SIAM J. OPTIM.
Vol. 20, No. 2, pp. 1002–1031

BARRIER METHODS FOR OPTIMAL CONTROL PROBLEMS WITH STATE CONSTRAINTS*

ANTON SCHIELA†

Abstract. We study barrier methods for state constrained optimal control problems with PDEs. 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.

Key words. interior point methods in function space, optimal control, state constraints

AMS subject classifications. 90C51, 49M05

DOI. 10.1137/070692789

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 was laid a couple of years ago, but only recently have there been attempts to construct and analyze 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 and coworkers [16, 18, 17] transform state constraints into 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 Hintermüller and Kunisch [13, 12], which mainly use an L_2-penalty regularization of the state constraints and a semismooth Newton corrector [15]. Both methods achieve a regularizing effect by weakening the constraints. The irregular Lagrange multipliers of the original problem are replaced by more regular quantities contained in L_2. Both methods yield infeasible iterates as a consequence.

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, and coworkers [28, 29, 31, 30, 20] and Ulbrich and Ulbrich [26]. These methods are to be distinguished from *affine scaling interior point methods* considered in [25, 27], which are strongly related to semismooth Newton methods [24, 11]. 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

*Received by the editors May 23, 2007; accepted for publication (in revised form) May 29, 2009; published electronically July 23, 2009. This research was supported by the DFG Research Center MATHEON "Mathematics for key technologies."
http://www.siam.org/journals/siopt/20-2/69278.html
†Zuse Institute Berlin, Takustr, 7, 14195 Berlin, Germany (schiela@zib.de).

variants of path-following methods conceivable whose analysis is, however, beyond the scope of this paper.

The results of this paper extend the ideas in [21], which sketches the analysis of barrier methods for state constraints on the basis of the author's dissertation [20]. 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., [7, 4]).

2. State constrained optimal control problems.

Consider a convex minimization problem of the following form, the details of which are fixed in the remaining section:

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) \quad \text{subject to (s.t.) } Ay - Bu &= 0, \\ \underline{u} \le u &\le \overline{u}, \\ \underline{y} \le y &\le \overline{y}. \end{aligned} \tag{2.1}$$

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

2.1. Function spaces.

Let Q_Y be a compact subset of $\mathbb{R}^d$, and Q_U be a subset of $\mathbb{R}^n$, both equipped with a positive, regular, and finite measure.

Readers should have in mind the following simple examples. Let $\Omega \subset \mathbb{R}^d$ be an open, bounded Lipschitz domain. Then we can choose $Q_Y := \overline{\Omega}$ and equip it with the Lebesgue measure. Further, we may choose $Q_U := \Omega$ equipped with the Lebesgue measure for distributed control, or for boundary control $Q_U := \partial\Omega$ equipped with the measure on $\partial\Omega$ that induces the usual boundary integral. However, our results remain true in a more general setting (cf. Appendix C).

Let the *space of states* Y be a closed subspace of $C(Q_Y)$, and for some $1 < p < \infty$ let $U := L_p(Q_U)$ be the *space of controls*. We set $X := Y \times U$ and denote the natural norm on X by

$$\|(y,u)\|_X := \|y\|_{C(Q_Y)} + \|y\|_{L_p(Q_U)}.$$

Clearly, Y and U are both Banach spaces, U is reflexive, and both are equipped with a natural partial ordering (which makes the inequality constraints well defined). Moreover, integrals on Q_Y and Q_U and their measurable subsets can be used to define (barrier) functionals on Y and U, respectively.

2.2. Equality constraints.

The equality constraint $Ay - Bu = 0$ is introduced to model a partial differential equation.

Assumption 2.1. Let Y and U be defined as above, and let R be a Banach space.

(i) Let $A : Y \supset \operatorname{dom} A \to R$ be a linear, densely defined, bijective, and closed operator.

(ii) Let $B : U \to R$ be a linear continuous operator.

Define E as the subspace of all $(y,u) \in \operatorname{dom} Y \times U$ that satisfy $Ay - Bu = 0$.

Because A usually models a differential operator, we do not assume that it is defined on all of Y but only on a dense subspace $\operatorname{dom} A$, which may be an appropriately chosen Sobolev space in applications. This concept allows us to decouple the algebraic properties of A from its analytic properties and the choice of a topology that is appropriate for the analysis of the control problem. We have gathered some basic results on closed operators in Appendix A. In section 3 we consider a class of examples.

Remark 2.2. Bijectivity of A is not a crucial assumption. It has been imposed for simplicity here and can be weakened to accommodate, for example, pure Neumann problems. This has been exemplified in [22].

2.3. Inequality constraints and convex functionals. Some basic background on convex functions can be found in Appendix B.

Assumption 2.3.

(i) $J : X \to \mathbb{R}$ is a lower semicontinuous, convex functional.

(ii) The state constraints $y \geq \underline{y}$ and $y \leq \overline{y}$ are imposed pointwise on closed subsets $\underline{Q}_Y \subset Q_Y$ and $\overline{Q}_Y \subset Q_Y$, respectively, with $\underline{y} \in C(\underline{Q}_Y)$ and $\overline{y} \in C(\overline{Q}_Y)$.
For $U = L_p(Q_U)$, the control constraints $u \geq \underline{u}$ and $u \leq \overline{u}$ are imposed pointwise almost everywhere on measurable subsets $\underline{Q}_U \subset Q_U$ and $\overline{Q}_U \subset Q_U$, respectively, with $\underline{u} \in L_p(\underline{Q}_U)$ and $\overline{u} \in L_p(\overline{Q}_U)$.
Define $G = G_Y \times G_U \subset Y \times U$ as the feasible set with respect to these inequality constraints.

(iii) There are $\breve{x} = (\breve{y}, \breve{u})$ and $\tau > 0$ such that $A\breve{y} - B\breve{u} = 0$ and the following inequalities are all satisfied:

$$\begin{aligned} \tau &\leq \operatorname{ess}\inf_{t\in\underline{Q}_U} (\breve{u}(t) - \underline{u}(t)), \qquad & \tau &\leq \operatorname{ess}\inf_{t\in\overline{Q}_U} (\overline{u}(t) - \breve{u}(t)), \\ \tau &\leq \min_{t\in\underline{Q}_Y} (\breve{y}(t) - \underline{y}(t)), \qquad & \tau &\leq \min_{t\in\overline{Q}_Y} (\overline{y}(t) - \breve{y}(t)). \end{aligned} \tag{2.2}$$

(iv) J is coercive on $G \cap E$; i.e., every unbounded feasible sequence x_k yields $\lim_{k\to\infty} J(x_k) = +\infty$. This holds trivially if the feasible set is bounded in X.

By Assumption 2.3(ii) any of the constraints may be dropped. The barrier functionals, introduced later will be defined only on the subsets, where the inequality constraints are imposed. Assumption 2.3(iii) is a *Slater condition*. It is essential for the structure of dual variables. By our assumptions, G is nonempty, closed, and convex. By our choice of topology G_U may have empty interior, while the interior of G_Y is nonempty. Finally, in section 6 we will impose some additional regularity assumptions on J.

Assumption 2.4. Let $\|\cdot\|$ be a norm on X that satisfies an estimate of the form

$$c_N \|(y,u)\|_X \geq \|(y,u)\| \geq C_N \left(\|y\|_{L_1(Q_Y)} + \|u\|_{L_1(Q_U)} \right).$$

(i) The subdifferential ∂J of J is bounded in $(X, \|\cdot\|)^*$ on bounded sets of $(X, \|\cdot\|)$.

(ii) A strong (or uniform) convexity condition holds on the feasible set $E \cap G$:

$$\exists \alpha > 0 \,:\, \alpha \|\tilde{x} - x\|^2 \leq J(\tilde{x}) + J(x) - 2J\left(\frac{1}{2}\tilde{x} + \frac{1}{2}x\right) \ \forall \tilde{x}, x \in E \cap G. \tag{2.3}$$

Such a condition holds for optimal control problems with L_2-Tikhonov regularization term and allows the passage from convergence of function values to norm convergence of the corresponding arguments. The introduction of the norm $\|\cdot\|$ gives us more flexibility to apply our results to a larger class of problems.

2.4. Outline. After demonstrating in section 3 how elliptic control problems fit into our theoretical framework we will devote section 4 to the analysis of barrier functions with rational derivatives, which yield barrier functionals by integration. To incorporate inequality constraints, barrier functionals are added to J. In section 5 we consider existence of minimizers of the resulting barrier problems, first order optimality conditions, and uniform bounds on the dual variables. Here Assumption 2.1 and 2.3 are used. In section 6 we study the properties of the path of minimizers, such as Lipschitz continuity and convergence, if additionally Assumption 2.4 holds.

One conclusion of sections 4 and 5 is that measure valued dual variables can appear in the presence of state constraints. Section 7 clarifies this situation and shows how to avoid it by choosing an appropriate barrier function.

3. Examples. In the following we will discuss how optimal control problems with elliptic partial differential equations fit into our framework. Since it is rather clear how to choose the functional J and the inequality constraints to satisfy Assumption 2.3, we will concentrate on the equality constraints $Ay - Bu = 0$. In particular, we will discuss how differential operators are modeled as closed operators.

3.1. Closedness of elliptic differential operators. Let Ω be a smoothly bounded domain of $\mathbb{R}^d$, $\kappa \in C(\Omega, \mathbb{R}^{d\times d})$, $a \in L_\infty(\Omega;\mathbb{R})$. Assume that κ is symmetric and uniformly positive definite and $0 \neq a \geq 0$.

First we consider the following class of elliptic differential operators in the weak form:

$$\begin{aligned} &A : H^1(\Omega) \to (H^1(\Omega))^*, \\ &\langle Ay, v\rangle = \int_\Omega \langle \kappa\nabla y, \nabla v\rangle + ayv\, dt \quad \forall v \in H^1(\Omega). \end{aligned} \tag{3.1}$$

It is well known that there is no continuous embedding $H^1(\Omega) \hookrightarrow C(\overline{\Omega})$ for $d \geq 2$. To define an operator $A : C(\overline{\Omega}) \supset \operatorname{dom} A \to R$ we have to employ advanced regularity results, which can be found in the literature in many variants. A concise account on regularity theory is [3, section 9]. For our class of problems [3, Theorem 9.2] states that for $\infty > s > 1$ and $s' = s/(s-1)$ the restricted mapping

$$A : W^{1,s}(\Omega) \leftrightarrow (W^{1,s'}(\Omega))^*$$

is an isomorphism. If $s > d$, then by the Sobolev embedding theorem $W^{1,s}(\Omega) \hookrightarrow C(\overline{\Omega})$ densely. Setting $\operatorname{dom} A := W^{1,s}(\Omega)$, we conclude by Lemma A.1 closedness and bijectivity of

$$A : C(\overline{\Omega}) \supset \operatorname{dom} A \to (W^{1,s'}(\Omega))^*. \tag{3.2}$$

The regularity requirements on $\partial\Omega$, and the coefficients can be weakened considerably. The case of discontinuous κ is particularly delicate and has been analyzed in [2]. Yet, A can still be declared as a closed, bijective operator and fits into our framework. In view of Lemma A.1 all the information we need is stated in [2, Theorem 2], which asserts continuity of

$$A^{-1} : (W^{1,s'}(\Omega))^* \to C(\overline{\Omega}) \cap H^1(\Omega).$$

However, as opposed to the regular case, $\operatorname{dom} A := \operatorname{ran} A^{-1}$ cannot be characterized as a Sobolev space.

3.2. Various control problems. Next we will consider some examples for the operator $B : U \to R$ with $R = (W^{1,s'}(\Omega))^*$ defined as above. Here we follow the discussion of [14].

For distributed control, let $d \le 3$ and consider $U = L_2(\Omega) = U^*$. Then there is $s' < d/(d-1)$ such that the Sobolev embedding $I : W^{1,s'}(\Omega) \to L_2(\Omega)$ is continuous. Hence, we may define $B := I^* : U \to R$, which has a representation via integrals

$$\langle Bu, v\rangle = \int_\Omega uI(v)\,dt \quad \forall v \in W^{1,s'}(\Omega). \tag{3.3}$$

Next let Q_U be a subset of $\overline{\Omega}$, for which a trace operator $\gamma : W^{1,s'}(\Omega) \to L_2(Q_U)$ is defined. A prominent example is $Q_U = \partial\Omega$ for $d \le 2$. Then $B := \gamma^* : U \to R$ with representation

$$\langle Bu, v\rangle = \int_{Q_U} u\gamma(v)\,dt \quad \forall v \in W^{1,s'}(\Omega) \tag{3.4}$$

can be used to capture Neumann boundary control in the weak form.

Finally, if $r_1, \ldots, r_n$ are finitely many elements of $W^{1,s}(\Omega)$, then we may choose $Q_U := \{1, \ldots, n\}$, equipped with the counting measure, $U := L_2(Q_U)$, and

$$\langle Bu, v\rangle = \sum_{i=1}^n u_i \langle r_i, v\rangle \quad \forall v \in W^{1,s'}(\Omega).$$

3.3. Other cases. In broad terms our framework can be used for convex optimal control problems for which continuous states can be guaranteed. This corresponds to many analytic results in the literature (cf., e.g., [5, 6, 2]).

In some cases this may require a stricter choice for U than $U = L_2(Q_U)$. For example, elliptic Neumann boundary control in three spacial dimensions requires $U = L_p(\partial\Omega)$ with $p > 2$, because then no continuous trace operator $\gamma : W^{1,s'}(\Omega) \to L_2(\partial\Omega)$ exists for any $s' < d/(d-1) = 3/2$, and a definition of B, similar to (3.4), is not possible for $U = L_2(\partial\Omega)$. This situation is similar for parabolic optimal control problems.

To cope with these examples, we may choose the feasible control set G_U bounded in $L_p(Q_U)$ and use a quadratic objective functional, which is not coercive on $L_p(Q_U)$ but (trivially) on G_U. In this case the results of section 6 hold for the norm $\|\cdot\|_{L_2(Q_U)}$ and not for $\|\cdot\|_{L_p(Q_U)}$. Alternatively, we may use Tikhonov regularization terms of the form $\alpha/2\,\|u\|^2_{L_2(Q_U)} + \tilde{\alpha}/2\,\|u\|^p_{L_p(Q_U)}$, which are coercive on $L_p(Q_U)$ and strongly convex on $L_2(Q_U)$. As a third alternative, we may consider $Y = Y_h$ as a finite element subspace of $C(Q_Y)$. Then our theory can be applied; and even more, although the solution operator $S_h : U \to Y_h$ becomes unbounded for $h \to 0$, the estimates from section 6 still hold uniformly in h in a standard setting (cf. section 7.4).

4. Barrier functionals. 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 (cf. section 7). In finite dimensional spaces a comprehensive study of classes of barrier functions has been performed in [10].

DEFINITION 4.1. *For all $q \ge 1$ and $\mu > 0$ the functions defined by*

$$l(\cdot;\mu;q) : \mathbb{R}_+ \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\},$$
$$z \mapsto l(z;\mu;q) := \begin{cases} -\mu\ln(z), & q = 1, \\ \dfrac{\mu^q}{(q-1)z^{q-1}}, & q > 1, \end{cases}$$

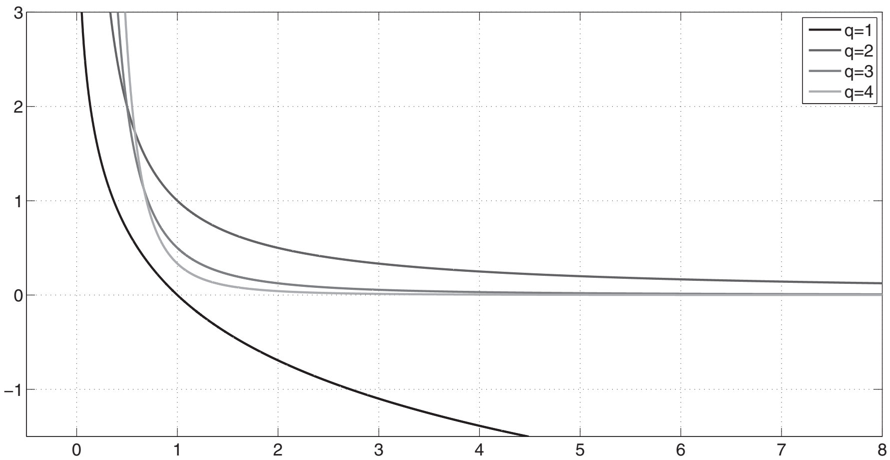

FIG. 4.1. *Barrier functions for various values of q.*

are called barrier functions of order q. *We extend their domain of definition to* $\mathbb{R}$ *by setting* $l(z;\mu;q) = \infty$ *for* $z \le 0$. *If we do not consider special values of* q *or* μ, *we may abbreviate the notation* $l(z;\mu;q)$ *by* $l(z;\mu)$ *or* $l(z)$.

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 (cf. Figure 4.1). Moreover, $\lim_{z\to 0} l(z) = \infty$ 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

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

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

Using barrier *functions* $l(z;\mu;q)$ we construct extended real valued barrier *functionals* $b(z;\mu;q)$.

DEFINITION 4.2. *Let Z be one of the two following spaces:*

(i) *$Z = L_p(Q)$ for $1 \le p < \infty$ and for $Q \subset \mathbb{R}^n$, equipped with a positive finite measure (cf. Appendix C).*

(ii) *Z is a closed subspace of $C(Q)$ for a compact set $Q \subset \mathbb{R}^d$, equipped with a positive finite and regular Borel measure (cf. Appendix C).*

Integrals and sets of zero sets are defined with respect to the choice of the corresponding measure. Define barrier functionals for $q \ge 1$ and $\mu > 0$ by

$$\begin{aligned} b(\cdot;\mu;q) : Z &\to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}, \\ z &\mapsto \int_Q l(z(t);\mu;q)\,dt. \end{aligned} \tag{4.1}$$

If $z < 0$ on a nonzero subset of Q, we set $b(z;\mu;q) = +\infty$.

To incorporate bounds like $y \ge \overline{y}$ and $y \le \underline{y}$, we insert $z := \overline{y} - y$ and $z := y - \underline{y}$ (and analogously for u) into the barrier functional. Of course, the choice of $L_p(Q)$ and $C(Q)$ as domain space necessitates that the arguments of b and thus also upper and lower bounds are contained in the corresponding space (cf. Assumption 2.3(ii)).

PROPOSITION 4.3. *Let $b(\cdot;\mu;q) : Z \to \overline{\mathbb{R}}$ be given as in Definition* 4.2.

(i) *$b(\cdot;\mu;q)$ is well defined, convex, lower semicontinuous, and bounded from below on bounded subsets of $(Z, \|\cdot\|_{L_1(Q)})$. For $q > 1$, $b(z;\mu;q) \geq 0$ for all $z \in Z$.*

(ii) *If $b(z;\mu;q) < \infty$, then $z > 0$ almost everywhere in Q.*

(iii) *If z_k is an unbounded sequence in Z, then $\liminf_{k\to\infty} b(z_k;\mu;q)/\|z_k\|_Z \geq 0$.*

Proof. Since Q has bounded measure, $\|z\|_{L_1(Q)} \leq c_Z \|z\|_Z$ in all cases, and corresponding continuous embeddings $Z \hookrightarrow L_1(Q)$ exist.

Although similar arguments as the following have been used elsewhere (cf. [20, 26]), we give a full proof of (i) for the sake of self-containedness. For $l(\cdot) = l(\cdot;\mu;q)$, consider the sequence of functions $l_k(\cdot) := \min\{l(\cdot), k\}$. Each l_k is continuous on $\mathbb{R}$, bounded from above by k and bounded from below by $l_k(z) \geq -|z|$ for $q = 1$ and $l_k(z) \geq 0$ for $q > 1$. Consequently, the corresponding Nemyckii (or superposition) operators $L_k : Z \to L_1(Q)$, defined by $L_k(z)(t) := l_k(z(t))$ a.e. in Q, are well defined and by continuity of l_k continuous (cf., e.g., [7, Proposition IV.1.1]). Hence, all functionals $b_k(z) : z \to \int_Q l_k(z)\,dt$ are well defined and continuous on Z and admit the common lower bound $b_k(z) \geq -\|z\|_{L_1(Q)}$ for $q = 1$ and $b_k(z) \geq 0$ for $q > 1$.

Now let $z \in Z$ be fixed. Then, after possibly choosing a representative,

$$l(z(t)) = \lim_{k\to\infty} l_k(z(t)) = \sup_{k\in\mathbb{N}} l_k(z(t)) \in \overline{\mathbb{R}}.$$

Thus, by the theorem of Beppo Levi, $b(z) = \sup_k b_k(z)$ exists as an integral (possibly taking the value $+\infty$), and the estimate $-\|z\|_{L_1(Q)} \leq b(z) \leq +\infty$ holds. Since each b_k is independent of the choice of the representative of z, so is the supremum $b(z)$. Hence, $b : Z \to \overline{\mathbb{R}}$, as given in (4.1), is well defined and bounded below on bounded sets of $(Z, \|\cdot\|_{L_1(Q)})$ for $q = 1$ and positive for $q > 1$. As the pointwise supremum of the continuous functionals b_k, b is lower semicontinuous. Finally, convexity of b follows immediately from the convexity of l.

Now (ii) follows from the fact that a positive integral is infinite if the integrand is infinite on a set of nonzero measure.

To prove (iii) we have only to consider $\mu = 1$ and (by positivity of b for $q > 1$) $q = 1$, i.e., $b(z;\mu;q) = -\int \ln(z)\,dt$. By Proposition 4.3(ii) it suffices to consider sequences z_k that satisfy $z_k(t) \geq 0$ almost everywhere in Q. If $b(z_k;\mu;q)$ is bounded from below for $\|z_k\|_Z \to \infty$, we are done. Otherwise, there are subsequences (which we again denote by z_k) such that $\lim_{k\to\infty} b(z_k) = -\infty$. Any such sequence has to be unbounded in $(Z, \|\cdot\|_{L_1(Q)})$ by Proposition 4.3(i), and we may assume $b(z_k) \leq 0$ for all $k \in \mathbb{N}$. It remains to show that $\liminf_{k\to\infty} b(z_k)/\|z_k\|_Z = 0$ for any such sequence.

To this end we define the average $\overline{z}_k$ of $z_k \geq 0$ by

$$\overline{z}_k := |Q|^{-1} \int_Q z_k\,dt = |Q|^{-1} \|z_k\|_{L_1(Q)}.$$

Because $-\ln$ is a convex function, we can apply Jensen's inequality (c.f. [8, Theorem VI.1.3] or [19, Theorem 3.3]), which yields

$$-\ln \overline{z}_k = -\ln\left(|Q|^{-1} \int_Q z_k\,dt\right) \leq |Q|^{-1} \int_Q -\ln z_k\,dt = |Q|^{-1} b(z_k).$$

Since $\lim_{k\to\infty} \|z_k\|_{L_1(Q)} = \infty$, $b(z_k) \leq 0$, and $|Q|\overline{z}_k = \|z_k\|_{L_1(Q)} \leq c_Z \|z_k\|_Z$, the following chain of inequalities shows (iii):

$$0 \geq \liminf_{k\to\infty} \frac{b(z_k)}{\|z_k\|_Z} \geq \liminf_{k\to\infty} \frac{c_Z b(z_k)}{\|z_k\|_{L_1(Q)}} \geq \lim_{k\to\infty} \frac{-c_Z|Q| \ln \overline{z}_k}{\|z_k\|_{L_1(Q)}} = \lim_{k\to\infty} \frac{-c_Z \ln \overline{z}_k}{\overline{z}_k} = 0. \qquad \square$$

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

4.1. Subdifferentiability in $L_p(Q)$ and $C(Q)$. The most important properties of barrier functionals are due to their first derivatives. A popular notion of a derivative in convex analysis is that of a subdifferential (cf. Appendix B).

Here we will consider the subdifferential $\partial b(z)$ of $b : Z \to \overline{\mathbb{R}}$. Subdifferentials of barrier functionals that incorporate upper bounds (like $z \le \overline{z}$) and lower bounds (like $z \ge \underline{z}$) are computed straightforwardly as $-\partial b(\overline{z} - z)$ and $\partial b(z - \underline{z})$, respectively.

The consideration of the same function in two different spaces Z_1 and Z_2 may lead to different subdifferentials in dual spaces Z_1^* and Z_2^*, respectively. In our case it will turn out that the subdifferential $\partial b(z)$ of b at z in $Z = L_p(Q)$ contains at most one element in $(L_p(Q))^*$. It is (if it exists) then given as expected by the *formal derivative* $b'(z;\mu;q)$ of b at z defined as

$$\langle b'(z;\mu;q), \delta z\rangle := \int_Q l'(z;\mu;q)\delta z \, dt = -\int_Q \frac{\mu^q}{z^q}\delta z \, dt. \tag{4.2}$$

Observe that (4.2) is not well defined in general but only if $l'(z;\mu;q)\delta z \in L_1(Q)$. Also here, we may use the abbreviated forms $b'(z;\mu)$ and $b'(z)$.

For $Z = C(Q)$—the case that appears in the presence of state constraints—the situation is more delicate because $C(Q)^*$ cannot be represented as a function space but only as a space of regular Borel measures [19, Theorem 6.19]. We approach subdifferentiability of $b(z)$ via directional derivatives $b'(z;\delta z)$ for directions δz.

LEMMA 4.4. *Consider $b : Z \to \overline{\mathbb{R}}$. Let $z, \delta z \in Z$, and assume that $b(z)$, $b(z+\delta z)$, and $\langle b'(z), \delta z\rangle$ are finite. Then b is directionally differentiable at z in the direction δz and*

$$b'(z;\delta z) = \langle b'(z), \delta z\rangle \ge \langle z^*, \delta z\rangle \qquad \forall\, z^* \in \partial b(z). \tag{4.3}$$

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

By convexity $l'(z)\delta z \le (l(z+h\delta z) - l(z))/h$ holds, $(l(z+h\delta z) - l(z))/h$ is monotonically increasing in h. Since $b(z)$, $b(z+\delta z)$, and $\langle b'(z), \delta z\rangle$ are finite, the function r_h defined by

$$r_h(t) := \frac{l(z(t) + h\delta z(t)) - l(z(t))}{h} - l'(z(t))\delta z(t)$$

is a positive element of $L_1(Q)$ for all $h \in]0;1]$. Moreover, by monotonicity of the difference quotient r_h is dominated by r_1.

Since $l(z)$ is differentiable if $z > 0$, and $z(t) > 0$ a.e. by Proposition 4.3(ii), $\lim_{h\to 0} r_h = 0$ pointwise a.e. in Q. Hence, we can apply the convergence theorem of Lebesgue to obtain $\lim_{h\to 0} \int_Q r_h(t)\, dt = 0$.

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

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

To prove the remaining part of (4.3) let $z^* \in Z^*$ and $\langle z^*, \delta z\rangle > \langle b'(z), \delta z\rangle$. Then there is $\varepsilon > 0$ such that

$$\langle z^*, h\delta z\rangle \ge \langle b'(z), h\delta z\rangle + \varepsilon h \ge b(z + h\delta z) - b(z) - o(h) + \varepsilon h.$$

But this implies $\langle z^*, h\delta z\rangle > b(z + h\delta z) - b(z)$ for sufficiently small h. This excludes z^* from the subdifferential and shows (4.3). ∎

PROPOSITION 4.5. *Consider $b : L_p(Q) \to \overline{\mathbb{R}}$, $1 \le p < \infty$, and let $p' = p/(p-1)$ be the dual exponent. If $l'(z) \in L_{p'}(Q)$, then $\partial b(z) = \{b'(z)\}$. Otherwise, $\partial b(z) = \emptyset$.*

Proof. If $b(z) = \infty$, then $\partial b(z) = \emptyset$ by definition, and we are done. Hence, we may assume that $b(z) < \infty$. Further, because $p < \infty$, we identify $L_{p'}(Q)$ and $(L_p(Q))^*$.

By convexity $l'(z(t))\delta z \le l(z(t)+\delta z) - l(z(t))$ holds pointwise almost everywhere in Q, as long as $z(t) > 0$ and $\delta z \in \mathbb{R}$. If $l'(z) \in L_{p'}(Q)$, then by monotonicity of the integral $\langle b'(z), \delta z\rangle \le b(z + \delta z) - b(z)$ for all $\delta z \in L_p(Q)$ and thus $b'(z) \in \partial b(z)$.

However, $l'(z) \in L_{p'}(Q)$ does not hold in general. Thus we proceed as follows: if $m \in \partial b(z) \subset L_{p'}(Q)$, we will show that $m = l'(z)$ in $L_{p'}(Q)$. Hence, either $\partial b'(z) = \emptyset$ or $\partial b(z) = \{b'(z)\}$. The latter case implies $l'(z) \in L_{p'}(Q)$. To prove $m = l'(z)$ we will show $\int_S l'(z)\, dt = \int_S m\, dt$ on all measurable subsets $S \subset Q$.

For fixed S define $S_d := \{t \in S : z(t) > d\}$ and denote by χ_{S_d} its characteristic function. Clearly, $\|l'(z(t))\|_{L_\infty(S_d)} < \infty$ for all $d > 0$. Let $z_k := \chi_{S_{1/k}}$ and $\varepsilon := 1/(2k)$. Then $|\langle b'(z), \pm\varepsilon z_k\rangle| < \infty$ and $b(z \pm \varepsilon z_k) < \infty$. Hence, Lemma 4.4 yields $\langle b'(z), \pm\varepsilon z_k\rangle \ge \langle m, \pm\varepsilon z_k\rangle$, because $m \in \partial b(z)$. Thus

$$\int_Q l'(z) z_k\, dt = \langle b'(z), z_k\rangle = \langle m, z_k\rangle = \int_Q m z_k\, dt \quad \forall k \in \mathbb{N}. \tag{4.4}$$

Since $b(z) < \infty$, the set $\{t \in S : z = 0\}$ has zero measure by Proposition 4.3. Thus $\chi_S(t) = \lim_{k\to\infty} z_k(t)$ pointwise almost everywhere in Q. Because $\|z_k\|_{L_\infty(S)} \le 1$ and $m \in L_{p'}(Q)$, $\chi_S m \in L_1(Q)$. Moreover, $\chi_S \ge z_k \ge z_{k-1} \ge 0$ by construction, and thus $|\chi_S m|$ dominates the sequence $z_k m$. Hence, by the convergence theorem of Lebesgue

$$\lim_{k\to\infty} \int_Q m z_k\, dt = \int_S m\, dt. \tag{4.5}$$

Further, because $-l'(z)$ is positive and $z_k \ge z_{k-1}$, the convergence theorem of Beppo Levi yields

$$\lim_{k\to\infty} \int_Q l'(z) z_k\, dt = \int_S l'(z)\, dt. \tag{4.6}$$

By (4.4) the sequences on the left-hand sides of (4.5) and (4.6) coincide and thus $\int_S m\, dt = \int_S l'(z)\, dt$. This completes the proof. ∎

At a crucial point in the proof we used the convergence theorem of Lebesgue to the sequence of *functions* $\chi_{S_{1/k}} m$. This was possible, because $m \in L_{p'}(Q)$. When we consider barrier functionals on $C(Q)$ (for a compact set Q), we cannot use this argument. The dual space $C(Q)^*$ is represented by the space of regular signed Borel measures $M(Q)$ (see [19, Theorem 6.19]). For $m \in M(Q)$, a similar sequence $\chi_{S_{1/k}} m$ is now a sequence of *measures*, and the theorem of Lebesgue is not available, of course. This spoils uniqueness of the subdifferential. However, by σ-additivity and regularity of these measures we can restrict nonuniqueness to that subset of Q where z touches the bound.

PROPOSITION 4.6. *Consider $b : C(Q) \to \overline{\mathbb{R}}$, and let $Q_F := \{t \in Q : z(t) = 0\}$ be the compact set, where z is not strictly feasible. Denote by $M_+(Q_F)$ the cone of regular positive Borel measures on Q_F.*

If $l'(z) \in L_1(Q)$, then

$$\partial b(z) = \{l'(z)\} - M_+(Q_F) \;\subset\; M(Q), \tag{4.7}$$

$$\|l'(z)\|_{L_1(Q)} = \min_{z^* \in \partial b(z)} \|z^*\|_{M(Q)}\,. \tag{4.8}$$

Otherwise, $\partial b(z) = \emptyset$. If z is strictly feasible, then $\partial b(z) = \{b'(z)\}$.

Proof. We may assume throughout the proof that $b(z) < \infty$. Otherwise, $l'(z) \notin L_1(Q)$ and $\partial b(z) = \emptyset$. Then our assertion holds trivially. Note also that, by continuity of z, Q_F is a closed subset of the compact set Q and thus compact.

Because b is monotonically decreasing, we have $b(\tilde z) - b(z) \le 0$ if $z \ge \tilde z$. If $z^* \in \partial b(z)$, then $\langle z^*, \tilde z - z\rangle \le b(\tilde z) - b(z) \le 0$, and thus z^* is a negative functional on $C(Q)$, which implies that it has a representation as a negative Radon measure ν. We will show that $\int_S l'(z)\,dt = \nu(S)$ for every (relatively) open subset of $Q \setminus Q_F$ and conclude by (outer) regularity of ν

$$\forall \text{ measurable } M \subset Q \setminus Q_F \;:\; \nu(M) = \inf\{\nu(O) : M \subset S \subset Q \setminus Q_F,\, S \text{ open}\}$$

that $l'(z) = \nu$ on $Q \setminus Q_F$.

Let S be an open subset of $Q \setminus Q_F$, and consider $S_d := \{t \in S : z(t) > d\}$, which is open in $Q \setminus Q_F$ by continuity of z. On S_d, $l'(z)$ is bounded and negative and can be interpreted as a measure there. Hence, $L_1(S_d, -\nu)$ and $L_1(S_d, -l'(z))$ are well defined, as well as $L_1(S_d, -\nu - l'(z))$. By positivity of $-\nu$ and $-l'(z)$ the norm of $L_1(S_d, -\nu - l'(z))$ is stronger than the norms of the other two spaces. We will show the following analogue to (4.4):

$$\int_{S_d} l'(z)\,dt = \nu(S_d). \tag{4.9}$$

Let χ_{S_d} be the characteristic function of S_d, which is an element of $L_1(S_d, -\nu - l'(z))$. By density (cf., e.g., [19, Theorem 3.14]) there is a sequence z_k of continuous functions with compact support $\operatorname{supp}(z_k) \subset S_d$ such that $z_k \to d/2\,\chi_{S_d}$ in $L_1(S_d, -\nu - l'(z))$ and thus simultaneously in $L_1(S_d, -\nu)$ and $L_1(S_d, -l'(z))$. Moreover, because S_d is open in Q, its relative boundary $\partial S_d := \overline{S_d} \setminus S_d$ and S_d are disjoint. Hence, since $\operatorname{supp}(z_k) \subset S_d$, $z_k = 0$ on ∂S_d, and thus z_k can be extended by 0 to an element of $C(Q)$.

From $b(z) < \infty$, $z_k = 0$ on $Q \setminus S_d$, and $\|z_k\|_\infty \le d/2$, we conclude $b(z \pm z_k) < \infty$ and $\langle b'(z), \pm z_k\rangle < \infty$. Hence, Lemma 4.4 can be applied to $\pm z_k$, and (4.3) yields

$$\int_{S_d} l'(z) z_k\,dt = \int_Q l'(z) z_k\,dt = \langle b'(z), z_k\rangle = \langle z^*, z_k\rangle = \int_Q z_k\,d\nu = \int_{S_d} z_k\,d\nu.$$

Since $z_k \to d/2\,\chi(S_d)$ in $L_1(S_d, -\nu)$ and $L_1(S_d, -l'(z))$, the integrals of this equation converge to the corresponding expressions in (4.9). This shows (4.9).

Because $S := \bigcup_{i\in\mathbb{N}} S_{1/i}$ it follows (for the left-hand side by the theorem of Beppo Levi and for the right-hand side by σ-additivity of ν) that

$$\int_S l'(z)\,dt = \lim_{i\to\infty} \int_{S_{1/i}} l'(z)\,dt = \lim_{i\to\infty} \nu(S_{1/i}) = \nu(S), \tag{4.10}$$

and thus $l'(z) = \nu$ on $Q \setminus Q_F$. This implies in particular that $l'(z) \in M(Q \setminus Q_F)$ and thus $l'(z) \in L_1(Q \setminus Q_F)$, because $l'(z)$ is a function on $Q \setminus Q_F$. Since Q_F is a set of zero measure so that $l'(z)\chi_{Q\setminus Q_F}$ is a representative of an element of $L_1(Q)$, we can write $l'(z) \in L_1(Q)$. Because ν is a negative measure on Q, and Q_F, has zero measure, $\nu|_{Q_F}$ has to be negative, too. This finally shows that each $z^* \in \partial b(z)$ satisfies (4.7).

Next we show that $b'(z) + m \in \partial b(z)$ if $m \in M(Q_F)$ is negative and $l'(z) \in L_1(Q)$. By convexity of l and by monotonicity of the integral, we have $\langle b'(z), \tilde{z} - z\rangle \leq b(\tilde{z}) - b(z)$ and thus $b'(z) \in \partial b(z)$. Since Q_F is a zero set and m is negative, this implies $\langle b'(z) + m, \tilde{z} - z\rangle \leq b(\tilde{z}) - b(z)$ if $\tilde{z} \geq z = 0$ on Q_F. But otherwise, $b(\tilde{z})$ is infeasible, and $b(\tilde{z}) = \infty$. But then trivially $\langle b'(z) + m, \tilde{z} - z\rangle \leq b(\tilde{z}) - b(z)$. Hence, $\langle b'(z) + m, \tilde{z} - z\rangle \leq b(\tilde{z}) - b(z)$ for all $\tilde{z} \in C(Q)$, and thus $b'(z) + m \in \partial b(z)$.

Clearly, by nonpositivity of m and $b'(z)$, $b'(z)$ is the element of $\partial b(z)$ of smallest norm. Further, if z is strictly feasible, then $Q_F = \emptyset$ and $b'(z) \in C(Q) \subset L_1(Q)$. Hence, $\partial b(z) = \{b'(z)\}$ in this case. ∎

For our next, easy observation we use the set $Q_d := \{t \in Q : z(t) > d\}$. For fixed d and $m \in \partial b(z)$, we show that $m|_{Q_d}$ is asymptotically negligible for $\mu \to 0$. This means that upper and lower bounds *decouple* for $\mu \to 0$ as long as there is a Slater point.

Corollary 4.7. *Consider $b : Z \to \overline{\mathbb{R}}$ for $Z = L_p(Q)$ with $1 \leq p < \infty$ or $Z = C(Q)$, and let $m \in \partial b(z; \mu; q)$. Then $m|_{Q_d} = l'(z; \mu; q)|_{Q_d}$ and*

$$\|l'(z; \mu; q)\|_{L_\infty(Q_d)} \leq \left(\frac{\mu}{d}\right)^q. \tag{4.11}$$

Further,

$$\left\|\left(\frac{\mu}{z}\right)^r\right\|_{L_1(Q)} \leq C(r, q, |Q|)\, \|l'(z; \mu; q)\|_{L_1(Q)} \quad \forall\, 0 \leq r \leq q. \tag{4.12}$$

Proof. The proof follows by definition $z \leq d$ on Q_d and $Q_d \subset Q$ as used in (4.7). The assertion (4.11) follows now directly from the definition of $b'(z)$.

For (4.12) we note that the case $r = 0$ is trivial. Since Q is bounded, this estimate can be derived for $r > 0$ from the following chain of inequalities:

$$\left\|\left(\frac{\mu}{z}\right)^r\right\|_{L_1(Q)}^{1/r} = \left\|\frac{\mu}{z}\right\|_{L_r(Q)} \leq C(r, q, |Q|) \left\|\frac{\mu}{z}\right\|_{L_q(Q)} = C(r, q, |Q|) \left\|\left(\frac{\mu}{z}\right)^q\right\|_{L_1(Q)}^{1/q}. \quad ∎$$

4.2. Other cases. We can now easily derive subdifferentials of barrier functions in two other cases, which often appear in the analysis.

First we characterize barrier subdifferentials on subspaces of $C(Q)$. This includes cases like $Z = C_0(Q)$ (continuous functions that vanish on the boundary) but also finite element subspaces of $C(Q)$. The key to their characterization is the chain-rule of convex analysis.

Proposition 4.8. *Let Z be a closed subspace of $C(Q)$. For $\underline{z} \in C(Q)$ consider the barrier functional $b(\cdot - \underline{z}; \mu; q) : C(Q) \to \overline{\mathbb{R}}$ and denote by $b|_Z$ its restriction to Z. Assume that there is $\check{z} \in Z$ such that $\min_{t\in Q} \check{z}(t) - \underline{z}(t) > 0$. Then the subdifferential of the restricted functional is the restriction of the subdifferential:*

$$\partial b|_Z(z - \underline{z}; \mu; q) = \partial b(z - \underline{z}; \mu; q)\Big|_Z \quad \forall z \in Z.$$

Proof. Let $I : Z \to C(Q)$ be the continuous injection $Iz = z$. Then its adjoint $I^* : C(Q)^* \to Z^*$ is the restriction operator $I^*z^* = z^*|_Z$. With these operators we can write $b|_Z(z - \underline{z}; \mu; q) = (b \circ I)(z - \underline{z}; \mu; q)$, and we have to show

$$\partial(b \circ I)(z - \underline{z}; \mu; q) = I^*\partial b(z - \underline{z}; \mu; q).$$

This follows from the chain-rule of convex analysis (cf. [7, Proposition I.5.7]) if there is $\breve{z}$ such that $b(I\breve{z}) = b(\breve{z})$ is continuous and finite. In $C(Q)$ this continuity assertion for b is guaranteed by our assumption. $\square$

For the case $C_0(Q)$ our assumptions require that $\underline{z} < 0$ on ∂Q. This corresponds nicely to the Slater condition Assumption 2.3(iii) that we need anyway to derive first order optimality conditions.

Finally, we sketch how to treat sums of barrier functionals to be able to cover upper and lower bounds simultaneously.

PROPOSITION 4.9. *Let $\underline{b}(z) := b(z - \underline{z})$ and $\overline{b}(z) := b(\overline{z} - z)$, and consider the following two cases:*

(i) $Z = L_p(Q)$, $1 \le p < \infty$, $\underline{z}, \overline{z} \in L_p(Q)$, and $\operatorname{ess\,inf}_{t\in Q}(\overline{z} - \underline{z}) > 0$.

(ii) *Z is a closed subspace of $C(Q)$, $\underline{z}, \overline{z} \in C(Q)$, and there is $\breve{z} \in Z$ such that $\overline{z} > \breve{z} > \underline{z}$.*

Then

$$\partial(\underline{b}(z) + \overline{b}(z)) = \partial\underline{b}(z) + \partial\overline{b}(z). \tag{4.13}$$

Proof. For the case $Z = L_p(Q)$, we note that $\partial(f(z) + g(z)) \subset \partial f(z) + \partial g(z)$ in general (cf. [7, equation II(5.22)]), and it remains to show the reverse inclusion of (4.13). Let $m \in \partial(\underline{b}(z) + \overline{b}(z)) \subset L_{p'}(Q)$. We have to show that $m = l'(z) := \underline{l}'(z) + \overline{l}'(z)$ holds in $L_{p'}(Q)$. This works analogously to the proof of Proposition 4.5 with the difference that $l'(z)$ was negative there, and now it may change signs on Q. However, we may split every measurable subset $S \subset Q$ into S_+ and S_-, according to the sign of $l'(z)$, and consider the limit process (4.6) separately on those two sets. Then the proof of Proposition 4.5 carries over. After having shown $l'(z) \in L_{p'}(Q)$ it follows from $\operatorname{ess\,inf}_{t\in Q}(\overline{z} - \underline{z}) > 0$ and Corollary 4.7 that $\underline{l}'(z), \overline{l}'(z) \in L_{p'}(Q)$, and thus by Proposition 4.5, $\overline{b}'(z) \in \partial\overline{b}(z)$ and $\underline{b}'(z) \in \partial\underline{b}(z)$.

Let Z be a subspace of $C(Q)$. Then $\overline{b}$ and $\underline{b}$ are continuous at $\breve{z}$, because $\breve{z}$ has positive distance from $\underline{z}$ and $\overline{z}$ by compactness of Q. Hence, we can apply the sum-rule of convex analysis (cf., e.g., [7, Proposition I.5.6]) to obtain (4.13). $\square$

5. Minimizers of barrier problems and their optimality conditions.

Having studied the properties of barrier functionals we want to use them to incorporate upper and lower bounds for optimal control problems. Thus we have to add one or several barrier functionals, the sum of which we denote by $b(x;\mu)$ or shorter by $b(x)$. We consider only pure state and control constraints and write $b(y)$ and $b(u)$ for those barrier functionals corresponding to the state and the control constraints, respectively. Moreover, we distinguish by $\overline{b}(\cdot)$ and $\underline{b}(\cdot)$ barrier functionals that incorporate upper and lower bounds, respectively. In view of Assumption 2.3(ii) the barrier integrals are all defined on the corresponding subsets of Q_Y and Q_U (which may also be empty, in which case corresponding barrier functionals are left away).

Our main concern will be the interaction of the several barrier functionals. Corollary 4.7 is a first hint that the situation may be comparable to the original problem.

It is a popular strategy in convex analysis to combine the functional J, the equality constraints $x \in E$ (cf. Assumption 2.1), and the inequality constraints $x \in G$ (cf. Assumption 2.3) into a single functional. This is done via indicator functions (cf. (B.1)). With their help we can rewrite (2.1) as an unconstrained minimization problem defined by the following functional:

$$F : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\},\qquad F := J + \iota_E + \iota_G. \tag{5.1}$$

It follows from standard arguments (taking into account Lemma A.1(i)) that F admits a minimizer if Assumptions 2.1 and 2.3 hold. Adding barrier functionals to F we obtain another convex functional F_μ defined by

$$\begin{aligned} F_\mu(x) := F(x) + b(x;\mu) &= J(x) + \iota_E(x) + \iota_G(x) + b(x;\mu) \\ &= J(x) + \iota_E(x) + b(x;\mu). \end{aligned} \tag{5.2}$$

The last equality holds because $\iota_G = 0$ in G and $\iota_G = b = +\infty$ in $X \setminus G$.

5.1. Existence of minimizers. We start with a lemma on coercivity of F_μ.

LEMMA 5.1. *Suppose that Assumptions* 2.1 *and* 2.3 *hold. Then F_μ is coercive for every $\mu > 0$ and every $q \geq 1$.*

Proof. Since F takes a finite value somewhere, we may assume without loss of generality (w.l.o.g.) that $F(0) = 0$. Further, by coercivity of F there is R such that $F(x) \geq 1$ for all x with $\|x\|_X \geq R$. For such an x, $x_R := R/\|x\|_X\, x$ satisfies $\|x_R\|_X = R$ and thus $F(x_R) \geq 1$. By convexity of F we obtain

$$1 \leq F(x_R) \leq R/\|x\|_X\, F(x) + (1 - R/\|x\|_X)F(0) = R/\|x\|_X\, F(x).$$

Hence, we conclude

$$F(x) \geq R^{-1}\|x\|_X \quad \forall x \in X : \|x\|_X \geq R. \tag{5.3}$$

Consider an unbounded sequence x_k in X. Then by (5.3) and Proposition 4.3(iii)

$$\begin{aligned} \liminf_{k\to\infty} F_\mu(x_k) &= \liminf_{k\to\infty}(F(x_k) + b(x_k;\mu)) \geq \liminf_{k\to\infty} \|x_k\|_X\,(R^{-1} + b(x_k;\mu)/\|x_k\|_X) \\ &\geq \lim_{k\to\infty} R^{-1}\|x_k\|_X + \liminf_{k\to\infty} b(x_k;\mu)/\|x_k\|_X \geq \lim_{k\to\infty} R^{-1}\|x_k\|_X = \infty. \end{aligned}$$

Hence, F_μ is coercive. ∎

THEOREM 5.2. *Suppose that Assumptions* 2.1 *and* 2.3 *hold. Let $F : X \to \overline{\mathbb{R}}$ be of the form* (5.1)*. Then F_μ as defined in* (5.2) *admits a minimizer $x(\mu) = (y(\mu), u(\mu))$ for each $\mu > 0$. Moreover, $y(\mu)$ and $u(\mu)$ are strictly feasible with respect to the inequality constraints almost everywhere in Q_Y and Q_U, respectively. Further, for every $\mu_0 > 0$, $\|x(\mu)\|_X$ is uniformly bounded for $\mu \in]0;\mu_0]$.*

Proof. Since b is convex and lower semicontinuous, F_μ inherits convexity and lower semicontinuity from F and b. Moreover, F_μ is coercive by Lemma 5.1.

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

$$\min\{F(x), F_{\mu_0}(x)\} \leq F_\mu(x) \leq \max\{F(x), F_{\mu_0}(x)\}. \tag{5.4}$$

It follows that the level sets of F_μ with $\mu \in [0;\mu_0]$ are uniformly bounded in X. By Assumption 2.3, there is a point $\breve{x}$, with $b(\breve{x}) < \infty$, and thus $F_\mu(\breve{x}) < \infty$. We can thus apply the main existence theorem for minimizers in convex optimization (cf., e.g., [7, Proposition I.1.2]) to obtain existence of a minimizer $x(\mu)$. Moreover, by (5.4)

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

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 Proposition 4.3(ii) a barrier minimizer can only touch the bounds on a set of zero measure. Otherwise, the barrier functional would be $+\infty$. ∎

Usually (and in particular under Assumption 2.4(ii)) F and thus F_μ are strictly convex. Then the minimizer is unique.

5.2. Optimality conditions. Next we study first order optimality conditions for barrier problems.

LEMMA 5.3. *Suppose that Assumptions* 2.1 *and* 2.3 *hold, and let* $\mu > 0$. *Then for all* $x \in X$

$$\partial F_\mu(x) = \partial J(x) + \partial\iota_E(x) + \partial b(x;\mu), \tag{5.5}$$

$$\partial b(x) = \partial\underline{b}(y) + \partial\overline{b}(y) + \partial\underline{b}(u) + \partial\overline{b}(u), \tag{5.6}$$

$$\partial b(x;\mu;q) = \mu^q \partial b(x;1;q) \quad \forall\mu > 0\ \forall q \geq 1. \tag{5.7}$$

Proof. To show (5.5) we apply the sum-rule of subdifferential calculus. By definition $F_\mu(x) := F(x) + b(x;\mu) = J(x) + \iota_E(x) + b(x;\mu)$. We will show in two steps

$$\partial(J + \iota_E + b)(x) = \partial J(x) + \partial(\iota_E + b)(x) = \partial J(x) + \partial\iota_E(x) + \partial b(x).$$

To apply the sum-rule to a sum $f + g$ the summands have to be convex, lower semicontinuous and satisfy an additional regularity condition such as the following (cf., e.g., [4, Theorem 4.3.3]):

$$0 \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} g). \tag{5.8}$$

Here the expression $\operatorname{dom} f - \operatorname{dom} g$ denotes the pointwise difference of the two sets where f and g are finite, respectively (cf. Appendix B). Now let B_V be the unit ball in a normed space V. We observe that showing (5.8) is equivalent to showing that there is $\varepsilon > 0$ such that each $x \in \varepsilon B_X$ can be written as a difference $x = x_1 - x_2$ with $x_1 \in \operatorname{dom} f$ and $x_2 \in \operatorname{dom} g$.

By Assumption 2.3(iii) there is a feasible point $\breve{x} = (\breve{y}, \breve{u})$ with positive minimal distance to the bounds, which implies in particular that $\breve{x} \in \operatorname{dom} F_\mu$. By Assumption 2.3(i), J is finite on X and thus also in the ball $\breve{x} + B_X$. Because $\breve{x} \in \operatorname{dom}(\iota_E + b)$ we can compute $B_X = (\breve{x} + B_X) - \breve{x} \subset \operatorname{dom} J - \operatorname{dom}(b + \iota_E)$, and we conclude that (5.8) is fulfilled for $\varepsilon = 1$. Thus $\partial(J + \iota_E + b) = \partial J + \partial(\iota_E + b)$.

Next we show that $\partial(\iota_E + b) = \partial\iota_E + \partial b$ by verifying (5.8) for b and ι_E. Here $Y \subset C(Q_Y)$ is crucial because it guarantees that $(\breve{y} + rB_Y, \breve{u}) \in \operatorname{dom} b$ for every $r < \tau$ via Assumption 2.3(iii).

By Lemma A.1 the control-to-state mapping $S = A^{-1}B : U \to Y$ is continuous. Thus there is $\delta > 0$ such that for each $v \in \delta B_U$ it follows that $Sv \in (r/2)B_Y$ with $ASv - Bv = 0$. Hence, $(\breve{y} + Sv, \breve{u} + v) \in \operatorname{dom}\iota_E$ and $(\breve{y} + Sv + w, \breve{u}) \in \operatorname{dom} b$ for all $w \in (r/2)B_Y$. Consequently, for sufficiently small ε and arbitrary $(w, v) \in \varepsilon B_X$ we have

$$\begin{array}{rcccc} w & = & \breve{y} + Sv & - & (\breve{y} + Sv + w), \\ v & = & \underbrace{\breve{u} + v}_{\in \operatorname{dom}\chi_E} & - & \underbrace{\quad\breve{u}\quad}_{\in \operatorname{dom} b} . \end{array}$$

This shows (5.8), and the sum-rule yields $0 \in \partial J(x) + \partial\iota_E(x) + \partial b(x)$.

Next we consider (5.6). Because $\operatorname{dom} b(u;\mu) \subset (Y, \breve{u})$ and $\operatorname{dom} b(y;\mu) \subset (\breve{y}, U)$, (5.8) can easily be verified for $b(y)$ and $b(u)$. Then the sum-rule yields

$$\partial b(x) = \partial(b(y) + b(u)) = \partial b(y) + \partial b(u).$$

By Proposition 4.9 we conclude

$$\partial b(y) + \partial b(u) = \partial \underline{b}(y) + \partial \overline{b}(y) + \partial \underline{b}(u) + \partial \overline{b}(u).$$

Because $b(x;\mu;q) = \mu^q b(x;1;q)$, (5.7) follows from the general relation $\partial(\lambda f) = \lambda \partial f$ for $\lambda > 0$ (cf., e.g., [7, I(5.21)]). ∎

THEOREM 5.4. *Suppose that Assumptions* 2.1 *and* 2.3 *hold, and let* $\mu > 0$. *Then* $x = (y, u)$ *is a minimizer of* F_μ *if and only if* x *is feasible with respect to the state equation and the inequality constraints and there are* $j = (j_y, j_u) \in \partial J(x)$, $\overline{m} \in \partial \overline{b}(y;\mu)$, $\underline{m} \in \partial \underline{b}(y;\mu)$, *and* $p \in \operatorname{dom} A^*$ *that satisfy*

$$\begin{aligned} j_y + \overline{m} + \underline{m} + A^* p &= 0, \\ j_u + \overline{b}'(u;\mu) + \underline{b}'(u;\mu) - B^* p &= 0. \end{aligned} \tag{5.9}$$

Proof. By definition of the subdifferential, x is a minimizer of F_μ if and only if $0 \in \partial F_\mu(x)$. By Lemma 5.3 this is equivalent to

$$0 \in \partial J(x) + \partial \iota_E(x) + \partial \underline{b}(y;\mu) + \partial \overline{b}(y;\mu) + \partial \underline{b}(u;\mu) + \partial \overline{b}(u;\mu).$$

Next consider the operator $T := (A, -B)$. By Lemma A.1(ii), T is closed, densely defined, and surjective, and $E = \ker T$. Hence, (B.2) yields $\partial \iota_E = \partial \iota_{\ker T} = \operatorname{ran} T^*$, with $T^* = (A^*, -B^*) : R^* \supset \operatorname{dom} T^* \to X^*$. So the following inclusion is equivalent to optimality of x:

$$0 \in \partial J(x) + \operatorname{ran} T^* + \partial \underline{b}(y;\mu) + \partial \overline{b}(y;\mu) + \partial \underline{b}(u;\mu) + \partial \overline{b}(u;\mu).$$

In other words, there are $j \in \partial J(x)$, $p \in \operatorname{dom} T^* \subset R^*$, $\underline{m} \in \partial \underline{b}(y;\mu)$, $\overline{m} \in \partial \underline{b}(y;\mu)$, $\underline{l} \in \partial \underline{b}(u;\mu)$, and $\overline{l} \in \partial \underline{b}(u;\mu)$ such that

$$0 = j + T^* p + \underline{m} + \overline{m} + \underline{l} + \overline{l} \text{ in } X^* \tag{5.10}$$

Clearly, $\underline{m}, \overline{m} \in Y^*$, $\underline{l}, \overline{l} \in U^*$, and $T^* p = (A^* p, -B^* p)$. Finally, Proposition 4.5 yields $\underline{l} = \underline{b}'(u;\mu)$ and $\overline{l} = \overline{b}'(u;\mu)$. Splitting (5.10) into two equations in Y^* and U^* yields (5.9). ∎

In many cases we obtain uniqueness results on the dual variables. This may be important for algorithms, because nonuniqueness of dual variables often implies singularity of Jacobians. We remark that similar uniqueness results for the original state constrained problem may not hold in general.

COROLLARY 5.5. *Under the hypothesis of Theorem* 5.4, *let* x *be a minimizer of* F_μ *for* $\mu > 0$ *and suppose that* J *is Gâteaux differentiable at* x. *Then* $(j_y, j_u) = J'(x)$ *in* (5.9), *and the following hold:*

(i) If B^* *is injective, then* $\underline{m}$, $\overline{m}$, *and* p *are unique.*

(ii) If y *is strictly feasible with respect to the state constraints, then* $\overline{m} = \overline{b}'(y;\mu)$, $\underline{m} = \underline{b}'(y;\mu)$, *and* p *are unique.*

Proof. By Theorem 5.4, (5.9) has a solution. Since J is Gâteaux differentiable, its subdifferential contains only one element that coincides with $j'(x)$ (cf. [7, Proposition I.5.3]).

To show (i) we remark that the second line of (5.9) yields uniqueness of p because B^* is injective. By the first line of (5.9), this implies uniqueness of $\underline{m}+\overline{m}$. Proposition 4.6 yields that $\underline{m}$ and $\overline{m}$ are unique up to the set where y touches the bounds. But since the corresponding sets for $\underline{m}$ and $\overline{m}$ are necessarily disjoint, $\underline{m}$ and $\overline{m}$ have to be unique.

By strict feasibility of y, (ii) follows from Proposition 4.6, which yields $\partial\overline{b}(y;\mu) = \overline{b}'(y;\mu)$ and $\partial\underline{b}(y;\mu) = \underline{b}'(y;\mu)$.

Because A is closed and bijective, it follows from the closed range theorem (cf. [9, Theorem IV.1.2]) that A^{-*} is also bijective, and thus p is unique. ∎

If F_μ is strictly convex, then x is the unique minimizer and Corollary 5.5 yields unique $(x,p,\overline{m},\underline{m})$ such that x is feasible and (5.9) is satisfied.

5.3. Bounds on the dual variables. Next we study the boundedness of the dual variables for $\mu \to 0$. This will be the basis for the estimates in section 6. We will show under reasonable assumptions that the dual variables remain uniformly bounded as μ tends to 0. Due to the importance of these estimates we provide accurate information, which may also be useful for later work.

PROPOSITION 5.6. *Suppose that Assumptions* 2.1 *and* 2.3 *hold, and let x be a minimizer of F_μ for $\mu > 0$. Denote by $j = (j_y, j_u)$, $\overline{m}$, and $\underline{m}$ corresponding solutions of* (5.9)*, and let τ and $\breve{x}$ be defined as in Assumption* 2.3*. Then*

$$\begin{aligned}
&\|\overline{m}\|_{M(Q_Y)} + \|\underline{m}\|_{M(Q_Y)} + \left\|\underline{b}'(u;\mu)\right\|_{L_1(Q_U)} + \|\overline{b}'(u;\mu)\|_{L_1(Q_U)} \\
&\quad\le \frac{2}{\tau}|\langle j, x-\breve{x}\rangle| + \frac{2^{q+2}\mu^q}{\tau^{q+1}}\,\|x-\breve{x}\|_{L_1(Q_Y)\times L_1(Q_U)} + \frac{2^{q+1}\mu^q}{\tau^q}(|Q_Y|+|Q_U|).
\end{aligned} \tag{5.11}$$

In particular, if Assumption 2.4(i) *holds, then* (5.11) *is bounded by a constant $\mathcal{C}_m$ that depends on the problem setting and on μ_0 but not on $\mu \in]0,\mu_0]$.*

Proof. In the following let $(y,u) = x$ be the minimizer of the barrier functional F_μ and $(\breve{y},\breve{u})$ a Slater point. Consider the optimality system (5.9). We multiply the first equation of (5.9) by $\delta y := y-\breve{y}$ and the second equation by $\delta u := u-\breve{u}$ to obtain

$$\begin{aligned}
0 &= \langle j_y + \overline{m} + \underline{m} + A^*p, \delta y\rangle + \langle j_u + \overline{b}'(u) + \underline{b}'(u) - B^*p, \delta u\rangle \\
&= \langle j_y + \overline{m} + \underline{m}, \delta y\rangle + \langle j_u + \overline{b}'(u) + \underline{b}'(u), \delta u\rangle + \langle p, A\delta y - B\delta u\rangle \\
&= \langle j_y + \overline{m} + \underline{m}, \delta y\rangle + \langle j_u + \overline{b}'(u) + \underline{b}'(u), \delta u\rangle,
\end{aligned}$$

since $A\delta y - B\delta u = 0$. By definition we have $\langle j_y,\delta y\rangle + \langle j_u,\delta u\rangle = \langle j,\delta x\rangle$, and we obtain by taking absolute values

$$\left|\langle \underline{m}, \delta y\rangle + \langle \overline{m}, \delta y\rangle + \langle \underline{b}'(u), \delta u\rangle + \langle \overline{b}'(u), \delta u\rangle\right| = |\langle j, \delta x\rangle|. \tag{5.12}$$

Next we show that even the absolute values of the summands in the left-hand side are bounded. For this let Q be one of $\underline{Q}_Y$, $\overline{Q}_Y$, $\underline{Q}_U$, or $\overline{Q}_U$ (cf. Assumption 2.3(ii)) and introduce for two functions v and w the following partition on Q:

$$\begin{aligned}
I(v,w) &:= \{t \in Q : |v(t) - w(t)| > \tau/2\}, \\
A(v,w) &:= \{t \in Q : |v(t) - w(t)| \le \tau/2\} = Q \setminus I(v,w),
\end{aligned}$$

with τ as defined in Assumption 2.3(iii). For a motivation of this splitting, consider, e.g., lower bounds $\underline{y}$ on $\underline{Q}_Y$. Then on $A(y,\underline{y})$ (the "almost active" part) we will apply positivity arguments, whereas on $I(y,\underline{y})$ (the "strongly inactive" part) bounds on $b'(y)$ are available.

In fact, according to Corollary 4.7, $\underline{m}$, $\overline{m}$, $\underline{b}'(u)$, and $\overline{b}'(u)$ are bounded pointwise a.e. on $I(y,\underline{y})$, $I(y,\overline{y})$, $I(u,\underline{u})$, and $I(u,\overline{u})$, respectively, by $(2\mu/\tau)^q$. Hence,

$$\left|\langle \underline{m}, \delta y\rangle_{I(y,\underline{y})}\right| \le (2\mu/\tau)^q \|\delta y\|_{L_1(Q_Y)},$$

and similar estimates hold for the other three similar terms. In (5.12) we shift these four terms to the right-hand side and obtain

$$\left|\langle \underline{m}, \delta y\rangle_{A(y,\underline{y})} + \langle \overline{m}, \delta y\rangle_{A(y,\overline{y})} + \langle \underline{b}'(u), \delta u\rangle_{A(u,\underline{u})} + \langle \overline{b}'(u), \delta u\rangle_{A(u,\overline{u})}\right| \le r,$$

with $r = |\langle j, x-\breve{x}\rangle| + 2\cdot(2\mu/\tau)^q \|x-\breve{x}\|_{L_1(Q_Y)\times L_1(Q_U)}$. It is checked easily that all four duality products on the left-hand side of this equation are positive. For example, we have

$$\underline{m} \le 0, \quad \delta y\Big|_{A(y,\underline{y})} = (y-\breve{y})\Big|_{A(y,\underline{y})} \le \tau/2 - \tau \le -\tau/2 \Rightarrow \langle \underline{m}, y-\breve{y}\rangle_{A(y,\underline{y})} \ge 0.$$

Hence, we can conclude

$$\left|\langle \underline{m}, \delta y\rangle_{A(y,\underline{y})}\right| + \left|\langle \overline{m}, \delta y\rangle_{A(y,\overline{y})}\right| + \left|\langle \underline{b}'(u), \delta u\rangle_{A(u,\underline{u})}\right| + \left|\langle \overline{b}'(u), \delta u\rangle_{A(u,\overline{u})}\right| \le r.$$

Finally, we prove bounds on the norms of $\underline{m}$, $\overline{m}$, $\underline{b}'(u)$, and $\overline{b}'(u)$. We perform the proof only for $\underline{m}$. The other cases are identical.

As already noted, $\underline{m} \le 0$ as a measure, and $\delta y < -\tau/2$ on $A(y,\underline{y})$. Hence,

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

and thus $\|\underline{m}\|_{M(A(y,\underline{y}))} \le 2/\tau |\langle \underline{m}, \delta y\rangle_{A(y,\underline{y})}|$. Summing up all similar expressions, we get

$$\|\underline{m}\|_{M(A(y,\underline{y}))} + \|\overline{m}\|_{M(A(y,\overline{y}))} + \left\|\underline{b}'(u)\right\|_{M(A(u,\underline{u}))} + \|\overline{b}'(u)\|_{M(A(u,\overline{u}))} \le \frac{2r}{\tau}.$$

Since $\|\underline{m}\|_{L_\infty(I(y,\underline{y}))} \le (2\mu/\tau)^q$, we get $\|\underline{m}\|_{M(I(y,\underline{y}))} \le (2\mu/\tau)^q |Q_Y|$ and a similar estimate for the three similar expressions. Now a short computation yields (5.11).

As for existence of $\mathcal{C}_m$, Theorem 5.2 yields existence and uniform boundedness of solutions $x(\mu)$ for $\mu \in [0;\mu_0]$ in X and thus also in $L_1(Q_Y)\times L_1(Q_U)$. Since, by Assumption 2.4(i), ∂J is bounded in $(X, \|\cdot\|)^*$ on bounded subsets of $(X, \|\cdot\|)$, the right-hand side of (5.11) is a uniform bound in $]0;\mu_0]$. ∎

5.4. Application to an elliptic control problem. To illustrate the application of our abstract optimality conditions we consider one of the optimal control problems from section 3. We choose distributed control in $\Omega \subset \mathbb{R}^d$ with $d \le 3$. Thus we choose $X = Y \times U := C(\overline{\Omega}) \times L_2(\Omega)$. For ease of writing we consider only lower bounds $\underline{y} = 0$ and Lipschitz continuous $\underline{u}$ on the control and on the state, which are chosen such that a Slater condition holds. Further, we choose a tracking type functional for $y_d \in L_2(\Omega)$:

$$J(y,u) := \frac{1}{2}\|y-y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Omega)}.$$

Clearly, J is continuous and even Gâteaux differentiable. Thus its subdifferential coincides with the derivative that we compute as $J'(y,u) = (j_y, j_u) = (y - y_d, \alpha u) \in Y^* \times U^*$, which is bounded on bounded subsets of X.

Our state equation is given by

$$\int_\Omega \langle \kappa \nabla y, \nabla v \rangle + ayv\, dt - \int_\Omega uv\, dt = 0 \quad \forall v \in H^1(\Omega),$$

and we assume for simplicity that κ, a, and $\partial\Omega$ are smooth. We have already demonstrated in section 3 how to fit this state equation into the framework of Assumption 2.1. In particular, we know that $y \in W^{1,s}(\Omega)$ for some $s > 3$ and thus $y \in C(\overline{\Omega})$. Hence, the above integrals are well defined not only for $v \in H^1(\Omega)$ but also for all $v \in W^{1,s'}(\Omega)$, with $s' = s/(s-1)$.

THEOREM 5.7. *The above optimal control problem satisfies Assumptions* 2.1*,* 2.3*, and* 2.4*. For every $q \geq 1$ barrier regularizations of the above problem possess a unique minimizer $x = (y,u)$ for each $\mu > 0$; y is strictly feasible with respect to the inequality constraints, except for a set Q_F of measure zero.*

For each μ there exists a unique adjoint state $p \in W^{1,s'}(\Omega)$ and a unique positive measure $m \in M(Q_F)$ such that the following equations are satisfied:

$$\int_\Omega (y - y_d)v - \frac{\mu^q}{y^q} v\, dt - \int_{Q_F} v\, dm + \int_\Omega \langle \kappa \nabla v, \nabla p \rangle + avp\, dt = 0 \;\forall v \in W^{1,s}(\Omega), \tag{5.13}$$

$$\alpha u(t) - \frac{\mu^q}{(u(t) - \underline{u}(t))^q} - p(t) = 0 \;\; a.e. \;\; in\; \Omega. \tag{5.14}$$

If y is strictly feasible with respect to the state constraints, then m vanishes. The following bounds hold uniformly in each positive interval $]0, \mu_0]$:

$$\|p\|_{W^{1,s'}} \leq C, \quad \|u\|_{W^{1,s'}} \leq C, \quad \|y\|_{W^{3,s'}} \leq C.$$

Here C may depend on the problem setting and μ_0 but not on $\mu \in]0, \mu_0]$.

Proof. Since Assumptions 2.1, 2.3(i)–(iii), and 2.4(i) have already been verified, it remains to consider Assumption 2.3(iv) (coercivity) and 2.4(ii) (strong convexity for the norm $\|\cdot\| := \|\cdot\|_X = \|\cdot\|_{C(\overline{\Omega})} + \|\cdot\|_{L_2(\Omega)}$). Because J is a quadratic functional, $J + b(x;\mu;q)$ is coercive and strongly convex in $L_2(\Omega) \times L_2(\Omega)$ for every $\mu > 0$. Continuity of $S = A^{-1}B : L_2(\Omega) \to C(\overline{\Omega})$ (cf. Lemma A.1) implies $\|(y,u)\|_X \leq (1 + \|S\|)\|u\|_{L_2(\Omega)}$ for (y,u) in the feasible subspace E (cf. Assumption 2.1). From this, coercivity and strong convexity of J on E, and thus on the feasible set in X, follow.

Now Theorem 5.2 yields existence of a unique minimizer for each $\mu > 0$. By Theorem 5.4 the optimality system (5.9) possesses a solution, which is unique by Corollary 5.5. The adjoint PDE (5.13) follows from the first line of (5.9), which reads in our case

$$j_y + A^* p + \underline{m} = 0. \tag{5.15}$$

Here we have to find a representation of the expression A^*p, with $A : C(\overline{\Omega}) \supset \operatorname{dom} A \to (W^{1,s'}(\Omega))^*$ with $s' < d/(d-1)$ as defined in section 3.1. Taking into account our choices of spaces, the definition $A^* : R^* \supset \operatorname{dom} A^* \to Y^*$ yields $A^* : W^{1,s'}(\Omega) \supset \operatorname{dom} A^* \to M(\overline{\Omega})$. Further, we know that $p \in \operatorname{dom} A^*$, and thus it follows by definition of the adjoint operator

$$\langle v, A^* p \rangle = \langle Av, p \rangle = \int_\Omega \langle \kappa \nabla v, \nabla p \rangle + avp\, dt \quad \forall v \in W^{1,s}(\Omega), \tag{5.16}$$

and this expression is continuous in v with respect to the norm $\|\cdot\|_{C(Q_Y)}$. Now (5.13) follows from (5.15) and from the representation of $\underline{m}$ via Proposition 4.6.

For the optimality condition (5.14), we take into account the definition of $B = I^*$, where $I : W^{1,s'}(\Omega) \to L_2(\Omega)$ is the Sobolev embedding. Hence, the second line of (5.9) reads in our case $\alpha u + b'(u;\mu) - Ip = 0$. Since $U^* = L_2(\Omega)$, we can write this more explicitly as an equation that holds pointwise almost everywhere and conclude (5.14). Since I is injective, Corollary 5.5 yields uniqueness of p and m.

Next we show the claimed uniform boundedness results on p, u, and y. Consider Proposition 5.6. It follows from Theorem 5.2 that y and u are uniformly bounded in $L_2(\Omega)$ (and thus also in $L_1(\Omega)$) as $\mu \to 0$, and thus $\langle j', x - \check{x}\rangle$ is uniformly bounded, too. Hence, obviously (5.11) is a uniform bound for $\mu \to 0$. Thus in (5.15) the summands $\underline{m}$ and j_y are uniformly bounded in $M(\overline{\Omega})$ and hence so is A^*p. From bijectivity of A we conclude bijectivity of A^* via the closed range theorem (cf. [9, Theorem IV.1.2]). Because all adjoint operators are closed, Lemma A.1 even yields continuous invertibility of A^* and thus uniform boundedness of p in $W^{1,s'}(\Omega)$ via (5.15). From (5.14) it follows from an elementary but tedious computation that the uniform bound for p in $W^{1,s'}(\Omega)$ implies a uniform bound for u in $W^{1,s'}(\Omega)$ if $\underline{u}$ is Lipschitz. We remark only that $u \to \max(p, \underline{u})$ pointwise a.e. and that the max-function defines a continuous mapping on $W^{1,s'}(\Omega)$. Inserting u into the state equation finally yields the stated bound for y if the coefficients of the PDE are sufficiently regular. ∎

6. The central path.

For a suitable interval $]0, \mu_0]$ Theorem 5.2 shows that there is a mapping from $]0, \mu_0]$ to X, which maps each value of μ to the unique minimizer $x(\mu)$ of the corresponding barrier problem. We will show now under Assumption 2.4 that this mapping actually defines a Lipschitz continuous and convergent path, which we will call the central path. It will turn out that there are precise estimates of the local Lipschitz constant and on the rate of convergence available. Simply speaking, the central path behaves like a square root function. Similar estimates have been for control constrained problems and under stronger smoothness assumptions on the functional (cf., e.g., [28, 29, 26]). 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 the subdifferential of the combined functional F_μ (cf. (5.2)). First we study convergence and continuity of the function values on the central path for $\mu \to 0$ and show in particular that the error in the function values is $O(\mu)$. Then we use the strong convexity assumption (2.3) to transfer these results to convergence and continuity estimates on (y, u).

Equipped with the insight of the preceeding 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 Q_Y. All results of this section remain valid for the general case, with the only exception that (6.1) and (6.14) have to be modified straightforwardly.

6.1. Convergence of the function values.

The following lemma is already the key to all results in section 6.

LEMMA 6.1. *Suppose that Assumptions* 2.1 *and* 2.3 *hold. Let* $x(\mu_0)$ *be a minimizer of* $F_{\mu_0} := F + b(\cdot;\mu_0;q)$. *Then for every* $x \in X$ *and for all* $\mu \in]0, \mu_0]$

$$F_\mu(x(\mu_0)) \leq F_\mu(x) + q(\mu_0 - \mu)\int_{y(\mu_0)>y} \frac{\mu_0^{q-1}}{y(\mu_0)^q}(y(\mu_0) - y)\,dt. \tag{6.1}$$

Let x_* be the minimizer of the original state constrained problem, given by F as defined in (5.1)*. If Assumption* 2.4(i) *holds or $q = 1$, then*

$$F(x(\mu)) \leq F(x_*) + C\mu. \tag{6.2}$$

Here $C = C(\mathcal{C}_m, |Q_Y|, q)$ ($\mathcal{C}_m$ as in Proposition 5.6) is independent of μ for $\mu \in]0; \mu_0]$.

Proof. We may assume $F_\mu(x) < \infty$ and hence feasibility of x. Otherwise, (6.1) is trivially fulfilled. By (5.5) and (5.7) we conclude for every $\mu > 0$ and $x \in X$

$$\partial F_\mu(x) = \partial(J + \iota_E)(x) + \partial b(y; \mu; q) = \partial(J + \iota_E)(x) + \mu^q \cdot \partial b(y; 1; q). \tag{6.3}$$

Let $x_0 := x(\mu_0)$ be a minimizer of F_{μ_0}, which is equivalent to $0 \in \partial F_{\mu_0}(x_0)$. Inserting $\mu = \mu_0$ and $x = x_0$ into (6.3) yields $m \in \partial b(y_0; 1; q)$ and $j^* \in \partial(J + \iota_E)$ such that

$$j^* + \mu_0^q \cdot m = 0. \tag{6.4}$$

Moreover, (5.7) yields $\mu^q \cdot m \in \mu^q \cdot \partial b(y_0; 1; q)$ for all $\mu > 0$. Thus, for every $\mu > 0$, (6.3) yields $j^* + \mu^q \cdot m \in \partial F_\mu(x_0)$, and we conclude from (6.4)

$$(\mu^q - \mu_0^q) \cdot m \in \partial F_\mu(x_0). \tag{6.5}$$

By definition of the subdifferential, $\langle x^*, x - x_0\rangle \leq F_\mu(x) - F_\mu(x_0)$ holds for every $x^* \in \partial F_\mu(x_0)$. Thus in particular for m as defined in (6.4) we have

$$F_\mu(x_0) \leq F_\mu(x) - (\mu^q - \mu_0^q)\langle m, y - y_0\rangle = F_\mu(x) + (\mu_0^q - \mu^q)\langle -m, y_0 - y\rangle. \tag{6.6}$$

By convexity of the function μ^q for $q \geq 1$ we have $\mu_0^q - \mu^q \leq q\mu_0^{q-1}(\mu_0 - \mu)$. Further, since we have only lower bounds, $-m$ is a positive measure. Hence, because $\mu_0 > \mu$ and regions with $y \leq y_0$ do not contribute positively to the integral, we can estimate

$$(\mu_0^q - \mu^q)\langle -m, y_0 - y\rangle \leq (\mu_0 - \mu)q\mu_0^{q-1}\langle -m, y_0 - y\rangle\Big|_{y_0 > y}. \tag{6.7}$$

By Proposition 4.6 and feasibility of y it follows that

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

The measure part of $-m$ does not appear, because $y_0 > y \Rightarrow y_0 > 0$. Inserting this estimate into (6.7) and then into (6.6) finally yields (6.1).

For the case $\mu = 0$, we observe that (6.4) yields $-\mu_0 \cdot m \in \partial(J + \chi_E)(x_0)$. Thus by definition of the subdifferential we obtain instead of (6.6)

$$F(x_0) = (J + \chi_E)(x_0) \leq (J + \chi_E)(x_*) + \mu_0^q\langle -m, y_0 - y_*\rangle = F(x_*) + \mu_0^q\langle -m, y_0 - y_*\rangle.$$

Again by Proposition 4.6 and feasibility of y_* we conclude

$$\mu_0^q\langle -m, y_0 - y_*\rangle\Big|_{y_0 > y_*} = \int_{y_0 > y_*} \frac{\mu_0^q}{y_0^q}(y_0 - y_*)\, dt \leq \mu_0 \int_{y_0 > y_*} \frac{\mu_0^{q-1}}{y_0^{q-1}}\, dt.$$

Here we used that $(y_0 - y_*)/y_0 \leq 1$. If $q = 1$, then (6.2) follows immediately. Otherwise, the last integral is uniformly bounded for $\mu \to 0$ due to Proposition 5.6 and (4.12). Hence, we obtain (6.2). $\square$

In the case of an upper bound, an estimate similar to (6.1) holds, where the domain of integration is then $\{t \in \Omega : y(\mu_0)(t) < y(t)\}$.

6.2. Convergence of the path. Convergence of the central path is a simple consequence of (6.2) and a strong convexity assumption. We obtain a rate of convergence, and even the corresponding constant is accessible via the estimate in Proposition 5.6.

LEMMA 6.2. *If J satisfies Assumption* 2.4(ii), *then F_μ satisfies a growth condition at its unique minimizer $x(\mu)$:*

$$\alpha \|x - x(\mu)\|^2 + b(x) + b(x(\mu)) - 2b\left(\frac{x + x(\mu)}{2}\right) \le F_\mu(x) - F_\mu(x(\mu)). \tag{6.8}$$

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

$$\alpha \|x - x_*\|^2 \le F(x) - F(x_*) \quad \forall\, x \in X. \tag{6.9}$$

Proof. We start with the computation:

$$\begin{aligned} F_\mu(x) &+ F_\mu(x(\mu)) - 2F_\mu\left(\frac{1}{2}x + \frac{1}{2}x(\mu)\right) \\ &\le F_\mu(x) + F_\mu(x(\mu)) - 2F_\mu(x(\mu)) = F_\mu(x) - F_\mu(x(\mu)). \end{aligned} \tag{6.10}$$

Now we take into consideration that $F_\mu = F + b(\cdot;\mu)$ and apply (2.3) to F while we leave b unchanged in (6.10). This yields (6.8), which in turn implies uniqueness of the minimizer. Since (6.10) also holds for F, (6.9) and thus uniqueness of x_* follow similarly. ∎

THEOREM 6.3. *Assume that Assumptions* 2.1, 2.3, *and* 2.4 *hold, and let $\mu_0 > 0$ be given. Denote by $x(\mu)$ the minimizer of F_μ and by x_* the minimizer of F. Then*

$$\|x(\mu) - x_*\| \le C\sqrt{\frac{\mu}{\alpha}}. \tag{6.11}$$

Here $C = C(\mathcal{C}_m, |Q_Y|, q)$ ($\mathcal{C}_m$ as in Proposition 5.6*) is independent of μ as long as $\mu \le \mu_0$.*

Proof. Let $\mu > 0$. Then by (6.2) and by the growth condition (6.9) we conclude (6.11). ∎

6.3. Lipschitz continuity of the path. Similar to our convergence result, we obtain quantitative information on the Lipschitz constant.

LEMMA 6.4. *For a barrier function $l(x)$ and $x_1 > x_2$ we have the following pointwise estimate on second order finite differences:*

$$\frac{1}{4}l''(x_1)(x_1 - x_2)^2 \le l(x_1) + l(x_2) - 2l\left(\frac{x_1 + x_2}{2}\right). \tag{6.12}$$

Proof. Let w.l.o.g. $x_2 > 0$. Otherwise, $l(x_2) = \infty$ and the inequality is trivially fulfilled.

Define $f(x) := l(x_1) + l(x) - 2l((x_1 + x)/2)$. Then $f_x(x) = l'(x) - l'((x_1 + x)/2)$ and $f_{xx}(x) = l''(x) - 1/2\, l''((x_1 + x)/2)$. Hence, $f(x_1) = f_x(x_1) = 0$. Setting $\Delta x := x_2 - x_1$ Taylor expansion of f at x_1 yields

$$f(x_2) = f(x_1) + f_x(x_1)\Delta x + \frac{1}{2}f_{xx}(x)(\Delta x)^2 = \frac{1}{2}f_{xx}(x)(\Delta x)^2$$

for some $x_2 \leq x \leq x_1$. Clearly, the positive rational function l'' is monotonically decreasing. This implies that $l''(x_1) \leq l''((x_1+x)/2) \leq l''(x)$. Hence,

$$f_{xx}(x) = l''(x) - 1/2\, l''((x_1+x)/2) \geq l''(x) - 1/2\, l''(x) = 1/2\, l''(x) \geq 1/2\, l''(x_1).$$

Thus we conclude (6.12) from the computation

$$l(x_1) + l(x_2) - 2l\left(\frac{x_1+x_2}{2}\right) = f(x_2) = \frac{1}{2} f_{xx}(x)(\Delta x)^2 \geq \frac{1}{4} l''(x_1)(\Delta x)^2. \quad \square$$

THEOREM 6.5. *Assume that Assumptions* 2.1, 2.3, *and* 2.4 *hold, and let* $\mu_0 > 0$ *be given. The central path is locally Lipschitz continuous in* $]0;\mu_0]$, *and for every* $\mu > 0$ *we have the estimate*

$$\|x(\mu) - x(\tilde{\mu})\| \leq \frac{C}{\sqrt{\alpha\mu}} |\mu - \tilde{\mu}| \quad \forall \tilde{\mu} \in]0;\mu_0]. \tag{6.13}$$

For $\tilde{\mu} \leq \mu \leq 2\tilde{\mu} \leq \mu_0$ *there is the additional estimate*

$$\left\| \sqrt{l''(y(\mu);\mu;q)}(y(\mu) - y(\tilde{\mu})) \right\|^2_{L_2(S)} \leq \frac{C}{\sqrt{\mu}} |\mu - \tilde{\mu}|. \tag{6.14}$$

In both cases $C = C(\mathcal{C}_m, |Q_Y|, q)$ *(*$\mathcal{C}_m$ *as in Proposition* 5.6*) is independent of* $\mu \in]0;\mu_0]$.

Proof. We first will show (6.13) 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 define $S := \{t \in \Omega : y_1(t) > y_2(t)\}$. First we observe that by (6.8) and (6.1)

$$\begin{aligned} \alpha \|\Delta x\|^2 + b(y_1;\mu_2) + b(y_2;\mu_2) - 2b\left(\frac{y_1+y_2}{2};\mu_2\right) &\leq F_{\mu_2}(x_1) - F_{\mu_2}(x_2) \\ &\leq q\Delta\mu \int_S \frac{\mu_1^{q-1}}{y_1^q} \Delta y \, dt = q\Delta\mu \left\| \frac{\mu_1^{q-1}}{y_1^q} \Delta y \right\|_{L_1(S)}. \end{aligned} \tag{6.15}$$

Next we apply Lemma 6.4 on S to obtain

$$\begin{aligned} \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|^2_{L_2(S)} &= \int_S \frac{\mu_2^q}{y_1^{q+1}} \Delta y^2 \, dt = \frac{1}{4} \int_S l''(y_1;\mu_2) \Delta y^2 \, dt \\ &\leq c\left(b(y_1;\mu_2) + b(y_2;\mu_2) - 2b\left(\frac{y_1+y_2}{2};\mu_2\right)\right). \end{aligned} \tag{6.16}$$

Combining (6.15) and (6.16), application of the Hölder inequality yields

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

Using (4.12) and Proposition 5.6 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

$$\alpha \left\|\Delta x\right\|^2 + \left\| \frac{\mu_2^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|^2_{L_2(S)} \le c\, \frac{\Delta \mu}{\sqrt{\mu_1}} \left\| \frac{\mu_1^{q/2}}{y_1^{(q+1)/2}} \Delta y \right\|_{L_2(S)}. \tag{6.17}$$

Subdividing this inequality by

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

yields the desired estimate (6.13) for our special case, taking into account that μ_1/μ_2 is bounded due to our assumption $\mu_1 \le 2\mu_2$. Further, we conclude (6.14), because $l''(y;\mu;q) = c\mu^q/y^{q+1}$.

For the general case first note that $\tilde{\mu} = 0$ is covered by Theorem 6.3. Otherwise, we can divide $[\tilde{\mu}, \mu]$ into finitely many subintervals $[\mu_{i-1}, \mu_i]$ with $\mu_{i-1} \le \mu_i \le 2\mu_{i-1}$ and apply our special case of (6.13) to each. We obtain the estimate

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

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 (6.13) via the computation

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

6.4. 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 5 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 $M(Q_Y)$. For state constrained problems without control constraints and functionals of the form $J(x) = J(y) + \alpha \|u\|^2_{L_2}$ we easily obtain $\|p(\mu) - p_*\|_{L_2} = O(\sqrt{\alpha\mu})$ via the second equation of the optimality system (5.9), which yields $p = j_u = 2\alpha u$ in that case.

7. Strict feasibility of the state. As we have seen in Proposition 4.6, 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 behavior can actually be observed, and we will discuss a way to avoid measures in subdifferentials. By Proposition 4.5 comparable issues do not appear for control constraints, because subdifferentials of barrier functionals for control constraints are in $L_{p'}(Q)$.

7.1. An example with lack of strict feasibility. We present a simple example where the logarithmic barrier method (i.e., $q = 1$) admits a minimizer for each $\mu \ge 0$ 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:

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

We use the unit ball in $\mathbb{R}^d$ as the domain Ω and impose homogeneous Dirichlet boundary conditions. In this setting y is a quadratic rotationally symmetric function of the form $y = uc_1 \left(1 - r^2\right), \quad r^2 := |t|^2$. Hence, $y \in C_0^\infty(\overline{\Omega})$ is perfectly regular. Application of the barrier method and elimination of the state y yield

$$J_\mu(u) = -u - \mu \int_\Omega \ln(1 - y(t))\, dt = -u - c\mu \int_{r=0}^1 r^{d-1} \ln(1 - uc_1 \left(1 - r^2\right))\, dr.$$

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 derivative of J yields

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

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

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

For $d > 2$ the value of this integral is finite, and we conclude for sufficiently small μ that $J_\mu'(u) = -1 + c\mu < 0$.

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.

7.2. Strict feasibility of the state. Our example shows that violation of strict feasibility of y may lead to a breakdown of methods that try to solve the *formal* optimality system in spite of Theorem 5.4 guaranteeing the existence of solutions of the optimality system. However, Proposition 4.6 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 compact set $Q_Y \subset \mathbb{R}^d$ satisfies a *cone property* (cf. [1, Definition IV.4.3]), which mainly states that there is a cone K (defined as the convex hull of a ball and a point in $\mathbb{R}^d$) such that each point t of Q_Y is a subset of a cone $K_t \subset Q_Y$, which is the image of a rigid motion of K. Clearly, Q_Y has nonempty interior in this case.

Define the following spaces. For $0 < \beta < 1$ let $C_H^\beta(Q_Y)$ be the space of Hölder continuous functions of order β (for $\beta = 1$, $C_H^\beta(Q_Y)$ is the space of Lipschitz functions). For $1 < \beta \leq 2$ let $C_H^\beta(Q_Y)$ be the space of functions with Hölder continuous first derivatives (for $\beta = 2$ the derivatives are Lipschitz functions).

LEMMA 7.1. *Let $Q_Y \subset \mathbb{R}^d$ be a compact set satisfying a cone property. Let $0 \leq y \in C_H^\beta(Q_Y)$, $1/y^q \in L_1(Q_Y)$, $0 < \beta \leq 1$. Assume that $q \geq d/\beta$. Then $1/y \in C(Q_Y)$.*

If additionally $y > 0$ on ∂Q_Y, then the same assertion holds for $1 < \beta \leq 2$.

Proof. We first consider the case $\beta \le 1$. Assume that without loss of generality $0 \in Q_Y$ and $y(0) = 0$. We will show that this contradicts the assumption $1/y^q \in L_1(Q_Y)$. By the Hölder continuity of y we infer that $y(t) < cr^\beta$ for all $t \in B(0,r) \cap Q_Y$. Using the cone property of Q_Y, there is a cone $K_0 \in Q_Y$, and we can compute for sufficiently small $R < 1$

$$\left\|y^{-q}\right\|_{L_1(\Omega)} \ge \int_{K_0} y(t)^{-q}\,dt \ge c\int_{[0,R]} r^{-\beta q} r^{d-1}\,dr \ge c\int_{[0,R]} r^{-1} dr = \infty,$$

which shows $1/y^q \notin L_1(Q_Y)$. Hence, $y > 0$ in Q_Y, which implies by compactness of Q_Y that there is $\psi > 0$ with $y \ge \psi$ and thus $1/y \in C(Q_Y)$.

If $1 < \beta \le 2$, then y is continuously differentiable, and by our assumption $y > 0$ on ∂Q_Y. Hence, if $y(t) = 0$, then y necessarily obtains a minimum at t, and $t \in \operatorname{int} Q_Y$. Assume again w.l.o.g. that $t = 0$. Then $\nabla y(0) = 0$, which implies by Taylor approximation $y(t) < cr^\beta$ for all $t \in B(0,r) \cap \overline{\Omega}$, even for our extended range $1 < \beta \le 2$. Hence, we can apply the same argument as above to this case. ∎

If we compute the integrals, then the assertion of Lemma 7.1 can be quantified in the following sense. Assume that $y(\mu)$ is uniformly bounded in $C_H^\beta(Q_Y)$ for $0 < \beta \le 2$. Then for every $q \ge d/\beta$ there is a function $\psi(\mu)$ that is *uniformly bounded* on each *positive* compact interval $[\underline{\mu}, \mu_0]$ such that $\min_{t\in\Omega} y(\mu)(t) \ge \psi(\mu)$. More specifically $\psi(\mu)$ must fulfill the following inequality:

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

Hence, we obtain $\|1/y(\mu)\|_{C(Q_Y)} \le \psi(\mu)^{-1}$ and analogously $\|b'(y(\mu))\|_{C(Q_Y)} \le \widetilde{\psi}(\mu)$.

7.3. The central path for an elliptic control problem. To illustrate the results of sections 6 and 7 we continue the study of the elliptic control problem in section 5.4. It has already been shown to satisfy Assumptions 2.1, 2.3, and 2.4.

THEOREM 7.2. *Consider the optimal control problem defined in section* 5.4*, and let* $\mu_0 > 0$*. In addition to the results of Theorem* 5.7 *we conclude the following assertions.*

(i) *The central path* $\mu \to x(\mu)$ *for* $\mu \in]0;\mu_0]$ *is bounded in* $W^{3,s'}(\Omega) \times W^{1,s'}(\Omega)$ *for all* $1 < s' < d/(d-1)$.
(ii) *It is locally Lipschitz continuous and convergent in* $H^2(\Omega) \times L_2(\Omega)$*. The Lipschitz constant behaves like* $O(\mu^{-1/2})$*, and the rate of convergence is* $O(\sqrt{\mu})$.
(iii) *The function values* $J(x(\mu))$ *converge with the linear rate* $O(\mu)$.
(iv) *If the barrier function is chosen of order* $q > d$*, then* $y(\mu)$ *is strictly feasible and uniformly bounded away from* $\underline{y}$ *on every positive compact interval* $[\underline{\mu}, \mu_0]$*. In this case the measure part in* (5.13) *vanishes for all* $\mu > 0$.

Proof. Boundedness of the set of solutions has already been shown in Theorem 5.7. Further, the results in section 6 directly apply for $X = C(\overline{\Omega}) \times L_2(\Omega)$ and yield Lipschitz continuity and convergence of the central path in these norms. Since the control to state mapping $S : A^{-1}B : u \to y(u)$ is continuous from $L_2(\Omega)$ to $H^2(\Omega)$, the same results hold even for these stronger norms. Linear convergence of the function values $J(x(\mu)) = F(x(\mu))$ is a direct consequence of Lemma 6.1.

Finally, if $s' < d/(d-1)$ is large enough, the Sobolev embedding theorems [1, Theorem 5.4] yield $W^{3,s'} \hookrightarrow C_H^\beta(\overline{\Omega})$ for all $\beta \le 2$, $\beta < 2$, and $\beta < 1$ in the cases $d = 1, 2$, and 3, respectively. In any case, if $q > d$, then $q \ge d/\beta$ for some β in these sets. Hence, we can apply the results from section 7.2, remarking that smoothness of $\partial\Omega$ implies the uniform cone condition. ∎

7.4. Discretized control problems. Finally, we briefly discuss the analytic properties of the central path for a sequence of finite element spaces Y_h, $h \to 0$, each of which is closed in $C(Q_Y)$. This holds, for example, for standard Lagrange elements. We do *not* assume that the continuous problem can be treated in $C(Q_Y)$. Examples for such a case would be the problem in section 5.4 for spacial dimension $d \geq 4$ or Neumann boundary control for $d \geq 3$.

In all those cases the control-to-state mapping S fails to be well defined and continuous as a mapping from $U = L_2(Q_U)$ to $C(Q_Y)$. Hence, continuity of the discrete mappings $S_h : U \to (Y_h, \|\cdot\|_{C(Q_Y)})$ degenerates with $h \to 0$. This is, however, usually not the case if we consider $(Y_h, \|\cdot\|_{L_2(Q_Y)})$ instead . Hence, if J is a quadratic functional, then discrete central path $x_h(\mu)$ remains bounded in $L_2(Q_Y) \times L_2(Q_U)$, independent of μ and h. A look at Proposition 5.6 shows that this is sufficient to prove uniform boundedness on the constant $\mathcal{C}_m$, provided the Slater condition holds with τ bounded from below as $h \to 0$.

This in turn shows that all constants used in section 6 remain bounded as $h \to 0$. So the results in this section hold independently of the mesh-size, even in cases where the limit problem is irregular. Of course, this is different in section 7. Although all $x_h(\mu)$ can be assumed Lipschitz continuous for each h, the corresponding Lipschitz constant and thus the distance to the bounded may degenerate as $h \to 0$. But, if $q > d$, no additional measures will appear in the subdifferential of the barrier terms, because all $x_h(\mu)$ are then strictly feasible.

8. Conclusion and outlook. Recapitulating our results, we conclude that on the side of the primal variables (y, u) barrier methods can rely on rather *strong and robust convergence and continuity properties* of the central path. On the dual side, 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 often 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 semilinear 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 [21] together with a proof of convergence. For an efficient variant we refer to the recent work [23].

Appendix A. Closed operators. We consider some basic concepts of the theory of unbounded operators that we have used in our work. Unbounded operators possess a rich theory that generalizes the theory of continuous operators in many respects while retaining a large number of important results. For a detailed exposition we refer to [9], but most textbooks of functional analysis contain an introduction to unbounded operators. Consider Banach spaces V and W. An unbounded operator

$$T : V \supset \operatorname{dom} A \to W$$

is usually not defined everywhere but has a *domain of definition* $\operatorname{dom} T$. We say that T is *injective*, *surjective*, or *bijective* if the mapping $T : \operatorname{dom} T \to W$ has this algebraic property. If $\operatorname{dom} T$ is dense in V, then T is called *densely defined*.

A standard regularity assumption for unbounded operators is *closedness*. T is called *closed* if $\operatorname{dom} T \supset v_k \to v$ and $Tv_k \to w$ imply $v \in \operatorname{dom} T$ and $Tv = w$.

Closedness has also a geometrical interpretation: T is closed if and only if

$$\operatorname{graph}(T) := \{(v,w) \in \operatorname{dom} T \times W : Tv = w\} \subset V \times W \tag{A.1}$$

is closed in $V \times W$. Continuous operators that are defined on the whole domain space are closed, because $v_k \to v$ already implies $Tv_k \to Tv$. Closed operators have closed kernels, which follows from considering sequences with $Tv_k = 0$.

LEMMA A.1. *A linear operator $T : V \supset \operatorname{dom} T \to W$ is closed and bijective if and only if it has a continuous inverse $T^{-1} : W \to \operatorname{dom} T \subset V$.*

Under Assumption 2.1 *the following assertions hold:*

(i) *There is a continuous control-to-state mapping $S := A^{-1}B : U \to Y$. The feasible subspace E is closed, and every bounded sequence in E has a weak accumulation point.*

(ii) *The operator $(A, -B) : X \supset \operatorname{dom} A \times U \to R$ is closed, surjective, and densely defined.*

Proof. Clearly, bijectivity of T is equivalent to existence of an algebraic inverse T^{-1}. Then the graphs of T and of T^{-1} are related by

$$\begin{aligned}\operatorname{graph}(T) &= \{(v,w) : Tv = w\} = \{(v,w) : v = T^{-1}w\}\\ &= \{(v,w) : (w,v) \in \operatorname{graph}(T^{-1})\}.\end{aligned}$$

If T^{-1} is continuous and defined on all of W, its graph is closed and thus the graph of T is closed, too. In converse, bijectivity and closedness of T and completeness of V and W admit the application of the open mapping theorem (cf., e.g., [32, Satz IV.4.4] or [9, Theorem II.1.8]) for closed operators, which yields continuity of T^{-1}.

Let us turn to (i). By completeness of Y and R and since A is closed and bijective, A^{-1} exists and is continuous. By continuity of B, $S := A^{-1}B$ is continuous, which implies that $E = \{(y,u) : Ay - Bu = 0\} = \{(y,u) : y = Su\}$ is closed. By reflexivity of U every bounded sequence u_k in U has a weak accumulation point u. Because S is continuous, the sequence $y_k := Su_k$ also has a weak accumulation point y. Hence, every sequence $(y_k, u_k) = (Su_k, u_k)$ in E has a weak accumulation point, which is contained in E by (weak)-closedness of E.

Next we show (ii). Because A is bijective and densely defined, $(A, -B)$ is surjective and densely defined. Moreover, by bijectivity of A,

$$\begin{aligned}\operatorname{graph}(A,-B) &= \{(y,u,r) : Ay - Bu = r\}\\ &= \{(y,u,r) : A^{-1}r + Su = y\} = \{(y,u,r) : (r,u,y) \in \operatorname{graph}(A^{-1}, S)\}.\end{aligned}$$

Since $(A^{-1}, S) : R \times U \to Y$ is continuous and thus closed, $(A, -B)$ is closed, too. ∎

Let us recapitulate the definition of the adjoint of a *densely defined* operator $T : V \supset \operatorname{dom} T \to W$, which generalizes the adjoint of a continuous operator. Define

$$\operatorname{dom} T^* := \{w^* \in W^* : \text{ the linear functional } \langle w^*, T\cdot\rangle \text{ is continuous on } \operatorname{dom} T\}.$$

If $w^* \in \operatorname{dom} T^*$, then $\langle w^*, T\cdot\rangle$ has a unique continuous extension to a functional $v^* = T^*w^* \in V^*$, because it is continuous on the dense subset $\operatorname{dom} T \subset V$. This yields the definition of $T^* : W^* \supset \operatorname{dom} T^* \to V^*$ and the relation

$$\langle T^*w^*, v\rangle = \langle w^*, Tv\rangle \quad \forall\, v \in \operatorname{dom} T\; \forall\, w^* \in \operatorname{dom} T^*.$$

Adjoint operators are always closed by [9, Theorem II.2.6].

Appendix B. Convex analysis. We will introduce some basic concepts and tools from convex analysis. For more details on convex analysis we refer to [7, Chapter I] or [4, Chapter 4].

We consider *extended real valued functions* $f : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$. This makes it possible to consider constrained and unconstrained optimization problems in one framework by setting $f = \infty$ for infeasible points. The set of points $\operatorname{dom} f$ where f takes a finite value is called the *effective domain* of f. Apart from *convexity*, standard assumptions on f are *lower semicontinuity* (i.e., the sets $\{x \in X : f(x) \le \alpha\}$ are closed for all $\alpha \in \overline{\mathbb{R}}$) and *properness*: $\operatorname{dom} f \neq \emptyset$.

The *indicator function* ι_M of a set $M \subset X$ is defined by

$$\iota_M(x) = \begin{cases} 0 & x \in M, \\ \infty & \text{otherwise.} \end{cases} \tag{B.1}$$

It is convex and lower semicontinuous if and only if M is convex and closed, respectively, and $\operatorname{dom} \iota_M = M$. It follows from the definition that $\iota_{M_1} + \iota_{M_2} = \iota_{M_1 \cap M_2}$.

In convex analysis the usual differentiability concept is replaced by subdifferentiability. The *subdifferential* $\partial f(x)$ of $f : X \to \overline{\mathbb{R}}$ at a point $x \in \operatorname{dom} f$ is the set of all $x^* \in X^*$ for which the relation $\langle x^*, \hat{x} - x\rangle \le f(\hat{x}) - f(x)$ holds for all $\hat{x} \in X$. If $f(x) = \infty$, then $\partial f(x)$ is defined to be the empty set. If f is convex and Gâteaux differentiable at x with derivative $f'(x)$, then $\partial f(x) = \{f'(x)\}$ (cf. [7, Proposition I.5.3]). Further, minimizers x_{opt} of f are characterized by $0 \in \partial f(x_{opt})$, since $0 \in \partial f(x_{opt}) \Leftrightarrow 0 \le f(x) - f(x_{opt}) \forall x \in X$.

Lemma B.1. *Let X and R be Banach spaces and $T : X \supset \operatorname{dom} T \to R$ a closed, densely defined, linear operator. Denote by $\chi_{\ker T}$ the indicator function of* $\ker T$. *If T has closed range, then*

$$\partial \iota_{\ker T}(x) = \operatorname{ran} T^* \quad \forall x \in \ker T. \tag{B.2}$$

Proof. By definition $\iota_{\ker T}(x) = 0$ if $x \in \ker T$ and $+\infty$ otherwise. It is easily verified that $x^* \in \partial \iota_{\ker T}(x)$ if and only if $\langle x^*, \tilde{x} - x\rangle \le 0$ for all $\tilde{x} \in \ker T$. Because $\ker T$ is a linear subspace of X, it follows that $\partial \iota_{\ker T}(x) = (\ker T)^\perp$. Now (B.2) is a consequence of the *closed range theorem* (cf. [9, Theorem IV.1.2]) which states $(\ker T)^\perp = \operatorname{ran} T^*$ if T is closed, densely defined with closed range. ∎

Appendix C. Measure and integration. For the definition and analysis of barrier functionals, basic results from the theory of measure and integration have been used (cf., e.g., [8, 19]). In section 2.1 and Definition 4.2 we consider sets Q, equipped with certain measures. Technically, this means the following:

For the control space U and the L_p-theory of barrier functionals, we consider a *finite measure space* $(Q, \Sigma, |\cdot|)$ [19, Definition 1.18] with $Q \subset \mathbb{R}^n$, a σ-algebra Σ, and a positive finite measure $|\cdot|$. In this setting, the convergence theorems of Beppo Levi (in its version for positive integrals, [19, Theorem 1.27]) and Lebesgue [19, Theorem 1.34] are available, as well as the construction of L_p spaces [19, Definition 3.6] and the characterization of their dual spaces [19, Theorem 6.16]. Finiteness of $|Q_U|$ is used for the relation $L_{p_1}(Q) \subset L_{p_2}(Q)$ for $p_1 \ge p_2$. This generality is useful for the consideration of various control spaces, such as the ones defined in section 3.2.

For the space of states Y and theory of barrier functionals in spaces of continuous functions, we need additional results concerning the interplay of continuous functions and measures and thus consider for the construction of barrier integrals the finite measure space $(Q, \Sigma, |\cdot|)$, where $Q \subset \mathbb{R}^d$ is *compact*, Σ is the *Borel* σ-algebra on Q, and $|\cdot|$ is a regular positive Borel measure [19, Definition 2.15].

In both settings, the measure $|\cdot|$ should be seen as a *natural* underlying measure, which is mainly used to define integrals, denoted by $\int \cdot\, dt$, and sets of zero measure.

Apart from these underlying measures, the *space* of all (signed) regular Borel measures $M(Q)$ appears in the characterization of the subdifferentials of barrier functionals, since this is the representation of the dual space $C(Q)^*$ [19, Theorem 6.19]. Regularity of these measures is used for the characterization of barrier subdifferentials in Proposition 4.6. Also density of continuous functions in $L_p(S)$ spaces [19, Theorem 3.14] on open subsets $S \subset Q$ is used.

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

REFERENCES

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

[2] J.-J. Alibert and J.-P. Raymond, *Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded controls*, Numer. Funct. Anal. Optim., 3/4 (1997), pp. 235–250.

[3] H. Amann, *Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems*, in Function Spaces, Differential Operators and Nonlinear Analysis, H. J. Schmeisser and H. Triebel, eds., Teubner, Stuttgart, Leipzig, 1993, pp. 9–126.

[4] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, CMS Books Math. Ouvrages Math. SMC, Springer, New York, 2005.

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

[6] E. Casas, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim., 31 (1993), pp. 993–1006.

[7] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.

[8] J. Elstrod, *Maß- und Integrationstheorie*, 3rd ed., Springer, Berlin, 2002.

[9] S. Goldberg, *Unbounded Linear Operators*, Dover, New York, 1966.

[10] C. Grossmann and M. Zadlo, *A general class of penalty/barrier path-following Newton methods for nonlinear programming*, Optimization, 54 (2005), pp. 161–190.

[11] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2002), pp. 865–888.

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

[13] M. Hintermüller and K. Kunisch, *Path-following methods for a class of constrained minimization problems in function space*, SIAM J. Optim., 17 (2006), pp. 159–187.

[14] M. Hinze and A. Schiela, *Discretization of Interior Point Methods for State Constrained Elliptic Optimal Control Problems: Optimal Error Estimates and Parameter Adjustment*, Technical report SPP1253-08-03, Priority Program 1253, German Research Foundation, Bonn, Germany, 2007.

[15] K. Ito and K. Kunisch, *Semi-smooth Newton methods for state-constrained optimal control problems*, Systems Control Lett., 50 (2003), pp. 221–228.

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

[17] C. Meyer, U. Prüfert, and F. Tröltzsch, *On two numerical methods for state-constrained elliptic control problems*, Optim. Methods Softw., 22 (2007), pp. 871–899.

[18] U. Prüfert, F. Tröltzsch, and M. Weiser, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints*, Comput. Optim. Appl., 39 (2008), pp. 183–218.

[19] W. Rudin, *Real and Complex Analysis*, 2nd ed., McGraw–Hill, New York, 1974.

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

[21] 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, Berlin, 2006.

[22] A. Schiela, *An extended mathematical framework for barrier methods in function space*, in Domain Decomposition Methods in Science and Engineering XVIII, Lect. Notes Comput. Sci. Eng. 70, M. Bercovier, M. J. Gander, R. Kornhuber, and O. Widlund, eds., Springer, Berlin, to appear.

[23] A. Schiela, *An Interior Point Method in Function Space for the Efficient Solution of State Constrained Optimal Control Problems*, ZIB report 07-44, Zuse Institute Berlin, Berlin, 2008.

[24] M. Ulbrich, *Nonsmooth Newton-like Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, Habilitation thesis, Fakultät für Mathematik, Technische Universität München, Munich, 2002.

[25] 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 (2000), pp. 1938–1984.

[26] M. Ulbrich and S. Ulbrich, *Primal-dual interior point methods for PDE-constrained optimization*, Math. Program., 117 (2009), pp. 435–485.

[27] 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 (1999), pp. 731–764.

[28] M. Weiser, *Function Space Complementarity Methods for Optimal Control Problems*, Ph.D. thesis, Free University of Berlin, Dept. Math. and Comp. Sci., Berlin, 2001.

[29] M. Weiser, *Interior point methods in function space*, SIAM J. Control Optim., 44 (2005), pp. 1766–1786.

[30] M. Weiser, T. Gänzler, and A. Schiela, *A control reduced primal interior point method for PDE constrained optimization*, Comput. Optim. Appl., 41 (2008), pp. 127–145.

[31] M. Weiser and A. Schiela, *Function space interior point methods for PDE constrained optimization*, Proc. Appl. Math. Mech., 4 (2004), pp. 43–46.

[32] D. Werner, *Funktionalanalysis*, 3rd ed., Springer, Berlin, 2000.