Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

Anton Schiela

Convergence of the Control Reduced Interior Point Method for PDE Constrained Optimal Control with State Constraints. [1]

[1]Supported by the DFG Research Center Matheon "Mathematics for key technologies"

ZIB-Report 06-16 (April 2006)

Convergence of the Control Reduced Interior Point Method for PDE Constrained Optimal Control with State Constraints. [†]

Anton Schiela

April 12, 2006

Abstract

We propose a variant of the control reduced interior point method for the solution of state constrained problems. We show convergence of the corresponding interior point path-following algorithm in function space. Morever, we provide error bounds for the iterates.

AMS MSC 2000: 90C51, 49M15

Keywords: interior point methods in function space, optimal control, state constraints

[†]Supported by the DFG Research Center MATHEON "Mathematics for key technologies"

1 Introduction

State constrained optimal control is widely considered as a difficult, but practically very important issue. Interior point methods applied to such problems usually show good numerical performance, which suggests that interior point methods may converge in function space. The purpose of this paper is the construction and proof of convergence of an interior point method for a model problem in state constrained optimal control.

We consider a method that is member of the class of *control reduced interior point methods* [14, 13]. This class of methods has been considered in the authors PhD-thesis [12] in the context of control constrained optimal control. The favourable theoretical properties of these methods led to the conjecture that a convergence proof for state constraints is possible.

Interior point methods for state constrained problems suffer from an important theoretical difficulty: if the state approaches the bound of the feasible region too quickly or even touches the bound, then convergence of the method may slow down or even come to a halt. If the classical *logarithmic* barrier method is used, then this problem cannot be excluded theoretically, except for special cases. We tackle this difficulty by strengthening the barrier functional. Then regularity of the state can be employed to show strict feasibility of the iterates. After strict feasibility has been established, a proof of convergence can be completed in the theoretical framework developed in [12].

As alternatives to interior point methods Lavrentiev-regularizations of state constrained problems [10, 11] or exterior penalty methods [8, 7] have been discussed recently. Convergence of the *solutions* of the regularized problems to the original solution has been shown in both. Moreover, for the case of Lavrentiev-regularization an interior point method has been shown to converge linearly to each regularized problem [11]. In the case of exterior penalty methods it has been shown that a semismooth Newton method can be applied to obtain local superlinear convergence for each subproblem. However, for the corresponding path-following algorithms no rigorous proof of convergence to the original state constrained problem seems to be available up to now.

The paper is organized as follows. In Section 2 we mainly collect general results on the analysis of barrier regularizations for state constrained optimal control problems. To keep this exposition concise we will heavily rely on the results from [12]. Section 3 is devoted to L_1-estimates for gradients of barrier functions. This is the first important step of the proof. The next – crucial – step is presented in Section 4, where strict feasibility of the state is shown, provided an appropriate barrier functional is chosen, depending on the regularity of the state. Section 5 concludes the discussion by showing convergence of an interior point path-following algorithm. Here we use strict feasibility to successfully apply the theoretical framework developed in [12].

Acknowledgement. The author want's to thank Prof. Dr. Peter Deuflhard, Prof. Dr. Fredi Tröltzsch, and Dr. Martin Weiser for the helpful discussions, the encouragement, and the support during the work on this topic.

2 Basic Existence and Convergence Results

We will consider the barrier method for convex optimization problems in an abstract setting and show first basic results. Our exposition uses the framework of convex analysis as for example presented in the first part of [5]. Our problem we consider in this paper will be stated as follows:

$$\min_{u\in L_2(\Omega),y\in L_2(\Omega)} \|y-y_d\|^2_{L_2(\Omega)}+\|u\|^2_{L_2(\Omega)} \qquad s.t. \tag{1}$$

$$Ly = u \tag{2}$$

$$y \geq 0 \tag{3}$$

We call y the state and u the control. We restrict ourselves to this model problem only for ease of presentation. The techniques used here also apply to more complicated problems in terms of bounds and type of the PDE. However, at least for the techniques used in Section 2 convexity is crucial.

We fix our framework by introducing the following basic assumptions:

Assumption 2.1. Let $\Omega \subset \mathbb{R}^d, d = 1,2,3$ be a smoothly bounded, open region and denote by $t \in \Omega$ the independent spacial variable. Let (2) a linear elliptic PDE with $Ly = -\operatorname{div}(a(t)\nabla y) + c(t)y$ with symmetric $a(t) \in \mathbb{R}^{d\times d}$ uniformly positive definite and $c(t) \in \mathbb{R}$ non-negative, and positive on a non-zero subset of Ω is a second order elliptic differential operator. Both $a(t)$ and $c(t)$ are assumed to be smooth. As boundary conditions we impose homgeneous Neumann conditions, or Robin conditions. Under these assumptions standard regularity results (c.f. e.g. [6]) hold.

As usual $L : H^1 \to (H^1)^*$, and hence, we can define an adjoint operator $L^* : H^1 \to (H^1)^*$. Moreover, by positivity of $c(t)$ we know that the equation $Ly = u$ is uniquely solvable for $u \in L_2(\Omega)$ and

$$L^{-1} : L_2(\Omega) \to H^2(\Omega) \hookrightarrow C(\overline{\Omega}).$$

To be able to apply an interior point method, we assume that there are $(\breve{u}, \breve{y}) \in L_2(\Omega) \times H^2(\Omega)$, that satisfy (2) and that are *strictly feasible* with respect to (3). This means that there is a $d_{\min} > 0$, such that $\breve{u} \in L_2(\Omega)$ and

$$\inf_{t\in\Omega} \breve{y} = d_{\min}. \tag{4}$$

This assumption can be interpreted as a Slater constraint qualification.

By the solvability assumption we can insert $y = L^{-1}u$ into (1)-(3) and rewrite our problem as an unconstrained convex minimization problem in u:

$$\min_{u \in L_2} F(u) : L_2 \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\} \tag{5}$$

This is possible, since we may set $F(u) = \infty$ for infeasible points. Note that F is strictly convex and coercive in L_2. Our assumptions imply that Problem (1)-(3) admits exactly one solution in $L_2 \times H^2$. We will call this solution (u_*, y_*).

Next we introduce a class of barrier functions, that has been analysed in [12].

Definition 2.2. Let $B \subset \Omega \times \mathbb{R}^n$ be measurable such that all slices $B(t), t \in \Omega$ are closed and convex with non-empty interior. Assume that the function $l(t, x) : \Omega \times \mathbb{R}^n \to \overline{\mathbb{R}}$ has the following properties:

(i) $l(\cdot, x)$ is a measurable function for any constant $x \in \mathbb{R}^n$.

(ii) for each $t \in \Omega$ $l(t, \cdot)$ is convex, continuous, and (Fréchet-)differentiable on $\mathrm{int}B(t)$

(iii) $x(t) \notin \mathrm{int}B(t) \Leftrightarrow l(t, x) = \infty$, and $lim_{x \to \partial B} l(t, x) = \infty \; \forall t \in \Omega$.

(iv) $l(t, x)$ can be minorized by a function $f(t, x) = a(t) - c \sum_{i=1}^{n} |x_i|$.

Then we call l a *barrier function* for B.

Further, for $1 \le p \le \infty$ we define the *barrier functional* corresponding to l by:

$$\begin{aligned} b : \Omega \times L_p(\Omega) &\to \overline{\mathbb{R}} \\ (t, u) &\mapsto \int_\Omega l(t, u(t))\, dt. \end{aligned}$$

Assumption (iv) is technical and assures that $\int_\Omega l(t, x(t))\, dt > -\infty \, \forall x \in L_1(\Omega)$ (c.f. [5, Remark IX.2.1]).

If we add a barrier functional b to the original objective function F, then we denote the resulting barrier problem by:

$$\min_{u \in U} (F + \mu b)(u), \tag{6}$$

and call its solution $(u(\mu), y(\mu))$. Our problem now reads:

$$(F + \mu b)(u) = F(L^{-1}u, u) + \mu b(L^{-1}u).$$

Note that b may also depend on μ. Then μb can be seen as a notation for $b(y; \mu)$. We recapitulate the most important results obtained in [12] for this class of barrier functions:

Theorem 2.3 (Solutions of Barrier Problems). *For each* $\mu > 0$ (6) *is well defined and admits a unique minimizer* $(u(\mu), y(\mu))$. *Moreover,* $(u(\mu)(t), y(\mu)(t))$ *are strictly feasible almost everywhere in* Ω.

Proof. This is [12, Theorem 2.2.2]. Mainly one has to show that $b(u)$ is well defined, convex, and lower semi-continuous. Then the classical theorem on minima of convex functions (e.g. [5, Proposition II.1.2]) yields the result. $\square$

Next we recapitulate convergence of barrier methods. Our theorem shows that barrier methods show favourable convergence behaviour under very general assumptions.

Theorem 2.4. *Let b be a barrier functional with $l(t,\mu) \le -C\ln(t,\mu)$. Then for $b(y(\mu)) > 0$:*

$$
\begin{aligned}
(F+\mu b)(u(\mu)) &\le F(u_*) + C\mu(1-\mu\ln(d_{\min}\mu)) \\
\|u(\mu)-u_*\|_{L_2(\Omega)} &\le \frac{C}{\sqrt{\alpha}}\sqrt{\mu(1-\mu\ln(d_{\min}\mu))}.
\end{aligned}
$$

Proof. This is again a consequence of the results in [12], namely Lemma 2.3.2 and Corollary 2.3.3. $\square$

By definition, a barrier function $l(t,\cdot)$ is differentiable in $\mathrm{int}B(t)$. Hence for $b(y) \neq +\infty$ the linear functional

$$
b'(y) : L_p(\Omega) \to \mathbb{R} : \left\langle b'(y), v\right\rangle := \int_\Omega l'(t,y(t))v(t)\,dt
$$

is a natural candidate for a derivative of b.

Lemma 2.5. *Let $b(y)$ and $b(y+\delta y)$ be finite. Then b is directionally differentiable at y in the direction δy and*

$$
b'(y;\delta y) = \left\langle b'(y);\delta y\right\rangle .
$$

If $p=\infty$ and y is strictly feasible, then b is Gâteaux differentiable at y with derivative $b'(y)$.

Proof. See [12, Lemma 2.1.10]. $\square$

In the following let $M(\overline{\Omega})$ denote the space of regular Borel measures, which comes into play by the Riesz representation theorem: $M(\overline{\Omega}) \cong C(\overline{\Omega})^*$.

Theorem 2.6 (First order necessary conditions for state constraints). *Assume that L^{-*} maps $M(\overline{\Omega})$ continously into some Banach space $W^\diamond \subset L_2(\Omega)$. Then the following first order necessary (and sufficient) conditions hold:*

$$
\begin{aligned}
y - y_d + m - L^*\lambda &= 0 \\
\alpha u + \lambda &= 0 \\
Ly + u &= 0,
\end{aligned}
\tag{7}
$$

for some $m \in \mu\partial b(y) \subset M(\overline{\Omega})$ and $\lambda \in W^\diamond$. If y is strictly feasible, then $m = \mu b'(y)$.

Proof. This is [12, Theorem 2.2.4]. The important result $m = b'(y)$ for strictly feasible y is a consequence of Gâteaux differentiability of $b(y)$ in $C(\overline{\Omega})$ as shown in Lemma 2.5. Analysis of the operator L^{-*} (c.f. [3, 2]) yields characterizations of $W^\diamond$ as a Sobolev space. □

If only $y \geq 0$ holds, then we cannot show $b'(y) = \partial b(y)$. There may occur additional terms (such as point functionals) at the touch points of the bounds and the solution. Assume, for example, that $y(0) = 0$, $b'(y) \in \partial b(y)$. Then

$$b(y + \delta y) \geq b(y) + b'(y)\delta y - \delta y(0).$$

Indeed, if $\delta y(0) < 0$, then by continuity of $y + \delta y$ it follows that $b(y + \delta y) = \infty$. For $\delta y(0) \geq 0$ this inequality holds, since $b'(y) \in \partial b(y)$. Thus $b'(y) - \delta_0 \in \partial b(y)$ with δ_0 being the point measure at 0.

The consequence for barrier methods applied to state constrained optimal control problems is unpleasant. If y is not strictly positive and $b'(y)$ is used in place of $\partial b(y)$ in the first order necessary conditions of the barrier problem, then this system is not necessarily solvable, although there exists a minimizer of the barrier problem. Thus, in such a case a barrier method in function space is likely to fail.

Hence we have to construct methods with *strictly feasible* iterates. This is the aim of the rest of this paper.

3 Uniform Bounds for Gradients

The first order necessary conditions state the qualitative result that there is $m(\mu) \in M(\overline{\Omega})$, such that (7) holds. We augment this by a quantification, that holds uniformly for all $\mu > 0$.

Lemma 3.1. *Let $l : \mathbb{R}_+ \to \mathbb{R}$ be a monotonically decreasing barrier function and b the associated barrier functional. Let $p \in C(\overline{\Omega})$ be non-negative, and $m \in \partial b(y)$. Then*

$$0 \geq \left\langle b'(y), p \right\rangle \geq \langle m, p \rangle . \tag{8}$$

In particular, m is a non-positive measure.

Proof. Since p is non-negative and $b(y) < \infty$, so is $b(y + p) \leq b(y)$ by monotonicity of b. Thus, $b(y)$ is directionally differentiable in the direction p and

$$b(y + hp) - b(y) \leq \left\langle b'(y), hp \right\rangle + o(h).$$

Assume now for contradiction that

$$\langle m, p \rangle - \left\langle b'(y), p \right\rangle = \varepsilon > 0.$$

Then

$$\langle m, hp \rangle \geq \left\langle b'(y), hp \right\rangle + h\varepsilon > b(y + hp) - b(y)$$

for sufficiently small h, which contradicts the hypothesis $m \in \partial b(y)$. □

Now we show that $\|m(y(\mu))\|_{M(\overline{\Omega})}$ is unifomly bounded in μ, depending only on the coercivity of F and on the Slater point $(\breve{u}, \breve{y})$.

Proposition 3.2. *Let b be a barrier functional corresponding to a monotonically decreasing barrier function l. Assume, there are $\breve{u} \in L_2$, $\breve{y} \geq d_{\min} > 0$ with $L\breve{y} = \breve{u}$. Then*

$$\left\|\mu b'(y(\mu))\right\|_{L_1(\Omega)} \leq \|m(\mu)\|_{M(\overline{\Omega})} \leq C$$

independently of μ. Moreover,

$$\|\lambda(\mu)\|_{W^\diamond} \leq C$$

independently of μ.

Proof. We eliminate $y = L^{-1}u$ from the optimal control problem. Then the first order necessary conditions for the minimum read:

$$\langle u(\mu), v\rangle + \left\langle y(\mu) - y_d, L^{-1}v\right\rangle + \left\langle m(\mu), L^{-1}v\right\rangle = 0 \qquad \forall v \in L_2.$$

By Lemma 2.5 $m(\mu)$ must be a non-positive measure. Inserting $v := \breve{u}, L^{-1}\breve{u} = \breve{y}$ we obtain

$$\alpha \langle u(\mu), \breve{u}\rangle + \langle y(\mu) - y_d, \breve{y}\rangle - \langle m(\mu), \breve{y}\rangle = 0.$$

Since $u(\mu), y(\mu), \breve{u}, \breve{y}, y_d$ are bounded in L_2 independent of μ we conclude

$$\langle m(\mu), \breve{y}\rangle = \int_{\overline{\Omega}} \breve{y}\, dm(\mu) \leq C$$

and thus by non-positivity of $m(\mu)$ and positivity of $\breve{y}$:

$$\|m(\mu)\|_{M(\overline{\Omega})} \leq \frac{|\langle m(\mu), \breve{y}\rangle|}{\min\{\breve{y}\}} \leq C/d_{\min}.$$

By (8) it follows that $\|\mu b'(y(\mu))\|_{L_1(\Omega)} \leq \|m(\mu)\|_{M(\overline{\Omega})}$. □

4 Barrier Functions with Rational Gradients

Now we consider a special barrier functional b:

$$b(y) := \int_\Omega -\ln(y(t)) + \frac{\mu^{q-1}}{(q-1)y(t)^{q-1}}\, dt \tag{9}$$

with $q > 1$. By $q = 1$ we denote the case of the pure logarithmic barrier functional. We say that the barrier functional $b(y)$ is of *order* q. Its gradient can then be computed as

$$\mu b'(y) = -\frac{\mu}{y} - \frac{\mu^q}{y^q}.$$

Observe that the rational term of the barrier function only adds significantly to the functional, if $y \ll \mu$, thus the difference between b and the logarithmic barrier function is small on most parts of the domain.

We will derive conditions, under which $b'(y(\mu)) \in L_\infty$. This will be crucial for the analysis of our Newton path-following scheme.

Lemma 4.1. *Let $\Omega \subset \mathbb{R}^d$, $0 \le y \in C^\beta(\Omega), 1/y^q \in L_1(\Omega)$, $\beta \le 2$. Assume that $\beta \ge d/q$. Then $1/y \in C(\Omega)$.*

Proof. Assume that without loss of generality $0 \in \Omega$, $y(0) = 0$. Since $y \in C^1$ for $\beta \ge 1$, and y necessarily obtains a minimum at 0, we conclude $\nabla y(0) = 0$ for $\beta \ge 1$. By the Hölder-continuity of y or ∇y we infer that $y(x) < cr^{1+\beta}$ for all $x \in B(0,r)$. Hence we can compute

$$\int_\Omega |1/y(t)|^q \, dt \ge c \int_{[0,R]} |1/(cr^\beta)|^q r^{d-1} \, dr = \infty,$$

which is a contradiction to $1/y^q \in L_1$. Hence $y > 0$ in $\overline{\Omega}$, which implies by compactness of $\overline{\Omega}$ that $y \ge c$, and thus $1/y \in C(\Omega)$. □

Since $\|m(\mu)\|_{M(\overline{\Omega})} \le C$, we can use well known regularity results to show that $\|y(\mu)\|_{C^\beta}$ is *uniformly bounded* for $\mu \to 0$ for some β. By [3] we know that $y \in W^{3,p}$ with $p < d/(d-1)$. By Sobolev embedding theorems [1] we thus conclude $\beta = 2, 2-\varepsilon, 1-\varepsilon$ for $d = 1,2,3$ respectively.

For sufficiently large q Lemma 4.1 yields that $y(\mu) \ge \varepsilon(\mu) > 0$ and thus strict feasibility. Consequently $m(\mu) = b'(y(\mu)) \in L_\infty(\Omega)$.

If we compute the integrals, then also the assertion of Lemma 4.1 can be quantified, and it can be shown that for sufficiently large q there is a function $\psi(\mu)$ that is *uniformly bounded* in $[\mu_0; \overline{\mu}]$ for each $\overline{\mu} > 0$, such that $\min_{t\in\Omega} y(\mu)(t) \ge \psi(\mu)$. More specifically $\psi(\mu)$ must fulfill the following inequality:

$$\int_\Omega \frac{r^{d-1}\mu^q}{(r^\beta + \psi(\mu))^q} \, dr \le C.$$

This uniformity property is essential for a convergence proof. Thus for barrier functionals with sufficiently large order q(e.g. $q = 1, 1+\varepsilon, 3+\varepsilon$ for $d = 1,2,3$ resp.) we obtain a bound $\|1/y(\mu)\|_{L_\infty} \le \psi(\mu)^{-1}$ and analogously $\|b'(y(\mu))\|_{L_\infty} \le \widetilde{\psi}(\mu)$.

All in all we have proven:

Theorem 4.2. *Let μ_0 be given. For sufficiently large q there are positive functions $\psi(\mu)$ and $\widetilde{\psi}(\mu)$, such that*

$$\min_{t\in\Omega} y(\mu)(t) \ge \psi(\mu)$$
$$\left\|b'(y(\mu))\right\|_{L_\infty} \le \widetilde{\psi}(\mu).$$

For each $\underline{\mu} > 0$ the functions $\psi(\mu)^{-1}$ and $\widetilde{\psi}(\mu)$ are uniformly bounded on the interval $[\underline{\mu}, \mu_0]$.

In the following we set the order of the barrier functional q sufficiently large. Then $\partial b = b'$ and the (control reduced) first order necessary conditions read:

$$\begin{aligned} (y - y_d) + \mu b'(y) + L^*\lambda &= 0 \\ Ly - \alpha^{-1}\lambda &= 0. \end{aligned} \tag{10}$$

Here we have eliminated the control using the equation $\alpha u + \lambda = 0$. By Theorem 2.6 and Theorem 4.2 this system has a strictly feasible and unique solution $(y(\mu), \lambda(\mu))$ for each $\mu > 0$. Since $b'(y(\mu)) \in L_\infty$ we also have $y(\mu), \lambda(\mu) \in H^2$.

Corollary 4.3. *For sufficiently large order q the gradient $b'(y(\mu))$ converges to the Lagrange multiplier m_* in the sense of weak* convergence in $M(\overline{\Omega})$:*

$$b'(y(\mu)) \rightharpoonup m_*.$$

Proof. The functional $\mu b'(y(\mu))$ is uniformly bounded in $M(\overline{\Omega}) = C^*(\overline{\Omega})$ and thus a sequence $\mu_k b'(y(\mu_k))$ contains a weakly* convergent subsequence with limit $(\mu b')_*$ for which (7) is satisfied. Moreover,

$$\langle \mu b'(y(\mu)), y(\mu) \rangle = \mu \int_\Omega 1 + \frac{\mu^{q-1}}{y^{q-1}}\, dt \leq C\mu,$$

since $\int \frac{\mu^{q-1}}{y^{q-1}}$ is bounded by boundedness of $\int \frac{\mu^q}{y^q}$. Since $y(\mu)$ converges strongly and $\mu b'(y(\mu))$ converges weakly, we obtain

$$\lim_{\mu \to 0} \langle \mu b'(y(\mu)), y(\mu) \rangle = \langle (\mu b')_*, y_* \rangle = 0.$$

Moreover, by non-positivity of b' and convexity of the negative cone we conclude $b'_* \leq 0$. Hence $b'_* = m_*$. By uniqueness of m_* the desired weak* convergence of the whole sequence follows. $\square$

5 A Primal Interior Point Path-Following Method

Strict feasibility of y enables us to study the analytic properties of the central path in the framework of [12, Chapter 3]. We define $x := (y; \lambda)$ and write (10) as

$$F(x; \mu) = 0,$$

which can be declared on the Banach space

$$X = H^2(\Omega) \times L_2(\Omega)$$

with $F(x; \mu) : X \to X^*$. We equip X with the norm

$$\|\delta x\|_X := \|\delta y\|_{H^2(\Omega)} + \alpha^{-1/2} \|\delta \lambda\|_{L_2(\Omega)}$$

Note that L^* can be extended by duality to $L^* : L_2 \to (H^2(\Omega))^*$. F is well defined and smooth in a neighbourhood of the central path defined by:

$$U_\mu(y(\mu)) = B\left(y(\mu); \psi(\mu)/2, \|y\|_{L_\infty}\right).$$

It is important that these balls do not degenerate, except for $\mu \to 0$. If we assume strict feasibility, then under suitable assumptions all following results can be carried over to non-convex problems.

Qualitatively speaking, the barrier problem is smooth inside this neighbourhood and thus Newton's method applied to F can be shown to converge locally. For a convergence proof we will have to quantify the relation between the radius of convergence and μ.

As pointed out in [12, Chapter 3] the concept of *inverse differentiability* is a convenient approach to the affine covariant analysis of Newton methods in function space. This can be seen as an extension of the ideas due to [4] to infinite dimensional normed spaces. For a function $F : X \to Y$ we introduce an *inverse derivative* L^- that has the role of an inverse of the Jacobian matrix. A criterion for inverse (Fréchet)-differentiability at x_0 is

$$\left\| x - x_0 - L^-(x_0)(F(x) - F(x_0)) \right\|_X = o(\|x - x_0\|_X).$$

For the quadratic convergence of Newton's method, one has to show for all x_0 in a neighbourhood of the solution x_*:

$$\left\| x - x_0 - L^-(x_0)(F(x) - F(x_0)) \right\|_X \leq \frac{1}{2}\omega \left\| x - x_0 \right\|_X^2 ,$$

which is equivalent to showing

$$\left\| F'(x_0)^{-1}(F'(x) - F'(x_0))(x - x_0) \right\|_X \leq \omega \left\| x - x_0 \right\|_X^2$$

for some formal Jacobian matrix satisfying the Fundamental Theorem of Calculus. Then the radius of quadratic convergence of Newton's method behaves proportional to ω^{-1}. Moreover, an affine invariant Implicit Function Theorem holds for zeros of $F(x; \mu)$.

In our case it is easy to construct the formal Jacobian matrix $F'(x; \mu)$ for the perturbed optimality system:

$$F'(x; \mu) = \begin{pmatrix} I + \frac{\mu}{y^2} + \frac{q\mu^q}{y^{q+1}} & L^* \\ L & -\alpha^{-1} \end{pmatrix}. \tag{11}$$

Similar as for the control constrained case it can be shown that the Fundamental Theorem of Calculus holds for this definition of $F'(x; \mu)$. We introduce the following norm:

$$\|\delta y\|_{x;\mu} = \left\| \alpha^{1/2} \delta y \right\|_{H^2} + \left\| \sqrt{1 + \mu/y + \mu^q/y^{q+1}} \delta y \right\|_{L_2} .$$

Proposition 5.1. *The Jacobian matrix F' for the perturbed optimality system has the following invertibility property. Let*

$$F'(x; \mu)\delta x = r := \begin{pmatrix} r_a \\ r_s \end{pmatrix}.$$

Then

$$\|\delta y\|_{x;\mu} + \alpha^{-1/2} \|\delta\lambda\|_{L_2} \leq \sup_{v \in H^2} \frac{\langle r_a, v \rangle}{\alpha^{1/2} \|v\|_{x;\mu}} + \sup_{w \in L_2} \frac{\langle r_s, w \rangle}{\alpha^{-1/2} \|w\|_{L_2}}.$$

Proof. This follows from [12, Theorem 4.2.3]. □

To avoid additional technicalities let the order q be an integer in the following.

Lemma 5.2. *Let $a, b \geq c/2 > 0$. Then*

$$|a^{-n} - b^{-n}| \leq 4nc^{-n-1}|a-b|.$$

Proof. This is a simple algebraic computation:

$$\begin{aligned} |a^{-n} - b^{-n}| &= |\frac{b^n - a^n}{(ab)^n}| = \frac{\sum_{i=1}^{n} b^{n-i}a^{i-1}}{(ab)^n}|b-a| \\ &= \sum_{i=1}^{n} b^{-i}a^{i-n-1}|a-b| \leq 4nc^{-n-1}|b-a|. \end{aligned}$$

□

Proposition 5.3. *The following affine covariant Lipschitz condition holds in $U_\mu(y(\mu))$:*

$$\left\|F'(x;\mu)^{-1}(F'(x;\mu) - F'(\widetilde{x};\mu))\Delta x\right\|_X \leq c\omega(\mu)\left\|\Delta y\right\|_{L_\infty}^2 \leq \omega(\mu)\left\|\Delta x\right\|_X^2 \tag{12}$$

together with the estimate

$$\omega(\mu) \leq C\psi(\mu)^{-3/2}.$$

In particular, in a neighbourhood of the central path the family of mappings

$$\begin{aligned} M_{\widetilde{\mu}} : X \times [\underline{\mu}; \mu_0] &\to X \\ (x;\mu) &\mapsto F'(x(\widetilde{\mu});\widetilde{\mu})^{-1}F'(x;\mu) \end{aligned}$$

is equi-continuous in $[\mu_0; \underline{\mu}]$.

Proof. We abbreviate the right hand side of (12) by $\delta x = (\delta y, \delta\lambda)$ with

$$F'(x;\mu)\delta x = \begin{pmatrix} r_a \\ r_s \end{pmatrix} = \begin{pmatrix} (F'(x;\mu) - F'(\widetilde{x};\mu))\Delta x \\ 0 \end{pmatrix}.$$

In view of application of Proposition 5.1 we compute the difference of two Jacobians in $U_\mu(y(\mu))$ multiplied with a test function $\varphi = (v, w)$. Since only the upper left block of the Jacobian depends on x we obtain

$$\left\langle (F'(x;\mu) - F'(\widetilde{x};\mu))\Delta y, \varphi\right\rangle = \left\langle \left(\frac{\mu}{y^2} + \frac{q\mu^q}{y^{q+1}} - \frac{\mu}{\widetilde{y}^2} - \frac{q\mu^q}{\widetilde{y}^{q+1}}\right)\Delta y, v\right\rangle.$$

Now we apply Lemma 5.2 pointwise setting $a = y, b = \widetilde{y}, c = y(\mu)$ and continue:

$$
\begin{aligned}
&\left|(F'(x;\mu) - F'(\widetilde{x};\mu))\Delta x\right| \le C\left|\left(\frac{\mu}{y(\mu)^3} + \frac{\mu^q}{y(\mu)^{q+2}}\right)\Delta y^2\right| \\
&= C\left|\sqrt{\frac{\mu}{y(\mu)^2} + \frac{\mu^q}{y(\mu)^{q+1}}}\right|\left|\frac{1}{y(\mu)^{3/2}}\right|\left|\sqrt{\frac{\mu}{y(\mu)} + \frac{\mu^q}{y(\mu)^q}}\right|\left|\Delta y^2\right| \\
&\le C\left|\sqrt{\frac{\mu}{y^2} + \frac{\mu^q}{y^{q+1}}}\right|\left|\frac{1}{y(\mu)^{3/2}}\right|\left|\sqrt{\frac{\mu}{y(\mu)} + \frac{\mu^q}{y(\mu)^q}}\right|\left|\Delta y^2\right|.
\end{aligned}
$$

The exchanges between $y(\mu)$ and y are possible, since $y \in U_\mu(y(\mu))$. Hence,

$$
\begin{aligned}
&\left\langle (F'(x;\mu) - F'(\widetilde{x};\mu))\Delta y, \varphi\right\rangle \\
&\le \left\langle \left|\frac{1}{y(\mu)^{3/2}}\right|\left|\sqrt{\frac{\mu}{y(\mu)} + \frac{\mu^q}{y(\mu)^q}}\right|\left|\Delta y^2\right|, \left|\sqrt{\frac{\mu}{y^2} + \frac{\mu^q}{y^{q+1}}}\right| v\right\rangle \\
&\le C\left\|\frac{1}{y(\mu)^{3/2}}\right\|_{L_\infty}\left\|\frac{\mu}{y(\mu)} + \frac{\mu^q}{y(\mu)^q}\right\|_{L_1}^{1/2}\|\Delta y\|_{L_\infty}^2\|v\|_{x,\mu} \\
&\le C\psi(\mu)^{-3/2}\|\Delta y\|_{L_\infty}^2\|v\|_{x,\mu}.
\end{aligned}
$$

For the last estimate we used the fact that $\frac{\mu}{y(\mu)} + \frac{\mu^q}{y(\mu)^q} = b'(y(\mu))$ is uniformly bounded in L_1. By Proposition 5.1 and the embedding $H^2 \to L_\infty$ we conclude

$$
\begin{aligned}
\|\delta x\|_X &\le \|\delta y\|_{x;\mu} + \alpha^{-1/2}\|\delta\lambda\|_{L_2} \\
&\le \psi(\mu)^{-3/2}\|\Delta y\|_{L_\infty}^2 \le c\psi(\mu)^{-3/2}\|\Delta y\|_{H^2}^2 \le c\psi(\mu)^{-3/2}\|\Delta x\|_X^2.
\end{aligned}
$$

Hence, $M_{\widetilde{\mu}}$ is uniformly continuous in an X-neighbourhood of $x(\widetilde{\mu})$. It is now easy to show continuity with respect to μ and thus continuity in $(x;\mu)$. Since the continuity-modules hold uniformly (we have shown Lipschitz continuity together with a bound on the Lipschitz constant) for $\widetilde{\mu} \in [\underline{\mu};\mu_0]$ the conclusion holds. □

The result of Proposition 5.3 suggests that interior point methods may converge sublinearly. However, the structure of the proof makes clear, why in practice linear convergence is observed: the norm of the Newton corrections is naturally measured in $\|\delta y\|_{x;\mu}$. The scaling in the L_2-term suggests that in regions, where y is small the Newton correction is also small. Then it may happen that

$$
\left\|\frac{1}{y(\mu)^{3/2}}\Delta y^2\right\|_{L_\infty} \ll \left\|\frac{1}{y(\mu)^{3/2}}\right\|_{L_\infty}\|\Delta y\|_{L_\infty}^2,
$$

and thus ω may be much smaller than our estimate.

Now we study differentiability of the central path. For this purpose let

$$
\eta(\mu) := \left\|\frac{d}{d\mu}y(\mu)\right\|_{x;\mu} + \alpha^{-1/2}\left\|\frac{d}{d\mu}\lambda(\mu)\right\|_{L_2}.
$$

Proposition 5.4. *The central path is differentiable and its slope is bounded by:*

$$\eta(\mu) \leq c\mu^{-1/2}.$$

Thus,

$$\alpha^{1/2} \|y(\mu) - y_*\|_{H^2} + \alpha^{-1/2} \|\lambda - \lambda(\mu)\|_{L_2} \leq c\sqrt{\mu}. \tag{13}$$

Proof. Proposition 5.3 assures strong inverse differentiability of F via [12, Lemma 3.2.2] in a neighbourhood of the central path. To show existence and differentiability of the central path via the affine invariant global existence Theorem in [12, Theorem 3.5.1], we merely have to show differentiability of $F'(x(\hat{\mu}); \hat{\mu})^{-1} F(x(\hat{\mu}); \mu)$ for fixed $\hat{\mu}$ with respect to μ. Since $y(\mu)$ is bounded and μ is scalar, differentiability is clear, and we can easily compute the derivative:

$$F_\mu(x(\mu); \mu) = \frac{1}{y(\mu)} + \frac{\mu^{q-1}}{y(\mu)^q}.$$

Inserting this as a right hand side into Proposition 5.1 we obtain the estimate:

$$\begin{aligned} \|y_\mu(\mu)\|_{x,\mu} + \alpha^{-1/2} \|\lambda_\mu(\mu)\|_{L_2} &\leq \left\| \mu^{-1/2} \left(1 + \frac{\mu^{q-1/2}}{y(\mu)^q} \frac{y(\mu)^{(q+1)/2}}{\mu^{q/2}} \right) \right\|_{L_2} \\ &\leq \mu^{-1/2} c \left(1 + \left\| \frac{\mu^{(q-1)/2}}{y(\mu)^{(q-1)/2}} \right\|_{L_2} \right). \end{aligned}$$

Since $\|\mu^q / y(\mu)^q\|_{L_1}$ is uniformly bounded in μ by Proposition 3.2, so is

$$\left\| \frac{\mu^{(q-1)/2}}{y(\mu)^{(q-1)/2}} \right\|_{L_2} = \left\| \frac{\mu^{q-1}}{y(\mu)^{q-1}} \right\|_{L_1}^{1/2}.$$

This shows the conclusion of the proposition. Integration over the slope shows the estimate for the length. □

Now we have collected all the estimates to show the convergence of a simple path-following algorithm, which is the main theorem of this paper. Recall that F is defined by (10), F' is defined by (11), and $x_k = (y_k, \lambda_k)$.

Algorithm 5.5.
select $\mu_0 > 0$, and x_0 with $y_0 \in U_\mu(y(\mu_0))$
for $k = 0, \ldots$
  $x_{k+1} := x_k + F'(x_k, \mu_k)^{-1} F(x_k, \mu_k)$
  select $0 \leq \sigma_k < 1$
  $\mu_{k+1} := \sigma_k \mu_k$

Theorem 5.6 (Convergence for State Constraints). *For some $\mu_0 < 0$ assume that y_0 sufficiently close to $y(\mu_0)$ is given. Then there is a sequence of step sizes σ_k, for which Algorithm* (5.5) *produces a sequence of iterates (y_k, λ_k), such that the sequence $(y_k, u_k := -\alpha^{-1}\lambda_k)$ converges to the optimal solution (y_*, u_*) of* (1)-(3). *Moreover,*

$$\|y_k - y_*\|_{H^2} + \|u_k - u_*\|_{L_2} \leq C\alpha^{-1/2}\sqrt{\mu_k}. \tag{14}$$

Proof. The estimates for $\omega(\mu), \eta(\mu)$, and for the radius $r(\mu)$ of $U_\mu(y(\mu))$ are sufficient to conclude with [12, Theorem 3.6.2] the existence of such a sequence. Without loss of generality the iterates can be assumed to be close enough to the central path, such that (14) holds due to (13). □

Remark 5.7. Since $\lambda(\mu)$ is uniformly bounded in some Sobolev space $W^{1,p}$ and admits a convergence rate in L_2, it follows by interpolation theory that $\lambda(\mu)$ also admits a convergence rate in Sobolev spaces $W^{s,q}$ that depends only on the choice of s and q. An analogous result is valid for $u(\mu)$ and $y(\mu)$.

6 Conclusion

We have shown convergence of an interior point method for a model problem of state constrained optimal control. Our aim was to fix the main ideas in a framework as simple as possible. We note, however, that an extension to more general problems (bilateral bounds, control constraints) is straightforward.

We regard this analysis as a first step to the understanding of interior point methods for state contraints. Its algorithmic implications are clear: other functions than the logarithmic barrier function are worth considering for this class of problems. In practice one will design an affine covariant adaptive path-following algorithm similar to the algorithm developed in [12, Chapter 9]. This is subject to current work.

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, 1975.

[2] H. Amann and P. Quittner. Elliptic boundary value problems involving measures: Existence, regularity, and multiplicity. *Advances in Differential Equations*, 3(6):753–813, 1998.

[3] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.

[4] P. Deuflhard. *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*, volume 35 of *Series Computational Mathematics.* Springer, 2004.

[5] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* Number 28 in Classics in Applied Mathematics. SIAM, 1999.

[6] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer, 1977.

[7] M. Hintermüller and K Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. Technical Report 1, Karl-Franzens-Universität Graz und Technische Universität Graz, 2005.

[8] M. Hintermüller and K. Kunisch. Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.*, accepted for publication.

[9] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. Appl.*, 30:45–63, 2005.

[10] C. Meyer, F. Tröltzsch, and A. Rösch. Optimal control problems of PDEs with regularized pointwise state constraints. Technical Report 14-2003, TU Berlin, 2003. Accepted by Computational Optimization and Applications.

[11] U. Prüfert, F. Tröltzsch, and M. Weiser. The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. ZIB Report 04-47, Zuse Institute Berlin, 2004.

[12] A. Schiela. *The Control Reduced Interior Point Method - A Function Space Oriented Algorithmic Approach.* PhD thesis, Free University of Berlin, Dept. Math. and Comp. Sci., 2006. under review.

[13] A. Schiela and M. Weiser. Superlinear convergence of the control reduced interior point method for PDE constrained optimization. ZIB Report 04-15, Zuse Institute Berlin, 2005.

[14] M. Weiser, T. Gänzler, and A. Schiela. A control reduced primal interior point method for PDE constrained optimization. ZIB Report 04-38, Zuse Institute Berlin, 2004.