sum_{j=0}^{m-k-1} \binom{m-j-2}{k-1} 2^j.$$

This is the volume of the Newton polytope of a polynomial with support in monomials

$$\lambda, 1, t_1, \ldots, t_{m-k-1}, t_1^2, t_1t_2, \ldots, t_{m-k-1}^2$$

$$y_1, y_1t_1, \ldots, y_1t_{m-k-1}, \ldots, y_k, y_kt_1, \ldots, y_kt_{m-k-1}$$

Proof. By Lemma 2.2.3 and as in the proof of Theorem 2.2.5 we need to count solutions to (2.18) plus a generic linear equation $ex = f$ in the torus $(\mathbb{C}^*)^{m+k+1}$. The solutions to the equations $Ax = b$ and $ex = f$ can again be parametrized as

$$x = v_0 + t_1v_1 + \cdots + t_{m-k-1}v_{m-k-1}$$

where $v_0, \ldots, v_{m-k-1}$ are generic vectors. Substituting this into the first m equations in (2.18) we obtain m equations in m variables $\lambda, y_1, \ldots, y_k, t_1, \ldots, t_{m-k-1}$. Furthermore, the genericity assumptions guarantee that each equation will have support equal to

$$\lambda, 1, t_1, \ldots, t_{m-k-1}, t_1^2, t_1t_2, \ldots, t_{m-k-1}^2$$

$$y_1, y_1t_1, \ldots, y_1t_{m-k-1}, \ldots, y_k, y_kt_1, \ldots, y_kt_{m-k-1}$$

The number of solutions to these m equations in $(\mathbb{C}^*)^m$ is bounded by the normalized volume of the Newton polytope of the above monomials. Since this is a pyramid of height one, we just need to compute the volume of the Newton polytope of the monomials except λ. This polytope has a staircase triangulation as for $\Delta_{m-k-1} \times \Delta_k$ where each simplex corresponds to a path as we described in the proof of Theorem 2.2.5, except that the volume of a simplex corresponding to a path which reaches the south edge of the grid j steps before the southeast corner is 2^j. Therefore $\psi_{QP}(m,k)$ is at most

$$\sum_{j=0}^{m-k-1} \binom{m-j-2}{k-1} 2^j.$$

□

Remark 2. *There are many more specializations of semidefinite programming that one may consider. For example, one can consider the degree of the central curve in second-order conic programming (SOCP) which can be reformulated as a sparse SDP. Another special case is convex quadratically constrained quadratic programs (QCQP) which can be written as an SOCP. However, the resulting SDP will no longer have generic data* $\{C, A_1, \ldots, A_k, b\}$*. Indeed, using computer software such as Macaulay2 one can compute these degrees for small cases using the appropriate KKT conditions and projecting onto the primal space. In doing so one can see that they do not generally coincide with the ML degree. Like in the linear programming and quadratic programming case above, one must find a different theory to determine the degree of the central curve in these specializations, an interesting potential future project.*

Chapter 3

Symmetry Adapted Gram Spectrahedra

In this chapter we study the spectrahedra that arise in the theory of symmetric and sums of squares (SOS) polynomials. For a finite group G, we are interested in sums of squares polynomials which are G-invariant. Section 3.1 offers precise definitions and background for sums of squares and representation theory. We then go on to define what is called a symmetry adapted basis and use this to construct the symmetry adapted PSD cone. While Section 3.2 goes into more detail, we briefly describe the process here.

We start with a representation of G on $\mathbb{R}^n$, extending by linear substitution to a representation $D : G \to GL(V)$ on $V = \mathbb{R}[x_1, \ldots, x_n]_d$, the vector space of degree d homogeneous polynomials in n indeterminates. The dimension of V is $N = \binom{n+d-1}{d}$, and we denote the cone of $N \times N$ positive semidefinite matrices by PSD_N. Choosing a basis for V gives matrices $D(g)$, and we obtain the *symmetry adapted* version of PSD_N, namely,

$$\mathrm{PSD}_N^G := \left\{ Q \in \mathrm{PSD}_N \;\middle|\; D(g)^T Q D(g) = Q, \text{ for all } g \in G \right\}.$$

For a given polynomial f of degree $2d$ which is invariant under the action of a group G, the *symmetry adapted Gram spectrahedron* of f is the closed, convex, semi-algebraic set

$$K_f^G := L_f \cap \mathrm{PSD}_N^G.$$

Here, L_f is the linear space of symmetric matrices Q which represent f as $f(x) = [x]^T Q[x]$, and $[x]$ is a column vector whose entries form a basis for V, usually chosen to be all monomials of degree d in the variables $x_1, \ldots, x_n$.

The *Gram spectrahedron* $K_f = L_f \cap \mathrm{PSD}_N$ for a polynomial f is a set parameterizing all ways to write f as a sum of squares. Its geometry is important for understanding sums of squares representations of f. For example, the matrices of lowest rank contained in K_f encode the ways to write f as a sum of a minimal number of squares. These matrices of lowest rank are extremal points of K_f. Characterizing the minimal number of squares is a topic that has been widely studied [15, 18, 19, 52, 66, 69, 77]. The symmetry adapted Gram spectrahedron K_f^G is a smaller and simpler convex set for which we can ask similar questions. It was introduced in [32] and has since been used in a variety of applications [2, 63, 62].

In this thesis, we mainly focus on the case $G = S_n$, the symmetric group, and the polynomials we consider will be the usual symmetric polynomials [47]. Section 3.1 offers a brief summary of the background needed from representation theory and SOS polynomials. In Section 3.2 we discuss what is called a symmetry adapted basis which will allow us to construct the symmetry adapted PSD cone. Section 3.3 then goes on to define the symmetry adapted Gram spectrahedron. In Section 3.4 we return to the symmetry adapted cone PSD_N^G. In particular, we compute the dimension of PSD_N^G, characterize its extremal rays, prove it is Terracini convex, and in the case of $G = S_n$, we present the block in any symmetric matrix $Q \in \mathrm{PSD}_N^{S_n}$ corresponding to the trivial representation.

The remainder of the chapter collects results about symmetric polynomials for particular n and d. Section 3.5 focuses on binary ($n = 2$) and quadratic ($2d = 2$) symmetric polynomials that are SOS. In the binary case, we compute the symmetry adapted matrix representations of all symmetric polynomials, and in the quadratic case, we do the same, and prove that, as the number of indeterminates tends to infinity, the ratio of SOS symmetric quadratic forms to all symmetric quadratic forms is $\frac{1}{8}$. Another interesting consequence obtained is that symmetric quadratic SOS polynomials in n variables can only be sums of 1, $n-1$ or n squares. In Section 3.6, we start with the classic case of ternary quartics, describing the associated symmetry adapted PSD cone. We then completely describe the geometric structure of the symmetry adapted Gram spectrahedron for a generic, smooth, positive, symmetric ternary quartic including the rank of the matrices on its boundary. Further, we provide necessary conditions on the coefficients for a symmetric ternary quartic to be SOS. We end the section by going up in degree and considering degree six symmetric polynomials in three variables. Here we show that the rank of a matrix in the symmetry adapted Gram spectrahedron of a generic symmetric ternary sextic will be at least 4.

3.1 Preliminaries

In this section, we review some basic facts about sums of squares and representation theory. This allows us to discuss invariant semidefinite programs that arise from proving nonnegativity of invariant polynomials. In particular, we will be interested in symmetric polynomials, which are invariant under the S_n action of permuting the variables.

Sums of Squares

We begin with a few definitions and a bit of notation. Let $\mathbb{R}[x]_{\leq d} = \mathbb{R}[x_1, \dots, x_n]_{\leq d}$ be the space of real polynomials in n variables of degree at most d. A polynomial p in $\mathbb{R}[x_1, \dots, x_n]_{\leq 2d}$ is said to be a *sums of squares* (SOS) polynomial if $p = q_1^2 + \dots + q_r^2$

where $q_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}$, $i = 1, \ldots, r$. That is, a polynomial is SOS if it can be written as a sum of polynomials squared.

As we discussed in Section 1.1, sums of squares is a common method for proving nonnegativity of a polynomial as an SOS polynomial must be nonnegative. Moreover, certifying SOS-ness can be reduced to a semidefinite program, as we see in the following theorem.

Theorem 3.1.1. *Let $f(x) \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ be a polynomial of degree $2d$ and let $[x]$ be a column vector consisting of a basis of $\mathbb{R}[x_1, \ldots, x_n]_{\leq d}$. Then $f(x)$ is a sum of squares if and only if there exists an $N \times N$ real positive semidefinite matrix Q where $N = \binom{n+d}{d}$ such that*

$$f(x) = [x]^t Q [x]. \tag{3.1}$$

Example 3.1.1. *Consider the polynomial $f(x) = 5x^2 - 2x + 2$. We take the monomial basis as our basis of $\mathbb{R}[x]_{\leq 1}$, letting $[x] = \begin{bmatrix} 1 & x \end{bmatrix}^t$. Then by Theorem 3.1.1 we must find a positive semidefinite matrix $Q = (q_{ij})$ such that*

$$5x^2 - 2x + 2 = \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}.$$

Expanding the right-hand side to get

$$5x^2 - 2x + 2 = q_{11} + 2q_{12}x + q_{22}x^2$$

and equating coefficients defines the affine subspace of symmetric matrices, $q_{11} = 5$, $2q_{12} = -2$, and $q_{22} = 2$. Indeed

$$Q = \begin{bmatrix} 5 & -1 \\ -1 & 2 \end{bmatrix}$$

is positive semidefinite and thus we certify that f is SOS.

The set of all positive semidefinite matrices Q that satisfy (3.1) for a given f is called the *Gram spectrahedron* of f. Gram spectrahedra have been studied intensively in [12, 18, 19, 32, 59], to name a few. By Theorem 3.1.1 it is precisely the affine subspace defined by (3.1) intersected with the PSD cone, i.e. the feasible region of an SDP. Thus we see that checking if a polynomial is SOS reduces to an SDP, more specifically, a feasibility problem.

We omit the proof for Theorem 3.1.1, but note that the key ingredient is that a positive semidefinite matrix can be factored as in Proposition 2.1.1 (5). Thus f can be written as $(U[x])^t(U[x])$ which gives us the sum of squares decomposition. While the factorization is not unique, given a Q of rank r, we will always get an SOS decomposition of f with r squares.

In the example above, this means we must get a sum of two squares:

$$\begin{aligned} 5x^2 - 2x + 2 &= \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} 5 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} \\ &= \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} \\ &= \begin{bmatrix} x-2 & x+1 \end{bmatrix} \begin{bmatrix} x-2 \\ x+1 \end{bmatrix} \\ &= (x-2)^2 + (x+1)^2. \end{aligned}$$

For the remainder of this thesis, we will assume our polynomials are homogeneous, meaning that each term is of the same degree. We can always homogenize a polynomial $f(x_1, \dots, x_n)$ of degree d by introducing a new variable x_{n+1} and defining a new polynomial

$$\bar{f} = x_{n+1}^d f(\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}}).$$

If f is positive, then so is $\bar{f}$. Likewise for SOS polynomials. Moreover, if $p = \sum_i q_i^2$ is homogeneous of degree $2d$, the polynomials q_i will be homogeneous of degree d and furthermore the vector $[x]$ in Equation (3.1) will only need to contain homogeneous polynomials of degree d. Thus we restrict our study to homogeneous polynomials, also called forms, and denote the space of forms of degree d by $\mathbb{R}[x]_d = \mathbb{R}[x_1, \dots, x_n]_d$.

Note that the dimension of $\mathbb{R}[x]_d$ is $\binom{n+d-1}{d}$. We quickly restate the theorems and definitions from earlier in the context of homogeneous polynomals.

Theorem 3.1.2. *Let $f(x) \in \mathbb{R}[x_1, \dots, x_n]_{2d}$ be a homogeneous polynomial and let $[x]$ be a column vector consisting of a basis of $\mathbb{R}[x_1, \dots, x_n]_d$. Then $f(x)$ is a sum of squares if and only if there exists an $N \times N$ real positive semidefinite symmetric matrix Q where $N = \binom{n+d-1}{d}$ and*

$$f(x) = [x]^T Q[x]. \tag{3.2}$$

Now the vector $[x]$ only contains homogeneous polynomials of degree d and N is the number of monomials of degree exactly d. As before the set of $N \times N$ real symmetric matrices $\mathcal{S}^N$ is a vector space isomorphic to $\mathbb{R}^{\binom{N+1}{2}}$. The subset of positive semidefinite matrices PSD_N is a full-dimensional closed convex cone in this vector space. It is a semi-algebraic set defined by $2^N - 1$ polynomial inequalities given by forcing the $2^N - 1$ principal minors of an $N \times N$ symmetric matrix to be nonnegative.

Definition 3.1.1. *Let $f \in \mathbb{R}[x_1, \dots, x_n]_{2d}$. The Gram spectrahedron of f is the spectrahedron*

$$K_f := L_f \cap \mathrm{PSD}_N,$$

where L_f is the affine subspace of symmetric matrices Q satisfying (3.2).

Proposition 3.1.1. *The Gram spectrahedron K_f is non-empty if and only if f is an SOS polynomial.*

Representation Theory

Representation theory is the study of symmetry. Specifically, it is concerned with the study of symmetries that arise from a group acting on a vector space. By linearizing the structure of a group in this way, we are able to study the group through the more familiar lens of vector spaces.

There is a great deal of theory required to fully understand and prove even the few key ideas in representation theory that appear in this section. As this is beyond the scope of this thesis, we simple provide the necessary statements and definitions along with examples. We are primarily interested in representations of S_n, the symmetric group, of permutations on n elements. As such most examples will be about S_n.

Definition 3.1.2. *Let V be a complex finite-dimensional vector space. A* representation *of a group G is a morphism $D : G \to GL(V)$.*

That is, D assigns to each element $g \in G$ a linear transformation, $D(g)$, such that

$$D(g)D(h) = D(gh), \ \forall g, h \in G \tag{3.3}$$

which in particular implies that $D(id)$, where id is the identity element in G, is the identity matrix and $D(s^{-1}) = D(s)^{-1}$. We say the *dimension* of the representation is the dimension of the underlying vector space V.

Example 3.1.2. *The trivial representation sends each element of a group to 1, i.e. $D(g) = 1$ for all $g \in G$, and therefore fixes a 1-dimensional vector space.*

Example 3.1.3. *Consider S_n and let $V = \mathbb{C}^n$ with standard basis vectors $e_1, \ldots, e_n$. To each permutation we assign the linear transformation that permutes the corresponding basis vectors. For example, the element (12) sends $e_1 \mapsto e_2$, $e_2 \mapsto e_1$, and fixes e_i for $i = 3, \ldots, n$. We can write this representation in matrix form. For a transposition $(i(i+1))$ we get*

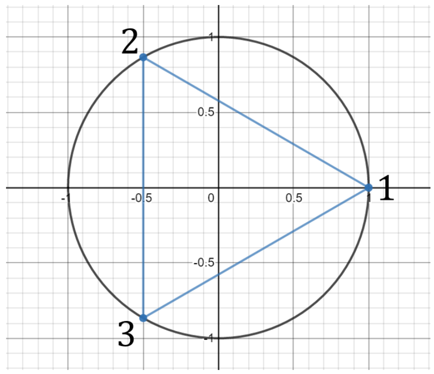
$$D((i(i+1))) = \begin{bmatrix} e_1 & e_2 & \cdots & e_{i+1} & e_i & \cdots & e_n \end{bmatrix}$$

that is, the identity matrix with the ith and $(i+1)$th columns swapped. By Equation (3.3), the remaining linear transformations can be generated by these matrices as transpositions generate S_n. This is sometimes called the defining representation of S_n.

Given a representation we say that G *acts on* the vector space V or call it the *action of G on V*. It is also common to refer to the vector space itself as the representation and we will at times do so in this thesis.

It is useful to think of two representations as being the “same” representation in some sense. We illustrate this with an example first before giving a precise definition.

Example 3.1.4. *Consider the symmetries of a triangle. Let $G = S_3$ and consider the representation which permutes the vertices of the below triangle in the unit circle:*



Transpositions correspond to reflections of the triangle, for example,

$$D((23)) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

with respect to the standard basis vectors. We also have rotations, such as

$$D((123)) = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$$

which together with reflections completely describe the symmetries of a triangle. On the other hand, consider these same symmetries in a different basis, letting b_1 be the vector going to vertex 2 and b_2 be the vector going to vertex 3. In this basis, we get the matrices

$$D'((23)) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad D'((123)) = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

That is, we have two representations that describe the same symmetries (i.e. linear transformations), but in different coordinate systems.

We think of these two representations as the same representation. To be more precise they are *equivalent*.

Definition 3.1.3. *Two n-dimensional representations $g \to D(g)$ and $g \to D'(g)$ of a group G are* equivalent *provided there exists a fixed non-singular $n \times n$ matrix T such that*

$$D'(g) = T^{-1} D(g) T \qquad \forall g \in G.$$

In the remainder of this section we wish to fully decompose a representation into its most basic components. With this goal in mind we start by defining subrepresentations.

Definition 3.1.4. *A* subrepresentation *of a representation V of a group G is a subspace $W \subset V$ such that $D(g)w \in W$ for all $w \in W$ and $g \in G$.*

That is, restricting the action of G to W is itself a representation. For example, $\{0\}$ and V are always subrepresentations of V.

Example 3.1.5. *Consider the defining representation of S_n on $\mathbb{C}^n$ from Example 3.1.3. It is easy to see that $W = \text{span}\{(1, 1, \dots, 1)^t\}$ is a subrepresentation.*

Definition 3.1.5. *A representation V is said to be* irreducible *if $\{0\}$ and V are the only subrepresentations of V.*

If V has a proper subspace which is also a subrepresentation, we say it is *reducible*.

Theorem 3.1.3 (Schur's Lemma)**.** *Let V and V' be irreducible representations of a group G. If $\phi : V \to V'$ is a morphism of representations then either ϕ is an isomorphism or $\phi = 0$. Further, any two such isomorphisms between V and V' differ by a scalar multiple.*

It is helpful to think in terms of matrix representations. Let D and D' be irreducible representations of vectors spaces V and V' respectively. Suppose we have a matrix T (not necessarily square) such that $TD(g) = D'(g)T$ for all $g \in G$. Schur's lemma tells us that either T is the null matrix or T is square, invertible, and D and D' are equivalent.

Theorem 3.1.4 (Maschke's Theorem)**.** *Let V be a representation of a finite group G. Then there exists a direct sum decomposition*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

where each V_i is an irreducible representation of G.

We will also write the decomposition into irreducible representations as

$$V = m_1 V_1 \oplus m_2 V_2 \oplus \cdots \oplus m_k V_k$$

where m_i, $i = 1, \dots, k$, is called the *multiplicity* of V_i. That is, there are m_i equivalent copies of this irreducible representation in the direct sum decomposition. We call the subspace $m_i V_i = \oplus_{j=1}^{m_i} V_i$ an isotypic component.

Not only can we decompose representations of finite groups into its irreducibles, using Schur's lemma we can produce a matrix representation with block diagonalized structure. That is, if $D(g)$ is a matrix representation for V and $D^{(i)}$, $i = 1, \dots, k$, are matrix representations for the irreducible representations V_i, then there exists an invertible matrix T such that

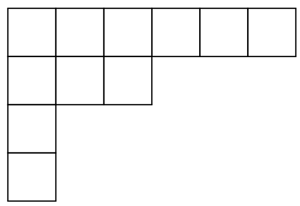
$$T^{-1} D(g) T = \begin{bmatrix} D^{(1)}(g) & 0 & \cdots & 0 \\ 0 & D^{(2)}(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D^{(k)}(g) \end{bmatrix}$$

for all $g \in G$.

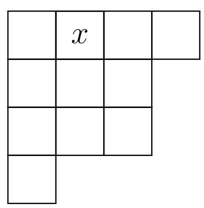
As this thesis is mainly interested in the symmetric group, we end with a discussion of the irreducible representations of S_n. Indeed, it turns out that the irreducible representations of S_n are in one-to-one correspondence with partitions of n.

Definition 3.1.6. *Given a positive integer n, a* partition λ *of n is a sequence of positive integers λ_i, $i = 1, \ldots, k$, in nonincreasing order such that $\sum_i \lambda_i = n$. By $\lambda \vdash n$ we mean λ is a partition of n. At times we also denote a partition by its sequence $(\lambda_1, \ldots, \lambda_k)$.*

We identify partitions with Young diagrams which have k rows of boxes with λ_i boxes in each row. For example, for partition $\lambda = (6, 3, 1, 1)$ the corresponding Young diagram is



One can define the hook length to any given box in the Young diagram as the number of boxes directly below and to the right of it (including itself). For example,



the hook length of the box with x inside is 5.

As previously stated, there is a natural bijection between the irreducible representations of S_n and partitions of n. It is beyond the scope of this thesis to go into the details of constructing these irreducible representations and showing the isomorphisms. The reader is encouraged to consult [31, Chapter 7].

Example 3.1.6. *The irreducible representations associated to $\lambda = (n)$ and $\mu = (1, 1, \ldots, 1)$ (both partitions of n) are both 1-dimensional. They are called the trivial and the alternating (or sign) irreducible representations, respectively. The trivial irreducible representation sends all elements to 1 while the alternating sends a permutation to its sign, 1 for even permutations, -1 for odd permutations.*

Example 3.1.7. *The irreducible representation labeled by $\lambda = (n-1, 1)$ is called the standard irreducible representation of S_n. Example 3.1.4 illustrates one way to construct this representation for $n = 3$. More generally for S_n, one starts with an $(n-1)$-simplex with vertices labeled $1, \ldots, n$ on the unit sphere in $\mathbb{R}^{n-1}$. The standard irreducible representation is defined by the linear transformations corresponding to the symmetries of the simplex.*

Until now, we have only considered $V = \mathbb{C}^n$. However, a very interesting vector space to consider is the space of polynomials. In this thesis we are mainly concerned with real polynomials, however, it is useful at times to consider $\mathbb{C}[x]_d$ as well. Let S_n act on $V = \mathbb{C}[x]_d$ by permuting the variables. By Maschke's Theorem, $\mathbb{C}[x]_d$ decomposes into the irreducible representations of S_n,

$$\mathbb{C}[x]_d = m_{\lambda_1} V_{\lambda_1} \oplus m_{\lambda_2} V_{\lambda_2} \oplus \cdots \oplus m_{\lambda_s} V_{\lambda_s}$$

where λ_j, $j = 1, \ldots, s$, are all partitions of n. Let $\lambda_1 = (n)$ so that V_{λ_1} is the trivial representation. The isotypic component associated to $\lambda_1 = (n)$ is the space of symmetric polynomials of degree d. These are precisely polynomials which are invariant under the action of S_n, i.e. $D(g)p(x) = p(x)$ for a symmetric polynomial p, and this subspace is denoted $\mathbb{C}[x]_d^{S_n}$.

It will be particularly important to be able to determine the multiplicity of each irreducible representation of S_n in the decomposition of $\mathbb{C}[x]_d$ and the remainder of this section is dedicated to such computations.

Multiplicities of irreducible representations for S_n acting on homogeneous polynomials

In this section, $V = \mathbb{R}[x_1, \ldots, x_n]_d \simeq \mathbb{R}^N$ or its complexification, the vector space of homogeneous degree d polynomials in $x_1, \ldots, x_n$. We begin with a representation of $G = S_n$ on $\mathbb{R}^n$ which extends to a representation on V by linear substitution of variables. Furthermore, we will also use the fact that the irreducible representations of S_n are indexed by partitions λ [67]. In other words,

$$V = m_{\lambda_1} V_{\lambda_1} \oplus \cdots \oplus m_{\lambda_s} V_{\lambda_s}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_s$ are partitions of n. Here we provide a simple way to determine the multiplicity of the irreducible representation V_λ. For this we need to compute $\langle \chi_\lambda, \chi_d \rangle$ where χ_λ is the irreducible character associated to V_λ and χ_d is the character of the representation $V = \mathbb{C}[x_1, \ldots, x_n]_d$. We will present a method which we have learned from Mark Haiman.

The space of complex-valued functions $\mathbb{C}^G$ on a group has a natural inner product $\mathbb{C}^G \times \mathbb{C}^G \to \mathbb{C}$ defined by

$$\langle f, g \rangle := \frac{1}{|G|} \sum_{\sigma \in G} \overline{f(\sigma)} g(\sigma).$$

The ring of symmetric functions Λ also has a natural inner product. This can be defined by specifying its values on pairs of basis vectors; for instance

$$\langle m_\lambda, h_\mu \rangle = \delta_{\lambda\mu}$$

where m_λ and h_μ are monomial and complete homogeneous symmetric functions associated to partitions λ and μ, respectively. Elsewhere in this thesis m_λ denotes the multiplicity of the irreducible representation V_λ, but in this section it denotes the monomial symmetric function associated to such a partition. A key tool for us will be the *Frobenius characteristic map* [74, p. 351]. This is a linear map between the subspace of functions $\chi : S_n \to \mathbb{C}$ constant on conjugacy classes and the ring Λ. It is defined by

$$\mathrm{ch}(\chi) = \frac{1}{n!} \sum_{\sigma \in S_n} \chi(\sigma) p_{\mathrm{par}(\sigma)}$$

where $\mathrm{par}(\sigma) = \mu$ is the partition given by the cycle type of σ, and $p_\mu = p_{\mu_1} \cdots p_{\mu_k}$ is the power sum symmetric polynomial [74, Section 7.7]. The characteristic map ch is an isometry [74, Proposition 7.18.1] between the subspace of functions constant on conjugacy classes and the space Λ_n of degree n symmetric functions, each equipped with their respective inner products. In the former, the irreducible characters χ_λ of S_n form an orthonormal basis, and in the latter, the Schur polynomials s_λ form an orthonormal basis. It is a standard fact in representation theory and the theory of symmetric functions that $\mathrm{ch}(\chi_\lambda) = s_\lambda$ [67, Section 4.7].

Theorem 3.1.5. *Let χ_d be the character of the representation of the symmetric group S_n acting on polynomials of degree d in n variables $V = \mathbb{C}[x_1, \ldots, x_n]_d$. Let $b(\lambda) = \sum_i (i-1)\lambda_i$ and let h_i be the hook length for the ith box in the Young diagram of λ. The multiplicity of the irreducible representation V_λ in V is equal to the number of solutions $y \in \mathbb{N}^n$ of the equation*

$$h_1 y_1 + \cdots + h_n y_n = d - b(\lambda).$$

Proof. We first compute the inner product

$$\begin{aligned}\langle s_\lambda(z), \sum_d \mathrm{ch}(\chi_d) q^d \rangle &= \langle s_\lambda(z), \sum_{\mu \vdash n} s_\mu(z) s_\mu(1, q, q^2, \ldots) \rangle \\ &= s_\lambda(1, q, q^2, \ldots).\end{aligned}$$

Here, the first equality is by [74, Exercise 7.73], while the second one is by orthonormality of the Schur basis for Λ. Thus we have shown that

$$\sum_d \langle \chi_\lambda, \chi_d \rangle q^d = \sum_d \langle \mathrm{ch}(\chi_\lambda), \mathrm{ch}(\chi_d) \rangle q^d = \langle s_\lambda(z), \sum_d \mathrm{ch}(\chi_d) q^d \rangle = s_\lambda(1, q, q^2, \ldots).$$

Let $f_\lambda(q)$ be the q-analogue of the number of standard Young tableaux of shape λ, which means that

$$f_\lambda(q) = \sum_{T \in SYT(\lambda)} q^{\mathrm{maj}(T)}$$

where $\mathrm{maj}(T)$ is the sum of the descents in T, i.e. it is the sum over all i such that $i+1$ appears in a lower row in T than i. We let $h(x)$ be the hook length for a box x in the Young diagram of λ. Using this, we obtain

$$s_\lambda(1, q, q^2, \ldots) = \frac{f_\lambda(q)}{(1-q)(1-q^2)\cdots(1-q^n)} = \frac{q^{b(\lambda)}}{\prod_{x \in \lambda}(1 - q^{h(x)})}$$

$$= q^{b(\lambda)}(1 + q^{h_1} + q^{2h_1} + \cdots)(1 + q^{h_2} + q^{2h_2} + \cdots) \cdots (1 + q^{h_n} + q^{2h_n} + \cdots)$$

where the first equality is [74, Proposition 7.19.11], the second equality is [74, Corollary 7.21.3], and h_i are all the hook lengths of λ. Expanding this out we see that the coefficient of the q^d term is the number of ways we can add multiples of the hook lengths $h_1, 2h_1, \ldots, h_2, 2h_2, \ldots, h_i, 2h_i, \ldots$ so they add up to $d - b(\lambda)$. □

Example 3.1.8. *Suppose we have S_4 acting on $V = \mathbb{C}[x_1, x_2, x_3, x_4]_4$. Consider the Young diagram of $\lambda = (2, 1, 1)$ with zeros filled in the first row, 1's in the second row and so on:*

0	0
1	
2	

Summing up the numbers gives us $b(\lambda) = 3$. The hook lengths are 4,2,1,1, and $d = 4$. Thus in order to calculate the multiplicity of the irreducible representation associated to $(2, 1, 1)$, by Theorem 3.1.5 we need to find all nonnegative integer solutions to the equation $x_1 + x_2 + 2x_3 + 4x_4 = 4 - 3 = 1$. In the table below we calculate this for all partitions of 4:

Partition	$b(\lambda)$	$h^T x = n - b(\lambda)$	*Number of solutions*
(4)	0	$x_1 + 2x_2 + 3x_3 + 4x_4 = 4$	5
		(0, 0, 0, 1)	
		(1, 0, 1, 0)	
		(0, 2, 0, 0)	
		(2, 1, 0, 0)	
		(4, 0, 0, 0)	
$(3,1)$	1	$x_1 + x_2 + 2x_3 + 4x_4 = 3$	6
		(0, 1, 1, 0)	
		(1, 0, 1, 0)	
		(0, 3, 0, 0)	
		(1, 2, 0, 0)	
		(2, 1, 0, 0)	
		(3, 0, 0, 0)	
$(2,2)$	2	$x_1 + 2x_2 + 2x_3 + 3x_4 = 2$	3
		(0, 0, 1, 0)	
		(0, 1, 0, 0)	
		(2, 0, 0, 0)	
$(2,1,1)$	3	$x_1 + x_2 + 2x_3 + 4x_4 = 1$	2
		(0, 1, 0, 0)	
		(1, 0, 0, 0)	
$(1,1,1,1)$	6	$4x_1 + 3x_2 + 2x_3 + x_4 = -2$	0

As a last example consider the partition $\lambda = (n)$ associated to the trivial irreducible representation of S_n for $V = \mathbb{C}[x]_d$. The hook lengths are $1, 2, \ldots, n$ and $b((n)) = 0$. To calculate the multiplicity we wish to know the positive integer solutions to the equation

$$x_1 + 2x_2 + \cdots + nx_n = d.$$

Notice that x_1 counts the number of 1's in our sum, x_2 counts the number of 2's and so on. Assuming $n \geq d$, this implies that the multiplicity of the trivial irreducible representation will be equal to the number of partitions of d. Moreover, for any n and d we will always have at least one copy of the trivial irreducible representation as $(d, 0, \ldots, 0)$ will always be a solution.

For some recent results on the interplay between n, d and the multiplicities of irreducible representation of S_n for $V = \mathbb{C}[x]_d$, the interested reader should consider [26].

3.2 Symmetry Adapted Basis

In this section we define and discuss a symmetry adapted basis of a vector space V given a particular representation. For completeness we also include Algorithm 1 at the end for concrete calculations of such a basis.

Recall from the previous section that a representation of a group G is a homomorphism $\rho : G \to GL(V)$ where $GL(V)$ is the group of invertible linear transformations of a vector space V. If V is finite-dimensional, we also write $GL(n)$ for $n = \dim V$. A subrepresentation of V is a subspace $U \subset V$ which is invariant under the action of G. If the only subrepresentations of V are $\{0\}$ and V, we say that V is irreducible. The character $\chi_\rho : G \to \mathbb{C}$ is defined by taking the trace of each $\rho(g)$ and is used to decompose representations. A representation which admits a direct sum decomposition $V = \oplus V_i$ with each V_i irreducible is said to be completely reducible. Representations of finite groups are completely reducible. When we decompose V into irreducibles $V_1, \ldots, V_s$, each V_i appears with multiplicity m_i:

$$V = m_1 V_1 \oplus \cdots \oplus m_s V_s.$$

This means that there exists a basis of V such that $\rho(g)$ becomes the matrix $D(g)$ for $g \in G$ and is block diagonal with m_i blocks corresponding to V_i where each block is $n_i \times n_i$ with $n_i = \dim V_i$. Here we denote by $D(g)$ the matrix written in a chosen basis for the linear map $\rho(g)$.

In general, these m_i matrices of size $n_i \times n_i$ corresponding to V_i are *not* identical. Fortunately, one can choose a different basis of V with respect to which the representation matrices $\tilde{D}(g)$ for all $g \in G$ are block diagonal where the m_i blocks corresponding to V_i are identical. See [28, Section 5.2] or [72, p. 23] for Algorithm 1 to compute such a basis. In other words, this algorithm constructs a change of basis matrix T such that $T^{-1}D(g)T$ is block diagonal with this extra nice property for all $g \in G$. Such a basis is known as a *symmetry adapted basis*. A symmetry adapted basis can also be used to simplify linear operators $P \in \mathrm{Hom}(V, V)$ which commute with the representation matrices $D(g)$ for all $g \in G$.

In this thesis, we are concerned with the field of real numbers $\mathbb{R}$, but to more easily and uniformly describe the representation theory involved we work with $\mathbb{C}$. Irreducible representations of finite groups over $\mathbb{C}$ come in three types [72, p. 108]. All of them give rise to representations of G over $\mathbb{R}$, although the dimension may stay the same (type 2) or double (types 1 or 3). The characters of the representations over $\mathbb{R}$ are either equal to the character χ of the representation over $\mathbb{C}$ (type 2) or equal to $\chi + \overline{\chi}$ or 2χ (types 1 or 3). By averaging over the group, an invariant inner product can be created which allows each of these real representations to be written using real orthogonal matrices. See Example 3.3.1. For many results, the orthogonality of the matrices is important. Therefore we will assume that all irreducibles appearing in the isotypic decompositions under consideration are of type 2. In the case $G = S_n$, all irreducibles are indeed of type 2, so this assumption is always justified. For other groups, to see if an irreducible representation is type 2, one needs check if $\frac{1}{|G|}\sum_{g\in G}\chi(g^2) = 1$ [72, p. 109]. Whenever we use the complexification $\mathbb{C} \otimes_{\mathbb{R}} V$, recall that adjustments can be made so that all the matrices are real, and the dimensions will not change.

Theorem 3.2.1. *[28, Theorem 2.5] Let $\rho : G \to GL(V)$ be a representation of the finite group G, and let*

$$V = m_1V_1 \oplus \cdots \oplus m_sV_s$$

be the direct sum decomposition into irreducible representations V_i with $\dim V_i = n_i$ and multiplicity m_i. Then every $P \in Hom(V, V)$ such that $D(g)P = PD(g)$ for all $g \in G$ has the following form in a symmetry adapted basis:

$$P = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_s \end{pmatrix}$$

where each P_i is an $(m_in_i) \times (m_in_i)$ matrix

$$P_i = \begin{pmatrix} \mu^i_{11}I_{n_i} & \mu^i_{12}I_{n_i} & \dots & \mu^i_{1m_i}I_{n_i} \\ \mu^i_{21}I_{n_i} & \mu^i_{22}I_{n_i} & \dots & \mu^i_{2m_i}I_{n_i} \\ \vdots & \vdots & \ddots & \vdots \\ \mu^i_{m_i1}I_{n_i} & \mu^i_{m_i2}I_{n_i} & \dots & \mu^i_{m_im_i}I_{n_i} \end{pmatrix}.$$

Proof. In a symmetry adapted basis we have

$$D(g) = \begin{pmatrix} D_1(g) & 0 & \dots & 0 \\ 0 & D_2(g) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D_s(g) \end{pmatrix}$$

where each $D_i(g)$ is an $(m_i n_i) \times (m_i n_i)$ block diagonal matrix with m_i identical $n_i \times n_i$ matrices along its diagonal:

$$D_i(g) = \begin{pmatrix} \Sigma_i(g) & 0 & \dots & 0 \\ 0 & \Sigma_i(g) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_i(g) \end{pmatrix}.$$

After partitioning P into $(m_i n_i) \times (m_j n_j)$ matrices P_{ij} for $i, j = 1, \dots, s$, we see that

$$D(g)P = PD(g)$$

implies $D_i(g)P_{ij} = P_{ij}D_j(g)$. We partition each P_{ij} further

$$P_{ij} = \begin{pmatrix} P_{ij}^{11} & P_{ij}^{12} & \dots & P_{ij}^{1m_j} \\ P_{ij}^{21} & P_{ij}^{22} & \dots & P_{ij}^{2m_j} \\ \vdots & \vdots & \ddots & \vdots \\ P_{ij}^{m_i 1} & P_{ij}^{m_i 2} & \dots & P_{ij}^{m_i m_j} \end{pmatrix}$$

and observe that $\Sigma_i(g)P_{ij}^{tu} = P_{ij}^{tu}\Sigma_j(g)$ for all $i, j = 1, \dots, s$ and $g \in G$. When we view P_{ij}^{tu} as an element of $\mathrm{Hom}(V_i, V_j)$, Schur's Lemma implies that $P_{ij}^{tu} = 0$ whenever $i \neq j$. Furthermore, $P_{ii}^{tu} = \mu_{tu}^i I_{n_i}$. □

Corollary 3.2.1. *Let $V = m_1V_1 \oplus \cdots \oplus m_sV_s$ be as in Theorem 3.2.1. Then the dimension of the subspace of linear operators $P \in Hom(V, V)$ such that $D(g)P = PD(g)$ for all $g \in G$ is $m_1^2 + m_2^2 + \cdots + m_s^2$.*

Proof. The above theorem implies that the dimension is at most $m_1^2 + m_2^2 + \cdots + m_s^2$. Every block diagonal matrix $P = \mathrm{diag}(P_1, \dots, P_s)$, with P_i as in the theorem, commutes with each $D(g)$. Since the m_i^2 scalars μ_{tu}^i are free parameters for $i = 1, \dots, s$, we get the result. □

A reordering of the symmetry adapted basis which block-diagonalizes the $D(g)$ matrices also leads to a more convenient block-diagonalization of commuting linear operators P such that $PD(g) = D(g)P$.

Corollary 3.2.2. *Given* $V = m_1V_1 \oplus \cdots \oplus m_sV_s$ *and* $P \in Hom(V,V)$ *such that* $D(g)P = PD(g)$ *for all* $g \in G$, *let*

$$\mathcal{B} = \bigcup_{i=1}^{s}\bigcup_{k=1}^{m_i}\mathcal{B}_{ik}$$

be an ordered basis that is symmetry adapted where $\mathcal{B}_{ik} = \{v_1^{ik}, v_2^{ik}, \ldots, v_{n_i}^{ik}\}$. *If one reorders the basis vectors in* $\bigcup_{k=1}^{m_i}\mathcal{B}_{ik}$ *as* $\bigcup_{\ell=1}^{n_i}\tilde{\mathcal{B}}_{i\ell}$ *with* $\tilde{\mathcal{B}}_{i\ell} = \{v_\ell^{i1}, v_\ell^{i2}, \ldots, v_\ell^{im_i}\}$ *then*

$$P = \begin{pmatrix} \tilde{P}_1 & 0 & \ldots & 0 \\ 0 & \tilde{P}_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \tilde{P}_s \end{pmatrix}$$

where

$$\tilde{P}_i = \begin{pmatrix} M_i & 0 & \ldots & 0 \\ 0 & M_i & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & M_i \end{pmatrix} \qquad \text{and} \qquad M_i = \begin{pmatrix} \mu_{11}^i & \mu_{12}^i & \ldots & \mu_{1m_i}^i \\ \mu_{21}^i & \mu_{22}^i & \ldots & \mu_{2m_i}^i \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m_i1}^i & \mu_{m_i2}^i & \ldots & \mu_{m_im_i}^i \end{pmatrix}.$$

Proof. The reordering of the symmetry adapted basis has the effect of reordering the rows and columns of P_i in Theorem 3.2.1 resulting in $\tilde{P}_i$. □

For completeness, we briefly summarize the algorithm in [28, p. 113] used to compute the change of basis matrix to get a symmetry adapted basis as in Corollary 3.2.2. This algorithm can also be found in [72, p. 23]. For each irreducible representation V_i of the finite group G, let $d^i(g)$ be the matrix representation for $g \in G$. The size of $d^i(g)$ is $n_i \times n_i$ where n_i is the dimension of V_i. We furthermore choose $d^i(g)$ to be real orthogonal matrices, which can easily be done when all irreducibles appearing are of type 2, as we assume throughout.

Algorithm 1: Computation of symmetry adapted change of basis matrix as in Corollary 3.2.2

- For each irreducible representation $i = 1, \dots, s$,
  1. Compute the matrix
  $$\pi^i = \sum_{g \in G} d^i_{11}(g^{-1}) D(g).$$
  2. The matrix π^i will be of rank m_i. Choose m_i linearly independent columns and label them
  $$v_1^{i1}, v_1^{i2}, \dots, v_1^{im_i}.$$
  If this set of vectors is not orthonormal, apply Gram-Schmidt (here we utilize a modification to the algorithm [28, Theorem 5.4]) and relabel each v_1^{ij}.
  3. For each $k = 2, \dots, n_i$,
     a) Compute the matrix
     $$P_{ik} = \frac{n_i}{|G|} \sum_{g \in G} d^i_{1k}(g^{-1}) D(g).$$
     b) Define new column vectors
     $$v_k^{ij} = P_{ik} v_1^{ij}$$
     for $j = 1, \dots, m_i$.
- The above generates a symmetry adapted basis for all m_i copies of V_i. Arrange these vectors,

$$\begin{array}{lcccc} \text{Basis } \mathcal{B}_{i1} \text{ for } V_i^1 : & v_1^{i1} & v_2^{i1} & \cdots & v_{n_i}^{i1} \\ \text{Basis } \mathcal{B}_{i2} \text{ for } V_i^2 : & v_1^{i2} & v_2^{i2} & \cdots & v_{n_i}^{i2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \text{Basis } \mathcal{B}_{im_i} \text{ for } V_i^{m_i} : & v_1^{im_i} & v_2^{im_i} & \cdots & v_{n_i}^{im_i} \end{array}$$

- Construct the change of basis matrix T: For each $i = 1, \dots, s$, the corresponding columns of T will be the $\{v_k^{ij}\}$ in the following order:

$$\begin{array}{cccc} v_1^{i1} & v_2^{i1} & \cdots & v_{n_i}^{i1} \\ v_1^{i2} & v_2^{i2} & \cdots & v_{n_i}^{i2} \\ \vdots & \vdots & \vdots & \vdots \\ v_1^{im_i} & v_2^{im_i} & \cdots & v_{n_i}^{im_i} \end{array}$$

starting with v_1^{i1} and going down each column of the array and ending with $v_{n_i}^{im_i}$.

3.3 Symmetry Adapted Gram Spectrahedron

In this section we consider SOS polynomials invariant under the linear action of a finite group G. Therefore we start with a representation of G on $\mathbb{R}^n$. A polynomial f is G-invariant if $f(g^{-1}x) = f(x)$ for all $g \in G$. The S_n-invariant polynomials are the usual symmetric polynomials. The vector space of G-invariant polynomials of degree $2d$ will be denoted $\mathbb{R}[x_1, \ldots, x_n]_{2d}^G$.

The action of G on $\mathbb{R}^n$ extends to a representation $D : G \to GL(V)$ for $V = \mathbb{R}[x_1, \ldots, x_n]_d$ with matrices $D(g)$ with respect to a chosen basis. Let $[x]$ be the column vector whose entries form a basis for V. For *any* (possibly non-invariant) polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]_{2d}$ we can write $f(x) = [x]^T Q[x]$ for some $Q \in \mathcal{S}^N$. Hence $g \cdot f(x) = [x]^T D(g)^T QD(g)[x]$ for all $g \in G$, and if f is G-invariant then $f(x) = [x]^T D(g)^T QD(g)[x]$ for all $g \in G$.

Proposition 3.3.1. *If f is a G-invariant polynomial in $\mathbb{R}[x_1, \ldots, x_n]_{2d}^G$ then there exists $Q \in \mathcal{S}^N$ such that $f(x) = [x]^T Q[x]$ where $Q = D(g)^T QD(g)$ for all $g \in G$.*

Proof. By Theorem 3.1.1 there exists $Q' \in \mathcal{S}^N$ such that $f(x) = [x]^T Q'[x]$. Since f is G-invariant, $f(x) = [x]^T D(g)^T Q' D(g)[x]$ for all $g \in G$. Now let

$$Q = \frac{1}{|G|} \sum_{g \in G} D(g)^T Q' D(g).$$

□

Definition 3.3.1. *Let $f \in \mathbb{R}[x_1, \ldots, x_n]_{2d}^G$ be a G-invariant polynomial for some representation of G on $\mathbb{R}^n$. Let $D : G \to GL(V)$ be the representation of G on $V = \mathbb{R}[x_1, \ldots, x_n]_d$ given by linear substitution. The symmetry adapted Gram spectrahedron of f is*

$$K_f^G := L_f \cap \mathrm{PSD}_N^G,$$

where

$$\mathrm{PSD}_N^G := \left\{ Q \in \mathrm{PSD}_N \;\middle|\; D(g)^T QD(g) = Q, \text{ for all } g \in G \right\}.$$

Here, L_f is the affine space of symmetric matrices Q satisfying $f(x) = [x]^T Q[x]$ for $[x]$ a column vector whose entries form a basis of V, and $D(g)$ are the matrices of D in this basis. The set PSD_N^G consists of all positive semidefinite matrices which are fixed by the action of G. We call this the symmetry adapted PSD cone.

Corollary 3.3.1. *Let $V = \mathbb{R}[x_1, \ldots, x_n]_d$ and let $D : G \to GL(V)$ be the representation of G on V obtained by linear substitution from a representation of G on $\mathbb{R}^n$. Assume that all irreducible representations appearing in the isotypic decomposition*

$$\mathbb{C} \otimes_{\mathbb{R}} V = m_1 V_1 \oplus \cdots \oplus m_s V_s$$

are of type 2, with $\dim V_i = n_i$ *and multiplicity* m_i*. Then there exists a basis for* V *such that a symmetric matrix* $Q \in \mathcal{S}^N$ *is in* PSD_N^G *if and only if*

$$Q = \begin{pmatrix} \tilde{Q}_1 & 0 & \dots & 0 \\ 0 & \tilde{Q}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{Q}_s \end{pmatrix} \quad \text{where} \quad \tilde{Q}_i = \begin{pmatrix} Q_i & 0 & \dots & 0 \\ 0 & Q_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_i \end{pmatrix} \tag{3.4}$$

with $Q_i \in \mathrm{PSD}_{m_i}$ *for all* $i = 1, \dots, s$ *and* n_i *identical copies in* Q_i*.*

Proof. By Corollary 3.2.2 an arbitrary matrix Q commutes with all $D(g)$ if and only if it has the stated block-diagonal form in a symmetry adapted basis. If all matrices $D(g)$ are orthogonal, requiring $D(g)^T Q D(g) = Q$ is the same as requiring $QD(g) = D(g)Q$. Since the irreducibles are of type 2, the matrices d^i, and therefore π^i and P_{ik}, can be chosen with real entries in Algorithm 1. Thus, the symmetry adapted basis can be written as real linear combinations of the original basis vectors. By using the invariant inner product

$$\langle v, w \rangle := v^T \left(\frac{1}{|G|} \sum_{g \in G} D(g)^T D(g) \right) w,$$

the symmetry adapted basis can further be adjusted so that the matrices $D(g)$ in that basis are orthogonal matrices. To carry this out, one can apply Gram-Schmidt using the invariant inner product above. It only remains to require symmetry and positive semidefiniteness. This is the condition stated above, that $Q_i \in \mathrm{PSD}_{m_i}$. □

As one might expect PSD_N^G and K_f^G are simpler, smaller, and more structured objects than their counterparts PSD_N and K_f when f is G-invariant. The rest of this chapter is devoted to convincing the reader that this is indeed the case.

Example 3.3.1. *The focus of this thesis is the case* $G = S_n$*. However, we include an example with the symmetry group* $G = I_h$ *of an icosahedron. All 10 irreducible representations of* I_h *are of type 2. We continue this example in Section 3.4 to demonstrate extremal rays of rank* > 1*. This group consists of* 120 *invertible* 3×3 *orthogonal matrices. Generators are, for instance,*

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & -\frac{1}{4}\sqrt{5} - \frac{1}{4} & \frac{1}{\sqrt{5}+1} \\ \frac{1}{4}\sqrt{5} + \frac{1}{4} & \frac{1}{\sqrt{5}+1} & -\frac{1}{2} \\ \frac{1}{\sqrt{5}+1} & \frac{1}{2} & \frac{1}{4}\sqrt{5} + \frac{1}{4} \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The action on $\mathbb{R}^3$ *extends to an action on* $V = \mathbb{R}[x_1, x_2, x_3]_2$*. The* 6×6 *matrices* $\widetilde{D(g)}$ *for all* 120 *elements* $g \in I_h$ *written in the monomial basis* $\{x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2\}$ *are*

$$\begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} \overset{\widetilde{D}}{\longmapsto} \begin{pmatrix} g_{11}^2 & g_{11}g_{21} & g_{11}g_{31} & g_{21}^2 & g_{21}g_{31} & g_{31}^2 \\ 2\,g_{11}g_{12} & g_{12}g_{21} + g_{11}g_{22} & g_{12}g_{31} + g_{11}g_{32} & 2\,g_{21}g_{22} & g_{22}g_{31} + g_{21}g_{32} & 2\,g_{31}g_{32} \\ 2\,g_{11}g_{13} & g_{13}g_{21} + g_{11}g_{23} & g_{13}g_{31} + g_{11}g_{33} & 2\,g_{21}g_{23} & g_{23}g_{31} + g_{21}g_{33} & 2\,g_{31}g_{33} \\ g_{12}^2 & g_{12}g_{22} & g_{12}g_{32} & g_{22}^2 & g_{22}g_{32} & g_{32}^2 \\ 2\,g_{12}g_{13} & g_{13}g_{22} + g_{12}g_{23} & g_{13}g_{32} + g_{12}g_{33} & 2\,g_{22}g_{23} & g_{23}g_{32} + g_{22}g_{33} & 2\,g_{32}g_{33} \\ g_{13}^2 & g_{13}g_{23} & g_{13}g_{33} & g_{23}^2 & g_{23}g_{33} & g_{33}^2 \end{pmatrix}.$$

The resulting 6×6 *matrices above will not be orthogonal matrices. However, we can create the matrix*

$$S := \frac{1}{|G|} \sum_{g \in G} \widetilde{D(g)}^T \widetilde{D(g)}$$

which we use to define the invariant inner product $\langle v, w \rangle := v^T S w$*. In this case,*

$$S = \begin{pmatrix} \frac{7}{5} & 0 & 0 & -\frac{1}{5} & 0 & -\frac{1}{5} \\ 0 & \frac{4}{5} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{4}{5} & 0 & 0 & 0 \\ -\frac{1}{5} & 0 & 0 & \frac{7}{5} & 0 & -\frac{1}{5} \\ 0 & 0 & 0 & 0 & \frac{4}{5} & 0 \\ -\frac{1}{5} & 0 & 0 & -\frac{1}{5} & 0 & \frac{7}{5} \end{pmatrix}.$$

Applying a modified Gram-Schmidt to the monomial basis we can create a new basis $u_1, \ldots, u_6$ *for which the representation matrices become orthogonal. Collecting the new basis vectors in the columns of a matrix* U *we create orthogonal matrices* $D(g) = U^{-1}\widetilde{D(g)}U$ *for all* $g \in I_h$*. A useful fact is that* $U^{-1} = U^T S$*. Consider* $[x]^T I [x]$ *for* $[x]$ *the column vector containing the monomials of degree 2. This would produce the polynomial*

$$x_1^4 + x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^4 + x_2^2 x_3^2 + x_3^4,$$

which is not I_h*-invariant. Proposition 3.3.2 below implies that, in the monomial basis, the identity matrix is not in* $PSD_N^{I_h}$*, as can also be checked directly. However, if we apply the change of basis and extract the polynomial corresponding to the identity matrix* $f = (U^T m)^T I (U^T m)$ *we obtain the* I_h*-invariant polynomial*

$$f = \frac{3}{4}\, x_1^4 + \frac{3}{2}\, x_1^2 x_2^2 + \frac{3}{4}\, x_2^4 + \frac{3}{2}\, x_1^2 x_3^2 + \frac{3}{2}\, x_2^2 x_3^2 + \frac{3}{4}\, x_3^4.$$

In the basis given by the column vector $U^T m$*, the 2-dimensional symmetry adapted PSD cone* $\mathrm{PSD}_6^{I_h}$ *is given by the 63 inequalities arising from the principal minors of the matrix given (to 5 digits) by*

$$\begin{pmatrix} \frac{13}{28}\, q_{55} + \frac{15}{28}\, q_{66} & 0 & 0 & -0.61859\, q_{55} + 0.61859\, q_{66} & 0 & -0.73193\, q_{55} + 0.73193\, q_{66} \\ 0 & q_{55} & 0 & 0 & 0 & 0 \\ 0 & 0 & q_{55} & 0 & 0 & 0 \\ -0.61859\, q_{55} + 0.61859\, q_{66} & 0 & 0 & \frac{2}{7}\, q_{55} + \frac{5}{7}\, q_{66} & 0 & -0.84515\, q_{55} + 0.84515\, q_{66} \\ 0 & 0 & 0 & 0 & q_{55} & 0 \\ -0.73193\, q_{55} + 0.73193\, q_{66} & 0 & 0 & -0.84515\, q_{55} + 0.84515\, q_{66} & 0 & q_{66} \end{pmatrix}.$$

We close this section with the observation that constructing SOS decompositions with symmetry adapted bases has another advantage. Namely, the partial sums of squares are G-invariant polynomials themselves when one groups them according to the isotopic components. This result was also pointed out in [32, pp. 107-112], but we would like to call attention to it, as well as provide a fully explicit proof. We then use this result to prove that every matrix in PSD_N^G produces a G-invariant polynomial.

Proposition 3.3.2. *Let $f \in \mathbb{R}[x_1, \dots, x_n]_{2d}^G$ be a G-invariant polynomial with real coefficients and let every irreducible appearing with nonzero multiplicity in $\mathbb{C}[x_1, \dots, x_n]_d = \mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}[x_1, \dots, x_n]_d$ be of type 2. If f is an SOS polynomial then*

$$f = \sum_{\alpha_1=1}^{r_1} q_{1,\alpha_1}^2 + \sum_{\alpha_2=1}^{r_2} q_{2,\alpha_2}^2 + \cdots + \sum_{\alpha_s=1}^{r_s} q_{s,\alpha_s}^2 \tag{3.5}$$

where each q_{i,α_i} is a polynomial of degree d appearing in the ith isotypic component m_iV_i of

$$\mathbb{C}[x_1, \dots, x_n]_d = m_1V_1 \oplus \cdots \oplus m_sV_s.$$

Further, each partial sum of squares $\sum_{\alpha_i=1}^{r_i} q_{i,\alpha_i}^2$ is a G-invariant polynomial, with $r_i = \operatorname{rank}(Q_i)$ as in Corollary 3.3.1. By choosing bases agreeing with the real representations corresponding to each isotypic component, each q_{i,α_i} may be chosen with real coefficients.

Proof. Let v_j^i be the column vector $[v_j^{i1}, v_j^{i2}, \dots, v_j^{im_i}]^T$ of basis polynomials chosen in Algorithm 1 as an orthonormal basis for the column space of the jth projection operator for the ith isotypic component. Since V_i is of type 2, these basis vectors can be chosen as polynomials with real coefficients, and such that the matrices $d^i(g)$ are orthogonal. Let Q_i be the matrices appearing in Corollary 3.3.1. Then the partial sum of squares for the ith isotypic component can be rewritten

$$\begin{aligned}\sum_{\alpha_i=1}^{r_i} q_{i,\alpha_i}^2 &= \sum_{j=1}^{n_i} (v_j^i)^T Q_i (v_j^i) \\ &= \left\langle Q_i, \sum_{j=1}^{n_i} (v_j^i)(v_j^i)^T \right\rangle \\ &= \langle Q_i, P_i(x) \rangle\end{aligned}$$

where $P_i(x)$ is an $m_i \times m_i$ matrix with polynomial entries and $r_i = \operatorname{rank}(Q_i)$. Specifically, the (k, ℓ) entry of the matrix $P_i(x)$ is given by

$$p_{k,\ell}^i = \sum_{j=1}^{n_i} v_j^{ik} v_j^{i\ell}. \tag{3.6}$$

Letting $d^i(g) = (d^i_{\alpha,\beta})$ for $g \in G$ be the orthogonal matrices for the real representation associated to the ith isotypic component, we have the relations

$$\sum_{j=1}^{n_i} (d^i_{\alpha j})(d^i_{\beta j}) = \delta_{\alpha\beta}.$$

Recall for each $k = 1, \dots, m_i$ the entry v_j^{ik} of the column vector v_j^i is a symmetry adapted basis polynomial which *transforms* like the jth basis vector of the ith irreducible representation:

$$g \cdot v_j^{ik} = \sum_{\alpha=1}^{n_i} d^i_{\alpha j} v_\alpha^{ik}.$$

Acting with the group element $g \in G$ we have

$$\sum_{j=1}^{n_i} (v_j^i)(v_j^i)^T \mapsto \sum_{j=1}^{n_i} \begin{bmatrix} \vdots \\ \sum_{\alpha=1}^{n_i} d_{\alpha j}^i v_\alpha^{ik} \\ \vdots \end{bmatrix} \begin{bmatrix} \cdots & \sum_{\beta=1}^{n_i} d_{\beta j}^i v_\beta^{i\ell} & \cdots \end{bmatrix}$$

$$= \sum_{j=1}^{n_i} \begin{bmatrix} \ddots & & \\ & \begin{matrix} (k,\ell) \text{ entry} = \\ \sum_{(\alpha,\beta)\in[n_i]\times[n_i]} d_{\alpha j}^i d_{\beta j}^i v_\alpha^{ik} v_\beta^{i\ell} \end{matrix} & \\ & & \ddots \end{bmatrix}.$$

Pulling the sum over $j = 1, \ldots, n_i$ inside to each individual entry of the matrix we see that the orthogonality relations zero out all terms except those giving the (k, ℓ) entry of $P_i(x)$. Therefore, each of the entries of $P_i(x)$ will be itself an invariant polynomial, and hence $\sum_{\alpha_i=1}^{r_i} q_{i,\alpha_i}^2 = \langle Q_i, P_i(x) \rangle$ is invariant. Note that a factorization of Q_i will still be required to find the r_i explicit squares q_{i,α_i}^2 as usual. □

Example 3.3.2. *Consider the polynomial* $(H_{21} - H_{111})(x_1^2, x_2^2, x_3^2) =$

$$\frac{1}{18}(x_1^4 + x_1^2x_2^2 + x_1^2x_3^2 + x_2^4 + x_2^2x_3^2 + x_3^4)(x_1^2 + x_2^2 + x_3^2) - \frac{1}{27}(x_1^2 + x_2^2 + x_3^2)^3$$
$$= \frac{1}{54}x_1^6 + \frac{1}{54}x_2^6 + \frac{1}{54}x_3^6 - \frac{1}{18}x_1^2x_2^2x_3^2$$

which is an S_3-invariant (symmetric) polynomial. We will define a family of such polynomials in Chapter 4. One matrix in its symmetry adapted Gram spectrahedron is

$$\frac{1}{108}\begin{pmatrix} 4 & -2\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2\sqrt{2} & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{pmatrix}.$$

The rows and columns of this matrix correspond to polynomials which form a symmetry *adapted basis, and using these we can write our polynomial as*

$$
\begin{aligned}
(H_{21}-H_{111})(x_1^2,x_2^2,x_3^2) &= \frac{1}{54}x_1^6+\frac{1}{54}x_2^6+\frac{1}{54}x_3^6-\frac{1}{18}x_1^2x_2^2x_3^2\\
&= \frac{1}{108}\Big(\big(\frac{2\sqrt{3}}{3}(x_1^3+x_2^3+x_3^3)-\frac{\sqrt{3}}{3}(x_1^2x_2+x_1^2x_3+x_1x_2^2+x_1x_3^2+x_2^2x_3+x_2x_3^2)\big)^2\\
&+\big(\frac{\sqrt{6}}{6}(2x_1^3-x_2^3-x_3^3)+\frac{\sqrt{6}}{6}(2x_1^2x_2+2x_1^2x_3-x_1x_2^2-x_1x_3^2-x_2^2x_3-x_2x_3^2)\big)^2\\
&+\big(\frac{\sqrt{2}}{2}(x_2^3-x_3^3)+\frac{\sqrt{2}}{2}(x_1x_2^2-x_1x_3^2+x_2^2x_3-x_2x_3^2)\big)^2\\
&+\big(x_1^2x_2-x_1^2x_3-x_1x_2^2+x_1x_3^2+x_2^2x_3-x_2x_3^2\big)^2\Big)
\end{aligned}
$$

where the first square comes from the rank one trivial block, the second and third squares from the two copies of the rank one standard block and the last square from the rank one alternating block. Clearly, the first and last squares are symmetric polynomials. Proposition 3.3.2 states that the sum of the second and third squares is also a symmetric polynomial. Although it is not immediately clear from the above representation, it is indeed so. We invite the reader to check.

Note that the proof of Proposition 3.3.2 can be applied to any matrix Q in the symmetry adapted PSD cone, which leads to the following results.

Corollary 3.3.2. *Let $[x]$ be a vector of polynomials comprising a fixed basis of $\mathbb{R}[x_1,\dots,x_n]_d$. Then every matrix $Q \in \mathrm{PSD}_N^G$, calculated using the representation matrices $D(g)$ written in this basis produces a G-invariant polynomial $f(x)=[x]^TQ[x]$.*

Corollary 3.3.3. *The symmetry adapted Gram spectrahedron K_f^G is non-empty if and only if the G-invariant polynomial f is SOS.*

3.4 Properties of symmetry adapted PSD cones and Gram spectrahedra

In this section we provide general results about PSD_N^G and K_f^G. We compute the dimension of PSD_N^G and give a characterization of its extreme rays, as well as describe the matrix block of $Q \in \mathrm{PSD}_N^{S_n}$ in a symmetry adapted basis corresponding to the trivial representation when G is the symmetric group.

Corollary 3.4.1. *The dimension of PSD_N^G is $\sum_{i=1}^s \binom{m_i+1}{2}$.*

Proof. Since the dimension of PSD_{m_i} is $\binom{m_i+1}{2}$ Corollary 3.3.1 implies the result. □

Extremal Rays

Every point $Q \in PSD_N^G$ gives rise to a *ray*, as in

$$\mathrm{ray}(Q) := \{cQ : c \in \mathbb{R}_{\geq 0}\}.$$

A ray r is *extremal* if it cannot be written as a non-trivial convex combination of other rays. We note that in the case of the usual cone of positive semidefinite matrices PSD_N, the Spectral Theorem for symmetric matrices implies that the extremal rays correspond to matrices of rank one. A *face* F of a convex set K is a convex subset such that if a convex combination of two points of K lies in F, then the points were already elements of F. In symbols, if $a, b \in K$ and $ta + (1-t)b \in F$ for some $t \in (0,1)$ then $a, b \in F$. Given a spectrahedron K, any matrix $Q \in K$ belongs to the relative interior of a unique face denoted by $F_K(Q)$. The face $F_K(Q)$ is the intersection of K with the subspace of all matrices whose kernel contains the kernel of Q; see [61].

Theorem 3.4.1. *[61, Theorem 1] Let $K \subset \mathrm{PSD}_k$ be a spectrahedron, and for $Q \in K$ define*

$$S(Q) = \{X \in \mathcal{S}^k \,:\, \ker(Q) \subset \ker(X)\}.$$

Then $F_K(Q) = S(Q) \cap K$.

Corollary 3.4.2. *Let $K = L \cap PSD_k$ be a spectrahedral cone for some linear subspace $L \subset \mathcal{S}^k$, and let $Q \in K$. Then Q is extremal if and only if the dimension of the affine hull of $F_K(Q)$ is one.*

This leads to the following theorem, further specialized to our case:

Theorem 3.4.2. *Let $Q_1, \ldots, Q_s$ be the symmetric matrices appearing in the blocks as in Corollary 3.3.1. Then the extremal rays of PSD_N^G are in bijection with the set of matrices $Q \in \mathrm{PSD}_N^G$ such that exactly one matrix Q_i has rank one, and the other Q_j, $j \neq i$ have rank zero, considered up to scaling by $\mathbb{R}_{\geq 0}$.*

Proof. Let $Q \in \mathrm{PSD}_N^G$ such that one Q_i has rank one and the others are zero matrices. The existence of such Q follows from Corollary 3.2.2. Without loss of generality we can assume that the $(1,1)$ entry of Q_i is nonzero. We denote this entry by a. Since the columns of Q_i are multiples of the first column and the rows are multiples of the first row we get

$$Q_i = \begin{pmatrix} a & s_2 a & \cdots & s_{m_i} a \\ s_2 a & s_2^2 a & \cdots & s_2 s_{m_i} a \\ \vdots & \vdots & \ddots & \vdots \\ s_{m_i} a & s_{m_i} s_2 a & \cdots & s_{m_i}^2 a \end{pmatrix}.$$

A basis for $\ker(Q_i)$ is

$$\begin{pmatrix} -s_2 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} -s_3 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} -s_{m_i} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

The only symmetric matrices whose kernel contains $\ker(Q_i)$ are scalar multiples of

$$\begin{pmatrix} 1 & s_2 & \cdots & s_{m_i} \\ s_2 & s_2^2 & \cdots & s_2 s_{m_i} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m_i} & s_{m_i} s_2 & \cdots & s_{m_i}^2 \end{pmatrix}.$$

This also shows that the only symmetric matrices whose kernel contains $\ker(Q)$ have the same block structure as Q where $\tilde{Q}_j = 0$ when $j \neq i$, and in $\tilde{Q}_i$ each block is a (possibly different) multiple of Q_i. But then by Theorem 3.4.1 $S(Q) \cap \mathrm{PSD}_N^G = F_{\mathrm{PSD}_N^G}(Q)$, and this consists of positive multiples of Q. Therefore the ray generated by Q is an extremal ray. Any other type of matrix in PSD_N^G is easily seen to be a conical combination of the above matrices. This proves the theorem. □

Corollary 3.4.3. *The ranks of extremal rays of* PSD_N^G *are precisely* $\{n_1, \dots, n_s\}$, $n_i = \dim V_i$. *In particular, the minimum rank attained by extremal matrices is* $\min(n_1, \dots, n_s)$, *and if no one-dimensional representation of* G *appears in* V *with positive multiplicity, this minimum rank is bigger than one.*

Note that this differs from PSD_N, whose extremal rays are defined by rank one matrices. We continue with Example 3.3.1.

Example 3.4.1. *Consider again the group* $G = I_h$ *of* 120 *symmetries of the icosahedron. The space of degree 3 polynomials has dimension* 10, *and can help us write the degree* 6 *icosahedral invariants as sums of squares. Using the Mulliken symbols for irreducible representations of* I_h *typical in chemistry [23, last page], we have that*

$$\begin{array}{lcl} \textit{vector spaces} & \mathbb{C}[x_1, x_2, x_3]_3 & = T_{1u} \oplus T_{2u} \oplus G_u \\ \textit{dimensions} & 10 & = 3 + 3 + 4. \end{array}$$

Since the minimum dimension of an irreducible in this decomposition is 3, *we can already conclude that the extremal rays of* $\mathrm{PSD}_{10}^{I_h}$ *will not be given by matrices of rank 1. The extremal rays correspond to matrices of rank at least* 3.

Similarly, since the degree 5 *polynomials decompose as*

$$\begin{array}{lcl} \textit{vector spaces} & \mathbb{C}[x_1, x_2, x_3]_5 & = 2T_{1u} \oplus 2T_{2u} \oplus G_u \oplus H_u \\ \textit{dimensions} & 21 & = 2(3) + 2(3) + 4 + 5, \end{array}$$

we know that the extremal rays are defined by matrices of rank exactly 3, 4, *and* 5 *in* $\mathrm{PSD}_{21}^{I_h}$.

Terracini convexity

In [68] the authors define Terracini convexity as a generalization of neighborliness to convex cones which are not necessarily polyhedral. Indeed, in the polyhedral case, k-Terracini is equivalent to k-neighborly. However, there are many families of non-polyhedral cones which are also Terracini convex, including the PSD cone. In what follows, we will introduce the necessary definitions and then prove that PSD_N^G is also Terracini convex.

Definition 3.4.1. *Let C be a closed, convex cone in $\mathbb{R}^d$. To a point $x \in C$, we associate the set of feasible directions into C,*

$$\mathcal{K}_C(x) = cone\{z - x : z \in C\}.$$

We call the closure of this cone, $\overline{\mathcal{K}_C(x)}$, the tangent cone *of C at x. The* convex tangent space *of C at x is*

$$\mathcal{L}_C(x) = \overline{\mathcal{K}_C(x)} \cap -\overline{\mathcal{K}_C(x)}.$$

Definition 3.4.2. *Let C be a closed, convex, pointed cone in $\mathbb{R}^d$. C is* k-Terracini convex *if for any k extremal rays $x^{(1)}, \dots, x^{(k)}$,*

$$\mathcal{L}_C\Big(\sum_{i=1}^{k} x^{(i)}\Big) = \sum_{i=1}^{k} \mathcal{L}_C(x^{(i)}) \tag{3.7}$$

Furthermore, if C is k-Terracini convex for all $k \in \mathbb{N}$, then it is Terracini convex.

Example 3.4.2. *[68, Example 1.9] The cone of positive semidefinite matrices is Terracini convex. This can be shown by proving that $\mathcal{L}_{\mathrm{PSD}_N}(X) = \{MX + XM : M \in \mathcal{S}^N\}$ and we will show something similar for the symmetry adapted PSD cone.*

Theorem 3.4.3. *Let C be a direct sum of PSD cones of varied sizes,*

$$C = \oplus_{i=1}^{s} \mathrm{PSD}_{k_i}.$$

Then C is Terracini convex.

Proof. First, we claim that $\mathcal{L}_C(X) = \{MX + XM : M \in \oplus_{i=1}^{s} \mathcal{S}^{k_i})\}$. To prove this we assume without loss of generality that $X = \mathrm{diag}(\tilde{D}_1, \dots, \tilde{D}_s)$ where

$$\tilde{D}_i = \begin{pmatrix} D_i & 0 \\ 0 & 0 \end{pmatrix}$$

is a $k_i \times k_i$ matrix and D_i is a diagonal matrix of rank r_i with positive diagonal entries. Consider the cone of feasible directions.

$$\begin{aligned}\mathcal{K}_C(X) &= \text{cone}\{Z - X : Z \in \oplus_{i=1}^{s} \text{PSD}_{k_i}\} \\ &= \text{cone}\{\text{diag}(\begin{pmatrix} Z_{11}^i - D_i & (Z_{12}^i)^t \\ Z_{12}^i & Z_{22}^i \end{pmatrix}) : \begin{pmatrix} Z_{11}^i & (Z_{12}^i)^t \\ Z_{12}^i & Z_{22}^i \end{pmatrix} \succeq 0\} \\ &= \text{cone}\{\text{diag}(\begin{pmatrix} Z_{11}^i - D_i & (Z_{12}^i)^t \\ Z_{12}^i & Z_{22}^i \end{pmatrix}) : Z_{22}^i \succeq 0, Z_{11}^i - Z_{12}^i (Z_{22}^i)^\dagger Z_{12}^i \succeq 0, \\ &\qquad (I - Z_{22}^i (Z_{22}^i)^\dagger) Z_{12}^i = 0\}\end{aligned}$$

where $Y^\dagger$ is the pseudo inverse of Y, I is the identity matrix of appropriate size, and the last equality follows from a characterization of positive semidefiniteness via the Schur complement. Note that we are coning over these matrices, so any positive scaling of matrices that satisfy these conditions is in the cone of feasible directions. This tells us that the top left block can be anything and we can further rewrite our cone as

$$\mathcal{K}_C(X) = \{\text{diag}(\begin{pmatrix} W_{11}^i & (W_{12}^i)^t \\ W_{12}^i & W_{22}^i \end{pmatrix}) : W_{22}^i \succeq 0, col(W_{12}^i) \subset col(W_{22}^i)\}$$

where *col* refers to column space. Now if we take the closure of this we get that

$$\overline{\mathcal{K}_C(X)} = \{\text{diag}(\begin{pmatrix} W_{11}^i & (W_{12}^i)^t \\ W_{12}^i & W_{22}^i \end{pmatrix}) : W_{22}^i \succeq 0\}$$

hence

$$\begin{aligned}\mathcal{L}_C(X) &= \overline{\mathcal{K}_C(X)} \cap -\overline{\mathcal{K}_C(X)} \\ &= \{\text{diag}(\begin{pmatrix} W_{11}^i & (W_{12}^i)^t \\ W_{12}^i & 0 \end{pmatrix})\} \\ &= \{MX + XM : M \in \oplus_{i=1}^{s} \mathcal{S}^{k_i}\}.\end{aligned}$$

With this characterization of the convex tangent space, it is easy to see that C is Terracini convex since

$$\begin{aligned}\mathcal{L}_C(\sum_i X^{(i)}) &= \{M(\sum_i X^{(i)}) + (\sum_i X^{(i)})M : M \in \oplus_{i=1}^{s} \mathcal{S}^{k_i}\} \\ &= \{\sum_i MX^{(i)} + X^{(i)}M : M \in \oplus_{i=1}^{s} \mathcal{S}^{k_i}\}\end{aligned}$$

and

$$\sum \mathcal{L}_C(X^{(i)}) = \sum \{MX^{(i)} + X^{(i)}M : M \in \oplus_{i=1}^{s} \mathcal{S}^{k_i}\}$$

for any k extreme points. □

Remark 3. *This proof was adapted from the proof that* $\mathcal{L}_{\mathrm{PSD}_N}(X) = \{MX+XM : M \in \mathcal{S}^N\}$ *which Professor Chandrasekaran was kind enough to share with me in our discussion of his paper [68].*

Corollary 3.3.1 tells us that PSD_N^G is exactly a direct sum of PSD cones. Hence we have the following Corollary to Theorem 3.4.3,

Corollary 3.4.4. PSD_N^G *is Terracini convex.*

Trivial Block

Here we turn to $G = S_n$ acting on $V = \mathbb{C}[x_1, \ldots, x_n]_d$ by permuting the indices of the indeterminates. For all n and d, the trivial representation appears in V with multiplicity equal to $p = p(n, d)$ where $p(n, d)$ is the number of partitions of d with at most n parts via Theorem 3.1.5. Therefore, in a symmetry adapted basis, there is one $p \times p$ diagonal block corresponding to the trivial representation, called the trivial block.

We now use Algorithm 1 to build the trivial block for any n and d in the case of S_n. Note that we may always use degree d monomials in n variables as a basis for V when $G = S_n$. To start, we order our monomial basis so that orbits of $G = S_n$ acting on the finite set of monomials are grouped together. For example, for degree 3 monomials in 3 variables, we could order our basis as

$$\{x_1^3, x_2^3, x_3^3, x_1^2x_2, x_1^2x_3, x_1x_2^2, x_1x_3^2, x_2^2x_3, x_2x_3^2, x_1x_2x_3\}$$

which has three orbits Gv for $v \in \{x_1^3, x_1^2x_2, x_1x_2x_3\}$. Note that in general the orbits can be labeled by partitions of d with $\leq n$ parts. Under this ordering, a general symmetric matrix will be described by the blocks indexed by the orbits of our monomials

$$Q = \begin{array}{c} \\ O(x^{\lambda^{(1)}}) \\ O(x^{\lambda^{(2)}}) \\ \vdots \\ O(x^{\lambda^{(p)}}) \end{array} \begin{array}{c} \begin{array}{cccc} O(x^{\lambda^{(1)}}) & O(x^{\lambda^{(2)}}) & \cdots & O(x^{\lambda^{(p)}}) \end{array} \\ \left(\begin{array}{cccc} & & & \\ & & & \\ & & & \\ & & & \end{array} \right) \end{array}.$$

Proposition 3.4.1. *Let* $Q \in \mathrm{PSD}_N^{S_n}$ *be an* $N \times N$ *symmetric matrix represented in the monomial basis ordered with respect to the orbits* $O(x^{\lambda^{(1)}}), O(x^{\lambda^{(2)}}), \ldots, O(x^{\lambda^{(p)}})$. *Let* $\Lambda_{i,j}$ *be the submatrix of* Q *indexed by* $O(x^{\lambda^{(i)}})$ *and* $O(x^{\lambda^{(j)}})$ *on the rows and columns, respectively. Let* $s_i = \sqrt{|O(x^{\lambda^{(i)}})|}$. *Then there exists an orthogonal change of basis matrix* T *such that the* *trivial block of T^TQT is*

$$Q^{\square\square\cdots\square} = \begin{pmatrix} \frac{s_1^2}{s_1^2} colsum(\Lambda_{1,1}) & \frac{s_2^2}{s_1 s_2} colsum(\Lambda_{1,2}) & \cdots & \frac{s_p^2}{s_1 s_p} colsum(\Lambda_{1,p}) \\ \frac{s_1^2}{s_1 s_2} rowsum(\Lambda_{1,2}) & \frac{s_2^2}{s_2^2} colsum(\Lambda_{2,2}) & \cdots & \frac{s_p^2}{s_2 s_p} colsum(\Lambda_{2,p}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_1^2}{s_1 s_p} rowsum(\Lambda_{1,p}) & \frac{s_2^2}{s_2 s_p} rowsum(\Lambda_{2,p}) & \cdots & \frac{s_p^2}{s_p^2} colsum(\Lambda_{p,p}) \end{pmatrix}$$

where $colsum(\Lambda_{i,j})$ is the sum of the entries of any column of $\Lambda_{i,j}$ and $rowsum(\Lambda_{i,j})$ is the sum of the entries of any row of $\Lambda_{i,j}$.

Proof. We follow Algorithm 1. Since $n_1 = 1$, only the very first step needs to be executed. Moreover, $d^1(g) = [1]$ for all $g \in S_n$, and $D(g)$ are block diagonal in the basis given by the orbits for each g. Hence π^1 is block diagonal with p blocks of size $s_i^2 \times s_i^2, i = 1, \dots p$, along the diagonal. It is not hard to see that block i is a multiple of the $s_i^2 \times s_i^2$ matrix with every entry equal to one. Therefore the first p columns of T are

$$T = \begin{matrix} O(x^{\lambda^{(1)}}) \\ O(x^{\lambda^{(2)}}) \\ \\ O(x^{\lambda^{(p)}}) \end{matrix} \begin{pmatrix} \overline{1/s_1} & \bar{0} & \cdots & \bar{0} & & \\ \bar{0} & \overline{1/s_2} & \cdots & \bar{0} & & \\ \vdots & \vdots & \ddots & \vdots & \cdots & \cdots \\ \bar{0} & \bar{0} & \cdots & \overline{1/s_p} & & \end{pmatrix}$$

with the bar indicating a column vector. Now in T^TQT the trivial block has the stated form. □

3.5 Binary and Quadratic Symmetric Polynomials

In this section we first fix the number of variables $n = 2$ and consider the structure of the symmetry adapted PSD_N^G cone. In this case, the matrices have size $N = d + 1$. We choose the monomial basis $\{x^d, x^{d-1}y, \dots, xy^{d-1}, y^d\}$, and the symmetric matrices will be $Q = (q_{ij})$. Moreover, we restrict to matrices Q such that $QD(\sigma) = D(\sigma)Q$ where $\sigma = (1\,2)$.

Corollary 3.5.1. *When $n = 2$ the dimension of the symmetry adapted PSD cone is*

$$\dim \mathrm{PSD}_{d+1}^{S_2} = \begin{cases} \frac{(d+1)(d+3)}{4} & d \text{ is odd} \\ \frac{(d+2)^2}{4} & d \text{ is even} \end{cases}$$

Proof. The hook lengths are $h_1 = 2$ and $h_2 = 1$ for both partitions of $n = 2$ corresponding to the trivial and alternating representations. Furthermore, $b(\square\square) = 0$ and $b(\boxminus) = 1$. Thus we can fill out the following table,

d	Partition	$h^T y = d - b(\lambda)$	m_λ
odd	$\yng(2)$	$y_1 + 2y_2 = d$	$\frac{d+1}{2}$
	$\yng(1,1)$	$y_1 + 2y_2 = d - 1$	$\frac{d+1}{2}$
even	$\yng(2)$	$y_1 + 2y_2 = d$	$\frac{d}{2} + 1$
	$\yng(1,1)$	$y_1 + 2y_2 = d - 1$	$\frac{d}{2}$

By Corollary 3.4.1 we need to compute

$$\dim \mathrm{PSD}_{d+1}^{S_2} = \binom{m_{\yng(2)} + 1}{2} + \binom{m_{\yng(1,1)} + 1}{2},$$

and this gives the result. □

Proposition 3.5.1. *There exists a change of basis matrix so that every $Q \in \mathrm{PSD}_N^{S_2}$ with $N = d+1$ is of the form*

$$\frac{1}{2}\begin{pmatrix} Q^{\yng(2)} & \\ & Q^{\yng(1,1)} \end{pmatrix}$$

where if d is odd

$$Q^{\yng(2)} = \begin{pmatrix} q_{11} + q_{1N} & q_{12} + q_{1(N-1)} & \cdots & q_{1\frac{N}{2}} + q_{1(\frac{N}{2}+1)} \\ q_{12} + q_{1(N-1)} & q_{22} + q_{2(N-1)} & \cdots & q_{2\frac{N}{2}} + q_{2(\frac{N}{2}+1)} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1\frac{N}{2}} + q_{1(\frac{N}{2}+1)} & q_{2\frac{N}{2}} + q_{2(\frac{N}{2}+1)} & \cdots & q_{\frac{N}{2}\frac{N}{2}} + q_{\frac{N}{2}(\frac{N}{2}+1)} \end{pmatrix}$$

and

$$Q^{\yng(1,1)} = \begin{pmatrix} q_{11} - q_{1N} & q_{12} - q_{1(N-1)} & \cdots & q_{1\frac{N}{2}} - q_{1(\frac{N}{2}+1)} \\ q_{12} - q_{1(N-1)} & q_{22} - q_{2(N-1)} & \cdots & q_{2\frac{N}{2}} - q_{2(\frac{N}{2}+1)} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1\frac{N}{2}} - q_{1(\frac{N}{2}+1)} & q_{2\frac{N}{2}} - q_{2(\frac{N}{2}+1)} & \cdots & q_{\frac{N}{2}\frac{N}{2}} - q_{\frac{N}{2}(\frac{N}{2}+1)} \end{pmatrix},$$

while if d is even there are an extra row and column in the trivial block

$$Q^{\yng(2)} = \begin{pmatrix} q_{11} + q_{1N} & q_{12} + q_{1(N-1)} & \cdots & q_{1\frac{N-1}{2}} + q_{1\frac{N+3}{2}} & \sqrt{2}q_{1\frac{N+1}{2}} \\ q_{12} + q_{1(N-1)} & q_{22} + q_{2(N-1)} & \cdots & q_{2\frac{N-1}{2}} + q_{2\frac{N+3}{2}} & \sqrt{2}q_{2\frac{N+1}{2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{1\frac{N-1}{2}} + q_{1\frac{N+3}{2}} & q_{2\frac{N-1}{2}} + q_{2\frac{N+3}{2}} & \cdots & q_{\frac{N-1}{2}\frac{N-1}{2}} + q_{\frac{N-1}{2}\frac{N+3}{2}} & \sqrt{2}q_{\frac{N-1}{2}\frac{N+1}{2}} \\ \sqrt{2}q_{1\frac{N+1}{2}} & \sqrt{2}q_{2\frac{N+1}{2}} & \cdots & \sqrt{2}q_{\frac{N-1}{2}\frac{N+1}{2}} & q_{\frac{N+1}{2}\frac{N+1}{2}} \end{pmatrix}$$

and

$$Q^{\yng(1,1)} = \begin{pmatrix} q_{11} - q_{1N} & q_{12} - q_{1(N-1)} & \cdots & q_{1\frac{N-1}{2}} - q_{1\frac{N+3}{2}} \\ q_{12} - q_{1(N-1)} & q_{22} - q_{2(N-1)} & \cdots & q_{2\frac{N-1}{2}} - q_{2\frac{N+3}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1\frac{N-1}{2}} - q_{1\frac{N+3}{2}} & q_{2\frac{N-1}{2}} - q_{2\frac{N+3}{2}} & \cdots & q_{\frac{N-1}{2}\frac{N-1}{2}} - q_{\frac{N-1}{2}\frac{N+3}{2}} \end{pmatrix}.$$

Proof. Again we follow Algorithm 1 where $d^1(g) = [1]$ and $d^2(g) = [\operatorname{sign}(g)]$ for $g \in S_2$, $D(\mathrm{id}) = I_{d+1}$ and

$$D(12) = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}.$$

Then $\pi^{\square\square} = I_{d+1} + D(12)$ and $\pi^{\substack{\square\\\square}} = I_{d+1} - D(12)$. The change of basis matrix T looks a little different depending on the parity of d:

d odd	d even
$\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 0 & & & 1 & 0 & & \\ 0 & 1 & & & 0 & 1 & & \\ & & \vdots & & & & \vdots & \\ & & & 1 & 0 & & & 1 \\ & & & 1 & 0 & & & -1 \\ & & \vdots & & & & \vdots & \\ 0 & 1 & & & 0 & -1 & & \\ 1 & 0 & & & -1 & 0 & & \end{pmatrix}$	$\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 0 & & & 1 & 0 & & \\ 0 & 1 & & & 0 & 1 & & \\ & & \vdots & & & & \vdots & 1 \\ & & & \sqrt{2} & & & & 0 \\ & & \vdots & & & & \vdots & -1 \\ 0 & 1 & & & 0 & -1 & & \\ 1 & 0 & & & -1 & 0 & & \end{pmatrix}$

By computing T^TQT we get $Q^{\square\square}$ and $Q^{\substack{\square\\\square}}$ in both cases. $\square$

Example 3.5.1. *Consider the symmetric polynomial inequality $P_4(x,y) \geq P_{1111}(x,y)$ where $P_4(x,y) = \frac{1}{2}(x^4+y^4)$ and $P_{1111}(x,y) = \frac{1}{16}(x+y)^4$. It is proven in [24] that this inequality holds over the nonnegative orthant. We can certify this inequality via sums of squares. First, define the polynomial,*

$$\begin{aligned} f(x,y) = (P_4 - P_{1111})(x^2,y^2) &= \frac{1}{2}(x^8+y^8) - \frac{1}{16}(x^2+y^2)^4 \\ &= \frac{7}{16}x^8 - \frac{1}{4}x^6y^2 - \frac{3}{8}x^4y^4 - \frac{1}{4}x^2y^6 + \frac{7}{16}y^8 \end{aligned}$$

and note that if f is SOS, then the above inequality holds for $x, y \geq 0$. Next, assume that $f = [x]^TQ[x]$ where $Q = (q_{ij})$ is a 5×5 symmetric matrix in the monomial basis $[x]^T = [x^4, x^3y, x^2y^2, xy^3, y^4]$. We equate the coefficients of $f(x,y)$ and

$$\begin{aligned} [x]^TQ[x] = q_{11}x^8 + 2q_{12}x^7y + (2q_{13}+q_{22})x^6y^2 + (2q_{14}+2q_{23})x^5y^3 + (2q_{15}+2q_{24}+q_{33})x^4y^4 \\ + (2q_{14}+2q_{23})x^3y^5 + (2q_{13}+q_{22})x^2y^6 + 2q_{12}xy^7 + q_{11}y^8 \end{aligned}$$

to find out $q_{11} = \frac{7}{16}$, $q_{12} = 0$, $q_{22} = -\frac{1}{4} - 2q_{13}$, $q_{23} = -q_{14}$, and $q_{33} = -\frac{3}{8} - 2q_{15} - 2q_{24}$. Substituting these into the computed matrices in Proposition 3.5.1 for $d = 4$ (even), our *matrix Q becomes*

$$\begin{pmatrix} q_{15}+\frac{7}{16} & q_{14} & \sqrt{2}q_{13} & 0 & 0 \\ q_{14} & -2\,q_{13}+q_{24}-\frac{1}{4} & -\sqrt{2}q_{14} & 0 & 0 \\ \sqrt{2}q_{13} & -\sqrt{2}q_{14} & -2\,q_{15}-2\,q_{24}-\frac{3}{8} & 0 & 0 \\ 0 & 0 & 0 & -q_{15}+\frac{7}{16} & -q_{14} \\ 0 & 0 & 0 & -q_{14} & -2\,q_{13}-q_{24}-\frac{1}{4} \end{pmatrix}.$$

Now we can run an SDP on this to certify that f is SOS. One rank two solution is

$$\begin{pmatrix} \frac{7}{8} & 0 & -\frac{7\sqrt{2}}{8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{7\sqrt{2}}{8} & 0 & \frac{7}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

This is indeed positive semidefinite and thus $f(x,y) = (P_4 - P_{1111})(x^2, y^2)$ is SOS.

In the rest of the section we consider symmetric quadratic polynomials ($d = 1$) in any number of variables n. Then $N = n$ and $D(g)$ are the $n \times n$ permutation matrices represented in the monomial basis $\{x_1, x_2, \ldots, x_n\}$. It is not hard to see that in this basis all $n \times n$ symmetric matrices which commute with all permutation matrices are of the form

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{12} \\ q_{12} & q_{11} & \cdots & q_{12} \\ \vdots & \vdots & \ddots & \vdots \\ q_{12} & q_{12} & \cdots & q_{11} \end{pmatrix}.$$

In a symmetry adapted basis the matrices look even simpler.

Proposition 3.5.2. *There is a change of basis matrix such that every $Q \in \mathrm{PSD}_n^{S_n}$ is of the form*

$$\begin{pmatrix} q_{11}+(n-1)q_{12} & 0 & \cdots & 0 \\ 0 & q_{11}-q_{12} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{11}-q_{12} \end{pmatrix}.$$

Proof. The representation in question is the permutation representation of S_n. By Theorem 3.1.5, the trivial representation and the standard representation both appear with multiplicity one. This tells us that there will be one 1×1 block associated to the trivial representation, and $n-1$ copies of a 1×1 block associated to the standard representation. The application of Algorithm 1 yields the desired diagonal matrix. ☐

Corollary 3.5.2. *An $n \times n$ symmetric matrix $Q = (q_{ij})$ which commutes with $D((1\,2))$ is in $\mathrm{PSD}_n^{S_n}$ if and only if $q_{12} \leq q_{11}$ and $q_{12} \geq \frac{-1}{n-1}q_{11}$. Hence $\mathrm{PSD}_n^{S_n}$ is a two-dimensional polyhedral cone defined by these linear inequalities.*

Corollary 3.5.3. *Let $f(x) = a\sum_i x_i^2 + b\sum_{i<j} x_i x_j$ be a symmetric quadratic form. Then f is SOS if and only if $\frac{-1}{n-1}a \leq b \leq a$. Moreover, the symmetry adapted Gram spectrahedron $K_f^{S_n}$ is either empty or an isolated point in $\mathrm{PSD}_n^{S_n}$.*

Proof. Observe that $a = q_{11}$ and $b = q_{12}$ by (3.2). Clearly, $K_f^{S_n} = \{(a,b)\}$ if and only if f is SOS. □

Corollary 3.5.4. *Symmetric quadratic SOS forms can only be written as a sum of one, $n-1$, or n squares.*

Proof. Let $f(x) = a\sum_i x_i^2 + b\sum_{i<j} x_i x_j$ be a symmetric quadratic form and consider (3.2) with $Q \in \mathrm{PSD}_N$,

$$a\sum_i x_i^2 + b\sum_{i<j} x_i x_j = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{12} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \cdots \\ q_{1n} & q_{2n} & \cdots & q_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Equating coefficients we see that $a = q_{ii}$ and $b = q_{ij}$ for $i \neq j$, the same structure as any invariant matrix. Thus $\mathrm{PSD}_n^{S_n}$ is in fact representative of all SOS decompositions of symmetric quadratic SOS forms. Now, if the point (a,b) is in the interior of $\mathrm{PSD}_n^{S_n}$, the corresponding matrix has full rank. If it is on the extreme ray defined by $q_{12} = q_{11}$, the matrix rank will be one as all the blocks $q_{12} - q_{11}$ will be zero. Lastly, if it is on the other extreme ray, we get a rank $n-1$ matrix. □

Finally, we consider what happens as the number of variables goes to infinity. In particular, note that the slope of $q_{12} = \frac{-1}{n-1}q_{11}$ goes to zero. This leads to the following result.

Theorem 3.5.1. *As n goes to infinity, the ratio of SOS symmetric quadratic forms in n variables to all symmetric quadratic forms in n variables is $\frac{1}{8}$.*

3.6 Ternary Symmetric Polynomials

In this section we consider the case where $n = 3$ and $N = \frac{1}{2}(d+2)(d+1)$.

Proposition 3.6.1. *Let $V = \mathbb{C}[x_1, x_2, x_3]_d$ be the representation of S_3 induced by permuting the variables. Then the multiplicities of the trivial, standard, and alternating irreducible representations are as in the following table*

Partition	$h^T y = d - b(\lambda)$	m_λ
(3)	$y_1 + 2y_2 + 3y_3 = d$	$Q(d)$
$(2,1)$	$y_1 + y_2 + 3y_3 = d - 1$	$P(d-1)$
$(1,1,1)$	$y_1 + 2y_2 + 3y_3 = d - 3$	$Q(d-3)$

where $Q(d)$ and $P(d)$ are quasi-polynomials as below:

$$Q(d) = \begin{cases} \frac{1}{12}d^2 + \frac{1}{2}d + 1 & d \equiv 0 \mod 6 \\ \frac{1}{12}d^2 + \frac{1}{2}d + \frac{5}{12} & d \equiv 1 \mod 6 \\ \frac{1}{12}d^2 + \frac{1}{2}d + \frac{2}{3} & d \equiv 2 \mod 6 \\ \frac{1}{12}d^2 + \frac{1}{2}d + \frac{3}{4} & d \equiv 3 \mod 6 \\ \frac{1}{12}d^2 + \frac{1}{2}d + \frac{2}{3} & d \equiv 4 \mod 6 \\ \frac{1}{12}d^2 + \frac{1}{2}d + \frac{5}{12} & d \equiv 5 \mod 6 \end{cases} \qquad P(d) = \begin{cases} \frac{1}{6}d^2 + \frac{5}{6}d + 1 & d \equiv 0 \mod 3 \\ \frac{1}{6}d^2 + \frac{5}{6}d + 1 & d \equiv 1 \mod 3 \\ \frac{1}{6}d^2 + \frac{5}{6}d + \frac{2}{3} & d \equiv 2 \mod 3 \end{cases}$$

Proof. The multiplicities are computed using Theorem 3.1.5. In all three cases, they are given by the Ehrhart quasi-polynomial [5] of a rational 2-simplex. For instance, for the trivial representation we wish to count the number of nonnegative integer solutions to the equation $y_1 + 2y_2 + 3y_3 = d$. This is the number of lattice points in the polytope defined by the hyperplane $y_1 + 2y_2 + 3y_3 = d$ and $y_1, y_2, y_3 \geq 0$. The vertices of this polytope are $(d, 0, 0), (0, d/2, 0), (0, 0, d/3)$, and it is the dth dilation of the polytope for $d = 1$. The lattice point count is given by the quasi-polynomial $Q(d)$ as in the statement. Similarly, for the multiplicity of the standard representation, the Ehrhart quasi-polynomial $P(d)$ of a different two-simplex is needed. □

Symmetric Ternary Quartics

Now we consider symmetric polynomials in three variables of degree four ($n = 3, d = 2$). The study of general ternary quartics has a long history. It is known that a smooth ternary quartic can always be written as $f = q_1^2 + q_2^2 + q_3^2$ where $q_i \in \mathbb{C}[x_1, x_2, x_3]_2$, and there are exactly 63 nonequivalent ways of doing that [21, Ch.1, §14]. There are always 28 bitangents to the smooth projective plane curve defined by f, and certain sixtuples of pairs of these bitangents, known as *Steiner complexes*, correspond to these 63 different representations; see [59, Section 5]. Moreover, for real smooth ternary quartics there are exactly 8 SOS representations with three squares [60]. This means that the usual Gram spectrahedron K_f has exactly 8 vertices corresponding to matrices of rank 3.

In this section, we want to study the symmetry adapted Gram spectrahedron $K_f^{S_3}$. The main objects of focus are the symmetric matrices $Q = (q_{ij}) \in \mathcal{S}^6$ such that $f(x) = [x]^T Q[x]$.

Proposition 3.6.2. *The symmetry adapted* $\mathrm{PSD}_6^{S_3}$ *is a six-dimensional cone consisting of positive semidefinite matrices of the form*

$$\begin{pmatrix} q_{11}+2q_{12} & 2q_{14}+q_{16} & 0 & 0 & 0 & 0 \\ 2q_{14}+q_{16} & q_{44}+2q_{45} & 0 & 0 & 0 & 0 \\ 0 & 0 & q_{11}-q_{12} & q_{14}-q_{16} & 0 & 0 \\ 0 & 0 & q_{14}-q_{16} & q_{44}-q_{45} & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{11}-q_{12} & q_{14}-q_{16} \\ 0 & 0 & 0 & 0 & q_{14}-q_{16} & q_{44}-q_{45} \end{pmatrix}.$$

Proof. Proposition 3.6.1 tells us that the multiplicities of the trivial and standard representations are each two, and that of the alternating representation is zero. By Corollary 3.4.1 the dimension of $\mathrm{PSD}_6^{S_3}$ is six. Using Algorithm 1, one can compute a 6×6 change of basis matrix such that the elements in $\mathrm{PSD}_6^{S_3}$ have the stated form. □

The next theorem is our main theorem in this section.

Theorem 3.6.1. *Let $f \in \mathbb{R}[x, y, z]$ be a smooth symmetric quartic. Then there are precisely* 3 *(possibly complex) symmetric matrices Q of rank* 3 *such that $f = [x]^T Q[x]$ and $D(\sigma)Q = QD(\sigma)$ for all $\sigma \in S_3$. Moreover, if f is SOS, there are exactly* 2 *such PSD matrices of rank* 3*. These correspond to the two vertices of the two-dimensional symmetry adapted Gram spectrahedron $K_f^{S_3}$. Furthermore, the boundary of $K_f^{S_3}$ is defined by two curves, a parabola and a hyperbola. Other than the two vertices, the points along the hyperbola give rank* 4 *matrices while those along the parabola are rank* 5 *matrices.*

Proof. Let

$$f(x_1, x_2, x_3) = a\sum_i x_i^4 + b\sum_{i\neq j} x_i^3 x_j + c\sum_{i<j} x_i^2 x_j^2 + d\sum_{i\neq j\neq k, j<k} x_i^2 x_j x_k$$

where a, b, c, d are fixed coefficients. Writing $f = [x]^T Q[x]$ and equating coefficients we get that $a = q_{11}$, $b = 2q_{14}$, $c = 2q_{12} + q_{44}$, and $d = 2q_{16} + 2q_{45}$. If we plug these into the block-diagonalized matrix in Proposition 3.6.2 we see that the symmetry adapted Gram spectrahedron $K_f^{S_3}$ consists of positive semidefinite matrices of the form

$$\begin{pmatrix} a+2q_{12} & b+q_{16} & 0 & 0 & 0 & 0 \\ b+q_{16} & c+d-2q_{12}-2q_{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & a-q_{12} & \frac{b}{2}-q_{16} & 0 & 0 \\ 0 & 0 & \frac{b}{2}-q_{16} & c-\frac{d}{2}-2q_{12}+q_{16} & 0 & 0 \\ 0 & 0 & 0 & 0 & a-q_{12} & \frac{b}{2}-q_{16} \\ 0 & 0 & 0 & 0 & \frac{b}{2}-q_{16} & c-\frac{d}{2}-2q_{12}+q_{16} \end{pmatrix}.$$

Hence, $K_f^{S_3}$ is the intersection of two spectrahedra:

$$K_1 \;=\; \{(q_{12}, q_{16}) : \begin{pmatrix} a+2q_{12} & b+q_{16} \\ b+q_{16} & c+d-2q_{12}-2q_{16} \end{pmatrix} \succeq 0\} \tag{3.8}$$

$$K_2 = \{(q_{12}, q_{16}) : \begin{pmatrix} a - q_{12} & \frac{b}{2} - q_{16} \\ \frac{b}{2} - q_{16} & c - \frac{d}{2} - 2q_{12} + q_{16} \end{pmatrix} \succeq 0\} \tag{3.9}$$

To prove the first statement in our theorem we ignore the condition that these matrices need to be positive semidefinite. The above 6×6 matrix has rank three if and only if the two 2×2 matrices have rank one. Thus their determinants must be zero. This gives us two quadratics in the variables q_{12} and q_{16} which we homogenize using a new variable q:

$$\begin{aligned} p_1 &= -4q_{12}^2 - 4q_{12}q_{16} - q_{16}^2 + q((-2a + 2c + 2d)q_{12} + (-2a - 2b)q_{16}) + q^2(ac + ad - b^2) \\ p_2 &= 2q_{12}^2 - q_{12}q_{16} - q_{16}^2 + q((-2a - c + d/2)q_{12} + (a + b)q_{16}) + q^2(ac - ad/2 - b^2/4). \end{aligned}$$

By Bezout's theorem, the projective plane curves defined by p_1 and p_2 intersect at exactly 4 complex points. Setting $q = 0$, we consider the solutions to the equations

$$\begin{aligned} 0 &= -4q_{12}^2 - 4q_{12}q_{16} - q_{16}^2 = -(2q_{12} + q_{16})^2 \\ 0 &= 2q_{12}^2 - q_{12}q_{16} - q_{16}^2 = (2q_{12} + q_{16})(q_{12} - q_{16}). \end{aligned}$$

We see that there is only one solution, giving us $[q_{12} : q_{16} : q] = [1 : -2 : 0]$ as the intersection point at the line at infinity. The remaining three points are obtained by setting $q = 1$ which gives us back the determinants of the two submatrices. This proves the first statement.

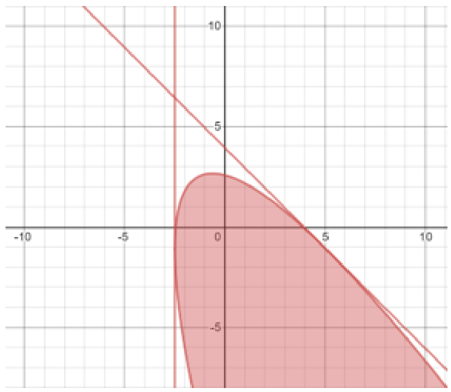
Next we consider the spectrahedra K_1 and K_2. For fixed a, b, c, d, K_1 is defined by the inequalities

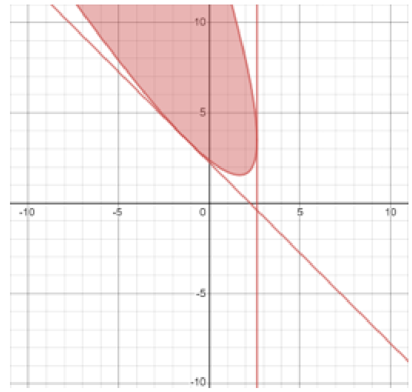
$$\begin{aligned} (a + 2q_{12})(c + d - 2q_{12} - 2q_{16}) - (b + q_{16})^2 &\geq 0 \\ a + 2q_{12} &\geq 0 \\ c + d - 2q_{12} - 2q_{16} &\geq 0. \end{aligned}$$

The first quadratic can be rewritten as

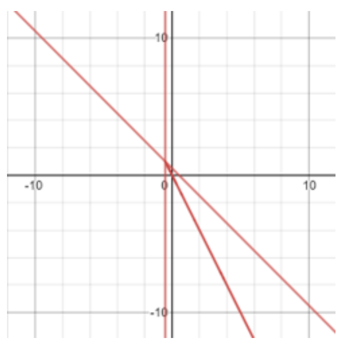
$$\begin{pmatrix} q_{12} & q_{16} & 1 \end{pmatrix} \begin{pmatrix} -4 & -2 & -a + c + d \\ -2 & -1 & -a - b \\ -a + c + d & -a - b & ac + ad - b^2 \end{pmatrix} \begin{pmatrix} q_{12} \\ q_{16} \\ 1 \end{pmatrix} \geq 0.$$

Since the determinant of the upper left 2×2 matrix is zero, the curve defined by this quadric is a parabola [33, Table 5.3]. Moreover, the lines $a + 2q_{12} = 0$ and $c + d - 2q_{12} - 2q_{16} = 0$ are tangent to the curve at the points $(-\frac{a}{2}, -b)$ and $(b + \frac{c}{2} + \frac{d}{2}, -b)$ respectively. As we vary a, b, c, d, the region defined by the first inequality moves between only two of the four connected components in the complement of the two lines as illustrated below:





Moreover, by the last two inequalities, K_1 is nonempty when the parabola is in the bottom region, as in the left most figure. It is worth noting that this is the generic case and that there is one more possibility. If the determinant of the above matrix is zero, i.e., $(a+2b+c+d)^2 = 0$, then the quadric defines a double line [33, Table 5.3], $(a - c - d + 4q_{12} + 2q_{16})^2 = 0$. This double line intersects the lines $a + 2q_{12} = 0$ and $c + d - 2q_{12} - 2q_{16} = 0$ at the same point. Thus K_1 is a ray, starting from this intersection point and going out to $(\infty, -\infty)$:



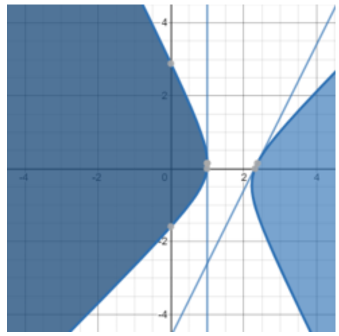
We can do a similar analysis of K_2 which is defined by the inequalities

$$\begin{aligned} (a - q_{12})(c - \frac{d}{2} - 2q_{12} + q_{16}) - (\frac{b}{2} - q_{16})^2 &\geq 0 \\ a - q_{12} &\geq 0 \\ c - \frac{d}{2} - 2q_{12} + q_{16} &\geq 0 \end{aligned}$$

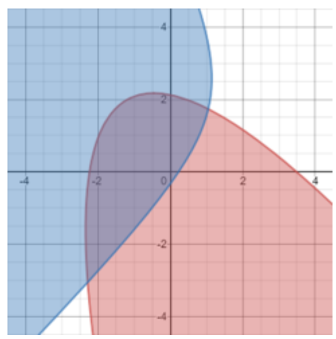
We rewrite the first quadratic as

$$\begin{pmatrix} q_{12} & q_{16} & 1 \end{pmatrix} \begin{pmatrix} 2 & -\frac{1}{2} & -a - \frac{c}{2} + \frac{d}{4} \\ -\frac{1}{2} & -1 & \frac{a}{2} + \frac{b}{2} \\ -a - \frac{c}{2} + \frac{d}{4} & \frac{a}{2} + \frac{b}{2} & -\frac{b^2}{4} + ac - \frac{ad}{2} \end{pmatrix} \begin{pmatrix} q_{12} \\ q_{16} \\ 1 \end{pmatrix}.$$

This is a hyperbola (or a pair of crossing lines) because the leading 2×2 minor is nonzero [33, Table 5.3]. Again the two additional inequalities define lines that are tangent to the curve and give K_2 as the left most component of the hyperbola:

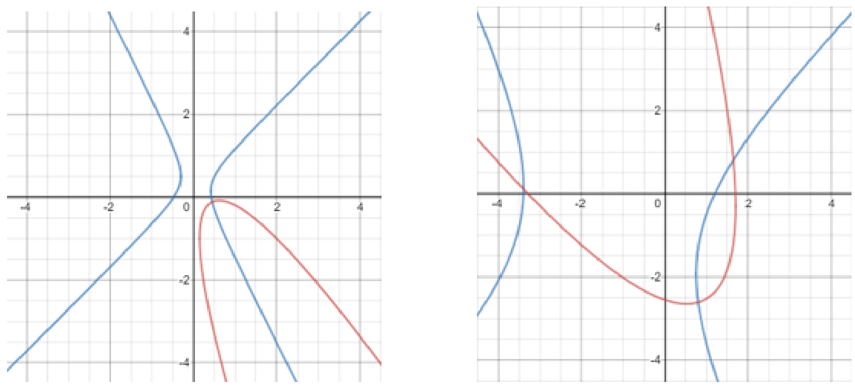


We now see that for a generic symmetric ternary quartic that is SOS, the symmetry adapted Gram spectrahedron $K_f^{S_3}$ is the intersection of the parabola and one component of the hyperbola.



The two points in $K_f^{S_3}$ where these curves intersect are the two vertices corresponding to rank three matrices. If we move along the boundary defined by the parabola, we get rank 5 matrices, because on these points the matrix block corresponding to the parabola has rank 1 while the two blocks corresponding to the hyperbola are each rank 2. A similar argument shows that matrices along the hyperbola have rank 4. □

Remark 4. *Theorem 3.6.1 illustrates one of three cases, namely, the case where f is SOS when the two quadrics defined by the determinants of the matrices in K_1 and K_2 intersect at three real points, two of which give PSD matrices. If f is not SOS, then we have two additional situations. The first is that the curves only intersect at one real point and two complex points, and the second case is when the curves have three real intersection points. In the latter, even though there are three real points, none of them correspond to a PSD matrix.*



As mentioned above, the Gram spectrahedron of an SOS ternary quartic f has 8 vertices of rank three. Let the *Steiner graph* be the graph on these vertices whose edges represent edges of the Gram spectrahedron. For a generic SOS ternary quartic the Steiner graph is $K_4 \sqcup K_4$, the disjoint union of two complete graphs on 4 vertices [59]. Moreover, the matrices along those edges are of rank at most 5. It is not known whether the Steiner graph coincides with all edges of the Gram spectrahedron. However, it is clear from Theorem 3.6.1 that, generically, there are no edges of the symmetry adapted Gram spectrahedron contributing to the edges of the Steiner graph.

Corollary 3.6.1. *The Steiner graph of the symmetry adapted Gram spectrahedron of a generic symmetric SOS ternary quartic f is the disjoint union of two vertices.*

Proof. By Theorem 3.6.1, $K_f^{S_n}$ has two vertices. Thus, either both vertices are in one complete graph K_4 or each graph contains one of the two vertices. If it were the former, then $K_f^{S_n}$ would also contain the corresponding edge. This is, however, the interior of the symmetry adapted Gram spectrahedron and all matrices there are rank 6. Thus no such edge of matrices of rank 5 exists, i.e. the vertices are each in different complete graphs. □

The vertices of the Gram spectrahedron of f or of its symmetry adapted version when f is G-invariant are not the end of the story. The boundary of these spectrahedra are very interesting and the work to unearth it is only starting. In the symmetric ternary quartics case, the boundary consists of the union of a piece of a parabola and a piece of a hyperbola. It is an interesting question how a typical SOS decomposition would look like if we used an SDP solver for $K_f^{S_n}$. It is not difficult to run simulations. Below are the results of such computations. We generated random symmetric ternary quartics and ran SDPs until we found 16 that were SOS. For each of these 16 polynomials we randomly generated 1000 objective functions and ran an SDP for each of them. The ranks of the corresponding 1000 optimal SOS matrices are shown in Figure 3.6.

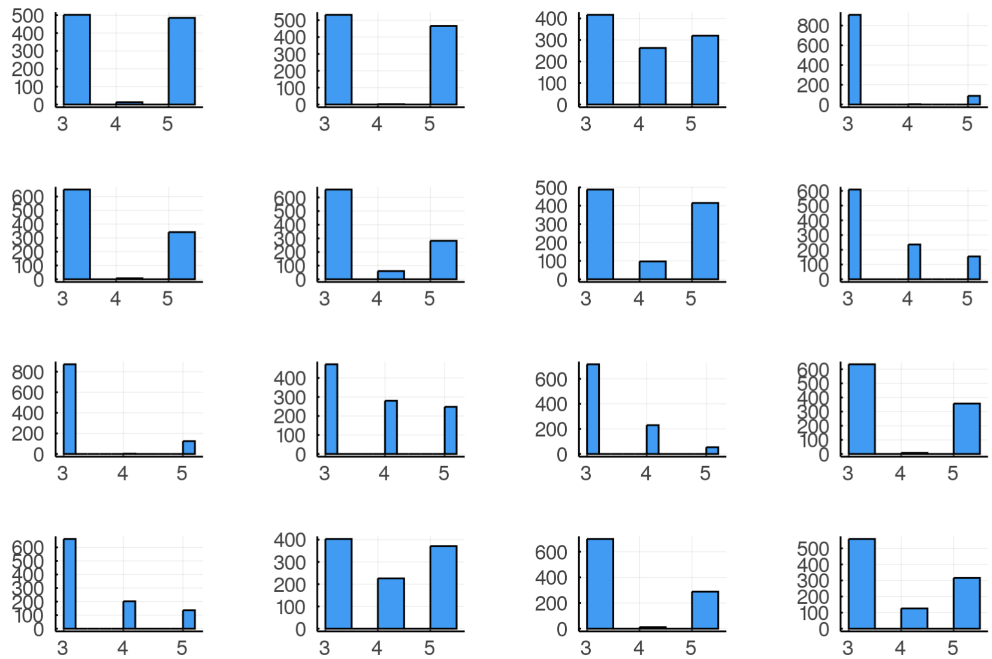
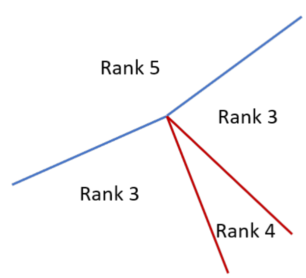


Figure 3.1: Distribution of ranks for SOS decomposition of 16 symmetric ternary quartics

Remark 5. *Computational data can provide some insight about the normal fan of the symmetry adapted Gram spectrahedron. In the generic case for a positive ternary quartic, the normal fan will be something like the following:*



Hence a random cost function is more likely to return a rank three or a rank five solution than a rank 4 solution, as reflected by the data.

We close this section with a characterization of all symmetric ternary quartics that are SOS. First we provide necessary linear conditions on the coefficients of such a polynomial. Then we report on a full characterization in a form which can be used to certify whether a symmetric ternary quartic is SOS.

Proposition 3.6.3. *If a symmetric ternary quartic*

$$f(x_1, x_2, x_3) = a\sum_i x_i^4 + b\sum_{i\neq j} x_i^3 x_j + c\sum_{i<j} x_i^2 x_j^2 + d\sum_{i\neq j\neq k, j<k} x_i^2 x_j x_k$$

with real coefficients a, b, c, d *is SOS, then*

a) $a \geq 0$,

b) $a + c \geq 0$,

c) $a + 2b + c + d \geq 0$.

Proof. The first two conditions follow from projecting the polyhedron defined by linear inequalities obtained from the four diagonals in (3.8) and (3.9). The third condition comes from applying quantifier elimination on the defining inequalities of K_1 in (3.8). □

Example 3.6.1. *As mentioned, the conditions in Proposition 3.6.3 are not sufficient. Let* $a = 1$, $b = 2$, $c = 1$, *and* $d = 0$. *Certainly,* a, $a + c$, *and* $a + 2b + c + d$ *are all nonnegative, but the corresponding polynomial,*

$$f(x_1, x_2, x_3) = \sum_i x_i^4 + 2\sum_{i\neq j} x_i^3 x_j + \sum_{i<j} x_i^2 x_j^2$$

is not SOS. In particular, $f(1, -2, 1) = -9$.

Additional conditions are not easy to find. The task is to project the spectrahedron $K_1 \cap K_2$ onto the (a, b, c, d)-space. Given that this 6-dimensional spectrahedron is a cone, one method is to consider the projection of an affine slice. We do this for $q_{16} = 1$. Then for any (a, b, c, d) in this projection, the corresponding polynomial is SOS and so is any positive scaling of that polynomial. However, for a complete description, we must also consider the projection when $q_{16} = 0$ and $q_{16} = -1$. In this way, we can find an exact description (up to positive scaling) of the semialgebraic set defined by the projection of the three affine slices when $q_{16} = 1$, $q_{16} = 0$, and $q_{16} = -1$ using quantifier elimination. The result is the union of 158 basic semialgebraic sets, each defined with polynomial inequalities and equations up to degree 4. We encourage the interested reader to visit

https://math.berkeley.edu/∼ishankar/SOSSymTernQuartic.html

for a code that will check if a given point (a, b, c, d) is contained in this set, and thus are the coefficients of an SOS polynomial. There one may also see the full description of the projected slices of the spectrahedron.

Symmetric Ternary Sextics

Here $V = \mathbb{R}[x_1, x_2, x_3]_3$ and we consider symmetric ternary sextics.

Proposition 3.6.4. *The symmetry adapted PSD cone* $\mathrm{PSD}_{10}^{S_3}$ *consists of* 10×10 *symmetric matrices of the form*

$$\begin{pmatrix} Q^{\ydiagram{3}} & & & \\ & Q^{\ydiagram{2,1}} & & \\ & & Q^{\ydiagram{2,1}} & \\ & & & Q^{\ydiagram{1,1}} \end{pmatrix}$$

where each

$$Q^{\ydiagram{3}} = \begin{pmatrix} q_{11} + 2q_{12} & \sqrt{2}(q_{14} + q_{16} + q_{18}) & \sqrt{3}q_{110} \\ \sqrt{2}(q_{14} + q_{16} + q_{18}) & q_{44} + q_{45} + q_{46} + 2q_{47} + q_{49} & \sqrt{6}q_{410} \\ \sqrt{3}q_{110} & \sqrt{6}q_{410} & q_{1010} \end{pmatrix}$$

$$Q^{\ydiagram{2,1}} = \begin{pmatrix} q_{11} - q_{12} & \frac{\sqrt{2}}{2}(2q_{14} - q_{16} - q_{18}) & \frac{\sqrt{6}}{2}(q_{16} - q_{18}) \\ \frac{\sqrt{2}}{2}(2q_{14} - q_{16} - q_{18}) & q_{44} + q_{45} - \frac{1}{2}q_{46} - q_{47} - \frac{1}{2}q_{49} & \frac{\sqrt{3}}{2}(q_{46} - q_{49}) \\ \frac{\sqrt{6}}{2}(q_{16} - q_{18}) & \frac{\sqrt{3}}{2}(q_{46} - q_{49}) & q_{44} - q_{45} + \frac{1}{2}q_{46} - q_{47} + \frac{1}{2}q_{49} \end{pmatrix}$$

$$Q^{\ydiagram{1,1}} = q_{44} - q_{45} - q_{46} + 2q_{47} - q_{49}$$

is positive semidefinite.

Proof. The multiplicities of the trivial, standard, and alternating irreducible representations are three, three, and one, respectively. Algorithm 1 provides a change of basis matrix T such that every positive semidefinite matrix $Q = (q_{ij})$ that commutes with $D(\sigma)$ for $\sigma \in S_3$ is of the above form after computing $T^T Q T$. □

It has been proved by Scheiderer [69, Corollary 3.5] that every generic ternary sextic that is SOS admits a representation using four squares; in other words, the corresponding Gram spectrahedron has extreme rays consisting of matrices of rank 4. Our main theorem in this section establishes four as the minimal rank for generic symmetric ternary sextics that are SOS using the technology of Gröbner bases.

Theorem 3.6.2. *Let* $f \in \mathbb{R}[x_1, x_2, x_3]_6$ *be a generic symmetric polynomial. If* f *is SOS, the symmetry adapted Gram spectrahedron has extreme points consisting of matrices of rank* 4.

Proof. The polynomial f is parametrized by 7 coefficients which we call $a_1, a_2, \ldots, a_7$. It is also represented as $f = [x]^T Q[x]$ by a 10×10 symmetric matrix $Q = (q_{ij})$. After equating coefficients and using a symmetry adapted basis we get a block-diagonal Q where

$$Q_f^{\ydiagram{3}} = \begin{pmatrix} a_1 + 2q_{12} & \sqrt{2}(\frac{a_2}{2} + q_{16} + q_{18}) & \sqrt{3}q_{110} \\ \sqrt{2}(\frac{a_2}{2} + q_{16} + q_{18}) & \alpha & \sqrt{6}q_{410} \\ \sqrt{3}q_{110} & \sqrt{6}q_{410} & a_7 - 6q_{49} \end{pmatrix}$$

$$Q_f^{\square\square} = \begin{pmatrix} a_1 - q_{12} & \frac{\sqrt{2}}{2}(a_2 - q_{16} - q_{18}) & \frac{\sqrt{6}}{2}(q_{16} - q_{18}) \\ \frac{\sqrt{2}}{2}(a_2 - q_{16} - q_{18}) & \beta_1 & \frac{\sqrt{3}}{2}(\frac{a_5}{2} - q_{12} - q_{49}) \\ \frac{\sqrt{6}}{2}(q_{16} - q_{18}) & \frac{\sqrt{3}}{2}(\frac{a_5}{2} - q_{12} - q_{49}) & \beta_2 \end{pmatrix}$$

$$Q_f^{\boxminus} = a_3 - \frac{a_4}{2} - \frac{a_5}{2} + a_6 + q_{12} - 2q_{16} - 2q_{18} + q_{110} - q_{49} - 2q_{410}$$

where

$$\begin{aligned} \alpha &= a_3 + \frac{a_4}{2} + \frac{a_5}{2} + a_6 - q_{12} - 2q_{16} - 2q_{18} - q_{110} + q_{49} - 2q_{410} \\ \beta_1 &= a_3 + \frac{a_4}{2} - \frac{a_5}{4} - \frac{a_6}{2} + \frac{q_{12}}{2} - 2q_{16} + q_{18} - q_{110} - \frac{q_{49}}{2} + q_{410} \\ \beta_2 &= a_3 - \frac{a_4}{2} + \frac{a_5}{4} - \frac{a_6}{2} - \frac{q_{12}}{2} - 2q_{16} + q_{18} + q_{110} + \frac{q_{49}}{2} + q_{410}. \end{aligned}$$

Now, to get a matrix of rank of 3, we have four cases:

a) Trivial block has rank 3 and all other blocks have rank zero.

b) Trivial block has rank 2 and the alternating block has rank 1.

c) Trivial block and standard block have rank 1 each.

d) Standard block and alternating block have rank 1 each.

In the first case we set all of the linear forms in the standard block and the alternating block to zero and eliminate q_{ij} from the ideal generated by these polynomials using a Gröbner basis. The elimination ideal contains

$$a_5 - 2a_1 - 2a_3 + 2a_2 = 0.$$

This means that a generic symmetric f will not have symmetry adapted representation of rank 3 as in the first case. The other three cases can be similarly investigated. For instance, in the second case we get the following relation on the coefficients:

$$\begin{aligned} &- 10a_1a_2^2 - \frac{5}{4}a_2^3 + 10a_1a_2a_3 - \frac{5}{2}a_2^2a_3 - 4a_1a_3^2 + \frac{5}{2}a_2^2a_4 - 3a_2a_3a_4 + \frac{3}{4}a_2a_4^2 - \frac{1}{2}a_3a_4^2 + 12a_1a_2a_5 \\ &+ \frac{1}{4}a_2^2a_5 - 6a_1a_3a_5 + 3a_2a_3a_5 - 2a_2a_4a_5 + a_3a_4a_5 - \frac{1}{4}a_4^2a_5 - 3a_1a_5^2 + \frac{5}{4}a_2a_5^2 - \frac{1}{2}a_3a_5^2 + \frac{1}{2}a_4a_5^2 \\ &- \frac{1}{4}a_5^3 - 2a_1a_2a_6 + a_2^2a_6 + 4a_1a_3a_6 + a_2a_4a_6 - a_2a_5a_6 - a_1a_6^2 - 3a_1a_2a_7 - a_2^2a_7 + 2a_1a_3a_7 + a_1a_5a_7 = 0. \end{aligned}$$

The fourth case yields one linear and six cubic relations in $a_1, \ldots, a_7$. In the third case, a lengthy computation in Macaulay 2 [36] gives a single polynomial of degree 14 with 6672 terms. Thus we see that SOS representations with three or fewer squares will only appear in very special cases of symmetric ternary sextics. □

This theorem establishes that we should expect to get a rank four SOS representation of symmetric ternary sextics. However, it is important to understand what one would get if an SDP were run on $K_f^{S_3}$. This question is related to the geometry of the boundary of $K_f^{S_3}$, and in order to shed some light on this geometry we present some experimental results.

Figure 3.6 is obtained as follows: After generating 100 random symmetric ternary sextics, we determined that only 12 of these were SOS according to our numerical SDP returning an optimal solution. For each of these 12 symmetric ternary sextics, we re-ran the SDP for 1000 distinct, randomly generated linear objective functions. Then we computed the rank of the output matrix by SVD with a cutoff tolerance of 10^{-7}. Each histogram shows the rank of the optimal matrix. This and other similar experiments we have conducted show that choosing a random linear functional to minimize resulted most commonly in a solution matrix of rank 6. However, for some polynomials other ranks were not unusual. For example, for several polynomials, over 100 of the 1000 objective functions picked out an optimal solution whose rank was judged to be 4.

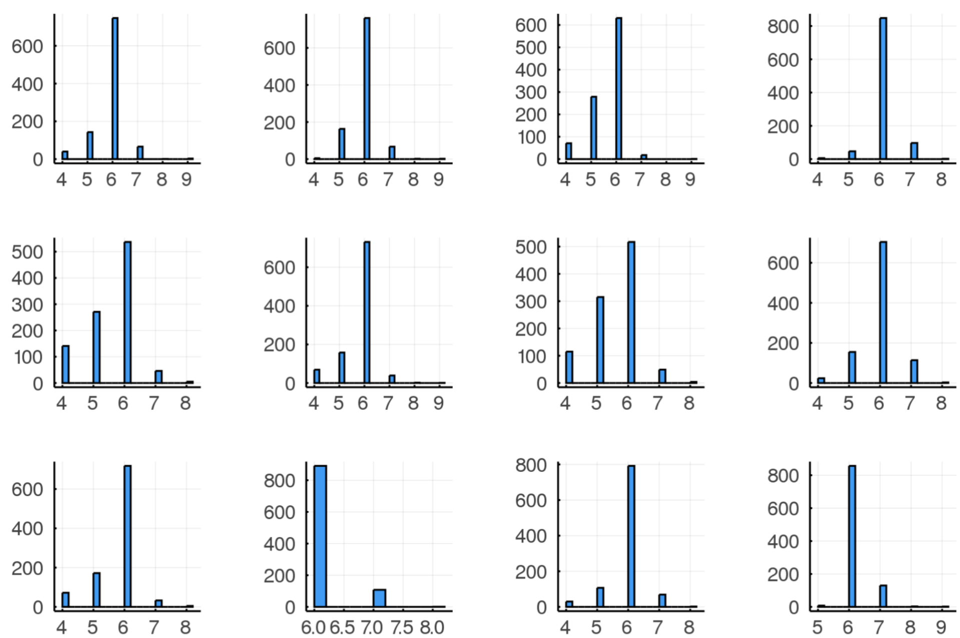


Figure 3.2: Distribution of ranks for SOS decomposition of symmetric ternary sextics

Chapter 4

An SOS counterexample to an inequality of symmetric functions

It is known that differences of symmetric functions corresponding to various bases are nonnegative on the nonnegative orthant exactly when the partitions defining them are comparable in dominance order. The only exception is the case of homogeneous symmetric functions where it is only known that dominance of the partitions implies nonnegativity of the corresponding difference of symmetric functions. It was conjectured by Cuttler, Greene, and Skandera in 2011 that the converse also holds, as in the cases of the monomial, elementary, power-sum, and Schur bases [24]. In this chapter we provide a counterexample, showing that homogeneous symmetric functions break the pattern. We use semidefinite programming to find an explicit sums of squares decomposition of the polynomial $H_{44} - H_{521}$ as a sum of 41 squares. This rational certificate of nonnegativity disproves the conjecture, since a polynomial which is a sum of squares cannot be negative, and since the partitions 44 and 521 are incomparable in dominance order.

4.1 Preliminaries

Symmetric Polynomials

Symmetric polynomials are of vital importance in representation theory and combinatorics. With respect to sums of squares, we saw in Chapter 3 that the structure of this invariant subspace can be exploited and it is for this reason that we are interested in studying symmetric polynomials. The space of symmetric polynomials has various commonly used bases and we begin this chapter by defining the symmetric polynomials which make up these bases, namely the monomial, elementary, power-sum, homogeneous and Schur symmetric polynomials. For more details the reader is encouraged to reference Chapter 7 of [74].

We begin with the monomial symmetric polynomials. Let $\lambda = (\lambda_1, \ldots, \lambda_k)$ be a partition of n. Then the monomial symmetric function with respect to λ is

$$m_\lambda(x) = \sum_{\alpha=\sigma(\lambda_1,\ldots,\lambda_k)} x^\alpha$$

where we sum over all distinct permutations of $(\lambda_1, \ldots, \lambda_k)$. For example,

$$\begin{aligned} m_{(1)}(x_1, \ldots, x_n) &= \sum_i x_i \\ m_{(1,1)}(x_1, x_2, x_3) &= x_1x_2 + x_1x_3 + x_2x_3 \\ m_{(4,2,2)}(x_1, \ldots, x_n) &= \sum_{i,j,k:i<k,j<k} x_i^4 x_j^2 x_k^2. \end{aligned}$$

Next we define the elementary symmetric functions as

$$e_\lambda = \prod_{i=1}^{\lambda_k} e_{\lambda_i}$$

where e_{λ_i} is the sum of all products of λ_i distinct variables. For example

$$\begin{aligned} e_{(3,1,1)}(x_1, x_2, x_3) &= (x_1x_2x_3)(x_1 + x_2 + x_3)^2 \\ e_{(2,2)}(x_1, x_2, x_3) &= (x_1x_2 + x_1x_3 + x_2x_3)^2. \end{aligned}$$

The (complete) homogeneous symmetric polynomials are similarly defined as

$$h_\lambda = \prod_{i=1}^{\lambda_k} h_{\lambda_i}$$

where h_{λ_i} is the sum of all monomials of λ_i variables. For instance,

$$\begin{aligned} h_{(3,1,1)}(x_1, x_2) &= (x_1^3 + x_1^2x_2 + x_1x_2^2 + x_2^3)(x_1 + x_2)^2 \\ h_{(2,1)}(x_1, x_2, x_3) &= (x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2)(x_1 + x_2 + x_3). \end{aligned}$$

Notice that in all the polynomials we have defined so far the degree is the positive integer that λ partitions. This is true as well for the power-sum symmetric polynomials which we define as follows,

$$p_\lambda = \prod_{i=1}^{\lambda_k} p_{\lambda_i}$$

where $p_{\lambda_i} = \sum_j x_j^{\lambda_i}$. Some examples are,

$$\begin{aligned} p_{(6)}(x_1, \ldots, x_n) &= x_1^6 + \ldots + x_n^6 \\ p_{(3,1)}(x_1, \ldots, x_n) &= (x_1^3 + \ldots + x_n^3)(x_1 + \ldots + x_3). \end{aligned}$$

Finally we end with the Schur polynomials, which are defined quite differently then the other symmetric polynomials thus far. There are various ways to define it, but we will use the follow definition. Let $\lambda = (\lambda_1, \ldots, \lambda_k)$. A semistandard Young tableaux T of shape λ is a Young diagram which has numbers $1, \ldots, n$ in the boxes such that these numbers weakly increase along each row and strictly increase down each column. Then

$$s_\lambda(x) = \sum_T x^T = \sum_T x_1^{t_1} \cdots x_n^{t_n}$$

where we sum over all semistandard Young tableaux T of shape λ and where each t_i counts how many times i appears in T. Take for example $s_{(2,2)}(x_1, x_2, x_3)$. The semistandard Young tableaux are

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 2 \\ \hline \end{array}\ , \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 3 \\ \hline \end{array}\ , \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 & 3 \\ \hline \end{array}\ , \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 & 3 \\ \hline \end{array}\ , \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 3 \\ \hline \end{array}\ , \text{ and } \begin{array}{|c|c|}\hline 2 & 2 \\ \hline 3 & 3 \\ \hline \end{array}.$$

The first tableaux has two 1's and two 2's and the corresponding monomial is $x_1^2 x_2^2$. Thus

$$s_{(2,2)}(x_1, x_2, x_3) = x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 + x_1^2 x_2 x_3 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2$$

and is indeed symmetric.

4.2 Symmetric Function Inequalities

In the article *Inequalities for Symmetric Means* [24], by Cuttler, Greene, and Skandera, Muirhead-type inequalities are classified for the different common bases of symmetric functions. We briefly provide some definitions in order to state our main Theorem 4.2.2. As done above, let m_λ, e_λ, p_λ, h_λ, and s_λ denote the monomial, elementary, power-sum, homogeneous, and Schur polynomials, respectively, associated to a partition λ. Given a symmetric polynomial $g(x)$, the *term-normalized symmetric polynomial* is

$$G(x) := \frac{g(x)}{g(\mathbf{1})}$$

where $g(\mathbf{1})$ is the symmetric polynomial evaluated on the all ones vector. By $G_\lambda \geq G_\mu$, we mean that $G_\lambda(x_1, \ldots, x_n) \geq G_\mu(x_1, \ldots, x_n)$ on the nonnegative orthant. That is, the inequality holds for any number of variables n, but only for $x_i \geq 0$, $i = 1, \ldots, n$. We denote the term-normalized symmetric polynomials for monomial, elementary, power-sum, homogeneous, and Schur polynomials by M_λ, E_λ, P_λ, H_λ, and S_λ, respectively. The following theorem is a summary of known results on Muirhead-type inequalities (special cases of which go back to Maclaurin, Muirhead, Newton, and Schur), which are proven in [24, 43, 54, 73]. In particular, the arithmetic-geometric mean inequality is a special case.

Theorem 4.2.1. *Let λ and μ be partitions such that $|\lambda| = |\mu|$. Then*

$$\begin{aligned} M_\lambda \le M_\mu &\iff \mu \succeq \lambda \\ E_\lambda \le E_\mu &\iff \lambda \succeq \mu \\ P_\lambda \le P_\mu &\iff \mu \succeq \lambda \\ S_\lambda \le S_\mu &\iff \mu \succeq \lambda \end{aligned}$$

whereas $\mu \succeq \lambda$ implies that $H_\lambda \le H_\mu$, i.e.,

$$H_\lambda \le H_\mu \impliedby \mu \succeq \lambda.$$

The converse for the homogeneous symmetric functions statement was conjectured in [24]. The authors also reported that for $d = |\lambda| = |\mu| = 1, 2, \ldots, 7$ their conjecture had been proven. For $d = 8$ and higher, the question was unresolved.

Theorem 4.2.2. *A degree-minimal counterexample exhibiting a polynomial $H_\mu - H_\lambda \ge 0$ with λ, μ incomparable in dominance order is provided by $H_{44} - H_{521}$.*

We certify the nonnegativity of this polynomial on $\mathbb{R}^3_{\ge 0}$ by writing a related polynomial explicitly as a sum of 41 squares with rational coefficients. Specifically, the polynomial we exhibit an SOS polynomials is $\big(H_{44} - H_{521}\big)(x_1^2, x_2^2, x_3^2) =$

$$\begin{aligned} \frac{1}{9450}\Big(& 17\,x_1^{16} + 9\,x_1^{14}x_2^2 + x_1^{12}x_2^4 + 18\,x_1^{10}x_2^6 + 60\,x_1^8x_2^8 + 18\,x_1^6x_2^{10} + x_1^4x_2^{12} + 9\,x_1^2x_2^{14} \\ & + 17\,x_2^{16} + 9\,x_1^{14}x_3^2 - 32\,x_1^{12}x_2^2x_3^2 - 6\,x_1^{10}x_2^4x_3^2 + 11\,x_1^8x_2^6x_3^2 + 11\,x_1^6x_2^8x_3^2 \\ & - 48\,x_1^4x_2^{10}x_3^2 - 32\,x_1^2x_2^{12}x_3^2 + 9\,x_2^{14}x_3^2 + x_1^{12}x_3^4 - 48\,x_1^{10}x_2^2x_3^4 - 22\,x_1^8x_2^4x_3^4 \\ & - 5\,x_1^6x_2^6x_3^4 - 22\,x_1^4x_2^8x_3^4 - 48\,x_1^2x_2^{10}x_3^4 + x_2^{12}x_3^4 + 18\,x_1^{10}x_3^6 + 11\,x_1^8x_2^2x_3^6 \\ & - 5\,x_1^6x_2^4x_3^6 - 5\,x_1^4x_2^6x_3^6 + 11\,x_1^2x_2^8x_3^6 + 18\,x_2^{10}x_3^6 + 60\,x_1^8x_3^8 + 11\,x_1^6x_2^2x_3^8 \\ & - 22\,x_1^4x_2^4x_3^8 + 11\,x_1^2x_2^6x_3^8 + 60\,x_2^8x_3^8 + 18\,x_1^6x_3^{10} - 48\,x_1^4x_2^2x_3^{10} - 48\,x_1^2x_2^4x_3^{10} \\ & + 18\,x_2^6x_3^{10} + x_1^4x_3^{12} - 32\,x_1^2x_2^2x_3^{12} + x_2^4x_3^{12} + 9\,x_1^2x_3^{14} + 9\,x_2^2x_3^{14} + 17\,x_3^{16}\Big). \end{aligned}$$

Remark 6. *Of course, there are other ways to show that a polynomial is nonnegative. However, Theorem 4.2.2 states something stronger. Not only is it nonnegative, but also a sum of squares. An ongoing research topic belonging to the general context of Hilbert's 17th Problem is to understand the difference between sums of squares and nonnegativity, see [7, 6, 8, 10, 16, 22] to name only a few. As an interesting example, the degrees of irreducible components of the boundary of the SOS cone are Gromov-Witten numbers (see [11, 51]).*

Remark 7. *We refer the reader to [74, Ch. 7] for details on symmetric functions, but we give some brief comments on the homogeneous symmetric functions, since they are the main focus in this chapter. The homogeneous and elementary symmetric functions h_λ and e_λ are in many ways dual. By the fundamental theorem of symmetric functions, the e_d are* *algebraically independent and generate the algebra of symmetric functions* Λ. *Therefore, an algebra endomorphism* $\omega : \Lambda \to \Lambda$ *is uniquely defined by specifying the images* $\omega(e_d)$. *Defining* $\omega(e_d) = h_d$ *gives an involution* $\omega^2 = id$ *of* Λ *sending* $\omega(e_\lambda) = h_\lambda$ *and* $\omega(h_\lambda) = e_\lambda$. *While the transition matrices between the bases* (m_λ) *and* (e_λ) *are matrices with entries in* $\{0, 1\}$, *those for* (m_λ) *and* (h_λ) *are matrices with entries in* $\mathbb{N}$. *This also reflects the combinatorial interpretations related to placing balls in boxes without or with repetition. See [74, Sections 7.5, 7.6] for more details.*

We leave the proof of Theorem 4.2.2 to the end of the chapter in Section 4.5. Instead we begin in Section 4.3 by describing our approach of finding a counterexample and extracting an exact rational SOS certificate. In Section 4.4, we provide many additional counterexamples that have been certified via numerical means. While our proof of Theorem 4.2.2 is provided by an explicit list of polynomials which square and sum to $(H_{44} - H_{521})\,(x_1^2, x_2^2, x_3^2)$, in fact we found many numerical counterexamples in degrees 8, 9, and 10. We expect that most of these numerical counterexamples, could, with some effort, be converted into provably-correct sums of squares using exact rational arithmetic, as we did for $H_{44} - H_{521}$. Section 4.4 displays a poset showing how the dominance partial order on partitions would need to be modified in order to correctly reflect nonnegativity, at least as suggested by our numerical counterexamples. It would be very interesting to see if some modification of dominance order could achieve the correct nonnegativity relationships.

4.3 Methods

In this section we outline the steps taken to find the counterexample given in Theorem 4.2.2. In order to find such a counterexample, we must search over pairs of partitions (μ, λ) that are incomparable in dominance order and such that $(H_\lambda - H_\mu) \geq 0$, specifically searching over partitions of 8 and polynomials in $\mathbb{R}[x_1, x_2, x_3]$. To this end, we first recognize that we can certify nonnegativity on the nonnegative orthant by replacing the variables with squares. That is, to certify that $(H_\lambda - H_\mu)(x_1, x_2, x_3) \geq 0$ for all $x \in \mathbb{R}^3_{\geq 0}$, we instead search for polynomials $(H_\lambda - H_\mu)(x_1^2, x_2^2, x_3^2)$ which are SOS. See Lemma 4.5.1 in Section 4.5 for a proof.

Now we may utilize the well-known machinery of sums of squares discussed in Chapter 3, the study of which has a long history. See, for example, [9] for more on the theory. We will specifically use Proposition 4.3.1 below, a well-known and extremely important fact in sums of squares. Indeed it is a slight specialization of Theorem 3.1.2. Let $\mathcal{S}^N_+$ denote the cone of $N \times N$ symmetric positive semidefinite matrices in the space of $N \times N$ symmetric matrices $\mathcal{S}^N$.

Proposition 4.3.1. *Let* h *be a homogeneous polynomial of degree* $2d$ *in* n *variables,* $h \in \mathbb{R}[x_1, \ldots, x_n]_{2d}$. *Let* $[x]$ *be a vector containing all* $N = \binom{n+d-1}{d}$ *monomials of degree* d. *Then* h *is a sum of squares exactly when there exists some* $Q \in \mathcal{S}^N_+$ *such that*

$$h(x_1, \ldots, x_n) = [x]^T Q [x].$$

Proposition 4.3.1 tells us that writing a polynomial as a sum of squares is equivalent to solving a semidefinite program (SDP). That is, we must find a positive semidefinite matrix Q that also satisfies the linear constraints defined by equating the coefficients of $h(x_1, \ldots, x_n)$ and $[x]^T Q[x]$. Unfortunately, SDP solvers return numerical solutions, i.e. a matrix with floating point entries. In particular, the matrix will (almost) never exactly reproduce the desired polynomial, a problem when searching for an exact counterexample. We illustrate this with the following running example.

Example 4.3.1. *An attempt to reproduce the polynomial $(H_{21} - H_{111})(x_1^2, x_2^2, x_3^2)$, whose nonnegativity follows from Theorem 4.2.1, produced the following polynomial with floating point coefficients*

$$\begin{aligned}\frac{1}{54}\,x_1^6+\frac{1}{54}\,x_2^6+\left(7.888609052210118\times 10^{-31}\right)\,x_1^3x_2^2x_3+\left(7.888609052210118\times 10^{-31}\right)\,x_1^2x_2^3x_3 \\ +\left(3.944304526105059\times 10^{-31}\right)\,x_1^3x_2x_3^2-0.05555555555555555\,x_1^2x_2^2x_3^2 \\ +\left(3.944304526105059\times 10^{-31}\right)\,x_1x_2^3x_3^2+\left(3.944304526105059\times 10^{-31}\right)\,x_1^2x_2x_3^3 \\ +\left(3.944304526105059\times 10^{-31}\right)\,x_1x_2^2x_3^3+\frac{1}{54}\,x_3^6\end{aligned}$$

even though the desired polynomial is

$$\frac{1}{54}\,x_1^6+\frac{1}{54}\,x_2^6-\frac{1}{18}\,x_1^2x_2^2x_3^2+\frac{1}{54}\,x_3^6.$$

In fact, Hurwitz proved nonnegativity of this polynomial via sums of squares [43].

Therefore, in order to find an *exact* sum of squares certificate of nonnegativity, we must make adjustments. To satisfy Proposition 4.3.1 we must replace the entries of the matrix itself, while staying in the PSD cone, and continuing to satisfy the requirements of $m^T Qm = h$ exactly. One approach to this problem is to use continued fractions to find the best rational approximation (with some user-specified bound B on the size of the denominator) to the entries of the matrix. Geometrically, the SDP may return a matrix on or near the boundary of the PSD cone. By rounding the floating point entries to rational numbers, we risk moving outside the cone, resulting in a matrix which is not positive semidefinite. Therefore, many times this rational rounding procedure will fail.

Remark 8. *In general, rational certificates for polynomials with rational coefficients do not always exist. This was shown by Scheiderer in [70] where he provided explicit minimal examples of degree 4 polynomials in 3 variables. Since the polynomial in Theorem 4.2.2 is of degree 16, there is no a priori reason to believe it must have a rational sum of squares representation.*

Several approaches to this rational rounding problem have been developed. The package `SOS` has a rational rounding procedure built-in, but for our polynomial the package returned an error stating that the rational rounding had failed. The software `RealCertify` [49], based on [48], uses a hybrid numeric-symbolic algorithm for finding rational approximations for polynomials lying in the interior of the SOS cone. In correspondence with Mohab Safey El Din, we learned that `RealCertify` failed to terminate for our problem. However, in [48] they also describe and compare complexity of several different algorithms, including geometric critical point methods. Safey El Din reported that the geometric critical point methods were successful on our problem, providing a second confirmation of the nonnegativity of our polynomial, although not of its SOS-ness.

Remark 9. *Another approach would be to search directly for the matrix using exact arithmetic as in [40] with the package* `SPECTRA` *for* `Maple`*. However, for a problem of our size, this approach is not promising. Indeed we let* `SPECTRA` *run for several days, and it did not terminate. Ours is a feasibility problem, but when optimizing a linear function for a rational SDP, the entries of the optimal solution matrix will be algebraic numbers. In [56] the algebraic degree of an SDP is introduced. For generic inputs, this degree depends only on the rank r of the solution matrix, the size n of the symmetric matrices, and the dimension m of the affine subspace. In [35] the authors give an exact formula for the algebraic degree. If $n = 45, m = 129, r = 41$ (which corresponds to our problem), their formula yields the following 74 digit number:*

27986928303724394857777762195272647267703276932951767224059513477726952420.

Example 4.3.2. ***(continued)*** *Returning to the example above, the output of the SDP for $(H_{21} - H_{111})(x_1^2, x_2^2, x_3^2)$ was the following matrix, for which we print only the first three of ten columns:*

$$\begin{pmatrix}
\frac{1}{54} & 0 & 0 & \\
0 & 0.018518456551295637 & 1.4472934340259067 \times 10^{-17} & \\
0 & 1.4472934340259067 \times 10^{-17} & 0.018518456551295637 & \\
-0.009259228275647818 & -3.1517672055168234 \times 10^{-14} & 3.1499353284307314 \times 10^{-14} & \\
-1.4472934340259067 \times 10^{-17} & 7.150822547960123 \times 10^{-18} & 7.150822547960123 \times 10^{-18} & \\
-0.009259228275647818 & 3.1499353284307314 \times 10^{-14} & -3.1517672055168234 \times 10^{-14} & \cdots \\
3.1517672055168234 \times 10^{-14} & -0.009259228275647818 & -3.1506504106855275 \times 10^{-14} & \\
-3.1506504106855275 \times 10^{-14} & 3.1499353284307314 \times 10^{-14} & -0.009259259448737624 & \\
-3.1506504106855275 \times 10^{-14} & -0.009259259448737624 & 3.1499353284307314 \times 10^{-14} & \\
3.1517672055168234 \times 10^{-14} & -3.1506504106855275 \times 10^{-14} & -0.009259228275647818 &
\end{pmatrix}$$

Using continued fractions with denominator bound $B = 150$ we obtain the following (preferable) matrix:

$$Q = \frac{1}{108}\begin{pmatrix}
2 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\
-1 & 0 & 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & -1 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & -1 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & -1 \\
0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 2 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 2
\end{pmatrix}$$

With $[x]^T = (x_1^3,\, x_1^2x_2,\, x_1^2x_3,\, x_1x_2^2,\, x_1x_2x_3,\, x_1x_3^2,\, x_2^3,\, x_2^2x_3,\, x_2x_3^2,\, x_3^3)$*, we can calculate* $[x]^TQ[x]$*, obtaining*

$$\frac{1}{54}\Big(x_1^6 + x_2^6 - 3\,x_1^2x_2^2x_3^2 + x_3^6\Big)$$

which is exactly $(H_{21} - H_{111})(x_1^2, x_2^2, x_3^2)$*, as desired. Since* Q *is positive semidefinite, by carrying out* LDL^T *factorization we obtain the following sum of squares representation of* $(H_{21} - H_{111})(x_1^2, x_2^2, x_3^2)$*:*

$$\frac{1}{216}\left(2\,x_1^3 - x_1x_2^2 - x_1x_3^2\right)^2 + \frac{1}{216}\left(2\,x_1^2x_2 - x_2^3 - x_2x_3^2\right)^2 + \frac{1}{72}\left(x_1x_2^2 - x_1x_3^2\right)^2 + \frac{1}{72}\left(x_2^3 - x_2x_3^2\right)^2 + \frac{1}{216}\left(2\,x_1^2x_3 - x_2^2x_3 - x_3^3\right)^2 + \frac{1}{72}\left(x_2^2x_3 - x_3^3\right)^2.$$

Of course, the expression of a polynomial as a sum of squares is not unique. For example, this same polynomial appears as a sum of squares of binomials in [43] and also as a sum of squares of binomials and one trinomial in [64].

$$\begin{aligned} x_1^6 + x_2^6 + x_3^6 - 3\,x_1^2x_2^2x_3^2 = &\ (x_1^3 - x_1x_2^2)^2 + (x_3^3 - x_2^2x_3)^2 \\ &+ \frac{1}{2}(x_1^2x_2 - x_2^3)^2 + \frac{1}{2}(x_2x_3^2 - x_2^3)^2 + \frac{3}{2}(x_1^2x_2 - x_2x_3^2)^2 \\ = &\ (x_1^2x_2 - x_2^3)^2 + (x_1^2x_3 - x_3^3)^2 + \frac{7}{4}(x_1x_2^2 - x_1x_3^2)^2 \\ &+ \frac{1}{4}(x_1x_2^2 + x_1x_3^2 - 2x_1^3)^2 \end{aligned}$$

As noted in the discussion above, for $H_{44} - H_{521}$, existing tools do not return a numerical matrix which can be successfully rounded. Our solution to this problem depended crucially on two things, using the real zeros of the polynomial, and using symmetry. In particular, we used the symmetry reduction techniques developed by Gatermann and Parrilo in [32] which we briefly describe now.

To start, we wish to take advantage of the fact that $H_{44} - H_{521}$ is a symmetric polynomial. That is to say, it is invariant under the action of the symmetric group. There is a great deal of literature on the subject of symmetric polynomials and sums of squares, including [10, 17, 22, 32, 34], to name only a few. In particular, we specialize Proposition 3.3.1 to our setting as follows:

Theorem 4.3.1. *Given an orthogonal linear representation of the symmetric group* S_n*,* $\sigma : S_n \to Aut(\mathcal{S}^N)$*, consider a semidefinite program whose objective and feasible matrices are invariant under the group action. Then the optimal value of the SDP is equal to the optimal value of the same SDP restricted to its fixed point subspace,* $\{X \in \mathcal{S}^N : X = \sigma(g)X,\ \forall g \in S_n\}$*.*

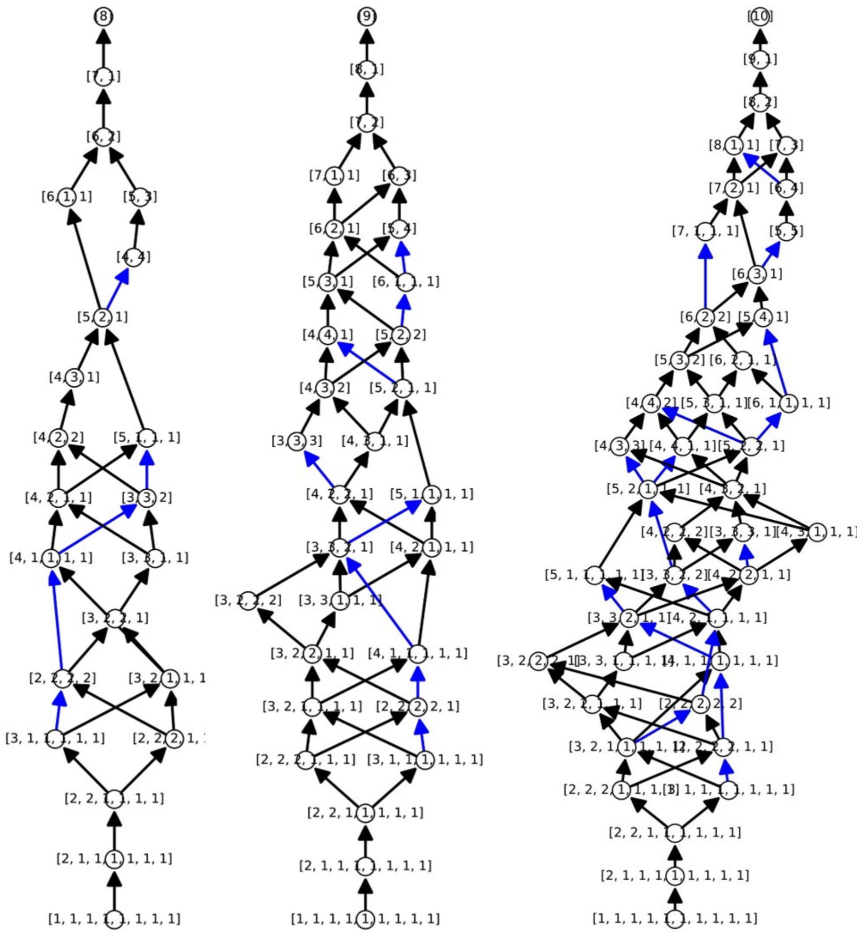
In our case, the S_3 action on the space of polynomials in three variables induces an action on the symmetric matrix of our SDP, where group elements act on the symmetric matrix by conjugation. More specifically, for each $g \in S_3$, let $\rho(g)$ be the associated matrix that permutes the monomials of degree 8 in 3 variables. Then the induced action sends a symmetric 45×45 matrix $X \to \rho(g)^T X \rho(g)$. Note that $\rho(g)$ is an orthogonal matrix. Then the fixed-point subspace for our particular SDP is

$$\mathcal{F} = \{X : X\rho(g) = \rho(g)X, \ \forall g \in S_3\}.$$

Theorem 4.3.1 guarantees that if a solution exists to our SDP, a solution also exists if we restrict to this fixed-point subspace. Thus we force our matrix Q to commute with the elements of our group, obtaining better constraints on our semidefinite program. Indeed, we further simplify our SDP with additional linear constrains via Lemma 4.5.2, but leave the details in Section 4.5. The resulting simplified SDP returns a positive semidefinite matrix to which we can successfully apply the rational rounding described above. Upon factoring the matrix using exact, rational arithmetic we obtain an explicit sum of squares representation for $H_{44} - H_{521}$ evaluated at (x_1^2, x_2^2, x_3^2).

4.4 Poset of SOS Certifications

Theorem 4.2.2 offers a counterexample in the pair of partitions (521, 44). We provide an exact, rational certificate of nonnegativity by applying rational rounding to the numerical solution returned by the SDP solver, and then factoring the resulting matrix to obtain a provably-correct expression of $(H_{44} - H_{521})(x_1^2, x_2^2, x_3^2)$ as a sum of squares. However, our search over other pairs of partitions returned many other counterexamples, in degrees 8, 9, and 10. We left these other counterexamples in floating point, though we found an exact, rational SOS certificate in the case of $H_{44} - H_{521}$. Below we provide a partially ordered set of all differences of term-normalized homogeneous symmetric polynomials of degree $8, 9, 10$ in three variables that are SOS. That is, for each arrow going from λ to μ, $\big(H_\mu - H_\lambda\big)(x_1^2, x_2^2, x_3^2)$ is an SOS polynomial, as certified numerically.



The black arrows coincide with the dominance order, and therefore certify that these polynomials are not only nonnegative as stated in Theorem 4.2.1, but in fact SOS. Additionally, the blue arrows (numerically) certify SOS-ness for incomparable pairs of partitions, i.e. each blue arrow is a counterexample to the conjecture.

Remark 10. *A natural question arises: Is there another partial order on the set of partitions which matches the poset of nonnegativity relations amongst the H_λ above? A first idea would be to modify the dominance (also called majorization) order slightly, to incorporate the correct relationships among the H_λ. Since dominance order is related to the cumulative sums produced by the vectors*

$$(1, 0, 0, 0, \dots),\ (1, 1, 0, 0, \dots),\ (1, 1, 1, 0, \dots),\ \dots$$

perhaps a modification of the components of these vectors might be a step in the correct direction. Such a modification might also be informed by the relationship of the homogeneous symmetric functions to the other usual bases.

Remark 11. *If we fix the degree and let the number of variables n go to infinity, the homogeneous symmetric functions behave like the elementary symmetric functions. This is because the number of square free terms will dominate the number of monomials with squares for very large n. Thus, as n increases, we will likely see fewer counterexamples. An interesting question is whether the conjecture is true asymptotically i.e. does $H_\lambda \leq H_\mu$ imply $\mu \succeq \lambda$ for large enough n?*

4.5 Proof of Theorem 4.2.2

As outlined in Section 4.3, our proof relies on Proposition 4.3.1 and on the following well-known fact.

Lemma 4.5.1. *Consider a polynomial $H(x_1, \ldots, x_n)$. Define another polynomial*

$$h(x_1, \ldots, x_n) = H(x_1^2, \ldots, x_n^2).$$

If h can be written as a sum of squares, then H is nonnegative on the nonnegative orthant.

Proof. Suppose

$$h = \sum d_i q_i^2$$

for positive $d_i > 0$ and polynomials $q_i(x_1, \ldots, x_n)$. Then h is nonnegative on all of $\mathbb{R}^n$. By way of contradiction, assume there is some point $(a_1, \ldots, a_n) \in \mathbb{R}^n_{\geq 0}$ where $H(a_1, \ldots, a_n) < 0$. This implies that there exist real numbers $\sqrt{a_1}, \ldots, \sqrt{a_n}$. But then $h(\sqrt{a_1}, \ldots, \sqrt{a_n}) = H(a_1, \ldots, a_n) < 0$, contradicting the nonnegativity of h. □

We also use one more lemma to impose additional constrains and help further reduce the size of our SDP, thus ensuring we can apply rational rounding.

Lemma 4.5.2. *If x^* is a (nonzero) real root of the polynomial $h = [x]^T Q[x]$, which we write as a sum of squares using the factorization of Q, then the monomial vector $[x]$ evaluated at x^* must be in the nullspace of Q.*

Proof. This follows from $0 = h(x^*) = [x]|^T_{(x^*)} Q[x]|_{(x^*)}$ and the fact that Q is positive semidefinite. □

Proof of the main Theorem 4.2.2. We ran a semidefinite program to find a symmetric positive semidefinite matrix whose factorization could produce a sums of squares representation for $H_{44} - H_{521}$ evaluated at x_1^2, x_2^2, x_3^2. By Lemma 4.5.1 this certifies the non-negativity of $H_{44} - H_{521}$ on the non-negative octant. By Proposition 4.3.1, the existence of a matrix which *exactly* reproduces our polynomial is equivalent to the existence of a sums of squares representation. Also by Proposition 4.3.1, we impose linear constraints on the entries of our unknown matrix Q to require that $[x]^T Q[x]$ exactly matches the coefficients of our desired polynomial. By Theorem 4.3.1, we impose linear constraints to force our unknown matrix

Q to commute with the action of the symmetric group on the space of symmetric 45×45 matrices. Finally, by Lemma 4.5.2, we incorporate 4 linearly independent vectors $[x]|_{(x^*)}$, coming from real zeros of our polynomial, to obtain more linear conditions on our SDP. The output is a numerical matrix sufficiently located in the PSD cone such that continued fractions rational approximation yields a matrix with exact, rational entries. This matrix remains positive semidefinite and produces our desired polynomial via $[x]^T Q[x]$. We pause to emphasize that without using symmetry, and without using real zeros, the numerical output of the SDP was not accurate enough to recover an exact solution. Only by using symmetry and real zeros were our methods successful. The entries of this 45×45 matrix include rational numbers with quite large denominators. Therefore we do not print the matrix here. Rather, we refer the reader to the co-author's website [39] for the explicit matrix.

To give the reader a feel for the matrix, the first row is:

$$\Big(\frac{17}{9450}, 0, 0, -\frac{4}{1433}, 0, -\frac{4}{1433}, 0, 0, 0, 0, \frac{5}{6699}, 0, \frac{1}{484}, 0, \frac{5}{6699}, 0, 0, 0, 0, 0, 0, \frac{1}{5516},$$
$$0, \frac{4}{6655}, 0, \frac{4}{6655}, 0, \frac{1}{5516}, 0, 0, 0, 0, 0, 0, 0, 0, \frac{157747519610069845105323375343}{780042543477736474894875053177040\,0},$$
$$0, -\frac{1}{2243}, 0, -\frac{490859542561433043273727488474533399}{1004640661046807224753241364163337033424}, 0, -\frac{1}{2243}, 0,$$
$$\frac{157747519610069845105323375343}{7800425434777364748948750531770400}\Big)$$

In order to produce an explicit sum of squares representation for $(H_{44} - H_{521})(x_1^2, x_2^2, x_3^2)$ it remains to factor this matrix. We factored the matrix using exact, rational arithmetic, obtaining from this factorization an explicit sum of squares representation. The first polynomial which is to be squared and linearly combined with other squares is displayed below. The others can be found in a `.txt` file at [39], along with the required (positive) coefficients. For those averse to squaring and summing by hand, we also provide open-source computer code that squares and sums them, verifying Theorem 4.2.2. We have also made this code available at the "mathrepo" hosted by Max Planck Institute for Mathematics in the Sciences at https://mathrepo.mis.mpg.de/soscounterexample/index.html.

$$\frac{460541924076360285607591683723404557053617981848633701831928139437729 5}{4799310937074435809326377636641953348906668422268214343278250076694421 6}x_1^7x_2$$

$$- \frac{130231403780304631325531179538254850114140223622887929762087169125707433659841 7}{736915035538518817374314455206846462046850100818942617698716932502122272548972 8}x_1^5x_2^3$$

$$- \frac{130231403780304631325531179538254850114140223622887929762087169125707433659841 7}{736915035538518817374314455206846462046850100818942617698716932502122272548972 8}x_1^3x_2^5$$

$$+ \frac{460541924076360285607591683723404557053617981848633701831928139437729 5}{4799310937074435809326377636641953348906668422268214343278250076694421 6}x_1x_2^7$$

$$- \frac{509842531246889292559585469138864624934685823550090365451356517319046 25}{191972437482977432373055105465678133956266736890728573731130003067776864}x_1^5x_2x_3^2$$

$$+ x_1^3x_2^3x_3^2$$

$$- \frac{509842531246889292559585469138864624934685823550090365451356517319046 25}{191972437482977432373055105465678133956266736890728573731130003067776864}x_1x_2^5x_3^2$$

$$- \frac{145016490586868172805265027238496682192202685262016624189299558387921997579 43}{741432108598005779762565967227432775403214265884634082892895675248306448603 52}x_1^3x_2x_3^4$$

$$- \frac{145016490586868172805265027238496682192202685262016624189299558387921997579 43}{741432108598005779762565967227432775403214265884634082892895675248306448603 52}x_1x_2^3x_3^4$$

$$+ \frac{543703581181487689919586312916849948429866258636425931996915135241636712338485 5}{648254945325290772158967247314772747081838448064163584006840051560460686632924 51}x_1x_2x_3^6$$

□

Bibliography

[1] Xavier Allamigeon et al. "Log-barrier interior point methods are not strongly polynomial". In: *SIAM J. Appl. Algebra Geom.* 2.1 (2018), pp. 140–178.

[2] Christine Bachoc et al. "Invariant semidefinite programs". In: *Handbook on semidefinite, conic and polynomial optimization*. Vol. 166. Internat. Ser. Oper. Res. Management Sci. Springer, New York, 2012, pp. 219–269.

[3] D. A. Bayer and J. C. Lagarias. "The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories". In: *Trans. Amer. Math. Soc.* 314.2 (1989), pp. 499–526.

[4] D. A. Bayer and J. C. Lagarias. "The nonlinear geometry of linear programming. II. Legendre transform coordinates and central trajectories". In: *Trans. Amer. Math. Soc.* 314.2 (1989), pp. 527–581.

[5] Matthias Beck and Sinai Robins. *Computing the continuous discretely*. Second. Undergraduate Texts in Mathematics. Integer-point enumeration in polyhedra, With illustrations by David Austin. Springer, New York, 2015, pp. xx+285. ISBN: 978-1-4939-2968-9; 978-1-4939-2969-6.

[6] Grigoriy Blekherman. "Nonnegative polynomials and sums of squares". In: *Semidefinite optimization and convex algebraic geometry*. Vol. 13. MOS-SIAM Ser. Optim. SIAM, Philadelphia, PA, 2013, pp. 159–202.

[7] Grigoriy Blekherman. "There are significantly more nonnegative polynomials than sums of squares". In: *Israel J. Math.* 153 (2006), pp. 355–380. ISSN: 0021-2172.

[8] Grigoriy Blekherman, Sadik Iliman, and Martina Juhnke-Kubitzke. "Dimensional differences between faces of the cones of nonnegative polynomials and sums of squares". In: *Combinatorial methods in topology and algebra*. Vol. 12. Springer INdAM Ser. Springer, Cham, 2015, pp. 69–77.

[9] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas, eds. *Semidefinite optimization and convex algebraic geometry*. Vol. 13. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013, pp. xx+476. ISBN: 978-1-611972-28-3.

[10] Grigoriy Blekherman and Cordian Riener. “Symmetric Non-Negative Forms and Sums of Squares”. In: *Discrete and Computational Geometry* (2012).

[11] Grigoriy Blekherman et al. “Algebraic boundaries of Hilbert’s SOS cones”. In: *Compos. Math.* 148.6 (2012), pp. 1717–1735. ISSN: 0010-437X.

[12] Grigoriy Blekherman et al. “Low-rank sum-of-squares representations on varieties of minimal degree”. In: *Int. Math. Res. Not. IMRN* 1 (2019), pp. 33–54. ISSN: 1073-7928.

[13] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004, pp. xiv+716. ISBN: 0-521-83378-7.

[14] Marcel K. de Carli Silva and Levent Tunçel. “Vertices of spectrahedra arising from the elliptope, the theta body, and their relatives”. In: *SIAM J. Optim.* 25.1 (2015), pp. 295–316. ISSN: 1052-6234.

[15] John W. S. Cassels, William J. Ellison, and Albrecht Pfister. “On sums of squares and on elliptic curves over function fields”. In: *J. Number Theory* 3 (1971), pp. 125–149. ISSN: 0022-314X.

[16] M. D. Choi, T. Y. Lam, and Bruce Reznick. “Even symmetric sextics”. In: *Math. Z.* 195.4 (1987), pp. 559–580. ISSN: 0025-5874.

[17] Man Duen Choi and Tsit Yuen Lam. “An old question of Hilbert”. In: *Conference on Quadratic Forms—1976 (Proc. Conf., Queen’s Univ., Kingston, Ont., 1976)*. 1977, 385–405. Queen’s Papers in Pure and Appl. Math., No. 46.

[18] Man-Duen Choi, Tsit Yuen Lam, and Bruce Reznick. “Sums of squares of real polynomials”. In: *K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992)*. Vol. 58. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 1995, pp. 103–126.

[19] Lynn Chua et al. “Gram spectrahedra”. In: *Ordered algebraic structures and related topics*. Vol. 697. Contemp. Math. Amer. Math. Soc., Providence, RI, 2017, pp. 81–105.

[20] Yairon Cid-Ruiz. *Equations and Multidegree for Inverse Symmetric Matrix Pairs*. To appear in Le Matematiche on Linear Spaces of Symmetric Matrices. arXiv preprint arxiv.org/abs/2011.04616. 2020. arXiv: 2011.04616.

[21] Arthur B. Coble. *Algebraic geometry and theta functions*. Vol. 10. American Mathematical Society Colloquium Publications. Reprint of the 1929 edition. American Mathematical Society, Providence, R.I., 1982, pp. vii+282. ISBN: 0-8218-1010-3.

[22] Logan Coe. *Sum of Squares Decomposition for Symmetric Polynomial Inequalities. San Francisco State University Master’s Thesis*. 2017.

[23] F. Albert Cotton. *Chemical Applications of Group Theory*. John Wiley & Sons, Inc., 1990. ISBN: 978-0-471-51094-9.

[24] Allison Cuttler, Curtis Greene, and Mark Skandera. “Inequalities for symmetric means”. In: *European J. Combin.* 32.6 (2011), pp. 745–761. ISSN: 0195-6698.

[25] Jesús A. De Loera, Bernd Sturmfels, and Cynthia Vinzant. "The central curve in linear programming". In: *Found. Comput. Math.* 12.4 (2012), pp. 509–540.

[26] Sebastian Debus and Cordian Riener. *Reflection groups and cones of sums of squares.* arXiv preprint arxiv.org/abs/2011.09997. 2020. arXiv: 2011.09997.

[27] Jean-Pierre Dedieu, Gregorio Malajovich, and Mike Shub. "On the curvature of the central path of linear programming theory". In: *Found. Comput. Math.* 5.2 (2005), pp. 145–171.

[28] Albert Fässler and Eduard Stiefel. *Group theoretical methods and their applications.* Translated from the German by Baoswan Dzung Wong. Birkhäuser Boston, Inc., Boston, MA, 1992, pp. xii+296. ISBN: 0-8176-3527-0.

[29] Anthony V. Fiacco and Garth P. McCormick. *Nonlinear programming: Sequential unconstrained minimization techniques.* John Wiley and Sons, Inc., New York-London-Sydney, 1968, pp. xiv+210.

[30] Anders Forsgren, Philip E. Gill, and Margaret H. Wright. "Interior Methods for Nonlinear Optimization". In: *SIAM Review* 44.4 (2002), pp. 525–597.

[31] William Fulton. *Young tableaux.* Vol. 35. London Mathematical Society Student Texts. With applications to representation theory and geometry. Cambridge University Press, Cambridge, 1997, pp. x+260. ISBN: 0-521-56144-2; 0-521-56724-6.

[32] Karin Gatermann and Pablo A. Parrilo. "Symmetry groups, semidefinite programs, and sums of squares". In: *J. Pure Appl. Algebra* 192.1-3 (2004), pp. 95–128. ISSN: 0022-4049.

[33] Christopher G. Gibson. *Elementary geometry of algebraic curves: an undergraduate introduction.* Cambridge University Press, Cambridge, 1998, pp. xvi+250. ISBN: 0-521-64140-3; 0-521-64641-3.

[34] Charu Goel, Salma Kuhlmann, and Bruce Reznick. "The analogue of Hilbert's 1888 theorem for even symmetric forms". In: *J. Pure Appl. Algebra* 221.6 (2017), pp. 1438–1448. ISSN: 0022-4049.

[35] Hans-Christian Graf von Bothmer and Kristian Ranestad. "A general formula for the algebraic degree in semidefinite programming". In: *Bull. Lond. Math. Soc.* 41.2 (2009), pp. 193–197. ISSN: 0024-6093.

[36] Daniel R. Grayson and Michael E. Stillman. *Macaulay2, a software system for research in algebraic geometry.* Available at http://www.math.uiuc.edu/Macaulay2/.

[37] Alexander Heaton, Serkan Hoşten, and Isabelle Shankar. "Symmetry Adapted Gram Spectrahedra". In: *SIAM J. Appl. Algebra Geom.* 5.1 (2021), pp. 140–164.

[38] Alexander Heaton and Isabelle Shankar. "An SOS counterexample to an inequality of symmetric functions". In: *J. Pure Appl. Algebra* 225.8 (2021), pp. 106656, 10. ISSN: 0022-4049.

[39] Alexander Heaton and Isabelle Shankar. *SOS counterexample*. https://github.com/alexheaton2/SOS-counterexample. Accessed: 2019-09-18.

[40] Didier Henrion, Simone Naldi, and Mohab Safey El Din. "Exact algorithms for linear matrix inequalities". In: *SIAM J. Optim.* 26.4 (2016), pp. 2512–2539. ISSN: 1052-6234.

[41] J. Herzog, W. V. Vasconcelos, and R. Villarreal. "Ideals with sliding depth". In: *Nagoya Math. J.* 99 (1985), pp. 159–172. ISSN: 0027-7630.

[42] Serkan Hoşten, Isabelle Shankar, and Angélica Torres. *The degree of the central curve in semidefinite, linear, and quadratic programming*. arXiv preprint arxiv.org/abs/2011.15115. 2020. arXiv: 2011.15115.

[43] Adolf Hurwitz. "Über den Vergleich des arithmetischen und des geometrischen Mittels." In: *Journal für die reine und angewandte Mathematik* 108 (1891), pp. 266–268.

[44] Boris V. Kotzev. "Determinantal ideals of linear type of a generic symmetric matrix". In: *J. Algebra* 139.2 (1991), pp. 484–504. ISSN: 0021-8693.

[45] Monique Laurent. "Cuts, matrix completions and graph rigidity". In: vol. 79. 1-3, Ser. B. Lectures on mathematical programming (ismp97) (Lausanne, 1997). 1997, pp. 255–283.

[46] Robert Lazarsfeld. *Positivity in algebraic geometry. I.* Vol. 48. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Springer-Verlag, Berlin, 2004, pp. xviii+387.

[47] Ian G. Macdonald. *Symmetric functions and Hall polynomials*. Second. Oxford Classic Texts in the Physical Sciences. With contribution by Andrey V. Zelevinsky and a foreword by Richard Stanley, Reprint of the 2008 paperback edition. The Clarendon Press, Oxford University Press, New York, 2015, pp. xii+475. ISBN: 978-0-19-873912-8.

[48] Victor Magron and Mohab Safey El Din. "On exact Polya and Putinar's representations". In: *ISSAC'18—Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*. ACM, New York, 2018, pp. 279–286.

[49] Victor Magron and Mohab Safey El Din. "RealCertify: A Maple package for certifying non-negativity". In: *arXiv: https://arxiv.org/abs/1805.02201* (2018). URL: https://arxiv.org/abs/1805.02201.

[50] Laurent Manivel et al. *Complete Quadrics: Schubert Calculus for Gaussian Models and Semidefinite Programming*. To appear in Le Matematiche on Linear Spaces of Symmetric Matrices. arXiv preprint arxiv.org/abs/2011.08791. 2020. arXiv: 2011.08791.

[51] Davesh Maulik and Rahul Pandharipande. "Gromov-Witten theory and Noether-Lefschetz theory". In: *A celebration of algebraic geometry*. Vol. 18. Clay Math. Proc. Amer. Math. Soc., Providence, RI, 2013, pp. 469–507.

[52] Laura Menini, Corrado Possieri, and Antonio Tornambè. "A linear algebra method to decompose forms whose length is lower than the number of variables into weighted sum of squares". In: *Internat. J. Control* 92.11 (2019), pp. 2647–2666. ISSN: 0020-7179.

[53] Mateusz Michałek, Leonid Monin, and Jarosław A. Wiśniewski. *Maximum Likelihood Degree, Complete Quadrics and $\mathbb{C}^*$–Action*. arXiv preprint arxiv.org/abs/2004.07735. 2020. arXiv: 2004.07735.

[54] Robert F. Muirhead. "Some Methods applicable to Identities and Inequalities of Symmetric Algebraic Functions of n Letters". In: *Proceedings of the Edinburgh Mathematical Society* 21 (1902), pp. 144–162.

[55] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. Vol. 13. SIAM Studies in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994, pp. x+405.

[56] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels. "The algebraic degree of semidefinite programming". In: *Math. Program.* 122.2, Ser. A (2010), pp. 379–405. ISSN: 0025-5610.

[57] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Second. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006, pp. xxii+664.

[58] Gabor Pataki. *On Positive Duality Gaps in Semidefinite Programming*. arXiv preprint arxiv.org/abs/1812.11796. 2020. arXiv: 1812.11796.

[59] Daniel Plaumann, Bernd Sturmfels, and Cynthia Vinzant. "Quartic curves and their bitangents". In: *J. Symbolic Comput.* 46.6 (2011), pp. 712–733. ISSN: 0747-7171.

[60] Victoria Powers et al. "A new approach to Hilbert's theorem on ternary quartics". In: *C. R. Math. Acad. Sci. Paris* 339.9 (2004), pp. 617–620. ISSN: 1631-073X.

[61] Motakuri Ramana and Alan J. Goldman. "Some geometric results in semidefinite programming". In: *J. Global Optim.* 7.1 (1995), pp. 33–50. ISSN: 0925-5001.

[62] Annie Raymond, Mohit Singh, and Rekha R. Thomas. "Symmetry in Turán sums of squares polynomials from flag algebras". In: *Algebr. Comb.* 1.2 (2018), pp. 249–274.

[63] Annie Raymond et al. "Symmetric sums of squares over k-subset hypercubes". In: *Math. Program.* 167.2, Ser. A (2018), pp. 315–354. ISSN: 0025-5610.

[64] Bruce Reznick. "A quantitative version of Hurwitz' theorem on the arithmetic-geometric inequality". English (US). In: *Journal fur die Reine und Angewandte Mathematik* 1987.377 (Jan. 1987), pp. 108–112. ISSN: 0075-4102.

[65] Joshua D. Rhodes. *Computing the Central Sheet in Linear, Quadratic, and Semidefinite Programs*. Masters Thesis. 2016.

[66] Raphael M. Robinson. "Some definite polynomials which are not sums of squares of real polynomials". In: *Selected questions of algebra and logic (collection dedicated to the memory of A. I. Mal'cev) (Russian)*. 1973, pp. 264–282.

[67] Bruce E. Sagan. *The symmetric group*. Second. Vol. 203. Graduate Texts in Mathematics. Representations, combinatorial algorithms, and symmetric functions. Springer-Verlag, New York, 2001, pp. xvi+238. ISBN: 0-387-95067-2.

[68] James Saunderson and Venkat Chandrasekaran. *Terracini Convexity.* arXiv preprint arxiv.org/abs/2010.00805. 2020. arXiv: 2010.00805.

[69] Claus Scheiderer. "Sum of squares length of real forms". In: *Math. Z.* 286.1-2 (2017), pp. 559–570. ISSN: 0025-5874.

[70] Claus Scheiderer. "Sums of squares of polynomials with rational coefficients". In: *J. Eur. Math. Soc. (JEMS)* 18.7 (2016), pp. 1495–1513. ISSN: 1435-9855.

[71] Dennis Schlief. *Degree for the Central Curve of Quadratic, Programing.* Masters Thesis. 2014.

[72] Jean-Pierre Serre. *Linear representations of finite groups.* Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977, pp. x+170. ISBN: 0-387-90190-6.

[73] Suvrit Sra. "On inequalities for normalized Schur functions". In: *European J. Combin.* 51 (2016), pp. 492–494. ISSN: 0195-6698.

[74] Richard P. Stanley. *Enumerative combinatorics. Vol. 2.* Vol. 62. Cambridge Studies in Advanced Mathematics. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge University Press, Cambridge, 1999, pp. xii+581. ISBN: 0-521-56069-1; 0-521-78987-7.

[75] Bernd Sturmfels and Caroline Uhler. "Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry". In: *Ann. Inst. Statist. Math.* 62.4 (2010), pp. 603–638. ISSN: 0020-3157.

[76] Joel A. Tropp. "Simplicial faces of the set of correlation matrices". In: *Discrete Comput. Geom.* 60.2 (2018), pp. 512–529. ISSN: 0179-5376.

[77] Paul Yiu. "The length of $x_1^4 + x_2^4 + x_3^4 + x_4^4$ as a sum of squares". In: *J. Pure Appl. Algebra* 156.2-3 (2001), pp. 367–373. ISSN: 0022-4049.