

SENSITIVITY ANALYSIS AND THE ANALYTIC  
CENTRAL PATH

by

Allen Holder

M.S., University of Southern Mississippi, 1993

B.S., University of Southern Mississippi, 1990

A thesis submitted to the  
University of Colorado at Denver  
in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Applied Mathematics

1998

This thesis for the Doctor of Philosophy

degree by

Allen Holder

has been approved

by

---

Harvey Greenberg

---

Stephen Billups

---

Gary Kochenberger

---

Rich Lundgren

---

Burt Simon

---

Date

Holder, Allen (Ph.D., Applied Mathematics)

Sensitivity Analysis and the Analytic Central Path

Thesis directed by Professor Harvey Greenberg

## ABSTRACT

The analytic central path for linear programming has been studied because of its desirable convergence properties. This dissertation presents a detailed study of the analytic central path under perturbation of both the right-hand side and cost vectors for a linear program. The analysis is divided into three parts: extensions of results required by the convergence analysis when the data is unperturbed to include that case of data perturbation, marginal analysis of the analytic center solution with respect to linear changes in the right-hand side, and parametric analysis of the analytic central path under simultaneous changes in both the right-hand side and cost vectors.

To extend the established convergence results when the data is fixed, it is first shown that the union of the elements comprising a portion of the perturbed analytic central paths is bounded. This guarantees the existence of subsequences that converge, but these subsequences are not guaranteed to have the same limit without further restrictions on the data movement. Sufficient conditions are provided to insure that the limit is the analytic center of the

limiting polytope. Furthermore, as long as the data converges and the parameter of the path is approaching zero, certain components of the analytic central path are forced to zero.

Since the introduction of the analytic center to the mathematical programming community, the analytic central path has been known to be analytic in both the right-hand side and cost vectors. However, since the objective function is a continuous, piece-wise linear function of the right-hand side, the analytic center solution is not differentiable. We show that this solution is continuous and is infinitely, continuously, one-sided differentiable. Furthermore, the analytic center solution is analytic if the direction of right-hand side change is contained in a particular space. Uniform bounds for the first order derivatives are presented.

The parametric analysis of the analytic central path follows as a consequence of characterizing when the parameterized analytic center converges. The development allows for simultaneous changes in the right-hand side and cost vectors, provided the cost perturbation is linear. Although the analytic center solution is not a continuous function of the cost vector, the analytic central path is continuous when viewed as a set.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed \_\_\_\_\_  
Harvey Greenberg

## DEDICATION

This dissertaion is dedicated to my wife for her constant love and support.

## ACKNOWLEDGMENTS

My deepest gratitude is given to Dr. Harvey Greenberg; for without his support and encouragement this dissertation would never have been completed. His mentoring over the last few years has been invaluable. Dr. Stephen Billups' meticulous review is much appreciated. In addition, I would like to thank: Dr. Richard J. Caron, Markus Emserman, Caroline Heberton, Dr. Gray Kochenberger, Dave Krutsinger, Dr. Richard Lundgren, Francis Newman, Dr. Kees Roos, Dr. Burt Simon, Dr. Jeff Stuart, Dr. Jos Sturm, Lexie Taylor, Dr. Tamas Terlaky, Dr. Radek Tezaur, Dr. Ronald VanIwaarden, John Wilson, and Dr. Shuzhong Zhang.

Last, but not least, I would like to thank all those who have made the Saturday afternoon flag football games possible.

## CONTENTS

<u>Figures</u> . . . . .	i
<u>Tables</u> . . . . .	i
1. Introduction . . . . .	1
1.1 Sensitivity Analysis . . . . .	1
1.2 Basic Notation . . . . .	2
1.3 Thesis Outline . . . . .	5
2. Central Paths . . . . .	7
2.1 A Brief Historical View . . . . .	7
2.2 Analytic Centers . . . . .	8
2.3 Basic Properties of the Omega Central Path . . . . .	21
2.4 Analytic Centers and Data Perturbations . . . . .	34
2.5 Chapter Summary . . . . .	47
3. Marginal Analysis of The Analytic Center Solution . . . . .	49
3.1 Introduction . . . . .	49
3.2 Notation, Terminology, and Preliminary Results . . . . .	51
3.3 Differentiating the Central Path . . . . .	53
3.4 Marginal Properties of the Analytic Center Solution with Respect to the Right-Hand Side . . . . .	63
3.5 Bounding the Derivatives of the Analytic Center Solution . . . . .	78
3.6 Chapter Summary . . . . .	87

4. The Omega Central Path and Simultaneous Rim Data Perturbations . . . . .	88
4.1 Introduction . . . . .	88
4.2 Convergence Under Linear Cost Changes . . . . .	89
4.3 Equivalent Dual Statements . . . . .	114
4.4 Central Path Convergence Under General Rim Data Changes . . . . .	134
4.5 Chapter Summary . . . . .	152
5. Conclusion and Avenues for Future Research . . . . .	154
<u>Notation Index</u> . . . . .	156
<u>References</u> . . . . .	160

## FIGURES

Figure		
2.1	A central path. . . . .	34
3.1	An example of tracking the analytic center and its marginal derivatives. . . . .	78
4.1	The geometry of central path convergence. . . . .	94
4.2	Two different interpretations of an analytic center. . . . .	100

## TABLES

### Table

2.1	A synopsis of analytic properties of the omega central path . . .	22
3.1	Some higher order limiting derivatives from example 2.12 . . . .	63

## 1. Introduction

### 1.1 Sensitivity Analysis

As defined in the *Mathematical Programming Glossary* [30], sensitivity analysis is

The concern with how the solution changes if some changes are made in either the data or in some of the solution values (by fixing their value). Marginal analysis is concerned with the effects of small perturbations, maybe measurable by derivatives. Parametric analysis is concerned with larger changes in parameter values that affect the data in the mathematical program, such as a cost coefficient or resource limit.

The first papers concerned with sensitivity analysis for linear programs were published in the middle to late 1950s. The original two papers were authored by S. Gass and T. Saaty [20, 77], and the topic was parametric behavior of the objective function. H. Mills published the first paper on marginal analysis and linear programs in 1956 [56], and T. Saaty had the second work on this topic in 1959 [76]. The first English monograph on the topic was published by T. Gal [19] in 1973.

The study of sensitivity analysis is paramount for many reasons; two of the most important are:

- (1) such investigations provide insight into both the mathematical statement of the problem and its solution, and
- (2) sensitivity analysis connects the certainty of the mathematical model with the uncertainty of the real world situation it is modeling.

With regard to the latter statement, H. Zimmermann [19] wrote that sensitivity analysis is “the bridge between pure dissemination of information and decision making.”

In the definition, the solution is usually interpreted as either the optimal value of a variable, or the optimal objective value. The differences in studying how the optimal objective function and a optimal value rely on the data are possibly significant. For example, suppose  $f(x)$  is the objective function and  $z^* = f(x^1) = f(x^2)$  is the optimal objective value. A data perturbation may cause  $x^1$  to remain optimal and  $x^2$  to become non-optimal. In such a case, the optimal objective function value may be invariant under such a perturbation, but the set of optimal elements is not. Such discrepancies lead to a plethora of questions for a practitioner of optimization.

## 1.2 Basic Notation

The following notation is used throughout this work. A notation index is included and the *Mathematical Programming Glossary* [30] is useful to

any reader when faced with an unknown concept. Some definitions from linear algebra [48] are presented first. Let  $A$  be an  $m \times n$  matrix. The row, column, null, and leftnull spaces of  $A$  are defined as follows:

$$\text{row}(A) \equiv \{yA : y \in \mathbb{R}^m\}$$

$$\text{col}(A) \equiv \{Ax : x \in \mathbb{R}^n\}$$

$$\text{null}(A) \equiv \{x \in \mathbb{R}^n : Ax = 0\}$$

$$\text{leftnull}(A) \equiv \{y \in \mathbb{R}^m : yA = 0\}.$$

It is well known [48] that  $\text{row}(A) \perp \text{null}(A)$  and  $\text{col}(A) \perp \text{leftnull}(A)$ . Generalized inverses are used to establish many results. A generalized inverse of  $A$  is any  $n \times m$  matrix, denoted  $A^+$ , for which

$$AA^+A = A \text{ and}$$

$$A^+AA^+ = A^+.$$

The linear system of equations  $Ax = b$  has a solution if, and only if,  $AA^+b = b$ .

Furthermore, the solutions have the form

$$x = A^+b + (I - A^+A)z,$$

for any  $z \in \mathbb{R}^n$ . If the generalized inverse has the additional property that  $AA^+$  and  $A^+A$  are symmetric, the generalized inverse is unique and is known

as the *Moore-Penrose* generalized inverse. The Moore-Penrose generalized inverse allows easy representations of the projections onto  $\text{row}(A)$  and  $\text{null}(A)$ , which are needed in Chapter 4. Specifically, the projection of  $x \in \mathbb{R}^n$  onto  $\text{row}(A)$  is  $\text{proj}_{\text{row}(A)}x = AA^+x$ , and the projection of  $x \in \mathbb{R}^n$  onto  $\text{null}(A)$  is  $\text{proj}_{\text{null}(A)}x = (I - A^+A)x$ . For any matrix  $A$ ,  $A^+$  is always assumed to be the Moore-Penrose generalized inverse.

Sequences are enumerated with superscripts. For example,  $\{x^k \in W\}$  is a sequence of elements from the set  $W$ . The convergence of such a sequence is shown by  $\{x^k \in W\} \rightarrow x \in V$ , which means the sequence converges to the element  $x$  in  $V$ . If the elements of the sequence are vectors, the components of the vectors are denoted with the corresponding superscript. For example,  $\{z^k \in \mathbb{R}^n\} = \{(w, v)^k \in \mathbb{R}^n\} = \{(w^k, v^k) \in \mathbb{R}^n\}$ , where  $w$  and  $v$  indicate a partitioning of  $z$ .

A *polyhedron* is a set of the form  $\{x : Ax \leq b\}$ , where  $A$  is an  $m \times n$  matrix and  $b$  is a vector of length  $m$ . A bounded polyhedron is called a *polytope*.

Much of this work relies on previous results. For completeness, proofs of these results are included, and the author and reference in which the result and proof are found is cited in the label of the statement. If multiple researchers proved the same result, these references are made in the few sentences precluding the statement of the result.

### 1.3 Thesis Outline

Chapter 2 begins by establishing the major definitions and any needed background results. The manner in which these results rely on the problem data has not been investigated, and the chapter proceeds by extending some of these results to the case when certain data perturbations are allowed. An example used to exemplify many concepts is developed.

Chapter 3 is concerned with the marginal analysis of the analytic center solution and the subsequent information that this analysis yields about the entire central path. Since differentiating the elements of the central path is crucial to the development of these results, a complete investigation of the differential properties of the central path is presented. This analysis was developed by Güler [63] and is extended to show that the analytic center solution is one-sided, infinitely, continuously differentiable. Furthermore, a set of directions along which the analytic center solution is analytic is classified. The chapter ends by providing uniform bounds for the various derivatives.

Chapter 4 contains an analysis of the central path under simultaneous changes in the cost coefficients and the right-hand side. The questions answered in this chapter show precisely how the central path relies on the rim data of the linear program. The chapter begins by presenting the analysis for linear

changes in the cost coefficients together with arbitrary changes in the right-hand side. In an attempt to extend these results to the situation of arbitrary, simultaneous changes in the cost coefficients and right-hand side vectors, two new types of convergence are defined. These types of convergence show that the central paths naturally induce an equivalence relation on the set of admissible cost vectors. The chapter concludes by presenting results showing the difficulty of guaranteeing convergence under arbitrary, simultaneous changes in the rim data.

## 2. Central Paths

### 2.1 A Brief Historical View

In 1979, Khachiyan [42] showed that the class of linear programming problems is solvable in polynomial time. The algorithm presented by Khachiyan was an *interior point* algorithm, meaning that it produced points in the relative interior of the feasible region. Although the theoretical, worst-case complexity of this algorithm was provably superior to that of the simplex algorithm, implementation showed no such realistic advantage [91]. This was not the first interior point algorithm presented for linear programming. In 1967, Huard [36] suggested using a method of centers to solve mathematical programs problems. Even though Huard proved that these algorithms converge to an optimal solution, the success of Dantzig's simplex method apparently thwarted any attempt of implementation for linear programming at the time. During the same time period in the Soviet Union, Dikin [10, 11] developed another interior point algorithm which was implemented and used to solve economic problems. Although Dikin's method has a simple interpretation as steepest descent on a scaled space, the polynomiality of this algorithm is still not proven and it is

not believed to be polynomial. In 1968 the seminal book by Fiacco and McCormick [13] was published and the term “interior point method” was coined. This text contains the first real development of the theory and use of penalty and barrier methods. In 1984, Karmarkar [41] developed another polynomial time algorithm for linear programs and claimed that this algorithm would be far superior computationally to the simplex algorithm. The mathematical programming community was rather shocked by these claims [46]. Some delved into the theory and implementation of this algorithm to investigate the validity of these claims. It is now understood that interior point methods are viable algorithms and appear computationally superior to simplex based approaches when the problems are large [49, 50, 51, 52].

In 1986, Sonnevend [80] introduced the mathematical programming community to the concept of the *analytic center*. In [81, 82, 83, 84, 85], Sonnevend connects this concept to Karmarkar’s algorithm, develops his own algorithm, demonstrates applications, and together with Stoer, shows several complexity results. It is with this idea that we begin our development.

## 2.2 Analytic Centers

Consider the polyhedron  $\mathcal{P} = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$  and its strict interior  $\mathcal{P}^\circ = \{x \in \mathbb{R}^n : Ax = b, x > 0\}$ . For any positive vector

$\omega \in \mathbb{R}^n$ , define

$$p_\omega(x) : \mathcal{P}^\circ \rightarrow \mathbb{R} : x \rightarrow \sum_{i=1}^n \omega_i \ln(x_i).$$

The standard notation that capital letters indicate the diagonal matrix of the corresponding vector is employed. For example,  $\Omega = \text{diag}(\omega_1, \omega_2, \dots, \omega_n)$  and  $X = \text{diag}(x_1, x_2, \dots, x_n)$ . Since,  $\nabla^2 p_\omega(x) = -\Omega X^{-2}$ , which is negative definite,  $p_\omega(x)$  is strictly concave. So, if  $\mathcal{P}$  is bounded, there exists a unique maximizer of  $p_\omega(x)$  over  $\mathcal{P}^\circ$ .

**Definition 2.1** Let  $\mathcal{P} = \{x : Ax = b, x \geq 0\}$  be a polytope with non-empty relative interior, and  $\omega \in \mathbb{R}_{++}^n$ . The unique maximizer of  $p_\omega(x)$  is the omega analytic center of  $\mathcal{P}$ . If  $\omega = e$ , which is the vector of ones, then the omega analytic center is simply called the analytic center of  $\mathcal{P}$ .

The concept of the analytic center is fundamental to algorithmically attaining the known polynomiality of the class of linear programs. In fact, all known polynomial algorithms for solving linear programs incorporate a centering component in their search directions. There are interior point methods that do not use a centering component in their search direction, such as Dikin's original algorithm, and the polynomiality of these algorithms is not known and generally not believed [33].

It has been shown by Atkinson and Vaidya [3], Freund [15], Goffin and Vial [21], and Roos and Hertog [72] that the omega analytic center can be found in polynomial time. The usefulness of adjusting  $\omega$  to solve linear programming problems and linear complementarity problems is investigated in [1, 55, 57]. The most important reason, from an algorithmic perspective, is that if *any* strictly interior point is known, it is possible to find an  $\omega$  that makes this feasible point a “good” starting point. Furthermore, the basic idea of weighting the terms of  $p_\omega(x)$  has led to what are known as “target following methods” for solving linear programs [37, 73]. Another by-product of the flexibility to adjust the components of  $\omega$  is found in theorem 2.2, which states that every point in the relative interior of a polytope is representable as an omega analytic center.

**Theorem 2.2** Let  $\mathcal{P} = \{x : Ax = b, x \geq 0\}$  be a polytope with non-empty strict interior, and let  $\bar{x} \in \mathcal{P}^\circ$ . Then there exists  $\omega \in \mathbb{R}_{++}^n$  such that

$$\{\bar{x}\} = \operatorname{argmax}\{p_\omega(x) : x \in \mathcal{P}^\circ\}.$$

**Proof:** Let  $\bar{x} \in \mathcal{P}^\circ$ . Since

$$\max\{p_\omega(x) : x \in \mathcal{P}^\circ\}$$

is a convex program, the Lagrange condition, that there exists  $y$  satisfying

$$\nabla p_\omega(x) = \omega^T X^{-1} - yA = 0,$$

is necessary and sufficient for  $x$  to be a maximizer of  $p_\omega$  over  $x \in \mathcal{P}_b^o$ . Noticing that with an appropriate choice of  $\omega$ ,  $\omega^T \bar{X}^{-1}$  can be any positive vector, all that must be shown is that  $\text{row}(A)$  contains a strictly positive element. The boundedness assumption implies there does not exist a solution to the system  $Ax = 0$ ,  $x \geq 0$ , and  $x \neq 0$ . Hence, Gordon's theorem of the alternative implies there does exist a solution to the system  $yA > 0$ . ■

Associated with any linear program, for which  $\mathcal{P}^o \neq \emptyset$ , is a geometric structure called the omega central path or omega central trajectory. As is now developed, the central path is an infinitely smooth path of analytic centers. Throughout, the following “standard form” linear program and its associated dual are considered,

$$LP : \min\{cx : Ax = b, x \geq 0\} \quad \text{and} \quad LD : \max\{yb : yA + s = c, s \geq 0\},$$

where  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ ,  $c \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ ,  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$ ,  $s \in \mathbb{R}^n$  and  $c$ ,  $y$ , and  $s$  are row vectors. The assumption that the  $\text{rank}(A) = m$  is made, which has the implications that  $y$  and  $s$  are related by a one-to-one linear mapping and that the mapping  $x \mapsto Ax$  is onto. Because  $y$  and  $s$  are related in a one-to-one fashion, a dual element may be referred to by  $(y, s)$ ,  $y$ , or  $s$ . The rank assumption is made without loss in generality because if  $A$  did not have full

row rank, row reduction could be used to form an  $A'$  and  $b'$ , such that  $A'$  has full row rank and  $\mathcal{P} = \{x : A'x = b', x \geq 0\}$ . The  $b$  vector is referred to as the *right-hand side* vector and the  $c$  vector is referred to as the *cost* vector. Together,  $b$  and  $c$  are called the *rim data* for the linear programs stated above, and notationally  $r = (b, c)$  is a rim data instance.

Defining  $z^*$  to be the common optimal value of  $LP$  and  $LD$ , the following set conventions are made:

$$\mathcal{P}_b \equiv \{x : Ax = b, x \geq 0\}$$

$$\mathcal{D}_c \equiv \{(y, s) : yA + s = c, s \geq 0\}$$

$$\mathcal{P}_b^o \equiv \{x : Ax = b, x > 0\}$$

$$\mathcal{D}_c^o \equiv \{(y, s) : yA + s = c, s > 0\}$$

$$\mathcal{P}_r^* \equiv \{x \in \mathcal{P} : cx = z^*\}$$

$$\mathcal{D}_r^* \equiv \{(y, s) \in \mathcal{D} : yb = z^*\}.$$

The optimality sets,  $\mathcal{P}_r^*$  and  $\mathcal{D}_r^*$ , have another important representation that relies upon the *optimal partition*. The optimal partition is induced by a *strictly complementary* optimal solution to  $LP$  and  $LD$ , which is any primal-dual pair,  $x^*$  and  $(y^*, s^*)$ , such that  $s^*x^* = 0$  and  $(s^*)^T + x^* > 0$ . Such solutions have been known to exist since 1956, when Goldman and Tucker published this result in [22]. Assuming that  $x^*$  and  $(y^*, s^*)$  are strictly complementary optimal

solutions for the rim data instance  $r$ , define the optimal partition by the index sets  $B(r)$  and  $N(r)$ :

$$B(r) = \{i : x_i^*(r) > 0\} \text{ and} \quad (2.1)$$

$$N(r) = \{1, 2, 3, \dots, n\} \setminus B(r). \quad (2.2)$$

The dependence of the optimal partition on the rim data is used only when clarity is warranted, and when the rim data instance is understood, the optimal partition is denoted  $(B|N)$ . Vectors and matrices are often decomposed into sub-vectors and sub-matrices corresponding to the optimal partition. To facilitate this, two standard notations are used in conjunction with set subscripts. First, a set subscript attached to a vector indicates the sub-vector corresponding to the elements contained in the set. Second, a set subscript on a matrix is used to denote the sub-matrix whose columns are indicated by the set. For example, if  $(B|N)$  is the optimal partition for the rim data instance  $r$ ,

$$Ax = A_B x_B + A_N x_N = b.$$

This notation and the complementarity condition imply that

$$\begin{aligned} \mathcal{P}_r^* &= \{x \in \mathcal{P} : x_N = 0\} \\ &= \{x : A_B x_B = b, x_B \geq 0, x_N = 0\} \end{aligned} \quad (2.3)$$

and

$$\mathcal{D}_r^* = \{(y, s) \in \mathcal{D} : s_B = 0\} \quad (2.4)$$

$$= \{(y, s) : yA_B = c_B, yA_N + s_N = c_N, s_N \geq 0\}. \quad (2.5)$$

For a fixed  $A$  matrix, define the following sets of *admissible rim data*,

$$\mathcal{G} \equiv \{r = (b, c) \in \mathbb{R}^m \times \mathbb{R}^n : \mathcal{P}_b^o \neq \emptyset, \mathcal{D}_c^o \neq \emptyset\},$$

$$\mathcal{G}_b \equiv \{b \in \mathbb{R}^m : \mathcal{P}_b^o \neq \emptyset\}, \text{ and}$$

$$\mathcal{G}_c \equiv \{c \in \mathbb{R}^n : \mathcal{D}_c^o \neq \emptyset\}.$$

The above definitions do not correspond to the traditional definition of admissible, which usually means that the defined linear programs have a finite optimal solution. Here the definition is more restrictive, in that only rim data for which  $\mathcal{P}_b^o$  and  $\mathcal{D}_c^o$  are not empty is included. The next theorem establishes that  $\mathcal{G}$  is open. This is important because this guarantees that a perturbed admissible rim data element remains admissible.

**Theorem 2.3**  $\mathcal{G}$  is an open set.

**Proof:** Let  $(\hat{b}, \hat{c}) \in \mathcal{G}$ . Then there exists  $\hat{x}$  and  $(\hat{y}, \hat{s})$  such that  $A\hat{x} = \hat{b}$ ,  $\hat{x} > 0$ ,  $\hat{y}A + \hat{s} = \hat{c}$ , and  $\hat{s} > 0$ . Let  $U$  be an open set in  $\mathbb{R}^n$  which contains  $\hat{x}$  and has the property that  $x \in U$  implies  $x > 0$ . Since the rank of  $A$  is  $m$ , the linear transformation  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \rightarrow Ax$  is onto. Furthermore, since

$T$  is a continuous mapping, the open mapping theorem [75] implies that  $T(U)$  is open. Let  $\epsilon = \min\{\hat{s}_i : i = 1, 2, \dots, m\}$ , and define  $V = \{c : \|c - \hat{c}\| < \epsilon\}$ . Then,  $(\hat{b}, \hat{c}) \in T(U) \times V \subset \mathcal{G}$ , and the result follows since  $T(U) \times V$  is open. ■

The next result shows that the elements of  $\mathcal{G}$  guarantee that the primal and dual elements having a bounded duality gap are bounded [64, 71]. Notationally, for any rim element  $r$ , and any  $M \in \mathbb{R}_+$ , define

$$\mathcal{L}(r, M) = \{(x, (y, s)) \in \mathcal{P}_b \times \mathcal{D}_c : sx \leq M\}.$$

**Theorem 2.4 (Roos and Vial [74])** Let  $r \in \mathcal{G}$ . Then,  $\mathcal{L}(r, M)$  is bounded for all  $M \geq 0$ .

**Proof:** Let  $M \in \mathbb{R}_+$  and  $(\hat{x}, (\hat{y}, \hat{s})) \in \mathcal{P}_b^o \times \mathcal{D}_c^o$ . Then, for any primal and dual elements,  $x$  and  $s$ , we have  $\hat{x} - x \in \text{null}(A)$ ,  $\hat{s} - s \in \text{row}(A)$ , and

$$0 = (\hat{s} - s)(\hat{x} - x) = \hat{s}\hat{x} - s\hat{x} - \hat{s}x + sx. \quad (2.6)$$

Hence, if  $(x, (y, s)) \in \mathcal{L}(r, M)$ ,

$$\begin{aligned} \hat{s}_i x_i &\leq \hat{s}x + s\hat{x} \\ &= \hat{s}\hat{x} + sx \\ &\leq \hat{s}\hat{x} + M, \end{aligned}$$

where the first inequality follows from non-negativity and the first equality follows from equation 2.6. This implies that  $0 \leq x_i \leq \frac{\hat{s}\hat{x} + M}{\hat{s}_i}$ . A similar argument

shows that  $s$  is bounded.

All that remains to be shown is that  $y$  is bounded. The full rank assumption of  $A$  implies

$$y = (c - s)(AA^T)^{-1},$$

and

$$\|y\| = \|(c - s)(AA^T)^{-1}\| \leq \|c - s\| \|(AA^T)^{-1}\|.$$

The boundedness of  $y$  now follows since  $\|c - s\|$  is bounded. ■

**Corollary 2.5** If  $r \in \mathcal{G}$ , then  $\mathcal{P}_r^*$  and  $\mathcal{D}_r^*$  are bounded.

The converse of theorem 2.4 is also true [53, 54]. Hence,  $\mathcal{G}$  is precisely the set of rim data that guarantees the boundedness of  $\mathcal{L}(r, M)$ . Corollary 2.5 implies that both  $\mathcal{P}_r^*$  and  $\mathcal{D}_r^*$  have omega analytic centers. These omega analytic centers are  $x^*(r)$  and  $(y^*(r), s^*(r))$ , respectively, and

$$\begin{aligned} \{x^*(r)\} &= \operatorname{argmax} \left\{ \sum_{i \in B} \omega_i \ln(x_i) : x \in \mathcal{P}_r^* \right\} \text{ and} \\ \{(y^*(r), s^*(r))\} &= \operatorname{argmax} \left\{ \sum_{i \in N} \omega_i \ln(s_i) : (y, s) \in \mathcal{D}_r^* \right\}. \end{aligned}$$

The development of the central path is completed by considering the penalized linear program,

$$\min\{cx - \mu p_\omega(x) : Ax = b, x > 0\}. \tag{2.7}$$

Since the Hessian of the objective function is  $\mu\Omega X^{-2}$ , this mathematical program is strictly convex. So, if a solution exists it is unique. To show that there exists a solution for each  $\mu > 0$ , consider the function

$$h(x) = s^0 x - \mu \sum_{i=1}^n \omega_i \ln(x_i) = cx - \mu p_\omega(x) - y^0 b,$$

where  $(y^0, s^0) \in \mathcal{D}_c$ . Elementary calculus calculations show for any  $\mu > 0$ ,  $h(x)$  attains its global minimum over  $\mathbb{R}_{++}^n$  at  $x = \mu(S^0)^{-1}\omega$ . Since the objective function of the mathematical program in 2.7 differs from  $h$  by a constant, the mathematical program in 2.7 has a unique solution for all  $\mu > 0$ .

The unique solution to 2.7 has a characterization as part of the unique solution to a system of equations, whose only nonlinearities are the bilinear complementarity equations (equation 2.10 below). This system of equations is developed by noticing that the necessary and sufficient Lagrange optimality conditions imply that the gradient of the objective function is contained in  $\text{row}(A)$ . Hence, for all  $\mu > 0$ , there exists a  $y$  such that

$$yA = c - \mu\omega^T X^{-1}.$$

Allowing  $s = \mu\omega^T X^{-1}$ , the necessary and sufficient conditions for  $x$  may be written as: there exists  $y$  and  $s$  such that

$$Ax = b \tag{2.8}$$

$$yA + s = c \tag{2.9}$$

$$Sx = \mu\omega. \tag{2.10}$$

Notice that equations 2.8, 2.9, and 2.10 imply that  $x$  is the unique solution of the mathematical program in 2.7 if, and only if, there exists a dual solution that allows the component-wise “near” complementarity shown in 2.10. An immediate implication of equation 2.10 is that if the primal problem is bounded, then the dual problem is unbounded. This is because equation 2.10 has a solution for all  $\mu > 0$ . So, if  $\mathcal{P}_b$  is bounded, as  $\mu \rightarrow \infty$  the dual variable must become arbitrarily large. In fact, equation 2.10 implies that if the primal problem is bounded, each component of  $s$  may become arbitrarily large. Also, notice that the full row rank assumption implies that the dual elements are unique.

The goal is to use this system of equations and the full rank assumption of  $A$  to show that  $x$ ,  $y$ , and  $s$  are analytic functions of  $\mu$ ,  $b$ , and  $c$ . This result follows as a direct consequence of the implicit function theorem.

**Theorem 2.6 (Sonnevend [80])** Fix  $A \in \mathbb{R}^{m \times n}$ ,  $m \leq n$ , with  $\text{rank}(A) = m$ . Given  $r \in \mathcal{G}$ ,  $\omega \in \mathbb{R}_{++}^n$ , and  $\mu > 0$ , let  $x$ ,  $y$ , and  $s$  satisfy equations 2.8, 2.9, and 2.10. Then,  $x$ ,  $y$ , and  $s$  are analytic functions of  $(\mu, b, c)$ .

**Proof:** Let  $\bar{r} \in \mathcal{G}$ , and  $\bar{\mu} > 0$ . Furthermore, let  $\hat{x}$  and  $(\hat{y}, \hat{s})$  satisfy  $A\hat{x} = \bar{b}$ ,  $\hat{x} > 0$ ,  $\hat{y}A + \hat{s} = \bar{c}$ ,  $\hat{s} > 0$ , and  $\hat{S}\hat{x} = \bar{\mu}\omega$ . Define

$$\Psi : \mathbb{R}^{3n+2m+1} \rightarrow \mathbb{R}^{2n+m} : (x, y, s, \mu, b, c) \rightarrow \begin{pmatrix} Ax - b \\ yA + s - c \\ Xs - \mu\omega \end{pmatrix}.$$

Then,  $\Psi(\hat{x}, \hat{y}, \hat{s}, \bar{\mu}, \bar{b}, \bar{c}) = 0$ . The Jacobian of  $\Psi$  with respect to  $x$ ,  $y$ , and  $s$  is

$$\nabla_{x,y,s} \Psi(x, y, s, \mu, b, c) = \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix},$$

and the rank assumption implies that this Jacobian is non-singular.

Using the fact that each component function is analytic, the general form of the implicit function theorem [12] implies the result. ■

An immediate consequence of Theorem 2.6 is that if  $\{(\mu, b, c)^k \in \mathbb{R}_{++} \times \mathcal{G}\} \rightarrow (\bar{\mu}, \bar{b}, \bar{c}) \in \mathbb{R}_{++} \times \mathcal{G}$ ,

$$\lim_{k \rightarrow \infty} x((\mu, b, c)^k) = x(\bar{\mu}, \bar{b}, \bar{c}).$$

Theorem 2.6 also shows that the omega central path, defined below, is an infinitely smooth curve over  $\mathbb{R}_{++} \times \mathcal{G}$ .

**Definition 2.7** Let  $\omega > 0$  and  $r \in \mathcal{G}$ . Then

$$CP_r \equiv \{(x(\mu, b, c), y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}$$

is called the primal-dual omega central path. The primal omega central path is

$$PCP_r \equiv \{x(\mu, b, c) : \mu > 0\}$$

and the dual omega central path is

$$DCP_r \equiv \{(y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}.$$

The primal omega central path is simply referred to as the omega central path and the primal, dual distinctions are used only when clarity is needed. Also, when  $\omega = e$ , the primal omega central path is called the central path.

Notice that  $x(\mu, b, c)$  is the omega analytic center of  $\mathcal{P}_b \cap \{x : cx = cx(\mu, b, c)\}$ , which is bounded from theorem 2.4. Hence, the omega central path has the interpretation of being composed of the omega analytic centers of constant-cost-slices of  $\mathcal{P}_b$ . Renegar's "conceptual" algorithm [67] is based on this perspective.

The omega central path has many properties, and in the next section it is shown that the omega central path converges to an optimal omega analytic center and that the objective function is either strictly monotonic or constant along the central path. Also, the special case when the primal or dual central path contains a single element is characterized.

### 2.3 Basic Properties of the Omega Central Path

Much more than the analytic property shown in theorem 2.6 is known. Properties of the central path were investigated by Fiacco and McCormick from the perspective of nonlinear programming [13]. The important fact that  $cx(\mu, b, c)$  is a strictly decreasing function of  $\mu$  is proven here. Hence, the previously mentioned constant-cost-slices are strictly monotonic as they approach the optimal face. Furthermore, as is shown in theorem 2.8, for any fixed  $r \in \mathcal{G}$ , the limit of  $x(\mu, b, c)$ , as  $\mu$  approaches zero, is the omega analytic center of the optimal face. This result is strengthened in Chapter 4, where necessary and sufficient conditions are given for  $x(\mu, b, c)$  to converge, even when both  $b$  and  $c$  are changing. Others have investigated limiting behavior of the derivatives, but a discussion of this topic is postponed until chapter 3. Table 2.1 gives a short guide to these and other analytic properties of the omega central path.

For our purposes, the omega central path is shown to have a unique limit that is optimal, is strictly complementary, and is the omega analytic center of the optimal set. These results are well established, and the techniques of proof are originally in [53]. The proof uses the *support set* of a vector, which is defined as the index set,  $\sigma(x) = \{i : x_i > 0\}$ , for any  $x \in \mathbb{R}_+^n$ .

<b>Author</b>	<b>Type of Results</b>
Fiacco and McCormick [13]	barrier function concept and convergence proofs
McLinden [53]	analytic centers, complementarity problems, and properties of limit points
Bayer and Lagarias [4, 5]	many analytic properties about the central path and the Legendre transformation
Megiddo [55]	primal-dual properties are developed through the use of the logarithmic barrier function
Zhao and Zhu [100], and Zhao [101]	properties for the curvature properties for the curvature integral of the central path
Adler and Monteiro [1] and Witzgall, Boggs, and Domich [90]	convergence of the first order limiting derivatives
Güler [63]	convergence of higher order derivatives
Roos, Terlaky, and Vial [73]	text on interior point algorithms; convergence rate properties of the central path and first order derivative convergence
Wright [95]	text on interior point algorithms; convergence rate properties, and infeasible interior point methods
bibliographies [47, 94]	large interior point method bibliographies on the internet

**Table 2.1.** A synopsis of analytic properties of the omega central path

**Theorem 2.8 (McLinden [53])** Let  $\omega \in \mathbb{R}_{++}^n$  and  $r \in \mathcal{G}$ . Then both

$$\lim_{\mu \rightarrow 0^+} x(\mu, b, c) \text{ and } \lim_{\mu \rightarrow 0^+} s(\mu, b, c)$$

exist and form a strictly complementary optimal solution. Furthermore, the above limits are the omega analytic centers of the primal and dual optimal sets, respectively.

**Proof:** Since,

$$\begin{aligned} & \{(x(\mu, b, c), y(\mu, b, c), s(\mu, b, c)) : \mu > 0\} \\ &= \{(x(\frac{\mu n}{e^T \omega}, b, c), y(\frac{\mu n}{e^T \omega}, b, c), s(\frac{\mu n}{e^T \omega}, b, c)) : \mu > 0\}, \end{aligned}$$

we assume without loss of generality that  $e^T \omega = n$ . For any  $\mu > 0$ , the duality gap is

$$s(\mu, b, c)x(\mu, b, c) = e^T S(\mu, b, c)x(\mu, b, c) = \mu e^T \omega = n\mu.$$

Hence, Theorem 2.4 implies that  $\{(x(\mu, b, c), s(\mu, b, c)) : 0 \leq \mu \leq \bar{\mu}\}$  is bounded for any fixed  $\bar{\mu} > 0$ . This means that there exists a sequence, say  $\{\mu^k\}$ , such that  $\lim_{k \rightarrow \infty} x(\mu^k, b, c) = \hat{x}$  and  $\lim_{k \rightarrow \infty} s(\mu^k, b, c) = \hat{s}$ . Since,  $x(\mu^k, b, c) - \hat{x} \in \text{null}(A)$  and  $s(\mu^k, b, c) - \hat{s} \in \text{row}(A)$ ,

$$s(\mu^k, b, c)x(\mu^k, b, c) + \hat{s}\hat{x} = s(\mu^k, b, c)\hat{x} + \hat{s}x(\mu^k, b, c). \quad (2.11)$$

Using that  $\hat{x}$  and  $\hat{s}$  are optimal, this is equivalent to

$$\mu^k n = \sum_{i \in \sigma(\hat{x})} s_i(\mu^k, b, c) \hat{x}_i + \sum_{i \in \sigma(\hat{s})} \hat{s}_i x_i(\mu^k, b, c). \quad (2.12)$$

Equation 2.10 implies that  $s_i(\mu^k, b, c) x_i(\mu^k, b, c) = \mu^k \omega_i$ , which means that equation 2.12 is equivalent to,

$$n = \sum_{i \in \sigma(\hat{x})} \frac{\omega_i \hat{x}_i}{x_i(\mu^k, b, c)} + \sum_{i \in \sigma(\hat{s})} \frac{\omega_i \hat{s}_i}{s_i(\mu^k, b, c)}.$$

As  $\mu^k \rightarrow 0$ , the equality holds if, and only if,  $\sigma(\hat{x}) \cup \sigma(\hat{s}) = \{1, 2, 3, \dots, n\}$ . Hence  $\hat{x}$  and  $\hat{s}$  are strictly complementary, and  $\sigma(\hat{x}) = B$  and  $\sigma(\hat{s}) = N$ . So, any cluster point of the omega central path induces the optimal partition.

The fact that  $\hat{x}$  and  $\hat{s}$  are the omega analytic centers of  $\mathcal{P}_r^*$  and  $\mathcal{D}_r^*$ , respectively, is now shown. The existence of the limits follows by the uniqueness of the analytic center. Let  $(x^*, y^*, s^*) \in \mathcal{P}_r^* \times \mathcal{D}_r^*$ . Then, since  $x(\mu^k, b, c) - x^* \in \text{null}(A)$  and  $s(\mu^k, b, c) - s^* \in \text{row}(A)$ , an analogous argument to that above shows

$$n = \sum_{i \in \sigma(x^*)} \frac{\omega_i x_i^*}{\hat{x}_i} + \sum_{i \in \sigma(s^*)} \frac{\omega_i s_i^*}{\hat{s}_i},$$

where the subset relations,  $\sigma(x^*) \subseteq \sigma(\hat{x})$  and  $\sigma(s^*) \subseteq \sigma(\hat{s})$ , are used. After dividing both sides by  $n$ , the arithmetic - geometric mean inequality implies

$$\prod_{i \in \sigma(x^*)} \left( \frac{x_i^*}{\hat{x}_i} \right)^{\omega_i} \prod_{i \in \sigma(s^*)} \left( \frac{s_i^*}{\hat{s}_i} \right)^{\omega_i} \leq \frac{1}{n} \left( \sum_{i \in \sigma(x^*)} \frac{x_i^*}{\hat{x}_i} + \sum_{i \in \sigma(s^*)} \frac{s_i^*}{\hat{s}_i} \right)^n \leq 1.$$

Allowing  $s^* = \hat{s}$  yields

$$\prod_{i \in \sigma(x^*)} (x_i^*)^{\omega_i} \leq \prod_{i \in \sigma(x^*)} (\bar{x}_i^*)^{\omega_i},$$

and  $\hat{x}_B$  solves  $\max\{\prod_{i \in B} x_i^{\omega_i} : x \in \mathcal{P}_r^*\}$ . Since this means  $\hat{x}_B$  also solves

$$\max\{\sum_{i \in B} \omega_i \ln(x_i) : x \in \mathcal{P}_r^*\},$$

$\hat{x}$  is the analytic center of  $\mathcal{P}_r^*$ . Upon replacing  $x^*$  with  $\hat{x}$ , a similar argument completes the dual statement. ■

The proof technique used in theorem 2.8 is now used to show that if either the primal or the dual polyhedron is bounded, then as  $\mu$  increases to infinity the omega central path terminates at the omega analytic center of the polytope.

**Theorem 2.9 (McLinden [53])** Let  $\omega \in \mathbb{R}_{++}$  and  $r \in \mathcal{G}$ . If  $\mathcal{P}_b$  is bounded and  $\bar{x}$  is the omega analytic center of  $\mathcal{P}_b$ , then

$$\lim_{\mu \rightarrow \infty} x(\mu, b, c) = \bar{x}.$$

If  $\mathcal{D}_c$  is bounded and  $(\bar{y}, \bar{s})$  is the analytic center of  $\mathcal{D}_c$ , then

$$\lim_{\mu \rightarrow \infty} (y(\mu, b, c), s(\mu, b, c)) = (\bar{y}, \bar{s}).$$

**Proof:** Recall that equation 2.10 implies that at most one of  $\mathcal{P}_b$  and  $\mathcal{D}_c$  is bounded. Assume that  $\mathcal{P}_b$  is bounded and that  $\mu^k$  is a sequence increasing to

infinity such that

$$\lim_{k \rightarrow \infty} x(\mu^k, b, c) = \tilde{x}.$$

Then, for any  $(\hat{x}, (\hat{y}, \hat{s})) \in \mathcal{P}_b \times \mathcal{D}_c$ , the orthogonal argument used in the proof of theorem 2.8 implies

$$\frac{\hat{s}\hat{x}}{\mu^k} + n = \sum_{i=1}^n \frac{\omega_i \hat{x}_i}{x_i(\mu^k, b, c)} + \sum_{i=1}^n \frac{\omega_i \hat{s}_i}{s_i(\mu^k, b, c)},$$

where it is assumed that  $e^T \omega = n$ . Allowing  $k \rightarrow \infty$ ,

$$n = \sum_{i=1}^n \frac{\omega_i \hat{x}_i}{\tilde{x}},$$

and the arithmetic-geometric mean inequality implies

$$\prod_{i=1}^n (\hat{x}_i)^{\omega_i} \leq \prod_{i=1}^n (\tilde{x}_i)^{\omega_i}.$$

Hence,  $\tilde{x} = \bar{x}$ . A similar argument follows for the dual statement. ■

Theorem 2.9 shows that for any  $c \in \mathcal{G}_c$ ,  $\lim_{\mu \rightarrow \infty} x(\mu, b, c)$  solves

$$\max \left\{ \sum_{i=1}^n \ln(x_i) : x \in \mathcal{P}_b \right\},$$

so long as  $\mathcal{P}_b$  is bounded and  $b \in \mathcal{G}_b$ . Hence, this limit does not depend on  $c$ , but only on  $b$ . Similarly, when  $\lim_{\mu \rightarrow \infty} (y(\mu, b, c), s(\mu, b, c))$  exists, this limit does not depend on  $b$ , but only on  $c$ . The “bar” notation is used to denote these limits:

$$\begin{aligned} \bar{x}(b) &\equiv \lim_{\mu \rightarrow \infty} x(\mu, b, c) \quad \text{and} \\ (\bar{y}(c), \bar{s}(c)) &\equiv \lim_{\mu \rightarrow \infty} (y(\mu, b, c), s(\mu, b, c)), \end{aligned}$$

provided the limits exist.

The next result shows the relationship between a constant primal objective function and the omega central path. In particular, this result demonstrates that the primal objective function is constant if, and only if, the primal omega central path contains a single element. The equivalence of the fourth statement, which shows how the cost vector indicates a single element omega central path, forces the inclusion of a special case in many of the results in Chapter 4. This is because a cost vector may indicate a single element omega central path, while a perturbed cost vector may not have this quality.

**Theorem 2.10 (Roos, Terlaky, and Vial [73])** The following are equivalent for the primal problem:

- (1)  $cx$  is constant on  $\mathcal{P}_b$
- (2)  $x(\mu^1, b, c) = x(\mu^2, b, c)$  for all  $0 < \mu^1 < \mu^2$
- (3)  $x(\mu^1, b, c) = x(\mu^2, b, c)$  for some  $0 < \mu^1 < \mu^2$
- (4)  $c \in \text{row}(A)$
- (5)  $s(\mu, b, c) = \mu s(1, b, c)$  for all  $0 < \mu$ .

Similarly, the following are equivalent for the dual problem:

- (1)  $yb$  is constant on  $\mathcal{D}_c$
- (2)  $(y(\mu^1, b, c), s(\mu^1, b, c)) = (y(\mu^2, b, c), s(\mu^2, b, c))$  for all  $0 < \mu^1 < \mu^2$
- (3)  $(y(\mu^1, b, c), s(\mu^1, b, c)) = (y(\mu^2, b, c), s(\mu^2, b, c))$  for some  $0 < \mu^1 < \mu^2$

(4)  $b = 0$ .

**Proof:** Consider the primal statements. We begin by showing  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ . If 1 holds, then the objective function in 2.7 is independent of both  $cx$  and  $\mu$ , and 1 implies 2. The result that 2 implies 3 is obvious. Assume that 3 is true. Then, because

$$s(\mu^1, b, c) = \mu^1 \omega^T X^{-1}(\mu^1, b, c)$$

and

$$s(\mu^2, b, c) = \mu^2 \omega^T X^{-1}(\mu^2, b, c),$$

the dual conditions are

$$y(\mu^1, b, c)A + \mu^1 \omega^T X^{-1}(\mu^1, b, c) = c \text{ and}$$

$$y(\mu^2, b, c)A + \mu^2 \omega^T X^{-1}(\mu^2, b, c) = c.$$

Since  $\omega^T X^{-1}(\mu^1, b, c) = \omega^T X^{-1}(\mu^2, b, c)$ , it follows that

$$\left( \frac{1}{\mu^1} y(\mu^1, b, c) - \frac{1}{\mu^2} y(\mu^2, b, c) \right) A = \left( \frac{1}{\mu^1} - \frac{1}{\mu^2} \right) c.$$

So,  $c \in \text{row}(A)$ . Assume 4 is true. Fixing  $\hat{x} \in \mathcal{P}_b$ , we have  $\mathcal{P}_b - \{\hat{x}\} \subseteq \text{null}(A)$ .

So, for all  $x \in \mathcal{P}_b$ ,  $c(x - \hat{x}) = 0$  and 1 follows.

The equivalence of the first four statements is now established. Note that  $5 \Rightarrow 2$  follows immediately from the equality,  $X(\mu, b, c)s(\mu, b, c) = \mu\omega$ . To

complete the result,  $4 \Rightarrow 5$  is shown. Assume that 4 is true and let  $\hat{y}A = c$ .

Then,  $s(1, b, c) \in \text{row}(A)$  because

$$s(1, b, c) = c - y(1, b, c)A = (\hat{y} - y(1, b, c))A.$$

Therefore,  $c - \mu s(1, b, c) \in \text{row}(A)$  for all  $\mu \in \mathbb{R}_{++}$  and there exists  $v_\mu$  such that

$$v_\mu A + \mu s(1, b, c) = c.$$

Recognizing that  $(x(1, b, c), (v_\mu, \mu s(1, b, c)))$  satisfy equations 2.8, 2.9, and 2.10, for all  $\mu \in \mathbb{R}_{++}$ , it follows that

$$(x(1, b, c), (v_\mu, \mu s(1, b, c))) = (x(\mu, b, c), (y(\mu, b, c), s(\mu, b, c))).$$

Hence, 5 is true.

The proof of the equivalence of the dual statements is similar in nature, and we show  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ . Since  $(y(\mu, b, c), s(\mu, b, c))$  is the unique solution to

$$\max\{yb + \mu \sum_{i=1}^n \omega_i \ln(s_i) : yA + s = c, s > 0\},$$

if  $yb$  is constant,  $(y(\mu, b, c), s(\mu, b, c))$  is independent of  $yb$  and  $\mu$ . Hence,  $1 \Rightarrow 2$ .

The fact that  $2 \Rightarrow 3$  is obvious. Assume that 3 is true. Then, there exists different  $\mu^1$  and  $\mu^2$  in  $\mathbb{R}_{++}$ , such that

$$(y(\mu^1, b, c), s(\mu^1, b, c)) = (y(\mu^2, b, c), s(\mu^2, b, c)).$$

The necessary and sufficient Lagrange conditions are

$$\begin{bmatrix} b \\ \mu^1 S^{-1}(\mu^1, b, c)e \end{bmatrix} = \begin{bmatrix} A \\ I \end{bmatrix} x(\mu^1, b, c)$$

and

$$\begin{bmatrix} b \\ \mu^1 S^{-1}(\mu^2, b, c)e \end{bmatrix} = \begin{bmatrix} A \\ I \end{bmatrix} x(\mu^2, b, c).$$

Since  $S^{-1}(\mu^1, b, c)e = S^{-1}(\mu^2, b, c)e$ ,

$$\frac{x(\mu^1, b, c)}{\mu^1} = \frac{x(\mu^2, b, c)}{\mu^2}.$$

So,

$$\frac{b}{\mu^1} = A \left( \frac{x(\mu^1, b, c)}{\mu^1} \right) = A \left( \frac{x(\mu^2, b, c)}{\mu^2} \right) = \frac{b}{\mu^2},$$

from which it follows that  $b = 0$ . The fact that 4 implies 1 is obvious, and the proof is complete. ■

The last result of this section establishes that the primal objective function is monotonically decreasing along the omega central path. This result is originally found in [13].

**Theorem 2.11 (Fiacco and McCormick [13])**  $cx(\mu^1, b, c) < cx(\mu^2, b, c)$ , for all  $0 \leq \mu^1 < \mu^2$ , if, and only if,  $c \notin \text{row}(A)$ . Also,  $y(\mu^1, b, c)b > y(\mu^2, b, c)b$ , for all  $0 \leq \mu^1 < \mu^2$ , if, and only if,  $b \neq 0$ .

**Proof:** If  $cx(\mu^1, b, c) < cx(\mu^2, b, c)$ , theorem 2.10 implies that  $c \notin \text{row}(A)$ .

Assume that  $c \notin \text{row}(A)$  and let  $0 \leq \mu^1 < \mu^2$ . Then the strict concavity of the objective function in 2.7 implies

$$\begin{aligned} & cx(\mu^1, b, c) - \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^1, b, c)) \\ & < cx(\mu^2, b, c) - \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^2, b, c)) \end{aligned}$$

and

$$\begin{aligned} & cx(\mu^2, b, c) - \mu^2 \sum_{i=1}^n \omega_i \ln(x_i(\mu^2, b, c)) \\ & < cx(\mu^1, b, c) - \mu^2 \sum_{i=1}^n \omega_i \ln(x_i(\mu^1, b, c)). \end{aligned}$$

Multiplying the first inequality by  $\mu^2$ , the second inequality by  $\mu^1$ , and adding the two inequalities produces

$$\begin{aligned} & \mu^2 cx(\mu^1, b, c) + \mu^1 cx(\mu^2, b, c) - \\ & \quad \left( \mu^2 \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^1, b, c)) + \mu^2 \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^2, b, c)) \right) \\ & < \mu^2 cx(\mu^2, b, c) + \mu^1 cx(\mu^1, b, c) - \\ & \quad \left( \mu^2 \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^1, b, c)) + \mu^2 \mu^1 \sum_{i=1}^n \omega_i \ln(x_i(\mu^2, b, c)) \right). \end{aligned}$$

So,

$$\mu^2 cx(\mu^1, b, c) + \mu^1 cx(\mu^2, b, c) < \mu^2 cx(\mu^2, b, c) + \mu^1 cx(\mu^1, b, c),$$

which implies that

$$(\mu^2 - \mu^1)cx(\mu^1, b, c) < (\mu^2 - \mu^1)cx(\mu^2, b, c).$$

The result now follows by the assumption that  $\mu^2 > \mu^1$ .

The dual result follows from an analogous argument. ■

This section ends with the development of a simple example that is used several times. The simplicity of this example allows the central path to be written in closed form, which is valuable when attempting to motivate results.

**Example 2.12** Consider the linear program  $\min\{cx : 0 \leq x \leq b\}$ , where we assume that the components of  $c$  are not all the same. Incorporating the slack vector,  $b - x$ , the penalized objective function, with  $\omega = e$ , is

$$cx - \mu \sum_{i=1}^n \ln(x_i) - \mu \sum_{i=1}^n \ln(b_i - x_i). \quad (2.13)$$

The  $i$ th component of the gradient is

$$c_i - \frac{\mu}{x_i} + \frac{\mu}{b_i - x_i}.$$

After finding the roots of this gradient, we see that the central path is given by

$$x_i(\mu, b, c) = \begin{cases} \frac{c_i b_i + 2\mu - \sqrt{b_i^2 c_i^2 + 4\mu^2}}{2c_i} & \text{if } c_i \neq 0 \\ \frac{b_i}{2} & \text{if } c_i = 0. \end{cases}$$

This means that

$$\begin{aligned}
 x_i^*(b, c) = \lim_{\mu \rightarrow 0^+} x_i(\mu, b, c) &= \begin{cases} \frac{b_i(c_i - |c_i|)}{2c_i} & \text{if } c_i \neq 0 \\ \frac{b_i}{2} & \text{if } c_i = 0 \end{cases} \\
 &= \begin{cases} 0 & \text{if } c_i > 0 \\ b_i & \text{if } c_i < 0 \\ \frac{b_i}{2} & \text{if } c_i = 0 \end{cases}
 \end{aligned}$$

This example shows that  $x^*(b, c)$  is *not* a continuous function of  $c$ . For example, fixing  $b = e$ ,

$$x_i^*(e, c) = \begin{cases} 0 & \text{if } c_i > 0 \\ 1 & \text{if } c_i < 0 \\ \frac{1}{2} & \text{if } c_i = 0 \end{cases}$$

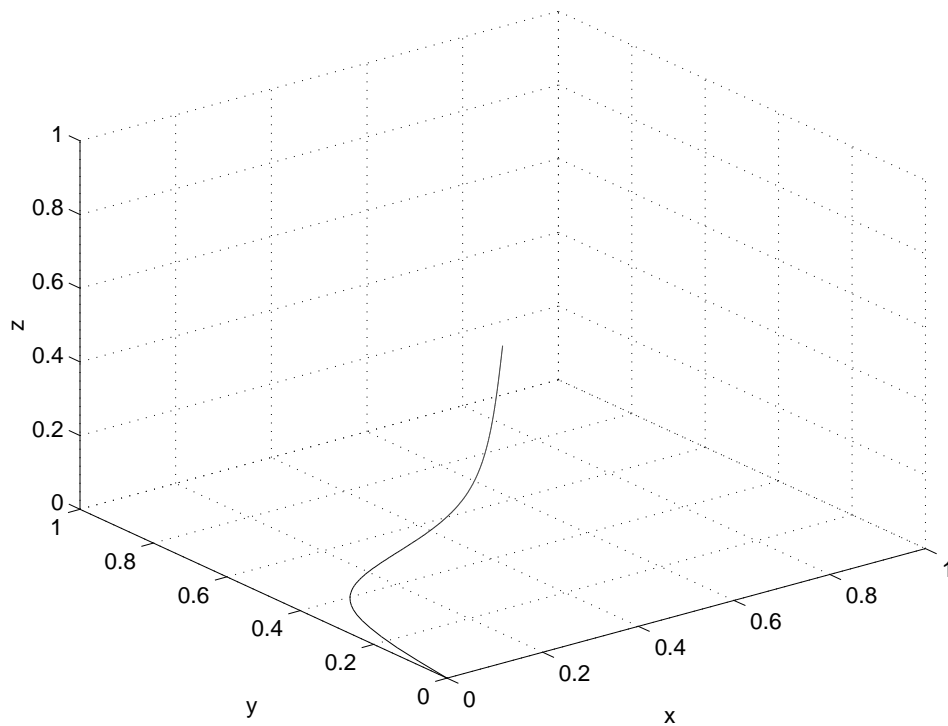
However,  $x^*(b, c)$  is a continuous function of  $b$  for this example. This property is in general true as shown in chapter 3.

The central path for

$$\min\{10x + y + 100z : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}$$

is shown in Figure 2.1. Notice that the central path terminates at the unique optimal solution. Furthermore, observe that the magnitudes of the components of  $c$  affect the path. Since  $c_3$  is the largest component, the path first attempts to drive  $x_3$  to zero. Once  $x_3$  becomes sufficiently small, it turns to decrease

$x_1$ . Finally, it turns again to move  $x_2$  towards zero. This idea that the central path makes somewhat sudden turns, and that between these sudden turns it is almost linear, was recently proven by Vavasis and Ye [89]. This geometric characterization leads to an algorithm for which the complexity analysis is not dependent upon  $b$  or  $c$ .



**Figure 2.1.** A central path.

## 2.4 Analytic Centers and Data Perturbations

How the central path and the analytic center solution react to changing data is the topic of interest for the remainder of this thesis. For any  $r \in \mathcal{G}$ ,

define  $\delta r = (\delta b, \delta c)$  to be a *direction of change*. Notice that theorem 2.3 guarantees that for any  $r \in \mathcal{G}$  and any  $\delta r$ , there exists  $\theta^* > 0$  such that for all  $\theta \in [0, \theta^*)$ ,  $r + \theta \delta r \in \mathcal{G}$ . Hence, any direction is admissible.

To facilitate linear changes in data, define for any  $r \in \mathcal{G}$  and any  $\delta r$ ,

$$b_\rho \equiv b + \rho \delta b \text{ and}$$

$$c_\tau \equiv c + \tau \delta c,$$

where  $\rho$  and  $\tau$  are always assumed non-negative. This notation is used only when a direction of change is understood. For general, nonlinear changes in rim data, sequences are used. For example, if a general change in  $b$  is desired, then a sequence,  $\{b^k\}$ , of admissible right-hand sides is used.

Directions of change for which the optimal partition does not change on the interval  $[r, r + \theta^* \delta r]$ , for some  $\theta^* > 0$ , are of particular interest. Define

$$\begin{aligned} \mathcal{H}(r) &= \{ \delta r : \text{there exists } \theta^* > 0 \text{ for which} \\ &\quad (B(r + \theta \delta r) | N(r + \theta \delta r)) = (B(r) | N(r)) \text{ for all } \theta \in [0, \theta^*) \}, \\ \mathcal{H}_b(r) &= \{ \delta b : (\delta b, 0) \in \mathcal{H}(r) \}, \text{ and} \\ \mathcal{H}_c(r) &= \{ \delta c : (0, \delta c) \in \mathcal{H}(r) \}. \end{aligned}$$

Recent investigations into the properties of these sets are found in [26], [28], and [31]. The next lemma shows that the optimal partition characterizes  $\mathcal{H}(r)$ ,  $\mathcal{H}_b(r)$ , and  $\mathcal{H}_c(r)$ .

**Lemma 2.13** Let  $r$  be admissible. Then,

(1)  $\mathcal{H}_b(r) = \text{col}(A_B)$ , and

(2)  $\mathcal{H}_c(r) = \{\delta c \in \mathcal{H}_c : \delta c_B \in \text{row}(A_B)\}$ .

**Proof:** For  $r \in \mathcal{G}$ , let  $x^*$  be an associated strictly complementary optimal solution, and  $\delta b \in \mathbb{R}^m$ . The optimality conditions defining the optimal partition for changes in the right-hand side are,

$$A_B x_B = b_\rho$$

$$y A_B = c_B$$

$$y A_N < c_N$$

$$x_B > 0,$$

where  $\rho$  is sufficiently small. If  $\delta b \in \text{col}(A_B)$ , then there exists  $x'$  such that  $A_B(\rho x') = \rho \delta b$ . Since  $x_B^* - \rho x' > 0$  for sufficiently small  $\rho$ , the above conditions hold, and  $\text{col}(A_B) \subseteq \mathcal{H}_b(r)$ .

If the optimal partition is invariant for sufficiently small  $\rho$ , then there exists  $x_B(\rho)$  such that  $A_B x_B(\rho) = b_\rho$ . Hence,

$$A_B(x_B(\rho) - x_B^*) = \rho \delta b.$$

and  $\delta b \in \text{col}(A_B)$ .

The argument for the second statement is similar, except the optimality conditions defining the optimal partition are

$$\begin{aligned} A_B x_B &= b \\ y A_B &= c_B + \tau \delta c_B \\ y A_N &< c_N + \tau \delta c_N \\ x_B &> 0, \end{aligned}$$

where  $\tau$  is sufficiently small. ■

In addition to the characterizations in lemma 2.13 for  $\mathcal{H}_b(r)$  and  $\mathcal{H}_c(r)$ , the following decoupling principle is shown in [26],

$$\mathcal{H}(r) = \mathcal{H}_b(r) \times \mathcal{H}_c(r). \tag{2.14}$$

The next theorem is an extension of theorem 2.4 to the cases when rim data elements are allowed to change. It establishes the boundedness of the union over a converging sequence of rim data of the primal and dual elements having a bounded duality gap.

**Theorem 2.14** Let  $\{r^k \in \mathcal{G}\} \rightarrow r \in \mathcal{G}$ . Then, for all  $M \in \mathbb{R}_+$ ,

$$\bigcup_k \mathcal{L}(r^k, M)$$

is bounded.

**Proof:** For any  $(x^k, s^k) \in \mathcal{P}_{b^k}^o \times \mathcal{D}_{c^k}^o$ , the proof of theorem 2.4 demonstrates that when  $(x, s) \in \mathcal{L}(r^k, M)$ ,

$$x_i \leq \frac{M + s^k x^k}{s_i^k}$$

and

$$s_i \leq \frac{M + s^k x^k}{x_i^k}.$$

Let  $\bar{\mu} > 0$ . Then, setting  $s^k = s(\bar{\mu}, b^k, c^k)$  and  $x^k = (\bar{\mu}, b^k, c^k)$ , we have  $s^k x^k = n\bar{\mu}$ . Since  $\{r^k \in \mathcal{G}\} \rightarrow r \in \mathcal{G}$ , the analytic properties of  $x(\mu, b, c)$  and  $s(\mu, b, c)$  shown in theorem 2.6 imply that  $s^k \rightarrow s(\bar{\mu}, \bar{b}, \bar{c}) > 0$  and  $x^k \rightarrow x(\bar{\mu}, \bar{b}, \bar{c}) > 0$ . This implies that there exists a natural number  $K$  such that  $s_i^k > \frac{1}{2}s_i(\bar{\mu}, \bar{b}, \bar{c})$  and  $x_i^k > \frac{1}{2}x_i(\bar{\mu}, \bar{b}, \bar{c})$ , for all  $k \geq K$  and  $i = 1, 2, \dots, n$ . Hence, for all  $(x, y, s) \in \mathcal{L}(r^k, M)$

$$x_i \leq \frac{M + s^k x^k}{s_i^k} < \frac{2(M + n\bar{\mu})}{s_i(\bar{\mu}, \bar{b}, \bar{c})}$$

and

$$s_i \leq \frac{M + s^k x^k}{x_i^k} < \frac{2(M + n\bar{\mu})}{x_i(\bar{\mu}, \bar{b}, \bar{c})},$$

when  $k \geq K$ . The fact that  $y$  is bounded follows from the one-to-one linear relationship between  $y$  and  $s$ . So,

$$\bigcup_{k \geq K} \mathcal{L}(r^k, M)$$

is bounded. Since

$$\bigcup_{k < K} \mathcal{L}(r^k, M)$$

is a finite union of bounded sets, the result follows.  $\blacksquare$

Theorem 2.14 has the favorable consequence that for any sequence,  $\{(\mu^k, r^k) \in \mathbb{R}_{++} \times \mathcal{G}\} \rightarrow (0, r) \in \{0\} \times \mathcal{G}$ , the sequences  $\{x(\mu^k, b^k, c^k)\}$  and  $\{s(\mu^k, b^k, c^k)\}$  both have cluster points. However, as is seen in chapter 4, the sequences  $\{x(\mu^k, b^k, c^k)\}$  and  $\{s(\mu^k, b^k, c^k)\}$  do not necessarily converge. This means a straight forward extension of theorem 2.8 is not available. Although the sequences  $\{x(\mu^k, b^k, c^k)\}$  and  $\{s(\mu^k, b^k, c^k)\}$  are not guaranteed to converge, the next lemma provides sufficient conditions for sub-vectors of these sequences to converge to zero.

**Lemma 2.15** Let  $\{r^k \in \mathcal{G}\} \rightarrow \bar{r} \in \mathcal{G}$ . If  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$ ,

$$\lim_{k \rightarrow \infty} x_{N(\bar{r})}(\mu^k, b^k, c^k) = 0 \text{ and}$$

$$\lim_{k \rightarrow \infty} s_{B(\bar{r})}(\mu^k, b^k, c^k) = 0.$$

**Proof:** Let  $\{\mu^k\}$  and  $\{r^k\}$  be as above. Theorem 2.14 implies that

$$\left\{ \left( x(\mu^k, b^k, c^k), \left( y(\mu^k, b^k, c^k), s(\mu^k, b^k, c^k) \right) \right) \right\}$$

is bounded, and hence there exists a convergent subsequence. Let

$$\lim_{i \rightarrow \infty} \left( x(\mu^{k_i}, b^{k_i}, c^{k_i}), \left( y(\mu^{k_i}, b^{k_i}, c^{k_i}), s(\mu^{k_i}, b^{k_i}, c^{k_i}) \right) \right) = (\hat{x}, (\hat{y}, \hat{s})).$$

Since,

$$\begin{aligned}
Ax(\mu^{k_i}, b^{k_i}, c^{k_i}) &= b^{k_i} \quad , \quad x(\mu^{k_i}, b^{k_i}, c^{k_i}) \geq 0, \\
y(\mu^{k_i}, b^{k_i}, c^{k_i})A + s(\mu^{k_i}, b^{k_i}, c^{k_i}) &= c^{k_i} \quad , \quad s(\mu^{k_i}, b^{k_i}, c^{k_i}) \geq 0, \text{ and} \\
s(\mu^{k_i}, b^{k_i}, c^{k_i})x(\mu^{k_i}, b^{k_i}, c^{k_i}) &= n\mu^{k_i},
\end{aligned}$$

allowing  $i$  to go to infinity implies that  $\hat{x} \in \mathcal{P}_{\bar{r}}^*$  and  $(\hat{y}, \hat{s}) \in \mathcal{D}_{\bar{r}}^*$ . The result follows because  $\mathcal{P}_{\bar{b}}^* = \{x \in \mathcal{P}_{\bar{b}} : x_N = 0\}$  and  $\mathcal{D}_{\bar{c}}^* = \{(y, s) \in \mathcal{D}_{\bar{c}} : s_B = 0\}$ . ■

Lemma 2.15 leaves open the question of what happens to  $x_{B(r)}(\mu^k, b^k, c^k)$  and  $s_{N(r)}(\mu^k, b^k, c^k)$ . The results in chapter 4 characterize the conditions needed to guarantee the convergence of these components, provided the cost vector perturbation is linear.

The last theorem of this chapter supplies sufficient conditions for  $x(\mu, b, c)$  to converge to the analytic center of a polytope. Before proving this result, we present two supporting lemmas.

**Lemma 2.16** For any matrix,  $Q \in \mathbb{R}^{p \times q}$ , and  $b \in \mathbb{R}^q$ , let

$$\mathcal{Q}_b = \{u : Qu = b, u \geq 0\}.$$

If  $\{b^k\} \rightarrow b$  and  $\mathcal{Q}_b$  is non-empty and bounded, then

$$\bigcup_k \mathcal{Q}_{b^k}$$

is bounded.

**Proof:** Suppose, for the sake of attaining a contradiction, that

$$\bigcup_k \mathcal{Q}_{b^k}$$

is unbounded. Then, there exists  $\{u^k\}$  such that

$$Q \frac{u^k}{\|u^k\|} = \frac{b^k}{\|u^k\|}$$

and  $\lim_{k \rightarrow \infty} \|u^k\| = \infty$ . Since  $\{\frac{u^k}{\|u^k\|}\}$  is bounded, there exists a convergent subsequence, say  $\{\frac{u^{k_i}}{\|u^{k_i}\|}\} \rightarrow \hat{u}$ . Since  $\{\frac{b^k}{\|u^{k_i}\|}\} \rightarrow 0$ ,  $Q\hat{u} = 0$ . However, since  $\hat{\mu} > 0$ , this leads to the contradiction that for any  $u \in \mathcal{Q}_b$  and  $\beta \in \mathbb{R}_{++}$ ,  $u + \beta\hat{u} \in \mathcal{Q}_b$ . ■

**Lemma 2.17** Define  $\mathcal{Q}_b$  as in lemma 2.16. Let  $\{b^k\} \rightarrow b$ , be such that  $\mathcal{Q}_{b^k} \neq \emptyset$ , for all  $k$ , and  $\mathcal{Q}_b \neq \emptyset$ . Then,

- (1) for any  $u \in \mathcal{Q}_b$ , there exists  $\{u^k \in \mathcal{Q}_{b^k}\} \rightarrow u$ , and
- (2) if  $\mathcal{Q}_b^o = \{u : Qu = b, u > 0\} \neq \emptyset$ , for any  $u \in \mathcal{Q}_b^o$ , there exists  $\{u^k \in \mathcal{Q}_{b^k}^o : k \geq K\} \rightarrow u$ , for some sufficiently large  $K$ .

**Proof:** Let  $\bar{u} \in \mathcal{Q}_b$ . Because the following system is consistent,

$$Q(\bar{u} + \delta u) = b + (b^k - b)$$

$$\bar{u} + \delta u \geq 0,$$

there exists  $\delta u$  such that

$$Q\delta u = b^k - b$$

$$\bar{u} + \delta u \geq 0.$$

All solutions of the equation  $Q\delta u = b^k - b$  have the form

$$\delta u = Q^+(b^k - b) + q,$$

where  $q \in \text{null}(Q)$ . Hence, there exists  $q \in \text{null}(Q)$  such that

$$\bar{u} + \delta u = \bar{u} + Q^+(b^k - b) + q \geq 0.$$

Define  $q^k = \min\{\|q\| : q \in \text{null}(Q), \bar{u} + \delta u \geq 0\}$ , and set

$$\delta u^k = Q^+(b^k - b) + q^k \text{ and}$$

$$u^k = \bar{u} + \delta u^k.$$

Then,  $u^k \geq 0$ , and

$$\begin{aligned} Qu^k &= Q(\bar{u} + \delta u^k) \\ &= Q\bar{u} + QQ^+(b^k - b) + Qq^k \\ &= b + (b^k - b) \\ &= b^k. \end{aligned}$$

So,  $u^k \in \mathcal{Q}_{b^k}$ . Since  $\{\bar{u} + QQ^+(b^k - b)\} \rightarrow \bar{u} \geq 0$ ,  $\{q^k\} \rightarrow 0$ , and

$$\begin{aligned} \lim_{k \rightarrow \infty} u^k &= \lim_{k \rightarrow \infty} \bar{u} + \delta u^k \\ &= \lim_{k \rightarrow \infty} \bar{u} + QQ^+(b^k - b) + q^k \\ &= \bar{u}. \end{aligned}$$

This proves the first statement of the lemma.

The second statement follows because if  $\bar{u} \in \mathcal{Q}_b^o$ , there exists  $K$  such that  $k \geq K$  implies  $\bar{u} + Q^+(b^k - b) > 0$ . Hence, for  $k \geq K$ ,  $q^k = 0$  and  $u^k \in \mathcal{Q}_{b^k}^o$ .

■

The next theorem provides sufficient conditions for  $\{x(\mu^k, b^k, c^k)\}$  to converge to the omega analytic center of a polytope.

**Theorem 2.18** Let  $\omega \in \mathbb{R}^n$ . Also let  $\{\mu^k \in \mathbb{R}_{++}\}$  and  $\{r^k \in \mathcal{G}\}$  be such that:

- (1)  $\{b^k\} \rightarrow \bar{b} \in \mathcal{G}_b$ ,
- (2)  $\mathcal{P}_{\bar{b}}$  is bounded,
- (3)  $\left\{\frac{c^k}{\mu^k}\right\}$  is bounded, and
- (4) every cluster point of  $\left\{\frac{c^k}{\mu^k}\right\}$  is contained in  $\text{row}(A)$ .

Then,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, c^k) = \bar{x}(\bar{b}).$$

Similarly, if

- (1)  $\{c^k\} \rightarrow \bar{c} \in \mathcal{G}_c$ ,
- (2)  $\mathcal{D}_{\bar{c}}$  is bounded, and
- (3)  $\left\{\frac{b^k}{\mu^k}\right\} \rightarrow 0$ ,

then,

$$\lim_{k \rightarrow \infty} \left( y(\mu^k, b^k, c^k), s(\mu^k, b^k, c^k) \right) = (\bar{y}(\bar{c}), \bar{s}(\bar{c}))$$

**Proof:** Let  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ ,  $\mathcal{P}_{\bar{b}}$  be bounded, and  $\{\frac{c^k}{\mu^k}\}$  be a bounded sequence such that every cluster point is contained in  $\text{row}(A)$ . Lemma 2.16 implies that  $\{x(\mu^k, b^k, c^k)\}$  is bounded, so there exists a subsequence such that

$$\begin{aligned} \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c^{k_i}) &= \hat{x} \text{ and} \\ \lim_{i \rightarrow \infty} \frac{c^{k_i}}{\mu^{k_i}} &= \hat{c}. \end{aligned}$$

Since,  $Ax(\mu^{k_i}, b^{k_i}, c^{k_i}) = b^{k_i}$  and  $x(\mu^{k_i}, b^{k_i}, c^{k_i}) \geq 0$ , allowing  $i \rightarrow \infty$  implies  $\hat{x} \in \mathcal{P}_{\bar{b}}$ . For any  $i$ , the necessary and sufficient conditions describing  $x(\mu^{k_i}, b^{k_i}, c^{k_i})$  are

$$\begin{aligned} Ax(\mu^{k_i}, b^{k_i}, c^{k_i}) &= b \\ y(\mu^{k_i}, b^{k_i}, c^{k_i})A + s(\mu^{k_i}, b^{k_i}, c^{k_i}) &= c^{k_i} \\ S(\mu^{k_i}, b^{k_i}, c^{k_i})x(\mu^{k_i}, b^{k_i}, c^{k_i}) &= \mu^{k_i}\omega \\ x(\mu^{k_i}, b^{k_i}, c^{k_i}) &> 0 \end{aligned}$$

which are equivalent to

$$\begin{aligned} Ax(\mu^{k_i}, b^{k_i}, c^{k_i}) &= b \\ -\frac{y(\mu^{k_i}, b^{k_i}, c^{k_i})}{\mu^{k_i}}A &= \omega^T X^{-1}(\mu^{k_i}, b^{k_i}, c^{k_i}) - \frac{c^{k_i}}{\mu^{k_i}} \\ x(\mu^{k_i}, b^{k_i}, c^{k_i}) &> 0. \end{aligned}$$

The full rank assumption implies

$$-\frac{y(\mu^{k_i}, b^{k_i}, c^{k_i})}{\mu^{k_i}} = \left( \omega^T X^{-1}(\mu^{k_i}, b^{k_i}, c^{k_i}) - \frac{c^{k_i}}{\mu^{k_i}} \right) A^T (AA^T)^{-1}.$$

If we prove  $\hat{x} > 0$ , this last equality implies that the sequence  $\left\{ \frac{y(\mu^{k_i}, b^{k_i}, c^{k_i})}{\mu^{k_i}} \right\}$  has a limit, say  $\hat{y}$ . Since  $\hat{c}$  is in  $\text{row}(A)$ , this implies the existence of a  $\tilde{y}$  such that

$$A\hat{x} = b$$

$$\tilde{y}A = \omega^T \hat{X}^{-1}$$

$$\hat{x} > 0.$$

Because these are the necessary and sufficient conditions describing  $\bar{x}(\bar{b})$ , the result would be established. We now show that  $\hat{x} > 0$ .

Lemma 2.17 implies the existence of a sequence,  $\{\tilde{x}^i \in \mathcal{P}_{b^{k_i}}^o\} \rightarrow \tilde{x} \in \mathcal{P}_{\bar{b}}^o$ . The optimality of  $x(\mu^{k_i}, b^{k_i}, c^{k_i})$  implies

$$\frac{c^{k_i}}{\mu^{k_i}} x(\mu^{k_i}, b^{k_i}, c^{k_i}) - \sum_{j=1}^n \omega_j \ln(x_j(\mu^{k_i}, b^{k_i}, c^{k_i})) \leq \frac{c^{k_i}}{\mu^{k_i}} \tilde{x}^i - \sum_{j=1}^n \omega_j \ln(\tilde{x}_j^i),$$

which is equivalent to

$$\frac{c^{k_i}}{\mu^{k_i}} (\tilde{x}^i - x(\mu^{k_i}, b^{k_i}, c^{k_i})) + \sum_{j=1}^n \omega_j \ln(x_j(\mu^{k_i}, b^{k_i}, c^{k_i})) \geq \sum_{j=1}^n \omega_j \ln(\tilde{x}_j^i). \quad (2.15)$$

Since  $\{\tilde{x}^i\}$  is bounded away from zero, the right hand side of this last inequality is bounded below. Suppose, for the sake of attaining a contradiction, that as

$i \rightarrow \infty$ ,  $x_j(\mu^{k_i}, b^{k_i}, c^{k_i}) \rightarrow 0$ , for some  $j$ . The boundedness of  $\{x(\mu^{k_i}, b^{k_i}, c^{k_i})\}$

implies

$$\sum_{j=1}^n \omega_j \ln \left( x_j(\mu^{k_i}, b^{k_i}, c^{k_i}) \right) \rightarrow -\infty.$$

Hence, 2.15 implies  $\frac{c^{k_i}}{\mu^{k_i}} \left( \tilde{x}^i - x(\mu^{k_i}, b^{k_i}, c^{k_i}) \right) \rightarrow \infty$ . However, since  $\hat{c} \in \text{row}(A)$

and  $(\tilde{x} - \hat{x}) \in \text{null}(A)$ ,

$$\frac{c^{k_i}}{\mu^{k_i}} \left( \tilde{x}^i - x(\mu^{k_i}, b^{k_i}, c^{k_i}) \right) \rightarrow \hat{c}(\tilde{x} - \hat{x}) = 0.$$

Hence, no such  $j$  exists, and  $\hat{x} > 0$ .

The dual argument is similar. Again, lemma 2.16 implies that

$$\left\{ \left( y(\mu^k, b^k, c^k), s(\mu^k, b^k, c^k) \right) \right\}$$

is bounded, and for some subsequence

$$\lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b^{k_i}, c^{k_i}), s(\mu^{k_i}, b^{k_i}, c^{k_i}) \right) = (\hat{y}, \hat{s}).$$

Clearly,  $(\hat{y}, \hat{s}) \in \mathcal{D}_{\bar{c}}$ . We show that  $(\hat{y}, \hat{s}) = (\bar{y}(\bar{c}), \bar{s}(\bar{c}))$ , from which the result

follows because of the uniqueness of  $(\bar{y}(\bar{c}), \bar{s}(\bar{c}))$ . Similar to the primal case, the

necessary and sufficient conditions describing  $\left( y(\mu^{k_i}, b^{k_i}, c^{k_i}), s(\mu^{k_i}, b^{k_i}, c^{k_i}) \right)$  are

rewritten as,

$$y(\mu^{k_i}, b^{k_i}, c^{k_i})A + s(\mu^{k_i}, b^{k_i}, c^{k_i}) = c$$

$$AS^{-1}(\mu^{k_i}, b^{k_i}, c^{k_i})\omega = \frac{b^{k_i}}{\mu^{k_i}}$$

$$s(\mu^{k_i}, b^{k_i}, c^{k_i}) > 0.$$

As before, the key element of the proof is showing that  $\hat{s} > 0$ . This is because

$S^{-1}(\mu^{k_i}, b^{k_i}, c^{k_i})\omega = \frac{x^i}{\mu^{k_i}}$ , and if  $\hat{s} > 0$ ,  $\left\{\frac{x^i}{\mu^{k_i}}\right\} \rightarrow \hat{x}$ . Hence,

$$\begin{aligned} A\hat{x} &= 0 \\ \hat{y}A + \hat{s} &= \bar{c} \\ \hat{S}\hat{x} &= \omega \\ \hat{s} &> 0 \end{aligned}$$

which implies that  $(\hat{y}, \hat{s}) = (\bar{y}(\bar{c}), \bar{s}(\bar{c}))$ . We proceed to show  $\hat{s} > 0$ .

Lemma 2.17 implies there exists a sequence  $\{(\tilde{y}^i, \tilde{s}^i) \in \mathcal{D}_{c^{k_i}}^o\}$  such that,  $\{\tilde{s}^i\} \rightarrow \tilde{s} \in \mathcal{D}_{\bar{c}}^o$ . The optimality of  $(y(\mu^{k_i}, b^{k_i}, c^{k_i}), s(\mu^{k_i}, b^{k_i}, c^{k_i}))$  implies

$$\tilde{y}^i \left( \frac{b^{k_i}}{\mu^{k_i}} \right) + \sum_{i=1}^n \omega_i \ln(\tilde{s}_i) \leq y(\mu^{k_i}, b^{k_i}, c^{k_i}) \left( \frac{b^{k_i}}{\mu^{k_i}} \right) + \sum_{i=1}^n \omega_i \ln(s_i(\mu^{k_i}, b^{k_i}, c^{k_i})),$$

or equivalently,

$$\sum_{i=1}^n \omega_i \ln(\tilde{s}_i) \leq (y(\mu^{k_i}, b^{k_i}, c^{k_i}) - \tilde{y}^i) \left( \frac{b^{k_i}}{\mu^{k_i}} \right) + \sum_{i=1}^n \omega_i \ln(s_i(\mu^{k_i}, b^{k_i}, c^{k_i})).$$

As before,  $\hat{s} > 0$  because the left-hand side of the above inequality is bounded below,  $\{s_i(\mu^{k_i}, b^{k_i}, c^{k_i})\}$  is bounded, and  $(y(\mu^{k_i}, b^{k_i}, c^{k_i}) - \tilde{y}^i) \left( \frac{b^{k_i}}{\mu^{k_i}} \right) \rightarrow 0$ . ■

## 2.5 Chapter Summary

The omega central path, primal omega central path, and dual omega central path were defined, and their analyticity with respect to  $(\mu, b, c)$  was

established. This result followed from a direct application of the implicit function theorem. The traditional convergence properties of  $x(\mu, b, c)$ , with  $b$  and  $c$  fixed, were developed after the boundedness of  $\{x(\mu, b, c) : 0 < \mu \leq \bar{\mu}\}$  was shown. This result was extended in theorem 2.14 to include the situation of converging rim data.

Because much of what is to follow pertains to linear changes in the rim data, notation to accommodate linear changes was introduced. The set of directions for which the optimal partition is invariant, for arbitrarily small amounts of change, was characterized in lemma 2.13. Theorem 2.18 provides sufficient conditions for a sequence  $\{x(\mu^k, b^k, c^k)\}$  to converge to the omega analytic center of a polytope.

### 3. Marginal Analysis of The Analytic Center Solution

#### 3.1 Introduction

This chapter develops the differential properties of the analytic center solution with respect to the rim data elements. Marginal properties of basic optimal solutions are well understood [19], and an analogy to this analysis is made. When dealing with a basic optimal solution, the marginal analysis is of penultimate concern for post optimal queries since the solution is linear along any given direction of right-hand side change. To see this, suppose that  $B$  is an optimal basis, which is compatible with  $\delta b$  [29]. Then, for sufficiently small  $\theta$ , this basic solution is

$$x^* = B^{-1}b + \theta B^{-1}\delta b. \tag{3.1}$$

So, the rate of change along  $\delta b$ , for this basic solution, is  $B^{-1}\delta b$ . Such rates of change are useful when anticipating the behavior of a particular solution under fluctuations in constraint levels.

When the primal and dual solutions are unique, the linearity of a basic solution along any given direction of right-hand side change makes the marginal analysis important and easy to interpret. However, when the primal

and dual solutions are not unique, a simplex based approach may have to pivot to find a compatible basis for the desired direction of change. In such situations, the degenerate pivots are needed to provide a basic variable that accommodates the sought after change.

The analysis presented in this chapter deals with the analytic center solution. Since this solution is unique, the analysis does not have to be divided into separate cases for unique and non-unique solutions. However, the analysis is more complicated because the linearity provided by a basic solution is lost. The main results of this chapter are:

- the analytic center solution is a continuous function of its right-hand side,
- the analytic center solution is analytic almost everywhere, and
- the analytic center solution is always infinitely, continuously one-sided differentiable.

The second property implies that the analytic center solution has a Taylor series expansion with a non-trivial radius of convergence, almost everywhere. Noticing that equation 3.1 is the Taylor expansion of a basic optimal solution, we see that such a Taylor expansion is the exact counterpart for the analytic center solution. Furthermore, the Taylor expansion is easy to calculate as seen in section 3.4.

### 3.2 Notation, Terminology, and Preliminary Results

Let  $f$  be a differentiable univariate function defined on the interval  $[x^0, x^1)$ . Define,

$$\begin{aligned}f'(x^0+) &\equiv \lim_{x \rightarrow x^0+} f'(x) \text{ and} \\f'_+(x^0) &\equiv \lim_{x \rightarrow x^0+} \frac{f(x) - f(x^0)}{x - x^0}.\end{aligned}$$

It is worth noting that even if  $f$  is continuous at  $x^0$ , the equality  $f'(x^0+) = f'_+(x^0)$ , is not guaranteed. For example, take  $f(x) = x \sin(x)$ , for  $x^0 = 0$ .

The “prime” notation is not precise enough when dealing with multi-variable functions. To accommodate the notational requirements needed for differentiable functions from  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ , define for a differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,

$$D_{x_i}^k f(x)$$

to be the vector whose  $j^{\text{th}}$  component is

$$\frac{\partial^k f_j(x)}{\partial x_i^k},$$

where  $f_j$  is the  $j^{\text{th}}$  component function. This notation is extended for the

right-sided limit of the derivative and the right-derivative as follows,

$$D_{x_i}^k f(\hat{x}+) \equiv \lim_{x_i \rightarrow \hat{x}_i+} D_{x_i}^k f \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_{i-1} \\ x_i \\ \hat{x}_{i+1} \\ \vdots \\ \hat{x}_n \end{pmatrix}$$

and

$$D_{x_i+}^k f(\hat{x}) \equiv \lim_{x_i \rightarrow \hat{x}_i+} \frac{1}{|x_i - \hat{x}_i|} \left( D_{x_i}^{k-1} f \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ x_i \\ \hat{x}_{i+1} \\ \vdots \\ \hat{x}_n \end{pmatrix} - D_{x_i}^{k-1} f(\hat{x}) \right).$$

The word *locally* is used to distinguish between one-sided analytic properties and analytic properties over a neighborhood. For example,  $x^*(b_\rho, c)$  is said to be *locally differentiable along*  $\mathfrak{d}$  if there exists a neighborhood about zero, say  $\mathcal{N}$ , such that  $D_\rho^1 x^*(b_\rho, c)$  exists for all  $\rho \in \mathcal{N}$ . Notice that to be locally differentiable along  $\mathfrak{d}$ , it is necessary that the directional derivatives corresponding

to  $\mathfrak{B}$  and  $-\mathfrak{B}$  agree when  $\rho$  is zero. In the next section, a set of directions with this property is classified.

The formula for the  $k$ th derivative of a composition function [24] is used to develop several results. Let  $h(x) = f(g(x))$ , where both  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$  are in  $\mathcal{C}^\infty$  on some suitable neighborhoods. Then  $h^{(k)}(x)$  is

$$\sum \frac{k!}{j_1!j_2!\dots j_k!} \cdot \frac{d^m f}{dy^m}(g(x)) \cdot \left[ \left( \frac{g'(x)}{1!} \right)^{j_1} \left( \frac{g''(x)}{2!} \right)^{j_2} \dots \left( \frac{g^{(k)}(x)}{k!} \right)^{j_k} \right], \quad (3.2)$$

where the sum is taken over all non-negative integer solutions to  $\sum_{i=1}^k ij_i = k$  and  $\sum_{i=1}^k j_i = m$ .

### 3.3 Differentiating the Central Path

The goal of this section is to establish the existence of the the limiting derivatives of the central path, with respect to  $\mu$ . The first order derivatives were addressed by Adler and Monteiro [1] and Witzgall, Boggs, and Domich [90]. The higher order derivatives were investigated by Güler [63], and this development is followed here. Although the marginal analysis of  $x^*(r)$  with respect to  $b$  and  $c$  is of primary interest, the later sections of this chapter show that this analysis follows once the existence of the limiting derivatives, with respect to  $\mu$ , is established.

The first task is to bound the derivatives with respect to  $\mu$ . Consider

the defining system of equalities for the central path:

$$Ax(\mu, b, c) = b \quad (3.3)$$

$$y(\mu, b, c)A + s(\mu, b, c) = c \quad (3.4)$$

$$S(\mu, b, c)x(\mu, b, c) = \omega\mu. \quad (3.5)$$

Using Leibnitz's rule, the  $k$ th differential of equation 3.5 is

$$\sum_{i=0}^k \binom{k}{i} D_\mu^{k-i} X(\mu, b, c) \cdot D_\mu^i s(\mu, b, c) = \begin{cases} \omega & \text{if } k = 1 \\ 0 & \text{if } k \geq 2, \end{cases} \quad (3.6)$$

where the ‘ $\cdot$ ’ simply means multiplication. For  $k \geq 2$ , define  $\wp^k(\mu, b, c)$  to be the above sum without the first and last terms:

$$\wp^k(\mu, b, c) = \sum_{i=1}^{k-1} \binom{k}{i} D_\mu^{k-i} X(\mu, b, c) \cdot D_\mu^i s(\mu, b, c).$$

Thus, when  $k \geq 2$ , the  $k$ th derivatives are iteratively the unique solution to

$$A \cdot D_\mu^k x(\mu, b, c) = 0 \quad (3.7)$$

$$D_\mu^k y(\mu, b, c) \cdot A + D_\mu^k s(\mu, b, c) = 0 \quad (3.8)$$

$$S(\mu, b, c) \cdot D_\mu^k x(\mu, b, c) + X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c) = -\wp^k(\mu, b, c). \quad (3.9)$$

Since 3.7 implies that  $D_\mu^k x(\mu, b, c) \in \text{null}(A)$  and 3.8 implies that  $D_\mu^k s(\mu, b, c) \in \text{row}(A)$ ,

$$D_\mu^k x(\mu, b, c)^T D_\mu^k s(\mu, b, c) = 0.$$

Recalling that  $\Omega = \text{diag}(\omega)$ , the equality  $X(\mu, b, c)S(\mu, b, c) = \mu\Omega$  implies,

$$\begin{aligned}
& (\Omega^{-\frac{1}{2}}S(\mu, b, c) \cdot D_\mu^k x(\mu, b, c))^T (\Omega^{-\frac{1}{2}}X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c)) \\
&= D_\mu^k x(\mu, b, c)^T \cdot S(\mu, b, c)\Omega^{-1}X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c) \\
&= \mu D_\mu^k x(\mu, b, c)^T \cdot D_\mu^k s(\mu, b, c) \\
&= 0.
\end{aligned}$$

Hence,  $\Omega^{-\frac{1}{2}}S(\mu, b, c) \cdot D_\mu^k x(\mu, b, c)$  and  $\Omega^{-\frac{1}{2}}X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c)$  are perpendicular. Now, for any two perpendicular vectors, say  $v$  and  $w$ ,

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

This, together with 3.9, show that,

$$\begin{aligned}
& \|\Omega^{-\frac{1}{2}}\varphi^k(\mu, b, c)\|^2 \\
&= \|\Omega^{-\frac{1}{2}}S(\mu, b, c) \cdot D_\mu^k x(\mu, b, c) + \Omega^{-\frac{1}{2}}X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c)\|^2 \\
&= \|\Omega^{-\frac{1}{2}}S(\mu, b, c) \cdot D_\mu^k x(\mu, b, c)\|^2 + \|\Omega^{-\frac{1}{2}}X(\mu, b, c) \cdot D_\mu^k s(\mu, b, c)\|^2.
\end{aligned}$$

Expressing the norms in the last equality as inner products yields

$$\begin{aligned}
& \|\Omega^{-\frac{1}{2}}\varphi^k(\mu, b, c)\|^2 \geq \\
& \min_i \left\{ \frac{1}{\omega_i} \right\} \left( \|S(\mu, b, c)D_\mu^k x(\mu, b, c)\|^2 + \|X(\mu, b, c)D_\mu^k s(\mu, b, c)\|^2 \right).
\end{aligned}$$

So,

$$\begin{aligned}
& \min_i \left\{ \frac{1}{\omega_i} \right\} \left( \|S(\mu, b, c) D_\mu^k x(\mu, b, c)\|^2 + \|X(\mu, b, c) D_\mu^k s(\mu, b, c)\|^2 \right) \\
& \leq \|\Omega^{-\frac{1}{2}} \wp^k(\mu, b, c)\|^2 \\
& \leq \|\Omega^{-\frac{1}{2}}\|^2 \|\wp^k(\mu, b, c)\|^2 \\
& = \max_i \left\{ \frac{1}{\omega_i} \right\} \|\wp^k(\mu, b, c)\|^2,
\end{aligned}$$

and

$$\begin{aligned}
& \|S(\mu, b, c) D_\mu^k x(\mu, b, c)\|^2 + \|X(\mu, b, c) D_\mu^k s(\mu, b, c)\|^2 \\
& \leq \frac{\min_i \{\omega_i\}}{\max_i \{\omega_i\}} \|\wp^k(\mu, b, c)\|^2.
\end{aligned} \tag{3.10}$$

This last inequality implies that  $D_\mu^1 x_N(\mu, b, c)$  and  $D_\mu^1 s_B(\mu, b, c)$  are uniformly bounded as  $\mu \rightarrow 0^+$ . This follows since  $\wp^1(\mu, b, c) = \omega$ ,  $x_B^*(r) > 0$ ,  $s_N^*(r) > 0$ , and

$$\begin{aligned}
& \|S(\mu, b, c) D_\mu^1 x(\mu, b, c)\|^2 + \|X(\mu, b, c) D_\mu^1 s(\mu, b, c)\|^2 \\
& = \left\| \begin{bmatrix} S_B(\mu, b, c) \cdot D_\mu^1 x_B(\mu, b, c) \\ S_N(\mu, b, c) \cdot D_\mu^1 x_N(\mu, b, c) \end{bmatrix} \right\|^2 + \left\| \begin{bmatrix} X_B(\mu, b, c) \cdot D_\mu^1 s_B(\mu, b, c) \\ X_N(\mu, b, c) \cdot D_\mu^1 s_N(\mu, b, c) \end{bmatrix} \right\|^2 \\
& \leq \frac{\min_i \{\omega_i\}}{\max_i \{\omega_i\}} \|\omega\|^2.
\end{aligned}$$

The next lemma gives a representation of  $D_\mu^1 x_B(\mu, b, c)$  in terms of  $D_\mu^k x_N(\mu, b, c)$ ,  $x_B(\mu, b, c)$ , and  $D_\mu^j x_B(\mu, b, c)$ , where  $j \leq k - 1$ . Such a representation allows the use of an induction argument to establish the convergence

of  $D_\mu^k x(\mu, b, c)$ . Before stating the needed representation of  $D_\mu^k x_B(\mu, b, c)$ , note that if  $f(x) = \frac{1}{x}$  and  $g(\nu) = x(\nu)$ , 3.2 shows that,

$$\begin{aligned} & \frac{d^k f(g(\nu))}{\nu^k} \\ &= \sum (-1)^m m! \frac{k!}{j_1! j_2! \dots j_k!} x(\nu)^{-m-1} \prod_{l=1}^k \left( \frac{x^{(l)}(\nu)}{l!} \right)^{j_l} \end{aligned} \quad (3.11)$$

$$= -\frac{x^{(k)}(\nu)}{x^2(\nu)} + \frac{1}{x(\nu)} h_{k-1} \left( x^{-m}(\nu), x'(\nu), x''(\nu), \dots, x^{k-1}(\nu) \right), \quad (3.12)$$

where  $-\frac{x^{(k)}(\nu)}{x^2(\nu)}$  corresponds to the index  $(j_1, j_2, \dots, j_k) = (0, 0, \dots, 0, 1)$ , and the function  $h_{k-1}$  defined by 3.12 is a polynomial. function.

The next lemma is found in [63]. The proof is included here for completeness.

**Lemma 3.1 (Güler [63])** Let  $r \in \mathcal{G}$  and define  $H_B(\mu, b, c) = \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c)$ .

Then, for every  $k \geq 1$  and  $\mu > 0$ ,

$$\begin{aligned} D_\mu^k x_B(\mu, b, c) = \\ -H_B(\mu, b, c) (A_B H_B(\mu, b, c))^\dagger A_N \cdot D_\mu^k x_N(\mu, b, c) + H_B(\mu, b, c) \cdot \pi. \end{aligned}$$

where  $\pi$  is the projection of

$$\Omega^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c) e_B, D_\mu^1 x_B(\mu, b, c), D_\mu^2 x_B(\mu, b, c), \dots, D_\mu^{k-1} x_B(\mu, b, c) \right),$$

onto  $\text{null}(A_B H_B(\mu, b, c))$  and  $h_{(k-1, B)}$  is the vector valued function such that the  $i^{\text{th}}$  component function, for  $i \in B$ , is the polynomial in 3.12.

**Proof:** Since  $s_B(\mu, b, c) - c_B \in \text{row}(A_B)$  and  $c_B = c_B - s_B^*(r) \in \text{row}(A_B)$ , we have that

$$\left( X_B^{-1}(\mu, b, c) \omega_B \right) \mu = s_B(\mu, b, c) \in \text{row}(A_B).$$

Hence,

$$X_B^{-1}(\mu, b, c) e_B \in \text{row}(\Omega_B^{-1} A_B)$$

and

$$D_\mu^k X_B^{-1}(\mu, b, c) e_B \in \text{row}(\Omega_B^{-1} A_B),$$

for all  $k \geq 1$ . Furthermore, 3.12 shows that,

$$\begin{aligned} & D_\mu^k X_B^{-1}(\mu, b, c) e_B \\ &= -X_B^{-2}(\mu, b, c) \cdot D_\mu^k x_B(\mu, b, c) \\ &\quad + X_B^{-1} e_B h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_\mu^1 x(\mu, b, c), \dots, D_\mu^{k-1} x(\mu, b, c) \right) \\ &\in \text{row} \left( \Omega_B^{-1} A_B \right). \end{aligned}$$

Multiplying on the left by  $\Omega_B^{\frac{1}{2}} X_B(\mu, b, c)$  provides,

$$\begin{aligned} & H_B^{-1}(\mu, b, c) \cdot D_\mu^k x_B(\mu, b, c) - \\ &\quad \Omega_B^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_\mu^1 x(\mu, b, c), \dots, D_\mu^{k-1} x(\mu, b, c) \right) \\ &\in \text{row} \left( A_B H_B(\mu, b, c) \right). \end{aligned}$$

For notational convince, let

$$\mathcal{V} = \text{row} \left( A_B H_B(\mu, b, c) \right) \text{ and } \mathcal{W} = \text{null} \left( A_B H_B(\mu, b, c) \right).$$

Since,  $\mathcal{V}$  and  $\mathcal{W}$  are perpendicular,

$$\begin{aligned}
& \text{proj}_{\mathcal{W}} \left( H_B^{-1}(\mu, b, c) \cdot D_\mu^k x_B(\mu, b, c) - \right. \\
& \quad \left. \Omega_B^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_\mu^1 x(\mu, b, c), \dots, D_\mu^{k-1} x(\mu, b, c) \right) \right) \\
&= \text{proj}_{\mathcal{W}} \left( H_B^{-1}(\mu, b, c) \cdot D_\mu^k(\mu, b, c) \right) \\
& \quad - \text{proj}_{\mathcal{W}} \left( \Omega_B^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_\mu^1 x(\mu, b, c), \dots, D_\mu^{k-1} x(\mu, b, c) \right) \right) \\
&= 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
& \text{proj}_{\mathcal{W}} \left( H_B^{-1}(\mu, b, c) \cdot D_\mu^k x_B(\mu, b, c) \right) \tag{3.13} \\
&= \text{proj}_{\mathcal{W}} \left( \Omega_B^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_\mu^1 x(\mu, b, c), \dots, D_\mu^{k-1} x(\mu, b, c) \right) \right).
\end{aligned}$$

To complete the orthogonal decomposition of

$$H_B^{-1}(\mu, b, c) \cdot D_\mu^k x_B(\mu, b, c),$$

the projection of this vector onto  $\mathcal{V}$  is needed. Since, equation 3.7 implies

$$A_B \cdot D_\mu^k x_B(\mu, b, c) + A_N \cdot D_\mu^k x_N(\mu, b, c) = 0,$$

we have

$$(A_B H_B) \left( H_B^{-1} \cdot D_\mu^k x_B(\mu, b, c) \right) = A_B \cdot D_\mu^k x_B(\mu, b, c) = -A_N \cdot D_\mu^k x_N(\mu, b, c).$$

Also, since  $(A_B H_B)^+(A_B H_B)$  is the projection operator onto  $\mathcal{V}$ , the last equality implies

$$\begin{aligned} & \text{proj}_{\mathcal{V}} \left( H_B^{-1} \cdot D_{\mu}^k x_B(\mu, b, c) \right) \\ &= -(A_B H_B)^+ A_N \cdot D_{\mu}^k x_N(\mu, b, c). \end{aligned} \quad (3.14)$$

Adding 3.13 and 3.14 shows that

$$\begin{aligned} & H_B^{-1}(\mu, b, c) \cdot D_{\mu}^k x_B(\mu, b, c) \\ &= \text{proj}_{\mathcal{W}} \left( \Omega_B^{\frac{1}{2}} h_{(k-1, B)} \left( X_B^{-m}(\mu, b, c), D_{\mu}^1 x(\mu, b, c), \dots, D_{\mu}^{k-1} x(\mu, b, c) \right) \right) \\ & \quad - (A_B H_B)^+ A_N \cdot D_{\mu}^k x_N(\mu, b, c). \end{aligned}$$

Multiplying both sides by  $H_B(\mu, b, c)$  proves the result. ■

The importance of lemma 3.1 lies not in the specific representation of  $D_{\mu}^k x_B(\mu, b, c)$ , but rather that this representation requires only lower order derivatives of  $x(\mu, b, c)$  and  $D_{\mu}^k x_N(\mu, b, c)$ . This allows the use of an induction argument to show that all limiting derivatives exist. The next theorem establishes this result and is found in [63].

**Theorem 3.2 (Güler [63])** Let  $r \in \mathcal{G}$ . Then,

$$\lim_{\mu \rightarrow 0^+} D_{\mu}^k x(\mu, b, c)$$

exists for all  $k \geq 1$ .

**Proof:** Assume  $k = 1$ . Then, the first derivative of  $x_B(\mu, b, c)$ , with respect to  $\mu$ , reduces to

$$\begin{aligned} D_\mu^1 x_B(\mu, b, c) = \\ -H_B(\mu, b, c) (A_B H_B(\mu, b, c))^+ A_N \cdot D_\mu^1 x_N(\mu, b, c). \end{aligned} \quad (3.15)$$

As already mentioned,  $D_\mu^1 x_N(\mu, b, c)$  is bounded as  $\mu \rightarrow 0^+$ . Furthermore, since  $x_B^*(r) > 0$ ,  $(A_B H_B(\mu, b, c))^+$  converges as  $\mu \rightarrow 0^+$ , see [63]. This implies that  $D_\mu^1 x_B(\mu, b, c)$  also remains bounded as  $\mu \rightarrow 0^+$ . From this boundedness and the fact that  $x_N^*(r) = 0$  and  $s_B^*(r) = 0$ , we see that

$$\begin{aligned} \lim_{\mu \rightarrow 0^+} X(\mu, b, c) D_\mu^1 s(\mu, b, c) + S(\mu, b, c) D_\mu^1 x(\mu, b, c) \\ = \begin{bmatrix} X_B^*(r) \cdot D_{\mu^+}^1 s_B^*(r) \\ S_N^*(r) \cdot D_{\mu^+}^1 x_N^*(r) \end{bmatrix} = \omega. \end{aligned} \quad (3.16)$$

So,

$$\begin{aligned} D_{\mu^+}^1 x_N^*(r) &= (S_N^*(r))^{-1} \omega_N \text{ and} \\ D_{\mu^+}^1 s_B^*(r) &= (X_B^*(r))^{-1} \omega_B. \end{aligned}$$

The existence of  $\lim_{\mu \rightarrow 0^+} D_\mu^1 x_B(\mu, b, c)$  is now implied by 3.15.

Suppose that the result is true for all  $j$  between 1 and  $k - 1$ . Since  $\wp^k(\mu, b, c)$  contains derivatives whose order is less than  $k$ , the induction hypothesis implies that

$$\lim_{\mu \rightarrow 0^+} \wp^k(\mu, b, c)$$

exists. Since,  $(x_B^*(r), s_N^*(r)) > 0$ , this implies that

$$D_\mu^k x_N(\mu, b, c) \text{ and } D_\mu^k s_B(\mu, b, c)$$

remain bounded as  $\mu$  approaches zero. Hence, the representations for

$$D_\mu^k x_B(\mu, b, c) \text{ and } D_\mu^k s_N(\mu, b, c),$$

as found in lemma 3.1, show that these derivative also stay bounded as  $\mu \rightarrow 0^+$ .

The result follows by letting  $\mu \rightarrow 0^+$  in 3.6, and applying the representations from lemma 3.1. ■

To illustrate theorem 3.2, consider the central path from example 2.12.

Supposing that  $c > 0$ ,

$$x_i(\mu, b, c) = \frac{c_i b_i + 2\mu - \sqrt{b_i^2 c_i^2 + 4\mu^2}}{2c_i},$$

and

$$\begin{aligned} \lim_{\mu \rightarrow 0^+} D_\mu^1 x_i(\mu, b, c) &= \lim_{\mu \rightarrow 0^+} \frac{1}{c_i} - \frac{\mu}{c_i^2 \sqrt{\frac{b_i^2}{4} + \frac{\mu^2}{c_i^2}}} \\ &= \frac{1}{c_i}. \end{aligned}$$

Some higher order limiting derivatives are found in table 3.1

Order	Derivative	Order	Derivative
1	$\frac{1}{c_i}$	6	$-\frac{2 \cdot 6!}{b_i^2 c_i^6}$
2	$-\frac{2}{b_i c_i^2}$	7	0
3	0	8	$\frac{5 \cdot 8!}{b_i^4 c_i^8}$
4	$\frac{4!}{b_i^3 c_i^4}$	9	0
5	0	10	$-\frac{14 \cdot 10!}{b_i^9 c_i^{10}}$

**Table 3.1.** Some higher order limiting derivatives from example 2.12

Now that the differential properties of  $x(\mu, b, c)$ , with respect to  $\mu$  are established, the next section shows how to exploit these properties to obtain similar differential properties with respect to the right-hand side.

### 3.4 Marginal Properties of the Analytic Center Solution with Respect to the Right-Hand Side

This section investigates the marginal properties of the analytic center solution with respect to linear changes in the right-hand side vector. It was seen in 2.3, 2.4, and lemma 2.13 that the optimal partition defines both the optimal sets and the cones of admissible directions for which the optimal partition is invariant, provided the perturbation is sufficiently small. The first result of this section establishes that the optimal partition also yields a set of right-hand

side directions along which the analytic center solution is locally analytic. The main result of this section is that even when an admissible direction is not a locally differentiable direction, one sided derivatives still exist.

Theorem 3.3 establishes the local analyticity of the analytic center solution along directions in  $\mathcal{H}_b$ . This result is quite simple, and all that is needed is to invoke the implicit function theorem.

**Theorem 3.3** Let  $r \in \mathcal{G}$ . Then, if  $\bar{d} \in \mathcal{H}_b$ ,  $x^*(b_\rho, c)$  is locally analytic along  $\bar{d}$ .

**Proof:** Lemma 2.13 implies  $\bar{d} \in \text{col}(A_B)$ . So, there exists a full rank matrix sub-matrix of  $A$ , say  $\bar{A}_B$ , with  $\text{rank}(\bar{A}_B) = m'$ , such that  $\{x : A_B x_B = b + \rho \bar{d}\} = \{x : \bar{A}_B x_B = \bar{b} + \rho \bar{d}\}$ . For sufficiently small  $\rho$ ,  $x_B^*(b_\rho, c)$  is defined by the following system of equations,

$$\bar{A}_B x_B(b_\rho, c) = \bar{b} + \rho \bar{d} \quad (3.17)$$

$$\gamma(b_\rho, c) \bar{A}_B + \varsigma(b_\rho, c) = 0 \quad (3.18)$$

$$X_B(b_\rho, c) \varsigma(b_\rho, c) = e. \quad (3.19)$$

Similar to the proof of theorem 2.6, define

$$\Psi : \mathbb{R}^{2|B|+m'+1} \rightarrow \mathbb{R}^{2|B|+m'} : (x_B, \gamma, \varsigma, \rho) \rightarrow \begin{bmatrix} \bar{A}_B x_B - \rho \bar{d} - \bar{b} \\ \gamma \bar{A}_B + \varsigma \\ X_B \varsigma^T - e \end{bmatrix}.$$

Then,  $\Psi$  is analytic in an open neighborhood of  $(x_B^*(\bar{b}_0, c), \gamma(\bar{b}_0, c), \varsigma(\bar{b}_0, c), 0)$  and  $\Psi\left(\left(x_B^*(\bar{b}_0, c), \gamma(\bar{b}_0, c), \varsigma(\bar{b}_0, c), 0\right)\right) = 0$ . Since the Jacobian of  $\Psi$  taken with respect to  $(x_B, \gamma, \varsigma)$  and evaluated at  $\left(x_B^*(\bar{b}_0, c), \gamma(\bar{b}_0, c), \varsigma(\bar{b}_0, c), 0\right)$  is non-singular, the implicit function theorem implies that  $x_B^*(b_\rho, c)$  is an analytic function of  $\rho$  in some sufficiently small neighborhood of zero.  $\blacksquare$

Notice that it is important for  $\bar{\mathfrak{b}} \in \mathcal{H}_b$ , since if it were not, row reduction could not be used to form equations 3.17, 3.18, and 3.19. Theorem 3.3 implies that along a direction of change that does not immediately alter the partition, not only is  $x^*(b_\rho, c)$  of class  $\mathcal{C}^\infty$ , but  $x^*(b_\rho, c)$  has a power series expansion. Differentiating 3.17, 3.18 and 3.19 with respect to  $\rho$  shows,

$$\begin{aligned} \bar{A}_B \cdot D_\rho^1 x_B^*(b_\rho, c) &= \bar{\mathfrak{b}} \\ D_\rho^1 \gamma(b_\rho, c) \cdot \bar{A}_B + D_\rho^1 \varsigma(b_\rho, c) &= 0 \\ \text{diag}(\varsigma(b_\rho, c)) \cdot D_\rho^1 x_B^* + X_B^* \cdot \left(D_\rho^1 \varsigma(b_\rho, c)\right)^T &= 0. \end{aligned} \quad (3.20)$$

This is a non-singular system of linear equations in  $D_\rho^1 x_B(b_\rho, c)$ ,  $D_\rho^1 \gamma(b_\rho, c)$  and  $D_\rho^1 \varsigma(b_\rho, c)$ . The following are seen to hold after a some algebraic manipulations,

$$D_\rho^1 x_B^*(b_\rho, c) = (X_B^*(b_\rho, c))^2 \bar{A}_B^T (\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{\mathfrak{b}}, \quad (3.21)$$

$$D_\rho^1 \varsigma(b_\rho, c) = \bar{A}_B^T (\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{\mathfrak{b}}, \text{ and}$$

$$D_\rho^1 \gamma(b_\rho, c) = (\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{\mathfrak{b}}. \quad (3.22)$$

Equation 3.20 is used iteratively to establish the higher order derivatives. Similar to the previous section, define

$$\varphi_B^k(b_\rho, c) = \sum_{i=1}^{k-1} \binom{k}{i} \text{diag} \left( D_\rho^i \zeta(b_\rho, c) \right) \cdot D_\rho^{k-i} x^*(b_\rho, c).$$

Then, for  $k \geq 2$ ,

$$\begin{aligned} \bar{A}_B \cdot D_\rho^k x_B^*(b_\rho, c) &= 0 \\ D_\rho^k \gamma(b_\rho, c) \bar{A}_B + D_\rho^k \zeta(b_\rho, c) &= 0 \\ \text{diag}(\zeta(b_\rho, c)) \cdot D_\rho^k x_B + X_B^* \cdot \left( D_\rho^k \zeta(b_\rho, c) \right)^T &= -\varphi_B^k(b_\rho, c), \end{aligned}$$

which implies

$$\begin{aligned} D_\rho^k x_B^*(b_\rho, c) &= \varphi_B^k(b_\rho, c) - (X_B^*(b_\rho, c))^2 \bar{A}_B^T (\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{A}_B (X_B^*(b_\rho, c))^2 \varphi_B^k(b_\rho, c), \\ D_\rho^k \zeta(b_\rho, c) &= \bar{A}_B^T (\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{A}_B (X_B^*(b_\rho, c))^2 \varphi_B^k(b_\rho, c), \text{ and} \\ D_\rho^k \gamma(b_\rho, c) &= -(\bar{A}_B (X_B^*(b_\rho, c))^2 \bar{A}_B^T)^{-1} \bar{A}_B (X_B^*(b_\rho, c))^2 \varphi_B^k(b, c). \end{aligned}$$

Since  $x_N^*(b_\rho, c) = 0$  for sufficiently small  $\rho$ ,  $D_\rho^k x_N^*(b_\rho, c) = 0$  for all  $k$  and sufficiently small  $\rho$ . The Taylor expansion for  $x^*(b_\rho, c)$  is

$$x^*(b_\rho, c) = \sum_{k=0}^{\infty} D_\rho^k x^*(b_0, c) \rho^k.$$

The radius of convergence of this power series still needs investigation. Currently, interior point solvers are not capable of using this technique to approximate a solutions behavior under right-hand side perturbation; however, the creator of LIPSOL wishes to add these capabilities [99]. Since all the derivatives rely on the same matrix factorization, the Taylor expansion is relatively cheap as compared to re-solving the problem for several right-hand sides.

The next aim is to establish that even when  $\mathfrak{b}$  is not in  $\mathcal{H}_b$ , one-sided derivatives still exist. Four lemmas are presented in support of this aim. The first of these lemmas was originally shown by Berge in ?? (also see [9, 34]), and states that the limit of any convergent sequence in  $\mathcal{P}_{r^k}^* \times \mathcal{D}_{r^k}^*$  is contained in  $\mathcal{P}_r^* \times \mathcal{D}_r^*$ , provided  $\{r^k\} \rightarrow r$ .

**Lemma 3.4 (Berge [6])** Let  $A$  be an  $m \times n$  matrix,  $r \in \mathcal{G}$ , and consider the point-to-set mapping

$$\mathcal{X}(r) = \{(x, y, s) : (x, y, s) \in \mathcal{P}_r^* \times \mathcal{D}_r^*\}.$$

Then  $\mathcal{X}$  is a closed map.

**Proof:** Let  $\{r^k \in \mathcal{G}\} \rightarrow r \in \mathcal{G}$ , and  $\{(x^k, y^k, s^k) \in \mathcal{P}_{r^k}^* \times \mathcal{D}_{r^k}^*\} \rightarrow (x, y, s)$ .

Then,

$$\begin{aligned} Ax^k &= b^k \\ y^k A + s^k &= c^k \end{aligned}$$

$$S^k x^k = 0$$

$$x^k \geq 0$$

$$s^k \geq 0$$

Allowing  $k \rightarrow \infty$  produces

$$Ax = b$$

$$yA + s = c$$

$$Sx = 0$$

$$x \geq 0$$

$$s \geq 0,$$

which implies  $(x, y, s) \in \mathcal{P}_r^* \times \mathcal{D}_r^*$ . ■

The next lemma provides a monotonicity for the optimal partition, with respect to the right-hand side. Results similar to lemma 3.5 are found in [2, 40], and are used in [35]. However, these results are only pertinent to linear changes in the right-hand side, and not general changes in  $b$ . Furthermore, the results in both [2, 40] are developed from a parametric analysis of the objective function value, while development below is a simple polyhedral argument.

**Lemma 3.5** Let  $c \in \mathcal{G}_c$  and  $\{b^k \in \mathcal{G}_b\} \rightarrow b \in \mathcal{G}_b$ . Furthermore, let  $(B^k, N^k)$  be the optimal partition  $LP_{(b^k, c)}$  and  $(B|N)$  be the optimal partition for  $LP_{(b, c)}$ .

Then, for sufficiently large  $k$ ,

$$B \subseteq B^k \text{ and } N^k \subseteq N.$$

**Proof:** We prove the equivalent statement that  $\mathcal{D}_{(b^k, c)}^* \subseteq \mathcal{D}^*(b, c)$ , for sufficiently large  $k$ . Suppose, for the sake of attaining a contradiction, that for all natural numbers  $K$ , there exists  $k \geq K$  such that

$$\mathcal{D}_{(b^k, c)}^* \not\subseteq \mathcal{D}^*(b, c).$$

Then there exists a subsequence  $\{b^{k_i}\} \rightarrow b$  such that, for all  $i$ ,

$$\mathcal{D}_{(b^{k_i}, c)}^* \not\subseteq \mathcal{D}^*(b, c).$$

Because there exists  $T$  such that,

$$\{\mathcal{D}_{(b^{k_i}, c)}^* : i = 1, 2, \dots\} = \{\mathcal{D}_{(b^{k_t}, c)}^* : t = 1, 2, \dots, T\},$$

This means there exists

$$\{(x^{k_i}, y^{k_i}, s^{k_i}) \in \mathcal{P}_{(b^{k_i}, c)}^* \times \mathcal{D}_{(b^{k_i}, c)}^*\},$$

which contains only a finite number of distinct elements, say

$$\{(x^{p_1}, y^{p_1}, s^{p_1}), (x^{p_2}, y^{p_2}, s^{p_2}), \dots, (x^{p_q}, y^{p_q}, s^{p_q})\},$$

and for  $j = 1, 2, \dots, q$ ,

$$(y^{p_j}, s^{p_j}) \in \mathcal{D}_{(b^{k_j}, c)}^* \setminus \mathcal{D}_{(b, c)}^*.$$

Since none of the cluster points of this sequence are in  $\mathcal{D}_{(b,c)}^*$ , this is in contradiction to lemma 3.4. Hence, for sufficiently large  $k$ ,  $\mathcal{D}_{(b^k,c)}^* \subseteq \mathcal{D}_{(b,c)}^*$ , which implies  $B \subseteq B^k$  and  $N^k \subseteq N$ . ■

The third lemma in support of showing that  $x^*(r)$  is one-sided, infinitely, continuously, differentiable along any  $\mathcal{D}$  states that  $x^*(r)$  is a continuous function of  $b$ . Continuity is the strongest analytic property allowed for  $x^*(b, c)$  over all  $b \in \mathcal{G}_b$ , when  $c$  is fixed. This is because the objective function parameterized with respect to the right-hand side is piecewise linear (see [2, 39] for a recent development using the analytic center solution). This means that the analytic center solution is not differentiable at the sharp corners of this parameterized objective function. The continuity result found in the following lemma is frequently used in the next chapter.

**Lemma 3.6** Let  $\{r^k = (b^k, c) \in \mathcal{G}\} \rightarrow r \in \mathcal{G}$ . Then,

$$\lim_{k \rightarrow \infty} x^*(r^k) = x^*(r).$$

**Proof:** We show that any convergent subsequence of  $\{x^*(b^k, c)\}$  converges to  $\{x^*(b, c)\}$ . From lemma 2.14,

$$\bigcup_k \mathcal{P}_{(b^k,c)}^* \times \mathcal{D}_{(b^k,c)}^*$$

is bounded. Without loss in generality, assume that  $\{x^*(b^k, c)\} \rightarrow \hat{x}$ . Let

$(B^k|N^k)$  be the optimal partition for  $LP_{(b^k,c)}$  and  $(B|N)$  be the optimal partition for  $LP_{(b,c)}$ . From lemma 3.5,  $B \subseteq B^k$ , for sufficiently large  $k$ . This means  $x_B^*(b^k, c) > 0$ , for sufficiently large  $k$ . Because  $x_B^*(b^k, c)$  is the omega analytic center of

$$\{x_B : A_B x_B = b^k - A_N x_N^*(b^k, c), x_B \geq 0\},$$

the following equation is feasible,

$$y A_B = e^T (X_B^*(b^k, c))^{-1}.$$

Set  $y^k = e^T (X_B^*(b^k, c))^{-1} A_B^+$ . Since  $\{(X_B^*(b^k, c))^{-1}\}$  converges,  $\{y^k\}$  converges, say to  $\hat{y}$ . From lemma 2.15  $\{x_N^*(b^k, c)\} \rightarrow 0$ . Hence, there exists  $y$  such that

$$\begin{aligned} A_B \hat{x}_B &= b \\ y A_B &= e^T \hat{X}_B^{-1} \\ \hat{x}_B &> 0. \end{aligned}$$

These are the necessary and sufficient Lagrange conditions for  $x_B^*(b, c)$ . So,  $\{x^*(b^k, c)\} \rightarrow x^*(b, c)$ . ■

The last of the four supporting lemmas for theorem 3.8 provides some one-sided analytic properties of functions and their inverses.

**Lemma 3.7** Consider the function  $f : [0, \alpha^*) \rightarrow [0, \beta^*)$  with the following properties:

- $f([0, \alpha^*)) = [0, \beta^*)$ ,
- $f$  is in  $\mathcal{C}^\infty$  on  $(0, \alpha^*)$ ,
- $f'(x) > 0$  for all  $x \in (0, \alpha^*)$ ,
- $f$  is continuous on  $[0, \alpha^*)$ ,
- $f^{(k)}(0+)$  exists for all  $k \geq 1$ , and
- $f'(0+) > 0$ .

Then,  $g = f^{-1}$  exists and has the property that  $g^{(k)}(0+)$  exists for all  $k \geq 1$ .

Furthermore,

$$f^{(k)}(0+) = f_+^{(k)}(0) \quad \text{and} \quad g^{(k)}(0+) = g_+^{(k)}(0).$$

**Proof:** Let  $f$  be as above. The general inverse function theorem [62] establishes that  $g = f^{-1}$  exists and is in  $\mathcal{C}^\infty$  on  $(0, \beta^*)$ . The mean value theorem guarantees that for all  $\alpha \in (0, \alpha^*)$ , there exists  $t(\alpha) \in (0, \alpha)$  such that,

$$\frac{f(\alpha) - f(0)}{\alpha} = f'(t(\alpha)).$$

From the assumption that  $f'(0+)$  exists,

$$\lim_{\mu \rightarrow 0^+} \left| \frac{f(\mu) - f(0)}{\mu} - f'(\mu) \right| = \lim_{\mu \rightarrow 0^+} |f'(t(\mu)) - f'(\mu)| = 0.$$

So,  $f'_+(0) = f'(0+)$  and  $f'$  is continuous at zero. Repeated use of the mean value theorem establishes that  $f$  is in  $\mathcal{C}^\infty$  on  $[0, \alpha^*)$ .

Because  $f$  is onto  $[0, \beta^*)$  and strictly increasing on  $[0, \alpha^*)$ ,  $f(0) = g(0) = 0$ . Since

$$g'(\beta) = \frac{1}{(f'(g(\beta)))}, \quad (3.23)$$

for all  $\beta \in (0, \beta^*)$ , the assumption that  $f'(0+) > 0$  immediately implies

$$g'(0+) = \frac{1}{(f'(0+))}.$$

Applying 3.11 to 3.23 shows that  $g^{(k)}(\beta)$  is expressible in terms of  $f'(g(\beta))$ ,  $f''(g(\beta))$ ,  $f'''(g(\beta)) \dots f^{(k)}(g(\beta))$ . So  $g^{(k)}(0+)$  exists by induction.

Now, suppose for the sake of attaining a contradiction that  $g$  is not continuous at 0. This means there exists a sequence  $\{\beta^k \in [0, \beta^*)\} \rightarrow 0$  such that  $\{g(\beta^k)\} \not\rightarrow 0$ . Since  $f$  is a bijection, there exists a unique sequence,  $\{x^k \in [0, \alpha^*)\}$ , such that  $f(x^k) = \beta^k$ . This means the supposition is equivalent to  $\{f(x^k)\} \rightarrow 0$  and  $\{x^k\} \not\rightarrow 0$ . So there exists  $\epsilon > 0$  and natural number  $K$  such that  $\{x^k : k \geq K\} \subset (\epsilon, \alpha^*)$ , which implies  $\{f(x^k)\} \subset f((\epsilon, \alpha^*))$ . Since  $g$  is continuous on  $(0, \beta^*)$  and  $g(f((\epsilon, \alpha^*))) = (\epsilon, \alpha^*)$ ,  $f((\epsilon, \alpha^*))$  is open. Furthermore,  $0 \notin f((\epsilon, \alpha^*))$ , and there exists  $\delta > 0$  such that  $[0, \delta) \cap f((\epsilon, \alpha^*)) = \emptyset$ . However, this implies the contradiction that  $\{f(x^k)\} \not\rightarrow 0$ .

Since  $g$  is continuous at 0, the mean value theorem is used as before to show that  $g^{(k)}(0+) = g_+^{(k)}(0)$ . ■

The tools are now in hand to prove the main result of this section.

Theorem 3.3 already established that  $x^*(r)$  is locally analytic along any direction of right-hand side change for which the optimal partition is not immediately altered, and the next result shows that even when a change in the optimal partition is forced, one sided derivatives still exist.

**Theorem 3.8** Let  $r \in \mathcal{G}$  and  $\delta \notin \mathcal{H}_b$ . Then,  $D_\rho^k x^*(b_{0+}, c)$  exists for all  $k$  and

$$D_{\rho^+}^k x^*(b_0, c) = D_\rho^k x^*(b_{0+}, c).$$

**Proof:** Let  $r$  and  $\delta$  be as above. From [28], there exists  $\bar{\rho} > 0$  such that  $(B((b + \rho\delta, c)) | N((b, +\rho\delta, c)))$  is unchanged for all  $\rho \in (0, \bar{\rho})$ . Denote this optimal partition by  $(B' | N')$ . Lemma 3.5 implies  $\Delta B \equiv B' \setminus B(r) \neq \emptyset$ . If  $i \in N \setminus \Delta B$ ,  $x_i^*(b_\rho, c) = 0$  for all  $\rho \in [0, \bar{\rho})$ . Consider the following linear program which has an optimal value of zero:

$$\min\{\rho : A_B z_B + A_{\Delta B} z_{\Delta B} - \rho\delta = b, z_B \geq 0, z_{\Delta B} \geq 0, \rho \geq 0\}.$$

Let  $\{(z_B(\mu), z_{\Delta B}(\mu), \rho(\mu)) : \mu > 0\}$  be the central path. Then

$$z_B(\mu) = x_B^*(b_{\rho(\mu)}, c) \quad , \quad z_{\Delta B}(\mu) = x_{\Delta B}^*(b_{\rho(\mu)}, c),$$

and from the continuity of  $x^*(r)$  in its right-hand side,

$$z_B(\mu) \rightarrow x_B^*(r) > 0 \quad , \quad z_{\Delta B}(\mu) \rightarrow x_{\Delta B}^*(r) = 0,$$

as  $\mu \rightarrow 0^+$ . Furthermore,  $\rho(\mu) \rightarrow 0$ , as  $\mu \rightarrow 0^+$ , and the optimal partition is  $(B|\Delta B \cup \{|B \cup \Delta B| + 1\})$ , where the index  $|B \cup \Delta B| + 1$  corresponds to  $\rho$ .

The proof of theorem 3.2 implies that  $\rho'_+(0) > 0$ , and hence there exists some interval, say  $[0, \mu^*)$ , where  $\rho(\mu)$  is invertible. Denote the inverse by  $\mu(\rho)$  and let the corresponding interval be  $[0, \rho^*)$ . Then for all  $\rho \in (0, \rho^*)$ ,

$$z_B(\mu(\rho)) = x_B^*(b_\rho, c) \quad , \quad \text{and} \quad z_{\Delta B}(\mu(\rho)) = x_{\Delta B}^*(b_\rho, c).$$

So, for all  $i \in B \cup \Delta B$ , equation 3.2 implies

$$\begin{aligned} & D_\rho^k x_i^*(b_\rho, c) \\ &= \sum \frac{k!}{j_1! j_2! \dots j_k!} \cdot \frac{d^m z_i}{d\mu^m}(\mu(\rho)) \cdot \left[ \left( \frac{\mu'(\rho)}{1!} \right)^{j_1} \left( \frac{\mu''(\rho)}{2!} \right)^{j_2} \dots \left( \frac{\mu^{(k)}(\rho)}{k!} \right)^{j_k} \right]. \end{aligned}$$

Lemma 3.7 implies that  $\mu(\rho) \rightarrow 0^+$  as  $\rho \rightarrow 0^+$ , and theorem 3.2 consequently shows the existence of

$$\lim_{\rho \rightarrow 0^+} \frac{d^m z_i}{d\mu^m}(\mu(\rho)).$$

Lemma 3.7 also shows the existence of

$$\lim_{\rho \rightarrow 0^+} \mu^{(i)}(\rho),$$

for  $i = 1, 2, \dots, k$ . So  $D_\rho^k x^*(b_{0^+}, c)$  exists. From lemmas 3.4 and 3.6,  $x^*(b_\rho, c)$  is a continuous function of  $\rho$ , from which lemma 3.7 shows  $D_{\rho^+}^k x^*(b_0, c) = D_\rho^k x^*(b_{0^+}, c)$ . ■

This section ends with an example that demonstrates many of the results just developed.

**Example 3.9** Consider the linear program  $\min\{-x_2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_1 + x_2 \leq b_3\}$ , which is shown in figure 3.1. After adding a slack vector,

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ b_3 \end{pmatrix}, \text{ and } c = (0, -1, 0, 0, 0).$$

The analytic center solution is easily verified as

$$(x_1^*(r), x_2^*(r)) = \begin{cases} \left( \frac{1}{3} \left[ b_3 + 2 - \sqrt{b_3^2 + b_3 + 7} \right], 1 \right) & \text{if } b_3 > 1 \\ (0, b_3) & \text{if } 0 \leq b_3 \leq 1 \end{cases}$$

The optimal partitions are

$$(\{1, 2, 3, 5\}, \{4\}), (\{2, 3\}, \{1, 4, 5\}), \text{ and } (\{2, 3, 4\}, \{1, 5\}),$$

for  $b_3 > 1$ ,  $b_3 = 1$ , and  $0 < b_3 < 1$ , respectively. Furthermore, since

$$\lim_{b_3 \rightarrow 1^+} \frac{1}{3} \left[ b_3 + 2 - \sqrt{b_3^2 + b_3 + 7} \right] = 0,$$

$x_1^*(r)$  is continuous in  $b_3$ .

Let  $\delta b = (0, 0, 1)^T$ . Then,  $\delta b \in \text{col}(A_B)$  when  $b_3 > 1$  and  $0 < b_3 < 1$ . When  $b_3 > 1$ ,

$$x_1^*(r) = \frac{1}{3} \left( b_3 + \rho + 2 - \sqrt{(b_3 + \rho)^2 + (b_3 + \rho) + 7} \right),$$

which is a locally analytic function of  $\rho$  on  $(1 - b_3, b_3 - 1)$ . Similarly, when  $0 < b_3 < 1$ ,

$$x_2^*(r) = b_3 + \rho,$$

which is locally analytic on  $(-b_3, 1 - b_3)$ . Hence, theorem 3.3 is validated for this example.

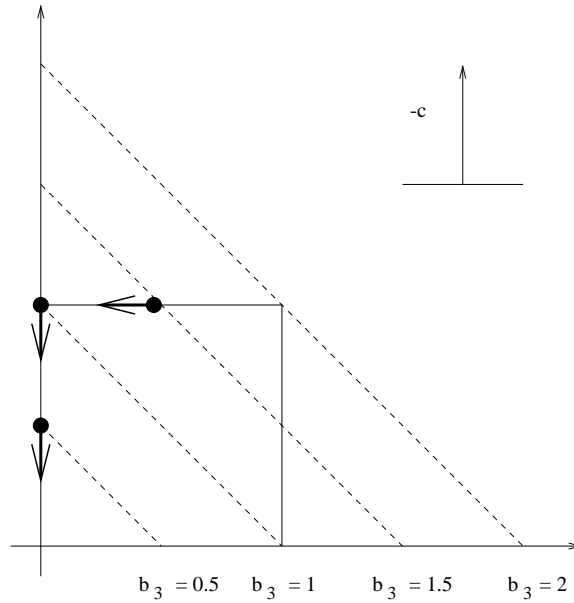
Set  $b_3 = 1$ . To illustrate the application of theorem 3.8, both limiting and one-sided derivatives need to be calculated. The limiting right-sided derivative is

$$\lim_{\rho \rightarrow 0^+} D_{b_3}^1 x^*(r) = \lim_{\rho \rightarrow 0^+} \begin{pmatrix} \frac{1}{3} \left[ 1 - \frac{2\rho+3}{2\sqrt{\rho^2+3\rho+9}} \right] \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ 0 \end{pmatrix}.$$

Since,  $D_{\rho^+}^1 x_2^*(b + \rho\delta, c) = 0$  and

$$\begin{aligned} D_{\rho^+}^1 x_1^*(b + \rho\delta, c) &= \lim_{\rho \rightarrow 0^+} \frac{3 + \rho - \sqrt{\rho^2 + 3\rho + 9}}{3\rho} \\ &= \frac{1}{6}, \end{aligned}$$

the equality,  $D_{b_3^+}^1 x^*(b + \rho\delta, c) = D_{b_3}^1 x^*(b + \rho\delta, c)$ , holds. In Figure 3.1, The darkened circles indicate how the analytic center solution moves as  $b_3$  decreases from infinity to 0. When  $b_3 = 1$ , the optimal partition changes under any perturbation. As just shown, the first order limiting derivatives agree with the one-sided derivatives when  $b_3 = 1$ .



**Figure 3.1.** An example of tracking the analytic center and its marginal derivatives.

### 3.5 Bounding the Derivatives of the Analytic Center Solution

The previous section showed that  $x^*(r)$  is a continuous function of  $b$  and is analytic along  $(\delta b, 0)$  when  $\delta b \in \text{col}(A_B)$ . Furthermore, theorem 3.8 established that even when  $x^*(r)$  is not differentiable along  $\delta b$ , it is one-sided, infinitely, continuously differentiable. The two main results of this section provide uniform bounds on these derivatives. The first of these results show that the derivatives of  $x^*(r)$  along  $\delta b$  are uniformly bounded. The second of

these results shows that  $D_{\mu^+}^1 x^*(r)$  is uniformly bounded over

$$\{w \in \mathbb{R}_{++}^n : \|\omega\| = 1\}.$$

Both of these results use an observation due to Stewart [86] (see also [10, 14, 87]), which is stated in lemma 3.10.

Define  $\mathbf{D}_{++}$  to be the set of positive diagonal matrices in  $\mathbb{R}^{m \times m}$ . For any  $D \in \mathbf{D}_{++}$  set  $A_D = D^{\frac{1}{2}}A$  and  $P_D = AA_D^+D^{\frac{1}{2}}$ . The following lemma shows that  $P_D$  is uniformly bounded over  $\mathbf{D}_{++}$ .

**Lemma 3.10 (Stewart [86])** There is a number  $k > 0$ , such that

$$\sup_{D \in \mathbf{D}_{++}} \|P_D\| \leq k.$$

**Proof:** Let

$$C = \{u \in \text{col}(A) : \|u\| = 1\}$$

and

$$K = \bigcup_{D \in \mathbf{D}_{++}} \text{leftnull}(DA).$$

Denoting the closure of  $K$  by  $\bar{K}$ , the first fact to establish is that  $\bar{K} \cap C = \emptyset$ .

This would imply

$$\inf_{(u,v) \in C \times K} \|v - u\| > 0,$$

which is then used to complete the result. Suppose for the sake of attaining a contradiction that  $\bar{K} \cap C \neq \emptyset$ . Then there exists a sequence  $\{v^k \in K\} \rightarrow u \in C$ . For every  $k$  there exists  $D^k \in \mathbf{D}_{++}$  such that  $v^k \in \text{leftnull}(D^k A)$ . So,  $(D^k)^{\frac{1}{2}} v^k \in \text{leftnull}((D^k)^{\frac{1}{2}} A)$  and since  $u \in \text{col}(A)$ ,  $(D^k)^{\frac{1}{2}} u \in \text{col}((D^k)^{\frac{1}{2}} A)$ . Hence,

$$0 = v^k D^k u = \sum_{u_i \neq 0} u_i d_i^k v_i.$$

Since,  $\|u\| = 1$ , some  $u_i \neq 0$ . However, for sufficiently large  $k$ ,  $u_i v_i^k > 0$  when  $u_i \neq 0$ , which implies the contradiction that the right-hand side of the last equality is positive. Hence,  $\bar{K} \cap C = \emptyset$ .

Define

$$k \equiv \frac{1}{\inf_{(u,v) \in C \times K} \|u - v\|} > 0.$$

For any fixed  $D \in \mathbf{D}_{++}$ ,  $\|P_D\| = \max_{\|w\|=1} \|P_D w\|$ . Let  $\|w\| = 1$  and  $w = u + v$ , where  $u \in \text{col}(A)$  and  $v \in \text{leftnull}(DA)$ . Then  $\|P_D w\| = \|u\|$  and to bound  $\|P_D\|$  we show that  $\|u\|$  is bounded. Without loss of generality, it is assumed that  $u \neq 0$ . Since  $\frac{u}{\|u\|} \in C$  and  $\frac{v}{\|u\|} \in \text{leftnull}(DA)$ ,

$$\begin{aligned} \frac{1}{\|u\|} &= \frac{\|w\|}{\|u\|} \\ &= \frac{\|u + v\|}{\|u\|} \\ &= \left\| \frac{u}{\|u\|} + \frac{v}{\|u\|} \right\| \\ &\leq \frac{1}{k}. \end{aligned}$$

So,  $\|P_D\| \leq k$ . Since  $k$  is independent of  $D$ ,

$$\sup_{D \in \mathbf{D}_{++}} \|P_D\| \leq k.$$

■

The following corollary contains the exact statements used to establish the desired bounds on the derivatives.

**Corollary 3.11 (Stewart [86])** If  $A$  has full column rank, there exists  $k$  such that

$$\sup_{D \in \mathbf{D}_{++}} \|(A^T D A)^{-1} A D\| \leq k$$

and

$$\sup_{D \in \mathbf{D}_{++}} \|(D^{\frac{1}{2}} A)^+ D^{\frac{1}{2}}\| \leq \|A^+\| k.$$

**Proof:** Since the full column rank of  $A$  implies

$$P_D = (A^T D A)^{-1} A D,$$

the proof of the first inequality is immediate from lemma 3.10, Furthermore, the full column rank assumption implies

$$\begin{aligned} A^+ P_D &= (A^T A)^{-1} A^T A (A^T D A)^{-1} A^T D \\ &= (A^T D A)^{-1} A^T D \\ &= (D^{\frac{1}{2}} A)^+ D^{\frac{1}{2}}. \end{aligned}$$

Hence,

$$\begin{aligned}
\sup_{D \in \mathbf{D}_{++}} \|(D^{\frac{1}{2}}A)^+ D^{\frac{1}{2}}\| &= \sup_{D \in \mathbf{D}_{++}} \|A^+ P_D\| \\
&\leq \|A^+\| \sup_{D \in \mathbf{D}_{++}} \|P_D\| \\
&\leq \|A^+\| k.
\end{aligned}$$

■

Corollary 3.11 allows the following definition for any full column rank matrix,

$$\chi_A \equiv \sup_{D \in \mathbf{D}_{++}} \|(A^T D A)^{-1} A D\|.$$

The next theorem establishes that the first order derivatives along any  $\delta b$  are uniformly bounded.

**Theorem 3.12** Let  $r \in \mathcal{G}$  and  $\delta b$  be fixed. Then,

$$\|D_\rho^1 x(\mu, b_\rho, c)\| \leq \chi_{A^T} \|\delta b\|.$$

Furthermore, if  $\delta b \in \mathcal{H}_b$ ,

$$\|D_\rho^1 x^*(b_\rho, c)\| \leq \chi_{\bar{A}^T} \|\bar{\delta b}\|,$$

and if  $\delta b \notin \mathcal{H}_b$ ,

$$\|D_{\rho^+}^1 x^*(b_0, c)\| \leq \chi_{\bar{A}^T} \|\bar{\delta b}\|,$$

where  $\bar{A}$  and  $\bar{\delta b}$  are defined as in the proof of theorem 3.3.

**Proof:** Let  $r$  and  $\delta$  be as stated.  $D_\rho^1 x(\mu, b_\rho, c)$  is defined by the following equations:

$$\begin{aligned} AD_\rho^1 x(\mu, b_\rho, c) &= \delta \\ D_\rho^1 y(\mu, b_\rho, c)A + D_\rho^1 s(\mu, b_\rho, c) &= 0 \\ S(\mu, b_\rho, c)D_\rho^1 x(\mu, b_\rho, c) + X(\mu, b_\rho, c)D_\rho^1 s(\mu, b_\rho, c) &= 0, \end{aligned}$$

which implies

$$D_\rho^1 x(\mu, b_\rho, c) = X^2(\mu, b_\rho, c)A^T \left( AX^2(\mu, b_\rho, c)A^T \right)^{-1} \delta.$$

Hence, corollary 3.11 and the assumption that  $A$  has full row rank establish the first bound. Similarly, 3.21, theorem 3.8, and corollary 3.11 provide the remaining two bounds. ■

In contrast to theorem 3.12, which provides uniform bounds on the directional derivatives, the forthcoming material shows that  $D_{\mu^+}^1 x(\mu, b, c)$  is uniformly bounded over  $\{\omega \in \mathbb{R}_{++}^n : \|\omega\| = 1\}$ . Recalling from theorem 2.2 that any element in the relative interior of a polytope is an omega analytic center, we have that such a result essentially shows that the bounds provided by theorem 3.12 are independent of a specific omega central path. The result relies on the assumption that the dual is not degenerate. This is because dual non-degeneracy is shown in the next lemma to guarantee that  $A_B$  has full row

rank, which is needed to invoke corollary 3.11. In the following lemma, the dimension of a set in  $\mathbb{R}^n$  is defined to be the dimension of the smallest affine subspace containing the set.

**Lemma 3.13 (Roos, Terlaky, and Vial [73])** Let  $r \in \mathcal{G}$ . Then,

$$\dim(\mathcal{P}_r^*) = |B| - \text{rank}(A_B) \text{ and}$$

$$\dim(\mathcal{D}_r^*) = m - \text{rank}(A_B).$$

**Proof:** Because  $\mathcal{P}_r^* = \{x : A_B x_B = b, x_B \geq 0, x_N = 0\}$ , the smallest affine subspace containing  $\mathcal{P}_r^*$  is  $\{x : A_B x_B = b, x_N = 0\}$ . The dimension of this affine space is equal to the dimension of the  $\text{null}(A_B)$  which is  $|B| - \text{rank}(A_B)$ .

Similarly,  $\mathcal{D}_r^* = \{(y, s) : yA + s = c, s_B = 0, s_N \geq 0\}$  and the smallest affine subspace containing  $\mathcal{D}_r^*$  is  $\{(y, s) : yA_B = c_B, yA_N + s_N = c_N, s_B = 0\}$ . Since  $s_N$  is uniquely determined by any  $y$  that solves  $yA_B = c_B$ , this affine space has the same dimension as the  $\text{leftnull}(A_B)$ . The proof is complete because the  $\dim(\text{leftnull}(A_B)) = m - \text{rank}(A_B)$ . ■

Lemma 3.13 shows that the dual solution is unique if, and only if,  $A_B$  has full row rank. An explanation of the required dual non-degeneracy follows. Recall that lemma 3.1 showed that

$$D_\mu^1 x_B(\mu, b, c) = -\Omega_B^{-\frac{1}{2}} X_B(\mu, b, c) \left( A_B \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c) \right)^+ A_N \cdot D_\mu^1 x_N(\mu, b, c).$$

The first problem that a dual non-degeneracy assumption solves is the convergence of  $\left(A_B \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c)\right)^+$  as  $\mu \rightarrow 0^+$ . The problem here is, that in general, a sequence of matrices  $\{A^k\}$ , which converges to  $A$ , does not guarantee  $\left\{\left(A^k\right)^+\right\} \rightarrow A^+$ . See [8] for an example of such a sequence. However, if  $A_B$  has full row rank,

$$\left(A_B \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c)\right)^+ = A_B \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c) \left(A_B \Omega_B^{-1} X_B^2(\mu, b, c) A_B^T\right)^{-1}.$$

Since invertible matrices have the desired convergence property,

$$\lim_{\mu \rightarrow 0^+} \left(A_B \Omega_B^{-\frac{1}{2}} X_B(\mu, b, c)\right)^+ = \left(A_B \Omega_B^{-\frac{1}{2}} X_B^*(r)\right)^+.$$

The dual non-degeneracy assumption overcomes a second problem. Because dual non-degeneracy implies that  $A_B$  has full row rank, this assumption allows the use of corollary 3.11. From 3.16,  $D_{\mu^+}^1 x_N^*(r) = (S_N^*(r))^{-1} \omega_N$ , and

$$D_{\mu^+}^1 x_B^*(r) = -\Omega_B^{-\frac{1}{2}} X_B^*(r) \left(A_B \Omega_B^{-\frac{1}{2}} X_B^*(r)\right)^+ A_N \left((S_N^*(r))^{-1} \omega_N\right). \quad (3.24)$$

Corollary 3.11 guarantees

$$\sup \left\{ \left\| \Omega_B^{-\frac{1}{2}} X_B^*(r) \left(A_B \Omega_B^{-\frac{1}{2}} X_B^*(r)\right)^+ \right\| : \omega \in \mathbb{R}_{++}^n, \|\omega\| = 1 \right\} \leq \|A_B^+\| \chi_{A_B^T}.$$

So, if  $(S_N^*(r))^{-1} \omega_N$  is uniformly bounded over  $\{\omega \in \mathbb{R}_{++}^n : \|\omega\| = 1\}$ ,  $D_{\mu^+}^1 x^*(r)$  is uniformly bounded over  $\{\omega \in \mathbb{R}_{++}^n : \|\omega\| = 1\}$ . The next lemma provides such a result, and the statement of the uniform bounds follows.

**Lemma 3.14** As in lemmas 2.16 and 2.17 let

$$\mathcal{Q} = \{u : Qu = b, u \geq 0\},$$

where here we assume that  $\mathcal{Q}$  is bounded and has a non-empty strict interior.

Let  $u(\omega)$  be the omega analytic center of  $\mathcal{Q}$ . Then,

$$\sup\{\|\omega^T U^{-1}(\omega)\| : \omega \in \mathbb{R}_{++}^n, \|\omega\| = 1\} < \infty.$$

**Proof:** Since  $\|\omega\| = 1$ ,  $e^T \omega \leq n$ . The necessary and sufficient conditions describing  $u(\omega)$  are the existence of  $(v(\omega), w(\omega))$  such that

$$Qu(\omega) = b$$

$$v(\omega)Q + w(\omega) = 0$$

$$U(\omega)w(\omega) = \omega.$$

As in the proof of theorem 2.4, let  $\hat{u} \in \mathcal{Q}^\circ$  and  $(\hat{v}, \hat{w})$  be such that  $\hat{v}Q + \hat{w} = 0$  and  $\hat{w} > 0$ . Then,

$$0 \leq w_i(\omega) \leq \frac{\hat{w}\hat{u} + e^T \omega}{\hat{u}_i} \leq \frac{\hat{w}\hat{u} + n}{\hat{u}_i}.$$

Since  $w_i(\omega) = \frac{\omega_i}{u_i(\omega)}$  and the last bound given is independent of  $\omega$ , the result is proven. ■

**Theorem 3.15** Let  $r \in \mathcal{G}$ . Then, if the dual solution is unique,

$$\sup\{\|D_{\mu^+}^1 x^*(r)\| : \omega \in \mathbb{R}_{++}^n, \|\omega\| = 1\} < \infty.$$

**Proof:** From equation 3.24,

$$D_{\mu^+}^1 x_B^*(r) = -\Omega_B^{-\frac{1}{2}} X_B^*(r) \left( A_B \Omega_B^{-\frac{1}{2}} X_B^*(r) \right)^+ A_N \left( (S_N^*(r))^{-1} \omega_N \right).$$

Hence,

$$\begin{aligned} \|D_{\mu^+}^1 x^*(r)\| &\leq \|D_{\mu^+}^1 x_B^*(r)\| + \|D_{\mu^+}^1 x_N^*(r)\| \\ &\leq \left\| \Omega_B^{-\frac{1}{2}} X_B^*(r) \left( A_B \Omega_B^{-\frac{1}{2}} X_B^*(r) \right)^+ \right\| \|A_N\| \left\| (S_N^*(r))^{-1} \omega_N \right\| \\ &\quad + \left\| (S_N^*(r))^{-1} \omega_N \right\|. \end{aligned}$$

The assumption of a unique dual solution, corollary 3.11, and lemma 3.14 provide the result. ■

### 3.6 Chapter Summary

The results of this chapter establish that the analytic center solution has the following properties:

- $x^*(r)$  is a continuous function of  $b$ ,
- $x^*(r)$  is locally analytic along any  $\mathfrak{d}b \in \mathcal{H}_b$ , and
- the one sided derivatives of  $x^*(r)$ , along any  $\mathfrak{d}b$ , always exist.

Furthermore, theorems 3.12 and 3.15 show that, under a dual non-degeneracy assumption, the first order derivatives are uniformly bounded, independent of which omega central path is of concern. Whether or not this result is true without the non-degeneracy assumption remains an open question.

## 4. The Omega Central Path and Simultaneous Rim Data Perturbations

### 4.1 Introduction

This chapter establishes a set convergence property for the closure of  $PCP_r$ , denoted  $\overline{PCP}_r$  and defined as  $PCP_r \cup \{x^*(r)\}$  if  $\mathcal{P}_b$  is unbounded and  $PCP_r \cup \{x^*(r), \bar{x}(b)\}$  if  $\mathcal{P}_b$  is bounded. For any  $r \in \mathcal{G}$  and  $\delta c$ , define  $PCP_{(r, \delta c)}^* \equiv \{z_B(\eta, b, \delta c) : \eta > 0\}$  to be the omega central path corresponding to

$$\min\{\delta c z_B : A_B z_B = b, z_B \geq 0\}, \quad (4.1)$$

where  $(B|N)$  is the optimal partition for the data instance  $r$ . For any  $\eta \in \mathbb{R}$ ,  $z_B(\eta, b, \delta c)$  is referred to as the  $\eta$  point of  $PCP_{(r, \delta c)}^*$ . Because

$$\{z_B(\eta, b, \delta c) : \eta > 0\}$$

is an omega central path, it has all the properties established in chapter 2. For example,  $z_B(\eta, b, \delta c)$  is an analytic function on  $\mathbb{R}_{++} \times \mathcal{G}_b \times \mathbb{R}^n$ . The omega analytic center solution for 4.1 is denoted by

$$z_B^*(b, \delta c) = \lim_{\eta \rightarrow 0^+} z_B(\eta, b, \delta c). \quad (4.2)$$

Recognizing that  $\{(z_B, 0) : A_B z_B = b, z_B \geq 0\} = \mathcal{P}_r^*$ , corollary 2.5 implies that the feasible set for 4.1 is bounded. This together with theorem 2.9 implies,

$$\bar{z}_B(b, \delta c) \equiv \lim_{\eta \rightarrow \infty} z_B(\eta, b, \delta c) = x_B^*(r). \quad (4.3)$$

The closure of  $PCP_{(r, \delta c)}^*$  is defined similarly to  $\overline{PCP}_r$ ;

$$\overline{PCP}_{(r, \delta c)}^* \equiv PCP \cup \{\bar{z}_B(b, \delta c), z_B^*(b, \delta c)\}.$$

Specifically, the main result of this chapter is that

$$\lim_{(b, \bar{c}_\tau) \rightarrow (\bar{b}, \bar{c})} \overline{PCP}_{(b, \bar{c})} = \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right),$$

where  $\mathbf{0}^{|N|}$  is the set containing the zero vector of length  $|N|$ . This result follows from a parametric analysis of the elements in  $PCP_r$ , with respect to  $\tau$  along a fixed  $\delta c$ . Surprisingly, the parametric analysis presented for  $\tau$  allows general changes in the right-hand side vector (notice that  $b \rightarrow \bar{b}$  in the above limit with no restrictions on  $\{b^k\}$ ). This means that simultaneous changes in the cost coefficients and right-hand side vectors are allowed, so long as the cost coefficient change is linear. As seen in theorems 4.5 and 4.14, extensions to theorem 2.8 are possible with such a parametric analysis.

## 4.2 Convergence Under Linear Cost Changes

The fact that  $x^*(r)$  is not a continuous function of  $c$  is demonstrated in example 2.12. Since discontinuities produce severe changes in  $x^*(r)$ , one

might believe that severe changes in the central path must occur. However, there is a continuity of central paths, when central paths are viewed as sets. To motivate this idea consider the following example.

**Example 4.1** Consider the linear program,

$$\min\{x_j : 0 \leq x_i \leq 1, x \in \mathbb{R}^n\}.$$

The optimal objective value is zero and the analytic center solution is  $\frac{1}{2}e - \frac{1}{2}e_j$ .

Furthermore, the optimal partition is

$$\begin{aligned} B &= \{1, 2, \dots, n\} \setminus \{j\} \text{ and} \\ N &= \{j\}. \end{aligned}$$

Let  $\delta c$  be such that  $\delta c_i \neq 0$  for some  $i \neq j$ . Using equation 2.13 in Example 2.12, the central path is defined by

$$x_i(\mu, e, c_\tau) = \begin{cases} \frac{\tau \delta c_i + 2\mu - \sqrt{\tau^2 \delta c_i^2 + 4\mu^2}}{2\tau \delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ \frac{1 + \tau \delta c_i + 2\mu - \sqrt{(1 + \tau \delta c_i)^2 + 4\mu^2}}{2(1 + \tau \delta c_i)} & \text{if } i = j \end{cases} .$$

The question posed is, ‘‘How does  $x_i(\mu, e, c_\tau)$  behave as  $\tau \rightarrow 0^+$ ?’’ Consider the three cases,

- (1)  $\mu = \eta\tau$ , for some positive real value  $\eta$ ,
- (2)  $\mu = \tau^2$ , and

$$(3) \mu = \sqrt{\tau}.$$

**Case 1** Let  $\eta$  be any positive real value. Then making the substitution  $\mu = \eta\tau$ ,

$$x_i(\eta\tau, e, \tau) = \begin{cases} \frac{\tau\delta c_i + 2\eta\tau - \sqrt{\tau^2\delta c_i^2 + 4(\eta\tau)^2}}{2\tau\delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ \frac{1 + \tau\delta c_i + 2\eta\tau - \sqrt{(1 + \tau\delta c_i)^2 + 4(\eta\tau)^2}}{2(1 + \tau\delta c_i)} & \text{if } i = j \end{cases}$$

$$= \begin{cases} \frac{\delta c_i + 2\eta - \sqrt{\delta c_i^2 + 4\eta^2}}{2\delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ \frac{1 + \tau\delta c_i + 2\eta\tau - \sqrt{(1 + \tau\delta c_i)^2 + 4(\eta\tau)^2}}{2(1 + \tau\delta c_i)} & \text{if } i = j. \end{cases}$$

Since,

$$\lim_{\tau \rightarrow 0^+} \frac{1 + \tau\delta c_i + 2\eta\tau - \sqrt{(1 + \tau\delta c_i)^2 + 4(\eta\tau)^2}}{2(1 + \tau\delta c_i)} = 0,$$

we have

$$\hat{x}_i = \lim_{\tau \rightarrow 0^+} x_i(\eta\tau, e, c_\tau) = \begin{cases} \frac{\delta c_i + 2\eta - \sqrt{\delta c_i^2 + 4\eta^2}}{2\delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ 0 & \text{if } i = j. \end{cases}$$

Notice that  $\sigma(\hat{x}) = B$ , and hence  $\hat{x} \in \mathcal{P}^*$ . Upon examining equations describing the central path for

$$\min\{\delta c_B x_B - \eta \sum_{i \in B} \ln(x_i) : A_B x_B = b, x_B > 0\},$$

$\hat{x}_B$  is found to be the  $\eta$  point of  $PCP_{(r,\delta c)}^*$ . Hence, for any  $\{\tau^k\} \rightarrow 0$  and  $x \in PCP_{(r,\delta c)}^* \times \mathbf{0}^{|N|}$ , there exists  $\{x^k \in PCP_{(b,c_{\tau^k})}\} \rightarrow x$ .

**Case 2** When  $\mu = \tau^2$ ,

$$x_i(\tau^2, e, \tau) = \begin{cases} \frac{\delta c_i + 2\tau - \sqrt{\delta c_i^2 + 4\tau^2}}{2\delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ \frac{1 + \tau\delta c_i + 2\tau^2 - \sqrt{(1 + \tau\delta c_i)^2 + 4\tau^4}}{2(1 + \tau\delta c_i)} & \text{if } i = j. \end{cases}$$

Allowing  $\tau \rightarrow 0^+$  produces,

$$\hat{x}_i = \lim_{\tau \rightarrow 0^+} x_i(\tau^2, e, \tau) = \begin{cases} 1 & \text{if } i \neq j \text{ and } \delta c_i < 0 \\ 0 & \text{if } i \neq j \text{ and } \delta c_i > 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ 0 & \text{if } i = j. \end{cases}$$

Hence, in this case  $\hat{x}$  is the analytic center solution to

$$\min\{c_\tau x : Ax = b, x \geq 0\},$$

where  $\tau > 0$  is sufficiently small. Furthermore,  $\hat{x}_B$  is the analytic center solution to

$$\min\{\delta c_B x_B : A_B x_B = b, x_B \geq 0\}.$$

**Case 3** If  $\mu = \sqrt{\tau}$ ,

$$x_i(\sqrt{\tau}, e, \tau) = \begin{cases} \frac{\tau\delta c_i + 2\sqrt{\tau} - \sqrt{\tau^2\delta c_i^2 + 4\tau}}{2\tau\delta c_i} & \text{if } i \neq j \text{ and } \delta c_i \neq 0 \\ \frac{1}{2} & \text{if } i \neq j \text{ and } \delta c_i = 0 \\ \frac{1 + \tau\delta c_i + 2\sqrt{\tau} - \sqrt{(1 + \tau\delta c_i)^2 + 4\tau}}{2(1 + \tau\delta c_i)} & \text{if } i = j. \end{cases}$$

When  $\tau \rightarrow 0^+$ ,

$$\hat{x} = \lim_{\tau \rightarrow 0^+} x_i(\sqrt{\tau}, e, \tau) = \begin{cases} \frac{1}{2} & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases}$$

which is precisely the analytic center solution of

$$\min\{cx : Ax = b, x \geq 0\}.$$

The idea here is that the perturbed central paths are converging not only to the central path corresponding to the linear program with no data perturbation, but also to the central path corresponding to the minimization of the perturbation direction over the optimal face. Figure 4.1 shows the central paths for

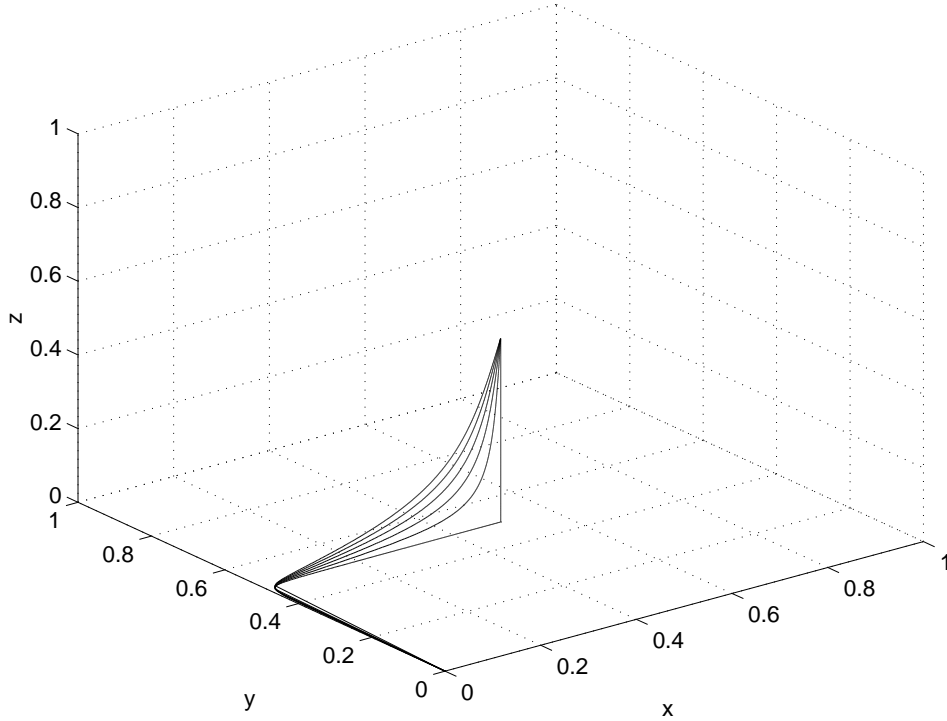
$$\min\left\{\frac{\tau}{4}x + \frac{\tau}{2000}y + z : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\right\},$$

where  $\tau = 1, 0.8, 0.6, 0.4, 0.2$ . The vertical line is the central path for

$$\min\{z : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\},$$

and the path contained in the  $x$ - $y$  plane is the central path for

$$\min\left\{\frac{1}{4}x + \frac{1}{2000}y : 0 \leq x \leq 1, 0 \leq y \leq 1\right\}.$$



**Figure 4.1.** The geometry of central path convergence.

The first lemma of this section shows that  $c_B x_B$  is constant on the null space of  $A_B$ . This fact is almost obvious once the optimal set is recognized as an affine transformation of the null space  $A_B$ .

**Lemma 4.2**  $c_B u = 0$  for all  $u \in \text{null}(A_B)$ .

**Proof:** If the columns of  $A_B$  are linearly independent, zero is the only element in  $\text{null}(A_B)$  and the result is trivial. Otherwise, let  $u_B \in \text{null}(A_B)$  and  $(u_B^*, 0) \in (\mathcal{P}_r^*)^o$ . Then there exists  $\alpha > 0$  such that  $(u_B^*, 0) + \alpha(u, 0) \in (\mathcal{P}_r^*)^o$ .

Since  $c_B(u_B^* + \alpha u_B) = c_B u_B^*, 0 = \alpha c_B u_B = c_B u_B$ . ■

Although the primary interest in what follows is the behavior of the omega central path under linear parametric changes in  $c$ , the addition of allowing general changes in  $b$  presents no difficulty. Historically, the properties of a solution under simultaneous changes have received less attention than the properties of a solution when only one rim data element is perturbed. A recent exception to this is found in [28], where it was shown, using the analytic center solution, that the objective value is a quadratic function of  $r$ . The ability of allowing simultaneous changes in  $b$  and  $c$  is a strength of the forthcoming analysis.

The next lemma demonstrates how the constant-cost-slices corresponding to each  $\mu > 0$  are modeled through right-hand side changes. A consequence of this modeling is that  $x_B(\mu, b, c)$ , for any fixed  $\mu > 0$ , is describable without the linear portion of the objective function. This follows because the omega central path's reliance on the objective function is implicitly expressed through the optimal partition. It is precisely this modeling scheme that allows for the simultaneous changes in the rim data.

**Lemma 4.3**  $c_B x_B$  is constant on

$$\{x_B : A_B x_B = b - A_N x_N(\mu, b, c_\tau) \text{ , } x_B \geq 0\},$$

and consequently  $x_B(\mu, b, c_\tau)$  is the unique solution to

$$\min\{\tau\delta c_B x_B - \mu \sum_{i \in B} \omega_i \ln(x_i) : A_B x_B = b - A_N x_N(\mu, b, c_\tau) , x_B > 0\}. \quad (4.4)$$

**Proof:** By definition,  $x(\mu, b, c_\tau)$  is the unique solution to

$$\min\{cx - \mu \sum_{i=1}^n \omega_i \ln(x_i) + \tau\delta cx : x \in \mathcal{P}_b\}.$$

Holding the components of  $x_N(\mu, b, c_\tau)$  constant,  $x_B(\mu, b, c_\tau)$  is the unique solution of

$$\min\{c_B x_B - \mu \sum_{i \in B} \omega_i \ln(x_i) + \tau\delta c_B x_B : A_B x_B = b - A_N x_N(\mu, b, c_\tau) , x_B > 0\}.$$

So, once it is shown that  $c_B x_B$  is constant on

$$\{x_B : A_B x_B = b - A_N x_N(\mu, b, c_\tau) , x_B \geq 0\},$$

the result follows. If the columns of  $A_B$  are linearly independent, the result is immediate because this set contains a single element. Otherwise, let  $x_B^1$  and  $x_B^2$  be in

$$\{x_B : A_B x_B = b - A_N x_N(\mu, b, c_\tau) , x_B \geq 0\}.$$

Then  $x_B^1 - x_B^2 \in \text{null}(A_B)$ , and lemma 4.2 implies

$$c_B(x_B^1 - \alpha(x_B^2 - x_B^1)) = c_B x_B^1,$$

for all  $0 \leq \alpha \leq 1$ , which proves the desired result. ■

Lemma 4.3 allows the following useful convergence result, which states that if  $\mu$  converges to zero, the right-hand side vector converges, and the cost vector is held fixed, then  $x(\mu, b, c)$  converges to the analytic center solution. The proof is similar to the proof of theorem 2.18.

**Lemma 4.4** Let  $\bar{r} \in \mathcal{G}$ ,  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ , and  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$ . Then,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, \bar{c}) = x^*(\bar{r}).$$

**Proof:** Let  $\{b^k\}$  and  $\{\mu^k\}$  be as above. From theorem 2.14,  $\{x(\mu^k, b^k, \bar{c})\}$  is bounded. Let  $\{x(\mu^{k_i}, b^{k_i}, \bar{c})\}$  be a subsequence such that

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, \bar{c}) = \hat{x}.$$

Establishing that  $\hat{x}$  is  $x^*(\bar{r})$  proves the result by of the uniqueness of  $x^*(\bar{r})$ .

Lemma 4.3 shows that  $x_B(\mu^{k_i}, b^{k_i}, \bar{c})$  solves

$$\min\left\{-\sum_{j \in B} \ln(x_j) : A_B x_B = b^{k_i} - A_N x_N(\mu^{k_i}, b^{k_i}, \bar{c}), x_B > 0\right\}.$$

The necessary and sufficient conditions for this math program are that, for any  $i$ , there exists  $y^i$  such that,

$$A_B x_B(\mu^{k_i}, b^{k_i}, \bar{c}) = b^{k_i} - A_N x_N(\mu^{k_i}, b^{k_i}, \bar{c})$$

$$y^i A_B = e^T X_B^{-1}(\mu^{k_i}, b^{k_i}, \bar{c})$$

$$x_B(\mu^{k_i}, b^{k_i}, \bar{c}) > 0.$$

Let

$$y^i = e^T X_B^{-1}(\mu^{k_i}, b^{k_i}, \bar{c}) A_B^+,$$

so that

$$y^i A_B = e^T X_B^{-1}(\mu^{k_i}, b^{k_i}, \bar{c}) A_B^+ A_B = e^T X_B^{-1}(\mu^{k_i}, b^{k_i}, \bar{c}).$$

If  $\hat{x} > 0$ ,  $\{y^i\}$  converges, say to  $\hat{y}$ , and because lemma 2.15 shows

$$\{x_N(\mu^{k_i}, b^{k_i}, \bar{c})\} \rightarrow 0,$$

we then have,

$$\begin{aligned} A\hat{x} &= b \\ \hat{y}A_B &= e^T \hat{X}_B^{-1} \\ \hat{x} &> 0. \end{aligned}$$

Since these are the necessary and sufficient conditions describing  $x_B^*(\bar{r})$ , the result would follow. The fact that  $\hat{x} > 0$  is forthcoming.

Lemma 2.17 implies the existence of a sequence  $\{\tilde{x}_B^i\}$  such that,

$$A_B \tilde{x}_B^i = b^{k_i} - A_N x_N(\mu^{k_i}, b^{k_i}, \bar{c}),$$

$$\tilde{x}_B^i > 0, \text{ and}$$

$$\lim_{i \rightarrow \infty} \tilde{x}_B^i = \tilde{x} > 0.$$

The optimality of  $x_B(\mu^{k_i}, b^{k_i}, \bar{c})$  implies that

$$\sum_{j \in B} \ln(x_j^i(\mu^{k_i}, b^{k_i}, \bar{c})) \geq \sum_{j \in B} \ln(\tilde{x}_j^i).$$

As  $i \rightarrow \infty$  the right-hand side of the above inequality converges to

$$\sum_{j \in B} \ln(\tilde{x}_j).$$

Hence,  $\{x(\mu^{k_i}, b^{k_i}, \bar{c})\}$  is bounded away from zero, and  $\hat{x} > 0$ . ■

Lemma 4.3 leads to the following observation:

$$x_B(\mu, b, c_\tau) = z_B\left(\frac{\mu}{\tau}, b - A_N x_N(\mu, b, c_\tau), \delta c\right), \quad (4.5)$$

which is crucial in establishing subsequent results. In words, each element of the central path is also an element of a central path contained in a constant-cost-slice and defined by the perturbation direction. Figure 4.2 illustrates these two interpretations of central path elements. The vertical line is the central path for

$$\min\{z : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\},$$

the curve in the  $x, y$ -plane is the central path for

$$\min\{x + \frac{1}{10}y : 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

and the curve from  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  to  $(0, 0, 0)$  is the central path for

$$LP^1 : \min\{x + \frac{1}{10}y + z : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}.$$

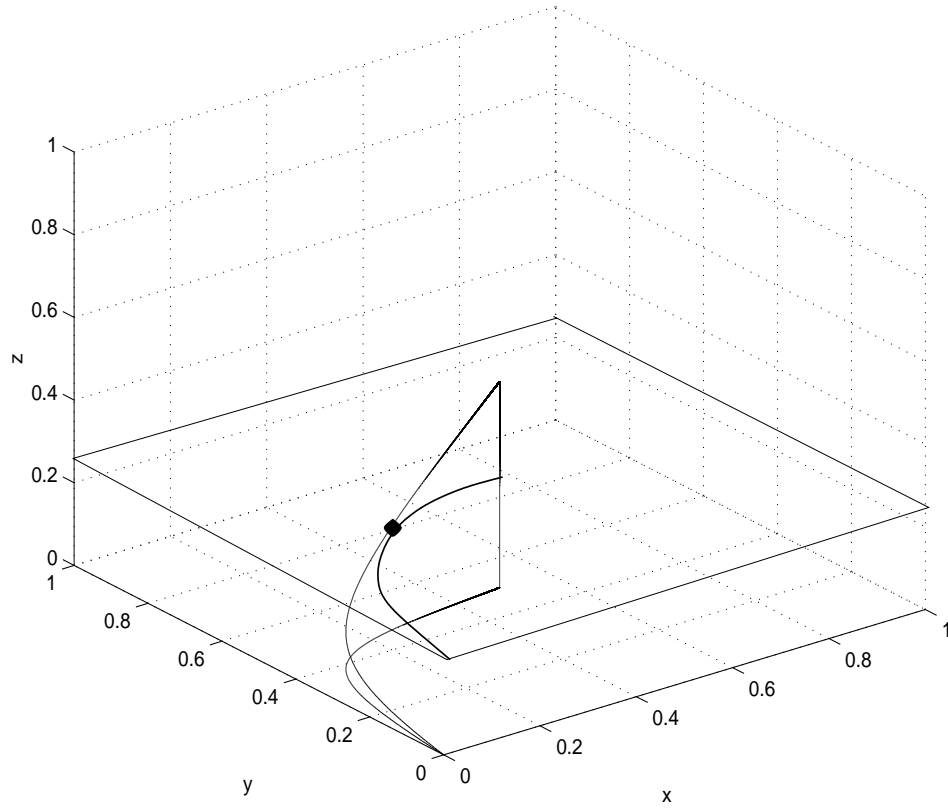
The constant-cost-slice is, for some  $\bar{\mu} > 0$ ,

$$\{(x, y, z) : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1, z = z(\bar{\mu}, b, c)\},$$

and the curve on this constant-cost-slice is the central path for

$$LP^2 : \min\{x + \frac{1}{10}y : 0 \leq x \leq 1, 0 \leq y \leq 1, z = z(\bar{\mu}, b, c)\}.$$

Equation 4.5 shows how the central paths for  $LP^1$  and  $LP^2$  intersect.



**Figure 4.2.** Two different interpretations of an analytic center.

From theorem 2.14,

$$\{z_B(\frac{\mu}{\tau}, b - A_N x_N(\mu, b, c_\tau), \delta c) : 0 \leq \frac{\mu}{\tau} \leq \frac{\bar{\mu}}{\tau}\}$$

is bounded, for any  $\bar{\mu} > 0$  and  $\bar{\tau} > 0$ , and hence has at least one cluster point. However, if  $\{\mu^k\}$  and  $\{\tau^k\}$  are positive sequences converging to 0, the sequence  $\{\frac{\mu^k}{\tau^k}\}$  does not have to have a limit point. Hence, guaranteeing the existence of

$$\lim_{(\mu, b, c_\tau) \rightarrow (0, \bar{b}, \bar{c})} x(\mu, b, c_\tau),$$

is, in general, not possible. This means extending theorem 2.8 to include shifting rim data is not straightforward. However, the next theorem does extend theorem 2.8 to include the case of general right-hand side changes and linear cost coefficient changes. The conditions stated for the extension are “nearly necessary and sufficient”, as shown in lemmas 4.6 and 4.7.

**Theorem 4.5** Let  $(\bar{\mu}, \bar{b}, \bar{c}) \in \mathbb{R}_+ \times \mathcal{G}$  and  $\delta c$  be fixed. Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow \bar{\mu}$ ,  $\{\tau^k \in \mathbb{R}_{++}\} \rightarrow 0$ , and  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b}$ . Then,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, \bar{c}_{\tau^k}) = \begin{cases} x(\bar{\mu}, \bar{b}, \bar{c}) & \text{if } \bar{\mu} > 0 \\ x^*(\bar{\tau}) & \text{if } \bar{\mu} = 0 \text{ and } \lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = \infty \\ (z_B(\eta, \bar{b}, \delta c), 0) & \text{if } \bar{\mu} = 0 \text{ and } \lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = \eta > 0 \\ (z_B^*(\bar{b}, \delta c), 0) & \text{if } \bar{\mu} = 0 \text{ and } \lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = 0. \end{cases}$$

**Proof:** The case when  $\bar{\mu} > 0$  is an immediate consequence of theorem 2.6.

So, assume  $\bar{\mu} = 0$  and consider the case when,

$$\lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = \bar{\eta} > 0.$$

Lemma 2.15 implies that

$$\lim_{k \rightarrow \infty} x_N(\mu^k, b^k, \bar{c}_{\tau^k}) = 0.$$

Since,  $\{\frac{\mu^k}{\tau^k}\}$  is bounded away from zero, theorem 2.6 implies that

$$\lim_{k \rightarrow \infty} z_B \left( \frac{\mu^k}{\tau^k}, b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c \right) = z_B(\bar{\eta}, \bar{b}, \delta c).$$

Using the fact that  $x_B(\mu^k, b^k, \bar{c}_{\tau^k}) = z_B \left( \frac{\mu^k}{\tau^k}, b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c \right)$ ,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, c^k) = (z_B(\bar{\eta}, \bar{b}, \delta c), 0).$$

The next situation considered is when,

$$\lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = 0.$$

Since lemma 2.15 implies the sequence

$$\{b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k})\} \rightarrow \bar{b},$$

lemma 4.4 implies that

$$\lim_{k \rightarrow \infty} x_B(\mu^k, b^k, \bar{c}_{\tau^k}) = \lim_{k \rightarrow \infty} z_B \left( \frac{\mu^k}{\tau^k}, b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c \right) = z_B^*(\bar{r}, \delta c),$$

which establishes the result in this case.

Lastly, assume that

$$\lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = \infty.$$

Once again, lemma 2.15 implies the sequence

$$\{b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k})\} \rightarrow \bar{b}$$

Furthermore,

$$\lim_{k \rightarrow \infty} \frac{\delta c}{\frac{\mu^k}{\tau^k}} = \lim_{k \rightarrow \infty} \frac{\tau^k \delta c}{\mu^k} = 0 \in \text{row}(A_B).$$

Since corollary 2.5 shows that  $\mathcal{P}_{\bar{r}}^*$  is bounded, the conditions to apply theorem 2.18 are satisfied, which states that

$$\begin{aligned} \lim_{k \rightarrow \infty} x_B(\mu^k, b^k, \bar{c}_{\tau^k}) &= \lim_{k \rightarrow \infty} z_B\left(\frac{\mu^k}{\tau^k}, b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c\right) \\ &= \bar{z}_B(b, \delta c) \\ &= x_B^*(\bar{r}). \end{aligned}$$

■

Together with definition ??, this last result is used to show

$$\lim_{(b, \bar{c}_\tau) \rightarrow (\bar{b}, \bar{c})} \overline{PCP}_{(b, \bar{c}_\tau)} = \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right).$$

This means that  $\{x^k \in PCP_{(b^k, c_{\tau^k})}\} \rightarrow x$  must imply

$$x \in \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right).$$

If  $x^k \in PCP_{(b^k, c_{\tau^k})}$ , there exists  $\mu^k$  such that  $x^k = x(\mu^k, b^k, c_{\tau^k})$ . Notice that characterizing the conditions of  $\{(\mu^k, b^k, c_{\tau^k})\}$ , for which  $\{x(\mu^k, b^k, c_{\tau^k})\}$  converges would be of great use. Theorem 4.5 is used to develop such necessary

and sufficient conditions; almost yielding enough information to state these conditions directly. The problem is that if  $\bar{\mu} = 0$ , not enough is shown about the last three limiting values stated in theorem 4.5. The next two lemmas add this needed information and theorem 4.8 states the necessary and sufficient conditions.

**Lemma 4.6** Let  $\bar{r} \in \mathcal{G}$ ,  $\delta c \notin \mathcal{H}_c$ , and  $\{b^k \in \mathcal{G}_B\} \rightarrow \bar{b}$ . Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$  and  $\{\tau^k \in \mathbb{R}_{++}\} \rightarrow 0$  be such that the sequence  $\left\{\frac{\mu^k}{\tau^k}\right\}$  does not converge. Furthermore, let  $\left\{\frac{\mu^{k_i}}{\tau^{k_i}}\right\}$  and  $\left\{\frac{\mu^{k_j}}{\tau^{k_j}}\right\}$  be two convergent subsequences, one of which possibly converges to infinity. Then,

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} \neq \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\tau^{k_j}},$$

implies

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) \neq \lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, c_{\tau^{k_j}}).$$

**Proof:** Without loss in generality, assume

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} < \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\tau^{k_j}}.$$

Theorem 4.5 implies that

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) = \begin{cases} (z_B^*(\bar{r}, \delta c), 0) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} = 0 \\ (z_B(\eta^1, \bar{b}, \delta c), 0) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} = \eta^1 > 0 \end{cases}$$

and

$$\lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, c_{\tau^{k_j}}) = \begin{cases} (z_b(\eta^2, \bar{b}, \delta c), 0) & \text{if } \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\tau^{k_j}} = \eta^2 < \infty \\ x^*(\bar{r}) & \text{if } \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\tau^{k_j}} = \infty. \end{cases}$$

Since  $\delta c \notin \mathcal{H}_c$ , lemma 2.13 implies that  $\delta c_B \notin \text{col}(A_B)$ . Theorem 2.11 now provides that for any  $\eta^1 < \eta^2$ ,

$$\delta c_B z_B^*(\bar{b}, \delta c) < \delta c_B z_B(\eta^1, \bar{b}, \delta c) < \delta c_B z_B(\eta^2, \bar{b}, \delta c) < \delta c_B x_B^*(\bar{r}),$$

and the result follows. ■

**Lemma 4.7** If  $\bar{r} \in \mathcal{G}$  and  $\delta c \in \mathcal{H}_c$ ,

$$\bar{z}_B(\bar{b}, \delta c) = z_B(\eta, \bar{b}, \delta c) = z_B^*(\bar{b}, \delta c),$$

for all  $\eta \in \mathbb{R}_{++}$ .

**Proof:** Lemma 2.13 implies that  $\delta c_B \in \text{row}(A_B)$  and theorem 2.10 implies that  $z_B(\eta^1, \bar{b}, \delta c) = z_B(\eta^2, \bar{b}, \delta c)$ , for all  $\eta^1$  and  $\eta^2$  in  $\mathbb{R}_{++}$ . Equalities 4.2 and 4.3 imply the result. ■

Theorem 4.8 contains necessary and sufficient conditions to assure the convergence of  $\{x(\mu^k, b^k, c_{\tau^k})\}$ .

**Theorem 4.8** Let  $\bar{r} \in \mathcal{G}$ . Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$ ,  $\{\tau^k \in \mathbb{R}_{++}\} \rightarrow 0$ , and  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ . If  $\delta c \in \mathcal{H}_c$ ,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, \bar{c}_{\tau^k}) = x^*(\bar{r}).$$

If  $\delta c \notin \mathcal{H}_c$ ,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, \bar{c}_{\tau^k})$$

exists if, and only if,  $\left\{\frac{\mu^k}{\tau^k}\right\}$  converges, possibly to infinity.

**Proof:** From lemma 4.7, if  $\delta c \in \mathcal{H}_c$ ,

$$\begin{aligned} x_B(\mu^k, b^k, \bar{c}_{\tau^k}) &= z_B\left(\frac{\mu^k}{\tau^k}, b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c\right) \\ &= z_B^*(b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c). \end{aligned}$$

Since lemma 3.6 shows that  $z_B^*$  is a continuous function of its right-hand side and lemma 2.15 shows that  $\lim_{k \rightarrow 0} x_N(\mu^k, b^k, \bar{c}_{\tau^k}) = 0$ ,

$$\begin{aligned} \lim_{x \rightarrow \infty} x(\mu^k, b^k, \bar{c}_{\tau^k}) &= \lim_{x \rightarrow \infty} \left( z_B^*(b^k - A_N x_N(\mu^k, b^k, \bar{c}_{\tau^k}), \delta c), x_N(\mu^k, b^k, \bar{c}_{\tau^k}) \right) \\ &= \left( z_B^*(\bar{b}, \delta c), 0 \right) \\ &= x^*(\bar{r}). \end{aligned}$$

Consider the case when  $\delta c \notin \mathcal{H}_c$ . If  $\left\{\frac{\mu^k}{\tau^k}\right\}$  converges, even to infinity, theorem 4.5 demonstrates that the limit exists. If  $\left\{\frac{\mu^k}{\tau^k}\right\}$  does not converge, this sequence has at least two cluster points, one of which may be infinity. Hence,

there exist at least two convergent subsequences of  $\{\frac{\mu^k}{\tau^k}\}$ , say  $\{\frac{\mu^{k_i}}{\tau^{k_i}}\}$  and  $\{\frac{\mu^{k_j}}{\tau^{k_j}}\}$ , one of which may converge to infinity, such that

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} \neq \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\tau^{k_j}}.$$

Theorem 4.5 implies that both

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}})$$

and

$$\lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, \bar{c}_{\tau^{k_j}})$$

exist, and lemma 4.6 implies that these limits are different. So,  $\{x(\mu^k, b^k, c_{\tau^k})\}$  does not converge and the result follows. ■

The following set convergence property is the main result of this chapter, and it shows exactly how the primal omega central path behaves as a set under simultaneous changes in the right-hand side vector and cost coefficient vector, provided that the cost coefficient change is linear.

**Theorem 4.9** If  $\bar{r} \in \mathcal{G}$  and  $\delta c$  is fixed,

$$\lim_{(b, \bar{c}_\tau) \rightarrow (\bar{b}, \bar{c})} \overline{PCP}_{(b, \bar{c}_\tau)} = \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right).$$

**Proof:** Let  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ , and  $\{\tau^k \in \mathbb{R}_{++}\} \rightarrow 0$ . The first part of the proof establishes

$$\lim_{(b, \bar{c}_\tau) \rightarrow (\bar{b}, \bar{c})} PCP_{(b, \bar{c}_\tau)} = \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right).$$

Suppose the sequence  $\{x^k \in PCP_{(b^k, \bar{c}_{\tau^k})}\} \rightarrow \hat{x}$ . For each  $k$ , let  $\mu^k$  be such that  $x^k = x(\mu^k, b^k, \bar{c}_{\tau^k})$ . Consider a subsequence of  $\{\mu^k\}$ , say  $\{\mu^{k_i}\}$ , that converges, possibly to infinity. To show that  $\hat{x} \in \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right)$ , three cases are considered.

**Case 1** If  $\{\mu^{k_i}\} \rightarrow \bar{\mu} > 0$ , theorem 2.6 implies

$$\hat{x} = \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) = x(\bar{\mu}, \bar{b}, \bar{c}).$$

Since  $x(\bar{\mu}, \bar{b}, \bar{c}) \in \overline{PCP}_{\bar{r}}$ , this situation is complete.

**Case 2** Suppose  $\{\mu^{k_i}\} \rightarrow 0$ . Theorem 4.8 shows that when  $\delta c \in \mathcal{H}_c$ ,

$$\hat{x} = \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) = x^*(\bar{r}) \in \overline{PCP}_{\bar{r}}.$$

Furthermore, theorem 4.8 shows that when  $\delta c \notin \mathcal{H}_c$ ,  $\left\{ \frac{\mu^{k_i}}{\tau^{k_i}} \right\}$  must converge, possibly to infinity. From theorem 4.5,

$$\hat{x} = \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}}) = \begin{cases} x^*(\bar{r}) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} = \infty \\ (z_B(\eta, \bar{b}, \delta c), 0) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} = \eta > 0 \\ (z_B^*(\bar{b}, \delta c), 0) & \text{if } \lim_{k \rightarrow \infty} \frac{\mu^{k_i}}{\tau^{k_i}} = 0. \end{cases}$$

Since  $x^*(\bar{r}) \in \overline{PCP}_{\bar{r}}$ , and both  $z_B(\eta, \bar{b}, \delta c)$  and  $z_B^*(\bar{b}, \delta c)$  are in  $\overline{PCP}_{(\bar{r}, \delta c)}^*$ , this case is validated.

**Case 3** Assume  $\{\mu^{k_i}\} \rightarrow \infty$ . This case is completed once theorem 2.18 is shown to apply, which in turn provides that

$$\hat{x} = \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) = \bar{x}(\bar{b}).$$

Since  $\{b^{k_i}\} \rightarrow \bar{b}$  and  $\{\frac{\bar{c}_{\tau^{k_i}}}{\mu^{k_i}}\} \rightarrow 0 \in \text{row}(A)$ , all that is left to establish is that  $\mathcal{P}_{\bar{b}}$  is bounded. Assume for the sake of attaining a contradiction that  $\mathcal{P}_{\bar{b}}$  is unbounded. Then there exists a sequence  $\{x^i \in \mathcal{P}_{b^{k_i}}\}$  such that  $\|x^i\| \rightarrow \infty$  and  $\lim_{i \rightarrow \infty} \frac{x_j^i}{\mu^{k_i}} = 0$ ,  $j = 1, 2, \dots, n$ . The optimality of  $x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}})$  implies

$$\bar{c}_{\tau^{k_i}} x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}}) - \mu^{k_i} \sum_{j=1}^n \omega_j \ln(x_j(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}})) \leq \bar{c}_{\tau^{k_i}} x^i - \sum_{j=1}^n \omega_j \ln(x_j^i),$$

or equivalently

$$\frac{\bar{c}_{\tau^{k_i}}}{\mu^{k_i}} \left( x^i - x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}}) \right) + \sum_{j=1}^n \omega_j \ln(x_j(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}})) \geq \sum_{j=1}^n \omega_j \ln(x_j^i).$$

However, since  $\{x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}})\} \rightarrow \hat{x}$  and  $\lim_{i \rightarrow \infty} \frac{x_j^i}{\mu^{k_i}} = 0$ , the contradiction that the left-hand side of this inequality is bounded above implies that  $\mathcal{P}_{\bar{b}}$  is bounded.

The proof has so far established that if a sequence  $\{x^k \in PCP_{(b^k, \bar{c}_{\tau^k})}\}$  converges, the limit of this sequence is in  $\overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right)$ . The next part of the proof shows that any element in  $\overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right)$  is the limit point of a sequence  $\{x^k \in PCP_{(b^k, \bar{c}_{\tau^k})}\}$ .

Let  $x \in \overline{PCP}_{\bar{r}}$ . In the case where  $\mathcal{P}_{\bar{b}}$  is bounded and  $x = \bar{x}(\bar{b})$ , theorem 2.18 shows that  $\{x(k, b^k, \bar{c}_{\tau^k})\} \rightarrow x$ . If  $x = x(\mu, \bar{b}, \bar{c})$  or  $x = x^*(\bar{r})$ , theorem 4.5 implies

$$\begin{aligned} \left\{x\left(\mu + \frac{1}{k}, b^k, \bar{c}_{\tau^k}\right)\right\} &\rightarrow x(\mu, \bar{b}, \bar{c}) \text{ and} \\ \left\{x\left(\sqrt{\tau^k}, b^k, \bar{c}_{\tau^k}\right)\right\} &\rightarrow x^*(\bar{r}). \end{aligned}$$

Let  $x \in \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|}$ . If  $x$  is either  $(z_B(\eta, \bar{b}, \delta c), 0)$  or  $(z_B^*(b, \delta c), 0)$ , lemma 2.15 and theorem 4.5 imply

$$\begin{aligned} \left\{x(\eta\tau^k, b^k, \bar{c}_{\tau^k})\right\} &\rightarrow (z_B(\eta, \bar{b}, \delta c), 0) \text{ and} \\ \left\{x((\tau^k)^2, b^k, \bar{c}_{\tau^k})\right\} &\rightarrow (z_B^*(b, \delta c), 0). \end{aligned}$$

The possibility that  $x = x^*(\bar{r})$  is taken care of in the previous paragraph.

The proof has now established that

$$\lim_{(b, \bar{c}_\tau) \rightarrow (\bar{b}, \bar{c})} PCP_{(b, \bar{c}_\tau)} = \overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right).$$

What remains to be shown is that if a convergent sequence,  $\{x^k \in \overline{PCP}_{(b^k, \bar{c}_{\tau^k})}\}$ , contains infinitely many times either  $x^*(b^k, \bar{c}_{\tau^k})$  or, in the case when  $\mathcal{P}_{\bar{b}}$  is bounded,  $\bar{x}(b^k)$ , the limit of this sequence is in  $\overline{PCP}_{\bar{r}} \cup \left( \overline{PCP}_{(\bar{r}, \delta c)}^* \times \mathbf{0}^{|N|} \right)$ .

If  $\mathcal{P}_{\bar{b}}$  is bounded,  $\mathcal{P}_{b^k}$  is bounded for sufficiently large  $k$  by lemma 2.16. Furthermore, lemma 3.6 shows that  $\bar{x}(b)$  is a continuous function of  $b$  (since

setting  $c = 0$  implies  $x^*(r) = \bar{x}(b)$ . So, if  $\{x^k \in \overline{PCP}_{(b^k, \bar{c}_{\tau^k})}\} \rightarrow x$  and contains  $\bar{x}(b^k)$  infinitely many times,  $x = \bar{x}(\bar{b})$ .

Suppose that  $\{x^k \in \overline{PCP}_{(b^k, \bar{c}_{\tau^k})}\} \rightarrow x$  and that the sequence contains infinitely many elements of the form  $x^*(b^k, c_{\tau^k})$ . The first step in handling this situation is to show that  $\{x_N^k\} \rightarrow 0$ . Let  $\epsilon > 0$ . By definition,

$$x_N^*(b^k, c_{\tau^k}) = \lim_{\mu \rightarrow 0^+} x_N(\mu, b^k, c_{\tau^k}).$$

So, when  $x^k = x^*(b^k, c_{\tau^k})$ , there exists  $\bar{\mu}^k > 0$  such that  $\mu \in (0, \bar{\mu}^k)$  implies

$$\|x_N(\mu, b^k, \bar{c}_{\tau^k}) - x_N^*(b^k, \bar{c}_{\tau^k})\| < \frac{\epsilon}{2}.$$

Let  $\{\mu^k \in (0, \bar{\mu}^k)\} \rightarrow 0$ . Lemma 2.15 shows that  $\{x_N(\mu^k, b^k, c_{\tau^k})\} \rightarrow 0$ . Hence, there exists natural number  $K$ , such that  $k \geq K$  implies  $\|x_N(\mu^k, b^k, c_{\tau^k})\| < \frac{\epsilon}{2}$ .

Hence, when  $k \geq K$ ,

$$\begin{aligned} \|x_N(b^k, \bar{c}_{\tau^k})\| &\leq \|x_N(b^k, \bar{c}_{\tau^k}) - x_N(\mu^k, b^k, c_{\tau^k})\| + \|x_N(\mu^k, b^k, c_{\tau^k})\| \\ &< \epsilon, \end{aligned}$$

and  $\{x_N^k\} \rightarrow 0$ . Now, using lemmas 3.6 and 4.4 to establish the third and fourth equalities respectively,

$$\begin{aligned} \lim_{k \rightarrow \infty} x_B^*(b^k, \bar{c}_{\tau^k}) &= \lim_{k \rightarrow \infty} \left( \lim_{\mu \rightarrow 0^+} x_B(\mu, b^k, \bar{c}_{\tau^k}) \right) \\ &= \lim_{k \rightarrow \infty} \left( \lim_{\mu \rightarrow 0^+} z_B\left(\frac{\mu}{\tau^k}, b^k - A_N x_N(\mu, b^k, \bar{c}_{\tau^k}), \delta c\right) \right) \end{aligned}$$

$$\begin{aligned}
&= \lim_{k \rightarrow \infty} z_B^*(\bar{b} - A_N x_N^*(b^k, \bar{c}_{\tau^k}), \delta c) \\
&= z_B^*(\bar{b}, \delta c).
\end{aligned}$$

Hence, if  $\{x^k \in \overline{PCP}_{(b^k, \bar{c}_{\tau^k})}\} \rightarrow x$  and contains infinitely many elements of the form  $x^*(b^k, c_{\tau^k})$ ,  $x = (z_B^*(\bar{b}, \delta c), 0) \in \overline{PCP}_r$  and the proof is complete.  $\blacksquare$

The proof of theorem 4.9 does not use any dual information. This is because it is possible that both the rim data and primal elements converge, while the dual elements diverge. For example, if  $c \in \text{row}(A)$  and  $\{\mu^k\} = \{1, 2, 1, 2, 1, 2, \dots\}$ ,  $x(\mu^k, b^k, c_{\tau^k})$  converges to the analytic center of  $\mathcal{P}_{\bar{b}}$ . However, theorem 2.10 implies that the corresponding dual sequence,  $\{s(\mu^k, b^k, c_{\tau^k})\}$ , has the two cluster points  $s(1, \bar{b}, \bar{c})$  and  $2s(1, \bar{b}, \bar{c}) = s(2, \bar{b}, \bar{c})$ . The problem here is that the near complementarity constraint implies that  $s_i^k = \frac{u^k}{x_i^k} \omega_i$ . The failure of the dual elements to converge is now easily seen from the divergence of the sequence  $\{\mu^k\}$ . In general, even if both  $\{\mu^k\}$  and  $\{x_i^k\}$  converge, the convergence of the sequence  $\left\{\frac{u^k}{x_i^k}\right\}$  is not guaranteed unless

$$\lim_{k \rightarrow \infty} x_i^k > 0.$$

Once the dual counterparts are established in the following section, example 4.19 shows exactly how this can happen. The next theorem does not completely resolve this issue, but it does show when the convergence of  $\{\mu^k\}$

is guaranteed.

**Theorem 4.10** Let  $\bar{r} \in \mathcal{G}$ . Also, let  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ , and  $\{\tau^k \in \mathbb{R}_{++}\} \rightarrow 0$ . Then, the convergence of  $\{x(\mu^k, b^k, \bar{c}_{\tau^k}) \in PCP_{(b^k, c_{\tau^k})}\}$  implies the convergence of  $\{\mu^k\}$  (possibly to infinity, if  $\mathcal{P}_{\bar{b}}$  is bounded) if, and only if,  $\bar{c} \notin \text{row}(A)$ .

**Proof:** Suppose that  $\bar{c} \in \text{row}(A)$ . Consider the situation of  $\mathcal{P}_{\bar{b}}$  being bounded,  $\{\mu^k\} = \{1, 2, 1, 2, \dots\}$ , and  $\{\tau^k\} = \{\frac{1}{k}\}$ . Then,

$$\lim_{k \rightarrow \infty} \frac{\mu^k}{\tau^k} = \infty,$$

and independent of whether or not  $\bar{c} \in \text{row}(A_B)$ ,

$$\lim_{k \rightarrow \infty} x(\mu^k, b^k, \bar{c}_{\tau^k}) = \bar{x}(\bar{b}).$$

Hence, if  $\bar{c} \in \text{row}(A)$ , the convergence of the sequence  $\{x(\mu^k, b^k, \bar{c}_{\tau^k})\}$  cannot guarantee the convergence of  $\{\mu^k\}$ .

Assume  $\bar{c} \notin \text{row}(A)$  and suppose for the sake of attaining a contradiction that  $\{\mu^k\}$  does not converge. Then there exist subsequences,  $\{\mu^{k_i}\}$  and  $\{\mu^{k_j}\}$ , such that

$$0 \leq \lim_{i \rightarrow \infty} \mu^{k_i} < \lim_{j \rightarrow \infty} \mu^{k_j} \leq \infty.$$

If  $0 < \lim_{i \rightarrow \infty} \mu^{k_i} = \mu^1$ , theorem 2.6 implies that

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c_{\tau^{k_i}}) = x(\mu^1, \bar{b}, \bar{c}).$$

Similarly, theorem 2.6 and theorem 2.9 imply

$$\lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, \bar{c}_{\tau^{k_j}}) = \begin{cases} x(\mu^2, \bar{b}, \bar{c}) & \text{if } \lim_{j \rightarrow \infty} \mu^{k_j} = \mu^2 < \infty \\ \bar{x}(\bar{r}) & \text{if } \lim_{j \rightarrow \infty} \mu^{k_j} = \infty. \end{cases}$$

However, theorem 2.11 implies

$$cx(\mu^1, \bar{b}, \bar{c}) < cx(\mu^2, \bar{b}, \bar{c}) < c\bar{x}(\bar{r})$$

for all  $0 < \mu^1 < \mu^2$ , where the last inequality is included only when  $\bar{x}$  exists.

This is a contradiction since this implies

$$\lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, \bar{c}_{\tau^{k_i}}) \neq \lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, \bar{c}_{\tau^{k_j}}).$$

The only situation left is when  $\lim_{i \rightarrow \infty} \mu^{k_i} = 0$ . However,  $\lim_{i \rightarrow \infty} \mu^{k_i} = 0$ , together with lemma 2.15, lead to the contradiction:

$$0 = \lim_{i \rightarrow \infty} x_N(\mu^{k_i}, b^{k_i}, c^{k_i}) \neq \lim_{j \rightarrow \infty} x_N(\mu^{k_j}, b^{k_j}, \bar{c}_{\tau^{k_j}}).$$

Hence,  $\{\mu^k\}$  must converge. ■

### 4.3 Equivalent Dual Statements

This section presents the dual counterparts of the previous section. There there are no new proof techniques, and consequently, the results of this section may be read without the proofs. Example 4.19 shows exactly how the convergence of both the rim data and primal elements does not guarantee the

convergence of the dual elements. The first Lemma is the dual statement of Lemma 4.2.

**Lemma 4.11**  $yb = 0$  for all  $(y, s) \in \text{leftnull} \left( \begin{bmatrix} A_N & A_B \\ I_N & 0 \end{bmatrix} \right)$ .

**Proof:** Let  $(\hat{y}, \hat{s}_N) \in \text{leftnull} \left( \begin{bmatrix} A_N & A_B \\ I_N & 0 \end{bmatrix} \right)$  and  $(y^*, s^*) \in (\mathcal{D}_r^*)^o$ . Then  $(y^*, (0, s_N^*)) + \theta(\hat{y}, (0, \hat{s}_N)) \in \mathcal{D}_r^*$ , for sufficiently small  $\theta$ . Since  $(y^* + \theta\hat{y})b = y^*b$ , for sufficiently small  $\theta$ ,  $\hat{y}b = 0$ . ■

The perpendicular property shown in lemma 4.11 is now used to show that  $(y(\mu, b_\rho, c), s_N(\mu, b_\rho, c))$  is uniquely optimal on a cut through the dual feasible region.

**Lemma 4.12** Let  $r \in \mathcal{G}$  and  $\delta b$  be fixed. Then,  $yb$  is constant on

$$\{(y, s) : yA_B = c_B - s_B(\mu, b_\rho, c), yA_N + s_N = c_N, s_N \geq 0\},$$

and consequently  $(y(\mu, b_\rho, c), s_N(\mu, b_\rho, c))$  is the unique solution to

$$\begin{aligned} & \max\{\rho y \delta b + \mu \sum_{i \in N} \omega_i \ln(s_i) : \\ & yA_B = c_B - s_B(\mu, b_\rho, c), yA_N + s_N = c_N, s_N > 0\}. \end{aligned}$$

**Proof:** By definition  $(y(\mu, b_\rho, c), s(\mu, b_\rho, c))$  is the unique solution to

$$\max\{yb + \rho y \delta b + \mu \sum_{i=1}^n \omega_i \ln(s_i) :$$

$$yA_B + s_B = c_B, yA_N + s_N = c_N, s > 0\}.$$

Fixing  $s_B(\mu, b_\rho, c)$ ,  $(y(\mu, b_\rho, c), s_N(\mu, b_\rho, c))$  is the unique solution to

$$\max\{yb + \rho y \delta b + \mu \sum_{i \in N} \omega_i \ln(s_i) : yA_B = c_B - s_B(\mu, b_\rho, c), yA_N + s_N = c_N, s_N \geq 0\}.$$

So, the result follows once it is shown that  $yb$  is constant on

$$\{(y, s_N) : yA_B = c_B - s_B(\mu, b_\rho, c), yA_N + s_N = c_N, s_N \geq 0\}.$$

If the rows of

$$\begin{bmatrix} A_N & A_B \\ I_N & 0 \end{bmatrix}$$

are linearly independent, then the result follows since this set contains a single

element. Otherwise, let  $(y^1, s_N^1)$  and  $(y^2, s_N^2)$  be in

$$\{(y, s_N) : yA_B = c_B - s_B(\mu, b_\rho, c), yA_N + s_N = c_N, s_N \geq 0\}.$$

Then,

$$(y^1, s_N^1) - (y^2, s_N^2) \in \text{leftnull} \left( \begin{bmatrix} A_N & A_B \\ I_N & 0 \end{bmatrix} \right),$$

and lemma 4.11 implies

$$(y^1 - \alpha(y^1 - y^2))b = y^1 b,$$

for all  $0 \leq \alpha \leq 1$ , which proves the result. ■

As a dual counterpart to  $PCP_{(r,\delta c)}^*$ , define

$$DCP_{(r,\delta b)}^* \equiv \{(p(\eta, \delta b, c), q_N(\eta, \delta b, c)) : \eta > 0\}$$

to be the central path corresponding to the linear program,

$$\max\{p\delta b : pA_B = c_B, pA_N + q = c_N, q \geq 0\}.$$

Both  $\overline{DCP}_r$  and  $\overline{DCP}_{(r,\delta b)}^*$  are defined as  $\overline{PCP}_r$  and  $\overline{PCP}_{(r,\delta b)}^*$  were defined. As in the previous section, the needed observations provided by lemma 4.12 are that

$$y(\mu, b_\rho, c) = p \left( \frac{\mu}{\rho}, \delta b, \begin{bmatrix} c_B - s_B(\mu, b_\rho, c) \\ c_N \end{bmatrix} \right) \quad (4.6)$$

and

$$s_N(\mu, b_\rho, c) = q_N \left( \frac{\mu}{\rho}, \delta b, \begin{bmatrix} c_B - s_B(\mu, b_\rho, c) \\ c_N \end{bmatrix} \right). \quad (4.7)$$

Similar to lemma 4.4, the next lemma provides a useful convergence result when  $b$  is held fixed.

**Lemma 4.13** Let  $\bar{r} \in \mathcal{G}$ ,  $\{c^k \in \mathcal{G}_c\} \rightarrow \bar{c} \in \mathcal{G}_c$ , and  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$ . Then,

$$\lim_{k \rightarrow \infty} (y(\mu^k, \bar{b}, c^k), s(\mu^k, \bar{b}, c^k)) = (y^*(\bar{r}), s^*(\bar{r})).$$

**Proof:** Let  $\{c^k\}$  and  $\{\mu^k\}$  be as above. Theorem 2.14 implies that

$$\{(y(\mu^k, \bar{b}, c^k), s(\mu^k, \bar{b}, c^k))\}$$

is bounded and hence contains a convergent subsequence. Let

$$\lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, \bar{b}, c^{k_i}), s(\mu^{k_i}, \bar{b}, c^{k_i}) \right) = (\hat{y}, \hat{s}).$$

Showing that  $(\hat{y}, \hat{s}) = (y^*(\bar{r}), s^*(\bar{r}))$  proves the result by the uniqueness of  $(y^*(\bar{r}), s^*(\bar{r}))$ . Clearly,  $(\hat{y}, \hat{s}) \in \mathcal{D}_{\bar{e}}$ . Lemma 4.12 shows that

$$\left( y(\mu^{k_i}, \bar{b}, c^{k_i}), s_N(\mu^{k_i}, \bar{b}, c^{k_i}) \right)$$

solves

$$\max \left\{ \sum_{j \in N} \ln(s_j) : yA_B = c_B - s_B(\mu^{k_i}, \bar{b}, c^{k_i}), yA_N + s_N = c_N, s_N > 0 \right\}.$$

The necessary and sufficient Lagrange conditions for this mathematical program are the existence of  $x^i$  such that

$$y(\mu^{k_i}, \bar{b}, c^{k_i})A_B = c_B - s_B(\mu^{k_i}, \bar{b}, c^{k_i})$$

$$y(\mu^{k_i}, \bar{b}, c^{k_i})A_N + s_N(\mu^{k_i}, \bar{b}, c^{k_i}) = c_N$$

$$Ax^i = 0$$

$$x_N^i = S_N^{-1}(\mu^{k_i}, \bar{b}, c^{k_i})e$$

$$s_N(\mu^{k_i}, \bar{b}, c^{k_i}) > 0.$$

If  $\hat{s}_N > 0$ ,  $\{x_N^i\}$  converges, say to  $\hat{x}_N$ . In this case, setting

$$x_B^i = -A_B^+ A_N x_N^i,$$

implies  $\{x_B^i\} = \{-A_B^+ A_N x_N^i\} \rightarrow -A_B^+ A_N \hat{x}_N$ , and

$$\begin{aligned} A\hat{x} &= -A_B A_B^+ A_N \hat{x}_N + A_N \hat{x}_N \\ &= -A_N \hat{x}_N + A_N \hat{x}_N \\ &= 0. \end{aligned}$$

Hence, because Lemma 2.15 shows  $\{s_B(\mu^{k_i}, \bar{b}, c^{k_i})\} \rightarrow 0$ ,  $\hat{s}_N > 0$  implies

$$\begin{aligned} \hat{y}A_B &= c_B \\ \hat{y}A_N + \hat{s}_N &= c_N \\ A\hat{x} &= 0 \\ \hat{x}_N &= \hat{S}^{-1}e, \end{aligned}$$

which are the necessary and sufficient conditions for  $(\hat{y}, \hat{s}) = (y^*(\bar{r}), s^*(\bar{r}))$ . The fact that  $\hat{s}_N > 0$  is now established.

Lemma 2.17 guarantees the existence of a sequence  $\{(\tilde{y}^i, \tilde{s}_N^i)\}$  that converges to  $(\tilde{y}, \tilde{s}_N)$ , where  $\tilde{s}_N > 0$  and

$$(\tilde{y}^i, \tilde{s}_N^i) \in \{(y, s_N) : yA_B = c_B - s_B(\mu^{k_i}, \bar{b}, c^{k_i}), yA_N + s_N = c_N, s_N > 0\}.$$

The optimality of  $(y(\mu^{k_i}, \bar{b}, c^{k_i}), s_N(\mu^{k_i}, \bar{b}, c^{k_i}))$  implies that

$$\sum_{j \in N} \ln(s_j(\mu^{k_i}, \bar{b}, c^{k_i})) \geq \sum_{j \in N} \ln(\tilde{s}_j^i).$$

Since the right-hand side of this inequality converges to  $\sum_{j \in N} \ln(\tilde{s}_j)$ ,  $\hat{s}_N > 0$ , and the result follows. ■

The next theorem is the dual extension of theorem 2.8 to include the case of general cost coefficient changes and linear right-hand side changes.

**Theorem 4.14** Let  $(\bar{\mu}, \bar{r}) \in \mathbb{R}_+ \times \mathcal{G}$  and  $\mathfrak{B}$  be fixed. Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow \bar{\mu}$ ,  $\{\rho^k \in \mathbb{R}_{++}\} \rightarrow 0$ , and  $\{c^k \in \mathcal{G}_c\} \rightarrow \bar{c}$ . Then,

$$\lim_{k \rightarrow \infty} \left( y(\mu^k, \bar{b}_{\rho^k}, c^k), s(\mu^k, \bar{b}_{\rho^k}, c^k) \right) = \begin{cases} \left( y(\bar{\mu}, \bar{b}, \bar{c}), s(\bar{\mu}, \bar{b}, \bar{c}) \right) & \text{if } \bar{\mu} > 0 \\ \left( y^*(\bar{r}), s^*(\bar{r}) \right) & \text{if } \bar{\mu} = 0, \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \infty \\ \left( p(\eta, \mathfrak{B}, \bar{c}), (0, q_N(\eta, \mathfrak{B}, \bar{c})) \right) & \text{if } \bar{\mu} = 0, \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \eta > 0 \\ \left( p^*(\mathfrak{B}, \bar{c}), (0, q_N^*(\mathfrak{B}, \bar{c})) \right) & \text{if } \bar{\mu} = 0, \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = 0 \end{cases}$$

**Proof:** If  $\bar{\mu} > 0$ , the result follows as a direct consequence of theorem 2.6.

Assume that  $\bar{\mu} = 0$  and consider the case when  $\lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \eta > 0$ . Lemma 2.15 implies

$$\lim_{k \rightarrow \infty} s_B(\mu^k, \bar{b}_{\rho^k}, c^k) = 0.$$

Since  $\frac{\mu^k}{\rho^k}$  is bounded away from zero, theorem 2.6, together with 4.6, and 4.7

imply that

$$\begin{aligned}\lim_{k \rightarrow \infty} y(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} p \left( \frac{\mu^k}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= p(\eta, \delta, \bar{c})\end{aligned}$$

and

$$\begin{aligned}\lim_{k \rightarrow \infty} s_N(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} q_N \left( \frac{\mu^k}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= q(\eta, \delta, \bar{c}).\end{aligned}$$

The next situation considered is when  $\lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = 0$ . Since lemma 2.15

implies

$$\left\{ \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right\} \rightarrow \bar{c},$$

lemma 4.13 implies that

$$\begin{aligned}\lim_{k \rightarrow \infty} y(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} p \left( \frac{\mu^k}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= p^*(\delta, \bar{c})\end{aligned}$$

and

$$\begin{aligned}\lim_{k \rightarrow \infty} s_N(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} q_N \left( \frac{\mu^k}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= q_N^*(\delta, \bar{c}).\end{aligned}$$

The last case to consider is when  $\lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \infty$ . Once again, lemma 2.15

implies that

$$\left\{ \left[ \begin{array}{c} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{array} \right] \right\} \rightarrow \bar{c},$$

Furthermore,

$$\lim_{k \rightarrow \infty} \frac{\delta}{\frac{\mu^k}{\rho^k}} = 0.$$

Since corollary 2.5 shows that  $\mathcal{D}_{\bar{r}}^*$  is bounded, the conditions to apply theorem 2.18 are met, which implies

$$\begin{aligned} \lim_{k \rightarrow \infty} y(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} p \left( \frac{\mu^k}{\rho^k}, \delta, \left[ \begin{array}{c} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_\tau^k \end{array} \right] \right) \\ &= y^*(\bar{r}) \end{aligned}$$

and

$$\begin{aligned} \lim_{k \rightarrow \infty} s_N(\mu^k, \bar{b}_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} q_N \left( \frac{\mu^k}{\rho^k}, \delta, \left[ \begin{array}{c} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{array} \right] \right) \\ &= s_N^*(\bar{r}). \end{aligned}$$

■

As in the previous section, the goal is to use theorem 4.14 to establish necessary and sufficient conditions for  $\{(\mu^k, b_{\rho^k}, c^k)\}$ , which guarantee the

convergence of  $\left\{ \left( p(\mu^k, b_{\rho^k}, c^k), q_N(\mu^k, b_{\rho^k}, c^k) \right) \right\}$ . The next two lemmas provide enough added information to state such conditions.

**Lemma 4.15** Let  $\bar{r} \in \mathcal{G}$ ,  $\delta \notin \mathcal{H}_b$ , and  $\{c^k \in \mathcal{G}_c\} \rightarrow \bar{c} \in \mathcal{G}_c$ . Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$  and  $\{\rho^k \in \mathbb{R}_{++}\} \rightarrow 0$  be such that the sequence  $\left\{ \frac{\mu^k}{\rho^k} \right\}$  does not converge. Let  $\left\{ \frac{\mu^{k_i}}{\rho^{k_i}} \right\}$  and  $\left\{ \frac{\mu^{k_j}}{\rho^{k_j}} \right\}$  be two convergent subsequences, one of which possibly converges to infinity. Then,

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\rho^{k_i}} \neq \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\rho^{k_j}},$$

implies

$$\begin{aligned} & \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right) \\ & \neq \lim_{j \rightarrow \infty} \left( y(\mu^{k_j}, b_{\rho^{k_j}}, c^{k_j}), s(\mu^{k_j}, b_{\rho^{k_j}}, c^{k_j}) \right). \end{aligned}$$

**Proof:** Assume without loss of generality that

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\rho^{k_i}} < \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\rho^{k_j}}.$$

From theorem 4.14,

$$\begin{aligned} & \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, \bar{b}_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, \bar{b}_{\rho^{k_i}}, c^{k_i}) \right) = \\ & \begin{cases} \left( p^*(\delta, \bar{c}), (0, q_N^*(\delta, \bar{c})) \right) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\rho^{k_i}} = 0 \\ \left( p(\eta^1, \delta, \bar{c}), (0, q_N(\eta^1, \delta, \bar{c})) \right) & \text{if } \lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\rho^{k_i}} = \eta^1 > 0, \end{cases} \end{aligned}$$

and

$$\lim_{j \rightarrow \infty} \left( y(\mu^{k_j}, \bar{b}_{\rho^{k_j}}, c^{k_j}), s(\mu^{k_j}, \bar{b}_{\rho^{k_j}}, c^{k_j}) \right) =$$

$$\begin{cases} (p(\eta^2, \mathfrak{d}, \bar{c}), (0, q_N(\eta^2, \mathfrak{d}, \bar{c}))) & \text{if } \lim_{j \rightarrow \infty} \frac{\mu^{kj}}{\rho^{kj}} = \eta^2 > 0 \\ (y^*(\bar{r}), s^*(\bar{r})) & \text{if } \lim_{j \rightarrow \infty} \frac{\mu^{kj}}{\rho^{kj}} = \infty. \end{cases}$$

Since  $\mathfrak{d} \notin \mathcal{H}_b$ , lemma 2.13 and theorem 2.11 imply that for any  $\eta^1 < \eta^2$ ,

$$p^*(\mathfrak{d}, \bar{c})\mathfrak{d} > p(\eta^1, \mathfrak{d}, \bar{c})\mathfrak{d} > p(\eta^2, \mathfrak{d}, \bar{c})\mathfrak{d} > y^*(\bar{r})\mathfrak{d},$$

which proves the result. ■

**Lemma 4.16** If  $\bar{r} \in \mathcal{G}$  and  $\mathfrak{d} \in \mathcal{H}_b$ ,

$$(p(\eta, \mathfrak{d}, \bar{c}), (0, q_N(\eta, \mathfrak{d}, \bar{c}))) = (p^*(\mathfrak{d}, \bar{c}), (0, q_N^*(\mathfrak{d}, \bar{c}))),$$

for all  $\eta \in \mathbb{R}_{++}$ .

**Proof:** Lemma 2.13 implies that  $\mathfrak{d} \in \text{col}(A_B)$  and theorem 2.10 implies that

$$(p(\eta^1, \mathfrak{d}, \bar{c}), (0, q_N(\eta^1, \mathfrak{d}, \bar{c}))) = (p(\eta^2, \mathfrak{d}, \bar{c}), (0, q_N(\eta^2, \mathfrak{d}, \bar{c}))) \text{ for all } \eta^1 \text{ and } \eta^2 \text{ in}$$

$\mathbb{R}_{++}$ . Since

$$(p^*(\mathfrak{d}, \bar{c}), (0, q^*(\mathfrak{d}, \bar{c}))) = \lim_{\eta \rightarrow 0^+} (p(\eta, \mathfrak{d}, \bar{c}), (0, q_N(\eta, \mathfrak{d}, \bar{c}))),$$

the proof is complete. ■

The next theorem characterizes the conditions for which the sequence

$$\left\{ \left( y(\mu^k, b_{\rho^k}, c^k), s(\mu^k, b_{\rho^k}, c^k) \right) \right\} \text{ converges.}$$

**Theorem 4.17** Let  $\bar{r} \in \mathcal{G}$ . Also, let  $\{\mu^k \in \mathbb{R}_{++}\} \rightarrow 0$ ,  $\{\rho^k \in \mathbb{R}_{++}\} \rightarrow 0$ , and  $\{c^k \in \mathcal{G}_c\} \rightarrow \bar{c} \in \mathcal{G}_c$ . If  $\mathfrak{B} \in \mathcal{H}_b$ ,

$$\lim_{k \rightarrow \infty} \left( y(\mu^k, b_{\rho^k}, c^k), s(\mu^k, b_{\rho^k}, c^k) \right) = (p^*(\mathfrak{B}, \bar{c}), (0, q_N^*(\mathfrak{B}, \bar{c}))).$$

If  $\mathfrak{B} \notin \mathcal{H}_b$ ,

$$\lim_{k \rightarrow \infty} \left( y(\mu^k, b_{\rho^k}, c^k), s(\mu^k, b_{\rho^k}, c^k) \right)$$

exists if, and only if,  $\left\{ \frac{\mu^k}{\rho^k} \right\}$  converges, possibly to infinity.

**Proof:** If  $\mathfrak{B} \in \mathcal{H}_b$ , lemma 4.16 implies that

$$\begin{aligned} y(\mu^k, b_{\rho^k}, c^k) &= p \left( \frac{\mu^k}{\rho^k}, \mathfrak{B}, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c^k \end{bmatrix} \right) \\ &= p^* \left( \mathfrak{B}, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c^k \end{bmatrix} \right) \end{aligned}$$

and

$$\begin{aligned} s_N(\mu^k, b_{\rho^k}, c^k) &= q_N \left( \frac{\mu^k}{\rho^k}, \mathfrak{B}, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= q_N^* \left( \mathfrak{B}, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right). \end{aligned}$$

From lemma 3.6, both  $p^*$  and  $q_N^*$  are continuous functions of their right hand side, and from lemma 2.15  $s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \rightarrow 0$ . So,

$$\begin{aligned} \lim_{k \rightarrow \infty} y(\mu^k, b_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} p^* \left( \delta b, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c^k \end{bmatrix} \right) \\ &= p^*(\delta b, \bar{c}) \end{aligned}$$

and

$$\begin{aligned} \lim_{k \rightarrow \infty} s_N(\mu^k, b_{\rho^k}, c^k) &= \lim_{k \rightarrow \infty} q_N^* \left( \delta b, \begin{bmatrix} c_B^k - s_B(\mu^k, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \\ &= q_N^*(\delta b, \bar{c}). \end{aligned}$$

Assume  $\delta b \notin \mathcal{H}_b$ . If  $\left\{ \frac{\mu^k}{\rho^k} \right\}$  converges, even to infinity, then theorem 4.14 shows the existence of  $\lim_{k \rightarrow \infty} \left( y(\mu^k, b_{\rho^k}, c^k), s(\mu^k, b_{\rho^k}, c^k) \right)$ . If  $\left\{ \frac{\mu^k}{\rho^k} \right\}$  does not converge, then this sequence has at least two cluster points, one of which may be infinity. Let  $\left\{ \frac{\mu^{k_i}}{\rho^{k_i}} \right\}$  and  $\left\{ \frac{\mu^{k_j}}{\rho^{k_j}} \right\}$  be two convergent subsequences such that

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\rho^{k_i}} \neq \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\rho^{k_j}}.$$

Theorem 4.14 implies that both

$$\lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right)$$

and

$$\lim_{j \rightarrow \infty} \left( y(\mu^{k_j}, b_{\rho^{k_j}}, c^{k_j}), s(\mu^{k_j}, b_{\rho^{k_j}}, c^{k_j}) \right)$$

exist and lemma 4.15 implies that these limits are different. ■

The last theorem is the dual analog to theorem 4.9, and it shows how the dual omega central path behaves under simultaneous perturbations in the rim date, provided that the change in  $b$  is linear. In the statement of the theorem,  $\overline{DCP}_{(\partial b, \bar{c})}^* \times \mathbf{0}^{|B|}$  is used to denote  $\{(y, (0, s_N)) : (y, s_N) \in \overline{DCP}_{(\partial b, \partial b)}^*\}$ .

**Theorem 4.18** If  $\bar{r} \in \mathcal{G}$  and  $\partial b$  is fixed,

$$\lim_{(\bar{b}, c) \rightarrow (\bar{b}, \bar{c})} \overline{DCP}_{(\bar{b}, c)} = \overline{DCP}_{\bar{r}} \cup \left( \overline{DCP}_{(\partial b, \bar{c})}^* \times \mathbf{0}^{|B|} \right).$$

**Proof:** Let  $\{c^k \in \mathcal{G}_c\} \rightarrow \bar{c} \in \mathcal{G}_c$ , and  $\{\rho^k \in \mathbb{R}_{++}\} \rightarrow 0$ . The first part of the proof establishes

$$\lim_{(\bar{b}, c) \rightarrow (\bar{b}, \bar{c})} DCP_{(b, c)} = \overline{DCP}_{\bar{r}} \cup \left( \overline{DCP}_{(\partial b, \bar{c})}^* \times \mathbf{0}^{|B|} \right).$$

Suppose the sequence  $\{(y^k, s^k) \in DCP_{(b_{\rho^k}, c^k)}\}$  converges to  $(\hat{y}, \hat{s})$ . For each  $k$ , let  $\mu^k$  be such that  $(y^k, s^k) = (y(\mu^k, b_{\rho^k}, c^k), s(\mu^k, b_{\rho^k}, c^k))$ . Consider a subsequence of  $\{\mu^k\}$ , say  $\{\mu^{k_i}\}$ , that converges, possibly to infinity. The task of showing that  $(\hat{y}, \hat{s}) \in \overline{DCP}_{\bar{r}} \cup \left( \overline{DCP}_{(\partial b, \bar{c})}^* \times \mathbf{0}^{|B|} \right)$  is divided into three cases.

**Case 1** If  $\{\mu^{k_i}\} \rightarrow \bar{\mu} > 0$ , theorem 2.6 implies

$$\begin{aligned} (\hat{y}, \hat{s}) &= \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right) \\ &= \left( y(\bar{\mu}, \bar{b}, \bar{c}), s(\bar{\mu}, \bar{b}, \bar{c}) \right) \in DCP_{\bar{r}}, \end{aligned}$$

and this case is complete.

**Case 2** Suppose  $\{\mu^{k_i}\} \rightarrow 0$ . When  $\mathfrak{d}b \in \mathcal{H}_b$ , theorem 4.17 shows that

$$\begin{aligned} (\hat{y}, \hat{s}) &= \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right) \\ &= (p^*(\mathfrak{d}b, \bar{c}), (0, q_N^*(\mathfrak{d}b, \bar{c}))) \in \overline{DCP}_{(\mathfrak{d}b, \bar{c})}^* \times \mathbf{0}^{|B|}. \end{aligned}$$

Also, when  $\mathfrak{d}b \notin \mathcal{H}_b$ , theorem 4.17 guarantees that  $\left\{ \frac{\mu^{k_i}}{\rho^{k_i}} \right\}$  converges, possibly to infinity. Then, from theorem 4.14,

$$\begin{aligned} (\hat{y}, \hat{s}) &= \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right) \\ &= \begin{cases} (y^*(\bar{r}), s^*(\bar{r})) & \text{if } \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \infty \\ (p(\eta, \mathfrak{d}b, \bar{c}), (0, q(\eta, \mathfrak{d}b, \bar{c}))) & \text{if } \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = \eta > 0 \\ (p^*(\mathfrak{d}b, \bar{c}), (0, q^*(\mathfrak{d}b, \bar{c}))) & \text{if } \lim_{k \rightarrow \infty} \frac{\mu^k}{\rho^k} = 0. \end{cases} \end{aligned}$$

Since  $(y^*(\bar{r}), s^*(\bar{r}))$ ,  $(p(\eta, \mathfrak{d}b, \bar{c}), (0, q(\eta, \mathfrak{d}b, \bar{c})))$ , and  $(p^*(\mathfrak{d}b, \bar{c}), (0, q^*(\mathfrak{d}b, \bar{c})))$  are all in  $\overline{DCP}_{(\mathfrak{d}b, \bar{c})}^* \times \mathbf{0}^{|B|}$ , this case is dismissed.

**Case 3** Assume  $\{\mu^{k_i}\} \rightarrow \infty$ . This case is completed once it is shown that theorem 2.18 is applicable, which in turn shows that

$$\begin{aligned} (\hat{y}, \hat{s}) &= \lim_{i \rightarrow \infty} \left( y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) \right) \\ &= (\bar{y}, \bar{s}) \in \overline{DCP}_{\bar{r}}. \end{aligned}$$

Since  $\{c^k\} \rightarrow \bar{c}$  and  $\left\{ \frac{b_{\rho^{k_i}}}{\mu^{k_i}} \right\} \rightarrow 0$ , all that is left to show is that  $\mathcal{D}_{\bar{c}}$  is bounded.

Assume for the sake of attaining a contradiction that  $\mathcal{D}_{\bar{c}}$  is unbounded. Since

$$\|y\| = \|(c - s)A^T(AA^T)^{-1}\| \leq \|c - s\| \|A^T(AA^T)^{-1}\|,$$

the unboundedness of  $\mathcal{D}_{\bar{c}}$  means there exists a sequence  $\{(y^i, s^i) \in \mathcal{D}_{c^{k_i}}\}$  such that  $\|s^i\| \rightarrow \infty$  and  $\left\{\frac{y^i}{\mu^{k_i}}\right\} \rightarrow 0$ . The optimality of

$$\left(y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}), s(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i})\right)$$

implies

$$y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i})b_{\rho^{k_i}} + \mu^{k_i} \sum_{j=1}^n \omega_j \ln(s_j(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i})) \geq y^i b_{\rho^{k_i}} + \mu^{k_i} \sum_{j=1}^n \omega_j \ln(s_j^i),$$

or equivalently,

$$\left(y(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i}) - y^i\right) \frac{b_{\rho^{k_i}}}{\mu^{k_i}} + \sum_{j=1}^n \omega_j \ln(s_j(\mu^{k_i}, b_{\rho^{k_i}}, c^{k_i})) \geq \sum_{j=1}^n \omega_j \ln(s_j^i).$$

The contradiction that the left hand side of this last inequality is bounded implies that  $\mathcal{D}_{\bar{c}}$  is bounded.

So far, the proof has shown that if a sequence  $\{(y^k, s^k) \in DCP_{(b_{\rho^k}, c^k)}\}$  converges, the limit point of this sequence must be an element of

$$\overline{DCP}_{\bar{r}} \cup \left(\overline{DCP}_{(\bar{b}, \bar{c})}^* \times \mathbf{0}^{|B|}\right).$$

The next piece of the proof shows that every element in

$$\overline{DCP}_{\bar{r}} \cup \left(\overline{DCP}_{(\bar{b}, \bar{c})}^* \times \mathbf{0}^{|B|}\right)$$

is the limit of a sequence  $\{(y^k, s^k) \in DCP_{(b_{\rho^k}, c^k)}\}$ .

Let  $(y, s) \in \overline{DCP}_{\bar{r}}$ . From theorem 2.18, if  $\mathcal{D}_{\bar{c}}$  is bounded and  $(y, s) = (\bar{y}, \bar{s})$ ,  $\left\{(y(k, b_{\rho^k}, c^k), s(k, b_{\rho^k}, c^k))\right\} \rightarrow (y, s)$ . Similarly, if

$$(y, s) = \left(y(\mu, \bar{b}, \bar{c}), s(\mu, \bar{b}, \bar{c})\right) \quad \text{or} \quad (y, s) = (y^*(\bar{r}), s^*(\bar{r})),$$

theorem 4.14 shows

$$\left\{ \left( y\left(\mu + \frac{1}{k}, b_{\rho^k}, c^k\right), s\left(\mu + \frac{1}{k}, b_{\rho^k}, c^k\right) \right) \right\} \rightarrow \left( y(\mu, \bar{b}, \bar{c}), s(\mu, \bar{b}, \bar{c}) \right)$$

and

$$\left\{ \left( y(\sqrt{\rho^k}, b_{\rho^k}, c^k), s(\sqrt{\rho^k}, b_{\rho^k}, c^k) \right) \right\} \rightarrow (y^*(\bar{r}), s^*(\bar{r})).$$

Let  $(y, s) \in \overline{DCP}_{(\delta b, \bar{c})}^* \times \mathbf{0}^{|B|}$ . If  $(y, s)$  is  $(p(\eta, \delta b, \bar{c}), (0, q_N(\eta, \delta b, \bar{c})))$  or  $(p^*(\delta b, \bar{c}), (0, q_N^*(\delta b, \bar{c})))$ , lemma 2.15 and theorem 4.14 demonstrate that

$$\left\{ \left( y(\eta \rho^k, b_{\rho^{k_i}}, c^k), s(\eta \rho^k, b_{\rho^{k_i}}, c^k) \right) \right\} \rightarrow (p(\eta, \delta b, \bar{c}), (0, q_N(\eta, \delta b, \bar{c})))$$

and

$$\left\{ \left( y((\rho^k)^2, b_{\rho^{k_i}}, c^k), s((\rho^k)^2, b_{\rho^{k_i}}, c^k) \right) \right\} \rightarrow (p^*(\delta b, \bar{c}), (0, q_N^*(\delta b, \bar{c}))).$$

The proof has now established that

$$\lim_{(b_\rho, c) \rightarrow (\bar{b}, \bar{c})} DCP_{(b_{\rho^k}, c^k)} = \overline{DCP}_{\bar{r}} \cup \left( \overline{DCP}_{(\delta b, \bar{c})}^* \times \mathbf{0}^{|B|} \right).$$

What remains to be shown is that if the convergent sequence  $\{(y^k, s^k) \in \overline{DCP}_{(\bar{b}_{\rho^k}, c^k)}\}$  contains infinitely many of either  $(y^*(\bar{b}_{\rho^k}, c^k), s^*(\bar{b}_{\rho^k}, c^k))$ , or in the case when  $\mathcal{D}_{\bar{c}}$  is bounded,  $(\bar{y}(c^k), \bar{s}(c^k))$ , the limit point of the sequence is contained in  $\overline{DCP}_{\bar{r}} \cup \left( \overline{DCP}_{(\delta b, \bar{c})}^* \times \mathbf{0}^{|B|} \right)$ . From lemma 2.16,  $\mathcal{D}_{c^k}$  is bounded when  $\mathcal{D}_{\bar{c}}$  is bounded. Since lemma 3.6 shows that  $(\bar{y}(c), \bar{s}(c))$  is a continuous function of  $c$ ,

$$\left\{ (y^k, s^k) \in \overline{DCP}_{(b_{\rho^k}, c^k)} \right\} \rightarrow (\bar{y}(\bar{c}), \bar{s}(\bar{c})),$$

when the sequence contains  $(\bar{y}(c^k), \bar{s}(c^k))$  infinitely many times.

Suppose  $\{(y^k, s^k) \in \overline{DCP}_{(b_{\rho^k}, c^k)}\} \rightarrow (y, s)$ , and that the sequence contains infinitely many elements of the form  $(y^*(\bar{b}_{\rho^k}, c^k), s^*(\bar{b}_{\rho^k}, c^k))$ . First, the fact that  $\lim_{k \rightarrow \infty} s_B^*(b_{\rho^k}, c^k) = 0$  is established. By definition,

$$s_B^*(b_{\rho^k}, c^k) = \lim_{\mu \rightarrow 0^+} s_B(\mu, b_{\rho^k}, c^k).$$

So, when  $(y^k, s^k) = (y^*(\bar{b}_{\rho^k}, c^k), s^*(\bar{b}_{\rho^k}, c^k))$ , there exists  $\bar{\mu}^k > 0$  such that  $\mu \in (0, \bar{\mu}^k)$  implies

$$\|s_B^*(b_{\rho^k}, c^k) - s_B(\mu, b_{\rho^k}, c^k)\| < \frac{\epsilon}{2}.$$

Let  $\{\mu^k \in (0, \bar{\mu}^k)\} \rightarrow 0$ . From lemma 2.15,  $\lim_{k \rightarrow \infty} s_B(\mu^k, b_{\rho^k}, c^k) = 0$ . So, there exists natural number  $K$ , such that  $k \geq K$  implies

$$\|s_B(\mu^k, b_{\rho^k}, c^k)\| < \frac{\epsilon}{2}.$$

Hence, when  $k \geq K$ ,

$$\begin{aligned} \|s_B^*(b_{\rho^k}, c^k)\| &\leq \|s_B^*(b_{\rho^k}, c^k) - s_B(\mu^k, b_{\rho^k}, c^k)\| + \|s_B(\mu^k, b_{\rho^k}, c^k)\| \\ &< \epsilon, \end{aligned}$$

and  $\lim_{k \rightarrow \infty} s_B^*(b_{\rho^k}, c^k) = 0$ . Now, using lemma 4.13 to establish the third equality,

$$\lim_{k \rightarrow \infty} (y^*(\bar{b}_{\rho^k}, c^k), s_N^*(\bar{b}_{\rho^k}, c^k))$$

$$\begin{aligned}
&= \lim_{k \rightarrow \infty} \left( \lim_{\mu \rightarrow 0^+} \left( y(\mu, \bar{b}_{\rho^k}, c^k), s_N(\mu, \bar{b}_{\rho^k}, c^k) \right) \right) \\
&= \lim_{k \rightarrow \infty} \left( \lim_{\mu \rightarrow 0^+} \left( p \left( \frac{\mu}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right), \right. \right. \\
&\quad \left. \left. q_N \left( \frac{\mu}{\rho^k}, \delta, \begin{bmatrix} c_B^k - s_B(\mu, \bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \right) \right) \\
&= \lim_{k \rightarrow \infty} \left( p^* \left( \delta, \begin{bmatrix} c_B^k - s_B^*(\bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right), q_N^* \left( \delta, \begin{bmatrix} c_B^k - s_B^*(\bar{b}_{\rho^k}, c^k) \\ c_N^k \end{bmatrix} \right) \right) \\
&= (p^*(\delta, \bar{c}), q_N^*(\delta, \bar{c})).
\end{aligned}$$

Hence, if  $\{(y^k, s^k) \in \overline{DCP}_{(b_{\rho^k}, c^k)}\} \rightarrow (y, s)$ , and the sequence contains infinitely many elements of the form  $(y^*(\bar{b}_{\rho^k}, c^k), s^*(\bar{b}_{\rho^k}, c^k))$ ,

$$(y, s) = (p^*(\delta, \bar{c}), (0, q_N^*(\delta, \bar{c}))) \in \overline{DCP}_{(\delta, cb)}^* \times \mathbf{0}^{|B|},$$

and the proof is complete. ■

As in the previous section, this last result did not use any primal information in the proof. This is because the primal elements may diverge even when both the rim data elements and  $\mu$  converge. This section ends with an example demonstrating such behavior.

**Example 4.19** Consider the linear program,

$$\min\{\tau\delta c_1 x_1 + \tau\delta c_2 x_2 + x_3 : 0 \leq x_i \leq 1 + \rho\delta b, i = 1, 2, 3\}.$$

From example 2.12,

$$\begin{aligned}
& x_i(\mu^k, b_{\rho^k}, c_{\tau^k}) \\
&= \begin{cases} \frac{\tau^k \delta c_i (1 + \rho^k \delta b) + 2\mu^k - \sqrt{(\tau^k \delta c_i)^2 (1 + \rho^k \delta b)^2 + (2\mu^k)^2}}{2\tau^k \delta c_i} & \text{if } i = 1, 2 \\ \frac{(1 + \rho^k \delta b) + 2\mu^k - \sqrt{(1 + \rho^k \delta b)^2 + (2\mu^k)^2}}{2} & \text{if } i = 3 \end{cases} \\
&= \begin{cases} \frac{1}{2}(1 + \rho^k \delta b) + \frac{\mu^k}{\tau^k \delta c_i} - \frac{1}{2}\sqrt{(1 + \rho^k \delta b)^2 + \left(\frac{2\mu^k}{\tau^k \delta c_i}\right)^2} & \text{if } i = 1, 2 \\ \frac{1}{2}(1 + \rho^k \delta b) + \mu^k - \frac{1}{2}\sqrt{(1 + \rho^k \delta b)^2 + (2\mu^k)^2} & \text{if } i = 3. \end{cases}
\end{aligned}$$

Using that  $x_i(\mu^k, b_{\rho^k}, c_{\tau^k})s_i(\mu^k, b_{\rho^k}, c_{\tau^k}) = -\mu$  (the negative is needed because the primal problem is not in symmetric form), it follows that the complementary dual variables are

$$\begin{aligned}
& s_i(\mu^k, b_{\rho^k}, c_{\tau^k}) \\
&= \begin{cases} \frac{-2\mu^k \tau^k \delta c_i}{\tau^k \delta c_i (1 + \rho^k \delta b) + 2\mu^k - \sqrt{(\tau^k \delta c_i)^2 (1 + \rho^k \delta b)^2 + (2\mu^k)^2}} & \text{if } i = 1, 2 \\ \frac{-2\mu^k}{(1 + \rho^k \delta b) + 2\mu^k - \sqrt{(1 + \rho^k \delta b)^2 + (2\mu^k)^2}} & \text{if } i = 3 \end{cases} \\
&= \begin{cases} 2\sqrt{\left(\frac{\mu^k}{1 + \rho^k \delta b}\right)^2 + (\tau^k \delta c_i)^2} - \frac{2\mu^k}{1 + \rho^k \delta b} - \tau^k \delta c_i & \text{if } i = 1, 2 \\ 2\sqrt{\left(\frac{\mu^k}{1 + \rho^k \delta b}\right)^2 + 1} - \frac{2\mu^k}{1 + \rho^k \delta b} - 1 & \text{if } i = 3. \end{cases}
\end{aligned}$$

Allowing  $\{\mu^k\} = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\}$ ,  $\{\tau^k\} = \{1, 1, \frac{1}{4}, \frac{1}{4}, \frac{1}{16}, \frac{1}{16}, \dots\}$ , and  $\{\rho^k \in \mathbb{R}_{++}\} \rightarrow 0$ , the sequence  $\{x(\mu^k, b_{\rho^k}, c_{\tau^k})\}$  has the two cluster points,

$$\left( \frac{1}{2} + \frac{1}{\delta c_1} - \frac{1}{2}\sqrt{1 + \left(\frac{2}{\delta c_1}\right)^2}, \frac{1}{2} + \frac{1}{\delta c_2} - \frac{1}{2}\sqrt{1 + \left(\frac{2}{\delta c_2}\right)^2}, 0 \right)$$

and

$$\left( \frac{1}{2} \sqrt{1 + \left( \frac{2}{2\delta c_1} \right)^2} - \frac{1}{2} - \frac{1}{2\delta c_1}, \frac{1}{2} \sqrt{1 + \left( \frac{2}{2\delta c_1} \right)^2} - \frac{1}{2} - \frac{1}{2\delta c_1}, 0 \right),$$

while the sequence  $\{s(\mu^k, b_{\rho^k}, c_{\tau^k})\}$  has the limit,  $(0, 0, 1)$ .

#### 4.4 Central Path Convergence Under General Rim Data Changes

The previous two sections discussed the convergence properties of the central path under linear changes in the objective function. Characterizing the convergence of  $\{x(\mu, b, c)\}$  and  $\{(y(\mu, b, c), s(\mu, b, c))\}$  was paramount in establishing the results of the last two sections. The restriction of allowing only linear changes in the objective function coefficients is relinquished in this section. The analysis becomes significantly more challenging, and characterizing the conditions of  $\{(\mu, b, c)\}$  for which the primal and dual elements converge remains an open question. However, this section does develop sufficient conditions on  $\{(\mu, b, c)\}$  to guarantee the convergence of  $\{x(\mu, b, c)\}$ . These conditions show that the cost coefficient vector need not converge. An example at the end of this section indicates further difficulties in the quest for establishing exactly when  $\{x(\mu, b, c)\}$  converges.

The sufficient conditions presented for  $\{(\mu, b, c)\}$  to guarantee the convergence of  $\{x(\mu, b, c)\}$  required that  $\mathcal{G}_c$  is partitioned into equivalence classes.

For any  $b \in \mathcal{G}_b$ , we say that  $c^1 \in \mathcal{G}_c$  and  $c^2 \in \mathcal{G}_c$  are “ $A$ -similar”, denoted  $c^1 \stackrel{A}{\sim} c^2$ , if

$$PCP_{(b,c^1)} \cap PCP_{(b,c^2)} \neq \emptyset.$$

The first goal of this section is to show that  $\stackrel{A}{\sim}$  is an equivalence relation on  $\mathcal{G}_c$ . To show this, the property that central paths may not intersect unless they are equal is proven. The first lemma gives sufficient conditions for the primal central paths to be equivalent.

**Lemma 4.20** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . Define

$$\begin{aligned} c_0^1 &= \text{proj}_{\text{null}(A)} c^1 \text{ and} \\ c_0^2 &= \text{proj}_{\text{null}(A)} c^2. \end{aligned}$$

Then,

$$\begin{aligned} PCP_{(b,c_0^1)} &= PCP_{(b,c^1)} \text{ and} \\ PCP_{(b,c_0^2)} &= PCP_{(b,c^2)}. \end{aligned}$$

Furthermore, if  $c_0^1 = \alpha c_0^2$ , for some  $\alpha \in \mathbb{R}_{++}$ ,

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}.$$

**Proof:** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . Also, let

$$\begin{aligned} c_R^1 &= \text{proj}_{\text{row}(A)} c^1 \text{ and} \\ c_R^2 &= \text{proj}_{\text{row}(A)} c^2, \end{aligned}$$

so that  $c^1 = c_0^1 + c_R^1$  and  $c^2 = c_0^2 + c_R^2$ . Let  $\alpha \in \mathbb{R}_{++}$  be such that  $c_0^1 = \alpha c_0^2$ .

Since  $c_R^1$  and  $c_R^2$  are perpendicular to  $\text{null}(A)$ ,  $c_R^1 x = c_R^2 x = 0$ , for all  $x \in \text{null}(A)$ . Consequently, theorem 2.10 shows that  $c_R^1 x$  and  $c_R^2 x$  are constant on  $\mathcal{P}_b$ . This means that  $x(\mu, b, c^1)$  is the unique solution to

$$\min\{c_0^1 x - \mu \sum_{i=1}^n \omega_i \ln(x_i) : x \in \mathcal{P}_b^o\}$$

and  $x(\mu, b, c^2)$  is the unique solution to

$$\min\{c_0^2 x - \mu \sum_{i=1}^n \omega_i \ln(x_i) : x \in \mathcal{P}_b^o\}.$$

Hence,

$$PCP_{(b, c_0^1)} = PCP_{(b, c^1)} \text{ and } PCP_{(b, c_0^2)} = PCP_{(b, c^2)}.$$

Multiplying the objective function of the first math program by  $\alpha$ , shows that

$$x(\mu, b, c^1) = x(\alpha\mu, b, c^2), \tag{4.8}$$

which implies,

$$PCP_{(b, c^1)} = PCP_{(b, c^2)}.$$

■

The following corollary is stated for future reference.

**Corollary 4.21** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . If

$$\text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c^2,$$

for some  $\alpha \in \mathbb{R}_{++}$ , then

$$x(\mu, b, c^1) = x(\mu, b, \text{proj}_{\text{null}(A)} c^1) = x(\alpha\mu, b, \alpha \text{proj}_{\text{null}(A)} c^1) = x(\alpha\mu, b, c^2).$$

**Proof:** The result is immediate from the proof of lemma 4.20. ■

The next theorem establishes the result that for any polyhedron, the central paths defined in this polyhedron are either the same or disjoint. Since the definition of  $PCP_r$  is concerned only with elements corresponding to positive  $\mu$  values, this does not say that two different central paths may not terminate at the same point. However, it does say that two different central paths may not cross en-route to either  $x^*(r)$  or  $\bar{x}(b)$ .

**Theorem 4.22** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . If there exists  $\mu^1$  and  $\mu^2$  in  $\mathbb{R}_{++}$  such that  $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$ , then

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}.$$

**Proof:** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . Lemma 4.20 implies that there is no loss of generality by assuming that  $c^1$  and  $c^2$  are in  $\text{null}(A)$ . Let  $\mu^1$  and  $\mu^2$  be in  $\mathbb{R}_{++}$ , such that  $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$ . Then,

$$c^1 - \mu^1 e^T X^{-1}(\mu^1, b, c^1) - y(\mu^1, b, c^1)A = 0,$$

and

$$c^2 - \mu^2 e^T X^{-1}(\mu^2, b, c^2) - y(\mu^2, b, c^2)A = 0.$$

Multiplying the first equation by  $\frac{1}{\mu^1}$ , the second equation by  $\frac{1}{\mu^2}$ , and subtracting yields

$$\left(\frac{1}{\mu^1}c^1 - \frac{1}{\mu^2}c^2\right) = \left(\frac{1}{\mu^1}y(\mu^1, b, c^1) - \frac{1}{\mu^2}y(\mu^2, b, c^2)\right)A.$$

Since the left-hand side of the above equality is in the  $\text{null}(A)$  and the right-hand side is in the  $\text{row}(A)$ , both must be zero. Hence,

$$c^1 = \frac{\mu^1}{\mu^2}c^2,$$

and lemma 4.20 implies

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}.$$

■

Two important corollaries follow. The first is used to show that  $\overset{A}{\sim}$  is an equivalence relation, and the second is used to establish the equivalence classes associated with  $\overset{A}{\sim}$ .

**Corollary 4.23** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . Then, if  $c^1 \overset{A}{\sim} c^2$ ,

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}.$$

**Proof:** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . If  $c^1 \overset{A}{\sim} c^2$ , there exists  $\mu^1$  and  $\mu^2$  in  $\mathbb{R}_{++}$  such that,  $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$ , and the result follows immediately from theorem 4.22. ■

**Corollary 4.24** Let  $b \in \mathcal{G}_b$ ,  $c^1 \in \mathcal{G}_c$ , and  $c^2 \in \mathcal{G}_c$ . Then

$$\text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c^2,$$

for some  $\alpha \in \mathbb{R}_{++}$ , if, and only if

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}.$$

**Proof:** The sufficiency is established in lemma 4.20. The necessity follows since

$$PCP_{(b,c^1)} = PCP_{(b,c^2)}$$

implies that there exist  $\mu^1$  and  $\mu^2$  in  $\mathbb{R}_{++}$  such that  $x(\mu^1, b, c^1) = x(\mu^2, b, c^2)$ , and the proof of theorem 4.22 shows that

$$\text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c^2,$$

for some  $\alpha \in \mathbb{R}_{++}$ . ■

The next result shows that  $\overset{A}{\sim}$  is indeed an equivalence relation.

**Theorem 4.25**  $\overset{A}{\sim}$  is an equivalence relation on  $\mathcal{G}_c$ . Furthermore, the equivalence class of  $c^1 \in \mathcal{G}_c$  is,

$$[c^1] = \{c : \text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c, \text{ for some } \alpha \in \mathbb{R}_{++}\}.$$

**Proof:** Let  $b \in \mathcal{G}_b$  and  $c^1, c^2, c^3 \in \mathcal{G}_c$ . Clearly  $c^1 \stackrel{A}{\sim} c^1$ , and  $c^1 \stackrel{A}{\sim} c^2$  implies  $c^2 \stackrel{A}{\sim} c^1$ . So  $\stackrel{A}{\sim}$  is reflexive and symmetric. Corollary 4.23 implies that if  $c^1 \stackrel{A}{\sim} c^2$  and  $c^2 \stackrel{A}{\sim} c^3$ ,

$$PCP_{(b,c^1)} = PCP_{(b,c^2)} = PCP_{(b,c^3)},$$

which implies  $c^1 \stackrel{A}{\sim} c^3$ . Hence  $\stackrel{A}{\sim}$  is transitive and is an equivalence relation. Theorem 4.22 and corollary 4.24 imply that the equivalence classes are as stated. ■

Notice that  $\text{row}(A)$  is equivalent to  $[0]$ . This equivalence class is problematic as demonstrated in example 4.30.

To demonstrate that the cost vector need not converge for  $\{x(\mu, b, c)\}$  to converge, two new types of convergence are defined. The first type of new convergence is called *class convergence*. To define this concept adequately, some notation is introduced. For a sequence  $\{x^k\}$ , let  $\mathbf{C}(\{x^k\})$  be the set of cluster points of  $\{x^k\}$ . Furthermore, for any sequence  $\{c^k \in \mathcal{G}_c\}$ , set

$$\mathcal{F}(\{c^k\}) \equiv \mathbf{C}(\{c^k\}) \cup \mathbf{C}\left(\left\{\frac{c^k}{\|c^k\|} : c^k \neq 0\right\}\right).$$

Notice that  $\mathcal{F}$  contains all of the limiting directions of a sequence of cost vectors. The concept behind class convergence is that all of these limiting directions are contained in the same equivalence class.

**Definition 4.26** The sequence,  $\{c^k \in \mathcal{G}_c\}$ , is class convergent to  $[\hat{c}]$ ,

for some  $\hat{c} \in \mathcal{G}_c$ , if  $\mathcal{F}(\{c^k\}) \subseteq [\hat{c}]$ .

The definition of class convergence does not imply that the sequence  $\{c^k\}$  actually converges; however, if  $\{c^k\}$  does converge then it is also class convergent.

The second type of convergence is called *proportional convergence*.

The idea is that the sequence  $\{(\mu^k, c^k)\}$  is proportionately convergent if the sets  $\mathbf{C}\left(\left\{\frac{\mu^k}{\|c^k\|} : c^k \neq 0\right\}\right)$  and  $\mathbf{C}\left(\{c^k\}\right)$  are, in some sense, proportional.

**Definition 4.27** The sequence,  $\{(\mu^k, c^k)\} \subset \mathbb{R}_{++} \times \mathcal{G}_c$ , is proportionately convergent if, whenever two subsequences of  $\{c^k\}$ , say  $\{c^{k_i}\}$  and  $\{c^{k_j}\}$ , have the properties that,

- (1)  $\lim_{i \rightarrow \infty} \frac{c^{k_i}}{\|c^{k_i}\|} = c^1$ ,
- (2)  $\lim_{j \rightarrow \infty} \frac{c^{k_j}}{\|c^{k_j}\|} = c^2$ , and
- (3)  $\text{proj}_{\text{null}(A)} c^1 = \alpha \text{proj}_{\text{null}(A)} c^2$ , for some  $\alpha \in \mathbb{R}_{++}$ ,

we also have

$$\lim_{i \rightarrow \infty} \frac{\mu^{k_i}}{\|c^{k_i}\|} = \alpha \lim_{j \rightarrow \infty} \frac{\mu^{k_j}}{\|c^{k_j}\|}.$$

The next lemma is a relaxation of theorem 2.14, and provides less stringent conditions for the boundedness of the sequence  $\{x(\mu, b, c)\}$ .

**Lemma 4.28** Let  $W$  be a closed subset of  $\mathcal{G}$  and  $\{(\mu^k, r^k)\} \subset \mathbb{R}_{++} \times W$  be a sequence with the following properties:

- (1)  $\left\{\frac{\mu^k}{\|c^k\|}\right\}$  is bounded, and

(2)  $\{b^k\}$  converges to  $\bar{b} \in \mathcal{G}_b$ .

Then  $\{x(\mu^k, b^k, c^k)\}$  is bounded.

**Proof:** Let  $\left| \frac{\mu^k}{\|c^k\|} \right| \leq M$ . Suppose, for the sake of attaining a contradiction, that  $\{x(\mu^k, b^k, c^k)\}$  is not bounded. Without loss in generality assume that

$$\lim_{k \rightarrow \infty} \|x(\mu^k, b^k, c^k)\| = \infty.$$

Let  $\left\{ \frac{c^{k_i}}{\|c^{k_i}\|} \right\}$  be a convergent subsequence of  $\left\{ \frac{c^k}{\|c^k\|} \right\}$ . Because

$$S \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right) x \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right) = \omega \frac{\mu^{k_i}}{\|c^{k_i}\|} \leq \omega M,$$

the following inner product inequality holds:

$$s \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right) x \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right) \leq e^T \omega M.$$

This implies

$$\left\{ \left( x \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right), \left( y \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right), s \left( \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right) \right) \right) \right\}$$

is a subset of

$$\bigcup_{k_i} \mathcal{L} \left( \left( b^{k_i}, \frac{c^{k_i}}{\|c^{k_i}\|} \right), e^T \omega M \right).$$

By assumption, any cluster point of  $\left\{ \frac{c^k}{\|c^k\|} \right\}$  is an element of  $\mathcal{G}_c$ . Hence, theorem 2.14 implies the above union is bounded. Since

$$x(\mu^k, b^k, c^k) = x \left( \frac{\mu^k}{\|c^k\|}, b^k, \frac{c^k}{\|c^k\|} \right),$$

we have a contradiction, and the result follows. ■

Theorem 4.29 gives sufficient conditions for  $\{(\mu^k, b^k, c^k)\}$  to guarantee the convergence of  $\{x(\mu^k, b^k, c^k)\}$  to an element of a central path. These conditions do not include that  $\{c^k\}$  converges, but instead that  $\{c^k\}$  is class convergent. As example 4.30 shows, this condition is still too restrictive for necessity. Characterizing the conditions of  $\{(\mu^k, b^k, c^k)\}$  for which  $\{x(\mu^k, b^k, c^k)\}$  converges remains an open question. Such a classification is needed before generalizing the set convergence results of the previous two sections.

**Theorem 4.29** Let  $\bar{r} \in \mathcal{G}$ , and assume  $\{c^k \in \mathcal{G}_c\}$  is a sequence such that  $\mathbf{C}(\{c^k\}) \subseteq \mathcal{G}_c \setminus \text{row}(A)$ . Then  $\{x(\mu^k, b^k, c^k)\} \rightarrow \hat{x} \in PCP_{(\bar{b}, \bar{c})}$  if,

- (1)  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ ,
- (2)  $\{c^k \in \mathcal{G}_c\}$  is class convergent to  $[\bar{c}]$ ,
- (3)  $\{(\mu^k, c^k) : c^k \neq 0\}$  is proportionately convergent, and
- (4) there exists  $M \in \mathbb{R}_{++}$  such that  $\frac{1}{M} \leq \left\{ \frac{\mu^k}{\|c^k\|} \right\} \leq M$ .

**Proof:** Because  $\bar{c} \notin \text{row}(A)$ , the sequence  $\{c^k\}$  contains zero at most a finite number of times. So, without loss in generality, zero is assumed not to be contained in  $\{c^k\}$ . The result is established by showing that all cluster points of  $\{x(\mu^k, b^k, c^k)\}$  are equal. Since  $\mathbf{C}(\{c^k\}) \subseteq \mathcal{G}_c \setminus \text{row}(A)$  and  $\{b^k \in \mathcal{G}_b\} \rightarrow \bar{b} \in \mathcal{G}_b$ , there exists a closed set  $W \subset \mathcal{G}$  such that  $\{(c^k, b^k)\} \subset W$ . Because of the

assumption that  $\left\{\frac{\mu^k}{\|c^k\|}\right\}$  is bounded,  $\{x(\mu^k, b^k, c^k)\}$  is bounded from lemma 4.28.

Hence, there exist subsequences such that

$$\{x(\mu^{k_i}, b^{k_i}, c^{k_i})\} \rightarrow \hat{x}^1,$$

$$\{x(\mu^{k_j}, b^{k_j}, c^{k_j})\} \rightarrow \hat{x}^2,$$

$$\left\{\frac{c^{k_i}}{\|c^{k_i}\|}\right\} \rightarrow \hat{c}^1, \text{ and}$$

$$\left\{\frac{c^{k_j}}{\|c^{k_j}\|}\right\} \rightarrow \hat{c}^2.$$

Clearly,  $\hat{x}^1$  and  $\hat{x}^2$  are in  $\mathcal{P}_{\bar{b}}$ . The class convergence implies the existence of  $\alpha^1, \alpha^2 \in \mathbb{R}_{++}$  such that,

$$\begin{aligned} \lim_{i \rightarrow \infty} \alpha^1 \text{proj}_{\text{null}(A)} \frac{c^{k_i}}{\|c^{k_i}\|} &= \alpha^1 \text{proj}_{\text{null}(A)} \hat{c}^1 \\ &= \text{proj}_{\text{null}(A)} \bar{c} \\ &= \alpha^2 \text{proj}_{\text{null}(A)} \hat{c}^2 \\ &= \lim_{j \rightarrow \infty} \alpha^2 \text{proj}_{\text{null}(A)} \frac{c^{k_j}}{\|c^{k_j}\|}. \end{aligned}$$

The proportional convergence assumption, together with the assumption that

$\left\{\frac{\mu^k}{\|c^k\|}\right\}$  is bounded away from zero, implies

$$0 < \bar{\mu} = \lim_{i \rightarrow \infty} \alpha^1 \frac{\mu^{k_i}}{\|c^{k_i}\|} = \lim_{j \rightarrow \infty} \alpha^2 \frac{\mu^{k_j}}{\|c^{k_j}\|}.$$

From corollary 4.21,

$$\begin{aligned} x(\mu^{k_i}, b^{k_i}, c^{k_i}) &= x\left(\alpha^1 \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \alpha^1 \text{proj}_{\text{null}(A)} \frac{c^{k_i}}{\|c^{k_i}\|}\right) \text{ and} \\ x(\mu^{k_j}, b^{k_j}, c^{k_j}) &= x\left(\alpha^2 \frac{\mu^{k_j}}{\|c^{k_j}\|}, b^{k_j}, \alpha^2 \text{proj}_{\text{null}(A)} \frac{c^{k_j}}{\|c^{k_j}\|}\right). \end{aligned}$$

This and theorem 2.6 yield,

$$\begin{aligned}
\hat{x}^1 &= \lim_{i \rightarrow \infty} x(\mu^{k_i}, b^{k_i}, c^{k_i}) \\
&= \lim_{i \rightarrow \infty} x\left(\alpha^1 \frac{\mu^{k_i}}{\|c^{k_i}\|}, b^{k_i}, \alpha^1 \text{proj}_{\text{null}(A)} \frac{c^{k_i}}{\|c^{k_i}\|}\right) \\
&= x(\bar{\mu}, \bar{b}, \bar{c}) \\
&= \lim_{j \rightarrow \infty} x\left(\alpha^2 \frac{\mu^{k_j}}{\|c^{k_j}\|}, b^{k_j}, \alpha^2 \text{proj}_{\text{null}(A)} \frac{c^{k_j}}{\|c^{k_j}\|}\right) \\
&= \lim_{j \rightarrow \infty} x(\mu^{k_j}, b^{k_j}, c^{k_j}) \\
&= \hat{x}^2.
\end{aligned}$$

Hence,  $\{x(\mu^k, b^k, c^k)\}$  converges to an element in  $PCP_{(\bar{b}, \bar{c})}$ . ■

Unlike theorems 4.5 and 4.14, the conditions required in theorem 4.29 are much too restrictive for necessity. The following example has the desirable property that  $\{c^k\}$  converges; but, even with this property, the convergence of  $\{x(\mu, b, c)\}$  requires the analysis of several nested linear programs. The example hints at what might be required to obtain the elusive necessary and sufficient conditions.

**Example 4.30** Let  $\mu^k = \frac{1}{k}$ , and consider the following linear programming

data for the linear program  $\min\{cx : Ax \leq b, x \geq 0\}$ ,

$$c^k = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{\sqrt{k}} \\ \frac{1}{\sqrt{k}} \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ and } A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

To analyze the behavior of  $\{x(\mu^k, b, c^k)\}$ , a sequence of problems is considered.

The idea is to iteratively reduce the original problem. This is done using the results for linear changes in the cost coefficients found in section 4.2, and at each step identifying an increasingly larger subset of variables that must be zero.

The ‘root’ problem uses the limiting rim data. So, zero is the cost vector of the root problem since  $\{c^k\} \rightarrow 0 \equiv \bar{c}[0]$ , and the root problem is

$$\min\{0x : Ax \leq b, x \geq 0\}.$$

The optimal partition for the root problem is  $B[0] = \{1, 2, 3\}$  and  $N[0] = \emptyset$ .

To define the first ‘subproblem’, set

$$\begin{aligned} \tau^k[1] &= \|c^k[0] - \bar{c}[0]\| = \frac{\sqrt{2k+1}}{k}, \\ \mu^k[1] &= \frac{\mu^k[0]}{\tau^k[1]} = \frac{1}{\sqrt{2k+1}}, \\ b^k[1] &= b[0] - A_{N[0]}x_{N[0]}(\mu^k[0], b[0], c^k[0]), \text{ and} \\ c^k[1] &= \frac{c^k[0] - \bar{c}[0]}{\tau^k[1]} = \frac{1}{\sqrt{2k+1}} \begin{pmatrix} 1 & \sqrt{k} & \sqrt{k} \end{pmatrix}, \end{aligned}$$

where the  $[0]$  indicates the original sequence data. With this notation,

$$x(\mu^k, b, c^k) = x(\mu^k, b, \bar{c}[0] + \tau^k[1]c^k[1]).$$

Notice that the sequence  $\{c^k[1]\}$  no longer converges to zero, but rather

$$\lim_{k \rightarrow \infty} c^k[1] = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Furthermore,  $\{b^k[1]\} \rightarrow b$  because  $N[0] = \emptyset$ . However, lemma 2.15 would imply this even when  $N[0] \neq \emptyset$ . Defining

$$\lim_{k \rightarrow \infty} c^k[1] = \bar{c}[1],$$

the first subproblem is

$$\min\{\bar{c}[1]z[1] : A_{B[0]}z[1] \leq b, z[1] \geq 0\},$$

or equivalently

$$\min\left\{\frac{1}{\sqrt{2}}z_2[1] + \frac{1}{\sqrt{2}}z_3[1] : 0 \leq z_i[1] \leq 1, i = 1, 2, 3\right\}.$$

The optimal partition for this problem is  $B[1] = \{1\}$  and  $N[1] = \{2, 3\}$ . More importantly, analogous to 4.5,

$$\begin{aligned} x_{B[0]}(\mu^k[0], b[0], c^k[0]) &= x_{B[0]}(\mu^k[0], b[0], \bar{c}[0] + \tau^k[1]c^k[1]) \\ &= z[1] \left( \frac{\mu^k[0]}{\tau^k[1]}, b^k[1], c^k[1] \right) \\ &= z[1] (\mu^k[1], b^k[1], c^k[1]). \end{aligned}$$

Similar to the first subproblem, the second ‘subproblem’ relies on

$$\begin{aligned}\tau^k[2] &= \|c^k[1] - \bar{c}[1]\| \\ \mu^k[2] &= \frac{\mu^k[1]}{\tau^k[2]} \\ b^k[2] &= b[1] - A_{N[1]}x_{N[1]}(\mu^k[1], b[1], c^k[1]), \text{ and} \\ c^k[2] &= \frac{c^k[1] - \bar{c}[1]}{\tau^k[2]},\end{aligned}$$

where  $\{c^k[2]\}$  is defined only for the elements in  $B[1]$  (so only the first coordinate of the original  $c$  remains for this subproblem). This notation allows,

$$z[1](\mu^k[1], b^k[1], c^k[1]) = z[1](\mu^k[1], b^k[1], \bar{c}[1] + \tau^k[2]c^k[2]).$$

It is easily checked that

$$\begin{aligned}\lim_{k \rightarrow \infty} \mu^k[2] &= 1 \text{ and} \\ \bar{c}[2] \equiv \lim_{k \rightarrow \infty} c^k[2] &= 1.\end{aligned}$$

The most important observation here is that  $\{\mu^k[2]\}$  does *not* converge to zero.

The second subproblem is

$$\min\{\bar{c}[2]z[2] : A_{B[1]}z[2] \leq b, z[2] \geq 0\},$$

which is same as

$$\min\{z[2] : 0 \leq z[2] \leq 1\}.$$

Using the 'z' notation recursively, 4.5 shows

$$\begin{aligned}
x_{B[1]}(\mu^k[0], b[0], c^k[0]) &= x_{B[1]}(\mu^k[0], b[0], \bar{c}[0] - \tau^k[1]c^k[1]) \\
&= z_{B[1]}[1] \left( \frac{\mu^k[0]}{\tau^k[1]}, b^k[1], c^k[1] \right) \\
&= z_{B[1]}[1] \left( \mu^k[1], b^k[1], c^k[1] \right) \\
&= z_{B[1]}[1] \left( \mu^k[1], b^k[1], \bar{c}[1] + \tau[2]^k c^k[2] \right) \\
&= z[2] \left( \frac{\mu^k[1]}{\tau^k[2]}, b^k[2], c^k[2] \right) \\
&= z[2] \left( \mu^k[2], b^k[2], c^k[2] \right).
\end{aligned}$$

Because each of these problems are variants of example 2.12, the convergence of each problem is easily examined. The central path for the original problem is defined by

$$x(\mu^k[0], b[0], c^k[0]) = \begin{pmatrix} \frac{\frac{1}{k} + \frac{2}{k} - \sqrt{\frac{1}{k^2} + \frac{4}{k^2}}}{\frac{2}{k}} \\ \frac{\frac{1}{\sqrt{k}} + \frac{2}{k} - \sqrt{\frac{1}{k} + \frac{4}{k^2}}}{\frac{2}{\sqrt{k}}} \\ \frac{\frac{1}{\sqrt{k}} + \frac{2}{k} - \sqrt{\frac{1}{k} + \frac{4}{k^2}}}{\frac{2}{\sqrt{k}}} \end{pmatrix} = \begin{pmatrix} \frac{3 - \sqrt{5}}{2} \\ \frac{1 + \frac{2}{\sqrt{k}} - \sqrt{1 + \frac{4}{k}}}{2} \\ \frac{1 + \frac{2}{\sqrt{k}} - \sqrt{1 + \frac{4}{k}}}{2} \end{pmatrix}.$$

Hence,

$$\lim_{k \rightarrow \infty} x(\mu^k, b, c^k) = \begin{pmatrix} \frac{3 - \sqrt{5}}{2} \\ 0 \\ 0 \end{pmatrix}.$$

This seems odd since the central path for the first root problem is  $\mathcal{P}_{(b,0)} = \{\frac{1}{2}e\}$ .

The problem is that  $\{\mu^k[0]\} = \{\frac{1}{k}\} \rightarrow 0^+$ . If the change in the cost coefficients

had been linear, theorems 4.5 and 4.8 would have applied to deduce that the limit of  $\{x(\mu^k[0], b[0], c^k[0])\}$  is an element of a central path defined on the optimal face of the root problem. The limit point of  $\{c^k[1]\}$  attempts to define this central path, even though the change is nonlinear.

Although  $\bar{c}[1]$  does not actually indicate this central path, lemma 2.15 implies that any components that are zero in the solution to the first subproblem, which uses  $\bar{c}[1]$  as the cost vector, are also zero in any cluster point of  $\{x(\mu^k[0], b[0], c^k[0])\}$ . The elements of the central path for the first subproblem are

$$\begin{aligned}
x_{B[0]}(\mu^k[0], b[0], c^k[0]) &= z[1](\mu^k[1], b[1], c^k[1]) \\
&= \begin{pmatrix} \frac{\frac{1}{\sqrt{2k+1}} + \frac{2}{\sqrt{2k+1}} - \sqrt{\frac{1}{2k+1} + \frac{4}{2k+1}}}{\frac{2}{\sqrt{2k+1}}} \\ \frac{\frac{1}{\sqrt{2k+1}} + \frac{2}{\sqrt{2k+1}} - \sqrt{\frac{k}{2k+1} + \frac{4}{2k+1}}}{2\sqrt{\frac{k}{2k+1}}} \\ \frac{\frac{1}{\sqrt{2k+1}} + \frac{2}{\sqrt{2k+1}} - \sqrt{\frac{k}{2k+1} + \frac{4}{2k+1}}}{2\sqrt{\frac{k}{2k+1}}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{3-\sqrt{5}}{2} \\ \frac{1}{2} \left( 1 + \frac{2}{\sqrt{k}} - \sqrt{1 + \frac{4}{k}} \right) \\ \frac{1}{2} \left( 1 + \frac{2}{\sqrt{k}} - \sqrt{1 + \frac{4}{k}} \right) \end{pmatrix}.
\end{aligned}$$

Once again, this shows

$$\lim_{k \rightarrow \infty} x(\mu^k, b, c^k) = \begin{pmatrix} \frac{3-\sqrt{5}}{2} \\ 0 \\ 0 \end{pmatrix}.$$

However, the problem of  $\{\mu^k[1]\} \rightarrow 0^+$  re-occurs in the first subproblem. Notice that none of the variables become zero in the root problem, and that the second and third variables become zero in the first subproblem. This is good news, since these are all the variables that must become zero. Once these variables are identified, the parameter for the second subproblem may not converge to zero, i.e.  $\{\mu^k[2]\} \not\rightarrow 0$ . Theorem 2.6 shows that

$$\begin{aligned} \lim_{k \rightarrow \infty} x_1(\mu^k[0], b^k[0], c^k[0]) &= \lim_{k \rightarrow \infty} z[2](\mu^k[2], b^k[2], c^k[2]) \\ &= z[2](1, b, \bar{c}[2]) \\ &= z[2](1, b, 1) \\ &= \frac{1 + 2 - \sqrt{1 + 4}}{2} \\ &= \frac{3 - \sqrt{5}}{2}. \end{aligned}$$

If the original problem were not so simple, the challenge of directly evaluating the limit point of  $\{x(\mu^k[0], b, c^k[0])\}$  is substantial. However, calculating the limit point of  $\{c^k[0]\}$  allows the formation of the root problem. This produces an optimal partition, which shows some of the variables that must go to zero.

If  $\{\mu^k[1]\}$  does not converge to zero, theorem 2.6 is used to calculate the components that are positive. If  $\{\mu^k[1]\}$  does converge to zero, the limit point of  $\{c^k[1]\}$  is calculated, and the first subproblem is solved. The process repeats until either all variables are forced to zero, or  $\{\mu^k[i]\}$  does not converge to zero.

## 4.5 Chapter Summary

The main result characterizes when

$$\{x(\mu, b, c_\tau)\} \text{ and } \{(y(\mu, b_\rho, c), s(\mu, b_\rho, c))\}$$

converge. The generality of allowing simultaneous changes in the rim data, so long as the objective function perturbation is linear, is obtained from a modeling scheme that shows how to model the changes in both rim data elements as only right-hand side perturbations. The new model retains  $\delta_{Bx_B}$  as a term in the objective function, and as a consequence the set convergence property for the omega central path shows that the central path converges to the union of the original central path and a central path on the optimal set defined by  $\delta_{Bx_B}$ .

The last section demonstrated the difficulty of extending these results when the restriction of linearity is removed from the objective change. Sufficient conditions are given to guarantee the convergence of  $\{x(\mu, b, c)\}$ , and

these conditions show that it is not necessary for the cost coefficients to converge. In support of these sufficient conditions, the collection of central paths was shown to induce an equivalence relation on  $\mathcal{G}_c$ . An interesting, unanswered question is whether  $\bigcup_{c \in \mathcal{G}_c} PCP_{(b,c)} = \mathcal{P}_b^o$ .

Example 4.30 alludes to an algorithm that is capable of finding the limit of  $\{x(\mu, b, c)\}$ , provided the limit exists. The algorithm is greedy in the sense that it creates an increasingly larger set of variables that are forced to zero. Finding conditions for which this algorithm converge seems equivalent to finding the necessary conditions for the convergence of  $\{x(\mu, b, c)\}$ .

## 5. Conclusion and Avenues for Future Research

The immediate question from chapter 3 is whether or not the unique dual assumption is removable from theorem 3.15. This assumption essentially provides a rank argument, and this assumption may be found unnecessary upon a closer examination of  $\chi_{AB}$ .

Many questions remain open from chapter 4; the most intriguing being the characterization of when  $x(\mu, b, c)$  converges. These conditions appear equivalent to establishing when the algorithm hinted at in example 4.30 converges. Another question from this chapter is, “Does the collection of central paths cover the relative interior of the polyhedron?” If this is true, every interior element of a polyhedron is associated with an equivalence class of cost vectors.

Several results from this work are finding themselves useful elsewhere. The parametric analysis has been experimentally used to find radiation plans in the field of radiation oncology. A question that has arisen from this work is whether or not the parametric analysis presented here may be used to find minimal support sets. More recently, R. Caron, H. Greenberg, and A. Holder

are working on the development of a theory of repelling constraints. Several of the techniques and insights gained from this work have proven themselves worthy in this endeavour.

## NOTATION INDEX

- $(B|N)$  is the optimal partition for  $LP_r$
- $B = \{i : x_i^*(r) > 0\}$
- $b_\rho = b + \rho\delta b$
- $C(\{x^k\})$  is the set of cluster points of  $\{x^k\}$
- $\text{col}(A) = \{z : Ax = z \text{ for some } x\}$
- $CP_r = \{(x(\mu, b, c), y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}$
- $\overline{CP}_r = \overline{PCP}_r \times \overline{DCP}_r$
- $c_\tau = c + \tau\delta c$
- $DCP_r = \{(y(\mu, b, c), s(\mu, b, c)) : \mu > 0\}$
- $\overline{DCP}_r = \begin{cases} DCP \cup \{(y^*(r), s^*(r))\} & \text{if } \mathcal{D}_c \text{ is unbounded} \\ DCP \cup \{(y^*(r), s^*(r)), (\bar{y}(r), \bar{s}(r))\} & \text{if } \mathcal{D}_c \text{ is bounded} \end{cases}$
- $DCP_{(r, \delta b)}^* = \{(p(\eta, r, \delta b), (0, q(\eta, r, \delta b))) : \eta > 0\}$
- $\overline{DCP}_{(r, \delta b)}^* = DCP_{(r, \delta b)}^* \cup \{(y^*(r), s^*(r)), (p^*(r, \delta b), (0, q^*(r, \delta b)))\}$
- $D_{x_i} f(x^0) = \lim_{\theta \rightarrow 0} \frac{f(x^0 + \theta \delta x) - f(x^0)}{\theta}$ , where  $\|\delta x\| = 1$
- $D_{x_i}^k f(x^0)$  is  $k$  applications of the derivative operator on  $f$  with respect to  $x_i$
- $D_{x_i}^k f(x^0+) = \lim_{x \rightarrow x^0+} D_{x_i}^k f(x)$

- $D_{x_i+}^k f(x^0) = \lim_{\theta \rightarrow 0+} \frac{f(x^0 + \theta \delta x_i) - f(x^0)}{\theta}$ , where  $\delta x_i$  is the  $i$ th unit vector
- $\mathcal{D}_c = \{(y, s) : yA + s = c, s \geq 0\}$
- $\mathcal{D}_c^o = \{(y, s) : yA + s = c, s > 0\}$
- $\mathcal{D}_r^* = \{(y, s) \in \mathcal{D}_c : yb \text{ is optimal to } LD\}$
- $\delta b$  is a direction of change for  $b$  in  $LP_r$
- $\delta c$  is a direction of change for  $c$  in  $LP_r$
- $\delta r$  is a direction of change for  $b$  and  $c$  in  $LP_r$
- $e = (1, 1, 1, \dots, 1)^T$
- $\mathcal{F}(\{c^k\}) \equiv \mathbf{C}(\{c^k\}) \cup \mathbf{C}\left(\left\{\frac{c^k}{\|c^k\|} : c^k \neq 0\right\}\right)$
- $\mathcal{G} = \{r : \mathcal{P}_b^o \neq \emptyset, \mathcal{D}_c^o \neq \emptyset\}$
- $\mathcal{G}_b = \{b : \mathcal{P}_b^o \neq \emptyset\}$
- $\mathcal{G}_c = \{c : \mathcal{P}_c^o \neq \emptyset\}$
- $\mathcal{H} = \{\delta r \in \mathcal{H} : \text{the optimal partition is invariant on } [r, r + \theta \delta r]$   
for some sufficiently small positive  $\theta\}$
- $\mathcal{H}_b = \{\delta b : (\delta b, 0) \in \mathcal{H}\}$
- $\mathcal{H}_c = \{\delta c : (0, \delta c) \in \mathcal{H}\}$
- $\text{leftnull}(A) = \{y : yA = 0\}$
- $\mathcal{L}(r, M) = \{(x, y, s) \in \mathcal{P}_b \times \mathcal{D}_c : sx \leq M\}$
- $LP_r \quad \min\{cx : Ax = b, x \geq 0\}$
- $LD_r \quad \max\{yb : yA + s = c, s \geq 0\}$

- $\text{null}(A) = \{x : Ax = 0\}$
- $N = \{i : x_i^*(r) = 0\}$
- $\Omega = \text{diag}(\omega)$  where  $\omega \in \mathbb{R}^n$
- $\text{opt}(b)$  is the optimal objective function value of  $LP$ , relative to the right-hand side  $b$
- $(p(\eta, r, \delta), q(\eta, r, \delta))$  is the unique solution to
 
$$\max\{p\delta + \eta \sum_{i \in N} \omega_i \ln(q_i) : pA_B = c_B, pA_N + q = c_N, q \geq 0\}$$
- $(\bar{p}(r), \bar{q}(r)) = \lim_{\eta \rightarrow \infty} (p(\eta, r, \delta), q(\eta, r, \delta))$
- $(p^*(r, \delta), q^*(r, \delta)) = \lim_{\eta \rightarrow 0} (p(\eta, r, \delta), q(\eta, r, \delta))$
- $\mathcal{P}_b = \{x : Ax = b, x \geq 0\}$
- $\mathcal{P}_b^o = \{x : Ax = b, x > 0\}$
- $\mathcal{P}_r^* = \{x \in \mathcal{P}_b : cx \text{ is optimal to } LP\}$
- $(\mathcal{P}_r^*)^o = \{x \in \mathcal{P}_r^* : x_B > 0\}$
- $PCP_r = \{x(\mu, b, c) : \mu > 0\}$
- $\overline{PCP}_r = \begin{cases} PCP_r \cup \{x^*(r)\} & \text{if } \mathcal{P}_b \text{ is unbounded} \\ PCP_r \cup \{x^*(r), \bar{x}(b)\} & \text{if } \mathcal{P}_b \text{ is bounded} \end{cases}$
- $PCP_{(r, \delta)}^* = \{(z_B(\eta, b, \delta c), 0) : \eta > 0\}$
- $\overline{PCP}_{(r, \delta)}^* = PCP_{(r, \delta)} \cup \{x^*(r), (z^*(b, \delta c), 0)\}$
- $\text{row}(A) = \{z : z = yA \text{ for some } y\}$
- $\mathbb{R}^n = \{(x_1, x_2, x_3, \dots, x_n) : x_i \in \mathbb{R}, i = 1, 2, 3, \dots, n\}$

- $\mathbb{R}_+^n = \{(x_1, x_2, x_3, \dots, x_n) : x_i \in \mathbb{R}, x_i \geq 0, i = 1, 2, 3, \dots, n\}$
- $\mathbb{R}_{++}^n = \{(x_1, x_2, x_3, \dots, x_n) : x_i \in \mathbb{R}, x_i > 0, i = 1, 2, 3, \dots, n\}$
- $r = (b, c)$
- $S = \text{diag}(s)$  where  $s \in \mathbb{R}^n$
- $\sigma(x) = \{i : x_i > 0\}$  where  $x \in \mathbb{R}_+^n$
- $X = \text{diag}(x)$  where  $x \in \mathbb{R}^n$
- $x(\mu, b, c)$  is the unique solution to
 
$$\min\{cx - \mu \sum_i^n \omega_i \ln(x_i) : Ax = b, x \geq 0\}$$
- $\bar{x}(b) = \lim_{\mu \rightarrow \infty} x(\mu, b, c)$ , if  $\mathcal{P}_b$  is bounded
- $x^*(r)$  is the analytic center of  $\mathcal{P}_r^*$
- $(y(\mu, b, c), s(\mu, b, c))$  is the unique solution to
 
$$\max\{yb + \mu \sum_i^n \omega_i \ln(s_i) : yA + s = c, s \geq 0\}$$
- $(\bar{y}(c), \bar{s}(c)) = \lim_{\mu \rightarrow \infty} (y(\mu, b, c), s(\mu, b, c))$ , if  $\mathcal{D}_c$  is bounded
- $(y^*(r), s^*(r)) = \lim_{\mu \rightarrow 0} (y(\mu, b, c), s(\mu, b, c))$
- $z_B(\eta, b, \delta c)$  is the unique solution to
 
$$\min\{\delta c_B z_B - \eta \sum_{i \in B} \omega_i \ln(z_i) : A_B z_B = b, z_B \geq 0\}$$
- $\bar{z}_B(b) = \lim_{\eta \rightarrow \infty} z(\eta, b, \delta c)$
- $z_B^*(b, \delta c) = \lim_{\eta \rightarrow 0^+} z(\eta, b, \delta c)$

## REFERENCES

- [1] I. Adler and R. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50:29–51, 1991.
- [2] I. Adler and R. Monteiro. A geometric view of parametric linear programming. *Algorithmica*, 8:161–176, 1992.
- [3] D. Atkinson and P. Vaidya. A scaling technique for finding the weighted analytic center of a polytope. *Mathematical Programming*, 57:163–192, 1992.
- [4] D. Bayer and J. Lagarias. The nonlinear geometry of linear programming I. *Transactions of the American Mathematical Society*, 314(2):499–526, 1989.
- [5] D. Bayer and J. Lagarias. The nonlinear geometry of linear programming II. *Transactions of the American Mathematical Society*, 314(2):527–581, 1989.
- [6] C. Berge. *Topological Spaces*. The McMillan Company, New York, NY, 1963.
- [7] J. Bonnans and F. Porta. Infeasible path following algorithms for linear complementarity problems. Rapport de recherche 2445, Institut National De Recherche en Informatique Et En Automatique, Rocquencourt, France, 1994.
- [8] S. Campbell and Jr. C. Meyer. *Generalized Inverses of Linear Transformations*. Fearon Pitman Publishers Inc., Belmont, CA, 1979.
- [9] G. Dantzig, J. Folkman, and N. Shapiro. On the continuity of the minimum set of a continuous function. *Journal of Mathematical Analysis and Applications*, 17:519–548, 1967.

- [10] I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 174:747–748, 1967.
- [11] I. Dikin. On the convergence of an iterative process. *Upravlyaemye Sistemy*, 12:54–60, 1974.
- [12] A. Fiacco. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Academic Press, New York, NY, 1983.
- [13] A. Fiacco and G. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968.
- [14] A. Forsgren. On linear least-squares problems with diagonally dominant weight matrices. Technical Report TRITA-MAT-1995-OS2, Royal Institute of Technology, Stockholm, Sweden, 1995.
- [15] R. Freund. Projective transformation for interior-point algorithms, and superlinearly convergent algorithm for the  $\omega$ -center problem. *Mathematical Programming*, 58:385–414, 1993.
- [16] R. Freund. Following a balanced trajectory from an infeasible point to an optimal linear programming solution with a polynomial-time algorithm. *Mathematics of Operations Research*, 21(4):839–859, 1996.
- [17] R. Freund, F. Jarre, and S. Mizuno. Convergence of a class of inexact interior-point algorithms for linear programs. Numerical Analysis Manuscript 97-3-11, Bell Laboratories, Murry Hill, New Jersey, 1997.
- [18] R. Freund and M. Nunez. Condition measures and properties of the central trajectory of a linear program. Working Paper, 1996.
- [19] T. Gal. *Postoptimal analysis, parametric programming, and related topics: degeneracy, multicriteria decision making, redundancy*. Walter de Gruyter & Co., New York, NY, 2nd edition, 1995.
- [20] S. Gass and T. Saaty. Parametric objective function (part 2) - generalization. *Operations Research*, 3(4):395–401, 1955.
- [21] J.-L. Goffin and J.-Ph. Vial. On the computation of weighted analytic centers and dual ellipsoids with the projective algorithm. *Mathematical Programming*, 60:81–92, 1993.

- [22] A. Goldman and A. Tucker. Theory of linear programming. In H. Kuhn and A. Tucker, editors, *Linear Inequalities and Related Systems*, volume 38, pages 53–97. Princeton University Press, Princeton, New Jersey, 1956.
- [23] C. Gonzaga. Path-following methods for linear programming. *SIAM Review*, 34(2):167–224, 1992.
- [24] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, NY, 1979.
- [25] H. Greenberg. *A Computer-Assisted Analysis System for Mathematical Programming Models and Solutions: A User's Guide for ANALYZE*. Kluwer Academic Publishers, Boston, MA, 1993.
- [26] H. Greenberg. Quantitative sensitivity analysis. Class notes used in Advanced Linear Programming, 1994.
- [27] H. Greenberg. The use of the optimal partition in a linear programming solution for postoptimal analysis. *Operations Research Letters*, 15(4):179–185, 1994.
- [28] H. Greenberg. Rim sensitivity analysis from an interior solution. Technical Report CCM No. 86, University of Colorado, at Denver, 1996.
- [29] H. Greenberg. An analysis of degeneracy. *Naval Logistics Research Quarterly*, 33:635–655, 86-1.
- [30] H. Greenberg. *Mathematical Programming Glossary*. World Wide Web, <http://www-math.cudenver.edu/~hgreenbe/glossary/glossary.html>, 1996-1998.
- [31] H. Greenberg, A. Holder, C. Roos, and T. Terlaky. On the dimension of the set of rim perturbations for optimal partition invariance. Technical Report CCM No. 94, University of Colorado, at Denver, 1996. To appear in *SIAM Journal of Optimization*.
- [32] B. Grunbaum. *Convex Polytopes*. Wiley-Interscience, New York, NY, 1967.

- [33] D. Den Hertog and C. Roos. A survey of search directions in interior point methods for linear programming. *Mathematical Programming*, 52(2):481–510, 1991.
- [34] W. Hogan. Point-to-set maps in mathematical programming. *SIAM Review*, 15(3):591–603, 1973.
- [35] A. Holder, J. Sturm, and S. Zhang. The analytic central path, sensitivity analysis, and parametric programming. Technical Report CCM No. 118, University of Colorado at Denver, 1997.
- [36] P. Huard. Resolution of mathematical programming with nonlinear constraints by the method of centres. In J. Abadie, editor, *Nonlinear Programming*, chapter 8, pages 209–219. John Wiley & Sons, Inc., New York, NY, 1967.
- [37] B. Jansen. *Interior Point Techniques in Optimization*. PhD thesis, Delft University of Technology, Delft, Netherlands, 1995.
- [38] B. Jansen, J.J. de Jong, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: Just be careful! Technical Report AMER.93.022, Royal/Shell Laboratories Amsterdam, Amsterdam, Netherlands, 1993.
- [39] B. Jansen, C. Roos, and T. Terlaky. An interior point approach to post optimal and parametric analysis in linear programming. Technical Report 92-21, Delft University of Technology, Faculty of Technical Mathematics and Computer Science, Delft, Netherlands, 1992 1992.
- [40] B. Jansen, C. Roos, T. Terlaky, and J.-Ph. Vial. Interior-point methodology for linear programming: duality, sensitivity analysis and computational aspects. Technical Report 93-28, Delft University of Technology, Faculty of Technical Mathematics and Computer Science, Delft, Netherlands, 1993.
- [41] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [42] L. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR*, 244:1093–1096, 1979.
- [43] M. Kojima. Basic lemmas in polynomial-time infeasible-interior-point

- methods for linear programs. *Annals of Operations Research*, 62:1–28, 1996.
- [44] M. Kojima, N. Megiddo, and S. Mizuno. A general framework of continuation methods for complementarity problems. *Mathematics of Operations Research*, 18:945–963, 1993.
- [45] M. Kojima, S. Mizuno, and A. Yoshi. A primal dual interior point method for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point Algorithms and Related Methods*. Springer-Verlag, New York, 1989.
- [46] A. Kozlov. The Karmarkar algorithm: Is it for real? *SIAM News*, 18(6):1,4, 1985.
- [47] E. Kranich. *Interior Point Method Bibliography*. World Wide Web, <http://netlib.org/bib/intbib.bib>, 1993.
- [48] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Inc., New York, NY, 1985.
- [49] I. Lustig, R. Marsten, and D. Shanno. Computational experience with a primal-dual interior point method for linear programming. Technical report, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1989.
- [50] I. Lustig, R. Marsten, and D. Shanno. Interior point methods for linear programming: Computational state of the art. *ORSA Journal on Computing*, 6:1–14, 1994.
- [51] R. Marsten, M. Saltzman, D. Shanno, G. Pierce, and J. Ballintijn. Implementation of a dual affine interior point algorithm for linear programming. *ORSA Journal on Computing*, 1(4):287–297, 1989.
- [52] R. Marsten, R. Subramanian, I. Lustig, and D. Shanno. Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick. *Interfaces*, 20(4):105–116, 1990.
- [53] L. McLinden. An analogue of Moreau’s proximation theorem, with applications to the nonlinear complementary problem. *Pacific Journal of Mathematics*, 88(1):101–161, 1980.

- [54] L. McLinden. The complementarity problem for maximal monotone multifunctions. In R. Cottle, R. Giannessi, and J. Lions, editors, *Variational Inequalities and Complementarity Problems*, pages 251–270. John Wiley & Sons, 1980.
- [55] N. Megiddo. Pathways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior-Point Algorithms and Related Methods*, pages 131–158. Springer-Verlag, New York, 1989.
- [56] H. Mills. Marginal values of matrix games and linear programs. In H. Kuhn and A. Tucker, editors, *Linear Inequalities and Related Systems*. Princeton University Press, Princeton, NJ, 1956.
- [57] S. Mizuno. A new polynomial time method for a linear complementarity problem. *Mathematical Programming*, 56:31–43, 1992.
- [58] S. Mizuno and F. Jarre. Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation. Research memorandum 605, The Institute of Statistical Mathematics, Tokyo, Japan, 1996.
- [59] S. Mizuno, M. Todd, and Y. Ye. A surface of analytic centers and primal-dual infeasible-interior point algorithms for linear programming. *Mathematics of Operations Research*, 20(1):135–162, 1995.
- [60] R. Monteiro and I. Adler. Interior path-following primal-dual algorithms, part I: Linear programming. *Mathematical Programming*, 44:27–41, 1989.
- [61] R. Monteiro and S. Mehrotra. A general parametric analysis approach and its implication to sensitivity analysis in interior point methods. *Mathematical Programming*, 72:65–82, 1996.
- [62] J. Munkres. *Analysis on Manifolds*. Addison-Wesley Publishing Company, Redwood City, CA, 1991.
- [63] O. Güler. Limiting behavior of weighted central paths in linear programming. *Mathematical Programming*, 65:347–363, 1994.
- [64] O. Güler and C. Roos and T. Terlaky and J.-Ph. Vial. A survey of the implications of the behavior of the central path for duality theory of linear programming. *Management Science*, 41(12):1922–1934, 1995.

- [65] F. Porta. A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points. *Mathematical Programming*, 67:383–406, 1994.
- [66] R. Tütüncü. An infeasible-interior-point potential-reduction algorithm for linear programming. Technical Report 1136, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1996.
- [67] J. Renegar. A polynomial-time algorithm, based on Newton’s method, for linear programming. *Mathematical Programming*, 40:59–93, 1988.
- [68] J. Renegar. Some perturbation theory for linear programming. *Mathematical Programming*, 65:73–91, 1994.
- [69] J. Renegar. Incorporating condition measures into the complexity theory of linear programming. *SIAM Journal of Optimization*, 5(3):506–524, 1995.
- [70] J. Renegar. Linear programming, complexity theory, and elementary elementary functional analysis. *Mathematical Programming*, 70(3):279–351, 1995.
- [71] C. Roos. Interior point approach to linear programming: Theory, algorithms & parametric analysis. In A. van der Burgh and J. Simonis, editors, *Topics in Engineering Mathematics*, pages 181–216. Kluwer Academic Publishers, Netherlands, 1992.
- [72] C. Roos and D. Hertog. A polynomial method of approximate weighted centers for linear programming. Technical Report 89-13, Delft University of Technology, Delft, Netherlands, 1989.
- [73] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach*. John Wiley & Sons, New York, NY, 1997.
- [74] C. Roos and J.-Ph. Vial. Interior-point methods for linear programming. Technical Report 94-77, Delft University of Technology, Delft, Netherlands, 1994.
- [75] W. Rudin. *Real and Complex Analysis*. McGraw-Hill Book Company, New York, NY, 1987.

- [76] T. Saaty. Coefficient perturbation of a constrained extremum. *Operations Research*, 7:294–302, 1959.
- [77] T. Saaty and S. Gass. Parametric objective function (part 1). *Operations Research*, 2(3):316–3–19, 1954.
- [78] A. Seifi and L. Tuncel. A constant-potential infeasible-start interior-point algorithm with computational experiments and applications. In *Computational Optimization and Applications*, pages 1–50. Kluwer Academic Publishers, Boston, MA, 1996.
- [79] E. Simantiraki and D. Shanno. An infeasible-interior-point method for linear complementarity problems. RUTCOR Research Report RRR 7-95, Rutgers Center for Operational Research, New Brunswick, NJ, 1996.
- [80] G. Sonnevend. An “analytic centre” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In A. Prekopa, J. Szelezsan, and B. Strazicky, editors, *Lecture Notes in Control and Information Sciences*, volume 84, pages 866–875. Springer-Verlag, Heidelberg, Germany, 1986.
- [81] G. Sonnevend. An implementation of the method of analytic centers. In A. Bensoussan and J. Lions, editors, *Lecture Notes in Control and Information Sciences*, volume 111, pages 297–308. Springer-Verlag, Heidelberg, Germany, 1988.
- [82] G. Sonnevend. New algorithms in convex programming based on a notion of “Centre” (for systems of analytic inequalities) and on rational extrapolation. In K. Hoffman and et al, editors, *International Series of Numerical Mathematics*, volume 84, pages 311–326. Birkhauser Verlag Basel, Germany, 1988.
- [83] G. Sonnevend. Applications of the notion of analytic center in approximation (estimation) problems. *Journal of Computational and Applied Mathematics*, 28:349–358, 1989.
- [84] G. Sonnevend and J. Stoer. Global ellipsoidal approximations and homotopy methods for solving convex analytic programs. *Applied Mathematics and Optimization*, 21:139–165, 1990.
- [85] G. Sonnevend, J. Stoer, and G. Zhao. On the complexity of following the

- central path of linear programs by linear extrapolation II. *Mathematical Programming*, 52:527–553, 1991.
- [86] B. Stewart. On scaled projections and pseudoinverses. *Linear Algebra and its Applications*, 112:189–193, 1989.
- [87] M. Todd. A Dantzig-Wolfe-like variant of Karmarkar’s interior-point linear programming algorithm. *Operations Research*, 38:1006–1018, 1990.
- [88] P. Tseng. Analysis of an infeasible interior path-following method for complementarity problems. Technical report, Department of Mathematics, University of Washington, Seattle, WA, 1997.
- [89] S. Vavasis and Y. Ye. A primal-dual interior point method whose running time depends only on the constraint matrix. *Mathematical Programming*, 74:79–120, 1996.
- [90] C. Witzgall, P. Boggs, and P. Domich. On the convergence behavior of trajectories for linear programming. *Contemporary Mathematics*, 114:161–187, 1990.
- [91] M. Wright. The interior-point revolution in constrained optimization. Technical Report 98-4-09, Bell Laboratories, Murray Hill, New Jersey, 1998.
- [92] S. Wright. An infeasible-interior-point algorithm for linear complementarity problems. *Mathematical Programming*, 67:29–52, 1994.
- [93] S. Wright. A path-following interior-point algorithm for linear and quadratic problems. *Annals of Operations Research*, 62:103–130, 1996.
- [94] S. Wright. *Interior Point Methods Online*. World Wide Web, <http://www.mcs.anl.gov/home/otc/InteriorPoint/index.html>, 1997.
- [95] S. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, PA, 1997.
- [96] S. Wright and D. Ralph. A superlinear infeasible-interior-point algorithm for monotone complementarity problems. *Mathematics of Operations Research*, 21(4):815–838, 1996.
- [97] S. Wright and Y. Zhang. A superquadratic infeasible-interior-point

method for linear complementarity problems. Research Report YZ94-03, Department of Mathematics and Statistics, University of Maryland, Baltimore County, MD, 1994.

- [98] Y. Zhang. On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem. *SIAM Journal of Optimization*, 4:208–227, 1994.
- [99] Y. Zhang. User's guide to LIPSOL. Technical report, University of Maryland Baltimore County, Department of Mathematics and Statistics, Baltimore, Maryland 21228-5398, 1995.
- [100] G. Zhao. On the relationship between the curvature integral and the complexity of path-following methods in linear programming. *SIAM Journal of Optimization*, 6(1):57–73, 1986.
- [101] G. Zhao and J. Zhu. Analytic properties of the central trajectory in interior point methods. In D. Du and J. Sun, editors, *Advances in Optimization and Approximation*, pages 362–375. Kluwer Academic Publishers, The Netherlands, 1994.