# DEGREE FOR THE CENTRAL CURVE OF QUADRATIC PROGRAMMING

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

Dennis Schlief

San Francisco, California

July 2014

CERTIFICATION OF APPROVAL

I certify that I have read *DEGREE FOR THE CENTRAL CURVE OF QUADRATIC PROGRAMMING* by Dennis Schlief 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 Arts in Mathematics at San Francisco State University.

Serkan Hosten
Professor of Mathematics

Matthias Beck
Professor of Mathematics

Federico Ardila
Professor of Mathematics

DEGREE FOR THE CENTRAL CURVE OF QUADRATIC PROGRAMMING

Dennis Schlief
San Francisco State University
2014

Our research begins with a standard quadratic programming problem from the field of optimization. One family of solution methods is called interior point methods. These methods trace a path through the feasible region until they converge to the optimal solution. This path, defined by polynomial equations, is a piece of the central curve. Recent work has been done studying the degree and total curvature of the central curve in the linear programming case. We broaden the scope and look to prove similar results in the quadratic case. Through computer experimentation we hypothesized that the degree can be considered just for the case that the objective function has a generic diagonal matrix. We prove the reduction to diagonal and consider the degree of the central curve only in this case. We proceed by constructing a monomial ideal resulting from the optimality conditions of our quadratic program. We prove the degree of this monomial ideal, and show that this degree is an upper bound for the degree of the quadratic central curve. While the proposed degree is in fact exact, we only prove an upper bound for this thesis. Our result has nice symmetry with the main result of its linear counterpart.

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

---

Chair, Thesis Committee Date

ACKNOWLEDGMENTS

I thank my advisor, Serkan Hosten. I would also like to thank my idols Goku, Leeroy Jenkins, Poseidon, and Tim Tebow.

TABLE OF CONTENTS

# LIST OF FIGURES

# Chapter 1

# Introduction

Much of the field of operations research is dedicated to solving optimization problems. Optimization problems are mathematical models that try to calculate the best possible outcome in a given scenario. A common example of this is given by airline schedules. Considering that any given airline is trying to maximize their profit, they obviously want to find the most effective way to have as many planes landing and taking off as possible. A mathematical model called a program is created which details the limitations of the airline and airport, and then this program is solved, yielding the most effective schedule an airline should follow in order to make profits. This is just one type of optimization problem, and the mathematics behind them is relatively new, making this field of study explode with new results and possibilities.

## 1.1 Central Curve in Quadratic Programming

One branch of optimization is quadratic programming. Quadratic programming problems form a subset of convex optimization problems, meaning that the objective function is convex, and the feasible region defined by the constraints is a convex set. In the general case, a quadratic programming problem is of the form

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^tQx + c^tx \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{1.1}$$

where $Q$ is an $n \times n$ symmetric positive-definite matrix, $A$ is a $d \times n$ matrix with real entries, $b$ is a vector in $\mathbb{R}^d$, and $c$ is a vector in $\mathbb{R}^n$. Here $n$ is the number of decision variables the problem has, while $d$ is the number of equality constraints to adhere to. Optimization problems tend to be large scale, so a lot of research has been conducted on finding the best solution method. When determining what solution method is best, one of the main factors considered is computational runtime. If a specific solving method is faster than the rest, then using it on large scale problems can become very efficient. One group of popular and relatively fast solution methods are called interior point methods (see 2.1.1).

One type of interior point method uses a barrier function. A barrier function is a new objective function that implies the nonnegativity conditions and employs a positive parameter $\lambda$. For each choice of $\lambda$, the new objective function has a unique solution. By taking a limit on $\lambda$, a solution is obtained for the original optimization problem.

This exact process is done using a logarithmic barrier function, which guarentees nonnegativity on the decision variables $x_i$. Given our general quadratic program, the new optimization problem is the following

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^tQx + c^tx - \lambda\sum_{i=1}^{n}\log(x_i) \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{1.2}$$

The interior point algorithm calculates the optimal solution starting when $\lambda$ is positive infinite. In practice the algorithm initially lets $\lambda$ be a large positive number. Here the logarithmic term is impactful and the original objective function is negligible, producing a solution in the interior of the feasible region that is far away from the coordinate hyperplanes defined by $x_i \geq 0$. Then $\lambda$ is iterated down to 0. As $\lambda$ gets closer to 0, the added $-\lambda\sum_{i=1}^{m}\log(x_i)$ terms become negligible, thus the optimal solution at this point is equal to the optimal solution of our original problem. As $\lambda$ changes from infinite to 0, each unique optimal solution creates a point in the interior of the original feasible region. The collection of these optimal solutions form a curve, known as the central path.

Each iteration is a Newton-type iteration. The complexity of the problem comes down to how many of these Newton-type iterations the interior point algorithm must make. The amount of iterations the algorithm needs to perform can be estimated using the total curvature of the central path. The curvature of the central path represents how curvy it is. An analogy is roads. A smooth road with few turns and bends could be driven in a short amount of time compared to a road that is windy, full of twists and turns. In linear programming, a subset of quadratic program where $Q$ is the zero matrix thus creating a linear objective function, an upper bound was found for the curvature of the central path [**?**, Theorem 2.1]. By knowing bounds on the curvature, one may estimate the computational runtime that these interior point methods may take, and one can use that information to decide which solution method to use.

Our research starts with a generic quadratic program, which is a program of the form (1.1) where $Q, A, c$, and $b$ have random real-valued entries. As we will show, the central path of a generic quadratic program can be defined using a system of polynomial equations (see equation (3.4)). The set of solutions to this system of equations forms an algebraic curve, called the central curve. The degree of the central curve provides an upper bound on the total curvature of the central path. By calculating the degree of the central curve we gain valuable information about the curvature, which illuminates aspects such as computational runtime. We are optimistic that our exploration into the degree of the central curve has paved a path that will lead to results for curvature in quadratic programming similar to those found in linear programming.

## 1.2 Results

In a paper by De Leora, Sturmfels, and Vinzant [**?**], the degree of the central curve of a generic linear program was computed to be $\binom{n-1}{d}$. We set out with the goal of deriving a formula for the degree of the central curve for a generic quadratic programming problem, which is the spiritual successor of their work. Following their work, we view a quadratic program and the central curve as a problem in algebraic geometry. Since we are working with quadratic programs with generic data $Q, A, c$, and $b$, it is expected that the degree of the central curve depends on $d$ and $n$

As mentioned before, we can define the central path in terms of polynomial equations. This is done using the Karush-Kuhn-Tucker (KKT) conditions [**?**]. The KKT conditions are first order necessary and sufficient conditions for a solution to be optimal. By writing the KKT conditions for (1.2), our conditions for optimality are

- $(Qx)_i + c_i - \frac{\lambda}{x_i} - (A^t y)_i = 0$, $i = 1, \ldots, n$,
- $Ax = b$,
- $x \geq 0$,

where $y = (y_1, \ldots, y_d)$. The KKT conditions as stated do not provide us with polynomials, due to the presence of $x_i$ in the denominator. By clearing the denominators we obtain a system of polynomials defining a curve.

$$(Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i = 0, i = 1, \ldots, n \tag{1.3}$$

$$Ax - b = 0.$$

We omit the nonnegativity constraints on $x$ so that we acquire polynomial equations we later form ideals with. The curve defined by these equations is a complex curve in $\mathbb{C}^{n+d+1}$, and the central curve $\mathcal{C}$ is the projection of this curve onto $\mathbb{C}^n$. In other words, the central curve is the projection of the curve defined by the above equations onto the $x$-variables. The projection is done by eliminating the $\lambda$ and $y_i$ variables from the equations (1.3). By performing this process whilst excluding the $Ax - b = 0$ equations, we obtain the central sheet. One can eliminate $\lambda$ and $y_i$ using Gröbner bases. Our elimination ideal $J_x$ is the following intersection, and is computed using an elimination term order [**?**, Chapter 7, Proposition 3].

$$J_x = \langle (Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i, i = 1, \ldots, n \text{ and } Ax = b \rangle \cap \mathbb{C}[x_1 \ldots x_n].$$

Computing $J_x$ is crucial in computing the degree of the central curve using Gröbner bases. We determined early on that this was impossible for computational experimentation, even when $d$ was as low as 2. When $Q$ was a diagonal matrix, however, finding $J_x$ became easier. Using computer simulation, we observed that for all cases of $d$ and $n$ where we could compute the degree, those degrees agreed regardless of whether $Q$ was symmetric positive-definite or diagonal. This inspired us to prove, using a deformation argument, that when computing the degree of the central curve of a generic quadratic program, it is enough to consider the diagonal case. This is shown in Chapter 3, Theorem 3.5. By considering the diagonal case we arrived at the following theorem, our main result.

**Theorem 1.1.** *Given a generic quadratic program of the form* (1.1)*, the degree of the central curve is at most*

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

**Conjecture 1.1.** *Given a generic quadratic program of the form* (1.1)*, the degree is exactly the formula provided in Theorem* 1.1

Conjecture 1.1 is supported by computer experimentation done using the program Singular [**?**]. In this thesis, however, it will remain only a conjecture, since the proof we have exceeds the scope of this paper. Our later writings will offer a proof of the conjecture, but for now we focus on Theorem 1.1, which has beautiful symmetry with the main result in [**?**] for linear programming. There, for a generic linear program, the degree is

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

At first glace the two formulae may seem unrelated, but through some simple manipulation we can see that

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

so the degree in both cases are the same except for a multiple of a power of two.

The outline of this paper is as follows: In Chapter 2, Section 1 we provide a detailed background of linear and convex optimization, including a discussion on solution methods that includes the interior point algorithm. In Section 2 of Chapter 2 we discuss algrebraic geometry and introduce important terms such as monomial ideals, projective varieties, and degree. Chapter 3 provides motivation and proofs for the reduction to the diagonal case. There, using algebraic geometry, we detail specific equations that define the central curve. Using these equations we introduce Lemma 3.1, the pivotal lemma in this paper which allows for the reduction to diagonal. This is utilized in Chapter 4, where we assume the diagonal case and construct an ideal that will allow us to bound the degree. There we also prove Theorem 1.1 using techniques discussed in Chapter 2 and the ideal created. In Chapter 5 we have a brief discussion on the future of our writings, with intended next steps for this research.

# Chapter 2

# Background

## 2.1 Optimization

We begin our explanation of optimization from the ground up, starting at linear programming. Linear programming solves the problem of minimizing a linear cost function subject to linear equality and inequality constraints. Essentially it is the mathematics behind finding the best possible solution to problems presented in a certain way. We will first introduce an example, without solution or context, that we will use to introduce many of the terms we define.

**Example 2.1.** Consider the following linear programming problem:

$$
\begin{array}{lrrrrcr}
\text{minimize} & 3x_1 & -x_2 & +2x_3 & & & \\
\text{subject to} & x_1 & -x_2 & & -x_4 & \leq & -2 \\
 & & 3x_2 & -x_3 & & = & 5 \\
 & & -3x_2 & +x_3 & & \leq & -5 \\
 & & & x_3 & +x_4 & \leq & 3 \\
 & x_1 & & & & \leq & 0 \\
 & & & -x_3 & & \leq & 0
\end{array}
$$

To solve this problem, one must find a solution $(x_1, x_2, x_3, x_4)$ such that the expression $3x_1 - x_2 + 2x_3$ attains the smallest value possible while each constraint on the variables is met. An objective function, or cost function, is the particular function that an optimization problem attempts to minimize. In linear programming this function is always linear, and in our example we can see that the objective function is $3x_1 - x_2 + 2x_3$. Decision variables are the variables being used, and are generally under constraints (but in some cases are not). In our case $x_1, x_2, x_3, x_4$ are our decision variables. Any solution $(x_1, x_2, x_3, x_4)$ that satisfies all constraints is called a feasible solution, and if that solution does minimize the objective function we say the solution is an optimal solution. The set of all feasible solutions is known as the feasible region.

We are going to use linear algebra to express all of our optimization problems, including linear programs. Let us view Example 2.1 through a linear algebra perspective. We let $x = (x_1, x_2, x_3, x_4)$ be the vector of decision variables. Given our cost vector $c = (3, -1, 2, 0)$, we seek to minimize the function $c^t x = \sum_{i=1}^{4} c_i x_i$.

Our constraints should also be expressed through linear algebra as well. In linear programming, the constraints are also always linear functions. We can see that all of our constraints are of the form $a^t x \leq b$, $a^t x = b$, or $a^t x \geq b$. Here $a = (a_1, a_2, a_3, a_4) \in \mathbb{R}^4$, and $b$ is a scalar. Consider our example's first constraint $x_1 - x_2 - x_4 \leq -2$. Here $a = (1, -1, 0, -1)$ and $b = -2$. We will also disregard any constraint of the form $a^t x = b$ and replace them with two constraints, $a^t x \leq b$ and $a^t x \geq b$ (which accomplish the same thing). In other words we have no need for the constraint $3x_2 - x_3 = 5$ if we introduce the constraints $3x_2 - x_3 \leq 5$ and $3x_2 - x_3 \geq 5$.

The reader may realize now that having both "$\leq$" and "$\geq$" constraints is unneccesary, since multiplying a constraint by $-1$ can conform the inequalities. For the purposes of this example, let us make every constraint of the form $a'x \geq b$. So instead of using $x_1 - x_2 - x_4 \leq -2$, we shall consider $-x_1 + x_2 + x_4 \geq -2$.

In doing this we can create a concise statement of our linear programming problem, namely:

$$\begin{aligned} \text{minimize} \quad & c^t x \\ \text{subject to} \quad & Ax \ \geq \ b, \end{aligned}$$

where $c = (3, -1, 2, 0)$, $b = (-2, 5, -5, 3, 0, 0)$, and

$$A = \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 3 & -1 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

Linear programs (and later convex optimization problems) can all be expressed in standard form. The following is a linear program in standard form:

$$\begin{aligned} \text{minimize} \quad & c^t x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

The geometry of optimization problems has many results that can be directly used to find solutions. Since each constraint in a linear program is linear, they give rise to rigid geometric boundaries. In order to discuss the geometry, we first introduce several definitions:

**Definition 2.1.** A polyhedron is a set that can be described in the form $\{x \in \mathbb{R}^n | Ax \geq b\}$, where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. If this set is bounded, it is called a polytope.

A convex set is a set $E \subseteq \mathbb{R}^n$ such that given $a, b \in E$ and $0 \leq \lambda \leq 1$, the convex combination $\lambda x + (1 - \lambda) y$ is in $E$. In layman's terms, a set is convex if every line segment joining two points in the set is entirely contained in the set. As a non-example, one could imagine a torus because of the missing center, a torus is not convex, while a disk is. It is not difficult to show that a polyhedron is a convex set.

### 2.1.1 Simplex Algorithm and Interior Point Methods

There are two main ways of solving linear programs: the simplex algorithm and interior point methods. We will demonstrate how the simplex algorithm works using an example. For a more detailed look at the simplex algorithm see Bertsimas [**?**], and for more details on interior point algorithms see Boyd [**?**].

**Example 2.2.** Assume we want to find the optimal solution to the following linear programming problem

$$\begin{array}{rcccl} \text{minimize} & -8x & - & -12y & \\ \text{subject to} & 10x & + & 20y & \leq 140 \\ & 6x & + & 8y & \leq 72 \\ & x & & & \geq 0 \\ & & & y & \geq 0. \end{array}$$

We can observe the feasible region in Figure 2.1. The feasible region is a polytope, and we call the "corners" of the polytope vertices. The simplex algorithm uses the fact that the optimal solution to a linear program will occur at one of these vertices. The algorithm then starts at a vertex and checks for optimality. If that solution is optimal then there is no need to continue, but if it is not then it moves to a neighboring vertex that improves the solution and repeats. In our example, if we start at the vertex $(0, 0)$ we see that the objective value would be 0. This is obviously not optimal, so we move to the vertex $(0, 7)$, producing an objective value of $-84$. With one more move to the vertex $(8, 3)$, we arrive at our optimal value of $-100$.

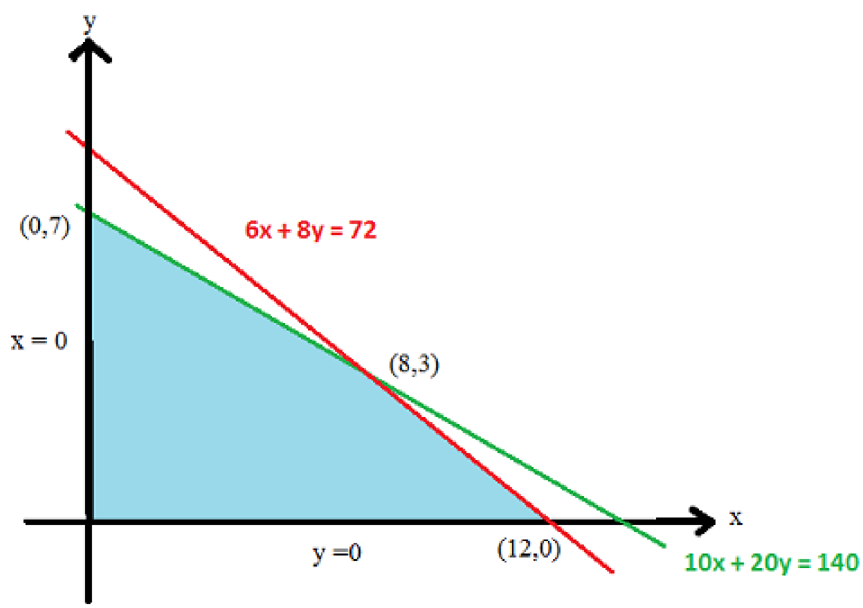


Figure 2.1: Feasible region for Example 2.2.

Interior point methods are an alternative family of solution methods, and they involve creating a path through the feasible region to the optimal function. Our research revolves around one aspect of interior point methods and is known as the central curve.

The construction of the curve comes from the following logic. Imagine we have a linear program in standard form:

$$\begin{aligned} \text{minimize} \quad & c^t x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

One thing to note is that while most of the constraints were changed to equality, the nonnegativity constraints are still inequalities. We want to incorporate a trick that will eliminate all inequalities. To that end, one can modify the objective function to imply nonnegativity of the decision variables while still insuring the optimal solution remains unchanged. Such a modified objective function is called a logarithmic barrier function. While a normal linear program minimizes the function $c^t x$, a logarithmic barrier function is of the form

$$f_\lambda(x) = c^t x - \lambda \textstyle\sum_{j=1}^{n} \log(x_i).$$

The barrier function is defined to be infinity if $x_i \leq 0$, thus eliminating the need for inequality constraints. Since the barrier function is strictly convex for any fixed $\lambda > 0$, $f_\lambda(x)$ has a unique optimal solution which we denote $x(\lambda)$. As $\lambda$ gets smaller, the expression $\lambda \sum_{i=1}^{n} \log x_i$ becomes insignificant. Because of this logic it can be shown that $\lim_{\lambda \to 0} x(\lambda)$ exists, and is an optimal solution to our original program with objective function $c^t x$. This gives rise to a new family of programming problems with no inequality constraints, called barrier problems, that for $\lambda > 0$ are of the form:

$$\begin{aligned} \text{minimize} \quad & f_\lambda(x) \\ \text{subject to} \quad & Ax = b. \end{aligned}$$

Since $x_i \neq 0$ in a barrier function, the decision variables are prevented from reaching the boundary. Thus for any given $\lambda > 0$, $x(\lambda)$ lies in the interior of the feasible region of the original linear program. As $\lambda$ approaches 0, $x(\lambda)$ gets closer to the boundary, specifically the optimal solution to $c^t x$. As this happens, $x(\lambda)$ forms the central path, which originates in the feasible region and travels to the optimal solution. Figure 2.2 shows the central path and its analytic center, which is characterised as the worst solution to the original problem (found by letting $\lambda = \infty$.)

### 2.1.2 Curvature of the Central Path

Interior point methods track the central path through the feasible region. The algorithm is impacted by how curvy the central path is. One way to quantify curviness is total curvature. Formally, the total curvature $\Phi$ over a closed interval $[a, b]$ measures the rotation of the unit tangent $T(s)$ as $s$ changes from $a$ to $b$.

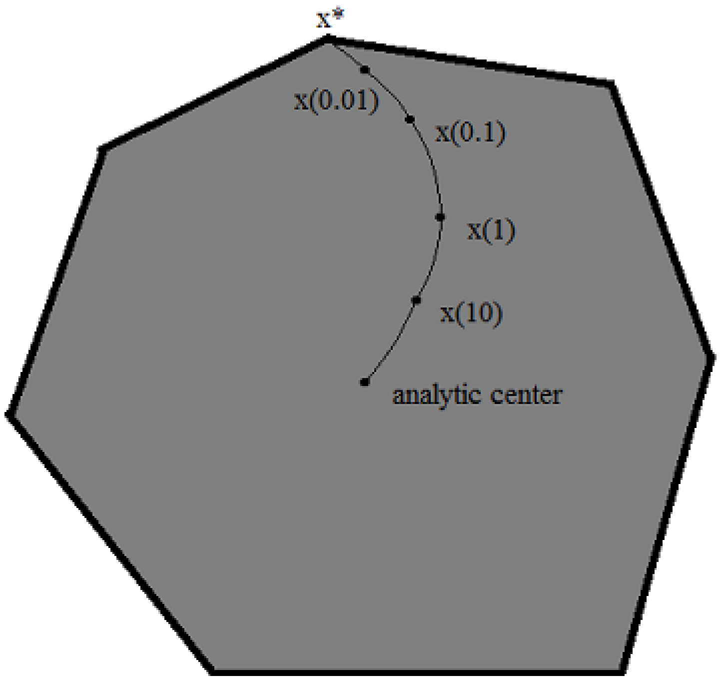


Figure 2.2: The central curve and its analytic center.

It is given by

$$\Phi(a, b) = \int_a^b \kappa \, ds,$$

where $\kappa$ is the curvature, given by $\kappa = \left|\frac{dT}{ds}\right|$. We were inspired to study the curvature of the quadratic case by a paper by Dedieu, Malajovich, and Shub [**?**] as well as the work in [**?**]. Theorem 2.1 below is from [**?**] and they consider different feasible regions depending on the sign of the $x$-variables. Since there are $2^n$ possibilities, they consider $2^n$ central paths, one for each region. The theorem proves a bound on curvature for the combination of all central paths

**Theorem 2.1.** *Let $n > d \geq 1$. Let $A$ be an $d \times n$ matrix of rank $d$, and let $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^d$, $c$ not in the row space of $A$, and $b$ nonzero. The sum over all $2^n$ sign conditions of the total curvature of the central path of the linear programs is less than or equal to $2\pi d\binom{n-1}{d}$.*

The above theorem and the work in [**?**] bring the question of whether we can bound the total curvature in a generic quadratic program. Up until this point the focus of our optimization background has been linear programming. Linear programs are the most basic optimization examples, but many of the results extend nicely to a quadratic programming setting. Quadratic programs are the concentration of this thesis, so we will now begin to dive into the larger branch of optimization, namely convex optimization.

The notion of convexity can be extended to functions. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain is a convex set and for all $x$ and $y$ in the domain of $f$, $0 \leq \lambda \leq 1$, we have $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$. One example to imagine is the quadratic function $f(x) = x^2$. Since the line segment connecting $(a, f(a))$ and $(b, f(b))$ over the interval $[a, b]$ lies above the graph of the function $f$

for all real $a, b$, then the function is convex.

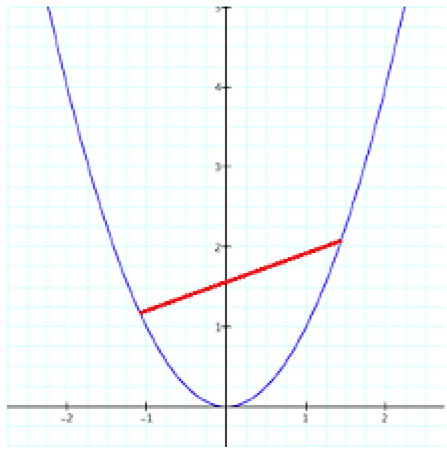

Figure 2.3: The function $f(x) = x^2$. Note how the line segment connecting two points lies above the graph of the function.

We have now set the stage to discuss convex optimization. Similar to how linear programs have a standard form, so do convex programs. A convex optimization problem is one of the form

$$\begin{array}{rlll} \text{minimize} & f_0(x) & & \\ \text{subject to} & f_i(x) & \geq & 0, \quad i = 1, \ldots, m \\ & a_i^T x & = & b_i, \quad i = 1, \ldots, p, \end{array}$$

where $f_0, \ldots, f_m$ are convex functions. One consequence of this definition is that the feasible set of a convex optimization problem will always be a convex set. Since linear functions are convex, then despite the difference of appearance, linear programs are always convex optimization problems. The converse is generally not true. Convex optimization is a giant family of problems, and one branch of that family is quadratic programming. A convex optimization problem is called a quadratic program if the objective function is quadratic (and of course still convex), but the constraint functions are linear. Quadratic programming problems are represented in the form

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x^t Q x + c^t x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned}$$

where $Q$ is an $n \times n$ symmetric positive-definitie matrix, $A$ is a real valued $d \times n$ matrix, $b$ is a vector in $\mathbb{R}^d$, and $c$ a vector in $\mathbb{R}^n$. When we mention a generic quadratic program, it will be of this form. We note that when $Q = 0$, we have a linear program, thus linear programming is a special case of quadratic programming.

The linear constraints make quadratic programs similar to their linear counterparts. The central curve can also be constructed in the same way as earlier. We employ another logarithmic barrier function, transforming our general quadratic program to a barrier problem of the form

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^tQx + c^tx - \lambda \sum_{i=1}^{m} \log(x_i) \\ \text{subject to} & Ax = b. \end{array}$$

For each value of $\lambda > 0$, the barrier problem has an optimal solution $x(\lambda)$. As $\lambda$ varies, the minimizers $x(\lambda)$ form the central path. When $\lambda$ is small, the logarithmic term is insignificant but still maintains the purpose of preventing the solutions from landing on the boundary, thus it can be shown that, just like in the linear case, $\lim_{\lambda \to 0} x(\lambda)$ exists and is the optimal solution to the original quadratic program.

This concludes the background information we need on optimization problems. Since the focus of this research is the degree of the central curve, it is imperative to have an understanding of degree as well. When we mention degree, we are viewing it from an algebraic geometry perspective. Much of our results about the central curve in general stem from an algebraic geometry standpoint. We will construct the notion of degree from the ground up using projective varieties, hyperplanes, and Gröbner bases.

## 2.2 Algebraic Geometry

### 2.2.1 Ideals and Affine Varieties

We start our exploration into algebraic geometry with a discussion of affine geometry. Algebraic geometry is usually done over a polynomial ring $k[x_1, \ldots, x_n]$ where $k$ is a general field. For the purposes of our work, it suffices if we only consider the ring $\mathbb{C}[x_1, \ldots .x_n]$, where $\mathbb{C}$ is the field of complex numbers. We now introduce several definitions vital to our work.

**Definition 2.2.** A nonempty subset $I \subset \mathbb{C}[x_1, \ldots, x_n]$ is an ideal if it satisfies:
(i) If $f, g \in I$, then $f + g \in I$.
(ii) If $f \in I$ and $h \in \mathbb{C}[x_1, \ldots, x_n]$, then $hf \in I$.

One basic geometric object in algebraic geometry is a variety.

**Definition 2.3.** Let $f_1, \ldots, f_s$ be polynomials in $\mathbb{C}[x_1, \ldots, x_n]$. We call $V(f_1, \ldots, f_s)$ the affine variety defined by $f_1, \ldots, f_s$, where

$$V(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in \mathbb{C}^n : f_i(a_1, \ldots, a_n) = 0 \text{ for all } 1 \leq i \leq s\}.$$

In other words the affine variety $V(f_1, \ldots, f_s) \in \mathbb{C}^n$ is the set of all solutions of the system of equations $f_1(x_1, \ldots, x_n) = \cdots = f_s(x_1, \ldots, x_n) = 0$. Let us view an example.

**Example 2.3.** We consider the example of the twisted cubic in $\mathbb{C}^3$, which is defined to be the variety $V(y-x^2, z-x^3)$. By definition the variety is the set of points $(x, y, z) \in \mathbb{C}^3$ such that $y-x^2 = 0$ and $z-x^3 = 0$. We observe that $V(y-x^2, z-x^3) = \{(x, x^2, x^3) : x \in \mathbb{C}\}$.

Since an ideal is a set of polynomials in $\mathbb{C}[x_1, \ldots, x_n]$, it makes sense to consider varieties of an ideal. The definition is intuitive. If $I$ is an ideal in $\mathbb{C}[x_1, \ldots, x_n]$, then the variety $V(I)$ is defined to be

$$V(I) = \{(x_1, \ldots, x_n) \in \mathbb{C}^n : f(x_1, \ldots, x_n) = 0 \text{ for all } f \in I\}.$$

There are many subclasses of ideals that come with an array of properties. One such type of ideal that is useful to us is a monomial ideal.

**Definition 2.4.** An ideal $I \subset \mathbb{C}[x_1, \ldots, x_n]$ is a monomial ideal if there is a subset $A \subset \mathbb{Z}^n_{\geq 0}$ (possibly infinite) such that $I$ consists of all polynomials which are finite sums of the form $\sum_{\alpha \in A} h_\alpha x^\alpha$, where $h_\alpha \in \mathbb{C}[x_1, \ldots, x_n]$. In this case, we write $I = \langle x^\alpha : \alpha \in A \rangle$.

Monomial ideals are special in that they possess many unique properties. One such property relates to how varieties of monomial ideals behave.

**Proposition 2.2.** *The variety of a monomial ideal in* $\mathbb{C}[x_1, \ldots, x_n]$ *is a finite union of coordinate subspaces of* $\mathbb{C}^n$.

*Proof.* We first note that if $x_{i_1}^{\alpha_1} \ldots x_{i_r}^{\alpha_r}$ is a monomial in $\mathbb{C}[x_1, \ldots, x_n]$ with $\alpha_j \geq 1$ for $1 \leq j \leq r$, then $V(x_{i_1}^{\alpha_1} \ldots \mathrm{x}_{i_r}^{\alpha_r}) = H_{x_{i_1}} \cup \cdots \cup H_{x_{i_r}}$, where $H_{x_k} = V(x_k)$. Thus, the variety defined by a monomial is a union of coordinate hyperplanes. Note also that there are only $n$ such hyperplanes.

Since a monomial ideal is generated by a finite collection of monomials, the variety corresponding to a monomial ideal is a finite intersection of unions of coordinate hyperplanes. By the distributive property of intersection of unions, any finite intersection of unions of coordinate hyperplanes can be rewritten as a finite union of intersections of coordinate hyperplanes. But the intersection of any collection of coordinate hyperplanes is a coordinate subspace. $\square$

**Example 2.4.** Let $I = \langle x^2y, x^3\rangle \in \mathbb{C}[x, y]$. We will denote $H_x = V(x)$, in other words $H_x$ is the line in $\mathbb{C}^2$ defined by $x = 0$. Similarily let $H_y = V(y)$. We observe that

$$\begin{aligned} \text{V(I)} &= V(x^2y) \cap V(x^3) \\ &= (H_x \cup H_y) \cap H_x \\ &= (H_x \cap H_x) \cup (H_y \cap H_x) \\ &= H_x, \end{aligned}$$

thus we conclude that $V(I) = H_x =$ the $y$-axis. Let us consider a second example. Let $I = \langle y^2z^3, x^5z^4, x^2yz^2\rangle \subset \mathbb{C}[x, y, z]$. Let $H_x$ be the plane $x = 0$, $H_y$, $H_z$ defined similarly, and let $H_{xy}$ be the line $x = y = 0$. Then

$$\begin{aligned} \text{V(I)} &= V(y^2z^3) \cap V(x^5z^4) \cap V(x^2yz^2) \\ &= (H_y \cup H_z) \cap (H_x \cup H_z) \cap (H_x \cup H_y \cup H_z) \\ &= H_z \cup H_{xy}, \end{aligned}$$

thus we conclude $V(I)$ is the union of the $(x, y)$-plane and the $z$-axis.

Since we express varieties of monomial ideals as a union of finitely many coordinate subspaces, we can omit certain subspaces, namely those that are contained in another union. Using this logic repeatedly we can always express the variety of a monomial ideal as

$$V(I) = V_1 \cup \cdots \cup V_p,$$

where $V_i \not\subset V_j$ for $i \neq j$. It can be shown that this decomposition is unique. Keeping in mind the result of Proposition 2.2, we define the dimension of a the variety of a monomial ideal in the following way:

**Definition 2.5.** Let $V$ be a variety that is a union of a finite number of linear subspaces of $\mathbb{C}^n$. Then the dimension of $V$, denoted dim $V$, is the largest of the dimensions of these subspaces. Furthermore we define the dimension of a monomial ideal as the dimension of the variety it defines.

For example the dimension of the union of two planes and a line is 2, since a plane is a subspace of dimension 2. The dimension of the union of seven lines is 1. This is true for the union of any finite amount of lines. This definition will be used often in our computations later, as well as the more general definition for the degree of a projective variety.

One class of monomial ideal that this paper takes great advantage of is an irreducible monomial ideal.

**Definition 2.6.** A monomial $I$ is called irreducible if it is of the form $I = \langle x_{i_1}^{u_{i_1}}, x_{i_2}^{u_{i_2}}, \ldots, x_{i_r}^{u_{i_r}} \rangle$, where $u_{i_1}, \ldots, u_{i_r}$ are positive.

These types of monomial ideals can be considered building blocks for all monomial ideals. In fact, every monomial ideal can be expressed as an intersection of irreducible monomial ideals [**?**, Lemma 5.8].

**Proposition 2.3.** *Every monomial ideal $M$ can be written as*

$$M = \bigcap Q_i$$

*where each $Q_i$ is an irreducible monomial ideal. Furthermore if $\bigcap_{j \neq i} Q_i \neq M$ then $\bigcap Q_i$ is called an irredundant irreducible decomposition of $M$.*

**Example 2.5.** Let $M = \langle x^2y, xz^3, xyz^2 \rangle$. An irredundant irreducible decomposition of $M$ is

$$\langle x^2y, xz^3, xyz^2 \rangle = \langle x \rangle \cap \langle x^2, z^2 \rangle \cap \langle y, z^3 \rangle.$$

### 2.2.2 Gröbner Bases and Eliminaton

A key idea we will use is that each $f \in \mathbb{C}[x_1, \ldots, x_n]$ has a unique leading term once a term order has been fixed.

**Definition 2.7.** A term order, or monomial ordering, on $\mathbb{C}[x_1, \ldots, x_n]$ is any relation $<$ on $\mathbb{Z}^n_{\geq 0}$, or equivalently, any relation on the set of monomials $x^\alpha$, $\alpha \in \mathbb{Z}^n_{\geq 0}$, satisfying:

(i) $<$ is a total ordering on $\mathbb{Z}^n_{\geq 0}$.

(ii) If $\alpha > \beta$ and $\gamma \in \mathbb{Z}^n_{\geq 0}$, then $\alpha + \gamma > \beta + \gamma$.

(iii) $<$ is a well-ordering on $\mathbb{Z}^n_{\geq 0}$. This means that every nonempty subset of $\mathbb{Z}^n_{\geq 0}$ has a smallest element under $<$.

The leading term of a polynomial is the greatest element under a certain ordering $<$. Let $f = \sum_\alpha a_\alpha x^\alpha$ be a nonzero polynomial in $\mathbb{C}[x_1, \ldots, x_n]$. Then the initial term of $f$ with respect to $<$ is $x^{f^*}$, where $f^*$ =max$(\alpha \in \mathbb{Z}^n_{\geq 0} : a_\alpha \neq 0)$. We can extend this concept to ideals as well.

**Definition 2.8.** Let $I \subset \mathbb{C}[x_1, \ldots, x_n]$ be an ideal that is not trivially $\{0\}$. Then the initial ideal $\text{in}_<(I)$ is the monomial ideal generated by the initial terms of all the polynomials in $I$.

It is not hard to show that every initial ideal is a monomial ideal. Similarily given that $I \subset \mathbb{C}[x_1, \ldots, x_n]$ is an ideal, there are $g_1, \ldots, g_t \in I$ such that $\text{in}_<(I) = \langle \text{in}_<(g_1), \ldots, \text{in}_<(g_t) \rangle$. This leads us to the definition of a Gröbner bases.

**Definition 2.9.** Fix a monomial order $<$. A finite subset $G = \{g_1, \dots, g_t\}$ of an ideal $I$ is said to be a Gröbner basis with respect to $<$ if

$$\langle \mathrm{in}_<(g_1), \dots, \mathrm{in}_<(g_t) \rangle = \mathrm{in}_<(I).$$

Gröbner bases can be computed using Buchberger's algorithm [**?**]. This algorithm is implemented into most computer algebra systems, including the one we use, Singular [**?**]. Earlier when discussing the central path we mentioned that we have to project our equations onto the $x$-variables, and we do that with elimination. Namely, we use elimination ideals.

**Definition 2.10.** Given $I = \langle f_1, \dots, f_s \rangle \subset \mathbb{C}[x_1, \dots, x_n]$, the $i$th elimination ideal $I_i$ is the ideal of $\mathbb{C}[x_{i+1}, \dots, x_n]$ defined by

$$I_i = I \cap \mathbb{C}[x_{i+1}, \dots, x_n].$$

To compute an $i$th elimination ideal, a term order must first be fixed. We must take a type of lexicographical order, so say $<$ is the lexicographical order with $x_1 > x_2 > \cdots > x_n$. The Gröbner basis with respect to $<$ is then computed, and the elimination ideal consists of those polynomials in the Gröbner basis that contain only the variables $x_{i+1}, \dots, x_n$ [**?**, Chapter 3 Theorem 2]. In our quadratic programming case our equations are in $x$ and $y$ variables, with $\lambda$ as well. Our projection to the $x$-varaibles is just an elimination ideal that eliminates the $y$ and $\lambda$ variables.

**Example 2.6.** Consider again the twisted cubic in $\mathbb{C}^3$, $I = \langle y - x^2, z - x^3\rangle$. To project the twisted cubic down to $\mathbb{C}^2$ we use an elimination ideal to eliminate $z$. Let $<$ be a lexicographical order with $y > z > x$. A Gröbner basis for $I$ with respect to $<$ is $G = \{y - x^2, z - x^3\}$. Of the two Gröbner basis elements, only $y - x^2$ does not contain the variable $z$, so our elimination ideal is $I_{x,y} = \langle y - x^2\rangle$.

### 2.2.3 Projective Varieties and Degree

The construction of the projective plane can be generalized to create projective spaces of any dimension. For this we let $k$ be a field and define an equivalence relation on $k^{n+1}$. Let $(x_0', x_1', \ldots, x_n')$ and $(x_0, x_1, \ldots, x_n)$ be nonzero points in $k^{n+1}$. Then we say that

$$(x_0', x_1', \ldots, x_n') \sim (x_0, x_1, \ldots, x_n)$$

if there is a nonzero element $\lambda \in k$ such that $(x_0', x_1', \ldots, x_n') = \lambda(x_0, x_1, \ldots, x_n)$. Using this equivalence class we form the following definition:

**Definition 2.11.** $\mathbb{P}^n(\mathbb{C})$, the n-dimensional projective space over the field $\mathbb{C}$, is the set of equivalence classes of $\sim$ on $\mathbb{C}^{n+1} - \{0\}$, where $\{0\}$ is the origin $(0, \ldots, 0)$ in $\mathbb{C}^{n+1}$.

One could think of $\mathbb{P}^n(\mathbb{C})$ as the set of lines passing through the origin in $\mathbb{C}^{n+1}$. For example the projective plane $\mathbb{P}^2(\mathbb{C})$ is isomorphic to the set of lines passing through the origin in $\mathbb{C}^3$. It can be shown easily that this relationship is one-to-one, so $\mathbb{P}^n(\mathbb{C})$ is isomorphic to the set of lines passing through the origin in $\mathbb{C}^{n+1}$. Note this implies that $\mathbb{C}^n \subset \mathbb{P}^n(\mathbb{C})$.

We defined affine varieties, but with projective space defined we can introduce the notion of a projective variety. First we introduce homogeneity.

**Definition 2.12.** A polynomial $f \in k[x_1, \ldots, x_n]$ is homogenous of total degree $d$ provided that every term appearing in $f$ has degree $d$.

Take for example the polynomial $f = x_1^3 x_2 + x_1^2 x_2^2 - 3x_2^4$. All three terms in the polynomial have degree 4, thus $f$ is homogenous in $\mathbb{C}[x_1, x_2]$. A polynomial not homogenous in $\mathbb{C}[x_1, x_2]$ is $g = x_1 - x_2^2$. We can extend homogeneity to ideals. The following definition is equivalent to saying that a homogenous ideal is generated by homogenous polynomials.

**Definition 2.13.** An ideal $I$ in $\mathbb{C}[x_0, \ldots, x_n]$ is said to be homogenous if for each $f \in I$, the homogenous components $f_i$ of $f$ are in $I$ as well.

But why do we care about homogenous polynomials when considering projective varieties? Homogenous coordinates in the projective plane create a problem when extending affine varieties to the projective case. Consider the variety $V(x_1 - x_2^2)$ in $\mathbb{P}^2(\mathbb{C})$. The point $p = (x_0, x_1, x_2) = (1, 4, 2)$ appears to be in the variety, since the components of $p$ satisfy $x_1 - x_2^2 = 0$. However in the projective plane, $p$ can be represented by the homogenous coordinate $p = 2(1, 4, 2) = (2, 8, 4)$. This new coordinates do not satisfy $x_1 - x_2^2 = 0$, since $8 - 4^2 \neq 0$. To avoid this type of problem, we limit ourself to using homogenous polynomials in $\mathbb{P}^n(\mathbb{C})$.

**Definition 2.14.** Let $f_1, f_2, \ldots, f_s \in \mathbb{C}[x_0, \ldots, x_n]$ be homogenous polynomials. We call $V(f_1, \ldots, f_s)$ the projective variety definied by $f_1, \ldots, f_s$, where

$$V(f_1, \ldots, f_s) = \{(a_0, \ldots, a_n) \in \mathbb{P}^n(\mathbb{C}) : f_i(a_0, \ldots, a_n) = 0 \text{ for all } 1 \leq i \leq s\}$$

**Example 2.7.** Consider the variety of the projective twisted cubic $V = (xz - y^2, yw - z^2, xw - yz)$ in $\mathbb{P}^3(\mathbb{C})$. We observe that a point of the form $(x, y, z, w) = (s^3, s^2t, st^2, t^3), s, t \in \mathbb{C}$ makes each equation vanish, so the variety is equal to the equivalence class $V = (xz - y^2, yw - z^2, xw - yz) = [s^3 : s^2t : st^2 : t^3]$.

Since ultimately the goal of this paper is to discuss the degree of the central path, which we view as a projective variety, we must consider how to compute the degree of a projective variety. There are several methods for this, but perhaps the one most useful to us relates to intersection points.

**Definition 2.15.** Let $K$ be an algebraically closed field and $V$ a projective variety embedded in $\mathbb{P}^n(k)$. The degree $d$ of $V$ is the number of intersection points of $V$ counted with multiplicity, defined over $K$, with a generic linear subspace $L$ with complementary dimension, i.e. $\dim(V) + \dim(L) = n$. The degree of a homogenous ideal $I$ is defined to be the degree of $V(I)$.

The following process is how we manually compute the degree for a projective variety $V(I)$ [**?**].

1. Pick a degree reverse lex term order $<$.
2. Compute a Gröbner basis of $I$ with respect to $<$.
3. Form the initial ideal $M = \text{in}_<(I)$.
4. Compute the degree of the initial ideal $M$.
5. The degree of $V(I)$ is equal to the degree of $M$ [**?**, Chapter 9, Theorem 8].

This process is fairly self explanatory, except for the fourth step, computing the degree of the initial ideal $M$. Let us walk through that process. Assume $M$ is the initial ideal of $I$ with respect to $<$. We first compute an irredundant irreducible decomposition of $M$. Recall that this will express $M$ as $M = \bigcap Q_i$ where $Q_i$ is an irreducible monomial ideal of the form $\langle x_{i_1}^{u_{i_1}}, \ldots, x_{i_r}^{u_{i_r}} \rangle$.

Since each generator of $Q_i$ is a single variable raised to a positive power, we can look at the top-dimensional components. Notice that the variety of $Q_i$ is easy to compute, since $V(Q_i) = \{(x_1, \ldots, x_n) : x_{i_1} = x_{i_2} = \cdots = x_{i_r} = 0\}$. Here we assume our ideal $M$ is in $\mathbb{C}^n$, meaning $V(Q_i) \subset \mathbb{C}^{n-r}$. We look only at the $Q_i$ that are in the highest dimension, i.e, the ones with the smallest $r$ value.

It is possible that more than one component of $M$ has the highest dimension. In this case we consider all $Q_i$ with top dimension. After the top-dimensional components are selected, we compute their degrees. Since $Q_i$ is irreducible, the degree is computed by taking the product of the positive powers of the generators, namely $u_{i_1} u_{i_2} \ldots u_{i_r}$. We sum the degree of all $Q_i$ with top dimension, and the resulting number is the degree of the initial ideal $M$.

**Example 2.8.** Again consider the variety of the projective twisted cubic $V = (xz - y^2, yw - z^2, xw - yz)$ in $\mathbb{P}^3(\mathbb{C})$. The ideal defining this variety is $I = \langle xz - y^2, yw - z^2, xw - yz\rangle$. We fix a degree reverse lexicographical order $<$ on the variables $x < y < w < z$. A reduced Gröbner basis with respect to $<$ is the set $G = \{xz - y^2, yw - z^2, xw - yz\}$.

We create the monomial ideal $\mathrm{in}_<(I) = \langle xz, yw, xw\rangle$ generated by the initial terms of $G$. An irredundant irreducible decomposition of $\mathrm{in}_<(I)$ is $\langle xz, yw, xw\rangle = \langle x, y\rangle \cap \langle x, w\rangle \cap \langle w, z\rangle$. Each component is top dimensional, so the sum of their degrees $1 + 1 + 1 = 3$ is the degree of $\mathrm{in}_<(I)$ and thus the degree of $V = (xz - y^2, yw - z^2, xw - yz)$.

# Chapter 3

# Reduction to Diagonal Matrices

## 3.1 Exploration and Examples

To briefly reiterate the logic described in Chapter 1 Section 2, we begin with a generic quadratic program of the form

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^tQx + c^tx \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{3.1}$$

where $Q$ is an $n \times n$ symmetric positive-definitie matrix, $A$ is a $d \times n$ matrix with real entries, $b$ is a vector in $\mathbb{R}^d$, and $c$ is a vector in $\mathbb{R}^n$.

Before considering the central curve we further convert this quadratic program to the barrier problem

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^tQx + c^tx - \lambda\sum_{i=1}^{n}\log(x_i) \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{3.2}$$

where $0 < \lambda < \infty$. For a given $\lambda \in (0, \infty)$, there is an optimal solution to (3.2). This is found by solving the KKT conditions for optimality given by

$$(Qx)_i + c_i - \frac{\lambda}{x_i} - (A^ty)_i = 0, i = 1, \ldots, n. \tag{3.3}$$

$$Ax = b$$

$$x \geq 0.$$

With the goal of forming polynomial equations in mind, we clear denominators and remove the $x \geq 0$ condition to obtain the polynomial equations

$$(Qx)_ix_i + c_ix_i - \lambda - (A^ty)_ix_i = 0, i = 1, \ldots, n \tag{3.4}$$

$$Ax - b = 0.$$

These equations are in the variables $x$, $y$, and $\lambda$, and they define a complex curve in $\mathbb{C}^{n+d+1}$. The projection of this curve onto the $x$-variables in $\mathbb{C}^n$ is precisely the central curve $\mathcal{C}$. Technically speaking, after clearing denominators, our KKT conditions give rise to an ideal. Let $J$ be the ideal generated by (3.4),

$$J = \langle (Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i, i = 1, \ldots, n \text{ and } Ax = b \rangle \subset \mathbb{C}[x_1 \ldots x_n, y_1, \ldots, y_d, \lambda].$$

We let $\mathcal{E}$ be the curve that is the affine variety $V(J) \subset \mathbb{C}^{n+d+1}$. The central curve is obtained by projecting $\mathcal{E}$ onto the $x$ variables, eliminating $y$ and $\lambda$. We denote this projection map as $\pi : \mathbb{C}^{n+d+1} \to \mathbb{C}^n$ and note that $\pi(\mathcal{E})$ is not necessarily an affine variety [**?**, Chapter 1, Section 2, Definition 1]. To remedy this, we take the smallest affine variety containing $\pi(\mathcal{E})$, which is equal to the variety defined by the elimination ideal $J_x = J \cap C[x_1, \ldots, x_n]$ [**?**]. This is also equal to the Zariski closure $\overline{\pi(\mathcal{E})}$. This produces a curve $\mathcal{C} \subset \mathbb{C}^n$, which is the central curve. Formally, we use the following definition.

**Definition 3.1.** The central curve $\mathcal{C}$ is equal to $\overline{\pi(\mathcal{E})}$, namely, the Zariski closure of the projection of $\mathcal{E}$ onto $\mathbb{C}^n$.

Our research aims to compute the degree of $\mathcal{C}$, or more accurately the degree of the projective closure of $\mathcal{C}$. Keeping in mind these observations, our first goal was to explore using computational software a potential pattern for the degree of the central curve of (3.1). The computations were done in the program Singular [**?**]. We applied the elimination techniques just described to specific examples.

**Example 3.1.** Beginning with a barrier problem of the form (3.2), let

$$Q = \begin{bmatrix} 11 & -2 & 6 & 15 & 8 \\ -2 & 4 & 9 & -13 & 18 \\ 6 & 9 & 1 & 5 & -7 \\ 15 & -13 & 5 & 30 & -10 \\ 8 & 18 & -7 & -10 & 3 \end{bmatrix} \qquad c^t = \begin{bmatrix} -2 \\ -3 \\ 4 \\ 9 \\ 1 \end{bmatrix}.$$

For the purposes of this example we picked random small integer values for the entries. In general they are selected randomly. Here $Q$ is a symmetric $5 \times 5$ matrix, since we are using the values $d = 2$ and $n = 5$. For our constraint matrices we let

$$A^t = \begin{bmatrix} 0 & -1 \\ -1 & 10 \\ 1 & 3 \\ -1 & 11 \\ 7 & -4 \end{bmatrix} \qquad b = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

To utilize Singular we explicitly write out the KKT conditions (3.4) for this system. The condition $Ax - b = 0$ is expressed by the following two polynomials

$$
\begin{aligned}
f_1 &= -x_2 + x_3 - x_4 + 7x_5 - 2, \\
f_2 &= -x_1 + 10x_2 + 3x_3 + 11x_4 - 4x_5 - 6,
\end{aligned}
$$

while our other KKT condition, $(Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i = 0$, $i = 1, \ldots, n$, is expressed with the polynomials

$$
\begin{aligned}
g_1 &= (-y_2 + 2)x_1 - (11x_1 - 2x_2 + 6x_3 + 15x_4 + 8x_5)x_1 - \lambda, \\
g_2 &= (-y_1 + 10y_2 + 3)x_2 - (-2x_1 + 4x_2 + 9x_3 - 13x_4 + 18x_5)x_2 - \lambda, \\
g_3 &= (y_1 + 3y_2 - 4)x_3 - (6x_1 + 9x_2 + 1x_3 + 5x_4 - 7x_5)x_3 - \lambda, \\
g_4 &= (-y_1 + 11y_2 - 9)x_4 - (15x_1 - 13x_2 + 5x_3 + 30x_4 - 10x_5)x_4 - \lambda. \\
g_5 &= (-4y_2 - 1)x_5 - (8x_1 + 18x_2 - 7x_3 - 10x_4 + 3x_5)x_5 - \lambda.
\end{aligned}
$$

These functions are the defining parameters of our central curve. We create an ideal generated by these equations and have Singular project that ideal down to the $x$-variables using elimination. The degree of the resulting elimination ideal is precisely the degree of our central curve. The following is example code taken from Singular, including both input and output.

```
ring r=0,(x1,x2,x3,x4,x5,y1,y2,z),dp;
poly f1=-x2+x3-x4+7*x5-2;
poly f2=-x1+10*x2+3*x3+11*x4-4*x5-6;
poly g1=(-y2+2)*x1-(11*x1-2*x2+6*x3+15*x4+8*x5)*x1-z;
poly g2=(-y1+10*y2+3)*x2-(-2*x1+4*x2+9*x3-13*x4+18*x5)*x2-z;
poly g3=(y1+3*y2-4)*x3-(6*x1+9*x2+1*x3+5*x4-7*x5)*x3-z;
poly g4=(-y1+11*y2-9)*x4-(15*x1-13*x2+5*x3+30*x4-10*x5)*x4-z;
poly g5=(7*y1-4*y2-1)*x5-(8*x1+18*x2-7*x3-10*x4+3*x5)*x5-z;
ideal I=f1,f2,g1,g2,g3,g4,g5;
ideal J=elim(I,y1*y2*z);
degree(J);
dimension (proj.)  = 3
degree (proj.)   = 11
```

Singular returns a degree of 11. Table 3.1 shows the different degrees we obtained through these types of computations for various $d$ and $n$. An early issue was that these computations were hard for Singular to compute, and for values of $n$ as low as 7 the degree could not be computed. These limited observations gave us no insight to a formula, or even a pattern. We proceeded to try and manually find equations that could reveal a formula, but with symmetric $Q$ this seemed impossible. We went back to Singular, only this time we limited $Q$ to be a diagonal matrix.

| d\n | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| 1 | 3 | 7 | 15 | 31 |
| 2 | 1 | 4 | 11 | 26 |
| 3 | | 1 | 5 | 16 |
| 4 | | | 1 | 6 |

Table 3.1: Degree of central curve with symmetric $Q$.

**Example 3.2.** Beginning with a barrier problem of the form (3.2), let

$$Q = \begin{bmatrix} 11 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \qquad c^t = \begin{bmatrix} -2 \\ -3 \\ 4 \\ 9 \\ 1 \end{bmatrix}.$$

We once again let $d = 2$ and $n = 5$, but use a diagonal $Q$ instead. We use the exact same constraints as in Example 3.1.

$$A^t = \begin{bmatrix} 0 & -1 \\ -1 & 10 \\ 1 & 3 \\ -1 & 11 \\ 7 & -4 \end{bmatrix} \qquad b = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

Our polynomials are set up using the KKT conditions in the same way as Example 3.1, only the $(Qx)_i x_i$ terms are simpler due to $Q$ being diagonal. Once again inputting our equations into Singular and using elimination we see that

```
ring r=0,(x1,x2,x3,x4,x5,y1,y2,z),dp;
poly f1=-x2+x3-x4+7*x5-2;
poly f2=-x1+10*x2+3*x3+11*x4-4*x5-6;
poly g1=(-y2+2)*x1-(11*x1)*x1-z;
poly g2=(-y1+10*y2+3)*x2-(4*x2)*x2-z;
poly g3=(y1+3*y2-4)*x3-(1*x3)*x3-z;
poly g4=(-y1+11*y2-9)*x4-(30*x4)*x4-z;
poly g5=(7*y1-4*y2-1)*x5-(3*x5)*x5-z;
```

```
ideal I=f1,f2,g1,g2,g3,g4,g5;
ideal J=elim(I,y1*y2*z);
degree(J);
dimension (proj.)  = 3
degree (proj.)   = 11
```

a degree of 11 is returned again. This exactly matches our observation for the same $d$ and $n$ in the symmetric case. In general, one would expect the degree of the central curve when $Q$ is diagonal to be smaller than or equal to the degree when $Q$ is arbitrary symmetric. We considered the above computational observation to be a coincidence, so we repeated this process for every $d$ and $n$ we tested in Table 3.1 and would always get matching results. Table 3.2 shows the degree of the central curve computed for diagonal $Q$ using various $d$ and $n$. Notice that because of the simplicity of the diagonal case, Singular was able to perform the computations easier, allowing us to extend our observations further.

| d\n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 7 | 15 | 31 | 63 | 127 | 255 | 511 |
| 2 | 1 | 4 | 11 | 26 | 57 | 120 | 247 | 502 |
| 3 | | 1 | 5 | 16 | 42 | 99 | 219 | 466 |
| 4 | | | 1 | 6 | 22 | 64 | 163 | 382 |

Table 3.2: Degree of the central curve for diagonal $Q$.

All of our computations seemed to imply that the degree was the same in the diagonal case as it was in the symmetric case. This served as our motivation to pursuing and proving Theorem 3.5 that shows our problem can be reduced to the diagonal case. When $Q$ is diagonal, manually finding equations to use as a Gröbner basis and prove Theorem 1.1 becomes possible. The reduction to diagonal paved the way for the formulation of Theorem 1.1 and breathed new life into our research.

## 3.2 Reduction to Diagonal

After our computations in Singular revealed the potential for reduction to diagonal, we pursued the idea. The reduction to diagonal is convenient to us because computing a Gröbner basis using an elimination term order used to prove Theorem 1.1 appeared to be difficult without assuming $Q$ was diagonal. The reduction to diagonal is done using a deformation argument, essentially showing that by starting with a symmetric $Q$, it could be shifted to a diagonal $Q$ without ever affecting the degree. We begin by adapting the algebraic geometry discussed in Chapter 2 to the central curve and then proceed with deformation.

After clearing denominators, our KKT conditions give rise to an ideal. Let $J$ be the ideal

$$J = \langle (Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i, i = 1, \ldots, n \rangle \cap \langle Ax = b \rangle \subset \mathbb{C}[x_1 \ldots x_n, y_1, \ldots, y_d, \lambda].$$

Theorem 3.4 provides us a way for computing the degree of the central curve. However since the theorem counts solutions in $\mathbb{C}^*$, it is important we rule out the possibility that a component of the central curve lie on a coordinate hyperplane. If $\mathcal{C}$ can be shown to be irreducible, then this result immediately follows. In the linear case this is exactly what is done [?]. Unfortunately we see no way to extend the techniques of [?] to the quadratic case, and we are unable to provide an alternative proof. For now we leave the irreducibility of quadratic central curve $\mathcal{C}$ to a conjecture that we believe to be true.

In lieu of proving irreducibility, we provide Lemma 3.1 below. Lemma 3.1 accomplishes the goal of proving that no components of $\mathcal{C}$ can lie in coordinate hyperplanes. The proof shows that by removing all parts of the curve that intersect the coordinate hyperplanes, then taking the closure, the entire curve is returned. This would not be possible if a component of the curve was lying on a coordinate hyperplane.

**Lemma 3.1.** $V(J : (x_1 \ldots x_n)^\infty) = V(J) = \mathcal{E}$.

*Proof.* It is trivial that $V(J : (x_1 \ldots x_n)^\infty) \subset V(J)$, so we need only show containment in the other direction. If we observe the difference of the two varieties, then every point $(x^*, y^*, \lambda^*) = (x_1^*, \ldots, x_n^*, y_1^*, \ldots, y_d^*, \lambda^*)$ in that difference must have at least one $x_i^* = 0$ for $1 \leq i \leq n$.

Without loss of generality let us assume $x_1^* = \cdots = x_k^* = 0$. We make two observations. The first is that the equations that generate $J$ imply that $\lambda^* = 0$.

Second we observe that $k \leq n-d$. Consider the point in the difference $(0, \bar{x}^*, y^*) = (0, \ldots, 0, x^*_{k+1}, \ldots, x^*_n, y^*_1, \ldots, y^*_d)$. Then $(\bar{x}^*, y^*)$ is a solution to the linear system

$$\begin{pmatrix} Q' & -A'^t \\ A' & 0 \end{pmatrix} \begin{pmatrix} \bar{x}^* \\ y^* \end{pmatrix} = \begin{pmatrix} -\bar{c}^t \\ b \end{pmatrix}$$

Where $\bar{c} = (c_{k+1}, \ldots, c_n)$, $Q'$ a submatrix of $Q$ that consists of the entries that are the last $n-k$ rows and columns of $Q$, and similarly $A'$ consisting of the last $n-k$ columns of $A$. Since $x^*_1 = \cdots = x^*_k = 0$ we can ignore the first $k$ columns of $Ax = b$ and the first $k$ polynomial equations generating $J$, giving us this linear system. Furthermore since $x^*_i \neq 0$ for $k+1 \leq i \leq n$, we can divide the $i$th equation of $(Q'\bar{x}^*)x^*_i + c_i x^*_i - (y^{*t}A)_i x^*_i = 0$ by $x^*_i$.

The above $(n-k+d) \times (n-k+d)$ matrix must have full rank as guarenteed by the genericity of $Q$ and $A$ as well as $(\bar{c}, b) \in \mathbb{C}^{n-k+d}$, since otherwise the system might not have a solution. However if $k > n-d$, then the rank of $A'$ is less than $d$, and the system would not have full rank. Therefore $k \leq n-d$.

Now we show that $Q_{ii}x^*_i + (Qx^*)_i + c_i - (y^{*t}A)_i \neq 0$ for $1 \leq i \leq k$. For the sake of contradiction let us assume $Q_{ii}x^*_i + (Qx^*)_i + c_i - (y^{*t}A)_i = 0$ for $i = 1$ (note we can let $i = 1$ without loss of generality). We append this equation to our previous linear system to obtain a new $(n-k+d+1)\times(n-k+d+1)$ matrix of full rank, due to $x_2^* = \cdots = x_n^* = 0$. Once again the genericity of $(\bar{c}, b) \in \mathbb{C}^{n-k+d}$ implies the system should have a unique solution with full support, and hence $x_1^* \neq 0$. This is a contradiction.

Let $f_1, \ldots, f_n$ be the quadratic equations generating $J$ while $g_1, \ldots, g_d$ the linear constraint generators of $J$. We let $x^* = (0, \bar{x}^*)$ and compute the Jacobian of these generators at $(x^*, y^*, 0)$. Computationally we see that $\frac{\partial f_i}{\partial x_i} = Q_{ii}x_i^* + (Qx^*)_i + c_i - (y^{*t}A)_i$ for $1 \leq i \leq n$, and $\frac{\partial f_i}{\partial x_j} = 0$ for $1 \leq i \leq k$ and $1 \leq j \neq i \leq n$ as well as for $k+1 \leq i \leq n$ and $1 \leq j \leq k$. For the other cases we simply have $\frac{\partial f_i}{\partial x_j} = Q_{ij}x_j^*$. Of course, $\frac{\partial g_i}{\partial x_j} = a_{ij}$ for $1 \leq i \leq d$ and $1 \leq j \leq n$. Also $\frac{\partial g_i}{\partial y_j} = 0$ for $1 \leq i \leq d$ and $1 \leq j \leq d$, as well as $\frac{\partial g_i}{\partial \lambda} = 0$ for $1 \leq i \leq d$. Moreover, $\frac{\partial f_i}{\partial y_j} = 0$ for $1 \leq i \leq k$ and $1 \leq j \leq d$ and $\frac{\partial f_i}{\partial y_j} = -a_{ji}$ for $k+1 \leq i \leq n$ and $1 \leq j \leq d$. Finally, $\frac{\partial f_i}{\partial \lambda} = -1$ for $1 \leq i \leq n$.

The Jacobian evaluated at $(x^*, y^*, 0)$ is an $(n+d)\times(n+d+1)$ matrix. We will show that its rank is $n+d$. To show this we consider just the part of this matrix corresponding to the $n+d$ variables $x$ and $y$.

The first $n$ rows coming from the quadratic generators is of the form

$$\begin{pmatrix} \tilde{Q} & 0 & 0 \\ 0 & Q^* & -A'^t \end{pmatrix}$$

where $\tilde{Q}$ is a diagonal $k \times k$ matrix with diagonal entires equal to $\frac{\partial f_i}{\partial x_i} \neq 0$ where $1 \leq i \leq k$. $Q^*$ is an $(n-k) \times (n-k)$ matrix with entries $\frac{\partial f_i}{\partial x_j}$ for $k+1 \leq i, j \leq n$ as computed above and $A'$ a matrix obtained by removing the first $k$ columns of $A$. $Q^*$ has rank $n-k$ since $Q$ is generic. If we examine the upper $n \times n$ block made up of $\tilde{Q}$ and $Q^*$ and the appropriate zero matricies, we see it has rank $n$. The matrix $A'$ has rank $d$, due to the genericty of $A$ and the fact that $n-k \geq d$. For this part of the Jacobian matrix, the last $d$ rows are of the form $\begin{pmatrix} A & 0 \end{pmatrix}$ where the zero block has size $d \times d$. By stacking both displayed matrices we get a new matrix with rank $n+d$.

The Implicit Function Theorem implies that in a neighborhood of $(x^*, y^*, 0)$, the curve $\mathcal{E}$ is parametrized by $(x(\lambda), y(\lambda), \lambda)$ where $\lambda \neq 0$. When we say neighborhood we mean the usual topology sense. Therefore this neighborhood of $\mathcal{E}$ consists of complex solutions to the KKT conditions displayed in (3.4), and hence $(x^*, y^*, 0)$ is in the closure of the complex solutions to (3.4), and thus is in $V(J : (x_1 \ldots x_n)^\infty)$. This proves $V(J : (x_1 \ldots x_n)^\infty) = V(J) = \mathcal{E}$. $\square$

Lemma 3.1 guarantees us that any component of $\mathcal{E}$ on a coordinate hyperplane could not exist.

**Corollary 3.2.** *No component of the central curve $\mathcal{C}$ lies in a coordinate hyperplane of $\mathbb{C}^n$.*

*Proof.* Assume there is a component of $\mathcal{C}$ that lies in a coordinate hyperplane of $\mathbb{C}^n$. This component then comes from a component of $\mathcal{E}$ that lies a coordinate hyperplane $x_i = 0$. This contradicts the result of Lemma 3.1, and thus no such coordinate could exist. □

Recall Definition 2.14 regarding the degree of a projective variety. Here, when we say the degree of $\mathcal{C}$, what we mean is the degree of the projective closure $\tilde{\mathcal{C}}$ of $\mathcal{C}$ in $\mathbb{P}^n(\mathbb{C})$.

**Corollary 3.3.** *The degree of the central curve $\mathcal{C}$ is the number of intersection points of $\mathcal{C}$ with a generic hyperplane $\mathcal{H} \in \mathbb{C}^n$. Since $\mathcal{C}$ arises from generic data, $\mathcal{C} \cap \mathcal{H} \subset (\mathbb{C}^*)^n$.*

*Proof.* Let $\tilde{\mathcal{C}}$ be the projective closure of $\mathbb{C}$ in $\mathbb{P}^n(\mathbb{C})$, $\tilde{\mathcal{H}}$ be a generic hyperplane in $\mathbb{P}^n(\mathbb{C})$, and $\mathcal{H}$ is the hyperplane obtained by restricting $\tilde{\mathcal{H}}$ to $\mathbb{C}^n$. The degree of $\mathcal{C}$ is equal to $|\tilde{\mathcal{C}} \cap \tilde{\mathcal{H}}|$. By taking the closure, only a finite amount of points can be added to $\mathcal{C}$ from the hyperplane at infinity. Due to the genericity of $\tilde{\mathcal{H}}$, there is a probability of 0 that these points will be in $\tilde{\mathcal{C}} \cap \tilde{\mathcal{H}}$. Therefore we conclude that $|\tilde{\mathcal{C}} \cap \tilde{\mathcal{H}}| = |\mathcal{C} \cap \mathcal{H}|$. By Corollary 3.2, every point in $\mathcal{C} \cap \mathcal{H}$ is also in $(\mathbb{C}^*)^n$, concluding the proof. □

The following theorem extends Corollary 3.3 to involve our KKT conditions, which when solved give solutions in $(\mathbb{C}^*)^{n+d+1}$.

**Theorem 3.4.** *The degree of $\mathcal{C}$ is equal to the number of solutions in $(\mathbb{C}^*)^{n+d+1}$ to the system* (3.4) *together with an extra equation of the form $gx = f$, where the coefficients of this last equation are generic.*

*Proof.* By the construction of the barrier problem we know that each solution to (3.4) together with the solutions to $gx = f$ in $(\mathbb{C}^*)^{n+d+1}$ project down to points in $\mathcal{C} \cap \mathcal{H}$. Conversely, since $\mathcal{H}$ is a generic hyperplane, we are guarenteed that each point in $\mathcal{C} \cap \mathcal{H}$ is the result of a projection from points in $(\mathbb{C}^*)^{n+d+1} \cap \mathcal{E}$ that satisfy the equation $gx = f$. Furthermore we claim that for each point $x^* \in \mathcal{C} \cap \mathcal{H}$, there is a unique point in $(\mathbb{C}^*)^{n+d+1}$ satisfying the equations (3.4) and $gx = f$. To show this, we assume this point is not unique, i.e, that there exists points $(x^*, y^*, \lambda^*)$ and $(x^*, z^*, \mu^*)$ satisfying these properties. Then for each $t \in \mathbb{C}^n$, we have that $(x^*, ty^*, t\lambda^* + (1-t)\mu^*)$ is also a solution satisfying the same properties. This then would produce uncountable many solutions with these properties, which is a contradiction to the fact that there is at most only countable preimages due to the genericty of the data. □

Theorem 3.5 allows us to prove, along with a homotopy, that our problem can be reduced to the diagonal case. In other words, given a quadratic program of the form (3.1) with generic $Q$,$A$,$b$, and $c$ with $Q$ symmetric, the degree of the central curve $\mathcal{C}$ is equal to the case where $Q_{dg}$ is a generic diagonal matrix with the same $A$,$b$, and $c$.

**Theorem 3.5.** *Let $\mathcal{C}$ be the central curve defined by* (3.4) *and $\mathcal{C}_{dg}$ be the central curve defined by* (3.4)*, only where $Q$ is diagonal. Then the degrees of $\mathcal{C}$ and $\mathcal{C}_{dg}$ are equal.*

*Proof.* Let $gx = f$ be a generic hyperplane. The result of Theorem 3.4 guarantees that the degrees of the central curves $\mathcal{C}$ and $\mathcal{C}_{dg}$ are equal to the number of solutions to their respective systems (with $gx = f$) in $(\mathbb{C}^*)^{n+d+1}$. In order to use a deformation argument, we construct the homotopy

$$((tQ_{dg} + (1-t)Q)x)_i x_i + c_i x_i - \lambda - (y^t A)_i x_i = 0,\ i = 1, \ldots, n,$$
$$Ax = b,\ gx = f \text{ for } 0 \leq t \leq 1.$$

Let $S_t^*$ denote the solutions to the corresponding systems in $(\mathbb{C}^*)^{n+d+1}$. To complete the deformation argument, we need to show that $|S_1^*| = |S_0^*|$. Immediately we know from Theorem 3.4 that every solution in $S_1^*$ is a unique continuation of a solution in $S_0^*$. To show equality, we must show that every solution in $S_0^*$ extends to a solution in $S_1^*$. This can only be true if some solutions in $S_0^*$ wander to infinity as $t$ approaches 1. To show this is impossible, consider the system when $t = 1$, and retain only the terms with highest degree in each equation. We are left with

$$(Q_{dg}x)_i x_i - (y^t A)_i x_i = 0,\ i = 1, \ldots, n,\ Ax = 0,\ gx = 0$$

Since we are tracking solutions in the algebraic torus, we can divide the $i$th equation by $x_i$. This gives us a homogenous linear system associated with the following $(n+d+1) \times (n+d)$ matrix

$$\begin{pmatrix} Q_{dg} & -A^t \\ A & 0 \\ g & 0 \end{pmatrix}.$$

To prove our claim, we must show that this matrix has rank $n + d$. However since the data is generic, this matrix must has full rank, proving no points wander to infinity, concluding the proof. $\square$

# Chapter 4

# Degree Bound

## 4.1 Elimination

As explored in the previous chapter, by assuming $Q$ is a generic symmetric matrix we can define a curve by considering $V(J)$, where

$$J = \langle (Qx)_i x_i + c_i x_i - \lambda - (A^t y)_i x_i, i = 1, \ldots, n \text{ and } Ax = b \rangle.$$

We pursue the degree of the central curve given by $\overline{\pi(\mathcal{E})} = \overline{\pi(V(J))}$. To do so we examine the ideal of the central curve $I_{\mathcal{C}}$ which consists of all polynomials that vanish on the central curve. By eliminating the $y$ and $\lambda$ variables from $J$, we obtain an ideal $J_x = J \cap \mathbb{C}[x_1, \ldots, x_n]$. Clearly $J_x \subset I_{\mathcal{C}}$. Chapter 3 proved that it is sufficient to consider just the case where $Q$ is a generic diagonal matrix in order to compute the degree when all data is generic. By letting $Q$ be diagonal, we can clear denominators and simplify what our ideal $J$ looks like. Since the $Ax = b$ equations contain no $y$ or $\lambda$ variables, we discard them for the moment and consider

$$\tilde{J} = \langle q_i x_i^2 - (y^t A_i - c_i)x_i - \lambda \text{ for } i = 1, \ldots, n\rangle.$$

Recall that we let $\pi$ be the projection map from $\mathbb{C}^{n+d+1}$ to $\mathbb{C}^n$ that eliminates the $y$ and $\lambda$ variables. Here we look at the object $\overline{\pi(V(\tilde{J}))}$. Since each defining equation of $\tilde{J}$ is quadratic in $x_i$, we let the two roots to be $w_i$ and $z_i$, $i = 1, \ldots, n$. Hence, we observe that

$$q_i x_i^2 - (y^t A_i - c_i)x_i - \lambda = 0 \text{ if and only if } q_i(w_i + z_i) = y^t A_i - c_i \text{ and}$$
$$q_i w_i z_i = -\lambda, i = 1, \ldots, n$$

Now we define a new ideal $K$ arising from this equivalent conditions,

$$K = \langle q_i(w_i + z_i) - (y^t A_i - c_i), \text{ and } q_i w_i z_i + \lambda \text{ for } i = 1, \ldots, n\rangle.$$

We will define another projection map similar to $\pi$. Let $\phi : \mathbb{C}^{2n+d+1} \to \mathbb{C}^n$. Here $\phi$ projects to the $w$ variables by eliminating $z$, $y$, and $\lambda$. Since $K$ contains $y$ and $\lambda$, we can consider the variety $V(K_w)$ where $K_w = \mathbb{C}[w_1, w_2 \ldots, w_n] \cap K$. We observe the following proposition.

**Proposition 4.1.** $\overline{\phi(V(K))} = \overline{\pi(V(\tilde{J}))}$.

Proposition 4.1 gives us a nice consequence in that $K_w \subset I_{\mathcal{C}}$, since $\overline{\phi(V(K))} = V(K_{\mathcal{C}}) = \overline{\pi(V(\tilde{J}))}$. Regardless, we still do not have a good generating set for $K_w$. To prove Theorem 1.1 however, we can use a subset of $K_w$. We will construct an ideal $T \subset K_{\mathcal{C}}$, where $T$ is generated by very specific polynomials. The rest of the section is dedicated to describing the generators of $T$ step by step alongside an example. Let us recall the problem in Example 3.2, where we have $d = 2$, $n = 5$, and

$$Q = \begin{bmatrix} 11 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \quad c^t = \begin{bmatrix} -2 \\ -3 \\ 4 \\ 9 \\ 1 \end{bmatrix} \quad A^t = \begin{bmatrix} 0 & -1 \\ -1 & 10 \\ 1 & 3 \\ -1 & 11 \\ 7 & -4 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

First we examine $A$. Since $A$ is a $d \times n$ matrix of rank $d$, with $d < n$, the genericity of $A$ implies $d$ randomly selected columns of $A$ will be linearly independent, while $d + 1$ randomly selected columns will be dependent. Following matroid theory terminology we will call such a subset of $d + 1$ columns a circuit. A generic matrix $A$

will contain $\binom{n}{d+1}$ circuits.

Let us take a look at our matrix $A$. Here $A$ has rank 2, so given 3 columns of $A$ we can find a dependence. If we examine the first three columns of $A$, we can see the following dependence,

$$13 \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} -1 \\ 10 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We will refer to this circuit as $I$ with $f_I = (f_{i_1}, f_{i_2}, f_{i_3}) = (13, 1, 1)$. Stating this more generally for each selection of $d+1$ columns of $A$ we let $I = \{i_1, i_2, \ldots, i_{d+1}\}$. Since $A$ is generic we can find a dependence $f_I = (f_{i_1}, f_{i_2}, \ldots, f_{i_{d+1}})$ where

$$f_{i_1} A_{i_1} + f_{i_2} A_{i_2} + \cdots + f_{i_{d+1}} A_{i_{d+1}} = (0, 0, \ldots, 0)^t.$$

For each circuit $I$ we can use the fact that our polynomials in $K$ satisfy the condition $q_i(w_i + z_i) - (y^t A_i - c_i) = 0$ for all $i = 1, \ldots, n$, and examine the following equation

$$f_{i_1}(q_{i_1}(w_{i_1} + z_{i_1}) - c_{i_1} + y^t A_{i_1}) + \cdots + f_{i_{d+1}}(q_{i_{d+1}}(w_{i_{d+1}} + z_{i_{d+1}}) - c_{i_{d+1}} + y^t A_{i_{d+1}}) = 0.$$

The left-hand side will be a linear polynomial in the $w$ and $z$ variables, since the combination of $f_{i_k} A_{i_k}$ is zero for $k = 1, \ldots, d+1$. We can create $\binom{n}{d+1}$ types of these polynomials. In our example, we will have ten of these equations

$$f_{I_1} = 13 \cdot 11(w_1 + z_1) + 1 \cdot 4(w_2 + z_2) + 1 \cdot 1(w_3 + z_3) + 43 = 0$$

$$f_{I_2} = -1 \cdot 11(w_1 + z_1) + 1 \cdot 4(w_2 + z_2) - 1 \cdot 30(w_4 + z_4) + 10 = 0$$

$$\vdots$$

$$f_{I_{10}} = -\tfrac{73}{25} \cdot 1(w_3 + z_3) + 1 \cdot 30(w_4 + z_4) + \tfrac{14}{25} \cdot 3(w_5 + z_5) + \tfrac{103}{25} = 0.$$

Finally we will use $-\lambda = q_1 w_1 z_1 = \cdots = q_n w_n z_n$. Simple manipulation means that we have $q_{i_n} w_{i_n} z_{i_n} = q_{i_m} w_{i_m} z_{i_m} \implies z_{i_n} = \frac{q_{i_m} w_{i_m} z_{i_m}}{q_{i_n} w_{i_n}}$ for $i_n, i_m \in [1, \ldots, n]$. Noting that $f_{I_k}(w_{i_1}, z_{i_1}, \ldots, w_{i_{d+1}}, z_{i_{d+1}})$ is a function of $2d + 2$ variables, we can use this relation to eliminate $z_{i_2}, \ldots, z_{i_{d+1}}$ (or more generally all $z$ variables except one).

We are specific in which $z$ variable we keep. Given our $\binom{n}{d+1}$ different $f_I$ equations, we choose all pairs that match on exactly $d$ indices. Recall that $I$ contains $d+1$ indices, thus we choose the pairs $I_a$ and $I_b$ such that $|I_a \cap I_b| = d$. We denote these equations as $f^*_{I_a}$ and $g^*_{I_b}$ after eliminating all $z_i$ except for $z_{i_1}$, the term with the lowest common index, and clearing denominators.

Let us walk through this process both in general and using our example. Consider two circuits $I_a$ and $I_b$. Let $\{i_1, i_2, \ldots, i_d\}$ be a subset of $\{1, 2, \ldots, n\}$ with $i_1 < i_2 < \cdots < i_d$. Assume $I_a = \{i_1, i_2, \ldots, i_d, i_j\}$ and $I_b = \{i_1, i_2, \ldots, i_d, i_l\}$, with $i_j, i_l \notin i_1, i_2, ..., i_d$, and $i_j < i_l$. Looking at our particular example we see that both $I_1$ and $I_2$ agree on $d = 2$ indices, so we shall consider the pair of them. In this case $I_a = \{1, 2, 3\}$ and $I_b = \{1, 2, 4\}$. In general we can express $f_{I_a}$ and $g_{I_b}$ as

$$f_{I_a} = f_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + f_{i_d} q_{i_d}(w_{i_d} + z_{i_d}) + f_{i_j} q_{i_j}(w_{i_j} + z_{i_j}) + const$$

$$g_{I_b} = g_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + g_{i_2} q_{i_d}(w_{i_d} + z_{i_d}) + g_{i_l} q_{i_l}(w_{i_l} + z_{i_l}) + const.$$

These equations in our example are

$$f_{I_1} = 143(w_1 + z_1) + 4(w_2 + z_2) + (w_3 + z_3) + 43$$

$$f_{I_2} = -11(w_1 + z_1) + 4(w_2 + z_2) - 30(w_4 + z_4) + 10.$$

Once the pair is chosen, both polynomials are then expressed in terms of the lowest common index, which is $i_1$. We use the fact that $q_{i_1} w_{i_1} z_{i_1} = \cdots = q_{i_n} w_{i_n} z_{i_n}$ and after substitution we get the following functions

$$f_{I_a} = f_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + f_{i_d} q_{i_d}(w_{i_d} + \tfrac{q_{i_1} w_{i_1} z_{i_1}}{q_{i_d} w_{i_d}}) + f_{i_j} q_{i_j}(w_{i_j} + \tfrac{q_{i_1} w_{i_1} z_{i_1}}{q_{i_j} w_{i_j}}) + const$$

$$g_{I_b} = g_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + g_{i_d} q_{i_d}(w_{i_d} + \tfrac{q_{i_1} w_{i_1} z_{i_1}}{q_{i_d} w_{i_d}}) + g_{i_l} q_{i_l}(w_{i_l} + \tfrac{q_{i_1} w_{i_1} z_{i_1}}{q_{i_l} w_{i_l}}) + const.$$

Performing these subsitutions on the two equations from our example we get

$$f_{I_1} = 143(w_1 + z_1) + 4(w_2 + \tfrac{11 w_1 z_1}{4 w_2}) + (w_3 + \tfrac{11 w_1 z_1}{w_3}) + 43$$

$$f_{I_2} = -11(w_1 + z_1) + 4(w_2 + \tfrac{11 w_1 z_1}{4 w_2}) - 30(w_4 + \tfrac{11 w_1 z_1}{30 w_4}) + 10.$$

The presence of denominators makes these functions no longer polynomials. We can multiply through and clear denominators. For $f_{I_a}$, we will multiply by $q_{i_j} w_{i_j} \prod_{k=2}^{d} q_{i_k} w_{i_k}$, while for $g_{I_b}$ we multiply by $q_{i_l} w_{i_l} \prod_{k=2}^{d} q_{i_k} w_{i_k}$. After clearing denominators we denote these polynomials $f_{I_a}^*, g_{I_b}^*$. Both $f_{I_a}^*$ and $g_{I_b}^*$ are linear in $z_{i_1}$.

$f_{I_a}^* = (\sum_{p=1}^{d} \frac{w_j}{w_{i_p}} \prod_{k=1}^{d} w_{i_k} + \prod_{k=1}^{d} w_{i_k}) z_{i_1} + (\sum_{p=1}^{d} w_{i_p} \prod_{k=2}^{d} w_{i_k} w_j + \prod_{k=2}^{d} w_{i_k} w_j^2) +$ $const \cdot q_{i_j} w_{i_j} \prod_{k=2}^{d} q_{i_k} w_{i_k}$.

$g_{I_b}^* = (\sum_{p=1}^{d} \frac{w_l}{w_{i_p}} \prod_{k=1}^{d} w_{i_k} + \prod_{k=1}^{d} w_{i_k}) z_{i_1} + (\sum_{p=1}^{d} w_{i_p} \prod_{k=2}^{d} w_{i_k} w_l + \prod_{k=2}^{d} w_{i_k} w_l^2) +$ $const \cdot q_{i_l} w_{i_l} \prod_{k=2}^{d} q_{i_k} w_{i_k}$.

To see a more simple representation of these functions, let us turn back to our example.

$f_{I_1}^* = (44w_1w_2 + 44w_1w_3 + 572w_2w_3)z_1 + (572w_1w_2w_3 + 16w_2^2w_3 + 4w_2w_3^2 + 172w_2w_3)$

$f_{I_2}^* = (-120w_1w_2 + 120w_1w_4 - 1320w_2w_4)z_1 + (-1320w_1w_2w_4 + 480w_2^2w_4 - 3600w_2w_4^2 +$ $1200w_2w_4)$

$f^*_{I_a}$ and $g^*_{I_b}$ will be used to construct our generators for $T$. Since both polynomials are linear in $z_{i_1}$, we can consider them as

$$f^*_{I_a} = Az_{i_1} + B \text{ and } g^*_{I_b} = Cz_{i_1} + D$$

where $A, B, C, D$ are polynomials in $w_1, \ldots, w_n$ only. We are now in a position to eliminate the last $z$ variable $z_{i_1}$ using the equation $C \cdot f^*_{I_a} - A \cdot g^*_{I_b} = BC - AD$. Turning to our general and example case this elimination produces

$$BC - AD = (\textstyle\sum_{p=1}^d w_{i_p} \prod_{k=2}^d w_{i_k} w_j + \prod_{k=2}^d w_{i_k} w_j^2 + \; const\, q_{i_j} w_{i_j} \prod_{k=2}^d q_{i_k} w_{i_k}) \cdot (\sum_{p=1}^d \frac{w_l}{w_{i_p}} \prod_{k=1}^d w_{i_k} + \prod_{k=1}^d w_{i_k}) - (\sum_{p=1}^d \frac{w_j}{w_{i_p}} \prod_{k=1}^d w_{i_k} + \prod_{k=1}^d w_{i_k}) \cdot (\sum_{p=1}^d w_{i_p} \prod_{k=2}^d w_{i_k} w_l + \prod_{k=2}^d w_{i_k} w_l^2 + const\, q_{i_l} w_{i_l} \prod_{k=2}^d q_{i_k} w_{i_k})$$

$$BC - AD = (572w_1w_2w_3 + 16w_2^2w_3 + 4w_2w_3^2 + 172w_2w_3) \cdot (-120w_1w_2 + 120w_1w_4 - 1320w_2w_4) - (44w_1w_2 + 44w_1w_3 + 572w_2w_3) \cdot (-1320w_1w_2w_4 + 480w_2^2w_4 - 3600w_2w_4^2 + 1200w_2w_4).$$

Each term of $BC - AD$ will contain a common factor of $\prod_{k=2}^d w_{i_k}$. We remove this factor, and call the remaining polynomial $R_{I_a, I_b}$. This polynomial will be one of the generators of T. More precisely,

$$T = \langle R_{I_a, I_b} : |I_a \cap I_b| = d \rangle.$$

## 4.2 Creating the Upper Bound

Using a degree reverse lexicographical term ordering we can identify the leading term of $R_{I_a,I_b}$. We will use one where $w_1 > w_2 > \cdots > w_n$. We can observe in Lemma 4.2 exactly what these leading terms will look like.

**Lemma 4.2.** *Given two circuits $I_a$ and $I_b$ that agree on d indices, the initial term of $R_{I_a,I_b}$ is $w_{k_1}^2 w_{k_2} \ldots w_{k_{d+1}}$ where $k_1 < k_2 < \cdots < k_{d+2}$ represent the indices of $I_a \cup I_b$.*

*Proof.* Since the initial term will not be affected by constants, this proof will leave out the scalars $q_{i_k}$ as well as $f_{i_k}$ and $g_{i_k}$ for all $k = 1, \ldots, n$. Furthermore we leave out added constants, because even after clearing denominators these terms cannot possibly be considered the leading term with the term order we have defined. Let $\{i_1, i_2, ..., i_d\}$ be a subset of $\{1, 2, ..., n\}$ such that $i_1 < i_2 < ... < i_d$ Let $i_j, i_l \in \{1, 2, ..., n\}$ such that $i_j, i_l \notin i_1, i_2, ..., i_d$, and $i_j < i_l$. Define $I_a = \{i_1, i_2, \ldots, i_d, i_j\}$ and $I_b = \{i_1, i_2, \ldots, i_d, i_l\}$. Let

$$f_{I_a} = f_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + f_{i_d} q_{i_d}(w_{i_d} + z_{i_d}) + f_{i_j} q_{i_j}(w_{i_j} + z_{i_j})$$

$$g_{I_b} = g_{i_1} q_{i_1}(w_{i_1} + z_{i_1}) + \cdots + g_{i_2} q_{i_d}(w_{i_d} + z_{i_d}) + g_{i_l} q_{i_l}(w_{i_l} + z_{i_l}).$$

From our original equations we have that for any $i, j \in \{1, 2, ..., n\}$, $i \neq j$, $z_i = \frac{z_j w_j}{w_i}$. Since $i_1$ is the lowest index of the $z$ varaibles that both $f_{I_a}$ and $g_{I_b}$ have in common, then we express both $f_{I_a}$ and $g_{I_b}$ as linear polynomials in $z_{i_1}$.

$$f_{I_a} = (z_{i_1} + w_{i_1}) + (\frac{z_{i_1}w_{i_1}}{w_{i_2}} + w_{i_2}) + \cdots + (\frac{z_{i_1}w_{i_1}}{w_{i_d}} + w_{i_d}) + (\frac{z_{i_1}w_{i_1}}{w_{i_j}} + w_{i_j}),$$

$$g_{I_b} = (z_{i_1} + w_{i_1}) + (\frac{z_{i_1}w_{i_1}}{w_{i_2}} + w_{i_2}) + \cdots + (\frac{z_{i_1}w_{i_1}}{w_{i_d}} + w_{i_d}) + (\frac{z_{i_1}w_{i_1}}{w_{i_l}} + w_{i_l}),$$

By cancelling denominators, we obtain two polynomials linear in $z_{i_1}$.

$$f^*_{I_a} = (\textstyle\sum_{p=1}^d \frac{w_j}{w_{i_p}} \prod_{k=1}^d w_{i_k} + \prod_{k=1}^d w_{i_k})z_{i_1} + (\sum_{p=1}^d w_{i_p} \prod_{k=2}^d w_{i_k} w_j + \prod_{k=2}^d w_{i_k} w_j^2).$$

$$g^*_{I_b} = (\textstyle\sum_{p=1}^d \frac{w_l}{w_{i_p}} \prod_{k=1}^d w_{i_k} + \prod_{k=1}^d w_{i_k})z_{i_1} + (\sum_{p=1}^d w_{i_p} \prod_{k=2}^d w_{i_k} w_l + \prod_{k=2}^d w_{i_k} w_l^2).$$

To see the initial term with respect to $w_1 > w_2 > \cdots > w_n$, we must consider several cases.

Case 1: $j < l < i_1$.

The initial term is $(w_{i_1}w_{i_2}\ldots w_{i_{d-1}}w_l)(w_{i_2}w_{i_3}\ldots w_{i_d}w_j^2) = w_j^2 w_l w_{i_1} w_{i_2}^2 w_{i_3}^2 \ldots w_{id-1}^2 w_{i_d}$

Case 2: $j < i_1 < l < i_d$.

The initial term is $(w_{i_1}w_{i_2}\ldots w_{i_{d-1}})(w_{i_2}w_{i_3}\ldots w_{i_d}w_j^2) = w_j^2 w_l w_{i_1} w_{i_2}^2 w_{i_3}^2 \ldots w_{i_{d-1}}^2 w_{i_d}$.

Case 3: $j < i_1 < i_d < l$.

The initial term is $(w_{i_1} w_{i_2} \dots w_{i_d})(w_{i_2} w_{i_3} \dots w_{i_d} w_j^2) = w_j^2 w_{i_1} w_{i_2}^2 w_{i_3}^2 \dots w_{id-1}^2 w_{i_d}^2$

Case 4: $i_1 < j < l < i_d$.

The initial term is $(w_{i_1} w_{i_2} \dots w_{i_{d-1}} w_l)(w_{i_1} w_{i_2} \dots w_{i_d} w_j) = w_j w_l w_{i_1}^2 w_{i_2}^2 w_{i_3}^2 \dots w_{i_{d-1}}^2 w_{i_d}$.

Case 5: $i_1 < j < i_d < l$.

The initial term is $(w_{i_1} w_{i_2} \dots w_{i_d})(w_{i_1} w_{i_2} \dots w_{i_d} w_j) = w_{i_1}^2 w_{i_2}^2 w_{i_3}^2 \dots w_{id-1}^2 w_{i_d}^2 w_j$.

Case 6: $i_d < j < l$.

The initial term is $(w_{i_1} w_{i_2} \dots w_{i_d})(w_{i_1} w_{i_2} \dots w_{i_d} w_j) = w_{i_1}^2 w_{i_2}^2 w_{i_3}^2 \dots w_{id-1}^2 w_{i_d}^2 w_j$.

We can see in each case there is a common factor of $\prod_{k=2}^{d} w_{i_k}$. By removing the common factor from each initial term and relabeling the indices $k_1, k_2, ..., k_{d+2}$ where $k_1 < k_2 < \cdots < k_{d+2}$ represent each index of $f$ in $f_{I_a}, g_{I_b}$ , then each initial term is precisely $w_{k_1}^2 w_{k_2} w_{k_3} \dots w_{k_{d+1}}$. □

One thing to note about the result of Lemma 4.1 is that there is a potential for repeated intial terms. We will drop redundant terms, but Lemma 4.2 calculates how many unqiue initial terms this process produces.

**Lemma 4.3.** *The number of unique initial terms of the generators of $T$ is precisely*

$$\sum_{j=d}^{n-2}\binom{j}{d}.$$

*Proof.* This is easy enough to see, since given any initial term, it is of the form $w_i^2 w_{j_1} w_{j_2} \dots w_{j_d}$, where $i = 1, 2, \dots, n-d-1$ and $i < j_1 < j_2 < \cdots < j_d < n$. Thus for various $i$ we observe that when,

$i = 1$, we have $\binom{n-2}{d}$ choices,

$i = 2$, we have $\binom{n-3}{d}$ choices,

$\vdots$

$i = n-d-1$, we have $\binom{d}{d}$ choices.

Therefore the total number of unique initial terms we observe is

$$\sum_{j=d}^{n-2}\binom{j}{d}.$$

□

Since Lemma 4.2 provides us with the initial terms of the generators of $T$, we will create a new ideal $M$ that is generated by these initial terms. To be more precise we say $M = \langle w_i^2 w_{j_1} w_{j_2} \dots w_{j_d} : i = 1, 2, \dots, n-d-1, i < j_1 < \cdots < j_d < n\rangle$. Recall that degree$(I_{\mathcal{C}})$ = degree$(\text{in}_<(I_{\mathcal{C}}))$. After constructing $T$, we have that $T \subset K_w$.

Through the construction of $M$ we also have that $M \subset \text{in}_<(T) \subset \text{in}_<(K_w)$. The theorem below will provide an irredundant irreducible decomposition of $M$, allowing us to compute degree($M$).

**Theorem 4.4.** *The monomial ideal $M$ has the irredundant irreducible decomposition*

$$M = \bigcap_{k=0}^{n-d-1} \bigcap_{S\subseteq\{k+2,\ldots,n\},|S|=n-d-k-1} (\langle w_j^2 : 0 < j < k+1\rangle + \langle w_s : s \in S\rangle).$$

*Proof.* "$\subseteq$" Let $w_i^2 w_{j_1} w_{j_2} \ldots w_{j_d}$ be in the LHS and fix a $k$ and $S$ in the RHS. We need to show that this element in RHS contains $w_i^2 w_{j_1} w_{j_2} \ldots w_{j_d}$. If $i < k+1$, then we are done, since that implies that one of the $w_j^2$ divides $w_i^2$, since $j$ ranges from 1 to $k$. Therefore assume $i > k$. It follows then that $w_i^2 w_{j_1} w_{j_2} \ldots w_{j_d}$ must have $d+1$ indices from the set $L = \{k+1, k+2, \ldots, n-1\}$. Since $|L| = n-k-2$, the amount of elements in $L$ not equal to the indices $i, j_1, \ldots, j_d$ is $n-k-2-d-1 = n-k-d-3$.

We observe that in the RHS, the ideal $\langle w_s : s \in S\rangle$ will have $n-d-1-k$ indices selected from the set $I = \{k+2, \ldots, n-1\}$. Since $(n-k-d-1)+2 = n-k-d-3$, and $|S|+1 = |L|$, then it follows that at least one index must appear in both our $M$ and RHS element. Therefore we have that $M\subseteq$ RHS.

"$\supseteq$" It suffices to show that for every element *not* in $M$, there exists a component in RHS that does *not* contain it. In order to show this, we characterize the elements not contained in $M$. Let us examine what an element in $M$ must look like. Given an element in $M$, it could have squarefree terms and terms with positive exponents greater than one. We order these separately, and say the element $w$ is of the form

$$w = w_{i_1} w_{i_2} \dots w_{i_n} w_{s_1}^{p_1} w_{s_2}^{p_2} \dots w_{s_m}^{p_m} w_n^p,$$

where $i_1 < i_2 < \cdots < i_n, s_1 < s_2 < \cdots < s_m \in \{1, 2, \dots, n\}$. Given this characterization of elements in $M$, any element $w'$ not in $M$ must either be squarefree or $m \leq d$. This is because clearly $M$ has no squarefree monomials, and given an element $w'$ with $m > d + 1$, we can create this element using the generators of $M$ and scalar multiplication. If $w'$ is squarefree, then the $k = n - d - 1$ element of RHS does not contain $w'$, since that element has no squarefree monomials. If $m \leq d$, then there are only at most $d$ different indices of the $w$ variables that are not squarefree (discluding $w_n$). This means the $k = d$ term of the RHS does not divide $w'$, as there will be some selection of the set $S$ which will avoid producing monomials that divide the remaining nonsquarefree monomials. Therefore $M \supseteq$ RHS. □

With this, we are almost able to prove the main result of the paper, Theorem 1.1. This next proposition establishes that the degree of $M$ is given by the formula used in Theorem 1.1.

**Proposition 4.5.**

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

*Proof.* Given the irredundant irreducible decomposition provided by Theorem 4.4, calculating the degree of $M$ is simple. We begin by noting that each ideal in the decomposition has the same number of variables, and thus they are all top dimensional. Therefore the degree of $M$ will be the sum of the products of the exponents in each monomial ideal in the decomposition. Referencing the form provided in Theorem 4.4, when $k = 0$ we have $\binom{n-2}{d-1}$ components in the decomposition that are squarefree, therefore each of these components contribute only 1 to the degree. When $k = 1$, we have $\binom{n-3}{d-1}$ components in the decomposition that have exactly one squared term, and the rest are squarefree. Thus for these terms, they each contribute 2 to the degree. When $k = 2$, we will similarly have $\binom{n-4}{d-1}$ components in the decomposition that have exactly two squared terms, and the rest are squarefree, making these terms contribute 4 to the degree.

This pattern continues until $k = n - d - 1$, which will produce one component with $2^k$ squared terms, contributing that amount to the degree. By taking the sum of all of these observations we compute the degree of $M$ as

$$\binom{n-2}{d-1} 2^0 + \binom{n-3}{d-1} 2^1 + \cdots + \binom{d-1}{d-1} 2^{n-d-1}.$$

Condensing this expression we get the desired result of

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

□

Finally, we are ready to prove Theorem 1.1.

**Thoerem** 1.1. *Given a generic quadratic program of the form* (1.1), *the degree of the central curve is at most*

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

*Proof.* We first recall $\tilde{J}$ and the Zariski closure $\overline{\pi(\tilde{J})}$. The latter is the central sheet, and is equal to the elimination ideal $\tilde{J}_{y,\lambda}$. We have also shown that the central sheet is defined by $V(K_w)$. By definition the central curve $\mathcal{C}$ is equal to the central sheet intersected with the linear variety defined by $Ax = b$. The genericty of $Ax = b$ guarentees us that the degree of the central sheet and the degree of the central curve are equal.

When the central sheet is intersected by the $d$ equations of $Ax = b$ we get a curve, which has dimension 1. This implies that the dimension of the central sheet is $d+1$. This is also equal to the dimension of $M$, since each component of the irredundant irreducible decomposition of $M$ is defined by setting exactly $n-d-1$ vairables to 0, making the degree $n-(n-d-1)=d+1$.

Because of their construction we know that $M \subset \text{in}_<(T) \subset \text{in}_<(K_w)$. [**?**] tells us that the dimension of a variety is equal to the dimension of any of its initial ideals. Therefore $d+1$, the dimension of the central sheet, is also equal to the dimension of $\text{in}_<(K_w)$. Thus $M$ has the same dimension as $\text{in}_<(K_w)$. If a monomial ideal is contained in another monomial ideal of the same dimension, then the degree of the first is at least the degree of the second. Therefore the degree of the central curve is less than or equal to the degree of $M$, which Proposition 4.5 proved to be the desired result. □

Theorem 1.1 provides us with some corollaries, namely in the cases when $d=1$ or $d=2$.

**Corollary 4.6.** *Given a generic quadratic program of the form* (1.1)*, the degree of the central curve when $d=1$ is at most*

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

This result follows directly from Theorem 1.1, since when $d = 1$, the binomial terms all become 1, just leaving the degree to be the sum of powers of 2. In the case where $d = 2$, the degree also has an alternative formula.

**Corollary 4.7.** *Given a generic quadratic program of the form* (1.1)*, the degree of the central curve when $d = 2$ is at most equal to the Eulerian number $A(n+1, 1)$.*

The closed form of the Eulerian numbers for $A(n + 1, 1)$ is

$$A(n+1,1) = \sum_{k=0}^{1} (-1)^k \binom{n+2}{k} (2-k)^{n+1}.$$

This is equivalent to the formula in Theorem 1.1 when $d = 2$.

# Chapter 5

# Conclusion

The main objective of our research was to prove Conjecture 1.1, which explicitly states the degree of the central curve of a genertic quadratic program. The scope of this thesis limited our exposition to only include the proof of Theorem 1.1, which provides only an upper bound. Chapter 2, Section 1 gave relevant background information on linear and quadratic programming. A main focus of the section was to detail solution methods, primarily interior point methods, since these create the central curve. Also in Section 1 we discussed curvature, and introduced the idea of providing a bound on the curvature of the central curve in quadratic programming. This was inspired by [**?**] , since they were able to achieve this result in the linear case.

Unfortunately, given the result of Theorem 1.1, it is clear that the bound on the curvature will be exponential. Since the degree upper bound is exponential itself, it is not practical for finding a bound on curvature. The only exception is the codimension 2 case, i.e, when $n-d=2$. There the formula in Theorem 1.1 becomes the linear polynomial $n$, since

$$\sum_{k=0}^{1}\binom{n-k-2}{n-2-1}2^k=\binom{n-2}{n-3}2^0+\binom{n-3}{n-3}2^1=n-2+2=n.$$

In Section 2 of Chapter 2 we discussed algebraic geometry. The majority of our proofs use many concepts from this field, and thus the background in this section is key. Since the central curve is viewed as a variety, this research has been heavily ingrained in algebraic geometry, even though at first glance it may seem like optimization based work.

Chapter 3 was dedicated to the task of proving the reduction to diagonal case. As we said, in the generic case the equations of the central curve were too difficult to work with. We were fortunate to discover the sufficiency of the diagonal case in proving our claims. As was the case with most aspects of this research, the quadratic case proved more difficult than the linear case. This meant the reduction to diagonal proof had to involve the implicit function theorem in order to bypass showing the irreducibility of $\mathcal{C}$.

Finally in Chapter 4 we were able to prove Theorem 1.1, our main result. This involved creating a custom elimination ideal $T$ from our central curve equations. Once $T$ was constructed we could create an initial ideal contained in the initial ideal of the central curve. This was a key step, as then a dimension argument was used to create the bound on degree. As we stated, we have a proof showing that our formula is indeed the exact degree, but it extends past the scope of this paper.

During the course of this research, we were inspired to employ mixed volumes. Using Bernstein's theorem on mixed volumes, our future paper will offer an alternative proof for not only for Theorem 1.1, but the main result of [**?**] as well. Both of our results are remarkably similar, and our mixed volume work is just another bridge between the linear and quadratic cases.

# Bibliography

[1] Dimitris Bertsimas, John N. Tsitsiklis, and John Tsitsiklis, *Introduction to linear optimization*, Athena scientific series, vol. 1, Athena, New York, NY, 1997.

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

[3] Little John Cox, David and Donal O'Shea, *Ideals, varieties, and algorithms*, vol. 2, Springer, NY, 1996.

[4] Jesús A. De Loera, Bernd Sturmfels, and Cynthia Vinzant, *The central curve in linear programming*, Found. Comput. Math. **12** (2012), no. 4, 509–540.

[5] Wolfram Decker, Gert-Martin Greuel, Gerhard Pfister, and Hans Schönemann, SINGULAR *3-1-6 — A computer algebra system for polynomial computations*, http://www.singular.uni-kl.de, 2012.

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

[7] David S. Dummit and Richard M. Foote, *Abstract algebra*, vol. 3, John Wiley Sons, Inc., NJ, 2004.

[8] Ezra Miller and Bernd Sturmfels, *Combinatorial commutative algebra*, Springer, NY, 2005.