INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
RICE UNIVERSITY

The Behavior of Newton's Method on Two Equivalent Systems from Linear and Nonlinear Programming

by

María Cristina Villalobos

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED. THESIS COMMITTEE:

Richard A. Tapia, Chairman
Noah A. Harding Professor of
Computational and Applied Mathematics

Yin Zhang, Co-Chairman
Associate Professor of Computational and
Applied Mathematics

John Polking
Professor of Mathematics

Amr El-Bakry
Adjunct Associate Professor of
Computational and Applied Mathematics

Houston, Texas
August, 1999
Abstract

The Behavior of Newton’s Method on Two Equivalent Systems from Linear and Nonlinear Programming

by

María Cristina Villalobos

Newton’s method is a fundamental technique for approximating solutions of nonlinear equations. However, it is often not fully appreciated that the method can produce significantly different behavior when applied to equivalent systems. In this thesis, we investigate differences in local and global behavior of two well-known methods for constrained optimization: the Newton logarithmic barrier function method and the Newton primal-dual interior-point method. As we shall show, these two methods can be viewed as applying Newton’s method to two different but equivalent systems. Through theoretical analysis and numerical experimentation, we show the Newton primal-dual method performs more effectively.
Acknowledgments

I would like to dedicate this dissertation to my parents who have always been loving and supportive.

Many people have contributed to my education. I acknowledge a few of them here. I am greatly indebted to the following professors: Dr. Amr El-Bakry, Dr. Richard Tapia, and Dr. Yin Zhang who supported me throughout the years. They invested much of their time in my education, and I learned a great deal of optimization from them. It is very lengthy to mention all the numerous ways that Dr. Tapia has contributed to my academic, professional, and personal life, and I sincerely thank him for the person he has helped shaped. A special thank you goes to Jean, Rich, and Becky Tapia for making me a part of their family. Dr. Zhang I thank for all his patience and frankness. I truly enjoyed all the conversations we had for these moments have helped me to become a more confident and critical colleague of his.

Next, I would like to thank my friends at Rice, especially my colleagues in the Rice optimization group, with whom these graduate years would not have been educational, fun, and rewarding. I sincerely thank my friends with whom I shared many debates/discussions and ideas. I name only a few friends here with whom I spent much of my graduate career with and from whom I learned much: Miguel Argáez, Luis Melara, Zeferino Parada, Leticia Velázques, Donald Williams, and Erica Zimmer Klampfl.

I would also like to thank God for helping me through the difficult times of my life and for allowing me to find happiness everywhere, especially in the moments of laughter.
Financial support from the following institutions is gratefully acknowledged: the National Research Council (Ford Foundation Fellowship), the Spend a Summer with a Scientist program, and the Sloan Foundation. Financial support was also provided by the Center for Research on Parallel Computation cooperative agreement CCR-8890615 and the Department of Energy under the cooperative agreement DOE-DE-FG03-93ER25178.
# Contents

Abstract ........................................... ii  
Acknowledgments ................................... iii  
List of Illustrations ............................... vii  
List of Tables ..................................... viii  

1 Introduction .................................. 1  

2 Nonlinear Programming ....................... 5  
   2.1 Optimality Conditions ................. 6  
   2.2 Barrier System ......................... 7  
   2.3 Perturbed System ....................... 9  
   2.4 Trajectory of Solutions ............... 10  
   2.5 Newton Interior-Point Path-Following Methods 12  

3 Sphere of Convergence ....................... 14  

4 Linear Programming ........................... 17  
   4.1 Optimality Conditions ................. 18  
   4.2 Barrier System ......................... 19  
   4.3 Perturbed System ....................... 20  
   4.4 Theory for the Radius of the Sphere of Convergence 22  
      4.4.1 Preliminaries ..................... 23  
      4.4.2 Barrier System ................... 28  
      4.4.3 Perturbed System ................. 33
4.5 Numerical Experiments Concerning the Sphere of Convergence ......................... 37
  4.5.1 Nondegenerate Problems ........................................ 40
  4.5.2 Degenerate Problems ............................................ 41
4.6 Efficiency of an Interior-Point Method ........................................... 44
  4.6.1 Numerical Results ................................................ 46
4.7 Conclusions ........................................................................ 54

5 Inequality Constrained Nonlinear Optimization 56
  5.1 Optimality Conditions .................................................. 57
  5.2 Barrier System .......................................................... 58
  5.3 Perturbed System ......................................................... 60
  5.4 Theory for the Radius of the Sphere of Convergence ....................... 61
    5.4.1 Preliminaries ...................................................... 62
    5.4.2 Barrier System ................................................... 68
    5.4.3 Perturbed System ................................................ 80
  5.5 Numerical Experiments Concerning the Sphere of Convergence .......... 82
  5.6 Conclusions .................................................................. 88

6 Conclusions ........................................................................ 89

Bibliography ........................................................................... 91
Illustrations

3.1 .............................................................. 16

4.1 .............................................................. 41
4.2 .............................................................. 42
4.3 .............................................................. 42
4.4 .............................................................. 43

5.1 .............................................................. 86
5.2 .............................................................. 86
Tables

4.1 Results for $tol = 10^{-6}$ and fixed $\mu tol = 10^{-2}$ .................. 50
4.2 Results for $tol = 10^{-6}$ and fixed $\mu tol = 10^{-6}$ .................. 51
4.3 Results for $tol = 10^{-2}$ and fixed $\mu tol = 10^{-2}$ .................. 51
4.4 Results for $tol = 10^{-2}$ and fixed $\mu tol = 10^{-6}$ .................. 52
4.5 Results for $tol = 10^{-6}$ with initial $\mu tol = 10^{-1}$ and updated .. 52

5.1 Test Problems ................................................. 85
Chapter 1

Introduction

It is often not fully appreciated that Newton's method can exhibit different behavior when applied to equivalent nonlinear systems. In the nonlinear programming literature, two equivalent systems have evolved and have been the focus of considerable study. The first of these systems consists of the optimality conditions of the logarithmic barrier function formulation of the nonlinear program. The equivalent system of interest consists of the perturbed optimality conditions for the nonlinear program. This thesis examines properties associated with two Newton interior-point methods: a Newton logarithmic barrier function method and a Newton primal-dual interior-point method. As we shall show, these two methods can be viewed as applying Newton's method to the nonlinear systems of equations that arise from the two equivalent systems.

In the 1960s and 1970s, much research focused on studying the Newton logarithmic barrier (log-barrier) function method. This method applies Newton's method to the system of equations (barrier system) that arises from the optimality conditions of the log-barrier formulation. In order to obtain feasible points, the Newton step is damped. However, the popularity of the method was limited due to the inherent ill-conditioning in the Jacobian of the barrier system. Then in 1984, Karmarkar [11] generated considerable excitement by introducing a polynomial-time interior-point algorithm for linear programming. A year later, Gill, Murray, Saunders, Tomlin, and M. Wright [8] showed a similarity between Karmarkar's method and classical Newton log-barrier methods for the linear program. As a result, there was a rebirth of research in Newton log-barrier function methods.
Recent studies have focused on studying the performance of Newton log-barrier function methods for the nonlinear program. The literature contains many papers, such as [16, 20], that describe Newton log-barrier methods and their properties for the nonlinear program with inequality constraints. For example, M. Wright [20] investigates the effectiveness of a damped Newton step obtained from the barrier system immediately upon decreasing the barrier parameter. Several papers [21, 24] have focused on showing that the inherent ill-conditioning in the Jacobian of the barrier system may be reduced.

In the 1980's, Megiddo's [14] and Kojima, Mizuno, and Yoshise's [12] work on Newton primal-dual interior-point methods in linear programming opened a new area of research. The underlining component of these methods involves applying Newton's method to the system of equations (perturbed system) that arises from the perturbed optimality conditions for the linear program. Newton primal-dual methods have been a success for solving linear programs and are much favored over Newton log-barrier methods. As a consequence, some current investigation [4, 6, 19, 18] has focused on extending these methods to solve the nonlinear program. However, the reasons for the success of Newton primal-dual methods in linear programming seem not to be fully understood.

In this thesis, we isolate some properties associated with Newton primal-dual interior-point methods which we believe have contributed to their success in linear programming. We extend this study to Newton primal-dual methods for the nonlinear program to explain why these methods should be the method of choice for solving the nonlinear program. The results from this thesis will provide insight into the efficient performance of Newton primal-dual methods. In addition, this research will contribute to the development of efficient algorithms for the nonlinear program.
We compare the Newton log-barrier function method and the Newton primal-dual interior-point method in terms of their local and global convergence behavior and their efficiency for the linear and nonlinear program. In order to accomplish this objective, we introduce the notion of the sphere of convergence of Newton’s method. We then establish lower- and upper-bounds for the radius of the sphere of convergence of Newton’s method applied to the barrier and perturbed systems as a function of the barrier parameter. For nondegenerate linear programs, the radius of the sphere of convergence of Newton’s method on the perturbed system is proven to be bounded away from zero. However, the radius of the sphere of convergence of Newton’s method applied to the barrier system is shown to decrease to zero with the same order as $\mu$ decreases to zero. Similar results are shown for the inequality constrained nonlinear optimization problem. In particular, our results are sharper than S. Wright’s [22] lower-bound result for the radius of the sphere of convergence of Newton’s method on the barrier system for the nonlinear program.

In the remaining work, we conduct numerical experiments on the efficiency of the Newton log-barrier method and the Newton primal-dual method for the linear program. The results show that the Newton primal-dual method is more efficient for solving the linear program. The inherent ill-conditioning of the Jacobian of the barrier system can preclude the convergence of the Newton log-barrier method. Furthermore, the Newton log-barrier method may require more iterations to converge to a solution than does the Newton primal-dual method.

The thesis is organized as follows. Chapter 2 lays the foundation for topics which will be developed in later chapters. We first discuss the general nonlinear program and the associated optimality conditions. Having presented the two equivalent nonlinear systems under consideration, we discuss the Newton log-barrier function method and the Newton primal-dual interior-point method. We end the chapter with a discus-
sion on the trajectory of solutions produced by the two equivalent nonlinear systems. Chapter 3 defines the sphere of convergence of Newton's method. Chapter 4 introduces the linear program and the barrier and perturbed systems under consideration. Next, we show theoretical and numerical results concerning the radius of the sphere of convergence of the pure Newton log-barrier and pure Newton primal-dual methods for nondegenerate problems. We also present numerical results for degenerate problems. Finally, numerical results are presented to determine the efficiency of a Newton log-barrier method and a Newton primal-dual method for the linear program. Chapter 5 introduces the barrier and perturbed systems of the inequality constrained optimization problem. Then we discuss some theoretical results for the radius of the sphere of convergence of the pure Newton log-barrier method and the pure Newton primal-dual method for the nonlinear program for nondegenerate problems. We present also numerical results supporting the theoretical findings for the radius of the sphere of convergence of the two methods. Finally, a summary of the results comprises Chapter 6.

Chapters 2, 4, and 5 contain references to the barrier and perturbed systems. It should be clear that reference is made only to the systems under discussion within the chapter, unless otherwise noted.
Chapter 2

Nonlinear Programming

Two commonly used Newton interior-point methods have evolved for approximating functions of nonlinear programs: the Newton logarithmic barrier (log-barrier) function method and the Newton primal-dual method. Each method consists of applying Newton's method to a nonlinear system of equations and dampening the step to produce feasible points. The optimality conditions of the log-barrier formulation of the nonlinear program form the foundation of the Newton log-barrier method. The kernel of the Newton primal-dual method consists of the perturbed optimality conditions of the nonlinear program. Though the two nonlinear systems are equivalent, El-Bakry, Tapia, Tsuchiya, and Zhang [4] show that the above two methods produce necessarily different iterates for the linear program, and one would expect the same in the case of the nonlinear program.

We begin by introducing the general nonlinear program and the associated optimality conditions. In Section 2.2, we discuss the optimality conditions for the log-barrier subproblem of the nonlinear program. The subsequent section discusses the perturbed optimality conditions for the nonlinear program.

Section 2.4 discusses the equivalence of the two nonlinear systems under consideration for the nonlinear program. We describe also the existence of a trajectory parameterized by the solutions of the nonlinear system. Finally, we discuss two well-known Newton interior-point methods that have been a subject of interest for solving the nonlinear program.
2.1 Optimality Conditions

We first introduce the nonlinear program and then discuss the optimality conditions. Consider the nonlinear program in the following form

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad h(x) = 0 \\
& \quad g(x) \geq 0
\end{align*}
\]  \hspace{1cm} (2.1)

where \( f : \mathbb{R}^p \to \mathbb{R} \), \( h : \mathbb{R}^p \to \mathbb{R}^m \), and \( g : \mathbb{R}^p \to \mathbb{R}^m \) are twice continuously differentiable.

The Lagrangian function for problem (2.1) is given by

\[
L(x, y, z) = f(x) + \sum_{i=1}^{p} y_i h_i(x) - \sum_{i=1}^{m} z_i g_i(x).
\]

where \( y \in \mathbb{R}^p \) and \( z \in \mathbb{R}^m \) are, respectively, the Lagrangian multipliers associated with the equality and inequality constraints. The gradient of the Lagrangian function with respect to \( x \) is

\[
\nabla_x L(x, y, z) = \nabla f(x) + \sum_{i=1}^{p} y_i \nabla h_i(x) - \sum_{i=1}^{m} z_i \nabla g_i(x)
\]

where \( \nabla h_i(x) \) denotes the gradient of \( h_i(x) \) and similarly for \( \nabla g_i(x) \). Let \( x^* \) be a solution of problem (2.1). Let \( B = \{ i : g_i(x^*) = 0 \} \), i.e., \( B \) denotes the set of indices corresponding to the inequality constraints that are binding or active at \( x^* \). Regularity holds at \( x^* \) if the set of active constraint gradients.

\[
\{ \nabla h_1(x^*), \ldots, \nabla h_p(x^*) \} \cup \{ \nabla g_i(x^*) : i \in B \}
\]  \hspace{1cm} (2.2)
is linearly independent. If the functions \( f, h, \) and \( g \) are differentiable and regularity holds at a minimizer \( x^* \) then the following first-order necessary conditions, otherwise known as the Karush-Kuhn-Tucker (KKT) conditions, must hold for multipliers \( y^* \) and \( z^* \)

\[
\nabla_x L(x^*, y^*, z^*) = 0 \\
h(x^*) = 0 \\
Z^* g(x^*) = 0 \\
g(x^*) \geq 0 \\
z^* \geq 0
\] (2.3)

where \( Z^* \) denotes the diagonal matrix \( \text{diag}(z^*) \). Strict complementarity holds if \( g_i(x^*) + z_i^* > 0 \) for \( i = 1, \ldots, m \).

By the second-order condition, we mean that for \( \eta \neq 0 \) and \( \nabla h_i(x^*)^T \eta = 0, i = 1, \ldots, p \) and \( \nabla g_i(x^*)^T \eta = 0, i \in B \) and \( \nabla g_i(x^*)^T \eta \geq 0, i \notin B \)

\[
\eta^T \nabla_x^2 L(x^*, y^*, z^*) \eta > 0.
\] (2.4)

If the functions \( f, h, \) and \( g \) are twice differentiable, then the sufficiency conditions for a unique local minimizer \( x^* \) are that the KKT conditions and the second-order condition hold. For further details on Lagrange multiplier theory, see Avriel [1].

2.2 Barrier System

In this section, we introduce the logarithmic barrier framework for the nonlinear program. Then we discuss the barrier system which is associated with the optimality conditions of the logarithmic barrier formulation for the nonlinear program.

Implicit in Frisch’s work [7], the logarithmic barrier (log-barrier) framework for the nonlinear program (2.1) consists of solving a sequence of equality constrained min-
imization problems that are parameterized by $\mu > 0$. For a given value of the barrier parameter $\mu > 0$, the log-barrier subproblem for problem (2.1) has the following form

$$\begin{align*}
\text{minimize} & \quad f(x) - \mu \sum_{i=1}^{m} \log g_i(x) \\
\text{subject to} & \quad h(x) = 0 \\
& \quad (g(x) > 0).
\end{align*}$$

The second term in the objective function is a penalty term, which penalizes the objective function at points close to the boundary of the feasible region and forces the inequality constraints to be satisfied directly. Thus, the framework must cope only with satisfying the equality constraints. Let $x_\mu^*$ denote the solution of the log-barrier problem for a given value of $\mu$ sufficiently small and positive. Then under certain conditions (see for example, Fiacco and McCormick [5]), as $\mu \to 0$ the sequence of iterates $\{x_\mu^*\}$ converges to a solution $x^*$ of problem (2.1).

To obtain the optimality conditions for the log-barrier subproblem we consider first the Lagrangian function

$$\mathcal{L}(x, y) = f(x) - \mu \sum_{i=1}^{m} \log g_i(x) + \sum_{i=1}^{p} y_i h_i(x).$$

where $y \in \mathbb{R}^p$ are the Lagrangian multipliers associated with the equality constraints. Then the optimality conditions are given by

$$F_B(x, y; \mu) \equiv \begin{bmatrix} \nabla_x \mathcal{L}(x, y) \\ h(x) \end{bmatrix} = 0, \quad (g(x) > 0) \quad (2.5)$$

where

$$\nabla_x \mathcal{L}(x, y) = \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) + \sum_{i=1}^{p} y_i \nabla h_i(x).$$
is the gradient of the Lagrangian function with respect to the primal variables $x$. We will refer to the equations in system (2.5) as the barrier system. The Jacobian of $F_B(x, y; \mu)$ is given by

$$F'_B(x, y; \mu) \equiv \begin{bmatrix} \nabla^2_x \mathcal{L}(x, y) & \nabla h(x) \\ \nabla^T h(x) & 0 \end{bmatrix}$$

where

$$\nabla^2_x \mathcal{L}(x, y) = \nabla^2 f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla^2 g_i(x) + \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) \nabla^T g_i(x) + \sum_{i=1}^{p} y_i \nabla^2 h_i(x)$$

is the Hessian of the Lagrangian function. $\nabla^2 g_i(x)$ denotes the Hessian of $g_i(x)$, and similarly for $\nabla^2 h_i(x)$. Let $(x^*_\mu, y^*_\mu)$ denote the solution to system (2.5). For small $\mu > 0$, the Jacobian is nonsingular at $(x^*_\mu, y^*_\mu)$ if regularity, strict complementarity, and the second-order condition (2.4) hold.

### 2.3 Perturbed System

We present a system equivalent to system (2.5) introduced in the previous section. The equivalence between the new system and system (2.5) will be discussed in Section 2.4.

Consider the introduction of an auxiliary variable $z \in \mathbb{R}^n$. Define $z_i$ to be the term $\mu/g_i(x)$ that appears in the gradient of the Lagrangian function in system (2.5). Then $z_i = \mu/g_i(x)$ can be written equivalently as $Zg(x) = \mu e$, where $Z = \text{diag}(z)$ and $e$ is a vector of all ones with appropriate dimension. These equivalent defining relations yield the system
\begin{equation}
F_p(x, y, z; \mu) \equiv \begin{bmatrix} \nabla_x \ell(x, y, z) \\ h(x) \\ Zg(x) - \mu e \end{bmatrix} = 0, \quad g(x).z \geq 0 \tag{2.6}
\end{equation}

where

$$\nabla_x \ell(x, y, z) = \nabla f(x) + \sum_{i=1}^{p} y_i \nabla h_i(x) - \sum_{i=1}^{m} z_i \nabla g_i(x)$$

is the gradient of the Lagrangian of problem (2.1). System (2.6) can also be obtained by considering the KKT conditions (2.3) of problem (2.1) and perturbing the complementarity equation. \(Zg(x) = 0\) by \(\mu e\). We will refer to the system \(F_p(x, y, z) = 0\) as the \textit{perturbed system}.

The Jacobian of \(F_p(x, y, z; \mu)\) is given by

$$F_p'(x, y, z; \mu) \equiv \begin{bmatrix} \nabla^2 \ell(x, y, z) & \nabla h(x) & -\nabla g(x) \\ \nabla^T h(x) & 0 & 0 \\ Z \nabla^T g(x) & 0 & G(x) \end{bmatrix}$$

where \(\nabla^2 \ell(x, y, z)\) denotes the Hessian of the Lagrangian and \(G(x) = \text{diag}(g(x))\). Let \((x^*_\mu, y^*_\mu, z^*_\mu)\) denote the solution to system (2.6). El-Bakry et al. [4] proved that for small \(\mu > 0\), \(F_p'(x^*_\mu, y^*_\mu, z^*_\mu; \mu)\) is nonsingular if regularity, strict complementarity, and the second-order condition (2.4) for the nonlinear problem (2.1) hold.

### 2.4 Trajectory of Solutions

The previous section showed how system (2.6) could be obtained from system (2.5) or from the optimality conditions of the nonlinear program (2.1). This section discusses the equivalence of systems (2.5) and (2.6) for the nonlinear program given that regularity and the sufficiency conditions hold. Finally, we introduce a trajectory of solutions produced by these two systems.
Systems (2.5) and (2.6) are parameterized by \( \mu > 0 \). For a particular value of \( \mu > 0 \), let \((x^*_\mu, y^*_\mu)\) denote the solution to system (2.5). and similarly let \((x^*_\mu, y^*_\mu, z^*_\mu)\) denote the solution to system (2.6). Under mild conditions, for \( \mu \) small and positive, the Jacobians of \( F_B(x, y; \mu) \) and \( F_P(x, y, z; \mu) \) are nonsingular at the solutions \((x^*_\mu, y^*_\mu)\) and \((x^*_\mu, y^*_\mu, z^*_\mu)\), respectively. In addition, the Jacobian \( F_P(x^*, y^*, z^*) \) is nonsingular at \( \mu = 0 \).

Assume \( f, h, \) and \( g \) are twice continuously differentiable. Let \( x^* \) satisfy the sufficiency conditions and regularity. Then for small \( \hat{\mu} > 0 \) and \( \mu \leq \hat{\mu} \) (see [5] for example), there exists a trajectory of solutions \((x^*_\mu, y^*_\mu)\) for system (2.5) which is described by

\[
C_B = \{(x^*_\mu, y^*_\mu) : F_B(x^*_\mu, y^*_\mu; \mu) = 0, g(x^*_\mu) > 0, \hat{\mu} \geq \mu > 0 \}. \tag{2.7}
\]

In particular, as \( \mu \to 0 \), the sequence \( \{(x^*_\mu, y^*_\mu)\} \) converges to \((x^*, y^*)\), where \( y^* \) is the vector of Lagrangian multipliers corresponding to the equality constraints.

Systems (2.5) and (2.6) are equivalent, in the sense that for \( 0 < \mu \leq \hat{\mu} \) and \((x^*_\mu, y^*_\mu) \in C_B\),

\[
F_B(x^*_\mu, y^*_\mu; \mu) = 0 \iff F_P(x^*_\mu, y^*_\mu, z^*_\mu; \mu) = 0
\]

for \((z^*_\mu)_i = \mu / g(x^*_\mu)_i, \ i = 1, \ldots, m\) (see for example [4]). This result states that systems (2.5) and (2.6) share the same set of solutions. For system (2.6), we also have \( \lim_{\mu \to 0}(x^*_\mu, y^*_\mu, z^*_\mu) = (x^*, y^*, z^*) \), where \( z^* \) are the Lagrangian multipliers corresponding to the inequality constraints for problem (2.1). For this system the trajectory of solutions consists of the set

\[
C_P = \{(x^*_\mu, y^*_\mu, z^*_\mu) : F_P(x^*_\mu, y^*_\mu, z^*_\mu; \mu) = 0, z^*_\mu, g(x^*_\mu) > 0, \hat{\mu} \geq \mu \geq 0 \}. \tag{2.8}
\]
We refer the reader to the proofs given by Fiacco and McCormick [5] for further
details on the existence and smoothness of the trajectory for systems (2.5) and (2.6).
McLinden [13] was the first to show the existence and smoothness of the trajectory
of solutions for a convex nonlinear program.

When strict complementarity fails to hold, the solution \((x^*, y^*)\) may not be locally
unique or isolated. Thus, many trajectories may exist leading to the solution \((x^*, y^*)\).
For the inequality constrained optimization problem, Jittorntrum and Osborne [10]
show that in the case where strict complementarity fails to hold, but regularity and
the sufficiency conditions hold at \(x^*\) and the sequence \(\{x^*_n\}\) converges to \(x^*\), then a
smooth trajectory of solutions exists for \(\mu\) sufficiently small.

2.5 Newton Interior-Point Path-Following Methods

A brief description of Newton interior-point path-following methods for solving the
nonlinear program will be presented.

As is well known, in general Newton's method is only locally convergent. In most
applications, a "globally convergent" method is desired. We can create a globally
convergent method by modifying Newton's method and applying it to different pa-
rameterized systems until a solution is obtained. Hence, we can apply Newton's
method to the barrier or perturbed systems with decreasing values of \(\mu\). In this
manner, we are following the trajectory (2.7) or (2.8) to a solution of the nonlinear
program. Interior-point path-following methods are "globally convergent" methods
of the type just described.

Two well-known Newton interior-point path-following methods are the Newton
log-barrier method and the Newton primal-dual method. These methods generate
iterates that follow the trajectory, equivalently defined by (2.7) or (2.8), to a solution
of the nonlinear program. For a given value of \(\mu\), the Newton log-barrier method
(Newton primal-dual method) applies Newton’s method to the barrier system (perturbed system). The Newton step is damped to drive the iterates towards the point on the trajectory corresponding to the given μ value and to keep the inequality constraints. \( g(x) \geq 0 \) in system (2.5) and \( g(x).z \geq 0 \) in system (2.6), strictly positive, that is, to keep the iterates in the interior of the feasible region. In practice, the two equivalent systems (2.5) and (2.6) are solved only approximately.

For nonlinear programming, the trajectory may exist only for small values of \( \mu > 0 \), that is, for points that lie close to a solution of the nonlinear program. In this case, it may not be clear how to implement algorithms to follow the trajectory. However, in the case of the linear program, Megiddo [14] showed that under mild conditions the trajectory of solutions exists for every \( \mu > 0 \) for the two equivalent systems (2.5) and (2.6).
Chapter 3

Sphere of Convergence

Under a set of standard assumptions, local convergence theory of Newton’s method (see for example [2]) provides the existence of a neighborhood about a solution where Newton’s method is well-defined. More importantly, starting from any point in the neighborhood, Newton’s method converges to a solution of the nonlinear system. The local convergence analysis of Newton’s method can be applied to the barrier and perturbed systems. In this manner, we also obtain a neighborhood about each point on the trajectory, equivalently given by (5.8) and (5.9). Starting from any point in the neighborhood, Newton’s method will converge to the solution of the nonlinear system.

The notion of a region of convergence is standard. Several references can be found in the literature where this notion or a similar concept is used, see [2, pg 91], for example. In this chapter, we introduce the notion of the sphere of convergence of Newton’s method. In subsequent chapters, we analyze the behavior of the radius of the sphere of convergence of Newton’s method applied to the barrier and perturbed systems as $\mu$ decreases to zero for the linear program and the inequality constrained optimization problem. We begin now by providing a definition for the sphere of convergence.

Definition The closed ball with radius $r$ centered at $v^*$ is defined as $B(v^*, r) = \{v : \|v - v^*\| \leq r\}$.

Definition For a given nonlinear system, $F(v) = 0$ and a solution $v^*$, the sphere of convergence of Newton’s method at $v^*$ is defined as the largest closed ball
centered at \( v^\ast \) such that starting from any point in the interior of the sphere.

other than \( v^\ast \), Newton's method is well-defined and generates a sequence that

converges to \( v^\ast \).

We note that the Jacobian of \( F(v) \) may be singular or undefined at \( v^\ast \) and therefore restrict Newton's method to be defined everywhere in the neighborhood except at \( v^\ast \).

Suppose that the solutions for systems (2.5) and (2.6) exist for \( 0 < \mu \leq \hat{\mu} \). Let \( v^\ast_\mu \)
denote the solution to the nonlinear system \( F(v; \mu) = 0 \) for \( 0 < \mu \leq \hat{\mu} \). We can think

of system \( F(v; \mu) = 0 \) with solution \( v^\ast_\mu \) as a special case of the solutions of systems

(2.5) or (2.6), where \( v^\ast_\mu \) denotes \( (x^\ast_\mu, y^\ast_\mu) \) or \( (x^\ast_\mu, y^\ast_\mu, z^\ast_\mu) \), respectively.

Suppose a sphere of convergence exists about each solution \( v^\ast_\mu \). For the purpose of
discussing and illustrating the sphere of convergence of Newton's method, we will assume the sphere of convergence has been determined for various values of \( \mu \). However, in Chapters 4 and 5, we analyze the behavior of the radius of the sphere of convergence of Newton's method as \( \mu \) decreases to zero.

Figure 3.1 shows the trajectory of points \( v^\ast_\mu \) near a solution \( v^\ast \) and the spheres of convergence of Newton's method at three points, \( v^\ast_\mu_1, v^\ast_\mu_2, \) and \( v^\ast_\mu_3 \), for three different values of \( \mu \) associated with a nonlinear system, \( F(v; \mu) = 0 \). To obtain a solution to

the KKT conditions (2.3), systems (5.6) and (5.7) are solved for a sequence of values

with \( \mu > 0 \). Then as \( \mu \) decreases to zero, a solution \( v^\ast \) is obtained.

We are primarily concerned with the behavior of the radii of the spheres of convergence of Newton's method applied to the barrier and perturbed systems as \( \mu \)
decreases to zero. The asymptotic behavior of the radii will provide an explanation

for the performance of the Newton log-barrier and the Newton primal-dual interior-
point method. As a reminder, we note that the former (latter) method can be viewed

as applying Newton's method to the barrier system (perturbed system). If the radius

for the sphere of convergence of Newton's method applied to \( F(v; \mu) = 0 \) is bounded
away from zero as $\mu$ decreases to zero, then this implies that few iterations would be required by Newton's method to converge to a solution $v_\mu^*$ on the trajectory. However, if the radius decreases to zero, then more iterations would be required by a Newton interior-point method to converge to a solution $v^*_\mu$ and, in general, to obtain a solution $v^*$.

\textbf{Figure 3.1} Spheres of convergence for three $\mu$ values.
Chapter 4

Linear Programming

In this chapter, we investigate theoretical and numerical results on properties associated with a Newton log-barrier function method and a Newton primal-dual interior-point method for the linear program. The Newton log-barrier method can be viewed as applying a damped Newton's method to the system of equations (barrier system) arising from the optimality conditions of the log-barrier formulation of the linear program. Similarly, the Newton primal-dual method can be viewed as applying Newton's method to the system of equations (perturbed system) arising from the perturbed optimality conditions of the linear program.

We study first the radius of the sphere of convergence (see Chapter 3) of Newton's method on the two equivalent systems as $\mu$ decreases to zero for nondegenerate problems. Our analysis shows that the radius of the sphere of convergence associated with the barrier system decreases to zero with the same order that the barrier parameter, $\mu$, decreases to zero. However, we prove that the radius of the sphere of convergence of Newton's method applied to the perturbed system is bounded away from zero. Numerical experimentation validates our theoretical findings. For degenerate problems, we conduct only numerical experiments of the radius of the sphere of convergence of Newton's method on the two equivalent systems.

Finally, the global behavior of a Newton log-barrier method and a Newton primal-dual interior-point method for the linear program is studied. In particular, we are interested in the efficiency of the two damped Newton interior-point methods for nondegenerate and degenerate problems. We will measure efficiency in terms of the number of iterations required to solve a system.
The chapter is organized as follows. In Section 4.1, we discuss the optimality conditions for the linear program and present terminology used in this chapter. Sections 4.2 and 4.3 describe the two equivalent nonlinear systems under consideration: the optimality conditions of the log-barrier formulation for the linear program and the perturbed optimality conditions for the linear program. In addition, we briefly discuss the unique trajectory of solutions for the two equivalent systems. Section 4.4 discusses our theoretical results for nondenegerate problems on the radius of the sphere of convergence of Newton's method on the two equivalent systems. Section 4.5 presents numerical experiments of the radius of the sphere of convergence of Newton's method on nondegenerate and degenerate problems. In Section 4.6, we discuss numerical results on the efficiency of a Newton log-barrier method and a Newton primal-dual method for the linear program. Finally, we present our concluding remarks.

4.1 Optimality Conditions

We present the optimality conditions for the linear program and conclude this section with some terminology. Consider the following linear programming problem

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

(4.1)

where \(c, x \in \mathbb{R}^n, b \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}\) with \(m \leq n\) and \(\text{rank}(A) = m\). The Lagrangian function for the linear program (4.1) is given by

\[L(x, y, z) = c^T x + y^T (Ax - b) - x^T z.\]
where \( y \in \mathbb{R}^n \) and \( z \in \mathbb{R}^n \) are, respectively, the vectors of Lagrange multipliers associated with the equality and the inequality constraints. Then the first-order optimality conditions for problem (4.1) are

\[
\begin{align*}
\nabla_z L(x, y, z) &= A^T y + z - c = 0 \\
\nabla_y L(x, y, z) &= Ax - b = 0 \\
XZe &= 0 \\
x.z &\geq 0
\end{align*}
\]

where \( X = \text{diag}(x), Z = \text{diag}(z), \) and \( e = (1, \ldots, 1)^T \in \mathbb{R}^n \).

**Terminology**

Let \( B = \{i : (x^*)_i > 0\} \) and \( N = \{i : (x^*)_i = 0\} \). **Primal nondegeneracy** holds when \( m \) components of \( x^* \) are positive and the remaining components are zero. Without loss of generality, let the first \( m \) components of \( x^* \) be positive, and the remaining \( (n - m) \) components be zero. Then \( B = \{1, \ldots, m\} \) and \( N = \{m + 1, \ldots, n\} \). **Dual nondegeneracy** holds when the first \( m \) components of \( z^* \) are zero, and the last \( (n - m) \) components are positive. Primal and dual nondegeneracy imply **strict complementarity**, that is, \( x^*_i + z^*_i > 0 \) for \( i = 1, \ldots, n \).

### 4.2 Barrier System

As described in Section 2.2, the linear program can be solved via a sequence of log-barrier subproblems as the barrier parameter, \( \mu \), decreases to zero. Thus, for a given \( \mu > 0 \), we solve
\begin{align*}
\text{minimize} & \quad c^T x - \mu \sum_{i=1}^{n} \log x_i \\
\text{subject to} & \quad Ax = b \\
& \quad (x > 0).
\end{align*}

The optimality conditions for the log-barrier subproblem are given by

\[ F_B(x, y; \mu) \equiv \begin{bmatrix} A^T y + \mu X^{-1} \epsilon - c \\ Ax - b \end{bmatrix} = 0 \quad (x > 0). \quad (4.3) \]

The system of equations in (4.3) will be referred to as the barrier system for the linear program. Observe that the Jacobian of \( F_B(x, y; \mu) \) is given by

\[ F'_B(x, y; \mu) \equiv \begin{bmatrix} -\mu X^{-2} & A^T \\ A & 0 \end{bmatrix}. \]

If \( \text{rank}(A) = m \) and \( x > 0 \), then \( F'_B(x, y; \mu) \) is nonsingular for \( \mu > 0 \). If primal and dual nondegeneracy hold at \( (x^*, z^*) \), then several primal variables \( x \) are zero at the solution. Consequently, near the solution the Jacobian necessarily becomes ill-conditioned for \( \mu \) close to zero.

### 4.3 Perturbed System

We derive a nonlinear system that is equivalent to system (4.3) for the linear program. By equivalence, we mean that the two systems have the same set of solutions for \( \mu > 0 \), see Section 2.4. Then we briefly present the perturbed system for the linear program.

As in Section 2.3, we introduce the auxiliary variable \( z \in \mathbb{R}^n \), and define \( z = \mu X^{-1} \epsilon \). Then substituting for \( z \) in system (4.3) and adding the defining relation \( XZe = \mu \epsilon \), we obtain the equivalent system
\[ F_p(x, y, z; \mu) \equiv \begin{bmatrix} A^T y + z - c \\ Ax - b \\ XZe - \mu e \end{bmatrix} = 0 \quad x, z \geq 0. \quad (4.4) \]

The system of equations in (4.4) will be referred to as the perturbed system. The Jacobian of \( F_p(x, y, z; \mu) \) is given by

\[ F_p'(x, y, z; \mu) \equiv \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix}. \]

The Jacobian is nonsingular if \( \text{rank}(A) = m \) and \( x, z > 0 \). In addition, if primal and dual nondegeneracy hold at \( x^* \) and \( z^* \) (see Section 4.4.1), respectively, and the columns of the matrix \( A \) associated with the set \( B \) are linearly independent, then the solution \( (x^*, y^*, z^*) \) is unique. Hence, the Jacobian is nonsingular at \( (x^*, y^*, z^*) \).

System (4.4) can be obtained also by considering the first-order optimality conditions (4.2) of the linear program and perturbing the complementarity equation \( XZe = 0 \) by \( \mu e \).

**Central Path**

Section 2.4 discusses the trajectory produced by the solutions of the two equivalent nonlinear systems: system (2.5) and system (2.6) for the nonlinear program. In general, solutions to the equivalent systems may exist only for small \( \mu > 0 \). However, for the linear program, Megiddo [14] proved that under mild conditions the trajectory exists for all \( \mu > 0 \). In linear programming, the trajectory is conveniently called the central path. The central path for system (4.3) is given by

\[ C_B = \{(x^*_\mu, y^*_\mu) : F_B(x^*_\mu, y^*_\mu; \mu) = 0, x^*_\mu > 0, \mu > 0\}. \quad (4.5) \]
The central path for system (4.4) is given by

\[ C_P = \{(x_\mu^*, y_\mu^*, z_\mu^*) : F_P(x_\mu^*, y_\mu^*, z_\mu^*; \mu) = 0, x_\mu^*, z_\mu^* > 0, \mu > 0\}. \tag{4.6} \]

An interesting result on the two equivalent systems of the linear program was proved by El-Bakry et al. [4]. They show that although system (4.3) and system (4.4) are equivalent, Newton's method applied to the barrier and perturbed systems necessarily generates different iterates for the two systems. The following section shows a new result on the behavior of Newton's method applied to the barrier and perturbed systems for the linear program.

### 4.4 Theory for the Radius of the Sphere of Convergence

This section analyzes the behavior of the radius of the sphere of convergence of Newton's method applied to the barrier and perturbed systems as \( \mu \) decreases to zero. Under the nondegeneracy assumption, the analysis shows that the radius of the sphere of convergence of Newton's method on the barrier system (4.3) decreases to zero in the same order as \( \mu \) decreases to zero. However, it is shown that the radius of the sphere of convergence of Newton's method applied to the perturbed system (4.4) has a lower-bound independent of \( \mu \). These results show Newton's method to be more efficient on the perturbed system than on the barrier system, at least for small values of \( \mu > 0 \). In the next section, we introduce preliminary results to be used in our analysis for the radii of the spheres of convergence of Newton's method applied to the two equivalent systems. Then, we present theory for the radius of the sphere of convergence associated with the barrier system in Section 4.4.2 and subsequently present theory associated with the perturbed system in Section 4.4.3.
4.4.1 Preliminaries

This section introduces notation, a nondegeneracy assumption, and lemmas that will be used for the analysis concerning the radius of the sphere of convergence of Newton's method applied to the barrier (4.3) and perturbed (4.4) systems. The definition for the sphere of convergence at a point $v^*$ for the general nonlinear system $F(v) = 0$ is given in Chapter 3.

Notation

Recall that $\mathcal{B} = \{i : (x^*)_i > 0\}$ and $\mathcal{N} = \{i : (x^*)_i = 0\}$. Let the matrix $A$ be partitioned into $A = [A_\mathcal{B} \ A_\mathcal{N}]$ where $A_\mathcal{B}$ denotes the submatrix consisting of the columns of $A$ with indices in $\mathcal{B}$ and similarly for $A_\mathcal{N}$. Note that rank($A_\mathcal{B}$) = $m$. If $u$ is a vector, then its uppercase counterpart $U$ will denote the diagonal matrix whose diagonal consists of the elements of $u$. For a vector $u \in \mathbb{R}^n$, $u_\mathcal{B}$ is the vector whose indices are in $\mathcal{B}$, and $u_\mathcal{N}$ is the vector indexed by $\mathcal{N}$. We will call $u_\mathcal{B}$ the vector of basic components of $u$, and $u_\mathcal{N}$ will be the vector of nonbasic components of $u$. The quantity $u^2$ represents the vector $u$ whose components are individually squared. All norms $\| \cdot \|$ are assumed to be the Euclidean norm unless otherwise noted.

The following assumption will be referred to in the subsequent lemmas and theorems for the analysis of the radius of the sphere of convergence for the barrier and perturbed systems.

Nondegeneracy Assumption The matrix $A$ has full rank $m$. The solution $(x^*, y^*, z^*)$ is primal and dual nondegenerate for system (4.2).

The following two lemmas describe bounds for the nonnegative points $x$ and $(x, z)$ that lie close to the central paths described by (4.5) and (4.6). These bounds will be
of particular use in proving our results on the radius of the sphere of convergence of Newton's method on the two equivalent systems.

**Lemma 4.1** Consider $\mu > 0$ and $(x^*_\mu, y^*_\mu, z^*_\mu)$ contained in $C_p$, which is given by (4.6). Then under the nondegeneracy assumption, there exists $\tilde{\mu} > 0$ so that for $\mu \leq \tilde{\mu}$ there is a ball $B((x^*_\mu, z^*_\mu); \delta_\mu)$ and constants $C_1, C_2, C_3, C_4 > 0$ such that for $(x, z) \in B((x^*_\mu, z^*_\mu); \delta_\mu)$, and $(x, z) > 0$, $(x, z)$ satisfies

\[
x_i \geq C_1 \quad \text{and} \quad z_i \leq C_2 \mu \quad i \in B
\]

\[
x_i \leq C_3 \mu \quad \text{and} \quad z_i \geq C_4 \quad i \in N.
\]

(4.7)

**Proof**

Let $\mathbb{R}^n_{\mu} = \{(x, z) : x \geq 0, z \geq 0\}$. From the last equation in the perturbed system (4.4), we have for each $\mu > 0$.

\[
(x^*_\mu)_i (z^*_\mu)_i = \mu. \quad \text{for } i = 1, \ldots, n.
\]

(4.8)

We have that $x^*_i, i \in B$ and $z^*_i, i \in N$ are strictly positive, and $(x^*_\mu, z^*_\mu) \to (x^*, z^*)$ as $\mu \to 0$. Then there exist $\tilde{\mu} > 0$ and $C > 0$ such that for $\mu \leq \tilde{\mu}$, $(x^*_\mu)_i \geq C$ for $i \in B$ and $(z^*_\mu)_i \geq C$ for $i \in N$.

Consider $\mu \leq \tilde{\mu}$. Since $(x^*_\mu, z^*_\mu)$ is an interior point of $\mathbb{R}^n_{\mu}$, there exists a ball centered at $(x^*_\mu, z^*_\mu)$ containing strictly positive points $(x, z) \in \mathbb{R}^n_{\mu}$ which satisfy

\[
\left\| \begin{pmatrix} x - x^*_\mu \\ z - z^*_\mu \end{pmatrix} \right\| \leq \delta_\mu.
\]

(4.9)

where $\delta_\mu > 0$ and $\delta_\mu \leq \{\beta \mu, C/2\}$ for $\beta > 0$. From (4.9), we obtain

\[
x_i - (x^*_\mu)_i \leq \delta_\mu \quad \text{and} \quad z_i - (z^*_\mu)_i \leq \delta_\mu \quad \text{for } i = 1, \ldots, n.
\]

(4.10)
First we show that \( x_i \) for \( i \in \mathcal{B} \) is bounded away from zero. From the first relation in (4.10) we have

\[
(x^*_\mu)_i - \delta_\mu \leq x_i \leq \delta_\mu + (x^*_\mu)_i \quad \text{for } i = 1, \ldots, n.
\]  

(4.11)

Since \((x^*_\mu)_i \geq C\) for \(i \in \mathcal{B}\) and \(\delta_\mu \leq C/2\), from (4.11) we obtain

\[
0 < C/2 \leq C - \delta_\mu \leq x_i \quad \text{for } i \in \mathcal{B}.
\]

Thus \( x_i \geq C_1 \) for \( i \in \mathcal{B} \) with \( C_1 = C/2 \). Similarly, we can show \( z_i \geq C_4 \) for \( i \in \mathcal{X} \) and \( C_4 > 0 \).

Now, we show the second part of the proof. By (4.8) and (4.11) we obtain

\[
x_i \leq \delta_\mu + \frac{\mu}{(z^*_\mu)_i} = (\delta_\mu / \mu + 1/(z^*_\mu)_i) \mu.
\]

Now \( \delta_\mu / \mu \leq \beta \), and for \( i \in \mathcal{X} \), \((z^*_\mu)_i \geq C\), hence

\[
x_i \leq (\beta + 1/C) \mu = C_3 \mu \quad \text{for } i \in \mathcal{X}
\]

where \( C_3 = \beta + 1/C > 0 \). Therefore

\[
x_i \leq C_3 \mu \quad i \in \mathcal{X}.
\]

Similarly, from the second relation in (4.10), we obtain that for some constant \( C_2 > 0 \), \( z_i \leq C_2 \mu \) for \( i \in \mathcal{B} \).

\[\square\]

**Lemma 4.2**  Consider \( \mu > 0 \) and \( (x^*_\mu, y^*_\mu) \) contained in \( \mathcal{C}_B \), which is given by (4.5). Then under the nondegeneracy assumption, there exists \( \tilde{\mu} > 0 \) so that for \( \mu \leq \tilde{\mu} \) there
is a ball $B(x^*_\mu; \delta_\mu)$ and constants $G_1, G_2 > 0$ such that for $x \in B(x^*_\mu; \delta_\mu)$, and $x > 0$, $x$ satisfies

$$x_i \geq G_1 \quad \text{for } i \in \mathcal{B}$$

$$x_i \leq G_2 \mu \quad \text{for } i \in \mathcal{N}.$$  \hfill (4.12)

**Proof** The proof is similar to that of Lemma 4.1 so we omit the proof. \hfill \Box

The next two lemmas provide bounds for submatrices associated with the barrier and perturbed systems. Note that the bounds depend on the parameter $\mu$. Lemma 4.3 will be used in the analysis of the radius of the sphere of convergence of Newton's method on the perturbed system, and the subsequent lemma will be used for the barrier system. The proofs in these two lemmas are similar to the proof given by Zhang, Tapia, and Dennis [26].

**Lemma 4.3** Let the nondegeneracy assumption hold. Also, let $\tilde{A} \equiv AW$ where $W = (X)^{\frac{1}{2}}(Z)^{-\frac{1}{2}}$. Define

$$P = \tilde{A}^T(\tilde{A}, \tilde{A}^T)^{-1}\tilde{A} \equiv \begin{bmatrix} P_{SS} & P_{SN} \\ P_{NS} & P_{NN} \end{bmatrix}$$ \hfill (4.13)

where $P_{SS} \in \mathbb{R}^{n \times m}$, $P_{SN} \in \mathbb{R}^{n \times (n-m)}$, $P_{NS} \in \mathbb{R}^{(n-m) \times m}$, and $P_{NN} \in \mathbb{R}^{(n-m) \times (n-m)}$. Then there exist $\hat{\mu} > 0$ and constants $C_6, C_7, C_8, C_9 > 0$ such that for $0 < \mu \leq \hat{\mu}$ and $(x, z) \in B((x^*_\mu, z^*_\mu); \delta_\mu)$, where $\delta_\mu$ is given in Lemma 4.1, it follows that

$$\|P_{SS} - I_m\| \leq C_6 \mu^2. \quad \|P_{SN}\| \leq C_7 \mu$$

$$\|P_{NS}\| \leq C_8 \mu. \quad \|P_{NN}\| \leq C_9 \mu^2.$$ \hfill (4.14)
Proof Consider \(0 < \mu \leq \hat{\mu}\) where \(\hat{\mu}\) is given by Lemma 4.1. Let \(W_8 = (X_8^{1/2}Z_8^{-1/2})\) and \(W_N = (X_N^{1/2}Z_N^{-1/2})\). Note that \(W_8\) and \(W_N\) are nonsingular. Let \(\tilde{A} \equiv AW = [A_8W_8 \ A_NW_N]\). Then substituting \(\tilde{A}\) in the definition of \(P\) we obtain

\[
P = [A_8W_8 \ A_NW_N]^T (A_8(W_8)^2A_8^T + A_N(W_N)^2A_N^T)^{-1} [A_8W_8 \ A_NW_N].
\]

Now, introduce the \(m \times (n - m)\) matrix \(R\) where

\[
R = (W_8)^{-1}A_8^{-1}A_NW_N.
\]

(4.15)

Then \(P\) can be partitioned as follows

\[
P = \begin{bmatrix}
(I_m + RR^T)^{-1} & (I_m + RR^T)^{-1}R \\
R^T(I_m + RR^T)^{-1} & R^T(I_m + RR^T)^{-1}R
\end{bmatrix}.
\]

(4.16)

Applying the bounds in (4.7) to (4.15), we obtain \(\|R\| \leq C_5\mu\) for a constant \(C_5 > 0\).

Since \(\|RR^T\| = O(\mu^2)\), then \(\|RR^T\| \to 0\) as \(\mu \to 0\). Then there exists \(\hat{\mu} > 0\) and \(\hat{\mu} \leq \hat{\mu}\) such that for all \(\mu \leq \hat{\mu}\) we obtain \(\|RR^T\| < 1\). Using the Neumann series expansion on \((I_m + RR^T)^{-1}\), we obtain from (4.16) and (4.13) that for \(\mu \leq \hat{\mu}\) and constants \(C_6, C_7, C_8, C_9 > 0\) we have

\[
\|P_{88} - I_m\| \leq C_6\mu^2, \quad \|P_{8N}\| \leq C_7\mu
\]

and

\[
\|P_N8\| \leq C_8\mu, \quad \|P_{NN}\| \leq C_9\mu^2.
\]

\(\square\)
Lemma 4.4  Let the nondegeneracy assumption hold. Also, let \( \bar{A} \equiv A \bar{X} \). Define

\[
P = \bar{A}^T (\bar{A} \bar{A}^T)^{-1} \bar{A} \equiv \begin{bmatrix} P_{BB} & P_{BN} \\ P_{NB} & P_{NN} \end{bmatrix}
\]

(4.17)

where \( P_{BB} \in \mathbb{R}^{n \times m} \), \( P_{BN} \in \mathbb{R}^{n \times (n-m)} \), \( P_{NB} \in \mathbb{R}^{(n-m) \times m} \), and \( P_{NN} \in \mathbb{R}^{(n-m) \times (n-m)} \).

Then there exist \( \bar{\mu} > 0 \) and constants \( G_3, G_4, G_5, G_6 > 0 \) such that for \( 0 < \mu \leq \bar{\mu} \) and \( x \in B(x^*_0; \delta_\mu) \), where \( \delta_\mu \) is given in Lemma 4.2, it follows that

\[
\begin{align*}
\|P_{BB} - I_m\| & \leq G_3 \mu^2, \\
\|P_{BN}\| & \leq G_4 \mu, \\
\|P_{NB}\| & \leq G_5 \mu, \\
\|P_{NN}\| & \leq G_6 \mu^2.
\end{align*}
\]

(4.18)

Proof  The proof is similar to that of Lemma 4.3 so we omit the proof. \( \square \)

4.4.2 Barrier System

We provide a sharp result proving that the radius of the sphere of convergence of Newton’s method applied to the barrier system (4.3) decreases to zero with the same order that \( \mu \) decreases to zero. Our result follows from showing that a lower-bound and an upper-bound of order \( \mu \) exist for the radius of the sphere of convergence.

Lemma 4.5  Under the nondegeneracy assumption, there exist \( \bar{\mu} > 0 \) and constant \( K_1 > 0 \) such that for any \( \mu \leq \bar{\mu} \) the radius of the sphere of convergence, \( r_B(\mu) \), of Newton’s method applied to the barrier system (4.3) satisfies

\[
K_1 \mu \leq r_B(\mu).
\]

Proof  We will prove the above result by showing that the sequence of Newton iterates converges to the solution \( (x^*_\mu, y^*_\mu) \) if the initial point \( x^0 \) satisfies

\[
\|x^0 - x^*_\mu\| \leq K_1 \mu.
\]

(4.19)
Consider \( \hat{\mu} \) given in Lemma 4.4. Suppose Newton's method is applied to the barrier system (4.3) for a particular value of \( \mu \leq \hat{\mu} \). Denote \((x, y)\) as the current Newton iterate where \( x \in B(x^*_\mu, \delta_\mu) \) and \( x \) satisfies the conditions given in (4.12). Now, consider the next Newton iteration to obtain

\[
\begin{pmatrix}
    x^+ \\
    y^+
\end{pmatrix} = \begin{pmatrix}
    x \\
    y
\end{pmatrix} - F'_B(x, y; \mu)^{-1} F_B(x, y; \mu).
\]

Subtracting \( x^*_\mu \) from both sides of the equation, noting that \( F_B(x^*_\mu, y^*_\mu; \mu) = 0 \) and evaluating \( F_B \) at the given points, we obtain

\[
\begin{pmatrix}
    x^+ - x^*_\mu \\
    y^+ - y^*_\mu
\end{pmatrix} = F'_B(x, y; \mu)^{-1} \left[ F_B(x^*_\mu, y^*_\mu; \mu) - F_B(x, y; \mu) - F'_B(x, y; \mu)(x^*_\mu - x, y^*_\mu - y) \right]
\]

\[
= F'_B(x, y; \mu)^{-1} \left[ \begin{array}{c}
    \mu (X^*_\mu)^{-1} \epsilon - \mu X^{-1} \epsilon + \mu X^{-2} (x^*_\mu - x) \\
    0
\end{array} \right]. \quad (4.21)
\]

By Taylor's Theorem,

\[
(X^*_\mu)^{-1} \epsilon = X^{-1} \epsilon - X^{-2} (x^*_\mu - x) + \hat{X}^{-3} (x - x^*_\mu) \quad \text{or} \quad \hat{X}^{-3} (x - x^*_\mu)^2 = (X^*_\mu)^{-1} \epsilon - X^{-1} \epsilon + X^{-2} (x^*_\mu - x) \quad (4.22)
\]

for some \( \hat{x}_i \in [\min\{(x^*_\mu)_i, x_i\}, \max\{(x^*_\mu)_i, x_i\}], i = 1 \ldots n \). Substitute (4.22) into (4.21) to obtain

\[
\begin{pmatrix}
    x^+ - x^*_\mu \\
    y^+ - y^*_\mu
\end{pmatrix} = F'_B(x, y; \mu)^{-1} \left[ \begin{array}{c}
    \mu \hat{X}^{-3} (x - x^*_\mu)^2 \\
    0
\end{array} \right]. \quad (4.23)
\]
where

\[
F_B'(x, y; \mu)^{-1} = \begin{bmatrix}
\frac{1}{\mu}[X^2A^T(AX^2A^T)^{-1}AX^2 - X^2] & X^2A^T(AX^2A^T)^{-1} \\
(AX^2A^T)^{-1}AX^2 & \mu(AX^2A^T)^{-1}
\end{bmatrix}.
\]

Making the above substitution for \((F_B')^{-1}\) in (4.23) and multiplying the right-hand side of (4.23) we obtain

\[
\begin{bmatrix}
(x^+ - x^*_\mu) \\
y^+ - y^*_\mu
\end{bmatrix} = \begin{bmatrix}
X[AX^2A^T(AX^2A^T)^{-1}AX - I]X\hat{x}^{-3}(x - x^*_\mu)^2 \\
\mu(AX^2A^T)^{-1}AX^2\hat{x}^{-3}(x - x^*_\mu)^2
\end{bmatrix}.
\]

Now substitute in the definition of \(P\) given by (4.17) to obtain

\[
\begin{bmatrix}
(x^+ - x^*_\mu) \\
y^+ - y^*_\mu
\end{bmatrix} = \begin{bmatrix}
X(P - I)X\hat{x}^{-3}(x - x^*_\mu)^2 \\
\mu(AX^2A^T)^{-1}AX^2\hat{x}^{-3}(x - x^*_\mu)^2
\end{bmatrix}.
\]

We will consider first the vector \((x^+ - x^*_\mu)\) in (4.24). If we partition \((x^+ - x^*_\mu)\) into its basic and nonbasic components and use the notation for \(P\) in (4.17), then

\[
\begin{align*}
(x^+ - x^*_\mu)_B &= X_B(P_{BB} - I_m)_BX_B\hat{x}^{-3}_B(x - x^*_\mu)_B^2 + X_BP_{BN}X_N\hat{x}^{-3}_N(x - x^*_\mu)_N^2 \\
(x^+ - x^*_\mu)_N &= X_NP_{NB}X_B\hat{x}^{-3}_B(x^*_\mu - x)_B^2 + X_N(P_{NN} - I_m)_NX_N\hat{x}^{-3}_N(x^*_\mu - x)_N^2
\end{align*}
\]

which leads to

\[
\begin{align*}
\|x^+ - x^*_\mu\|_B &\leq (\|X_B(P_{BB} - I_m)_B\|X_B\hat{x}^{-3}_B\| + \|X_BP_{BN}X_N\hat{x}^{-3}_N\|)\|x - x^*_\mu\|^2 \\
\|x^+ - x^*_\mu\|_N &\leq (\|X_NP_{NB}X_B\|\hat{x}^{-3}_B\| + \|X_N(P_{NN} - I_m)_N\|X_N\hat{x}^{-3}_N\|)\|x - x^*_\mu\|^2.
\end{align*}
\]

Applying the bounds given in (4.18) and (4.12), we obtain
\[ ||(x^+ - x^*_\mu)|| \leq C \frac{1}{\mu} ||x - x^*_\mu||^2. \]  
(4.25)

\[ ||(x^+ - x^*_\mu, s)|| \leq C \frac{1}{\mu} ||x - x^*_\mu||^2 \]

for some constant \( C > 0 \). Since \( ||x - x^*_\mu|| \leq \delta_\mu \) then using (4.25), the \( x \) portion of the Newton sequence will converge to \( x^*_\mu \) if the initial Newton point \( x^0 \) satisfies

\[ ||x^0 - x^*_\mu|| \leq \min\{\delta_\mu, \frac{1}{C\mu}\}. \]  
(4.26)

Now, consider the remaining \( m \) components of (4.24). Taking norms and partitioning matrices, we obtain

\[ \|y^+ - y^*_\mu\| \leq \|\mu(A_B X_B^2 A_B^T + A_N X_N^2 A_N^T)^{-1}(A_B X_B^2 \hat{\chi}_B^3 + A_N X_N^2 \hat{\chi}_N^3)\| \|x - x^*_\mu\|^2\|. \]

Applying (4.12), we obtain

\[ \|y^+ - y^*_\mu\| \leq \mu \left[ O(1) + O(\mu^2) \right]^{-1} \left[ O(1) + O(1/\mu) \right] \|x - x^*_\mu\|^2. \]

Now, for a constant \( \hat{C} > 0 \), we obtain

\[ \|y^+ - y^*_\mu\| \leq \hat{C} \|x - x^*_\mu\|^2. \]  
(4.27)

Thus, the \( y \) portion of the Newton sequence converges to \( y^*_\mu \) if (4.26) holds. Hence, using (4.26) and (4.27), for \( \mu \leq \hat{\mu} \) and a constant \( \hat{K}_1 > 0 \). Newton’s method is guaranteed to converge to \((x^*_\mu, y^*_\mu)\) if the initial iterate \( x^0 \) satisfies

\[ ||x^0 - x^*_\mu|| \leq \hat{K}_1 \mu. \]  
(4.28)
The above lemma shows that the radius of the sphere of convergence of Newton’s method applied to the barrier system (4.3) satisfies $K_1 \mu \leq r_B(\mu)$. It establishes only a lower-bound result for the radius of the sphere of convergence. To establish that the radius of the sphere of convergence decreases to zero at exactly the same rate as $\mu$ decreases to zero, we need an upper-bound of the same order. The following lemma establishes such an upper-bound.

**Lemma 4.6** Consider Newton’s method applied to the barrier system (4.3). Then there exist constants $\hat{\mu} > 0$ and $K_2 > 0$ such that for any given $\mu \leq \hat{\mu}$, the radius of the sphere of convergence, $r_B(\mu)$, corresponding to this $\mu$ satisfies

$$r_B(\mu) \leq K_2 \mu.$$  

**Proof** It suffices to show the existence of a point $x \geq 0$ with $\|x - x^*_\mu\| \leq K_2 \mu$ from which Newton’s method does not converge or is not defined. From Lemma 4.2 there exist $\hat{\mu} > 0$ and constant $K_2 > 0$ such that for $\mu \leq \hat{\mu}$ and for $i \in \mathcal{N}$, $(x^*_\mu)_i \leq K_2 \mu$. Consider an $i \in \mathcal{N}$, and let

$$x = x^*_\mu - (x^*_\mu)_i e_i,$$

where $e_i$ is the vector with 1 in the $i$th component and zeros elsewhere. Obviously, $\|x - x^*_\mu\| \leq K_2 \mu$. Moreover, because $x_i = 0$. Newton’s method is not defined at $x$. Therefore, $r_B(\mu) \leq K_2 \mu$. \qed

**Theorem 4.1** There exist constants $\hat{\mu} > 0$ and $K_1, K_2 > 0$ such that for $\mu \leq \hat{\mu}$, the radius of the sphere of convergence, $r_B(\mu)$, of Newton’s method applied to the barrier system (4.3) satisfies

$$K_1 \mu \leq r_B(\mu) \leq K_2 \mu.$$
Proof An application of Lemma 4.5 and Lemma 4.6 produces the result.

4.4.3 Perturbed System

We now provide a lower-bound result for the radius of the sphere of convergence of Newton's method applied to the perturbed system (4.4). We show that the lower-bound is independent of the value of \( \mu \). This result establishes that the radius of the sphere of convergence is bounded away from zero as \( \mu \) decreases to zero.

**Theorem 4.2** Under the nondegeneracy assumption, there exist constants \( D > 0 \) and \( \bar{\mu} > 0 \) such that for any \( \mu \leq \bar{\mu} \), the radius of the sphere of convergence, \( r_p(\mu) \), of Newton's method applied to the perturbed system (4.4) satisfies

\[
D \leq r_p(\mu).
\]

**Proof** We will show that Newton's method applied to the perturbed system (4.4) generates iterates that converge to the solution \((x^*_n, y^*_n, z^*_n)\) if the initial point \((x^0, z^0)\) satisfies

\[
\left\| \begin{pmatrix} x^0 - x^*_\mu \\ z^0 - z^*_\mu \end{pmatrix} \right\| \leq D, \tag{4.29}
\]

which then shows that \( r_p(\mu) \geq D > 0 \). Consider \( \mu \leq \bar{\mu} \) where \( \bar{\mu} \) is given by Lemma 4.3. At a given value of \( \mu \), let \((x, y, z)\) and \((x^*_\mu, y^*_\mu, z^*_\mu)\) denote respectively the current iterate and the solution for system (4.4). We have by Lemma 4.1 that \((x, z) \in B((x^*_\mu, z^*_\mu); \delta_\mu)\) and satisfies the conditions given in (4.7).

At the subsequent iteration, the Newton iterates are of the form

\[
\begin{pmatrix} x^+ \\ y^+ \\ z^+ \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - F'_p(x, y, z; \mu)^{-1} F_p(x, y, z; \mu).
\]
Subtracting \((x^*_\mu, y^*_\mu, z^*_\mu)\) from both sides of the equation above and noting that 
\[ F_P(x^*_\mu, y^*_\mu, z^*_\mu; \mu) = 0, \]
we obtain

\[
\begin{pmatrix}
x^+ - x^*_\mu \\
y^+ - y^*_\mu \\
z^+ - z^*_\mu
\end{pmatrix} = F'_P(x, y, z; \mu)^{-1} \left[ F_P(x^*_\mu, y^*_\mu, z^*_\mu; \mu) - F_P(x, y, z; \mu) \right]
\]

\[ - F'_P(x, y, z; \mu) \{(x^*_\mu, y^*_\mu, z^*_\mu) - (x, y, z)\} \]

\[ = F'_P(x, y, z; \mu)^{-1} \begin{bmatrix}
0 \\
0 \\
(X - X^*_\mu)(Z - Z^*_\mu)e
\end{bmatrix} \quad (4.30)\]

\[ = \begin{bmatrix}
Z^{-1} \left[ I_n - X.A^T/AZ^{-1}.A^T/AZ^{-1} \right] (X - X^*_\mu)(Z - Z^*_\mu)e \\
- (AZ^{-1}.X.A^T)AZ^{-1}(X - X^*_\mu)(Z - Z^*_\mu)e \\
A^T(AZ^{-1}.X.A^T)AZ^{-1}(X - X^*_\mu)(Z - Z^*_\mu)e
\end{bmatrix} \]

where the last equality is obtained by multiplying the right-hand side of (4.30).

First, we consider convergence in the \(x\) variables. Let \(W = Z^{-\frac{1}{2}}X^\frac{1}{2}\) be a diagonal matrix, then we obtain

\[
(x^+ - x^*_\mu) = W \left[ I_n - W.A^T/AW^{-2}A^T \right] W.X^{-1}(X - X^*_\mu)(Z - Z^*_\mu)e. \quad (4.31)
\]

Define \(P = (W.A^T/AW^{-2}A^T)^{-1}.AW\) where \(P\) has the submatrix representation given in (4.13). If we partition the vectors and matrices in (4.31) into their basic and nonbasic components, we obtain
\[
\begin{bmatrix}
(x^+ - x^*_\mu)_B
\end{bmatrix}
= \begin{bmatrix}
W_B(I_m - P_{BG})W_B X_B^{-1} & W_B P_{BV} W_N X_N^{-1} \\
W_N P_{NV} W_B X_B^{-1} & W_N(I_n-m - P_{BV}) W_N X_N^{-1}
\end{bmatrix}
\begin{bmatrix}
(X - X^*_\mu)_B(z - z^*_\mu)_B \\
(X - X^*_\mu)_N(z - z^*_\mu)_N
\end{bmatrix}.
\]

Now, consider the basic components of \((x^+ - x^*_\mu)_B\). Substitute for \(W\) to obtain
\[
(x^+ - x^*_\mu)_B = Z_B^{-1/2} X_B^{1/2} (I_m - P_{BG}) Z_B^{1/2} X_B^{-1/2} X_N^{-1} (X - X^*_\mu)_B (z - z^*_\mu)_B
+ Z_B^{-1/2} X_B^{1/2} P_{BV} Z_N^{-1/2} X_N^{1/2} X_N^{-1} (X - X^*_\mu)_N (z - z^*_\mu)_N
\]
and take the norm of \((x^+ - x^*_\mu)_B\) to obtain
\[
\| (x^+ - x^*_\mu)_B \| \leq (\| Z_B^{-1/2} X_B^{1/2} (I_m - P_{BG}) Z_B^{1/2} X_B^{-1/2} X_N^{-1} \| + \| Z_B^{-1/2} X_B^{1/2} P_{BV} Z_N^{-1/2} X_N^{1/2} X_N^{-1} \|) \| (X - X^*_\mu)_B (z - z^*_\mu) \|.
\]
Apply bounds (4.7) and (4.14) to the above inequality. Then for some constants \(D_1, D_2, D_3 > 0\), we obtain
\[
\| (x^+ - x^*_\mu)_B \| \leq (D_1 \mu + D_3) \| (X - X^*_\mu)_B (z - z^*_\mu) \|
\leq D_3 \max\{\| x - x^*_\mu \|^2, \| z - z^*_\mu \|^2\}. \quad (4.32)
\]

If we do a similar analysis for \((x^+ - x^*_\mu)_N\), then
\[
(x^+ - x^*_\mu)_N = Z_N^{-1/2} X_N^{1/2} P_{NV} Z_B^{-1/2} X_B^{1/2} X_N^{-1} (X - X^*_\mu)_B (z - z^*_\mu)_B
+ Z_N^{-1/2} X_N^{1/2} (I_n-m - P_{BV}) Z_N^{-1/2} X_N^{1/2} X_N^{-1} (X - X^*_\mu)_N (z - z^*_\mu)_N.
\]

And hence for constants \(D_4, D_5, D_6, D_7 > 0\), we obtain
\[
\| (x^+ - x^*_\mu)_N \| \leq (D_1 \mu + D_5 + D_6 \mu) \max\{\| x - x^*_\mu \|^2, \| z - z^*_\mu \|^2\}
\leq D_7 \max\{\| x - x^*_\mu \|^2, \| z - z^*_\mu \|^2\}. \quad (4.33)
\]
Combining (4.32) and (4.33) yields

\[ \|x^+ - x^*_\mu\| \leq D_8 \max\{\|x - x^*_\mu\|^2, \|z - z^*_\mu\|^2\}. \]  \hspace{1cm} (4.34)

where \( D_8 = \max\{D_3, D_7\} \). Through a similar argument on \((y^+ - y^*_\mu)\) and \((z^+ - z^*_\mu)\), we obtain for some constants \( D_9, D_{10} > 0 \) that

\[ \|y^+ - y^*_\mu\| \leq D_9 \max\{\|x - x^*_\mu\|^2, \|z - z^*_\mu\|^2\} \]  \hspace{1cm} (4.35)

and

\[ \|z^+ - z^*_\mu\| \leq D_{10} \max\{\|x - x^*_\mu\|^2, \|z - z^*_\mu\|^2\}. \]  \hspace{1cm} (4.36)

It follows from (4.34) and (4.36) that if the initial iterate \((x^0, z^0)\) satisfies

\[ \|x^0 - x^*_\mu\| \leq 1/D_8 \]

and

\[ \|z^0 - z^*_\mu\| \leq 1/D_{10} \]

then Newton's method converges to the solution \((x^*_\mu, y^*_\mu, z^*_\mu)\). Thus, a lower bound estimate for the radius of the sphere of convergence is

\[ \left\| \begin{pmatrix} x^0 - x^*_\mu \\ y^0 - y^*_\mu \\ z^0 - z^*_\mu \end{pmatrix} \right\| \leq D \text{ for } \mu \leq \tilde{\mu} \]  \hspace{1cm} (4.37)

where \( D = \min\{1/D_8, 1/D_{10}\} \).

The analysis shows that the radius of the sphere of convergence of Newton's method applied to the perturbed system (4.4) is independent of \( \mu \) and stays bounded.
away from zero as $\mu$ decreases to zero. This result indicates that the sphere of convergence of the perturbed system must eventually be larger than the sphere of convergence associated with the barrier system (4.3), at least for small $\mu$ values, so that $\frac{r_B(\mu)}{r_P(\mu)} \leq 1$. In the next section, we show numerically that this is indeed the case.

4.5 Numerical Experiments Concerning the Sphere of Convergence

In Section 4.4, we provided bounds on the radii of the spheres of convergence of Newton's method applied to the barrier (4.3) and the perturbed (4.4) systems under the nondegeneracy assumption. The analysis shows that at least for small values of $\mu$, the sphere of convergence associated with the perturbed system is larger than that for the barrier system. This result arises from the fact that the barrier system is not well-defined at $\mu = 0$. Therefore, as $\mu$ decreases to zero, we expect the radius of the sphere of convergence to decrease to zero. However, it is not clear what will occur if the half-sphere of convergence for the barrier system is considered. The half-sphere of convergence contains only points $x > 0$ from where Newton's method converges.

In this section, we will obtain numerical upper-bound estimates on the radius of the half-sphere of convergence for Newton's method applied to the barrier system and on the radius of the sphere of convergence for Newton's method applied to the perturbed system.

The upper-bound estimates are based on the following simple idea. In order to apply Newton's method with the same initial point to the barrier and perturbed systems, the initial points were selected so that the corresponding equations in the barrier and perturbed systems produced the same residual, and the complementarity equation was satisfied for the perturbed system. So let $x_0 \in \mathbb{R}^c$ be an arbitrary
vector with norm one and $\lambda > 0$ be a scalar. Consider applying Newton’s method to the barrier and perturbed systems starting from initial points of the form

\[ x^0 = x^*_\mu + \lambda x_\alpha, \quad y^0 = 0. \]  

(4.38)

and $z^0 = \mu(X^0)^{-1}e$ for the perturbed system. If for $\lambda = \lambda_0 > 0$, Newton’s method does not converge to $v^*_\mu = (x^*_\mu, y^*_\mu, z^*_\mu)$ for the perturbed system (and to $v^*_\mu = (x^*_\mu, y^*_\mu, \mu(X^*_\mu)^{-1}e)$ for the barrier system), then obviously $\lambda_0$ is an upper-bound for the radius of the sphere of convergence of Newton’s method at $v^*_\mu$. This upper-bound is the tightest possible one in this particular direction if Newton’s method converges to $v^*_\mu$ for any $\lambda \in (0, \lambda_0)$. Numerically, this upper-bound $\lambda_0$ can be approximated by gradually increasing $\lambda$ from zero by a small increment until Newton’s method fails to converge. A sharper upper-bound can be generated by calculating $\lambda_0$ for a set of random vectors $\{x_\alpha\}$ with norm one and then taking $\min, \{\lambda_0\}$ as an upper-bound.

Under the nondegeneracy assumption, Newton’s method is well-defined for the perturbed system in a neighborhood of the solution to the linear program. Since some components of $x^*$ and $z^*$ may be zero, then a ball exists centered at $(x^*, y^*, z^*)$, corresponding to $\mu = 0$, which contains points $x$ and $z$ whose components may be negative. Therefore, $x_\alpha$ can be chosen to be any random vector. For this purpose, we use the Matlab function `randn` to select ten random vectors $x_\alpha$ having norm one.

As we mentioned earlier, because of the presence of the term $X^{-1}$, the barrier system is not well-defined for $\mu = 0$, nor is Newton’s method at $\mu = 0$. This implies that the radius of the sphere of convergence of Newton’s method shrinks to zero as $\mu$ decreases to zero. However, it is not clear at all that the largest half-sphere inside the positive orthant should also shrink to zero as $\mu$ decreases to zero. To be fair to the barrier system, only positive random vectors $x_\alpha$ are used. In this way, we actually obtain upper-bound estimates for the radius of the half-sphere of convergence instead
of the sphere of convergence. For this purpose, ten random vectors \( x_\alpha \) of norm one are selected using the Matlab function \texttt{rand} to ensure \( x_\alpha > 0 \).

To observe the behavior of the radii of the half-sphere of convergence for the barrier system and for the sphere of convergence for the perturbed system as \( \mu \) decreases to zero, the numerical procedure described above was implemented for various values of \( \mu > 0 \):

\[
\mu = 50, \ 25, \ 1, \ 0.45, \ 0.25, \ 0.10, \ 0.05, \ 0.01, \ 0.0075, \ 0.005, \ 0.0005, \ 0.00005. \quad (4.39)
\]

The parameter \( \lambda \) given in (4.38) was given an initial value of \( 10^{-10} \) and was incremented when the convergence criteria

\[
\frac{||v^*_\mu - v^k||}{||v^*_\mu||} < tol
\]

was satisfied at some iteration \( k \), where \( v^k = (x^k, y^k, \mu(X^k)^{-1}e) \) for the barrier system and \( v^k = (x^k, y^k, z^k) \) for the perturbed system. Nonconvergence was recorded for a particular system with a given \( \mu \) value and initial point of the form given in (4.38) if the maximum number of iterations, which we set to 50, was reached. The convergence tolerance was set to \( tol = 10^{-8} \). The numerical solution \( v^*_\mu \) was obtained by solving system (4.1) for a given value of \( \mu \) in (4.39).

An upper-bound estimate for the radius of the half-sphere (sphere) of convergence associated with the barrier system (perturbed system) was determined in terms of the primal variables \( x \) only. We eliminated recording any information concerning the contribution of the \( z \) variables for the perturbed system. Therefore, a numerical upper-bound estimate for the radius of the half-sphere of convergence associated with the barrier system and for the radius of the sphere of convergence associated with the perturbed system was recorded as

\[
\min_{\alpha} \{ \lambda_\alpha \} \text{ for } \lambda_\alpha = ||x^0 - x^*_\mu||.
\]
We emphasize that in these experiments the pure Newton's method was used, that is, unit steplength was always taken.

Test problems consisted of six random nondegenerate problems r1-r6, the Netlib nondegenerate problems: scagr7, sc50b, share1b and the Netlib degenerate problems: adlittle, afinity, blend, sc50a, and share2b. The random data was generated from a uniform distribution on the interval (0,1). For a given problem, the same ten unit vectors \( x_\alpha \) were used for all values of \( \mu \) in (4.39). The problems were run on a Sun Ultra Sparc workstation using Matlab version 5.1. Test problem dimensions can be found in Table 4.1.

In our experiments, we observed that negative components in the iterates \((x^k, z^k)\) for the perturbed system did not preclude convergence; on the other hand, for the barrier system negative components in \( x^k \) always led to nonconvergence. The former phenomenon is not a surprise; however the latter was not completely expected.

4.5.1 Nondegenerate Problems

Experiments were performed on the set of six random nondegenerate problems and the three Netlib nondegenerate problems. Figures 4.1 and 4.2 show the radii of the half-sphere of convergence associated with the barrier system (4.3) and the sphere of convergence associated with the perturbed system (4.4) graphed against the values of \( \mu \) given in (4.39). Figure 4.1 contains the graph for a random problem, and the remaining graphs show results for the Netlib problems. The results show that the radius of the sphere of convergence of Newton's method on the perturbed system is bounded away from zero even for \( \mu \) sufficiently small, but the radius of the half-sphere of convergence of Newton's method on the barrier system decreases to zero as \( \mu \) decreases to zero. Furthermore, the tests show a larger radius for the sphere of convergence of Newton's method on the perturbed system than on the barrier
system as $\mu$ decreases to zero. In some instances, the radius for the barrier system is larger than that for the perturbed system for some values of $\mu$ greater than one. This phenomenon was observed in four of the random problems and in problem sc50b-

Figure 4.2a, where the two radii slightly differ.

For the two equivalent systems, we observed that if Newton's method failed to converge for an initial point $v^0$ with parameter $\lambda_n$, the final Newton iterate had negative components. For the barrier system, if any components of $x^k$ were negative the Jacobian matrix was highly ill-conditioned and convergence was precluded for Newton's method.

![Random problem 12](image1.png)

![Netlib Nondegenerate scagr7](image2.png)

(a) (b)

**Figure 4.1** Radii of the half-sphere and sphere of convergence for Newton's method applied to the barrier and perturbed systems

### 4.5.2 Degenerate Problems

In addition, we investigated the numerical behavior of the radius of the half-sphere and sphere of convergence, respectively, of Newton's method applied to the barrier
Figure 4.2 Radii of the half-sphere and sphere of convergence for Newton’s method applied to the barrier and perturbed systems.

Figure 4.3 Radii of the half-sphere and sphere of convergence for Newton’s method applied to the barrier and perturbed systems.
Figure 4.4  Radii of the half-sphere and sphere of convergence for
Newton's method applied to the barrier and perturbed systems
(4.3) and perturbed (4.4) systems for degenerate problems for decreasing values of
\( \mu > 0 \). No theory is provided for the behavior of the radii of the spheres of convergence
of Newton's method for degenerate problems. The same values of \( \mu \) given in (4.39)
were used. Numerical results are shown in Figures 4.3 and 4.4.

The numerical experiments indicate that the radius of the half-sphere of con-
vergence of Newton's method applied to the barrier system decreases to zero as \( \mu 
\) approaches zero, as in the case with the nondegenerate problems. Contrary to the
results obtained for the nondegenerate problems, the radius of the sphere of convergence
associated with the perturbed system decreases to zero as \( \mu \) decreases to zero
for the degenerate problems. However, the radius associated with the perturbed sys-
tem stays well above that for the barrier system, by at least an order of magnitude
of ten, as \( \mu \) decreases to zero.

When Newton's method failed to converge, the final Newton iterate contained
negative components. Also, if any component of the iterate \( x^k \) became negative for
the barrier system, the Jacobian became highly ill-conditioned and convergence of Newton's method was precluded.

4.6 Efficiency of an Interior-Point Method

The previous two sections investigated differences in the behavior of the radius of the sphere of convergence of Newton's method applied to the barrier (4.3) and perturbed (4.4) systems as $\mu$ decreased to zero. The notion of the sphere/half-sphere of convergence is a local feature of Newton's method. As such, our theoretical and numerical investigations were both performed for fixed $\mu$ values, even though we did consider the behavior of the radius of the sphere of convergence as $\mu$ approached zero. Now, we study the global behavior of a Newton log-barrier function method and a Newton primal-dual interior-point method.

Section 2.5 described the Newton log-barrier function method and the Newton primal-dual interior-point method for the nonlinear program. The same concepts extend when we apply these two methods for solving the linear program. The Newton log-barrier method can be viewed as applying Newton's method to the barrier system. Similarly, the Newton primal-dual method can be viewed as applying Newton's method to the perturbed system. To keep the variables $x$ and ($x, z$) strictly positive for system (4.3) and system (4.4), respectively, the Newton step is damped. Then $\mu$ is decreased and a new nonlinear system is solved. This procedure is continued for decreasing values of $\mu$ until an approximate solution for the linear program is obtained.

A very simple interior path-following algorithm is given below as Algorithm IPF. The algorithm can be viewed as a Newton log-barrier method for the linear program. In the application of Algorithm IPF, we let $F(v^k; \mu)$ be defined as $F_B(v^k; \mu)$ where
\( v^k = (x^k, y^k, z^k) \). Similarly, Algorithm IPF becomes a Newton primal-dual method for the linear program when we select \( F(v^k; \mu) \sim F_p(v^k; \mu) \) where \( v^k = (x^k, y^k, z^k) \).

We study the global behavior of a Newton log-barrier method and a Newton primal-dual method. Since damped Newton is the underlining method of Algorithm IPF, any performance discrepancy in Algorithm IPF should be due to the different behavior of the Newton log-barrier method and the Newton primal-dual method. In our numerical experiments, we will pay particular attention to the efficiency of Algorithm IPF when the iterates are required to follow the central path loosely or closely.

Algorithm IPF is not a particularly efficient algorithm for a Newton primal-dual method since there exist more efficient interior-point path-following algorithms, for example, Kojima, Mizuno, Yoshise’s primal-dual interior-point algorithm [12] and Mehrotra’s predictor-corrector algorithm [15]. For a Newton log-barrier method, however, it has not been shown that one can do significantly better than Algorithm IPF.

**Algorithm IPF** Let \( F(v^k; \mu) \equiv F_B(x^k, y^k; \mu) \) or \( F(v^k; \mu) \equiv F_p(x^k, y^k, z^k; \mu) \). Given initial iterate \( v^0 \) with positive components corresponding to the nonnegative variables, and parameters \( \mu, \mu_{tol} > 0, \alpha^k, \tau_{tol} \in (0, 1], \) and \( \sigma, \delta(\mu) \in (0, 1) \).

While \( \| F(v^k; \mu) \| + \mu > \tau_{tol} \)

1) Solve for \( \Delta v^k \) from \( F'(v^k; \mu) \Delta v^k = -F(v^k; \mu) \)

2) Form the new iterate
\[
 v^{k+1} \leftarrow v^k + \alpha^k \Delta v^k
\]

3) Update \( \mu \) and \( \mu_{tol} \)
\[
\text{if } \| F(v^{k+1}; \mu) \| < \mu_{tol}
\]
\[
\mu \leftarrow \sigma \mu
\]
\[
\mu_{tol} \leftarrow \delta(\mu)
\]
end

4) Increment iteration count

\[ k \leftarrow k + 1 \]

End

The algorithm's objective is to decrease \( \mu \) to zero and to guide the iterates along the central path to a solution for the linear program. The Newton step is damped to keep the primal iterates, \( x^k \), greater than zero for system (4.3) and \( x^k, z^k \) greater than zero for system (4.4); otherwise a unit step length is taken. Damping is performed by selecting \( \alpha_k \) to be a given fraction of the step length to the boundary of the positive orthant from the point \( v^k \) along the direction \( \Delta v^k \). The parameter \( \mu \) and possibly the tolerance \( \mu_{tol} \) are updated once an iterate lies within a given neighborhood of the solution \( v^*_\mu \) for the system \( F(v; \mu) = 0 \), which is described by \( \| F(v^{k+1}; \mu) \| < \mu_{tol} \). When this update occurs, \( \mu \) is decreased to \( \mu^+ \) and another nonlinear system is solved having \( \mu^+ \) as its parameter. As a consequence of decreasing \( \mu_{tol} \), the subsequent iterates will follow the central path more closely. An approximate solution to the linear program is one that lies close to the solution \( (x^*, y^*, z^*) \), corresponding to \( \mu = 0 \) for system (4.4). In our algorithm, an approximate solution is obtained when a convergence test is satisfied, for example, when the sum of \( \mu \) and the Euclidean norms of the residuals of \( F(v^k; \mu) \) fall below a tolerance that is close to zero.

4.6.1 Numerical Results

We study some properties associated with the Newton log-barrier method and the Newton primal-dual method. Our study focuses on the role of the centrality tolerance, \( \mu_{tol} \), as well as the effect of ill-conditioning on the two methods. A Newton log-barrier and Newton primal-dual method are applied to a set of test problems. Depending on the value of \( \mu_{tol} \) and on the ill-conditioning of the nonlinear system, the iterates
may not converge to a solution of prescribed accuracy, as will be demonstrated. We remark again that Algorithm IPF can be viewed as a Newton log-barrier method for the linear program if \( F(v^k; \mu) \equiv F_B(x^k, y^k; \mu) \). Similarly, it can be viewed as a Newton primal-dual algorithm when we define \( F(v^k; \mu) \) as \( F_P(x^k, y^k, z^k; \mu) \).

Now, we explain the parameters and tolerances used in Algorithm IPF. As mentioned earlier, \( \| F(v^k; \mu) \| \) denotes the sum of the Euclidean norms of the residuals of \( F(v^k; \mu) \). Let \( \sigma \) denote the amount of reduction in the updated value of \( \mu \), and let \( \sigma = 0.2 \). To keep the new iterates positive, the following steplength calculation of \( \alpha^k \) is chosen

\[
\alpha^k = \min(0.95, \frac{-1}{\min((I^k)^{-1}\Delta u^k, -0.5)}), 1)
\]

where

\[
I^k = \begin{cases} X^k & \text{for system (4.3)}, \\ \left( \begin{array}{c} X^k \\ Z^k \end{array} \right) & \text{for system (4.4)} \end{cases}
\]

and similarly for \( \Delta u^k \).

Two sets of experiments were conducted to observe any difference in the behavior of Algorithm IPF. In the first set, \( \mu_{tol} \) was fixed for Algorithm IPF, i.e. \( \delta(\mu) \equiv \mu_{tol} \). Three tests were conducted which depended on the value of \( \mu_{tol} \) given by

\[
\{10^{-6}, 10^{-4}, 10^{-2}\}.
\]

When \( \mu_{tol} = 10^{-6} \), the iterates were required to follow the central path closely, but when \( \mu_{tol} = 10^{-2} \) the iterates followed the central path loosely. In the second set of
experiments. $\mu_{tol}$ was initialized with $\mu_{tol} = 0.1$ and updated as follows

$$
\delta(\mu) = \frac{1}{10} \min(1, \mu).
$$

That is, when $\mu \geq 1$, then $\mu_{tol} = 0.1$: when $\mu < 1$, $\mu_{tol}$ is set to one-tenth the current $\mu$ value. As a consequence of decreasing $\mu$ to zero, the iterates generated by Algorithm IPF will follow the central path more and more closely.

Nonconvergence of Algorithm IPF was recorded if the maximum number of iterations, which was set to 300, was reached.

Test problems consisted of six randomly generated, nondegenerate problems $\texttt{r1-r6}$, three nondegenerate problems from Netlib: $\texttt{scagr7}$, $\texttt{sc50b}$, $\texttt{share1b}$, and five degenerate problems from Netlib: $\texttt{adlittle}$, $\texttt{afiro}$, $\texttt{blend}$, $\texttt{sc50a}$, and $\texttt{share2b}$. For the random test problems, the data was generated from a uniform distribution on the interval $(0, 1)$; the initial point $x^0$ was randomly generated, $z^0 = x^0$ and $y^0 = 0$. For the Netlib problems, the initial point was supplied by LIPSOL [25], an interior-point solver for linear programs. For each problem, the same initial point was used for all numerical tests performed. The parameter $\mu$ was given an initial value of $\|x^0\|^2/n$. Our results are shown in Tables 4.1-4.5 where the abbreviation NPDM refers to the Newton primal-dual method, and NLBM denotes the Newton log-barrier method.

In the first set of experiments, presented in Tables 4.1-4.4, the parameter $\mu_{tol}$ is fixed at one of the values listed in (4.40). Tables 4.1 and 4.2 show results for $\mu_{tol} = 10^{-2}$ and $\mu_{tol} = 10^{-6}$, respectively, with $tol = 10^{-6}$. Tables 4.3 and 4.4 show results for a larger convergence tolerance of $tol = 10^{-2}$ with $\mu_{tol} = 10^{-2}$ and $\mu_{tol} = 10^{-6}$, respectively.

From Tables 4.1-4.4, we see that our Newton primal-dual algorithm terminated successfully for all the test problems and for all the tested $tol$ and $\mu_{tol}$ values. These
results suggest that the Newton primal-dual method does not need to follow the central path closely in order to obtain high accuracy solutions, which are characterized by a small convergence tolerance of about $10^{-6}$. In fact, fewer iterations are required if the iterates are allowed to follow the central path only loosely. Furthermore, for most problems the condition number of the Jacobian matrix $F'_D(v^k,\mu)$ maintained a moderate value near the solution as $\mu$ approached zero.

The Newton log-barrier method exhibited very different behavior from the Newton primal-dual method. With $tol = 10^{-6}$ and fixed $\mu \text{tol} = 10^{-2}$, we observe from Table 4.1 that Algorithm IPF reached the prescribed accuracy in only one test case. Similar results were obtained for $\mu \text{tol} = 10^{-4}$. Nonconvergence was due to failure in reaching the prescribed accuracy in dual feasibility, i.e. $\|A^T y^k + \mu (X^k)^{-1} - c\| = 0$, which usually fell between $10^{-3}$ and $10^{-6}$ at the end, short of the required accuracy of $10^{-6}$. In addition, we observed that the condition number of the Jacobian matrix $F'_D(v^k,\mu)$ significantly increased at the end. It is worth noting, however, that when $\mu \text{tol} = 10^{-6}$ the Jacobian matrix exhibited better conditioning near the solution, and the algorithm was able to reach the accuracy of $tol = 10^{-6}$. These results strongly suggest that the severe ill-conditioning of $F'_D(v,\mu)$ near the solution can be alleviated if the Newton log-barrier method is forced to follow the central path closely.
Table 4.1 Results for $tol = 10^{-6}$ and fixed $\mu tol = 10^{-2}$

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Dimensions</th>
<th>NPDM</th>
<th>NLBM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m, n</td>
<td>Iters</td>
<td>cond($F_p^r$)</td>
</tr>
<tr>
<td>r1</td>
<td>13, 22</td>
<td>24</td>
<td>7.3e+03</td>
</tr>
<tr>
<td>r2</td>
<td>34, 45</td>
<td>28</td>
<td>8.6e+04</td>
</tr>
<tr>
<td>r3</td>
<td>13, 15</td>
<td>15</td>
<td>2.5e+03</td>
</tr>
<tr>
<td>r4</td>
<td>14, 16</td>
<td>16</td>
<td>2.7e+03</td>
</tr>
<tr>
<td>r5</td>
<td>23, 27</td>
<td>27</td>
<td>2.3e+04</td>
</tr>
<tr>
<td>r6</td>
<td>26, 34</td>
<td>20</td>
<td>1.3e+04</td>
</tr>
<tr>
<td>scagr7</td>
<td>130, 140</td>
<td>63</td>
<td>1.9e+07</td>
</tr>
<tr>
<td>share1b</td>
<td>118, 225</td>
<td>75</td>
<td>1.4e+11</td>
</tr>
<tr>
<td>sc50b</td>
<td>51, 48</td>
<td>32</td>
<td>1.3e+05</td>
</tr>
<tr>
<td>adlittle</td>
<td>57, 97</td>
<td>57</td>
<td>5.8e+12</td>
</tr>
<tr>
<td>afiro</td>
<td>28, 32</td>
<td>34</td>
<td>4.9e+12</td>
</tr>
<tr>
<td>blend</td>
<td>75, 83</td>
<td>37</td>
<td>1.7e+08</td>
</tr>
<tr>
<td>sc50a</td>
<td>51, 48</td>
<td>32</td>
<td>1.4e+07</td>
</tr>
<tr>
<td>share2b</td>
<td>97, 79</td>
<td>45</td>
<td>1.5e+11</td>
</tr>
<tr>
<td>Total Iters</td>
<td></td>
<td>505</td>
<td>4083</td>
</tr>
</tbody>
</table>

† refers to the maximum number of iterations reached.
* Inf refers to Matlab's representation of positive infinity.
<table>
<thead>
<tr>
<th>Problem Number</th>
<th>NPDM</th>
<th></th>
<th>NLBM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iters</td>
<td>cond($F'_p$)</td>
<td>iters</td>
<td>cond($F'_p$)</td>
</tr>
<tr>
<td>r1</td>
<td>30</td>
<td>7.3e+03</td>
<td>79</td>
<td>1.2e+11</td>
</tr>
<tr>
<td>r2</td>
<td>34</td>
<td>8.6e+04</td>
<td>81</td>
<td>2.0e+11</td>
</tr>
<tr>
<td>r3</td>
<td>24</td>
<td>2.5e+03</td>
<td>64</td>
<td>2.2e+08</td>
</tr>
<tr>
<td>r4</td>
<td>22</td>
<td>2.7e+03</td>
<td>73</td>
<td>4.3e+10</td>
</tr>
<tr>
<td>r5</td>
<td>32</td>
<td>2.3e+04</td>
<td>83</td>
<td>1.3e+12</td>
</tr>
<tr>
<td>r6</td>
<td>29</td>
<td>1.3e+04</td>
<td>70</td>
<td>2.9e+08</td>
</tr>
<tr>
<td>scagr7</td>
<td>79</td>
<td>1.9e+07</td>
<td>152</td>
<td>8.0e+15</td>
</tr>
<tr>
<td>share1b</td>
<td>91</td>
<td>1.4e+11</td>
<td>138</td>
<td>2.6e+14</td>
</tr>
<tr>
<td>sc50b</td>
<td>43</td>
<td>1.3e+05</td>
<td>87</td>
<td>1.4e+07</td>
</tr>
<tr>
<td>adlittle</td>
<td>68</td>
<td>5.8e+12</td>
<td>145</td>
<td>1.1e+23</td>
</tr>
<tr>
<td>afro</td>
<td>44</td>
<td>4.9e+12</td>
<td>103</td>
<td>8.8e+18</td>
</tr>
<tr>
<td>blend</td>
<td>55</td>
<td>1.7e+08</td>
<td>108</td>
<td>3.0e+14</td>
</tr>
<tr>
<td>sc50a</td>
<td>66</td>
<td>1.4e+07</td>
<td>89</td>
<td>6.4e+08</td>
</tr>
<tr>
<td>share2b</td>
<td>60</td>
<td>1.8e+11</td>
<td>102</td>
<td>2.4e+19</td>
</tr>
<tr>
<td>Total Iters</td>
<td>657</td>
<td></td>
<td>1374</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>NPDM</th>
<th></th>
<th>NLBM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iters</td>
<td>cond($F'_p$)</td>
<td>iters</td>
<td>cond($F'_p$)</td>
</tr>
<tr>
<td>r1</td>
<td>19</td>
<td>7.3e+03</td>
<td>36</td>
<td>4.1e+07</td>
</tr>
<tr>
<td>r2</td>
<td>23</td>
<td>8.6e+04</td>
<td>38</td>
<td>6.5e+07</td>
</tr>
<tr>
<td>r3</td>
<td>10</td>
<td>2.4e+03</td>
<td>24</td>
<td>1.6e+05</td>
</tr>
<tr>
<td>r4</td>
<td>11</td>
<td>2.7e+03</td>
<td>27</td>
<td>2.8e+06</td>
</tr>
<tr>
<td>r5</td>
<td>22</td>
<td>2.3e+04</td>
<td>40</td>
<td>4.2e+08</td>
</tr>
<tr>
<td>r6</td>
<td>15</td>
<td>1.5e+04</td>
<td>31</td>
<td>1.1e+05</td>
</tr>
<tr>
<td>scagr7</td>
<td>58</td>
<td>1.7e+08</td>
<td>93</td>
<td>2.6e+13</td>
</tr>
<tr>
<td>share1b</td>
<td>69</td>
<td>3.4e+13</td>
<td>81</td>
<td>2.8e+13</td>
</tr>
<tr>
<td>sc50b</td>
<td>26</td>
<td>3.4e+05</td>
<td>32</td>
<td>8.3e+03</td>
</tr>
<tr>
<td>adlittle</td>
<td>52</td>
<td>1.9e+09</td>
<td>91</td>
<td>1.2e+16</td>
</tr>
<tr>
<td>afro</td>
<td>28</td>
<td>3.1e+08</td>
<td>51</td>
<td>3.7e+10</td>
</tr>
<tr>
<td>blend</td>
<td>30</td>
<td>8.0e+04</td>
<td>52</td>
<td>1.0e+06</td>
</tr>
<tr>
<td>sc50a</td>
<td>26</td>
<td>9.2e+05</td>
<td>32</td>
<td>2.2e+04</td>
</tr>
<tr>
<td>share2b</td>
<td>37</td>
<td>8.1e+06</td>
<td>50</td>
<td>2.5e+12</td>
</tr>
<tr>
<td>Total Iters</td>
<td>426</td>
<td></td>
<td>678</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.4  Results for $tol = 10^{-2}$ and fixed $\mu tol = 10^{-6}$

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>NPDM</th>
<th></th>
<th>NLBM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lters</td>
<td>cond($F_p^2$)</td>
<td>lters</td>
<td>cond($F_p^2$)</td>
</tr>
<tr>
<td>r1</td>
<td>23</td>
<td>7.3e+03</td>
<td>39</td>
<td>4.1e+07</td>
</tr>
<tr>
<td>r2</td>
<td>27</td>
<td>8.6e+04</td>
<td>41</td>
<td>6.5e+07</td>
</tr>
<tr>
<td>r3</td>
<td>15</td>
<td>2.5e+03</td>
<td>28</td>
<td>1.6e+05</td>
</tr>
<tr>
<td>r4</td>
<td>15</td>
<td>2.7e+03</td>
<td>30</td>
<td>2.8e+06</td>
</tr>
<tr>
<td>r5</td>
<td>26</td>
<td>2.3e+04</td>
<td>42</td>
<td>4.2e+08</td>
</tr>
<tr>
<td>r6</td>
<td>20</td>
<td>1.1e+04</td>
<td>34</td>
<td>1.1e+05</td>
</tr>
<tr>
<td>scagr7</td>
<td>70</td>
<td>1.8e+08</td>
<td>111</td>
<td>2.6e+13</td>
</tr>
<tr>
<td>share1b</td>
<td>81</td>
<td>3.4e+13</td>
<td>101</td>
<td>2.8e+13</td>
</tr>
<tr>
<td>sc50b</td>
<td>33</td>
<td>3.1e+05</td>
<td>43</td>
<td>8.6e+03</td>
</tr>
<tr>
<td>adlittle</td>
<td>61</td>
<td>1.9e+09</td>
<td>104</td>
<td>1.2e+16</td>
</tr>
<tr>
<td>afro</td>
<td>35</td>
<td>3.1e+08</td>
<td>60</td>
<td>3.7e+10</td>
</tr>
<tr>
<td>blend</td>
<td>39</td>
<td>8.0e+04</td>
<td>64</td>
<td>1.0e+06</td>
</tr>
<tr>
<td>sc50a</td>
<td>36</td>
<td>8.9e+05</td>
<td>45</td>
<td>2.3e+04</td>
</tr>
<tr>
<td>share2b</td>
<td>46</td>
<td>8.2e+06</td>
<td>58</td>
<td>4.2e+10</td>
</tr>
<tr>
<td>Total lters</td>
<td>527</td>
<td></td>
<td>800</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5  Results for $tol = 10^{-6}$ with initial $\mu tol = 10^{-1}$ and updated

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>NPDM</th>
<th></th>
<th>NLBM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lters</td>
<td>cond($F_p^2$)</td>
<td>lters</td>
<td>cond($F_p^2$)</td>
</tr>
<tr>
<td>r1</td>
<td>24</td>
<td>7.3e+03</td>
<td>74</td>
<td>1.2e+11</td>
</tr>
<tr>
<td>r2</td>
<td>29</td>
<td>8.6e+04</td>
<td>77</td>
<td>2.0e+11</td>
</tr>
<tr>
<td>r3</td>
<td>15</td>
<td>2.5e+03</td>
<td>59</td>
<td>2.2e+08</td>
</tr>
<tr>
<td>r4</td>
<td>17</td>
<td>2.7e+03</td>
<td>66</td>
<td>8.6e+09</td>
</tr>
<tr>
<td>r5</td>
<td>27</td>
<td>2.3e+04</td>
<td>77</td>
<td>1.3e+12</td>
</tr>
<tr>
<td>r6</td>
<td>23</td>
<td>1.3e+04</td>
<td>67</td>
<td>2.9e+08</td>
</tr>
<tr>
<td>scagr7</td>
<td>64</td>
<td>1.9e+07</td>
<td>127</td>
<td>8.0e+15</td>
</tr>
<tr>
<td>share1b</td>
<td>75</td>
<td>1.4e+11</td>
<td>112</td>
<td>2.6e+14</td>
</tr>
<tr>
<td>sc50b</td>
<td>31</td>
<td>1.3e+05</td>
<td>73</td>
<td>1.4e+07</td>
</tr>
<tr>
<td>adlittle</td>
<td>55</td>
<td>5.8e+12</td>
<td>128</td>
<td>1.1e+23</td>
</tr>
<tr>
<td>afro</td>
<td>33</td>
<td>4.9e+12</td>
<td>89</td>
<td>8.8e+18</td>
</tr>
<tr>
<td>blend</td>
<td>44</td>
<td>1.7e+08</td>
<td>92</td>
<td>3.0e+14</td>
</tr>
<tr>
<td>sc50a</td>
<td>35</td>
<td>1.4e+07</td>
<td>72</td>
<td>6.4e+08</td>
</tr>
<tr>
<td>share2b</td>
<td>46</td>
<td>1.8e+11</td>
<td>89</td>
<td>2.4e+19</td>
</tr>
<tr>
<td>Total lters</td>
<td>518</td>
<td></td>
<td>1202</td>
<td></td>
</tr>
</tbody>
</table>
In order to minimize the effects of ill-conditioning for the barrier system, the algorithm was implemented with a much relaxed stopping tolerance of $tol = 10^{-2}$ to observe the behavior of Algorithm IPF away from the solution. In this experiment, Algorithm IPF terminated in all test cases with low accuracy solutions, characterized by a large convergence tolerance of $tol = 10^{-2}$. As can be seen from Tables 4.3 and 4.4, the Jacobian matrix exhibited smaller condition numbers compared to the previous results when $tol = 10^{-6}$, especially for the barrier system. In this case, still considerably fewer iterations are required by the Newton primal-dual method.

In the second set of experiments, $\mu tol$ was initialized to $\mu tol = 0.1$ and updated as given in (4.41). Thus, the iterates were required to follow the central path more closely as $\mu$ decreased to zero. In this instance, the Newton log-barrier and the Newton primal-dual methods converged for all the test cases. Again, the Newton primal-dual method required fewer iterations, as demonstrated in Table 4.5.

In summary, we observed that for the tested values of $tol$ or $\mu tol$, the Newton primal-dual method converged for all problems. Hence, the Newton primal-dual method is more efficient and does not need to follow the central path closely, requiring consistently less iterations to converge to a solution. We believe that this phenomenon can be partly explained by the behavior of the radius of the sphere of convergence for the two equivalent systems, as studied in Sections 4.4 and 4.5.

In the case of the Newton log-barrier method, our numerical results suggest that the algorithm needs to follow the central path closely to alleviate the problem of ill-conditioning and to obtain a solution to the linear program. For a related work on ill-conditioning of primal-dual systems, we refer the reader to the recent paper by M. Wright [21].
4.7 Conclusions

In this chapter, we studied the local and global behavior of Newton's method on two equivalent systems from linear programming: the optimality system for the log-barrier formulation (4.3) of the linear program and the perturbed optimality system (4.4) for the linear program itself.

On the issue of local behavior, the radius of the sphere of convergence of Newton's method applied to the barrier system decreases to zero at exactly the same rate as $\mu$ decreases to zero. On the other hand, the radius of the sphere of convergence associated with the perturbed system stays bounded away from zero as $\mu$ decreases to zero. The numerical experiments confirm these results. In addition, numerical results were provided for the case of degenerate problems. The results show that the radius of the sphere of convergence associated with the barrier and perturbed systems seems to decrease to zero as $\mu$ decreases to zero; however, the radius associated with the perturbed system is still larger than that of the barrier system.

It is then not surprising that the superior local behavior of Newton's method on the perturbed system will be reflected in a global setting. To test the global behavior of (damped) Newton's method on the two equivalent systems, a simple interior-point path-following algorithm for solving linear programs was applied to the two systems. We considered the following two interior-point methods: the Newton log-barrier method and the Newton primal-dual method. The numerical results show the Newton log-barrier method needs to follow the central path closely to reduce the ill-conditioning in the Jacobian matrix to obtain a solution for the linear program. However, the Newton primal-dual method always converges to a solution. Moreover, even when both methods follow the central path closely and converge, the Newton primal-dual method requires fewer iterations: thus it is more efficient. In fact, far
more efficient algorithms exist for the Newton primal-dual method than the simple path-following algorithm used in our test (see the algorithms by Kojima et al. [12] and Mehrotra [15]).

The results in this chapter for the linear program not only confirm that Newton primal-dual methods should be the methods of choice for an interior-point path-following framework but also provide an explanation for the efficient performance of these methods.
Chapter 5

Inequality Constrained Nonlinear Optimization

In the previous chapter, we showed differences in the local and global behavior of two Newton interior-point methods on two commonly used equivalent systems for the linear program. We extend the analysis of the local behavior of Newton’s method to the inequality constrained optimization problem. We focus also on studying properties of the radius of the sphere of convergence of Newton’s method applied to two equivalent systems associated with the inequality constrained nonlinear optimization problem: the barrier and the perturbed systems. The barrier system arises from the optimality conditions of the log-barrier formulation of the optimization problem and the perturbed system arises from the perturbed optimality conditions of the optimization problem. Recently, much research has involved studying properties associated with Newton log-barrier methods, which incorporate the barrier system, but not much explanation has been given for the effective performance of interior-point methods that employ the perturbed system. In this work, we show why Newton primal-dual methods should be preferred over Newton log-barrier methods to solve optimization problems.

S. Wright [22] provides a lower-bound for the radius of the sphere of convergence of Newton’s method applied to the barrier system as $\mu$ decreases to zero. In particular, he shows the lower-bound is dependent on $\mu$ but has order greater than one. However, we prove a stronger result showing that the radius of the sphere of convergence of Newton’s method for the barrier system decreases to zero with the same order as $\mu$ decreases to zero.
The chapter is organized as follows. Section 5.1 provides a brief discussion of the optimality conditions of the inequality constrained nonlinear optimization problem. Then, in Section 5.2 and Section 5.3, we present the barrier and perturbed systems for the optimization problem. Section 5.4 discusses theoretical results on the radius of the sphere of convergence of Newton's method applied to the two equivalent systems. Numerical results on the radius of the sphere of convergence of Newton's method applied to the two equivalent systems are presented in Section 5.5. Finally, we present some concluding remarks.

5.1 Optimality Conditions

We consider the following nonlinear inequality constrained optimization problem

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g(x) \geq 0
\end{align*}
\]

where \( f : \mathbb{R}^n \to \mathbb{R} \) and \( g : \mathbb{R}^n \to \mathbb{R}^m \) are twice Lipschitz continuously differentiable functions. The optimality conditions of problem (5.1) are formed by considering the Lagrangian function

\[
L(x, z) = f(x) - \sum_{i=1}^{m} z_i g_i(x),
\]

where \( z \in \mathbb{R}^n \) and nonnegative, are the Lagrangian multipliers associated with the inequality constraints.

Let \( x^* \) be a local solution of problem (5.1). Let \( B = \{ i : g_i(x^*) = 0 \} \) denote the set of indices of the active constraints. Regularity holds at \( x^* \) if the set of active constraint gradients, that is, \( \{ \nabla g_i(x^*) : i \in B \} \) is linearly independent. Strict complementarity
holds if for all \( i \). \( z_i^* + g_i(x^*) > 0 \). If \( x^* \) is a regular point, then the following KKT conditions are satisfied

\[
\nabla_x L(x^*, z^*) = 0, \quad Z^* g(x^*) = 0, \quad g(x^*) \geq 0, \quad z^* \geq 0
\]  

(5.2)

where \( Z^* = \text{diag}(z^*) \) and \( \nabla_x L(x^*, z^*) = \nabla f(x^*) - \sum_{i=1}^{m} z_i^* \nabla g_i(x^*) \) is the gradient of the Lagrangian function. We collect the equations in system (5.2) to form the system

\[
F(x, z) \equiv \begin{bmatrix}
\nabla f(x) - \sum_{i=1}^{m} z_i \nabla g_i(x) \\
Z g(x)
\end{bmatrix} = 0. \quad (5.3)
\]

The second-order condition for problem (5.1) states

\[
\eta^T \nabla^2_z L(x^*, z^*) \eta > 0 \text{ for } \nabla g_i(x^*)^T \eta = 0, \quad i \in \mathcal{B} \text{ and } \eta \neq 0. \quad (5.4)
\]

Then the sufficiency conditions are satisfied at \( x^* \) if the KKT conditions and second-order condition (5.4) hold.

## 5.2 Barrier System

In this section, we introduce the barrier system, one of two equivalent systems under consideration for the analysis of the radius of the sphere of convergence of Newton’s method. As stated earlier, the barrier system is the kernel of the Newton log-barrier method, which is an interior-point method for solving the nonlinear program (5.1).

As in Section 2.2, we first introduce the log-barrier framework for problem (5.1). The optimization problem is solved through a sequence of log-barrier subproblems with decreasing values of \( \mu > 0 \). For a given \( \mu > 0 \), the log-barrier subproblem is given by
minimize \[ J(x) \equiv f(x) - \mu \sum_{i=1}^{m} \log g_i(x) \]
\[ (g(x) > 0) \] (5.5)

where \( J(x) \) is the barrier function and \( \mu \) is the barrier parameter. The first-order optimality conditions for the log-barrier formulation of the inequality constrained optimization problem are

\[ F_B(x; \mu) \equiv \nabla_x J(x) = 0 \quad (g(x) > 0) \] (5.6)

where

\[ \nabla_x J(x) = \nabla f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) \]

is the gradient of the barrier function with respect to \( x \). We will call the nonlinear system of equations in (5.6) the barrier system for problem (5.1). The Jacobian of \( F_B(x; \mu) \) is given by

\[ F'_B(x; \mu) \equiv \nabla^2 f(x) - \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla^2 g_i(x) + \sum_{i=1}^{m} \frac{\mu}{g_i^2(x)} \nabla g_i(x) \nabla^T g_i(x). \]

Let \( x^* \) denote a local solution for the inequality constrained problem (5.1), and let \( x^*_\mu \) denote a solution for system (5.6), whenever it exists, for \( \mu > 0 \). The Jacobian is positive definite at \( x^*_\mu \) for small \( \mu > 0 \) when strict complementarity and the second-order condition (5.4) hold at \( x^* \). If these conditions hold, then \( x^*_\mu \) is an unconstrained minimizer of the log-barrier subproblem (5.5). Under mild conditions [5], the solutions \( x^*_\mu \) converge to \( x^* \) as \( \mu \) decreases to zero.
5.3 Perturbed System

Now we present a new system equivalent to system (5.6) discussed previously. This new system will form the basis of a Newton primal-dual method for the nonlinear program (5.1). The equivalence of the two systems is described in Section 2.4.

Once again, the new equivalent system is derived by introducing an auxiliary variable \( z \in \mathbb{R}^n \). Now, let \( z_i = \mu / g_i(x) \), which is equivalent to \( Zg(x) = \mu e \) where \( Z = \text{diag}(z) \). Substituting for \( z \) in the barrier system and adding the equivalent relations produces the following nonlinear system

\[
F_P(x, z; \mu) \equiv \begin{bmatrix}
\nabla f(x) - \sum_{i=1}^{m} z_i \nabla g_i(x) \\
Zg(x) - \mu e
\end{bmatrix} = 0 \quad (z, g(x) > 0).
\] (5.7)

The system \( F_P(x, z; \mu) \) will be referred to as the perturbed system for problem (5.1).

The Jacobian of \( F_P(x, z; \mu) \) is given by

\[
F_P'(x, z; \mu) \equiv \begin{bmatrix}
\nabla^2_x f(x, z) & -\nabla g(x) \\
Z \nabla^T g(x) & \mathbf{G}(x)
\end{bmatrix}
\]

where \( \mathbf{G}(x) = \text{diag}(g(x)) \). \( \nabla g(x) \) is an \( n \times m \) matrix whose \( i \)th column consists of the gradient of \( g_i(x) \), and \( \nabla^2_x f(x, z) \) denotes the Hessian of the Lagrangian function with respect to \( x \)

\[
f(x, z) = f(x) - \sum_{i=1}^{m} z_i g_i(x).
\]

The Jacobian \( F_P'(x^*, z^*; 0) \) is nonsingular if regularity, strict complementarity, and the second-order condition (5.1) for problem (5.1) are satisfied.
Trajectory of Solutions

In Section 2.4, we discussed the existence of an isolated trajectory of solutions produced by the equivalent systems (2.5) and (2.6) for the general nonlinear program (2.1). In this section, we present briefly the trajectory for systems (5.6) and (5.7) associated with the inequality constrained nonlinear program (5.1).

We assume all conditions are met as described in Section 2.4 by the functions $f$ and $g$ and a solution $x^*$ of problem (5.1). Then under these mild assumptions (see [5] and Section 2.4), a trajectory of solutions for the equivalent systems is guaranteed to exist for small $\mu$. We will assume that the solutions $x^*_\mu$ and $(x^*_\mu, z^*_\mu)$ exist for systems (5.6) and (5.7), respectively, for $\mu \leq \hat{\mu}$. A trajectory of solutions produced by system (5.6) is defined as follows

$$C_B = \{ x^*_\mu : F_B(x^*_\mu; \mu) = 0, \ g(x^*_\mu) > 0, \ \hat{\mu} \geq \mu > 0 \}. \hspace{1cm} (5.8)$$

By the equivalence of systems (5.6) and (5.7), a trajectory of solutions produced by system (5.7) is given by

$$C_P = \{ (x^*_\mu, z^*_\mu) : F_P(x^*_\mu, z^*_\mu; \mu) = 0, \ z^*_\mu, \ g(x^*_\mu) > 0, \ \hat{\mu} \geq \mu \geq 0 \}. \hspace{1cm} (5.9)$$

5.4 Theory for the Radius of the Sphere of Convergence

We analyze the behavior of the radius of the sphere of convergence of Newton's method applied to the barrier and perturbed systems of the inequality constrained optimization problem (5.1) as $\mu$ decreases to zero. For $\mu$ sufficiently small, S. Wright [22] establishes a lower-bound of $O(\mu^\alpha)$ where $\alpha > 1$ for the radius of the sphere of convergence of Newton's method applied to the barrier system. His analysis shows that the Newton step $s$ is of $O(\mu^\alpha)$. 
In Section 5.4.2, we provide a sharper result for the radius of the sphere of convergence of Newton's method on the barrier system showing that the radius is bounded below and above by a bound which is $O(\mu)$. This result implies that the radius of the sphere of convergence associated with the barrier system decreases to zero in the same order as $\mu$ decreases to zero. Our proof uses the same assumptions on the functions $f$ and $g$ of problem (5.1) as S. Wright [22]. We also consider range and null subspace information of the active constraint gradients. But, we approach the proof by considering the following step

$$x_+ - x_\mu^* = F_B(x; \mu)^{-1}[F_B(x_\mu^*: \mu) - F_B(x; \mu) - F_B^r(x; \mu)(x_\mu^* - x)].$$

(5.10)

where $x$ is the initial Newton iterate, $x_+$ is the subsequent iterate after having taken a full Newton step, and $x_\mu^*$ is the solution to the barrier system. The proofs in this section produce bounds for the step $(x_+ - x_\mu^*)$.

Finally in Section 5.4.3, we conduct the same analysis on the radius of the sphere of convergence of Newton's method applied to the perturbed system as $\mu$ decreases to zero. The reader is referred to Chapter 3 for a description of the sphere of convergence of Newton's method.

We note that Lemma 5.4 provides the existence of an upper-bound $\hat{\mu}$ for the values of $\mu$ that we consider. In some cases, the subsequent lemmas and theory may hold for values of $\mu \leq \mu'$, where $\mu' < \hat{\mu}$. Without loss of generality, we will assume the lemmas and theorems hold for $\mu \leq \hat{\mu}$.

### 5.4.1 Preliminaries

We present some preliminary results and lemmas to be used in the analysis of the radius of the sphere of convergence of Newton's method applied to the barrier system.
A nondegeneracy assumption is presented initially which is assumed to hold for the lemmas and theorems in the remainder of the chapter. We prove first some general results on functions given by Lemmas 5.1 and 5.2. We will make much use of Lemma 5.3, which is taken from [2], and is presented here for completeness and used in the proof of Lemma 5.7 for the barrier system.

**Notation**

Let $\mathcal{B} = \{i : g_i(x^*) = 0\}$ denote the set of indices of the active constraints. Now, let $\mathcal{U}_B(x)$ denote an orthonormal matrix that spans the range space of the active constraint gradients at $x$. Let $\mathcal{U}_N(x)$ denote an orthonormal matrix that spans the null space of the range of $\mathcal{U}_B(x)$. Then $[\mathcal{U}_B(x) \quad \mathcal{U}_N(x)]$ is orthogonal.

**Nondegeneracy Assumption** At a local unique solution $x^*$, regularity, strict complementarity, and the second-order condition (5.4) are satisfied.

**Lemma 5.1** Given $c : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$, where $D$ is a compact set. If $c$ is Lipschitz continuous over $D$ with Lipschitz constant $M_1 > 0$, then $cc^T$ is Lipschitz continuous over $D$ with Lipschitz constant

$$M = 2M_1 \tau_c$$

where $\tau_c = \max\{\|c(x)\| : x \in D\}$.

**Proof**

Let $z, y \in D$. Then

$$\|c(z)c(z)^T - c(y)c(y)^T\| = \|c(z)c(z)^T - c(z)c(y)^T + c(z)c(y)^T - c(y)c(y)^T\|$$

$$= \|c(z)(c(z)^T - c(y)^T) + (c(z) - c(y))c(y)^T\|$$

$$\leq \|c(z)\|\|c(z)^T - c(y)^T\| + \|c(z) - c(y)\|\|c(y)^T\| (5.11)$$
Since $c$ is bounded and is Lipschitz continuous with constant $M_1 > 0$, then from (5.11) we obtain

$$\| c(z)c(z)^T - c(y)c(y)^T \| \leq M \| z - y \|,$$

where $M = 2M_1 \tau_c$. \hfill \Box

**Lemma 5.2**

Given $h : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ and $v : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, where $D$ is a compact and convex set. Assume $h$ is Lipschitz continuous over $D$ with Lipschitz constant $M_1 > 0$ and $v$ is continuously differentiable with $v(x) \neq 0$ on $D$. Then for a positive integer $p$, $h(x)/v^p(x)$ is Lipschitz continuous in $D$ with Lipschitz constant

$$M = \frac{M_1}{\kappa^p} + \frac{p \eta \tau \gamma^{p-1}}{\kappa^{2p}},$$

where $\eta = \max\{\|\nabla v(x)\| : x \in D\}$, $\tau = \max\{\|h(x)\| : x \in D\}$, $\gamma = \max\{|v(x)| : x \in D\}$, and $\kappa = \min\{|v(x)| : x \in D\}$.

**Proof**

Let $z, y \in D$. Then

$$\left\| \frac{h(z)}{v^p(z)} - \frac{h(y)}{v^p(y)} \right\| = \left\| \frac{v^p(y)h(z) - v^p(z)h(y)}{v^p(z) v^p(y)} \right\| = \left\| \frac{v^p(z)(h(z) - h(y)) - h(z)(v^p(z) - v^p(y))}{v^p(z) v^p(y)} \right\| \leq \frac{v^p(z)}{v^p(z) v^p(y)} \left\| h(z) - h(y) \right\| + \frac{v^p(z)}{v^p(z) v^p(y)} \left\| v^p(z) - v^p(y) \right\| \ (5.12)$$

Using the Lipschitz continuity of $h$ and the first-order Taylor Series of $v^p(z)$, we obtain from (5.12) that for $\xi$ between $z$ and $y$, and $M_1 > 0$

$$\left\| \frac{h(z)}{v^p(z)} - \frac{h(y)}{v^p(y)} \right\| \leq \frac{M_1 \| z - y \|}{|v^p(y)|} + \frac{\| h(z) \| | p v^{p-1}(\xi) \nabla v(\xi)^T (z - y) |}{|v^p(z) v^p(y)|}. $$
Since $\nabla v$ is bounded above, then $|\nabla v(\xi)^T(z - y)| \leq \eta \| z - y \|$, and since $h$ and $v$ are bounded, we obtain

$$\left\| \frac{h(z)}{v^p(z)} - \frac{h(y)}{v^p(y)} \right\| \leq M \| z - y \|$$

for constant $M = M_1/\kappa^p + p\eta\tau\gamma^{p-1}/\kappa^{2p}$.

\[ \square \]

Lemma 5.3

Let $F : \mathbb{R}^c \rightarrow \mathbb{R}^n$ be continuously differentiable in the open convex set $D \subset \mathbb{R}^c$. Let $F'$ be Lipschitz continuous at $x \in D$ under a vector norm and the induced matrix operator norm with Lipschitz constant $\gamma$. Then, for any $x + p \in D$,

$$\|F(x + p) - F(x) - F'(x)p\| \leq \frac{\gamma}{2}\|p\|^2.$$  \hspace{1cm} (5.13)

Proof

$$F(x + p) - F(x) - F'(x)p = \left[ \int_0^1 F'(x + tp)p \, dt \right] - F'(x)p$$

$$= \int_0^1 [F'(x + tp) - F'(x)]p \, dt.$$  \hspace{1cm} (5.13)

Then by the definition of a matrix operator norm, and the Lipschitz continuity of $F'$ at $x$ in neighborhood $D$, we obtain that for $t \in [0, 1]$

$$\|F(x + p) - F(x) - F'(x)p\| \leq \int_0^1 \|F'(x + tp) - F'(x)\| \|p\| \, dt$$

$$\leq \int_0^1 \gamma \|tp\| \|p\| \, dt$$

$$= \gamma \|p\|^2 \int_0^1 t \, dt$$

$$= \frac{\gamma}{2} \|p\|^2.$$  \hspace{1cm} (5.14)
The following lemma provides bounds on the values that the inequality constraints $g_i(x)$ can take for points $x$ that lie close to the trajectory (5.9), which is parameterized by $\mu$. Recall $C_\rho = \{(x_*^*, z_*^*) : F_\rho(x_*^*, z_*^*) = 0, z_*^* \cdot g(x_*^*) > 0 \, \mu \leq \hat{\mu}\}$ (see (5.9)).

**Lemma 5.4**

Consider $\mu > 0$ and $(x_*^*, z_*^*)$ contained in $C_\rho$. Assume $g : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable on the compact, convex set $D$. Then under the nondegeneracy assumption, there exist $\hat{\mu} > 0$ and constants $K, J_1, J_2, J_3 > 0$ so that for $\mu \leq \hat{\mu}$ and $\|x - x_*^*\| \leq K\mu$ where $x \in D$, we have

$$J_1\mu \leq g_i(x) \leq J_2\mu \quad i \in \mathcal{B}. \quad (5.15a)$$

$$g_i(x) \geq J_3 \quad i \notin \mathcal{B}. \quad (5.15b)$$

**Proof**

For $(x_*^*, z_*^*) \in C_\rho$, we have

$$g_i(x_*^*) (z_*^*)_i = \mu \quad i = 1, \ldots, m. \quad (5.16)$$

Under our assumption, we have $(x_*^*, z_*^*) \to (x^*, z^*)$ as $\mu \to 0$ (see [5]). Because $g$ is continuous,

$$g_i(x_*^*) \to g_i(x^*).$$

Then, without loss of generality for $\mu \leq \hat{\mu}$, we have

$$(z_*^*)_i \leq 2z_i^*. \quad i \in \mathcal{B}. \quad (5.17)$$
Along with the strict complementarity condition at $x^*$, there exist $\hat{\mu} > 0$ and $J > 0$ such that for $\mu \leq \hat{\mu}$

$$g_i(x_\mu^*) \geq J \quad i \notin B \text{ and } (z_\mu^*)_i \geq J \quad i \in B. \quad (5.18)$$

Let

$$M = \max\{\|\nabla g_i(x)\| : x \in D, i \notin B\}. \quad (5.19)$$

$$\omega = \max\{z_i^* : i \in B\},$$

$$K \leq \frac{1}{2M\omega^2}, \text{ and }$$

$$\hat{\mu} \leq \frac{J}{2MK}. \quad (5.19)$$

Consider $\mu \leq \hat{\mu}$. Let $x$ be such that $\|x - x_\mu^*\| \leq K\mu$. By Taylor's Theorem, there exists $\xi_i$ between $x$ and $x_\mu^*$ such that

$$g_i(x) = g_i(x_\mu^*) + \nabla g_i(\xi_i)^T(x - x_\mu^*) \quad i = 1, \ldots, m. \quad (5.20)$$

Applying the Cauchy-Schwarz inequality to (5.20) and the quantities in (5.17) and (5.19), we have for $i \in B$

$$g_i(x) \geq g_i(x_\mu^*) - \|\nabla g_i(\xi_i)\| \|x - x_\mu^*\|$$

$$\geq \frac{\mu}{(z_\mu^*)_i} - M\|x - x_\mu^*\|$$

$$\geq \frac{\mu}{2z_i^*} - MK\mu$$

$$\geq (\frac{1}{2\omega} - MK)\mu$$

$$= J_1\mu$$
where \( J_1 = \left( \frac{1}{2\omega} - MK \right) > 0 \).

Now we establish the second part of the inequality in (5.15a). In a similar manner, we obtain from (5.20) that for \( i \in B \)

\[
g_i(x) \leq g_i(x^*_\mu) + \| \nabla g_i(x_i) \| \| x - x^*_\mu \| \\
\leq \mu/(z^*_\mu)_i + MK\mu. \tag{5.21}
\]

Applying (5.18) to (5.21), we have that

\[
g_i(x) \leq J_2\mu
\]

for \( J_2 = (1/J + MK) > 0 \).

Now consider \( i \notin B \). By (5.18) and (5.20), we have

\[
g_i(x) \geq J + \nabla g_i(x_i)^T(x - x^*_\mu) \\
\geq J - \| \nabla g_i(x_i) \| \| x - x^*_\mu \| \\
\geq J - MK\mu \\
\geq J - MK\hat{\mu} \\
\geq 1/2 J.
\]

where the last inequality arises from the upper-bound on \( \hat{\mu} \) in (5.19). Then, we have for \( i \in B \) that \( g_i(x) \geq J_3 \) for \( J_3 = 1/2 J > 0 \).

\[\square\]

### 5.4.2 Barrier System

In this section, we prove additional results required for the analysis of the behavior of the radius of the sphere of convergence of Newton's method applied to the barrier
system. Then we derive lower- and upper-bound results for the radius of the sphere of convergence in Lemmas 5.8 and 5.9, respectively.

S. Wright [22] provides a representation for \( F'_B(x; \mu)^{-1} \) considering that \( \mu \) is sufficiently small and that \( x \) is close to \( x^*_\mu \). We have provided additional information on an upper-bound for \( \mu \), given in Lemma 5.4, to achieve a similar lemma. The proof for Lemma 5.5 below is omitted, and the reader is referred to S. Wright's proof [22] for details.

**Lemma 5.5**

*Let the nondegeneracy assumption hold. Then for \( \mu \leq \hat{\mu} \) and \( \|x - x^*_\mu\| \leq K\mu \), where \( \hat{\mu} \) and \( K \) are given in Lemma 5.4, \( F'_B(x; \mu)^{-1} \) can be written in the following manner*

\[
F'_B(x; \mu)^{-1} = \begin{bmatrix} U_B(x) & U_N(x) \end{bmatrix} \begin{bmatrix} H_{11}(x; \mu) & H_{12}(x; \mu) \\ H_{12}^T(x; \mu) & H_{22}(x; \mu) \end{bmatrix} \begin{bmatrix} U_B^T(x) \\ U_N^T(x) \end{bmatrix}
\]

where \( H_{11}(x; \mu) = O(\mu), H_{12}(x; \mu) = O(\mu), \) and \( H_{22}(x; \mu) = O(1) \).

S. Wright and Jarre [24] use the implicit function theorem to describe a continuous trajectory of solutions for the system given in (5.3) with the right-hand side modified, that is.

\[
F(x, z; \lambda, \xi) = \begin{bmatrix} \lambda \\
\xi \end{bmatrix}
\]

where \( \lambda \in \mathbb{R}^c \) and \( \xi \in \mathbb{R} \). We make use of their lemma, modifying it with specific choices of \( \lambda \) and \( \xi \) to suit our purposes. An important result (5.24) will be used in Lemma 5.9 to show that the radius of the sphere of convergence of Newton's method applied to the barrier system is bounded above by \( O(\mu) \) for \( \mu \) small. For a proof of Lemma 5.6, we refer the reader to S. Wright and Jarre [24].
Lemma 5.6  Under the nondegeneracy assumption, let the vector pair \((\tilde{x}(\mu), \tilde{z}(\mu))\) denote the solution of the nonlinear system

\[
F(\tilde{x}, \tilde{z}) = \begin{bmatrix} 0 \\ \mu e \end{bmatrix}
\]

for given \(\mu \geq 0\) and \(F\) defined as in (5.3). Then there are positive constants \(\hat{\mu} > 0\) and \(M > 0\) such that

(i) \((\tilde{x}(\mu), \tilde{z}(\mu))\) is a \(C^2\) function of \(\mu\) in the neighborhood defined by

\[\mathcal{N}_{\hat{\mu}} = \{\mu : 0 \leq \mu \leq \hat{\mu}\}\]

(ii) For \(\mu_1, \mu_2 \in \mathcal{N}_{\hat{\mu}}\), we have

\[
\begin{bmatrix}
\tilde{x}(\mu_1) \\
\tilde{z}(\mu_1)
\end{bmatrix}
- \begin{bmatrix}
\tilde{x}(\mu_2) \\
\tilde{z}(\mu_2)
\end{bmatrix}
= F'(\tilde{x}(\mu_1), \tilde{z}(\mu_1))^{-1}
\begin{bmatrix}
0 \\
(\mu_1 - \mu_2)e
\end{bmatrix}
+ r.  \quad (5.23)
\]

where

\[\|r\| \leq M(\mu_1 - \mu_2)^2.\]

Since \(F'(\tilde{x}, \tilde{z})\) is continuous and nonsingular close to \((x^*, z^*)\), then for \(\hat{\mu}\) small, we obtain

\[
\left\| \begin{bmatrix}
\tilde{x}(\mu_1) \\
\tilde{z}(\mu_1)
\end{bmatrix}
- \begin{bmatrix}
\tilde{x}(\mu_2) \\
\tilde{z}(\mu_2)
\end{bmatrix}
\right\| = O(\|\mu_1 - \mu_2\|).
\]

Therefore, if we let \(\mu_1 = \mu\) and \(\mu_2 = 0\), then

\[
(\tilde{x}(\mu_1), \tilde{z}(\mu_1)) = (x^*_\mu, z^*_\mu)
\]

\[
(\tilde{x}(\mu_2), \tilde{z}(\mu_2)) = (x^*, z^*).
\]
It follows that for some constant $M > 0$

$$\left\| \begin{pmatrix} x^*_\mu - x^* \\ z^*_\mu - z^* \end{pmatrix} \right\| \leq M\mu. \tag{5.24}$$

The following notation and quantities will be used in the remaining lemmas. The next result, Lemma 5.7, will be needed in the proof of Lemma 5.8 to obtain a lower-bound for the radius of the sphere of convergence of Newton's method on the barrier system.

In our analysis, we obtain upper-bounds on the quantities in the right-hand side of (5.10). So define $R$ as the vector on the right-hand side of (5.10), that is.

$$R(x; \mu) \equiv F_B(x^*_\mu; \mu) - F_B(x; \mu) - F_B'_{\mu}(x; \mu)(x^*_\mu - x) \equiv R_1(x; \mu) + R_2(x; \mu) + R_3(x; \mu) \tag{5.25}$$

where

$$R_1(x; \mu) = \nabla f(x^*_\mu) - \nabla f(x) - \nabla^2 f(x)(x^*_\mu - x).$$

$$R_2(x; \mu) = \sum_{i \in B} \left( -\frac{\mu}{g_i(x^*_\mu)} \nabla g_i(x^*_\mu) + \frac{\mu}{g_i(x)} \nabla g_i(x) - \frac{\mu}{g_i^2(x)} \nabla^2 g_i(x)(x^*_\mu - x) \right). \tag{5.26}$$

$$R_3(x; \mu) = \sum_{i \in B} \left( -\frac{\mu}{g_i(x^*_\mu)} \nabla g_i(x) + \frac{\mu}{g_i(x)} \nabla g_i(x) - \frac{\mu}{g_i^2(x)} \nabla^2 g_i(x)(x^*_\mu - x) \right).$$

Let $\hat{\mu}$ and $K$ be given as in the proof of Lemma 5.4. For a constant $\varepsilon > 0$, define $B(x^*; \varepsilon) = \{ x : \| x - x^* \| \leq \varepsilon \}$. For all $0 < \mu \leq \hat{\mu}$ and for some constant $\rho > 0$, let

$$\| x - x^*_\mu \| \leq K\mu \subset B(x^*; \varepsilon).$$
We define the following terms for \( i = 1, \ldots, m \) and \( p \equiv (x^*_n - x) \).

\[
\tau_i = \max \left\{ \| \nabla g_i(x) \nabla g_i(x)^T \| : x \in B(x^*; z) \right\},
\]

\[
\alpha_i = \max \{ \| \nabla^2 g_i(x) \| : x \in B(x^*; z) \},
\]

\[
\beta_i = \max \{ \| \nabla g_i(x) \| : x \in B(x^*; z) \}.
\]

\[
\gamma_i = \max \{ g_i(x) : x \in B(x^*; z) \}, \quad \text{and}
\]

\[
\kappa_i(\mu) = \min \{ g_i(x) : \| x - x^*_n \| \leq K \mu \}.
\]

**Lemma 5.7** Assume \( f \) and \( g \) are twice Lipschitz continuously differentiable over \( B(x^*; z) \). Let \( R \) be given as defined in (5.25) with domain \( B(x^*; z) \). Assume \( \mu \) and \( K \) are given as in the proof of Lemma 5.4. Then under the nondegeneracy assumption, for \( \mu \leq \hat{\mu} \) and constants \( C_1, C_2 > 0 \), it follows that for \( \| x - x^*_n \| \leq K \mu \).

\[
\| U_{\mathcal{B}}^T(x) R(x; \mu) \| \leq C_1 \frac{1}{\mu^2} \| x - x^*_n \|^2,
\]

\[
\| U_{\mathcal{N}}^T(x) R(x; \mu) \| \leq C_2 \frac{1}{\mu} \| x - x^*_n \|^2.
\]

**Proof** Consider \( \mu \leq \hat{\mu} \) and \( \| x - x^*_n \| \leq K \mu \). Let all functions \( R_1, R_2, R_3 \) defined in (5.26) and \( U_{\mathcal{N}}^T \) and \( U_{\mathcal{B}}^T \) be evaluated at \( x \) and at a given parameter \( \mu \leq \hat{\mu} \).

Multiplying \( U_{\mathcal{B}}^T(x) \) with \( R \), we have

\[
\| U_{\mathcal{B}}^T(x) R(x; \mu) \| \leq \| U_{\mathcal{B}}^T \| \left( \| R_1 \| + \| R_2 \| + \| R_3 \| \right).
\]

We will obtain an upper-bound for \( \| U_{\mathcal{B}}^T R \| \) by obtaining upper-bounds for each of the four terms in (5.30).

Since \( \| U_{\mathcal{B}}^T \| \) is independent of \( \mu \), we have \( \| U_{\mathcal{B}}^T \| \leq C \) for a constant \( C > 0 \). Now, we will obtain upper-bounds for \( \| R_1 \|, \| R_2 \| \), and \( \| R_3 \| \).

\[
\| R_1 \| = \| \nabla f(x^*_n) - \nabla f(x) - \nabla^2 f(x)(x^*_n - x) \|.
\]
Applying Lemma 5.3 to the above term yields

\[ ||R_1|| \leq \frac{D_1}{2} ||x^*_\mu - x||^2 \]  \hspace{1cm} (5.31)

where \( D_1 > 0 \) is the Lipschitz constant of \( \nabla^2 f \).

To determine upper-bounds for the terms, \( ||R_2|| \) and \( ||R_3|| \), we will proceed as in the proof of Lemma 5.3 to obtain the Lipschitz constants for \( \nabla g \nabla g^T / g^2 \) and \( \nabla^2 g / g \) that depend on \( \mu \). Let \( x_t = x + tp \) for \( t \in [0, 1] \) and let \( p \) be the direction \( (x^*_\mu - x) \). Applying Lemma 5.3 to \( R_2 \), we obtain

\[
||R_2|| \leq \sum_{i \in \mathcal{G}} \int_0^1 \left[ \frac{\mu}{\eta_i(x_t)} \nabla g_i(x_t) \nabla g_i(x_t)^T - \frac{\mu}{\eta_i(x)} \nabla^2 g_i(x) \right] \|x^*_\mu - x\| \, dt \\
-\frac{\mu}{\eta_i(x)} \nabla g_i(x) \nabla g_i(x)^T + \frac{\mu}{\eta_i(x)} \nabla^2 g_i(x) \right] \|x^*_\mu - x\| \, dt \\
\leq \sum_{i \in \mathcal{G}} \int_0^1 \left[ \frac{\mu}{\eta_i(x)} \nabla g_i(x_t) \nabla g_i(x_t)^T - \frac{\mu}{\eta_i(x)} \nabla g_i(x) \nabla g_i(x) \nabla g_i(x)^T \right] \|x^*_\mu - x\| \, dt \\
+ \sum_{i \in \mathcal{G}} \int_0^1 \left[ \frac{\mu}{\eta_i(x)} \nabla^2 g_i(x_t) - \frac{\mu}{\eta_i(x)} \nabla^2 g_i(x) \right] \|x^*_\mu - x\| \, dt. \hspace{1cm} (5.32)
\]

Since \( \nabla g \nabla g^T \) is Lipschitz continuous by Lemma 5.1, then applying Lemma 5.2 to (5.32) yields

\[
||R_2|| \leq \sum_{i \in \mathcal{G}} \int_0^1 \left[ \frac{\mu M_i}{\kappa_i(\mu)^2} + \frac{2 \mu \eta_i \tau_i \gamma_i}{\kappa_i(\mu)^4} \right] \|x_t - x\| \|x^*_\mu - x\| \, dt \\
+ \sum_{i \in \mathcal{G}} \int_0^1 \left[ \frac{\mu F_i}{\kappa_i(\mu)} + \frac{\mu \rho_i \alpha_i}{\kappa_i(\mu)^2} \right] \|x_t - x\| \|x^*_\mu - x\| \, dt
\]

where \( \eta_i, \rho_i > 0 \) and \( \tau_i, \alpha_i, \gamma_i \) and \( \kappa_i(\mu) \) are given in (5.27), and \( M_i \) and \( F_i \) are respectively. the Lipschitz constants for \( \nabla g_i \nabla g_i^T \) and \( \nabla^2 g_i \). By (5.15), \( J_3 \leq \kappa_i(\mu) \) for
\[ i \notin \mathcal{B} \text{ and we obtain} \]
\[
\| R_2 \| \leq \sum_{i \in \mathcal{B}} \int_0^1 \left[ \frac{\mu_i M_i}{J^2_2} + \frac{2 \mu_i \eta_i \tau_i}{J^3_3} \right] \| x_t - x \| \| x^*_\mu - x \| \, dt \tag{5.33}
\]
\[ + \sum_{i \in \mathcal{B}} \int_0^1 \left[ \frac{\mu_i F_i}{J_3} + \frac{\mu_i \rho_i \alpha_i}{J^3_3} \right] \| x_t - x \| \| x^*_\mu - x \| \, dt. \]

From the definition of \( x_t \) and \( \rho \), we obtain that \( \| x_t - x \| = t \| x^*_\mu - x \| \). Making the substitution of \( \| x_t - x \| \) in (5.33) and integrating with respect to \( t \), we obtain

\[
\| R_2 \| \leq \sum_{i \in \mathcal{B}} \left[ \frac{\mu_i M_i}{2J^2_3} + \frac{2 \mu_i \eta_i \tau_i \gamma_i}{2J^4_3} \right] \| x^*_\mu - x \|^2 + \left[ \frac{\mu_i F_i}{2J_3} + \frac{\mu_i \rho_i \alpha_i}{2J^3_3} \right] \| x^*_\mu - x \|^2 \tag{5.34}
\]

for some constant \( D_2 > 0 \).

Similarly, we obtain for \( R_3 \) that

\[
\| R_3 \| \leq \sum_{i \in \mathcal{B}} \int_0^1 \left[ \frac{\mu_i M_i}{\kappa^2_i} + \frac{2 \mu_i \eta_i \tau_i \gamma_i}{\kappa^4_i} \right] \| x_t - x \| \| x^*_\mu - x \| \, dt \tag{5.35}
\]
\[ + \sum_{i \in \mathcal{B}} \int_0^1 \left[ \frac{\mu_i F_i}{\kappa_i} + \frac{\mu_i \rho_i \alpha_i}{\kappa^2_i} \right] \| x_t - x \| \| x^*_\mu - x \| \, dt. \]

Now, by (5.15a), \( J_i \mu \leq \kappa_i(\mu) \) and \( \gamma_i \leq J_2 \mu \) for \( i \in \mathcal{B} \). Making these substitutions in (5.35) and integrating we obtain

\[
\| R_3 \| \leq \sum_{i \in \mathcal{B}} \left[ \frac{\mu_i M_i}{2J^2_1 \mu^2} + \frac{2 \mu_i \eta_i \tau_i J_2 \mu}{2J^4_1 \mu^4} \right] \| x^*_\mu - x \|^2 + \left[ \frac{\mu_i F_i}{2J_1 \mu} + \frac{\mu_i \rho_i \alpha_i}{2J^3_1 \mu^2} \right] \| x^*_\mu - x \|^2.
\]

Therefore,

\[
\| R_3 \| \leq D_3 \frac{1}{\mu^2} \| x^*_\mu - x \|^2
\]

for some constant \( D_3 > 0 \). Collecting the upper-bounds obtained on the terms \( \| U^T \mathcal{B} R(x; \mu) \| \).

\[ \| R_1 \|, \| R_2 \|, \text{ and } \| R_3 \|, \text{ we obtain} \]

\[
\| U^T \mathcal{B} R(x; \mu) \| \leq C_1 \frac{1}{\mu^2} \| x^*_\mu - x \|^2
\]
for a constant $C_1 > 0$.

For the second part of the proof, we work with $U_N$ and again will determine an upper-bound for $\|U_N^T R\|$. We have

$$\|U_N^T R\| \leq \|U_N^T R_1\| + \|U_N^T R_2\| + \|U_N^T R_3\|$$

$$\leq \|U_N^T\| (\|R_1\| + \|R_2\|) + \|U_N^T R_3\|.$$ 

Since $\|U_N^T\|$ is independent of $\mu$, we have $\|U_N^T\| \leq G$ for some constant $G > 0$. We can use the upper-bound estimates obtained previously for $\|R_1\|$ and $\|R_2\|$. However, the situation changes for estimating $\|U_N^T R_3\|$, the term for the active constraints. Following the proof as in Lemma 5.3, we have

$$U_N^T R_3 = \sum_{i \in B} \left[ \int_0^1 \frac{\mu}{g_i(x_t)^2} U_N^T(x) \nabla g_i(x_t) \nabla^T g_i(x_t) - \frac{\mu}{g_i(x_t)} U_N^T(x) \nabla^2 g_i(x_t) \right] (x - x_\mu) dt.$$

from which follows

$$\|U_N^T R_3\| \leq \sum_{i \in B} \int_0^1 \frac{\mu}{g_i(x_t)^2} U_N^T(x) \nabla g_i(x_t) \nabla^T g_i(x_t) - \frac{\mu}{g_i(x_t)} U_N^T(x) \nabla^2 g_i(x_t)$$

$$- \frac{\mu}{g_i(x_t)^2} U_N^T(x) \nabla g_i(x_t) \nabla^T g_i(x_t) + \frac{\mu}{g_i(x)} U_N^T(x) \nabla^2 g_i(x_t) ||x - x_\mu|| dt.$$

$$\leq \sum_{i \in B} (T_{1t} + T_{2t})$$

where
\[ T_{1i} = \int_0^1 \left\| \frac{\mu}{g_i(x_1)^2} U_N^T(x) \nabla g_i(x_1) \nabla g_i(x_1)^T - \frac{\mu}{g_i(x)^2} U_N^T(x) \nabla g_i(x) \nabla g_i(x)^T \right\| ||x - x_n^*|| dt \]

and

\[ T_{2i} = \int_0^1 \left\| \frac{\mu}{g_i(x_1)} U_N^T(x) \nabla^2 g_i(x_1) - \frac{\mu}{g_i(x)} U_N^T(x) \nabla^2 g_i(x) \right\| ||x - x_n^*|| dt. \]

We will work with \( T_{1i} \) and \( T_{2i} \) individually and obtain upper-bounds for each. Since \( U_N(x) \) is orthogonal to the set of active constraint gradients at \( x \), we obtain for \( T_{1i} \) that

\[ T_{1i} = \int_0^1 \left\| \frac{\mu}{g_i(x_1)^2} U_N^T(x) \nabla g_i(x_1) \nabla g_i(x_1)^T \right\| ||x - x_n^*|| dt \]

\[ = \int_0^1 \left\| \frac{\mu}{g_i(x_1)^2} U_N^T(x) \left[ \nabla g_i(x_1) \nabla g_i(x_1)^T - \nabla g_i(x) \nabla g_i(x)^T \right] \right\| ||x - x_n^*|| dt. \]

By Lemma 5.1, \( g_i(x_1) \geq J_1 \mu \). Hence, we obtain

\[ T_{1i} \leq \frac{\mu}{J_1 \mu^2} ||U_N^T(x)|| \int_0^1 \left\| \nabla g_i(x_1) \nabla g_i(x_1)^T - \nabla g_i(x) \nabla g_i(x)^T \right\| ||x - x_n^*|| dt. \]

Because \( \nabla g_i(x) \nabla g_i(x)^T \) is Lipschitz continuous by Lemma 5.1, we obtain for \( S > 0 \)

\[ T_{1i} \leq \frac{1}{J_1 \mu} ||U_N^T(x)|| \int_0^1 2S \beta_1 ||x_t - x|| ||x - x_n^*|| dt \]

\[ \leq \frac{1}{J_1 \mu} S \beta_1 ||U_N^T(x)|| ||x - x_n^*||^2 \]

\[ \leq \frac{C_1}{\mu} ||x - x_n^*||^2 \]

for a constant \( C_1 > 0 \).

Now, we work with the second term \( T_{2i} \) which is bounded by
\[
T_{2i} \leq \|U^T\| \int_0^1 \left\| \frac{\mu}{g_i(x_t)} \nabla^2 g_i(x_t) - \frac{\mu}{g_i(x)} \nabla^2 g_i(x) \right\| \|x - x^*_\mu\| dt.
\]

After applying Lemma 5.2 to the middle term of \(T_{2i}\) and the bounds (5.4) given in Lemma 5.1, we obtain

\[
T_{2i} \leq \|U^T\| \int_0^1 \left[ \frac{\mu F_i}{J_i \mu} + \frac{\mu \rho_i \alpha_i}{J_i \mu^2} \right] \|x_t - x\| \|x - x^*_\mu\| dt
\leq \|U^T\| \left[ \frac{\mu F_i}{2J_i \mu} + \frac{\mu \rho_i \alpha_i}{2J_i^2 \mu^2} \right] \|x - x^*_\mu\|^2
\leq \frac{G_2}{\mu} \|x - x^*_\mu\|^2
\]

where \(F_i\) is the Lipschitz constant for \(\nabla^2 g_i\), and \(\rho_i, G_2 > 0\).

If we collect terms \(T_{1i}\) and \(T_{2i}\), for \(i \in \mathcal{B}\), we obtain that for some \(G_3 > 0\).

\[
\|U^T R_3\| \leq \frac{G_3}{\mu} \|x - x^*_\mu\|^2.
\]  
(5.36)

Using the upper-bounds on \(\|R_1\|\) and \(\|R_2\|\) given in (5.31) and (5.34) and the upper-bound for \(\|U^T R_3\|\) in (5.36), we obtain

\[
\|U^T R(x; \mu)\| \leq \|U^T\| \left[ \frac{D_1}{2} \|x^*_\mu - x\|^2 + D_2 \mu \|x^*_\mu - x\|^2 \right] + \frac{G_3}{\mu} \|x - x^*_\mu\|^2
\leq C_2 \frac{1}{\mu} \|x - x^*_\mu\|^2
\]

for some constant \(C_2 > 0\).

The next two lemmas establish lower- and upper-bound results for the radius of the sphere of convergence of Newton's method applied to the barrier system. The
two results are combined to produce Theorem 5.1, which proves that the radius of the sphere of convergence of Newton's method on the barrier system decreases to zero in the same order as $\mu$ decreases to zero. Our result shows that the radius of the sphere of convergence decreases to zero linearly as $\mu$ decreases to zero. It is stronger than S. Wright's result [22] which shows that the radius of the sphere of convergence is bounded below by $C\mu^\alpha$, where $C > 0$ and $\alpha > 1$.

**Lemma 5.8**

Under the nondegeneracy assumption, there exists $\hat{\mu} > 0$ and $K_1 > 0$ such that the radius of the sphere of convergence, $r_B(\mu)$, of Newton's method applied to the barrier system $F_B(x; \mu) = 0$, given by (5.6), satisfies

$$K_1\mu \leq r_B(\mu)$$

for all $\mu \leq \hat{\mu}$.

**Proof**

Let $\hat{\mu}$ and $K$ be given by the proof of Lemma 5.4. Consider $\mu \leq \hat{\mu}$ and let $x$ satisfy $\|x - x_\mu^*\| \leq K\mu$. We will prove the above result by showing that whenever the initial Newton iterate $x$ satisfies

$$\|x - x_\mu^*\| \leq K_1\mu$$

for $K_1 > 0$ and $\mu \leq \hat{\mu}$, then Newton's method will converge to $x_\mu^*$. As previously mentioned,

$$x_+ - x_\mu^* = F'(x; \mu)^{-1} \left[ F(x_\mu^*; \mu) - F(x; \mu) - F'(x; \mu)(x_\mu^* - x) \right]$$

$$= F'(x; \mu)^{-1} R(x; \mu).$$

where $R$ is defined in (5.25). Substituting (5.22) for the inverse of the Jacobian and multiplying it with $R$ produces
\[ x_+ - x_\mu = \begin{bmatrix} U_B(x) & U_N(x) \end{bmatrix} \begin{bmatrix} O(\mu) & O(\mu) \\ O(\mu) & O(1) \end{bmatrix} \begin{bmatrix} U_B^T(x)R(x;\mu) \\ U_N^T(x)R(x;\mu) \end{bmatrix}, \]

and taking norms yields

\[ \|x_+ - x_\mu^*\| \leq \|O(\mu)U_B^T R + O(\mu)U_N^T R\| + \|O(\mu)U_B^T R + O(1)U_N^T R\|. \]

Substituting upper-bounds (5.28) and (5.29) for \(\|U_B^T R\|\) and \(\|U_N^T R\|\), respectively, we obtain

\[ \|x_+ - x_\mu^*\| \leq O(\mu)C_1 \frac{1}{\mu} \|x_\mu^* - x\|^2 + O(\mu)C_2 \frac{1}{\mu} \|x_\mu^* - x\|^2 + GC_2 \frac{1}{\mu} \|x_\mu^* - x\|^2 \]
\[ \leq W \frac{1}{\mu} \|x_\mu^* - x\|^2 \]

for constants \(G, W > 0\). Therefore, if the initial iterate \(x^0\) satisfies

\[ \|x^0 - x_\mu^*\| < K_1 \mu, \]

where \(K_1 = \min\{1/W, K\}\), we obtain convergence of the Newton sequence to \(x_\mu^*\).

\[ \square \]

**Lemma 5.9**

Assume \(B \neq \emptyset\). Under the nondegeneracy assumption, we have that for \(\mu \leq \hat{\mu}\), there exists \(K_2 > 0\) such that the radius of the sphere of convergence, \(r_B(\mu)\), of Newton's method applied to the barrier system \(F_B(x;\mu)\), given by (5.6), satisfies

\[ r_B(\mu) \leq K_2 \mu. \]

**Proof**

It suffices to show that there exists an \(x\) satisfying \(\|x - x_\mu^*\| \leq K_1 \mu\), where \(K_1\) is given in Lemma 5.4. From where Newton's method does not converge or is not
defined. From S. Wright’s result (5.24), we have that \( \| x^* - x^*_\mu \| \leq M \mu \) for small \( \mu \) and \( M > 0 \). We will demonstrate two cases regarding the relationship of \( K \) and \( M \) to prove our result. If \( M < K \), then \( \| x^* - x^*_\mu \| \leq K \mu \) for \( \mu \leq \mu \) and \( g_i(x^*) = 0 \). \( i \in \mathcal{B} \). Thus, we can take \( K_2 = M \). Now, assume \( K < M \). Choose \( \hat{\mu} \leq \mu \) such that \( \| x^* - x^*_\mu \| \leq K \hat{\mu} \). Then for all \( \mu \leq \hat{\mu} \), we have \( \| x^* - x^*_\mu \| \leq K \mu \) and \( g_i(x^*) = 0 \). \( i \in \mathcal{B} \).

In this case, we let \( K_2 = K \). Clearly, \( F_B(x; \mu) \) and \( F'_B(x; \mu) \) are undefined whenever \( g_i(x) = 0 \) for any \( i \).

\[\square\]

**Theorem 5.1** Under the nondegeneracy assumption, there exist constants \( \hat{\mu} \) and \( K_1, K_2 > 0 \) such that for \( \mu \leq \hat{\mu} \), the radius of the sphere of convergence, \( r_B(\mu) \), of Newton’s method applied to the barrier system \( F_B(x; \mu) = 0 \), given by (5.6), satisfies

\[ K_1 \mu \leq r_B(\mu) \leq K_2 \mu. \]

**Proof** Application of Lemmas 5.8 and 5.9 produces the result with \( \mu \leq \hat{\mu} \). \[\square\]

**5.4.3 Perturbed System**

In the previous section, we derived lower- and upper-bound results for the radius of the sphere of convergence of Newton’s method applied to the barrier system. Now, we investigate the behavior of the radius of the sphere of convergence of Newton’s method applied to the perturbed system as \( \mu \) decreases to zero.

Lemma 5.6 provides the existence of \( F'_B(x^*_\mu, z^*_\mu; \mu)^{-1} \) for \( 0 \leq \mu \leq \hat{\mu} \). Since \( (x^*_\mu, y^*_\mu) \) exists, and \( F'_B(x, z; \mu) \) is Lipschitz continuous and nonsingular at \( (x^*_\mu, z^*_\mu) \) the standard assumptions for Newton’s method are satisfied at \( (x^*_\mu, z^*_\mu) \). Hence, there exists a ball centered at \( (x^*_\mu, z^*_\mu) \) such that starting from any point in the ball Newton’s method will converge to the solution \( (x^*_\mu, z^*_\mu) \). The following theorem states a new result
showing that the radius of the sphere of convergence of Newton's method applied to the perturbed system does not decrease to zero as \( \mu \) decreases to zero.

**Theorem 5.2** Under the nondegeneracy assumption, there exist constants \( \hat{\mu}, D > 0 \) such that for all \( 0 \leq \mu \leq \hat{\mu} \), the radius of the sphere of convergence, \( r_P(\mu) \), of Newton's method applied to the perturbed system \( F_P(x, z; \mu) = 0 \), given by (5.7), is bounded away from zero, that is,

\[
D \leq r_P(\mu).
\]

**Proof** We will prove the above result by showing that the Newton iterates converge to the solution \( (x^*_\mu, z^*_\mu) \) if the initial Newton iterate \( (x^0, z^0) \) satisfies

\[
\left\| \begin{pmatrix} x^0 - x^*_\mu \\ z^0 - z^*_\mu \end{pmatrix} \right\| \leq D \tag{5.37}
\]

for some constant \( D > 0 \). Consider \( 0 \leq \mu \leq \hat{\mu} \). Assume Newton's method is applied to the perturbed system for a fixed \( \mu \). Let \( (x, z) \) and \( (x^*_\mu, z^*_\mu) \) respectively denote the current Newton iterate and the solution for system (5.7) for a given value of \( \mu \). Then at the subsequent Newton iteration, we have

\[
\begin{pmatrix} x^* - x^*_\mu \\ z^* - z^*_\mu \end{pmatrix} = F'_P(x, z; \mu)^{-1} \left[ F_P(x^*_\mu, z^*_\mu; \mu) - F_P(x, z; \mu) \right.
\]

\[
- F'_P(x, z; \mu) \{(x^*_\mu, z^*_\mu) - (x, z)\} \tag{5.38}
\]

Taking norms on both sides of (5.38), we have
\[
\left\| \begin{pmatrix} x^+ - x^*_{\mu} \\ z^+ - z^*_{\mu} \end{pmatrix} \right\| \leq \left\| F'_p(x, z; \mu)^{-1} \right\| \times \\
\left\| F_p(x^*_\mu, z^*_\mu; \mu) - F_p(x, z; \mu) - F'_p(x, z; \mu) \left\{ \begin{pmatrix} x^*_{\mu} - x \\ z^*_{\mu} - z \end{pmatrix} \right\} \right\|.
\]

Since \(\left\| F'_p(x, z; \mu)^{-1} \right\| \leq M\) for some \(M > 0\) then a straightforward application of Lemma 5.3 yields that for some constant \(N > 0\)

\[
\left\| \begin{pmatrix} x^+ - x^*_{\mu} \\ z^+ - z^*_{\mu} \end{pmatrix} \right\| \leq \frac{MN}{2} \left\| \begin{pmatrix} x - x^*_{\mu} \\ z - z^*_{\mu} \end{pmatrix} \right\|^2.
\]

To obtain convergence of the Newton sequence in \((x, z)\) to \((x^*_{\mu}, z^*_{\mu})\), it is sufficient to have

\[
\left\| \begin{pmatrix} x^0 - x^*_{\mu} \\ z^0 - z^*_{\mu} \end{pmatrix} \right\| \leq \frac{2}{MN}.
\] (5.39)

Therefore, \(D \leq r_p(\mu)\) where \(0 < D \leq \frac{1}{MN}\).

\[
\square
\]

5.5 Numerical Experiments Concerning the Sphere of Convergence

We have established that the radius of the sphere of convergence of Newton's method applied to the barrier system (5.6) for the nonlinear program (5.1) decreases to zero as \(\mu\) decreases to zero. However, the radius of the sphere of convergence of Newton's
method for the equivalent perturbed system (5.7) stays bounded away from zero as \( \mu \) decreases to zero. In this section, we obtain numerical upper-bound estimates for the radius of the sphere of convergence of Newton's method on the two equivalent systems for various values of \( \mu > 0 \).

Our numerical experiments were conducted similar to the manner described in Section 4.5 for the linear program. We briefly describe the new modifications, as well as, the unaltered portions of the algorithm to present a complete description of the algorithm.

The numerical upper-bound estimates were conducted as follows. We considered the sphere of convergence of Newton's method for the two equivalent systems, that is, the half-sphere of convergence was not considered for the barrier system as in the linear programming case. Initial points for Newton's method were of the form

\[ v^0 = v^*_\mu + \lambda v_N, \]

where \( v_N \) was chosen using the Matlab function \texttt{rand}, which selects random entries from a uniform distribution on the unit interval, \((0,1)\). The point \( v^*_\mu \) is the solution \( x^*_\mu \) and \((x^*_\mu, z^*_\mu)\) for systems (5.6) and (5.7), respectively. The initial point \( v^0 \) was merely \( x^0 \) for Newton's method on the barrier system. The first \( n \) components of the initial point \( v^0 \) for the perturbed system composed the point \( x^0 \) and the remaining components comprised the point \( z^0 \) for Newton's method.

The experiments were begun with initial points \( v^0 \) close to \( v^*_\mu \), that is, \( \lambda \) was initialized close to zero. Full Newton steps were always taken. If for a given \( \lambda > 0 \) Newton's method generates an iterate that satisfies a convergence criterion, then \( \lambda \) was incremented and Newton's method was started with a new initial point \( v^0 \). The convergence criterion was
\[
\frac{\|x_\mu^* - x_k\|}{1 + \|x_\mu^*\|} + \frac{\|z_\mu^* - z_k\|}{1 + \|z_\mu^*\|} \leq tol
\] (5.40)

where \((z_k)_i = \mu / g_i(x_k)\) for the barrier system for \(i = 1, \ldots, m\), and the convergence tolerance was set to \(tol = 10^{-8}\). Nonconvergence of Newton's method was recorded when the maximum number of iterations, which was set to 100, was reached. In this case, we observed that the final iterate was infeasible for the two equivalent systems. As in the linear programming case, we observed also that infeasible Newton iterates did not preclude convergence to the solution of the perturbed system. However, convergence was precluded for Newton's method on the barrier system if the iterates became infeasible: as a result, the Jacobian became increasingly ill-conditioned.

Ten unit random directions \(v_\alpha\) having positive and negative components were generated. For each direction, \(v_\alpha\), Newton's method was applied to the barrier and perturbed systems until the convergence criteria (5.40) was met at some \(\lambda_\alpha\). For a given value of \(\mu > 0\), the numerical upper-bound estimate for the radius of the sphere of convergence of Newton's method applied to the barrier or perturbed system was recorded as the minimum of the ten \(\lambda_\alpha\)'s, that is,

\[
\min_\alpha \{\lambda_\alpha\},
\]

where \(\lambda_\alpha = \|x_0 - x_\mu^*\|\) for the barrier system and \(\lambda_\alpha = \|v_0 - v_\mu^*\|\) for the perturbed system. Newton's method was applied to the barrier and perturbed systems with the following parameterized values of \(\mu\):

\[
\mu = 0.45, 0.25, 0.10, 0.05, 0.01, 0.0075, 0.005, 0.0005, 0.00005 \quad (5.41)
\]
as in the experiments for the linear program (see Section 4.5).
A subset of nondegenerate problems were considered from the Hock and Schittkowski test set [9]. Table 5.1 shows the problem numbers in the first column, followed by the number of variables and inequality constraints. The problems were run on a Sun Ultra Sparc workstation using Matlab version 5.2.

<table>
<thead>
<tr>
<th>Problem</th>
<th>n</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS 10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>HS 11</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>HS 12</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>HS 16</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>HS 17</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>HS 22</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>HS 33</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>HS 34</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>HS 43</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

The theoretical results establish that the radius of the sphere of convergence of Newton's method applied to the perturbed system is bounded away from zero as \( \mu \) decreases to zero, however the radius associated with the barrier system decreases to zero with the same order as \( \mu \) decreases to zero. Now, we present several numerical experiments conducted for the radius of the sphere of convergence of Newton's method. We show only four problems since the results are similar for the other problems that are not presented. Figures 5.1 and 5.2 show the radius of the sphere of convergence associated with the two equivalent systems graphed against the values of \( \mu \) given in (5.41). The numerical experiments show the radius of the sphere of convergence of Newton's method on the barrier system decreasing towards zero for small values of \( \mu > 0 \). However, the radius associated with the perturbed system clearly stays away from zero. In addition, the radius associated with the perturbed system is also larger.
Figure 5.1  Radii of the spheres of convergence for Newton’s method applied to the barrier and perturbed systems

Figure 5.2  Radii of the spheres of convergence for Newton’s method applied to the barrier and perturbed systems
than the radius associated with the barrier system for all $\mu$ values. These numerical results confirm our theory for the radius of the sphere of convergence of Newton's method on these two equivalent systems for nondegenerate problems.
5.6 Conclusions

In this chapter, we studied an aspect of the local behavior of Newton's method on two commonly used equivalent systems: the barrier (5.6) and perturbed (5.7) systems. In particular, we focused our attention on analyzing the radius of the sphere of convergence of Newton's method applied to these two equivalent systems for nondegenerate problems. Our theoretical results establish that the radius of the sphere of convergence of Newton's method on the barrier system decreases to zero with the same order as \( \mu \) decreases to zero. Previous results [22] showed that the radius of the sphere of convergence of Newton's method on the barrier system was bounded below by \( O(\mu^{\alpha}) \) for \( \alpha > 1 \), but no upper-bound was obtained.

In addition, we presented numerical experiments concerning the radius of the sphere of convergence of Newton's method applied to the barrier and perturbed systems for nondegenerate problems. Our experiments reinforce the theoretical results obtained for the radius of the sphere of convergence of Newton's method. The numerical and theoretical results imply that the radius of the sphere of convergence of Newton's method for the perturbed system is larger than the radius associated with the barrier system, at least for small \( \mu \) values close to a solution at \( \mu = 0 \). As a consequence, we expect fewer Newton iterations to obtain a good approximate solution to the nonlinear program. Our theoretical and numerical results favor the use of the perturbed system. Thus, it is advantageous to use Newton primal-dual interior-point methods for solving the nonlinear program.
Chapter 6

Conclusions

Linear and nonlinear programming problems are often solved by formulating and solving a sequence of parameterized systems of nonlinear equations associated with the optimization problem. Thus, it is important to analyze differences between methods which incorporate these equivalent system formulations to determine the most desirable method. Two popular Newton interior-point methods have evolved in the optimization community for solving the linear and nonlinear program: the Newton log-barrier method and the Newton primal-dual interior-point method.

The two methods apply Newton's method to two different but equivalent systems. The Newton log-barrier method can be viewed as applying Newton's method to the system of equations (barrier system) arising from the log-barrier formulation of the optimization problem. Similarly, the Newton primal-dual method can be viewed as applying Newton's method to the system of equations (perturbed system) obtained from the perturbed optimality conditions of the optimization problem. Though both methods deal with different but equivalent systems, El-Bakry et al. [4] showed that Newton's method behaves differently for the two different applications.

In this thesis, we extended the study of these two Newton interior-point methods to observe additional differences in behavior between the two methods. We investigated the global and local behavior of the Newton log-barrier and Newton primal-dual methods. On the issue of local behavior, we studied the radius of the sphere of convergence of Newton's method applied to the barrier and perturbed systems as the barrier parameter, $\mu$, decreased to zero for nondegenerate problems. For the linear and inequality constrained optimization problems, the following was shown. The
radius of the sphere of convergence of Newton's method applied to the perturbed system is bounded away from zero for $\mu > 0$. However, the radius of the sphere of convergence associated with the barrier system decreases to zero with the same order as $\mu$ decreases to zero. The latter result is sharper than S. Wright's result [22], who provides a lower-bound of $O(\mu^\alpha)$ for $\alpha > 1$ for the inequality constrained optimization problem. Numerical results were also presented which were in alignment with our theory.

In addition, we showed through numerical experiments that the Newton log-barrier and Newton primal-dual methods exhibit different behavior on the two equivalent systems for the linear program. In particular, the Newton log-barrier method requires significantly more iterations to solve the linear program. The iterates must also follow the central path closely in order to converge to a solution. On the contrary, the Newton primal-dual method always obtained a solution in a fewer number of iterations. Furthermore, fewer iterations were required if the iterates did not follow the central path closely.

Altogether, these results show that Newton primal-dual interior-point methods are the method of choice for solving nonlinear programs. Our study on the radius of the sphere of convergence of Newton's method on the perturbed system and the efficient performance of our Newton primal-dual algorithm results in an important implication: fewer iterations are required for solving the nonlinear program using a Newton primal-dual method.

Our results provide insight into the effective performance of Newton primal-dual methods, especially in the case of the linear program. The results from this thesis will contribute to the understanding and development of efficient Newton primal-dual interior-point algorithms for the nonlinear program.
Bibliography


