On the Convergence of the Iteration Sequence in Primal-Dual Interior-Point Methods

R.A. Tapia
Y. Zhang
Y. Te

CRPC-TR91171
August, 1991

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

On the Convergence of the Iteration Sequence in Primal-Dual Interior-Point Methods *

Richard A. Tapia †, Yin Zhang ‡ and Yinyu Ye §

August, 1991

Abstract

This research is concerned with the convergence of the iteration sequence generated by a primal-dual interior-point method for linear programming. It is known that this sequence converges when both the primal and the dual problems have unique solutions. However, convergence for general problems has been an open question now for quite some time. In this work we demonstrate that for general problems, under mild conditions, the iteration sequence converges.

Keywords: Linear programming, Primal-dual interior-point algorithms, Convergence of iteration sequence.

Abbreviated Title: Convergence of Interior-point Methods

*This research was supported in part by NSF Coop. Agr. No. CCR-8809615. A part of this research was performed in June, 1991 while the second and the third authors were at Rice University as visiting members of the Center for Research in Parallel Computation.

†Department of Mathematical Sciences and the Center for Research in Parallel Computation, Rice University, Houston, Texas 77251-1892. Research supported in part by AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.

‡Department of Mathematics and Statistics, University of Maryland, Baltimore County Campus, Baltimore, Maryland 21228. Research supported in part by NSF DMS-9102761 and DOE DE-FG05-91ER25100.

§Department of Management Sciences, The University of Iowa, Iowa City, Iowa 52242. Research supported in part by NSF DDM-8922636.

1 Introduction

This paper considers linear programs in the standard form:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned} \tag{1.1}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n} (m < n)$ and A has full rank m.

The first-order optimality conditions for (1.1) can be written as the following $2n$ by $2n$ nonlinear system with non-negativity constraints on the variables:

$$F(x,y) = \begin{pmatrix} Ax - b \\ By - Bc \\ XYe \end{pmatrix} = 0, \quad (x,y) \geq 0, \tag{1.2}$$

where $B \in \mathbf{R}^{(n-m) \times n}$ is any matrix such that the columns of B^T form a basis for the null space of A, $X = diag(x)$, $Y = diag(y)$ and e is the n-vector of all ones.

The feasibility set of problem (1.2) is defined as

$$\mathcal{F} = \{(x,y) : x, y \in \mathbf{R}^n, Ax = b, By = Bc, (x,y) \geq 0\}.$$

A feasible pair $(x,y) \in \mathcal{F}$ is said to be strictly feasible if it is positive. In this work we assume that the relative interior of $\mathcal{F}$ is nonempty, i.e., strictly feasible points exist. We denote the solution set of problem (1.2) by

$$\mathcal{S} = \{(x^*, y^*) : F(x^*, y^*) = 0, (x^*, y^*) \geq 0\}.$$

The algorithms under consideration in this study are primal-dual interior-point algorithms. These algorithms can be motivated in various ways, e.g., path-following or potential reduction, but in essence they are all variants of Newton's method. The majority of these primal-dual interior-point algorithms are described by the following generic algorithmic framework.

Algorithm 1 (Generic Primal-Dual Algorithm)

Given a strictly feasible pair (x^0, y^0). *For* $k = 0, 1, 2, \ldots,$ *do*

(1) *Choose* $\sigma^k \in [0, 1)$ *and set* $\mu^k = \sigma^k \frac{(x^k)^T y^k}{n}$.

(2) *Solve the following system for* $(\Delta x^k, \Delta y^k)$:

$$F'(x^k, y^k) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -F(x^k, y^k) + \mu^k \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

(3) *Choose a step-length* $\alpha^k = \min(1, \tau^k \hat{\alpha}^k)$ *for* $\tau^k \in (0, 1)$ *and*

$$\hat{\alpha}^k = \frac{-1}{\min((X^k)^{-1}\Delta x^k, (Y^k)^{-1}\Delta y^k)}. \tag{1.3}$$

(4) *Form the new iterate*

$$(x^{k+1}, y^{k+1}) = (x^k, y^k) + \alpha^k(\Delta x^k, \Delta y^k).$$

Note that the choice of step-length α^k guarantees $(x^{k+1}, y^{k+1}) > 0$. Moreover, it can be easily verified that

$$(x^{k+1})^T y^{k+1} = (1 - \alpha^k(1 - \sigma^k))(x^k)^T y^k < (x^k)^T y^k. \tag{1.4}$$

One of the first primal-dual interior-point algorithms was proposed by Kojima, Mizuno and Yoshise [9] in 1987, based on an idea suggested by Megiddo [13] in 1986. Other early primal-dual interior point algorithms include Monteiro and Adler [14] and Todd and Ye [16]. A main feature of these algorithms is their polynomial complexity; or more specifically, their polynomial bound on the number of iterations required to reduce the duality gap to a specified tolerance.

While one group of researchers was mainly concerned with choices of the parameters in Algorithm 1 that led to good complexity bounds, another group was concerned with choices of these parameters that led to numerically effective algorithms. Practical implementations and numerical experiments performed by MacShane, Monma and Shanno [12] and by Lustig, Marsten and Shanno [10], for example, demonstrated that primal-dual algorithms can be computationally very effective. While these formulations exhibited reasonably fast convergence in practice, theoretical convergence of the duality gap sequence, let alone polynomial complexity, remains an open question.

The fast convergence of the latter algorithms, in part, motivated the recent flurry of papers concerned with establishing superlinear convergence of the duality gap to zero. For linear programming applications, these papers include Zhang, Tapia and Dennis [20] and Zhang and Tapia [19]. In addition, Ye, Tapia and Zhang [18], McShane [11] and Ye, Güler, Tapia and Zhang [17] recently studied the superlinear convergence of the Mizuno-Todd-Ye predictor-corrector algorithm that takes $\sigma^k = 1$ and $\sigma^k = 0$ alternatively. For linear complementarity problems, these papers include Zhang, Tapia and Potra [21], Kojima, Kurita and Mizuno [8] and Ji, Potra, Tapia and Zhang [7]. In the context of the present work, it is important to observe that in establishing superlinear convergence of the duality gap (or complementarity) to zero, all of the above authors, except Ye *et al* [17], assumed the convergence of the iteration sequence. For the Mizuno-Todd-Ye predictor-corrector algorithm, Ye *et al* [17] obtained the impressive result of quadratic convergence, when counting the predictor and corrector steps as a single step, without the assumption of nondegeneracy or the assumption of convergence of the iteration sequence. Whether such convergence can be obtained for Algorithm 1 without these assumptions is an important question and is the subject of current research. It is also of concern to us that while in practice the convergence of the iteration sequence is always observed, there is, to our knowledge, no theory in the literature implying that this should be expected for problems that do not have unique solutions. It is the objective of the present paper to fill this conspicuous gap.

This paper is organized as follows. Section 2 contains notation and some known results that will be used in the development of this paper. In Section 3, we first establish several technical results. These technical results then are used to derive a bound on the distance from the $(k+1)$-th iterate to a solution point in terms of the distance from the k-th iterate to this solution point and the centering parameter σ_k. The parametric form of this inequality bears a striking similarity to the form of the bounded deterioration principle for secant updates used by Broyden, Dennis and Moré [1] to establish convergence for several secant methods. For this reason, we shall refer to the distance inequality established in Theorem 3.1 as the bounded deterioration of the iteration sequence. We then derive Theorem 3.2, a useful bound on the search steps of Algorithm 1, from the bounded deterioration of the iteration sequence and the Hoffman lemma [6]. We present the main theorem of this paper, Theorem 4.1, in Section 4. It states that, under mild assumptions, the iteration sequence generated by Algorithm 1 converges. Finally, we make some concluding remarks in Section 5.

2 Preliminaries

In this section, we introduce notation and some known results. For brevity, we use the notation $z = (x, y) \in \mathbf{R}^{2n}$, i.e.,

$$z_i = \begin{cases} x_i, & 1 \leq i \leq n, \\ y_i, & n < i \leq 2n. \end{cases} \tag{2.1}$$

Let $\hat{e} = (0 \dots 0 \quad 1 \dots 1)^T \in \mathbf{R}^{2n}$ where the numbers of zeros and ones are both n. Then Algorithm 1 can be written as the following iterative process:

$$z^{k+1} = z^k - \alpha^k [F'(z^k)]^{-1}(F(z^k) - \mu^k \hat{e}). \tag{2.2}$$

A straightforward calculation gives

$$F'(x, y) = \begin{bmatrix} A & 0 \\ 0 & B \\ Y & X \end{bmatrix}. \tag{2.3}$$

The following proposition is known; however we could not locate a simple proof for the case of linear programming in the literature and therefore, for the sake of completeness, we include a short proof. Observe that no assumptions are made on the solution set $\mathcal{S}$ of Problem (1.2). Recall that by the distance from a point z to a set $\mathcal{S}$, we mean

$$dist(z, \mathcal{S}) = \inf\{\|z - s\| : s \in \mathcal{S}\}.$$

Proposition 2.1 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1. Then*

(i) *$\{z^k\}$ is bounded.*

(ii) *If $(x^k)^T y^k \to 0$, then $dist(z^k, \mathcal{S}) \to 0$.*

Proof: (i) Since $A(x^k - x^0) = 0$, $B(y^k - y^0) = 0$ and $AB^T = 0$,

$$0 = (x^k - x^0)^T (y^k - y^0) = (x^k)^T y^k + (x^0)^T y^0 - (y^0)^T x^k - (x^0)^T y^k.$$

From (1.4),

$$(y^0)^T x^k + (x^0)^T y^k = (x^k)^T y^k + (x^0)^T y^0 \leq 2(x^0)^T y^0.$$

The boundedness of (x^k, y^k) now follows from this inequality and the fact that both (x^0, y^0) and (x^k, y^k) are positive.

(ii) Let $\mathcal{L}$ denote the set of limit points of $\{z^k\}$. Since the duality gap converges to zero, $\mathcal{L} \subset \mathcal{S}$. Given any $\epsilon > 0$, cover $\mathcal{L}$ with

$$\mathcal{B}(\epsilon) = \bigcup_{z^* \in \mathcal{L}} \{z : \|z - z^*\| < \epsilon\}.$$

Since $\{z^k\}$ is bounded, all but a finite number of z^k's are in $\mathcal{B}(\epsilon)$; otherwise, $\{z^k\}$ would have a new limit point that is not in $\mathcal{L}$, and this would be a contradiction. Hence, there exists a smallest number N such that $z^k \in \mathcal{B}(\epsilon)$ for all $k \geq N$. This implies that $dist(z^k, \mathcal{S}) < \epsilon$ for all $k \geq N$ and completes the proof. □

Lemma 2.1 below is Lemma 2 of Güler and Ye [4], tailored to fit the needs of this paper. Many interesting properties of interior-point algorithms follow from this simple fact.

Lemma 2.1 (Güler-Ye)

Let $\{z^k\} = \{(x^k, y^k)\}$ be generated by Algorithm 1. Assume

A1 *$(x^k)^T y^k$ converges to zero.*

A2 $\min(X^k Y^k e)/(x^k)^T y^k \geq \gamma/n$ *for all k and some $\gamma \in (0,1)$.*

Let z^* be a limit point of $\{z^k\}$. Then

(i) z^* *is a strictly complementary solution of Problem 1.2 and*

(ii) $\lim_{k\to\infty} \inf z_i^k > 0$ *for every* i *such that* $z_i^* > 0$.

Recently, El-Bakry, Tapia and Zhang [2] demonstrated that the relative interior of $\mathcal{S}$ and the set of solutions satisfying strict complementarity coincide. Furthermore, the zero-nonzero structure of solutions in the relative interior of $\mathcal{S}$ (equivalently solutions satisfying strict complementarity) is invariant. Hence, under Assumption A2 of Lemma 2.1, the limit points of the iteration sequence generated by Algorithm 1 have considerable structure.

For any $z^* = (x^*, y^*)$ in the relative interior of $\mathcal{S}$, define

$$I_x^+ = \{i : x_i^* > 0, 1 \le i \le n\} \text{ and } I_y^+ = \{i : y_i^* > 0, 1 \le i \le n\}. \tag{2.4}$$

Since the zero-nonzero structure of the relative interior of $\mathcal{S}$ is invariant, the above two index sets are independent of the choice of z^*. By strict complementarity of z^*,

$$I_x^+ \bigcup I_y^+ = \{1, 2, ..., n\} \text{ and } I_x^+ \bigcap I_y^+ = \emptyset. \tag{2.5}$$

3 Bounded Deterioration

We consider the following three matrices which were introduced in Zhang, Tapia and Dennis [20]:

$$\begin{aligned} P &= I - X^{\frac{1}{2}}Y^{-\frac{1}{2}}A^T(AXY^{-1}A^T)^{-1}AX^{\frac{1}{2}}Y^{-\frac{1}{2}}, & (3.1)\\ H_p &= (XY)^{-\frac{1}{2}}P(XY)^{-\frac{1}{2}}, & (3.2)\\ H_d &= (XY)^{-\frac{1}{2}}(I-P)(XY)^{-\frac{1}{2}}. & (3.3)\end{aligned}$$

It is worth noting that P and $I - P$ are orthogonal projection matrices.

Lemma 3.1 *Let* H_p *and* H_d *be defined by (3.2) and (3.3), respectively, then*

$$F'(x,y)^{-1}\begin{pmatrix} 0 \\ w \end{pmatrix} = \begin{pmatrix} XH_pw \\ YH_dw \end{pmatrix}.$$

Proof: The proof follows from Proposition 2.1 in Zhang, Tapia and Dennis [20]. □

Lemma 3.2 *Let $(x^*, y^*) \in \mathcal{S}$ and $(x, y) \in \mathcal{F}$. Then*

$$F'(x,y)\begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} - F(x,y) = \begin{pmatrix} 0 \\ (X - X^*)(Y - Y^*)e \end{pmatrix}.$$

Proof: The proof follows from direct substitution. □

Lemma 3.3 *Let $z^* = (x^*, y^*) \in \mathcal{S}$, $z = (x, y) \in \mathcal{F}$ and $z^+ = (x^+, y^+)$ be given by*

$$z^+ = z - \alpha[F'(z)]^{-1}(F(z) - \mu\hat{e}).$$

Then

$$z^+ - z^* = \begin{pmatrix} XH_p w \\ YH_d w \end{pmatrix},$$

where

$$w = (X - X^*)(Y - Y^*)e + (1 - \alpha)XYe + \alpha\mu e. \tag{3.4}$$

Proof: Using Lemma 3.2, we have

$$\begin{aligned} z^+ - z^* &= z - z^* - \alpha[F'(z)]^{-1}(F(z) - \mu\hat{e}) \\ &= [F'(z)]^{-1}[F'(z)(z - z^*) - F(z) + (1 - \alpha)F(z) + \alpha\mu\hat{e}] \\ &= [F'(z)]^{-1}\begin{pmatrix} 0 \\ w \end{pmatrix}. \end{aligned}$$

Now the lemma follows from Lemma 3.1. □

Lemma 3.4 *Let P_i be the i-th row (or column) of an orthogonal projection matrix P. Then*

$$\|P_i\|_2 \leq 1.$$

Proof: Since $I - P$ is symmetric and positive semi-definite, $P_{ii} \leq 1$. Now $P^2 = P$ leads to $\|P_i\|_2^2 = P_{ii} \leq 1$. □

We now derive our bounded deterioration relationship between the distances from two consecutive iterates, z^k and z^{k+1}, to any solution z^*.

Theorem 3.1 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1 and let $z^* = (x^*, y^*)$ be any solution to Problem 1.2. Assume*

A1 $(x^k)^T y^k \to 0$.

A2 $\min(X^k Y^k e)/(x^k)^T y^k \geq \gamma/n$ *for all k and some $\gamma \in (0,1)$.*

Then there exist constants $\beta_i' > 0$ and $\beta_i'' > 0$ such that

$$|z_i^{k+1} - z_i^*| \leq (1 + z_i^k \beta_i')\|z^k - z^*\| + z_i^k \beta_i'' \sigma^k, \quad i = 1, 2, \ldots, 2n. \tag{3.5}$$

Proof: For simplicity, we will replace the superscript $(k+1)$ by $+$ and drop the superscript k. We let $\hat{z} = z + \Delta z$. Since $z^+ = z + \alpha \Delta z$ and $\alpha \in (0, 1]$, for every index i we have

$$z_i^+ \in [\min(z_i, \hat{z}_i), \max(z_i, \hat{z}_i)].$$

Therefore, for z^*

$$|z_i^+ - z_i^*| \leq \max(|z_i - z_i^*|, |\hat{z}_i - z_i^*|). \tag{3.6}$$

Without loss of generality, let us consider $i \leq n$, i.e., $z_i = x_i$, $z_i^* = x_i^*$ and so on. The proof for $i > n$ is the same.

Letting $\alpha = 1$ in Lemma 3.3 and using the definition of H_p and w, we obtain

$$\hat{x}_i - x_i^* = x_i \sum_{j=1}^{n} (x_i y_i)^{-\frac{1}{2}} P_{ij} (x_j y_j)^{-\frac{1}{2}} w_j = x_i (S_i' + S_i''), \tag{3.7}$$

where

$$S_i' = \sum_{j=1}^{n} (x_i y_i)^{-\frac{1}{2}} P_{ij} (x_j y_j)^{-\frac{1}{2}} (x_j - x_j^*)(y_j - y_j^*),$$

and

$$S_i'' = \sigma \sum_{j=1}^{n} (x_i y_i)^{-\frac{1}{2}} P_{ij} (x_j y_j)^{-\frac{1}{2}} \frac{x^T y}{n}.$$

In S_i'' we used the relation $\mu = \sigma x^T y / n$. Let us consider the above two sums. Notice that

$$(x_j - x_j^*)(y_j - y_j^*) = \begin{cases} y_j (x_j - x_j^*), & \text{if } j \in I_x^+, \\ x_j (y_j - y_j^*), & \text{if } j \in I_y^+. \end{cases}$$

By strict complementarity, for each index j either $x_j^* > 0$ or $y_j^* > 0$. Therefore,

$$S_i' = \left[\sum_{j \in I_x^+} P_{ij} \left(\frac{x_j y_j}{x_i y_i} \right)^{\frac{1}{2}} \frac{x_j - x_j^*}{x_j} + \sum_{j \in I_y^+} P_{ij} \left(\frac{x_j y_j}{x_i y_i} \right)^{\frac{1}{2}} \frac{y_j - y_j^*}{y_j} \right].$$

Notice that

$$\frac{x_j y_j}{x_i y_i} = \frac{x_j y_j}{x^T y} \frac{x^T y}{x_i y_i} \le \frac{n}{\gamma}$$

is bounded above and $\{x_j,\ j \in I_x^+\}$ and $\{y_j,\ j \in I_x^+\}$ are bounded away from zero (see (ii) of Lemma 2.1). Moreover, the matrix P is an orthogonal projection matrix; by Lemma 3.4, its rows are bounded (independent of k). Therefore, we can find $\beta_i' > 0$ such that

$$|S_i'| \le \beta_i' \|z - z^*\|.$$

In addition, by Assumption A2 and Lemma 3.4 again, we can find $\beta_i'' > 0$ such that

$$|S_i''| = \sigma \left(\frac{x^T y / n}{x_i y_i} \right)^{\frac{1}{2}} \left| \sum_{j=1}^{n} P_{ij} \left(\frac{x^T y / n}{x_j y_j} \right)^{\frac{1}{2}} \right| \le \beta_i'' \sigma.$$

It follows from (3.7) that

$$|\hat{x}_i - x_i^*| \le x_i (|S_i'| + |S_i''|) \le x_i (\beta_i' \|z - z^*\| + \beta_i'' \sigma).$$

Finally, (3.5) follows from (3.6) and the above inequality. This completes the proof. □

The following inequality, a consequence of the bounded deterioration and the Hoffman Lemma [6], will be a critical ingredient in the establishment of our convergence theory.

Theorem 3.2 *Let $\{z^k = (x^k, y^k)\}$ and $\{\Delta z^k\}$ be generated by Algorithm 1. Assume*

A1 $(x^k)^T y^k \to 0$.

A2 $\min(X^k Y^k e)/(x^k)^T y^k \ge \gamma/n$ *for all k and some $\gamma \in (0,1)$.*

Then there exist constants $\beta' > 0$ and $\beta'' > 0$ such that

$$\|\Delta z^k\| \le \beta' (x^k)^T y^k + \beta'' \sigma^k. \tag{3.8}$$

and, in particular,

$$\|z^{k+1} - z^k\| \le \beta' (x^k)^T y^k + \beta'' \sigma^k. \tag{3.9}$$

Proof: Since $z^{k+1} = z^k + \alpha^k \Delta z^k$ and $\alpha^k \in (0,1]$, it suffices to prove (3.8). Notice that the bounded deterioration (3.5) holds for any $\alpha^k \in (0,1]$. When $\alpha^k = 1$, from bounded deterioration (3.5) we have for any $z^* \in \mathcal{S}$

$$\|\Delta z^k\| \leq \|(z^k + \Delta z^k) - z^*\| + \|z^k - z^*\| \leq \hat{\beta}\|z^k - z^*\| + \beta''\sigma^k. \tag{3.10}$$

It follows from the Hoffman Lemma [6] that there exist a constant $\beta > 0$ and $w^k \in \mathcal{S}$ for every k such that

$$\|z^k - w^k\| \leq \beta (x^k)^T y^k.$$

Since $z^* \in \mathcal{S}$ is arbitrary in (3.10), it follows that (3.8) holds with $\beta' = \beta\hat{\beta}$. This proves the theorem. □

Observe that when $\sigma^k = 0$, Theorem 3.2 reduces to Lemma 3.2 and Theorem 3.1 of Ye *et al* [17], i.e., a bound on the Newton step. Theorem 3.2 can also be proved using this bound on the Newton step and the bound on $|S_i''|$ in the proof of Theorem 3.1.

4 Convergence Results

Theorem 4.1 below represents our main contribution to the convergence theory for the iteration sequence generated by primal-dual interior-point methods for linear programming.

Given any z^* in the relative interior of $\mathcal{S}$, let

$$I_0 = \{i : z_i^* = 0, \quad 1 \leq i \leq 2n\}.$$

From the remarks above, I_0 is invariant with respect to the choices of z^*.

Recall that a set in $\mathbf{R}^n$ is said to be a continuum if it is closed and cannot be written as the disjoint union of two proper closed subsets. In $\mathbf{R}^1$ the continua are closed intervals. Clearly, a continuum is uncountable.

The following assumptions will be selectively used in Theorem 4.1. Their compatibility will be demonstrated in Proposition 4.1.

A1 $(x^k)^T y^k \to 0$;

A2 $\min(X^kY^ke)/(x^k)^Ty^k \geq \gamma/n$ for all k and some $\gamma \in (0,1)$.

A3 $\sigma^k \to 0$.

A4 $\sigma^k \to 0$ at least R-linearly.

A5 $\sigma^k = O((x^k)^Ty^k)$.

A6 $\tau^k \geq \tau$ for some $\tau \in (0,1)$.

Theorem 4.1 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1. Then*

(i) *Under Assumptions A1-A2, $\{z_i^k\} \to 0$ for each $i \in I_0$.*

(ii) *Under Assumptions A1-A3, either*

- *$\{z^k\}$ converges to some z^* in the relative interior of $\mathcal{S}$, or*
- *the limit points of $\{z^k\}$ form a continuum.*

(iii) *Under Assumptions A1, A2, A4 and A6, $\{z^k\}$ converges to some z^* in the relative interior of $\mathcal{S}$.*

(iv) *Under Assumptions A1, A2, A5 and A6, any Q or R convergence behavior present in the convergence of $\{(x^k)^Ty^k\}$ to zero is reflected as R-convergence behavior in the convergence of $\{z^k\}$ to z^*.*

Proof: (i) Consider the component sequence $\{z_i^k\}$ for $i \in I_0$. It is bounded from (i) of Proposition 2.1. By Lemma 2.1 and El-Bakry, Tapia and Zhang [2], all limit points of $\{z^k\}$ are in the relative interior of $\mathcal{S}$ and have identical zero-nonzero structure; thus all limit points of $\{z_i^k\}$ are zero. This leads to $z_i^k \to 0$ for all $i \in I_0$.

(ii) Again let $\mathcal{L}$ denote the set of limit points of $\{z^k\}$. Clearly, $\mathcal{L}$ is compact. It is obviously closed and it is bounded since $\{z^k\}$ is bounded by (i) of Proposition 2.1. If $\mathcal{L}$ is the singleton set $\{z^*\}$, then $\{z^k\}$ converges to z^* since it is a bounded sequence with only one limit point.

Now suppose that $\mathcal{L}$ has more than one point and is not a continuum. Then we can write $\mathcal{L} = \mathcal{L}_1 \bigcup \mathcal{L}_2$ where $\mathcal{L}_i$, $i = 1, 2$, is nonempty and compact. By Hausdorff separation we can find sets $\mathcal{O}_i$, $i = 1, 2$, such that $\mathcal{O}_i$ is open, $\mathcal{L}_i \subset \mathcal{O}_i$ and the closures of $\mathcal{O}_1$ and $\mathcal{O}_2$ are disjoint. Clearly, $\mathcal{O}_i$ contains an infinite number of z^k's, or else, $\mathcal{L}_i$ would be empty; also only a finite number of z^k's are not in $\mathcal{O}_1 \bigcup \mathcal{O}_2$. This latter statement follows from the observation that if an infinite number of z^k's are not contained in $\mathcal{O}_1 \bigcup \mathcal{O}_2$, then they would have a limit point not in $\mathcal{O}_1 \bigcup \mathcal{O}_2$; and therefore not in $\mathcal{L}$. This would be a contradiction. It follows that we can select a subsequence $\{z^{k_j}\}$ with the property that $z^{k_j} \in \mathcal{O}_1$ and $z^{k_j+1} \in \mathcal{O}_2$. We can choose a thinner subsequence, which we also call $\{z^{k_j}\}$, with the property that $z^{k_j} \longrightarrow z^*$ for some $z^* \in \mathcal{L}_1$ and $z^{k_j+1} \in \mathcal{O}_2$ for all j. However, this contradicts our bounded deterioration inequality (3.5) of Theorem 3.1.

(iii) Since the assumptions of Theorem 3.2 hold, we have inequality (3.8) at our disposal. It is known (see Zhang, Tapia and Dennis [20], for example) that A2 and A6 imply $\alpha^k \geq \alpha > 0$ for all k and some α. From (1.4) it follows that the duality gap sequence converges Q-linearly to zero. It now follows from (3.8) and A4 that there exist positive constants η and $\delta < 1$ and a positive integer K such that

$$\|z^{k+1} - z^k\| \leq \eta\delta^k, \quad k \geq K. \tag{4.1}$$

In a standard manner, (4.1) implies that for all positive integers N

$$\|z^{k+N} - z^k\| \leq \frac{\eta}{1-\delta}\delta^k, \quad k \geq K. \tag{4.2}$$

Hence, $\{z^k\}$ is a Cauchy sequence and is therefore convergent.

(iv) The proof of (iv) follows from the derivation of the appropriate inequality analogous to (4.2) and letting $N \rightarrow \infty$. This proves the theorem. □

It is worth pointing out the implications of Theorem 4.1 on the superlinear and quadratic convergence theory of Zhang, Tapia and Dennis [20]. First, (iii) implies that the nondegeneracy assumption for quadratic convergence is no longer necessary. Second, (ii) implies that the assumption of the convergence of the iteration sequence made for superlinear convergence of the duality gap to zero is no longer necessary, due to the fact that in order to establish superlinear convergence one only needs $\|z^{k+1} - z^k\| \to 0$. These facts will be documented in a subsequent report currently in preparation.

It is perhaps interesting to observe that in the proof of (ii) by partitioning $\mathcal{L}$ so that $\mathcal{L}_1$ is a singleton set we have established that either $\{z^k\}$ converges or $\mathcal{L}$ has no isolated points. In analysis, a closed set in which every point is a limit point of the set is called a perfect set. Clearly, $\mathcal{L}$ is a perfect set in the case of nonconvergence. It is known that in a separable metric space, every perfect set must be uncountable. See Gelbaum and Olmsted [3] for a discussion and see Hausdorff [5] for a proof and further discussion. The interval $[0, 1]$ is a perfect set in $\mathbf{R}^1$. One might conjecture that a perfect set in $\mathbf{R}^1$ contains a continuum, or at least has positive Lebesgue measure. However, in $\mathbf{R}^1$ the celebrated Cantor set is a perfect set of Lebesgue measure zero; hence it contains no continuum. It follows that this line of reasoning would produce a weaker result in (ii).

It is also interesting to note that (ii) of Theorem 4.1 has the flavor of a result established by Ostrowski [15] in the context of the gradient method.

The quantities τ^k and σ^k are under direct control of the algorithm designer. This is an extremely important point and often missed by some readers. Even without the knowledge of (iii), it seems that it would be quite pathological for one to choose $\sigma^k \to 0$ in any manner that did not reflect R-linear convergence.

We now demonstrate the compatibility of Assumptions A1-A6 by showing that choices for σ^k and α^k exist such that these assumptions are simultaneously satisfied.

Proposition 4.1 *For Algorithm 1 there exist parameter sequences $\{\sigma^k\}$ and $\{\alpha^k\}$, and the corresponding iteration sequence $\{(x^k, y^k)\}$ such that Assumptions A1-A6 are satisfied.*

Proof: It suffices to demonstrate that Assumptions A1, A2, A5 and A6 are simultaneously satisfied. Let

$$\sigma^k = \min(\sigma, \rho(x^k)^T y^k) \tag{4.3}$$

where $\sigma \in (0, 1)$ and $\rho > 0$. For each k choose τ^k so that $\alpha^k = \tau^k \hat{\alpha}^k$ is the largest number in $(0,1]$ such that

$$\gamma \le \frac{x_i^{k+1} y_i^{k+1}}{(x^{k+1})^T y^{k+1}/n} \le \Gamma, \quad i = 1, 2, \ldots, n, \tag{4.4}$$

for some fixed constants $\gamma \in (0, 1/2)$ and $\Gamma \in (2, n)$. It has been shown (see Lemma 3.4 of Zhang and Tapia [19]) that

$$\alpha^k \ge \min\left(1, \frac{\sigma^k (x^k)^T y^k/n}{2 \max |\Delta x_i^k \Delta y_i^k|}\right). \tag{4.5}$$

The proposition will be proved if we can show that $\{\alpha^k\}$ is bounded away from zero.

When $\sigma^k = \sigma$, it is known that $\alpha^k = \Omega(\frac{1}{n})$ (see [19], for example). Otherwise, $\sigma^k = \rho(x^k)^T y^k$. By Theorem 3.2, $|\Delta x_i^k \Delta y_i^k| = O(((x^k)^T y^k)^2)$. Hence it follows from (4.5) that α^k is also bounded away from zero. This completes the proof. □

5 Concluding Remarks

The main result of this paper is Theorem 4.1. Among other things, it says that $\sigma^k \to 0$ will guarantee that the step $z^{k+1} - z^k$ converges to zero; but this alone does not imply the convergence of the iteration sequence. However, this convergence is guaranteed if the step converges to zero fast enough, i.e., at least R-linearly, and this in turn is guaranteed if σ^k converges to zero R-linearly. This requirement can be viewed as saying that in order to ensure the convergence of the iteration sequence, eventually any commitment to the central path should be phased out in an effective manner.

We now argue the mildness of the assumptions made in Theorem 4.1. Assumption A1, the convergence of the duality gap sequence to zero, is necessary if we are to consider convergence to a solution. Assumption A2 is a crucial ingredient for polynomial convergence of the duality gap sequence to zero. Moreover, Assumptions A2, A3 and $\tau^k \to 1$ are fundamental in establishing Q-superlinear convergence (see Zhang, Tapia and Dennis [20]).

In summary, we find it extremely satisfying that the only price one pays for obtaining convergence of the iteration sequence, in addition to superlinear convergence of the duality gap to zero, is to guarantee at least R-linear behavior while making $\sigma^k \to 0$. This is indeed very mild. If this simple requirement is not implemented, then, at least in theory and at this juncture, the iteration sequence could (but seems to be highly unlikely) exhibit the unusual behavior that it generates a continuum of limit points.

Acknowledgment

We thank Jorge Moré for referring us to the Ostrowski result.

References

[1] C. G. Broyden, and J. E. Dennis Jr. and J. J. Moré. On the local and superlinear convergence of Quasi-Newton methods. *J. Inst. Math. Appl.*, 12:223-245, 1973.

[2] A. El-Bakry, R. A. Tapia, and Yin Zhang. On the use of indicators in identifying zero variables for interior point methods. Technical Report TR91-15, Dept. Mathematical Sciences, Rice University, 1991.

[3] B. R. Gelbaum and John M. H. Olmsted. *Counterexamples in Analysis.* Holden-Day Inc., San Francisco, 1964.

[4] O. Güler and Y. Ye. Convergence behavior of some interior-point algorithms. Working Paper Series No. 91-4, Dept. of Management Sciences, The University of Iowa, 1991.

[5] F. Hausdorff. *Mengenlehre.* Dover Publications, New York, 1944.

[6] A. J. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49:263-265, 1952.

[7] J. Ji, F. Potra, R. A. Tapia, and Y. Zhang. An interior-point method for linear complementarity problems with polynomial complexity and superlinear convergence. Technical Report TR91-23, Dept. Mathematical Sciences, Rice University, 1991.

[8] M. Kojima, Y. Kurita and S. Mizuno. Large-step interior point algorithms for linear complementarity problems. Research Report on Information Sciences B-243, Dept. of Information Sciences, Tokyo Institute of Technology, 1991.

[9] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point method for linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 29–47. Springer-Verlag, New York, 1989.

[10] I.J. Lustig, R.E. Marsten, and D.F. Shanno. Computational experience with a primal-dual interior point method for linear programming. Technical Report SOR 89-17, Dept. Civil Eng. and O.R., Princeton University, 1989.

[11] K. McShane. A superlinearly convergent $O(\sqrt{n}L)$-iteration primal-dual linear programming algorithm. Manuscript, 2534 Villanova Drive, Vienna, Virginia, 1991.

[12] K.A. McShane, C.L. Monma, and D.F. Shanno. An implementation of a primal-dual interior point method for linear programming. *ORSA J. Computing*, 1:70–83, 1989.

[13] N. Megiddo. Pathways to the optimal set in linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 131–158. Springer-Verlag, New York, 1989.

[14] R.D.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part I: linear programming. *Math. Prog.*, 44:27–41, 1989.

[15] A. M. Ostrowski. *Solution of Equations and Systems of Equations.* Academic Press, New York, 1966.

[16] M.J. Todd and Y. Ye. A centered projective algorithm for linear programming. *Math. of O.R.*, 15:508–529,1990.

[17] Y. Ye, O. Güler, R. A. Tapia and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-26, Dept. Mathematical Sciences, Rice University, 1991.

[18] Y. Ye, R. A. Tapia and Y. Zhang. A superlinearly convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-22, Dept. Mathematical Sciences, Rice University, 1991.

[19] Y. Zhang and R. A. Tapia. A quadratically convergent polynomial primal-dual interior-point algorithm for linear programming. Technical Report No. 90-40, Dept. of Mathematical Sciences, Rice University, 1990.

[20] Y. Zhang, R. A. Tapia and J. E. Dennis. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. Technical Report TR90-6, Dept. Mathematical Sciences, Rice University, 1990, to appear in *SIAM Journal on Optimization.*

[21] Y. Zhang, R. A. Tapia and F. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. Technical Report TR90-9, Dept. Mathematical Sciences, Rice University, 1990, to appear in *SIAM Journal on Optimization.*