Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

A. Schiela and W. Wollner

Barrier Methods for Optimal Control Problems with Convex Nonlinear Gradient Constraints [1]

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

ZIB-Report 08-47 (Dezember 2008)

Barrier Methods for Optimal Control Problems with Convex Nonlinear Gradient Constraints †

A. Schiela and W. Wollner‡

December 18, 2008

Abstract

In this paper we are concerned with the application of interior point methods in function space to gradient constrained optimal control problems, governed by partial differential equations. We will derive existence of solutions together with first order optimality conditions. Afterwards we show continuity of the central path, together with convergence rates depending on the interior point parameter.

AMS MSC 2000: 90C51, 49M05

Keywords: interior point method, necessary optimality conditions, convergence of the central path, gradient constrained optimization

1 Introduction

In a large number of processes that are modeled using partial differential equations bounds on the gradient of the state variable are of vital importance for the underlying model: large temperature gradients during cooling or heating processes may lead to destruction of the object, that is being cooled or heated; in elasticity the gradient of the deformation determines the change between elastic and plastic material behavior. In any attempt to optimize such processes the gradient therefore has to be regarded. However, not much attention was given to constraints of gradient type, see [4–7, 11, 25]

Problems with constraints on the state (pointwise or regarding the gradient) form a class of highly nonlinear and non-smooth problems. A popular approach for their efficient solution are path-following methods, which solve a sequence of easier to tackle problems. These methods are constructed in a way such that the sequence of the solutions converges to the solution of the original problem. Among these

†Supported by the DFG Research Center Matheon "Mathematics for key technologies"

‡Institut für Angewandte Mathematik, Ruprecht-Karls-Universität Heidelberg, INF 294, 69120 Heidelberg, Germany (winnifried.wollner@iwr.uni-heidelberg.de)

methods one can distinguish three main lines of research. Lavrentiev regularization methods due to Tröltzsch et al. [8,18,19,24], Moreau-Yoshida approximation methods due to Hintermüller and Kunisch [1,2,16,17] and interior point methods [22,23]. While the first two candidates abandon feasibility to improve the regularity of the dual variables, interior point methods yield feasible solutions and aim towards smooth systems of equations.

Application of interior point methods to gradient bounds has been proposed in [25] together with a posteriori error estimates with respect to the interior point parameter and the discretization error.

In this paper we perform the analysis of the homotopy path generated by barrier methods to problems with gradient bounds. We approach this problem on the base of the analysis in [23], where pointwise state constraints are considered. Although we can build up on techniques and results established there, it will turn out that a number of interesting, additional issues arise in the case of gradient bounds. For example, the topological framework has to be chosen differently with a C^1-norm, and in contrast to pointwise state constraints the gradient bounds considered here are nonlinear.

Our paper is structured as follows. In Section 2 we establish an abstract theoretical framework for our analysis and illustrate the application of the framework to some PDE constrained optimal control problems. In Section 3 we consider barrier functionals for gradient bounds and characterize their subdifferentials. Then existence of minimizers and first order optimality conditions are established, together with uniform bounds on the barrier gradients. Finally we consider the convergence of the path of minimizers and derive an order of convergence for a typical case.

2 Gradient Constrained Optimal Control Problems

Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$, $\emptyset \neq \Omega_C \subseteq \Omega$ be an open subset, and let $\overline{\Omega}_C$ be its closure. Define the space of states U as a closed subspace of $C^1(\overline{\Omega}_C) \times L^2(\Omega \setminus \overline{\Omega}_C)$, which is clearly a Banach space, and let $W \subset U$ be a dense subspace of U. Consider $W = W^{2,p}(\Omega) \subset U = C^1(\overline{\Omega}_C) \times L^2(\Omega \setminus \overline{\Omega}_C)$ with $p > d$ for an example.

Further, consider two reflexive Banach spaces Q and Z, which will denote the space of controls and the space for the adjoint state, respectively. We denote the corresponding dual spaces by U^*, Q^*, and Z^*. Consider the following abstract linear partial differential equation on Ω:

$$Au = Bq \tag{2.1}$$

where we require the following properties:

Assumption 1. *Assume that $A : U \supset \operatorname{dom} A = W \to Z^*$ is densely defined and possesses a bounded inverse. Further let $B : Q \to Z^*$ be a continuous operator.*

We will see later that continuous invertibility of A is equivalent to closedness and bijectivity. The distinction between the state space U and the domain of definition W of A allows us to consider our optimal control problem in a convenient topological framework (the topology of U), while being able to model differential operators by A, which are only defined on a dense subspace W.

To define an optimal control problem, we specify an objective functional J with some basic regularity assumptions:

Assumption 2. *Let $J = J_1 + J_2$. We assume that $J_1 : U \to \mathbb{R}$ and $J_2 : Q \to \mathbb{R}$ are lower semi-continuous, convex and Gâteaux differentiable. In addition let J_1 be bounded from below and J_2 be strictly convex. Assume that the derivatives are uniformly bounded on bounded sets. This means that there exists a continuous $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\|J_1'(u)\|_{U^*} \le g(\|u\|_U)$ and $\|J_2'(q)\|_{Q^*} \le g(\|q\|_Q)$.*

We now consider the following minimization problem

$$\min_{Q^{\mathrm{ad}} \times W} \; J(q,u) = J_1(u) + J_2(q), \tag{2.2a}$$

$$\text{s.t.} \quad Au = Bq, \tag{2.2b}$$

$$\text{and} \quad |\nabla u(x)|^2 \le \psi(x) \text{ on } \overline{\Omega}_C \tag{2.2c}$$

where $\psi \in C(\overline{\Omega}_C)$ with $\psi \ge \delta > 0$ and $Q^{\mathrm{ad}} \subset Q$ closed and convex.

In order to ensure that there exists a solution we require that the following assumption holds

Assumption 3. *We assume that at least one of the following holds:*

(1) *Q^{ad} is bounded in Q.*

(2) *J_2 is coercive on Q.*

For the discussion of interior point methods for the gradient constraint we have to require an additional property, which is of Slater type

Assumption 4. *Assume there exists a feasible control $\check{q} \in Q^{ad}$, such that the corresponding state $\check{u}$ given by $A\check{u} = B\check{q}$ is strictly feasible, that is, $|\nabla \check{u}|^2 < \psi$.*

Lemma 2.1. *Let U be a Banach space. An operator $A : U \supset W \to Z^*$ has a continuous inverse if and only if A is closed and bijective.*

If Assumption 1 holds, then there exists a continuous "control-to-state" mapping

$$S : Q \to U, \quad S := A^{-1}B.$$

Proof. For our first assertion, cf. [22]. By Assumption 1 both A^{-1} and B exist and are continuous, and thus $S := A^{-1}B$, too. □

Using the Assumptions 1–4 it follows by standard arguments (coercivity, weak seq. compactness, convexity) that (2.2) admits a unique solution $(\overline{q},\overline{u}) \in Q^{\mathrm{ad}} \times W$.

For the discussion of the adjoint operator A^* of A we exploit density of W in U and reflexivity of Z. A^* possesses a domain of definition $\operatorname{dom} A^*$, given by

$$\operatorname{dom} A^* = \{z \in Z \,|\, \exists\, c_z \,:\, \langle Au, z\rangle \le c_z \|u\|_U \quad \forall\, u \in \operatorname{dom} A = W\}.$$

Because W is dense in U for each $z \in \operatorname{dom} A^*$ the linear functional $\langle A\,\cdot\,, z\rangle$ has a unique continuous extension to a functional on the whole space U. This defines a linear operator $A^* : Z \supset \operatorname{dom} A^* \to U^*$ and it holds

$$\langle u, A^* z\rangle = \langle Au, z\rangle \quad \forall\, u \in \operatorname{dom} A, z \in \operatorname{dom} A^*.$$

Lemma 2.2. *The operator A^* defined above has a continuous inverse, and it holds*

$$\left(A^{-1}\right)^* = (A^*)^{-1}. \tag{2.3}$$

Proof. Since Z^* is complete and A is surjective, we can apply [14, Theorem II.3.13], which states that A^* has a bounded inverse under these conditions. Hence, both $(A^{-1})*$ and $(A^*)^{-1}$ exist, and by [14, Theorem II.3.9] they are equal. □

2.1 Examples

Let us apply our abstract framework to optimal control problems with PDEs. First we consider two variants of modelling an elliptic partial differential operator of second order: via the strong form and via the weak form. It will turn out that the strong form yields a more convenient representation of A^* and is thus preferable.

Example 2.1. [Second Order Elliptic PDE in Strong Form] *Let $\overline{\Omega}_C = \overline{\Omega} \subset \mathbb{R}^d$, $U = C^1(\overline{\Omega}) \cap H^1_0(\Omega)$, $p > d$, and $Z = L^{p'}(\Omega)$ with $\frac{1}{p} + \frac{1}{p'} = 1$. Consider $A = -\Delta$ as a mapping from $\operatorname{dom} A = W = W^{2,p}(\Omega) \cap H^1_0(\Omega)$ to $L^p(\Omega)$. This means that A is a differential operator in strong form. We can write this as integral equation in the following form:*

$$\langle Au, z\rangle = \int_\Omega -\Delta u z\, dx \quad \forall\, u \in W, z \in Z.$$

Assume that the boundary of $\Omega \subset \mathbb{R}^d$ is either of class $C^{1,1}$ or that $\Omega \subset \mathbb{R}^n$ is convex and has a polygonal (or polyhedral) boundary for $n \le 3$. Then there exists p with $d < p < \infty$ such that A is an isomorphism from W onto Z^*, see, e.g., [13, Theorem 9.15] for the case of a $C^{1,1}$ boundary or [15] for the polygonal case. In particular, A has a continuous inverse from Z^* onto W. By Sobolev embedding W is continuously embedded into U and thus A^{-1} can also be defined as a continuous mapping from Z^* into U. Because W is dense in U the requirements on A from Assumption 1 are fulfilled.

A simple choice for the control space is $Q^{ad} = Q = L^p(\Omega) = Z^*$. Then $B = \mathrm{Id}$ is a continuous operator. This corresponds to distributed control. As a second setting *for the control we may consider* $Q = \mathbb{R}^n$ *and* $f_i \in L^p(\Omega)$, $i = 1 \dots n$. *Then the operator* B *defined by* $Bq = \sum_{i=1}^n f_i q_i$ *satisfies Assumption 1 on* B.

In the case of distributed control a simple cost functional might be

$$J(q,u) = J_1(u) + J_2(q) = \frac{1}{2}\|u - u^d\|^2_{L^2(\Omega)} + \frac{1}{p}\|q\|^p_{L^p(\Omega)}.$$

with given $u^d \in L^2$, $p > d$. *It is easily seen that* J_2 *is coercive on* Q. *Thus Assumption 3 is satisfied. By simple calculations Assumption 2 on* J *is verified.*

Since the gradient bound ψ *is assumed to be strictly positive, taking* $\breve{q} = 0$ *yields the required Slater condition from Assumption 4.*

The adjoint operator $A^* : Z \supset \operatorname{dom} A^* \to U^*$ *can be interpreted as a very weak form of the Laplace operator, i.e.*

$$\langle u, A^* z\rangle = \langle Au, z\rangle = \int_\Omega -\Delta u z\, dx \quad \forall\, u \in W, z \in \operatorname{dom} A^*.$$

Lemma 2.2 already yields the continuous invertibility of A^*.

Example 2.2. [Second Order Elliptic PDE in Weak Form] *Let us discuss an alternative approach to Example 2.1: the weak form of the "same" elliptic operator. Usually one defines the differential operator* $A = -\Delta : H^1_0(\Omega) \to H^{-1}(\Omega)$ *by:*

$$\langle Au, z\rangle = \int_\Omega \nabla u^T \nabla z\, dx \quad \forall\, z \in H^1_0(\Omega).$$

Our aim is to redefine the spaces for this operator such that Assumption 1 holds. To this end we have to restrict the image space from $H^{-1}(\Omega)$ *to* $L^p(\Omega)^*$. *Then the space* W *is given by*

$$W = \left\{ u \in H^1_0 \;\middle|\; \int_\Omega \nabla u^T \nabla z\, dx \le c_u \|z\|_{L^p} \quad \forall\, z \in H^1_0(\Omega) \right\}.$$

Observe that the integral in this expression is not defined for all $z \in L_p$, *but only for* $z \in H^1(\Omega)$. *However, if* $u \in W$ *then by definition of* W *it follows, that* Au *has a unique continuous extension to an element of* $L^p(\Omega)^*$. *It is given canonically by*

$$\langle Au, z\rangle = \lim_{\substack{z_k \in H^1_0,\\ z_k \to z \text{ in } L^p}} \langle Au, z_k\rangle. \tag{2.4}$$

Under the same regularity assumptions as in Example 2.1 we obtain that $W \subset C^1(\overline{\Omega})$ *and* $\|u\|_{C^1} \le c\|Au\|_{(L^p)^*}$, *thus Assumption 1 is fulfilled.*

In spite of the complicated representation of A *via* (2.4), *we may represent the equation* $Au = f$ *conveniently in the form*

$$\int_\Omega \nabla u^T \nabla \varphi\, dx = \int_\Omega f \varphi\, dx \quad \forall \varphi \in H^1_0(\Omega) \tag{2.5}$$

via density.

However, since the linear functional Au is defined in $L^p(\Omega)^*$ by continuous extension (2.4)*, the representation of the adjoint operator A^* is quite cumbersome. It is given by*

$$\langle u, A^* z\rangle = \lim_{\substack{z_k \in H_0^1,\\ z_k \to z \text{ in } L^p}} \langle Au, z_k\rangle = \lim_{\substack{z_k \in H_0^1,\\ z_k \to z \text{ in } L^p}} \int_\Omega \nabla u^T \nabla z_k \, dx.$$

and has to be used in the adjoint PDE. In contrast to the weak formulation of the primal equation (2.5)*, where the limit formulation for the* test functions *can be dropped by density, now the limit formulation applies to elements of the* ansatz space*, and thus cannot be neglected. Continuous invertibility of A^*, which follows from our abstract considerations only applies to its correct representation. A naive formulation of the adjoint PDE would yield wrong results. This is the reason why we prefer the strong formulation for optimal control problems of second order equations with gradient bounds.*

Example 2.3. [Fourth Order Elliptic PDE] *As a different example we consider once again $\overline{\Omega}_C = \overline{\Omega}$ but choose different spaces. Let $U = \{v \in C^1(\overline{\Omega}) \mid v(x) = |\nabla v(x)| = 0 \ \forall x \in \partial\Omega\}$, $Z = W_0^{2,p'}(\Omega)$. We consider the biharmonic operator $A = \Delta^2$ as a mapping from $\operatorname{dom} A = W = W_0^{2,p}(\Omega)$ to $Z^* = W^{-2,p}(\Omega)$ with $\frac{1}{p} + \frac{1}{p'} = 1$.*

Assume that the domain $\Omega \subset \mathbb{R}^2$ is convex with polygonal boundary, then it is well known [3, Theorem 2] that A has a continuous inverse from Z^* onto W. As it has already been remarked for $d < p < \infty$ the embedding from W into U exists and is dense.

Note that in this case both dual and primal operator can be represented by

$$\langle Au, z\rangle = \langle u, A^* z\rangle = \int_\Omega \Delta u \Delta z \, dx \quad \forall\, u \in W_0^{2,p}(\Omega),\, z \in W_0^{2,p'}(\Omega).$$

By the choice $Q = L^2(\Omega)$ with B the embedding from L^2 into $W^{-2,p}$ we see that Assumption 1 is fulfilled.

3 Barrier Functional and its Subdifferentiability

In this section we are concerned with the analysis of barrier functionals for the problem under consideration. We proceed as in [23]:

Definition 3.1. *For $r \geq 1$ and $\mu > 0$ we define barrier functions l of order r by*

$$l(v;\mu;r) \; : \; \mathbb{R}_+ \to \overline{\mathbb{R}},$$

$$l(v;\mu;r) := \begin{cases} -\mu \ln(v) & r = 1, \\ \frac{\mu^r}{(r-1)v^{r-1}} & r > 1. \end{cases}$$

We extend their domain of definition to $\mathbb{R}$ by setting $l(v;\mu;r) = \infty$ for $x \le 0$. We denote the pointwise derivative of $l(v;\mu;r)$ by $l'(v;\mu;r)$ if $v > 0$. This yields

$$l'(v;\mu;r) = \frac{-\mu^r}{v^r}.$$

With this we define a barrier functional b for the constraint $v \ge 0$ by:

$$\begin{aligned} b(\,\cdot\,;\mu;r) \;&:\; C(\overline{\Omega}_C) \to \overline{\mathbb{R}}, \\ v &\mapsto \int_{\overline{\Omega}_C} l(v(x);\mu;r)\,dx. \end{aligned}$$

Its formal derivative $b'(v,\mu;r) \in C(\overline{\Omega}_C)^*$, is defined as

$$\langle b'(v;\mu;r), \delta v\rangle := \int_{\overline{\Omega}_C} l'(v(x);\mu;r)\delta v(x)\,dx$$

if the right hand side exists.

Obviously, if $0 < \varepsilon \le v \in C(\overline{\Omega}_C)$, then b is differentiable with respect to v, and b' is the Fréchet derivative of b. If $v(x) = 0$, for some $x \in C(\overline{\Omega}_C)$, then the situation is more involved, and techniques of sub-differential calculus have to be applied.

In contrast to the case of state constraints, we may not use $\psi = 0$ to ease notation. This is due to the fact that in this case $u = 0$ would be the only admissible state. Therefore we introduce the following shifted barrier functional.

Definition 3.2. *We define the barrier functional for the constraint $|\nabla u|^2 \le \psi$ on a compact set $\overline{\Omega}_C \subseteq \overline{\Omega}$ by*

$$\begin{aligned} b_\psi(\,\cdot\,;\mu;r) \;&:\; C^1(\overline{\Omega}_C) \to \overline{\mathbb{R}}, \\ u \mapsto b_\psi(u;\mu;r) &:= b(\psi - |\nabla u|^2;\mu;r). \end{aligned} \tag{3.1}$$

In several cases we are only interested in a barrier functional of a fixed given order r, and sometimes even for only one fixed value of μ, in those cases we write $b(\cdot;\mu)$ or even $b(\cdot)$ if no confusion can occur.

Lemma 3.3. *The barrier functional b_ψ defined in (3.1) is well defined, convex, and lower-semicontinuous.*

Proof. By [23, Lemma 3.2] the outer function $b(\cdot;\mu;r)$ is well defined and lower semi-continuous. Since the inner function $\psi - |\nabla u|^2$ is well defined and continuous on U, the composition of both functions is well defined and lower semi-continuous.

Moreover, we know that $b(\cdot;\mu;r)$ is convex and monotonically decreasing. Further, the mapping $T(u) := \psi - |\nabla u|^2$ is pointwise concave. With these properties we can proof convexity of $b_\psi = b \circ T$ by the following computation which holds for every x in $\overline{\Omega}_C$:

$$l\big(T(\lambda u + (1-\lambda)\tilde{u})\big)(x) \le l\big(\lambda T(u) + (1-\lambda)T(\tilde{u})\big)(x) \le \lambda l\big(T(u)\big)(x) + (1-\lambda) l\big(T(\tilde{u})\big)(x).$$

By monotonicity of the integral we obtain that b_ψ is convex. □

We approach subdifferentiability of $b_\psi = b \circ (\psi - |\nabla \cdot|^2)$ via the following chain rule.

Lemma 3.4. *Let U, V be Banach spaces, $f : V \to \overline{\mathbb{R}}$ be a convex, lower-semicontinuous function, and $T : U \to V$ a continuously differentiable mapping with first derivative T'. Assume that the composite mapping $f \circ T$ is also convex.*

Let u be given and let $T'(u)$ be bounded. Assume that there is $\breve{u} \in U$, such that f is bounded above in a neighbourhood of $T(u) + T'(u)\breve{u}$. Then

$$\partial(f \circ T)(u) = (T'(u))^* \partial f(T(u)). \tag{3.2}$$

Proof. This is a slight extension of the well known chain rule of convex analysis (cf. [12, Prop. I.5.7]), which is, however, hard to find in the literature. We thus derive this result from a more general theorem from non-smooth analysis due to Clarke and Rockafellar (cf. [9, Thm. 2.9.9] or [20, Thm. 3]). Although the construction of the corresponding generalized differential is rather complicated in general, it reduces to the convex subdifferential in the case of convex functions (cf. [21, Thm. 5]).

First of all, we may assume that $f(T(u))$ is finite. Otherwise, $\partial(f \circ T)(u) = \partial(f(T(u)) = \emptyset$ holds trivially, because $\partial g(u) := \emptyset$ in case $g(u) = +\infty$ for every convex function g.

Otherwise we may argue as in [20, Cor. 1], which shows that the chain rule [20, Thm. 3] can be applied to show our assertion under the additional assumption that T is linear. However, inspection of its (short) proof shows that the same argumentation is still true in the case that T is "strictly differentiable" at u and $f \circ T$ is convex, as long as $\breve{u}$ exists that satisfies our assumptions. Now the Corollary subsequent to [9, Prop. 2.2.1] asserts that "strict differentiability" is implied by continuous differentiability, and our assertion follows. □

Remark 3.1. *Lemma 3.3 and Lemma 3.4 are also useful in the context of pointwise state constraints of the form $g(y(x), x) \leq 0$, if g is convex and differentiable in y.*

With the help of this lemma we can now characterize the subdifferential for barrier functionals with respect to gradient bounds in terms of the known subdifferential of a barrier functional in $C(\overline{\Omega}_C)$, see [23].

Proposition 3.5. *Assume that $\psi \geq \delta > 0$. Define*

$$\begin{aligned} b_\psi \,:\, C^1(\overline{\Omega}_C) &\to \overline{\mathbb{R}} \\ u &\mapsto b(\psi - |\nabla u|^2) \end{aligned}$$

as in Definition 3.2. Then the subdifferential $\partial b_\psi(u)$ has the following representation:

$$\partial b_\psi(u) = (-2\nabla u^T \nabla)^* \partial b(\psi - |\nabla u|^2). \tag{3.3}$$

This means, $\tilde{m} \in \partial b_\psi(u)$, if and only if there is $m \in \partial b(\psi - |\nabla u|^2)$, such that

$$\langle \delta u, \tilde{m} \rangle_{C^1(\overline{\Omega}_C), C^1(\overline{\Omega}_C)^*} = -2\langle \nabla u^T \nabla \delta u, m \rangle_{C(\overline{\Omega}_C), C(\overline{\Omega}_C)^*} \quad \forall \delta u \in C^1(\overline{\Omega}_C).$$

If u is strictly feasible, then $m = b'(\psi - |\nabla u|^2)$.

Proof. Let $T : C^1(\overline{\Omega}_C) \to C(\overline{\Omega}_C)$ be defined by $T(u) := \psi - |\nabla u|^2$. Obviously, the mapping $\psi - |\nabla u|^2 : C^1(\overline{\Omega}_C) \to C(\overline{\Omega}_C)$ is continuously differentiable with bounded derivative $(T'(u)\delta u)(x) = -2(\nabla u(x))^T \nabla \delta u(x)$.

We are going to apply Lemma 3.4 to the function $b_\psi : U \to \overline{\mathbb{R}}$, $b_\psi(u) = b \circ T$. By [23, Lemma 3.2], b is convex and lower semi-continuous and by Lemma 3.3 b_ψ is convex, too. Setting $\breve{u} := -0.5u$, we have $T'(u)\breve{u} = |\nabla u|^2$, and $\tilde{v} := T(u) + T'(u)\breve{u} = \psi$. Since $\psi \geq \delta > 0$, b is bounded from above in a $C(\overline{\Omega}_C)$-neighbourhood of $\tilde{v}$. Hence, Lemma 3.4 can be applied and yields our representation formula (3.3). Finally [23, Prop. 3.5] shows that $\partial b(v) = \{b'(v)\}$ if v is strictly feasible. □

The barrier functional b_ψ can also be analyzed on closed subspaces $\tilde{U}$ of $C^1(\overline{\Omega}_C)$. To this end let $E : \tilde{U} \to C^1(\overline{\Omega}_C)$ be the continuous embedding operator. Then its adjoint $E^* : C^1(\overline{\Omega}_C)^* \to \tilde{U}^*$ is the restriction operator for linear functionals. If $\breve{u}$ in Assumption 4 can be chosen from $\tilde{U}$, then the chain-rule of convex analysis applied to $b_\psi \circ E$ yields a characterization of the subdifferential of the restriction of b_ψ to $\tilde{U}$ as restriction of the subdifferential:

$$\partial(b_\psi \circ E)(u) = E^* \partial b_\psi(Eu).$$

Closed subspaces of $C^1(\overline{\Omega}_C)$ may for example be spaces that incorporate Dirichlet boundary conditions on $\overline{\Omega}_C \cap \Omega$ or finite dimensional subspaces.

4 Minimizers of Barrier Problems

With the preparations made in the previous sections we will now show that there exists a unique solution for the barrier problem, and later on some first order necessary conditions that are fulfilled by these.

Theorem 4.1. *(Existence of Solutions to Barrier Problems)*
Let Assumptions 1–4 be fulfilled. Then the Problem

$$\begin{aligned} \min J_\mu(q,u) \quad &:= J(q,u) + b_\psi(u;\mu), \\ s.t. \qquad & Au = Bq \end{aligned} \tag{4.1}$$

admits a unique minimizer (q_μ, u_μ). Moreover u_μ is strictly feasible almost everywhere in $\overline{\Omega}_C$.

Proof. By Assumption 4 $J_\mu(\breve{q}, \breve{u}) < \infty$. Further, J_μ is bounded from below by Assumption 3, by the required lower bound for J_1, and because b_ψ is bounded from below, since ψ is bounded above.

Taking a minimizing sequence $(q_k, u_k) = (q_k, Sq_k)$ (recall that $S = A^{-1}B$ is continuous by Lemma 2.1), we obtain from Assumption 3 that w.l.o.g. q_k converges weakly to some $q_\mu \in Q^{\mathrm{ad}}$. From Lemma 2.1 together with Assumption 1 we obtain that w.l.o.g. the sequence u_k converges to u_μ weakly in W where u fulfills equation (2.2b). From lower semi-continuity of J and b_ψ (cf. Lemma 3.3), we obtain that the limit (q_μ, u_μ) solves (4.1) and since $J_\mu(q_\mu, u_\mu) < \infty$ it follows that u is strictly feasible almost everywhere in Ω_C. Furthermore, the limit (q_μ, u_μ) is unique, since J is strictly convex with respect to the control variable, and the mapping $q_\mu \mapsto u_\mu$ is injective. □

The next theorem shows that the regularity of the solutions doesn't degenerate as $\mu \to 0$:

Theorem 4.2. *Let Assumption 1–4 be fulfilled. Then for every $\mu_0 > 0$ the solutions $(q_\mu, u_\mu) \in Q \times W$ of (4.1) are uniformly bounded on $(0, \mu_0]$.*

Proof. First note that due to Lemma 2.1 in combination with Assumption 1 it is sufficient to show that q_μ is uniformly bounded. To see this we note that, cf. [23],

$$J_\mu(q_\mu, u_\mu) \le J_\mu(q_{\mu_0}, u_{\mu_0}) \le J_{\mu_0}(q_{\mu_0}, u_{\mu_0}).$$

From $J(q_\mu, u_\mu) \le J_\mu(q_\mu, u_\mu)$ together with Assumption 3 we obtain, that q_μ is bounded, which concludes the proof. □

Usually, if $W \subset C^1(\overline{\Omega}_C)$ the state satisfies the additional regularity $W \subset C^{1,\beta}(\overline{\Omega}_C) \subset C^1(\overline{\Omega}_C)$. This means the gradients are even Hölder continuous of order β. Then we obtain for a sufficiently high order r of the barrier method that the state is in fact strictly feasible everywhere in Ω_C, as the following theorem shows.

Theorem 4.3. *Let $\Omega_C \subset \mathbb{R}^d$ be compact with Lipschitz boundary and for some $\beta \in (0,1)$ let $\psi \in C^{0,\beta}(\overline{\Omega}_C)$ be given. Let Assumptions 1–4 be satisfied. If the state has the additional regularity $u_\mu \in C^{1,\beta}(\overline{\Omega}_C)$, then for $r - 1 > \frac{d}{\beta}$ the state u_μ is strictly feasible in Ω_C.*

Proof. By Theorem 4.1 we obtain $0 \le \psi - |\nabla u_\mu|^2 \in C^{0,\beta}(\Omega_C)$. From [23, Lemma 6.1] we obtain that therefore $(\psi - |\nabla u_\mu|^2)^{-1} \in C(\Omega_C)$ which concludes the proof. □

We are now prepared to derive first order necessary conditions for the minimizer of the barrier problem (4.1).

Theorem 4.4. *Let the Assumptions 1–4 be fulfilled. Then $(q_\mu, u_\mu) \in Q^{ad} \times U$ is a solution to (4.1) if and only if there exist $m_\mu \in \partial b(\psi - |\nabla u_\mu|^2) \subset C(\overline{\Omega}_C)^*$ and $z_\mu \in Z$, $q_\mu^* \in Q^*$ such that the following holds:*

$$\begin{aligned}
Au_\mu &= Bq_\mu && \text{in } Z^* && (4.2a)\\
A^* z_\mu &= J_1'(u_\mu) + (-2(\nabla u_\mu)^T \nabla)^* m_\mu && \text{in } U^* && (4.2b)\\
J_2'(q_\mu) &= -B^* z_\mu - q_\mu^* && \text{in } Q^* && (4.2c)\\
\langle q - q_\mu, q_\mu^* \rangle &\le 0 && \forall\, q \in Q^{ad} && (4.2d)
\end{aligned}$$

Proof. We consider the following minimization problem where we omit the dependence on the parameter μ:

$$\min_{q\in Q} F(q) = \chi_{Q^{\mathrm{ad}}}(q) + j_\mu(q) := \chi_{Q^{\mathrm{ad}}}(q) + J_\mu(q, Sq) \tag{4.3}$$

where $\chi_{Q^{\mathrm{ad}}}$ is the indicator function for the admissible set of the controls, and S is the control to state mapping defined by (2.2b). Clearly $(q_\mu, u_\mu) = (q_\mu, Sq_\mu)$ is a solution to (4.1) if and only if q_μ is a solution to (4.3), which is in turn equivalent to $0 \in \partial F(q_\mu)$. In order to utilize this we will split the subdifferential by the sum-rule of convex analysis:

$$\partial F(q_\mu) = \partial(\chi_{Q^{\mathrm{ad}}})(q_\mu) + \partial j_\mu(q_\mu). \tag{4.4}$$

For its application note that Assumption 4 asserts the existence of a point

$$\breve{q} \in \operatorname{dom}\chi_{Q^{\mathrm{ad}}} \cap \operatorname{dom} j_\mu$$

such that j_μ is continuous in $\breve{q}$. In addition the function $\chi_{Q^{\mathrm{ad}}}$ is convex and lower semicontinuous, thus it coincides with its "Γ-regularization" [12, Chapter I, Prop. 3.1]. We can therefore apply the sum-rule of convex analysis, cf. [12, Chapter I, Prop. 5.6] to obtain (4.4).

Since j is continuous in q_μ we obtain by the same argument that:

$$\partial j_\mu(q_\mu) = \partial j(q_\mu) + \partial(b_\psi \circ S)(q_\mu)$$

where we recall the definition $b_\psi(u) = b(\psi - |\nabla u|^2)$. Now we note that

$$j(q) = J \circ (1, S)(q)$$

with the linear mapping

$$(1, S) \,:\, Q \to Q \times U, \qquad q \mapsto (q, Sq).$$

Together with Assumption 4 we are able to apply the linear chain rule and obtain

$$\begin{aligned} \partial j(q_\mu) &= (1, S^*)\partial J(q_\mu, u_\mu),\\ \partial(b_\psi \circ S)(q_\mu) &= S^*\partial b_\psi(Sq_\mu). \end{aligned}$$

Inserting the representation for the subdifferential of the barrier function b_ψ in Proposition 3.5 our computations have shown so far that

$$0 \in \partial(\chi_{Q^{\mathrm{ad}}})(q_\mu) + (1, S^*)\partial J(q_\mu, u_\mu) + S^*(-2(\nabla u_\mu)^T\nabla)^*\partial b(\psi - |\nabla u_\mu|^2) \tag{4.5}$$

is equivalent to (q_μ, u_μ) being a solution to (4.1). Since the cost functional is differentiable we obtain, cf. [12, Chapter I, Prop. 5.3]:

$$\partial J(q_\mu, u_\mu) = \{J_1'(u_\mu) + J_2'(q_\mu)\}.$$

Equation (4.5) means there exist $q_\mu^* \in \partial\chi_{Q^{\mathrm{ad}}}(q_\mu)$, and $m_\mu \in \partial b(\psi - |\nabla u_\mu|^2)$ such that

$$0 = q_\mu^* + J_2'(q_\mu) + S^*\big(J_1'(u_\mu) + (-2(\nabla u_\mu)^T\nabla)^* m_\mu\big) \in Q^*. \tag{4.6}$$

Note that $S^* = (A^{-1}B)^* = B^*(A^{-1})^* = B^*(A^*)^{-1}$, where $A^* : Z \supset \operatorname{dom} A^* \to U^*$ is well defined with continuous inverse due to Lemma 2.2. Define

$$z_\mu = (A^*)^{-1}\big(J_1'(u_\mu) + (-2(\nabla u_\mu)^T\nabla)^* m_\mu\big). \tag{4.7}$$

Then $z_\mu \in \operatorname{dom} A^* \subset Z$ and satisfies (4.2b) by definition. Equation (4.2c) now follows from (4.6). Further note that q_μ^* fulfills, see, e.g. [12, Chapter I, Prop. 5.1]

$$\sup_{q \in Q^{\mathrm{ad}}} \langle q, q_\mu^* \rangle = \langle q_\mu, q_\mu^* \rangle \tag{4.8}$$

which is equivalent to (4.2d). □

Example 4.1. *Let us apply our abstract results to Example 2.1 in the case of distributed control ($B = \mathrm{Id}$). Using the notation from there the first order optimality conditions have the following form. Let (q_μ, u_μ) be a solution to (4.1), then there exists $z_\mu \in Z$, $m_\mu \in \partial b(\psi - |\nabla u_\mu|^2; \mu)$ such that:*

$$\int_\Omega -\Delta u_\mu \, \varphi \, dx = \int_\Omega q_\mu \varphi \, dx \qquad \forall\, \varphi \in Z, \tag{4.9a}$$

$$\int_\Omega -\Delta\varphi \, z_\mu \, dx = \int_\Omega (u_\mu - u^d)\varphi \, dx - 2\int_{\overline{\Omega}} (\nabla u_\mu)^T \nabla\varphi \, dm_\mu \qquad \forall\, \varphi \in W, \tag{4.9b}$$

$$|q_\mu|^{p-2} q_\mu = -z_\mu \qquad a.e.\ in\ \Omega. \tag{4.9c}$$

For a discussion of the first two equations and in particular the representation of A and A^* we refer to Example 2.1. The barrier gradient m_μ is an element of $\partial b(u_\mu; \mu; r)$, and a measure in general. If u_μ is strictly feasible, which can usually be guaranteed a priori by a proper choice of the order r, then $m_\mu = b'(y; \mu; r)$ and thus a function, cf. [23, Prop. 3.5].

Equation 4.9c holds pointwise almost everywhere since it holds in L^p. The multiplier q_μ^* does not appear due to the fact that $Q^{ad} = Q$.

After having studied the necessary optimality conditions we will now discuss the behavior of the dual variables. The hard part is showing the boundedness of the measure obtained form the subdifferential of the barrier functional.

Theorem 4.5. *Let the assumptions of Theorem 4.4 be fulfilled. Then for each $\mu_0 > 0$*

$$\sup_{\mu \in (0, \mu_0]} \|m_\mu\|_{C(\overline{\Omega}_C)^*} \leq C.$$

Proof. Let (q_μ, u_μ) be the solution to (4.1) and $(\breve{q}, \breve{u})$ be a Slater point, e.g., let $\psi - |\nabla \breve{u}|^2 \geq d > 0$. Then, following [23], we multiply (4.2b) with $\delta u = u_\mu - \breve{u}$ and (4.2c) with $\delta q = q_\mu - \breve{q}$ and obtain

$$\begin{aligned} 0 &= \langle \delta u, -A^* z_\mu + J_1'(u_\mu) + (-2(\nabla u_\mu)^T \nabla)^* m_\mu \rangle + \langle \delta q, J_2'(q_\mu) + B^* z_\mu + q_\mu^* \rangle \\ &= \langle \delta u, J_1'(u_\mu) + (-2(\nabla u_\mu)^T \nabla)^* m_\mu \rangle + \langle \delta q, J_2'(q_\mu) + q_\mu^* \rangle + \langle A\delta u - B\delta q, -z_\mu \rangle. \end{aligned}$$

As $(\delta q, \delta u)$ fulfills the state equation (2.2b) this simplifies to

$$0 = \langle \delta u, J_1'(u_\mu) \rangle + \langle \delta q, J_2'(q_\mu) \rangle - 2\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu \rangle + \langle \delta q, q_\mu^* \rangle. \tag{4.10}$$

From the uniform boundedness of the primal variable, see Theorem 4.2 together with Assumption 2, we obtain that

$$|\langle \delta u, J_1'(u_\mu) \rangle + \langle \delta q, J_2'(q_\mu) \rangle| \leq C$$

with a constant C independent of μ. Inserting this estimate into (4.10) yields

$$|-2\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu \rangle + \langle \delta q, q_\mu^* \rangle| \leq C. \tag{4.11}$$

We would like to split this into the sum of the absolute values. To do so we will show that both terms have essentially the same sign. First, we now define the 'almost' active set

$$\mathcal{A} = \{x \in \overline{\Omega}_C \,|\, \psi - |\nabla u_\mu|^2 \leq 0.5d\}.$$

This is motivated by the fact, see [23, Corollary 3.6], that

$$|\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu|_{\overline{\Omega}_C \setminus \mathcal{A}} \rangle| \leq \|m_\mu\|_{L^1(\overline{\Omega}_C \setminus \mathcal{A})} \|(\nabla u_\mu)^T \nabla \delta u\|_{L^\infty} \leq C. \tag{4.12}$$

Thus it remains to take a look at the behavior of $\langle m_\mu|_{\mathcal{A}}, (\nabla u_\mu)^T \nabla \delta u \rangle$. We will now show that $0 < c \leq (\nabla u_\mu)^T \nabla \delta u$ holds on $\mathcal{A}$. For this we apply Young's-inequality and obtain

$$2|(\nabla u_\mu)^T \nabla \breve{u}| \leq |\nabla u_\mu|^2 + |\nabla \breve{u}|^2 \leq |\nabla u_\mu|^2 + \psi - d$$

leading to the following pointwise estimate on $\mathcal{A}$:

$$0.25d \leq 0.5(|\nabla u_\mu|^2 - \psi) + 0.5d \leq 0.5|\nabla u_\mu|^2 - 0.5|\nabla \breve{u}|^2 \leq (\nabla u_\mu)^T \nabla \delta u.$$

From [23, Prop. 3.5] we obtain that $m_\mu \leq 0$ as a measure thus leading to $-2\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu|_{\mathcal{A}} \rangle \geq 0$. Now we take a look on (4.2d) to see that $\langle q_\mu - \breve{q}, q_\mu^* \rangle \geq 0$. Together with (4.12) we obtain from (4.11) that

$$|\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu|_{\mathcal{A}} \rangle| \leq C.$$

Finally we note that due to $m_\mu \leq 0$ the following holds:

$$\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu|_{\mathcal{A}} \rangle \leq \min_{\mathcal{A}} \big((\nabla u_\mu)^T \nabla \delta u \big) \langle 1, m_\mu|_{\mathcal{A}} \rangle \leq -\frac{d}{4} \|m_\mu\|_{C(\mathcal{A})^*}.$$

This implies

$$\|m_\mu\|_{C(\mathcal{A})^*} \leq \frac{4}{d} |\langle (\nabla u_\mu)^T \nabla \delta u, m_\mu|_{\mathcal{A}} \rangle| \leq C$$

and completes the proof. □

Corollary 4.6. *Under the Assumptions 1–4 the following holds for every given* $\mu_0 > 0$*:*

$$\sup_{\mu\in(0,\mu_0]} \|z_\mu\|_Z \le C,$$
$$\sup_{\mu\in(0,\mu_0]} \|q_\mu^*\|_{Q^*} \le C.$$

Proof. First we note that the right hand side of (4.2b) is bounded due to Assumption 2, boundedness of u_μ, m_μ, and continuity of $((\nabla u_\mu)^T\nabla)^* : C(\overline{\Omega}_C)^* \to U^*$. The bound for z_μ follows from the boundedness of the right hand side of (4.2b) and continuity of $(A^*)^{-1}$. The bound for q_μ^* then follows from the bound on z_μ and q_μ using (4.2c) and Assumption 2 and continuity of B^*. □

5 Properties of the Central Path

We will now show convergence of the cost functional with rate μ. Before that we require the following simple geometric lemma.

Lemma 5.1. *Let* $x, y \in \mathbb{R}^d$ *with* $x^T(x-y) < 0$*. Then* $|x| < |y|$*.*

Proof. Choose $\alpha \in \mathbb{R}$ and $\hat{x} \in \operatorname{span}(x)^\perp$ such that $y = \alpha x + \hat{x}$. Then

$$x^T(x-y) = (1-\alpha)|x|^2 - x^T\hat{x} = (1-\alpha)|x|^2$$

hence $x^T(x-y) < 0$ implies $\alpha > 1$ thus

$$|x| < |\alpha x| \le |y|.$$

□

Theorem 5.2. *Let Assumptions 1–4 be fulfilled, and* (q_μ, u_μ) *be a solution of the barrier problem* (4.1) *for* $\mu > 0$*. Then the following holds for the minimizer* $(\overline{q}, \overline{u})$ *of* (2.2)*:*

$$J(q_\mu, u_\mu) \le J(\overline{q}, \overline{u}) + C\mu. \tag{5.1}$$

Proof. The proof follows the lines of [23, Lemma 5.1], however since we consider nonlinear constraints on the gradient of the states we have to modify the argumentation concerning the multiplier coming from the subdifferential of the barrier functional.

From the proof of Theorem 4.4 together with the relation

$$\partial b(\psi - |\nabla u_\mu|^2; \mu) = \mu^r \partial b(\psi - |\nabla u_\mu|^2; 1),$$

cf. [12, Chaper I, (5.21)], we obtain that there exists $m \in \partial b(\psi - |\nabla u_\mu|^2; 1)$ and $\varphi \in \partial\chi_{Q^{\mathrm{ad}}}(q_\mu) + \partial j(q_\mu) = \partial(\chi_{Q^{\mathrm{ad}}} + j)(q_\mu)$ such that:

$$\varphi - 2\mu^r S^*((\nabla u_\mu)^T\nabla)^* m = 0.$$

This shows that
$$2\mu^r S^*((\nabla u_\mu)^T\nabla)^* m \in \partial(\chi_{Q^{\mathrm{ad}}} + j)(q_\mu).$$
From convexity of $\chi_{Q^{\mathrm{ad}}}+j$ we obtain that for every $l \in \partial(\chi_{Q^{\mathrm{ad}}}+j)(q_\mu)$ the following holds:
$$j(q_\mu) \le j(\overline{q}) + \langle l, q_\mu - \overline{q}\rangle.$$
Applied to $2\mu^r S^*((\nabla u_\mu)^T\nabla)^* m$ we obtain:
$$J(q_\mu, u_\mu) \le J(\overline{q}, \overline{u}) + 2\mu^r \langle m, (\nabla u_\mu)^T \nabla(u_\mu - \overline{u})\rangle.$$
Since b is monotonically decreasing, the measure m is negative, cf. [23, Prop. 3.5]. Thus we can estimate further
$$2\mu^r \langle m, (\nabla u_\mu)^T \nabla(u_\mu - \overline{u})\rangle \le 2\mu^r \langle m|_{\Omega_S}, (\nabla u_\mu)^T \nabla(u_\mu - \overline{u})\rangle$$
where we define $\Omega_S := \{x \in \overline{\Omega}_C \,|\, (\nabla u_\mu)^T\nabla(u_\mu - \overline{u}) < 0\}$. From Lemma 5.1 it follows that $|\nabla u_\mu(x)| < |\nabla \overline{u}(x)| \le \psi(x)$ on Ω_S and thus $\Omega_S \subset \{x \in \overline{\Omega}_C \,|\, |\nabla u_\mu|^2 < \psi\}$. Hence we obtain from [23, Prop. 3.5.]
$$2\mu^r \langle m|_{\Omega_S}, \nabla u_\mu \nabla(u_\mu - \overline{u})\rangle = -2\int_{\Omega_S} \frac{\mu^r}{(\psi - |\nabla u_\mu|^2)^r} (\nabla u_\mu)^T \nabla(u_\mu - \overline{u})\, dx.$$
From $(\nabla u_\mu)^T \nabla \overline{u} \le |\nabla u_\mu|\, |\nabla \overline{u}| \le \psi$ we see that
$$\frac{-(\nabla u_\mu)^T\nabla(u_\mu - \overline{u})}{\psi - |\nabla u_\mu|^2} = \frac{(\nabla u_\mu)^T \nabla\overline{u} - |\nabla u_\mu|^2}{\psi - |\nabla u_\mu|^2} \le 1$$
and thus
$$2\mu^r \langle m|_{\Omega_S}, (\nabla u_\mu)^T \nabla(u_\mu - \overline{u})\rangle \le 2\mu \int_{\Omega_S} \frac{\mu^{r-1}}{(\psi - |\nabla u_\mu|^2)^{r-1}}\, dx. \tag{5.2}$$
From Theorem 4.5 and boundedness of the domain $\overline{\Omega}_C$ we obtain for the function $f := \mu/(\psi - |\nabla u_\mu|^2)$ that
$$\|f^{r-1}\|^{1/(r-1)}_{L^1(\overline{\Omega}_C)} = \|f\|_{L^{r-1}(\overline{\Omega}_C)} \le C\, \|f\|_{L^r(\overline{\Omega}_C)} = C\, \|f^r\|^{1/r}_{L^1(\overline{\Omega}_C)} \le C.$$
Thus the integral on the right hand side of (5.2) is bounded independent of μ. Hence the assertion follows. □

Theorem 5.3. *Let $\mu > 0$, (q_μ, u_μ) be a solution to the barrier problem (4.1) and $(\overline{q}, \overline{u})$ be the solution to the minimization problem (2.2). Further assume that there exist $c > 0$, $p \ge 2$ and a norm $\|\cdot\|$ such that*
$$c\|q_1 - q_2\|^p \le J_2(q_1) + J_2(q_2) - 2J_2\left(\frac{q_1+q_2}{2}\right).$$
Then the following estimate holds:
$$\|q_\mu - \overline{q}\| = O(\mu^{1/p}). \tag{5.3}$$

Proof. By assumption and convexity of J_1 the following proves the assertion

$$\begin{aligned} c\|q_\mu - \overline{q}\|^p &\le J_2(q_\mu) + J_2(\overline{q}) - 2J_2\left(\frac{q_\mu + \overline{q}}{2}\right) \\ &\le J(q_\mu, u_\mu) + J(\overline{q}, \overline{u}) - 2J((q_\mu + \overline{q})/2, (u_\mu + \overline{u})/2) \\ &\le J(q_\mu, u_\mu) + J(\overline{q}, \overline{u}) - 2J(\overline{q}, \overline{u}) = O(\mu). \end{aligned}$$

□

Remark 5.1. *By an analogous assumption on J_1 a similar result for the state u_μ can be obtained. In addition, if $\|\cdot\|$ is stronger than $\|\cdot\|_Q$ the convergence of u_μ in U (with the same rate $O(\mu^{1/p})$) follows by continuity of S.*

Example 5.1. *We finally return to Example 2.1. We apply the Clarkson inequality [10, Theorem 2 (3)] for L^p-spaces with $p > 2$, which yields*

$$\left\|\frac{f-g}{2}\right\|_{L^p}^p \le \frac{1}{2}\|f\|_{L^p}^p + \frac{1}{2}\|g\|_{L^p}^p - \left\|\frac{f+g}{2}\right\|_{L^p}^p$$

from this we see that $\|q\|_{L^p}^p$ matches the assumption of Theorem 5.3.

With the same techniques as in Theorem 5.2 it is possible to show for $\mu_0 > \mu > 0$ that $J_\mu(q_{\mu_0}, u_{\mu_0}) \le J_\mu(q_\mu, u_\mu) + C(\mu_0 - \mu)$. Then continuity of the central path follows via Theorem 5.3.

Acknowledgment

The first author is supported by the DFG Research Center MATHEON "Mathematics for key technologies". The second author is supported by the DFG priority program 1253 "Optimization with Partial Differential Equations".

References

[1] Maïtine Bergounioux, Mounir Haddou, Michael Hintermüller, and Karl Kunisch. A comparison of interior point methods and a Moreau-Yosida based active set strategy for constrained optimal control problems. *SIAM J. Optim.*, 11(2):495–521, 2000.

[2] Maïtine Bergounioux and Karl Kunisch. Primal-dual strategy for state-constrained optimal control problems. *Comp. Optim. Appl.*, 22:193–224, 2002.

[3] Heribert Blum and Rolf Rannacher. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Meth. in the Appl. Sci.*, 2(2):556–581, 1980.

[4] Eduardo Casas and Joseph Frédéric Bonnans. Contrôle de systèmes elliptiques semilinéares comportant des contraintes sur l'état. In H Brezzis and JL Lions, editors, *Nonlinear Partial Differential Equations and their Applications 8*, pages 69–86. Longman, New York, 1988.

[5] Eduardo Casas and Luis Alberto Fernández. Corrigendum: Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state. *Appl. Math. Optim.*, 28:337–339, 1993.

[6] Eduardo Casas and Luis Alberto Fernández. Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state. *Appl. Math. Optim.*, 27:35–56, 1993.

[7] Eduardo Casas, Mariano Mateos, and Jean-Pierre Raymond. Pontryagin's principle for the control of parabolic equations with gradient state constraints. *Nonlinear Anal.*, 46:933–956, 2001.

[8] Svetlana Cherednichenko and Arnd Rösch. Error estimates for the regularization of optimal control problems with pointwise control and state constraints. *Z. Anal. Anwendungen*, 27(2):195–212, 2008.

[9] F.H. Clarke. *Optimization and nonsmooth analysis*. John Wiley & Sons Inc., 1983.

[10] James A. Clarkson. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40(3):396–414, 1936.

[11] Klaus Deckelnick, Andreas Günther, and Michael Hinze. Finite element approximation of elliptic control problems with constraints on the gradient. *Numer. Math.*, 2008.

[12] Ivar Ekeland and Roger Temam. *Convex Analysis and Variational Problems*, volume 1 of *Studies in Mathematics and its Applications*. North-Holland Publishing Company, Amsterdam - Oxford, 1972.

[13] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*, volume 224 of *Grundlehren der mathematischen Wissenschaften*. Springer, revised 3. edition, 2001.

[14] Seymour Goldberg. *Unbounded Linear Operators*. McGraw-Hill Book Company, 1996.

[15] Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains*. Monographs and studies in Mathematics. Pitman, Boston, 1. edition, 1985.

[16] M. Hintermüller and K. Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. *SIAM J. Control Optim.*, 45(4):1198–1221, 2006.

[17] 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.

[18] C. Meyer, F. Tröltzsch, and A. Rösch. Optimal control problems of PDEs with regularized pointwise state constraints. *Comput. Optim. Appl.*, 33:206–228, 2006.

[19] Christian Meyer and Fredi Tröltzsch. On an elliptic optimal control problem with pointwise mixed control-state constraints. *Lecture Notes in Econom. and Math. Systems*, 563:187–204, 2006.

[20] R.T. Rockafellar. Directionally lipschitzian functions and subdifferential calculus. *Proc. London Math. Soc.*, 39(3):331–355, 1979.

[21] R.T. Rockafellar. Generalized directional derivatives and subgradients of nonconvex functions. *Can. J. Math.*, 32(2):257–280, 1980.

[22] A. Schiela. An interior point method in function space for the efficient solution of state constrained optimal control problems. Preprint 07-44, ZIB Report, 2008.

[23] Anton Schiela. Barrier methods for optimal control problems with state constraints. Preprint 07-07, ZIB Report, 2007.

[24] Fredi Tröltzsch. Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints. *SIAM J. Optim.*, 15(2):616–634, 2005.

[25] Winnifried Wollner. A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints. *Comput. Optim. Appl.*, 2008.