Control and Cybernetics

vol. **32** (2003) No. 3

High-order long-step methods for solving semidefinite linear complementarity problems

by

M. Preiß and J. Stoer

Institut für Angewandte Mathematik und Statistik,
Universität Würzburg, Germany

Abstract: The authors studied in Preiß and Stoer (2003) the analyticity properties of infeasible-interior-point paths encountered in the context of semidefinite linear complementarity problems. It will be shown that these results allow for the design of infeasible-interior-point methods of long-step type with an arbitrarily high order of local convergence for solving such problems.

Keywords: semidefinite linear complementarity problems, infeasible-interior-point methods, long-step methods.

1. Introduction

In this paper we consider semidefinite linear complementarity problems of the form

$$(SDLCP) \qquad \begin{aligned} P(X)+Q(Y) &= q,\\ XY &= 0,\\ X,Y &\succeq 0. \end{aligned}$$

Here, X, Y range in the space $\mathcal{S}^n := \{A \in \mathbb{R}^{n\times n} \mid A = A^T\}$ of real symmetric $n\times n$-matrices equipped with the half-order $\succeq$ of positive semidefinite matrices X, $X \succeq 0$, the scalar product $\langle X,Y\rangle := \operatorname{tr} XY$, and the associated norm $\|X\|_F := (\langle X,X\rangle)^{1/2}$, the familiar Frobenius norm. P, Q: $\mathcal{S}^n \to \mathbb{R}^{\bar n}$ are linear operators mapping $\mathcal{S}^n$ to the space $\mathbb{R}^{\bar n}$, where $\bar n := n(n+1)/2$ is the dimension of $\mathcal{S}^n$, and $q \in \mathbb{R}^{\bar n}$. Hence P and Q have the form $P(X) = (\langle P_1,X\rangle,\ldots,\langle P_{\bar n},X\rangle)^T$ resp. $Q(Y) = (\langle Q_1,Y\rangle,\ldots,\langle Q_{\bar n},Y\rangle)^T$ with P_i, $Q_i \in \mathcal{S}^n$ for all $i = 1,\ldots,\bar n$. By $\mathcal{S}^n_{++}$ (resp. $\mathcal{S}^n_{+}$) we denote the set of positive (semi)definite matrices in $\mathcal{S}^n$.

The null space of the operator pair $[P,Q]$ is defined by

$$\mathcal{N}([P,Q]) := \{(X,Y) \in \mathcal{S}^n \times \mathcal{S}^n \mid P(X)+Q(Y) = 0\},$$

and its range space by

$$\mathcal{R}([P,Q]) := \{P(X) + Q(Y) \mid X,\, Y \in \mathcal{S}^n\}.$$

We call $\mathcal{N}([P,Q])$ resp. $(SDLCP)$ *monotone*, if

$$\langle X, Y\rangle \geq 0 \quad \text{for all } (X,Y) \in \mathcal{N}([P,Q]). \tag{1}$$

The set of all solutions of $(SDLCP)$ is denoted by $\mathcal{F}^\star$, and the set $\mathcal{F}$ of *feasible points* of $(SDLCP)$ is defined by

$$\mathcal{F} := \{(X,Y) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+ \mid P(X) + Q(Y) = q\}.$$

A feasible point (X,Y) is called *strictly feasible* if $X \succ 0$, $Y \succ 0$.

A major problem in solving a general $(SDLCP)$ is that the nonlinear map

$$\varphi : \mathcal{S}^n \times \mathcal{S}^n \to \mathrm{I\!R}^{\bar{n}} \times \mathrm{I\!R}^{n\times n}, \quad \varphi(X,Y) := \begin{bmatrix} P(X) + Q(Y) - q \\ XY \end{bmatrix}$$

is a map between spaces of different dimension, since in general the product of two symmetric matrices X and Y is not symmetric. This prevents the direct use of Newton's method to find the approximate solutions of $(SDLCP)$. For this reason, many proposals have been made (see e.g. Todd, 1999) to symmetrize the problem. We will use one of the first proposals due to Alizadeh, Haeberly and Overton (1998) and replace XY by $(XY + YX)/2$. Then, $(SDLCP)$ is equivalent to

$$\begin{aligned} P(X) + Q(Y) &= q, \\ \frac{1}{2}(XY + YX) &= 0, \\ X, Y &\succeq 0. \end{aligned}$$

Interior-point methods for solving strictly feasible SDLCP's are connected with the path of solutions $Z = (X,Y)(r,M)$, $r > 0$, of the system

$$\begin{aligned} P(X) + Q(Y) &= q, \\ \frac{1}{2}(XY + YX) &= rM, \\ X, S &\succ 0. \end{aligned}$$

Here, $M \succ 0$ is a positive definite weight matrix. In particular, interior-point methods try to follow the *central path* $Z(r,I)$, $r \downarrow 0$, which belongs to the weight matrix $M = I$. A drawback is that these paths do not exist if $(SDLCP)$ has no strictly feasible points, even if $\mathcal{F}^\star \neq \emptyset$.

If $\mathcal{F}^\star \neq \emptyset$, but strictly feasible solutions do not exist or are not known, one chooses an arbitrary $r_0 > 0$, matrices X_0, $Y_0 \succ 0$ (e.g. $X_0 = Y_0 = I$) with $X_0Y_0 + Y_0X_0 \succ 0$ and $\bar{q} := (P(X_0) + Q(Y_0) - q)/r_0$ so that $\bar{q} \in \mathcal{M}(r_0)$, where

$$\mathcal{M}(r_0) := \{(P(X) + Q(Y) - q)/r_0 \mid X, Y \succ 0,\; XY + YX \succ 0\}. \tag{2}$$

Infeasible interior-point methods are then connected with the weighted infeasible interior-point path of solutions $Z = (X,Y)(r,M,\bar{q})$, $r > 0$, $M \succ 0$, of the system

$$(SDLCP)_{r,M,\bar{q}} \qquad \begin{array}{rcl} P(X) + Q(Y) &=& q + r\bar{q}, \\ \frac{1}{2}(XY + YX) &=& rM, \\ X, Y &\succ& 0. \end{array}$$

The definitions of $\bar{q}$ and $\mathcal{M}(r_0)$ ensure that this system is at least solvable for $r = r_0$, $M_0 := (X_0Y_0 + Y_0X_0)/r_0$. If these paths $Z(r,M,\bar{q})$ and their limits $Z^\star := \lim_{r\downarrow 0}(X,Y)(r,M,\bar{q})$ exist, then $Z^\star \in \mathcal{F}^\star$ is a solution to $(SDLCP)$.

In Preiß and Stoer (2003), we studied the theoretical properties of the path function $Z(.)$ under the following rather general assumption:

Assumption 1.1

(a) $\mathcal{R}([P,Q]) = \mathbb{R}^{\bar{n}}$,
(b) $(SDLCP)$ is monotone (see (1)),
(c) $(SDLCP)$ is solvable, $\mathcal{F}^* \neq \emptyset$.

Note that it is not required that $(SDLCP)$ has only one solution.

For the results to follow we have to sharpen Assumption 1.1 slightly by requiring the existence of a strictly complementary solution of $(SDLCP)$:

Assumption 1.2

(a) $\mathcal{R}([P,Q]) := \{P(X) + Q(Y) : X, Y \in \mathcal{S}^n\} = \mathbb{R}^{\bar{n}}$,
(b) $(SDLCP)$ is monotone (see (1)),
(c) $(SDLCP)$ has at least one solution $(X^\star, Y^\star) \in \mathcal{F}^*$ that is strictly complementary, that is, a solution with $X^\star + Y^\star \succ 0$.

Note that this assumption still allows the degenerate problems with more than one solution.

We use the following three results of Preiß and Stoer (2003):

Lemma 1.1 *Assume $\mathcal{R}([P,Q]) = \mathbb{R}^{\bar{n}}$, that $(SDLCP)$ is monotone, and let X, $Y \in \mathcal{S}^n_{++}$ be positive definite matrices satisfying $XY + YX \succ 0$. Then the linear map $\Psi: \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}^{\bar{n}} \times \mathcal{S}^n$ defined by*

$$\Psi(A,B) := \begin{bmatrix} P(A) + Q(B) \\ H_Y(A) + H_X(B) \end{bmatrix},$$

is a bijective linear map. Here, H_X (and likewise H_Y) is the Liapunov map belonging to X: $H_X(S) := XS + XS$ for $S \in \mathcal{S}^n$.

Theorem 1.1 *Let Assumption 1.2 be satisfied, let $r_0 > 0$ and $\bar{q} \in \mathcal{M}(r_0)$. Then for all $0 < r \le r_0$, $M \succ 0$ the value $Z = (X,Y)(r,M,\bar{q})$ of the path function $Z(.)$ is uniquely defined as the solution of $(SDLCP)_{r,M,\bar{q}}$. $Z = (X,Y)(r,M,\bar{q})$ is an analytic function of $(r,M,\bar{q})$ for all $(r,M,\bar{q}) \in (0,r_0] \times \mathcal{S}^n_{++} \times \mathcal{M}(r_0)$.*

Remark 1.1

This theorem is still true if condition (c) of Assumption 1.2 is replaced by the weaker condition that $(SDLCP)$ has a solution. The result itself is due to Monteiro and Pang (1998).

Theorem 1.2 *Let Assumption* 1.2 *be satisfied, let* $r_0 > 0$ *and* $\bar{q} \in \mathcal{M}(r_0)$*. Then the path function* $Z(r, M, \bar{q})$ *defined for* $(r, M, \bar{q}) \in (0, r_0] \times \mathcal{S}^n_{++} \times \mathcal{M}(r_0)$ *can be extended to a function that is an analytic function of* $(r, M, \bar{q})$ *on a larger set* $[-\varepsilon, r_0] \times \mathcal{S}^n_{++} \times \mathcal{M}(r_0)$*, where* $\varepsilon > 0$ *is some positive number.*

In what follows, we will write also $Z = (X, Y)(r, M, \bar{q})$ for the extension of the path function.

A special case of this theorem, namely the analytic extendability of the central path belonging to the standard semidefinite linear program, was shown by Halická (2002).

Similar (and even stronger) results were proved for the standard monotone linear complementarity problem (LCP) in Stoer and Wechs (1999). This led to the construction of superlinearly convergent infeasible-interior-point methods for solving (LCP) (see Stoer, Wechs and Mizuno, 1998), also of long-step type (see Stoer, 1999, 2001). A description of the various types of interior point methods for solving linear programs and standard linear complementarity problems (LCP) is given in Wright (1997).

We will show in this paper that one can use also the results of Theorems 1.1 and 1.2 for the design of infeasible-interior-point methods of long-step type for solving semidefinite linear complementarity problems $(SDLCP)$. These methods will locally converge with an arbitrarily high order of convergence, provided $(SDLCP)$ satisfies Assumption 1.2.

2. A long-step method

In this section, we generalize the high-order long-step algorithm of Stoer (1999) for solving (LCP) to a method for solving $(SDLCP)$. We suppose that Assumption 1.2 is satisfied, so that according to Theorem 1.2 the weighted infeasible interior-point path $(X, Y)(r, M, \bar{q})$ is an analytic function of all its arguments for all $(r, M, \bar{q}) \in [-\varepsilon, r_0] \times \mathcal{S}^n_{++} \times \mathcal{M}(r_0)$ for some $\varepsilon > 0$. We will construct an infeasible interior-point method that generates a sequence of points $(X_k, Y_k) = (X, Y)(r_k, M_k, \bar{q}_k)$, $k = 1, 2, \ldots$, belonging to parameters $(r_k, M_k, \bar{q}_k)$ with $r_k > 0$, which lie for all k in a certain compact subset $\mathcal{C} \subset [0, r_0] \times \mathcal{S}^n \times \mathcal{M}(r_0)$ of the domain of analyticity $[-\varepsilon, r_0] \times \mathcal{S}^n_{++} \times \mathcal{M}(r_0)$. This will have the following consequences:

1. The $(r_k, M_k, \bar{q_k})$ are bounded. Hence, if $r_k \downarrow 0$, every cluster point of the $(r_k, M_k, \bar{q_k})$ will belong to $\mathcal{C}$. Then also the sequence (X_k, Y_k), $k \geq 0$, will have cluster points, and each of its cluster points will be a solution of $(SDLCP)$.

2. Since the function $(X, Y)(r, M, \bar{q})$ is analytic on $\mathcal{C}$, there are uniform bounds
$$\sup_{(r,M,\bar{q})\in\mathcal{C}} \|D^j(X, Y)(r, M, \bar{q})\|_F \le c_j$$
for all derivatives of any order $j \ge 0$ with respect to any of the variables $r, M, \bar{q}$. In particular, we will have
$$\sup_k \|D^j(X, Y)(r_k, M_k, \bar{q}_k)\|_F \le c_j$$
This will be crucial for the investigation of the rate of convergence of the parameters r_k.

We will describe an infeasible-interior-point method of long-step type belonging to a particular compact set $\mathcal{C}$, which is explained next. First, we will only consider normalized weight matrices $M \succ 0$ by requiring $\operatorname{tr} M = n$. Then $(XY + YX)/2 = rM$ implies that the parameter r is given by

$$r = \mu := \langle X, Y\rangle/n.$$

Our long-step method uses this normalization of M and identification of r with μ. It is based on a compact set $\mathcal{C}$ of the form

$$\mathcal{C} := [0, \mu_0] \times T_{\underline{\gamma}} \times Q_{\underline{\theta}}, \qquad 0 < \underline{\gamma} < 1, \quad 0 < \underline{\theta} < 1,$$

where

$$\begin{aligned} T_\gamma &:= \{M \in \mathcal{S}^n_{++} \mid \operatorname{tr} M = n,\ M \succeq \gamma I\}, \quad 0 < \gamma < 1, \\ Q_{\underline{\theta}} &:= \{\alpha\bar{q}_0 \mid \underline{\theta} \le \alpha \le \underline{\theta}^{-1}\}, \qquad 0 < \underline{\theta} < 1. \end{aligned}$$

The parameter $\underline{\theta}$ is determined so that $0 < \underline{\theta} < 1$ and

$$Q_{\underline{\theta}} \subset \mathcal{M}(r_0) = \mathcal{M}(\mu_0).$$

This is possible, since $\mathcal{M}(\mu_0)$ is open. Then some constant $\underline{\gamma}$ with $0 < \underline{\gamma} < 1$ is chosen such that $e^{-(1-\underline{\gamma})} \ge \underline{\theta}$. Next, with an arbitrary γ_0 with $\underline{\gamma} < \gamma_0 < 1$, further constants γ_k satisfying

$$\underline{\gamma} < \cdots < \gamma_{k+1} < \gamma_k < \cdots < \gamma_0$$

are computed recursively by $\delta_0 := (\gamma_0 - \underline{\gamma})/2$ and $\gamma_{k+1} = \gamma_k - \delta_k$, $\delta_{k+1} := \delta_k/2$ for $k \ge 0$.

This construction allows the following estimates for the infinite products

$$\begin{aligned} \textstyle\prod_{j=0}^\infty (1 + \delta_j) = \prod_{j=0}^\infty (1 + 2^{-j}\delta_0) \le e^{2\delta_0} < e^{1-\underline{\gamma}} \le \underline{\theta}^{-1}, \\ \textstyle\prod_{j=0}^\infty (1 - \delta_j) = \prod_{j=0}^\infty (1 - 2^{-j}\delta_0) \ge e^{-2\delta_0} > e^{-(1-\underline{\gamma})} \ge \underline{\theta}, \end{aligned} \tag{3}$$

which will be used later on. Then, $\mathcal{C}_{\gamma_k} := [0, \mu_0] \times T_{\gamma_k} \times Q_{\underline{\theta}}$, $k \ge 0$, is a sequence of compact sets with

$$\mathcal{C}_{\gamma_0} \subset \cdots \subset \mathcal{C}_{\gamma_k} \subset \mathcal{C}_{\gamma_{k+1}} \subset \cdots \subset \mathcal{C}_{\underline{\gamma}} = \mathcal{C}.$$

The algorithm will be such that its iterates $(X_k, Y_k) = (X, Y)(\mu_k, M_k, \bar{q}_k)$ belong to parameters $(\mu_k, M_k, \bar{q}_k) \in \mathcal{C}_{\gamma_k}$, $k \geq 0$, which means that (X_k, Y_k) and the associated quantities μ_k, M_k, $\bar{q}_k$ satisfy

$$\begin{aligned} P(X_k) + Q(Y_k) &= q + \mu_k \bar{q}_k, \quad X_k \succ 0,\ Y_k \succ 0, \\ \tfrac{1}{2}(X_k Y_k + Y_k X_k) &= \mu_k M_k, \\ \mu_k = \langle X_k, Y_k \rangle / n, &\quad M_k \succeq \gamma_k I, \\ \bar{q}_k = \alpha_k \bar{q}_0, &\quad \text{for some } \alpha_k \text{ with } \underline{\theta} \leq \alpha_k \leq \underline{\theta}^{-1}. \end{aligned} \tag{4}$$

The second and third of these relations are usually expressed by

$$(X_k, Y_k) \in \mathcal{N}_{-\infty}(\gamma_k),$$

where for $0 < \gamma < 1$, $\mathcal{N}_{-\infty}(\gamma)$ is the set

$$\mathcal{N}_{-\infty}(\gamma) := \left\{ (X, Y) \mid X,\, Y \succ 0,\ \tfrac{XY + YX}{2} \succeq \gamma \mu I,\ \mu := \tfrac{\langle X, Y \rangle}{n} \right\} \tag{5}$$

defining a large neighborhood of the central path (for $\gamma = 1$ it would reduce to the central path).

We now define the algorithm with the properties (4) for all $k \geq 0$. For the convenience of presentation we will assume $\mu_0 \leq 1$. As starting values we choose any μ_0 with $0 < \mu_0 \leq 1$ and then

$$X_0 = Y_0 := I, \quad \mu_0 = 1, \quad M_0 = I, \quad \bar{q}_0 := P(X_0) + Q(Y_0) - q,$$

which obviously satisfy (4) for $k = 0$.

For the description of the iteration step

$$(X_k, Y_k) \longrightarrow (X_{k+1}, Y_{k+1})$$

we assume inductively that (4) is satisfied for k and use the following abbreviations

$$Z := (X, Y) := (X_k, Y_k), \quad \mu := \frac{\langle X, Y \rangle}{n}, \quad M := \frac{1}{2}(XY + YX)/\mu,$$

$$\bar{q} := \bar{q}_k = (P(X) + Q(Y) - q)/\mu,$$

$$\gamma := \gamma_k, \quad \delta := \delta_k, \quad \gamma_+ := \gamma_{k+1} = \gamma - \delta.$$

Then by induction hypothesis $(\mu, M, \bar{q}) \in \mathcal{C}_\gamma$. Define for $0 \leq \nu \leq \mu$ the function $\bar{Z}(\nu) := (\bar{X}, \bar{Y})(\nu)$ as the solution of the nonlinear system

$$\begin{aligned} P(\bar{X}) + Q(\bar{Y}) &= q + \nu \bar{q}, \quad \bar{X} \succ 0,\ \bar{Y} \succ 0, \\ \tfrac{1}{2}(\bar{X}\bar{Y} + \bar{Y}\bar{X}) &= \nu M(\nu), \end{aligned} \tag{6}$$

where

$$M(\nu) := M + (\mu - \nu)(I - M)$$

is a linear function of ν. Because of $M(\nu) \succ 0$ for $0 \le \nu \le \mu(\le 1)$ and $\bar{Z}(\mu) = (X, Y)$, $M(\mu) = M$, the solution $\bar{Z}(\nu)$ is well-defined by Theorem 1.2 and an analytic function of ν for all $\nu \in [0, \mu]$ by Theorem 1.2. We note the following properties for all μ, ν with $0 \le \nu \le \mu \le \mu_0 \le 1$:

Property 2.1

(1) $M(\nu) = (1 - (\mu - \nu))M + (\mu - \nu)I \in \text{conv}\{I, M\}$, so that $M(\nu) \succ 0$.
(2) $M(\nu) - \gamma_+ I \succeq (1 - (\mu - \nu))\gamma I + (\mu - \nu)I - (\gamma - \delta)I$
$= (\mu - \nu)(1 - \gamma)I + \delta I$
$\succeq \delta I.$
(3) $\operatorname{tr}(M(\nu)) = n,\ \bar{\mu}(\nu) := \langle \bar{X}(\nu), \bar{Y}(\nu)\rangle / n = \nu \operatorname{tr}(M(\nu))/n = \nu.$

We now approximate the path $\bar{Z}(\nu)$ by a matrix-valued Taylor polynomial $\hat{Z}(\nu)$ of degree $p \ge 1$ in ν around $\nu = \mu$:

$$\hat{Z}(\nu) = (\hat{X}(\nu), \hat{Y}(\nu)) := Z + (\nu - \mu)Z_1 + \cdots + (\nu - \mu)^p Z_p\,,$$

where the coefficient matrices Z_j are defined as the scaled derivatives

$$Z_j = (X^{[j]}, Y^{[j]}) := \frac{1}{j!}\frac{\partial^j}{\partial \nu^j}\,(\bar{X}, \bar{Y})(\nu)\Big|_{\nu=\mu}\,. \tag{7}$$

The coefficients Z_j can be obtained by differentiating (6) repeatedly with respect to ν at $\nu = \mu$, e.g.,

$$\begin{array}{rcl} P(X^{[1]}) + Q(Y^{[1]}) & = & \bar{q} \\ \frac{1}{2}(H_Y(X^{[1]}) + H_X(Y^{[1]})) & = & M - \mu(I - M), \\ P(X^{[2]}) + Q(Y^{[2]}) & = & 0 \\ H_Y(X^{[2]}) + H_X(Y^{[2]}) & = & -H_{X_1}(Y_1) - 2(I - M)\,. \end{array} \tag{8}$$

Note that these equations are uniquely solvable by Lemma 1.1; since the same linear equation in $\mathcal{S}^n$ with different right hand sides has to be solved, only one linear operator in $\mathcal{S}^n$ needs to be inverted.

Because of Property 2.1, (1) – (3) and Theorem 1.2, the function

$$\bar{Z}(\nu) = (X, Y)(\nu, M(\nu), \bar{q}), \quad 0 \le \nu \le \mu,$$

belongs to parameters $(\nu, M(\nu), \bar{q}) \in \mathcal{C}_{\gamma_+} \subset \mathcal{C}$, so that by Assumption 1.2 all derivatives

$$\left\| \frac{\partial^j}{\partial \nu^j} \bar{Z}(\nu) \right\|_F \le d_j \tag{9}$$

are bounded for all μ, ν with $0 \le \nu \le \mu$ by constants d_j that depend only on $\mathcal{C}$.

In order to determine the next iterate (X_+, Y_+) and the associated μ_+, M_+, and $\bar{q}_+$, we first determine the number ν_+ by the curvilinear search

$$\begin{aligned}\nu_+ &:= \inf\{\nu > 0 \mid \hat{Z}(\nu) \in \mathcal{N}_{-\infty}(\gamma_+),\ |\hat{\mu}(\nu) - \nu| \le \delta\nu,\ \hat{\mu}(\nu) \le \mu\} \\ &= \inf\{\nu > 0 \mid \hat{Z}(\nu) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}, \\ &\qquad (\hat{X}(\nu)\hat{Y}(\nu) + \hat{Y}(\nu)\hat{X}(\nu))/2 \succeq \gamma_+\hat{\mu}(\nu)I, \\ &\qquad |\hat{\mu}(\nu) - \nu| \le \delta\nu,\ \hat{\mu}(\nu) \le \mu\},\end{aligned} \tag{10}$$

where, of course, $\hat{\mu}(\nu) := \langle \hat{X}(\nu), \hat{Y}(\nu)\rangle / n$.

Then, the new iterates are defined by

$$(X_{k+1}, Y_{k+1}) = (X_+, Y_+) := (\hat{X}(\nu_+), \hat{Y}(\nu_+)),$$

$$\mu_{k+1} = \mu_+ := \hat{\mu}(\nu_+), \quad M_{k+1} = M_+ := (X_+Y_+ + Y_+X_+)/(2\mu_+),$$

$$\bar{q}_{k+1} = \bar{q}_+ := (P(X_+) + Q(Y_+) - q)/\mu_+.$$

(These definitions assume $\mu_+ > 0$: if $\mu_+ = 0$, we are finished because (X_+, Y_+) then solves $(SDCLP)$.)

We show next that the new iterates again satisfy (4), $(\mu_+, M_+, \bar{q}_+) \in \mathcal{C}_{\gamma_+} \subset \mathcal{C}$. First, note that ν_+ is well-defined, since $\hat{\mu}(\mu) = \mu$. Also, because of $0 < \delta < \delta_0 < 1/2$ the curvilinear search implies

$$\nu_+(1 - \delta) \le \hat{\mu}(\nu_+) \le \nu_+(1 + \delta)\,, \tag{11}$$

so that $\hat{\mu}(\nu_+) \ge 0$. Moreover, by (7) and (8)

$$P(X_+) + Q(Y_+) = q + \nu_+\bar{q}\,.$$

As noted above, we may assume $\mu_+ > 0$. Then

$$P(X_+) + Q(Y_+) = q + \mu_+\left(\frac{\nu_+}{\mu_+}\bar{q}\right).$$

It follows

$$\bar{q}_+ = \frac{\nu_+}{\mu_+}\bar{q} \quad \text{and} \quad \frac{1}{1+\delta} \le \frac{\nu_+}{\mu_+} \le \frac{1}{1-\delta},$$

so that by induction

$$\bar{q}_+ = \left(\prod_{l=1}^{k+1} \frac{\nu_l}{\mu_l}\right)\bar{q}^{(0)}.$$

Then (3) gives

$$\underline{\theta} < \prod_{l=0}^{\infty}(1 + \delta_l)^{-1} \le \prod_{l=1}^{k+1}\frac{\nu_l}{\mu_l} \le \prod_{l=0}^{\infty}(1 - \delta_l)^{-1} < \underline{\theta}^{-1}\,,$$

that is: $\bar{q}_+ \in Q_{\underline{\theta}}$.

Since $\mu_+ \in [0, \mu] \subset [0, \mu_0]$, $M_+ \in T_{\gamma_+} \subset T_{\underline{\gamma}}$, we get $(\mu_+, M_+, \bar{q}_+) \in \mathcal{C}_{\gamma_+} \subset \mathcal{C}$, that is, (4) is satisfied for the index $k+1$. Hence, by induction, (4) holds for all $k \geq 0$.

We now study the local convergence behavior of the algorithm under Assumption 1.2, i.e., its behavior for small $\mu_0 > 0$. Consider the k-th iteration, where we are given $(\mu, M, \bar{q}) := (\mu_k, M_k, \bar{q}_k) \in \mathcal{C}_{\gamma_k} \subset \mathcal{C}$. Then $(\nu, M(\nu), \bar{q}) \in \mathcal{C}$ for all $0 \leq \nu \leq \mu$. Then, by Taylor's formula,

$$\begin{aligned} \bar{X}(\nu) &= \hat{X}(\nu) + \frac{1}{(p+1)!}(D_\nu^{p+1}\bar{x}_{st}(\xi_{st}))(\nu-\mu)^{p+1},\ \xi_{st} \in (\nu, \mu), \\ \bar{Y}(\nu) &= \hat{Y}(\nu) + \frac{1}{(p+1)!}(D_\nu^{p+1}\bar{y}_{st}(\theta_{st}))(\nu-\mu)^{p+1},\ \theta_{st} \in (\nu, \mu). \end{aligned}$$

Because of (9) we may write

$$\hat{X}(\nu) = \bar{X}(\nu) + O(|\nu-\mu|^{p+1}), \quad \hat{Y}(\nu) = \bar{Y}(\nu) + O(|\nu-\mu|^{p+1}),$$

where the O-terms represent symmetric matrices in ν (because of the symmetry of $\hat{X}(\nu)$, $\bar{Y}(\nu)$) satisfying a bound

$$\left\|O(|\nu-\mu|^{p+1})\right\| \leq c_1|\nu-\mu|^{p+1} \quad \text{for all } 0 \leq \nu \leq \mu,$$

with a constant c_1 that depends only on $\mathcal{C}$ and on p. (It was remarked by M. Halická (private communication) that the analyticity of $\bar{Z}(\nu)$ at $\nu = 0$ even implies that one may assume that the constant c_1 only depends on $\mathcal{C}$ and not on p.) Hence

$$\begin{aligned} \frac{1}{2}(\hat{X}(\nu)\hat{Y}(\nu) + \hat{Y}(\nu)\hat{X}(\nu)) &= \frac{1}{2}(\bar{X}(\nu)\bar{Y}(\nu) + \bar{Y}(\nu)\bar{X}(\nu)) \\ &\quad + O(|\nu-\mu|^{p+1}) \\ &= \nu M(\nu) + O(|\nu-\mu|^{p+1}), \\ \hat{\mu}(\nu) &= \frac{\langle \hat{X}(\nu), \hat{Y}(\nu) \rangle}{n} = \bar{\mu}(\nu) + O(|\nu-\mu|^{p+1}) \\ &= \nu + O(|\nu-\mu|^{p+1}) = \nu + O(\mu^{p+1}). \end{aligned}$$

Therefore, there is a constant $c = c(\mathcal{C})$ only depending on $\mathcal{C}$ so that for all $0 \leq \nu \leq \mu$

$$\begin{aligned} \frac{1}{2}(\hat{X}(\nu)\hat{Y}(\nu) + \hat{Y}(\nu)\hat{X}(\nu)) - \gamma_+\hat{\mu}(\nu)I &= \nu M(\nu) - \gamma_+\nu I \\ &\quad + O(|\nu-\mu|^{p+1}) \\ &\succeq \nu(M(\nu) - \gamma_+ I) \\ &\quad - c(\mu-\nu)^{p+1}I \\ &\succeq \nu\delta I - c\mu^{p+1}I, \\ |\hat{\mu}(\nu) - \nu| &\leq c(\mu-\nu)^{p+1} \leq c\mu^{p+1}. \end{aligned}$$

For all ν satisfying

$$\frac{c}{\delta}\mu^{p+1} \le \nu \le \mu \tag{12}$$

there follows

$$\frac{1}{2}(\hat{X}(\nu)\hat{Y}(\nu) + \hat{Y}(\nu)\hat{X}(\nu)) - \gamma_+\hat{\mu}(\nu)I \succeq 0\,, \tag{13}$$

$$|\hat{\mu}(\nu) - \nu| \le c\mu^{p+1} \le \delta\nu\,. \tag{14}$$

Moreover,

$$\begin{aligned} \mu - \hat{\mu}(\nu) \quad &\ge \quad \mu - \nu - c(\mu-\nu)^{p+1} \\ &= \quad (\mu-\nu)(1 - c(\mu-\nu)^p) \ge (\mu-\nu)(1-c\mu^p) \\ &\ge \quad (\mu-\nu)(1-c\mu_0^p) \ge 0\,, \end{aligned} \tag{15}$$

if $\mu_0 > 0$ is small enough, namely if

$$\mu_0 \le \min\{1, c^{-1/p}\}. \tag{16}$$

Suppose now that (12) has a solution ν, that is: $(c/\delta)\mu^p \le 1$. Hence (14) and (15) imply $\mu \ge \hat{\mu}(\nu) > 0$ for these ν. Then, by (13), $(\hat{X},\hat{Y})(\mu) = (X,Y) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ and the continuity of $(\hat{X},\hat{Y})(\nu)$, also $(\hat{X},\hat{Y})(\nu) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}$ for all $\nu \in [(c/\delta)\mu^{p+1},\, \mu]$. According to the definition (10) of ν_+ we therefore get $\nu_+ \le (c/\delta)\mu^{p+1}$, and, by $0 < \delta < 1/2$ and (14),

$$\mu_{k+1} = \mu_+ \le \nu_+(1+\delta) \le 2\nu_+ \le \frac{2c}{\delta}\mu^{p+1} = \frac{2c}{\delta_k}\mu_k^{p+1}\,. \tag{17}$$

By induction, we show next that for all $k \ge 0$

$$\frac{c}{\delta_k}\mu_k^p \le \frac{1}{2^{k+2}}, \tag{18}$$

provided $\mu_0 > 0$ is small enough, viz., if

$$\frac{c}{\delta_0}\mu_0^p \le \frac{1}{8}\,.$$

The assertion is trivially true for $k = 0$. If it holds for some $k \ge 0$, then also $(c/\delta_k)\mu_k^p \le 1$, so that (17) applies. Using $\delta_{k+1} = \delta_k/2$ and $p \ge 1$ there follows

$$\begin{aligned} \mu_{k+1}^p\frac{c}{\delta_{k+1}} = 2\frac{c}{\delta_k}\mu_{k+1}^p \quad &\le \quad \frac{2c}{\delta_k}\left(2\frac{c}{\delta_k}\mu_k^{p+1}\right)^p = \left(2\frac{c}{\delta_k}\mu_k^p\right)^{p+1} \\ &\le \quad \left(\frac{1}{2^{k+1}}\right)^{p+1} \le \frac{1}{2^{2k+2}} \le \frac{1}{2^{k+3}}. \end{aligned}$$

The last inequality is true for $k \ge 1$. The estimate $\frac{\mu_1^p c}{\delta_1} \le \frac{1}{8}$ can be seen from the first inequality. Hence, (18) holds for all $k \ge 0$.

For the local convergence analysis we sharpen (16) by assuming that $\mu_0 > 0$ even satisfies

$$\begin{array}{rcl}\mu_0 & \leq & \min\left\{1, (\delta_0/8)^{1/p} c^{-1/p}\right\} \\ & = & \min\left\{1, ((\gamma_0 - \underline{\gamma})/16)^{1/p} c^{-1/p}\right\} =: \beta(\underline{\gamma}).\end{array} \tag{19}$$

Then, in particular, (17) and (18) imply

$$\mu_{k+1} \leq \frac{1}{2^{k+1}} \mu_k \quad \text{for all } k \geq 0. \tag{20}$$

Note that, because of $\delta_k \to 0$, (17) does not imply the Q-superlinear convergence of the μ_k with order $p+1$. However, we can show that the order of convergence is "almost" $p+1$:

Theorem 2.1 *Suppose that Assumption* 1.2 *holds and the starting value* μ_0 *satisfies* $0 < \mu_0 \leq \beta(\underline{\gamma})$ (see (19)). *Then, the algorithm defined by the curvilinear search* (10) *generates parameters* μ_k, $k \geq 0$, *that converge Q-superlinearly to zero with convergence order* $p + 1 - 0$*: that is, for each* $\epsilon \in (0, 1/2]$ *there exists an integer* $k(\epsilon) \geq 0$ *and a constant* $d(\epsilon) > 0$ *such that for all* $k \geq k(\epsilon)$

$$\mu_{k+1} \leq d(\epsilon) \mu_k^{p+1-\epsilon}.$$

Proof. The Q-superlinear convergence of the μ_k already follows from (20) and $0 < \mu_0 \leq 1$. Moreover, (20) implies by induction

$$\mu_{k+1} \leq \left(\prod_{j=1}^{k+1} 2^{-j} \right) \mu_0 \leq 2^{-\frac{1}{2}(k+1)(k+2)} \quad \text{for all } k \geq 0.$$

Hence, because of (17) and $\delta_k = 2^{-k}\delta_0$,

$$\mu_{k+1} \leq \left(\frac{2c}{\delta_k} \mu_k^{\epsilon} \right) \mu_k^{p+1-\epsilon} \leq \left(\frac{c}{\delta_0} 2^{-\frac{1}{2}\epsilon k^2 + \frac{1}{2}(2-\epsilon)k + 1} \right) \mu_k^{p+1-\epsilon}.$$

Since for $\epsilon > 0$

$$\lim_{k \to \infty} 2^{-\frac{1}{2}\epsilon k^2 + \frac{1}{2}(2-\epsilon)k+1} = 0,$$

existence of a constant $d(\epsilon)$ with the required properties follows for $\epsilon \in (0, \frac{1}{2}]$. ■

References

Alizadeh, F., Haeberly, J.-P. and Overton, M. (1998) Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM Journal on Optimization*, **8**, 746-768.

Halická, M. (2002) Analyticity of the central path at the boundary point in semidefinite programming. *EJOR*, **143**, 311-324.

Monteiro, R.D.C. and Pang, J.-S. (1998) On two interior-point mappings for nonlinear semidefinite complementarity problems. *Math. Oper. Res.*, **23**, 39 - 60.

Preiss, M. and Stoer, J. (2003) Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems. To appear in *Math. Programming.*

Stoer, J. (2001) High Order Long-Step Methods for Solving Linear Complementary Problems. *Annals of Operations Research*, **103**, 149–159.

Stoer, J. (1999) Improved High Order Long-Step Methods for Solving Linear Complementary Problems. Technical report, Universität Würzburg, Würzburg.

Stoer, J. and Wechs, M. (1999) On the analyticity properties of infeasible-interior-point paths for monotone linear complementarity problems. *Numer. Mathematik*, **81**, 631-645.

Stoer, J., Wechs, M. and Mizuno, S. (1998) High order infeasible-interior-point methods for solving sufficient linear complementarity problems. *Math. of Operations Research*, **23**, 832-862.

Todd, M.J. (1999) A study of search directions in primal-dual interior-point methods for semidefinite programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Wright, S. (1997) *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia.