September 30, 1996

SUPERLINEAR CONVERGENCE OF AN ALGORITHM FOR MONOTONE LINEAR COMPLEMENTARITY PROBLEMS, WHEN NO STRICTLY COMPLEMENTARY SOLUTION EXISTS

Jos F. Sturm

Report 9656/A

Econometric Institute, Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
URL: http://www.eur.nl/few/ei/~sturm
E-mail: sturm@few.eur.nl

ABSTRACT

A new predictor–corrector interior point algorithm for solving monotone linear complementarity problems (LCP) is proposed, and it is shown to be superlinearly convergent with at least order 1.5, even if the LCP has no strictly complementary solution. Unlike Mizuno's recent algorithm [16], the fast local convergence is attained without any need for estimating the optimal partition. In the special case that a strictly complementary solution does exist, the order of convergence becomes quadratic. The proof relies on an investigation of the asymptotic behavior of first and second order derivatives that are associated with trajectories of weighted centers for LCP.

AMS 1991 subject classification: 90C33.

Key words. monotone linear complementarity problem, primal-dual interior point method, superlinear convergence, central path.

1. Introduction

Given $n \times n$ real matrices Q and R and a real vector b of order n, the horizontal linear complementarity problem (LCP) is the problem of finding a member of the following set:

$$\mathcal{F}^* = \{(x,s) \in \Re^n_+ \times \Re^n_+ \mid Xs = 0,\, Qx + Rs = b\}, \tag{1}$$

where capital X denotes the diagonal matrix with entries $X_{ii} = x_i$ for $i = 1, 2, \ldots, n$. In this paper, we will only consider the *monotone* LCP, i.e. we assume that

$$Qu + Rv = 0 \Rightarrow u^{\mathrm{T}} v \geq 0. \tag{2}$$

The monotone LCP covers convex quadratic programming (QP), since the Karush-Kuhn-Tucker (KKT) optimality conditions of QPs form monotone LCPs.

It was shown by Kojima, Mizuno and Yoshise [10] that the monotone LCP is polynomially solvable. Kojima et al. [8] developed a unified framework in which interior point methods for LCPs can be analyzed. Interior point algorithms solve LCP by generating a sequence $(x^0, s^0), (x^1, s^1), \ldots$ in

$$\{(x,s) > 0 \mid Qx + Rs = b\},$$

(the set of interior points), with

$$\lim_{k\to\infty} (x^k)^{\mathrm{T}} s^k = 0.$$

The quantity $x^{\mathrm{T}} s$ is commonly referred to as the *duality gap*, due to the interpretation of the monotone LCP as a KKT-system for QP. Moreover, $x_1, x_2, \ldots, x_n$ are called *primal* variables, and the variables $s_1, s_2, \ldots, s_n$ are called *dual*.

More recently, research has been devoted to the design of interior point algorithms for the monotone LCP that are demonstrably superlinearly convergent under certain conditions. Ji, Potra and Huang [7] showed that an extension of the Mizuno-Todd-Ye predictor-corrector algorithm [17] to LCP is superlinearly convergent if the iteration sequence converges to a *strictly* complementary solution. A strictly complementary solution is a solution (x, s) satisfying

$$(x,s) \in \mathcal{F}^*, \qquad x + s > 0.$$

In Ye and Anstreicher [26], the same algorithm was shown to have even a quadratic order of convergence, assuming only the *existence* of a strictly complementary solution. By adding improvement steps to each main iteration of the algorithm, Wright and Zhang [24] obtain superquadratic convergence under the strict complementarity assumption. The largest step path-following algorithm, which is slightly different from the Mizuno-Todd-Ye predictor-corrector algorithm, has gone through a similar development: McShane [14] proved the superlinear convergence under the same hypothesis as in [7]. Then Bonnans and Gonzaga [1] obtained quadratic and — if additional improvement steps are used — superquadratic convergence, assuming only the existence of a strictly complementary solution.

In the absence of strict complementarity however, neither the Mizuno-Todd-Ye nor the largest step algorithm will exhibit superlinear convergence, see Monteiro and Wright [19] and El-Bakry, Tapia and Zhang [3].

Most interestingly, Mizuno [16] proposed a modification of the Mizuno-Todd-Ye predictor-corrector algorithm achieving superlinear convergence of the duality gap with Q-order 1.25, without assuming strict complementarity. Potra and Sheng [22, 21] extended the algorithm to $P_*(\kappa)$-LCPs and also included some practical enhancements. A drawback of Mizuno's algorithm and its variants is their need to estimate the optimal partition in order to achieve superlinear convergence.

Just like Mizuno's algorithm, the algorithm to be proposed in this paper has quadratic convergence if a strictly complementary solution exists. However, the new algorithm does not use estimates of the optimal partition, but still achieves superlinear convergence without strict complementarity. Moreover, it generates iterates that converge Q-superlinearly to the solution set $\mathcal{F}^*$ with at least order 1.5, which is clearly an improvement over the local convergence properties of Mizuno's algorithm. The new algorithm is obtained naturally, by modifying the Mizuno-Todd-Ye predictor-corrector algorithm along two lines, viz. we will use

- a different parameterization of weighted trajectories, and
- a second order predictor step.

The use of a second order predictor step was first proposed by Mehrotra [15] for linear programming. Zhang and Zhang [27] described a variant of Mehrotra's algorithm with polynomial time convergence for monotone LCPs. However, the local convergence of Zhang and Zhang's algorithm is only linear when no strictly complementary solution exists.

This paper is organized as follows. After stating some preliminary results in Section 2, we shall prove in Section 3 that the new parameterization makes the first- and second order derivatives along weighted trajectories bounded. The error-bound of Mangasarian and Shiau [13], later also obtained by Luo and Pang [11], will play a crucial role there. Some of the results in Section 3 were known already from Kojima, Mizuno and Noma [9] and Monteiro and Tsuchiya [18], but are re-established with a new and clarifying proof. Second order derivatives along weighted trajectories for LCP are treated for the first time. This investigation of the asymptotic properties of weighted trajectories is algorithm independent, and of interest in itself. We will prove the polynomial and superlinear convergence of the new algorithm in Sections 4.1 and 4.2 respectively. We conclude the paper in Section 5, by addressing some topics of future research.

Notation. The set of order n vectors with real entries is denoted $\Re^n$. We write $x \in \Re^n_+$ ($x \in \Re^n_{++}$) if for all $i = 1, 2, \ldots, n$ there holds $x_i \geq 0$ ($x_i > 0$). Given a vector $x \in \Re^n$, the corresponding upper case symbol X denotes the diagonal matrix with entries $X_{11} = x_1, X_{22} = x_2, \ldots, X_{nn} = x_n$. The Euclidean norm is denoted $\|x\|$, and $\|x\|_p$ is the l_p-norm of x. We let e be the all-one vector.

Given a finite index set J, we let $\mid J \mid$ denote its cardinality. For $x \in \Re^n$, we denote the subvector $[x_i]_{i \in J}$ of order $\mid J \mid$ by x_J. Given a matrix $A \in \Re^{m \times n}$, the submatrix $A_J \in \Re^{m \times |J|}$ has columns $A_{\cdot j}$ for $j \in J$. The kernel and the image of A are denoted by Ker A and Img A respectively.

We use the standard Landau notation ("big O"). In particular, if $\{u(t) \mid t > 0\}$ and $\{w(t) \mid t > 0\}$ are real sequences with $w(t) > 0$, then $u(t) = O(w(t))$ means that $u(t)/w(t)$ is bounded, independent of t; if also $u(t) > 0$, we write $u(t) = \Theta(w(t))$ whenever $u(t)/w(t)$ and $w(t)/u(t)$ are both bounded. Moreover, if $\{u(t) \mid t > 0\}$ is a sequence of vectors in $\Re^n$, then $u(t) = O(w(t))$, ($u(t) = \Theta(w(t))$) means that $u_i(t) = O(w(t))$ ($u_i(t) = \Theta(w(t))$) for each $i = 1, 2, \ldots, n$.

2. Preliminaries

The monotonicity assumption (2) has some important implications. The following result is a particularly useful one; similar results were given in Bonnans and Gonzaga [2] and, for the standard form LCP, in Kojima et al. [8], among others.

Lemma 2.1. *Given a positive integer m, let $T : \Re^m \times \Re^m \to \Re^m$ be a linear mapping with the monotonicity property*

$$T(u, v) = 0 \Rightarrow u^T v \geq 0.$$

Let C and $D \in \Re^{m\times m}$ be invertible matrices. Then the linear mapping $\hat{T} : \Re^m \times \Re^m \to \Re^m \times \Re^m$, defined by

$$\hat{T}(u, v) = \begin{bmatrix} CD^{-T}u + CDv \\ T(u, v) \end{bmatrix},$$

is invertible.

Proof:

Let (u, v) be an element of the kernel of $\hat{T}$, then

$$u = -D^{\mathrm{T}} Dv, \quad T(u, v) = 0.$$

However, from the monotonicity property of T we have

$$0 \leq u^{\mathrm{T}} v = -v^{\mathrm{T}} D^{\mathrm{T}} Dv = -\|Dv\|^2 \leq 0,$$

which implies that $Dv = 0$. Since D is invertible, we conclude that

$$u = -D^{\mathrm{T}}(Dv) = 0, \quad v = D^{-1}(Dv) = 0.$$

Hence, the kernel contains only the zero element, which means that $\hat{T}$ is invertible.

$\square$

Remark. As a consequence of Lemma 2.1, we observe that the matrix $\begin{bmatrix} Q & R \end{bmatrix}$ has full row rank. Relabeling pairs (x_i, s_i) as (s_i, x_i) when necessary, we can therefore reformulate the horizontal LCP (1) as a *standard form* LCP: find $(x, s) \in \Re^n_+ \times \Re^n_+$ satisfying

$$Xs = 0, \quad s = Mx + h,$$

with $h \in \Re^n$ and $M \in \Re^{n\times n}$. This equivalence between the horizontal LCP and the standard form LCP is also shown in Sznajder and Gowda [23], Güler [5], Monteiro and Tsuchiya [18] and Bonnans and Gonzaga [2]. Since the monotonicity (2) implies that the matrix $M + M^{\mathrm{T}}$ is positive semidefinite, it follows that the monotone LCP is equivalent to the QP

$$\min\{x^{\mathrm{T}}Mx + h^{\mathrm{T}}x \mid Mx + h \geq 0,\, x \geq 0\},$$

see also Murty [20].

Throughout this paper, we make the standard assumption that there exist interior solutions, viz.

$$\{(x, s) \in \Re^n_{++} \times \Re^n_{++} \mid Qx + Rs = b\} \neq \emptyset. \tag{3}$$

This assumption can be enforced by an embedding of the original LCP, see Kojima et al. [8] and Ye [25]. As is well known, Assumption 3 guarantees that $\mathcal{F}^*$ is nonempty and bounded, see e.g. Mangasarian [12].

3. Trajectory of weighted centers

For a given vector $w \in \Re^n_{++}$, the trajectory of w-weighted centers is defined as the trajectory $\{(x(t; w), s(t; w)) \mid t > 0\}$ satisfying

$$x_i(t; w)s_i(t; w) = t^2 w_i^2 \text{ for } i = 1, 2, \ldots, n, \tag{4}$$

$$Qx(t; w) + Rs(t; w) = b, \tag{5}$$

$$x(t; w) \geq 0, \quad s(t; w) \geq 0. \tag{6}$$

We remark that (4) differs from the usual parameterization [8], which requires

$$x_i(t; w)s_i(t; w) = tw_i^2 \text{ for } i = 1, 2, \ldots, n, \tag{7}$$

instead of (4). At first sight, this subtle difference may seem irrelevant, since t^2 is a monotone transformation of t for $t > 0$. However, the change in parameterization has an essential impact on the derivatives of the trajectory of w-weighted centers, as we will see in the sequel.

For each $t > 0$, the system (4)–(6) has a unique solution $(x(t;w), s(t;w))$, see Kojima, Mizuno and Noma [9], Kojima et al. [8] and Monteiro and Tsuchiya [18]. Applying Lemma 2.1 with

$$C = tW, \quad D = \frac{1}{t}W^{-1}X(t;w),$$

it follows that the Jacobian of the system (4)–(5) is nonsingular for $t > 0$. Hence, we can apply the classical inverse function theorem (or alternatively, the implicit function theorem) to conclude that the trajectory $\{(x(t;w), s(t;w)) \mid t > 0\}$ is smooth, i.e. infinitely many times continuously differentiable. This fact is known from Theorem 4.4 in Kojima et al. [8].

For simplicity, we will restrict ourselves to those $w \in \Re^n$ for which

$$\frac{1}{2n} \leq w_i \leq 2n \text{ for } i = 1, 2, \ldots, n. \tag{8}$$

In fact, there is nothing magical to the number $2n$ above: its single purpose is to allow for bounds on $(x(t;w), s(t;w))$ and its derivatives that do not depend on w. This becomes important in Section 4, where the weight vector w is altered iteratively, but it will always satisfy (8).

As usual, $(\dot{x}(t;w), \dot{s}(t;w))$ and $(\ddot{x}(t;w), \ddot{s}(t;w))$ will denote the trajectories of first and second order derivatives along the w-weighted path. Differentiating (4) with respect to t yields

$$x_i(t;w)\dot{s}_i(t;w) + s_i(t;w)\dot{x}_i(t;w) = 2tw_i^2, \tag{9}$$

and differentiating once more, it follows that

$$2\dot{x}_i(t;w)\dot{s}_i(t;w) + x_i(t;w)\ddot{s}_i(t;w) + s_i(t;w)\ddot{x}_i(t;w) = 2w_i^2. \tag{10}$$

Similarly, we obtain from (5) that

$$Q\dot{x}(t;w) + R\dot{s}(t;w) = 0, \quad Q\ddot{x}(t;w) + R\ddot{s}(t;w) = 0. \tag{11}$$

We have already pointed out that the derivatives exist and are continuous for $t > 0$, since the Jacobian of the system (4)–(5) is smooth and nonsingular.

Let (x^*, s^*) be in the relative interior of $\mathcal{F}^*$. Since $X^*s^* = 0$, we have for each $i \in \{1, 2, \ldots, n\}$ that either $x_i^* = 0$ or $s_i^* = 0$ or both. Relabeling appropriate pairs of complementary variables (x_i^*, s_i^*) as (s_i^*, x_i^*), we can assume without loss of generality that

$$s^* = 0,$$

see also Bonnans and Gonzaga [2]. Now, let B denote the support of x^*, and T its complement, i.e.

$$x_B^* > 0, \quad x_T^* = 0, \quad B \cup T = \{1, 2, \ldots, n\}.$$

Remark here that definition (1) implies

$$Q_B x_B^* = b. \tag{12}$$

We will refer to B as the *basic* index set, and to T as the *double root* index set. Using the monotonicity of the LCP, it is easily verified that $\mathcal{F}^*$ is a convex set. Hence, the fact that (x^*, s^*) is in the *relative interior* of $\mathcal{F}^*$ implies that

$$s = 0, \quad x_T = 0 \text{ for all } (x, s) \in \mathcal{F}^*. \tag{13}$$

In other words, (x^*, s^*) has the largest number of nonzeroes, viz. $\mid B \mid$, for all solutions $(x, s) \in \mathcal{F}^*$. This explains why the relative interior of $\mathcal{F}^*$ is sometimes called the set of *maximal complementary* solutions. The w-weighted analytic center of the solution set $\mathcal{F}^*$ is the vector $x(0; w)$ defined by

$$x(0; w) := \arg\max_x \{\sum_{i\in B} w_i^2 \log x_i \mid Q_B x_B = b,\, x_B > 0,\, x_T = 0\}.$$

From the strict convexity of the barrier $\sum_{i\in B} w_i^2 \log x_i$, it follows that $x_B(0; w)$ is characterized by the Karush-Kuhn-Tucker system

$$\begin{cases} X_B(0; w)^{-1} W_B^2 e - Q_B^{\mathrm{T}} \lambda = 0, \\ Q_B x_B(0; w) = b, \\ x_B(0; w) > 0. \end{cases} \tag{14}$$

It is known from Kojima, Mizuno and Noma [9] and Monteiro and Tsuchiya [18] that $x(0; w) = \lim_{t\downarrow 0} x(t; w)$. We will re-establish this result in Theorem 3.1.

Some interesting properties of weighted central paths follow easily from the error-bound result of Mangasarian and Shiau [13] for LCP, which was also proved in Luo and Pang [11]. Namely, it is shown in [13, 11] that if (x, s) satisfies the feasibility requirements

$$Qx + Rs = b, \quad x \geq 0, \quad s \geq 0,$$

then

$$\operatorname{dist}((x, s), \mathcal{F}^*) = O(\sqrt{x^{\mathrm{T}} s}) \text{ if } x^{\mathrm{T}} s < 1. \tag{15}$$

Combining this with (13), it follows that

$$s = O(\sqrt{x^{\mathrm{T}} s}), \quad x_T = O(\sqrt{x^{\mathrm{T}} s}) \text{ if } x^{\mathrm{T}} s < 1.$$

Now consider the weighted central path $\{(x(t; w), s(t; w)) \mid 0 < t < 1\}$. Since

$$x(t; w)^{\mathrm{T}} s(t; w) = t^2 \|w\|^2 = \Theta(t^2),$$

we have $x_T(t;w) = O(t)$ and $s_T(t;w) = O(t)$. However, we also have

$$x_i(t;w)s_i(t;w) = t^2 w_i^2 = \Theta(t^2) \text{ for } i = 1, 2, \ldots, n,$$

so that

$$x_T(t;w) = \Theta(t), \quad s_T(t;w) = \Theta(t). \tag{16}$$

From Güler and Ye [6] it is known that $x_B(t;w) = \Theta(1)$ for $0 < t < 1$, so that we further obtain

$$x_B(t;w) = \Theta(1), \quad s_B(t;w) = \Theta(t^2). \tag{17}$$

Letting

$$s(0;w) := \lim_{t \downarrow 0} s(t;w), \quad x_T(0;w) := \lim_{t \downarrow 0} x_T(t;w),$$

it follows that $s(0;w) = 0$ and $x_T(0;w) = 0$. We want to stress here that it is due to assumption (8) that we can bound $x(t;w)$ and $s(t;w)$ independently of w. Bonnans and Gonzaga [2] observed that the estimations (16) and (17) also follow from Monteiro and Tsuchiya [18]; an alternative proof is provided in the recent article of Mizuno [16].

In the sequel of this section, we will analyze the limiting behavior of the w-weighted trajectory and its first- and second-order derivatives for $t \downarrow 0$, in the 'new' parameterization (4). Our findings will be summarized in Theorem 3.1. For an analysis of the first order derivatives in the 'old' parameterization (7), we refer to Monteiro and Tsuchiya [18]. In an attempt to simplify notations, we will drop the argument 'w' in the remainder of this section, i.e.

$$x(t) = x(t;w), \quad \dot{x}(t) = \dot{x}(t;w), \quad \ldots,$$

if no confusion is possible.

3.1. Behavior of the dual basic variables

Let $0 < t < 1$, and consider the trajectory of derivatives $\{(\dot{x}(t), \dot{s}(t)) \mid 0 < t < 1\}$. Dividing (9) by tw_i yields, using (4),

$$\frac{tw_i \dot{x}_i(t)}{x_i(t)} + \frac{x_i(t)\dot{s}_i(t)}{tw_i} = 2w_i \text{ for } i = 1, 2, \ldots, n,$$

so that due to monotonicity, we have $\dot{x}(t)^{\mathrm{T}} \dot{s}(t) \geq 0$ and

$$\begin{aligned} 4\|w\|^2 &= \left\| tWX(t)^{-1}\dot{x}(t) + \frac{1}{t}W^{-1}X(t)\dot{s}(t) \right\|^2 \\ &\geq \left\| tWX(t)^{-1}\dot{x}(t) \right\|^2 + \left\| \frac{1}{t}W^{-1}X(t)\dot{s}(t) \right\|^2. \end{aligned}$$

Combining this with (16)-(17), we obtain

$$\dot{x}_B(t) = O(\frac{1}{t}), \quad \dot{s}_B(t) = O(t), \quad \dot{x}_T(t) = O(1), \quad \dot{s}_T(t) = O(1). \tag{18}$$

Letting

$$\dot{s}_B(0) := \lim_{t \downarrow 0} \dot{s}_B(t),$$

it follows that $\dot{s}_B(0) = 0$. With a similar argument, we will show that $\ddot{s}_B(t) = O(1)$ for $t \downarrow 0$. First, divide (10) by tw_i to obtain

$$\frac{tw_i \ddot{x}_i(t)}{x_i(t)} + \frac{x_i(t)\ddot{s}_i(t)}{tw_i} = \frac{2}{t}(w_i - \frac{\dot{x}_i(t)\dot{s}_i(t)}{w_i}) \text{ for } i = 1, 2, \ldots, n,$$

so that due to monotonicity, we have $\ddot{x}(t)^{\mathrm{T}}\ddot{s}(t) \geq 0$ and

$$\begin{aligned} \frac{2}{t}\left\| w - W^{-1}\dot{X}(t)\dot{s}(t)\right\| &= \left\| tWX(t)^{-1}\ddot{x}(t) + \frac{1}{t}W^{-1}X(t)\ddot{s}(t)\right\| \\ &\geq \sqrt{\|tWX(t)^{-1}\ddot{x}(t)\|^2 + \left\|\frac{1}{t}W^{-1}X(t)\ddot{s}(t)\right\|^2}. \end{aligned}$$

Using (17) and (18), it follows from the above inequality that

$$\ddot{s}_B(t) = O(1). \tag{19}$$

The following lemma summarizes (17)–(19).

Lemma 3.1. *Let $w \in \Re^n_{++}$ satisfy (8), and consider the w-weighted path $\{(x(t;w), s(t;w)) \mid t > 0\}$. For $0 < t < 1$, we have*

$$s_B(t;w) = \Theta(t^2), \quad \dot{s}_B(t;w) = O(t), \quad \ddot{s}_B(t;w) = O(1).$$

3.2. Behavior of the double root variables

Let K be a matrix whose columns span Ker Q_B^{T}. The next lemma shows that

$$\{(u, v) \mid K^{\mathrm{T}}Q_T u + K^{\mathrm{T}}R_T v = 0\}$$

is a monotone linear space.

Lemma 3.2. *If $(u, v) \in \Re^{|T|} \times \Re^{|T|}$ satisfies*

$$K^T Q_T u + K^T R_T v = 0, \tag{20}$$

then $u^T v \geq 0$.

Proof:

Relation (20) can be rewritten as

$$Q_T u + R_T v = Q_B y \text{ for some } y \in \Re^{|B|}.$$

Due to the monotonicity assumption (2), this implies that

$$0 \leq \begin{bmatrix} -y^{\mathrm{T}} & u^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} 0 \\ v \end{bmatrix} = u^{\mathrm{T}} v.$$

□

We will now use Lemma 3.2 to study the behavior of the double root variables.

Lemma 3.3. *Let $w \in \Re^n_{++}$ satisfy (8), and consider the w-weighted path $\{(x(t;w), s(t;w)) \mid t > 0\}$. For $0 < t < 1$, we have*

$$x_T(t;w) = \Theta(t), \quad \dot{x}_T(t;w) = O(1), \quad \ddot{x}_T(t;w) = O(1),$$

and

$$s_T(t;w) = \Theta(t), \quad \dot{s}_T(t;w) = O(1), \quad \ddot{s}_T(t;w) = O(1).$$

Proof:

For $t \in (0,1)$, define

$$\bar{x}(t) := x(t)/t, \quad \bar{s}(t) := s(t)/t.$$

Differentiating with respect to t yields

$$\frac{\mathrm{d}\bar{x}(t)}{\mathrm{d}t} = -\frac{x(t) - t\dot{x}(t)}{t^2} \tag{21}$$

and

$$\begin{aligned} \frac{\mathrm{d}^2\bar{x}(t)}{\mathrm{d}t^2} &= \frac{\ddot{x}(t)}{t} + 2\frac{x(t) - t\dot{x}(t)}{t^3} \\ &= \frac{1}{t}\left[\ \ddot{x}(t) - 2\mathrm{d}\bar{x}(t)/\mathrm{d}t\ \right]. \end{aligned} \tag{22}$$

Similar identities obviously hold for the derivatives of $\bar{s}(t)$ as well. Hence, the bounds of Lemma 3.1 imply that

$$\bar{s}_B(t) = \Theta(t), \quad \frac{\mathrm{d}\bar{s}_B(t)}{\mathrm{d}t} = O(1), \quad \frac{\mathrm{d}^2\bar{s}_B(t)}{\mathrm{d}t^2} = O(\frac{1}{t}). \tag{23}$$

Introducing $\bar{s}_B(0) := \lim_{t\downarrow 0} \bar{s}_B(t)$, we immediately obtain that $\bar{s}_B(0) = 0$.

Consider the mapping $F : \Re^{|T|} \times \Re^{|T|} \to \Re^{|T|} \times \Re^{|T|}$ defined by

$$F(u, v) := \begin{bmatrix} Uv - W_T^2 e \\ K^{\mathrm{T}} Q_T u + K^{\mathrm{T}} R_T v \end{bmatrix}.$$

From (4)–(5), we have

$$F(\bar{x}_T(t), \bar{s}_T(t)) + \begin{bmatrix} 0 \\ K^{\mathrm{T}} R_B \bar{s}_B(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{t} K^{\mathrm{T}} b \end{bmatrix} = 0, \tag{24}$$

where we used (12). Due to (16), the sequence $\{(\bar{x}_T(t), \bar{s}_T(t)) \mid 0 < t < 1\}$ is bounded. Consequently, it has a cluster point for $t \downarrow 0$. Since $\bar{s}_B(t) \to 0$, any such cluster point (u, v) satisfies the following relations:

$$F(u, v) = 0, \tag{25}$$

$$u \geq 0, \qquad v \geq 0. \tag{26}$$

Using Lemma 3.2, it follows that the system (25)-(26) cannot have multiple solutions, cf. the uniqueness of w-weighted centers (4)-(6) shown in [9, 8]. This implies that the following limits exist:

$$\dot{x}_T(0) := \lim_{t \downarrow 0} \bar{x}_T(t), \qquad \dot{s}_T(0) := \lim_{t \downarrow 0} \bar{s}_T(t).$$

Moreover, $\dot{X}_T(0)\dot{s}_T(0) = W_T^2 e$, because $F(\dot{x}_T(0), \dot{s}_T(0)) = 0$. Due to Lemma 3.2, we can apply Lemma 2.1 with

$$C = W_T, \qquad D = W_T^{-1} \dot{X}_T(0),$$

to conclude that the Jacobian $\nabla F(\dot{x}_T(0), \dot{s}_T(0))$ is nonsingular. Using also (24), it follows from the classical inverse function theorem that there exists a smooth mapping $G(\cdot)$, defined in an open neighborhood of $\bar{s}_B(0) = 0$, such that

$$(\bar{x}_T(t), \bar{s}_T(t)) = G(\bar{s}_B(t)),$$

for t sufficiently close to zero. Using (23) and the fact that $G(\cdot)$ is smooth in a neighborhood of $\bar{s}_B(0)$, it follows that

$$\frac{\mathrm{d}\bar{x}_T(t)}{\mathrm{d}t} = O(1), \qquad \frac{\mathrm{d}^2 \bar{x}_T(t)}{\mathrm{d}t^2} = O(\frac{1}{t}),$$

and

$$\frac{\mathrm{d}\bar{s}_T(t)}{\mathrm{d}t} = O(1), \qquad \frac{\mathrm{d}^2 \bar{s}_T(t)}{\mathrm{d}t^2} = O(\frac{1}{t}),$$

for $t \downarrow 0$. Using the identities (21)–(22) and the bounds (16), we obtain

$$\dot{x}_T(t) = O(1), \quad \ddot{x}_T(t) = O(1),$$

and similarly,

$$\dot{s}_T(t) = O(1), \quad \ddot{s}_T(t) = O(1).$$

Remark that the bounds $\dot{x}_T(t) = O(1)$ and $\dot{s}_T(t) = O(1)$ were already known from (18), and from (16) we have $x_T(t) = \Theta(t)$ and $s_T(t) = \Theta(t)$.

□

Remark. The above proof also shows that $(\dot{x}_T(0), \dot{s}_T(0))$ is a w_T-weighted center of a homogeneous LCP.

3.3. Behavior of the primal basic variables

Introduce a linear subspace $\mathcal{S}_B$ of $\Re^{|B|}$ as follows:

$$\mathcal{S}_B := \{v \in \Re^{|B|} \mid R_B v \in \mathrm{Img} \begin{bmatrix} Q & R_T \end{bmatrix}\}.$$

Due to (12), we have for any (x, s) satisfying $Qx + Rs = b$ that

$$Q_B(x_B - x_B^*) + Q_T x_T + Rs = 0,$$

and therefore

$$\lambda s_B(t) \in \mathcal{S}_B \text{ for all } \lambda \in \Re. \tag{27}$$

The following result is similar to Lemma 3.4 in [26], Lemma 3.3 in [18] and Lemma 5.1 in [2].

Lemma 3.4. *There holds*

$$\mathcal{S}_B \subseteq \mathit{Img}\ Q_B^T.$$

Proof:

Let $v_B \in \mathcal{S}_B$. Then there exist v_T and u such that

$$Qu + R_B v_B + R_T v_T = 0.$$

Therefore, we have for any $q \in \mathrm{Ker}\ Q_B$ and $\lambda \in \Re$ that

$$Q_B(u_B + \lambda q) + Q_T u_T + R_B v_B + R_T v_T = 0,$$

and due to monotonicity,

$$(u_B + \lambda q)^{\mathrm{T}} v_B + u_T^{\mathrm{T}} v_T \geq 0.$$

Since λ is arbitrary, this implies $v_B \perp q$, for any $q \in \text{Ker } Q_B$. It follows that $v_B \in \text{Img } Q_B^{\mathrm{T}}$.

$\square$

Let l denote the dimension of $\mathcal{S}_B$. We know from Lemma 3.4 that there exists a matrix L of order $n \times l$ such that

$$\mathcal{S}_B = \text{Img } Q_B^{\mathrm{T}} L.$$

Lemma 3.5. *Let $w \in \Re^n_{++}$ satisfy (8), and consider the w-weighted path $\{(x(t;w), s(t;w)) \mid t > 0\}$. The path converges to the w-weighted analytic center of $\mathcal{F}^*$, i.e.*

$$\lim_{t \downarrow 0} x(t;w) = x(0;w).$$

Moreover, for $0 < t < 1$ we have

$$x_B(t;w) = \Theta(1), \quad \dot{x}_B(t;w) = O(1), \quad \ddot{x}_B(t;w) = O(1),$$

and if $T = \emptyset$, then

$$\dot{x}(t;w) = O(t).$$

Proof:

Consider the mapping $F : \Re^{|B|} \times \Re^l \to \Re^{|B|} \times \Re^l$ defined by

$$F(u,v) := \begin{bmatrix} U^{-1} W_B^2 e - Q_B^{\mathrm{T}} L v \\ L^{\mathrm{T}} Q_B u - L^{\mathrm{T}} b \end{bmatrix}.$$

As is known from (27), $s_B(t) \in \mathcal{S}_B = \text{Img } Q_B^{\mathrm{T}} L$, which means that there exists a sequence $y(t)$ in $\Re^l$ such that

$$\frac{1}{t^2} s_B(t) = Q_B^{\mathrm{T}} L y(t).$$

From (4)–(6), we have

$$F(x_B(t), y(t)) + \begin{bmatrix} 0 \\ L^{\mathrm{T}} Q_T x_T(t) + L^{\mathrm{T}} R s(t) \end{bmatrix} = 0, \tag{28}$$

$$Q_B x_B(t) = b,$$

$$x_B(t) \geq 0, \quad s_B(t) = t^2 Q_B^{\mathrm{T}} L y(t) \geq 0.$$

Due to (17), the sequence $\{(x_B(t), y(t)) \mid 0 < t < 1\}$ is bounded, and hence it has a cluster point (u, v) for $t \downarrow 0$, satisfying the relations

$$F(u, v) = 0,$$

$$Q_B u = b, \quad u \geq 0.$$

By (14), this means that $u = x_B(0)$, the w-weighted analytic center of $\mathcal{F}^*$. Since the matrix $Q_B^{\mathrm{T}} L$ has full rank, viz. rank l, it follows that v is unique as well. Hence, $\lim_{t\downarrow 0} x_B(t) = x_B(0)$ and the limit

$$y(0) := \lim_{t\downarrow 0} y(t)$$

exists. It is easily verified that the Jacobian $\nabla F(x_B(0), y(0))$ is nonsingular. Using (28), it follows from the classical inverse function theorem that there exists a smooth mapping $G(\cdot, \cdot)$, defined in an open neighborhood of $(s(0), x_T(0)) = (0, 0)$, such that

$$(x_B(t), y(t)) = G(s(t), x_T(t)),$$

for t sufficiently close to zero. Using Lemma 3.1, Lemma 3.3 and the fact that $G(\cdot, \cdot)$ is smooth in a neighborhood of $(s(0), x_T(0))$, it follows that

$$\dot{x}_B(t) = O(1), \quad \ddot{x}_B(t) = O(1),$$

for $t \downarrow 0$.

Finally, consider the special case $T = \emptyset$. Then $G(\cdot)$ is a smooth function, defined in a neighborhood of $s(0) = 0$, with

$$(x_B(t), y(t)) = G(s(t)),$$

for t sufficiently close to zero. However, we know from Lemma 3.1 that $\dot{s}(t) = O(t)$, and therefore $\dot{x}_B(t) = O(t)$ as well, if $T = \emptyset$.

□

3.4. Summary of the limiting behavior

The results obtained in Lemma 3.1, Lemma 3.3 and Lemma 3.5 are summarized in the following theorem.

Theorem 3.1. *Let $w \in \Re^n_{++}$ satisfy (8), and consider the w-weighted path $\{(x(t;w), s(t;w)) \mid t > 0\}$. The path converges to the w-weighted analytic center of $\mathcal{F}^*$, i.e.*

$$\lim_{t \downarrow 0} x(t;w) = x(0;w), \quad \lim_{t \downarrow 0} s(t;w) = 0.$$

Moreover, we have for $0 < t < 1$ that

$$x_B(t;w) = \Theta(1), \quad s_B(t;w) = \Theta(t^2), \quad x_T(t;w) = \Theta(t), \quad s_T(t;w) = \Theta(t),$$

and the derivatives are bounded, viz.

$$\dot{x}(t;w) = O(1), \quad \dot{s}_B(t;w) = O(t), \quad \dot{s}_T(t;w) = O(1),$$

$$\ddot{x}(t;w) = O(1), \quad \ddot{s}(t;w) = O(1).$$

If $T = \emptyset$, then

$$\dot{x}(t;w) = O(t).$$

It should be noticed that the new parameterization (4) is responsible for the boundedness of the trajectory of derivatives; in the old parameterization (7), the derivatives were shown to be unbounded by Monteiro and Tsuchiya [18].

As a corollary to Theorem 3.1, we get an accurate estimate of the order of the error in $(x(t), s(t))$ with respect to $\mathcal{F}^*$, in terms of t.

Corollary 3.1. *If $T = \emptyset$, then*

$$dist((x(t), s(t)), \mathcal{F}^*) = \Theta(t^2)$$

and if $T \neq \emptyset$ then

$$dist((x(t), s(t)), \mathcal{F}^*) = \Theta(t).$$

Proof:

Since $(x(t), s(t))$ is continuously differentiable for $t > 0$, it follows from the mean value theorem that for any $\bar{t} > 0$, there exist $t_1, t_2, \ldots, t_{2n} \in (0, \bar{t})$ such that

$$x_i(\bar{t}) - x_i(0) = \bar{t}\dot{x}_i(t_i),$$

$$s_i(\bar{t}) = \bar{t}\dot{s}_i(t_{n+i}),$$

for $i = 1, 2, \ldots, n$. From the bounds of Theorem 3.1, it thus follows that

$$\mathrm{dist}((x(t), s(t)), \mathcal{F}^*) = O(t^2) \text{ if } T = \emptyset,$$

$$\operatorname{dist}((x(t), s(t)), \mathcal{F}^*) = O(t) \text{ if } T \neq \emptyset.$$

Since $s_B(t) = \Theta(t^2)$ and $s_T(t) = \Theta(t)$, this implies the error bound of the theorem.

□

4. A second-order predictor-corrector algorithm

The central path is by definition the e-weighted trajectory, i.e. the trajectory $\{(x(t;e), s(t;e)) \mid t > 0\}$ where

$$x_i(t;e)s_i(t;e) = t^2 \text{ for } i = 1, 2, \ldots, n.$$

For a fixed parameter t, the solution $(x(t;e), s(t;e))$ is known as the *t-center*. A neighborhood of the t-center is given as

$$N(\beta) := \{(x, s, t) > 0 \mid \left\|Xs - t^2 e\right\| \leq \beta t^2,\ Qx + Rs = b\},$$

with parameter $\beta \in (0, 1)$. Suppose that in iteration k, a solution

$$(x^k, s^k, t_k) \in N(\frac{1}{4}) \tag{29}$$

is given. We like to construct a solution $(x^{k+1}, s^{k+1}, t_{k+1}) \in N(1/4)$ with $t_{k+1} < t_k$. Let

$$w_i^{(k)} := \frac{1}{t_k}\sqrt{x_i^{(k)} s_i^{(k)}} \text{ for } i = 1, 2, \ldots, n, \tag{30}$$

and remark from (29) that

$$\left\|(W^{(k)})^2 e - e\right\| \leq \frac{1}{4}. \tag{31}$$

Now consider the w^k-weighted trajectory

$$\{(x(t;w^k), s(t;w^k)) \mid t > 0\},$$

for which we construct at t_k a second-order approximation as

$$x^k(t) := x^k + (t - t_k)\dot{x}(t_k; w^k) + \frac{(t - t_k)^2}{2}\ddot{x}(t_k; w^k), \tag{32}$$

$$s^k(t) := s^k + (t - t_k)\dot{s}(t_k; w^k) + \frac{(t - t_k)^2}{2}\ddot{s}(t_k; w^k). \tag{33}$$

We propose to choose $t_{k+1} < t_k$ in such a way that $(x^k(t_{k+1}), s^k(t_{k+1}))$ is still close to the w^k-weighted trajectory. This is called the *predictor* step. The specific choice of t_{k+1} is

$$t_{k+1} = \min_{\bar{t}}\{\bar{t} \mid \left\|X^k(t)s^k(t) - (tW^k)^2 e\right\| \leq \frac{t^2}{4} \text{ for all } t \in [\bar{t}, t_k]\}. \tag{34}$$

It is a simple application of the triangle inequality to verify that

$$(x^k(t_{k+1}), s^k(t_{k+1}), t_{k+1}) \in N(\frac{1}{2}).$$

Based on $(x^k(t_{k+1}), s^k(t_{k+1}), t_{k+1}) \in N(1/2)$, we would like to construct a solution (x^{k+1}, s^{k+1}) such that $(x^{k+1}, s^{k+1}, t_{k+1}) \in N(1/4)$. It is well known [4] that this can be accomplished by a standard *centering* or *corrector* step, i.e. choose (x^{k+1}, s^{k+1}) as the solution (x, s) of the linear system

$$X^k(t_{k+1})s + S^k(t_{k+1})x = t_{k+1}^2 e + X^k(t_{k+1})s^k(t_{k+1}), \tag{35}$$

$$Qx + Rs = b. \tag{36}$$

In this way, we have

$$(x^{k+1}, s^{k+1}, t_{k+1}) \in N(\frac{1}{4}),$$

and (29) holds for $k = 0$ by assumption, and for $k = 1, 2, \ldots$ by induction. The algorithm is summarized as follows:

Algorithm 1. *Input:* $(x^0, s^0, t^0) \in N(1/4)$.

Step 0 Let $k = 0$.

Step 1 Predictor. *Compute* $(x^k(t_{k+1}), s^k(t_{k+1}), t_{k+1}) \in N(1/2)$ *using formulas (30) and (32)–(34).*

Step 2 Corrector. *Solve the linear system (35)–(36), yielding* (x^{k+1}, s^{k+1}).

Step 3 Let $k = k + 1$ *and return to Step 1.*

Remark.

We refer to Kojima et al. [8] and Ye [25] for initializing Algorithm 1 by means of an embedding.

Remark.

The derivatives $(\dot{x}(t_k; w^k), \dot{s}(t_k; w^k))$ and $(\ddot{x}(t_k; w^k), \ddot{s}(t_k; w^k))$, which are used in the predictor step, can be computed by solving the linear system (9)–(11). In each predictor step, the Jacobian has to be factored only once for solving both the first and second order derivatives.

In order to further simplify notations, we write

$$\dot{x}^k := \dot{x}(t_k; w^k), \quad \ddot{x}^k := \ddot{x}(t_k; w^k), \quad \dot{s}^k := \dot{s}(t_k; w^k), \quad \ddot{s}^k := \ddot{s}(t_k; w^k).$$

Notice from (32)-(33) that for all $i = 1, 2, \ldots, n$, there holds

$$\begin{aligned} x_i^{(k)}(t)s_i^{(k)}(t) &= (x_i^{(k)} + (t-t_k)\dot{x}_i^{(k)} + \frac{(t-t_k)^2}{2}\ddot{x}_i^{(k)}) \\ &\quad \times (s_i^{(k)} + (t-t_k)\dot{s}_i^{(k)} + \frac{(t-t_k)^2}{2}\ddot{s}_i^{(k)}) \\ &= t^2(w_i^{(k)})^2 + \frac{(t-t_k)^3}{2}(\dot{x}_i^{(k)}\ddot{s}_i^{(k)} + \dot{s}_i^{(k)}\ddot{x}_i^{(k)}) + \frac{(t-t_k)^4}{4}\ddot{x}_i^{(k)}\ddot{s}_i^{(k)}, \end{aligned} \tag{37}$$

where we used (9) and (10). We shall use (37) in Section 4.1 to establish the polynomiality of Algorithm 1, and in Section 4.2 to demonstrate the superlinear convergence of this algorithm.

4.1. Polynomiality

We will now show that in each iteration of Algorithm 1, the parameter t can sufficiently be reduced. Especially, we will demonstrate its global convergence, i.e. $\lim_{k\to\infty} t_k = 0$. But first, we need a technical lemma to shed some light on the properties of the neighborhood $N(\beta)$.

Lemma 4.1. *Let $y \in \Re^n$. If $y > 0$ then*

1. $\min_{1\le i\le n} y_i \ge (1 - \|y-e\|)e^T y/n$,
2. $\|y\|_1 = e^T y \le n + \sqrt{n}\,\|y-e\|$.

Proof:

1. Let

$$\alpha := \frac{(e^{\mathrm{T}}y)^2}{n\,\|y\|^2}, \qquad u := \frac{n}{e^{\mathrm{T}}y}y,$$

and remark that $0 < \alpha \le 1$. Since

$$0 = y^{\mathrm{T}}\left[\ \frac{e^{\mathrm{T}}y}{\|y\|^2}y - e\ \right] = y^{\mathrm{T}}(\alpha u - e),$$

we have

$$\|y-e\|^2 = \|\alpha u - e\|^2 + \left[\ \frac{e^{\mathrm{T}}y}{\|y\|^2} - 1\ \right]^2 \|y\|^2 \ge \|\alpha u - e\|^2.$$

Hence,

$$\frac{n}{e^{\mathrm{T}}y}\min_{1\le i\le n} y_i = \min_{1\le i\le n} u_i \ge \frac{1}{\alpha}(1 - \|y-e\|).$$

Since $0 < \alpha \le 1$ and $y > 0$, this implies that

$$\min_{1\le i\le n} y_i \ge (1 - \|y-e\|)\frac{e^{\mathrm{T}}y}{n}.$$

2. Using the Cauchy-Schwartz inequality, we have

$$e^{\mathrm{T}}y = n + e^{\mathrm{T}}(y-e) \le n + \sqrt{n}\,\|y-e\|.$$

$\square$

To facilitate our further analysis, we define the primal-dual scaling vector

$$d_i^{(k)} := \sqrt{\frac{x_i^{(k)}}{s_i^{(k)}}} \text{ for } i = 1, 2, \ldots, n.$$

From (30) it easily follows that

$$d^k = (t_k W^{(k)})^{-1} x^{(k)} = t_k (S^{(k)})^{-1} w^k. \tag{38}$$

Lemma 4.2. *Consider Algorithm 1. For $k = 1, 2, \ldots$, there holds*

1. $\left\|(D^{(k)})^{-1}\dot{x}^k\right\|^2 + \left\|D^{(k)}\dot{s}^k\right\|^2 \le 4n + \sqrt{n}$,
2. $\left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2 + \left\|D^{(k)}\ddot{s}^k\right\|^2 \le 4n(4n+\sqrt{n})/(3t_k^2)$.

Proof:

1. Due to monotonicity, we have $(\dot{x}^k)^{\mathrm{T}}\dot{s}^k \ge 0$ and

$$\left\|(D^{(k)})^{-1}\dot{x}^k\right\|^2 + \left\|D^{(k)}\dot{s}^k\right\|^2 \le \left\|(D^{(k)})^{-1}\dot{x}^k + D^{(k)}\dot{s}^k\right\|^2.$$

However, combining (9) and (38) yields

$$\left\|(D^{(k)})^{-1}\dot{x}^k + D^{(k)}\dot{s}^k\right\| = 2\left\|w^k\right\|, \tag{39}$$

where we have from Lemma 4.1 and using (31) that

$$\left\|w^k\right\|^2 = \left\|(W^{(k)})^2 e\right\|_1 \le n + \frac{1}{4}\sqrt{n}. \tag{40}$$

Combining the above relations, we conclude

$$\left\|(D^{(k)})^{-1}\dot{x}^k\right\|^2 + \left\|D^{(k)}\dot{s}^k\right\|^2 \le 4n + \sqrt{n}.$$

2. Similarly, using monotonicity, we have $(\ddot{x}^k)^{\mathrm{T}}\ddot{s}^k \ge 0$ and

$$\begin{aligned} \left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2 + \left\|D^{(k)}\ddot{s}^k\right\|^2 &\le \left\|(D^{(k)})^{-1}\ddot{x}^k + D^{(k)}\ddot{s}^k\right\|^2 \\ &= \frac{4}{t_k^2}\left\|w^k - (W^{(k)})^{-1}\dot{X}^k\dot{s}^k\right\|^2, \end{aligned} \tag{41}$$

where we used (10) and (38). Now define

$$p^k := \frac{1}{2}((D^{(k)})^{-1}\dot{x}^k - D^{(k)}\dot{s}^k), \tag{42}$$

so that, using (9) and (38),

$$(D^{(k)})^{-1}\dot{x}^k = w^k + p^k, \quad D^{(k)}\dot{s}^k = w^k - p^k,$$

and therefore

$$\dot{x}_i^{(k)}\dot{s}_i^{(k)} = (w_i^{(k)})^2 - (p_i^{(k)})^2 \text{ for } i = 1, 2, \ldots, n.$$

Applying this to (41), we obtain

$$\begin{aligned} t_k^2(\left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2 + \left\|D^{(k)}\ddot{s}^k\right\|^2) &\leq 4\left\|w^k - (W^{(k)})^{-1}\dot{X}^k\dot{s}^k\right\|^2 \\ &\leq 4\left\|w^k - (W^{(k)})^{-1}\dot{X}^k\dot{s}^k\right\|_1^2 \\ &= 4\left[\ \textstyle\sum_{i=1}^n (p_i^{(k)})^2/w_i^{(k)}\ \right]^2 \\ &\leq 4\frac{\left\|p^{(k)}\right\|^4}{(\min_{1\leq i\leq n} w_i^{(k)})^2}. \end{aligned} \tag{43}$$

Due to monotonicity, $(\dot{x}^k)^{\mathrm{T}}\dot{s}^k \geq 0$ and it follows from definition (42) that

$$\left\|p^k\right\| \leq \frac{1}{2}\left\|(D^{(k)})^{-1}\dot{x}^k + D^{(k)}\dot{s}^k\right\| = \left\|w^k\right\|, \tag{44}$$

where we used (39). Using (31) and Lemma 4.1, we have

$$(\min_{1\leq i\leq n} w_i^{(k)})^2 \geq \frac{3}{4n}\left\|w^k\right\|^2. \tag{45}$$

Applying (40), (44) and (45) to (43), we conclude that

$$t_k^2(\left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2 + \left\|D^{(k)}\ddot{s}^k\right\|^2) \leq \frac{4n(4n+\sqrt{n})}{3}.$$

□

It will now be shown that Algorithm 1 converges at least linearly with a guaranteed reduction rate $1 - 1/(1 + 3\sqrt{n})$.

Theorem 4.1 (Polynomiality). *For each iteration $k = 0, 1, \ldots$ in Algorithm 1, we have*

$$t_{k+1} \leq (1 - \frac{1}{1+3\sqrt{n}})t_k \leq (1 - \frac{1}{1+3\sqrt{n}})^{k+1}t_0.$$

Hence, Algorithm 1 has globally linear convergence.

Proof:

It follows from (37) that

$$\begin{aligned}\left\|X^{(k)}(t)s^{(k)}(t)-t^2(W^{(k)})^2e\right\| &\leq \frac{|\, t-t_k\,|^3}{2}\left\|\dot{X}^{(k)}\ddot{s}^{(k)}+\dot{S}^{(k)}\ddot{x}^{(k)}\right\| \\ &\quad +\frac{|\, t-t_k\,|^4}{4}\left\|\ddot{X}^{(k)}\ddot{s}^{(k)}\right\|. \qquad (46)\end{aligned}$$

However, there holds

$$\left\|\dot{X}^{(k)}\ddot{s}^{(k)}+\dot{S}^{(k)}\ddot{x}^{(k)}\right\| \leq \left\|\dot{X}^{(k)}\ddot{s}^{(k)}\right\|_1+\left\|\dot{S}^{(k)}\ddot{x}^{(k)}\right\|_1,$$

so that, using the Cauchy-Schwartz inequality,

$$\begin{aligned}\left\|\dot{X}^{(k)}\ddot{s}^{(k)}+\dot{S}^{(k)}\ddot{x}^{(k)}\right\|^2 &\leq (\left\|(D^{(k)})^{-1}\dot{x}^k\right\|^2+\left\|D^{(k)}\dot{s}^k\right\|^2) \\ &\quad \times(\left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2+\left\|D^{(k)}\ddot{s}^k\right\|^2).\end{aligned}$$

By applying Lemma 4.2 to the above relation, we get

$$t_k\left\|\dot{X}^{(k)}\ddot{s}^{(k)}+\dot{S}^{(k)}\ddot{x}^{(k)}\right\| \leq 2\sqrt{\frac{n}{3}}(4n+\sqrt{n}) \leq \frac{10}{\sqrt{3}}n^{3/2}, \qquad (47)$$

where we used the rough inequality $4n+\sqrt{n}\leq 5n$. We also deduce from Lemma 4.2 that

$$\begin{aligned}t_k^2\left\|\ddot{X}^{(k)}\ddot{s}^{(k)}\right\| &\leq \frac{t_k^2}{2}(\left\|(D^{(k)})^{-1}\ddot{x}^k\right\|^2+\left\|D^{(k)}\ddot{s}^k\right\|^2) \\ &\leq \frac{2n(4n+\sqrt{n})}{3} \\ &\leq \frac{10n^2}{3}. \qquad (48)\end{aligned}$$

Combining (47) and (48) with (46), we have for $|\, t-t_k\,|\leq t_k/(1+3\sqrt{n})$ that

$$\begin{aligned}\left\|X^{(k)}(t)s^{(k)}(t)-t^2(W^{(k)})^2e\right\| &\leq \frac{5}{\sqrt{3}}n^{3/2}\frac{|\, t-t_k\,|^3}{t_k}+\frac{5}{6}n^2\frac{|\, t-t_k\,|^4}{t_k^2} \\ &< 0.107t_k^2+0.011t_k^2 \\ &= 0.118t_k^2 < 0.210t^2,\end{aligned}$$

where we used the rough inequality $t\geq 3t_k/4$. From (34), the definition of t_{k+1}, it follows that

$$t_{k+1}<(1-\frac{1}{1+3\sqrt{n}})t_k \text{ for } k=0,1,\ldots,$$

which concludes the proof.

□

4.2. Superlinear convergence

We finally arrive at the main result of the paper.

Theorem 4.2. *The iterates (x^k, s^k) of Algorithm 1 converge to $\mathcal{F}^*$ Q-superlinearly with at least order 1.5. The convergence is even Q-quadratic if $T = \emptyset$.*

Proof:

From (37) we have

$$\begin{aligned}\left\| X^{(k)}(t)s^{(k)}(t) - t^2(W^{(k)})^2 e \right\| \quad \le \quad & \frac{\mid t_k - t \mid^3}{2} \left\| \dot{X}^{(k)}\ddot{s}^{(k)} + \dot{S}^{(k)}\ddot{x}^{(k)} \right\| \\ & + \frac{(t_k - t)^4}{4} \left\| \ddot{X}^{(k)}\ddot{s}^{(k)} \right\|. \end{aligned} \tag{49}$$

From Theorem 3.1, we know that there exists a constant γ, independent of k, such that

$$2 \left\| \dot{X}^{(k)}\ddot{s}^{(k)} + \dot{S}^{(k)}\ddot{x}^{(k)} \right\| + \left\| \ddot{X}^{(k)}\ddot{s}^{(k)} \right\| \le \gamma \text{ for } k = 0, 1, \ldots. \tag{50}$$

It follows from (49)-(50) that for $0 \le t \le t_k \le 1$ we have

$$\left\| X^{(k)}(t)s^{(k)}(t) - t^2(W^{(k)})^2 e \right\| \le \frac{\gamma t_k^3}{4}$$

Therefore, we obtain from definition (34) that

$$t_{k+1} \le \sqrt{\gamma} t_k^{1.5}.$$

This shows that t_k is order 1.5 Q-superlinearly convergent to 0. Using Corollary 3.1, it follows that the iterates (x^k, s^k) converge to $\mathcal{F}^*$ superlinearly with the same Q-order.

Now consider $T = \emptyset$. Then Theorem 3.1 implies that

$$\frac{2}{t_k} \left\| \dot{X}^{(k)}\ddot{s}^{(k)} + \dot{S}^{(k)}\ddot{x}^{(k)} \right\| + \left\| \ddot{X}^{(k)}\ddot{s}^{(k)} \right\| \le \gamma' \text{ for } k = 0, 1, \ldots,$$

for some constant γ'. Combining this with (49), it follows that

$$\left\| X^{(k)}(t)s^{(k)}(t) - t^2(W^{(k)})^2 e \right\| \le \frac{\gamma' t_k^4}{4},$$

for all $0 < t \le t_k \le 1$. Hence,

$$t_{k+1} \le \sqrt{\gamma'} t_k^2,$$

which means that t_k is Q-quadraticly convergent to zero, and consequently (x^k, s^k) converges Q-quadraticly to $\mathcal{F}^*$ due to Corollary 3.1.

$\square$

5. Subjects of future research

A second–order predictor–corrector algorithm has been discussed. It is a natural idea to extend the algorithm to an order $r > 2$ predictor–corrector algorithm, and to investigate whether orders higher than 2 can improve the order of superlinear convergence. Another interesting question is whether we can use a single matrix factorization per main iteration, rather than computing two factorizations: one for the predictor and one for the corrector. Finally, it would be most interesting to see whether some of our ideas can be extended to more general classes of complementarity problems.

Acknowledgement. I am highly indebted to my supervisor Shuzhong Zhang, who gave many useful comments.

References

[1] J.F. Bonnans and C.C. Gonzaga. Fast convergence of the simplified largest step path following algorithm. Technical Report 2433, INRIA, France, 1994.

[2] J.F. Bonnans and C.C. Gonzaga. Convergence of interior point algorithms for the linear complementarity problem. *Mathematics of Operations Research*, 21:1–25, 1996.

[3] A.S. El-Bakry, R.A. Tapia, and Y. Zhang. On the convergence rate of Newton interior point methods in the absence of strict complementarity. Technical report, Rice University, Houstan, Texas, USA, 1993.

[4] C.C. Gonzaga. Interior path following algorithms. In J.R. Birge and K.G. Murty, editors, *Mathematical Programming, State of the Art 1994*, pages 93–101, Ann Arbor, 1994. University of Michigan.

[5] O. Güler. Generalized linear complementarity problems and interior point algorithms for their solutions. Technical report, Faculty of Technical Mathematics and Computer Science, TU Delft, The Netherlands, 1992. Revised 1993.

[6] O. Güler and Y. Ye. Convergence behavior of interior-point algorithms. *Mathematical Programming*, 60:215–228, 1993.

[7] J. Ji, F.A. Potra, and S. Huang. A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence. Technical Report 18, Department of Mathematics, University of Iowa, Iowa, 1991.

[8] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*. Springer-Verlag, Berlin, 1991.

[9] M. Kojima, S. Mizuno, and T. Noma. Limiting behavior of trajectories by a continuation method for monotone complementarity problems. *Mathematics of Operations Research*, 15:662–675, 1990.

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

[11] Z.-Q. Luo and J.-S. Pang. Error bounds for analytic systems and their applications. *Mathematical Programming*, 67:1–28, 1994.

[12] O.L. Mangasarian. Simple computable bounds for solutions of linear complementarity problems and linear programs. *Mathematical Programming Study*, 25:1–12, 1985.

[13] O.L. Mangasarian and T.-H. Shiau. Error bounds for monotone linear complementarity problems. *Mathematical Programming*, 36:81–89, 1986.

[14] K. McShane. Superlinearly convergent $O(\sqrt{n}L)$-iteration interior-point algorithms for LP and the monotone LCP. *SIAM Journal on Optimization*, 4:247–261, 1994.

[15] S. Mehrotra. On the implementation of a primal–dual interior point method. *SIAM Journal on Optimization*, 2:575–601, 1992.

[16] S. Mizuno. A superlinearly convergent infeasible-interior-point algorithm for geometical LCPs without a strictly complementarity condition. *Mathematics of Operations Research*, 2:382–400, 1996.

[17] S. Mizuno, M.J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18:964–981, 1993.

[18] R.D.C. Monteiro and T. Tsuchiya. Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem. Technical Report 92-28, School of Industrial and System Engineering, Georgia Tech, Atlanta, Georgia, USA, 1992. To appear in *Mathematics of Operations Research.*

[19] R.D.C. Monteiro and S. Wright. Local convergence of interior–point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3:131–155, 1993.

[20] K.G. Murty. *Linear complementarity, linear and nonlinear programming.* Heldermann Verlag, Berlin, 1988.

[21] F.A. Potra and R. Sheng. A path following method for LCP with superlinearly convergent iteration sequence. Technical report, Department of Mathematics, University of IOWA, Iowa City, IA, USA, 1995.

[22] F.A. Potra and R. Sheng. A superlinearly convergent infeasible-interior-point algorithm for degenerate LCP. Technical report, Department of Mathematics, University of IOWA, Iowa City, IA, USA, 1995.

[23] R. Sznajder and M.S. Gowda. Generalization of P_0– and P–properties; extended vertical and horizontal LCPs. *Linear Algebra and its Applications*, 223/224:716–720, 1995.

[24] S. Wright and Y. Zhang. A superquadratic infeasible-interior-point method for linear complementarity problems. *Mathematical Programming*, 73:269–289, 1996.

[25] Y. Ye. On homogeneous and self-dual algorithm for LCP. Technical report, Department of Management Sciences, The University of Iowa, Iowa City, Iowa 52242, USA, 1994. To appear in *Mathematical Programming.*

[26] Y. Ye and K. Anstreicher. On quadratic and O($\sqrt{n}L$) convergence of a predictor-corrector algorithm for LCP. *Mathematical Programming*, 62:537–551, 1993.

[27] Y. Zhang and D. Zhang. On polynomiality of the Mehrotra–type predictor–corrector interior–point algorithms. *Mathematical Programming*, 68:303–318, 1995.