On the Superlinear Convergence of Interior Point Algorithms for a General Class of Problems

Y. Zhang
R. A. Tapia
F. Porta

CRPC-TR90111
March, 1990

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

On the Superlinear Convergence of Interior Point Algorithms for a General Class of Problems *

Yin Zhang †, Richard Tapia ‡and Florian Potra §

March, 1990

Abstract

In this paper, we extend the Q-superlinear convergence theory recently developed by Zhang, Tapia and Dennis for a class of interior point linear programming algorithms to similar interior point algorithms for quadratic programming and for linear complementarity problems. Our unified approach consists of viewing all these algorithms as the damped Newton method applied to perturbations of a general problem. We show that under appropriate assumptions, Q-superlinear convergence can be achieved by asymptotically taking the step to the boundary of the positive orthant and letting the barrier (or path-following) parameter approach zero at a specific rate.

Keywords: Interior point algorithms, Linear programming, Quadratic programming, Linear complementarity problems, perturbed and damped Newton's method, Q-superlinear convergence.

*Research supported in part by NSF Coop. Agr. No. CCR-8809615, AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.

†Department of Mathematical Sciences, Rice University, Houston, Texas, 77251-1892

‡Department of Mathematical Sciences, Rice University, Houston, Texas, 77251-1892

§Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242

Abbreviated Title: Superlinear convergence for interior point algorithms

1 Introduction

Consider the general nonlinear system

$$F(x,y) = \begin{pmatrix} Mx + Ny - h \\ XYe \end{pmatrix} = 0, \quad (x,y) \geq 0, \tag{1.1}$$

where $x, y, h, e \in \mathbf{R}^n$, $M, N \in \mathbf{R}^{n \times n}$, $X = diag(x)$, $Y = diag(y)$ and e has all components equal to one.

We call the following set the feasibility set of problem (1.1):

$$\Omega = \{(x,y) : x, y \in \mathbf{R}^n, Mx + Ny = h, (x,y) \geq 0\}.$$

A feasible pair $(x,y) \in \Omega$ is said to be strictly feasible if it is positive. In this work we tacitly assume that the relative interior of Ω is nonempty, i.e., strictly feasible points exist.

Problem (1.1) is sufficiently general to include linear complementarity problems, quadratic programming problems and linear programming problems. We will now demonstrate this fact. To begin with observe that if $N = -I$, then this problem is the standard linear complementarity problem (LCP). Moreover, the assumption that M is positive semi-definite will be sufficient to guarantee that the algorithms under investigation produce well-defined iterates (Corollary 2.1).

Now consider the quadratic programming problem (QP)

$$\begin{array}{ll} \text{minimize} & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array} \tag{1.2}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n} (m < n)$ and has full row rank, and $Q \in \mathbf{R}^{n \times n}$ is symmetric. In Corollary 2.1, we will demonstrate that iterates produced by the algorithms ~~under investigation~~ are well-defined if Q is positive semi-definite on the null space of A.

The role that the positive semi-definiteness of Q on the null space of A plays within the context of the quadratic program (1.2) is often misrepresented or confused. It therefore merits further discussion. A well-known argument from convexity theory can be used to show that, as long as (1.2) has at least one strictly feasible point, then QP (1.2) is also a convex program if and only if Q is positive semi-definite on the null space of A. Hence, it seems reasonable to refer to this situation as convex quadratic programming. However, recently, some researchers in this area have chosen to use this terminology to describe the more restrictive situation where Q is positive semi-definite on the entire space. Consequently, in this work, we will not use the term convex quadratic program and will always delineate the assumptions that have been made on Q.

Recall that when the QP (1.2) is also a convex program, i.e., Q is positive semi-definite on the null space of A, the first order conditions are both necessary and sufficient for optimality. The first-order conditions for (1.2) can be transformed into the form of (1.1). To see this, let $B \in \mathbf{R}^{(n-m)\times n}$ be any matrix such that the columns of B^T form a basis for the null space of A. The first-order conditions for the quadratic program (1.2) are (see Dantzig [2])

$$\begin{pmatrix} Ax - b \\ A^T\lambda - Qx + y - c \\ XYe \end{pmatrix} = 0, \quad (x, y) \geq 0, \tag{1.3}$$

where λ and y are the dual variables. To eliminate the dual variables λ from the above system, we pre-multiply the second equation by the nonsingular matrix $[A^T \; B^T]^T$. Noticing that $BA^T = 0$, we obtain

$$0 = \begin{bmatrix} A \\ B \end{bmatrix} (A^T\lambda - Qx + y - c) = \begin{pmatrix} AA^T\lambda - A(Qx - y + c) \\ -BQx + By - Bc \end{pmatrix}.$$

Since AA^T is nonsingular, λ is uniquely determined once x and y are known. Removing the equation for λ, we arrive at the following $2n$-dimensional nonlinear system with non-negativity constraints for (x, y)

$$\begin{pmatrix} Ax - b \\ -BQx + By - Bc \\ XYe \end{pmatrix} = 0, \quad (x, y) \geq 0. \tag{1.4}$$

Clearly, (1.4) is in the form of (1.1) with

$$M = \begin{bmatrix} A \\ -BQ \end{bmatrix}, \quad N = \begin{bmatrix} 0 \\ B \end{bmatrix} \quad \text{and} \quad h = \begin{bmatrix} b \\ Bc \end{bmatrix}. \tag{1.5}$$

When $Q = 0$, the quadratic programming problem (1.2) reduces to the linear programming problem (LP)

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b, \\ &&& x \geq 0. \end{aligned} \tag{1.6}$$

Hence (1.2) also includes the linear program. However, because of the importance of linear programming in optimization, we will state results for linear programming separately; fully aware that they are special cases of quadratic programming. We have shown that the framework of problem (1.1) is quite general.

It is well known (see [1] or [3, p.250]) that for the inequality constrained QP, i.e., in (1.2) we have $Ax \geq b$ instead of $Ax = b$, the first-order conditions for the QP can be formulated as a linear complementarity problem of dimension $n + m$ with $N = -I_{n+m}$ and

$$M = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix}.$$

As was mentioned above, the positive semi-definiteness of M is sufficient to guarantee that the algorithms under investigation produce well-defined iterates (Corollary 2.1). Observe that M is positive semi-definite if and only if Q is positive semi-definite. Under the assumption that Q is positive semi-definite, Megiddo [13] used the linear complementarity problem formulation to show that central paths exist for the inequality constrained QP.

Our formulation (1.4) is for the equality constrained QP (1.2). We stress that no comparison should be made between assumptions on Q for inequality constrained QP and those for equality constrained QP. However, it is worth noting that we only need to require Q to be positive semi-definite on the null space of A to have well-defined iterates (see Corollary 2.1). This is in contrast to the assumption that Q is positive semi-definite on the entire space as was assumed by a number of authors in their studies of similar interior point algorithms for equality constrained quadratic programming (1.2) (see [17], for example). In addition, it is not difficult to see that under the assumption that Q is positive semi-definite on the null space of A, since $F'(x, y)$ is nonsingular for all $(x, y) > 0$ as shown in Corollary 2.1, there exists a central path for the quadratic program (1.2), defined by

$$F(x,y) = \mu \begin{pmatrix} 0 \\ e \end{pmatrix}, \quad (x,y) > 0 \text{ and } \mu > 0.$$

The objective of this work is to analyze the asymptotic behavior of a general interior point algorithm for solving problem (1.1). More specifically, we will study the Q-convergence rate of this general algorithm. The issues of global convergence and complexity are not of concern here.

Recently, Zhang, Tapia and Dennis [22, see Theorem 3.1] established a Q-superlinear convergence theory for a class of primal-dual interior point algorithm for linear programming. In this paper, we extend their result to the general problem (1.1) and therefore extend the result to quadratic programming and linear complementarity problems. In spite of its close connection to [22], we have made this paper self-contained.

We will use the notation:

$$\min(v) = \min_{1 \le i \le n} [v]_i \quad \text{and} \quad \max(v) = \max_{1 \le i \le n} [v]_i$$

for $v \in \mathbf{R}^n$, where $[v]_i$ denotes the i-th component of v.

The paper is organized as follows. In Section 2, we describe a general interior point algorithmic framework for (1.1). Then in Section 3, we present our superlinear convergence rate result. Concluding remarks are given in Section 4.

2 Algorithm

It is now fairly well understood how a class of interior point algorithms can be viewed as damped Newton methods and that the inclusion of the logarithmic barrier term (so-called centering) can be viewed as perturbing the right-hand side of the Newton system. Indeed, Zhang, Tapia and Dennis [22] focused on issues concerning how fast the damped Newton method could approach the Newton method (i.e., step-length approach one), and how fast the perturbation term (barrier parameter) should be phased out so that the fast convergence of Newton's method would not be compromised. Their work covered linear programming applications. As previously mentioned, the objective of the present work is to extend a particular nice part of their superlinear convergence theory to quadratic programming and linear complementarity problems. Our vehicle for accomplishing this objective is the use of the general problem (1.1). We assume that the reader is familiar with the above algorithmic considerations and therefore present our algorithmic framework with no further motivation or explanation.

Recall that $F(x, y)$ is given by (1.1).

Algorithm 1 *Given a pair* $(x_0, y_0) > 0$. *For* $k = 0, 1, 2, \ldots,$ *do*

(1) *Choose* $\sigma_k \in [0, 1)$ *and* $\tau_k \in (0, 1)$. *Set* $\mu_k = \sigma_k x_k^T y_k / n$.

(2) *Solve the following system for* $(\Delta x_k, \Delta y_k)$:

$$F'(x_k, y_k) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -F(x_k, y_k) + \begin{pmatrix} 0 \\ \mu_k e \end{pmatrix}. \tag{2.1}$$

(3) *Compute the step-length:*

$$\alpha_k = \min\left(1, \frac{-\tau_k}{\min(X_k^{-1}\Delta x_k)}, \frac{-\tau_k}{\min(Y_k^{-1}\Delta y_k)}\right). \tag{2.2}$$

(4) *Update:* $x_{k+1} = x_k + \alpha_k \Delta x_k$ *and* $y_{k+1} = y_k + \alpha_k \Delta y_k$.

Notice that in Algorithm 1, we do not require that the starting point (x_0, y_0) be feasible. Also notice that without the perturbation term $\mu_k e$ in the right-hand side of (2.1), the search direction $(\Delta x_k, \Delta y_k)$ is the Newton step. We should expect that only in rare cases would the full Newton step lead to a strictly positive iterate; hence we should expect in most cases to have $\alpha_k < 1$ where α_k is given by (2.2). The choice $\tau_k = 1$ corresponds to allowing steps to the boundary of the positive orthant and a loss of strict feasibility. Therefore, it is natural to view Algorithm 1 as a perturbed and damped Newton's method. We see that if (x_0, y_0) is in Ω, then the iteration sequence $\{(x_k, y_k)\}$ will be strictly feasible. In the case of linear programming, there are no linear equations in $F(x, y)$ that involve both x and y. If $(x_k, y_k) \in \Omega$, then different step-lengths can be used to update x_k and y_k and still retain strictly feasible (x_{k+1}, y_{k+1}). This strategy has been shown to be more efficient in practice (see Lustig, Marsten and Shanno [12], for example). However, it will not affect our results since our analysis will show that as long as $\tau_k \to 1$ both step-lengths will converge to one.

Algorithm 1 covers or is closely related to a wide range of existing interior point algorithms for linear programming, quadratic programming and linear complementarity problems. In particular, it covers most of the existing primal-dual interior point algorithms for linear programming as well as quadratic programming, including Kojima, Mizuno and Yoshise [10], Todd and Ye [19], Monteiro and Adler [16, 17], Lustig [11], Gonzaga and Todd [4], Mizuno, Todd and Ye [14, 15]. Algorithms for linear complementarity problems that are covered by Algorithm 1 include Kojima, Mizuno and Yoshise [9, 8], Kojima, Megiddo and Noma [5], Kojima, Megiddo and Ye [6], Kojima, Mizuno and Noma [7].

Although these algorithms have been motivated and presented in various ways including path-following (homotopy or continuation), potential reduction or affine scaling algorithms, most of them fit into the framework of the perturbed and damped Newton's method applied to the general problem (1.1). Due to the extensive activity in this area, our list of references is not complete. For a more complete list of references, especially in the cases of quadratic programming and linear complementarity problems, we refer the reader to two recent survey papers by Ye [20, 21].

The following proposition gives a condition which guarantees that the iterates produced by Algorithm 1 are well-defined.

Proposition 2.1 *The iterates produced by Algorithm 1 are well-defined if for any positive diagonal matrix $D \in \mathbf{R}^{n\times n}$, the matrix $N - MD$ is nonsingular.*

Proof: Notice that

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

Since (x_0, y_0) is positive, the matrix

$$G = \begin{bmatrix} I & -MY_0^{-1} \\ 0 & I \end{bmatrix}$$

is nonsingular. Thus, the nonsingularity of $F'(x_0, y_0)$ is equivalent to that of

$$GF'(x_0, y_0) = \begin{bmatrix} 0 & N - MY_0^{-1}X_0 \\ Y_0 & X_0 \end{bmatrix}.$$

This latter matrix is nonsingular if and only if $N - MY_0^{-1}X_0$ is nonsingular. By our assumption, (x_1, y_1) is well-defined. An induction argument completes the proof. □

Corollary 2.1 *The iterates produced by Algorithm 1 are well-defined for*

1. *the linear complementarity problem ($N = -I$) with M positive semi-definite,*

2. *the quadratic programming problem (1.2) with Q positive semi-definite on the null space of A,*

3. *the linear programming problem (1.6).*

Proof: We will verify that the condition in Proposition (2.1) is satisfied in the three cases. Let D be any positive diagonal matrix.

For LCP, $N = -I$ leads to $N - MD = -(D^{-1} + M)D$. If M is positive semi-definite, then $N - MD$ is clearly nonsingular because $D^{-1} + M$ is positive definite.

For QP, from (1.5)

$$N - MD = \begin{bmatrix} -AD \\ B(I + QD) \end{bmatrix}.$$

Since $N - MD$ is nonsingular if and only if $(N - MD)D^{-1} = ND^{-1} - M$ is nonsingular, we will prove the nonsingularity of

$$ND^{-1} - M = \begin{bmatrix} -A \\ B(D^{-1} + Q) \end{bmatrix}.$$

Suppose $(ND^{-1} - M)u = 0$ for some $u \in \mathbf{R}^n$. $Au = 0$ implies $u = B^T v$ for some $v \in \mathbf{R}^{n-m}$. Therefore, $B(D^{-1} + Q)u = B(D^{-1} + Q)B^T v = 0$. If Q is positive semi-definite in the null space of A, then $B(D^{-1} + Q)B^T$ is positive definite. This leads to $v = 0$ and consequently $u = B^T v = 0$. So $N - MD$ is nonsingular.

For LP, the conclusion follows immediately from the fact that $Q = 0$ is positive semi-definite. $\square$

We should mention that we have stated Algorithm 1 in the current form purely for the purposes of obtaining a unified theory and notational convenience. By directly applying the perturbed and damped Newton method to the first order conditions for the quadratic program (1.2), it is not difficult to see that an identical iteration sequence $\{(x_k, y_k)\}$ will be generated without eliminating the dual variable λ and introducing the matrix B.

3 Superlinear Convergence

The literature contains numerous studies directed at investigating the convergence properties of interior point algorithms covered by or closely related to Algorithm 1. However, most of these studies were concerned only with the issues of global convergence and complexity. The issue of convergence rate, which is certainly important, has not been thoroughly studied for many interior point algorithms. One of the few papers that studied asymptotic behavior (local convergence) of interior point algorithms is Kojima, Megiddo and Noma [5]. In their paper, Kojima, Megiddo and Noma proved that for a class of complementarity problems, Q-linear, superlinear and quadratic convergence can be achieved by interior point algorithms of the form of Algorithm 1. However, all their convergence rate results were obtained under the restrictive assumption that the Jacobian matrix $F'(x,y)$ was nonsingular at the solution. In this section, we establish a superlinear convergence theory for Algorithm 1 applied to the general problem (1.1). Moreover, our theory does not require the nonsingularity of $F'(x,y)$ at solutions.

It is satisfying that we are able to obtain a superlinear convergence rate without the assumption of nonsingularity of the Jacobian matrix at the solution. In the case of linear programming, this allows us to avoid restrictive nondegeneracy assumptions. The motivation for this theory came from numerical experiments that demonstrated superlinear convergence even for highly degenerate linear programs.

At the k-th iteration of Algorithm 1, let

$$\eta_k = \frac{x_k^T y_k/n}{\min(X_k Y_k e)}.$$

Since $x_k^T y_k/n$ is the average value of the elements of $X_k Y_k e$, it is clear that $\eta_k \geq 1$.

Theorem 3.1 *Let $\{(x_k, y_k)\}$ be generated by Algorithm 1 and $(x_k, y_k) \to (x_*, y_*)$. Suppose the following assumptions hold:*

(i) *strict complementarity,*

(ii) *the sequence* $\{\eta_k\}$ *is bounded,*

(iii) $\tau_k \to 1$ *and* $\sigma_k \to 0$,

(iv) *there exists* $\rho \in [0, 1)$ *such that for k sufficiently large*

$$\Delta x_k^T \Delta y_k \geq -\frac{\rho}{2}(\Delta x_k^T (X_k^{-1} Y_k) \Delta x_k + \Delta y_k^T (X_k Y_k^{-1}) \Delta y_k).$$

Then (x_*, y_*) *solves problem (1.1) and the sequence* $\{F(x_k, y_k)\}$ *component-wise converges to zero Q-superlinearly. Furthermore, the sequence* $\{F(x_k, y_k)\}$ *is Q-superlinearly convergent, i.e., for any norm*

$$\lim_{k\to\infty} \sup \frac{\|F(x_{k+1}, y_{k+1})\|}{\|F(x_k, y_k)\|} = 0.$$

Before we prove Theorem 3.1, we would like to comment on the assumptions of Theorem 3.1. First, Assumption (iv) is not particularly restrictive since we will see later that in the context of linear programming, quadratic programming with Q positive semi-definite on the null space of A and linear complementarity problems with M positive semi-definite, we have the stronger result that $\Delta x_k^T \Delta y_k \geq 0$ for $(x_k, y_k) \in \Omega$. On the other hand, the compatibility of Assumptions (ii) and (iii) may be a cause for concern. It seems as if letting $\tau_k \to 1$ and $\sigma_k \to 0$ might force $\eta_k \to \infty$. However, our numerical experience has shown this not to be the case for linear programming. In our numerical studies with Netlib problems for linear programming, we let $\tau_k \to 1$ and $\sigma_k \to 0$ and always observed strict complementarity and bounded $\{\eta_k\}$. While on occasion we saw some rather large values for η_k's, they eventually leveled off or actually started to decrease as the iterates approached a solution. We did not observe continued growth in the values of η_k as our algorithm converged. Moreover, the observed convergence was clearly Q-superlinear and $\alpha_k \to 1$. Of course, the behavior of $\{\eta_k\}$ varies with several factors including how fast $\{\tau_k\}$ converges to one and $\{\sigma_k\}$ to zero. We do not mean to imply that unbounded $\{\eta_k\}$ cannot occur. Instead, we feel that it appears to be more the exception than the rule in linear programming. It still remains to be seen whether or not this same phenomenon exists in quadratic programming and linear complementarity problems. There is no doubt that this topic merits further study.

To prove Theorem 3.1, we need the following lemma.

Lemma 3.1 *Under the assumptions of Theorem 3.1,*

$$\lim_{k\to\infty} \alpha_k = 1. \tag{3.1}$$

Proof: Define at each iteration

$$p_k = X_k^{-1}\Delta x_k, \quad \text{and} \quad q_k = Y_k^{-1}\Delta y_k. \tag{3.2}$$

At iteration k, from (2.1) and (2.3) we have

$$Y_k\Delta x_k + X_k\Delta y_k = -X_kY_ke + \mu_k e,$$

or equivalently, recalling that $\mu_k = \sigma_k x_k^T y_k/n$ (see Step (1) of Algorithm 1)

$$p_k + q_k = -e + \mu_k(X_kY_k)^{-1}e = -e + \sigma_k T_k e, \tag{3.3}$$

where $T_k = (x_k^T y_k/n)(X_kY_k)^{-1}$. Since $\eta_k = \|T_ke\|_\infty$, Assumptions (ii) and (iii) imply

$$\lim_{k\to\infty}(p_k + q_k) = -e. \tag{3.4}$$

Multiply both sides of (3.3) by $(X_kY_k)^{\frac{1}{2}}$ and consider the square of the ℓ_2-norm. We have the following equality

$$\|(X_kY_k)^{\frac{1}{2}}p_k\|_2^2 + \|(X_kY_k)^{\frac{1}{2}}q_k\|_2^2 + 2\Delta x_k^T\Delta y_k = x_k^Ty_k\left(1 - 2\sigma_k + \sigma_k^2\frac{x_k^Ty_k}{n}\frac{e^T(X_kY_k)^{-1}e}{n}\right).$$

Note that

$$\|(X_kY_k)^{\frac{1}{2}}p_k\|_2^2 = \Delta x_k^T(X_k^{-1}Y_k)\Delta x_k \quad \text{and} \quad \|(X_kY_k)^{\frac{1}{2}}q_k\|_2^2 = \Delta y_k^T(X_kY_k^{-1})\Delta y_k.$$

By Assumption (iv),

$$(1-\rho)(\|(X_kY_k)^{\frac{1}{2}}p_k\|_2^2 + \|(X_kY_k)^{\frac{1}{2}}q_k\|_2^2) \le x_k^Ty_k\left(1 - 2\sigma_k + \sigma_k^2\frac{x_k^Ty_k}{n}\frac{e^T(X_kY_k)^{-1}e}{n}\right).$$

Dividing the above inequality by $x_k^T y_k/n$, we obtain

$$(1-\rho)(\|T_k^{-\frac{1}{2}}p_k\|_2^2 + \|T_k^{-\frac{1}{2}}q_k\|_2^2) \le n(1 - 2\sigma_k + \sigma_k^2 \frac{e^T T_k e}{n}). \tag{3.5}$$

Assumption (ii) implies that $\{\|T_k\|\}$ is bounded above and $\{\|T_k^{-\frac{1}{2}}\|\}$ is bounded away from zero. Therefore, from (3.5) both $\{p_k\}$ and $\{q_k\}$ are bounded. It now follows from (2.2) that $\{\alpha_k\}$ is bounded away from zero.

Now assume $[x_*]_i > 0$. Obviously,

$$1 = \lim_{k\to\infty} \frac{[x_{k+1}]_i}{[x_k]_i} = \lim_{k\to\infty}(1 + \alpha_k [p_k]_i).$$

This implies $[p_k]_i \to 0$, because $\{\alpha_k\}$ is bounded away from zero. From (3.4) we have $[q_k]_i \to -1$. On the other hand, if $[x_*]_i = 0$, then $[y_*]_i > 0$ by strict complementarity. The same argument, interchanging the roles of p_k and q_k, gives $[q_k]_i \to 0$ and $[p_k]_i \to -1$. Therefore, the components of p_k and q_k converge to either 0 or -1. Consequently, from (3.2), (2.2) and $\tau_k \to 1$ it follows that $\alpha_k \to 1$. This completes the proof. □

Now we are ready to prove Theorem 3.1.

Proof of Theorem 3.1: Let

$$F_1(x,y) = Mx + Ny - h \quad \text{and} \quad F_2(x,y) = XYe.$$

We will prove that both $\{F_1(x_k,y_k)\}$ and $\{F_2(x_k,y_k)\}$ component-wise converge to zero Q-superlinearly. This will imply that $\{F(x_k,y_k)\}$ component-wise converges to zero Q-superlinearly. It is not difficult to see that component-wise Q-superlinear convergence of a vector sequence implies its Q-superlinear convergence.

First we show that $\{F_1(x_k,y_k)\}$ component-wise converges to zero Q-superlinearly. If $F_1(x_0,y_0) = 0$ (i.e., (x_0,y_0) is a feasible starting point), then it is easy to see that $F_1(x_k,y_k) = 0$ for all k. Therefore, we need only consider the case where $F_1(x_0,y_0) \neq 0$. Note that Newton's method solves linear equations in one step. If for some integer $p \ge 0$, $\alpha_p = 1$, then we have $F_1(x_k,y_k) = 0$ for all $k > p$. Therefore, we need only consider the case where $\alpha_k < 1$ for all k. It is easy to see from Steps (2) and (4) of Algorithm 1 that

$$F_1(x_{k+1}, y_{k+1}) = (Mx_k + Ny_k - h) + \alpha_k(M\Delta x_k + N\Delta y_k) = (1-\alpha_k)F_1(x_k, y_k).$$

Since $\alpha_k \to 1$, we have that $\{F_1(x_k, y_k)\}$ component-wise converges to zero Q-superlinearly.

Next, we show that $\{F_2(x_k, y_k)\}$ component-wise converges to zero Q-superlinearly. From Step (4) of Algorithm 1,

$$X_k^{-1}x_{k+1} = e + \alpha_k p_k \quad \text{and} \quad Y_k^{-1}y_{k+1} = e + \alpha_k q_k.$$

Adding the above two equations, we have

$$X_k^{-1}x_{k+1} + Y_k^{-1}y_{k+1} = 2e + \alpha_k(p_k + q_k).$$

It follows from (3.4) and $\alpha_k \to 1$ that

$$\lim_{k\to\infty}(X_k^{-1}x_{k+1} + Y_k^{-1}y_{k+1}) = e. \tag{3.6}$$

If $[x_*]_i = 0$, then by strict complementarity, $[y_*]_i > 0$ and $[y_{k+1}]_i/[y_k]_i \to 1$. It follows from (3.6) that $[x_{k+1}]_i/[x_k]_i \to 0$. Therefore, $[x_k]_i \to 0$ Q-superlinearly. By the symmetry of the relation (3.6), we have $[y_k]_j \to 0$ Q-superlinearly if $[y_*]_j = 0$. Thus, all variables that converge to zero do so Q-superlinearly. That is, for each index i either

$$\lim_{k\to\infty}\frac{[x_{k+1}]_i}{[x_k]_i} = 0 \quad \text{and} \quad \lim_{k\to\infty}\frac{[y_{k+1}]_i}{[y_k]_i} = 1$$

or

$$\lim_{k\to\infty}\frac{[x_{k+1}]_i}{[x_k]_i} = 1 \quad \text{and} \quad \lim_{k\to\infty}\frac{[y_{k+1}]_i}{[y_k]_i} = 0.$$

In either case, for every index i,

$$\lim_{k\to\infty}\frac{[x_{k+1}]_i[y_{k+1}]_i}{[x_k]_i[y_k]_i} = \lim_{k\to\infty}\frac{[X_{k+1}Y_{k+1}e]_i}{[X_kY_ke]_i} = 0. \tag{3.7}$$

We have proved that $\{[X_kY_ke]_i\}$ converges to zero Q-superlinearly for every index i. As was mentioned above, the component-wise Q-superlinear convergence of $\{F(x_k, y_k)\}$ implies its Q-superlinear convergence. This completes the proof. □

A key idea in the proof of Theorem 3.1 can be traced back to a work by Tapia in 1980 [18]. In that paper, Tapia pointed out [18, Theorem 3] that an algorithm which at each iteration satisfies the Taylor linearization of the complementarity equation has the property that the variables that converge to zero do so Q-superlinearly. This result assumed strict complementarity and step-length one. Observe that (3.6) is equivalent to

$$X_kY_ke + Y_k(x_{k+1} - x_k) + X_k(y_{k+1} - y_k) \to 0.$$

We see that the Taylor linearization of complementarity is satisfied asymptotically in our situation.

The following theorem deals with the Q-superlinear convergence of Algorithm 1 applied to linear complementarity problems, quadratic programming and linear programming.

Theorem 3.2 *Let $\{(x_k, y_k)\}$ be generated by Algorithm 1 and $(x_k, y_k) \to (x_*, y_*)$. Under Assumptions (i)-(iii) of Theorem 3.1, if $(x_p, y_p) \in \Omega$ for some p, then (x_*, y_*) sovles problem (1.1) and the sequence $\{F(x_k, y_k)\}$ component-wise converges to zero Q-superlinearly for the following three cases:*

1. *the linear complementarity problem ($N = -I$) with M positive semi-definite,*
2. *the quadratic programming problem (1.2) with Q positive semi-definite on the null space of A,*
3. *the linear programming problem (1.6),*

Proof: We need to prove that Assumption (iv) of Theorem 3.1 is satisfied for each of the above three cases. Observe that for all $k \geq p$ we have $(x_k, y_k) \in \Omega$ and $M\Delta x_k + N\Delta y_k = 0$ (see (2.1)). It suffices to prove that $u^Tv \geq 0$ for all $u, v \in \mathbf{R}^n$ satisfying $Mu + Nv = 0$.

In the first case ($N = -I$), $Mu + Nv = 0$ is equivalent to $v = Mu$. Hence $u^Tv = u^TMu \geq 0$ because M is positive semi-definite.

In the second case (see (1.5)), $Mu + Nv = 0$ is equivalent to $Au = 0$ and $BQu = Bv$. Using the representations $u = B^T u_2$ and $v = A^T v_1 + B^T v_2$, where $v_1 \in \mathbf{R}^m$ and $u_2, v_2 \in \mathbf{R}^{n-m}$, and noticing that $A^T \perp B^T$, we have $u^T v = u_2^T B B^T v_2$. Moreover, $BQu = Bv$ is equivalent to $BQB^T u_2 = BB^T v_2$. Hence, if Q is positive semi-definite in the null space of A, then

$$u^T v = u_2^T B B^T v_2 = u_2^T (BQB^T) u_2 \geq 0.$$

The third case follows immediately from the fact that $Q = 0$ is positive semi-definite. $\square$

4 Concluding Remarks

The generality of problem (1.1) and the perturbed and damped Newton's method viewpoint have enabled us to analyze a class of interior point algorithms for linear programming, quadratic programming and linear complementarity problems in a unified approach.

We developed a Q-superlinear convergence theory that does not assume any information on the Jacobian matrix at the solution. This theory was used to establish Q-superlinear convergence for a class of interior point algorithms for linear programming, quadratic programming (with Q positive semi-definite on the null space of A) and positive semi-definite linear complementarity problems.

References

[1] R. W. Cottle and G. B. Dantzig. Complementarity pivot theory of mathematical programming. *Linear Algebra Appl.*, 1:103–125, 1968.

[2] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, NJ, 1963.

[3] R. Fletcher. *Practical Methods of Optimization.* John Wiley & Sons, New York, 1988.

[4] C. C. Gonzaga and M.J. Todd. An $O(\sqrt{n}L)$-iteration large-step primal-dual affine algorithm for linear programming. Technical Report 862, School of Operations Research and Industrial Engineering, Cornell University, 1989.

[5] M. Kojima, N. Megiddo, and T. Noma. Homotopy continuation methods for complementarity problems, 1988. manuscript, IBM Almaden Research Center, San Jose, California.

[6] M. Kojima and Ye Y. Megiddo, N. An interior point potential reduction algorithm for the linear complementarity problems. Technical report, IBM Almaden Research Center, San Jose, California, 1988.

[7] M. Kojima, S. Mizuno, and T. Noma. A new continuation method for complementarity problems with uniform p-functions. *Mathematical Programming*, 43:107–113, 1989.

[8] M. Kojima, S. Mizuno, and A. Yoshise. An $o(\sqrt(n)l)$ iteration potential reduction algorithm for linear complementarity problems. Technical report, Tokyo Institute of Technology, Japan, 1988. Research Report on Information Sciences B-217.

[9] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44:1–26, 1989.

[10] 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.

[11] I.J. Lustig. A generic primal-dual interior point algorithm. Technical Report SOR 88-3, Dept. Civil Eng. and O.R., Princeton University, 1988.

[12] 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.

[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] S. Mizuno, M.J. Todd, and Y. Ye. Anticipated behavior of long-step algorithms for linear programming. Technical Report Working paper Series No. 89-33, College of Business Administration, The University of Iowa, 1989.

[15] S. Mizuno, M.J. Todd, and Y. Ye. Anticipated behavior of path-following algorithms for linear programming. Technical Report 878, School of Operations Research and Industrial Engineering, Cornell University, 1989.

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

[17] R.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part II: convex quadratic programming. *Math. Prog.*, 44:43–66, 1989.

[18] R. A. Tapia. On the role of slack variables in quasi-Nweton methods for constrained optimization. In L. C. W. Dixon and G. P. Szegõ, editors, *Numerical Optimization of dynamic systems*, pages 235–246. North-Holland, 1980.

[19] M.J. Todd and Yinyu Ye. A centered projective algorithm for linear programming. Technical Report No. 763, School of Operations Research and Industrial Engineering, Cornell University, 1987 (revised 1989). To appear in *Math. Prog.*

[20] Yinyu Ye. Interior point algorithms for quadratic programming. Technical report, Dept. of Management Sciences, The University of Iowa, 1989. Working Paper Series No. 89-29.

[21] Yinyu Ye. Interior point algorithms for global optimization. Technical report, Dept. of Management Sciences, The University of Iowa, 1990. Working Paper Series No. 90-2.

[22] Yin 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.