Newton's Method For
Parametric Center Problems

by

Kok Choon Tan
BSc.(Hon), National University of Singapore
(1984)

Submitted to the Sloan School of Management in
Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Operations Research

at the

Massachusetts Institute of Technology

May 1990

©Kok Choon Tan, 1990

The author hereby grants to MIT the permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author

Operations Research Center
May 4, 1990

Certified by

Professor Robert M. Freund
Elisha Gray II Career Development
Associate Professor of Management Science
Thesis Supervisor

Accepted by

Professor Thomas L. Magnanti
George Eastman Professor of Management Science
Co-director, Operations Research Center
Newton's Method For Parametric Center Problems

Kok Choon Tan
BSc.(Hon), National University of Singapore, 1984

May 1990
Newton’s Method For Parametric Center Problems

by

Kok Choon Tan

BSc.(Hon), National University of Singapore, 1984

Submitted to the Sloan School of Management on May 4, 1990
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Operations Research.

Abstract

In this thesis, we consider the problem of tracing the analytic center of a linear inequality system $Ax \leq b$ as the system is parametrically deformed. Specifically, we are interested in generating approximate centers of the parametric family of systems $(A + \alpha B)x \leq b + \alpha d$, as the parameter $\alpha$ is varied over a specified range. This problem is very closely related to the generalised linear fractional programming problem (GLFP). We propose algorithms, based on Newton’s method, for generating a piecewise linear path of approximate centers.

When only the right hand side is varied parametrically, we obtain, as a by-product, a polynomial-time path-following algorithm for the linear programming problem (LP). To achieve a fixed objective improvement, the algorithm requires $O(m)$ iterations, where $m$ is the number of constraints and each iteration involves the solution of an $n \times n$ system of linear equations. The salient feature of this algorithm is its strict monotonicity, which apparently was not achieved before by other path-following algorithms.

For the case where all the constraints vary parametrically, we obtain, as a by-product, an algorithm for GFLP which achieves a fixed improvement in $O((m + k)k)$ iterations, where $m$ is the total number of constraints, and $k$ is the number of linear fractional functions in the objective.

Thesis Supervisor: Professor Robert M. Freund
Title: Elisha Gray II Career Development Associate Professor of Management Science
Acknowledgements

I would like to thank my thesis advisor, Professor Robert Freund, for introducing this thesis topic to me, for his guidance and for his patience. Without his continuous encouragement and help, none of this would have been possible. I would also like to thank Professor Thomas Magnanti and Dr. Paul Tseng, who despite their tight schudules, had kindly agreed to serve on my thesis committee.

Special thanks go to Professors Chong Chi-Tat and Chew Kim-Lin of the National University of Singapore (NUS) who encouraged me to come to MIT to do Operations Research and kindly served as my guarantors for the NUS Overseas Graduate Scholarship. I gratefully acknowledge the funding by NUS for my entire stay at MIT.

Friends at the Operations Research Center have made my stay there a pleasant experience. My involvement with the Singapore Students Society always make me feel closer to home. My friends in the MIT Tae Kwon Do Club, especially Anthony D’ Amico, Young Soo Ha, Scott Thomson, Ali Kazeroonian, Rubina Arif, Janice Taitel and Karen Needels, who have sweated with me throughout our workouts, also taught me many lessons about life. All these friends have, in many ways, enriched my experience at MIT tremendously. To all of them, I would like to say “thank you”, “kam-sah-mi-dah”, “xie-xie”.

Last, but not the least, I thank my parents, from the bottom of my heart, for all they have given me, and my brothers and sisters, for giving me an enjoyable childhood.
Dedication

To my parents,
who never had a chance to go to school
but fully understand the importance and benefits of educating their children
to the best of their capabilities.
Contents

1 Introduction ............................................................ 6
  1.1 Scope of Thesis .................................................... 7
  1.2 Motivation: Theory ............................................. 8
  1.3 Motivation: Applications ..................................... 10
    1.3.1 Linear Fractional Programming ....................... 10
    1.3.2 von Neumann Model of Economic Expansion .......... 11
    1.3.3 Generalized Linear Fractional Programming .......... 12
  1.4 Methodology of Thesis ...................................... 13
  1.5 Notation ......................................................... 14
  1.6 Outline of Thesis ............................................. 15

2 Literature Review .................................................. 16
  2.1 Projective Scaling Algorithms ............................... 17
  2.2 Affine Scaling Algorithm .................................... 24
  2.3 Path-following Algorithms .................................. 26
    2.3.1 Barrier Function Approach ............................. 27
    2.3.2 Method-of-Centers Approach ......................... 32
  2.4 Center Finding Algorithms .................................. 40
3 The Center Problem 42
3.1 The Analytic Center 42
3.1.1 Properties of the Analytic Center 43
3.2 Approximate Centers and Measures of Closeness 44
3.2.1 First Measure of Closeness 44
3.2.2 Second Measure of Closeness 46
3.2.3 Third Measure of Closeness 48
3.2.4 Property of Approximate Centers 50
3.3 Newton’s Method for Center Problem 50

4 Right-Hand-Side Parametric Center Problem 52
4.1 Introduction 52
4.2 Properties of RHS Parametric Centers 58
4.3 Translations 61
4.4 Strict Improvement Theorem 63
4.5 Lower and Upper Bounds 70
4.6 A RHS Parametric Center Algorithm 77
4.6.1 Algorithm RHSPCP 78
4.6.2 Algorithmic Performance 81
4.7 A Strictly Monotonic Algorithm for Linear Programming 85

5 Multiple Constraints Parametric Center Problem 89
5.1 Introduction 89
5.2 Assumptions and Notation 91
5.3 Main Results and Algorithm 93
5.3.1 A Parametric Center Algorithm 95
5.4 Preliminary Lemmas ........................................... 98
5.5 Proofs of Main Theorems ................................. 105

6 Applications and Conclusion .............................. 113
6.1 Linear Programming ........................................... 113
6.2 Linear Fractional Programming ...................... 117
6.3 The von Neumann Model of Economic Expansion .... 122
6.4 Solution of the von Neumann Model by a Method of Parametric Centers ........................................ 127
6.5 Generalized Linear Fractional Programming .......... 132
6.6 Conclusion ....................................................... 134
Chapter 1

Introduction

Consider a system of linear inequalities $Ax \leq b$, where $A$ is an $m \times n$ matrix and $b$ is a column m-vector. Assuming that the interior of the polyhedral set $\mathcal{X} \equiv \{x \in \mathbb{R}^n | Ax \leq b\}$ is nonempty and bounded (therefore $m > n$), the (analytic) center of the system of linear inequalities is the unique optimal solution $\hat{x}$ of the following nonlinear program (see Sonnevend[60,61]), which we call the center problem.

$$CP: \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i$$
$$\text{subject to} \quad Ax + s = b$$
$$s > 0$$

Observe that the objective function is strictly concave so that the center $\hat{x}$ is uniquely characterised by the Karush-Kuhn-Tucker conditions:

$$\hat{s} = b - A\hat{x} \quad > \quad 0 \quad \quad (1.1)$$
$$A^T\hat{s}^{-1}e \quad = \quad 0, \quad \quad (1.2)$$
where $\hat{S} = \text{diag}(\hat{s})$ denotes the diagonal matrix with diagonal entries equal to the components of $\hat{s}$ and $e$ denotes the vector of all ones, that is, $e := (1, 1, \ldots, 1)^T$.

1.1 Scope of Thesis

Since the seminal work of Karmarkar [40], it has been shown by various researchers that the concept of centers plays an important role in the development of efficient interior-point algorithms for linear and convex quadratic programming problems (Anstreicher [2], Bayer and Lagarias [5,6], Freund [17, 21], Jarre [38], Mehrotra and Sun [50], Renegar [57], Sonnevend [60,61] and Vaidya [66], among many others). For example, Karmarkar’s projective scaling algorithm transforms the problem at each iteration so as to place the current point at the center of the transformed feasible set. The research for this thesis was, therefore, carried out with the goal of gaining a better insight into the behavior of interior-point algorithms through the study of the path of centers as the system of linear inequalities is parametrically deformed.

In this thesis, we are primarily interested in tracing the path of centers as the system of linear inequalities changes parametrically. Specifically, let $\hat{x}_\alpha$ denote the center of the system $(A + \alpha B)x \leq b + \alpha d$, where $B$ is an $m \times n$ matrix, $d$ is a column $m$-vector and $\alpha$ is a scalar parameter. As the parameter $\alpha$ increases (or decreases) monotonically, we are interested in generating a (piecewise linear) path of approximate centers $\bar{x}_\alpha$ such that $\bar{x}_\alpha$ is close (with respect to some suitable norm, depending on $\alpha$ and $\bar{x}_\alpha$) to $\hat{x}_\alpha$ for all values of $\alpha$ over a specified range. We call this problem the Parametric Center
Problem. Our emphasis will be on the analysis of the algorithmic complexity of the algorithms we propose, which are based on Newton's method.

1.2 Motivation: Theory

For the case when $A = \begin{bmatrix} A' \\ c \end{bmatrix}$, $B = 0$, $b = \begin{bmatrix} b' \\ 0 \end{bmatrix}$ and $d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, the path of centers $\Gamma : \alpha \mapsto \hat{x}_\alpha$ has been studied analytically by Megiddo [48], Bayer and Lagarias [5,6] and Anstreicher [2]. These studies have, to a great extent, provided much insight into the theoretical behavior of interior-point algorithms for the linear programming problem (LP). In Bayer and Lagarias [5,6], the authors specified a linear objective function $c^T x$, and then examined the vector field of directions (assigned to points in the interior of the feasible region) induced by an interior point algorithm which seeks to maximize (or minimize) the objective over the feasible region. They showed that the vector field generates trajectories leading from any interior point to a boundary point which solves the given linear program. The trajectory that passes through the center of the feasible region is called the central trajectory, and it can be shown that it is exactly the path of centers $\Gamma$.

Megiddo [48] studied the infinitesimal version of the logarithmic barrier function approach (see Gill et al. [28]) to maximize the linear objective function over the feasible region. This and the infinitesimal versions of both Karmarkar's projective algorithm and the affine scaling algorithm (Dikin [12], Barnes [3], Vanderbei et al. [68]), when started at the center, follow the central trajectory exactly. In addition, Barnes et al. [4] showed that the affine
scaling algorithm can be made polynomial-time by adding “centering” steps, so that the resulting algorithm in effect tries to follow the central path. Consequently, these and most other interior-point algorithms may be viewed as algorithms that follow the path of centers.

In Anstreicher [2], rather than fixing the objective function $c^T x$ and considering the resulting trajectories on $\mathcal{X}$, one of which is central, the author instead considered simultaneously the central trajectories for all possible cost vectors $c$. He showed that the vector field of negative Newton directions for the logarithmic barrier function (with no objective component) flows from the center of $\mathcal{X}$, along central trajectories, to solutions of every possible linear program on $\mathcal{X}$.

Renegar [57] presented the first complexity analysis for a linear programming algorithm which uses Newton’s method to explicitly follow the central path. This algorithm is reminiscent of the method of centers for solving convex programming problems proposed by Huard [36] in 1967. However, there was no complexity analysis in Huard’s work. Renegar’s path-following algorithm requires $O(\sqrt{m})$ iterations to achieve a given objective decrease, compared to $O(m)$ iterations for Karmarkar’s algorithm.

Motivated by the above, we tried using Newton’s method to explicitly follow the path of centers for the slightly more general case where $d$ is arbitrary and $B = 0$. We call this the Right-Hand-Side Parametric Center Problem. We obtained an algorithm which generalizes Renegar’s algorithm, and that when applied to the linear programming problem, yields a strictly monotonic algorithm which requires $O(m)$ iterations to achieve a given objective value improvement. In many applications the strict monotonicity property
is desirable (see Huard [36]).

Further encouraged by our success with the Right-Hand-Side Parametric Center Problem, we therefore tried the same methodology to the general Parametric Center Problem. We obtain an algorithm which achieves a fixed improvement in the parametric value in $O(m(\sqrt{l} + k))$ iterations, where $m$ is the total number of constraints, $l$ is the number of varying constraints, and $k$ is the number of nonzero rows in the matrix $B$.

### 1.3 Motivation: Applications

Parametric center problems arise very naturally. We have seen that the linear programming problem (LP) is closely related with the right-hand-side parametric center problem (RHSPCP), and an algorithm for RHSPCP can also be used for solving LP. In this section, we shall see three problems that are closely related to the general parametric center problem, namely, the linear fractional programming problem (LFP), the von Neumann model of economic expansion, and the generalized linear fractional programming problem (GLFP). The solutions of these problems using the method of parametric centers are given in Chapter 6.

#### 1.3.1 Linear Fractional Programming

Suppose we are interested in solving the following linear fractional program

$$LFP : \max_{x} \frac{f - c^{T}x}{d^{T}x - h} \quad s.t. \ Ax \leq b.$$
Then it is easy to see that $LFP$ is equivalent to

$$\begin{align*}
\max_{x, \alpha} & \quad \alpha \\
\text{s. t.} & \quad Ax \le b \\
& \quad (c + \alpha d)^T x \le (f + \alpha h),
\end{align*}$$

which leads to the parametric center problem

$$\begin{align*}
\max & \quad \sum_{i=1}^{m} \ln s_i + \ln g \\
\text{s. t.} & \quad Ax + s = b \\
& \quad s > 0 \\
& \quad (c + \alpha d)^T x + g = (f + \alpha h) \\
& \quad g > 0.
\end{align*}$$

### 1.3.2 von Neumann Model of Economic Expansion

The von Neumann model of economic expansion may be cast as the following optimization problem (see Section 6.3)

$$\begin{align*}
EEP : \max_{x} \min_{i} \frac{B_i x}{A_i x} \\
\text{s. t.} & \quad e^T x = 1 \\
& \quad x \ge 0,
\end{align*}$$

where $A, B \in \mathbb{R}_{+}^{m \times n}$ are given (non-negative) input and output matrices of the economy. We can easily show that $EEP$ is equivalent to the following optimization problem
\[
\begin{align*}
\max_{x, \alpha} \quad & \alpha \\
\text{s. t.} \quad & (B - \alpha A)x \geq 0 \\
\quad & e^T x = 1 \\
\quad & x \geq 0.
\end{align*}
\]

This leads to the parametric center problem

\[
\begin{align*}
\max \quad & \sum_{i=1}^{m} \ln s_i + \sum_{i=1}^{n} \ln x_i \\
\text{s. t.} \quad & (B - \alpha A)x - s = 0 \\
\quad & e^T x = 1 \\
\quad & s, x > 0.
\end{align*}
\]

1.3.3 Generalized Linear Fractional Programming

Suppose we are interested in solving a generalized linear fractional program in the following format

\[
GLFP : \quad \max_x \min_i \quad \frac{f_i - C_i x}{D_i x - h_i} \quad (i = 1, 2, \ldots, k)
\]

\[
\text{s. t.} \quad A x \leq b.
\]

We see that GLFP is equivalent to

\[
\begin{align*}
\max_{x, \alpha} \quad & \alpha \\
\text{s. t.} \quad & A x \leq b \\
\quad & (C + \alpha D)x \leq f + \alpha h,
\end{align*}
\]
which leads to the parametric center problem

$$\max \sum_{i=1}^{m} \ln s_i + \sum_{j=1}^{k} \ln g_j$$

s. t. \hspace{1cm} Ax + s = b \\
\hspace{1cm} s > 0 \\
(C + \alpha D)^T x + g = (f + \alpha h) \\
\hspace{1cm} g > 0.$$

1.4 Methodology of Thesis

The methodology of this thesis is as follows. We shall first show that if the parameter $\alpha$ increases by a small quantity $\beta > 0$, then the two centers $\hat{x}_\alpha$ and $\hat{x}_{\alpha+\beta}$ will be close to each other. The closeness is with respect to some appropriately chosen measures. Three equivalent measures of closeness will be developed in this thesis. (See Chapter 3.)

Therefore, if $\bar{x}$ is close to the center $\hat{x}_\alpha$, and if the increase $\beta$ is small, $\bar{x}$ will be sufficiently close to the center $\hat{x}_{\alpha+\beta}$ so that a Newton step from $\bar{x}$ will bring us to a point $\bar{x}_{\text{new}}$ which is again close to the (new) center $\hat{x}_{\alpha+\beta}$. Again, appropriate measures of closeness to the center are used.

As our emphasis is on the complexity analysis of our algorithms, our results differ from classical results; instead of using arbitrary $\epsilon$ and $\delta$, we quantify the increase $\beta$ such that the two centers $\hat{x}_\alpha$ and $\hat{x}_{\alpha+\beta}$ are close, in terms of the number of constraints, the size of the input data and other specific constants.
1.5 Notation

The following notation is used throughout this thesis.

For any vector $s \in \mathbb{R}^k$, we let the corresponding upper-case letter $S$ denote the $k \times k$ diagonal matrix with $i^{th}$ diagonal entry equals to $s_i$. We write $S := diag(s)$. If $Q$ is a positive definite matrix, the $Q$-norm $||v||_Q$ is given by

$$||v||_Q = \sqrt{v^TQv}.$$ 

The usual $l_1-$, $l_2-$ and $l_\infty$-norms will be denoted by $|| \cdot ||_1$, $|| \cdot ||_2$ and $|| \cdot ||_\infty$ respectively. Given a matrix $M$, we let $M_i$ denote the $i^{th}$ row of $M$ and $M_i^T$ denote the transpose of $M_i$. The usual (Euclidean) matrix norm of $M$ is given by

$$||M|| = \sup_{||x||=1} ||Mx||.$$ 

Note that if $M$ is a diagonal matrix, then $||M|| = \max_i |m_{ii}|$. Similarly, the $Q$-norm of $M$ is given by

$$||M||_Q = \sup_{||x||=1} ||Mx||_Q.$$ 

The matrix $P_M := I - M(M^TM)^{-1}M^T$ denotes the orthogonal projection matrix which projects onto the null space of $M^T$. The vector of all ones (of appropriate dimension) shall be denoted by $e$, that is, $e := (1, 1, \ldots, 1)^T$.

We shall let $\hat{x}$ denote the center of the system $Ax \leq b$ and let $\hat{x}_\alpha$ denote the center of the system $(A + \alpha B)x \leq b + \alpha d$. With respect to the system $(A + \alpha B)x \leq b + \alpha d$, for $x$ satisfying

$$s_\alpha := (b + \alpha d) - (A + \alpha B)x > 0,$$
we let $Q_\alpha(x)$ denote the negative of the Hessian of the logarithmic barrier function

$$f_\alpha(x) = \sum_{i=1}^{m} \ln[(b + \alpha d) - (A + \alpha B)x]_i$$

for the center problem on the system $(A + \alpha B)x \leq b + \alpha d$.

1.6 Outline of Thesis

The rest of this thesis is organized as follows. In Chapter 2, we give an extensive literature review. In Chapter 3, we present some results concerning the center problem on an inequality system and give three equivalent measures of closeness to the center. In Chapter 4, we present an algorithm for the case when only the right-hand-side (RHS) of the linear inequality system is parametrically changed (Right-Hand-Side Parametric Center Problem) and describe an application of the algorithm to the linear programming problem. In Chapter 5, we present an algorithm for the case when one or more constraints are parametrically changed (Multiple Constraints Parametric Center Problem). This case is closely related to the generalised linear fractional programming problem (GLFP). Finally, in Chapter 6, we present applications of the results of Chapter 5 to the von Neumann model of economic expansion, to the linear programming problem (LP) to obtain an algorithm for LP that requires $O(\sqrt{m})$ iterations to achieve a fixed improvement in the objective value, to the linear fractional programming problem (LFP), and the generalised linear fractional programming problem (GLFP).
Chapter 2

Literature Review

The Simplex method, developed by Dantzig in 1947, has been used for solving linear programming problems since its inception and, when implemented in commercial codes, has proven to be very efficient in the solution of real world problems. However, in the worst case its computational complexity is exponential (Klee-Minty [43]). Until now the existence of a pivoting rule that will make the Simplex method polynomial-time is an open question.

The exponential behavior of the Simplex method raised the question whether linear programming is solvable in polynomial time. This question was answered in the affirmative by Khachian [42] with the Ellipsoid method in 1979. On the other hand, attempts to implement the Ellipsoid method efficiently more or less failed in comparison with the Simplex method.

In 1984, Karmarkar [40] proposed a Projective Scaling algorithm for solving linear programming problems. Its worst case complexity is even better than the Ellipsoid method under suitable assumptions. Furthermore, he claimed that his method outperforms the Simplex method by a large fac-
tor, and this created much controversy in the mathematical programming community. On the other hand, this has also stimulated much research in interior-point methods for linear and convex quadratic programming problems, linear complementarity problems, and certain nonlinear programming problems.

2.1 Projective Scaling Algorithms

Karmarkar’s algorithm applies directly only to linear programs in a special form:

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = 0, \\
& \quad e^T x = n, \\
& \quad x \geq 0.
\end{align*}
\]

But Karmarkar showed that any linear program can be transformed to this required form without destroying the polynomial complexity of the algorithm. Assuming that the optimal value is known and is zero, the basic projective scaling method obtains a feasible rational solution \( x \) to the linear system

\[
Ax = 0, \ e^T x = n, \ x \geq 0, \ c^T x \leq 0,
\]

for rational \( m \times n \) matrix \( A \), \( n \)-vector \( c \), with \( Ae = 0 \). Starting from the initial solution \( e \), a sequence of interior solutions \( x \in \Omega \cap \Sigma_+ \) is generated, where \( \Omega := \{ x \mid Ax = 0 \} \), \( \Sigma := \{ x \mid e^T x = n, \ x \geq 0 \} \), and \( \Sigma_+ := \{ x \mid e^T x = n, \ x > 0 \} \). At each iteration, the problem is transformed by a
projective transformation (hence the name Projective Scaling) $T_x : \Sigma \rightarrow \Sigma$ which depends on the current iterate $x > 0$ and is defined by

$$T_x(y) := \frac{n}{e^T X^{-1} y} X^{-1} y,$$

where $X = \text{diag}(x)$ is a diagonal matrix with diagonal entries $x_i$, $i = 1, \ldots, n$. The inverse of $T_x$ is given by

$$T_x^{-1} := \frac{n}{e^T X y} X y.$$

The transformation $T_x$ maps the current iterate $x$ to $e$, the center of the simplex $\Sigma$, and $\Omega$ to $\bar{\Omega} \equiv \{ y \mid \bar{A} y = 0 \}$, where $\bar{A} := AX$. The transformed problem becomes

$$\begin{align*}
\text{minimize} & \quad \bar{c}^T y \\
\text{subject to} & \quad \bar{A} y = 0, \\
& \quad e^T y = n, \\
& \quad y \geq 0,
\end{align*}$$

where $\bar{c} := X c$, with $e$ being a feasible solution. Note that the transformed problem has the same structure as the original problem.

Under the assumption that the optimal value is zero, the transformed problem is equivalent to the original problem in the sense that $y \in \Omega \cap \Sigma$ is an optimal solution of the original problem if and only if $T_x(y) \in \bar{\Omega} \cap \Sigma$ is an optimal solution of the transformed problem.

In order to drive $e^T x$ to 0, at each iteration, $\bar{c}^T y$ is locally minimized over the set $B(e, \alpha r) \cap \bar{\Omega}$, where

$$B(e, \alpha r) := \{ x \mid e^T x = n, \| x - e \| \leq \alpha r \}$$

18
is a ball of radius $\alpha r$ around the center $e \in \Sigma$ in the affine hull of the simplex $\Sigma$, and $r := \sqrt{\frac{n}{n-1}}$, $0 < \alpha < 1$ (for example, Karmarkar set $\alpha = 1/4$). In fact, $B(e, r)$ is the largest sphere centered at $e$ inside the simplex $\Sigma$. The minimization on $B(e, \alpha r) \cap \tilde{\Omega}$ is easier than on $\Sigma \cap \tilde{\Omega}$ and is achieved by taking a step to the boundary of the ball in the direction of the negative projection of the transformed cost vector $\bar{c}$ onto the nullspace $\{ y \mid A\bar{y} = 0; e^T \bar{y} = 0 \}$. Its solution $\bar{y}$ is then projectively transformed back using $T_x^{-1}$ to the original space to obtain the next iterate $\bar{x} = T_x^{-1}(\bar{y})$. Each iteration consists of $O(n^3)$ arithmetic operations, with the bulk of the computational effort spent on the computation of the projection of the transformed cost vector.

In order to derive a polynomial bound on the number of iterations, Karmarkar suggested using a logarithmic potential function to monitor the progress of the algorithm, as the objective function does not necessarily decrease in each iteration. The exact potential function used is the following.

$$ f(x) := n \ln(c^T x) - \sum_{i=1}^n \ln x_i $$

Observe that this potential function is invariant under the projective transformation used in the transformation of the space. That is, under $T_x$, $f$ is mapped to $\bar{f}$ given by

$$ \bar{f}(y) := n \ln(c^T y) - \sum_{i=1}^n \ln y_i + constant. $$

Note that $f(x)$ approaches $\infty$ as $x$ approaches the boundary of the feasible solution set except when it approaches any point $x$ for which $c^T x = 0$, in which case $f(x)$ approaches $-\infty$. Karmarkar showed that at each iteration, the potential function must decrease by at least a fixed constant if there
is a feasible solution satisfying (2.1). Therefore, the algorithm will drive the potential function to $-\infty$ if there is a feasible solution satisfying (2.1). Assuming without loss of generality that $A$ and $c$ are integral and letting $L$ denote the size of the problem instance, the algorithm stops after $O(nL)$ iterations with either a solution $\hat{x} \in \Omega \cap \Sigma$ such that $c^T\hat{x} = 0$ or it will verify that $\min \{ c^T x \mid x \in \Omega \cap \Sigma \} > 0$. (Actually, it stops with a solution $\hat{x}$ such that $c^T\hat{x} < 2^{-L}$. However, since the data is assumed to be integral, we can easily round the solution off to get $\hat{x}$ such that $c^T\hat{x} = 0$.)

Karmarkar also devised an intricate modification of his projective method which reduces the complexity of the method by a factor of $\sqrt{n}$. The key to this theoretical improvement is the observation that the ball $B(e, \alpha r)$ may be replaced by a suitable ellipsoid

$$E(e, \alpha r/2) := \{ y \mid e^T y = n, (y - e)^T \Xi^{-2}(y - e) \leq (\alpha r/2)^2 \}$$

where $\Xi := \text{diag}(\xi)$, with $(\xi_i)^2 \geq 1/2$. Then

$$B(e, \alpha r/4) \subseteq E(e, \alpha r/2) \subseteq B(e, \alpha r)$$

implies that the number of iterations remains essentially unchanged at $O(nL)$, but the projection (of the cost vector) in each iteration can be computed much more easily using rank-one updates, each rank-one update requiring $O(n^2)$ arithmetic operations. Karmarkar was able to show that in $K$ iterations, the total number of rank-one updates necessary is $O(\sqrt{n}K)$. Therefore, on average each iteration needs only $O(n^{2.5})$ arithmetic operations compared to $O(n^3)$ in the original approach.

Todd and Burrell [63] showed that, in the transformed space, the objective function moves by a substantial fraction of the way towards a lower bound
given by dual variables, rather than zero, and consequently ensures that the potential function decreases by a constant at each iteration. This also applies when the optimal value is non-zero and unknown, thus allowing them to extend the basic projective scaling algorithm, using dual variables that are naturally generated, to an algorithm that applies even in degenerate cases and when the optimal value is unknown.

Anstreicher [1] demonstrates that Karmarkar’s projective algorithm is fundamentally an algorithm for linear fractional programs on the simplex:

$$FLP: \quad z^* := \min \left\{ \frac{c^T x}{d^T x} \mid Ax = 0, \; e^T x = n, \; x \geq 0 \right\}.$$  

The fact that a standard form linear program can be easily converted into a linear fractional program was actually noted by Karmarkar in his original unpublished version of [40]. Associated with $FLP$ is a parametric family of linear programs

$$LP(z): \quad \theta(z) := \min \left\{ (c - zd)^T x \mid Ax = 0, \; e^T x = n, \; x \geq 0 \right\}$$

defined for each $z \in \mathbb{R}$. It is easy to see that $x^*$ solves $FLP$ if and only if $x^*$ solves $LP(z^*)$ and that $\theta(z^*) = 0$. At each iteration $k \geq 0$, starting with a feasible $x^k > 0$ and a lower bound $z^k \leq z^*$ (assumed $Ae = 0$ and $z^0 \leq z^*$ is known), the problem is transformed by Karmarkar’s projective transformation $T_{x^k}$. Note that $FLP$ is transformed into another fractional linear program

$$FLP^\prime: \quad \min \left\{ \frac{\bar{c}^T y}{d^T y} \mid \bar{A}y = 0, \; e^T y = n, \; y \geq 0 \right\}$$

where $X := \text{diag}(x^k)$, $\bar{c} := Xc$, $\bar{d} := Xd$ and $\bar{A} := AX$. Therefore, the same iterative procedure can be repeated. A revised lower bound $\bar{z}$, $z^k \leq \bar{z} \leq z^*$,
is computed by applying either one of the two following relaxations of $LP(z)$.

**Method I: Ellipsoid Bound** (Anstreicher)

$$
\theta_1(z) := \min\{ \ (c - zd)^T x \mid Ax = 0 \ e^T x = n, \ ||x - e|| \leq R \}
$$

where $R := \sqrt{n(n - 1)}$ is the radius of the smallest ball in the affine hull of the simplex $\Sigma$ centered at $e$ that circumscribe $\Sigma$. Note that $R/r = n - 1$.

**Method II: Simplex Bound** (Todd and Burrell)

$$
\theta_2(z) := \min\{ \ (c_q - zd_q)^T x \mid x \in \Sigma \}
$$

where $c_q$ and $d_q$ denote the projections of $c$ and $d$ onto the subspace $\{ x \mid Ax = 0 \}$.

Note that both $\theta_1(z)$ and $\theta_2(z)$ are concave, $\theta_2(z)$ is piecewise linear, and $\theta_1(z) \leq \theta_2(z) \leq \theta(z)$ for every $z \in \mathbb{R}$. Using either of the two methods, $\bar{z}$ is obtained such that $\theta_1(\bar{z}) \leq 0$. The new point

$$
y(\alpha) := e - \alpha \frac{(\bar{c}_p - \bar{z} \bar{d}_p)}{||\bar{c}_p - \bar{z} \bar{d}_p||},
$$

is then computed with $\alpha = (\frac{n-1}{2n-2})r$, where $\bar{c}_p$ and $\bar{d}_p$ are the projections of $\bar{c}$ and $\bar{d}$ onto the subspace $\{ y \mid \bar{A} y = 0, \ e^T y = 0 \}$. (When $\bar{z} = z^* = 0$ and $\alpha = r/4$, this corresponds exactly to the step taken in Karmarkar’s algorithm.) Moving from $e$ to $y(\alpha)$ corresponds to taking a projected gradient step in the problem $LP(\bar{z})$ which reduces the “gap” $(c - \bar{z}d)^T x - \theta(\bar{z})$ in $LP(\bar{z})$ by at least a factor of $1 - \alpha/R$

$$
\frac{(\bar{c} - \bar{z} \bar{d})^T y(\alpha) - \theta(\bar{z})}{(\bar{c} - \bar{z} \bar{d})^T e - \theta(\bar{z})} \leq 1 - \frac{\alpha}{R}.
$$
Finally, after setting \( x^{k+1} := T_{z}^{-1}(y(\alpha)) \), \( z^{k+1} := \tilde{z} \), the iterative step is repeated. At each iteration, the potential function

\[
f(x; z) := n \ln((c - zd)^T x) - \sum_{i=1}^{n} \ln x_i
\]

in uniformly reduced by at least a constant \( > 1 - \ln 2 > 0.3 \).

Observing that the constraint \((c - vd)^T x \leq 0\) can trivially be added to \(FLP\), where \( v = c^T e/d^T e \) denotes the objective value at \( e \), Anstreicher was able to modify his algorithm to ensure the monotonicity of the objective value. Essentially, when \((c - vd)^T (c - \tilde{z}d) \geq 0\) the usual step is taken, and when \((c - vd)^T (c - \tilde{z}d) < 0\) a modified step is taken, corresponding to a centering step which holds the objective value constant while the potential function is reduced. This fact was noted by Goldfarb and Xiao [34], and using this fact they were able to extend this even further to obtain a strictly monotonic algorithm for \(FLP\) and thus a strictly monotonic polynomial time algorithm for linear programming.

Other variants of the projective method for linear programming include Barnes [3], Barnes et al. [4], Cavalier and Soyster [9], Freund [21], Gay [25] and Vanderbei et al. [68]. Freund [21] develops an analog of Karmarkar's algorithm for inequality constrained linear programs

\[
\max \{ c^T x \mid Ax \leq b \}
\]

which works directly in the space of linear inequalities. The main idea was the recognition of the importance of centering (i.e. mapping the current iterate to the analytic center) in Karmarkar’s algorithm. Based on classical polarity theory for convex sets (see Rockafellar [59] and Grunbaum [35]), a class of projective transformations for centering a polytope was also obtained.
Gay [25] develops a version of the projective method directly for linear programs in standard form

$$\min \{ c^T x \mid Ax = b, \ x \geq 0 \}$$

without a priori knowledge of the optimal function value. He uses a variation of Todd and Burrell's approach to compute a sequence of improving bounds on the optimal objective value after applying, at each iteration, a projective transformation $\Phi : \mathbb{R}^n \to \mathbb{R}^{n+1}$, defined by

$$\Phi(y) := \frac{1}{e^T X^{-1} y + 1} \begin{bmatrix} X^{-1} y \\ 1 \end{bmatrix},$$

where $x > 0$ is the current solution and $X := \text{diag} (x)$, which was first suggested by Karmarkar[40] to transform standard form linear programs to the form required for his projective algorithm.

### 2.2 Affine Scaling Algorithm

A most notable variant of the projective scaling algorithm is the so called affine scaling algorithm proposed independently by various authors (Barnes [3], Cavalier and Soyster [9] and Vanderbei et al. [68]) soon after the development of the projective scaling algorithm. The affine scaling algorithm works directly on standard form linear programs and is conceptually much simpler.

Starting from some interior point, to improve the objective function value, one can move in the opposite direction of the objective function gradient projected onto the affine hull of the feasible polyhedron. However, a substantial improvement is possible only if the current solution is sufficiently distant from
the boundary of the feasible polyhedron. Therefore, the current solution will
have to be "centered" in the polyhedron in some way.

Inspired by the projective scaling method, the various authors of the affine
scaling algorithm proposed to map the current solution $x > 0$ to $e$ using the
linear transformation $T \equiv T_x : \mathbb{R}^n \to \mathbb{R}^n$, defined by

$$T(y) := X^{-1}y$$

where $X := \text{diag} (x)$. Then, a projected gradient step is taken in the
transformed space, and the solution $\hat{y}$ is then transformed back to the original
space. As in the projective scaling method, the main effort at each iteration
is in the computation the projected gradient. The objective function value
strictly improves at each iteration and under primal and dual non-degeneracy
assumptions, the algorithm was shown to converge. However, there is strong
evidence to support the belief that this algorithm is in fact exponential in
the worst case (see Megiddo and Shub [49]).

It is interesting to note that the affine scaling algorithm in fact predated
by almost 20 years the projective scaling method. A Soviet mathematician,
I. I. Dikin [12,13], had proposed the basic affine scaling algorithm in 1967
and given a proof of convergence in 1974 which does not require the dual
non-degeneracy assumption. (See also Vanderbei and Lagarias [67].)

The linear transformation in the affine scaling algorithm corresponds
actually to a rescaling of the variables rather than centering. Barnes et
al. [4] modified the basic affine scaling algorithm by introducing centering
steps that actually keep the iterates centered in a precise mathematical way,
and so obtained a variant which is polynomial-time.
2.3 Path-following Algorithms

While the polynomial time bounds of Karmarkar's algorithm and all its projective variants were derived by using some appropriately chosen potential functions to monitor the algorithm's performance, (hence they have also been called potential reduction algorithms by Ye [74]) there is another class of interior-point algorithms for linear programming and convex quadratic programming which are based on the path-following approach proposed by various authors. The path being followed by all the algorithms in this class is, in their idealized infinitesimal versions, the path of centers or the central trajectory. The study of paths or trajectories in the interior of the feasible region of a linear program, including the central trajectory, has been carried out by Bayer and Lagarias [5,6,7], Megiddo [48], Megiddo and Shub [49], and Anstreicher [2].

There seems to be two not too apparently distinct approaches taken by the various researchers who propose these path-following algorithms. One approach is the logarithmic barrier function approach (Gill et al. [28], Kojima et al. [45], Gonzaga [29], Megiddo [48], and Monteiro and Adler [52,53]) and the other is the method of centers approach (Huard [36], Sonnevend [60,61], Renegar [57], Vaidya [66], Jarre [38] and Mehrotra and Sun [50]). However, all these path-following algorithms share the following common features: Newton's method is used in the solution of the subproblems and some concept of closeness to the central path is necessary.

It should be noted that not only are the path-following algorithms closely related to Newton's method, the projective algorithms are also closely
related to Newton’s method. Gill et al. [28] showed an equivalence between Karmarkar’s algorithm and the projected Newton barrier method they proposed. Bayer and Lagarias [7] showed that Karmarkar’s algorithm (with line search) may be viewed as a Newton’s method (with line search) to find the center of an unbounded polyhedron in a projectively transformed space where the set of optimal solutions of the linear program is projectively transformed to the hyperplane at infinity.

2.3.1 Barrier Function Approach

Since the introduction of Karmarkar’s interior-point algorithm for linear programs, Gill et al. [28] were the first to consider the logarithmic barrier function method for linear programs. This method was first proposed by Frish [23] in 1955. To solve a linear program

\[
\min \{ c^T x \mid Ax = b, \ x \geq 0 \},
\]

the barrier function method consists of solving a sequence of nonlinear sub-problems

\[
P_\mu : \min \{ c^T x - \mu \sum_{i=1}^n \ln x_i \mid Ax = b \}
\]

for a sequence of barrier parameters \( \mu > 0 \), with \( \mu \to 0 \). Assuming the sequence of solutions \( x(\mu) \) generated converges, the rate of convergence is \( O(\mu) \) even in the case of non-unique optima (see e.g., Jittorntrum [39]). Gill et al. proposed to solve \( P_\mu \) using a projected Newton’s method, in which the search direction \( d \) at the current iterate \( x > 0 \) is found from solving the quadratic program
\[
\min \left\{ g^T d + \frac{1}{2} d^T H d \mid Ad = 0 \right\},
\]
where \( g = g(x) := c - \mu X^{-1} e \) and \( H = H(x) := uX^{-2} \) are the gradient and the Hessian of the objective function in \( P_\mu \) respectively. They also showed that the search directions of Karmarkar's projective algorithm and their projected Newton barrier method are equivalent if the barrier parameter is chosen appropriately.

Megiddo [48] studied the infinitesimal version of the (weighted) barrier function method to solve linear programs in standard form (2.2) and in the following symmetric form

\[
(P) : \quad \text{maximize } c^T x \\
subject to \ Ax \leq b \\
x \geq 0.
\]

The dual of \((P)\) is

\[
(D) : \quad \text{minimize } b^T y \\
subject to \ A^T y \geq c \\
y \geq 0.
\]

For fixed \( \mu > 0 \), \((P)\) and \((D)\) can be approximated respectively by

\[
(P_\mu) : \quad v_\mu := \text{maximize } c^T x + \mu (\sum_{i=1}^n \ln x_i + \sum_{i=1}^m \ln u_i) \\
subject to \quad Ax + u = b \\
x, \ u > 0
\]
and

\[(D_\mu) : \quad z_\mu := \text{minimize} \quad b^T y - \mu (\sum_{i=1}^{m} \ln y_i + \sum_{i=1}^{n} \ln v_i) \]

subject to \quad A^T y + v = c
\]

\[y, \quad v > 0.\]

The Karush-Kuhn-Tucker stationary conditions for both \((P_\mu)\) and \((D_\mu)\) are symmetrical and can be written as

\[
\begin{align*}
\mu X^{-2} x - A^T y &= -c \\
Ax + \mu Y^{-2} y &= b \\
x, \quad y &> 0
\end{align*}
\]

(KKT)

i.e. \(x > 0 \quad (y > 0)\) is an optimal solution of \((P_\mu) \quad ((D_\mu))\) if and only if there exist some \(y > 0 \quad (x > 0)\) such that (KKT) holds. If \(x\) and \(y\) are optimal solutions for \((P_\mu)\) and \((D_\mu)\), then the gap between the optimal values of \((P_\mu)\) and \((D_\mu)\)

\[z_\mu - x_\mu = (m + n)\mu (1 - \ln \mu)\]

and the duality gap between the optimal values of the pair of dual linear programs \((P)\) and \((D)\)

\[b^T y - c^T x = (m + n)\mu.\]

Note that the two gaps depend only on the dimensions \(m\) and \(n\) and not on the size of the data. It can then be shown that if the paths of solutions \(x(\mu)\) and \(y(\mu)\) exist for every \(\mu > 0\), then they (not just their objective values \(c^T x(\mu)\) and \(b^T y(\mu)\)) converge to an optimal solution as \(\mu \to 0\).
Megiddo also showed that the only barrier or penalty functions that give the *primal-dual symmetric optimality conditions* \((KKT)\) are of the form \(f(\xi) := \kappa \ln(|\xi|)\), where \(\kappa\) is some constant. Such functions are appropriate only as barrier functions (i.e. for interior point methods) and not as penalty functions (for exterior point methods). Differential properties of the solution paths and behavior near vertices were also studied.

We observe that this approach can be easily generalized to the case where the constraints are weighted for the barrier function. That is, given a vector of weights \(w = (w^P, w^D) > 0\), the weighted logarithmic barrier function approximation of \((P)\) and \((D)\) are

\[
(P_\mu(w)) : \begin{align*}
\text{maximize} & \quad c^T x + \mu \left( \sum_{i=1}^n w_i^P \ln x_i + \sum_{i=1}^m w_i^D \ln u_i \right) \\
\text{subject to} & \quad Ax + u = b \\
& \quad x, u > 0
\end{align*}
\]

and

\[
(D_\mu(w)) : \begin{align*}
\text{minimize} & \quad b^T y - \mu \left( \sum_{i=1}^m w_i^D \ln y_i + \sum_{i=1}^n w_i^P \ln v_i \right) \\
\text{subject to} & \quad A^T y + v = c \\
& \quad y, v > 0.
\end{align*}
\]

Then the Karush-Kuhn-Tucker stationary conditions become

\[
X(A^T y - c) = \mu w^P
\]

\[
Y(b - Ax) = \mu w^D.
\]

30
Based on the primal-dual barrier function framework proposed by Megiddo, Kojima et al. [45] developed an algorithm that works simultaneously with a pair of primal and dual linear programs

\[(P) : \quad \max \{ c^T x \mid Ax = b, \ x > 0 \},\]

\[(D) : \quad \min \{ b^T y \mid A^T y + z = c, \ z \geq 0 \}.\]

Their algorithm requires $O(nL)$ iterations. Subsequently, Monteiro and Adler [52] improved on the convergence rate, so that the algorithm requires $O(\sqrt{n}L)$ iterations, by working closer to the path of solutions $w(\mu) = (x(\mu), y(\mu), z(\mu))$, $\mu > 0$, characterized by the Karush-Kuhn-Tucker conditions

\[Z X e - \mu e = 0,\]
\[A x - b = 0, \ x > 0 \quad (KKT)\]
\[A^T y + z - c = 0,\]

(where $z = \mu X^{-1}e$ is a dual slack vector) for the logarithmic barrier subproblem

\[(P_\mu) : \quad \min \ c^T x - \mu \sum_{i=1}^{n} \ln x_i \quad \text{s.t.} \ Ax = b, \ x > 0.\]

The direction taken is the Newton direction associated with the system (KKT). Using rank-one update techniques (presented in Karmarkar [40] and Gonzaga [29]) to exploit the special structure of the matrix to be inverted in each iteration, Monteiro and Adler further reduce the average work per iteration to $O(n^{2.5})$ arithmetic operations. Thus their algorithm overall requires
$O(n^3L)$ arithmetic operations. The same computational complexity was also achieved by a primal algorithm proposed by Gonzaga [29]. This is also based on following the central trajectory using a barrier function approach which works only in the primal space and has the following characteristics: there is no need for a priori knowledge of the optimal value and (for minimization problems) no lower bounds to the optimal value are used in the algorithm.

### 2.3.2 Method-of-Centers Approach

The method of centers was first proposed by Huard [36] in 1967 for solving convex programming problems

\[
\text{minimize} \quad f(x) \\
\text{subject to} \quad g_i(x) \leq 0, \quad i = 1, \ldots, m
\]

where $f, g_i, i = 1, \ldots, m$ are convex functions on $\mathbb{R}^n$. In Huard's method of centers, a sequence of interior points \(\{x^k\}_{k=1}^\infty\) are generated, where $x^{k+1}$ is the "center" of $X_k := \{x \mid g_i(x) < 0, \ i = 1, \ldots, m, \ f(x) < f(x^k)\}$ and $x^0 := \text{"center" of } X := \{x \mid g_i(x) < 0, \ i = 1, \ldots, m\}$ is assumed given. (See Figure 1.) The "center" is defined to be the unique point that maximizes its "distance" from the boundary of the feasible region. Two distance functions were suggested in [36], one of which is

\[
D(x) := (f(x^k) - f(x)) \prod_{i=1}^m (-g_i(x)).
\]
Figure 1. Huard's method of centers
For a linear inequality system $Ax \leq b$, the analytic center was defined by Sonnevend [60,61] to be the solution to the nonlinear optimization problem

$$CP: \quad \text{maximize} \quad \sum_{i=1}^{m} \ln(b - Ax);$$
$$\text{subject to} \quad Ax < b.$$

Assuming \( \{x \in \mathbb{R}^n | Ax < b \} \) is non-empty and bounded, the center is uniquely defined. Observe that the objective function in the center problem \( CP \) is strictly concave, and therefore the center is uniquely characterized by the Karush-Kuhn-Tucker conditions

$$A^T \Delta^{-1}(x)e = 0,$$
$$\Delta(x) := \text{diag}(b - Ax) > 0.$$

An algorithm to follow the path of analytic centers explicitly, as a linear objective function constraint is shifted in parallel, was proposed but no complexity analysis was given.

Renegar [57] independently proposed a similar algorithm and presented the first complexity analysis for a path-following algorithm which uses Newton's method to explicitly follow the path of analytic centers as the objective function is shifted in parallel. To solve a linear program in inequality form

$$\text{maximize} \quad c^T x$$
$$\text{subject to} \quad Ax \geq b$$

where \( x \in \mathbb{R}^n \) and \( A \) is an \( m \times n \) matrix, Renegar's algorithm follows approximately the centers of a sequence of systems \( \bar{A}x \geq b^{(k)} \), where \( \bar{A} \).
and \( b^{(k)} \) denote the \( 2m \times n \) matrix and \( 2m \)-vector

\[
\tilde{A} := \begin{bmatrix} A \\ c^T \\ \vdots \\ c^T \end{bmatrix}, \quad b^{(k)} := \begin{bmatrix} b \\ \delta^{(k)} \\ \vdots \\ \delta^{(k)} \end{bmatrix}
\]

and \( \{\delta^{(k)}\} \) is an increasing sequence of parameters. Therefore, Renegar's algorithm is essentially a parametric center algorithm where only the right-hand-side varies parametrically. The interior of \( \{x|\tilde{A}x \geq b^{(0)}\} \) is assumed to be non-empty and bounded, and the center of the initial system is assumed given. Renegar showed that any linear program can be reformulated such that these assumptions are satisfied without destroying the computational complexity.

Starting from a point \( \bar{x}(k) \) sufficiently close (with an appropriate definition of closeness) to the center \( \hat{x}(k) \) of the system \( \tilde{A}x \geq b^{(k)} \), the parameter \( \delta^{(k+1)} \) is chosen in the range \( \delta^{(k)} < \delta^{(k+1)} < c^T \bar{x}(k) \) to ensure that \( \{x|\tilde{A}x \geq b^{(k+1)}\} \) is non-empty and such that the two centers \( \hat{x}(k) \) and \( \hat{x}(k+1) \) are sufficiently close so that Newton's method when applied will converge, and then a Newton step is taken from \( \bar{x}(k) \) to find the center of the system \( \tilde{A}x \geq b^{(k+1)} \). Having obtained \( \bar{x}(k+1) \) which is close to the center \( \hat{x}(k+1) \), the process is repeated. (See Figures 2 and 3.) In fact, the algorithm works with \( \delta^{(k+1)} := \frac{1}{13\sqrt{m}}c^T \bar{x}(k) + (1 - \frac{1}{13\sqrt{m}})\delta^{(k)} \), so that

\[
\delta^* - c^T \bar{x}(k) \leq \frac{47}{92} (1 - \frac{1}{28\sqrt{m}})^k (\delta^* - \delta^0),
\]
Figure 2. Renegar's Algorithm
Figure 3. One Step in Renegar's Algorithm.
where $\delta^*$ denotes the optimal value of the linear program, and therefore in $O(\sqrt{mL})$ iterations, the algorithm finds a solution $x$ such that $\delta^* - c^T x < 2^{-L}$. Each iteration requires $O(m^3)$ arithmetic operations to solve a system of linear equations (to find the Newton step). Therefore, the total complexity is $O(m^{3.5}L)$ arithmetic operations. Subsequently, Vaidya [66] improved it using rank-one update techniques to reduce the average work per iteration to $O(m^{2.5})$ arithmetic operations, so that the total complexity is $O(m^3L)$ arithmetic operations.

This approach was also taken by Jarre [38] and Mehrotra and Sun [50] to develop polynomial-time algorithms for convex quadratic programming problems. The two algorithms are very similar and are essentially generalization of Renegar's algorithm to the cases where the objective function or the constraints are convex quadratic functions. The concept of analytic center for a convex quadratic system generalizes that for linear system in a natural way, i.e. maximizes the logarithm of the product of the constraint slacks.

Note that the method-of-centers approach is conceptually very simple and direct, and algorithms from this approach are primal or feasible-point algorithms. In the idealized version of the method of centers, where exact centers are used as in Huard [36], the objective function value strictly improves at each iteration. However, in using approximate centers so as to obtain polynomial time bounds as in Renegar [57] and Vaidya [66], it is conceivable that the objective function value may not necessarily improve at each iteration. (See Figure 4.)
Figure 4. Parameter $\alpha$ not necessarily increases in Renegar's Algorithm.
2.4 Center Finding Algorithms

Having recognized the importance of analytic centers for polynomial-time linear programming algorithms, Vaidya [65] and later Freund [17] developed algorithms for finding the analytic center of a linear inequality system. Vaidya [65] constructs the Newton direction from the current iterate and then performs an inexact line search in this direction. He showed that his algorithm converges linearly.

Freund [17] generalized the concept of the analytic center to the weighted center, or simply w-center, for a polyhedral system and obtained results regarding contained and containing ellipsoids centered at the w-center. He also exhibited an elementary projective transformation that transforms the current point to the w-center of the polyhedral system in the transformed space. He then applied these results to obtain a projective transformation algorithm, analogous to but more general than Karmarkar's algorithm, for finding the w-center of a polyhedral system. The algorithm was shown to be superlinearly convergent. At each iteration, the algorithm either improves the weighted logarithmic barrier function value by a fixed amount, or at a linear rate (which approaches 1) of improvement. The direction taken at each iteration was shown to be positively proportional to the Newton direction. Therefore, with a line search, Freund's algorithm specializes to Vaidya's algorithm, where all the weights on the constraints are the same. An important and useful characteristics of Freund's algorithm is that after a fixed number of iterations, the current iterate of the algorithm may be taken to be an approximate w-center in the sense that one can easily construct

40
well-scaled contained and containing ellipsoids centered at the current iterate whose scale factor depends only on the weights used but not the data.

Another algorithm for the center finding problem $CP$ was developed by Censor and Lent [10] while working on the entropy maximization problem. However, there was no complexity analysis of their algorithm.
Chapter 3

The Center Problem

Recall that the center problem on a system of linear inequalities $Ax \leq b$ is the following optimization problem.

$$CP: \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i$$
$$\text{subject to} \quad Ax + s = b$$
$$s > 0.$$ 

In this chapter, we present some known results concerning the center problem $CP$, and give three equivalent measures of closeness to the center which we will use in this thesis.

3.1 The Analytic Center

Assuming that the set $\{x | Ax < b\}$ is nonempty and bounded, the solution $\hat{x}$ of the center problem $CP$ exists uniquely, and is called the analytic center of the system $Ax \leq b$. (Sonnevend [60,61].) We shall refer to it simply as
the center. Since the objective function is strictly concave, the center \( \hat{x} \) is uniquely characterised by the Karush-Kuhn-Tucker conditions

\[
\hat{s} = b - A\hat{x} > 0 \quad \quad (3.1)
\]

\[
A^T\hat{S}^{-1}e = 0. \quad \quad (3.2)
\]

For \( x \) satisfying \( s = b - Ax > 0 \), let \(-Q(x)\) be the Hessian of the barrier function for the center problem \( CP \) at \( x \), that is,

\[
Q(x) = A^T S^{-2} A. \quad \quad (3.3)
\]

### 3.1.1 Properties of the Analytic Center

One particularly important property of the center is the following. (See Sonnevend [60,61], and also Freund [17].)

**Lemma 3.1.1** ([60,61], [17] Theorem 2.1) Let \( \hat{x} \) denote the center of the linear inequality system \( Ax \leq b \). Let \( \mathcal{X} \) := \{ \( x \in \mathbb{R}^n \mid Ax \leq b \) \}, and define the ellipsoids \( E_{in} \) and \( E_{out} \) by

\[
E_{in} := \{ x \in \mathbb{R}^n \mid \| x - \hat{x} \|_{Q(\hat{x})} \leq \sqrt{\frac{m}{m - 1}} \},
\]

\[
E_{out} := \{ x \in \mathbb{R}^n \mid \| x - \hat{x} \|_{Q(\hat{x})} \leq \sqrt{m(m - 1)} \}.
\]

Then \( E_{in} \subset \mathcal{X} \subset E_{out} \).

That is, we can construct contained and containing ellipsoids centered at the center. Note that \( E_{out} \) is an enlargement of \( E_{in} \) with an enlargement factor of \( m - 1 \), i.e., \( (E_{out} - \hat{x}) = (m - 1)(E_{in} - \hat{x}) \). Next, the following
lemma shows that the slacks of all feasible points of the linear inequality system $Ax \leq b$ are contained in a simplex. Therefore, as a corollary, we can bound the slacks of any feasible point $x \in \mathcal{X}$.

**Lemma 3.1.2 ([57] Proposition 3.1, [17] Proposition 2.1)** Suppose $\hat{x}$ is the center of the linear inequality system $Ax \leq b$. Let $\hat{s} = b - A\hat{x}$. For any $x$ satisfying $Ax \leq b$, let $s = b - Ax$. Then $e^T \hat{S}^{-1}s = m$, $s \geq 0$.

**Corollary 3.1.1** With the same conditions and definitions as Lemma 3.1.2, we have $0 \leq s_i \leq m \hat{s}_i$ for all $i = 1, 2, \ldots, m$.

### 3.2 Approximate Centers and Measures of Closeness

There are various ways to measure the closeness of a point $\bar{x}$ to the center $\hat{x}$ of an inequality system $Ax \leq b$. We shall give three equivalent measures in this section. A direct way is to use some norm. In fact, this way of measure was used by many authors of path-following algorithms (Renegar[57], Gonzaga[29], Kojima et al.[45], Monteiro and Adler[52,53], Jarre[38], and Mehrotra and Sun[50], among others). The first measure of closeness to the center is defined in a similar way as follows.

#### 3.2.1 First Measure of Closeness

For all $v \in \mathbb{R}^n$, define the $Q(x)$-norm (Hessian norm) of $v$ by

$$||v||_{Q(x)} := \sqrt{v^T Q(x) v}. \quad (3.4)$$
Definition: We say that \( \bar{x} \) is a \( \delta \)-approximate center if \( ||\bar{x} - \hat{x}||_{Q(\bar{x})} \leq \delta \).

(In [22], and in this thesis, we use \( \delta = 1/21 \).)

We have the following lemma which gives some basic inequalities.

Lemma 3.2.1 ([22] Lemma 3.2) Suppose \( \bar{x} \in \mathbb{R}^n \) is given such that \( \bar{s} = b - A\bar{x} > 0 \) and let \( Q(\bar{x}) \) be defined by (3.3). Then for any \( \tilde{x} \in \mathbb{R}^n \) such that \( ||\bar{x} - \tilde{x}||_{Q(\bar{x})} \leq \delta < 1 \), we have

(i) \( \bar{s} = b - A\tilde{x} > 0 \),

(ii) \( ||\bar{S}^{-1}\bar{S}|| \leq 1/(1 - \delta) \),

(iii) \( ||\bar{S}^{-1}\tilde{S}|| \leq 1 + \delta \),

(iv) \( ||v||_{Q(\bar{x})} \leq \frac{1}{1 - \delta} ||v||_{Q(\tilde{x})} \), for all \( v \in \mathbb{R}^n \),

(v) \( ||v||_{Q(\bar{x})} \leq (1 + \delta) ||v||_{Q(\tilde{x})} \), for all \( v \in \mathbb{R}^n \),

(vi) \( ||\bar{x} - \tilde{x}||_{Q(\bar{x})} \leq \frac{\delta}{1 - \delta} \),

where \( Q(\bar{x}) \) is defined by (3.3).

Proof: Observe that

\[
||\bar{x} - \tilde{x}||_{Q(\bar{x})} = ||(\bar{x} - \tilde{x})^T A^T \bar{S}^{-2} A(\bar{x} - \tilde{x})||^{1/2}
= ||(\bar{s} - \tilde{s})^T \bar{S}^{-2} (\bar{s} - \tilde{s})||^{1/2}
= ||\bar{S}^{-1} (\bar{s} - \tilde{s})|| \leq \delta < 1.
\]

Therefore, for each \( i = 1, 2, \ldots, m \),

\[
\left| \frac{\bar{s}_i - \tilde{s}_i}{\bar{s}_i} \right| \leq \delta,
\]

45
and hence,

$$(1 - \delta)\bar{s}_i \leq \bar{s}_i \leq (1 + \delta)\bar{s}_i.$$ 

Parts (i)–(iii) follow immediately. To prove Part (iv), we observe that

$$||v||_{Q(\bar{x})} = \sqrt{v^T A^T \bar{S}^{-2} A v} = ||\bar{S}^{-1} A v||.$$ 

Therefore, from Part (ii),

$$||v||_{Q(\bar{x})} = ||\bar{S}^{-1} A v|| \leq ||\bar{S}^{-1} \bar{S}|| \cdot ||\bar{S}^{-1} A v|| \leq \frac{1}{1 - \delta} ||v||_{Q(\bar{x})}.$$ 

The proof of Part (v) is the same as Part (iv) but uses Part (iii), and Part (vi) follows from Part (iv) immediately. \[\Box\]

### 3.2.2 Second Measure of Closeness

The second measure of closeness to the center is defined as follows. For \(\bar{x} \in \{x | Ax < b\}\), let \(\bar{S} := \text{diag}(b - A\bar{x})\) be the diagonal matrix of the slacks at \(\bar{x}\). Define

$$Q := (1/m)A^T \bar{S}^{-2} A, \tag{3.5}$$

$$y = y(\bar{x}) := (1/m)A^T \bar{S}^{-1}e, \tag{3.6}$$

and

$$\gamma = \gamma(\bar{x}) := \sqrt{\frac{(m - 1)y^T Q^{-1} y}{1 - y^T Q^{-1} y}}. \tag{3.7}$$

Note that \(Q := (1/m)Q(\bar{x})\) and \(y^T Q^{-1} y < 1\) for \(\bar{x} \in \{x | Ax < b\}\), so that \(\gamma(\bar{x}) ((3.7))\) is well-defined, and from the Karush-Kuhn-Tucker conditions
((3.1)-(3.2)), \( y(\tilde{x}) = 0 \) and so \( \gamma(\tilde{x}) = 0 \). In [17], the scalar \( \gamma(\tilde{x}) \) is used to measure the closeness of \( \tilde{x} \) to the center \( \hat{x} \). We found this measure to be very convenient because we do not need to know the exact center. To see the equivalence between the two measures of closeness, we have the following lemmas. The proofs can be found in the Appendix of [22]. First we need to refer to two functions defined in [17] (equations (6.3) and (6.4) of [17] respectively). Define, for \( h > 0 \),

\[
  p(h) := \frac{h - \ln(1 + h)}{h^2},
\]

\[
  q(h) := \frac{1}{2} \left(1 + hp(h) - \sqrt{1 + (hp(h))^2}\right). 
\]

**Lemma 3.2.2 ([17] Lemma 8.1, [22] Lemma 3.3).**

Let \( \hat{x} \) denote the center of the system \( Ax \leq b \). Let \( h > 0 \) be a given parameter. Suppose \( \gamma = \gamma(\tilde{x}) \leq q(h) \), where \( q(h) \) is given in (3.8)—(3.9). Then

\[
  ||\tilde{x} - \hat{x}||_{Q(\mu)}^2 \leq \left(\frac{m}{m-1}\right) \frac{h^2(1+\gamma^2)}{(1-h\gamma)^2}. \quad \Box
\]

That is, if \( \gamma(\tilde{x}) \) is small then \( \tilde{x} \) is a \( \delta \)-approximate center for some small \( \delta \). For example, if \( \gamma(\tilde{x}) \leq .0072 \) (taking \( h = 0.03 \)), then \( ||\tilde{x} - \hat{x}||_{Q(\mu)} \leq 1/21 \).

On the other hand, if \( \tilde{x} \) is a \( \delta \)-approximate center then we have the following lemma which says that \( \gamma(\tilde{x}) \) should be small.

**Lemma 3.2.3 ([22], Lemma 3.4)** Suppose \( ||\tilde{x} - \hat{x}||_{Q(\mu)} \leq \delta < 1/2 \). Then

\[
  \gamma(\tilde{x}) \leq a + \sqrt{2a}, \text{ where } a = \frac{\delta^2}{2(1 - \delta)(1 - 2\delta)}. \quad \Box
\]

For example, when \( \delta = 1/21 \), then \( \gamma(\tilde{x}) \leq 0.0527 \).
3.2.3 Third Measure of Closeness

Observe that in the definition of $\gamma(\bar{x})$ ((3.7)), $y^TQ^{-1}y$ may be expressed as

$$y^TQ^{-1}y = \frac{1}{m}e^T S^{-1} A Q(\bar{x})^{-1} A^T S^{-1} e$$

$$= \frac{1}{m}||A^T S^{-1} e||_{Q(\bar{x})^{-1}}^2.$$ 

We note that $A^T S^{-1} e$ is the gradient of the logarithmic barrier function in the center problem $CP$ at $\bar{x}$. Therefore, when $y^TQ^{-1}y$ is sufficiently small, $\gamma(\bar{x})$ is almost the same as the size of the gradient of the logarithmic barrier function measured in the norm of the Hessian inverse. We therefore define the third measure of closeness as follow.

For any $\bar{x} \in \mathbb{R}^n$ satisfying $\bar{s} = b - A\bar{x} > 0$, let $Q(\bar{x})$ be defined by (3.3). Define $\tau = \tau(\bar{x})$ by

$$\tau = \tau(\bar{x}) := ||A^T S^{-1} e||_{Q(\bar{x})^{-1}}. \quad (3.10)$$

Remark: As we shall see in the next section, this measure of closeness is also closely assssociated with Newton's method for the center problem $CP$.

It is easy to see that $\tau(\bar{x}) < \sqrt{m}$ for $\bar{x} \in \{x|Ax < b\}$, and

$$\gamma^2 = \frac{(m - 1)\tau^2}{m - \tau^2}; \quad \tau^2 = \frac{m\gamma^2}{m - 1 + \gamma^2}.$$ 

Therefore, the following lemma, which shows the equivalence between the measures $\gamma$ and $\tau$, follows easily.
Lemma 3.2.4 Let $\gamma = \gamma(\bar{x})$ be defined by (3.5)–(3.7) and $\tau = \tau(\bar{x})$ be defined by (3.10). Then

(i) if $\tau \leq 1$, then $\gamma \leq \tau$,

(ii) $\tau \leq \gamma \sqrt{m/(m-1)}$. $\square$

We observe that the factor of $\sqrt{m/(m-1)}$ in the above lemma is almost equal to 1, even for moderate values of $m$. For example, $\sqrt{m/(m-1)} \leq 1.05$ if $m > 10$.

From Lemma 3.2.2 and Lemma 3.2.4, we have

Lemma 3.2.5 Let $\tau = \tau(\bar{x})$ be defined by (3.10). Suppose $\tau \leq 1/76$. Then $||\bar{x} - \hat{x}||_{Q(\bar{x})} \leq 1/12$.

Proof: From Lemma 3.2.4(i), we have $\gamma \leq \tau \leq 1/76$. Thus from Lemma 3.2.2 with $h = 1/18$,

$$||\bar{x} - \hat{x}||_{Q(\bar{x})}^2 \leq \frac{2h^2(1 + \gamma^2)}{(1 - h\gamma)^2} \leq \left(\frac{1}{12}\right)^2. \square$$

Lemma 3.2.6 Assume that $m > 10$. Let $\tau = \tau(\bar{x})$ be defined by (3.10). Suppose $||\bar{x} - \hat{x}||_{Q(\bar{x})} \leq 1/21$. Then $\tau \leq 0.056$.

Proof: From Lemma 3.2.3 we conclude that $\gamma = \gamma(\bar{x}) \leq a + \sqrt{2a}$, where $a = \frac{(1/21)^2}{2(1 - (1/21))(1 - 2(1/21))} = \frac{1}{1760}$. Thus by Lemma 3.2.4(ii), we have $\tau \leq \sqrt{m(m-1)} \gamma \leq (1.05)(0.053) < 0.056$. $\square$

49
3.2.4 Property of Approximate Centers

Next, analogous to Lemma 3.1.1, we have the following property of an approximate center.

Lemma 3.2.7 ([17] Theorem 8.1, [22] Lemma 5.2).

Let \( \hat{x} \) denote the center of the linear inequality system \( Ax \leq b \). Suppose \( \bar{x} \in \{ x \in \mathbb{R}^n \mid Ax < b \} \) is given such that \( ||\bar{x} - \hat{x}||_{Q(\bar{z})} \leq \delta < 1 \). Define the ellipsoids \( F_{in} \) and \( F_{out} \) by

\[
F_{in} := \{ x \in \mathbb{R}^n : ||x - \bar{x}||_{Q(\bar{z})} \leq 1 \},
\]

\[
F_{out} := \{ x \in \mathbb{R}^n : ||x - \bar{x}||_{Q(\bar{z})} \leq (1 + \delta)\sqrt{m(m-1) + \delta} \}.
\]

Then \( F_{in} \subset \mathcal{K} \subset F_{out} \).

That is, we can construct contained and containing ellipsoids centered at a \( \delta \)-approximate center. Note that \( F_{out} \) is an enlargement of \( F_{in} \) with an enlargement factor of \( O(m) \). As we shall see in the following chapters, the elliptical bounds of Lemma 3.2.7 are very useful in our derivation of complexity bounds.

3.3 Newton’s Method for Center Problem

Also, we will need the following important useful result of Renegar[57] on the convergence of Newton’s method, giving a region and rate of convergence of Newton’s method for the center problem \( CP \).

Lemma 3.3.1 ([57] Theorem 3.2) Suppose \( \bar{x} \) satisfies \( \bar{s} - A\bar{x} > 0 \) and \( \varepsilon := ||\bar{x} - \hat{x}||_{Q(\bar{z})} < 1 \), where \( \hat{x} \) denotes the center of the system \( Ax \leq b \). Let
\( \tilde{\eta} := -Q(\bar{x})^{-1}A^T\bar{S}^{-1}e \) be the Newton step from \( \bar{x} \) in the center problem \( CP \) and let \( \tilde{y} := \bar{x} + \tilde{\eta} \). Then

\[
||\tilde{y} - \hat{x}||_{Q(\bar{x})} \leq \frac{(1 + \epsilon)^2}{1 - \epsilon}\epsilon^2.
\]

**Remarks:** The Newton step \( \tilde{\eta} \) is the solution of an unconstrained quadratic approximation to the center problem \( CP \),

\[
\tilde{\eta} = \arg \max_{\eta \in \mathbb{R}^n} \{ e^T \bar{S}^{-1} A\eta - \frac{1}{2} \eta^T Q(\bar{x}) \eta \},
\]

where \( A^T\bar{S}^{-1}e \) and \( Q(\bar{x}) \) are the gradient and the negative of the Hessian of the logarithmic barrier function \( \sum_{i=1}^{m} \ln(b_i - A_i x) \) at \( \bar{x} \). The solution \( \tilde{\eta} \) can be obtained by solving an \( n \times n \) system of linear equations

\[
Q(\bar{x}) \eta = -A^T\bar{S}^{-1}e.
\]

With respect to the third measure of closeness to the center, we see that

\[
\tau(\bar{x}) = ||\tilde{\eta}||_{Q(\bar{x})}.
\]

In other words, \( \bar{x} \) is close to the center if the Newton step from \( \bar{x} \) for the center problem \( CP \) is "small", that is, measured in an appropriate norm – the Hessian norm.

51
Chapter 4

Right-Hand-Side Parametric Center Problem

4.1 Introduction

Consider a system of $m$ linear inequalities in $\mathbb{R}^n$ of the form $Ax \leq b$ and $k$ linear equations of the form $Mx = g$. The analytic center of the system, referred to as $(A, b, M, g)$, is the optimal solution $\hat{x}$ of the program

$$
CP^= : \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i \\
\text{subject to} \quad Ax + s = b \\
\quad s > 0 \\
\quad Mx = g.
$$

(See Sonnevend [60,61] and also Freund [17].)

\footnote{This chapter has been written up as a paper [22] which has been accepted for publication in Mathematics of Operations Research.}

52
Assume that the set \( \{ x \in \mathbb{R}^n | Ax < b, \; Mx = g \} \) is nonempty and bounded, then the (analytic) center \( \hat{x} \) of \((A, b, M, g)\) is uniquely defined. Our interest lies in tracing the center as the right-hand-side (RHS) of the system \((A, b, M, g)\) varies parametrically. In particular, we are interested in generating the parametric family of optimal solutions to the programs

\[
PCP^\alpha(\alpha) : \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i \\
\text{subject to} \quad Ax + s = b + \alpha d \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad s > 0 \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad Mx = g + \alpha h.
\]

We develop an algorithm for generating a piecewise-linear path of approximate solutions to \(PCP^\alpha(\alpha)\) (\(a\)-approximate centers) as the parameter \(\alpha\) is varied strictly monotonically over a prespecified range.

We note that there is no loss of generality in assuming that the equations \(Mx = g + \alpha h\) are not present. To see this, we can assume without loss of generality that \(M = [B, N]\) is a \(k \times n\) matrix, where \(B\) is \(k \times k\) and nonsingular. By suitably partitioning \(A = [C, D]\) and \(x^T = (y^T, z^T)\), we can eliminate the \(y\) variables to obtain the equivalent programs

\[
\overline{PCP}(\alpha) : \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i \\
\text{subject to} \quad \bar{A}z + s = \bar{b} + \alpha \bar{d} \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad s > 0,
\]

where \(\bar{A} = D - CB^{-1}N, \quad \bar{b} = b - CB^{-1}g, \) and \(\bar{d} = d - CB^{-1}h\). It is straightforward to show that \(\hat{z}_\alpha\) solves \(\overline{PCP}(\alpha)\) if and only if \(\hat{x}_\alpha = (\hat{y}_\alpha^T, \hat{z}_\alpha^T)^T\)
solves $PCP^\varepsilon(\alpha)$, where $\hat{y}_\alpha = B^{-1}(g + \alpha h - N \hat{z}_\alpha)$. We can thus concentrate on the more convenient problem

\[
PCP(\alpha) : \quad \text{maximize } \sum_{i=1}^{m} \ln s_i \\
\text{subject to } \ Ax + s = b + \alpha d \\
s > 0.
\]

Therefore, in this chapter we consider the problem of tracing the centers $\hat{x}_\alpha$ (i.e. solutions to $PCP(\alpha)$) of the parametric family of linear inequality systems

\[
P(\alpha) : \quad Ax \leq b + \alpha d,
\]

where $A$ is a rational $m \times n$ matrix, $b$, $d$ are rational $m$-vectors and $\alpha$ is a scalar parameter.

Let $\mathcal{X} := \{x \in \mathbb{R}^n \mid Ax \leq b\}$, and let $\mathcal{X}_\alpha := \{x \in \mathbb{R}^n \mid Ax \leq b + \alpha d\}$ be the set of feasible solutions of system $P(\alpha)$. We assume that our initial value of $\alpha$ is $\alpha^0 = 0$, and the set $\mathcal{X}^+ := \{x \in \mathbb{R}^n \mid Ax < b\}$ in nonempty and bounded. In this case, it is straightforward to show that $\mathcal{X}_\alpha$ is bounded for all values of $\alpha$. Let $\mathcal{A} := \{\alpha \mid \mathcal{X}_\alpha^+ \neq \emptyset\}$, where $\mathcal{X}_\alpha^+ := \{x \in \mathbb{R}^n \mid Ax < b + \alpha d\}$. Then $\mathcal{A}$ is an open interval, and it is straightforward to extend the analysis in Megiddo [48] to show that the path of centers

\[
\Gamma : \alpha \mapsto \hat{x}_\alpha, \quad \alpha \in \mathcal{A},
\]

is continuous and differentiable.

Let $\alpha^{\max}$ ($\alpha^{\min}$) be the largest (smallest) value of $\alpha$ such that system $P(\alpha)$ has a feasible solution, that is, $\alpha^{\max} := \sup \{\alpha \mid \alpha \in \mathcal{A}\}$ and $\alpha^{\min} := $
\( \inf \{ \alpha \mid \alpha \in \mathcal{A} \} \). It is also straightforward to show that \( \alpha^{\text{max}} = \infty \) and \( \alpha^{\text{min}} = -\infty \) if and only if \( d = Ar \) for some \( r \in \mathbb{R}^n \), in which case, the sets \( \mathcal{X}_\alpha \) are just a family of translations of the set \( \mathcal{X} \) by the translation vector \( ar \), that is, \( \mathcal{X}_\alpha = \mathcal{X} + ar \). We therefore assume the following throughout this chapter.

**Assumption:** \( d \neq Ar \) for any \( r \in \mathbb{R}^n \).

In this case, either \( \alpha^{\text{max}} < \infty \), or \( \alpha^{\text{min}} > -\infty \), or both. (See Figure 5.)

The algorithm we propose in this chapter will trace a piecewise-linear path of \( \delta \)-approximate centers (with \( \delta = 0.59 \)) \( \tilde{\Gamma} : \alpha \mapsto \tilde{x}_\alpha \) over a prespecified range \( \alpha \in [0, \alpha^{\text{up}}] \). (See Section 3.2.1 for the definition of a \( \delta \)-approximate center.) At each iteration of the algorithm, the value of \( \alpha \) is strictly increased; thus the algorithm is strictly monotone in the parameter \( \alpha \). Furthermore, at each iteration, the magnitude of increase in \( \alpha \) is bounded (from below) in at least one of two ways, as follows. Suppose \( \alpha^j \) is the value of \( \alpha \) at the start of iteration \( j \). The algorithm will produce, at iteration \( j \), either a finite lower bound \( LB \leq \alpha^{\text{min}} \), or a finite upper bound \( UB \geq \alpha^{\text{max}} \), or both. If \( UB \) is produced, then \( \alpha^{j+1} \), the next value of \( \alpha \), will satisfy

\[
\alpha^{j+1} - \alpha^j \geq \frac{1}{128m}(UB - \alpha^j),
\]

so that \((\alpha^{\text{max}} - \alpha) \) decreases geometrically, that is,

\[
(\alpha^{\text{max}} - \alpha^{j+1}) \leq \left(1 - \frac{1}{128m}\right)(\alpha^{\text{max}} - \alpha^j). \tag{4.1}
\]
Figure 5. The four cases

Case 1: Bounded above and below.

Case 2: Bounded below.

Case 3: Bounded above.

Unbounded: $d = Ar$. 
If $LB$ is produced, then $\alpha^{j+1}$ will satisfy

$$\alpha^{j+1} - \alpha^j \geq \frac{1}{128m}(\alpha^j - LB),$$

so that $(\alpha - \alpha^{min})$ increases geometrically, that is,

$$(\alpha^{j+1} - \alpha^{min}) \geq \left(1 + \frac{1}{128m}\right)(\alpha^j - \alpha^{min}). \quad (4.2)$$

Under the Assumption of this chapter, at least one of the two bounds must be satisfied. Note that in either of the two cases, the geometric rate of change of $\alpha$ relative to a bound is at least $O(\frac{1}{m})$, where $m$ is the number of inequality constraints.

The rest of this chapter is organized as follow. In Section 2, we present some analytical properties of the path of parametric centers. In Section 3, we look at the effects of translation on the parametric centers. In Section 4, we prove the Strict Improvement Theorem, which gives the quantum of increase $\beta$ in the parametric value such that two successive centers $\hat{x}_\alpha$ and $\hat{x}_{\alpha+\beta}$ will be sufficiently close to each other (in an appropriate measure) so that Newton’s method when applied will work well. In Section 5, we derive bounds on the maximal and/or minimal values of $\alpha$ such that system $P(\alpha)$ is feasible. In Section 6, we present our algorithm (Algorithm RHSPCP) for the right-hand-side parametric center problem, which is based on the Strict Improvement Theorem of Section 4 and the bounds of Section 5. Finally, in Section 7, we describe an application of Algorithm RHSPCP of Section 6 to the linear programming problem (LP). We obtain a strictly monotone polynomial-time path-following algorithm for LP which requires $O(mL)$ iterations, where $L$ is the size of an encoding of the given problem instance.
4.2 Properties of RHS Parametric Centers

We shall first study the properties of the path of parametric (analytic) centers. For each \( \alpha \in \mathcal{A} \), let \( \hat{x}_\alpha \) be the center of system \( P(\alpha) \), and let \( \hat{s}_\alpha = b + \alpha d - A\hat{x}_\alpha \). The Karush-Kuhn-Tucker conditions for \( PCP(\alpha) \), which characterize the center \( \hat{x}_\alpha \), are as follow.

\[
\hat{s}_\alpha = b + \alpha d - A\hat{x}_\alpha > 0, \quad (4.3)
\]

\[
e^T\hat{S}_\alpha^{-1}A = 0. \quad (4.4)
\]

Also, let \( u_\alpha \) be the projection of \( \hat{S}_\alpha^{-1}d \) on to the range space of \( \hat{S}_\alpha^{-1}A \), that is,

\[
u_\alpha = P_{\hat{S}_\alpha^{-1}A}(\hat{S}_\alpha^{-1}d), \quad (4.5)
\]

where \( P_M := I - M(M^TM)^{-1}M^T \). Then we observe that

\[
d = \hat{S}_\alpha u_\alpha + Ar_\alpha, \quad (4.6)
\]

for some \( r_\alpha \in \mathbb{R}^n \). We now define the path indicator function \( \varphi(\alpha) \) as

\[
\varphi(\alpha) = e^Tu_\alpha, \quad (4.7)
\]

and note immediately that \( \varphi(\alpha) = e^Tu_\alpha = e^T\hat{S}_\alpha^{-1}(d - Ar_\alpha) = e^T\hat{S}_\alpha^{-1}d \), since \( e^T\hat{S}_\alpha^{-1}A = 0 \). Therefore, an alternative equivalent definition of \( \varphi(\alpha) \) is

\[
\varphi(\alpha) = e^T\hat{S}_\alpha^{-1}d. \quad (4.8)
\]

The motivation for considering the path indicator function \( \varphi(\alpha) \) is the following. It is clear that \( \alpha^{\max} \) is the optimal value of the following linear
program:

\[
LP^0: \quad \alpha^{\text{max}} = \max_{x,\alpha} \alpha \\
\text{s. t. } Ax - d\alpha \leq b.
\]

Suppose \(\varphi(\tilde{\alpha}) < 0\) for some given \(\tilde{\alpha}\). Then for any feasible solution \((x, \alpha)\) of \(LP^0\),

\[
\alpha = \frac{e^T \hat{S}_a^{-1} d\alpha}{\varphi(\tilde{\alpha})} \leq \frac{e^T \hat{S}_a^{-1} (Ax - b)}{\varphi(\tilde{\alpha})} = \frac{e^T \hat{S}_a^{-1} (Ax - A\hat{x}_a - \hat{s}_a + d\tilde{\alpha})}{\varphi(\tilde{\alpha})} = \frac{-m + \tilde{\alpha}\varphi(\tilde{\alpha})}{\varphi(\tilde{\alpha})},
\]

whereby \(\alpha^{\text{max}} \leq \tilde{\alpha} - \frac{m}{\varphi(\tilde{\alpha})}\). Similarly, if \(\varphi(\tilde{\alpha}) > 0\), then \(\alpha^{\text{min}} \geq \tilde{\alpha} - \frac{m}{\varphi(\tilde{\alpha})}\).

Now suppose \(\varphi(\alpha^*) = 0\) for some given value \(\alpha^*\) of \(\alpha\). Then \(e^T \hat{S}_a^{-1} A = 0\) and \(e^T \hat{S}_a^{-1} d = 0\), whereby \((\hat{x}_{\alpha^*}, \alpha^*)\) is the center of the (extended) system

\[
[A, -d] \begin{bmatrix}
x \\ \alpha
\end{bmatrix} \leq b.
\]

Thus, the set

\[
\mathcal{X}^0 = \{(x, \alpha) \in \mathbb{R}^{n+1} | Ax - d\alpha \leq b\}
\]

is bounded, and in particular, \(\alpha^{\text{max}}\) and \(\alpha^{\text{min}}\) are finite. We say that \(\mathcal{X}^0\) is bounded in \(\alpha\) if both \(\alpha^{\text{max}}\) and \(\alpha^{\text{min}}\) are finite. If \(\alpha^{\text{max}}\) (\(\alpha^{\text{min}}\)) is finite, we say that \(\mathcal{X}^0\) is bounded above (below) in \(\alpha\).

Returning to the path indicator function \(\varphi(\alpha)\) defined earlier, we say that \(\hat{x}_\alpha\) is on the upper path if \(\varphi(\alpha) \leq 0\), and is on the lower path if \(\varphi(\alpha) \geq 0\). The intuition behind this definition is provided by the next two propositions.
Proposition 4.2.1 The path indicator function \( \varphi(\alpha) \) is strictly decreasing for \( \alpha \in (\alpha_{\text{min}}, \alpha_{\text{max}}) \).

Proof: The Karush-Kuhn-Tucker conditions for \( \text{PCP}(\alpha) \) are

\[ A\dot{x}_\alpha + \dot{s}_\alpha = b + \alpha d, \quad \text{and} \quad e^T \dot{S}_\alpha^{-1} A = 0. \]

Let \( \dot{s}_\alpha \) and \( \dot{x}_\alpha \) denote the vectors of derivatives of \( s_\alpha \) and \( x_\alpha \) with respect to \( \alpha \), respectively. Then differentiating the above expressions yields

\[ A\dot{x}_\alpha + \dot{s}_\alpha = d, \quad \text{and} \quad \dot{s}_\alpha^T \dot{S}_\alpha^{-2} A = 0. \]

Furthermore, since \( \varphi(\alpha) = e^T \dot{S}_\alpha^{-1} d \),

\[ \varphi'(\alpha) = -\dot{s}_\alpha^T \dot{S}_\alpha^{-2} d = -\dot{s}_\alpha^T \dot{S}_\alpha^{-2} (A\dot{x}_\alpha + \dot{s}_\alpha) = -\dot{s}_\alpha^T \dot{S}_\alpha^{-2} \dot{s}_\alpha < 0, \]

unless \( \dot{s}_\alpha = 0 \), in which case \( d = A\dot{x}_\alpha \). But this last possibility is ruled out by the Assumption. \( \square \)

Proposition 4.2.2 (Upper and Lower Paths).
(i) Both \( \alpha_{\text{max}} \) and \( \alpha_{\text{min}} \) are finite if and only if there exists \( \alpha^* \in \mathbb{R} \) such that \( \varphi(\alpha^*) = 0 \), and

\[
\varphi(\alpha) > 0 \text{ for all } \alpha \in (\alpha_{\text{min}}, \alpha^*), \\
\varphi(\alpha) < 0 \text{ for all } \alpha \in (\alpha^*, \alpha_{\text{max}}); 
\]

(ii) \( \alpha_{\text{max}} \) is finite and \( \alpha_{\text{min}} = -\infty \) if and only if \( \varphi(\alpha) < 0 \) for all \( \alpha < \alpha_{\text{max}} \);

(iii) \( \alpha_{\text{min}} \) is finite and \( \alpha_{\text{max}} = \infty \) if and only if \( \varphi(\alpha) > 0 \) for all \( \alpha > \alpha_{\text{min}} \).

Proof: (i) We have seen earlier that if \( \varphi(\alpha^*) = 0 \), then \((\dot{x}_{\alpha^*}, \alpha^*)\) is the center of the system (in \( \mathbb{R}^{n+1} \)) \( Ax - d\alpha \leq b \). That being the case, \( \mathcal{X}^0 \) is bounded
and so $\alpha^{max}$ and $\alpha^{min}$ are both finite. Conversely, if $\alpha^{max}$ and $\alpha^{min}$ are both finite, then the center $(x^*, \alpha^*)$ of the system $Ax - d\alpha \leq b$ existed uniquely. The Karush-Kuhn-Tucker conditions that characterize the center $(x^*, \alpha^*)$ require that $e^T \hat{S}_\alpha^{-1} A = 0$ and $e^T \hat{S}_\alpha^{-1} d = 0$, where $\hat{s}_{\alpha^*} = b + d\alpha^* - Ax^*$, i.e., $\varphi(\alpha^*) = 0$. The rest follows from Proposition 4.2.1.

(ii) Suppose $\alpha^{max} < \infty$ and $\alpha^{min} = -\infty$. Let $\bar{\alpha} < \alpha^{max}$. We have seen earlier that $\varphi(\bar{\alpha}) > 0$ implies that $\alpha^{min}$ is finite. Thus, $\varphi(\bar{\alpha}) \leq 0$. But from part (i), if $\varphi(\bar{\alpha}) = 0$ then $\alpha^{min}$ is finite, which is a contradiction. Thus, $\varphi(\bar{\alpha}) < 0$. Conversely, if $\varphi(\bar{\alpha}) < 0$ for some $\bar{\alpha}$, then $\alpha^{max} < \infty$. Consequently, if $\varphi(\alpha) < 0$ for all $\alpha < \alpha^{max}$, then $\alpha^{min} = -\infty$ follows from part (i).

The proof of part (iii) is similar to that of part (ii). □

4.3 Translations

Now suppose $d = u + Ar$ for some $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$. Then we observe that for each $\alpha \in \mathbb{R}$,

$$\{x \in \mathbb{R}^n \mid Ax \leq b + \alpha d\} = \{x \in \mathbb{R}^n \mid Ax \leq b + \alpha u\} + \alpha r.$$  

That is, the feasible sets of the two systems,

$$Ax \leq b + \alpha d \quad \text{and} \quad Ax \leq b + \alpha u,$$

differ only by the translation vector $\alpha r$.

Let $s_\alpha(x; d) = b + \alpha d - Ax$ and $Q_\alpha(x; d)$ denote respectively the slacks and the negative of the Hessian of the logarithmic barrier function

$$\sum_{i=1}^m \ln[b + \alpha d - Ax]_i.$$  

61
at $x$ for the system $Ax \leq b + \alpha d$. Similarly, let $s_\alpha(x; u)$ and $Q_\alpha(x; u)$ denote
the corresponding entities for the system $Ax \leq b + \alpha u$. We immediately
observe that $s_\alpha(x; d) = s_\alpha(x - \alpha r; u)$ and hence $Q_\alpha(x; d) = Q_\alpha(x - \alpha r; u),$
since $Q_\alpha(x; d) = A^T S_\alpha^{-2} (x; d) A$. Also, let $\hat{x}_\alpha(d)$ ($\hat{x}_\alpha(u)$) denote the center of
the system $Ax \leq b + \alpha d$ ($Ax \leq b + \alpha u$). The next lemma therefore follows
easily from the above observations.

**Lemma 4.3.1** Suppose $d = u + Ar$ for some $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$. Then for
each $\alpha \in \mathcal{A},$

(i) $\hat{x}_\alpha(d) = \hat{x}_\alpha(u) + \alpha r,$

(ii) $||x - \hat{x}_\alpha(d)||_{Q_\alpha(\hat{x}_\alpha(d); d)} = ||(x - \alpha r) - \hat{x}_\alpha(u)||_{Q_\alpha(\hat{x}_\alpha(u); u)}$ for any $x \in \mathbb{R}^n,$

and

(iii) $\max \{\alpha | Ax \leq b + \alpha d \text{ for some } x\} = \max \{\alpha | Ax \leq b + \alpha u \text{ for some } x\}.$

Therefore, at each $\alpha$, the values of all the slacks, at corresponding points,
for the two systems are the same and thus the (Hessian) $Q$-norms, at cor-
responding points, are exactly the same, and the centers differ by only a
translation vector $\alpha r$. So, changing the right-hand-side by $\alpha d$ is the same as
translating the system by $\alpha r$ and then changing the right-hand-side by $\alpha u.$

In the sequel, we shall be working with appropriate choices of $u \in \mathbb{R}^m$
and $r \in \mathbb{R}^n$ instead of $d$. As we shall see in Section 5, this in fact is cen-
tral in our derivation of bounds on $\alpha_{\text{max}}$ and $\alpha_{\text{min}}.$
4.4 Strict Improvement Theorem

In this section, we shall first consider increasing the parameter from $\alpha = 0$ to $\alpha = \beta$ for some $\beta > 0$ for the parametric family of systems $Ax \leq b + \alpha u$, where $d = u + Ar$ for some appropriate choice of $u \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$. For $x \in \mathbb{R}^n$ satisfying $s_\alpha = b + \alpha u - Ax > 0$, let $Q_\alpha(x)$ denote the negative of the Hessian of the logarithmic barrier function $\sum_{i=1}^m \ln[b + \alpha u - Ax]$; at $x$ for the system $Ax \leq b + \alpha u$, that is,

$$Q_\alpha(x) = A^T S_\alpha^{-2} A. \tag{4.9}$$

First, we have the following

**Theorem 4.4.1** Suppose $\hat{x}$ is the center of the inequality system $Ax \leq b$.
Let $\hat{s} = b - A\hat{x}$ and let $\hat{x}_\alpha$ denote the center of the system $Ax \leq b + \alpha u$. Let $\hat{u} = \hat{S}^{-1}u$. Then for any $\alpha$ satisfying $|\alpha| \leq \frac{1}{76||\hat{u}||}$, we have

$$||\hat{x} - \hat{x}_\alpha||_{Q_\alpha(x)} < \frac{1}{12},$$

where $Q_\alpha(\hat{x})$ is as defined by (4.9).

**Proof:** The proof makes use of the third measure of closeness $\tau$ defined in Section 3.2.3. With respect to the system $Ax \leq b + \alpha u$, let the slacks be $\hat{s}_\alpha = b + \alpha u - A\hat{x}$. Note that $A^T \hat{S}^{-1} e = 0$ (from (3.2)) and $\hat{s}_\alpha = \hat{s} + \alpha u$, and so by definition of $\hat{u}$,

$$A^T \hat{S}_\alpha^{-1} e = A^T \hat{S}_\alpha^{-1} \hat{S}^{-1} [\hat{S} - \hat{S}_\alpha] e$$

$$= A^T \hat{S}_\alpha^{-1} \hat{S}^{-1} (\hat{s} - \hat{s}_\alpha)$$

$$= -\alpha A^T \hat{S}_\alpha^{-1} \hat{S}^{-1} u$$

$$= -\alpha A^T \hat{S}_\alpha^{-1} \hat{u}.$$ 

63
Note that $Q_\alpha(\hat{x}) = A^T\hat{S}_\alpha^{-2}A$. Hence, with respect to the inequality system $Ax \leq b + \alpha u$, for $\alpha$ satisfying $|\alpha| \leq \frac{1}{76}\|u\|$, (see (3.10))

$$
\tau_\alpha = \tau_\alpha(\hat{x}) := \|A^T\hat{S}_\alpha^{-1}e\|_{Q_\alpha^{-1}(\hat{x})} \leq |\alpha| \sqrt{\hat{u}^T\hat{S}_\alpha^{-1}AQ_\alpha^{-1}(\hat{x})A^T\hat{S}_\alpha^{-1}\hat{u}}.
$$

where the last inequality follows because the eliminated matrix is a projection matrix. Hence, from Lemma 3.2.5, we have

$$
\|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x})} < \frac{1}{12}.
$$

From Lemma 3.2.1, the two following corollaries are immediate.

**Corollary 4.4.1** Under the same conditions as Theorem 4.4.1,

$$
\|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x})} < \frac{1}{11}.
$$

**Corollary 4.4.2** Suppose $\hat{x}$ is the center of the inequality system $Ax \leq b$. Suppose $\bar{x}$ is given such that $\bar{s} = b - A\bar{x} > 0$ and $\|\bar{x} - \hat{x}\|_{Q_\alpha(\bar{x})} < \frac{1}{21}$. For any given $\bar{u} \in \mathbb{R}^m$, let $u = \bar{S}\bar{u}$. Then for any $\alpha$ satisfying $\|\alpha\bar{u}\| \leq \frac{1}{80}$, we have

(i) $\|\alpha\hat{S}^{-1}u\| \leq 1/76,$

(ii) $\|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x})} < 1/12,$

(iii) $\|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x})} < 1/11,$
where \( \hat{x}_\alpha \) is the center of the system \( Ax \leq b + \alpha u \) and \( Q_\alpha(x) \) is defined by (4.9).

**Proof:** Let \( \hat{u} = \hat{S}^{-1}u \). Then \( \hat{u} = \hat{S}^{-1}\bar{S}u \), so from Lemma 3.2.1(ii),

\[
||\hat{u}|| \leq ||\bar{S}^{-1}\bar{S}|| ||u|| \leq \left( \frac{21}{20} \right) ||u||.
\]

Thus, \( ||\alpha \bar{u}|| \leq \frac{1}{80} \) implies

\[
||\alpha \hat{u}|| \leq \left( \frac{21}{20} \right) ||\alpha \bar{u}|| \leq \frac{1}{76}.
\]

This proves part (i). Parts (ii) and (iii) follow immediately from Theorem 4.4.1 and Corollary 4.4.1. ☐

**Remark:** Therefore, if the increase in the parameter value is \( O(\frac{1}{||u||}) \), the two centers \( \hat{x} \) and \( \hat{x}_\alpha \), of the respective systems \( Ax \leq b \) and \( Ax \leq b + \alpha u \) (\( u = \bar{S}u \)), will be close to each other in a precise mathematical sense. Note that the theorem and corollaries proven above are valid for arbitrary \( u \) and \( \bar{u} \). As we shall see in the next section, there is an appropriate \( \bar{u} \), specifically, \( \bar{u} = \bar{P}_{\bar{S}^{-1}A}\bar{S}^{-1}d \), which we can use to derive bounds on \( \alpha^{\text{max}} \) and \( \alpha^{\text{min}} \) in terms of \( \frac{1}{||u||} \). In the following theorem, we therefore take \( \bar{u} \) to be \( \bar{u} = \bar{P}_{\bar{S}^{-1}A}\bar{S}^{-1}d \).

**Theorem 4.4.2 (Strict Improvement Theorem).**

Suppose \( \hat{x} \) is the center of the inequality system \( Ax \leq b \). Suppose \( \bar{x} \) is given such that \( \bar{s} = b - A\bar{x} > 0 \) and \( ||\bar{x} - \hat{x}||_{Q_0(\bar{s})} \leq \frac{1}{21} \) (i.e., \( \bar{x} \) is a \( \delta \)-approximate center of the system \( Ax \leq b \) with \( \delta = \frac{1}{21} \)). Define

\[
\bar{u} = \bar{P}_{\bar{S}^{-1}A}\bar{S}^{-1}d, \quad \text{(projection of } \bar{S}^{-1}d)\]

65
\begin{align*}
\bar{r} &= (A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-2} d, \quad \text{(translation vector)} \\
\beta &= \frac{1}{80\|\bar{u}\|}, \quad \text{(step-length)} \\
\bar{s}_\beta &= b + \beta \bar{S} \bar{u} - A \bar{x}, \quad \bar{S}_\beta = \text{diag}(\bar{s}_\beta) \\
\bar{\eta} &= -(A^T \bar{S}^{-2} A)^{-1} A^T \bar{S}^{-1} e, \quad \text{(Newton step)} \\
\bar{x}_{\text{new}} &= \bar{x} + \bar{\eta} + \beta \bar{r}.
\end{align*}

Then \( \|\bar{x}_{\text{new}} - \hat{x}_\beta\|_{Q_\beta(x_{\text{new}})} \leq \frac{1}{21} \) (i.e., \( \bar{x}_{\text{new}} \) is again a \( \delta \)-approximate center of the (new) system \( Ax \leq b + \beta d \) with \( \delta = \frac{1}{21} \)).

**Proof:** First note that \( d = u + A \bar{r} \), where \( u = \bar{S} \bar{u} \). Thus by Lemma 4.3.1 we may consider the system \( Ax \leq b + \beta u \) instead of the system \( Ax \leq b + \beta d \). Then by Lemma 3.2.1(vi) it suffices to show that

\[ ||\bar{x}_\beta - \hat{x}_\beta||_{Q_\beta(x_{\text{new}})} \leq 1/22, \]

where \( \bar{x}_\beta = \bar{x} + \bar{\eta} \), \( \hat{x}_\beta \) is the center of the system \( Ax \leq b + \beta u \) and \( Q_\beta(\hat{x}_\beta) \) is defined by (4.9). In Lemma 4.4.1(iii) in the following, we show that

\[ \varepsilon := ||\bar{x} - \hat{x}_\beta||_{Q_\beta(x_{\text{new}})} \leq 0.147, \]

which, by Lemma 3.3.1, implies that \( ||\bar{x}_\beta - \hat{x}_\beta||_{Q_\beta(x_{\text{new}})} < 0.034 \leq 1/22. \)

**Remarks:** In this theorem, the current iterate \( \bar{x} \) is a \( \delta \)-approximate center of the system \( Ax \leq b \), and after increasing the value of \( \alpha \) by the amount \( \beta = \frac{1}{80\|\bar{u}\|} \), the new point \( \bar{x}_{\text{new}} \), obtained by taking a Newton step from \( \bar{x} \) in the problem of finding the new center \( \hat{x}_\beta \), is again a \( \delta \)-approximate center of the new system \( Ax \leq b + \beta d \). Therefore, the process can be repeated. This therefore suggests an iterative algorithm, which we will present in Section 6.
The increase in $\alpha$ is $\beta = \frac{1}{80\|\tilde{u}\|}$, a function of $\|\tilde{u}\|$. Ideally, we would like this increase to be as large as possible to speed algorithmic convergence. However, the fact that it is proportional to $1/\|\tilde{u}\|$ makes good intuitive sense. The quantity $\|\tilde{u}\|$ is the (weighted) distance of $d$ from the column space of $A$ in the weighted norm. Suppose $d$ is very close to the range space of $A$, then $\|\tilde{u}\|$ is very small. So, we can take a large step. If, however, $\|\tilde{u}\|$ is large, then "most of $d$" lies outside the range space of $A$, so that a change in the right-hand-side of $\alpha d$ changes substantially the "shape" of the polyhedral set $X_\alpha$ and hence the system $P(\alpha)$, not just translating the system. Thus, the step will be small.

**Lemma 4.4.1** Under the conditions of Theorem 4.4.2, define $\tilde{u}$, $\bar{r}$, and $\beta$ as in Theorem 4.4.2. Let $u = \tilde{S} \tilde{u}$, let $\hat{x}_\alpha$ be the center of the system $Ax \leq b + \alpha u$, and let $Q_\alpha(\hat{x}_\alpha)$ be defined by (4.9). Then for all $\alpha \in [0, \beta]$ and $v \in \mathbb{R}^n$,

\begin{align*}
(i) \quad & \|v\|_{Q_\alpha(\hat{x}_\alpha)} \leq \left(\frac{11}{10}\right)\left(\frac{77}{76}\right)\|v\|_{Q_\alpha(\hat{x}_\alpha)}; \\
(ii) \quad & \|v\|_{Q_\alpha(\hat{x}_\alpha)} \leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)\|v\|_{Q_\alpha(\hat{x}_\alpha)}; \\
(iii) \quad & \|\tilde{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} \leq 0.147.
\end{align*}

**Proof:** From Corollary 4.4.2,

\[ \|\alpha \tilde{S}^{-1}u\| \leq 1/76 \quad \text{and} \quad \|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x}_\alpha)} < 1/11. \quad (4.10) \]

Thus, by Lemma 3.2.1(ii) and (iii), with respect to system $Ax \leq b + \alpha u$,

\[ \|\tilde{S}_\alpha^{-1}\hat{S}_\alpha\| \leq \frac{12}{11} \quad \text{and} \quad \|\tilde{S}_\alpha^{-1}\hat{S}_\alpha\| \leq \frac{11}{10}, \quad (4.11) \]
where $\hat{s}_\alpha = b + \alpha u - A\hat{x}$ and $\hat{s}_\alpha = b + \alpha u - A\hat{x}$. Also,

$$
||\hat{S}^{-1}\hat{s}_\alpha|| = \max_i \left| \frac{b_i + \alpha u_i - A_i\hat{x}}{b_i - A_i\hat{x}} \right|
= \max_i \left| 1 + \alpha \left( \frac{u_i}{b_i - A_i\hat{x}} \right) \right|
\leq 1 + ||\alpha\hat{S}^{-1}u|| \leq \frac{77}{76},
$$

because $||\alpha\hat{S}^{-1}u|| \leq 1/76$ from (4.10). Hence,

$$
||v||_{\mathcal{Q}_0(\tilde{x})} = ||\hat{S}^{-1}Av||
\leq ||\hat{S}^{-1}\hat{s}_\alpha|| \cdot ||\hat{S}^{-1}\hat{s}_\alpha|| \cdot ||\hat{S}^{-1}Av||
\leq \left(\frac{77}{76}\right)\left(\frac{11}{10}\right)||v||_{\mathcal{Q}_0(\tilde{x}_\alpha)}.
$$

This proves part (i). Similarly,

$$
||\hat{S}^{-1}\hat{s}_\alpha|| = \max_i \left| \frac{b_i - A_i\hat{x}}{b_i + \alpha u_i - A_i\hat{x}} \right|
= \max_i \left| \frac{1}{1 + \alpha \left( \frac{u_i}{b_i - A_i\hat{x}} \right)} \right|
\leq \frac{1}{1 - ||\alpha\hat{S}^{-1}u||} \leq \frac{76}{75}.
$$

Hence the proof of part (ii) follows similarly as that of part (i). Next, to show part (iii), we have by the triangle inequality,

$$
||\bar{x} - \hat{x}_\alpha||_{\mathcal{Q}_0(\tilde{x}_\alpha)} \leq ||\bar{x} - \hat{x}||_{\mathcal{Q}_0(\tilde{x}_\alpha)} + ||\hat{x} - \hat{x}_\alpha||_{\mathcal{Q}_0(\tilde{x}_\alpha)}
\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)||\bar{x} - \hat{x}||_{\mathcal{Q}_0(\tilde{x})} + \frac{1}{11}
\leq \left(\frac{12}{11}\right)\left(\frac{76}{75}\right)(\frac{1}{20}) + \frac{1}{11} < 0.147,
$$

where the second inequality follows from part (ii) and (4.10) and the third inequality follows because $||\bar{x} - \hat{x}||_{\mathcal{Q}_0(\tilde{x})} \leq 1/20$ by Lemma 3.2.1(vi). □
Next, we also have the following theorem which allows us to generate a linear segment of \( \delta \)-approximate centers, with \( \delta = 0.59 \).

**Theorem 4.4.3** Under the conditions and definitions of Theorem 4.4.2, for all \( \alpha \in [0, \beta] \), define \( \tilde{x}_\alpha := \bar{x} + \frac{(\alpha)}{(\beta)}(\tilde{x}_{new} - \bar{x}) \). Then, for all \( \alpha \in [0, \beta] \),

\[
||\tilde{x}_\alpha - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq 0.59.
\]

**Proof:** Let \( \alpha \in [0, \beta] \). By Lemma 3.2.1(vi), it suffices to show that

\[
||\tilde{x}_\alpha - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq 0.37.
\]

We have, from the proof of Theorem 4.4.2 above and Lemma 4.4.1(iii),

\[
||\tilde{x}_\beta - \bar{x}||_{Q_\beta(\tilde{x}_\beta)} \leq ||\tilde{x}_\beta - \hat{x}_\beta||_{Q_\beta(\tilde{x}_\beta)} + ||\bar{x} - \hat{x}_\beta||_{Q_\beta(\tilde{x}_\beta)} < 0.034 + 0.147 = 0.181.
\]

Hence, by Lemma 4.4.1(i)–(ii),

\[
||\tilde{x}_\beta - \bar{x}||_{Q_\alpha(\tilde{x}_\alpha)} \leq \left( \frac{12}{11} \right) \left( \frac{76}{73} \right) ||\tilde{x}_\beta - \bar{x}||_{Q_\alpha(\tilde{x}_\alpha)}
\]

\[
\leq \left( \frac{12}{11} \right) \left( \frac{76}{75} \right) \left( \frac{11}{10} \right) \left( \frac{77}{76} \right) ||\tilde{x}_\beta - \bar{x}||_{Q_\beta(\tilde{x}_\beta)} < 0.223.
\]

Next, we observe that \( \tilde{x}_\alpha - \bar{x} = \left( \frac{\alpha}{\beta} \right)(\tilde{x}_\beta - \bar{x}) \). Therefore,

\[
||\tilde{x}_\alpha - \bar{x}||_{Q_\alpha(\tilde{x}_\alpha)} = \left( \frac{\alpha}{\beta} \right) ||\tilde{x}_\beta - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} < \left( \frac{\alpha}{\beta} \right)(0.223) \leq 0.223.
\]

Thus, by the triangle inequality and Lemma 4.4.1(iii),

\[
||\tilde{x}_\alpha - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq ||\tilde{x}_\alpha - \bar{x}||_{Q_\alpha(\tilde{x}_\alpha)} + ||\bar{x} - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} < 0.223 + 0.147 = 0.37. \quad \square
\]
4.5 Lower and Upper Bounds

In this section, we shall derive upper and/or lower bounds for the values of
\( \alpha \) such that the inequality system \( P(\alpha) \) has a feasible solution. As in the
previous section, we assume the current value of \( \alpha \) is \( \alpha = 0 \). Suppose we
have a \( \delta \)-approximate center \( \ddot{x} \) of system \( P(0) \) (i.e., \( \|\ddot{x} - \dddot{x}\|_{Q_0(\dddot{x})} \leq 1/21 \)).
We decompose \( d \) (as in Theorem 4.4.2) into \( d = u + A\ddot{r} \), where

\[
  u = \dddot{S}\dddot{u} = \dddot{S}P_{\dddot{S}^{-1}A}\dddot{S}^{-1}d \quad \text{and} \quad \dddot{r} = (A^T\dddot{S}^{-2}A)^{-1}A^T\dddot{S}^{-2}d.
\]

Note that by the Assumption, \( d \) does not lie in the range space of \( A \), and so
\( \dddot{u} \neq 0 \). We shall prove the following in this section.

**Theorem 4.5.1 (Upper and Lower Bounds)**.

*Under the conditions and definitions of Theorem 4.4.2,*

1. if \( \frac{e^T \dddot{u}}{||\dddot{u}||} > 1/20 \), then

   \[
   \alpha^{\min} \geq LB = -\frac{22\sqrt{m(m + 1)} + 1}{21||\dddot{u}||};
   \]

2. if \( \frac{e^T \dddot{u}}{||\dddot{u}||} < -1/20 \), then

   \[
   \alpha^{\max} \leq UB = \frac{22\sqrt{m(m + 1)} + 1}{21||\dddot{u}||};
   \]

3. if \( |\frac{e^T \dddot{u}}{||\dddot{u}||}| \leq 1/20 \), then

   \[
   \alpha^{\min} \geq \bar{LB} = \frac{1.6\sqrt{m(m - 1)} + 0.6}{||\dddot{u}||}.
   \]
and
\[
\alpha^{\text{max}} \leq UB = \frac{1.6\sqrt{m(m-1)} + 0.6}{||\bar{u}||}.
\]

**Remarks:** Note that one of the three cases must be satisfied, so that either a finite lower bound on \(\alpha^{\text{min}}\) or a finite upper bound on \(\alpha^{\text{max}}\) is produced, or both are produced. Case 1 corresponds to being (approximately) on the lower path at \(\alpha = 0\), i.e., \(\varphi(0) \approx 0\). Case 2 corresponds to being (approximately) on the upper path at \(\alpha = 0\), i.e., \(\varphi(0) \approx 0\). Case 3 corresponds to \(\varphi(0) \approx 0\), so that \((\bar{x}, 0)\) is close to the center of the (extended) system

\[
[A, -d]
\begin{bmatrix}
\bar{x} \\
\alpha
\end{bmatrix}
\leq b.
\]

Note also that these bounds are \(O\left(\frac{m}{||\bar{u}||}\right)\), and thus even though the increase in \(\alpha\) (from Theorem 4.4.2) is \(\frac{1}{\beta_0||\bar{u}||}\), the ratio of the increase to one or both bounds is at least \(O\left(\frac{1}{m}\right)\). Therefore, repeated increases in \(\alpha\) using the methodology of Theorem 4.4.2 will result in geometric growth either in the quantity \((\alpha - \alpha^{\text{min}})\), with a rate of at least \((1 + O(\frac{1}{m}))\), or geometric decrease in the quantity \((\alpha^{\text{max}} - \alpha)\), with a rate of \((1 - O(\frac{1}{m}))\). (See the next section for the detailed complexity analysis.)

We begin with two fundamental lemmas. We observe that
\[
||\hat{x} - \bar{x}||_{Q_0(\bar{x})} = ||\hat{S}^{-1}(\bar{s} - \hat{s})|| \quad \text{and} \quad ||\hat{x} - \bar{x}||_{Q_0(\hat{x})} = ||\hat{S}^{-1}(\bar{s} - \hat{s})||.
\]

**Lemma 4.5.1** \(|e^T\hat{S}^{-1}u - e^T\bar{u}| \leq ||\bar{u}||/20\).
Proof: Using a Cauchy-Schwartz inequality and Lemma 3.2.1(vi),

\[ |e^T \tilde{S}^{-1} u - e^T \tilde{u}| = |e^T (\tilde{S}^{-1} - \tilde{S}^{-1}) u| \]
\[ = |(\tilde{s} - \hat{s}) \tilde{S}^{-1} \tilde{S}^{-1} u| \]
\[ \leq ||\tilde{S}^{-1} (\tilde{s} - \hat{s})|| ||\tilde{S}^{-1} u|| \]
\[ \leq (1/20)||\tilde{u}||, \]

where the first inequality is a Cauchy-Schwartz inequality and the last inequality follows from Lemma 3.2.1(vi). \( \square \)

Remark: Recall that \( \varphi(0) = e^T \tilde{S}^{-1} u \) is the path indicator function at \( \alpha = 0 \). If \( e^T \tilde{u} > \frac{1}{20}||\tilde{u}|| \), then from Lemma 4.5.1, \( e^T \tilde{S}^{-1} u \geq e^T \tilde{u} - \frac{1}{20}||\tilde{u}|| > 0 \). Therefore, the current center \( \hat{x} \) is on the lower path and \( \alpha^{\text{min}} \) is bounded from below by \( \frac{-m}{e^T \tilde{S}^{-1} u} \) (as we have seen in Section 4.2). Similarly, if \( e^T \tilde{u} < -\frac{1}{20}||\tilde{u}|| \), then \( \alpha^{\text{max}} \) is bounded from above by \( \frac{-m}{e^T \tilde{S}^{-1} u} \). However, we typically cannot deduce the exact value of \( e^T \tilde{S}^{-1} u \) from an approximate center. Therefore, we derive in this section an alternate bound using the containing ellipsoid given by a \( \delta \)-approximate center (see Lemma 3.2.7).

The next lemma concerns the classical least-square or minimum-norm problem.

Lemma 4.5.2 Given an \( m \times n \) matrix \( M \) and an \( m \)-vector \( d \), for all \( t \in \mathbb{R} \),

\[ ||Mx - td|| \geq |t| ||P_M d|| \] for all \( x \in \mathbb{R}^n \),

where \( P_M = I - M(M^T M)^{-1} M^T \). \( \square \)

72
Now we shall prove Theorem 4.5.1 in two parts as the followings.

**Proposition 4.5.1** Under the conditions and definitions of Theorem 4.5.1,

(i) if \( \frac{e^T \bar{u}}{||\bar{u}||} < -\frac{1}{20} \), then \( \alpha_{\text{max}} \leq UB = \frac{22\sqrt{m(m+1)}}{21||\bar{u}||} \);

(ii) if \( \frac{e^T \bar{u}}{||\bar{u}||} > \frac{1}{20} \), then \( \alpha_{\text{min}} \geq LB = -\frac{22\sqrt{m(m+1)}}{21||\bar{u}||} \).

**Proof:** Suppose \( \frac{e^T \bar{u}}{||\bar{u}||} < -\frac{1}{20} \). We first show that \((\bar{x}, 0)\) is close to the center of the following extended system in \( \mathbb{R}^{n+1} \) with one additional variable and one additional constraint:

\[
\begin{bmatrix}
A & -u \\
0 & \theta
\end{bmatrix}
\begin{bmatrix}
x \\
\alpha
\end{bmatrix}
\leq
\begin{bmatrix}
b \\
1
\end{bmatrix},
\]  

(4.12)

where \( \theta = e^T \hat{S}^{-1}u \). Note that from the remark following Lemma 4.5.1, \( \theta < 0 \).

Therefore, by Lemma 4.3.1(iii),

\[
\alpha_{\text{max}} = \sup_{\alpha} \{ \alpha | \ Ax \leq b + \alpha d \text{ for some } x \}
\]

\[
= \sup_{\alpha} \{ \alpha | \ Ax \leq b + \alpha u \text{ for some } x \}
\]

\[
= \sup_{\alpha} \left\{ \alpha | \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix} \begin{bmatrix} x \\ \alpha \end{bmatrix} \leq \begin{bmatrix} b \\ 1 \end{bmatrix} \text{ for some } x \right\}.
\]

Now, for system (4.12), define

\[
\Delta(x, \alpha) := \text{diag} \left( \begin{bmatrix} b \\ 1 \end{bmatrix} - \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix} \begin{bmatrix} x \\ \alpha \end{bmatrix} \right)
\]

and

\[
Q(x, \alpha) := \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix}^T \left[ \Delta(x, \alpha) \right]^{-2} \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix}.
\]

73
Then, \( \Delta(\hat{x}, 0) = \begin{bmatrix} \hat{S} & 0 \\ 0 & 1 \end{bmatrix} \), \( \Delta(\bar{x}, 0) = \begin{bmatrix} \bar{S} & 0 \\ 0 & 1 \end{bmatrix} \), and
\[
(e^T, 1)[\Delta(\hat{x}, 0)]^{-1} \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix} = [e^T \hat{S}^{-1}A, -e^T \hat{S}^{-1}u + \theta] = 0.
\]

Therefore, \((\hat{x}, 0)\) is the center of system (4.12), and
\[
\left\| \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{Q(\bar{x}, 0)} = \|\hat{x} - \bar{x}\|_{Q(\bar{x}, 0)} \leq \frac{1}{21}.
\]

Let \( X^0 := \left\{ \begin{bmatrix} x \\ \alpha \end{bmatrix} \mid \begin{bmatrix} A & -u \\ 0 & \theta \end{bmatrix} \begin{bmatrix} x \\ \alpha \end{bmatrix} \leq \begin{bmatrix} b \\ 1 \end{bmatrix} \right\} \), and
\[
F_{out} := \left\{ \begin{bmatrix} x \\ \alpha \end{bmatrix} \left| \left| \begin{bmatrix} x \\ \alpha \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right|_{Q(\bar{x}, 0)} \leq \frac{22}{21} \sqrt{m(m + 1)} + \frac{1}{21} \right\}.
\]

Then, from Lemma 3.2.7, \( X^0 \subset F_{out} \). Furthermore, because \( A^T \bar{S}^{-2}u = A^T \bar{S}^{-1}\bar{u} = 0 \), we have
\[
Q(\bar{x}, 0) = \begin{bmatrix} A^T \bar{S}^{-2}A & -A^T \bar{S}^{-2}u \\ -u^T \bar{S}^{-2}A & \theta^2 + u^T \bar{S}^{-2}u \end{bmatrix} = \begin{bmatrix} A^T \bar{S}^{-2}A & 0 \\ 0 & \theta^2 + ||\bar{u}||^2 \end{bmatrix}.
\]

Thus
\[
\left\| \begin{bmatrix} x \\ \alpha \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{Q(\bar{x}, 0)}^2 = ||\bar{S}^{-1}A(x - \bar{x})||^2 + \alpha^2(\theta^2 + ||\bar{u}||^2) \geq \alpha^2 ||\bar{u}||^2.
\]

Also, for any \( \begin{bmatrix} x \\ \alpha \end{bmatrix} \in F_{out} \),
\[
\left\| \begin{bmatrix} x \\ \alpha \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{Q(\bar{x}, 0)} \leq \frac{22}{21} \sqrt{m(m + 1)} + \frac{1}{21},
\]

whereby \( \alpha \leq \frac{22\sqrt{m(m + 1)} + 1}{21||\bar{u}||} \). It then follows from Lemma 4.3.1 and the
fact that $\theta < 0$ that

$$x_{\max} = \sup_{\alpha} \{ \alpha | Ax - u\alpha \leq b \text{ for some } x \}$$

$$= \sup_{\alpha} \{ \alpha | Ax - u\alpha \leq b \text{ for some } x, \theta \alpha < 1 \}$$

$$= \sup_{\alpha} \{ \alpha | (x^T, \alpha)^T \in X^0 \}$$

$$\leq \max_{\alpha} \{ \alpha | (x^T, \alpha)^T \in F_{out} \}$$

$$\leq \frac{22\sqrt{m(m+1)} + 1}{21||\bar{u}||}.$$ 

This completes the proof of part (i). Proof of part (ii) exactly parallels that of part (i). \( \Box \)

**Proposition 4.5.2** Suppose $|e^T\bar{u}| \leq \frac{1}{20}||\bar{u}||$, then

(i) $x_{\max} \leq \overline{UB} = \frac{1.6\sqrt{m(m-1)} + 0.6}{||\bar{u}||}$; 

(ii) $x_{\min} \geq \underline{LB} = -\frac{1.6\sqrt{m(m-1)} + 0.6}{||\bar{u}||}$.

**Proof:** We shall show that $(\bar{x}, 0)$ is a $\delta$-approximate center for the system in $\mathbb{R}^{n+1}$ (see Section 3.2.3)

$$[A, -u] \begin{bmatrix} x \\ \alpha \end{bmatrix} \leq b. \quad (4.13)$$

Let $Q(x, \alpha)$ be the negative of the Hessian of the logarithmic barrier function $f(x, \alpha) = \sum_{i=1}^{m} \ln [b - Ax + u\alpha]$. Observe that

$$b - [A, -u] \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} = b - A\bar{x} = \bar{s} > 0.$$
Therefore, since $\bar{u} = \bar{S}^{-1}u$ and $A^T\bar{S}^{-1}\bar{u} = 0$,

$$Q(\bar{x}, 0) = [A, -u]^T\bar{S}^{-2}[A, -u] = \begin{bmatrix} Q_0(\bar{x}) & 0 \\ 0 & ||\bar{u}||^2 \end{bmatrix}$$

and

$$e^T\bar{S}^{-1}[A, -u] = [e^T\bar{S}^{-1}A, -e^T\bar{u}].$$

Therefore,

$$\tau^2(\bar{x}, 0) := \|A, -u]^T\bar{S}^{-1}e\|_{Q^{-1}(x, \alpha)}^2 = \|A^T\bar{S}^{-1}e\|_{Q^{-1}(x, \alpha)}^2 + \frac{(e^T\bar{u})^2}{||\bar{u}||^2} = \tau^2(\bar{x}) + \frac{(e^T\bar{u})^2}{||\bar{u}||^2}.$$

Recalling that $||\bar{x} - \hat{x}||_{Q_0(\bar{x})} \leq 1/21$, we conclude from Lemma 3.2.6 that $\tau(\bar{x}) \leq 0.056$. Therefore, $\tau^2(\bar{x}, 0) \leq (0.056)^2 + (0.05)^2 \leq 0.0057$, and so $\gamma = \gamma(\bar{x}, 0) \leq 0.0755$. Taking $h = 0.43$ in Lemma 3.2.2, we have

$$\left\| \begin{bmatrix} x^* \\ \alpha^* \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{Q(\bar{x}, 0)} \leq (1.05)\sqrt{\frac{h^2(1 + \gamma^2)}{(1 - h\gamma)^2}} < 0.6,$$

where $(x^*, \alpha^*)$ is the center of the system (4.13). Note that

$$\left\| \begin{bmatrix} x \\ \alpha \end{bmatrix} - \begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} \right\|_{Q(\bar{x}, 0)} = ||\bar{S}^{-1}A(x - \bar{x}) - \alpha\bar{u}||.$$

Let $\mathcal{X}' := \{(x, \alpha) \in \mathbb{R}^{n+1} | Ax - u\alpha \leq b\}$ and

$$F^'_{out} := \{(x, \alpha) \in \mathbb{R}^{n+1} | ||\bar{S}^{-1}A(x - \bar{x}) - \alpha\bar{u}|| \leq 1.6\sqrt{m(m-1)} + 0.6\}.$$

Then by Lemma 3.2.7, $\mathcal{X}' \subset F^'_{out}$. By Lemma 4.5.2, for all $(x, \alpha) \in F^'_{out}$,

$$1.6\sqrt{m(m-1)} + 0.6 \geq ||\bar{S}^{-1}A(x - \bar{x}) - \alpha\bar{u}|| \geq ||\alpha\bar{u}||,$$

76
since $A^T\tilde{S}^{-1}\tilde{u} = 0$. Thus,

$$
\sup_{\alpha} \{ |\alpha| \mid \alpha \in \mathcal{A} \} = \sup_{\alpha} \{ |\alpha| \mid (x, \alpha) \in \mathcal{X}' \text{ for some } x \} = \sup_{\alpha} \{ |\alpha| \mid (x, \alpha) \in F'_{\text{out}} \text{ for some } x \} \leq \frac{1.6\sqrt{m(m - 1)} + 0.6}{||\tilde{u}||}.
$$

This completes the proof. \(\square\)

Theorem 4.5.1 follows immediately from Proposition 4.5.1 and Proposition 4.5.2.

### 4.6 A RHS Parametric Center Algorithm

Now suppose we are interested in tracing the path of parametric centers $\hat{x}_\alpha$ for the systems $P(\alpha)$ as the parameter $\alpha$ ranges over an interval $\alpha \in [0, \alpha^{up}]$ (we can assume without loss of generality that the initial value of the parameter is $\alpha = 0$). Suppose we have an (interior) point $x^0 \in \mathbb{R}^n$ which is a $\delta$-approximate center of the system $Ax \leq b$, with $\delta = 1/21$ (i.e., $||x^0 - \hat{x}||_{Q_0(x^0)} \leq 1/21$). Such a point can be obtained using a center-finding algorithm (of Vaidya [65] or Freund [17]). Based on Theorems 4.4.2, 4.4.3 and 4.5.1 of the previous two sections, we propose the following iterative algorithm for the right-hand-side parametric center problem. We call it **Algorithm RHSPCP**. (*RHSPCP* for Right-Hand-Side Parametric Center Problem.) We shall analyse the performance of the algorithm after stating the steps of the algorithm.
4.6.1 Algorithm RHSPCP

The input to the algorithm includes an \( m \times n \) matrix \( A \), \( m \)-vectors \( b, d \), a scalar \( \alpha^{\text{up}} \), and an (interior) point \( x^0 \in \mathbb{R}^n \) which is a \( \delta \)-approximate center of the system \( Ax \leq b \), with \( \delta = 1/21 \). At STEP 0, initial lower and upper bounds are set to their extreme values. The initial value of \( \alpha \) is \( \alpha^0 = 0 \), the counter \( j \) for the number of iterations is set to zero, the right-hand-side is \( RHS = b \) and the initial iterate is \( \bar{x} = x^0 \).

The iteration starts at STEP 1. At the \( j \)-th iteration, the current value of \( \alpha \) is \( \alpha^j \), the current iterate is \( \bar{x} = x^j \) and the current right-hand-side is \( RHS = b + \alpha^j d \). At STEP 1, the vectors \( \bar{u} \) and \( \bar{v} \) are computed as defined in Theorem 4.4.2. In STEP 2, the constant \( \beta \), which is the increase in \( \alpha \) at the iteration, is computed as defined in Theorem 4.4.2, and the next \( \delta \)-approximate center \( \bar{x}_{\text{new}} \) of system \( P(\alpha^j + \beta) \) is then computed according to Theorem 4.4.2. In STEP 3, a linear segment for \( \alpha \in [\alpha^j, \alpha^{j+1}] \), where \( \alpha^{j+1} = \alpha^j + \beta \), is generated by using \( \bar{x} \) and \( \bar{x}_{\text{new}} \) as endpoints and interpolating, according to Theorem 4.4.3. In STEP 4, the current bounds on \( \alpha^{\text{max}} \) and \( \alpha^{\text{min}} \) are updated, in accordance with Theorem 4.5.1. In STEP 5, the algorithm checks if \( \alpha^j + \beta \geq \alpha^{\text{up}} \). If so, it stops. Otherwise, it updates the counter, the right-hand-side, the current iterate, and then returns to STEP 1.

The output of the algorithm is the piecewise-linear path \( \bar{x}_\alpha \) and the sequence \( \{\alpha^j\} \) of incremental values of \( \alpha \).
Algorithm RHSPCP:

**INPUT:** \( A, b, d, \alpha^{up}, x^0 \) (\( ||x^0 - \hat{x}||_{Q_0(x^0)} \leq 1/21 \)).

**STEP 0:** [Initialization]
Set \( UB = \infty, LB = -\infty, j = 0, \alpha^0 = 0, RHS = b \).

**STEP 1:** [Decomposition of \( d \)]
Set \( \tilde{s} = RHS - A\tilde{x}, \quad \tilde{S} = \text{diag}(\tilde{s}) \). Compute
\[
\tilde{u} = P_{S^{-1}}A\tilde{S}^{-1}d, \quad \text{(projection of } \tilde{S}^{-1}d) \\
\tilde{r} = (A^T\tilde{S}^{-2}A)^{-1}A^T\tilde{S}^{-2}d, \quad \text{(translation vector)}
\]

**STEP 2:** [Steplength and New Approximate Center]
Compute
\[
\beta = \frac{1}{80||\tilde{u}||}, \quad \text{(steplength)} \\
\tilde{s}_\beta = b + \beta \tilde{S}\tilde{u} - A\tilde{x}, \quad \tilde{S}_\beta = \text{diag}(\tilde{s}_\beta) \\
\tilde{\eta} = -(A^T\tilde{S}_\beta^{-2}A)^{-1}A^T\tilde{S}_\beta^{-1}e, \quad \text{(Newton step)} \\
\tilde{x}_{new} = \tilde{x} + \tilde{\eta} + \beta \tilde{r}.
\]

**STEP 3:** [Extend Piecewise Linear Path]
Set \( \alpha^{j+1} = \alpha^j + \beta \), and for \( \alpha \in [\alpha^j, \alpha^{j+1}] \), define
\[
\tilde{x}_\alpha = \tilde{x} + \left( \frac{\alpha - \alpha^j}{\beta} \right) (\tilde{x}_{new} - \tilde{x}).
\]
**STEP 4:** [Update Lower and/or Upper Bounds]

(i) If \( \frac{e^T\tilde{u}}{||\tilde{u}||} > \frac{1}{20} \), then set

\[
LB = \max \left\{ LB, \alpha^j - \frac{22\sqrt{m(m+1)} + 1}{21||\tilde{u}||} \right\};
\]

(ii) If \( \frac{e^T\tilde{u}}{||\tilde{u}||} < -\frac{1}{20} \), then set

\[
UB = \min \left\{ UB, \alpha^j + \frac{22\sqrt{m(m+1)} + 1}{21||\tilde{u}||} \right\};
\]

(iii) If \( \frac{|e^T\tilde{u}|}{||\tilde{u}||} \leq \frac{1}{20} \), then set

\[
LB = \max \left\{ LB, \alpha^j - \frac{1.6\sqrt{m(m-1)} + 0.6}{||\tilde{u}||} \right\}
\]

and

\[
UB = \min \left\{ UB, \alpha^j + \frac{1.6\sqrt{m(m-1)} + 0.6}{||\tilde{u}||} \right\}.
\]

**STEP 5:** [Termination Check and Update]

If \( \alpha^{j+1} \geq \alpha^{up} \), then STOP.

Else set \( j = j + 1 \), \( RHS = RHS + \beta d \), \( \bar{x} = \bar{x}_{new} \), and

GO TO STEP 1.
According to Theorem 4.4.2, if \( x^0 \) is a \( \delta \)-approximate center of system \( P(0) \), then for each \( j = 1, 2, \ldots, \) \( x^j = \tilde{x}_\alpha \) will be a \( \delta \)-approximate center of system \( P(\alpha^j) \), with \( \delta = 1/21 \), and according to Theorem 4.4.3, for each \( \alpha \in [0, \alpha^{up}] \), \( \tilde{x}_\alpha \) will be a \( \delta \)-approximate center of system \( P(\alpha) \), with \( \delta = 0.59 \).

### 4.6.2 Algorithmic Performance

We shall now discuss the performance of Algorithm RHSPCP. Suppose \( \{\alpha^j\} \) is the sequence of parametric values generated by Algorithm RHSPCP, that is, \( \alpha^j \) is the value of \( \alpha \) at the start of the \( j \)-th iteration.

**Proposition 4.6.1 (Geometric Rate of Change)**. At least one of the following is satisfied.

\[ \begin{align*}
    (i) & \quad (\alpha^{j+1} - \alpha^{\text{min}}) \geq (1 + \frac{1}{128m})(\alpha^j - \alpha^{\text{min}}); \\
    (ii) & \quad (\alpha^{\text{max}} - \alpha^{j+1}) \geq (1 - \frac{1}{128m})(\alpha^{\text{min}} - \alpha^j).
\end{align*} \]

**Proof**: Suppose a lower bound, either \( LB \) or \( \widetilde{LB} \), is generated through Theorem 4.5.1 in the \( j \)-th iteration. Notice that \( \widetilde{LB} \leq LB \) if \( m \geq 2 \), so in either case,

\[ \alpha^{\text{min}} \geq \widetilde{LB} = \alpha^j - \frac{1.6\sqrt{m(m-1)} + 0.6}{||\tilde{u}||}. \]

Thus,

\[ \frac{\alpha^j - \alpha^{\text{min}}}{\alpha^{j+1} - \alpha^j} \leq \frac{(1.6\sqrt{m(m-1)} + 0.6)}{||\tilde{u}||} \left( \frac{1}{80||\tilde{u}||} \right) \leq \frac{1.6m}{||\tilde{u}||}(80||\tilde{u}||) \leq 128m. \]
Rearranging terms,

\[(\alpha^{j+1} - \alpha^\text{min}) \geq (1 + \frac{1}{128m})(\alpha^j - \alpha^\text{min}).\]

A parallel analysis demonstrates (ii) if an upper bound is generated. \(\square\)

According to Proposition 4.6.1, we obtain at each iteration either a geometric decrease in the gap \((\alpha - \alpha^\text{min})\) or a geometric increase in the gap \((\alpha^{\text{max}} - \alpha)\). We can thus measure algorithmic performance according to the change in \((\alpha - \alpha^\text{min})\), or \((\alpha^{\text{max}} - \alpha)\), or both, depending on whether we are approximating the upper path \((\varphi(\alpha) \leq 0)\), the lower path \((\varphi(\alpha) \geq 0)\), or both. Suppose that in the course of running Algorithm RHSPCP, a lower bound on \(\alpha^\text{min}\) is never generated. Then all iterates will satisfy criterion (ii) at STEP 4, so that all iterates will generate an upper bound, and all iterates will lie approximately on the upper path.

**Lemma 4.6.1 (Performance based only on \(\alpha^{\text{max}}\)).**

If \(\alpha^\text{min} = -\infty\) or none of the iterates of Algorithm RHSPCP generate a lower bound on \(\alpha^\text{min}\), then the sequence of \(\alpha\) values \(\{\alpha^j\}\) will satisfy

\[\alpha^{\text{max}} - \alpha^j \leq (1 - \frac{1}{128m})^j \alpha^{\text{max}}.\]

In particular, if \(\alpha^{\text{up}} < \alpha^{\text{max}}\), the algorithm will stop after at most \(K = [128m \ln(\frac{\alpha^{\text{max}}}{\alpha^{\text{up}}})]\) iterations.

**Proof:** Under the hypothesis of the Lemma, the algorithm must satisfy criterion (ii) at STEP 4. Thus, by Proposition 4.6.1,

\[\alpha^{\text{max}} - \alpha^{j+1} \leq (1 - \frac{1}{128m})(\alpha^{\text{max}} - \alpha^j).\]
Thus, we obtain the geometric decrease of the Lemma. If \( \alpha^{up} < \alpha^{max} \), let \( K = \lceil 128m \ln(\frac{\alpha^{max}}{\alpha^{max} - \alpha^{up}}) \rceil \). Then

\[
\ln (\alpha^{max} - \alpha^K) \leq K \ln (1 - \frac{1}{128m}) + \ln \alpha^{max} \\
\leq -(\frac{K}{128m}) + \ln \alpha^{max} \\
\leq -\ln (\frac{\alpha^{max}}{\alpha^{max} - \alpha^{up}}) + \ln \alpha^{max} \\
= \ln (\alpha^{max} - \alpha^{up}).
\]

Thus, \( \alpha^{max} - \alpha^K \leq \alpha^{max} - \alpha^{up} \), whereby \( \alpha^K \geq \alpha^{up} \). Thus the algorithm will stop. \( \Box \)

Suppose instead that none of the iterates of Algorithm RHSPCP generate an upper bound at STEP 4. Analogous to Lemma 4.6.1, we have

**Lemma 4.6.2 (Performance based only on \( \alpha^{min} \))**.

If \( \alpha^{max} = \infty \) or none of the iterates of Algorithm RHSPCP generate an upper bound on \( \alpha^{max} \), then the sequence of \( \alpha \) values \( \{\alpha^j\} \) will satisfy

\[
\alpha^j - \alpha^{min} \geq (1 + \frac{1}{128m})^j(-\alpha^{min}).
\]

In particular, the algorithm will stop after at most \( K = \lceil 128m \ln(\frac{\alpha^{up} - \alpha^{min}}{-\alpha^{min}}) \rceil \) iterations.

We next examine the case when the algorithm generates both upper and lower bounds. We need first the following result.

**Lemma 4.6.3** If criterion (ii) of STEP 4 of Algorithm RHSPCP is satisfied at iteration \( j \), then in all subsequent iterations, criterion (i) of STEP 4 will not be satisfied.
Proof: Suppose $e^T \bar{u} < -\frac{1}{20} \|\bar{u}\|$ in iteration $j$. Then by Lemma 4.5.1, $\varphi(\alpha^j) = e^T \bar{S}^{-1} u < 0$. Then by Proposition 4.2.1, $\varphi(\alpha) \leq \varphi(\alpha^j) < 0$ for all $\alpha \geq \alpha^j$. Therefore, in all subsequent iterations, criterion (i) will not be satisfied, as it would imply that $\varphi(\alpha) \geq 0$ for some $\alpha \geq \alpha^j$. □

The significance of Lemma 4.6.3 is as follows: if at the $j$-th iteration, an upper bound is generated, then an upper bound is generated at every subsequent iteration.

Lemma 4.6.4 (Performance based on upper and lower bounds).

If Algorithm RHSPCP generates both lower and upper bounds, then there is some $j^*$ such that for $j < j^*$, the algorithm generates lower bounds only, and for all $j \geq j^*$, the algorithm generates upper bounds, and

(i) for all $j < j^*$, $\alpha^j - \alpha^\text{min} \geq (1 + \frac{1}{128m})^j(-\alpha^\text{min})$;

(ii) for all $j \geq j^*$, $\alpha^\text{max} - \alpha^j \leq (1 - \frac{1}{128m})^{j-j^*}\alpha^\text{max}$.

Furthermore, if $\alpha^\text{up} < \alpha^\text{max}$, then the algorithm will stop after at most $K = [256m \ln(\frac{\alpha^\text{max} - \alpha^\text{min}}{2})] - [128m(\ln(\alpha^\text{max} - \alpha^\text{up}) + \ln(-\alpha^\text{min}))]$ iterations.

Proof: The existence of $j^*$ is guaranteed by Lemma 4.6.3. To see this, suppose the algorithm satisfies criterion (ii) or (iii) of STEP 4 at the first iteration, then we may take $j^* = 1$. Otherwise, the algorithm satisfies only criterion (i) of STEP 4 initially. By assumption, the problem is bounded, so that both $\alpha^\text{max}$ and $\alpha^\text{min}$ are finite. Then criterion (ii) or (iii) must be satisfied at some iteration $j > 1$. The geometric convergence rates are then a consequence of Proposition 4.6.1. Finally, suppose $\alpha^\text{up} < \alpha^\text{max}$, and let $K$
be as defined in the Lemma. Let $\alpha^* = \alpha^{j*}$ and note that

$$\frac{1}{2}(\ln(\alpha^{\max} - \alpha^*) + \ln(\alpha^* - \alpha^{\min})) \leq \ln\left(\frac{\alpha^{\max} - \alpha^{\min}}{2}\right),$$

from the arithmetic mean-geometric mean inequality. Thus,

$$K \geq 256m \ln\left(\frac{\alpha^{\max} - \alpha^{\min}}{2}\right) - 128m(\ln(\alpha^{\max} - \alpha^{up}) + \ln(-\alpha^{\min}))$$

$$\geq 128m \ln\left(\frac{\alpha^{\max} - \alpha^*}{\alpha^{\max} - \alpha^{up}}\right) + 128m \ln\left(\frac{\alpha^* - \alpha^{\min}}{-\alpha^{\min}}\right).$$

According to Lemma 4.6.2, with $\alpha^{up}$ replaced by $\alpha^*$, $\alpha^j \geq \alpha^*$ after at most \lfloor 128m \ln \left(\frac{\alpha^* - \alpha^{\min}}{-\alpha^{\min}}\right) \rfloor$ iterations. Furthermore, according to Lemma 4.6.1, with $\alpha^j$ replaced by $\alpha^j - \alpha^*$, and $\alpha^{up}$ replaced by $\alpha^{up} - \alpha^*$, we get $\alpha^j \geq \alpha^{up}$ after at most $K$ iterations. \hfill \Box

### 4.7 A Strictly Monotonic Algorithm for Linear Programming

Algorithm RHSPCP can be applied to solve the linear programming problem as follows. Suppose we wish to solve the linear program

$$\overline{LP} : \quad \begin{array}{rl}
\text{maximize} & \tilde{c}^T \tilde{x} \\
\text{subject to} & \tilde{A} \tilde{x} \leq b,
\end{array}$$

where $\tilde{A}$ is an $m \times (n + 1)$ matrix, and $\tilde{c}, \tilde{x} \in \mathbb{R}^{n+1}$. If $\tilde{c} = 0$, then any feasible solution of the program is optimal. Therefore, we assume that $\tilde{c} \neq 0$. Upon setting $\tilde{c}^T \tilde{x} = \alpha$ and eliminating one of the $(n + 1)$ variables of $\tilde{x}$, $\overline{LP}$
is transformed to the equivalent problem

\[
LP: \quad \alpha^{\text{max}} := \text{maximize } \alpha
\]
\[
\text{subject to } Ax \leq b + d\alpha,
\]

where \( x \in \mathbb{R}^n, \quad A \in \mathbb{R}^{m \times n}, \quad b, d \in \mathbb{R}^m, \) and the data \((A, \ b, \ d)\) is a linear transformation of the data \((\tilde{A}, \ \tilde{b}, \ \tilde{c})\). We can then solve \( LP \) by using Algorithm RIISPCP to trace the path of centers of the parametric family of linear inequality systems \( Ax \leq b + \alpha d \) for \( \alpha < \alpha^{\text{max}} \).

Suppose \( x^0 \) is a given starting point for which \( (x^0, \ 0) \) satisfies the starting criterion of Algorithm RIISPCP, namely \( ||x^0 - \tilde{x}_0||_{Q_0(x^0)} \leq 1/21 \). Then we can use Algorithm RIISPCP to generate the path \( \tilde{x}_\alpha \) for \( \alpha \in [0, \ \alpha^{\text{max}}] \). The sequence \{\( \alpha^j \)\} of \( \alpha \) values will be strictly increasing, according to Theorem 4.4.2, i.e., the objective value will be strictly increasing at each iteration. Let \( L \) be the total number of bits in a binary encoding of the problem instance. In order to analyze the complexity of Algorithm RIISPCP when applied to \( LP \), we consider three cases.

**Case 1:** The linear program is unbounded. In this case, the algorithm never generate a finite upper bound on \( \alpha^{\text{max}} \), which equals infinity. After \( K = O(mL) \) iterations, \( \alpha^K \geq -\alpha^{\text{min}}(1 + \frac{1}{128m})^K \) will exceed \( 2^L \), and we can conclude that \( LP \) is unbounded.

**Case 2:** The linear program is bounded and the path indicator function at \( \alpha = 0 \) is negative (i.e., \( \varphi(0) < 0 \)). This implies that criterion (ii) or (iii) at STEP 4 of Algorithm RIISPCP is satisfied. This being the case, the algorithm will always generate upper bounds, and after \( K = O(mL) \) iterations, \( (\alpha^{\text{max}} - \alpha^K) \leq \alpha^{\text{max}}(1 - \frac{1}{128m})^K \) will be less than \( 2^{-L} \), whereby \( \tilde{x}_{\alpha^K} \) can be rounded to an optimal solution of \( LP \) (see Karmarkar [40]).

86
Case 3: The linear program is bounded and $\varphi(0) > 0$. In this case, we can show as in Case 1 that after $K = O(mL)$ iterations, $\varphi(\alpha^K) \leq 0$, for otherwise LP would be unbounded. Furthermore, after an additional $O(mL)$ iterations, we will obtain via Case 2 an approximate solution that we can round to an optimal solution.

We observe that in either of the three cases, Algorithm RHISPCP will process LP, by detecting unboundedness or producing an optimal solution, after $O(mL)$ iterations.

This algorithm falls into the class of central trajectory-based algorithms but is inferior in term of complexity; the bound of $O(mL)$ iterations is worse that the bound of $O(\sqrt{mL})$ iterations for algorithms (such as Renegar [57] and Vaidya [66]) that trace the weighted center of the system

$$
Ax \leq b
$$

$$
c^T x \geq \alpha,
$$

as $\alpha$ is increased, or to the bound of $O(\sqrt{mL})$ iterations for algorithms based on the barrier penalty method that trace the solution to

$$
\max \quad c^T x + \varepsilon \sum_{i=1}^{m} \ln s_i
$$

$$
s.t. \quad Ax + s = b
$$

$$
s > 0,
$$

as $\varepsilon \to 0^+$. (See, for example, Gonzaga [29] and Monteiro and Adler [52], among others.)
All three methods follow the same path in their idealized version. Yet the latter two converge in $O(\sqrt{mL})$ iterations, which is superior to Algorithm RHSPCP. However, these other algorithms do not guarantee strict improvement in the objective value (but do guarantee strict improvement in the duality gap). In contrast, Algorithm RHSPCP will guarantee strict improvement of $\frac{\alpha_{\text{max}} - c^T x}{128m}$ in the objective value at each iteration. Perhaps it is the implicit imposition of the strict improvement in objective value that increases the iteration bound by a factor of $O(\sqrt{m})$. Furthermore, Algorithm RHSPCP does not assume that $LP$ is bounded. Instead, it will detect unboundedness of $LP$ directly.

As a final note, we observe that Algorithm RHSPCP can be used to mimic the algorithm of Renegar [57], tracing the center of the system

$$Ax \leq b$$

$$-c^T x \leq 0 - \alpha$$

as $\alpha$ is increased. Thus, Renegar's set-up is a special case of the right-hand-side parametric center problem ($RHSPCP$). However, we see no way to cast $RHSPCP$ as a special case of the set-up used by Renegar [57]; we allow all RHS values to vary simultaneously, which is apparently more general than in [57].

88
Chapter 5

Multiple Constraints

Parametric Center Problem

5.1 Introduction

In this chapter, we consider the general case of the Parametric Center Problem, where some or all of the constraints are allowed to change parametrically. Specifically, we consider the parametric family of linear inequality systems

\[ P(\alpha) : \quad (A + \alpha B)x \leq b + \alpha d, \]

where \( A, B \in \mathbb{R}^{m \times n} \) are \( m \times n \) matrices, \( b, d \in \mathbb{R}^m \) are (column) \( m \)-vectors, and \( \alpha \) is a scalar parameter. Suppose there are \( k \) nonzero rows in the matrix \( B \) and \( l \) nonzero rows in the matrix \([B, d]\). That is, we allow \( l \) constraints to vary with the parameter, \( l - k \) of which involve varying only the right-hand-side. We are interested in tracing the path of centers of \( P(\alpha) \) as \( \alpha \) varies over some specified range. Specifically, our goal is an algorithm to generate
a piecewise linear path of approximate centers of $P(\alpha)$, that is, approximate solutions to the parametric family of programs

$$PCP(\alpha): \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i$$

subject to \quad \(A + \alpha B)x + s = b + \alpha d$$

$$s > 0.$$  

This problem is very closely related to the generalized linear fractional programming problem (GLFP). Let $\alpha^*$ be the maximal $\alpha$ such that the system $P(\alpha)$ is feasible. Then, under the following assumptions, it is easy to see that $\alpha^*$ equals the maximal value of the following GLFP program.

$$GLFP: \quad \alpha^* := \max_{x} \min_{i} \left\{ \frac{b_i - A_i x}{B_i x - d_i} \right\}$$

subject to \quad Ax \leq b.$$

On the other hand, suppose we are interested in solving the following GLFP program.

$$GLFP: \quad \tilde{\alpha}^* := \max_{x} \min_{i} \left\{ \frac{f_i - C_i x}{D_i x - h_i} \right\}$$

subject to \quad \tilde{A}x \leq \tilde{b},$$

where $\tilde{A} \in \mathbb{R}^{m \times n}$, $\tilde{b} \in \mathbb{R}^{m}$, $C, D \in \mathbb{R}^{k \times n}$ and $f, h \in \mathbb{R}^{k}$, and $[C_i, f_i] \neq 0$, and $[D_i, h_i] \neq 0$ for $i = 1, 2, \ldots, k$. Then let

$$A = \begin{bmatrix} \tilde{A} \\ C \end{bmatrix}, \quad B = \begin{bmatrix} O \\ D \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{b} \\ f \end{bmatrix}, \quad d = \begin{bmatrix} 0 \\ d \end{bmatrix}.$$  

Then we see that

$$\tilde{\alpha}^* := \max_{x, \alpha} \alpha$$

s.t. \quad (A + \alpha B)x \leq b + \alpha d,$$
where $A, B \in \mathbb{R}^{m \times n}$ and $b, \bar{a} \in \mathbb{R}^m$, with $m = \tilde{m} + k$.

The organization of the rest of this chapter is as follows. We state the assumptions and notation of this chapter in the next section. In Section 3, we state the main results of this chapter and present our algorithm for the multiple constraints parametric center problem. The proofs of the main results are given in the next two sections. In Section 4, we prove some preliminary lemmas, and in the final section, we give the proofs of the main theorems.

\section{5.2 Assumptions and Notation}

We shall make the following assumptions in this chapter.

**Assumption 1.** The set $\{x | Ax < b\}$ is nonempty and bounded;
**Assumption 2.** For every $x$ satisfying $Ax \leq b$, we have $Bx \geq d$; $Bx \neq d$.

Under Assumption 1, the center of system $P(0)$ exists uniquely, and Assumption 2 implies that the solution set of $P(\alpha)$,

$$\mathcal{X}_{\alpha} := \{x | (A + \alpha B)x \leq b + \alpha d\},$$

is shrinking for increasing values of $\alpha$. Also, under these Assumptions, it is easy to see that $0 < \alpha^* < \infty$, and for all $\alpha, \ 0 \leq \alpha < \alpha^*$, the interior of the set $\mathcal{X}_{\alpha}$ is nonempty and bounded. Finally, note that the Assumptions imply that the number of nonzero rows in the matrix $[B, d]$ is $l \geq 1$.  

91
Notation

We let $\hat{x}_\alpha$ denote the center of system $P(\alpha)$. That is, $\hat{x}_\alpha$ is the unique solution of the nonlinear program (called parametric center problem)

\[
PCP(\alpha) : \quad \text{maximize} \quad \sum_{i=1}^{m} \ln s_i \\
\text{subject to} \quad (A + \alpha B)x + s = b + \alpha d \\
s > 0.
\]

The Karush-Kuhn-Tucker conditions, which characterize the center $\hat{x}_\alpha$ of system $P(\alpha)$ are

\[
\hat{s}_\alpha = b + \alpha d - (A + \alpha B)\hat{x}_c > 0, \quad (5.2)
\]

\[
(A + \alpha B)^T \hat{S}_\alpha^{-1}e = 0. \quad (5.3)
\]

Given an interior point $\bar{x}$ satisfying $\bar{s} = b - A\bar{x} > 0$, we let

\[
\bar{\alpha} = \alpha_{\bar{x}} := 1/||\hat{S}^{-1}(B\bar{x} - d)||_\infty \quad (5.4)
\]

and, for each $\alpha \in [0, \bar{\alpha})$, we let

\[
\hat{s}_\alpha := (b + \alpha d) - (A + \alpha B)\bar{x} > 0 \quad (5.5)
\]

and

\[
Q_\alpha(\bar{x}) := (A + \alpha B)^T \hat{S}_\alpha^{-2}(A + \alpha B). \quad (5.6)
\]

Remark: Note that we may express $\bar{\alpha}$ as

\[
\bar{\alpha} = \alpha_{\bar{x}} = \frac{1}{||\hat{S}^{-1}(B\bar{x} - d)||_\infty} = \min_i \left\{ \frac{b_i - A_i \bar{x}}{B_i \bar{x} - d_i} \right\}.
\]
Therefore, $\tilde{\alpha}$ corresponds to the value of the generalised linear fractional program $GLFP$ evaluated at $\tilde{x}$. Also, $\tilde{\alpha}$ is the value of $\alpha$ such that the polytope $\mathcal{X}_\alpha$ just touches the point $\tilde{x}$, so that $\mathcal{X}_\alpha^+ \overset{\text{def}}{=} \{ x \mid (A + \alpha B)x < b + \alpha d \}$.

Given a matrix $M$, we let $M_i$ denote the $i^{th}$ row of $M$ and $M_i^T$ denote the transpose of $M_i$. If $M$ is positive definite, $\|v\|_M$ denotes the norm defined by $M$ that is given by $\|v\|_M = \sqrt{v^TMv}$.

### 5.3 Main Results and Algorithm

In this chapter, we propose an algorithm, based on Newton’s method, for tracing the path of parametric centers. The main results of this chapter are the following two theorems, which form the basis for the iterative algorithm that we shall present here. For simplicity, we shall consider increasing $\alpha$ from $\alpha = 0$ to some $\alpha > 0$. That is, we shall consider one iteration of the iterative algorithm.

Suppose we have a $\delta$-approximate center $\tilde{x}$ of system $P(0)$. We need to determine the magnitude of increase $\beta$ in the parametric value such that by taking a Newton iterate from $\tilde{x}$, in the problem to find the center of system $P(\beta)$, we will get a new point $\tilde{x}_{\text{new}}$ which is again a $\delta$-approximate center of the new system $P(\beta)$.

Recall that there are $k$ nonzero rows in the matrix $B$ and $l \geq 1$ nonzero rows in the matrix $[B, \ b]$. In Section 5, we shall prove the following results.

**Theorem 5.3.1 (Improvement Theorem).**

*Suppose $\tilde{x}$ satisfies $\tilde{s} = b - A\tilde{x} > 0$ and is near the center of system $P(0)$*
in the sense that \( ||\tilde{x} - \tilde{x}||_{Q_0(x)} \leq 1/21 \), where \( Q_0(\tilde{x}) \) is given by (5.5)-(5.6).

Let \( \tilde{\alpha} \) be defined by (5.4). Let \( \beta = \frac{\tilde{\alpha}}{88(\sqrt{l + k})} \), and let \( \tilde{x}_{\text{new}} = \tilde{x} + \eta \), where \( \eta = -Q^{-1}_\beta(\tilde{x})(A + \beta B)^T \tilde{S}^{-1}_\beta e \) (i.e., \( \tilde{x}_{\text{new}} \) is a Newton iterate from \( \tilde{x} \) in center problem \( PCP(\beta) \)). Then \( ||\tilde{x}_{\text{new}} - \hat{x}_\beta||_{Q_\beta(\tilde{x}_{\text{new}})} \leq 1/21 \), where \( Q_\beta(x) \) is defined by (5.5)-(5.6), and \( \hat{x}_\beta \) denotes the center of system \( P(\beta) \).

Hence, letting \( \beta = \frac{\tilde{\alpha}}{88(\sqrt{l + k})} \), we may repeat the procedure with \( \tilde{x}_{\text{new}} \) replacing \( \tilde{x} \), \( A + \beta B \) replacing \( A \) and \( b + \beta d \) replacing \( b \). Note that the increase in \( \alpha \) is \( \frac{\tilde{\alpha}}{88(\sqrt{l + k})} \), which is a fraction of \( \tilde{\alpha} \), the value of GLFP at \( \tilde{x} \). It makes good sense that the fraction depends on, and is inversely proportional to, the total number of varying constraints through the quantities \( k \) and \( l \).

Next, we have the following result which allows us to extend and generate a piecewise linear path of approximate centers. In the following theorem, we increase \( \alpha \) from \( \alpha = 0 \) to \( \alpha = \beta \). We then take a step from \( \tilde{x} \) to \( \tilde{x}_{\text{new}} = \tilde{x} + \eta \), where \( \eta \) is a Newton step from \( \tilde{x} \) in problem \( PCP(\beta) \). We then extend the path of approximate centers by linearly interpolating between \( \tilde{x} \) and \( \tilde{x}_{\text{new}} \).

**Theorem 5.3.2 (Path Extension Theorem)**

Under the same conditions and definitions as Theorem 5.3.1, define \( \tilde{x}_\alpha \), for all \( \alpha \in [0, \beta] \), by

\[
\tilde{x}_\alpha := \tilde{x} + \left( \frac{\alpha}{\beta} \right)(\tilde{x}_{\text{new}} - \tilde{x}).
\]

Then

\[
||\tilde{x}_\alpha - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq 0.38.
\]
5.3.1 A Parametric Center Algorithm

Based on Theorem 5.3.1 and Theorem 5.3.2, we propose the following Newton method-based algorithm for tracing the path of centers of \( P(\alpha) \) as \( \alpha \) varies over the range \( \alpha \in [0, \alpha^\text{upper}] \). The input of the algorithm include the \( m \times n \) matrices \( A \) and \( B \), where we know that \( k \) of the rows of \( B \) are nonzero, the \( m \)-vectors \( b \) and \( d \), such that there are \( l \) nonzero rows in the matrix \([B, d]\), a scalar \( \alpha^\text{upper} \), \( 0 < \alpha^\text{upper} \leq \alpha^* \), and an approximate center \( x^0 \in \{x | Ax < b\} \) of system \( P(0) \) in the sense that \( ||x^0 - \hat{x}||_{Q_0(x^0)} \leq 1/21 \). Such an approximate center can be obtained by using an algorithm for the center problem (see Vaidya [65] or Freund [17]). The output includes a sequence of breakpoints \( \{(\tilde{x}^j, \alpha^j)\}, \tilde{x}^j = \tilde{x}_{\alpha^j}, \) and a piecewise linear path of approximate centers \( \tilde{x}_\alpha \) (in the sense that \( ||\tilde{x}_\alpha - \tilde{x}_{\alpha}||_{Q_0(x^0)} \leq 0.38 \) ) for \( \alpha \in [0, \alpha^\text{upper}] \). We note that, of course, the algorithm may not terminate if \( \alpha^\text{upper} > \alpha^* \).

Algorithm PCP

INPUT: \( A, B, b, d, \alpha^\text{upper}, k, l, x^0 (||x^0 - \hat{x}||_{Q_0(x^0)} \leq 1/21) \);

INITIALIZATION:
Set \( j = 0, \alpha^0 = 0, \tilde{x} = x^0, \tilde{A} = A, \tilde{b} = b. \)

ITERATION: Repeat the following steps until \( \alpha^j \geq \alpha^\text{upper} \).
Step 1. Set
\[
\tilde{\alpha} = \frac{1}{||S^{-1}(B\tilde{x} - d)||_{\infty}}, \quad \beta = \frac{\tilde{\alpha}}{88(\sqrt{l} + k)}.
\]
Step 2. Compute the Newton step $\bar{\eta}$ from $\bar{x}$ in problem $PCP(\beta)$:

$$\bar{\eta} = -Q_\beta^{-1}(\bar{x})(\bar{A} + \beta B)^T\bar{s}_\beta^{-1}e,$$

where (see (5.5)-(5.6))

$$\bar{s}_\beta = (\bar{b} + \beta d) - (\bar{A} + \beta B)\bar{x}$$

and

$$Q_\beta(\bar{x}) = (\bar{A} + \beta B)^T\bar{s}_\beta^{-2}(\bar{A} + \beta B).$$

Step 3. Set $\bar{x}_{\text{new}} = \bar{x} + \bar{\eta}$ and $\alpha^{j+1} = \alpha^j + \beta$,

and for $\alpha \in [\alpha^j, \alpha^{j+1}]$, define

$$\bar{x}_\alpha = \bar{x} + \left(\frac{\alpha - \alpha^j}{\beta}\right)(\bar{x}_{\text{new}} - \bar{x}).$$

Step 4. Update

$$j \leftarrow j + 1, \quad \bar{x} \leftarrow \bar{x}_{\text{new}}, \quad \bar{A} \leftarrow \bar{A} + \beta B, \quad \bar{b} \leftarrow \bar{b} + \beta d,$$

and go to Step 1.

Now, from a property of the center of the system $Ax \leq b$ (Corollary 3.1.1) and the Assumptions, we have an upper bound on $\alpha^*$ in terms of $\bar{\alpha}$ and $m$ as follow. Let $u$ and $v$ be the following constants.

$$u = \max_i \max_{\bar{x}} \{B_ix - d_i \mid Ax \leq b\} \quad (5.7)$$

96
\[ v = \min_x \{ \max_i B_i x - d_i \mid Ax \leq b \} \]  

(5.8)

[Note that \( u \) can be obtained by solving \( k \) linear programs, and \( v \) can be obtained by solving one linear program.] Since \( \{ x \mid Ax \leq b \} \) is compact under Assumption 1, we have \( \max_x \{ B_i x - d_i \mid Ax \leq b \} < \infty \) for each \( i = 1, 2, \ldots, m \). Thus, \( u < \infty \). By Assumption 2, for each \( x \) satisfying \( Ax \leq b \), there exists an \( i \) such that \( B_i x > d_i \), therefore \( \max_i B_i x - d_i > 0 \). Since \( \{ x \mid Ax \leq b \} \) is compact under Assumption 1, we have \( v > 0 \). It is easy to see that for all \( x \) satisfying \( Ax \leq b \), \( B_i x - d_i \leq u \) for all \( i = 1, 2, \ldots, m \) and \( 0 < v \leq \max_i B_i x - d_i \leq u < \infty \).

Now, define the constant \( c \) by
\[ c := \frac{u}{v}. \]  

(5.9)

Then we see that \( 0 < c < \infty \) and we have the following.

**Theorem 5.3.3** Suppose \( \hat{x} = \hat{x}_0 \) is the center of system \( P(0) \). Let \( \hat{\alpha} = \alpha_{\hat{x}} \) be given by (5.4). Then \( \alpha^* \leq cm \hat{\alpha} \), where \( c \) is the constant given by (5.9).

**Proof:** For \( i \in \{ 1, 2, \ldots, m \} \) and \( x \in \{ x \mid Ax \leq b \} \), we have
\[ 0 \leq b_i - A_i x \leq m \hat{s}_i, \]
from Corollary 3.1.1. Therefore,
\[ \frac{b_i - A_i x}{B_i x - d_i} \leq \left( \frac{m \hat{s}_i}{B_i \hat{x} - d_i} \right) \left( \frac{B_i \hat{x} - d_i}{B_i x - d_i} \right). \]

Hence,
\[ \alpha^* = \max_x \left\{ \min_i \frac{b_i - A_i x}{B_i x - d_i} \mid Ax \leq b, \right\} \]

97
\[
\leq \max_x \left\{ \min_i \left( \frac{m \hat{s}_i}{B_i \hat{x} - d_i} \right) \left( \frac{B_i \hat{x} - d_i}{B_i \hat{x} - d_i} \right) \mid Ax \leq b \right\} \\
\leq \frac{m(u/v)}{\|\hat{S}^{-1}(B \hat{x} - d)\|_\infty} \\
= cm\hat{\alpha}. \quad \square
\]

Now, taking $\delta = 1/21$, let $\bar{c}$ be the following constant.

\[
\bar{c} := (1.1)\left(\frac{u}{v}\right). \quad (5.10)
\]

Note that $\bar{c} = (1.1)c$, where $c$ is the constant defined by (5.9). From Theorem 5.3.3, we have the following.

**Corollary 5.3.1** Suppose $\bar{x}$ is a $\delta$-approximate center of system $P(0)$ with $\delta = 1/21$. Let $\hat{\alpha}$ be given by (5.4). Then $\alpha^* \leq \bar{c}m\hat{\alpha}$, where $\bar{c}$ is the constant defined by (5.7)–(5.10).

**Proof:** In Lemma 5.4.9 of the next section, we show that $\hat{\alpha} \leq \left(\frac{1+\delta}{1-\delta}\right)\hat{\alpha} = (1.1)\hat{\alpha}$. The proof then follows from Theorem 5.3.3. \quad \square

Therefore, the increase $\beta$ in the $\alpha$ value at each iteration of Algorithm PCP is a fraction of $O\left(\frac{1}{m(\sqrt{1+k})}\right)$ of the maximal value $\alpha^*$ of $\alpha$ such that the system $P(\alpha)$ is feasible. Hence, the algorithm will require $O(m(\sqrt{1+k}))$ iterations to achieve a fixed increase in the $\alpha$ value.

### 5.4 Preliminary Lemmas

Let us begin with some preliminary lemmas.
Lemma 5.4.1 Let \( Q \) be a (symmetric) positive definite matrix, and \( d \) be a given nonzero \( n \)-vector. Suppose \( \bar{x} \in \mathbb{R}^n \) satisfies \( \bar{h} = d^T \bar{x} - h > 0 \). Then we have
\[
\min_{x} \{ \| x - \bar{x} \|_Q^2 \mid d^T x = h \} = \frac{\bar{h}^2}{d^T Q^{-1} d}.
\]

**Proof:** Follows directly from the fact that
\[
x = \bar{x} - \left( \frac{\bar{h}}{d^T Q^{-1} d} \right) Q^{-1} d
\]
is the solution of the minimization program. \(\square\)

That is, the distance (in the \( Q \)-norm) of the point \( \bar{x} \) to the plane \( \{ x \mid d^T x = h \} \) is inversely proportional to the \( Q^{-1} \)-norm of the vector \( \frac{1}{\bar{h}} d \). Under Assumption 2, each (interior) point of the polytope \( \{ x \mid A x < b \} \) is at least some positive distance away from each of the planes \( \{ x \mid B_i^T x = d_i \} \), \( i = 1, 2, \ldots, m \). Therefore, each \( B_i \) should be bounded in some sense which we will make precise shortly in the next lemma.

Let \( \bar{s} \) satisfying \( \bar{s} = b - A \bar{x} > 0 \) be a given (interior) point. Define \( \bar{\alpha} \) by (5.4). Let \( \bar{s}_\alpha \) be given by (5.5) and let \( Q_\alpha(\bar{x}) \) be defined by (5.6).

Lemma 5.4.2 Under the Assumptions, suppose \( 0 \leq \alpha < \bar{\alpha}, \) then
\[
(i) \quad \| (\tilde{S}^{-1} B)_{i}^T \|_{Q_\alpha^{-1}(\bar{x})} \leq \| \tilde{S}^{-1}(B \bar{x} - d) \|_\infty, \quad i = 1, 2, \ldots, m;
(ii) \quad \| (\tilde{S}_\alpha^{-1} B)_{i}^T \|_{Q_\alpha^{-1}(\bar{x})} \leq \| \tilde{S}_\alpha^{-1}(B \bar{x} - d) \|_\infty, \quad i = 1, 2, \ldots, m;
(iii) \quad \| B^T S_\alpha^{-1} e \|_{Q_\alpha^{-1}(\bar{x})} \leq k \| \tilde{S}_\alpha^{-1}(B \bar{x} - d) \|_\infty,
\]
where \( k \) is the number of non-zero rows in the matrix \( B \).

**Proof:** (i) Let \( Q_\alpha = Q_\alpha(\bar{x}) \). We first note that \( Q_\alpha \) is positive definite for each \( \alpha \) satisfying \( 0 \leq \alpha < \bar{\alpha} \), so that the norm \( \| \cdot \|_{Q_\alpha^{-1}} \) is well-defined.
Let $F_{in} = \{x|\|x - \bar{x}\|_{Q_{\alpha}} \leq 1\} \text{ and let } \mathcal{X}_{\alpha} \text{ be given by (5.1). Then, from Lemma 3.2.7, } F_{in} \subset \mathcal{X}_{\alpha}. \text{ Under the Assumptions, } B_i x < d_i \text{ implies that } x \notin \mathcal{X}_{\alpha}, \text{ whereby } x \notin F_{in}. \text{ Therefore, using Lemma 5.4.1, we have}

$$\frac{(B_i\bar{x} - d_i)^2}{B_iQ_{\alpha}^{-1}B_i^T} \leq \min \{\|x - \bar{x}\|_{Q_{\alpha}}^2 | B_i x = d_i\} \geq 1.$$ 

Therefore,

$$B_iQ_{\alpha}^{-1}B_i^T \leq (B_i\bar{x} - d_i)^2. \quad (5.10)$$

Hence

$$\|((S^{-1}B_i)^T)\|_{Q_{\alpha}^{-1}} = \left(\frac{B_iQ_{\alpha}^{-1}B_i^T}{\bar{s}_i^2}\right)^{1/2} \leq \frac{|(B_i\bar{x} - d_i)|}{\bar{s}_i} \leq \|S^{-1}(B\bar{x} - d)\|_{\infty}.$$ 

Part (ii) follows similarly from (5.10). Using the triangle inequality, part (iii) follows from part (ii) immediately. \(\square\)

The next lemma shows the close relationship between the two norms $\|\cdot\|_{Q}$ and $\|\cdot\|_{Q^{-1}}$.

**Lemma 5.4.3** Let $Q_1$ and $Q_2$ be two (symmetric) positive definite $n \times n$ matrices, and let $\|\cdot\|_{Q_i}$ and $\|\cdot\|_{Q_i^{-1}}$ denote the norms defined by $Q_i$ and $Q_i^{-1}$ respectively. Suppose there exists a constant $\kappa > 0$ such that $\|v\|_{Q_1} \leq \kappa\|v\|_{Q_2}$ for all $v \in \mathbb{R}^n$. Then, for any $p \in \mathbb{R}^n$, we have $\|p\|_{Q_2^{-1}} \leq \kappa\|p\|_{Q_1^{-1}}$.

**Proof:** Let $p \in \mathbb{R}^n$ be given. Without loss of generality, we may assume that $p \neq 0$. Then there exists $\hat{u} \in \mathbb{R}^n$ such that $p^T\hat{u} = 1$. Therefore, using
Lemma 5.4.1 and by the hypothesis, we have

\[
\frac{1}{||p||_{Q^{-1}}} = (p^T Q^{-1} p)^{-1/2} = \min_u \{||u - \hat{u}||_{Q_1} \mid p^T u = 0\} \\
\leq \min_u \{\kappa||u - \hat{u}||_{Q_2} \mid p^T u = 0\} \\
= \kappa(p^T Q_2^{-1} p)^{-1/2} = \kappa/||p||_{Q_2^{-1}}
\]

Hence,

\[
||p||_{Q_2^{-1}} \leq \kappa ||p||_{Q_1^{-1}}.
\]

The next lemma show that we can bound the norm of a matrix in term of a bound on the transposes of the rows of the matrix (taken as vectors).

**Lemma 5.4.4** Let \( Q \) be a positive definite \( n \times n \) matrix and let \( || \cdot ||_Q \) denote the norm defined by \( Q \). Suppose \( M \) is a \( k \times n \) matrix and \( c > 0 \) is a constant such that \( ||M_i^T||_Q \leq c, \quad i = 1, 2, \ldots, k \), where \( M_i^T \) denotes the transpose of the \( i^{th} \) row of \( M \). Then

\[
||M^T||_Q := \max_{||v|| = 1} ||M^T y||_Q \leq c\sqrt{k}.
\]

**Proof:** Using the triangle inequality,

\[
||M^T y||_Q = ||\sum_{i=1}^k y_i M_i^T||_Q \leq \sum_{i=1}^k |y_i| ||M_i^T||_Q \leq \sqrt{k}||y||_c,
\]

where the last inequality follows from the fact that

\[
\sum_{i=1}^k |y_i| = ||y||_1 \leq \sqrt{k}||y||.
\]

**Remark:** It can be easily shown that \( ||M^T||_Q = ||MQM^T||^{1/2} \), and if
\( M^T = [N^T, 0] \), then \( ||M^T||_Q = ||N^T||_Q \).

The following lemma is an immediate corollary of Lemma 5.4.4 and Lemma 5.4.2.

**Lemma 5.4.5** Under the same conditions and definitions as Lemma 5.4.2, we have

\[ ||B^T \tilde{S}^{-1}||_{Q_{\alpha}^1(x)} \leq \sqrt{k} \||\tilde{S}^{-1}(B\bar{x} - d)||_{\infty} = \sqrt{k}/\bar{\alpha}. \]

**Proof:** Follows immediately from Lemma 5.4.2 and Lemma 5.4.4. \( \square \)

Next, we want to show the relationship between the two norms \( \| \cdot \|_{Q_{\alpha}(x)} \) and \( \| \cdot \|_{Q_{\alpha}(x)} \), where \( \| \cdot \|_{Q_{\alpha}(x)} \) is the norm defined using the Hessian at \( \bar{x} \) of the logarithmic barrier function in PCP(\( \alpha \)). Before that, we need the following

**Lemma 5.4.6** Suppose \( 0 \leq \alpha \leq \xi \bar{\alpha}, \ 0 \leq \xi < 1 \). Then

\[
\begin{align*}
(i) & \quad \| \tilde{S}^{-1} \tilde{S}_\alpha \| \leq 1; \quad \text{and} \\
(ii) & \quad \| \tilde{S}^{-1} \tilde{S} \| \leq 1/(1 - \xi).
\end{align*}
\]

**Proof:** (i) Follows easily from the following

\[ 0 \leq \tilde{S}^{-1} \tilde{s}_\alpha = e - \alpha \tilde{S}^{-1}(B\bar{x} - d) \leq e. \]

(ii) We have \( \tilde{S}_\alpha^{-1} \tilde{s} = e + \alpha \tilde{S}_\alpha^{-1}(B\bar{x} - d) \). Therefore,

\[
\begin{align*}
\| \tilde{S}_\alpha^{-1} \tilde{S} \| & \leq 1 + \alpha \| \tilde{S}_\alpha^{-1}(B\bar{x} - d) \|_{\infty} \\
& \leq 1 + \alpha \| \tilde{S}_\alpha^{-1} \tilde{S} \| \| \tilde{S}^{-1}(B\bar{x} - d) \|_{\infty} \\
& \leq 1 + (\alpha/\bar{\alpha}) \| \tilde{S}_\alpha^{-1} \tilde{S} \|
\end{align*}
\]

102
Rearranging, we get
\[ \|\tilde{S}^{-1}\tilde{S}\| \leq \frac{1}{1 - \alpha/\tilde{\alpha}} \leq \frac{1}{1 - \xi}. \]
\[ \square \]

Now we can show the relationship between the two norms \( \| \cdot \|_{Q_0(\tilde{x})} \) and \( \| \cdot \|_{Q_\alpha(\tilde{x})} \).

**Lemma 5.4.7** Under the Assumptions, suppose \( 0 \leq \alpha \leq \xi \tilde{\alpha}, \ 0 \leq \xi < 1 \), where \( \tilde{\alpha} \) is given by (5.4). Let \( Q_\alpha(\tilde{x}) \) be defined by (5.6). Then, for all \( v \in \mathbb{R}^n \),

(i) \[ \|v\|_{Q_\alpha(\tilde{x})} \leq \left( \frac{1 + \xi \sqrt{k}}{1 - \xi} \right) \|v\|_{Q_0(\tilde{x})}; \]

(ii) \[ \|v\|_{Q_\alpha(\tilde{x})} \leq (1 + \xi \sqrt{k})\|v\|_{Q_\alpha(\tilde{x})}. \]

**Proof:** (i) Let \( Q_\alpha = Q_\alpha(\tilde{x}) \). Using Lemma 5.4.6, the triangle inequality, the Cauchy-Schwartz inequality and Lemma 5.4.5, we have

\[ \|v\|_{Q_\alpha} = \|\tilde{S}^{-1}(A + \alpha B)v\| \]
\[ \leq \|\tilde{S}^{-1}\tilde{S}\| \cdot \|\tilde{S}^{-1}(A + \alpha B)v\| \]
\[ \leq \left( \frac{1}{1 - \xi} \right) \left( \|\tilde{S}^{-1}Av\| + \alpha \|\tilde{S}^{-1}Bv\| \right) \]
\[ \leq \left( \frac{1}{1 - \xi} \right) \left( \|v\|_{Q_0} + \alpha \|B^T \tilde{S}^{-1}||_{Q_0^{-1}} \|v\|_{Q_0} \right) \]
\[ \leq \left( \frac{1}{1 - \xi} \right) \left( 1 + \alpha \sqrt{k}/\tilde{\alpha} \right) \|v\|_{Q_0} \]
\[ \leq \left( \frac{1}{1 - \xi} \right) \left( 1 + \xi \sqrt{k} \right) \|v\|_{Q_\alpha}. \]
(ii) Similarly, we have

\[ ||v||_{Q_0} = ||\tilde{S}^{-1}Av|| \]
\[ \leq ||\tilde{S}^{-1}(A + \alpha B)v|| + \alpha||\tilde{S}^{-1}Bv|| \]
\[ \leq ||\tilde{S}^{-1}\tilde{S}_\alpha|| \cdot ||\tilde{S}^{-1}_\alpha(A + \alpha B)v|| + \alpha||B^T\tilde{S}^{-1}||_{Q_0^{-1}} ||v||_{Q_\alpha} \]
\[ \leq (1 + \alpha\sqrt{k}/\tilde{\alpha})||v||_{Q_\alpha} \]
\[ \leq (1 + \xi\sqrt{k})||v||_{Q_\alpha}. \]

As an immediate corollary of Lemma 5.4.7 and Lemma 5.4.3, we have the next lemma.

**Lemma 5.4.8** Under the same conditions and definitions as Lemma 5.4.7, we have for all \( p \in \mathbb{R}^n \),

(i) \[ ||p||_{Q_0^{-1}(x)} \leq (1 + \xi\sqrt{k})||p||_{Q_0^{-1}(\tilde{x})}; \]

(ii) \[ ||p||_{Q_0^{-1}(x)} \leq \left( \frac{1 + \xi\sqrt{k}}{1 - \xi} \right) ||p||_{Q_0^{-1}(\tilde{x})}. \]

**Proof:** Follows immediately from Lemma 5.4.3 and Lemma 5.4.7.

Finally, suppose \( \hat{x} = \hat{x}_0 \) is the center of system \( P(0) \) and suppose \( \tilde{x} \), satisfying \( \tilde{s} = b - A\tilde{x} > 0 \), is a \( \delta \)-approximate center of system \( P(0) \). Let \( \hat{\alpha} = \alpha_{\tilde{x}} \) and \( \tilde{\alpha} = \alpha_{\hat{x}} \) be defined by (5.4). The next lemma shows the relationship between \( \hat{\alpha} \) and \( \tilde{\alpha} \).

**Lemma 5.4.9** Suppose \( \hat{x} = \hat{x}_0 \) is the center of system \( P(0) \). Suppose \( \tilde{x} \), satisfying \( \tilde{s} = b - A\tilde{x} > 0 \), is near the center of system \( P(0) \) in the sense that

\[ ||\tilde{x} - \hat{x}||_{Q_0(\varepsilon)} \leq \delta < 1, \]
where $Q_0(\bar{x})$ is given by (5.5)—(5.6). Let $\hat{\alpha} = \alpha_{\bar{x}}$ and $\bar{\alpha} = \alpha_{\hat{x}}$ be defined by (5.4). Then
\[
\left(\frac{1 - \delta}{1 + \delta}\right) \bar{\alpha} \leq \hat{\alpha} \leq \left(\frac{1 + \delta}{1 - \delta}\right) \bar{\alpha}.
\]

**Proof:** First we note that $||\hat{S}^{-1}\hat{S}|| \leq 1 + \delta$ (from Lemma 3.2.1), and for $i = 1, 2, \ldots, m$,
\[
||\hat{S}^{-1}B(\bar{x} - \hat{x})|| \leq ||(\hat{S}^{-1}B)_i^T||_{Q_0^{-1}(\hat{x})} ||\bar{x} - \hat{x}||_{Q_0(\hat{x})} \leq \frac{1}{\hat{\alpha}} \left(\frac{\delta}{1 - \delta}\right),
\]
where the first inequality is a Cauchy-Schwartz inequality and the last inequality follows from Lemma 5.4.2 and Lemma 3.2.1. Therefore
\[
||\hat{S}^{-1}B(\bar{x} - \hat{x})||_{\infty} \leq \frac{1}{\hat{\alpha}} \left(\frac{\delta}{1 - \delta}\right),
\]
and hence
\[
1/\hat{\alpha} = ||\hat{S}^{-1}(B\bar{x} - d)||_{\infty}
\leq ||\hat{S}^{-1}\hat{S}|| \left(||\hat{S}^{-1}(B\hat{x} - d)||_{\infty} + ||\hat{S}^{-1}B(\bar{x} - \hat{x})||_{\infty}\right)
\leq (1 + \delta) \left(1 + \frac{\delta}{1 - \delta}\right) \frac{1}{\bar{\alpha}}
= \left(\frac{1 + \delta}{1 - \delta}\right) \frac{1}{\bar{\alpha}}.
\]
This proves the first part. The second part is proved in a similar manner.

\[\square\]

### 5.5 Proofs of Main Theorems

In this section, we shall prove Theorem 5.3.1 and Theorem 5.3.2. We shall first show that if the increase in parametric value is not too large, the two
successive centers of the corresponding systems will be close to each other such that Newton's method will work well (i.e., converge quadratically). For simplicity, we shall consider increasing the parameter from $\alpha = 0$ to some $\alpha > 0$.

Suppose $\hat{x} = \hat{x}_0$ is the center of system $P(0)$. That is, we have

\[ \hat{s} = b - A\hat{x} > 0, \quad (5.11) \]

\[ A^T\hat{S}^{-1}e = 0. \quad (5.12) \]

Define $\hat{\alpha} = \alpha_{\hat{x}}$, $\hat{s}_\alpha$ and $Q_\alpha(\hat{x})$ by (5.4)–(5.6). That is,

\[ \hat{\alpha} = \alpha_{\hat{x}} = 1/||\hat{S}^{-1}(B\hat{x} - d)||_{\infty} \quad (5.13) \]

and, for $0 \leq \alpha < \hat{\alpha}$,

\[ \hat{s}_\alpha := (b + \alpha d) - (A + \alpha B)\hat{x} > 0 \quad (5.14) \]

\[ Q_\alpha(\hat{x}) = (A + \alpha B)^T\hat{S}^{-2}_\alpha(A + \alpha B). \quad (5.15) \]

We have the following

**Theorem 5.5.1** Suppose $0 \leq \alpha \leq \frac{\hat{\alpha}}{80(\sqrt{l} + k)}$, where $\hat{\alpha}$ is given by (5.13).

Then $\hat{x}$ is near the center of system $P(\alpha)$, in the sense that

\[ \tau_\alpha(\hat{x}) := ||(A + \alpha B)^T\hat{S}^{-1}_\alpha e||_{Q^{-1}_\alpha(\hat{x})} \leq \frac{1}{78}, \]

where $\hat{s}_\alpha$ and $Q_\alpha(\hat{x})$ are given by (5.14)–(5.15).
**Remark:** Note that \( y \equiv y_\alpha := (A + \alpha B)^T \tilde{S}_\alpha^{-1} e \) is the gradient and \( Q_\alpha(\hat{x}) \) is the negative of the Hessian of the logarithmic barrier function for the center problem \( PCP(\alpha) \),

\[ f_\alpha(x) = \sum_{i=1}^{m} \ln \left[ (b + ad) - (A + \alpha B)x_i \right], \]

evaluated at \( \hat{x} \).

This result also implies (by Lemma 3.2.5) that \( \hat{x} \) is a \( \delta \)-approximate center of system \( P(\alpha) \) with \( \delta = 1/12 \). That is,

**Corollary 5.5.1** Under the same conditions and definitions as Theorem 5.5.1, \( \|\hat{x} - \hat{x}_\alpha\|_{Q_\alpha(\hat{x})} \leq 1/12 \), where \( \hat{x}_\alpha \) denotes the center of system \( P(\alpha) \).

In fact, we shall prove the following slightly more general result.

**Theorem 5.5.2** Suppose \( \bar{x} \), satisfying \( \bar{s} = b - A\bar{x} > 0 \), is near the center of system \( P(0) \) in the sense that

\[ \tau := \sqrt{e^T \tilde{S}^{-1} A[A^T \tilde{S}^{-2} A]^{-1} A^T \tilde{S}^{-1} e} \leq \varepsilon. \]

Let \( \bar{\alpha}, \; \bar{s}_\alpha \) and \( Q_\alpha(\bar{x}) \) be given by (5.4)–(5.6). Suppose \( 0 \leq \alpha \leq \frac{\bar{\alpha}}{80(\sqrt{l} + k)} \). Then \( \bar{x} \) is near the center of system \( P(\alpha) \) in the sense that

\[ \tau_\alpha(\bar{x}) := ||(A + \alpha B)^T \tilde{S}_\alpha^{-1} e||_{Q_\alpha^{-1}(\bar{x})} \leq \varepsilon_1 := \frac{81 \varepsilon}{80} + \frac{1}{78}. \]
Proof: Let $Q_\alpha = Q_\alpha(\bar{x})$. First we note that
\[
A^T \tilde{S}^{-1}e - A^T \tilde{S}^{-1}e = A^T \tilde{S}^{-1}(\bar{s} - \bar{s}_\alpha) = \alpha A^T \tilde{S}^{-1}(B\bar{x} - d).
\]
Therefore, using the triangle inequality, we get
\[
\|A^T \tilde{S}^{-1}e\|_{Q_\alpha^{-1}} \leq \|A^T \tilde{S}^{-1}e - A^T \tilde{S}^{-1}e\|_{Q_\alpha^{-1}} + \|A^T \tilde{S}^{-1}e\|_{Q_\alpha^{-1}} = \alpha \|A^T \tilde{S}^{-1}(B\bar{x} - d)\|_{Q_\alpha^{-1}} + \tau \leq \alpha \|\tilde{S}^{-1}(B\bar{x} - d)\| + \varepsilon \leq \alpha (80/79) \sqrt{l}/\bar{\alpha} + \varepsilon,
\]
where the second inequality follows from the fact that (since $Q_0 = A^T \tilde{S}^{-2}A$)
\[
\|A^T \tilde{S}^{-1}\|_{Q_\alpha^{-1}}^2 = \|\tilde{S}^{-1}A(A^T \tilde{S}^{-2}A)^{-1}A^T \tilde{S}^{-1}\| \leq 1,
\]
because the matrix is a projection matrix, and the last inequality follows from Lemma 5.4.6 (with $\xi = 1/80$) and the fact that if there are $l$ non-zero rows in the matrix $[B, d]$ (i.e. the number of varying constraints is $l$), then
\[
\|\tilde{S}^{-1}(B\bar{x} - d)\| \leq \sqrt{l}\|\tilde{S}^{-1}(B\bar{x} - d)\|_{\infty}.
\]
Hence, by Lemma 5.4.8 (with $\xi = \frac{1}{80(\sqrt{l}+k)}$), we have
\[
\|A^T \tilde{S}^{-1}e\|_{Q_\alpha^{-1}} \leq \frac{81}{80} \left( \frac{80}{79} \frac{\sqrt{l}\alpha}{\bar{\alpha}} + \varepsilon \right) = \frac{81}{79} \frac{\sqrt{l}\alpha}{\bar{\alpha}} + \frac{81}{80} \varepsilon.
\]
On the other hand, we have from Lemma 5.4.2 and Lemma 5.4.6
\[
\|B^T \tilde{S}^{-1}e\|_{Q_\alpha^{-1}} \leq k\|\tilde{S}^{-1}(B\bar{x} - d)\|_{\infty} \leq k\left( \frac{80}{79} \right) \|\tilde{S}^{-1}(B\bar{x} - d)\|_{\infty} = \left( \frac{80}{79} \right) \frac{k}{\bar{\alpha}}.
\]
108
Using the triangle inequality again, we get

\[
\tau_\alpha(\bar{x}) := ||(A + \alpha B)^T \bar{S}_\alpha^{-1} e||_{Q_\alpha^{-1}} \leq ||A^T \bar{S}_\alpha^{-1} e||_{Q_\alpha^{-1}} + \alpha ||B^T \bar{S}_\alpha^{-1} e||_{Q_\alpha^{-1}} \\
\leq \frac{81}{79} \sqrt{I(\alpha/\bar{\alpha})} + \frac{81}{80} \varepsilon + \frac{80}{79} k(\alpha/\bar{\alpha}) \\
\leq \frac{81}{79} (\sqrt{I + k})(\alpha/\bar{\alpha}) + \frac{81}{80} \varepsilon \\
\leq \frac{1}{78} + \frac{81}{80} \varepsilon.
\]

\[\Box\]

**Proof of Theorem 5.5.1:** Since \(A^T \bar{S}_\alpha^{-1} e = 0\), we have \(\tau = 0\) and the result follows immediately from Theorem 5.5.2. \[\Box\]

Next, using Theorem 5.5.1 and Lemma 5.4.9, we have the following

**Theorem 5.5.3** Suppose \(\bar{x}\), satisfying \(\bar{s} = b - A\bar{x} > 0\), is near the center of system \(P(0)\) in the sense that

\[||\bar{x} - \hat{x}||_{Q_0(\bar{e})} \leq 1/21,\]

where \(Q_0(\bar{e})\) is given by (5.5)—(5.6). Let \(\bar{\alpha}\) be defined by (5.4). Suppose \(0 \leq \alpha \leq \frac{\bar{\alpha}}{88(\sqrt{I + k})}\). Then \(\bar{x}\) is near the center of system \(P(\alpha)\) in the sense that

\[||\bar{x} - \hat{x}_\alpha||_{Q_\alpha(\hat{x}_\alpha)} \leq 0.148,\]

where \(Q_\alpha(\hat{x}_\alpha)\) is defined as in (5.5)—(5.6).

**Proof:** First note that by Lemma 5.4.9, \(0 \leq \alpha \leq \frac{\bar{\alpha}}{88(\sqrt{I + k})}\) implies that \(0 \leq \alpha \leq \frac{\hat{\alpha}}{80(\sqrt{I + k})}\). Therefore Theorem 5.5.1 implies that

\[||\hat{x} - \hat{x}_\alpha||_{Q_\alpha(\hat{x})} \leq 1/12\] (5.16)

109
and, by Lemma 3.2.1(iv),

\[ ||\tilde{x} - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq 1/11. \] (5.17)

Also, Lemma 5.4.7 (with \( \xi = \frac{1}{88(\sqrt{1+k})} \leq \frac{1}{88} \)) implies that

\[ ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x})} \leq \left( \frac{88}{87} \right) \left( \frac{89}{88} \right) ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x})} < 0.049. \] (5.18)

and therefore, by Lemma 3.2.1(vi),

\[ ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x})} \leq \frac{0.049}{1 - 0.049} < 0.052. \] (5.19)

Now, Lemma 3.2.1(iv), together with (5.16), implies that

\[ ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x}_\alpha)} \leq \frac{12}{11} ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x})} < 0.057. \] (5.20)

Hence (using the triangle inequality) we have

\[ ||\tilde{x} - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq ||\tilde{x} - \hat{x}||_{Q_\alpha(\tilde{x}_\alpha)} + ||\hat{x} - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} < 0.057 + 1/11 < 0.148. \quad \Box \]

Hence \( \tilde{x} \) is close enough to \( \hat{x}_\alpha \) (the center of system \( P(\alpha) \)) so that we may apply a Newton step to solve for \( \hat{x}_\alpha \) starting from \( \tilde{x} \). The result is given in the following theorem, which is a consequence of Theorem 5.5.3 and Lemma 3.3.1.

**Theorem 5.5.4** Under the same conditions and definitions as Theorem 5.5.3, let \( \tilde{x}_{\text{new}} = \tilde{x} + \eta \), where \( \eta = -Q_\alpha^{-1}(\tilde{x})(A + \alpha B)^T\tilde{S}_\alpha^{-1}e \) is a Newton step from \( \tilde{x} \) in parametric center problem \( \text{PCP}(\alpha) \). Then

\[ ||\tilde{x}_{\text{new}} - \hat{x}_\alpha||_{Q_\alpha(\tilde{x}_\alpha)} \leq 0.034. \]
Proof: We have, from Theorem 5.5.3,
\[
\varepsilon := ||\bar{x} - \hat{x}_\alpha||_{Q_\alpha(\bar{x}_\alpha)} < 0.148. \tag{5.21}
\]
Therefore, by Lemma 3.3.1,
\[
||\bar{x}_{new} - \hat{x}_\alpha||_{Q_\alpha(\bar{x}_\alpha)} \leq \frac{(1 + 0.148)^2(0.148)^2}{1 - 0.148} < 0.034. \quad \Box
\]

Corollary 5.5.2 Under the same definitions and conditions as Theorem 5.5.4,
\[
||\bar{x}_{new} - \hat{x}_\alpha||_{Q_\alpha(\bar{x}_{new})} \leq 1/21.
\]

Proof: From Lemma 3.2.1(vi),
\[
||\bar{x}_{new} - \hat{x}_\alpha||_{Q_\alpha(\bar{x}_{new})} \leq \frac{0.034}{1 - 0.034} < \frac{1}{21}. \quad \Box
\]

Now, Theorem 5.3.1 follows immediately from Corollary 5.5.2. Next, we give the proof of Theorem 5.3.2.

Proof of Theorem 5.3.2:
Using the triangle inequality, and by Theorem 5.5.3 and Theorem 5.5.4,
\[
||\bar{x}_{new} - \bar{x}||_{Q_{\rho}(\bar{x}_\beta)} \leq ||\bar{x}_{new} - \hat{x}_\beta||_{Q_{\rho}(\hat{x}_\beta)} + ||\bar{x} - \hat{x}_\beta||_{Q_{\rho}(\hat{x}_\beta)} \leq 0.034 + 0.148 = 0.182.
\]
Also from Corollary 5.5.1, \(||\hat{x} - \hat{x}_\alpha||_{Q_\alpha(\hat{x})} \leq 1/12\), which in turn implies (by applying Lemma 3.2.1(iv)–(v) to system \(P(\alpha)\)) that
\[
||v||_{Q_\alpha(\hat{x})} \leq \frac{13}{12}||v||_{Q_\alpha(\hat{x}_\alpha)} \tag{5.22}
\]
\[ \|\nu\|_{Q_{\alpha}(\bar{x}_\alpha)} \leq \frac{12}{11} \|\nu\|_{Q_0(\bar{x})}. \quad (5.23) \]

Next, using Lemma 5.4.7 (note that \(\xi \leq 1/80\)), we have
\[ \|\nu\|_{Q_0(\bar{x})} \leq \frac{81}{79} \|\nu\|_{Q_0(\bar{x})} \quad (5.24) \]

and
\[ \|\nu\|_{Q_0(\bar{x})} \leq \frac{81}{80} \|\nu\|_{Q_0(\bar{x})}. \quad (5.25) \]

Finally, observe that \(\bar{x}_\alpha - \bar{x} = (\alpha/\beta)(\bar{x}_{new} - \bar{x})\). Hence, using (5.22)—(5.25), we have
\[
\|\bar{x}_\alpha - \bar{x}\|_{Q_{\alpha}(\bar{x}_\alpha)} \leq \left( \frac{12}{11} \right) \|\bar{x}_\alpha - \bar{x}\|_{Q_{\alpha}(\bar{x})}
\leq \left( \frac{12}{11} \right) \left( \frac{81}{79} \right) \|\bar{x}_\alpha - \bar{x}\|_{Q_0(\bar{x})}
\leq \left( \frac{12}{11} \right) \left( \frac{81}{79} \right) \left( \frac{81}{80} \right) \|\bar{x}_\alpha - \bar{x}\|_{Q_0(\bar{x})}
\leq \left( \frac{12}{11} \right) \left( \frac{81}{79} \right) \left( \frac{81}{80} \right) \left( \frac{13}{12} \right) \|\bar{x}_\alpha - \bar{x}\|_{Q_0(\bar{x}_\alpha)}
\leq \left( \frac{12}{11} \right) \left( \frac{81}{79} \right) \left( \frac{81}{80} \right) \left( \frac{13}{12} \right) \left( \frac{\alpha}{\beta} \right) \|\bar{x}_{new} - \bar{x}\|_{Q_0(\bar{x}_\alpha)}
< 0.23.
\]

Thus,
\[
\|\bar{x}_\alpha - \hat{x}_\alpha\|_{Q_{\alpha}(\bar{x}_\alpha)} \leq \|\bar{x}_\alpha - \bar{x}\|_{Q_{\alpha}(\bar{x}_\alpha)} + \|\bar{x} - \hat{x}_\alpha\|_{Q_{\alpha}(\bar{x}_\alpha)}
< 0.23 + 0.148 < 0.38. \quad \square
\]
Chapter 6

Applications and Conclusion

In this chapter, we give four applications of the results of the previous chapter and present our conclusion. The first application is to the linear programming problem (LP), the second is to the linear fractional programming problem (LFP), the third is to the von Neumann model of economic expansion, and the fourth is to the generalized linear fractional programming problem (GLFP).

6.1 Linear Programming

In this section, we show that Algorithm PCP specializes to the algorithm of Renegar [57], for the linear programming problem (LP), which requires $O(\sqrt{mL})$ iterations to solve an LP, where $m$ is the total number of constraints and $L$ is the number of bits in a binary encoding of the LP.

Suppose we are interested in solving a linear program in the following
format:

\[ \hat{LP} : \quad z^* = \max c^T x \]
\[ \text{s.t. } \hat{A} x \leq \hat{b}, \]

where \( x \in \mathbb{R}^n \) and \( \hat{A} \) is an \( m \times n \) matrix. Non-negativity constraints and lower and upper bounds are not distinguished from other inequalities. We assume that \( c \neq 0 \), for otherwise any feasible solution will be optimal.

Then it is easy to see that \( \hat{LP} \) is equivalent to the following program.

\[ LP : \quad z^* = \max \alpha \]
\[ \text{s.t. } A x \leq b + \alpha d, \]

where \( A, b \) and \( d \) are the following \( 2m \times n \) matrix and \( 2m \)-vectors

\[
A = \begin{bmatrix}
\hat{A} \\
-c^T \\
\vdots \\
-c^T
\end{bmatrix}, \quad b = \begin{bmatrix}
\hat{b} \\
-v \\
\vdots \\
-v
\end{bmatrix}, \quad d = \begin{bmatrix}
0 \\
-1 \\
\vdots \\
-1
\end{bmatrix},
\]

(6.1)

for some lower bound \( v < z^* \) such that the set \( \{ x | Ax < b \} \) is nonempty and bounded. It can be shown that any linear program can be reformulated such that these assumptions are satisfied (see Renegar [57], for example). We may then assume that we have an approximate center \( \bar{x} \) of \( Ax \leq b \) (in the sense that \( \|x - \bar{x}\|_{Q_0(x)} \leq 1/21 \)), perhaps after applying a centering algorithm of Vaidya [65] or Freund [17].

We apply Algorithm PCP to solve \( LP \) in the following way. Note that \( LP \) corresponds to a parametric center problem where the total number of
constraints is $2m$, the number of varying constraints $l = m$, and the matrix $B = 0$ (therefore $k$ of Theorem 5.3.1 is zero). Suppose at the start of iteration $j$ the value of $\alpha$ is $\alpha^j$ and we have a $\delta$-approximate center $\bar{x}^j$ of system $Ax \leq b + \alpha^j d$. We take $\delta = 1/21$. (Recall that $\bar{x}$ is a $\delta$-approximate center of a system $Ax \leq b$ if $||\bar{x} - x||_{Q_0(x)} \leq \delta$. See Subsection 3.2.1.) We observe that for this case, $\check{\alpha}$ according to (5.4) is $\check{\alpha} = c^T \bar{x}^j - \alpha^j - v$. We therefore set $\beta = \frac{c^T \bar{x}^j - \alpha^j - v}{88 \sqrt{m}}$, according to Theorem 5.3.1. Then we change the right-hand-side to $(b + \alpha^j d) + \beta d$. Then we take a Newton iterate from $\bar{x}^j$ in the problem to find the center of system $Ax \leq b + \alpha^{j+1} d$, where $\alpha^{j+1} = \alpha^j + \beta$. Let $\bar{x}^{j+1}$ be the Newton iterate. Then according to Theorems 5.3.1 of the previous chapter, $\bar{x}^{j+1}$ is again a $\delta$-approximate center of system $Ax \leq b + \alpha^{j+1} d$, and we enter iteration $j + 1$ and repeat the procedure. Note that

$$\alpha^{j+1} - \alpha^j = \beta = \frac{1}{88 \sqrt{m}}(c^T \bar{x}^j - \alpha^j - v).$$ (6.2)

On the other hand, using a property of the center (Lemma 3.1.2), we have the following upper bounds.

**Lemma 6.1.1** Suppose $\hat{x}$ is the center of $Ax \leq b$, where $A$, $b$ are given by (6.1). For any $x$ satisfying $Ax \leq b$, we have

$$0 \leq c^T x - v \leq 2(c^T \hat{x} - v).$$

**Proof:** Let $\hat{s} = b - A\hat{x}$ and $s = b - Ax$. Then from Lemma 3.1.2, we have

$$2m = c^T \hat{s}^{-1}s = m \left( \frac{c^T x - v}{c^T \hat{x} - v} \right) + \sum_{i=1}^{m} \frac{b_i - \hat{A}_i x}{b_i - \hat{A}_i \hat{x}} \geq m \left( \frac{c^T x - v}{c^T \hat{x} - v} \right).$$

The Lemma follows immediately.  $$\square$$

115
Corollary 6.1.1 Suppose \( \bar{x} \) is a \( \delta \)-approximate center of \( Ax \leq b \), with \( \delta = 1/21 \). For all \( x \) satisfying \( Ax \leq b \), we have

\[
0 \leq c^T x - v \leq (2.2)(c^T\bar{x} - v).
\]

Proof: Suppose \( \hat{x} \) is the center of \( Ax \leq b \). From Lemma 5.4.9, we see that

\[
(0.9)(c^T\bar{x} - v) \leq c^T\hat{x} - v \leq (1.1)(c^T\bar{x} - v).
\]

The Corollary then follows immediately from the previous Lemma.

Therefore (replacing \( b \) with \( b + \alpha^j d \) and \( \bar{x} \) with \( \bar{x}^j \) in the above Corollary),

\[
0 \leq z^* - \alpha^j - v \leq (2.2)(c^T\bar{x}^j - \alpha^j - v). \tag{6.3}
\]

Combining (6.2) and (6.3), we have

\[
\alpha^{j+1} - \alpha^j \geq \frac{1}{194\sqrt{m}}(z^* - \alpha^j - v). \tag{6.4}
\]

Rearranging terms, we get

\[
z^* - \alpha^{j+1} - v \leq \left(1 - \frac{1}{194\sqrt{m}}\right)(z^* - \alpha^j - v). \tag{6.5}
\]

Hence, at each iteration, the gap \((z^* - \alpha - v)\) decreases by at least a factor of \((1 - \frac{1}{194\sqrt{m}})\). Therefore, we can show, as in [57], that Algorithm PCP requires \(O(\sqrt{m}L)\) iterations to solve LP. Summarizing the discussion of this section, we have

Theorem 6.1.1 Algorithm PCP, if properly initiated, solves the LP problem in \(O(\sqrt{m}L)\) iterations, where \( m \) is the number of constraints and \( L \) is the number of bits in a binary encoding of the problem instance.
6.2 Linear Fractional Programming

In this section, we show that Algorithm PCP can be used to solve the linear fractional programming problem (LFP), and analyze its complexity as an algorithm for LFP.

Suppose we are interested in solving a linear fractional program in $\mathbb{R}^n$ in the following format

\[ LFP : \quad \alpha^* = \max_{x} \frac{f - \tilde{c}^T x}{d^T x - h} \]
\[ s. t. \quad \tilde{A}x \leq \tilde{b}, \]

where $x \in \mathbb{R}^n$, $\tilde{A}$ is an $m \times n$ matrix, $\tilde{b}$ is an $m$-vector, $\tilde{c}$ and $\tilde{d}$ are $n$-vectors and $f$ and $h$ are scalars. We assume that

(1) the set $\mathcal{X}_0^+ := \{x | \tilde{A}x < \tilde{b}, \tilde{c}^T x < f\}$ is nonempty and bounded,

(2) $\tilde{d}^T x - h \geq 0$ for every $x \in \mathcal{X} := \{x | \tilde{A}x \leq \tilde{b}\}$, and

(3) $\tilde{d}^T x - h > 0$ for every $x \in \mathcal{X}_0 := \{x | \tilde{A}x \leq \tilde{b}, \tilde{c}^T x \leq f\}$.

Note that under these assumptions, $0 < \alpha^* < \infty$, and if $\tilde{d} = 0$ then $LFP$ is just a linear program, so we may assume that $\tilde{d} \neq 0$. (We note that Anstreicher [1] has a similar set of assumptions.)

It is easy to see that $LFP$ is equivalent to the following program in $\mathbb{R}^{n+1}$

\[ LFP : \quad \alpha^* = \max_{x, \alpha} \alpha \]
\[ s. t. \quad \tilde{A}x \leq \tilde{b} \]
\[ (\tilde{c} + \alpha \tilde{d})^T x \leq f + \alpha h. \]

We can then apply Algorithm PCP to solve $LFP$ by tracing the centers of the parametric family of systems

\[ P(\alpha) : \quad (A + \alpha B)x \leq b + \alpha d, \]

117
where $\alpha$ is taken as a parameter and is increased strictly monotonically, and

$$A = \begin{bmatrix} \tilde{A} \\ \tilde{c} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \tilde{d} \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{b} \\ f \end{bmatrix}, \quad d = \begin{bmatrix} 0 \\ h \end{bmatrix}. $$

Note that the total number of constraints in system $P(\alpha)$ is $m + 1$, and the number of varying constraints is $l = 1$, since $[\tilde{d}, h] \neq 0$ by our assumptions, and the number of nonzero rows in the matrix $B$ is $k = 1$. We can assume that we have an interior point $\bar{x}^0$ satisfying the starting criterion of Algorithm PCP, that is, $\bar{x}^0$ is a $\delta$-approximate center of the system $P(0)$ with $\delta = 1/21$, perhaps by using an algorithm (of Vaidya [65] or Freund [17]) for finding the center of the system $P(0)$.

When we apply Algorithm PCP to trace the center of $P(\alpha)$ as $\alpha$ increases strictly monotonically over the range $\alpha \in [0, \alpha^* - \varepsilon]$, we have the following. At the start of iteration $j$, the value of $\alpha$ is $\alpha^j$ and the current iterate is $\bar{x} = \bar{x}^j$, which is a $\delta$-approximate center of the system $P(\alpha^j)$ with $\delta = 1/21$. We observe that $\bar{\alpha}$, as defined by (5.4), is

$$\bar{\alpha} = \frac{(f + \alpha^j h) - (\tilde{c} + \alpha^j \tilde{d})^T \bar{x}^j}{\tilde{d}^T \bar{x}^j - h},$$

since the only varying constraint is the last constraint, and in the $j^{th}$ iteration, the last constraint is

$$(\tilde{c} + \alpha^j \tilde{d})^T x \leq f + \alpha^j h.$$ 

Therefore, $\beta$ is set, in accordance with Theorem 5.3.1, to be

$$\beta = \frac{1}{176} \bar{\alpha} = \frac{(f + \alpha^j h) - (\tilde{c} + \alpha^j \tilde{d})^T \bar{x}^j}{176(\tilde{d}^T \bar{x}^j - h)}.$$ 

Next $\alpha^{j+1}$ is set to be $\alpha^{j+1} = \alpha^j + \beta$, and a Newton iterate from $\bar{x}^j$ is taken (in the program to find the center of the system $P(\alpha^{j+1})$). The next iterate $\bar{x}^{j+1}$
is set to be the Newton iterate from $\bar{x}^j$. Then according to Theorem 5.3.1, $\bar{x}^{j+1}$ is again a $\delta$-approximate center of the system $P(\alpha^{j+1})$ with $\delta = 1/21$, and we enter iteration $j + 1$. Note that

$$\tilde{\alpha} = \frac{(f + \alpha^j h) - (\bar{c} + \alpha^j \bar{d})^T \bar{x}^j}{d^T \bar{x}^j - h} = \frac{f - \bar{c}^T \bar{x}^j}{d^T \bar{x}^j - h} + \alpha^j > 0. \quad (6.6)$$

Therefore, $\alpha^j$ is a strict lower bound on the objective value of LFP at $\bar{x}^j$.

Also,

$$\alpha^{j+1} - \alpha^j = \beta = \frac{1}{176} \tilde{\alpha}. \quad (6.7)$$

On the other hand, using a property of the center of a linear inequality system (Lemma 3.1.2), we can derive an upper bound on $\alpha^*$ in term of $\tilde{\alpha}$ and $m$ as follows. First, by Lemma 3.1.2, (note that the total number of constraints in the system $P(\alpha)$ is $m + 1$) we have

**Lemma 6.2.1** Suppose $\hat{x}_\alpha$ is the center of the system $P(\alpha)$. Then, for all $x \in \mathcal{X}_\alpha := \{x \mid (A + \alpha B)x \leq b + \alpha d\}$,

$$0 \leq (f + \alpha h) - (\bar{c} + \alpha \bar{d})^T x \leq (m + 1)[(f + \alpha h) - (\bar{c} + \alpha \bar{d})^T \hat{x}_\alpha].$$

Next, by Lemma 3.2.1 and the above Lemma, we have

**Lemma 6.2.2** Suppose $\bar{x}$ is a $\delta$-approximate center of the system $P(\alpha)$. Then, for all $x \in \mathcal{X}_\alpha := \{x \mid (A + \alpha B)x \leq b + \alpha d\}$,

$$0 \leq (f + \alpha h) - (\bar{c} + \alpha \bar{d})^T x \leq (1 + \delta)(m + 1)[(f + \alpha h) - (\bar{c} + \alpha \bar{d})^T \bar{x}].$$
Let $u$ and $v$ be the following constants.

$$u := \max \{ \tilde{d}^T x - h \mid \tilde{A} x \leq \tilde{b} \}$$

$$v := \min \{ \tilde{d}^T x - h \mid \tilde{A} x \leq \tilde{b}, \tilde{c}^T x \leq f \}$$

For this section, define the constant $\bar{c}$ by

$$\bar{c} := \left( \frac{22}{21} \right) \left( \frac{u}{v} \right). \tag{6.8}$$

Under the assumptions in this section, for all $x$ satisfying $Ax \leq b$, we have

$$0 < v \leq \tilde{d}^T x - h \leq u < \infty.$$ 

Therefore, we see that $0 < \bar{c} < \infty$. We have

**Lemma 6.2.3** Let $\bar{\alpha}$ be given by (6.6), and let $\bar{c}$ be the constant defined by (6.8). Then $\alpha^* - \alpha^i \leq \bar{c}(m + 1)\bar{\alpha}$.

**Proof:** Note that $\bar{x} = \bar{x}^i$ is a $\delta$-approximate center of system $P(\alpha^i)$ with $\delta = 1/21$. For any $x$ satisfying $Ax \leq b$,

$$\frac{f - \bar{c}^T x}{\tilde{d}^T x - h} - \alpha^i = \frac{(f + \alpha^i h) - (\bar{c} + \alpha^i \tilde{d})^T \bar{x}}{\tilde{d}^T x - h}$$

$$\leq \frac{22}{21}(m + 1) \left( \frac{f + \alpha^i h}{\tilde{d}^T x - h} - (\bar{c} + \alpha^i \tilde{d})^T \bar{x} \right)$$

$$= \frac{22}{21}(m + 1) \left( \frac{\tilde{d}^T \bar{x} - h}{\tilde{d}^T x - h} \right) \left( \frac{f + \alpha^i h}{\tilde{d}^T x - h} - (\bar{c} + \alpha^i \tilde{d})^T \bar{x} \right)$$

$$\leq \frac{22}{21}(m + 1) \left( \frac{u}{v} \right) \left( \frac{f + \alpha^i h}{\tilde{d}^T \bar{x} - h} - (\bar{c} + \alpha^i \tilde{d})^T \bar{x} \right)$$

$$= \bar{c}(m + 1)\bar{\alpha},$$

120
where the first inequality follows from Lemma 6.2.2. Also,
\[
\alpha^* := \max \left\{ \frac{f - \bar{c}^T x}{d^T x - h} \mid \bar{A} x \leq \bar{b} \right\} = \max \left\{ \frac{f - \bar{c}^T x}{d^T x - h} \mid \bar{A} x \leq \bar{b}, \bar{c}^T x \leq f \right\} = \max \left\{ \frac{f - \bar{c}^T x}{d^T x - h} \mid A x \leq b \right\}
\]
Hence, \( \alpha^* - \alpha^j \leq \bar{c}(m + 1)\bar{\alpha} \). \( \square \)

Lemma 6.2.3 and (6.7) implies that
\[
\alpha^{j+1} - \alpha^j \geq \left( \frac{1}{176\bar{c}(m + 1)} \right)(\alpha^* - \alpha^j).
\]  
(6.9)

Rearranging terms, we get
\[
\alpha^* - \alpha^{j+1} \leq (1 - \frac{1}{176\bar{c}(m + 1)}) (\alpha^* - \alpha^j).
\]  
(6.10)

Therefore, the gap \((\alpha^* - \alpha)\) decreases geometrically with a rate of at least 
\((1 - O(\frac{1}{m+1})) = (1 - O(\frac{1}{m}))\). Hence, we can show that the algorithm requires 
\(O(m)\) iterations to decrease the optimality gap \((\alpha^* - \alpha)\) by a fixed quantity. 
Summarizing the discussion in this section, we have

**Theorem 6.2.1** Suppose Algorithm PCP of Chapter 5 is applied to solve 
LFP. Then it produces a solution \(\bar{x}\) such that
\[
\alpha^* - \varepsilon \leq \frac{f - \bar{c}^T \bar{x}}{d^T \bar{x} - h} \leq \alpha^*
\]

after at most \(K = [176\bar{c}(m+1)(\ln \alpha^* - \ln \varepsilon)]\) iterations, where \(\bar{c}\) is a constant 
defined by (6.8).

**Proof:** Let \(K\) be as defined in the Theorem. From (6.10), we get
\[
\alpha^* - \alpha^j \leq (1 - \frac{1}{176\bar{c}(m + 1)})^j \alpha^*.
\]

121
Therefore, for all \( j \geq K \),

\[
\ln(\alpha^* - \alpha^j) \leq j \ln(1 - \frac{1}{176\bar{c}(m+1)}) + \ln \alpha^* \\
\leq \frac{-j}{176\bar{c}(m+1)} + \ln \alpha^* \\
\leq -(\ln \alpha^* - \ln \varepsilon) + \ln \alpha^*,
\]

whereby \( \alpha^* - \alpha^j \leq \varepsilon \). Hence, \( \alpha^* - \varepsilon \leq \alpha^j < \frac{-\varepsilon T \tilde{x}^j}{d^T \tilde{x}^j - h} \), from (6.6). \( \square \)

6.3 The von Neumann Model of Economic Expansion

In 1932, the mathematician John von Neumann developed a linear model of an expanding economy which was published in 1937, and an English translation was published in 1945 [55]. The model involves \( n \) productive processes \( P_1, P_2, \ldots, P_n \) producing \( m \) economic goods \( G_1, G_2, \ldots, G_m \). At unit intensity of operation, each process \( P_j \) will consume an amount \( a_{ij} \geq 0 \) and produce an amount \( b_{ij} \geq 0 \) of each good \( G_i \). The nonnegative \( m \times n \) matrices \( A = [a_{ij}] \) and \( B = [b_{ij}] \) are respectively called the input and output matrices of the model. (See Gale[24].)

Suppose each process \( P_j \) is operated at an intensity \( x_j \geq 0 \), and let the vector \( x := (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}_+^n \) denote the intensity vector. Then the components of the vector \( Ax \) give the amounts of inputs used up in production, and the components of the vector \( Bx \) give the amounts of outputs produced, during a unit time period. The model with input matrix \( A \) and output matrix \( B \) is referred to symbolically as \((A, B)\).
Given an intensity vector $x$, let $\alpha_x$ be defined by

$$\alpha_x := \max \{ \alpha \mid Bx \geq \alpha Ax \}$$

$$= \min \left\{ \frac{B_i x}{A_i x} \mid A_i x > 0 \right\},$$

where $A_i$ denotes the i-th row of $A$. Then $\alpha_x$ represents the expansion factor of the economy operating at intensity $x$. Thus the output of each good $G_i$ is at least $\alpha_x$ times as great as its input. The technological expansion problem (TEP) for an economic model $(A, B)$ is to find an intensity vector $x$ such that $\alpha_x$ is maximal. That is, we may write this as a nonlinear program:

$$\text{TEP} : \quad \alpha^\text{max} := \max \alpha$$

subject to \quad \begin{align*}
(B - \alpha A)x & \geq 0 \\
x & \geq 0, \quad x \neq 0.
\end{align*}$$

Associated with the technological expansion problem is a dual problem called the economic expansion problem, which is

$$\text{EEP} : \quad \beta^\text{min} := \min \beta$$

subject to \quad \begin{align*}
(B - \beta A)^T y & \leq 0 \\
y & \geq 0, \quad y \neq 0.
\end{align*}$$

Here $y \in \mathbb{R}_+^m$ is a vector of prices, and $B^T y$ and $A^T y$ are the corresponding revenue and cost vectors. For a given price vector $y$,

$$\beta_y := \max_j \left\{ \frac{B_j^T y}{A_j^T y} \right\}$$

represents the interest factor at prices $y$. (For an interpretation, see Gale[24].)
Observe that both problems are homogeneous, so that if \( \tilde{x} \) or \( \tilde{y} \) is optimal then also so is any positive multiple of either. Therefore, \( TEP \) and \( EEP \) may be expressed equivalently as the following. (Remember that \( e \) denotes the vector of all ones of the appropriate dimension.)

\[
TEP : \quad \alpha^{\text{max}} := \text{maximize} \quad \alpha \\
\text{subject to} \quad (B - \alpha A)x \geq 0 \\
\quad \quad \quad \quad e^T x = 1, \ x \geq 0.
\]

\[
EEP : \quad \beta^{\text{min}} := \text{minimize} \quad \beta \\
\text{subject to} \quad (B - \beta A)^T y \leq 0 \\
\quad \quad \quad \quad e^T y = 1, \ y \geq 0.
\]

It is clear that in order for the model to correspond to economic reality, some conditions must be imposed on the input and output coefficients. Consider the following uninteresting and unrealistic economies. (See Kemeny et. al. [41].)

**Example 1:** Let \( a_{ij} = 1 \) and \( b_{ij} = 0 \) for all \( i \) and \( j \). Then \( \alpha^{\text{max}} = \beta^{\text{min}} = 0 \) and any intensity vector \( x \) and price vector \( y \) is optimal. This is an economy which consumes inputs but produces nothing.

**Example 2:** Let \( a_{ij} = 0 \) and \( b_{ij} = 1 \) for all \( i \) and \( j \). Then \( \alpha^{\text{max}} = \beta^{\text{min}} = \infty \) and any intensity vector \( x \) and price vector \( y \) is optimal. This is an economy which produces goods without using any input.
Therefore, the following conditions, which eliminate the above cases, are assumed. (See Gale [24], Kemeny et. al. [41].)

**Assumption 1:** For every \( i \), there exists some \( j \) such that \( b_{ij} > 0 \).
That is, every good \( G_i \) is produced by some process \( P_j \).

**Assumption 2:** For every \( j \), there exists some \( i \) such that \( a_{ij} > 0 \).
That is, every productive process \( P_j \) consumes some input \( G_i \).

Under these assumptions, the problem becomes very structured. For example, it is easy to see that the set of feasible intensity vectors

\[
\mathcal{X}_\alpha := \{ x \in \mathbb{R}^n_+ \mid (B - \alpha A)x \geq 0, \ e^Tx = 1 \}
\]

is shrinking as the expansion factor \( \alpha \) increases. That is, \( \mathcal{X}_{\alpha_1} \subset \mathcal{X}_{\alpha_2} \) whenever \( \alpha_1 > \alpha_2 \). The following are some of the known results on the von Neumann model, and can be found in, for example, Gale[24].

**Theorem 6.3.1 (Existence Theorem (see [24]))** For models satisfying Assumptions 1 and 2, \( \alpha_{\text{max}} \) exists and is positive. (i.e. \( 0 < \alpha_{\text{max}} < \infty \)).

**Lemma 6.3.1** For models satisfying Assumptions 1 and 2, \( \beta_{\text{min}} \leq \alpha_{\text{max}} \).

The equality may not hold (see example in [24]).

**Theorem 6.3.2 (von Neumann [55])** Suppose model \( (A, B) \) satisfies A-
sumptions 1 and 2, then there exists \( \hat{x} \in \mathbb{R}^n_+ \), \( \hat{y} \in \mathbb{R}^m_+ \) and \( \gamma \in \mathbb{R} \) such that

\[
B_\hat{x} \geq \gamma A_\hat{x} \quad \text{(i)}
\]

and

\[
\text{if } B_i \hat{x} > \gamma A_i \hat{x} \text{ then } \hat{y}_i = 0; \quad \text{(ii)}
\]

\[
B^T \hat{y} \leq \gamma A^T \hat{y} \quad \text{(iii)}
\]

and

\[
\text{if } B^T_j \hat{y} < \gamma A^T_j \hat{y} \text{ then } \hat{x}_j = 0. \quad \text{(iv)}
\]

**Remark:** Although the pair of dual problems here seem very similar to dual linear programming problems, there is a fundamental difference: the constraints are not linear. In fact, even if all the coefficients are integers, the optimal expansion rate may be irrational (see example in [24]).

**Definition:** An economic model \((A, B)\) is said to be **reducible** if there exists a set of goods \( S \subset \{1, \ldots, m\} \), \( S \neq \{1, \ldots, m\} \) and a set of productive processes \( T \subset \{1, \ldots, n\} \) such that

1. for any \( i \not\in S \), \( a_{ij} = 0 \) for all \( j \in T \); and

2. for every \( i \in S \), there exists \( j \in T \) such that \( b_{ij} > 0 \).

That is, a model is reducible if there exists a proper subset of goods \( S \) that can be produced from themselves without consuming any good not in \( S \). A model is **irreducible** if it is not reducible.

**Theorem 6.3.3 (Duality Theorem (see [24]))** If an economic model \((A, B)\) satisfying Assumptions 1 and 2 is irreducible, then \( \beta^{\text{min}} = \alpha^{\text{max}} \).
6.4 Solution of the von Neumann Model by a Method of Parametric Centers

Here we describe a method for solving the von Neumann model using parametric centers. Suppose we are interested in solving the following technological expansion problem for an economic model \((A, B)\).

\[
\text{TEP: } \quad \alpha^{\max} := \maximize \alpha \\
\text{subject to} \quad (\tilde{B} - \alpha \tilde{A})\tilde{x} \geq 0 \\
e^T\tilde{x} = 1 \\
\tilde{x} \geq 0,
\]

where \(A\) and \(B\) are non-negative \(m \times (n+1)\) matrices, and \(\tilde{x} \in \mathbb{R}^{n+1}\). Without loss of generality, we assume that \(A = [\bar{A}, \bar{A}_{n+1}], B = [\bar{B}, \bar{B}_{n+1}]\), where \(\bar{A}_{n+1}\) and \(\bar{B}_{n+1}\) are column vectors, and \(\tilde{x} = [x^T, \tilde{x}_{n+1}]^T\), with \(\bar{A}, \bar{B} \in \mathbb{R}_+^{m \times n}\), \(\bar{A}_{n+1}, \bar{B}_{n+1} \in \mathbb{R}_+^{m}\) and \(x \in \mathbb{R}^{n}\). Then, by using the constraint \(e^T\tilde{x} = 1\) to eliminate the \((n + 1)^{st}\) variable \(\tilde{x}_{n+1}\), we see that, for each \(\alpha\), the following two linear systems are equivalent in the sense that there exists \(\tilde{x} \in \mathbb{R}^{n+1}\) satisfying system (I)

\[
(\tilde{B} - \alpha \tilde{A})\tilde{x} \geq 0 \\
\quad (I) \\
e^T\tilde{x} = 1 \\
\tilde{x} \geq 0
\]

if and only if there exists \(x \in \mathbb{R}^{n}\) satisfying system (II)
\[(\tilde{B}_{n+1}e^T - \tilde{B}) + \alpha(\tilde{A} - \tilde{A}_{n+1}e^T)x \leq \tilde{B}_{n+1} + \alpha(-\tilde{A}_{n+1})\]

\[(II) \quad e^T x \leq 1 \quad -x \leq 0,\]

where the obvious transformation is

\[x^T = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)^T \text{ and } \tilde{x}^T = (x_1, x_2, \ldots, x_n, 1 - e^T x)^T.\]

Now, system (II) may be expressed as

\[P(\alpha): \quad (A + \alpha B)x \leq b + \alpha d,\]

where

\[A = \begin{bmatrix} \tilde{B}_{n+1}e^T - \tilde{B} \\ e^T \\ -I \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{B}_{n+1} \\ 1 \\ 0 \end{bmatrix}, \quad (6.11)\]

\[B = \begin{bmatrix} \tilde{A} - \tilde{A}_{n+1}e^T \\ 0 \\ O \end{bmatrix}, \quad d = \begin{bmatrix} -\tilde{A}_{n+1} \\ 0 \\ 0 \end{bmatrix}. \quad (6.12)\]

Hence, \(\overline{TEP}\) is equivalent to the following problem.

\[\overline{TEP}: \quad \alpha^{\text{max}} := \text{maximize} \quad \alpha \quad \text{subject to} \quad (A + \alpha B)x \leq b + \alpha d,\]

where \(A, B \in \mathbb{R}^{(m+n+1) \times n}\) and \(b, d \in \mathbb{R}^{m+n+1}\) are given by (6.11)-(6.12).

Observe that the set \(\{x | Ax \leq b\}\) is bounded because of the existence of the constraints \(e^T x \leq 1, \ x \geq 0\). Also, it is straightforward to verify that
\[ x = \left( \frac{1}{n+1} \right) e \in \{ x | Ax < b \}. \] Therefore, \( \{ x | Ax < b \} \) is nonempty and bounded. Next, we see that the system

\[ Ax \leq b \]

is equivalent to the system

\[ \tilde{B}\tilde{x} \geq 0, \ e^T \tilde{x} = 1, \ \tilde{x} \geq 0, \]

and for \( \tilde{x} \) satisfying \( e^T \tilde{x} = 1, \ \tilde{x} \geq 0 \), the system

\[ \tilde{A}\tilde{x} \geq 0, \ \tilde{A}\tilde{x} \neq 0 \]

is equivalent to the system

\[ Bx \geq d, \ Bx \neq d. \]

Therefore, if the model \( (\tilde{A}, \tilde{B}) \) satisfies Assumptions 1 and 2 of Section 6.3, then the parametric family of linear inequality systems \( P(\alpha) \) also satisfies the Assumptions of Chapter 5.

Hence, we may apply Algorithm PCP of Chapter 5 to problem \( \overline{TEP} \) by tracing the center of system \( P(\alpha) \) as \( \alpha \) is increased strictly monotonically. Starting from \( \left( \frac{1}{n+1} \right) e \), we use a center finding algorithm of Vaidya [65] or Freund [17] to get an approximate center \( \bar{x} \) (satisfying \( \|\bar{x} - \hat{x}\|_{Q_0(\bar{x})} \leq 1/21 \)) for the system \( Ax \leq b \). Next, we apply Algorithm PCP with \( \alpha_{\text{upper}} = \alpha_{\text{max}} - \epsilon \), where \( \epsilon \) is the specified error tolerance. We have the following result concerning the efficiency of Algorithm PCP when applied to the problem \( \overline{TEP} \).

Let \( \alpha^j \) be the value of \( \alpha \) and let \( \bar{x}^j \) be the iterate in iteration \( j \). Observe that the total number of constraints in the problem \( \overline{TEP} \) is \( (m + n + 1) \).
and the numbers of non-zero rows in the matrices $B$ and $[B, d]$ are both $m$, whereby we can set $k = l = m$. From Theorem 5.3.1,

$$\alpha^{j+1} - \alpha^j = \frac{1}{88(\sqrt{m} + m)} \tilde{\alpha} \geq \frac{1}{176m} \tilde{\alpha}, \quad (6.13)$$

where $\tilde{\alpha}$ is defined by (5.4), and from Corollary 5.3.1,

$$\alpha^{\max} \quad \alpha^j \leq \tilde{c}(m + n + 1) \tilde{\alpha}, \quad (6.14)$$

for some constant $\tilde{c} < \infty$ defined by (5.10). Combining (6.13) and (6.14),

$$\alpha^{j+1} - \alpha^j \geq \frac{1}{176\tilde{c}(m + n + 1)m} (\alpha^{\max} - \alpha^j). \quad (6.15)$$

Rearranging terms, we see that the optimality gap $(\alpha^{\max} - \alpha^j)$ at each iteration $j$ of Algorithm PCP satisfies

$$\alpha^{\max} - \alpha^{j+1} \leq (1 - \frac{1}{176\tilde{c}(m + n + 1)m})(\alpha^{\max} - \alpha^j). \quad (6.16)$$

Therefore, we have the following.

**Theorem 6.4.1** Suppose Algorithm PCP of Chapter 5 is applied to solve problem $T\overline{EP}$. Then it will produce an intensity vector $\tilde{x}$ and an expansion factor $\tilde{\alpha}$ such that

$$\alpha^{\max} - \epsilon \leq \tilde{\alpha} \leq \alpha^{\max}$$

after at most $K = [176\tilde{c}(m + n + 1)m(\ln \alpha^{\max} - \ln \epsilon)]$ iterations, where $\tilde{c}$ is a constant defined by (5.10).

**Proof:** From (6.16), we get

$$\alpha^{\max} - \alpha^j \leq (1 - \frac{1}{176\tilde{c}(m + n + 1)m})^j \alpha^{\max}.$$
Let $K$ be as defined in the Theorem. Then for all $j \geq K$,

$$
\ln(\alpha^{\text{max}} - \alpha^j) \leq j \ln(1 - \frac{1}{176\tilde{c}(m + n + 1)m}) + \ln \alpha^{\text{max}}
$$

$$
\leq \frac{-j}{176\tilde{c}(m + n + 1)m} + \ln \alpha^{\text{max}}
$$

$$
\leq -(\ln \alpha^{\text{max}} - \ln \varepsilon) + \ln \alpha^{\text{max}}
$$

$$
= \ln \varepsilon,
$$

whereby $\alpha^{\text{max}} - \alpha^j \leq \varepsilon$. Hence, $\alpha^{\text{max}} - \varepsilon \leq \alpha^j \leq \alpha^{\text{max}}$. \hfill \Box

### 6.5 Generalized Linear Fractional Programming

We have already seen in Chapter 5 that the generalized linear fractional programming problem (GLFP) is very closely related to the multiple constraints parametric center problem (PCP). We shall analyze the complexity of Algorithm PCP as an algorithm for GLFP in this section.

Suppose we are interested in solving the following GLFP program.

$$
\text{GLFP} : \quad \alpha^* = \max_x \min_i \left\{ \frac{f_i - C_i x}{D_i x - h_i} \right\}
$$

subject to \quad \tilde{A}x \leq \tilde{b},

where $\tilde{A} \in \mathbb{R}^{m \times mn}$, $\tilde{b} \in \mathbb{R}^m$, $C$, $D \in \mathbb{R}^{k \times n}$ and $f$, $h \in \mathbb{R}^k$, and $[C_i, f_i] \neq 0$ and $[D_i, h_i] \neq 0$ for $i = 1, 2, \ldots, k$. Let

$$
A = \begin{bmatrix} \tilde{A} \\ C \end{bmatrix}, \quad B = \begin{bmatrix} O \\ D \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{b} \\ f \end{bmatrix}, \quad d = \begin{bmatrix} 0 \\ d \end{bmatrix}.
$$

131
Then it is easy to see that $GLFP$ is equivalent to the following program.

$$GLFP: \quad \alpha^* = \max_{x,\alpha} \alpha$$

$$s. t. \quad (A + \alpha B)x \leq b + \alpha d,$$

where $A, B \in \mathbb{R}^{(m+k) \times n}$ and $b, d \in \mathbb{R}^{m+k}$.

We can then apply Algorithm PCP to solve $GLFP$ by tracing the parameteric centers of the family of systems

$$P(\alpha): \quad (A + \alpha B)x \leq b + \alpha d,$$

as $\alpha$, taken as a parameter, is increased strictly monotonically over the range $\alpha \in [0, \alpha^* - \varepsilon]$, where $\varepsilon$ is the given error tolerance. We note that the total number of constraints in system $P(\alpha)$ is $(m + k)$ and the number of varying constraints is $k$.

Suppose $\alpha^j$ is the value of $\alpha$ at the start of iteration $j$. From Theorem 5.3.1 we get

$$\alpha^{j+1} - \alpha^j = \left( \frac{1}{88(\sqrt{k} + k)} \right) \bar{\alpha}, \quad (6.17)$$

where $\bar{\alpha}$ is defined by (5.4). Also from Corollary 5.3.1 we get

$$\alpha^* - \alpha^j \leq \bar{c}(m + k)\bar{\alpha}, \quad (6.18)$$

for some constant $\bar{c}$ defined by (5.10). Therefore, combining (6.17) and (6.18), we get

$$\alpha^{j+1} - \alpha^j \geq \frac{1}{88\bar{c}(m + k)(\sqrt{k} + k)}(\alpha^* - \alpha^j) \geq \frac{1}{176\bar{c}(m + k)}(\alpha^* - \alpha^j), \quad (5.19)$$

since $\sqrt{k} + k \leq 2k$. Rearranging terms, we get

$$\alpha^* - \alpha^{j+1} \leq (1 - \frac{1}{176\bar{c}(m + k)\bar{\alpha}})(\alpha^* - \alpha^j). \quad (6.20)$$

Therefore, we can show as in the previous section,
Theorem 6.5.1 Suppose Algorithm PCP of Chapter 5 is applied to solve GLFP. Then it produces a solution \( \tilde{x} \) such that

\[
\alpha^* - \varepsilon \leq \min_i \left\{ \frac{f_i - C_i \tilde{x}}{D_i \tilde{x} - h_i} \right\} \leq \alpha^*
\]

after at most \( K = [176 \bar{c}(m + k)k(\ln \alpha^* - \ln \varepsilon)] \) iterations, where \( \bar{c} \) is a constant defined by (5.10).

6.6 Conclusion

In this thesis, we have aimed to gain a better understanding and a deeper insight into the behavior of interior point algorithms through the study of parametric centers, as a system of linear inequalities is parametrically deformed. We have seen that the problem context for parametric centers is quite general and goes considerably beyond the context of interior point algorithms for the linear programming problem.

We have achieved the aim of analyzing the efficiency of Newton's method for the parametric center problems, of changing the right-hand-side and of changing simultaneously multiple constraints. We have thus shown that Newton's method is an efficient method for solving problems that are suitable for the method of centers (see Huard [36]), such as the linear programming problem, the linear fractional programming problem, the problem of finding the optimal expansion factor in the von Neumann model of economic expansion and the generalized linear fractional programming problem.

In using Newton's method for the parametric center problems, we learned that the choice of an appropriate measure of closeness (i.e., the definition of
an approximate center) is crucial in the theoretical development of efficient algorithms for the parametric center problems and other related problems.

In conclusion, this thesis contributes to the knowledge, and provides a better insight into the behavior, of interior point algorithms, in a more general problem context than that of the linear programming problem, through the study of the path of parametric centers. It also documents the efficiency of Newton's method for a class of problems solvable by the method of centers. We have considered only linear systems in this thesis. However, it seems very likely that the approach taken here can be extended to other more general systems, such as convex quadratic systems, where the concept of analytic centers can be defined as for linear systems. This, with applications to the linear complementarity problem, will undoubtedly be an area of future research.
This page is intentionally left blank.
Bibliography


