Following a "Balanced" Trajectory from an Infeasible Point to an Optimal Linear Programming Solution with a Polynomial-time Algorithm

Robert M. Freund

WP# 3453-92-MSA July, 1992

Following a "Balanced" Trajectory from an Infeasible Point to an Optimal Linear Programming Solution with a Polynomial-time Algorithm

July 1992

Robert M. Freund

Research supported in part by NSF, AFOSR, and ONR through NSF grant DMS-8920550, and in part by the MIT-NTU Collaboration Research Fund. Part of this research was performed when the author was a Visiting Scientist at the Center for Applied Mathematics, Cornell University, in 1991.

Abstract:

This paper is concerned with the problem of following a trajectory from an infeasible "warm start" solution of a linear programming problem, directly to an optimal solution of the linear programming problem. A class of trajectories for the problem is defined, based on the notion of a β-balanced solution to the "warm start" problem. Given a prespecified positive balancing constant β, an infeasible solution x is said to be β-balanced if the optimal value gap is less than or equal to β times the infeasibility gap. Mathematically, this can be written as $c^T x - z^* \leq \beta \xi^T x$, where the linear form $\xi^T x$ is the Phase I objective function. The concept of a β-balanced solution is used to define a class of trajectories from an infeasible points to an optimal solution of a given linear program. Each trajectory has the property that all points on or near the trajectory (in a suitable metric) are β-balanced. The main thrust of the paper is the development of an algorithm that traces a given β-balanced trajectory from a starting point near the trajectory to an optimal solution to the given linear programming problem in polynomial-time. More specifically, the algorithm allows for fixed improvement in the bound on the Phase I and the Phase II objectives in $O(n)$ iterations of Newton steps.

Key Words: Linear program, interior-point algorithm, polynomial-time complexity, trajectory method, Newton's method.

Running Header: Balanced Trajectory for Linear Programming

1. Introduction:

This paper is concerned with following a trajectory from an infeasible "warm start" solution of a linear programming problem, directly to an optimal solution of the linear programming problem. By a "warm start" solution, we mean a solution that is hopefully close to being feasible and is hopefully close to being optimal. Like other research on the "warm start" problem, this paper is motivated by the need in practice to solve many amended versions of the same base-case linear programming problem. In this case, it makes sense to use the optimal solution to the previous version of the problem as a "warm start" solution to the current amended version of the problem. Whereas this strategy has been successfully employed in simplex-based algorithms for solving linear programming problems, there is no guarantee that it will improve solution times, due to the inherent combinatorial nature of the simplex algorithm. However, in the case of interior-point algorithms, there is no limiting combinatorial structure, and the potential for establishing guaranteed results abounds.

Most of the research on "warm start" algorithms for solving a linear programming problem has been part of the research on combined Phase I – Phase II methods for linear programming. These methods attempt to simultaneously solve the Phase I and Phase II problem over a sequence of iterations. The starting point for these methods need not be feasible, and a "warm start" solution should serve as an excellent starting point for these methods. Most of the research on combined Phase I – Phase II methods has concentrated on potential reduction algorithms, see, e.g. de Gellinck and Vial [10], Anstreicher [1] and [2], Todd [16] and [18], Todd and Wang [17], and [6]. Typically, these approaches are based on trying to reduce the values of two potential functions, one for the Phase I problem, and a second potential function for the Phase II problem. In [9], the concept of a β- balanced point was introduced, and this allows the solution of the combined Phase I – Phase II problem with only one potential function. All of these methods achieve a complexity bound involving a fixed improvement in the two goals of attaining feasibility and attaining optimality in $O(n)$ iterations, where n is the number of inequalities in the underlying problem.

Research on trajectory-following approaches for the combined Phase I – Phase II problem has not been as successful. Gill et. al. [11] as well as Polyak [14] have studied general shifted-barrier approaches to solving the "warm start" linear programming problem by examining properties of parameterized trajectories underlying the shifted-barrier problem. In [5], complexity bounds on the use of Newton's method for solving this problem were developed, but they depend very much on the availability of very good dual information. The potential reduction algorithms for the combined Phase I – Phase II problem have not in general lent themselves to analyzing trajectories, since there is no one natural trajectory associated with the two potential functions.

In this paper, we develop a class of trajectories for the combined Phase I – Phase II problem that borrows heavily from the development of the notion of a β-balanced infeasible solution from [9]. Given a prespecified positive balancing constant β, an infeasible solution x is said to be β-balanced if the optimal value gap is less than or equal to β times the infeasibility gap. Mathematically, this can be written as

$$c^T x - z^* \leq \beta \xi^T x,$$

where c is the linear programming objective function vector, z^* is the optimal value of the linear program, and the linear form $\xi^T x$ is the Phase I objective function. As discussed in [9], there are some practical linear programming problems where it may be very appropriate to set β large, and other practical problems where it may be very appropriate to set β to be quite small. [9] contains a further discussion and motivation for the β-balancing criterion.

In this paper, the concept of a β-balanced solution is used to define a class of trajectories from an infeasible points to an optimal solution of a given linear program. Each trajectory has the property that all points on or near the trajectory (in a suitable metric) are β-balanced. The main thrust of the paper is the development of an algorithm that traces a given β-balanced trajectory from a starting point near the trajectory to an optimal solution to the given linear programming problem in polynomial-time. More specifically, the algorithm allows for fixed improvement in the bound on the Phase I and the Phase II objectives in O(n) iterations of Newton steps.

The paper is organized as follows. Section 2 develops notation, standard forms, and assumptions for the problem at hand. In Section 3, the β-trajectories are defined, and the metric for measuring the closeness of points to a given trajectory is developed. In Section 4, the algorithm for tracing a given β-trajectory is presented, and basic properties of this algorithm are proved. In Section 5, the main complexity result regarding convergence of the algorithm is stated and proved. Section 6 contains an analysis of a particular set of almost-linear equations that is used in

Section 4. Section 7 comments on how to satisfy one of the assumptions needed to start the algorithm. Section 8 contains open questions.

2. Notation, Standard Form, and Assumption for a Combined Phase I – Phase II Linear Program:

Notation. The notation used is standard for the most part. The vector of ones is denoted by $e = (1, 1, \ldots, 1)T$, where the dimension is n. For any vector x,, X denotes the diagonal matrix whose diagonal components correspond to x. If $v \in R^n$, $\|v\|$ denotes the Euclidean norm.

The Phase I - Phase II Problem. Following the convention established by Anstreicher [1], this paper will work with a linear program of the form:

LP:

$$\begin{array}{llll} & \underset{x}{\text{minimize}} & c^T x & \\ \text{s.t.} & Ax & = & b \\ & \xi^T x & = & 0 \\ & x \geq 0, & & \end{array}$$

where we are given an infeasible "warm start" vector x^0 that is feasible for all constraints of (LP) except $\xi^T x^0 > 0$, and that x^0 satisfies the inequality constants strictly. Thus $Ax^0 = b$, $\xi^T x^0 > 0$, and $x^0 > 0$. The Phase I problem associated with (LP) is to solve:

$$\begin{array}{llll} & \underset{x}{\text{minimize}} & \xi^T x & \\ \text{s.t.} & Ax & = & b \\ & \xi^T x & \geq & 0 \\ & x \geq 0, & & \end{array}$$

and note that x^0 is feasible and lies in the relative interior of the feasible region for this Phase I problem.

(There are a number of straightforward ways to convert an arbitrary linear programming problem with an initial infeasible "warm start" solution into an instance of (LP) above, by placing all of the infeasibilities of the initial "warm start" solution into the single constraint $\xi^T x = 0$, see, e.g., Anstreicher [2], or [9]. At the end of this section, we will discuss this issue further.)

If x is feasible for the Phase I problem and z^* is the optimal value of the Phase II problem (LP), then $\xi^T x$ measures the feasiblility gap and $(c^T x - z^*)$

measures the optimal value gap. Note that the feasibility gap is always nonnegative; but due to the infeasibility of x in (LP), the optimal value gap $(c^T x - z^*)$ can be positive, negative, or zero. The combined Phase I – Phase II approach to solving (LP) is to generate values of $x = x^k$ that are feasible for the Phase I problem and for which $\xi^T x^k \to 0$ and $c^T x^k - z^* \to 0$ as $k \to \infty$, where $\{x^k\}$ is the sequence of iteration values.

β-Balance. Let $\beta = 0$ be a given parameter. If x is feasible for the Phase I problem, we say that x is "β-balanced" if x satisfies:

$$c^T x - z^* \leq \beta \xi^T x , \tag{2.1}$$

and an algorithm for solving (LP) is "β-balanced" if all iteration values x^k satisfy (2.1).

Inequality (2.1) has the following obvious interpretation. The left-hand-side is the optimal value gap, and the right-hand-side is β times the feasibility gap. Therefore x is β-balanced if the optimal value gap is less than or equal to the feasibility gap times the parameter β.

We call β the "balancing parameter" because it measures the balance or trade-off between the twin goals of attaining feasibility (Phase I) and of obtaining optimality (Phase II). Suppose that (2.1) is enforced for all iteration values of a given algorithm. If the value of β is set very high, then even near feasible solutions to (LP) can have a possibly large optimal value gap. Then if the algorithm is set to stop when the feasibility gap is reduced to a given tolerance ε, the optimal value gap might possibly still be quite large. If the value of β is set very low, then even very infeasible solutions cannot have a large positive optimal value gap. In this case, even if the tolerance value ε is not very small, (2.1) will ensure that the optimal value gap is very small (positive) or negative. As discussed in [9], there are some practical instances of (LP) where it may be very appropriate to set β large, and other practical instances of (LP) where it may be very appropriate to set β to be quite small. [9] contains a further discussion and motivation for the β-balancing criterion.

In order to enforce (2.1) in an algorithm for solving (LP), the value of z^* needs to be known, which is not generally the case in practice. Instead, suppose that we are given a lower bound B on the optimal value z^* of (LP), and that we impose the following constraint on iteration values of x:

$$(-\beta\xi + c)^T x = B . \tag{2.2}$$

Then because $B \le z^*$, (2.1) will be satisfied automatically as a consequence of (2.2). If $\{x^k\}$ and $\{B^k\}$ are a sequence of primal iteration values and lower bounds on z^* and that (2.2) is satisfied for all iterations k, then (2.1) will be satisfied for all iterations k as well. In particular, we will assume that we are given an initial lower bound B^0 and that the initial value x^0 together with B^0 satisfies (2.2). (At the end of this section, we will discuss how to convert any linear program into the form of (LP) that will also satisfy (2.2) for the initial primal values x^0 and the initial lower bound B^0.)

Based on this discussion, we can now state our assumptions regarding the linear programming problem (LP), as follows:

A(i) The given data for LP is the array $(A, \xi, b, c, x^0, B^0, \beta)$.

A(ii) $Ax^0 = b,\ \xi^T x^0 > 0,\ x^0 > 0,\ B^0 < z^*$.

A(iii) $\beta > 0$ and $(-\beta\xi + c)^T x^0 = B^0$.

A(iv) $\{d \in R^n \mid Ad = 0,\ \xi^T d = 0,\ d \ge 0,\ e^T d > 0,\ c^T d \le 0\} = \phi$.

A(v) $b \ne 0$, and the matrix

$$M = \begin{bmatrix} A \\ \xi^T \\ c^T \end{bmatrix}$$

has full row rank.

Assumptions A(i), A(ii), and A(iii) are based on the discussion up to this point. Assumption A(iv) states that the set of optimal solutions to (LP), if it exists, is a bounded set. This assumption is necessary in order to process many interior-point algorithms for linear programming, although there are ways to avoid the assumption, see, e.g. Anstreicher [3] or Vial [21]. Assumption A(v) ensures that the rows of the equations of (LP) are linearly independent, and that the objective function is not constant over the entire feasible region of (LP), in which case it is not important to distinguish between Phase I and Phase II. The linear independence of the rows of (LP) is assumed for convenience, and could be removed with the additional burden of more cumbersome notation in the use of projections, etc.

<u>Conversions</u>: Here we briefly describe how to convert a linear programming problem into an instance of (LP) satisfying the five assumptions A(i) – A(v). Suppose that the given linear program is of the form:

$\widehat{LP}$:
$$\begin{array}{llcl} & \underset{\hat{x}}{\text{minimize}} & \hat{c}^T\hat{x} & \\ \text{s.t.} & \hat{A}\hat{x} & = & b \\ & \hat{x} & \geq & 0\,, \end{array}$$

where $\hat{x}^0$ is a given "warm start" vector that is hopefully near-feasible and near-optimal. Also suppose that $\hat{B}^0$ is a given known lower bound on z^*, and that the set of optimal solutions of $(\widehat{LP})$ is assumed to be a bounded (possible empty) set. Then Section 2 of [9] describes a method for converting $(\widehat{LP})$ into an instance of (LP) satisfy Assumptions A(i), A(ii), A(iv), and

$$(-\beta\xi + c)^T x^0 < B^0 \,.$$

Then since $B^0 \leq z^*$, B^0 can be replaced by $B^0 = (-\beta\xi + c)^T x^0$, and this B^0 is a valid lower bound on z^*, and Assumption A(iii) is satisfied. Finally, Assumption A(v) can be satisfied by checking the linear independence of the matrix M and eliminating rows as necessary.

3. The β-Trajectory for a Linear Program, and Approximate Solutions Near the Trajectory

In this section we consider solving the "standard form" problem

LP:
$$\min_x \; c^T x \tag{3.1}$$
$$\text{s.t.} \quad \begin{aligned} Ax &= b \\ \xi^T x &= 0 \\ x &\geq 0 \end{aligned}$$

whose dual is

LD:
$$\max_{\pi, \theta, s} \; b^T \pi \tag{3.2}$$
$$\text{s.t.} \quad \begin{aligned} A^T\pi + \theta\xi + s &= c \\ s &\geq 0 \, . \end{aligned}$$

It is assumed that the data array $(A, \ \xi, b, c, x^0, B^0, \beta)$ satisfies assumptions A(i) - A(v) of the previous section.

We now consider the parametric family of augmented Phase I problems:

P_B:
$$z_B = \underset{x}{\text{minimize}} \; \xi^T x \tag{3.3}$$
$$\text{s.t.} \quad \begin{aligned} Ax &= b \\ (-\beta\xi + c)^T x &= B \\ x &\geq 0 \end{aligned}$$

for a given lower bound $B \leq z^*$ on the optimal value z^* of LP.

Note that P_B is a linear program parameterized by the RHS-coefficient parameter B of the last equation.

The following are elementary properties of P_B :

<u>Proposition 3.1 (Properties of P_B</u>) If $z^* < \infty$ and $B \in [B^0, z^*]$,

(i) P_B is feasible, and if $B < z^*$, P_B admits a strictly positive feasible solution.

(ii) $z_B > 0$ if $B < z^*$

(iii) $z_B = 0$ if $B = z^*$

(iv) The set of optimal solutions of P_B is nonempty and bounded.

(v) For all x feasible for P_B,

$$c^T x - z^* \leq \beta \xi^T x \tag{3.4}$$

<u>Proof</u>: (i) From assumption A(ii) and A(iii), x^0 is feasible for P_{B^0}. If $z^* < \infty$, then there is an optimal solution x^* to LP, and so x^* is feasible for P_{z^*}. Also, $x_B = (z^* - B^0)^{-1}\left((z^* - B)x^0 + (B - B^0)x^*\right)$ is feasible for P_B for all $B \in [B^0, z^*]$. If $B < z^*$, then $x_B > 0$ because $x^0 > 0$ and $x^* \geq 0$.

(ii) Suppose $B < z^*$ but $z_B \leq 0$. Then there exists some $\hat{x}$ that is feasible for P_B and for which $\xi^T \hat{x} \leq 0$. Let $\bar{B} = (\xi^T x^0 - \xi^T \hat{x})^{-1}\left((\xi^T x^0)B - (\xi^T \hat{x})B^0\right)$ and $\bar{x} = (\xi^T x^0 - \xi^T \hat{x})^{-1}\left((\xi^T x^0)\hat{x} - (\xi^T \hat{x})x^0\right)$. Then $\bar{x}$ is feasible for $P_{\bar{B}}$ and $\xi^T \bar{x} = 0$, and so $\bar{x}$ is feasible for LP. But $c^T \bar{x} = B\xi^T \bar{x} + \bar{B} = \bar{B} < z^*$, a contradiction. Thus $z_B > 0$.

(iii) Suppose $B = z^*$. Let x^* be an optimal solution to LP. Then x^* is feasible for P_{z^*}, and so $z_B \le \xi^T x^* = 0$. But if $z_B < 0$, an argument identical to that in (ii) shows a contradiction. Thus $z_B = 0$.

(iv) From the optimality properties of linear programs, the set of optimal solutions of P_B is nonempty, since P_B is feasible and z_B is finite. It remains to show that this set is bounded. If not, there exists a vector $d \ge 0$ with $d \ne 0$ that satisfies $Ad = 0, \left(-\beta\xi + c\right)^T d = 0, \xi^T d = 0$, whereby $c^T d = 0$ as well. This contradicts Assumption A(iv).

(v) If x is feasible for P_B, $c^T x = \beta\xi^T x + B \le \beta\xi^T x + z^*$. ■

From Proposition 3.1 (iii), zero is a lower bound on z_B as long as $B < z^*$. Therefore consider the following potential function reduction problem parameterized by the bound $B < z^*$.

$$PR_B: \qquad v(B) = \min_x \; F(x) = q\, ln\, \xi^T x - \sum_{j=1}^{n} ln\, x_j \tag{3.5}$$

$$\begin{aligned} \text{s.t.} \qquad Ax &= b \\ \left(-\beta\xi + c\right)^T x &= B \\ x > 0, \quad \xi^T x &> 0. \end{aligned}$$

where

$$q \triangleq n + 1 \tag{3.6}$$

for the remainder of this paper. Let $x(B)$ denote the optimal solution of PR_B if such a solution exists.

For the given and fixed balancing parameter $\beta > 0$, we can think of the set of values of $x(B)$ as B ranges from B^0 to z^* as a trajectory, and we denote this trajectory as

$$T_\beta = \left\{ x(B) \,\middle|\, x(B) \text{ solves } PR_B,\ B \in \left[B^0, z^*\right) \right\} . \tag{3.7}$$

We will refer to T_β as the "β-trajectory" for the linear program LP.

In this study, we develop an algorithm that will trace the β-trajectory T_β, i.e., the algorithm will generate a sequence of approximate solutions x to PR_B for a sequence of strictly increasing values of B that converge linearly to z^*. More specifically, the algorithm will generate a sequence $\{x^k\}$ of approximate solutions to PR_{B^k} for a sequence of strictly increasing lower bounds B^k with the property that x^k is feasible for the problem PR_{B^k} and x^k nearly solves the optimality conditions of PR_{B^k} (where the sense of approximation is developed later in this section.) Furthermore, we will show that if LP has a feasible solution (i.e., $z^* < +\infty$), then $B^k \to z^*$, $\xi^T x^k \to 0$, and $c^T x^k \to z^*$, as $k \to \infty$, and the convergence is linear, with fixed improvement in O (n) iterations. Therefore the algorithm is of the O (nL)-iteration variety of algorithms for linear programming.

Although our goal is to trace the set of optimal solutions $x(B)$ of PR_B in R^n for $B \in \left[B^0, z^*\right)$, it will be convenient from a mathematical point of view to make a transformation of program PR_B to an equivalent problem in R^{n+1}, and instead perform our analysis on this transformed problem, as follows.

Suppose $B \in \left[B^0, z^*\right)$ and x is feasible for PR_B. Then from Proposition 3.1 (ii), $\xi^T x > 0$, and so consider the following elementary (nonscaled) projective transformation:

$$(z, y) = h(x) = \left(\frac{x}{\xi^T x}, \frac{1}{\xi^T x} \right) \tag{3.8a}$$

$$x = h^{-1}(z, y) = \frac{z}{y} \tag{3.8b}$$

If x is feasible for PR_B, then $h(x)$ is well-defined (by Proposition 3.1) and PR_B transforms to

$$TR_B: \qquad w(B) = \underset{z,y}{\text{minimum}} \; G(z, y) = -\sum_{j=1}^{n} \ln z_j - \ln y \tag{3.9}$$

$$\begin{array}{llcl} \text{s.t.} & -Az \quad +by & = & 0 \\ & (-\beta\xi + c)^T z \quad -By & = & 0 \\ & \xi^T z & = & 1 \\ & z > 0,\ y > 0\ . & & \end{array}$$

Let $z(B)$, $y(B)$ denote the optimal solution of TR_B if such a solution exists. It is then easy to verify the following elementary proposition.

<u>Proposition 3.2 (Properties of PR_B and TR_B)</u> Under the transformations (3.8a) and (3.8b),

(i) x is feasible for PR_B if and only if (z, y) is feasible for TR_B .

(ii) $F(x) = G(z, y)$

(iii) $v(B) = w(B)$

(iv) $x = x(B)$ solves PR_B if and only if $z = z(B), y = y(B)$ solves TR_B .

<u>Proof</u>: Follows from direct substitution. ■

From Proposition 3.2, tracing the path of solutions $x(B)$ of PR_B is equivalent to tracing the path of solutions $(z(B), y(B))$ of TR_B . The advantage of working with TR_B lies in the fact that TR_B is an analytic center problem, and Newton's method is a powerful tool for finding and tracing approximate (analytic) centers for such a problem, as the data is deformed, see e.g., Vaidya [20], as well as [8] and [7]. (Of course, one complication in working with TR_B is that the parameter B is part of the constraint matrix of TR_B , as opposed to being part of the RHS-vector of PR_B , and this is more difficult to deal with.) Program TR_B is a strictly convex program, and so will have a unique optimal solution, as the next result states.

Proposition 3.3. If $z^* < \infty$ and $B \in [B^0, z^*)$, then PR_B and TR_B have unique optimal solutions $x(B)$ and $(z(B), y(B))$, respectively.

Proof: Suppose $B \in [B^0, z^*)$. From Proposition 3.1(i), PR_B has a feasible solution, so TR_B does as well. Now, TR_B is an analytic center problem, and so will have a unique optimal solution if the feasible region is bounded, see for example [8]. We thus only need to show that the feasible region of TR_B is bounded. If the feasible region of TR_B is not bounded, there exists $(d, v) \in R^n \times R$, with $d \geq 0$, $v \geq 0$, $(d, v) \neq 0$, and $-Ad + bv = 0$, $(-\beta\xi + c)^T d - Bv = 0$, $\xi^T d = 0$. If $v > 0$, we can presume $v = 1$. Then $Ad = b$, $d \geq 0$, $\xi^T d = 0$, and $c^T d = B < z^*$, which is a contradiction, since d is feasible for LP and so $c^T d \geq z^*$. Therefore $v = 0$, and $d \neq 0$. Then $Ad = 0$, $\xi^T d = 0$, $c^T d = 0$, $d \geq 0$, $d \neq 0$, which contradicts assumption A(iv). Thus, the feasible region of TR_B is bounded, and so TR_B and PR_B each have a unique optimal solution. ■

Concentrating our attention on TR_B , suppose (z, y) is feasible for TR_B . Then the Karush-Kuhn-Tucker (K-K-T) conditions are both necessary and sufficient for (z, y) to be an optimal solution to TR_B . The K-K-T conditions are equivalent (after arithmetic substitution) to the existence of (z, y) and multiplier vectors $(\pi, \delta, \theta, s, g)$ that satisfy the following equations:

$$-Az + by = 0 \tag{3.10a}$$

$$(-\beta\xi + c)^T z - By = 0 \tag{3.10b}$$

$$\xi^T z = 0 \tag{3.10c}$$

$$z > 0, \quad y > 0 \tag{3.10d}$$

$$A^T \pi + s = \theta(-\beta\xi + c) + \delta\xi \tag{3.10e}$$

$$-b^T\pi + g = -\beta\theta \tag{3.10f}$$

$$Zs - e = 0, \quad yg - 1 = 0 \tag{3.10g}$$

Therefore (z, y) solves TR_B if and only if there exists $(\pi, \delta, \theta, s, g)$ so that (3.10a-g) is satisfied. Now let $\gamma < 1$ be a positive constant. Just as in Tseng [19] and Roos and Vial [15], we will say (z, y) is a γ-approximate solution of TR_B if the following slight relaxation of (3.10a-g) is satisfied for some $(\pi, \delta, \theta, s, g)$:

$$-Az + by = 0 \tag{3.11a}$$

$$(-\beta\xi + c)^T z - By = 0 \tag{3.11b}$$

$$\xi^T z = 0 \tag{3.11c}$$

$$z > 0, \quad y > 0 \tag{3.11d}$$

$$A^T\pi + s = \theta(-\beta\xi + c) + \delta\xi \tag{3.11e}$$

$$-b^T\pi + g = -\beta\theta \tag{3.11f}$$

$$\sqrt{\sum_{j=1}^{n}(1 - s_j z_j)^2 + (1 - yg)^2} \leq \gamma \tag{3.11g}$$

Note that (3.11a-f) is the same as (3.10a-f) and that (3.11g) is just a relaxation of (3.10g). Therefore if (3.11a-g) is satisfied, (z, y) "almost" satisfies the K-K-T conditions, if γ is small. (Furthermore, one can show that if (z, y) is a γ-approximate solution of TR_B and $\gamma < 1$, then $v(B) \leq G(z, y) \leq v(B) + \gamma^2/(2(1-\gamma))$, and so $G(z, y)$ is very close to the optimal value $v(B)$ of TR_B, see Roos and Vial [15].)

Finally, we end this section with a property of LP that will be needed in the next two sections.

Proposition 3.5 (Infeasibility of LP) Suppose $B \leq z^*$. If there exists a solution $(\hat{\pi}, \hat{\delta}, \hat{\theta}, \hat{s}, \hat{g})$ to the system:

$$A^T \pi + \delta \xi + s \quad = \quad \theta c, \quad s > 0 \;, \tag{3.12a}$$

$$-b^T \pi + g \quad = \quad -\theta B \,,\; g > 0 \;, \tag{3.12b}$$

and $\hat{\theta} \leq 0$, then LP is infeasible.

Proof: Because $z^* \geq B^0$, LP has an optimal solution if it is feasible.

Suppose LP is feasible. Then LP has some optimal solution x^*. Then from (3.1) and (3.12),

$$b^T \hat{\pi} + x^{*T} \hat{s} \quad = \quad \hat{\theta} c^T x^* = \hat{\theta} z^*$$

and

$$-b^T \pi + g \quad = \quad -\hat{\theta} B \;.$$

Therefore $0 < x^{*T} \hat{s} + \hat{g} = \hat{\theta}(z^* - B)$, and so $\hat{\theta} > 0$, a contradiction.

Therefore LP is infeasible. ■

4. An Algorithm for Tracing the β-Trajectory, and Basic Properties

In this section, an algorithm for tracing the β-Trajectory is developed and basic properties of this algorithm are derived. The idea in the algorithm is to generate an increasing sequence of values of the lower bound $B^0 < B^1 \ldots < B^k < \ldots < z^*$, and for each B^k to produce a feasible solution x^k to PR_{B^k} that is an approximation to $x(B^k)$. Since we are working with the program TR_B rather than PR_B, the algorithm will produce a feasible solution (z^k, y^k) to TR_{B^k} that is a γ-approximate solution to TR_{B^k} for some fixed constant $\gamma \in (0, 1)$. (We will use $\gamma = 1/9$ in our analysis to follow.)

In order to motivate the algorithm, suppose $\bar{B} < z^*$ is given and $(\bar{z}, \bar{y})$ is given, and $(\bar{z}, \bar{y})$ is a γ-approximate solution to $TR_{\bar{B}}$, i.e., $(\bar{z}, \bar{y})$ and $(\bar{\pi}, \bar{\delta}, \bar{\theta}, \bar{s}, \bar{g})$ satisfy (3.11a-g) for $B = \bar{B}$, for some $(\bar{\pi}, \bar{\delta}, \bar{\theta}, \bar{s}, \bar{g})$. We will first increase the lower bound $\bar{B}$ to a larger value $\hat{B} = \bar{B} + \hat{\Delta}$, where $\hat{\Delta} > 0$ is computed (somehow), and so that $\hat{B} = \bar{B} + \hat{\Delta} < z^*$, i.e., $\hat{B}$ is a valid lower bound on z^*. Secondly, we will computer the Newton step $(\hat{d}, \hat{y})$ by taking a quadratic approximation to the objective function of $TR_{\hat{B}}$ at $(z, y) = (\bar{z}, \bar{y})$. Thirdly, we will update the values of (z, y) to $(\hat{z}, \hat{y}) = (\bar{z} + \hat{d}, \bar{y} + \hat{v})$. Fourth, the value of x is updated to $\hat{x} = \hat{z}/\hat{y}$. More formally, the algorithm is as follows:

<u>Algorithm 1</u> $(A, b, \xi, c, \beta, x^0, B^0, \gamma) \quad (0 < \gamma < 1)$ (4.1)

<u>Step 0 (Initialization)</u> $q = n + 1, \quad k = 0$

<u>Step 1 (Transform to (z, y)-space)</u> $z^0 = x^0/(\xi^T x^0), \quad y^0 = 1/(\xi^T x^0)$ (4.2)

<u>Step 2</u> (<u>Rename Variables</u>) $(\bar{B}, \bar{z}, \bar{y}, \bar{x}) = (B^k, z^k, y^k, x^k)$ (4.3)

<u>Step 3</u> (<u>Compute Increase in Lower Bound $\hat{\Delta}$ and Compute the Step $(\hat{d}, \hat{v})$</u>)

Solve the following equations for the values $(\Delta, d, v, \pi, \delta, \theta, s, g) = (\hat{\Delta}, \hat{d}, \hat{v}, \hat{\pi}, \hat{\delta}, \hat{\theta}, \hat{s}, \hat{g})$:

$$-Ad + bv = 0 \qquad (4.4a)$$

$$\left(-\beta\xi + c\right)^T d - \left(\overline{B} + \Delta\right) v \quad = \quad \Delta \overline{y} \tag{4.4b}$$

$$\xi^T d \quad = \quad 0 \tag{4.4c}$$

$$\hat{s} = \overline{Z}^{-1} e - \overline{Z}^{-2} d, \quad g = \overline{y}^{-1}\left(1 - \overline{y}^{-1} v\right) \tag{4.4d}$$

$$A^T \pi + s \quad = \quad \theta\left(-\beta\xi + c\right) + \delta\xi \tag{4.4e}$$

$$-b^T \pi + g \quad = \quad -\theta\left(\overline{B} + \Delta\right) \tag{4.4f}$$

$$\sqrt{d^T \overline{Z}^{-2} d + \overline{y}^{-2} v^2} \quad = \quad \sqrt{\gamma} \tag{4.4g}$$

$$\Delta > 0 \ . \tag{4.4h}$$

If no solution exists, stop. LP is infeasible.

If $\hat{\theta} \le 0$, stop. LP is infeasible.

<u>Step 4</u> (<u>New Values in (z, y)-space</u>)

$$\hat{B} = \overline{B} + \hat{\Delta} \tag{4.5a}$$

$$\left(\hat{z}, \hat{y}\right) = \left(\overline{z} + \hat{d}, \overline{y} + \hat{v}\right) \tag{4.5b}$$

<u>Step 5</u> (<u>New Value in x-space</u>)

$$\hat{x} = \hat{z}/\hat{y} \tag{4.6}$$

<u>Step 6</u> (<u>Redefine all Variables and Return</u>)

$$\left(B^{k+1}, z^{k+1}, y^{k+1}, \pi^{k+1}, \delta^{k+1}, \theta^{k+1}, s^{k+1}, g^{k+1}\right) = \left(\hat{B}, \hat{z}, \hat{y}, \hat{\pi}, \hat{\delta}, \hat{\theta}, \hat{s}, \hat{g}\right) \tag{4.7a}$$

$$x^{k+1} = \hat{x} \tag{4.7b}$$

$k \leftarrow k + 1$

Go to Step 2.

Here we review the steps in more detail. The data for the algorithm is given in (4.1). The data for LP is (A, b, ξ, c) . The balancing parameter $\beta > 0$ is also data. Furthermore, an initial point x^0 and lower bound B^0 are inputs as well, and are assumed to satisfy Assumptions A(i) – A(v). Also, the approximateness constant γ is data as well. (In Section 5, we will analyze Algorithm 1 with $\gamma = 1/9$). At Step 1, x^0 is transformed to (z^0, y^0) in (z, y)-space as in (3.8). Step 2 is just a renaming of variables for notational convenience. At Step 3, the incremental increase $\hat{\Delta}$ in the lower bound $\bar{B}$ is computed, as is the Newton step $(\hat{d}, \hat{v})$ for the variables (z, y). The seven equation systems (4.4a-g) is almost a linear system. We will prove below that the work needed to solve (4.4a-g) is no more than the work involved in computing projections with the matrix $[-A, b]$, and can be accomplished in $O(n^3)$ arithmetic operations. At Step 4, the new lower bound $\hat{B}$ and new values of (z, y) are computed, and in Step 5, the new value of x is computed using (3.8). Finally, the variables are redefined in Step 6.

At the start of the algorithm, it is assumed that the starting values (z^0, y^0) are a γ-approximate solution to TR_B , stated formally as:

Assumption A (vi): The values $(z^0, y^0) = \left(x^0/(\xi^T x^0), 1/(\xi^T x^0)\right)$ are a γ-approximate solution of TR_B , for the given prespecified constant γ . The following four properties of the algorithm will be proved:

Theorem 4.1 (Iterative γ-approximate Solutions): If $(\bar{z}, \bar{y})$ is a γ-approximate solution of $TR_{\bar{B}}$ at Step 2, then $(\hat{z}, \hat{y})$ is a γ-approximate solution to $TR_{\hat{B}}$ at Step 4. ■

Proposition 4.1 (Demonstrating Infeasibility via $\hat{\theta}$). If, at Step 3, (4.4a-h) has a solution and $\hat{\theta} \le 0$, then LP is infeasible. If $\hat{\theta} > 0$, then $\hat{B}$ is a valid lower bound on z^* . ■

Theorem 4.1 states that if the algorithm starts with a γ-approximate solution (z°, y°) of $TR_{B^{\circ}}$, then by induction, for all iterate values of k, (z^k, y^k) will be a γ-approximate solution to TR_{B^k}. Proposition 4.1 validates the second stopping criterion of Step 3 of the algorithm and validates the criterion that $\hat{B}$ must be a lower bound on z^*. The next Theorem validates the first stopping criterion of Step 3.

Theorem 4.2 (Solving (4.4a-h)).

(a) If (4.4a-h) has no solution, then LP is infeasible.

(b) Determining a solution to (4.4a-h) or demonstrating that there is no solution can be accomplished by computing three weighted projections onto the null-space of $(A, -b)$, and by using the quadratic formula, in $O(n^3)$ operations. ■

Finally, the next result states that if LP is feasible, the value of the potential function $G(z, y)$ (equivalently $F(x)$) decreases by at least a fixed constant at each iteration:

Theorem 4.3 (Decrease in Potential Function). If $\bar{x}$ and $(\bar{z}, \bar{y})$ is given at Step 2 of the algorithm and $\hat{x}$ and $(\hat{z}, \hat{y})$ is a given at Steps 4 and 5, then

$$\text{(i)} \qquad G(\hat{z}, \hat{y}) \leq G(\bar{z}, \bar{y}) - \gamma + \frac{\gamma}{2(1-\sqrt{\gamma})}$$

$$\text{(ii)} \qquad F(\hat{x}) \leq F(\bar{x}) - \gamma + \frac{\gamma}{2(1-\sqrt{\gamma})}. \qquad ■$$

Theorem 4.1, Proposition 4.1, and Theorem 4.3 are proved below. The proof of Theorem 4.2 is deferred until Section 6.

Proof of Theorem 4.1: We will show that $(\hat{B}, \hat{z}, \hat{y}, \hat{\pi}, \hat{\delta}, \hat{\theta}, \hat{s}, \hat{g})$ satisfies (3.11a-g) at $B = \hat{B}$. Since $(\bar{z}, \bar{y})$ is a γ-approximate solution to $TR_{\bar{B}}$, $(\bar{z}, \bar{y})$ satisfies (3.11a, b, c) for $B = \bar{B}$. Then (4.4a, b, c) combined with (4.5a, b) combine to show that $(\hat{z}, \hat{y})$ satisfies (3.11a, b, c) for $B = \hat{B}$ by arithmetic substitution. To see (3.11d), note that

$\hat{z} = \bar{Z}\left(e + \bar{Z}^{-1}\hat{d}\right) > 0$ since $\left\|\bar{Z}^{-1}\hat{d}\right\| \le \gamma < 1$ from (4.4g), and that $\hat{y} = \bar{y}\left(1 + \bar{y}^{-1}\hat{v}\right) > 0$ since $\left|\bar{y}^{-1}\hat{v}\right| \le \gamma < 1$ from (4.4g). Also, (4.4e, f) and (4.5a, b) imply that (3.11e, f) are satisfied. Thus it only remains to show (3.11g). We proceed as follows:

Note that we can write $\hat{z}, \hat{y}, \hat{s}$, and $\hat{g}$ as

$$\hat{z} = \bar{Z}\left(e + \bar{Z}^{-1}\hat{d}\right), \qquad \hat{s} = \bar{Z}^{-1}\left(e - \bar{Z}^{-1}\hat{d}\right) \tag{4.8a}$$

$$\hat{y} = \bar{y}\left(e + \bar{y}^{-1}\hat{v}\right), \qquad \hat{g} = \bar{y}^{-1}\left(1 - \bar{y}^{-1}\hat{v}\right) \tag{4.8b}$$

from (4.5b) and (4.4d).

Thus $\quad 1 - \hat{z}_j\hat{s}_j = 1 - \left(1 + \hat{d}_j/\bar{z}_j\right)\left(1 - \hat{d}_j/\bar{z}_j\right) = \hat{d}_j^2/\bar{z}_j^2, \quad j = 1, \ldots, n,$

and $\quad 1 - \hat{y}\hat{g} = 1 - \left(1 + \hat{v}/\bar{y}\right)\left(1 - \hat{v}/\bar{y}\right) = \bar{y}^{-2}\hat{v}^2.$

Thus we have

$$\begin{aligned}
\left(\sum_{j=1}^{n}\left(1 - \hat{s}_j\hat{z}_j\right)^2 + \left(1 - \hat{y}\hat{g}\right)^2\right)^{\frac{1}{2}} &= \left(\sum_{j=1}^{n}\left(\hat{d}_j/\bar{z}_j\right)^4 + \left(\bar{y}^{-1}\hat{v}\right)^4\right)^{\frac{1}{2}} \\
&= \left[\left(\sum_{j=1}^{n}\left(\hat{d}_j/\bar{z}_j\right)^4 + \left(\bar{y}^{-1}\hat{v}\right)^4\right)^{\frac{1}{4}}\right]^2 \\
&\le \left[\left(\sum_{j=1}^{n}\left(\hat{d}_j/\bar{z}_j\right)^2 + \left(\bar{y}^{-1}\hat{v}\right)^2\right)^{\frac{1}{2}}\right]^2 \\
&= \left(\sqrt{\gamma}\right)^2 \qquad \text{from (4.4g)} \\
&= \gamma
\end{aligned}$$

Therefore (3.11g) is satisfied. ■

It should be noted that this proof, which is based on the quadratic convergence of the underlying Newton process, is based on constructs used in Newton-method proofs in Tseng [19] and Roos and Vial [15].

Proof of Proposition 4.1. If $\hat{\theta} \le 0$, LP is infeasible. If $\hat{\theta} > 0$, $\hat{B} = \bar{B} + \hat{\Delta}$ is a valid lower bound on z^*.

Proof: From (4.4d) and (4.4g), we have $\hat{s} > 0$ and $\hat{g} > 0$. Also from (4.4e, f),

$$A^T\hat{\pi} + \left(-\hat{\delta} + \hat{\theta}\beta\right)\xi + \hat{s} = \hat{\theta}c$$

$$-b^T\hat{\pi} + \hat{g} = -\hat{\theta}\hat{B}.$$

Thus (3.12a, b) are satisfied for $(\pi, \delta, \theta, s, g) = (\hat{\pi}, -\hat{\delta} + \beta\hat{\theta}, \hat{s}, \hat{g})$. Thus from Proposition 3.5, LP is infeasible if $\hat{\theta} \le 0$. If $\hat{\theta} > 0$, rearranging the above yields $(\pi, \theta, s) = \left(\hat{\pi}/\hat{\theta}, \left(-\hat{\delta} + \hat{\theta}\beta\right)/\hat{\theta}, \hat{s}/\hat{\theta}\right)$ is feasible for LD (3.2), and so $z^* \ge b^T\hat{\pi}/\hat{\theta} = \hat{B} + \hat{g}/\hat{\theta} > \hat{B}$. ■

Proof of Theorem 4.3: From Proposition 3.2, it suffices to prove (i), since (i) and (ii) are equivalent.

$$G(\bar{z}, \bar{y}) - G(\hat{z}, \hat{y}) = \sum_{j=1}^{n} \ln \hat{z}_j + \ln \hat{y} - \sum_{j=1}^{n} \ln \bar{z}_j - \ln \bar{y}$$

$$= \sum_{j=1}^{n} \ln\left(1 + \left(\bar{Z}^{-1}\hat{d}\right)_j\right) + \ln\left(1 + \bar{y}^{-1}\hat{v}\right) \qquad \textbf{(from 4.8a, b)}$$

$$\ge e^T\bar{Z}^{-1}\hat{d} + \bar{y}^{-1}\hat{v} - \sum_{j=1}^{n} \frac{\left(\bar{Z}^{-1}\hat{d}\right)_j^2}{2\left(1 - \sqrt{\gamma}\right)} - \frac{\left(\bar{y}^{-1}\hat{v}\right)^2}{2\left(1 - \sqrt{\gamma}\right)} \qquad \textbf{(from Proposition A.2 and (4.4g))}$$

$$= e^T\bar{Z}^{-1}\hat{d} + \bar{y}^{-1}\hat{v} - \frac{\gamma}{2\left(1 - \sqrt{\gamma}\right)} \qquad \textbf{(from 4.4g)}$$

$$\geq \gamma - \frac{\gamma}{2\left(1-\sqrt{\gamma}\right)},$$

so long as $\quad e^T \bar{Z}^{-1} \hat{d} + \bar{y}^{-1} \hat{v} \geq \gamma.$

To demonstrate this inequality, note that

$$e^T \bar{Z}^{-1} \hat{d} = \hat{s}^T \hat{d} + \hat{d}^T \bar{Z}^{-2} \hat{d} \qquad \text{(from 4.4d)}$$

$$= -\hat{d}^T A^T \hat{\pi} + \hat{\theta}\left(-\beta\xi + c\right)^T \hat{d} + \hat{\delta}\xi^T \hat{d} + \hat{d}^T \bar{Z}^{-2} \hat{d}, \qquad \text{(from 4.4e)}$$

$$= -\hat{v} b^T \hat{\pi} + \hat{\theta} \bar{B} \hat{v} + \hat{\theta} \hat{\Delta} \hat{v} + \hat{\theta} \hat{\Delta} \bar{y} + \hat{d}^T \bar{Z}^{-2} \hat{d}, \qquad \text{(from 4.4a, b, c)}$$

and

$$\bar{y}^{-1} \hat{v} = \hat{g} \hat{v} + \bar{y}^{-2} \hat{v}^2, \qquad \text{(from 4.4d)}$$

$$= \hat{v} b^T \hat{\pi} - \hat{\theta} \bar{B} \hat{v} - \hat{\theta} \hat{\Delta} \hat{v} + \bar{y}^{-2} \hat{v}^2. \qquad \text{(from 4.4f)}$$

Thus $\quad e^T \bar{Z}^{-1} \hat{d} + \bar{y}^{-1} \hat{v} = \hat{\theta} \hat{\Delta} \bar{y} + \hat{d}^T \bar{Z}^{-2} \hat{d} + \bar{y}^{-2} \hat{v}^2$

$$= \hat{\theta} \hat{\Delta} \bar{y} + \gamma \qquad \text{(from 4.4g)}$$

$$\geq \gamma,$$

since $\hat{\Delta} > 0$, $\bar{y} > 0$, and $\hat{\theta} > 0$. ■

5. Convergence Properties of Algorithm 1

Let $\gamma = \frac{1}{9}$ be the closeness parameter used in Algorithm 1, and let (x^k, B^k), $k=0, \ldots,$ be the sequence of iterate values determined by Algorithm 1. In this section we prove the following theorem:

Theorem 5.1 (Convergence). Suppose Algorithm 1 is executed with parameter $\gamma = \frac{1}{9}$. Also, suppose that $z^* < \infty$ and let (π^*, θ^*, s^*) be any optimal solution to LD (3.2). Then

(i) $$F(x^{k+1}) \leq F(x^k) - 1/36 \qquad \text{for all } k = 0, \ldots. \tag{5.1}$$

(ii) There exists a constant $\alpha^* > 0$ with the property that if

$$k > 36q\left(ln\left(\xi^T x^0\right) + ln\, \alpha^*\right), \text{ then} \tag{5.2}$$

(a) $$0 < \xi^T x^k \leq \left(\xi^T x^0\right) \alpha^* e^{-k/36q} \tag{5.3}$$

(b) $$-\left|\theta^*\right| \left(\xi^T x^0\right) \alpha^* e^{-k/36q} \leq c^T x^k - z^* \leq \beta \left(\xi^T x^0\right) \alpha^* e^{-k/36q} \tag{5.4}$$

(c) $$0 \leq z^* - B^k \leq \left(\beta + \left|\theta^*\right|\right) \left(\xi^T x^0\right) \alpha^* e^{-k/36q} \tag{5.5}$$

Furthermore, it suffices to set

$$\alpha^* = \max\left\{1, \hat{\alpha}\right\}, \tag{5.6}$$

where $\hat{\alpha}$ is the optimal objective value of the linear program:

$$\hat{L}: \qquad \hat{\alpha} = \max_{x,u} \left(\frac{1}{n}\right) e^T (X^0)^{-1} x \tag{5.7}$$

$$\text{s.t.} \qquad Ax = ub$$

$$(-\beta\xi + c)^T x \leq uz^*$$

$$0 \le \xi^T x \le 1$$

$$u \le 1$$

$$x \ge 0, \quad u \ge 0 . \quad \blacksquare$$

Theorem 5.1(i) states that the potential function $F(x)$ decreases by at least $1/36$ at each iteration of the algorithm. Theorem 5.1(ii) states that $\xi^T x^k \to 0$, $c^T x^k \to z^*$, and $B^k \to z^*$. Furthermore, for k sufficiently large, the values of $\xi^T x^k$, $c^T x^k - z^*$, and $B^k - z^*$ are bounded by a geometric sequence, thus showing linear convergence of these series. Furthermore, since $q = n+1$, there will be a fixed improvement in the bounding sequence in $O(n)$ iterations. Therefore, Algorithm 1 exhibits fixed improvement in the values $\xi^T x^k$, $c^T x^k - z^*$, and $z^* - B^k$, in $O(n)$ iterations, for k sufficiently large.

Let L be the bit-size representation of the data $(A, b, \xi, \beta, x^0, B^0)$.

We now argue that if z^* is finite, then $\alpha^* \le 2^{O(L)}$

and that (5.2) is satisfied after $O(L)$ iterations. To see this, note from (5.7) that the data for $\hat{L}$ has bit-size representation $O(L)$, since z^* can be stored with $O(L)$ bits (see Papadimitriou and Steiglitz [13]) and all other data are $O(L)$. Therefore $\hat{\alpha}$ (which we show below is finite) satisfies $\hat{\alpha} \le 2^{O(L)}$, and so from (5.7), $\alpha^* \le 2^{O(L)}$. Therefore (5.2) will be satisfied within $O(L)$ iterations.

Before proving the theorem, we first prove the following proposition regarding α^* and $\hat{\alpha}$:

Proposition 5.1. Suppose $z^* < \infty$. Then $0 < \hat{\alpha} < \infty$.

Proof: If $z^* < \infty$, then LP attains its optimum at some point x^*. Then $x = x^*$, $u = 1$, is feasible for the linear program $\hat{L}$ (5.8), and since $b \ne 0$, then $e^T (X^0)^{-1} x^* > 0$. Thus $\hat{\alpha} > 0$. Also, $\hat{\alpha} < +\infty$ unless the linear program $\hat{L}$

defining $\hat{\alpha}$ is unbounded. But if $\hat{L}$ is unbounded, there exists $(d, v) \geq 0$, with $(d, v) \neq 0$ satisfying $Ad = vb$, $(-\beta\xi + c)^T d \leq vz^*$, $\xi^T d = 0$, $v = 0$, and so $d \neq 0$, $d \geq 0$, $Ad = 0$, $\xi^T d = 0$, $c^T d \leq 0$, which contradicts Assumption A(iv). Thus $\hat{\alpha} < +\infty$. ■

(It should be noted that the construction of $\hat{\alpha}$ and α^* in (5.6) and (5.7) is a modification of the construction of bounds in Todd [18], but is more streamlined in that is uses a linear program and a smaller implicit feasible region for the optimization problem underlying $\hat{\alpha}$.

<u>Proof of Theorem 5.1</u>: Note that if $\gamma = \frac{1}{9}$, then $\gamma - \frac{\gamma}{2\left(1-\sqrt{\gamma}\right)} = \frac{1}{36}$, and so (5.1) follows from Theorem 4.3(ii). After k iterations, (5.1) implies that

$$F\left(x^k\right) \leq F\left(x^0\right) - k/36. \tag{5.8}$$

Then substituting the formula for $F(x)$ in (3.5) and exponentiating, (5.8) becomes

$$\begin{aligned} \xi^T x^k \leq \quad & \left(\xi^T x^0\right) e^{-k/36q} \left(\prod_{j=1}^{n} \frac{x_j^k}{x_j^0} \right)^{\frac{1}{q}} \\ \leq \quad & \left(\xi^T x^0\right) e^{-k/36q} \left(\frac{1}{n} e^T \left(X^0\right)^{-1} x^k \right)^{n/q}, \end{aligned} \tag{5.9}$$

where the latter inequality follows from arithmetic-geometric mean inequality. Thus in order to prove (5.3), it is sufficient to show that if

$$k > 36q\left(\ln\left(\xi^T x^0\right) + \ln \alpha^*\right),$$

then

$$\frac{1}{n} e^T \left(X^0\right)^{-1} x^k \leq \alpha^*. \tag{5.10}$$

since $\alpha^* \ge 1$ from (5.6).

It follows that if k satisfies (5.2) then

$$\left(\xi^T x^0\right) e^{-k/36q} < \frac{1}{\alpha^*}. \tag{5.11}$$

Suppose that for some k satisfying (5.2) that (5.10) does not hold. Let $\alpha = \frac{e^T\left(X^0\right)^{-1} x^k}{n}$. Then $\alpha > \alpha^* \ge 1$. Let $x = \left(\frac{\alpha^*}{\alpha}\right) x^k$ and $u = \frac{\alpha^*}{\alpha}$. Then $0 < u < 1$, and $Ax = uAx^k = ub$, and also:

$$\left(-\beta\xi + c\right)^T x = u\left(-\beta\xi + c\right)^T x^k = uB^k < uz^*$$

$$\xi^T x = u\xi^T x^k > 0.$$

$$\xi^T x = u\xi^T x^k \le u\left(\xi^T x^0\right) e^{-k/36q}\left(\frac{1}{n} e^T\left(X^0\right)^{-1} x^k\right)^{\frac{n}{q}} \qquad \text{(from (5.9))}$$

$$< \left(\frac{\alpha^*}{\alpha}\right)\left(\frac{1}{\alpha^*}\right)(\alpha)^{\frac{n}{q}} \qquad \text{(from (5.11))}$$

$$\le 1.$$

Thus, since $x > 0$ as well, (x, u) is feasible for $\hat{L}$, and satisfies all inequalities strictly.

Therefore $\hat{\alpha} > \frac{e^T(X^0)^{-1} x}{n} = \left(\frac{\alpha^*}{\alpha}\right)\frac{e^T(X^0)^{-1} x^k}{n} = \frac{\alpha^*}{\alpha} \cdot \alpha = \alpha^*$, a contradiction since from (5.6) we have $\alpha^* \ge \hat{\alpha}$. Therefore (5.10) holds, and hence so does (5.3).

Next, note that since x^k is feasible for PR_{B^k}, then

$$c^T x^k - z^* \leq c^T x^k - B^k = \beta \xi^T x^k \leq \beta\left(\left(\xi^T x^0\right) \alpha^* e^{-k/36q}\right) \tag{5.12}$$

which proves the right side of (5.4). To prove the left side, we use Proposition A.3 of the Appendix, which states that for each x^k,

$$c^T x^k - z^* \geq \theta^* \xi^T x^k \geq -\left|\theta^*\right| \xi^T x^k \geq -\left|\theta^*\right| \left(\xi^T x^0\right) \alpha^* e^{-k/36q} \tag{5.13}$$

where the last inequality follows from (5.3). Thus (5.4) is proved.

To prove (5.5), note that

$$0 \leq z^* - B^k = z^* - c^T x^k + c^T x^k - B^k \leq \left(\left|\theta^*\right| + \beta\right) \left(\xi^T x^0\right) \alpha^* e^{-k/36q}$$

from (5.12) and (5.13). ■

6. Analysis of Equations (4.4a-h)

In this section, equations (4.4a-h) of Step 3 of the algorithm are analyzed, basic properties are derived, and Theorem 4.2 is proved.

In order to motivate equations (4.4a-h) of Step 3 of the algorithm, suppose we are at the start of Step 3 of the algorithm. We have on hand a lower bound $\bar{B} < z^*$, and we have on hand a γ-approximate solution $(\bar{z}, \bar{y})$ to $TR_{\bar{B}}$. That is, we have the following array of vectors and scalars:

$$\left(\bar{B}, \bar{z}, \bar{y}, \bar{\pi}, \bar{\delta}, \bar{\theta}, \bar{s}, \bar{g}\right)$$

and these values satisfy the following system (from 3.11):

$$-A\bar{z} + b\bar{y} = 0 \tag{6.1a}$$

$$(-\beta\xi + c)^T\bar{z} - \bar{B}\bar{y} = 0 \tag{6.1b}$$

$$\xi^T\bar{z} = 1 \tag{6.1c}$$

$$\bar{z}, \bar{y} > 0 \tag{6.1d}$$

$$A^T\bar{\pi} + \bar{s} = \bar{\theta}(-\beta\xi + c) + \bar{\delta}\xi \tag{6.1e}$$

$$-b^T\bar{\pi} + \bar{g} = -\bar{B}\bar{\theta} \tag{6.1f}$$

$$\bar{\gamma} \triangleq \sqrt{\sum_{j=1}^{n} (1 - \bar{s}_j\bar{z}_j)^2 + (1 - \bar{y}\bar{g})^2} \le \gamma \tag{6.1g}$$

$$\bar{x} = \bar{z}/\bar{y}. \tag{6.1h}$$

Our interest lies in increasing the lower bound from $B = \bar{B}$ to $B = \bar{B} + \Delta$ for some $\Delta > 0$, and then taking a Newton step in the problem $TR_{\bar{B}+\Delta}$. Specifically, we want to find a direction array (d, v) so that $(\bar{z} + d, \bar{y} + v)$ is feasible for $TR_{\bar{B}+\Delta}$ and that $(\bar{z} + d, \bar{y} + v)$ is a γ-approximate solution to $TR_{\bar{B}+\Delta}$. The issue then is how to choose the increment value $\Delta > 0$ in order to ensure that the Newton step will result in a point that is a γ-approximate solution to the new problem $TR_{\bar{B}+\Delta}$.

As we change B from $B = \bar{B}$ to $B = \bar{B} + \Delta$, the constraint matrix of TR_B will change. For a given value of Δ, the Newton direction (d, v) will be the optimal solution to the following quadratic program:

$$Q_\Delta: \quad \min_{d,v} -e^T\bar{Z}^{-1}d - \bar{y}^{-1}v + \frac{1}{2}d^T\bar{Z}^{-2}d + \frac{1}{2}\bar{y}^{-2}v^2 \tag{6.2a}$$

$$\text{s.t.} \qquad -Ad + bv \quad = \quad 0 \qquad (\pi) \tag{6.2b}$$

$$(-\beta\xi + c)^T d - (\bar{B} + \Delta)v \quad = \quad \bar{y}\Delta \qquad (\theta) \tag{6.2c}$$

$$\xi^T d \quad = \quad 0 \qquad (\delta) \tag{6.2d}$$

It is a matter of checking arithmetic to see that if (d, v) satisfy (6.2b, c, d), then $(\bar{z} + d, \bar{y} + v)$ satisfies the equations for $TR_{\bar{B}+\Delta}$, see (3.9). Also note that the objective function of Q_Δ is just the first two terms of the Taylor series expansion of $G(z, y)$ at the point $(z, y) = (\bar{z}, \bar{y})$. Furthermore, from Assumption A(v), the rows of equations (6.2b, c, d) must be linearly independent, and therefore, Q_Δ will always have a feasible solution for any Δ. In fact, because the objective function is a strictly convex quadratic form, then Q_Δ will always have a unique solution, and that this solution can be computed in closed form. The Karush-Kuhn-Tucker (KKT) conditions for Q_Δ are both necessary and sufficient for optimality, and are satisfied if there exists $(d_\Delta, v_\Delta, \pi_\Delta, \delta_\Delta, \theta_\Delta, s_\Delta, g_\Delta)$ for which the following equations are satisfied:

$$-Ad_\Delta + bv_\Delta \quad = \quad 0 \tag{6.3a}$$

$$(-\beta\xi + c)^T d_\Delta - (\bar{B} + \Delta)v_\Delta \quad = \quad \bar{y}\Delta \tag{6.3b}$$

$$\xi^T d_\Delta \quad = \quad 0 \tag{6.3c}$$

$$s_\Delta \triangleq \bar{Z}^{-1}e - \bar{Z}^{-2}d_\Delta, \qquad g_\Delta \triangleq \bar{y}^{-1}(1 - \bar{y}^{-1}v_\Delta) \tag{6.3d}$$

$$A^T \pi_\Delta + s_\Delta \quad = \quad \theta_\Delta\left(-\beta\xi + c\right) + \delta_\Delta \xi \tag{6.3e}$$

$$-b^T \pi_\Delta + g_\Delta \quad = \quad -\theta_\Delta\left(\overline{B} + \Delta\right) \tag{6.3f}$$

Summarizing the discussion so far, we see that for any value of the parameter Δ, equations (6.3a-f) represent the optimality conditions for the quadratic approximation of the problem $TR_{\overline{B}+\Delta}$ at the point $(z, y) = (\bar{z}, \bar{y})$. These equations will have a unique solution for any value of Δ. Comparing (6.3a-f) to (4.4a-h), note that (6.3a-f) are identical to (4.4a-f). Then if we define

$$k(\Delta) = \sqrt{d_\Delta \bar{Z}^{-2} d_\Delta + \bar{y}^{-2} v_\Delta^2} \tag{6.4}$$

we see that (4.4g-h) is satisfied if we can find $\Delta > 0$ satisfying

$$k(\Delta) = \sqrt{\gamma} \tag{6.5}$$

We now show how to solve (6.5) as well as (6.3a-f), thus solving (4.4a-h). Structurally, the program Q_Δ is equivalent to:

$$P_\Delta: \quad \begin{array}{lrcl} \underset{x}{\text{minimize}} & q^T x + \frac{1}{2} x^T x & & \\ \text{s.t.} & Mx & = & 0 \\ & (f + w\Delta)^T x & = & \Delta \bar{y}, \end{array}$$

$$\begin{array}{lrcl} \text{where} & x & = & \left(\bar{Z}^{-1} d, \bar{y}^{-1} v\right), \end{array} \tag{6.6}$$

$$M = \begin{bmatrix} -A\bar{Z} & , & b\bar{y} \\ \xi^T \bar{Z} & , & 0 \end{bmatrix},$$

$$w^T = \left[\left(-\beta\xi^T + c^T\right)\bar{Z}, -\overline{B}\bar{y}\right],$$

$$w^T = \left[\quad 0^T, \ -\bar{y}\right],$$

and $q^T = \left[-e^T, -1\right]$.

The program P_Δ has the property that the data in the last row depends linearly on the parameter Δ . Because of this simple linear dependence, we can solve for the unique solution x_Δ of P_Δ with relative ease, as follows. Let P be the projection matrix that projects vectors x onto the null space of M , i.e.,

$$P = \left[I - M^T(M M^T)^{-1} M\right] . \tag{6.7}$$

Then let

$$\bar{q} = Pq, \bar{f} = Pf \text{ , and } \bar{w} = Pw . \tag{6.8}$$

The K-K-T conditions for P_Δ can be written as:

$$\bar{q} + x_\Delta = \left(\bar{f} + \bar{w}\Delta\right)\lambda_\Delta \tag{6.9a}$$

$$\left(\bar{f} + \bar{w}\Delta\right)^T x_\Delta = \Delta\bar{y} . \tag{6.9b}$$

This is easily solved as:

$$\lambda_\Delta = \frac{-\bar{q}^T(\bar{f} + \bar{w}\Delta) - \bar{y}\Delta}{(\bar{f} + \bar{w}\Delta)^T(\bar{f} + \bar{w}\Delta)} \tag{6.10a}$$

$$x_\Delta = -\bar{q} - \lambda_\Delta\left(\bar{f} + \bar{w}\Delta\right) \tag{6.10b}$$

and $\left(\bar{Z}^{-1} d_\Delta, \ \bar{y}^{-1} v_\Delta\right) = x_\Delta$ (6.10c)

Therefore (6.10a, b, c) gives the optimal solution to P_Δ , and hence Q_Δ , for any value of Δ . However, we seek a solution to Q_Δ for a value of Δ that satisfies (6.5) (through 6.4). Under the scale transformation (6.6), equation (6.5) becomes

$$(x_\Delta)^T(x_\Delta) = \gamma . \tag{6.11}$$

Substituting for x_Δ in (6.10a, b) equation (6.11) becomes

$$\gamma = \bar{q}^T\bar{q} + \frac{\left(-\bar{q}^T(\bar{f} + \bar{w}\Delta) - \bar{y}\Delta\right)\left(\bar{q}^T(\bar{f} + \bar{w}\Delta) - \bar{y}\Delta\right)}{(\bar{f} + \bar{w}\Delta)^T(\bar{f} + \bar{w}\Delta)} . \tag{6.12}$$

Now note that (6.12) can be solved analytically for Δ by clearing the denominator and using the quadratic formula on the resulting quadratic equation in Δ . If there is a solution to (6.12) with $\Delta > 0$, then this procedure will find such a solution; likewise, if there is no solution with $\Delta > 0$, this will be uncovered as well. The bulk of the work effort lies in computing the three projections $\bar{q}$, $\bar{f}$, and $\bar{w}$, which can be accomplished in $O(n^3)$ operations.

One further point must be made, namely that the denominator of (6.10a) is never zero. To see why this is true, note from Assumption A(v) that the following matrix must have full row rank, since $\bar{z} > 0$ and $\bar{y} > 0$:

$$\begin{bmatrix} -A\bar{Z} & , & b\bar{y} \\ \xi^T\bar{Z} & , & 0 \\ \left(-\beta\xi^T + c^T\right)\bar{Z} & , & -\bar{B}\bar{y} \\ 0 & , & -\bar{y} \end{bmatrix} .$$

But this matrix is simply

$$\begin{bmatrix} M \\ f^T \\ w^T \end{bmatrix} ,$$

which then has full row rank. Therefore the projections $\bar{f}$ and w (onto the null space of M) are linearly independent, and so the denominator of (6.10a) does not vanish.

Because the denominator of (6.10a) does not vanish, we have:

Proposition 6.1 (Continuity of Q_Δ). The functions λ_Δ, x_Δ, d_Δ, v_Δ, and $k(\Delta)$ are all continuous in Δ.

Proof: Follows immediately from the above discussion. ■

Proposition 6.2. At $\Delta = 0$, $k(\Delta) < \gamma$.

Proof: Let (d_0, v_0) be the optimal solution to Q_Δ at $\Delta = 0$. Then from (6.3a-f), there exists $(\pi_0, \delta_0, \theta_0, s_0, g_0)$ together with (d_0, v_0) satisfying

$$-A d_0 + b v_0 = 0 \tag{6.13a}$$

$$(-\beta\xi + c)^T d_0 - \bar{B} v_0 = 0 \tag{6.13b}$$

$$\xi^T d_0 = 0 \tag{6.13c}$$

$$s_0 = \bar{Z}^{-1} e - \bar{Z}^{-2} d_0, \quad g_0 = \bar{y}^{-1}\left(1 - \bar{y}^{-1} v_0\right) \tag{6.13.d}$$

$$A^T \pi_0 + s_0 = \theta_0(-\beta\xi + c) + \delta_0 \xi \tag{6.13e}$$

$$-b^T \pi_0 + g_0 = -\theta_0 \bar{B} \tag{6.13f}$$

Now consider the problem

$$\text{R:} \quad \underset{\pi, \delta, \theta, s, g}{\text{minimize}} \quad \frac{1}{2}\sum_{j=1}^{n} \left(1 - \bar{z}_j s_j\right)^2 + \frac{1}{2}\left(1 - \bar{y} g\right)^2$$

$$\text{s.t.} \quad -A^T \pi + \theta(-\beta\xi + c) + \delta\xi - s = 0$$

$$b^T \pi - \theta \bar{\beta} - g = 0$$

The K-K-T conditions for this convex problem are:

$$-A^T \pi + \theta(-\beta\xi + c) + \delta\xi - s = 0 \tag{6.14a}$$

$$b^T\pi - \theta\bar{B} - g \quad = \quad 0 \tag{6.14b}$$

$$-Ad + bv \quad = \quad 0 \tag{6.14c}$$

$$\left(-\beta\xi + c\right)^T - \bar{B}v \quad = \quad 0 \tag{6.14d}$$

$$\xi^T d \quad = \quad 0 \tag{6.14e}$$

$$s = \bar{Z}^{-1}\left(e - \bar{Z}^{-1}d\right), \quad g = \bar{y}^{-1}\left(1 - \bar{y}^{-1}v\right) \tag{6.14f}$$

Now note that $\left(d_0, v_0, \pi_0, \delta_0, \theta_0, s_0, g_0\right)$ satisfy (6.14a-f) and so $\left(\pi_0, \delta_0, \theta_0, s_0, g_0\right)$ solves the program R. Also note, however, that $\left(\bar{z}, \bar{y}\right)$ is a γ-approximate solution to $TR_{\bar{B}}$, so that there exists $\left(\bar{\pi}, \bar{\delta}, \bar{\theta}, \bar{s}, \bar{g}\right)$ together with $\left(\bar{z}, \bar{y}\right)$ that satisfy (3.11a-g) at $B = \bar{B}$. Therefore $\left(\bar{\pi}, \bar{\delta}, \bar{\theta}, \bar{s}, \bar{g}\right)$ is feasible for the program R. Therefore

$$k(0) \quad = \quad \sqrt{d_0 \bar{Z}^{-2} d_0 + \bar{y}^{-2} v_0^2} \qquad \text{(from (6.4)}$$

$$= \quad \sqrt{\sum_{j=1}^{n} \left(1 - \bar{z}_j (s_0)_j\right)^2 + \left(1 - \bar{y} g_0\right)^2} \qquad \text{(from 6.13d)}$$

$$\leq \quad \sqrt{\sum_{j=1}^{n} \left(1 - \bar{z}_j \bar{s}_j\right)^2 + \left(1 - \bar{y}\bar{g}\right)^2} \quad \leq \quad \gamma \quad \leq \quad \sqrt{\gamma}\,. \qquad \blacksquare$$

We now are ready to prove Theorem 4.2.

<u>Proof of Theorem 4.2</u>: Part (b) of Theorem 4.2 follows from the discussion on solving (4.4a-h) in this section. It only remains to prove part (a). From Proposition 6.2, $k(0) < \sqrt{\gamma} < 1$, and from Proposition 6.1, the continuity of $k(\cdot)$ ensures that if $k(\Delta) = \sqrt{\gamma}$ has no solution for $\Delta > 0$, then $k(\Delta) < \sqrt{\gamma} < 1$ for all values of $\Delta \geq 0$. Now suppose LP is feasible. Then z^* is finite, and let Δ be set to $\Delta = z^* - \bar{B} \geq 0$. Let $\left(d_\Delta, v_\Delta, \pi_\Delta, \delta_\Delta, \theta_\Delta, s_\Delta, g_\Delta\right)$ be the solution to (6.3a-f). Since $k(\Delta) < 1$, $s_\Delta > 0$ and $g_\Delta > 0$.

Therefore $A^T \pi_\Delta + (\theta_\Delta \beta - \delta_\Delta)\xi + s_\Delta = \theta_\Delta c$, $s_\Delta > 0$,

and $-b^T \pi_\Delta + g_\Delta = -\theta_\Delta(\overline{B} + \Delta) = -\theta_\Delta z^*$, from (6.3e, f).

Therefore if $\theta_\Delta \leq 0$, Proposition 3.5 states that LP is infeasible. So suppose $\theta_\Delta > 0$. Then from (6.3e, f) we obtain a feasible solution $(\pi, \theta, s) = (\pi_\Delta / \theta_\Delta, \beta - \delta_\Delta / \theta_\Delta, (s_\Delta / \theta_\Delta))$ to the dual LD, with objective value

$$b^T\pi = b^T(\pi_\Delta / \theta_\Delta) = \overline{B} + \Delta + g_\Delta / \theta_\Delta > \overline{B} + \Delta = z^* .$$

This contradicts the definition of z^* , and so LP must be infeasible. ■

7. Comments on Satisfying Assumption A(vi).

Assumption A(vi) states that the starting values $(z^0, y^0) = \left(x^0/\left(\xi^T x^0\right)\right), 1/\left(\xi^T x^0\right)$ are a γ-approximate solution of TR_{B°, for the given prespecified constant γ. Here we show how to satisfy this assumption. For convenience of notation, let $\bar{x} = x^0$. Then $(\bar{z}, \bar{y}) = \left(\bar{x}/\left(\xi^T \bar{x}\right), 1/\left(\xi^T \bar{x}\right)\right)$ is feasible for TR_{B°, with objective value $F(\bar{x}) = G(\bar{z}, \bar{y})$. Then using the potential reduction methodology of Ye [22], Gonzaga [12], or [4], it can be shown that in finitely many iterations of Newton's method augmented with a line-search, that a γ-approximate solution of TR_{B° will be generated. In fact, using analysis in [8], one can bound the number of iterations of the Newton process by

$$\left| \frac{F(\bar{x}) - v\left(B^\circ\right)}{1 + \gamma - \sqrt{1 + 2\gamma}} \right| .$$

8. Open Questions

Monotonicity. Traditional trajectory-following interior-point algorithms for linear programming parameterize the trajectory with a parameter μ that is related monotonically to the duality gap. One can show that along the idealized path of primal values $x(\mu)$ and dual values $\pi(\mu)$, $s(\mu)$, that the duality gap is equal to $n\mu$, where n is the number of inequalities, and hence is linear and monotone in trajectory parameter μ. The Trajectory T_β defined in the paper (3.7) is parameterized by the lower bound B on z^*. Theorem 5.1 indicates that as $B \to z^*$, that $c^T x(B) \to z^*$ and $\xi^T x(B) \to 0$. Furthermore, intuition suggests that $\xi^T x(B)$ is decreasing in B as B approaches z^* from below. However, after much mathematical effort, we have not been able to demonstrate any such monoticity.

Complexity. The lack of provable monoticity of the feasibility gap has forced the complexity analysis of the algorithm to rely on analysis of the potential functions $F(x)$ or $G(z, y)$. And using these potential functions, we are only able to prove constant improvement in the feasibility gap in $O(n)$ iterations. Most other trajectory-following algorithms can be shown to achieve constant improvement in $O(\sqrt{n})$ iterations. In fact, the bulk of the theoretical research on combined Phase I – Phase II methods has led to $O(n)$ iteration algorithms for constant improvement, and only quite recently has Ye et. al [23] shown $O(\sqrt{n})$ iteration algorithm for problems of this type.

Good Values of β. As argued in [9], there are many instances of linear programming problems for which the intuitive choice of the balancing constant β might be very high, or very low, depending on certain aspects of the problem. An open question is whether computationally there is a heuristically natural optimal choice of the balancing parameter β.

Appendix

Proposition A.1. If $x > -1$, $ln(1+x) \le x$.

Proposition A.2. If $|x| \le \alpha < 1$, then $ln(1+x) \ge x - \dfrac{x^2}{2(1-\alpha)}$.

Proofs of the above two inequalities can be found in [4], among other places.

Proposition A.3. Consider the dual linear programs:

$$
\begin{array}{llll}
LP_r: & z^*(r) = \min\limits_{x} c^T x & \qquad LD_r: & \max\limits_{\pi, \theta} b^T\pi + r\theta \\
 & \text{s.t} \quad Ax = b & & \text{s.t.} \quad A^T\pi + \xi\theta \le c \\
 & \qquad \xi^T x = r & & \\
 & \qquad x \ge 0. & &
\end{array}
$$

Suppose (π^*, θ^*) solves LD_0, and let $z^* = z^*(0)$. Then for any x feasible for LP_r, $c^T x \ge z^* + \theta^* r$.

Proof: Because (π^*, θ^*) is feasible for the dual LD_r for any r,

$$z^*(r) \ge b^T\pi^* + r\theta^* = z^*(0) + r\theta^* = z^* + r\theta^* .$$

Therefore, if x is feasible for LP_r, $c^T x \ge z^*(r) \ge z^* + r\theta^*$. ■

References

[1] K.M. Anstreicher, "A combined phase I - phase II projective algorithm for linear programming," Mathematical Programming 43 (1989) 209-223.

[2] K.M. Anstreicher, "A combined phase I - phase II scaled potential algorithm for linear programming," CORE Discussion Paper 8939, CORE Catholic University of Louvain, Belgium, 1989.

[3] K.M. Anstreicher, "On interior algorithms for linear programming with no regularity assumptions," Manuscript, Dept. of Operations Research, Yale University, New Haven (February 1991).

[4] R.M. Freund, "Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function," Mathematical Programming 51 (1991) 203-222.

[5] R.M. Freund, "Theoretical efficiency of a shifted barrier function algorithm for linear programming," Linear Algebra and its Applications 152 (1991) 19-41.

[6] R.M. Freund, "A potential-function reduction algorithm for solving a linear program directly from an infeasible 'warm start'," Mathematical Programming 52 (1991) 441-466.

[7] R.M. Freund, and K.C. Tan, "A Method for the Parametric Center Problem, with a Strictly Monotone Polynomial-Time Algorithm for Linear Programming," Mathematics of Operations Research 16 (1991), 775-801.

[8] R.M. Freund, "Projective Transformations for Interior Point Algorithms, and a Superlinearly Convergent Algorithm for the W-Center Problem," to appear in Mathematical Programming.

[9] R.M. Freund, "A Potential Reduction Algorithm with User-specified Phase I - Phase II Balance, for Solving a Lnear Program from an Infeaible Warm Start," Operations Research Center Paper 260-91, M.I.T. Operations Research Center, 1991.

[10] G. de Ghellinck and J.-P. Vial, "A polynomial Newton method for linear programming," Algorithmica 1 (1986) 425-453.

[11] P.E. Gill, W. Murray, M.A. Saunders and M.H. Wright, "Shifted barrier methods for linear programming," Technical Report SOL 88-9, Department of Operations Research, Stanford University, 1988.

[12] C.C. Gonzaga, "Polynomial affine algorithms for linear programming," Mathematical Programming 49 (1990) 7-21.

[13] C.H. Papadimitriou and K. Steiglitz, Combinatorial Optimization, Prentice-Hall, New Jersey, 1982.

[14] R.A. Polyak, "Modified barrier functions (theory and methods," Mathematical Programming 54 (1992) 177-222.

[15] C. Roos and J.P. Vial, "A polynomial method of approximate centers for linear programming," Mathematical Programming 54 295-305, 1992.

[16] M.J. Todd, "On Anstreicher's combined phase I - phase II projective algorithm for linear programming," Technical Report No. 776, School of OR&IE, Cornell University, Ithaca, NY, 1988, to appear in Mathematical Programming.

[17] M.J. Todd and Y. Wang, "On combined phase 1 - phase 2 projective methods for linear programming," Technical Report No. 877, School of OR&IE, Cornell University, Ithaca, NY. 1989, to appear in Algorithmica.

[18] M.J. Todd, "Combining phase I and phase II in a potential reduction algorithm for linear programming," Technical Report No. 907, School of OR&IE, Cornell University, Ithaca, NY, 1990, to appear in SIAM Journal on Optimization.

[19] P. Tseng, "A simple complexity proof for a polynomial-time linear programming algorithm," Operations Research Letters 8 155-159, 1989.

[20] P.M. Vaidya, "A locall well-behaved potentail function and a simple Newton-tpe method for finding the center of a plytope," in N. Megiddo, editor, Progress in Mathematical Programming: Interior Point and Related Methods, pp. 79-90, Springer Verlag, NewYork, 1989.

[21] J.P. Vial, "A projective algorithm for linear programming with no regularity condition," University of Geneva, Geneva, Switzerland, 1991.

[22] Y. Ye, "An $(O(n^3L))$ potential reduction algorithm for linear programming," Mathematical Programming 50 (1991) 239-258.

[23] Y. Ye, M.J. Todd, and S. Mizuno, "An $O(\sqrt{n}\,L)$-Iteration Homogeneous and Self-Dual Linear Programming Algorithm", Technical Report No. 1007, School of Operations Research and Industrial Engineering, Cornell University, 1992.