



ANTON SCHIELA

Barrier Methods for Optimal Control Problems with State Constraints¹

¹Supported by the DFG Research Center MATHEON "Mathematics for key technologies"

Barrier Methods for Optimal Control Problems with State Constraints [†]

Anton Schiela

March 6, 2008

Abstract

We study barrier methods for state constrained optimal control problems with PDEs. In the focus of our analysis is the path of minimizers of the barrier subproblems with the aim to provide a solid theoretical basis for function space oriented path-following algorithms. We establish results on existence, continuity, and convergence of this path. Moreover, we consider the structure of barrier subdifferentials, which play the role of dual variables.

AMS MSC 2000: 90C51, 49M05

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

1 Introduction

State constrained optimal control problems are an important and challenging topic, especially in the context of PDEs. The theoretical foundation for their understanding has been laid a couple of years ago, but only recently there have been attempts to construct and analyse function space oriented methods for their solution.

Because the original state constrained problem is difficult to tackle directly, the main algorithmic approaches to state constrained problems can all be classified as path-following methods. The author is aware of three main lines of research. *Lavrentiev regularization* methods due to TRÖLTZSCH ET AL. [11, 12, 10] transform state constraints to mixed state-control constraints, which can then be addressed by known algorithms. Related to this are the so called *primal-dual path-following* methods due to KUNISCH AND HINTERMÜLLER [9, 8], which mainly use an L_2 -penalty regularization of the state constraints. Both methods produce infeasible iterates in general. Opposed to this *interior point* or *barrier methods* always stay inside the feasible domain. Most publications on interior point path-following methods in function space consider control constraints such as WEISER, SCHIELA ET AL. [20, 21, 23, 22, 14] and ULBRICH AND ULBRICH [18]. These methods are to

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

be distinguished from *affine scaling interior point methods* considered in [17, 19], which are strongly related to semi-smooth Newton methods [16, 7]. We will study barrier methods for state constraints more closely in this paper.

The simple idea of barrier methods is to replace box constraints by a smooth barrier functional, which tends to infinity if the solution approaches the bounds. The set of barrier minimizers is often called the central path. Our main concern is to put interior point methods for state constraints on a firm theoretical basis by studying the properties of the central path. Based on this insight there are several different variants of path-following methods conceivable the analysis of which is, however, beyond the scope of this paper.

The results of this paper extend the ideas in [15], which sketches the analysis of barrier methods for state constraints on the base of the authors dissertation [14]. Our main interest lies in the derivation of the most important structural results of the central path in terms of classical convex analysis (cf. e.g. [4, 2]).

In Section 2 we describe the class of optimal control problems under consideration. In Section 3 we consider barrier functions with rational gradients, which yield barrier functionals by integration. Multiplication of a barrier function by a positive factor μ and adding it to a convex functional F results in another convex functional F_μ . In Section 4 we consider existence of minimizers of F_μ and first order optimality conditions. In Section 5 we study the properties of the path of minimizers of F_μ for $\mu \in [0; \mu_0]$, such as Lipschitz continuity and convergence.

One conclusion of Sections 3 and 4 is that measure valued dual variables can appear in the presence of state constraints. This may be an obstacle for the construction of Newton path-following methods. Section 6 clarifies this situation and shows how to avoid this problem by choosing an appropriate barrier function.

Acknowledgement. The author wishes to thank Prof. Dr. Peter Deuffhard, Dr. Martin Weiser, and Prof. Dr. Fredi Trötzsch for helpful discussions and various hints concerning the mathematical content of this paper and its presentation.

2 State Constrained Optimal Control Problems

Let Ω be an open and bounded Lipschitz domain in \mathbb{R}^d and $\bar{\Omega}$ its closure. Let $Y := C(\bar{\Omega})$ or $Y = C_0(\bar{\Omega})$ and $U := L_2(\Omega)$ or $U := L_2(\partial\Omega)$. Other variants are conceivable for which the results of this paper can be shown by the same or a slightly modified argumentation.

Define $X := Y \times U$ with $x := (y, u)$ and consider the following convex minimization problem, the details of which are fixed in the remaining section.

$$\begin{aligned} \min_{x \in X} J(x) \quad \text{s.t.} \quad & Ly - u = 0 \\ & \underline{u} \leq u \leq \bar{u} \\ & \underline{y} \leq y \leq \bar{y}. \end{aligned} \tag{1}$$

We will now specify our abstract theoretical framework and collect a couple of basic results about this class of problems.

Convex Functionals. Let $J : X \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous, convex functional, which is coercive on the feasible subset of X defined by the equality and inequality constraints. We assume that there is a feasible point \check{x} (cf. (4) below) where J is continuous. Further, we need some regularity of the subdifferential ∂J of J . We assume that ∂J is uniformly bounded in X^* on bounded sets of X .

Equality Constraints. The equality constraint $Ly - u = 0$ is introduced to model a partial differential equation. We start by stating our set of assumptions. Then we give a simple example where these assumption can be verified.

We assume that $L : Y \supset \text{dom } L \rightarrow U$ is a linear operator, which is *densely defined, surjective, and closed*. Its domain $\text{dom } L$, the subspace of Y defined by all y with $Ly \in U$, is dense in U by definition. For basic results on closed operators we refer to [24, Kapitel IV.4] or [13]. The subspace $E \subset \text{dom } L \times U$ of all (y, u) that satisfy $Ly - u = 0$ is called the *graph* of L . By definition of L as a closed operator E is a *closed subspace* of $Y \times U$.

Surjectivity implies (cf. [24, Satz IV.4.4]) that L is an *open mapping*. This implies that L maps neighbourhoods of 0 in Y onto neighbourhoods of 0 in U . We use surjectivity for the analysis of first order optimality conditions (cf. Theorem 4.3).

Remark 2.1. In most applications $L : \text{dom } L \rightarrow U$ is bijective. Then (cf. [24, Satz IV.4.4]) L is continuously invertible and a continuous solution operator L^{-1} exists – a standard assumption in optimal control. In other cases, such as pure Neumann problems, L is neither injective nor surjective but has a closed range and L is still an open mapping onto its range. Remark 4.4 explains how to weaken the surjectivity assumption to include pure Neumann into our framework.

Because E is a closed subspace of a Banach space, it is a Banach space itself. In order to prove existence of minimizers we have to strengthen this assumption slightly. We assume that E is *weakly sequentially compact* (bounded sequences in E have a weak accumulation point in E). If L is injective, this assumption follows from reflexivity of U , continuous invertibility of L , and closedness of E .

We exploit density of $\text{dom } L$ in U to define an adjoint operator L^* by the following standard construction. For every $l \in U^*$ the mapping $y \rightarrow \langle l, Ly \rangle$ is a linear functional on $\text{dom } L$. We define $\text{dom } L^*$ as the subspace of all $l \in U^*$ for which $y \rightarrow \langle l, Ly \rangle$ is continuous on $\text{dom } L$ and can thus by density be extended uniquely to a continuous functional on Y . Hence, for all $l \in \text{dom } L^*$ there is a unique $L^*l \in Y^*$ for which

$$\langle l, Ly \rangle = \langle L^*l, y \rangle \quad \forall y \in \text{dom } L. \quad (2)$$

This yields the definition of $L^* : U^* \supset \text{dom } L^* \rightarrow Y^*$.

An Example PDE Constraint. In order to become acquainted with our abstract framework, let Ω be a Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d = 1 \dots 3$, $U = L_2(\Omega)$, $Y = C_0(\Omega)$, and $L = -\Delta$, the Laplace operator. For many applications it is natural to consider $L : H_0^1(\Omega) \rightarrow (H_0^1(\Omega))^*$ in the weak formulation as a continuous operator. In our approach we consider $L : C_0(\overline{\Omega}) \supset \text{dom } L \rightarrow L_2(\Omega)$. In this setting L is not continuous and not defined on the whole space $C_0(\overline{\Omega})$.

But, as we will see, L is densely defined, surjective, and closed. This can be proved as follows, using unique solvability of the Poisson equation and the following norm estimate (cf. e.g. [5, Theorem 8.16.]):

$$\|y\|_\infty \leq c \|u\|_{L_2}. \quad (3)$$

Consider a convergent sequence of pairs $(y_k, u_k := Ly_k) \in \text{dom } L \times U$. Let $u_* = \lim_{k \rightarrow \infty} u_k$. By unique solvability there is $y_* \in C_0(\overline{\Omega})$ for which $Ly_* = u_*$. Moreover, by (3)

$$\lim_{k \rightarrow \infty} \|y_k - y_*\|_\infty \leq \lim_{k \rightarrow \infty} c \|u_k - u_*\|_{L_2} = 0$$

which shows closedness of L . By density of $C_0^\infty(\Omega)$ in $C_0(\overline{\Omega})$, $\text{dom } L \supset C_0^\infty$ is dense in $C_0(\overline{\Omega})$. We stress that it was not necessary for this argumentation to specify $\text{dom } L$ exactly. In particular, we are not restricted to $\text{dom } L = H^2(\Omega) \cap H_0^1(\Omega)$.

Concerning the adjoint operator, the definition of $\text{dom } L^* \subset U^* = L_2(\Omega)$ now reads

$$l \in \text{dom } L^* \quad \Leftrightarrow \quad \sup_{v \in \text{dom } L} \frac{|\int_\Omega l(-\Delta v) dt|}{\|v\|_\infty} < \infty.$$

Then L^*l is defined as the unique continuous extension of $v \rightarrow \int_\Omega l(-\Delta v) dt$ from $\text{dom } L$ to $C_0(\overline{\Omega})$. Structural results concerning $\text{dom } L^*$, in particular that $l \in \text{dom } L^*$ is a function that vanishes on $\partial\Omega$, are derived in [3, Theorem 4] under additional regularity conditions on Ω .

Estimates of the form $\|y\|_\infty \leq c \|u\|_{L_2}$ are available for a variety of problems with partial differential equations. Closedness of the corresponding differential operators can then be established analogously.

Inequality Constraints. The inequality constraints in (1) are interpreted to hold pointwise almost everywhere. We assume that there is a *strictly feasible point* $\check{x} = (\check{y}, \check{u})$, which satisfies $L\check{y} - \check{u} = 0$ and

$$0 < d_{\min} := \text{ess inf}_{t \in \overline{\Omega}} \min\{\check{u}(t) - \underline{u}(t), \overline{u}(t) - \check{u}(t), \check{y}(t) - \underline{y}(t), \overline{y}(t) - \check{y}(t)\}. \quad (4)$$

We call such a condition a (uniform pointwise) *Slater condition* and \check{x} a Slater point. This condition is essential for the structure of dual variables and subdifferentials. If some of the bounds or the variables are only defined on subsets of $\overline{\Omega}$, then this assumption has to be modified accordingly. Some of the inequality constraints may not be present. Throughout the paper we will not use any boundedness assumptions that might hold due to inequality constraints.

By our assumptions the feasible set $G = G_Y \times G_U \subset Y \times U$ defined by the inequality constraints is non-empty, closed, and convex. By our choice of topology, namely $\|\cdot\|_Y = \|\cdot\|_\infty$ and $\|\cdot\|_U = \|\cdot\|_{L_2}$, G_U has empty interior while the interior of G_Y is non-empty.

Combined Functionals. It is a popular strategy in convex analysis to combine the functional J and the constraints $x \in E$ and $x \in G$ to a single functional. This is done via indicator functions. The *indicator function* $\chi_C(x)$ of a set $C \subset X$ is defined by

$$\chi_C(x) := \begin{cases} 0 & : x \in C \\ \infty & : \text{otherwise.} \end{cases}$$

If C is non-empty, convex, and closed, then χ_C is a proper ($\chi_C \not\equiv +\infty$), convex, and lower-semi-continuous function. In particular, χ_E and χ_G enjoy these properties.

With the help of indicator functions we can rewrite (1) as an unconstrained minimization problem defined by the following functional:

$$\begin{aligned} F : X &\rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \\ F &:= J + \chi_E + \chi_G. \end{aligned} \tag{5}$$

By our assumptions F is a proper, lower semi-continuous, convex, and coercive functional and does thus admit a minimizer by weak compactness of E . For use in Section 5 we impose the following *strong convexity condition* on F :

$$\exists \alpha > 0 : \alpha \|x_1 - x_2\|^2 \leq F(x_1) + F(x_2) - 2F\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) \quad \forall x_1, x_2 \in X. \tag{6}$$

Such a condition holds for optimal control problems with regularization term and allows the passage from convergence of function values to norm convergence of the corresponding arguments. We stress that we impose this condition on F and not on J .

3 Barrier Functionals and their Subdifferentials

This section is devoted to the analysis of barrier functionals. Standard calculus results do not apply directly because barrier functionals are unbounded in the regions of interest. Rather than considering the classical logarithmic barrier function only we will consider the following class of barrier functions. This generality is useful to address phenomena that occur in function space only (cf. Section 6). In finite dimensional spaces a comprehensive study of classes of barrier functions has been performed in [6].

Definition 3.1. For all $q \geq 1$ and $\mu > 0$ the functions $l(x; \mu; q) : \mathbb{R}_+ \rightarrow \overline{\mathbb{R}}$ defined by

$$l(x; \mu; q) := \begin{cases} -\mu \ln(x) & : q = 1 \\ \frac{\mu^q}{(q-1)x^{q-1}} & : q > 1 \end{cases}$$

are called *barrier functions of order q* . We extend their domain of definition to \mathbb{R} by setting $l(x; \mu; q) = \infty$ for $x \leq 0$.

In many cases we do not have to consider special values of q or μ . In these cases we may abbreviate the notation $l(x; \mu; q)$ by $l(x; \mu)$ or even $l(x)$.

Obviously, all l are convex and monotonically decreasing functions, which are continuously differentiable in $]0, \infty[$ and bounded from above on each closed positive interval. Moreover, $\lim_{x \rightarrow 0} l(x) = \infty$ and all l are bounded from below by a linear function (for $q > 1$ they are even positive).

Our theory will not depend so much on the properties of l , but on the properties of their first derivatives. These can be computed as

$$l'(x; \mu; q) = -\frac{\mu^q}{x^q}.$$

All l' are negative on $]0; \infty[$ and $\lim_{x \rightarrow 0} l'(x) = -\infty$.

Using these barrier *functions* $l(x; \mu; q)$ we construct barrier *functionals* $b(x; \mu; q)$ to implement constraints of the form $x \geq 0$ on a measurable set $B \subset \overline{\Omega}$ by computing the integral over l :

$$\begin{aligned} b(\cdot; \mu; q) : L_p(B) &\rightarrow \overline{\mathbb{R}} & 1 \leq p < \infty \\ x &\mapsto \int_B l(x(t); \mu; q) dt. \end{aligned}$$

We are very flexible in the choice of B , for example $B = \Omega$ or $B = \partial\Omega$, as long as we can construct an appropriate integral on B . The notion “almost everywhere” is defined with respect to the choice of the corresponding measure. In the case $p = \infty$ we assume B to be compact and define the domain space of b to be $C(B)$ rather than $L_\infty(B)$ because the dual space $C(B)^*$ enjoys better structural properties than $L_\infty(B)^*$.

To incorporate bounds like $y \geq \overline{y}$ and $y \leq \underline{y}$, we insert $x := \overline{y} - y$ and $x := y - \underline{y}$ (and analogously for u) into the barrier functional. Bilateral bounds are straightforwardly incorporated by adding the corresponding barrier functions for the unilateral bounds. Of course, the choice of $L_p(B)$ and $C(B)$ as domain space necessitates that the arguments of b , and thus also upper and lower bounds are contained in the corresponding space. We will assume this tacitly.

Lemma 3.2. *Let $V = L_p(B)$ on a measurable set B with $1 \leq p < \infty$ or $V = C(B)$ on a compact set B . The barrier functional $b : V \rightarrow \overline{\mathbb{R}}$ is well-defined, convex, and lower semi-continuous. Moreover,*

$$b(x) < \infty \implies x > 0 \text{ a.e. in } B.$$

Proof. Because similar arguments as the following have been used elsewhere (cf. [14, 18]), we merely sketch the proof of this Lemma. By inclusion it is sufficient to consider $V = L_1(B)$.

First we show that for $x \in L_1(B)$ the function $L(x) : t \rightarrow l(x(t))$ is measurable. Obviously, l can be constructed as the pointwise supremum of Carathéodory functions, e.g., $l(x) = \sup_{n \in \mathbb{N}} l_n(x) := \max\{l(x), n\}$. Inserting x into each of these functions, we obtain a sequence of measurable functions. Their pointwise supremum, which is $L(x)$ by construction, is measurable because measurability is stable with respect to taking pointwise suprema.

It follows from the same argumentation that l is lower semi-continuous because lower semi-continuity is also stable with respect to taking pointwise suprema.

Next we show that $b(x)$ is well defined in $\overline{\mathbb{R}}$ for all $x \in L_1(B)$. This holds because l is bounded from below by a linear function f . Moreover, since $(l + f)(x)$ is a positive function, we can apply the Lemma of Fatou to show lower semi-continuity of b . Convexity of b follows immediately from the convexity of l . \square

By the relation $l'(x; q) = -q/\mu \cdot l(x; q + 1)$ we infer that $l'(x; q)$ is a measurable function, too.

Subdifferentiability of Barrier Functionals. The most important properties of barrier methods are due to their first derivatives. A popular notion of a derivative in convex analysis is that of a subdifferential (cf. [4, Section I.5]).

Recall that the subdifferential $\partial b(x)$ of $b : V \rightarrow \overline{\mathbb{R}}$ at x is the set of all continuous linear functionals $m \in V^*$ that satisfy $\langle m, (v - x) \rangle \leq b(v) - b(x)$ for all $v \in V$. In the case $b(x) = \infty$ one defines $\partial b(x) = \emptyset$. Subdifferentials of barrier functionals that incorporate upper bounds (like $y \leq \overline{y}$) and lower bounds (like $y \geq \underline{y}$) are computed straightforwardly as $-\partial b(\overline{y} - y)$ and $\partial b(y - \underline{y})$, respectively.

The consideration of the same function in two different spaces V_1 and V_2 may lead to different subdifferentials in dual spaces V_1^* and V_2^* . In our case it will turn out that the subdifferential $\partial b(x)$ of b at x in $V = L_p(B)$ is well behaved in the sense that it contains at most one element in $(L_p)^*$. It is (if it exists) then given by the formal derivative $b'(x)$ of b at x defined as

$$\langle b'(x; \mu; q), \delta x \rangle := \int_B l'(x; \mu; q) \delta x \, dt = - \int_B \frac{\mu^q}{x^q} \delta x \, dt. \quad (7)$$

Observe that (7) is not well defined in general, but only if $l'(x; \mu; q) \delta x \in L_1(B)$.

For $V = C(B)$ – the case that appears in the presence of state constraints – the situation is more delicate because $C(B)^*$ cannot be represented as a function space, but only as a space of regular measures.

We approach subdifferentiability of $b(x)$ via directional derivatives $b'(x; \delta x)$ for directions δx .

Lemma 3.3 (Directional Differentiability). *Consider $b : V \rightarrow \overline{\mathbb{R}}$ for $V = L_p(B)$ on a measurable set B with $1 \leq p < \infty$ or $V = C(B)$ on a compact set B . Let $x, \delta x \in V$ and assume that $b(x)$, $b(x + \delta x)$, and $\langle b'(x), \delta x \rangle$ are finite. Then b is directionally differentiable at x in the direction δx and*

$$b'(x; \delta x) = \langle b'(x), \delta x \rangle \geq \langle m, \delta x \rangle \quad \forall m \in \partial b(x). \quad (8)$$

Proof. To show that $\langle b'(x), \delta x \rangle$ is the directional derivative of $b(x)$ in direction δx , we have to show that h -finite differences of b converge to $\langle b'(x), \delta x \rangle$ for $h \rightarrow 0$.

By convexity $l'(x)\delta x \leq (l(x+h\delta x) - l(x))/h$ holds pointwise almost everywhere and $(l(x+h\delta x) - l(x))/h$ is monotonically increasing in h . Since $b(x)$, $b(x+\delta x)$, and $\langle b'(x), \delta x \rangle$ are finite, the function

$$(\delta x, t) \rightarrow r(\delta x, t; h) := \frac{(l(x(t) + h\delta x) - l(x(t)))}{h} - l'(x(t))\delta x$$

is well defined and non-negative almost everywhere and an in $L_1(B)$ for all $h \in]0; 1]$. Moreover, by monotonicity of the difference quotient $r(\delta x, t; h)$ is dominated by $r(\delta x, t; 1)$.

Since $l(x)$ is differentiable with respect to x a.e., $\lim_{h \rightarrow 0} r(\delta x, t; h) = 0$ pointwise almost everywhere. Hence, we can apply the convergence theorem of Lebesgue to obtain $\lim_{h \rightarrow 0} \int_B r(\delta x, t; h) dt = 0$.

This shows directional differentiability of b and the relation $b'(x; \delta x) = \langle b'(x), \delta x \rangle$. In particular,

$$b(x + h\delta x) - b(x) = \langle b'(x), h\delta x \rangle + o(h).$$

To prove the remaining part of (8) let $m \in X^*$ and $\langle m, \delta x \rangle > \langle b'(x), \delta x \rangle$. Then there is $\varepsilon > 0$ such that

$$\langle m, h\delta x \rangle \geq \langle b'(x), h\delta x \rangle + \varepsilon h \geq b(x + h\delta x) - b(x) - o(h) + \varepsilon h.$$

But this implies $\langle m, h\delta x \rangle > b(x + h\delta x) - b(x)$ for sufficiently small h . This excludes m from the subdifferential and shows (8). \square

Proposition 3.4 (Subdifferentiability in $L_p(B)$). *Consider $b : L_p(B) \rightarrow \overline{\mathbb{R}}$, $1 \leq p < \infty$ on a measurable set B . Then either $\partial b(x) = \emptyset$, or $\partial b(x) = \{b'(x)\}$.*

Proof. Let x be a measurable function on B and a representative of an element of $L_p(B)$. By convexity $l'(x(t))\delta x \leq l(x(t) + \delta x) - l(x(t))$ holds for $x(t) > 0$ and $\delta x \in \mathbb{R}$. If we assume that $b'(x) \in L_p^*$, then by monotonicity of the integral $\langle b'(x), \delta x \rangle \leq b(x + \delta x) - b(x)$ for all $\delta x \in L_p$ and $b'(x) \in \partial b(x)$. This assumption not valid in general. However, since $l'(x)$ is negative, its integral over a measurable subset of B is always well-defined, but possibly $-\infty$.

If $p < \infty$, then $L_p(B)^* \cong L_{p'}(B)$ for $p^{-1} + p'^{-1} = 1$. Consider $m \in L_{p'}(B)$ with $m \neq b'(x)$. We will prove that $m \notin \partial b(x)$. Then from $\partial b(x) \neq \emptyset$ it follows that $\partial b(x) = \{b'(x)\}$ and in particular $b'(x) \in L_p^*(B)$.

If $b(x) = \infty$, then $\partial b(x) = \emptyset$ and we are done. Hence, we may assume $x > 0$ and thus $l'(x) > -\infty$ almost everywhere in B . Moreover, $m \neq b'(x)$ on a set of non-zero measure.

Thus, there is a non-zero set $S \subset B$ such that $l'(x(t)) > -\infty$ for $t \in S$ and $\int_S b'(x) dt \neq \int_S m dt$. Define $S_d := \{t \in S : x(t) \geq d\}$ and denote by ι_{S_d} its characteristic function. Then $S = \bigcup_{k \in \mathbb{N}} S_{1/k}$ and $S_{1/(k-1)} \subset S_{1/k}$. Hence,

$$\lim_{k \rightarrow \infty} \int_S \iota_{S_{1/k}} m dt = \int_S m dt$$

by the convergence theorem of Lebesgue (using $m \in L_{p'}(S)$) and

$$\lim_{k \rightarrow \infty} \int_S \iota_{S_{1/k}} l'(x) dt = \int_S l'(x) dt$$

by the convergence theorem of Beppo-Levi (using positivity of $-l'(x)$ and monotonicity of the sequence).

Consequently there is $\varepsilon > 0$ such that $\sup_{t \in S_{2\varepsilon}} |l'(x(t))| < \infty$ and either $\delta x = \iota_{S_{2\varepsilon}}$ or $\delta x = -\iota_{S_{2\varepsilon}}$ yield

$$-\infty < \langle b'(x), \delta x \rangle < \langle m, \delta x \rangle \quad (9)$$

and $b(x + \varepsilon \delta x) < \infty$.

Therefore, by Lemma 3.3 b is directionally differentiable in the direction $\varepsilon \delta x$ and $\langle b'(x), \varepsilon \delta x \rangle \geq \langle \tilde{m}, \varepsilon \delta x \rangle$ for all $\tilde{m} \in \partial b(x)$. Consequently, due to (9) $m \notin \partial b(x)$. Since $m \in L_{p'}(B)$ was arbitrary, this completes the proof. \square

Proposition 3.4 can be extended straightforwardly to sums of barrier functionals used to incorporate bilateral bounds as long as their minimal distance is positive (cf. (4)).

At a crucial point in the proof we used $L_p(B)^* \cong L_{p'}(B)$ to conclude that $m \neq b'(x)$ on a non-zero set. Here we used the *absolute continuity* of integrable functions with respect to the underlying measure.

When we consider b on $C(B)$ (for a compact set B) we cannot use this structure. The dual space $C(B)^*$ is given by the space of regular measures $M(B)$, which are *not* necessarily absolutely continuous. Thus, the subdifferential is not necessarily unique. However, by σ -*additivity* of regular measures we can restrict non-uniqueness to that subset of B where x touches the bound.

Proposition 3.5 (Subdifferentiability in $C(B)$). *Consider $b : C(B) \rightarrow \overline{\mathbb{R}}$ on a compact set B and assume $\emptyset \neq \partial b(x) \subset M(B)$. Then on the set of strictly feasible points $S := \{t \in B : x(t) > 0\}$ we have*

$$m|_S = b'(x)|_S \quad \forall m \in \partial b(x). \quad (10)$$

In particular, $\partial b(x) \cap L_1(B) = \{b'(x)\}$. Moreover,

$$\langle m, \delta x \rangle \leq \langle b'(x), \delta x \rangle \leq 0 \quad \forall 0 \leq \delta x \in C(B) \quad (11)$$

and

$$\|b'(x)\|_{L_1(B)} = \min_{m \in \partial b(x)} \|m\|_{M(B)}. \quad (12)$$

Proof. Since every $m \in \partial b(x)$ is a measure on B we can write for each measurable subset $M \subset B$

$$\langle m|_M, \delta x \rangle = \int_M \delta x dm.$$

Consider the set $S_d = \{t \in B : x(t) \geq d\}$, which is closed in B by continuity of x . Let δx be an arbitrary element of $C(S_d)$. Then $x + h\delta x$ is strictly feasible on S_d

for sufficiently small h . By continuity x is still strictly feasible on a neighbourhood of S_d relative to B and we can extend $h\delta x \in C(S_d)$ to a function $\delta\tilde{x} \in C(B)$ such that $b(x + \delta\tilde{x})$ and $\langle b'(x), \delta\tilde{x} \rangle$ are finite.

Hence, the inequality in (8) holds for all $\delta x \in C(S_d)$. By changing the sign of δx it becomes an equality and (10) follows on S_d . Since $S = \bigcup_{k \in \mathbb{N}} S_{1/k}$ and m is σ -additive, (10) follows on S .

Because $b(x) < \infty$, $B \setminus S$ has zero measure and hence, (recalling that $l'(x)$ is negative)

$$0 \geq \int_B l'(x) dt = \int_S l'(x) dt = \int_S dm > -\infty.$$

Therefore, $b'(x) \in L_1(B)$ and thus $b'(x) \in \partial b(x)$. In particular, we can apply (8) for all $\delta x \geq 0$ (then $b(x + \delta x) \leq b(x) < \infty$), which yields (11).

To show (12) we use (11) and the positivity of $-b'(x)$ to compute for $m \in \partial b(x)$:

$$\begin{aligned} \|m\|_{M(B)} &= \sup_{\|\delta x\|_\infty \leq 1} \langle -m, \delta x \rangle \geq \sup_{\delta x \geq 0, \|\delta x\|_\infty \leq 1} \langle -m, \delta x \rangle \\ &\geq \sup_{\delta x \geq 0, \|\delta x\|_\infty \leq 1} \langle -b'(x), \delta x \rangle = \sup_{\|\delta x\|_\infty \leq 1} \langle -b'(x), \delta x \rangle = \|b'(x)\|_{L_1(B)}. \end{aligned}$$

□

For our last, easy observation we use again the set $S_d := \{t \in B : x(t) \geq d\}$. The next corollary states that $m \in \partial b(x)$ implies that $m|_{S_d}$ is asymptotically negligible for $\mu \rightarrow 0$. This means that upper and lower bounds *decouple* for $\mu \rightarrow 0$ as long as there is a Slater point.

Corollary 3.6. *For a barrier functional $b(x; \mu; q)$ and $m \in \partial b(x; \mu; q)$ we have $m|_{S_d} = b'(x; \mu; q)|_{S_d}$ and*

$$\|b'(x; \mu; q)\|_{L_\infty(S_d)} \leq c \left(\frac{\mu}{d}\right)^q. \quad (13)$$

Proof. By definition $x \leq d$ on S_d and $S_d \subset S$ as used in (10). The assertion (13) follows now directly from the definition of $b'(x)$. □

Subdifferentiability in $C_0(B; O)$. Cases like $Y = C_0(\overline{\Omega})$ (continuous functions that vanish on the boundary) are not yet covered by our considerations. We will now sketch how the argumentation carries over for such cases. Let O be a closed subset of the compact set B . Define

$$C_0(B; O) := \{x \in C(B) : y(t) = 0 \forall t \in O\}.$$

This is a closed subspace of $C(B)$ and its dual can be represented as $M(B \setminus O)$. Just as in the case $C(B)$ we may insert $x \in C_0(B; O)$ into barrier functionals $b(x - \underline{x})$ or $b(\overline{x} - x)$. In general \underline{x} and \overline{x} are not in $C_0(B; O)$, but only in $C(B)$. However, $b(\cdot - \underline{x})$ and $b(\overline{x} - \cdot)$ are convex functionals on $C_0(B; O)$.

With these preparations Proposition 3.5 carries over straightforwardly, if we replace $C(B)$ by $C_0(B; O)$ and $M(B)$, $L_1(B)$ by $M(B \setminus O)$, $L_1(B \setminus O)$, respectively. Moreover, we have to modify the definition of set S of strictly feasible points, taking into account the shift by \bar{x} or \underline{x} , e.g., $S := \{t \in B : x > \underline{x}\}$. Then (10) reads

$$m|_{S \setminus O} = b'(x)|_{S \setminus O} \quad \forall m \in \partial b(x).$$

Similar modifications are necessary in the proof. Analogous to S we have to modify the definition of S_d by taking into account the shifts. Then we have $\delta x \in C_0(S_d; O)$ and hence, fewer directions. This, however, fits to the representation of the dual space as $M(B \setminus O)$.

4 Minimizers of Barrier Problems and their Optimality Conditions

Having studied the properties of barrier functionals we now add them to a convex functional F of the form (5) as defined in Section 2. We want to incorporate upper and lower bounds and thus, we have to add several barrier functionals, the sum of which we denote (with slightly sloppy notation) by $b(x; \mu)$. We only consider pure state and control constraints and write $b(y; \mu)$ and $b(u; \mu)$ for those barrier functionals corresponding to the state and the control constraints, respectively. Moreover, we distinguish by $\bar{b}(\cdot; \mu)$ and $\underline{b}(\cdot; \mu)$ barrier functionals that incorporate upper and lower bounds, respectively.

Our main concern will be the interaction of the several barrier functionals. Corollary 3.6 is a first hint that the situation may be comparable to the original problem.

Adding barrier functionals to F we obtain another convex functional F_μ defined by

$$\begin{aligned} F_\mu(x) &:= F(x) + b(x; \mu) = J(x) + \chi_E(x) + \chi_G(x) + b(x; \mu) \\ &= J(x) + \chi_E(x) + b(x; \mu). \end{aligned} \tag{14}$$

The last equality holds because $\chi_G = 0$ in G and $\chi_G = b = +\infty$ in $X \setminus G$. Our definition implies $F_0 = F$. Since $l(x; \mu; q) > 0$ for $q > 1$, coercivity of F_μ follows from coercivity of F in this case. If $q = 1$ we have to assume that F_μ is coercive for some μ .

Theorem 4.1 (Solutions of Barrier Problems). *Let $F : X \rightarrow \overline{\mathbb{R}}$ be a proper, convex, lower semi-continuous, coercive functional of the form (5), where E is a weakly sequentially compact subspace of X . Assume that F_{μ_0} is coercive for some $\mu_0 > 0$.*

Then (14) admits a unique minimizer $x(\mu) = (u(\mu), y(\mu))$ for each $\mu \in]0; \mu_0]$. Moreover, $x(\mu)$ is strictly feasible almost everywhere in Ω and bounded in X uniformly in $\mu \in [0, \mu_0]$.

Proof. For each $\mu \in]0, \mu_0]$ we can write $F_\mu(x) = (1 - \mu^q/\mu_0^q)F(x) + \mu^q/\mu_0^q F_{\mu_0}(x)$ and hence

$$\min \{F(x), F_{\mu_0}(x)\} \leq F_\mu(x) \leq \max \{F(x), F_{\mu_0}(x)\}. \quad (15)$$

It follows that, since F and F_{μ_0} are coercive, all F_μ with $\mu \in [0; \mu_0]$ are coercive and all their level-sets are uniformly bounded in X . Since b is strictly convex and lower semi-continuous, and by (4), F_μ inherits strict convexity, lower semi-continuity and properness from F and b . We can thus apply the main existence theorem for minimizers in convex optimization (cf. e.g. [4, Proposition I.1.2]) to obtain existence and uniqueness of a minimizer $x(\mu)$. Moreover, by (15)

$$F_\mu(x(\mu)) \leq F_\mu(x(\mu_0)) \leq \max \{F(x(\mu_0)), F_{\mu_0}(x(\mu_0))\}$$

and hence, $F_\mu(x(\mu))$ are uniformly bounded from above. This and uniform boundedness of the level-sets implies uniform boundedness of $x(\mu)$.

By Lemma 3.2 a barrier minimizer cannot touch the bounds on a non-zero set. Otherwise, the barrier functional would be $+\infty$. \square

Next we study first order optimality conditions for barrier problems. For this purpose we first have to study the subdifferential of χ_E , the characteristic function for the equality constraints $Ly - u = 0$.

Proposition 4.2. *The subdifferential $\partial\chi_E(x)$ at a point $x \in E$ is given by the closed subspace of all $(\nu, \lambda) \in Y^* \times \text{dom } L^* \subset Y^* \times U^*$ that satisfy $L^*\lambda + \nu = 0$ in Y^* .*

Proof. The subdifferential $\partial\chi_E$ of χ_E at some $x \in E$ consist of all $l \in X^*$, say $l = (\nu, \lambda) \in Y^* \times U^*$ that satisfy $\langle l, \delta x \rangle = 0$ for all $\delta x \in E$. Hence,

$$0 = \langle \nu, \delta y \rangle + \langle \lambda, \delta u \rangle = \langle \nu, \delta y \rangle + \langle \lambda, L\delta y \rangle \quad \forall (\delta y, \delta u) \in E. \quad (16)$$

If $\lambda \notin \text{dom } L^*$, then there is a bounded sequence $\delta y_k \in \text{dom } L$ such that

$$\lim_{k \rightarrow \infty} |\langle \lambda, L\delta y_k \rangle| = \infty$$

while $\langle \nu, \delta y_k \rangle$ remains bounded due to $\nu \in Y^*$. This contradicts (16) and we conclude $\lambda \in \text{dom } L^*$. Consequently, (2) can be applied to $\langle \lambda, L\delta y \rangle$ and $\langle L^*\lambda + \nu, \delta y \rangle = 0$ follows for all $\delta y \in \text{dom } L$. By density this relation extends to $L^*\lambda + \nu = 0$ in Y^* .

If in converse $(\nu, \lambda) \in Y^* \times \text{dom } L^*$ and $L^*\lambda + \nu = 0$ in Y^* , then by (2) and (16) we directly conclude $(\lambda, \nu) \in \partial\chi_E$.

Obviously $L^*\lambda + \nu = 0$ defines a linear space and by [4, Corollary I.5.1] subdifferentials are always closed \square

Theorem 4.3 (First Order Optimality Conditions). *Assume that the general assumptions of Section 2 hold. For $\mu \geq 0$ let x be the unique minimizer of F_μ .*

Then there are $(j_y, j_u) = j \in \partial J(x)$, $m \in \partial b(y; \mu) \subset Y^*$ and $\lambda \in \text{dom } L^*$ such that

$$\begin{aligned} j_y + m + L^* \lambda &= 0 \\ j_u + b'(u; \mu) - \lambda &= 0 \\ Ly - u &= 0 \end{aligned} \tag{17}$$

holds. If y is strictly feasible, then $\partial b(y; \mu) = \{b'(y; \mu)\}$ and m is unique. If $\partial J(x)$ contains only one element and $\mu > 0$, then m and λ are unique in Y^* and U^* , respectively.

Proof. First of all, the third equation of (17) holds because y and u satisfy the state equation.

To show that the first two equations hold we apply techniques of subdifferential calculus. We set $(y, u) = x$ and consider

$$F_\mu(x) := F(x) + b(x; \mu) = J(x) + \chi_E(x) + b(x; \mu).$$

Let x be a minimizer of F_μ , which exists due to Theorem 4.1. Then $0 \in \partial(J + \chi_E + b)$. To show that (17) has a solution we have to apply the sum-rule twice:

$$0 \in \partial(J + \chi_E + b) = \partial J + \partial(\chi_E + b) = \partial J + \partial\chi_E + \partial b.$$

To be able to apply the sum-rule to a sum $f + g$ the summands have to be convex, lower semi-continuous, and satisfy an additional regularity condition, such as the following (cf. e.g. [2, Theorem 4.3.3]):

$$0 \in \text{int}(\text{dom } f - \text{dom } g). \tag{18}$$

The difference of two sets $V_1, V_2 \in X$ is defined here by

$$V_1 - V_2 := \{x \in X : \exists x_1 \in V_1, x_2 \in V_2 : x_1 - x_2 = x\}.$$

Let now B_V be the unit ball in a normed space V . We observe that showing (18) is equivalent to showing that there is $\varepsilon > 0$ such that each $x \in \varepsilon B_X$ can be written as a difference $x_1 - x_2$ with $x_1 \in \text{dom } f$ and $x_2 \in \text{dom } g$.

By (4) there exists a feasible point $\check{x} = (\check{y}, \check{u})$ with positive minimal distance to the bounds, which implies in particular that $\check{x} \in \text{dom } F_\mu$. Our assumptions on J include continuity at \check{x} and hence boundedness in some ball $\check{x} + \varepsilon B_X$. Since $\check{x} \in \text{dom}(\chi_E + b)$ and

$$\varepsilon B_X = (\check{x} + \varepsilon B_X) - \check{x} \subset \text{dom } J - \text{dom}(b + \chi_E)$$

we conclude that (18) is fulfilled and thus $\partial(J + \chi_E + b) = \partial J + \partial(\chi_E + b)$.

Next we show that $\partial(\chi_E + b) = \partial\chi_E + \partial b$ by verifying (18) for b and χ_E . Here $Y \subset C(\bar{\Omega})$ is crucial because it guarantees that $(\check{u}, \check{y} + rB_Y) \in \text{dom } b$ for some $r < d_{\min}$ via the Slater condition (4).

Because L is an open mapping (cf. Section 2), there is $\delta > 0$ such that for each $v \in \delta B_U$ we find an $y(v) \in (r/2)B_Y$ with $Ly(v) - v = 0$. Thus we have $(\check{u} + v, \check{y} + y(v)) \in \text{dom } \chi_E$ and $(\check{u}, \check{y} + y(v) + w) \in \text{dom } b$ for all $w \in (r/2)B_Y$.

Consequently, for sufficiently small ε and arbitrary $(v, w) \in \varepsilon B_X$ we have

$$v = (\check{u} + v) - \check{u} \quad (19)$$

$$w = (\check{y} + y(v)) - (\check{y} + y(v) + w) \quad (20)$$

with $(\check{u} + v, \check{y} + y(v)) \in \text{dom } \chi_E$ and $(\check{u}, \check{y} + y(v) + w) \in \text{dom } b$. This finally shows (18) and the sum-rule yields $0 \in \partial J + \partial \chi_E + \partial b$.

This is an inclusion in $Y^* \times U^*$. It implies that there are $(j_y, j_u) \in \partial J(x)$, $(\nu, \lambda) \in \partial \chi_E(x)$, $m \in \partial b(y; \mu)$, and $l \in \partial b(u; \mu)$, such that

$$j_y - \nu + m = 0 \quad (21)$$

$$j_u - \lambda + l = 0 \quad (22)$$

By Proposition 4.2 it follows that $L^* \lambda + \nu = 0$, and $l = b'(u; \mu)$ by Proposition 3.4. This yields the first two equations of (17). If ∂J contains only one element, then λ can be computed uniquely from (22), which yields uniqueness of $\nu = -L^* \lambda$ and via (21) uniqueness of m . If y is strictly feasible, then $m = b'(y; \mu)$ by Proposition 3.5. \square

In contrast to the case $\mu > 0$, uniqueness of the dual variables may not hold for the original problem ($\mu = 0$). In this case $\partial \chi_G(u)$ usually contains more than one element (a convex cone in U^*).

Remark 4.4. By (4) we know that $\check{u} + \check{v} \in \text{dom } b(u)$ if $\|\check{v}\|_\infty < d_{\min}$. This gives us more flexibility in (19), in case that L is *not surjective*, but with closed range, as for example in pure Neumann problems (cf. Remark 2.1). A close look shows that it is sufficient to assume $U = \text{ran } L \oplus L_\infty$, topologically. In case of pure Neumann problems this holds because $U = \text{ran } L \oplus V$, where V is the space of constant functions on Ω .

Next, we study uniform bounds on the dual variables for $\mu \rightarrow 0$. The difficulty is to derive a uniform bound for $\|m\|_{Y^*}$. All other estimates follow then simply by bootstrapping.

Proposition 4.5 (Uniform Bounds on the Dual Variables). *Assume that the general assumptions of Section 2 hold. Then for each $\mu_0 > 0$*

$$\sup_{\mu \in [0; \mu_0[} \|m\|_{Y^*} \leq C.$$

Proof. In the following let (y, u) be the minimizer of the barrier functional F_μ and (\check{y}, \check{u}) a Slater point. We multiply the first equation of (17) by $\delta y := y - \check{y}$ and the second equation by $\delta u := u - \check{u}$ to obtain:

$$\begin{aligned} 0 &= \langle j_y + m + L^* \lambda, \delta y \rangle + \langle j_u + b'(u) - \lambda, \delta u \rangle \\ &= \langle j_y + m, \delta y \rangle + \langle j_u + b'(u), \delta u \rangle + \langle \lambda, L \delta y - \delta u \rangle \end{aligned}$$

Since the pair $(\delta y, \delta u)$ satisfies the state equation, we have $L\delta y - \delta u = 0$ and the equality

$$0 = \langle j_y, \delta y \rangle + \langle j_u, \delta u \rangle + \langle m, \delta y \rangle + \langle b'(u), \delta u \rangle$$

remains. By our uniform boundedness assumption on ∂J and because $\|\delta y\|$ and $\|\delta u\|$ are uniformly bounded in μ , we have

$$|\langle j_y, \delta y \rangle + \langle j_u, \delta u \rangle| \leq C \quad \Rightarrow \quad |\langle m, \delta y \rangle + \langle b'(u), \delta u \rangle| \leq C.$$

Next we show that the absolute values of the summands in the last expression are bounded in terms of the absolute value of the sum itself. For this purpose we introduce for two functions v and w the following partition of $\bar{\Omega}$:

$$\begin{aligned} I(v, w) &:= \{t \in \bar{\Omega} : |v(t) - w(t)| > d_{\min}/2\} \\ A(v, w) &:= \{t \in \bar{\Omega} : |v(t) - w(t)| \leq d_{\min}/2\} = \bar{\Omega} \setminus I(v, w), \end{aligned}$$

with d_{\min} being the minimal distance between \check{x} and the upper and lower bounds as defined in (4).

Application of the sum-rule (using $Y \subset C(\bar{\Omega})$ and the Slater condition) yields the splitting $m = \underline{m} + \bar{m}$ with $\underline{m} \in \partial \underline{b}(y)$ and $\bar{m} \in \partial \bar{b}(y)$. Moreover, we can split $b'(u) = \underline{b}'(u) + \bar{b}'(u)$ by a simple pointwise computation.

According to the result of Corollary 3.6 \underline{m} , \bar{m} , $\underline{b}'(u)$, and $\bar{b}'(u)$ are small on $I(y, \underline{y})$, $I(y, \bar{y})$, $I(u, \underline{u})$, and $I(u, \bar{u})$, respectively. In particular, they can be bounded a-priori by a constant depending only on d_{\min} and some $\mu_0 \geq \mu > 0$. The remaining terms thus still fulfill:

$$\left| \langle \underline{m}, \delta y \rangle_{A(y, \underline{y})} + \langle \bar{m}, \delta y \rangle_{A(y, \bar{y})} + \langle \underline{b}'(u), \delta u \rangle_{A(u, \underline{u})} + \langle \bar{b}'(u), \delta u \rangle_{A(u, \bar{u})} \right| \leq C.$$

It is checked easily that all four duality products are positive. For example, we have:

$$\begin{aligned} \underline{m} \leq 0, \quad \delta y \Big|_{A(y, \underline{y})} &= (y - \check{y}) \Big|_{A(y, \underline{y})} \leq d_{\min}/2 - d_{\min} \leq -d_{\min}/2 \\ \implies \langle \underline{m}, y - \check{y} \rangle_{A(y, \underline{y})} &\geq 0. \end{aligned}$$

Hence, we can conclude:

$$\left| \langle \underline{m}, \delta y \rangle_{A(y, \underline{y})} \right| + \left| \langle \bar{m}, \delta y \rangle_{A(y, \bar{y})} \right| + \left| \langle \underline{b}'(u), \delta u \rangle_{A(u, \underline{u})} \right| + \left| \langle \bar{b}'(u), \delta u \rangle_{A(u, \bar{u})} \right| \leq C.$$

Now we prove bounds on the norms of \underline{m} , and \bar{m} , and thus on $m = \underline{m} + \bar{m}$. We perform the proof only for \underline{m} . The other case is identical.

As already noted, $\underline{m} \leq 0$ as a measure, and $\delta y < -d_{\min}/2$ on A . Hence,

$$\left| \langle \underline{m}, \delta y \rangle_{A(y, \underline{y})} \right| \geq \left| \langle \underline{m}, 1 \rangle_{A(y, \underline{y})} \right| \cdot \operatorname{ess\,inf}_{t \in A(y, \underline{y})} |\delta y(t)| \geq \|\underline{m}\|_{M(A(y, \underline{y}))} d_{\min}/2,$$

and thus

$$\|\underline{m}\|_{M(A(y, \underline{y}))} \leq 2/d_{\min} \left| \langle \underline{m}, \delta y \rangle_{A(y, \underline{y})} \right| \leq 2C/d_{\min}.$$

We have already noted that $\|\underline{m}\|_{M(I(y, \underline{y}))} \leq C$ and hence, $\|\underline{m}\|_{M(\bar{\Omega})} \leq C$ follows. \square

Corollary 4.6. *If b is a barrier function of order q corresponding to the constraint $y \geq 0$, then the following bound holds independently of μ for a minimizer y of the barrier problem:*

$$\left\| \left(\frac{\mu}{y} \right)^r \right\|_{L_1(\Omega)} \leq C \quad \forall 0 \leq r \leq q. \quad (23)$$

Proof. The case $r = 0$ is trivial. Since Ω is bounded, this estimate can be derived for $r > 0$ from the following chain of inequalities, the minimization property (12) and Proposition 4.5:

$$\left\| \left(\frac{\mu}{y} \right)^r \right\|_{L_1(\Omega)}^{1/r} = \left\| \frac{\mu}{y} \right\|_{L_r(\Omega)} \leq C \left\| \frac{\mu}{y} \right\|_{L_q(\Omega)} = C \left\| \left(\frac{\mu}{y} \right)^q \right\|_{L_1(\Omega)}^{1/q} \leq C.$$

□

Regularity of the Adjoint State. An analysis of the adjoint operator L^* and its domain usually reveals additional boundedness and regularity results for λ and $b'(u)$. Of course, these results depend on L and thus on the underlying partial differential equation. We state a simple abstract prototype theorem.

Corollary 4.7. *Assume that there is a normed space $W \supset \text{dom } L^*$, continuously embedded in U^* , such that $L^* : \text{dom } L^* \rightarrow Y^*$ has the following property:*

$$\|L^*v\|_{Y^*} \geq \|v\|_W. \quad (24)$$

Then $\|\lambda\|_W$ and $\|b'(u)\|_{U^}$ are uniformly bounded for $\mu \rightarrow 0$.*

Proof. This is a simple bootstrapping argument. Inserting the results of Proposition 4.5 into the first equation of (17), $\|L^*\lambda\|_{Y^*}$ is uniformly bounded for $\mu \rightarrow 0$. Uniform boundedness of $\|\lambda\|_W$ follows directly from (24). Inserting this into the second equation of (17) we obtain our last assertion. □

Regularity of the adjoint state λ (which is in our framework represented by the space W) has been considered in a couple of publications. For example, if L is an elliptic differential operator, then W is a Sobolev space $W^{1,p}(\Omega)$ with p depending on the dimension of the domain Ω (cf. e.g. [3]). Regularity issues will become important in Section 6.

5 Length and Lipschitz Continuity of the Central Path

Now we study convergence and continuity properties of the central path. Similar estimates are available for control constrained problems and under stronger smoothness assumptions on the data (cf. e.g. [20, 21, 18]). The corresponding approaches usually rely on the implicit function theorem, which breaks down in the case of measure valued subdifferentials.

In this section we use an alternative two stage approach, based mainly on the properties of ∂F_μ . First we study convergence and continuity of the function values on the central path for $\mu \rightarrow 0$ and show in particular that the error in the function values is $O(\mu)$. Then we use the strong convexity assumption (6) to transfer these results to convergence and continuity estimates on (u, y) .

Equipped with the insight of the preceding section, namely that barrier functionals are essentially decoupled for small μ , we concentrate in our proofs – without loss of generality and for notational reasons only – on a single barrier functional, corresponding to the constraint $y \geq 0$ on $\bar{\Omega}$. Since the boundedness results of the previous sections are used, we impose again the assumptions of Section 2, in particular a Slater condition. Under these assumptions the assertions of the following theorems also hold in the general state and control constrained case.

Lemma 5.1. *Assume that the feasible set contains a point \check{x} that satisfies (4). Let $x(\mu_0)$ be a barrier minimizer for $\mu_0 \geq 0$. Then the following bound holds for every $x \in X$ and every $\mu \in [0; \mu_0]$:*

$$F_\mu(x(\mu_0)) \leq F_\mu(x) + q(\mu_0 - \mu) \int_{y(\mu_0) > y} \frac{\mu_0^{q-1}}{y(\mu_0)^q} (y(\mu_0) - y) dt. \quad (25)$$

In particular, let x_* be the minimizer of F . Then

$$F(x(\mu_0)) \leq F(x_*) + C\mu_0. \quad (26)$$

Proof. In the following we need the relation that $\partial b(y; \mu; q) = \mu^q \partial b(y; 1; q)$. This is a simple consequence of the rules of subdifferential calculus (cf. [4, I(5.21)]) and the fact that $b(y; \mu; q) = \mu^q b(y; 1; q)$.

We may assume that $F_\mu(x) < \infty$ and hence feasibility of x . Otherwise, (25) is trivially fulfilled.

Let $x_0 := x(\mu_0)$ be a minimizer of $F_{\mu_0} := F + b(\cdot; \mu_0)$. By Theorem 4.3 we conclude that there are $m \in \partial b(y_0; 1)$ and $\phi \in \partial F(x_0)$ such that

$$\phi + \mu_0^q \cdot m = 0. \quad (27)$$

Moreover, $\mu^q \cdot m \in \mu^q \cdot \partial b(y_0; 1)$. By the sum-rule (cf. the proof of Theorem 4.3) we conclude

$$\partial F_\mu(x_0) = \partial F(x_0) + \partial b(y_0; \mu) = \partial F(x_0) + \mu^q \cdot \partial b(y_0; 1)$$

and thus $\phi + \mu^q \cdot m = (\mu^q - \mu_0^q) \cdot m \in \partial F_\mu(x_0)$. By convexity of F_μ the following inequality holds for every $v \in \partial F_\mu(x_0)$:

$$F_\mu(x) \geq F_\mu(x_0) + \langle v, x - x_0 \rangle,$$

or equivalently

$$F_\mu(x_0) \leq F_\mu(x) + \langle v, x_0 - x \rangle.$$

Thus, in particular for m as defined in (27) we have

$$F_\mu(x_0) \leq F_\mu(x) + (\mu^q - \mu_0^q) \langle m, y_0 - y \rangle. \quad (28)$$

Since we specialize on lower bounds, m is a negative measure and hence, $-m$ is positive. Then, since $\mu < \mu_0$, we can estimate:

$$(\mu^q - \mu_0^q) \langle m, y_0 - y \rangle = (\mu_0^q - \mu^q) \langle -m, y_0 - y \rangle \leq (\mu_0^q - \mu^q) \langle -m, y_0 - y \rangle \Big|_{y_0 > y}$$

because regions with $y \leq y_0$ do not contribute positively to the integral. Moreover, by convexity of the function μ^q for $q \geq 1$ we have $q\mu_0^{q-1}(\mu - \mu_0) \leq \mu^q - \mu_0^q$. Defining $S := \{t \in \Omega : y_0(t) > y(t)\}$ our considerations have shown

$$(\mu_0^q - \mu^q) \langle -m, y_0 - y \rangle \leq (\mu_0 - \mu)q\mu_0^{q-1} \langle -m, y_0 - y \rangle \Big|_S.$$

By Proposition 3.5 and feasibility of y it follows that (using $S \subset \{t \in \Omega : y_0(t) > 0\}$):

$$\mu_0^q \langle -m, y_0 - y \rangle \Big|_S = \int_S \frac{\mu_0^{q-1}}{y_0^q} (y_0 - y) dt.$$

Inserting this estimate into (28) finally yields (25).

Our computation is still valid for $\mu = 0$ and $x = x_*$, which we now use to show (26). We have

$$\mu_0^q \langle -m, y_0 - y \rangle \Big|_S = \int_S \frac{\mu_0^q}{y_0^q} (y_0 - y) dt \leq \mu_0 \int_S \frac{\mu_0^{q-1}}{y_0^{q-1}} dt \leq C\mu_0.$$

Here we use that $(y_0 - y)/y_0 \leq 1$ and uniform boundedness of the last integral for $\mu \rightarrow 0$ due to Corollary 4.6. Hence, we obtain (26). \square

Lemma 5.2. *If F satisfies the strong convexity condition (6) on X , then F_μ satisfies a growth condition at its minimizer $x(\mu)$:*

$$\alpha \|x - x(\mu)\|^2 + b(x) + b(x(\mu)) - 2b\left(\frac{x + x(\mu)}{2}\right) \leq F_\mu(x) - F_\mu(x(\mu)). \quad (29)$$

In particular, for the minimizer x_ of F we have the estimate*

$$\alpha \|x - x_*\|^2 \leq F(x) - F(x_*) \quad \forall x \in X. \quad (30)$$

Proof. This follows from the computation:

$$\begin{aligned} F_\mu(x) + F_\mu(x(\mu)) - 2F_\mu\left(\frac{1}{2}x + \frac{1}{2}x(\mu)\right) \\ \leq F_\mu(x) + F_\mu(x(\mu)) - 2F_\mu(x(\mu)) = F_\mu(x) - F_\mu(x(\mu)). \end{aligned} \quad (31)$$

Now we take into consideration that $F_\mu = F + b(\cdot; \mu)$ and apply (6) to F while we leave b unchanged in (31). This yields (29). \square

Theorem 5.3 (Convergence of the Central Path). *Denote by $x(\mu)$ the minimizer of F_μ and by x_* the minimizer of F . Under the assumptions of Section 2 we conclude*

$$\|x(\mu) - x_*\| \leq C \sqrt{\frac{\mu}{\alpha}}. \quad (32)$$

Proof. Let $\mu > 0$. Then by Equation (26) and by the growth condition (30) we conclude (32). \square

Lemma 5.4. *For a barrier function $l(x)$ and $x_1 > x_2$ we have the following point-wise estimate on second order finite differences:*

$$l(x_1) + l(x_2) - 2l\left(\frac{x_1 + x_2}{2}\right) \geq \frac{1}{4}l''(x_1)(x_1 - x_2)^2. \quad (33)$$

Proof. Let w.l.o.g. $x_2 > 0$. Otherwise, $l(x_2) = \infty$ and the inequality is trivially fulfilled. Setting $\Delta x := x_2 - x_1$ Taylor expansion of the left hand side in (33) yields:

$$\begin{aligned} & l(x_1) + l(x_2) - 2l\left(\frac{x_1 + x_2}{2}\right) \\ &= l(x_1) + l(x_1) + l'(x_1)\Delta x + r(x_1, \Delta x) - 2l(x_1) - 2l'(x_1)\frac{\Delta x}{2} - 2r(x_1, \Delta x/2) \\ &= r(x_1, \Delta x) - 2r(x_1, \Delta x/2). \end{aligned}$$

By the Lagrange formulation for remainder terms and a simple parameter transformation we obtain

$$\begin{aligned} & r(x_1, \Delta x) - 2r(x_1, \Delta x/2) \\ &= \frac{1}{2} \int_0^1 l''(x_1 + s\Delta x) \Delta x^2 ds - \int_0^1 l''\left(x_1 + \tilde{s}\left(\frac{\Delta x}{2}\right)\right) \left(\frac{\Delta x}{2}\right)^2 d\tilde{s} \\ &= \frac{1}{2} \int_0^1 l''(x_1 + s\Delta x) \Delta x^2 ds - \frac{1}{2} \int_0^{1/2} l''(x_1 + s\Delta x) \Delta x^2 ds \\ &= \frac{1}{2} \int_{1/2}^1 l''(x_1 + s\Delta x) ds \Delta x^2. \end{aligned}$$

To estimate the remaining integral term we exploit the fact that l'' is a positive and monotonically decreasing function and use our assumption that $x_1 > x_2$ to conclude:

$$\int_{1/2}^1 l''(x_1 + s\Delta x) ds \geq \frac{1}{2} \min_{s \in [1/2, 1]} l''(x_1 + s\Delta x) \geq \frac{1}{2} l''(x_1).$$

Collecting all our estimates finally yields

$$l(x_1) + l(x_2) - 2l\left(\frac{x_1 + x_2}{2}\right) = r(x_1, \Delta x) - 2r(x_1, \Delta x/2) \geq \frac{1}{4}l''(x_1)\Delta x^2,$$

which is (33). \square

Theorem 5.5 (Lipschitz Continuity of the Central Path). *If F satisfies the assumptions of Section 2, then the central path is locally Lipschitz continuous. For every $\mu > 0$ we have the estimate*

$$\|x(\mu) - x(\tilde{\mu})\| \leq \frac{c}{\sqrt{\alpha\mu}} |\mu - \tilde{\mu}| \quad \forall \tilde{\mu} \geq 0. \quad (34)$$

Proof. We first will show (34) for $\mu_2 \leq \mu_1 \leq 2\mu_2$ (for $\mu_2 = \mu_1$ there is nothing to prove). We fix the notation $\Delta x := x_1 - x_2 = x(\mu_1) - x(\mu_2)$ and $\Delta\mu := \mu_1 - \mu_2 > 0$ and recall $S = \{t \in \Omega : y_1(t) > y_2(t)\}$. First we observe that by (29) and (25)

$$\begin{aligned} \alpha \|\Delta x\|^2 + b(y_1; \mu_2) + b(y_2; \mu_2) - 2b\left(\frac{y_1 + y_2}{2}; \mu_2\right) &\leq F_{\mu_2}(x_1) - F_{\mu_2}(x_2) \\ &\leq q\Delta\mu \int_S \frac{\mu_1^{q-1}}{y_1^q} \Delta y \, dt = q\Delta\mu \left\| \frac{\mu_1^{q-1}}{y_1^q} \Delta y \right\|_{L_1(S)}. \end{aligned} \quad (35)$$

Next we we apply Lemma 5.4 on S to obtain

$$b(y_1) + b(y_2) - 2b\left(\frac{y_1 + y_2}{2}\right) \geq \frac{1}{4} \int_S b''(y_1) \Delta y^2 \, dt = c \int_S \frac{\mu_2^q}{y_1^{q+1}} \Delta y^2 \, dt. \quad (36)$$

Now we combine (35) and (36) and apply the Hölder-inequality:

$$\begin{aligned} \alpha \|\Delta x\|^2 + \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}^2 \\ \leq c\Delta\mu \left\| \frac{\mu_1^{q-1}}{y_1^q} \Delta y \right\|_{L_1(S)} \leq c\Delta\mu \left\| \frac{\mu_1^{q/2-1}}{y_1^{(q-1)/2}} \right\|_{L_2(S)} \left\| \frac{\mu_1^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}. \end{aligned}$$

Using Corollary 4.6 we estimate

$$\left\| \frac{\mu_1^{q/2-1}}{y_1^{(q-1)/2}} \right\|_{L_2(S)} = \frac{1}{\sqrt{\mu_1}} \left\| \frac{\mu_1^{(q-1)/2}}{y_1^{(q-1)/2}} \right\|_{L_2(S)} = \frac{1}{\sqrt{\mu_1}} \sqrt{\left\| \frac{\mu_1^{q-1}}{y_1^{q-1}} \right\|_{L_1(S)}} \leq \frac{c}{\sqrt{\mu_1}}$$

and thus we obtain

$$\alpha \|\Delta x\|^2 + \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}^2 \leq c \frac{\mu_1 - \mu_2}{\sqrt{\mu_1}} \left\| \frac{\mu_1^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}. \quad (37)$$

Subdividing this inequality by

$$\sqrt{\alpha \|\Delta x\|^2 + \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}^2} \geq \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)} = \frac{\mu_2^{q/2}}{\mu_1} \left\| \frac{\mu_1^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}$$

yields the desired estimate (34) for our special case, taking into account that μ_1/μ_2 is bounded due to our assumption $\mu_1 \leq 2\mu_2$.

For the general case note first that the case $\tilde{\mu} = 0$ is covered by Theorem 5.3. For the remaining case divide $[\tilde{\mu}, \mu]$ into finitely many subintervals $[\mu_{i-1}, \mu_i]$ with $\mu_{i-1} \leq \mu_i \leq 2\mu_{i-1}$ and apply our special case of (34) to each. We obtain the estimate

$$\|x(\mu) - x(\tilde{\mu})\| \leq \sum_i \|x(\mu_i) - x(\mu_{i-1})\| \leq \sum_i \frac{c}{\sqrt{\alpha\mu_i}}(\mu_i - \mu_{i-1}).$$

Because $\mu_i^{-1/2} = \min_{[\mu_{i-1}, \mu_i]} \mu^{-1/2}$, this finite Riemann sum can be estimated from above by an integral, and we obtain (34) via the computation

$$\begin{aligned} \sum_i \frac{c}{\sqrt{\alpha\mu_i}}(\mu_i - \mu_{i-1}) &\leq \frac{c}{\sqrt{\alpha}} \int_{[\tilde{\mu}, \mu]} \frac{1}{\sqrt{m}} dm = \frac{c}{\sqrt{2\alpha}} \left| \sqrt{\mu} - \sqrt{\tilde{\mu}} \right| \\ &= \frac{c}{\sqrt{2\alpha}} \frac{|\mu - \tilde{\mu}|}{\sqrt{\mu} + \sqrt{\tilde{\mu}}} \leq \frac{c}{\sqrt{2\alpha\mu}} |\mu - \tilde{\mu}|. \end{aligned}$$

□

Convergence of the Dual Variables. In contrast to the strong convergence properties of the primal variables the convergence properties of the dual variables are rather poor in general. This is not surprising because the optimality system to the original problem may not even have a unique solution.

However, by the uniform bounds derived in Section 4 and under a uniqueness assumption on the dual variables for the original problem, we can apply standard arguments to show weak* convergence of m in $M(\bar{\Omega})$ and λ in W .

For state constrained problems without control constraints and functionals of the form $J(x) = J(y) + \alpha \|u\|_{L_2}^2$ we easily obtain $\|\lambda(\mu) - \lambda_*\|_{L_2} = O(\sqrt{\alpha\mu})$ via the second equation of the optimality system (17), which yields $\lambda = j_u = 2\alpha u$.

6 Strict Feasibility of the State

As we have seen in the preceding sections, measure valued subdifferentials of barrier functionals cannot be excluded in general in the presence of state constraints. Now we will shed more light on this issue. We will give an example, where such behaviour can actually be observed, and we will discuss a way to avoid measures in subdifferentials.

Comparable issues do not appear for control constraints because subdifferentials of barrier functionals for control constraints are in L_p and hence unique by Proposition 3.4.

An Example with Lack of Strict Feasibility. We present a simple example where the logarithmic barrier method admits a minimizer for each $\mu \geq 0$, but eventually fails to admit $\partial b(y) \equiv b'(y)$ at this minimizer. To formulate an example

as easy as possible we use a single scalar control parameter u and consider the following problem:

$$\min J(u) := -u \quad \text{s.t.} \quad -\Delta y - u = 0, \quad y \leq 1.$$

We use the unit ball in \mathbb{R}^d as the computational domain Ω and impose homogeneous Dirichlet boundary conditions. In this setting y is a quadratic rotationally symmetric function of the form:

$$y = uc_1(1 - r^2), \quad r^2 := |t|^2.$$

Hence, $y \in C^\infty(\overline{\Omega})$ is perfectly regular. Application of the barrier method and elimination of the state y yield:

$$\begin{aligned} J(u; \mu) &= -u - \mu \int_{\Omega} \ln(1 - y(t)) dt \\ &= -u - c\mu \int_{r=0}^1 r^{d-1} \ln(1 - uc_1(1 - r^2)) dr. \end{aligned}$$

By convexity this problem in \mathbb{R} always possesses a minimum. The requirement $y \leq 1$ implies $uc_1 \leq 1$. Formal computation of the gradient of J yields

$$J'(u; \mu) = -1 + c\mu \int_{r=0}^1 r^{d-1} \frac{c_1(1 - r^2)}{1 - uc_1(1 - r^2)} dr.$$

Inserting the extreme value $u = 1/c_1$ we obtain $y(0) = 1$ and the estimate

$$J'(u; \mu) \leq -1 + c\mu \int_{r=0}^1 r^{d-3}(1 - r^2) dr.$$

For $d > 2$ the value of this integral is finite and we conclude for sufficiently small μ :

$$J'(u; \mu) = -1 + C\mu < 0.$$

This shows that there is no feasible solution available for which J' vanishes and thus there is no point that satisfies the *formal* first order optimality conditions. This seeming contradiction is resolved if we take into account that $\partial b(y)$ may contain a positive Dirac measure at the point $t = 0$ where y touches the bound.

Strict Feasibility of the State. Our example shows that violation of strict feasibility of y may lead to a break down of methods that try to solve the *formal* optimality system in spite of Theorem 4.3 guaranteeing the existence of solutions of the optimality system. However, Theorem 4.3 also states that strict feasibility of the state cures this issue. Hence, we will discuss the construction of methods with *strictly feasible* iterates. We will achieve strict feasibility by choosing rational barrier functions of sufficiently high order q .

In the following we assume that the domain Ω satisfies a *cone property* (cf. [1, Definition IV.4.3]), which mainly states that there is a cone C (defined as the convex hull of a ball and a point in \mathbb{R}^d) such that each point t of Ω lies inside a cone $C_t \subset \overline{\Omega}$, which is the image of a rigid motion of C .

Lemma 6.1. *Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain, satisfying a cone property. Let $0 \leq y \in C^\beta(\overline{\Omega})$, $1/y^q \in L_1(\Omega)$, $\beta \leq 2$. Assume that $q \geq d/\beta$. Then $1/y \in C(\overline{\Omega})$.*

Proof. Assume that without loss of generality $0 \in \overline{\Omega}$ and $y(0) = 0$. Since $y \in C^1$ for $\beta \geq 1$ and y necessarily obtains a minimum at 0, we conclude $\nabla y(0) = 0$ for $\beta \geq 1$. By the Hölder-continuity of y or ∇y we infer that $y(t) < cr^{1+\beta}$ for all $t \in B(0, r)$. Using the cone property of Ω , there is a cone $C_0 \in \overline{\Omega}$, and we can compute

$$\|y^{-q}\|_{L_1(\Omega)} \geq \int_{C_0} y(t)^{-q} dt \geq c \int_{[0,R]} |r^{-\beta q}| r^{d-1} dr \geq c \int_{[0,R]} |r^{-1}| dr = \infty,$$

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 there is $\psi > 0$ with $y \geq \psi$ and thus $1/y \in C(\Omega)$. \square

By Proposition 4.5 $\|m\|_{M(\overline{\Omega})} \leq C$ for $\mu \rightarrow 0$. We can insert this information into the optimality system (17) and invoke known regularity results to show that $\|y(\mu)\|_{C^\beta}$ is *uniformly bounded* for $\mu \rightarrow 0$ for some β . For the homogeneous Dirichlet problem for an elliptic PDE on a smoothly bounded domain in \mathbb{R}^d it is shown in [3] that $y \in W^{3,p}$ with $p < d/(d-1)$. By Sobolev embedding theorems [1, Theorem 5.4] we thus conclude $\beta = 2, 2 - \varepsilon, 1 - \varepsilon$ for $d = 1, 2, 3$, respectively.

Inserting these values for β into Lemma 6.1 we obtain strict feasibility of $y(\mu)$ and hence uniqueness of $\partial b(y; \mu) \in L_\infty(\Omega)$ for $q \geq 1, 1 + \varepsilon, 3 + \varepsilon$ for $d = 1, 2, 3$, respectively. If we compute the integrals, then the assertion of Lemma 6.1 can be quantified in the following sense. For sufficiently large q there is a function $\psi(\mu)$ that is *uniformly bounded* on each *positive* compact interval $[\underline{\mu}; \mu_0]$, such that $\min_{t \in \Omega} y(\mu)(t) \geq \psi(\mu)$. More specifically $\psi(\mu)$ must fulfill the following inequality:

$$\int_{[0,R]} \frac{r^{d-1} \mu^q}{(r^\beta + \psi(\mu))^q} dr \leq C.$$

Hence, we obtain $\|1/y(\mu)\|_{L_\infty} \leq \psi(\mu)^{-1}$ and analogously $\|b'(y(\mu))\|_{L_\infty} \leq \tilde{\psi}(\mu)$.

For properly chosen barrier functions the state stays inside the feasible domain and keeps a certain distance from the bounds. This holds uniformly in each closed positive interval.

7 Conclusion and Outlook

Recapitulating our main results we conclude that on the side of the primal variables (y, u) barrier methods can rely on rather *strong convergence and continuity properties* of the central path. Difficulties arise if the state touches the bounds of the feasible set. Then the subdifferential of the barrier functional may contain *measures*. The use of a sufficiently strong barrier functional avoids this situation, exploiting the inherent regularity of the state. In this case the *uniqueness of the dual variables* on the central path is a useful property.

We conjecture that our results can be generalized from linear to semi-linear PDEs as long as appropriate assumptions are imposed. Usually one needs the existence of a compact solution operator of the PDE and second order sufficient conditions at the barrier minimizer.

Based on our insight on the central path the next step will be to construct a Newton path-following scheme and show its convergence to an optimal solution. A simple idea for this has been sketched in [15] together with a proof of convergence. Several variants and extensions are conceivable.

References

- [1] R.A. Adams. *Sobolev Spaces*. Academic Press, 1975.
- [2] J.M. Borwein and Q.J. Zhu. *Techniques of Variational Analysis*. CMS Books in Mathematics. Springer, 2005.
- [3] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.
- [4] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. Number 28 in Classics in Applied Mathematics. SIAM, 1999.
- [5] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1977.
- [6] C. Grossmann and M. Zadlo. A general class of penalty/barrier path-following Newton methods for nonlinear programming. *Optimization*, 54:161–190, 2005.
- [7] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optim.*, 13:865–888, 2002.
- [8] 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. To appear in *SIAM J. Control and Optim.*
- [9] M. Hintermüller and K. Kunisch. Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.*, 17:159–187, 2006.
- [10] C. Meyer, U. Prüfert, and F. Tröltzsch. On two numerical methods for state-constrained elliptic control problems. Technical Report 5-2005, TU Berlin, 2005. Submitted.
- [11] C. Meyer, F. Tröltzsch, and A. Rösch. Optimal control problems of PDEs with regularized pointwise state constraints. *Computational Optimization and Applications*, 33:206–228, 2006.

- [12] 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. to appear at Computational Optimization and Applications.
- [13] W. Rudin. *Functional Analysis*. McGraw-Hill, 1974.
- [14] 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.
- [15] A. Schiela. Convergence of the control reduced interior point method for PDE constrained optimal control with state constraints. ZIB Report 06-16, Zuse Institute Berlin, 2006.
- [16] M. Ulbrich. Nonsmooth Newton-like Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces. Habilitation thesis, Fakultät für Mathematik, Technische Universität München, 2002.
- [17] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. *SIAM J. Control Optim.*, 38(6):1938–1984, 2000.
- [18] M. Ulbrich and S. Ulbrich. Primal-dual interior point methods for PDE-constrained optimization. Technical report, TU Darmstadt, 2006.
- [19] M. Ulbrich, S. Ulbrich, and M. Heinkenschloss. Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds. *SIAM J. Control Optim.*, 37(3):731–764, 1999.
- [20] M. Weiser. *Function Space Complementarity Methods for Optimal Control Problems*. PhD thesis, Free University of Berlin, Dept. Math. and Comp. Sci., 2001.
- [21] M. Weiser. Interior point methods in function space. *SIAM J. Control Optim.*, 44(5):1766–1786, 2005.
- [22] 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.
- [23] M. Weiser and A. Schiela. Function space interior point methods for PDE constrained optimization. *PAMM*, 4(1):43–46, 2004.
- [24] D. Werner. *Funktionalanalysis*. Springer, 3rd edition, 2000.