COMPUTING THE CENTRAL SHEET IN LINEAR, QUADRATIC, AND SEMI-DEFINITE PROGRAMS

A thesis presented to the faculty of
San Francisco State University
In partial fulfilment of
The Requirements for
The Degree

Master of Arts
In
Mathematics

by

Joshua Daniel Rhodes

San Francisco, California

August 2016

CERTIFICATION OF APPROVAL

I certify that I have read *Computing the Central Sheet in Linear, Quadratic, and Semi-definite Programs* by Joshua Daniel Rhodes and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirements for the degree: Master of Artsin Mathematicsat San Francisco State University.

Serkan Hoşten
Professor of Mathematics

Joseph Gubeladze
Professor of Mathematics

Federico Ardila
Professor of Mathematics

COMPUTING THE CENTRAL SHEET IN LINEAR, QUADRATIC, AND SEMI-DEFINITE PROGRAMS

Joshua Daniel Rhodes
San Francisco State University
2016

Convex optimization is the study of finding optimal solutions to various convex objective functions given a system of constraints defining a convex solution set. Interior point algorithms are utilized to obtain optimal solutions to a given convex program by navigating the solution space through Newton-type iterations. For linear programs, one can obtain an algebraic handle on the path of the interior point algorithm by finding the Zariski closure of such a path, called the central curve. Previous work has related computational complexity of these algorithms to the degree of the central curve. This degree has been computed for a generic linear program. This thesis will explore how to compute the degree of the central curve for the transportation problem over complete bipartite graphs. We will propose new methods for computing the central sheet—a higher dimensional object of the same degree as the central curve—for quadratic and semi-definite programs. We will also present conjectures about the degree of the central sheet for semidefinite programs supported by computational evidence.

I certify that the Abstract is a correct representation of the content of this thesis.

Chair, Thesis Committee Date

ACKNOWLEDGMENTS

I would like to thank my advisor Serkan Hoşten as well as the mathematics faculty at SFSU.

TABLE OF CONTENTS

Chapter 1

Introduction

The field of optimization is interested in finding optimal solutions to a given function within the confines of particular constraints. There are many ways to solve the many separate classes of optimization problems. The method of most interest for this thesis is a class of algorithms known as interior point methods. These are algorithms designed to arrive at an optimal solution through a Newton-type iteration process. However, because of the inherent imprecision in this algorithm, it is in the interest of mathematical soundness that the "path" these interior point methods follow be understood. This thesis establishes results for such paths in three separate areas of convex optimization: linear, quadratic, and semidefinite programming.

The standard form of a linear program is:

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{Subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0, \end{aligned} \tag{1.1}$$

where $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, A is a $d \times n$ matrix, and $\mathbf{b} \in \mathbb{R}^d$. There exists interior point methods for finding the optimal solution to such a system through the log-barrier approach, which augments the standard form to

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \mathbf{x} - \lambda \textstyle\sum_{i=1}^{n} \log(x_i) \\ \text{Subject to} \quad & A\mathbf{x} = \mathbf{b}. \end{aligned} \tag{1.2}$$

The optimal solutions to this new log-barrier system are in the interior of the feasible region of solutions for (1.1) for all values of $\lambda > 0$. Paired with log-barrier problem is the Karush-Kuhn-Tucker (KKT) conditions that state that a vector $(\mathbf{x}, \lambda)$ is optimal to the log-barrier program if and only if there exists a $\mathbf{y} \in \mathbb{R}^d$ such that

$$c_i - \lambda x_i^{-1} - \mathbf{y}^T A_i = 0, \;\; i = 1, \ldots, n$$
$$A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geq 0$$

The coordinate projection of all optimal $(\mathbf{x}, \lambda, \mathbf{y})$, where $\lambda \geq 0$, onto $\mathbb{R}^n$ forms the *central path* of the linear program (1.1). The Zariski closure of the central path in $\mathbb{C}^n$ is an algebraic curve known as the *central curve* of the linear program (1.1). This central curve was introduced in [8]. This thesis proves results on the degree and defining equations of the central curve for linear, quadratic, and semidefinite programs.

In Chapter 2, we give some background. This includes the definition of critical objects to this thesis, such as the *central path, central curve, and central sheet*. We will give some background for each type of optimization we look at and how they are connected. We will mention classical tools and results that will be called upon. This includes the Karush-Kuhn-Tucker conditions in optimization and hyperplane arrangements, matroids, and their Tutte polynomials for the combinatorial side of our work.

In Chapter 3 we utilize previous work to find an upper bound on the central curve for a class of transportation problems. It is known that the path the interior point method follows has complexity tied to the degree the central curve. In particular, we prove the following:

Theorem 1.1. *For a generic* $\mathbf{c} \in \mathbb{R}^n$ *and generic feasible* $\mathbf{b} \in \mathbb{R}^d$*, the central curve for the transportation problem over the graph* $K_{2,n}$ *has degree*

$$(n-2)2^{n-1}+1.$$

This theorem can be further generalized to the transportation problem over the graph $K_{m,n}$, where we state the explicit degree.

In Chapter 4 we extend methods from Chapter 3 to quadratic programming. The general form of a quadratic program is

$$\begin{aligned} \text{Minimize} \quad & \tfrac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{1.3}$$

where Q is an $n \times n$ positive-definite matrix. This problem comes with its own log-barrier problem whose KKT conditions are the following: Any $(\mathbf{x},\lambda)$ is an optimal solution to the log-barrier program if and only if there exists a $\mathbf{y}\in\mathbb{R}^d$ such that

$$\begin{aligned}(Q\mathbf{x})_i + c_i - \lambda x_i^{-1} - \mathbf{y}^T A_i = 0,\ i = 1,\dots,n \\ A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq 0.\end{aligned} \tag{1.4}$$

For $\lambda \geq 0$, the projection of $(\mathbf{x},\lambda,\mathbf{y})$ onto $\mathbb{R}^n$ forms the central path of the quadratic program (1.3), and its Zariski closure again provides us an algebraic curve in $\mathbb{C}^n$, the *central curve*. Results on the degree of the central curve of (1.3) have been obtained by [10] and we expand on those results. We propose a method for computing the equations of the central curve for a given quadratic program, and show that this method works for the case where Q is diagonal.

Then we further explore convex programming through semidefinite programs. A generic semidefinite program is of the form:

$$\begin{aligned}\text{Minimize}\quad & \operatorname{Tr}(C\mathbf{X}) \\ \text{subject to}\quad & \operatorname{Tr}(A_i\mathbf{X}) = b_i, i = 1,\dots d \\ & \mathbf{X} \succeq 0\end{aligned} \tag{1.5}$$

where $\operatorname{Tr}(\mathbf{X})$ represent the trace of a matrix $\mathbf{X}$. Here C and A_i are $n\times n$ symmetric matrices and $b_i \in \mathbb{R}$. Also the program requires $\mathbf{X}$ to be a positive-definite matrix which is denoted by $\mathbf{X}\succeq 0$. Using a log-barrier approach in a similar light, the KKT conditions for such a program states that an $n \times n$ matrix $\mathbf{X}$ is optimal to the log barrier problem with parameter λ if and only if there exists a $\mathbf{y} \in \mathbb{R}^d$ such that

$$\begin{aligned} & C - \lambda \mathbf{X}^{-1} - \textstyle\sum_{i=1}^{d} y_i A_i = 0 \\ & \operatorname{Tr}(A_i \mathbf{X}) = b_i \;\; i = 1, \ldots, d \quad . \\ & \mathbf{X} \succ 0. \end{aligned} \tag{1.6}$$

We again use these conditions to obtain an algebraic handle on the central path which is the projection of $(\mathbf{x}, \lambda, \mathbf{y})$ onto $\mathbb{R}^{n \times n}$ for all $\lambda \geq 0$. The Zariski closure of such points gives us the central curve of the semidefinite program (1.5). We also utilize a new method for computing the generators of this curve. There we observe that the central curve for a generic semidefinite program can be described combinatorially, and the degree for a fixed n appears to be symmetric in terms of the constraints d as well as exponential with base $(n-1)$.

In Chapter 5 we discuss open conjectures and possible routes for exploring such conjectures.

Chapter 2

Background

2.1 Tools for Linear Programming

The first set of results in this paper are in the field of linear programming. Linear programs are optimization problems where the cost function is a linear function and the constraints form a convex polytope. The standard notation for a general linear program is of the form

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}^T \mathbf{x} \\ \text{Subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{2.1}$$

where $\mathbf{c}^T$ is the transpose of the cost vector in $\mathbb{R}^n$, A is a $d \times n$ matrix, and $\mathbf{b}$ is a vector in $\mathbb{R}^d$ that, along with the rows of A, hold the information of our constraints. We have n

decision variables and d constraints.

In practice, many optimization problems have large numbers of variables, so methods for finding optimal solutions can be computationally challenging. Various methods are utilized depending on the quantity and type of data one may have. This paper is concerned with interior point algorithms [2]. Among the interior point algorithms is a logarithmic barrier approach, where, for $\lambda \in \mathbb{R}_{\geq 0}$, one considers a modified problem

$$\begin{aligned} &\text{Minimize} \quad \mathbf{c}^T\mathbf{x} - \lambda \textstyle\sum_{i=1}^n \log(x_i) \\ &\text{Subject to} \quad A\mathbf{x} = \mathbf{b}. \end{aligned} \tag{2.2}$$

Here, we see that the condition $x_i \geq 0$ is implied, and, if $\lambda = 0$, we have the original problem (2.1). For large positive λ, there is a unique solution sitting in the interior of the solution space as the $-\lambda \sum_{i=1}^n \log(x_i)$ forces the optimal solution away from the hyperplanes formed by $x_i = 0$. As λ iteratively goes to 0, we see that the $\lambda \sum_{i=1}^n \log(x_i)$ portion becomes negligible. Thus, when $\lambda = 0$, the solution to the general linear program and the solution to our logarithmic barrier problem agree.

The collection of all unique solutions to the logarithmic barrier function for every positive λ forms a curve called the central path. In practice, walking down the central path is done in Newton-type iterations. Thus, in order to compute the complexity of a problem before dedicating valuable computation time, it behooves one to know the number of such Newton-type steps required to find the optimal solution. The number of such steps can be measured in terms of the total curvature of the central path, so finding an upper bound on the curvature can give insight into the complexity before dedicating computing assets. Simply put, the more curvature along a path, the smaller and more frequent steps are and the longer an optimal solution will take to compute.

It has been shown, and will be reviewed below, that points on the central path can be represented as the solution to a system of polynomial equations where particular solutions to this system overlap with the central path [8]. The solutions to such a system of polynomial equations forms a curve called the central curve. It has been shown in [8] that knowing the degree of the central curve provides us with an upper bound to the curvature of the central path. This chapter will provide some results on the central curve of the optimization problems that we will study.

2.1.1 Computing the Central Sheet

A classical result from convex optimization is the *Karush-Kuhn-Tucker* (KKT) conditions—following from 2.2 and its dual:

Theorem 2.1 (Karush-Kuhn-Tucker)**.** *A solution* $\mathbf{x}$ *is optimal to* (2.2) *if and only if there exists a* $\mathbf{y}$ *such that:*

$$\begin{array}{ll} (i) & c_i - \lambda x_i^{-1} - \mathbf{y}^T A_i = 0, \;\; i = 1, \ldots, n \\ (ii) & A\mathbf{x} = \mathbf{b} \\ (iii) & \mathbf{x} \geq 0 \end{array}$$

where A_i *is the ith column of A and* $\mathbf{y}$ *is a vector with variable entries* y_i *for* $i = 1, \ldots, d$.

One method for computing the central curve, following work from [8] which will be referred to as the *elimination method* throughout the paper, is constructed through the following steps:

1. Clear denominators in each of the equations from (i) of the KKT conditions by right-multiplying all terms with a generic solution vector $\mathbf{x}$. This results in n equations of the form $c_i x_i - \lambda - \mathbf{y}^T A_i x_i = 0$ for $i = 1, \dots, n$

2. Let I be the ideal generated by these equations as well as the d linear equations from (ii) where $A\mathbf{x} - \mathbf{b} = 0$.

3. Eliminate the variables $\lambda, y_1, \dots, y_d$ from the ideal I. The resulting ideal will be called I_e.

4. The affine variety of I_e is the central curve in $\mathbb{R}^n$.

Remark. It is worth noting that we ignored the non-negativity constraints $\mathbf{x} \geq 0$. Therefore the central curve contains all central paths of all bounded regions formed from all combinations of halfspaces that come from these non-negativity constraints.

Definition 2.1. The *central sheet* of the linear program (2.1) is the variety of the ideal formed from steps $1-3$ without having the $A\mathbf{x} - \mathbf{b} = 0$ equations from step 2 in our ideal I.

Another method for computing the ideal that describes the central sheet is outlined in [8]. While the details can be explored more deeply in their paper, here is a brief outline of the methods and reasoning behind how they compute the central sheet and central curve. In (2.2), we can reinterpret

$$c_i - \lambda x_i^{-1} - \mathbf{y}^T A_i = 0 \ \text{ for } \ i = 1, \ldots, n$$

as $c_i - y^T A_i = \lambda x_i^{-1}$, $i = 1, \ldots, n$, or, more precisely,

$$\begin{array}{cccc} c_1 & -\mathbf{y}^T A_1 & = & \lambda x_1^{-1} \\ c_2 & -\mathbf{y}^T A_2 & = & \lambda x_2^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ c_n & -\mathbf{y}^T A_n & = & \lambda x_n^{-1} \end{array},$$

and expanding terms where $a_{i,j}$ represent the ith row, jth column entry of A, we arrive at

$$\begin{array}{cccccccc} c_1 & -y_1 a_{1,1} & -y_2 a_{2,1} & \ldots & -y_d a_{d,1} & = & \lambda x_1^{-1} \\ c_2 & -y_1 a_{1,2} & -y_2 a_{2,2} & \ldots & -y_d a_{d,2} & = & \lambda x_2^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_n & -y_1 a_{1,n} & -y_2 a_{2,n} & \ldots & -y_d a_{d,n} & = & \lambda x_n^{-1} \end{array}.$$

Arranging the conditions in such a way allows one to see that, because the y_i are real valued numbers, the coordinate-wise inverse of $\mathbf{x}$ lies in the row span of A and $\mathbf{c}$. More succinctly — if we let $\mathbf{x}^{-1} = \begin{bmatrix} \frac{1}{x_1} & \frac{1}{x_2} & \dots & \frac{1}{x_n} \end{bmatrix}$, then

$$\mathbf{x}^{-T} \in \operatorname{span}\{A^T, \mathbf{c}\}.$$

So let us construct the matrix

$$\binom{c}{A} = \begin{bmatrix} c_1 & c_2 & \dots & c_n \\ a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,n} \end{bmatrix}.$$

We can see that for a generic A and c, $\operatorname{rank}\binom{c}{A} = d+1$, and because $\mathbf{x}^{-T} \in \operatorname{span}\{A^T, \mathbf{c}\}$,

$$\operatorname{rank}\left(\begin{bmatrix} c_1 & c_2 & \dots & c_n \\ a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,n} \\ x_1^{-1} & x_2^{-1} & \dots & x_n^{-1} \end{bmatrix} \right) = d+1.$$

This means that every $(d+2)$-minor of the above matrix is equal to 0. However, these minors are not polynomials due to the x_i^{-1} terms. Because A is generic, forming the generators for the central sheet is simple—for each $(d+2)$-minor, associate a set C containing all $(d+2)$ column indices for the given $(d+2)$-minor, then multiply the determinant of the given minor by the monomial

$$\prod_{i\in C} x_i.$$

The following example should shed some light on both the elimination method and the method outlined above.

Example 2.1. Consider a randomly generated linear program of the form (1.1) where

$$\mathbf{c} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 5 \\ 7 \\ \frac{8}{3} \end{bmatrix}, A = \begin{bmatrix} \frac{1}{8} & \frac{9}{10} & 2 & \frac{5}{3} & \frac{7}{3} & 4 \\ \frac{5}{4} & 6 & 1 & 3 & \frac{5}{2} & 6 \\ \frac{7}{9} & \frac{9}{4} & 6 & \frac{7}{8} & 1 & \frac{5}{2} \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} \frac{857}{60} \\ \frac{4043}{168} \\ \frac{19711}{1512} \end{bmatrix}.$$

The elimination method for this example starts with the first of the KKT conditions where we clear denominators and form an ideal with n equations of the form

$$c_i x_i - \lambda - \mathbf{y}^T A_i x_i = 0.$$

Then, to calculate the central sheet, we eliminate the variables (λ, y_1, y_2, y_3) from this ideal.

The resulting ideal has generators of the form

$$4095x_2x_3x_4x_5 - 187516x_2x_3x_4x_6 + 270236x_2x_3x_5x_6 + 13912x_2x_4x_5x_6 - 63400x_3x_4x_5x_6,$$
$$492915x_1x_3x_4x_5 + 785396x_1x_3x_4x_6 - 2143876x_1x_3x_5x_6 - 201152x_1x_4x_5x_6 + 1369440x_3x_4x_5x_6,$$
$$605769x_1x_2x_4x_5 - 2112996x_1x_2x_4x_6 + 1934772x_1x_2x_5x_6 - 1005760x_1x_4x_5x_6 + 1502496x_2x_4x_5x_6,$$
$$5598681x_1x_2x_3x_5 - 7482996x_1x_2x_3x_6 - 967386x_1x_2x_5x_6 - 5359690x_1x_3x_5x_6 + 14592744x_2x_3x_5x_6,$$
$$1885758x_1x_2x_3x_4 - 3741498x_1x_2x_3x_6 - 528249x_1x_2x_4x_6 - 981745x_1x_3x_4x_6 + 5062932x_2x_3x_4x_6,$$

which results in the generators of the central sheet.

The circuit method begins with the matrix

$$\begin{bmatrix} 3 & 1 & 1 & 5 & 7 & \frac{8}{3} \\ \frac{1}{8} & \frac{9}{10} & 2 & \frac{5}{3} & \frac{7}{3} & 4 \\ \frac{5}{4} & 6 & 1 & 3 & \frac{5}{2} & 6 \\ \frac{7}{9} & \frac{9}{4} & 6 & \frac{7}{8} & 1 & \frac{5}{2} \\ \frac{1}{x_1} & \frac{1}{x_2} & \frac{1}{x_3} & \frac{1}{x_4} & \frac{1}{x_5} & \frac{1}{x_6} \end{bmatrix}.$$

To construct the generators of the central sheet we take the determinant of each 5-minor then multiply with each x_i corresponding to the indices of the minor. For example, the one generator would come from the minor excluding the second column of the above matrix.

This can be explicitly written as

$$\det \begin{bmatrix} 3 & 1 & 5 & 7 & \frac{8}{3} \\ \frac{1}{8} & 2 & \frac{5}{3} & \frac{7}{3} & 4 \\ \frac{5}{4} & 1 & 3 & \frac{5}{2} & 6 \\ \frac{7}{9} & 6 & \frac{7}{8} & 1 & \frac{5}{2} \\ \frac{1}{x_1} & \frac{1}{x_3} & \frac{1}{x_4} & \frac{1}{x_5} & \frac{1}{x_6} \end{bmatrix} x_1x_3x_4x_5x_6$$

which is

$$\frac{1}{10368}(-492915x_1x_3x_4x_5 - 785396x_1x_3x_4x_6 + 2143876x_1x_3x_5x_6 + 201152x_1x_4x_5x_6 - 1369440x_3x_4x_5x_6)$$

and this is the second generator from our elimination method above. After computing the generators for the central sheet through both the elimination method and the circuit method in *Macaulay2* [6], we find that the two ideals are the same!

While the circuit method provides a new approach to calculating the central sheet, we are only able to follow the example above when A and $\mathbf{c}$ are generic. However not all optimization problems are generic, so obtaining a method for calculating the generators of the central sheet for a non-generic linear program would be pertinent. Therefore, in order to be able to calculate the generators of the central curve explicitly through the circuit method, as they do in [8], it is worth reviewing the necessary ingredients.

2.1.2 A Glance at Matroids and Related Tools

A *matroid* is a generalization of the notion of linear independence in linear algebra. So we begin with a definition:

Definition 2.2. A Matroid (G,I) is a pair of sets G and I such that, for a *ground set* G, the set of *independent sets* I is a collection of subsets of G that satisfy the following:

1. $\emptyset \in I$.

2. If $S' \subset S$ and $S \in I$, then $S' \in I$.

3. If $S, S' \in I$ and $|S| < |S'|$, then there is an $x \in S'$ such that $\{x\} \cup \{S\} \in I$ and $|\{x\} \cup \{S\}| > |S|$.

It is also worth noting that a subset of G is called *dependent* if it is not independent.

Another definition for matroids can also be given through constructing a set of *circuits* for a given set instead of the independent sets.

Definition 2.3. A Matroid (G,C) is a pair of sets G and C such that, for a ground set G, the set of *circuits* C is a collection of subsets of G such that:

1. For any pair of circuits C_1, C_2, $C_1 \subsetneq C_2$ or $C_2 \subsetneq C_1$.

2. If $x \in C_1 \cap C_2$, then there exists a $C_3 \subset (C_1 \cup C_2) \setminus \{x\}$.

Remark. It is worth noting that, for any $x \in C$ from our second definition, $(C \setminus \{x\}) \in I$ from our first definition. That is, it is sufficient to define a matroid's independence relationships by the sets of minimally dependent sets.

Because this thesis does not require such heavy machinery to construct a proof, it is sufficient to know that a *realizable matroid* is given by a matrix and the independent sets of a realizable matroid are sets of independent columns of the matrix. In this thesis, any use of a matroid interpretation will come with an explanation in terms of the pertinent matrix, so an in-depth understanding of matroids is not necessary and a firm grasp of undergraduate linear algebra is sufficient.

Example 2.2. A quick example can be shown for the matroid with ground set $\{1,2,3,4,5,6\}$ and maximal independent sets

$$\{1,2,3,4\},\quad \{1,2,3,5\},\quad \{1,2,3,6\},\quad \{2,3,4,6\},\quad \{2,3,5,6\},\quad \{1,3,4,5\},$$
$$\{1,3,5,6\},\quad \{1,2,4,6\},\quad \{1,2,5,6\},\quad \{1,4,5,6\},\quad \{2,4,5,6\},\quad \{3,4,5,6\}$$

which is realizable through the matrix

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

It is sufficient to define only the maximal independent sets as any subset of these sets will also be independent. We can also construct all circuits for this example, which includes all circuits of five elements

$$\{1,2,3,4,5\},\{1,2,3,4,6\},\{1,2,3,5,6\},$$
$$\{1,2,4,5,6\},\{1,3,4,5,6\},\{2,3,4,5,6\}$$

as well as the circuits with only four elements

$$\{1,2,4,5\}, \quad \{1,3,4,6\} \quad \{2,3,5,6\} \ .$$

We can also see $\{1,2,4,5\}$ is a circuit because the set $\{1,2,4,5\}$ corresponds to the submatrix

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

formed from columns 1,2,4,5, which is minimally linearly dependent, as removal of any of the columns would result in an independent set.

2.1.3 Arrangements and the Tutte Polynomial

Another pertinent object of study in matroids is the *Tutte Polynomial*, which is an invariant of matroids. Because of the many applications of matroids to graph theory, the Tutte polynomial has been studied extensively in [4].

Definition 2.4. If we let A be a matroid, B a subset of A, $r(X)$ denote the number of elements in the largest independent subset of X, and $|X|$ denote the number of elements in a subset X, then the Tutte polynomial of a matroid A is given by

$$T_A(x,y) = \sum_{B \subset A} (x-1)^{r(A)-r(B)}(y-1)^{|B|-r(B)}.$$

To restate this definition in a realizable matroid context, one can think of A as a matrix, B as a submatrix of A, $r(X)$ as the rank of the submatrix X, and $|X|$ as the number of columns of a matrix X.

A particularly interesting invariant related to the Tutte polynomial is the Möbius invariant μ. This thesis will use a critical relationship between μ and *hyperplane arrangements*. Hyperplane arrangements and their relation to μ can be explored in great detail in [11].

One can define a hyperplane in $\mathbb{R}^n$ as the set

$$H_i = \{\mathbf{x} \in \mathbb{R}^n | A_i\mathbf{x} = b_i\}.$$

The hyperplanes this thesis is concerned with are given from a line by line reading of the system $A\mathbf{x} = \mathbf{b}$ in (2.1) as well as the hyperplanes $x_i = 0$ and $\mathbf{cx} = c_0$ for some random $c_0 \in \mathbb{R}$. Namely, we are concerned with the number of bounded regions formed by the hyperplanes $\mathbf{x}_i = 0$ in the affine space $\binom{c}{A}\mathbf{x} = \binom{c_0}{\mathbf{b}}$ [8, pg. 522]. Thus, the hyperplane arrangement we are interested in can be defined as the set

$$(\cup_{i=1,\dots,n}\{x_i = 0\}) \cap \{\mathbf{x} | \binom{c}{A} \mathbf{x} = \binom{c_0}{\mathbf{b}}\}$$

The reasoning for this is that the degree of the central curve of (2.1) is related to the number of bounded regions formed by H.

It is known that the number of bounded regions in a hyperplane arrangement is the Möbius number of the lattice of flats of the arrangement [11]. It is also known that the lattice of flats is isomorphic to the lattice of flats to an appropriate matroid. The Möbius number is also known to be an evaluation of the Tutte polynomial of this matroid at $(1,0)$ [1]. Therefore, in an effort to obtain the Möbius number as a tool for finding the degree of the central curve or central sheet, this thesis will utilize the Tutte polynomial.

Example 2.2. (continued) Suppose we want to compute the Tutte polynomial for our previous realizable matroid, with ground set $\{1,2,3,4,5,6\}$ and representation

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

Recall that the Tutte polynomial is

$$T_A(x,y) = \sum_{B \subset A} (x-1)^{r(A)-r(B)}(y-1)^{|B|-r(B)}.$$

This example will not show the steps for all 2^6 different possible subsets B, however we will show calculations for the circuit $\{1,2,4,5\}$. For

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

we see the contribution to the sum for the Tutte polynomial is

$$(x-1)^3(y-1)^1 = x^3y - 3x^2y + 3xy - y - x^3 + 3x^2 - 3x + 1.$$

Software allows for a faster computation of the entire Tutte polynomial, and gives us

$$T_{A_e}(x,y) = x^4 + 2x^3 + 3x^2 + 3xy + y^2 + x + y.$$

To further elaborate on this thesis' methods, we also find that the Möbius number for our example's hyperplane arrangement

$$H_e = \{\mathbf{x}|A_e\mathbf{x} = \mathbf{b}_e\} \text{ for a randomly generated } \mathbf{b}_e = \begin{bmatrix} 14 \\ 17 \\ 11 \\ 12 \end{bmatrix}$$

through the evaluation of the Tutte polynomial

$$\mu_{A_e} = T_{A_e}(1,0) = 7.$$

Such a $\mathbf{b}_e$ was given so that one can confirm visually that the number of bounded regions is indeed 7, see figure 2.2.

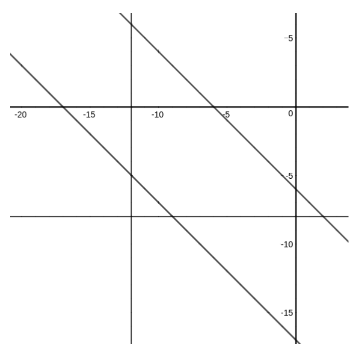

Figure 2.1: A projection of the hyperplanes $x_i = 0$ in $A_e\mathbf{x} = \mathbf{b}_e$ onto the coordinates x_5, x_6.

2.1.4 Degree of the Central Sheet of Linear Programs

In [8] a method is proposed to construct the generators of the ideal defining the central sheet of a linear program. Let $\mathscr{I}$ represent a given circuit of $\binom{c}{A}$. Also let $\begin{pmatrix} c \\ A \\ \mathbf{x}^{-1} \end{pmatrix}_{\mathscr{I}}$ be the submatrix of $\begin{pmatrix} c \\ A \\ \mathbf{x}^{-1} \end{pmatrix}$ corresponding to the column indices in $\mathscr{I}$. Then, we define one of the generators $g_{\mathscr{I}}$ to the ideal of the central sheet as

$$g_{\mathscr{I}} = \det \begin{pmatrix} c \\ A \\ \mathbf{x}^{-1} \end{pmatrix} \prod_{i \in \mathscr{I}} x_i.$$

Theorem 2.2 (Proudfoot-Speyer [9])**.** *The ideal describing the central sheet is*

$$J_c = \langle g_{\mathscr{I}} | \mathscr{I} \text{ is a circuit of } \tbinom{c}{A} \rangle.$$

An example of these calculations is given in the beginning of Chapter 3 of this thesis.

Theorem 2.3. *Let I be the ideal described in definition 2.1 and let J_c be ideal described above, then*

$$I = J_c.$$

In order to create the ideal for the central curve, we add the constraint equations (2.1) as additional generators to the ideal I_c. This leads us to further mathematical grounding to [8, Lemma 11]:

Proposition 2.4. *For a linear program as in* (2.1) *with generic cost vector* **c** *and generic* **b***, the ideal constructed through the elimination method I_e is equal to $I_c + \langle A_i\mathbf{x} = b_i \rangle$ for*

$i = 1,\ldots,d$, the ideal constructed through the method above in addition to the constraint equations from (2.1)*. Further, the ideal I_e defines an irreducible curve.*

Proof. $I_e = I_c + \langle A_i\mathbf{x} = b_i, i = 1,\ldots,d\rangle$ by [8, Lemma 11]. The ideal I_c is irreducible by [8, Lemma 10]. Each hyperplane of the form $H_i = \{x|A_i\mathbf{x} = b_i\}$ is a general hyperplane, and $V(I_c) \not\subset H_i \;\; i = 1,\ldots,d$. Then [7, Prop 18.10] shows that $I_c + \langle A_1\mathbf{x} = b_1\rangle$ is irreducible and non-degenerate. The ideal

$$I_c + \langle A_1\mathbf{x} = b_1\rangle$$

is an irreducible variety of one dimension lower than I_c, but because H_2 is generic, then

$$I_c + \langle A_1\mathbf{x} = b_1\rangle + \langle A_2\mathbf{x} = b_2\rangle$$

is irreducible and non-degenerate, again by [7]. This argument will continue until we arrive at the one dimensional variety

$$I_c + \langle A_1\mathbf{x} = b_1\rangle + \ldots + \langle A_d\mathbf{x} = b_d\rangle = I_e$$

which will also be irreducible. □

In [8], it is shown that an appropriate upper bound on the total curvature of the central path is given by the *degree* of the central curve. This can be read from the *Hilbert Polynomial*, which [11] provides many more details.

A particularly advantageous aspect of computing the central sheet is that the degree of the central sheet is the same as the degree of the central curve [8, Lemma 11] [5]. Therefore, if one would like to compute an upper bound on the curvature of the central path, it is equivalent to computing the degree of the central sheet. This leads us to the following theorem:

Theorem 2.5 (Proudfoot-Speyer [9]). *The degree of the central sheet J_c, regarded as a variety in complex projective space, coincides with the Möbius number* $\mu_{\binom{c}{A}}$.

This results in the following corollary:

Corollary 2.6 (Theorem 13 [8]). *For a linear program* (2.1) *with generic* $A \in \mathbb{R}^{d\times n}$, $\mathbf{c} \in \mathbb{R}^n$, *and* $\mathbf{b} \in \mathbb{R}^d$*, the degree of the primal central curve is at most*

$$\binom{n-1}{d}.$$

This corollary leads one to the question, "What is the degree for a non-generic linear program?" which this thesis explores in Chapter 3.

2.2 Tools for Quadratic Programming

Another branch of convex optimization is Quadratic Programming. The general form for a quadratic program is of the form

$$\begin{aligned} \text{Minimize} \quad & \tfrac{1}{2}\mathbf{x}^T Q\mathbf{x} + c^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{2.3}$$

where Q is an $n \times n$ symmetric positive-definite matrix, A is a $d \times n$ real matrix, and $\mathbf{b}$ is a feasible vector in $\mathbb{R}^n$. There is an interior point method for quadratic programs as well, and so again it is our interest to have an algebraic handle on the path that such an algorithm follows. In order to do this, we consider the log barrier function and its dual, to arrive at the KKT conditions for quadratic programming. The KKT conditions state that a solution $(\mathbf{x}, \lambda)$ is an optimal solution to (2.2) if and only if there exists a $\mathbf{y} \in \mathbb{R}^d$ such that:

$$\begin{gathered} (Q\mathbf{x})_i + c_i - \lambda x_i^{-1} - (\mathbf{y}^T A_i) = 0, \;\; i = 1, \ldots, n \\ A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq 0, \end{gathered} \tag{2.4}$$

where $\mathbf{y} \in \mathbb{R}^d$ and A_i is the ith column of A. We will utilize this fact to obtain results in Chapter 4, but we hope to confirm our methods with the elimination method for quadratic programs as well. The elimination method is as follows:

1. Take the first set of equations from (2.4) and right multiply by each respective term x_i to clear denominators.

2. This forms a polynomial ideal $\mathscr{I}$ in terms of x_i, y_j, λ as variables.

3. Eliminate the variables y_j, λ from $\mathscr{I}$, which results in the ideal defining the central sheet of our quadratic program. Let this new ideal be denoted as I_{qe}.

4. The variety of $(I_{qe} + \langle A_i\mathbf{x} - b_i,\ \ i = 1, \ldots, d\rangle)$ forms the central curve.

Some preliminary results on the structure of such ideals can be found in [10], in particular the following theorem:

Theorem 2.7. *For a generic program of the form 2.2, the degree of the central curve is equal to*

$$\sum_{k=0}^{n-d-1} \binom{n-k-2}{d-1} 2^k.$$

2.3 Tools For Semidefinite Programming

Semidefinite programming is a field of recent interest, especially since interior point methods can be used to solve these problems in polynomial time [2]. Thus, it is in the interest of mathematical soundness to get an algebraic handle on the path that an algorithm may take to find an optimal solution.

A semidefinite program (SDP) is generally given as follows: Let $\mathbb{R}^{n\times n}_{\succeq 0}$ be the set of all $n \times n$ positive semidefinite matrices and let $\mathbf{X} \succeq 0$ denote that the matrix $\mathbf{X}$ is positive semidefinite. For some symmetric objective matrix $C \in \mathbb{R}^{n\times n}$, some decision matrix $\mathbf{X} \in \mathbb{R}^{n\times n}_{\succeq 0}$, and d many given symmetric constraint matrices $A_i \in \mathbb{R}^{n\times n}$ the SDP is

$$\begin{aligned} \text{Minimize} \quad & \mathrm{Tr}(C\mathbf{X}) \\ \text{subject to} \quad & \mathrm{Tr}(A_i\mathbf{X}) = b_i, i = 1,\dots,d \\ & \mathbf{X} \succeq 0 \end{aligned} \tag{2.5}$$

where $\mathrm{Tr}(\mathbf{X})$ represent the trace of a matrix $\mathbf{X}$ and $b_i \in \mathbb{R}$.

It is worth noting that an SDP is a generalization of a linear program, namely, if we insist the that $\mathbf{X} \succeq 0$ are diagonal, then an SDP specializes to a linear program. There is another set of Karush-Kuhn-Tucker conditions found in [2] that state that any optimal solution $\mathbf{X}$ to (2.5) must satisfy the following: For $\lambda \geq 0$, there exists $y_i \in \mathbb{R},\ i = 1,\dots,d$

$$\begin{aligned} & C - \lambda\mathbf{X}^{-1} - \textstyle\sum_{i=1}^{d} y_i A_i = 0 \\ & \mathrm{Tr}(A_i\mathbf{X}) = b_i,\ 1,\dots,d \\ & \mathbf{X} \succeq 0. \end{aligned} \tag{2.6}$$

One should note the similarity between these conditions and the KKT conditions for linear programs. Because of this, one can construct another *elimination method* to acquire explicit algebraic equations for the central curve. The method is to start by letting

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ x_{1,2} & x_{2,2} & \ddots & x_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ x_{1,n} & x_{2,n} & \dots & x_{n,n} \end{bmatrix}$$

where it is worth nothing that we have $\binom{n+1}{2}$ many variables because $\mathbf{X}$ is symmetric. Then, starting with

$$C - \lambda \mathbf{X}^{-1} = \sum_{i=1}^{m} y_i A_i$$

from 2.6, we can multiply both sides with $\mathbf{X}$ and get

$$C\mathbf{X} - \lambda \mathrm{I} = \sum_{i=1}^{d} y_i A_i \mathbf{X}$$

where I represents the $n \times n$ identity matrix. This means that

$$C\mathbf{X} - \sum_{i=1}^{d} y_i A_i \mathbf{X} - \lambda \mathrm{I} = 0$$

which translates directly to $\binom{n}{2}$ explicit equations, each in the entry of the resultant matrix, i.e. the entries of the $n \times n$ symmetric matrix $(C\mathbf{X} - \sum_{i=1}^{m} y_i A_i \mathbf{X} - \lambda \mathrm{I})$ are equations in terms of the variables $\mathbf{X}_{ij}, y_i, \lambda \in \mathbb{R}, 1, \leq i \leq j \leq n$, and each of these equations must be equal to 0. Thus, we can construct an ideal by:

$$ \tag{2.7} $$

1. Form an ideal I generated by each of these equations.
2. We can then project I onto $\mathbb{R}[X_{ij}]$ by eliminating the y_i and λ variables from our ideal. Let the resultant ideal be I_s
3. We would like to see non-singular solutions, so we will saturate with respect to the determinant of $\mathbf{X}$, or formally, we define $I_{se} = I_s : (det\mathbf{X})^\infty$.
4. The variety of I_{se} gives us the central sheet of our SDP.

Let us see a simple example of this method for a randomly generated example with 2 constraints.

Example 2.3. Consider an SDP of the form (2.5) where

$$C = \begin{bmatrix} 194 & 136 & 74 \\ 136 & 146 & 80 \\ 74 & 80 & 54 \end{bmatrix},\ A_1 = \begin{bmatrix} 170 & 91 & 41 \\ 91 & 50 & 19 \\ 41 & 19 & 17 \end{bmatrix},\ b_1 = 5,\ A_2 = \begin{bmatrix} 101 & 25 & 41 \\ 25 & 30 & 54 \\ 41 & 54 & 98 \end{bmatrix},\ \text{and } b_2 = 6.$$

We begin by forming the $n \times n$ symmetric matrix from the equations in the KKT conditions.

There are 6 separate equations to form the ideal I:

$$-170x_{11}y_1 - 91x_{12}y_1 - 41x_{13}y_1 - 101x_{11}y_2 - 25x_{12}y_2 - 41x_{13}y_2 + 194x_{11} + 136x_{12} + 74x_{13} - \lambda,$$
$$-91x_{11}y_1 - 50x_{12}y_1 - 19x_{13}y_1 - 25x_{11}y_2 - 30x_{12}y_2 - 54x_{13}y_2 + 136x_{11} + 146x_{12} + 80x_{13},$$
$$-41x_{11}y_1 - 19x_{12}y_1 - 17x_{13}y_1 - 41x_{11}y_2 - 54x_{12}y_2 - 98x_{13}y_2 + 74x_{11} + 80x_{12} + 54x_{13},$$
$$-91x_{12}y_1 - 50x_{22}y_1 - 19x_{23}y_1 - 25x_{12}y_2 - 30x_{22}y_2 - 54x_{23}y_2 + 136x_{12} + 146x_{22} + 80x_{23} - \lambda,$$
$$-41x_{12}y_1 - 19x_{22}y_1 - 17x_{23}y_1 - 41x_{12}y_2 - 54x_{22}y_2 - 98x_{23}y_2 + 74x_{12} + 80x_{22} + 54x_{23},$$
$$-41x_{13}y_1 - 19x_{23}y_1 - 17x_{33}y_1 - 41x_{13}y_2 - 54x_{23}y_2 - 98x_{33}y_2 + 74x_{13} + 80x_{23} + 54x_{33} - \lambda.$$

Next, we eliminate the variables y_1, y_2, and λ from I to obtain I_s. For the readers confirmation, the leading generator (in a reverse lex ordering) is

$$37569x_{13}^2 - 73511x_{13}x_{22} + 73511x_{12}x_{23} + 41709x_{13}x_{23} + 29006x_{23}^2 - 37569x_{11}x_{33} - 41709x_{12}x_{33} - 29006x_{22}x_{33}.$$

Next, we saturate I_s with respect to $\det \mathbf{X}$, giving us the ideal I_{se}, which has generators:

$$37569x_{13}^2 - 73511x_{13}x_{22} + 73511x_{12}x_{23} + 41709x_{13}x_{23} + 29006x_{23}^2 - 37569x_{11}x_{33} - 41709x_{12}x_{33} - 29006x_{22}x_{33},$$

$$37569x_{12}x_{13} + 86935x_{13}x_{22} - 37569x_{11}x_{23} - 86935x_{12}x_{23} + 5434x_{13}x_{23} - 13859x_{23}^2 - 5434x_{12}x_{33} + 13859x_{22}x_{33},$$

$$1411429761x_{11}x_{13} + 423395045x_{13}x_{22} + 4833026436x_{11}x_{23} + 4691762058x_{12}x_{23} + 3830919508x_{13}x_{23} + 1502847661x_{22}x_{23} + 2316651587x_{23}^2 + 3874716384x_{11}x_{33} + 5521419691x_{12}x_{33} + 3691204342x_{13}x_{33} + 1944349255x_{22}x_{33} + 3228654814x_{23}x_{33} + 740948341x_{33}^2,$$

$$5367x_{12}^2 - 5367x_{11}x_{22} - 15510x_{13}x_{22} + 15510x_{12}x_{23} - 7351x_{13}x_{23} - 731x_{23}^2 + 7351x_{12}x_{33} + 731x_{22}x_{33},$$

$$1411429761x_{11}x_{12} + 4833026436x_{11}x_{22} + 5115157103x_{12}x_{22} + 5521419691x_{13}x_{22} + 1502847661x_{22}^2 + 3874716384x_{11}x_{23} + 3830919508x_{12}x_{23} + 5866952711x_{13}x_{23} + 4261000842x_{22}x_{23} + 3366769961x_{23}^2 - 2175748369x_{12}x_{33} - 138115147x_{22}x_{33} + 740948341x_{23}x_{33},$$

$$\begin{aligned}
&7575143527287x_{11}^2 - 61366930804311x_{11}x_{22} - 85939112368889x_{12}x_{22} \\
&\quad - 97273932704147x_{13}x_{22} - 27618904928812x_{22}^2 - 92222856805623x_{11}x_{23} \\
&\quad - 127097156761291x_{12}x_{23} - 146195399752309x_{13}x_{23} \\
&\quad - 100449984097592x_{22}x_{23} - 90873679540929x_{23}^2 \\
&\quad - 37278270436518x_{11}x_{33} - 42107002289569x_{12}x_{33} - 50408480399069x_{13}x_{33} \\
&\quad - 31241985135175x_{22}x_{33} - 61187014468044x_{23}x_{33} - 10916921156768x_{33}^2,
\end{aligned}$$

This illustrates how to calculate the generators for I_{se}.

Chapter 3

Linear Programming Results

A specific linear optimization problem is the transportation problem, where one is given a complete bipartite graph. This graph has one part with n nodes and another with m nodes. Every node in the part with n nodes is connected with every node in the part with m nodes, but there is no connection between any two nodes of the same part.

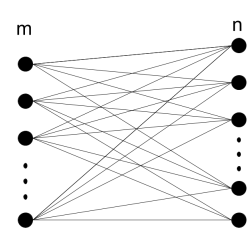

Figure 3.1: The graph of the complete bipartite graph $K_{m,n}$.

This can be thought of as a delivery system, where there are n warehouses with a given supply and m houses with a given demand to be delivered to. If the cost of delivery from each warehouse to each specific home is fixed, what would be the most cost-efficient way to deliver to all the homes? While there are many algorithms that compute optimal solutions to this problem quickly, it is in the interest of mathematical soundness to know the complexity of such problems. It is known that the upper bound for the curvature of the central path of a generic linear program with n decision variables and d linear constraints is $\binom{n-1}{d}$, as stated in (2.6) and [8], so we will establish the upper bound to the curvature of the central path of the transportation problem over the graph $K_{2,n}$ and $K_{m,n}$ by computing the degree of its central sheet.

An example over the graph $K_{2,4}$ will demonstrate a few critical components to our proof.

Example 3.1. Consider the following transportation problem over $K_{2,4}$:

$$\begin{aligned} \text{Minimize} \quad & \mathbf{c}_e^T \mathbf{x} \\ \text{subject to} \quad & A_e \mathbf{x} = \mathbf{b}_e \\ & \mathbf{x} \geq 0 \end{aligned} \tag{3.1}$$

where

$$\mathbf{x} = \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \\ x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \end{bmatrix}, \mathbf{c}_e = \begin{bmatrix} 1 \\ \frac{4}{9} \\ \frac{1}{5} \\ \frac{1}{2} \\ \frac{7}{2} \\ \frac{9}{5} \\ \frac{7}{6} \\ \frac{1}{5} \end{bmatrix}, A_e = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and } b_e = \begin{bmatrix} \frac{179}{24} \\ \frac{655}{168} \\ 1 \\ \frac{20}{3} \\ \frac{19}{14} \end{bmatrix}.$$

Note that the bottom row of the edge-incidence matrix of $K_{2,4}$ is removed to make the matrix full rank. This is a preference of the author, and not critical to the computation of the ideal.

First, we construct our matrix $\binom{c}{A}_e$, and note that $\text{rank}(A) = 5$, thus $\text{rank}\binom{c}{A}_e = 6$ as $\mathbf{c}_e$ is generic.

Now we consider the circuits of our matrix $\binom{c}{A}_e$. Because A_e is not generic, we can see that not all circuits are represented by a submatrix whose rank is equal to the rank of $\binom{c}{A}_e$ e.g. the circuit $1,2,3,5,6,7$ has rank 5. However, for this example problem the only circuits without rank 6 are of a similar form.

Inputting our example into the algebraic geometry software *Macaulay2*, we find that the ideal for the central sheet has generators:

$$35x_{12}x_{13}x_{14}x_{22}x_{23} - 149x_{12}x_{13}x_{14}x_{22}x_{24} + 114x_{12}x_{13}x_{14}x_{23}x_{24} - 35x_{12}x_{13}x_{22}x_{23}x_{24} + 149x_{12}x_{14}x_{22}x_{23}x_{24} - 114x_{13}x_{14}x_{22}x_{23}x_{24}$$

$$23x_{11}x_{13}x_{14}x_{21}x_{23} - 42x_{11}x_{13}x_{14}x_{21}x_{24} + 19x_{11}x_{13}x_{14}x_{23}x_{24} - 23x_{11}x_{13}x_{21}x_{23}x_{24} + 42x_{11}x_{14}x_{21}x_{23}x_{24} - 19x_{13}x_{14}x_{21}x_{23}x_{24}$$

$$103x_{11}x_{12}x_{14}x_{21}x_{22} - 252x_{11}x_{12}x_{14}x_{21}x_{24} + 149x_{11}x_{12}x_{14}x_{22}x_{24} - 103x_{11}x_{12}x_{21}x_{22}x_{24} + 252x_{11}x_{14}x_{21}x_{22}x_{24} - 149x_{12}x_{14}x_{21}x_{22}x_{24}$$

$$103x_{11}x_{12}x_{13}x_{21}x_{22} - 138x_{11}x_{12}x_{13}x_{21}x_{23} + 35x_{11}x_{12}x_{13}x_{22}x_{23} - 103x_{11}x_{12}x_{21}x_{22}x_{23} + 138x_{11}x_{13}x_{21}x_{22}x_{23} - 35x_{12}x_{13}x_{21}x_{22}x_{23}$$

It is worth noting that these generators come from the non-full rank circuits mentioned above. A computation in *Macaulay2* shows that the degree is 17.

3.1 Computing the Degree Through the Tutte Polynomial

A helpful method for computing the degree of the central sheet for such a problem, suggested in [8], is through the use of the Tutte polynomial.

For a general transportation problem, we can construct the node incident matrix A from the graph $K_{n,m}$, where each edge e_{ij} is given a column $A_{e_{ij}}$ with a 1 in the ith row representing the edge's exiting node, and a 1 in the $(n+j)$th row representing the edge's entering node. An example of the $K_{2,n}$ incidence matrix is

$$A = \begin{bmatrix} \overbrace{1 \quad 1 \quad 1 \quad \dots \quad 1}^{\text{n}} & 0 & 0 & 0 & \dots & 0 \\ 0 \quad 0 \quad 0 \quad \dots \quad 0 & 1 & 1 & 1 & \dots & 1 \\ 1 \quad 0 \quad 0 \quad \dots \quad 0 & 1 & 0 & 0 & \dots & 0 \\ 0 \quad 1 \quad 0 \quad \dots \quad 0 & 0 & 1 & 0 & \dots & 0 \\ \quad \vdots \quad 0 \quad \vdots \qquad\quad & & \vdots & & \ddots & 0 \\ 0 \quad 0 \quad 0 \quad \dots \quad 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}. \tag{3.2}$$

We will use the structure of (3.2) in an upcoming proof. However, before we begin the proof, it is important for one to understand some of the working pieces.

To compute the upper bound on the curvature of the central path of the transportation problem on $K_{2,n}$, this paper follows in the direction of [8]. In order to find which components of the easy to construct polynomial equations describe the central curve, we are interested in which components do not lay in the coordinate hyperplanes. These components lay in the bounded regions formed from the hyperplane arrangement

$$(\cup_{i=1,\ldots,n}\{x_i = 0\}) \cap \{\mathbf{x} | \binom{c}{A} \mathbf{x} = \binom{c_0}{\mathbf{b}},\ i = 1, \ldots, 2n\}. \tag{3.3}$$

The vertices of this arrangement are in bijection with the bases of matroid represented by $\binom{c}{A}$. Thus, following work in [3], we compute the number of bounded regions in (3.3), each of which gives a feasibility region of (2.1).

As stated earlier in section 2.1.2, from [12] it is known that the number of bounded regions formed from such hyperplanes is equal to the absolute value of the Möbius number on the poset of the intersection lattice formed from the given hyperplanes. [8] suggests one method to compute this is through the Tutte polynomial, specifically, we want to compute $T_{\binom{c}{A}}(1,0)$.

Through this machinery, this paper is able to establish an upper bound to the curvature of the central curve of the transportation problem on $K_{2,n}$. By constructing the matrix whose independence structure allows us to combinatorially count the number of bounded regions, we are able to get a handle on the upper bound of curvature for the central curve of the problem at hand.

3.2 Transportation problem over $K_{2,n}$

Theorem 3.1. *For a generic linear objective vector* $\mathbf{c} \in \mathbb{R}^{2n}$*, for the matrix A associated to* $K_{2,n}$ *as in* (3.2)*, and a constraint vector* $\mathbf{b}$ *such that* $A\mathbf{x} = \mathbf{b}$ *is feasible, the degree of the central curve of the transportation problem over* $K_{2,n}$ *is* $(n-2)2^{n-1}+1$.

Proof. Let $\binom{c}{A}$ represent the matrix (3.2) with the objective vector $\mathbf{c}$ appended as the top row. The matrix $\binom{c}{A}$ has rank $n+2$, as can be seen from the rowrank$(A) = n+1$ and the random cost vector $\mathbf{c}$ contributing 1 to the rank

In the evaluation of $T_{\binom{c}{A}}(1,0)$, we have

$$T_{\binom{c}{A}}(1,0) = \sum_{B \subset \binom{c}{A}} (0)^{r\binom{c}{A} - r(B)} (-1)^{|B| - r(B)}$$

We see that the only components that will contribute to computations are submatrices B where $r(B) = r\binom{c}{A} = n+2$. Any other type of submatrix will contribute a 0 term.

To this end, we take the following combinatorial approach: We find

$$T_{\binom{c}{A}}(1,0) = \sum_{\substack{B \subset \binom{c}{A} \\ r(B)=n+2}} (-1)^{|B| - r(B)}$$

by evaluating the sum over submatrices B such that $|B| = n+i$ and $r(B) = n+2$.

Each column of $\binom{c}{A}$ contributes uniquely to the row rank simultaneously in two possible places: corresponding to an edge's exiting and entering node in the bipartite graph.

For fixed $(n+i)$, we need to have at least one coefficient from each row chosen. This is bijective to the following counting problem: Having n urns with 2 colored balls in each urn with the condition that each ball is unique – e.g. a red ball chosen from the first urn is considered different than a red ball chosen from the second urn. The bijection is that the ith urn represents which $(2+i)$th row in $\binom{c}{A}$ we are choosing a coefficient from, and the color in each urn represents which of the first 2 rows we are choosing a coefficient from. We need to know how many ways can we select $n+i$ balls making sure we have chosen at least 1 ball of each color and 1 ball out of each urn.

To do this, we first choose the extra i balls out of separate urns, of which there are $\binom{n}{i}$ choices. Then, we choose the n remaining balls, one out of each urn. For every urn that has not had a ball chosen already, there are 2 choices remaining, thus we have 2^{n-i} choices total as the balls that already have been chosen from the initial i urns will not have a choice as to which ball to choose (there is only 1 ball remaining!).

So if we are to construct a full rank matrix from this construction, we must note that for each i, $|B| = n+i$ and $r(B) = n+2$ and so $|B| - r(B) = 2+i$.

Thus, we have

$$T_{\binom{c}{A}}(1,0) = \sum_{B \subset \binom{c}{A}} (-1)^{|B|-r(B)} = \sum_{i=2}^{n} (-1)^{2+i} \binom{n}{i} 2^{n-i} = \sum_{i=2}^{n} (-1)^{i} \binom{n}{i} 2^{n-i}$$

which is a binomial expansion without the first 2 terms, so then we have

$$T_{\binom{c}{A}}(1,0) = \sum_{i=2}^{n}(-1)^i\binom{n}{i}2^{n-i} = (2-1)^n + \underbrace{n2^{n-1}}_{i=1} - \underbrace{2^n}_{i=0} = (n-2)2^{n-1}+1.$$

Thus

$$|\mu_{\binom{c}{A}}| = T_{\binom{c}{A}}(1,0) = (n-2)2^{n-1}+1$$

□

These tools can be reused to find a formula for the upper bound of curvature on the central curve on the general bipartite matching problem over $K_{m,n}$. However the construction will be slightly different.

3.3 Transportation problem over $K_{m,n}$

Theorem 3.2. *For a generic linear objective function* $\mathbf{c} \in \mathbb{R}^{mn}$ *and, for a given A as in* (2.1)*, a* $\mathbf{b}$ *such that* $Ax = \mathbf{b}$ *is feasible, then the upper bound on the curvature of the central path in the bipartite transportation problem on the graph* $K_{m,n}$ *is*

$$\sum_{i=0}^{mn-m-n}(-1)^i\begin{pmatrix}\text{The number of}\\ \text{compositions of}\\ m+n+i \text{ into}\\ \text{exactly } m \text{ parts}\\ \text{each part less}\\ \text{than } n\end{pmatrix}\left(\sum_{\substack{v_1+\ldots+v_n=m+n+i\\ 1\le v_j\le m}}\frac{(m+n+i)!}{v_1!v_2!\ldots v_n!}\right)$$

Proof. The incidence matrix for the bipartite graph will represent the constraint matrix A.

For a general bipartite graph $K_{m,n}$ we have

$$A = \begin{array}{c} m\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \\ n\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \end{array} \left[\begin{array}{cccc|cccc|c|cccc} \multicolumn{4}{c|}{\overbrace{1 \quad 1 \quad \dots \quad 1}^{n}} & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\ & & \vdots & & & & \vdots & & \vdots & & & \vdots & \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \\ \hline 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 & 1 & \dots & 0 \\ & & \ddots & & & & \ddots & & \ddots & & & \ddots & \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 0 & 0 & \dots & 1 \end{array}\right] \tag{3.4}$$

We begin again by constructing the matrix $\binom{c}{A}$ by adding the objective vector as the top row to the constraint matrix A. However, we will again only be concerned with the structure of A as, for a generic $\mathbf{c}$, the row $\mathbf{c}$ will contribute once to the rank of $\binom{c}{A}$.

In order to contribute to the rank of A, a row must have a value in at least 1 entry. We see that $\text{rank}(A) = m+n-1$, so then one can conclude that $\text{rank}\binom{c}{A} = m+n$. Again we wish to compute $T_{\binom{c}{A}}(1,0)$ and will do so by considering all submatrices of $\binom{c}{A}$ with full rank, i.e., submatrices whose rank is $m+n$.

In an effort to construct all full rank submatrices, we will consider the sum of all entries in a row of a submatrix. If a matrix is full rank, then the sum of entries in each row must be greater than or equal to 1.

Let u_i be the sum of the entries in the ith row of a given submatrix of A, and v_i be the sum of the entries in the $(m+i)$th row of a given submatrix of A. Due to the structure of A, we have natural constraints on the integer values of u_i and v_i. This can be visualized through a columns selection process in the following:

$$A = \begin{array}{c} m\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \\ n\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \end{array} \left[\begin{array}{cccc|cccc|c|cccc} \overbrace{1 \quad 1 \quad \ldots \quad 1}^{n} & & & & 0 & 0 & \ldots & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 1 & 1 & \ldots & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & & \vdots & & & & \vdots & & \vdots & & & \vdots & \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots & 1 & 1 & \ldots & 1 \\ \hline 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & 1 & \ldots & 0 & \ldots & 0 & 1 & \ldots & 0 \\ & & \ddots & & & & \ddots & & \ddots & & & \ddots & \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 1 & \ldots & 0 & 0 & \ldots & 1 \end{array}\right] \begin{array}{ccc} 1 \leq & u_1 & \leq n \\ 1 \leq & u_2 & \leq n \\ & \vdots & \\ 1 \leq & u_m & \leq n \\ 1 \leq & v_1 & \leq m \\ 1 \leq & v_2 & \leq m \\ & \vdots & \\ 1 \leq & v_n & \leq m \end{array} \tag{3.5}$$

To get a numerical handle on $T_{\binom{c}{A}}(1,0)$, we again consider a submatrix B where $|B| = m+n+i$. To construct a full rank submatrix, we can then see that

$$u_1 + u_2 + \ldots + u_m = m+n+i = v_1 + v_2 + \ldots + v_n$$

for any given submatrix B. If we let $\{u_1, u_2, \ldots, u_m\}$ be a composition of $m+n+i$ into exactly m parts with each part less than or equal to n, and let $\{v_1, v_2, \ldots, v_n\}$ be a composition of $m+n+i$ into exactly n parts with each part less than or equal to m, then one can construct unique full rank submatrices of $\binom{c}{A}$ through the following column selection process:

1. Each u_i determines the number of columns we will select in the ith column block of length n as seen in (3.5)

2. To construct the $(m+1)$th row of our submatrix, we see that we have $m+n+i$ columns to choose from to place v_1 1's. So there are $\binom{m+n+i}{v_1}$ choices.

3. For v_2, we choose which of the remaining $m+n+i-v_1$ columns to place v_2 1's, or $\binom{m+n+i-v_1}{v_2}$ choices.

4. We continue this process, where for each v_j for $1 \leq j \leq i$, we choose which of the $m+n+i-\sum_{k=1}^{i} v_k$ remaining columns to place v_i 1's on the $(m+i)$th row.

For all the v_i terms, we have

$$\binom{m+n+i}{v_1}\binom{m+n+i-v_1}{v_2}\binom{m+n+i-v_1-v_2}{v_3}\cdots\binom{v_{n-1}+v_n}{v_n} = \frac{(m+n+i)!}{v_1!\,v_2!\,v_3!\ldots v_n!}$$

Due to this construction, one ensures that each submatrix is full rank, so then only these terms will contribute to our evaluation of $T_{\binom{c}{A}}(1,0)$. We can then see that

$$\mu_{\binom{c}{A}} = T_{\binom{c}{A}}(1,0) = \sum_{i=0}^{mn-m-n} (-1)^i \begin{pmatrix} \text{The number of} \\ \text{compositions of} \\ \text{m+n+i into} \\ \text{exactly m parts} \\ \text{each part less} \\ \text{than } n \end{pmatrix} \left(\sum_{\substack{v_1+\ldots+v_n=m+n+i \\ 1\leq v_j\leq m}} \frac{(m+n+i)!}{v_1!\,v_2!\ldots v_n!} \right)$$

□

Chapter 4

Central Curve in Quadratic and Semidefinite Programming

In this chapter our intended goal is to extend the method and argument used in the linear case to quadratic and semidefinite programs. That is, we first attempt to find a new way to create and explore the equations that describe the central curve in quadratic and semidefinite programming as well. We prove half of a conjecture as well as mention our claim is supported by evidence. One of the primary reasons for this pursuit is that the elimination method does not necessarily illuminate any particular structure of the ideals. We will again note the similarity of the KKT conditions in linear programs and semidefinite programs. This similarity motivates the following constructive argument, which we will refer to as the *circuit method*.

4.1 Circuit Method in Quadratic Programming

Our goal in this section is to propose and prove half of the circuit method for quadratic programs as well as to give another proof of Theorem 4.4 in [10]. One can follow an argument similar to [8] and section 2.1.1 to understand the general direction of this section. We begin with the KKT conditions for quadratic programs from (2.4). Recall that they state that a $(\mathbf{x}, \lambda)$ is optimal to the log-barrier problem for a quadratic program if and only if there exists a $\mathbf{y} \in \mathbb{R}^d$ such that

$$\begin{aligned} (Q\mathbf{x})_i + c_i - \lambda x_i^{-1} - (\mathbf{y}^T A_i) = 0,\ \ i = 1, \ldots, n \\ A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq 0, \end{aligned} \tag{4.1}$$

Consider the condition

$$(Q\mathbf{x})_i - c_i - \lambda x_i^{-1} - (A^T\mathbf{y})_i \; i = 1, \ldots, n.$$

Then one can conclude that, for any optimal solution $\mathbf{x}$, there exists a $\mathbf{y}$ such that

$$(Q\mathbf{x})_i - c_i - (A^T\mathbf{y})_i = \lambda x_i^{-1} \; i = 1, \ldots, n.$$

These equations are equivalent for $\mathbf{x} \neq 0$ to the condition that $x^{-T} \in \text{span}\{A^T, Qx - c\}$, giving us similar conditions to that in the linear programming case in section 2.1.1. This again allows us to construct a $(d+2) \times n$ matrix

$$\begin{bmatrix} (Q\mathbf{x}-\mathbf{c})^T \\ A \\ \mathbf{x}^{-1} \end{bmatrix}. \tag{4.2}$$

One can argue that, because A, Q, and $\mathbf{c}$ are generic, the rank of the matrix (4.2) is $d+1$, and thus, every $(d+2)$-minor is equal to 0. This also supports the argument that, in the context of matroids, the matrix (4.2) is generic, and thus all circuits are any $(d+2)$ subset. We set to propose a new method for computing the central sheet as well as the central curve for quadratic programs:

1. Construct the matrix (4.2).

2. If we let G be a $(d+2)$ subset of $[n]$, we calculate the $(d+2)$-minor of (4.2) whose columns are corresponding to G, then multiply said determinant by $\prod_{i \in G} x_i$.

3. The ideal whose generators are formed from the above steps for all possible subsets G forms the ideal I_{qc}.

Observational evidence leads us to the following:

Conjecture 4.1. *The ideal I_{qc}, as constructed above, is equal to the ideal I_{qe} of polynomials vanishing on the central sheet for quadratic programs* (2.2). *The central curve of a generic quadratic program is given by the ideal $(I_{qc} + \langle A\mathbf{x} - \mathbf{b} \rangle)$.*

Every computation we were able to carry out show that these two ideals are equal, though we do not have a complete proof of Conjecture (4.1). However, we are able to prove the following.

Theorem 4.1. *Let the ideal I_{qe} be the ideal of polynomials vanishing on the central sheet of* (2.2) *and let I_{qc} be the ideal constructed above from the $(d+2)$-minors of matrix* (4.2)*. Then*

$$I_{qc} \subset I_{qe}.$$

Proof. Our goal is to show that the generators of I_{qc} are in I_{qe}. Let K be a given d-subset of $[n]$, K_i be the ith term in K, A_K be the corresponding $d \times d$ submatrix of A with columns in K, and A_{Ki} be the product of the diagonal terms of the matrix A_k with the ith row and ith column missing. Also let x_{Kj} be the product of all x_i where $i, j \in K$ excluding x_j e.g. if $K = \{1,2,3\}$ then $x_{K2} = x_1x_3$. By computing determinants, a given generator of I_{qc} is of the form

$$((Q\mathbf{x})_{K_1} + c_{K_1})A_{K1}x_{K1} - ((Q\mathbf{x})_{K_2} + c_{K_2})A_{K2}x_{K2} + \ldots \pm ((Q\mathbf{x})_{K_d} + c_{K_d})A_{Kd}x_{Kd}.$$

Our goal is to show that there exists a $\mathbf{y} \in \mathbb{R}^d$ such that (2.4) is satisfied. The first set of equations from (2.4) state

$$\begin{gathered}
(Q\mathbf{x})_1 + c_1 - \lambda x_1^{-1} - y_1 a_{11} - y_2 a_{21} - \ldots - y_d a_{d,1} = 0 \\
(Q\mathbf{x})_2 + c_2 - \lambda x_2^{-1} - y_1 a_{12} - y_2 a_{22} - \ldots - y_d a_{d,2} = 0 \\
\vdots \\
(Q\mathbf{x})_n + c_n - \lambda x_n^{-1} - y_1 a_{1n} - y_2 a_{2,n} - \ldots - y_d a_{d,n} = 0.
\end{gathered}$$

Let us multiply each of the $K_j th$ equations above with the term $A_{Kj} x_{Kj}$, then add the result of these products together along with the equations indexed by terms outside K. By combining like terms, we arrive at the following:

$$\begin{pmatrix}
((Q\mathbf{x})_{K_1} + c_{K_1}) A_{K1} x_{K1} + ((Q\mathbf{x})_{K_2} + c_{K_2}) A_{K2} x_{K2} + \ldots + ((Q\mathbf{x})_{K_d} + c_{K_d}) A_{Kd} x_{Kd} \\
-y_1 (a_{1K_1} A_{K1} x_{K1} - a_{1K_2} A_{K2} x_{K2} - \ldots - a_{1K_d} A_{Kd} x_{Kd} - \sum_{h \in [n] \setminus K} a_{1h}) \\
-y_2 (a_{2K_2} A_{K2} x_{K2} - a_{2K_2} A_{K2} x_{K2} - \ldots - a_{2K_d} A_{Kd} x_{Kd} - \sum_{h \in [n] \setminus K} a_{2h}) \\
\vdots \\
-y_d (a_{dK_d} x_{K1} A_{K1} - a_{dK_2} x_{K2} A_{K2} - \ldots - a_{dK_d} x_{Kd} A_{Kd} - \sum_{h \in [n] \setminus K} a_{2h}) \\
-\lambda x_{K_1}^{-1} A_{K1} x_{K1} - \lambda x_{K_2}^{-1} A_{K2} x_{K2} - \ldots - \lambda x_{K_d}^{-1} A_{Kd} x_{kd} - \sum_{h \in [n] \setminus K} \lambda x_h^{-1} \\
- \sum_{h \in [n] \setminus K} (Q\mathbf{x})_h + c_h
\end{pmatrix} = 0.$$

For $\lambda > 0$, any optimal solution to the log-barrier problem of a quadratic program will have all $x_i \neq 0$. Thus, we can think of the above as a linear equation in $\mathbb{R}[x_i, \lambda, \; i = 1 \ldots n][y_i, \; i = 1, \ldots, d]$. Thus, because this is a linear equation, there exists a $\mathbf{y} \in \mathbb{R}^d$ such that

$$\begin{pmatrix} -y_1(a_{1K_1}A_{K1}x_{K1} - a_{1K_2}A_{K2}x_{K2} - \ldots - a_{1K_d}A_{Kd}x_{Kd} - \sum_{h\in[n]\setminus K} a_{1h}) \\ -y_2(a_{2K_2}A_{K2}x_{K2} - a_{2K_2}A_{K2}x_{K2} - \ldots - a_{2K_d}A_{Kd}x_{Kd} - \sum_{h\in[n]\setminus K} a_{2h}) \\ \vdots \\ -y_d(a_{dK_d}x_{K1}A_{K1} - a_{dK_2}x_{K2}A_{K2} - \ldots - a_{dK_d}x_{Kd}A_{Kd} - \sum_{h\in[n]\setminus K} a_{2h}) \\ -\lambda x_{K_1}^{-1}A_{K1}x_{K1} - \lambda x_{K_2}^{-1}A_{K2}x_{K2} - \ldots - \lambda x_{K_d}^{-1}A_{Kd}x_{kd} - \sum_{h\in[n]\setminus K} \lambda x_h^{-1} \\ -\sum_{h\in[n]\setminus K}(Q\mathbf{x})_h + c_h \end{pmatrix} = 0$$

which implies that there exists a $\mathbf{y} \in \mathbb{R}^d$ such that

$$((Q\mathbf{x})_{K_1} + c_{K_1})A_{K1}x_{K1} + ((Q\mathbf{x})_{K_2} + c_{K_2})A_{K2}x_{K2} + \ldots + ((Q\mathbf{x})_{K_d} + c_{K_d})A_{Kd}x_{Kd} = 0$$

which implies that, because the above is the generator from I_{qc}, then all $\mathbf{x} \in V(I_{qc})$ will vanish on the Zariski closure of points satisfying the KKT conditions. □

Utilizing some of the tools in this proof, we are able to make a parallel proof to [10][Lemma 4.2], which shows

Lemma 4.2. *The central curve of a generic quadratic program* (2.2) *for the case where Q is diagonal, the ideal describing the central sheet I_{qe} has an initial ideal (with respect to reverse lex) with generators of the form*

$$x_{k_1}^2 x_{k_2} \ldots x_{k_{d+1}}$$

where K is a $(d+2)$ subset of $[n]$ with $K = \{k_1 < k_2 < \ldots < k_{d+1} < k_{d+2}\}$.

We will use this lemma to prove the following theorem:

Theorem 4.3. *For a generic quadratic program of the form* (2.2) *where Q is diagonal,*

$$I_{qe} = I_{qc}.$$

Proof. The goal of this proof is to show that the initial ideal of I_{qe} and the initial ideal of I_{qc} are equal. We do this by constructing the leading terms of the generators for I_{qc} in the diagonal case, and showing they are the same generators as [10][Lemma 4.2]. Let K be a $(d+2)$ subset of $[n]$. Because Q is diagonal, the entries of the top row of (4.2) will be linear in terms of a single variable. Then the submatrix of (4.2) corresponding to K is of the form

$$\begin{bmatrix} q_{k_1}x_{k_1} - c_{k_1} & q_{k_2}x_{k_2} - c_{k_2} & \cdots & q_{k_{d+2}}x_{k_{d+2}} - c_{k_{d+2}} \\ A_{k_1} & A_{k_2} & \cdots & A_{k_{d+2}} \\ \frac{1}{x_{k_1}} & \frac{1}{x_{k_2}} & \cdots & \frac{1}{x_{k_{d+2}}} \end{bmatrix}$$

and through genericity, this is the same as multiplying the last row of the above matrix by $x_{k_1}x_{k_2}\ldots x_{k_{d+2}}$, which gives us, using the notation from above, where $x_{Ki} = \prod_{j\in K\setminus i} x_{k_j}$

$$\begin{bmatrix} q_{k_1}x_{k_1} - c_{k_1} & q_{k_2}x_{k_2} - c_{k_2} & \cdots & q_{k_{d+2}}x_{k_{d+2}} - c_{k_{d+2}} \\ A_{k_1} & A_{k_2} & \cdots & A_{k_{d+2}} \\ x_{K1} & x_{K2} & \cdots & x_{K(d+2)} \end{bmatrix}$$

which, in reverse lex, would read the leading term from the diagonal of such a matrix. Thus, the leading term comes from

$$(q_{k_1}x_{k_1} - c_{k_1})(\prod_{j\in K\setminus\{k_1,k_{d+2}\}} A_{j,k_j})x_{K(d+2)}$$

and ignoring the lower degree term coming from c_{k_1}, we have the leading term

$$= \underbrace{(q_{k_1} \prod_{j\in K\setminus\{k_1,k_{d+2}\}} A_{j,k_j})}_{\text{constant}} x_{k_1}x_{K(d+2)} = (q_{k_1} \prod_{j\in K\setminus\{k_1,k_{d+2}\}} A_{j,k_j})x_{k_1}^2 x_{k_2}x_{k_3}\ldots x_{k_{d+1}}$$

giving us the leading term of the form

$$x_{k_1}^2 x_{k_2}x_{k_3}\ldots x_{k_{d+1}}.$$

□

4.2 The Circuit Method in Semidefinite Programming

This section will follow the footsteps of the previous section, but in the field of semidefinite programming. We begin with the following lemma:

Lemma 4.4. *An* $\mathbf{X}$ *satisfies the KKT conditions* (2.6) *to the program* (2.5) *if and only if* $\mathbf{X}^{-1}$ *is in span*$\{C, A_i, i = 1, \ldots, d\}$.

Proof. Let $\mathbf{X}$ be an optimal solution to 2.5. Then $\mathbf{X}$ must satisfy all conditions in 2.6. Thus, one can see

$$C - \sum_{i=1}^{n} y_i A_i = \lambda \mathbf{X}^{-1}$$

which, because $\lambda \geq 0$, we have that $\mathbf{X}^{-1} \in \text{span}\{A_i, C\}$. □

However, in order to obtain an algebraic handle on such possible $\mathbf{X}$, we consider the entries of $\mathbf{X}^{-1}$. If we let $X_{i,j}$ denote the entry in the ith row, jth column of a matrix X and let X^{ij} denote the matrix X with row i and column j removed, then linear algebra tells us that

$$\mathbf{X}_{i,j}^{-1} = \frac{1}{\det(\mathbf{X})} \det(\mathbf{X}^{ji})$$

and because $\mathbf{X} \succeq 0$, then $\det(\mathbf{X}) \neq 0$, so then we can further conclude that $\det(\mathbf{X})\mathbf{X}^{-1} \in \text{span}\{A_i, C\}$. Because of this, we then construct a particular matrix, analogous to our matrix $\binom{c}{A}$ in linear program, by taking the matrix C, A_i, and $\det(\mathbf{X})\mathbf{X}^{-1}$ and flattening their rows then addend them vertically. That is, we are making an obvious flattening map $F : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$, and then constructing a matrix where the rows are the image of our matrices C, A_i, and $\mathbf{X}^{-1}$ under F. More explicitly we define,

$$\mathbf{M} = \begin{bmatrix} F(C)^T \\ F(A_1)^T \\ \vdots \\ F(A_m)^T \\ F(\det(\mathbf{X})\mathbf{X}^{-1})^T \end{bmatrix}. \tag{4.3}$$

Example 4.1 (continued from 2.3)**.** Our example from before would have us construct the matrix

$$\mathbf{M}_e = \begin{bmatrix} 194 & 136 & 74 & 136 & 146 & 80 & 74 & 80 & 54 \\ 170 & 91 & 41 & 91 & 50 & 19 & 41 & 19 & 17 \\ 101 & 251 & 41 & 25 & 30 & 54 & 41 & 54 & 98 \\ \multicolumn{9}{c}{F(\det(\mathbf{X})\mathbf{X}^{-1})} \end{bmatrix}$$

where

$$F(\det(\mathbf{X})\mathbf{X}^{-1}) = \begin{bmatrix} -x_{23}^2 + x_{22}x_{33} \\ x_{13}x_{23} - x_{12}x_{33} \\ -x_{13}x_{22} + x_{12}x_{23} \\ x_{13}x_{23} - x_{12}x_{33} \\ -x_{13}^2 + x_{11}x_{33} \\ x_{12}x_{13} - x_{11}x_{23} \\ -x_{13}x_{22} + x_{12}x_{23} \\ x_{12}x_{13} - x_{11}x_{23} \\ -x_{12}^2 + x_{11}x_{22} \end{bmatrix}^T$$

It is worth noting in our computations that $F(\det(\mathbf{X})\mathbf{X}^{-1})$ does not need the determinant component as a divisor of each term. This is because we assume our $\mathbf{X}$ is such that $\det(\mathbf{X}) \neq 0$, which is critical for the following method.

Because $\det(\mathbf{X})\mathbf{X}^{-1} \in \text{span}\{A_i, C\}$, then the determinant of every $m+2$ minor of $\mathbf{M}$ is equal to 0. Following in the footsteps of [8] and their matroidal arguments, we assume that, for generic C and A_i, every $m+2$ minor of $\mathbf{M}$ is rank $m+1$ and thus the set of circuits on $\mathbf{M}$ is uniform in number of elements. Thus we can construct an ideal I_{sc} through the following steps:

1. Let L be a $m+2$ subset of $[n]$.

2. Compute the determinant of the $m+2$ submatrix of (4.3) with columns corresponding to L

3. Form an ideal I composed of all such determinants for all subsets L.

4. Let $I_{sc} = I : (\det \mathbf{X})^{\infty}$

Using this process, we propose another conjecture.

Conjecture 4.2. *For I_{se} as seen in* (2.7)*, and for I_{sc} as constructed above,*

$$I_{se} = I_{sc}.$$

Every computation we have performed confirms that the ideal formed from the elimination method and the circuit method are the same. This has been computationally helpful at times, as particular matrix sizes and number of constraints may compute ideals faster using different methods. However, an explicit proof of this conjecture has not yet been formulated.

4.3 Some Generic Cases

Following the same argument as in the linear programming case, it would be mathematically relevant to know the degrees of the varieties that describe such central paths. Some computations are shown below, where n represents the size of the matrices in question, and d is the number of constraints.

$d \backslash n$	1	2	3	4	5
2	1	1	1	1	1
3	2	4	4	2	1
4	3	9			
5	4	16			

Table 4.1: The degree of the central curve of a generic SDP with $\binom{n}{2}$ variables and d constraints.

From this data, we form two separate conjectures about the degree of semidefinite programs.

Conjecture 4.3. *For a generic SDP of the form* (2.5)*, the degree of the central curve exponential with base* $(d-1)$ *in terms of* n *until* $n = d$.

We also have

Conjecture 4.4. *The degree of the central curve for a generic SDP of the form* (2.5) *for a fixed* d *is symmetric about* $(n-1)$.

Both of these conjectures met computational limitations by the author.

Chapter 5

Conclusion

This thesis begins by exploring the necessary tools for understanding the central curve as an object in both optimization and algebra. Chapter 2 outlines current methods related to interior point algorithms and their computational complexity. It also outlines the core methods that are used for obtaining an algebraic handle on the central curve for linear, quadratic, and semidefinite programs. Each section shows the elimination method for each representative field of optimization which obtains the ideal defining the central curve as well as examples allowing a reader to follow and confirm their own computations. This chapter also reviews the circuit method for linear programs, first shown in [8].

Once these concepts are understood, Chapter 3 begins by exploring the degree of the central curve, specifically for the transportation problem over the graph $K_{2,n}$, which was shown to be

$$(n-2)2^{n-1}+1$$

as well as state the explicit degree for the generic transportation problem over the graph

$K_{m,n}$.

Chapter 4 then proposes a circuit method for a generic quadratic program as well as prove that $I_{qc} \subset I_{qe}$. We utilized Lemma 4.2 in [10] to prove that $I_{qc} = I_{qe}$ for the case where Q is diagonal. Then a proposed circuit method for semidefinite program is outlined as well as two proposed conjectures on the degree of the central curve of generic semidefinite programs.

The overall direction of this thesis began with understanding the central curve, first in linear programs, and then using our understanding of such methods in an attempt to extend them to more complex convex optimization problems in quadratic and semidefinite programs. While our results build on results from [10] and our computations show that the proposed circuit method does indeed accurately describe the central sheet, our proof is only sufficient for one direction of inclusion. The technical language used to describe the proof was a product of trying to prove equality and the author is believes that such a proof is possible but out of the scope of this thesis.

The proof for the equality $I_{se} = I_{sc}$ is also expected to be difficult because the inherent geometric structure of the space of feasible solutions, namely over the space $\{\mathbf{X} \in \mathbb{R}_{\succ 0} | \det \mathbf{X} \neq 0\}$. We expect this proof does not entirely follow the same arguments as seen in [9], which is interested in bounded regions formed from the intersection of hyperplanes in an affine space. Instead the author believes that, for a generic SDP, the degree is still related to the number of bounded regions formed by the intersection of the determinantal cone in the affine space formed by the constraints of an SDP.

Bibliography

[1] Anders Bjrner, *Homology and shellability of matroids and geometric lattices*, Matroid Applications (Neil White, ed.), Cambridge University Press, 1992, Cambridge Books Online, pp. 226–283.

[2] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.

[3] Thomas Brylawski and James Oxley, *The Tutte polynomial and its applications*, Matroid applications, Encyclopedia Math. Appl., vol. 40, Cambridge Univ. Press, Cambridge, 1992, pp. 123–225.

[4] Henry H. Crapo, *The Tutte polynomial*, Aequationes Math. **3** (1969), 211–229.

[5] David Eisenbud, *Commutative algebra*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995, With a view toward algebraic geometry.

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

[7] Joe Harris, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1995, A first course, Corrected reprint of the 1992 original.

[8] Jesús A. Loera, Bernd Sturmfels, and Cynthia Vinzant, *The central curve in linear programming*, Foundations of Computational Mathematics **12** (2012), no. 4, 509–540.

[9] Nicholas Proudfoot and David Speyer, *A broken circuit ring*, Beiträge Algebra Geom. **47** (2006), no. 1, 161–166.

[10] Dennis Schlief, *Degree for the central curve or quadratic programming*, Master's thesis, San Francisco State University, United States, 2014.

[11] Richard P. Stanley, *Combinatorics and commutative algebra*, second ed., Progress in Mathematics, vol. 41, Birkhäuser Boston, Inc., Boston, MA, 1996.

[12] Thomas Karl Zaslavsky, *Facing up to Arrangements: Face-count Formulas For Partitions of Space by Hyperplanes*, ProQuest LLC, Ann Arbor, MI, 1974, Thesis (Ph.D.)–Massachusetts Institute of Technology.