A SUPERLINEAR INFEASIBLE-INTERIOR-POINT AFFINE SCALING ALGORITHM FOR LCP

R. D. C. MONTEIRO* AND S. J. WRIGHT†

Abstract. We present an infeasible-interior-point algorithm for monotone linear complementarity problems in which the search directions are affine scaling directions and the step lengths are obtained from simple formulae that ensure both global and superlinear convergence. By choosing the value of a parameter in appropriate ways, polynomial complexity and convergence with Q-order up to (but not including) two can be achieved. The only assumption made to obtain the superlinear convergence is the existence of a solution satisfying strict complementarity.

Key words. infeasible-interior-point methods, monotone linear complementarity problems, superlinear convergence

1. Introduction. The monotone linear complementarity problem (LCP) is to find a vector pair $(x, y) \in \mathsf{R}^n \times \mathsf{R}^n$ that satisfies the following conditions:

$$(1.1a) \qquad y = Mx + q,$$

$$(1.1b) \qquad x \geq 0, \;\; y \geq 0,$$

$$(1.1c) \qquad x^T y = 0,$$

where M is a positive semidefinite matrix. We use $\mathcal{S}$ to denote the solution set of (1.1) and $\mathcal{S}_c$ to denote the set of solutions that satisfy strict complementarity, that is,

$$\mathcal{S}_c = \{(x^*, y^*) \in \mathcal{S} \,|\, x^* + y^* > 0\}.$$

We say that (1.1) is *nondegenerate* if $\mathcal{S}_c \neq 0$.

It is well known that problems in linear programming and convex quadratic programming can be posed in the form (1.1). This fact partly accounts for the extensive effort that has gone into designing interior-point algorithms for this class of problems.

In this paper, we develop a superlinearly convergent infeasible-interior-point algorithm that uses only affine scaling directions. Several infeasible-interior-point algorithms have been developed to solve problem (1.1), or the linear programming (LP) subclass. For LP problems, these include works by Kojima, Megiddo, and Mizuno [1], Kojima, Megiddo, and Todd [3], Mizuno [2], Potra [6], and Zhang and Zhang [12]. For LCP, we mention Wright [7, 8], while Zhang [11] considers a generalized form of the monotone LCP.

In a companion paper [4], we describe a method in which all iterates (x^k, y^k) are *strictly feasible* (that is, $(x^k, y^k) > 0$ and $y^k = Mx^k + q$) and search directions are generated by solving the linear system

$$(1.2) \qquad \begin{bmatrix} M & -I \\ Y^k & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta y^k \end{bmatrix} = \begin{bmatrix} 0 \\ -X^k Y^k e \end{bmatrix},$$

where $X^k = \text{diag}(x_1^k, x_2^k, \ldots, x_n^k)$, $Y^k = \text{diag}(y_1^k, y_2^k, \ldots, y_n^k)$, and $e = (1, 1, \ldots, 1)^T$. The direction $(\Delta x^k, \Delta y^k)$ is known as the primal-dual affine scaling direction, and one can easily see that it is simply the Newton direction for the nonlinear mapping $F(x, y) = (Mx + q - y, XYe) = (0, 0)$ at the point (x^k, y^k). In [4], we choose a parameter $\delta \in (0, 1/2)$ and define the scalars ϕ_k and χ_k by

$$(1.3) \qquad \phi_k \triangleq \frac{\Delta x^{k^T} \Delta y^k}{x^{k^T} y^k}, \qquad \chi_k \triangleq \max \left(0, \max_{\substack{i=1,2,\ldots,n \\ \Delta x_i^k \Delta y_i^k < 0}} \left(\frac{|\Delta x_i^k \Delta y_i^k|}{x_i^k y_i^k} \right) \right)$$

and the step length α_k by

$$(1.4) \qquad \alpha_k \triangleq \frac{\delta}{\delta + \chi_k + (1 + \delta)\phi_k}.$$

The step from one iterate to the next is defined as

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

The resulting simple algorithm is globally and superlinearly convergent. It even exhibits polynomial complexity for certain choices of the parameter δ and starting point (x^0, y^0).

In this paper, we generalize this algorithm by allowing the iterates to be *infeasible*. Given an initial point $(x^0, y^0) > 0$, the modified algorithm generates iterates (x^k, y^k) in such a way that $x^{k^T} y^k \downarrow 0$ and $\|y^k - Mx^k - q\| \to 0$ as $k \to \infty$. Moreover, the convergence of $\{x^{k^T} y^k\}$ to zero is superlinear. Much of the elegance of the analysis for the feasible algorithm [4] does not carry through to this infeasible case and the results are slightly weaker here, so we feel that a separate report is needed.

Our analysis makes use of the fact that all iterates generated by our algorithm lie in a set of the form

$$\mathcal{N}(\delta, \eta, \rho) = \{(x, y) > 0 \mid \rho x^T y \geq \|r\|^{1+\delta}, \ x_i y_i \geq \eta (x^T y)^{1+\delta}, \ i = 1, 2, \ldots, n\},$$

where $\delta > 0$, $\eta > 0$, and $\rho \geq 0$, and $r = y - Mx - q$. Note that $\mathcal{N}(\delta, \eta, 0)$ corresponds to the neighborhood $\mathcal{N}(\delta, \eta)$ that was used in [4].

We formally specify our algorithm as follows.

Algorithm IAS
initially: Choose $\delta \in (0, 1]$ and $(x^0, y^0) > 0$;
for $k = 0, 1, 2, \ldots$
    Let $(\Delta x^k, \Delta y^k)$ denote the solution of the linear system

$$(1.5a) \qquad M\Delta x - \Delta y = r,$$
$$(1.5b) \qquad Y\Delta x + X\Delta y = -YXe,$$

    where $r = r^k = y^k - Mx^k - q$, $X = X^k$, and $Y = Y^k$;
    Calculate ϕ_k and χ_k as in (1.3) and set

$$(1.6) \qquad \alpha_k = \frac{\delta}{\delta + \chi_k + (1 + \delta)\phi_k^+},$$

where $\phi_k^+ = \max(0, \phi_k)$;

Set $(x^{k+1}, y^{k+1}) = (x^k + \alpha_k \Delta x^k, y^k + \alpha_k \Delta y^k)$;

end for

Note that in the case of (x^k, y^k) feasible, we have

$$\Delta x^{k^T} \Delta y^k = \Delta x^{k^T} M \Delta x^k \geq 0.$$

Therefore $\phi_k \geq 0$ for all k, so the definition of α_k in (1.6) reduces to (1.4). Hence, the algorithm in [4] is a special case of Algorithm IAS.

In the remainder of the paper, we prove that Algorithm IAS produces a sequence of iterates $\{(x^k, y^k)\}$ such that the complementarity gap $x^{k^T} y^k$ and the residual $\|r^k\|$ converge to zero. Moreover, we show that this convergence is Q-superlinear provided that $\mathcal{S}_c \neq \emptyset$ and $\delta < 1/7$. An interesting feature of our superlinear convergence result is that we do not assume any condition that would imply the boundedness of the sequence $\{(x^k, y^k)\}$ such as (i) existence of a strictly feasible point or (ii) convergence of $\{x^{k^T} y^k\}$ and $\{\|r^k\|\}$ to 0 at a similar rate. Most superlinear convergence results for infeasible-interior-point algorithms require the first condition. Exceptions are the results of Potra [6] and Wright [8], which require only the second condition; specifically, for some $\gamma > 0$ we have $x^{k^T} y^k \geq \gamma \|r^k\|$ for all $k \geq 0$. In this paper, we show first that $\{x^{k^T} y^k\}$ converges superlinearly to zero then, as a consequence, that the sequence of iterates $\{(x^k, y^k)\}$ converges and hence is bounded. Also, for certain values of δ and for initial points (x^0, y^0) satisfying $(x^0, y^0) \geq (x^*, y^*)$, for some $(x^*, y^*) \in \mathcal{S}$, we show that the number of iterations required to reduce the gap $x^{k^T} y^k$ below a certain $\epsilon > 0$ is polynomially bounded.

The following notational conventions are used in the remainder of the paper. When $\mathcal{S}_c \neq \emptyset$, we can partition the index set $\{1, 2, \ldots, n\}$ into subsets B and N such that

$$i \in B \quad \Leftrightarrow \quad x_i^* > 0, \ \forall (x^*, y^*) \in \mathcal{S}_c, \qquad i \in N \quad \Leftrightarrow \quad y_i^* > 0, \ \forall (x^*, y^*) \in \mathcal{S}_c.$$

Unless otherwise specified, $\|.\|$ denotes the the Euclidean norm. For a general vector $z \in \mathsf{R}^n$ and index set $L \subseteq \{1, 2, \ldots, n\}$, z_L denotes the vector made up of components z_i for $i \in L$. If $M \in \mathsf{R}^{n \times n}$ and $I, L \subseteq \{1, 2, \ldots, n\}$, then M_{IL} refers to the submatrix of M consisting of the elements M_{ij} for $i \in I$ and $j \in L$. If $D \in \mathsf{R}^{n \times n}$ is diagonal, then D_B denotes the diagonal matrix constructed from D_{ii} for $i \in B$.

2. Global Convergence. In this section, we prove the global convergence of Algorithm IAS. We show that for certain values of δ, Algorithm IAS has a polynomial bound on the number of iterations. We also obtain certain bounds on the directions generated by the algorithm, which will be used in the next section to prove superlinear convergence.

For much of the analysis, it is convenient to drop the iteration index k that appears on quantities such as (x^k, y^k), ϕ_k, and α_k. Accordingly, we define some index-free notation. For an arbitrary positive vector pair $(x, y) > 0$, let $(\Delta x, \Delta y)$ be obtained from (1.5), and define

$$(x(\alpha), y(\alpha)) \quad \triangleq \quad (x, y) + \alpha(\Delta x, \Delta y),$$

$$r(\alpha) \quad \triangleq \quad y(\alpha) - M x(\alpha) - q,$$

(2.1) $$\phi \quad \triangleq \quad \frac{\Delta x^T \Delta y}{x^T y},$$

$$\chi \quad \stackrel{\triangle}{=} \quad \max\left(0, \max_{\substack{i=1,2,\ldots,n \\ \Delta x_i \Delta y_i < 0}} \left(\frac{|\Delta x_i \Delta y_i|}{x_i y_i}\right)\right),$$

where $\alpha \in [0,1]$.

We start by stating and proving some basic results for quantities in (2.1).

LEMMA 2.1. $\phi \geq -\chi$.

Proof. We have

$$\phi = \frac{\Delta x^T \Delta y}{x^T y} \geq \sum_{\substack{i=1,2,\ldots,n \\ \Delta x_i \Delta y_i < 0}} \frac{\Delta x_i \Delta y_i}{x^T y} \geq -\frac{\chi}{x^T y} \sum_{\substack{i=1,2,\ldots,n \\ \Delta x_i \Delta y_i < 0}} x_i y_i \geq -\chi.$$

∎

LEMMA 2.2. *Let $(x,y) > 0$ be given, and let $(\Delta x, \Delta y)$ be the direction determined by (1.5). Then, for all $\alpha \in [0,1]$,*

(a) $x(\alpha)^T y(\alpha) = x^T y(1 - \alpha + \alpha^2 \phi)$;

(b) $x_i(\alpha) y_i(\alpha) \geq x_i y_i (1 - \alpha - \alpha^2 \chi)$;

(c) $r(\alpha) = (1 - \alpha) r$;

(d) $\Delta x_i \Delta y_i = 0$ *for all $i = 1, 2, \ldots, n$ if and only if $(x(1), y(1))$ is a solution of (1.1) that satisfies strict complementarity.*

Proof. Statements (a), (b), and (c) follow immediately from (1.5) and (2.1). The proof of (d) is essentially given by Monteiro and Wright [4, Lemma 3.1(d)], if we note in addition that $r(1) = 0$. ∎

We now show that $\{x^{k^T} y^k\}$ is a decreasing sequence and that the property of membership of $\mathcal{N}(\delta, \eta, \rho)$ is inherited by successive iterates of the algorithm. In the proof of this result, we make use of the inequality

(2.2) $$(1 - u)^{1+\delta} \leq 1 - (1 + \delta)u + \delta u^2, \qquad \forall u \leq 1, \qquad \forall \delta \in (0,1],$$

which is proved by Monteiro and Wright [4, Lemma 3.2].

LEMMA 2.3. *Suppose that $(x,y) \in \mathcal{N}(\delta, \eta, \rho)$ for some $\delta \in (0,1]$, $\eta > 0$, and $\rho \geq 0$ and that $(\Delta x, \Delta y)$ is determined by (1.5). Assume that $\Delta x_i \Delta y_i \neq 0$ for some $i \in \{1, 2, \ldots, n\}$. Then, for all*

(2.3) $$\alpha \in J \stackrel{\triangle}{=} \left[0, \frac{\delta}{\delta + \chi + (1+\delta)\phi^+}\right],$$

where $\phi^+ \equiv \max(0, \phi)$, we have

(a) $x(\alpha)^T y(\alpha) \leq [1 - \alpha/(1+\delta)] x^T y$; *and*

(b) $(x(\alpha), y(\alpha)) \in \mathcal{N}(\delta, \eta, \rho)$.

Proof. To show statement (a), note that for $\alpha \in J$, we have

(2.4) $$\alpha \phi^+ \leq \frac{\delta \phi^+}{\delta + \chi + (1+\delta)\phi^+} \leq \frac{\delta}{1+\delta}.$$

Thus in view of Lemma 2.2(a), we obtain

$$x(\alpha)^T y(\alpha) \leq x^T y(1 - \alpha(1 - \alpha\phi^+)) \leq x^T y \left(1 - \frac{\alpha}{1+\delta}\right).$$

For statement (b), we start by showing that $x(\alpha)^T y(\alpha) > 0$ for all $\alpha \in J$. Observe first that the condition that $\Delta x_i \Delta y_i \neq 0$ for some $i \in \{1, 2, \ldots, n\}$ implies that either $\chi > 0$ or $\phi > 0$. Hence, $\chi + (1 + \delta)\phi^+ > 0$, and therefore

$$(2.5) \qquad \alpha \leq \frac{\delta}{\delta + \chi + (1+\delta)\phi^+} \leq \frac{1}{1 + \chi + (1+\delta)\phi^+} < 1, \qquad \alpha \leq \frac{1}{1+\chi},$$

for all $\alpha \in J$. Observe that if $\phi \geq 0$, then, in view of Lemma 2.2(a) and (2.5), we have $x(\alpha)^T y(\alpha) \geq (1-\alpha)x^T y > 0$, for all $\alpha \in J$. Assume that $\phi < 0$. In this case, we must have $\chi > 0$. This together with (2.5), Lemma 2.1, and Lemma 2.2(a) then imply

$$
\begin{aligned}
x(\alpha)^T y(\alpha) &= x^T y(1 - \alpha + \alpha^2 \phi) \\
&\geq x^T y(1 - \alpha - \alpha^2 \chi) \\
&> x^T y \left(1 - \alpha(1 + \chi)\right) \geq 0
\end{aligned}
$$

for all $\alpha \in J$ with $\alpha > 0$. We have thus shown that $x(\alpha)^T y(\alpha) > 0$ for all $\alpha \in J$. Now, using Lemma 2.2(a), (2.2), and (2.4), we obtain for all $\alpha \in J$ that

$$
\begin{aligned}
0 &< \eta[x(\alpha)^T y(\alpha)]^{1+\delta} \\
&\leq \eta(x^T y)^{1+\delta}[1 - \alpha(1 - \alpha\phi^+)]^{1+\delta} \\
&\leq \eta(x^T y)^{1+\delta}[1 - (1+\delta)\alpha(1 - \alpha\phi^+) + \delta\alpha^2(1 - \alpha\phi^+)^2] \\
&\leq \eta(x^T y)^{1+\delta}[1 - (1+\delta)\alpha(1 - \alpha\phi^+) + \delta\alpha^2] \\
(2.6) \qquad &\leq \eta(x^T y)^{1+\delta}[1 - \alpha - \alpha^2 \chi],
\end{aligned}
$$

where in the last inequality we used the fact that the interval J is exactly the set of all $\alpha \geq 0$ for which

$$1 - \alpha(1 + \delta)(1 - \alpha\phi^+) + \delta\alpha^2 \leq 1 - \alpha - \alpha^2 \chi.$$

From (2.6), Lemma 2.2(b), and the fact that $(x, y) \in \mathcal{N}(\delta, \eta, \rho)$, we have

$$
\begin{aligned}
x_i(\alpha)y_i(\alpha) &\geq x_i y_i(1 - \alpha - \alpha^2 \chi) \\
&\geq \eta(x^T y)^{1+\delta}(1 - \alpha - \alpha^2 \chi) \\
&\geq \eta[x(\alpha)^T y(\alpha)]^{1+\delta} > 0.
\end{aligned}
$$

Thus, the second condition for membership of $\mathcal{N}(\delta, \eta, \rho)$ is satisfied by $(x(\alpha), y(\alpha))$. To show the first condition, we observe that

$$(2.7) \qquad \rho x^T y \geq \|r\|^{1+\delta}.$$

From Lemma 2.2(c), we know that $\|r(\alpha)\| = (1 - \alpha)\|r\|$. By (2.2) and (2.7), it follows that

$$(2.8) \qquad \|r(\alpha)\|^{1+\delta} = (1 - \alpha)^{1+\delta}\|r\|^{1+\delta} \leq [1 - (1+\delta)\alpha + \delta\alpha^2]\rho x^T y.$$

For $\alpha \geq 0$, the inequality

$$(2.9) \qquad 1 - (1 + \delta)\alpha + \delta\alpha^2 \leq 1 - \alpha + \alpha^2 \phi$$

is equivalent to $\delta \geq \alpha(\delta - \phi)$, which is satisfied by every $\alpha \in [0, \delta/(\delta + \phi^-)]$, where $\phi^- \equiv -\min(0, \phi)$. Now, using Lemma 2.1, we can easily see that $J \subseteq [0, \delta/(\delta + \phi^-)]$. We then

conclude that (2.9) is satisfied by every $\alpha \in J$. Hence, in view of (2.8) and Lemma 2.2(a), we have

$$\|r(\alpha)\|^{1+\delta} \leq (1 - \alpha + \alpha^2 \phi)\rho x^T y = \rho x(\alpha)^T y(\alpha).$$

Therefore, the first condition for membership of $(x(\alpha), y(\alpha))$ in $\mathcal{N}(\delta, \eta, \rho)$ is also satisfied, and we have the result. ∎

From now on, given the initial point (x^0, y^0) and $\delta \in (0, 1]$, we let

$$\pi_0 = \frac{\min_{i=1,2,\ldots,n}(x_i^0 y_i^0)}{x^{0T} y^0 / n}, \qquad \rho = \frac{\|r^0\|^{1+\delta}}{x^{0T} y^0},$$

(2.10)
$$\eta = \frac{\min_{i=1,2,\ldots,n}(x_i^0 y_i^0)}{(x^{0T} y^0)^{1+\delta}} = \frac{\pi_0}{n(x^{0T} y^0)^\delta},$$

Note that π_0 (a measure of the centrality of the initial point) is at most 1 and $(x^0, y^0) \in \mathcal{N}(\delta, \eta, \rho)$.

The following result is an immediate consequence of Lemma 2.2(c) and Lemma 2.3.

COROLLARY 2.4. *The sequence $\{(x^k, y^k)\}$ generated by Algorithm IAS satisfies*
(a) $x^{k+1T} y^{k+1} \leq [1 - \alpha_k/(1+\delta)] x^{kT} y^k$ for all $k \geq 0$;
(b) $r^k = \nu_k r^0$ for all $k \geq 0$, where $\nu_k \equiv \Pi_{l=0}^{k-1}(1 - \alpha_l)$;
(c) $(x^k, y^k) \in \mathcal{N}(\delta, \eta, \rho)$ for all $k \geq 0$, where η and ρ are as defined in (2.10).

Our main goal now is to derive estimates for the quantities χ and ϕ. These estimates will then be used to establish global convergence of Algorithm IAS, as well as polynomial complexity for certain values of δ. The following inequality is exploited in a number of proofs that follow.

LEMMA 2.5. *Let (x^0, y^0) and (x, y) be points such that $y - Mx - q = \nu(y^0 - Mx^0 - q)$ for some $\nu \in \mathsf{R}^n$, and let $(\bar{x}, \bar{y})$ be a point such that $\bar{y} - M\bar{x} - q = 0$. Then,*

$$
\begin{aligned}
0 \quad \leq \quad & \nu^2 x^{0T} y^0 + (1 - \nu)^2 \bar{x}^T \bar{y} + x^T y + \nu(1 - \nu)(x^{0T}\bar{y} + \bar{x}^T y^0) \\
& -\nu(x^{0T} y + x^T y^0) - (1 - \nu)(\bar{x}^T y + x^T \bar{y}).
\end{aligned}
$$
(2.11)

Proof. See, for example, Wright [8, Lemma 3.2]. ∎

As an immediate consequence of this inequality, we obtain the following result.

LEMMA 2.6. *Let $(x^0, y^0) > 0$ be given. Then there exists a constant $C_0 \geq 0$ satisfying the following property: If $(x, y) \geq 0$ is a point such that $y - Mx - q = \nu(y^0 - Mx^0 - q)$ for some $\nu \in [0, 1]$, then*

(2.12)
$$\max\{\nu\|x\|_\infty, \nu\|y\|_\infty\} \leq C_0 \left[x^T y + \nu\right].$$

Proof. Assume that $(\bar{x}, \bar{y})$ is a solution of (1.1). Using inequality (2.11) and the facts that $\nu \in [0, 1]$, $(x^0, y^0) > 0$, $(x, y) \geq 0$, $(\bar{x}, \bar{y}) \geq 0$, and $\bar{x}^T \bar{y} = 0$, we obtain

$$
\begin{aligned}
x^{0T}(\nu y) + y^{0T}(\nu x) \quad &\leq \quad \nu x^{0T} y^0 + x^T y + \nu(x^{0T}\bar{y} + y^{0T}\bar{x}) \\
&\leq \quad \max\left\{1, x^{0T} y^0 + x^{0T}\bar{y} + y^{0T}\bar{x}\right\}(x^T y + \nu).
\end{aligned}
$$

Relation (2.12) now follows by letting

$$C_0 \triangleq \frac{\max\left\{1, x^{0T} y^0 + x^{0T}\bar{y} + y^{0T}\bar{x}\right\}}{\min_{i=1,2,\ldots,n}\{\min(x_i^0, y_i^0)\}}.$$

LEMMA 2.7. *Let the vectors $u^0 \in \mathsf{R}^n$ and $v^0 \in \mathsf{R}^n$ and the parameters $\eta > 0$ and $\delta \geq 0$ be given. Then, for every positive vector $(x, y) \in \mathsf{R}^{2n}$ and scalar $\nu \geq 0$ such that*

$$(2.13) \qquad y - Mx - q \;=\; \nu(v^0 - Mu^0),$$

$$(2.14) \qquad \min_{i=1,2,\ldots,n} x_i y_i \;\geq\; \eta(x^T y)^{1+\delta},$$

the corresponding direction $(\Delta x, \Delta y)$ determined by (1.5) satisfies

$$(2.15) \quad \max\left(\|D^{-1}\Delta x\|, \|D\Delta y\|\right) \leq (x^T y)^{1/2} + \frac{2\nu}{\eta^{1/2}(x^T y)^{(1+\delta)/2}} \left[\|Y u^0\| + \|X v^0\|\right],$$

where $D \equiv X^{1/2} Y^{-1/2}$.

Proof. We will first show that

$$(2.16) \qquad \max\left(\|D^{-1}\Delta x\|, \|D\Delta y\|\right) \leq (x^T y)^{1/2} + 2\nu\left[\|D^{-1}u^0\| + \|Dv^0\|\right].$$

Indeed, it follows from (1.5a) and (2.13) that

$$M\Delta x - \Delta y = r = y - Mx - q = \nu(v^0 - Mu^0).$$

Therefore,

$$M(\Delta x + \nu u^0) = \Delta y + \nu v^0,$$

which by monotonicity of (1.1) implies that

$$(\Delta x + \nu u^0)^T(\Delta y + \nu v^0) \geq 0.$$

Hence, we obtain

$$(2.17) \qquad \Delta x^T \Delta y \geq -\nu u^{0^T}\Delta y - \nu v^{0^T}\Delta x - \nu^2 u^{0^T} v^0.$$

If we multiply (1.5b) by $(XY)^{-1/2}$ and square both sides, we obtain

$$\|D^{-1}\Delta x\|^2 + \|D\Delta y\|^2 + 2\Delta x^T \Delta y \;=\; x^T y,$$

which, in view of (2.17), implies

$$\|D^{-1}\Delta x - \nu Dv^0\|^2 + \|D\Delta y - \nu D^{-1}u^0\|^2 - \nu^2\|Dv^0 + D^{-1}u^0\|^2 \leq x^T y.$$

Therefore, using the inequality $(\beta^2 + \gamma^2)^{1/2} \leq |\beta| + |\gamma|$ and the triangle inequality, we have

$$
\begin{aligned}
\|D^{-1}\Delta x\| &\leq \left[x^T y + \nu^2\|Dv^0 + D^{-1}u^0\|^2\right]^{1/2} + \nu\|Dv^0\| \\
&\leq (x^T y)^{1/2} + 2\nu\|Dv^0\| + \nu\|D^{-1}u^0\| \\
(2.18) \qquad &\leq (x^T y)^{1/2} + 2\nu\left[\|Dv^0\| + \|D^{-1}u^0\|\right].
\end{aligned}
$$

A similar bound can be derived for $\|D\Delta y\|$, and hence relation (2.16) follows.

We now derive (2.15) from (2.16) and (2.14). From (2.14) we obtain

$$(2.19) \qquad \|D^{-1}u^0\|^2 = \sum_{i=1}^{n} \frac{y_i(u_i^0)^2}{x_i} = \sum_{i=1}^{n} \frac{(y_i u_i^0)^2}{x_i y_i} \leq \frac{\|Y u^0\|^2}{\eta(x^T y)^{1+\delta}}.$$

Similarly, we also have

$$(2.20) \qquad \|Dv^0\|^2 \leq \frac{\|Xv^0\|^2}{\eta(x^T y)^{1+\delta}}.$$

Relation (2.15) now follows by substituting (2.19) and (2.20) into (2.16). ■

We are now ready to show global convergence of Algorithm IAS without imposing any condition on the initial point (x^0, y^0).

THEOREM 2.8. *For any initial point $(x^0, y^0) > 0$, the sequence $\{(x^k, y^k)\}$ generated by Algorithm IAS satisfies*

(a) $x^{k^T} y^k \to 0$; and

(b) $r^k \to 0$.

Proof. First note that (b) follows as a consequence of (a) from Corollary 2.4(c). To show (a), we assume for contradiction that $x^{k^T} y^k$ does not converge to 0. Thus, in view of Corollary 2.4(a), it follows that there exists $\epsilon > 0$ such that

$$(2.21) \qquad x^{k^T} y^k \geq \epsilon, \qquad \forall k \geq 0.$$

Hence, in view of Corollary 2.4(c), we have

$$(2.22) \qquad \min_{i=1,2,\ldots,n} \{x_i^k y_i^k\} \geq \eta(x^{k^T} y^k)^{1+\delta} \geq \eta\epsilon^{1+\delta}, \qquad \forall k \geq 0.$$

Since the point (x^k, y^k) satisfies the assumptions of Lemma 2.6 from Corollary 2.4(b), we obtain

$$(2.23) \quad \max\{\nu^k\|x^k\|_\infty, \nu^k\|y^k\|_\infty\} \leq C_0\left(x^{k^T} y^k + \nu^k\right) \leq C_0\left(x^{0^T} y^0 + 1\right), \qquad \forall k \geq 0.$$

Let $(u^0, v^0) \in \mathbb{R}^{2n}$ be such that $v^0 - Mu^0 = r^0$ (e.g., $u^0 = 0$ and $v^0 = r^0$). From Corollary 2.4, (x^k, y^k) and ν_k satisfy the assumptions of Lemma 2.7 with η given by (2.10). Hence, inequality (2.15) gives

$$\max\left(\|(D^k)^{-1}\Delta x^k\|, \|D^k \Delta y^k\|\right)$$

$$\leq \left(x^{k^T} y^k\right)^{1/2} + \frac{2\nu^k}{\eta^{1/2}(x^{k^T} y^k)^{(1+\delta)/2}}\left[\|Y^k u^0\| + \|X^k v^0\|\right]$$

$$\leq \left(x^{k^T} y^k\right)^{1/2} + \frac{2}{\eta^{1/2}\epsilon^{(1+\delta)/2}}\max\left(\|\nu^k y^k\|_\infty, \|\nu^k x^k\|_\infty\right)\left(\|u^0\| + \|v^0\|\right)$$

$$(2.24) \qquad \leq \left(x^{0^T} y^0\right)^{1/2} + \frac{2C_0\left(x^{0^T} y^0 + 1\right)}{\eta^{1/2}\epsilon^{(1+\delta)/2}}\left(\|u^0\| + \|v^0\|\right) \triangleq \hat{C}_0.$$

Using (2.22) and (2.24), we obtain for all $k \geq 0$ that

$$\max\{\phi_k^+, \chi_k\} \leq \max_{i=1,2,\ldots,n}\left\{\frac{|\Delta x_i^k \Delta y_i^k|}{x_i^k y_i^k}\right\}$$

$$\leq \frac{\|\mathrm{diag}(\Delta x^k)\Delta y^k\|}{\min_{i=1,2,\ldots,n}\{x_i^k y_i^k\}}$$

$$\leq \frac{\|(D^k)^{-1}\Delta x^k\|\|D^k\Delta y^k\|}{\eta\epsilon^{1+\delta}}$$

$$\leq \frac{\hat{C_0}^2}{\eta\epsilon^{1+\delta}},$$

where the first inequality follows by using a similar argument as in the proof of Lemma 2.1. Hence, from (1.6) we have $\inf_k \alpha_k > 0$, which implies $x^{k^T}y^k \to 0$ from Corollary 2.4(a). We have thus obtained a contradiction, and statement (a) follows. ■

We note that Theorem 2.8 does not imply that the sequence $\{(x^k, y^k)\}$ is convergent or bounded. In Section 3, we will show that $\{(x^k, y^k)\}$ is convergent under the assumptions that there is a solution of (1.1) satisfying strict complementarity and that $\delta < 1/7$.

Our next aim is to derive an upper bound on the number of iterations required by Algorithm IAS to reduce the duality gap below a given tolerance $\epsilon > 0$. We will see that for certain values of δ, Algorithm IAS has a polynomial bound on this number of iterations. We restrict our analysis to certain choices of the initial point (x^0, y^0).

Given some conditions on the initial point (x^0, y^0), the next result gives bounds on the left-hand side of (2.15) in terms of the initial gap $x^{0^T}y^0$, the current gap x^Ty, and the measure of initial centrality π_0.

LEMMA 2.9. *Suppose that the initial point $(x^0, y^0) > 0$ is chosen so that $(x^0, y^0) \geq (x^*, y^*)$ for some $(x^*, y^*) \in \mathcal{S}$ and that π_0, η, and ρ are defined by (2.10). Let (x, y) be any vector pair in $\mathcal{N}(\delta, \eta, \rho)$ with $x^Ty \leq x^{0^T}y^0$ and $y - Mx - q = \nu(y^0 - Mx^0 - q)$ for some $\nu \in (0, 1]$. Let $(\Delta x, \Delta y)$ be the direction determined by (1.5). Then, we have*

$$(2.25) \qquad \max\left(\|D^{-1}\Delta x\|, \|D\Delta y\|\right) \leq 9\left(\frac{n}{\pi_0}\right)^{1/2}\left(\frac{x^{0^T}y^0}{x^Ty}\right)^{3\delta/2}(x^Ty)^{1/2}.$$

Proof. We start by defining a vector pair (u^0, v^0) by

$$(2.26) \qquad\qquad (u^0, v^0) = (x^0, y^0) - (x^*, y^*) \geq 0,$$

for $(x^*, y^*) \in \mathcal{S}$. It is easy to check that

$$\nu(v^0 - Mu^0) = \nu(y^0 - Mx^0 - q) = y - Mx - q.$$

Therefore Lemma 2.7 applies, and as a consequence the inequality (2.15) holds. Now, using (2.26), we obtain

$$(2.27) \quad \|Yu^0\| + \|Xv^0\| \leq y^Tu^0 + x^Tv^0 = y^T(x^0 - x^*) + x^T(y^0 - y^*) \leq y^Tx^0 + x^Ty^0.$$

By substituting this last relation into (2.15), we obtain

$$(2.28) \quad \max\left(\|D^{-1}\Delta x\|, \|D\Delta y\|\right) \leq (x^Ty)^{1/2} + 2\nu\eta^{-1/2}(x^Ty)^{-(1+\delta)/2}[y^Tx^0 + x^Ty^0].$$

We now invoke the inequality (2.11) with $(\bar{x}, \bar{y}) = (x^*, y^*)$, noting that $\nu \in (0, 1]$, $(x^*, y^*) \geq 0$, $x^{*^T}y^* = 0$, and $(x, y) > 0$, to deduce that

$$(2.29) \qquad \nu(x^{0^T}y + x^Ty^0) \leq \nu^2x^{0^T}y^0 + x^Ty + \nu(x^{0^T}y^* + x^{*^T}y^0).$$

Now, by (2.26), we have

$$(2.30) \quad \frac{x^{0T}y^* + x^{*T}y^0}{x^{0T}y^0} = \frac{x^{0T}(y^0 - v^0) + (x^0 - u^0)^T y^0}{x^{0T}y^0} = 2 - \frac{x^{0T}v^0 + u^{0T}y^0}{x^{0T}y^0} \le 2.$$

In addition, by (2.10) and the definition of $\mathcal{N}(\delta, \eta, \rho)$, we have

$$(2.31) \quad \nu = \frac{\|r\|}{\|r^0\|} \le \frac{[\rho x^T y]^{1/(1+\delta)}}{\|r^0\|} = \left(\frac{x^T y}{x^{0T}y^0} \right)^{1/(1+\delta)}.$$

Combining (2.29), (2.30), and (2.31), we get

$$\nu(x^{0T}y + x^T y^0) \le x^{0T}y^0 \left[\nu^2 + \left(\frac{x^T y}{x^{0T}y^0} \right) + 2\nu \right]$$

$$\le x^{0T}y^0 \left(\frac{x^T y}{x^{0T}y^0} \right)^{1/(1+\delta)} \left[\nu + \left(\frac{x^T y}{x^{0T}y^0} \right)^{\delta/(1+\delta)} + 2 \right]$$

$$(2.32) \qquad\qquad \le 4x^{0T}y^0 \left(\frac{x^T y}{x^{0T}y^0} \right)^{1/(1+\delta)}.$$

If we substitute (2.32) into (2.28) and replace η by its definition (2.10), we obtain

$$\max \left(\|D^{-1}\Delta x\|, \|D\Delta y\| \right) \le (x^T y)^{1/2} + 8\eta^{-1/2}(x^T y)^{-(1+\delta)/2}(x^{0T}y^0)^{\delta/(1+\delta)}(x^T y)^{1/(1+\delta)}$$

$$= (x^T y)^{1/2} \left[1 + 8 \left(\frac{n}{\pi_0} \right)^{1/2} \left(\frac{x^{0T}y^0}{x^T y} \right)^{(\delta/2)(3+\delta)/(1+\delta)} \right]$$

$$\le 9 \left(\frac{n}{\pi_0} \right)^{1/2} \left(\frac{x^{0T}y^0}{x^T y} \right)^{3\delta/2} (x^T y)^{1/2},$$

where the last inequality follows from

$$\frac{x^{0T}y^0}{x^T y} \ge 1, \qquad \frac{3+\delta}{1+\delta} < 3, \qquad \pi_0 \le 1.$$

∎

We next derive uniform bounds for the quantities $|\phi_k|$ and χ_k defined by (1.3).

LEMMA 2.10. *Suppose that the initial point $(x^0, y^0) > 0$ satisfies $(x^0, y^0) \ge (x^*, y^*)$ for some $(x^*, y^*) \in \mathcal{S}$ and that η and ρ are defined by (2.10). Then for all iterates (x^k, y^k) generated by Algorithm IAS, we have*

$$(2.33a) \qquad\qquad |\phi_k| \le 81 \left(\frac{n}{\pi_0} \right) \left(\frac{x^{0T}y^0}{x^{kT}y^k} \right)^{3\delta},$$

$$(2.33b) \qquad\qquad 0 \le \chi_k \le 81 \left(\frac{n}{\pi_0} \right)^2 \left(\frac{x^{0T}y^0}{x^{kT}y^k} \right)^{4\delta}.$$

Proof. By Corollary 2.4, it follows that (x^k, y^k) satisfies the assumptions of Lemma 2.9 for all $k \geq 0$. Hence, it follows from (2.25) that

$$|\phi_k| = \frac{|\Delta x^{k^T} \Delta y^k|}{x^{k^T} y^k} \leq \frac{\|D^k \Delta x^k\| \|(D^k)^{-1} \Delta y^k\|}{x^{k^T} y^k} \leq 81 \left(\frac{n}{\pi_0}\right) \left(\frac{x^{0^T} y^0}{x^{k^T} y^k}\right)^{3\delta}.$$

Similarly, by (2.10), (2.25), and the definition of $\mathcal{N}(\delta, \eta, \rho)$, we have

$$\begin{aligned}
\chi_k \leq \max_{i=1,2,\ldots,n} \frac{|\Delta x_i^k \Delta y_i^k|}{x_i^k y_i^k} &\leq \frac{\|D^k \Delta x^k\| \|(D^k)^{-1} \Delta y^k\|}{\eta (x^{k^T} y^k)^{1+\delta}} \\
&\leq 81 \left(\frac{n}{\pi_0}\right) \left(\frac{x^{0^T} y^0}{x^{k^T} y^k}\right)^{3\delta} (x^{k^T} y^k)^{-\delta} \eta^{-1} \\
&= 81 \left(\frac{n}{\pi_0}\right)^2 \left(\frac{x^{0^T} y^0}{x^{k^T} y^k}\right)^{4\delta},
\end{aligned}$$

as required. ∎

We can now state a global convergence result.

THEOREM 2.11. *Suppose that the initial point* $(x^0, y^0) > 0$ *satisfies* $(x^0, y^0) \geq (x^*, y^*)$ *for some* $(x^*, y^*) \in \mathcal{S}$. *Let* $\delta \in (0, 1]$ *be given, and let* η *and* ρ *be defined by (2.10). Then, given* $\epsilon \in (0, x^{0^T} y^0)$, *Algorithm IAS will generate an iterate* (x^k, y^k) *with* $x^{k^T} y^k \leq \epsilon$ *in at most*

$$(2.34) \qquad O\left(\frac{1}{\delta} \left(\frac{n}{\pi_0}\right)^2 \left(\frac{x^{0^T} y^0}{\epsilon}\right)^{4\delta} \log(x^{0^T} y^0 / \epsilon)\right)$$

iterations, and the sequence $\{x^{k^T} y^k\}$ *will be strictly monotonically decreasing.*

Proof. To simplify notation, let $\tau = (x^{0^T} y^0 / \epsilon)$. If the iterate (x^k, y^k) satisfies $x^{k^T} y^k \geq \epsilon$ then, in view of (2.33), we have

$$|\phi_k| \leq 81(n/\pi_0)\tau^{3\delta}, \qquad \chi_k \leq 81(n/\pi_0)^2 \tau^{4\delta},$$

and hence

$$\begin{aligned}
\alpha_k &= \frac{\delta}{\delta + \chi_k + (1+\delta)\phi_k^+} \\
&\geq \frac{\delta}{\delta + 81(n/\pi_0)^2 \tau^{4\delta} + (1+\delta)81(n/\pi_0)\tau^{3\delta}} \\
&\geq \frac{\delta}{\delta + 243(n/\pi_0)^2 \tau^{4\delta}} \\
&\geq \frac{\delta}{244} \left(\frac{\pi_0}{n}\right)^2 \tau^{-4\delta} \\
(2.35) \qquad &\triangleq \bar{\alpha}.
\end{aligned}$$

Strictly monotonic decrease of the sequence $\{x^{k^T} y^k\}$ now follows from Corollary 2.4(a) and the fact that the inclusion $(x^k, y^k) \in \mathcal{N}(\delta, \eta, \rho)$ implies that $x^{k^T} y^k > 0$. Now let K be

the smallest nonnegative integer for which $x^{K+1^T} y^{K+1} \leq \epsilon$. Since (2.35) holds for every $0 \leq k \leq K$, we obtain from Lemma 2.3(a) that

$$(2.36) \qquad x^{K^T} y^K \leq [1 - \bar{\alpha}/(1+\delta)]^K x^{0^T} y^0.$$

Assume for contradiction that

$$K \geq \frac{(1+\delta)\log\tau}{\bar{\alpha}}.$$

Using the inequality $\log(1-\beta) \leq -\beta$ for $\beta < 1$ and the previous relation, we can easily show that

$$[1 - \bar{\alpha}/(1+\delta)]^K \leq \tau^{-1},$$

which, in view of the definition of τ and relation (2.36), implies

$$x^{K^T} y^K \leq \tau^{-1} x^{0^T} y^0 = \epsilon.$$

This relation contradicts our definition of K. Therefore, using (2.35) and the definition of τ, we have

$$K < \frac{(1+\delta)\log\tau}{\bar{\alpha}} = O\left(\frac{1}{\delta}\left(\frac{n}{\pi_0}\right)^2 \left(\frac{x^{0^T} y^0}{\epsilon}\right)^{4\delta} \log(x^{0^T} y^0/\epsilon)\right).$$

∎

Note in particular that if

$$\epsilon = e^{-O(L)}, \qquad x^{0^T} y^0 = e^{O(L)}, \qquad \pi_0 = O(1), \qquad \delta = 1/O(L),$$

then the bound in (2.34) is $O(n^2 L^2)$, that is, there is a polynomial bound on the number of iterations of Algorithm IAS.

3. Local Convergence. Under the conditions that problem (1.1) is nondegenerate and $\delta \in (0, 1/7)$, we prove in this section that Algorithm IAS started from any initial point $(x^0, y^0) > 0$ generates a sequence of iterates $\{(x^k, y^k)\}$ such that $\{x^{k^T} y^k\}$ converges superlinearly to zero. Moreover, under the same conditions, we also show that the sequence $\{(x^k, y^k)\}$ converges to a solution of (1.1). We emphasize that the convergence of $\{(x^k, y^k)\}$ is obtained without assuming the existence of a strictly feasible point for (1.1). (This assumption was used by the local convergence analysis in Wright [7], for example.)

The following assumption is imposed throughout this section.

ASSUMPTION 1. (Nondegeneracy) $\mathcal{S}_c \neq \emptyset$.

We show in [5] that Assumption 1 is in fact necessary for superlinear convergence of Newton-based primal-dual algorithms.

An important feature of this section is that we are able to prove boundedness of the steps $(\Delta x^k, \Delta y^k)$ in terms of $x^{k^T} y^k$ without an intermediate result that the sequence (x^k, y^k) is bounded. The key result — that $\|\Delta x_B^k\|$ and $\|\Delta y_N^k\|$ are both $O(x^{k^T} y^k)$ — is obtained by combining Lemmas 3.3 and 3.4 below. In [7], the same result was obtained by combining Lemma 3.4 with a more elementary result in which boundedness of (x^k, y^k) was an essential ingredient. In fact, we can use the techniques of this section to drop the strict feasibility

assumption from the analysis of [7] without affecting the conclusions of the paper (see Wright [9]).

After finding the bounds on the steps $(\Delta x^k, \Delta y^k)$ in terms of the complementarity gap ${x^k}^T y^k$, we bound the scalars $|\phi_k|$ and χ_k. Subsequently, we show that $\alpha_k \to 1$, from which Q-superlinear convergence follows. We also find a lower bound on the Q-order of convergence in terms of δ.

The following result yields bounds on the nonbasic components of x and y.

LEMMA 3.1. *Suppose that Assumption 1 holds. Assume that a point $(x^0, y^0) > 0$ and that scalars $\delta \geq 0$, $\eta > 0$, and $\rho \geq 0$ are given. Then, there exists a constant $C_1 \geq 0$ with the following property: For every vector $(x, y) \in \mathcal{N}(\delta, \eta, \rho)$ such that*

$$(3.1) \qquad x^T y \;\leq\; {x^0}^T y^0,$$

$$(3.2) \qquad y - Mx - q \;=\; \nu(y^0 - Mx^0 - q), \text{ for some } \nu \in [0, 1/2],$$

we have

$$(3.3a) \qquad x_i \;\leq\; C_1(x^T y)^{1/(1+\delta)}, \qquad \forall i \in N,$$

$$(3.3b) \qquad y_i \;\leq\; C_1(x^T y)^{1/(1+\delta)}, \qquad \forall i \in B.$$

Proof. Consider first the infeasible case in which $r^0 = y^0 - Mx^0 - q \neq 0$. Setting $(\bar{x}, \bar{y}) = (x^*, y^*)$ for some $(x^*, y^*) \in \mathcal{S}_c$ in inequality (2.11) and using the fact that $\nu \in [0, 1/2]$, $(x, y) > 0$, $(x^*, y^*) \geq 0$ and ${x^*}^T y^* = 0$, we obtain

$$(3.4) \qquad {x^*}^T y + x^T y^* \;\leq\; \frac{\nu}{1 - \nu} {x^0}^T y^0 + \frac{1}{1 - \nu} x^T y + \nu({x^0}^T y^* + {x^*}^T y^0).$$

Using (3.2) and the fact that $(x, y) \in \mathcal{N}(\delta, \eta, \rho)$, we get

$$(3.5) \qquad \nu = \frac{\|r\|}{\|r^0\|} \leq \frac{(\rho x^T y)^{1/(1+\delta)}}{\|r^0\|},$$

where, as usual, $r = y - Mx - q$. From (3.1), (3.5), (3.4), and the fact that $1 - \nu \geq 1/2$, we get

$${x^*}^T y + x^T y^*$$

$$\leq \quad \frac{2{x^0}^T y^0}{\|r^0\|} \rho^{1/(1+\delta)} (x^T y)^{1/(1+\delta)} + 2x^T y + \frac{1}{\|r^0\|} \rho^{1/(1+\delta)}({x^0}^T y^* + {x^*}^T y^0)(x^T y)^{1/(1+\delta)}$$

$$\leq \quad (x^T y)^{1/(1+\delta)} \left[\frac{2{x^0}^T y^0}{\|r^0\|} \rho^{1/(1+\delta)} + 2({x^0}^T y^0)^{\delta/(1+\delta)} + \frac{1}{\|r^0\|} \rho^{1/(1+\delta)}({x^0}^T y^* + {x^*}^T y^0) \right]$$

$$= \quad \bar{C}_1 (x^T y)^{1/(1+\delta)},$$

where $\bar{C}_1$ is defined in an obvious way. This last relation immediately implies (3.3) when we define

$$C_1 \triangleq \bar{C}_1 \max \left\{ \max_{i \in N} \frac{1}{y_i^*}, \; \max_{i \in B} \frac{1}{x_i^*} \right\}.$$

For the feasible case $r^0 = 0$, note from the condition $y - Mx - q = \nu(y^0 - Mx^0 - q)$ that $r = y - Mx - q = 0$ for all (x, y) satisfying the assumptions of the theorem. Hence we can take $\nu = 0$ in (3.4) and obtain

$$x^{*T}y + x^T y^* \leq x^T y.$$

A simple argument produces the required bound with C_1 defined by

$$C_1 = (x^{0T}y^0)^{\delta/(1+\delta)} \max\left\{ \max_{i \in N} \frac{1}{y_i^*},\ \max_{i \in B} \frac{1}{x_i^*} \right\}.$$

∎

We now use this lemma to find bounds on the nonbasic components of $(\Delta x, \Delta y)$.

LEMMA 3.2. *Suppose that Assumption 1 holds and that $(x^0, y^0) > 0$ is given. Assume that scalars $\delta \in [0, 1/3]$, $\eta > 0$, and $\rho \geq 0$ are given. Then, there exists a constant $C_2 \geq 0$ with the following property: For any vector pair $(x, y) \in \mathcal{N}(\delta, \eta, \rho)$ such that*

$$(3.6) \qquad\qquad x^T y \quad \leq \quad \min(1, x^{0T}y^0),$$

$$(3.7) \qquad\qquad y - Mx - q \quad = \quad \nu(y^0 - Mx^0 - q),\ \textit{for some } \nu \in [0, 1/2],$$

we have

$$(3.8a) \qquad\qquad |\Delta x_i| \quad \leq \quad C_2(x^T y)^{1-3\delta}, \qquad \forall i \in N,$$

$$(3.8b) \qquad\qquad |\Delta y_i| \quad \leq \quad C_2(x^T y)^{1-3\delta}, \qquad \forall i \in B.$$

Proof. We first show that there exists a constant $\bar{C}_2$ such that

$$(3.9) \qquad\qquad \max\left\{ \|D^{-1}\Delta x\|, \|D\Delta y\| \right\} \leq \bar{C}_2(x^T y)^{(1-3\delta)/2}.$$

We give the proof only for the case in which $r^0 \equiv y^0 - Mx^0 - q \neq 0$, since the proof for $r^0 = 0$ is trivial. The assumptions of the lemma imply that (2.12) and (3.5) hold. These two relations together with (3.6) then imply

$$
\begin{aligned}
\max\{\nu\|x\|_\infty, \nu\|y\|_\infty\} \quad &\leq \quad C_0\left[x^T y + \nu\right] \\
&\leq \quad C_0\left[1 + \frac{\rho^{1/(1+\delta)}}{\|r^0\|}\right](x^T y)^{1/(1+\delta)} \\
&= \quad \hat{C}_2(x^T y)^{1/(1+\delta)} \\
(3.10) \qquad\qquad &\leq \quad \hat{C}_2(x^T y)^{1-\delta},
\end{aligned}
$$

where $\hat{C}_2$ is defined in an obvious way. Since the assumptions of Lemma 2.7 are satisfied, inequality (2.15) together with (3.10) and (3.6) then yield

$$
\begin{aligned}
\max\left(\|D^{-1}\Delta x\|, \|D\Delta y\|\right) & \\
\leq \quad (x^T y)^{1/2} &+ \frac{2\nu}{\eta^{1/2}(x^T y)^{(1+\delta)/2}}\left[\|Yu^0\| + \|Xv^0\|\right] \\
\leq \quad (x^T y)^{1/2} &+ \frac{2}{\eta^{1/2}(x^T y)^{(1+\delta)/2}}\left(\|u^0\| + \|v^0\|\right)\max\left(\|\nu y\|_\infty, \|\nu x\|_\infty\right)
\end{aligned}
$$

$$\leq \quad (x^T y)^{1/2} + \frac{2\hat{C}_2}{\eta^{1/2}(x^T y)^{(1+\delta)/2}} \left(\|u^0\| + \|v^0\| \right) (x^T y)^{1-\delta}$$

$$\leq \quad \left[1 + \frac{2\hat{C}_2}{\eta^{1/2}} \left(\|u^0\| + \|v^0\| \right) \right] (x^T y)^{(1-3\delta)/2}$$

$$= \quad \bar{C}_2 (x^T y)^{(1-3\delta)/2},$$

where $\bar{C}_2$ is defined in an obvious way. We have thus shown that (3.9) holds.

Considering $i \in N$, we have from the definition $D = X^{1/2} Y^{-1/2}$, the formulae (3.9) and (3.3a), and the inclusion $(x, y) \in \mathcal{N}(\delta, \eta, \rho)$ that

$$|\Delta x_i| \quad \leq \quad \left(\frac{x_i}{y_i} \right)^{1/2} \bar{C}_2 (x^T y)^{(1-3\delta)/2}$$

$$= \quad \frac{x_i}{(x_i y_i)^{1/2}} \bar{C}_2 (x^T y)^{(1-3\delta)/2}$$

$$\leq \quad \bar{C}_2 \frac{C_1 (x^T y)^{1/(1+\delta)}}{\eta^{1/2}(x^T y)^{(1+\delta)/2}} (x^T y)^{(1-3\delta)/2}$$

$$(3.11) \qquad\qquad\qquad = \quad (\bar{C}_2 C_1 \eta^{-1/2})(x^T y)^{(1-2\delta-2\delta^2)/(1+\delta)}.$$

Since $\delta \in [0, 1/3]$, we have

$$\frac{1 - 2\delta - 2\delta^2}{1 + \delta} \geq (1 - 2\delta - 2\delta^2)(1 - \delta) = 1 - 3\delta + 2\delta^3 \geq 1 - 3\delta \geq 0,$$

and since $x^T y \leq 1$, we conclude from (3.11) that

$$|\Delta x_i| \leq (\bar{C}_2 C_1 \eta^{-1/2})(x^T y)^{1-3\delta} = C_2 (x^T y)^{1-3\delta},$$

for C_2 defined in an obvious way. Hence, we have proved (3.8a). The proof of (3.8b) is identical. $\blacksquare$

To find bounds on the remaining components of $(\Delta x, \Delta y)$, we cite the following two results, which have been proved in earlier reports.

LEMMA 3.3. ([4, Lemma 2.2]) Let $f \in \mathsf{R}^q$ and $H \in \mathsf{R}^{p \times q}$ be given. Then there exists a constant $L = L(f, H)$ with the property that for any positive definite diagonal matrix D and any vector $h \in Range(H)$, the (unique) optimal solution $\bar{w} = \bar{w}(D, h)$ of

$$(3.12) \qquad\qquad \min_w \; f^T w + \frac{1}{2} \|Dw\|^2, \qquad \text{subject to } Hw = h,$$

satisfies

$$\|\bar{w}\|_\infty \leq L \left\{ |f^T \bar{w}| + \|h\|_\infty \right\}.$$

LEMMA 3.4. ([7, Lemma 5.2]) Let $(x, y) > 0$ be given, and consider the direction $(\Delta x, \Delta y)$ determined by (1.5). Then $(\Delta x_B, \Delta y_N)$ solves the convex quadratic programming problem

$$\min_{(w, z)} \; \tfrac{1}{2} \|D_B w\|^2 + \tfrac{1}{2} \|D_N^{-1} z\|^2$$

subject to

$$
\begin{aligned}
M_{BB}w &= r_B - M_{BN}\Delta x_N + \Delta y_B \\
M_{NB}w - z &= r_N - M_{NN}\Delta x_N.
\end{aligned}
$$

Lemma 3.4 is an extension to the infeasible case of a result due to Ye and Anstreicher [10, Lemma 3.5].

As a consequence of Lemmas 3.3 and 3.4, we can find bounds on Δx_B and Δy_N.

LEMMA 3.5. *Suppose the assumptions of Lemma 3.2 hold. Then there exists a constant* $C_3 \geq 0$ *such that*

$$
\begin{aligned}
(3.13a) \qquad |\Delta x_i| &\leq C_3(x^T y)^{1-3\delta}, & \forall i \in B, \\
(3.13b) \qquad |\Delta y_i| &\leq C_3(x^T y)^{1-3\delta}, & \forall i \in N.
\end{aligned}
$$

Proof. From Lemmas 3.2, 3.3, and 3.4, we have

$$
\begin{aligned}
\|(\Delta x_B, \Delta y_N)\|_\infty &\leq L \left[\|r\|_\infty + \|M\|_\infty \|\Delta x_N\|_\infty + \|\Delta y_B\|_\infty \right] \\
&\leq L \left[(\rho x^T y)^{1/(1+\delta)} + (\|M\|_\infty + 1)C_2(x^T y)^{1-3\delta} \right].
\end{aligned}
$$

Since $1/(1+\delta) \geq 1 - 3\delta$ when $\delta \in [0, 1/3]$, we obtain the desired result by setting

$$
C_3 = L \left[\rho^{1/(1+\delta)} + (\|M\|_\infty + 1)C_2 \right].
$$

∎

As a immediate consequence of Lemmas 3.2 and 3.5, we obtain the following result.

THEOREM 3.6. *Suppose that Assumption 1 holds and that* $\delta \in [0, 1/3]$ *is given. Then, there exists a constant* $C_4 \geq 0$ *such that for all* k *sufficiently large, we have*

$$
\begin{aligned}
(3.14a) \qquad |\Delta x_i^k| &\leq C_4(x^{k^T} y^k)^{1-3\delta}, & \forall i = 1, 2, \ldots, n, \\
(3.14b) \qquad |\Delta y_i^k| &\leq C_4(x^{k^T} y^k)^{1-3\delta}, & \forall i = 1, 2, \ldots, n.
\end{aligned}
$$

Proof. Since the assumptions of Theorem 2.8 are satisfied, we know that $x^{k^T} y^k \downarrow 0$ and $\|r^k\| \to 0$. Hence, in view of Corollary 2.4(b), we have that $\nu_k \to 0$. Thus, there exists an index $K \geq 0$ such that $\nu_k \leq 1/2$ and $x^{k^T} y^k \leq 1$ for all $k \geq K$. The above observations together with Corollary 2.4 imply that the assumptions of Lemmas 3.2 and 3.5 are satisfied by all points (x^k, y^k) with $k \geq K$. Thus, we obtain (3.14) by choosing $C_4 = \max(C_2, C_3)$. ∎

COROLLARY 3.7. *Suppose the assumptions of Theorem 3.6 are satisfied. Then there is a constant* $C_5 \geq 0$ *such that for all* k *sufficiently large, we have*

$$
(3.15) \qquad |\phi_k| \leq C_5(x^{k^T} y^k)^{1-6\delta}, \qquad 0 \leq \chi_k \leq C_5(x^{k^T} y^k)^{1-7\delta}.
$$

Proof. For sufficiently large k, we have from (3.14) that

$$
|\phi_k| = \frac{|\Delta x^{k^T} \Delta y^k|}{x^{k^T} y^k} \leq n C_4^2 (x^{k^T} y^k)^{1-6\delta},
$$

as required. Since $(x^k, y^k) \in \mathcal{N}(\delta, \eta, \rho)$, we also have that

$$\chi_k \leq \max_{i=1,2,\ldots,n} \frac{|\Delta x_i^k \Delta y_i^k|}{x_i^k y_i^k} \leq \frac{C_4^2 (x^{k^T} y^k)^{2-6\delta}}{\eta (x^{k^T} y^k)^{1+\delta}} = (C_4^2/\eta)(x^{k^T} y^k)^{1-7\delta}.$$

The result follows if we set $C_5 = \max(n C_4^2, C_4^2/\eta)$.   ∎

We are now ready to prove the main local convergence result.

THEOREM 3.8. *Suppose that Assumption 1 holds and that $\delta \in (0, 1/7)$ is given. Then the sequence $\{x^{k^T} y^k\}$ converges Q-superlinearly to zero with Q-order at least $2 - 7\delta$.*

Proof. Suppose the index K is chosen sufficiently large that the relation $x^{k^T} y^k \leq 1$ and the bounds in (3.15) hold for all $k \geq K$. Then we have

$$
\begin{aligned}
1 - \alpha_k &= \frac{\chi_k + (1+\delta)\phi_k^+}{\delta + \chi_k + (1+\delta)\phi_k^+} \\
&\leq \frac{\chi_k + (1+\delta)\phi_k^+}{\delta} \\
&\leq \frac{C_5 \left[1 + (1+\delta)(x^{k^T} y^k)^\delta\right](x^{k^T} y^k)^{1-7\delta}}{\delta} \\
&\leq \frac{3C_5}{\delta}(x^{k^T} y^k)^{1-7\delta}.
\end{aligned}
$$

Hence, from Lemma 2.2(a), we have for all $k \geq K$ that

$$
\begin{aligned}
x^{k+1^T} y^{k+1} &\leq x^{k^T} y^k (1 - \alpha_k + \alpha_k^2 \phi_k) \\
&\leq x^{k^T} y^k (1 - \alpha_k + |\phi_k|) \\
&\leq x^{k^T} y^k \left(\frac{3C_5}{\delta}(x^{k^T} y^k)^{1-7\delta} + C_5(x^{k^T} y^k)^{1-6\delta}\right) \\
&\leq \frac{4C_5}{\delta}(x^{k^T} y^k)^{2-7\delta},
\end{aligned}
$$

as required.   ∎

Convergence (and therefore boundedness) of the sequence $\{(x^k, y^k)\}$ follows as a corollary of Theorem 3.8.

COROLLARY 3.9. *Suppose the assumptions of Theorem 3.8 hold. Then the sequence $\{(x^k, y^k)\}$ is convergent.*

Proof. It suffices to show that the sequence is Cauchy. From Theorem 3.8, there exists an index $\bar{K}$ such that $x^{k+1^T} y^{k+1} \leq (1/2)^{1/(1-3\delta)} x^{k^T} y^k$ for all $k \geq \bar{K}$. Hence, from (3.14) we have for k_1, k_2 with $\bar{K} \leq k_1 < k_2$ that

$$
\begin{aligned}
\|(x^{k_2}, y^{k_2}) - (x^{k_1}, y^{k_1})\|_\infty &\leq \sum_{k=k_1}^{k_2-1} \|(\Delta x^k, \Delta y^k)\|_\infty \\
&\leq C_4 \sum_{k=k_1}^{k_2-1} \left(x^{k^T} y^k\right)^{1-3\delta} \\
&\leq C_4 \left(x^{k_1^T} y^{k_1}\right)^{1-3\delta} \sum_{k=k_1}^{k_2-1} \frac{1}{2^{k-k_1}}
\end{aligned}
$$

$$\leq \quad 2C_4 \left(x^{k_1 T} y^{k_1} \right)^{1-3\delta} .$$

Since $x^{k T} y^k \downarrow 0$, the above relation shows that, given any $\epsilon > 0$, we can choose $\hat{K} \geq \bar{K}$ such that

$$\|(x^{k_2}, y^{k_2}) - (x^{k_1}, y^{k_1})\|_\infty \leq \epsilon \qquad \forall k_2 > k_1 \geq \hat{K},$$

and so the sequence $\{(x^k, y^k)\}$ is Cauchy.   ∎

REFERENCES

[1] M. KOJIMA, N. MEGIDDO, AND S. MIZUNO, *A primal-dual infeasible-interior-point point algorithm for linear programming*, Mathematical Programming, Series A, 61 (1993), pp. 261–280.

[2] S. MIZUNO, *Polynomiality of Kojima-Meggido-Mizuno infeasible-interior-point algorithm for linear programming*, Tech. Rep. 1006, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, N.Y., January 1993.

[3] S. MIZUNO, M. KOJIMA, AND M. TODD, *Infeasible interior-point primal-dual potential-reduction algorithms for linear programming*, Tech. Rep. 1023, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, August 1992. To appear in SIAM Journal on Optimization.

[4] R. D. C. MONTEIRO AND S. J. WRIGHT, *Superlinear primal-dual affine scaling algorithms for LCP*, Preprint MCS–P360–0693, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., June 1993. To appear in Mathematical Programming, Series A.

[5] ———, *Local convergence of interior-point algorithms for degenerate monotone LCP*, Computational Optimization and Applications, 3 (1994), pp. 131–155.

[6] F. A. POTRA, *A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points*, Reports on Computational Mathematics 28 & 34, University of Iowa, Iowa City, December 1993. To appear in Mathematical Programming.

[7] S. J. WRIGHT, *An infeasible-interior-point algorithm for linear complementarity problems*, Preprint MCS–P331–1092, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., October 1992. To appear in Mathematical Programming, Series A.

[8] ———, *A path-following infeasible-interior-point algorithm for linear complementarity problems*, Optimization Methods and Software, 2 (1993), pp. 79–106.

[9] ———, *A path-following interior-point algorithm for linear and quadratic optimization problems*, Preprint MCS–P401–1293, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., December 1993.

[10] Y. YE AND K. ANSTREICHER, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP*, Mathematical Programming, Series A, 62 (1993), pp. 537–551.

[11] Y. ZHANG, *On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem*, SIAM Journal on Optimization, 4 (1994), pp. 208–227.

[12] Y. ZHANG AND D. ZHANG, *Superlinear convergence of infeasible interior-point methods for linear programming*, Research Report 92-15, Department of Mathematics and Statistics, University of Maryland Baltimore County, September 1992. To appear in Mathematical Programming, Series A.