
An Extended Mathematical Framework for Barrier Methods in Function Space

Anton Schiela¹

Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany, schiela@zib.de

1 Convex State Constrained Optimal Control

In this note, we extend the mathematical framework in [7] of barrier methods for state constrained optimal control problems with PDEs to a more general setting. In [7] we modelled the state equation by $Ly = u$ with L a closed, densely defined, surjective operator. This restricts the applicability of our theory mainly to certain distributed control problems. Motivated by the discussion in [6], we consider in this work operator equations of the more general form $Ay - Bu = 0$, where A is closed, densely defined and with closed range and B is continuous. While this change in framework only necessitates minor modifications in the theory, it extends its applicability to large additional classes of control problems, such as boundary control and finite dimensional control.

To make this paper as self contained as possible, assumptions and results of [7] are recapitulated, but for brevity proofs and more detailed information are only given when there are differences to [7]. This is possible, because our extension has only a very local effect.

Let Ω be an open and bounded Lipschitz domain in \mathbb{R}^d and $\overline{\Omega}$ its closure. Let $Y := C(\overline{\Omega})$ and $U := L_2(Q)$ for a measurable set Q , equipped with an appropriate measure. Standard examples are $Q = \Omega$ with the Lebesgue measure for distributed control, $Q = \partial\Omega$ with the boundary measure for boundary control and $Q = \{1, 2, \dots, n\}$ with the counting measure for finite dimensional controls.

Define $X := Y \times U$ with $x := (y, u)$ and consider the following convex minimization problem, the details of which are fixed in the rest of Section 1.

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

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

1.1 Linear Equality Constraints

By the equality constraint $Ay - Bu = 0$ we model a partial differential equation (cf. Section 1.3 below).

Assumption 1 *Let R be a Banach space. Assume that $B : U \rightarrow R$ is a continuous linear operator and that $A : Y \supset \text{dom}A \rightarrow R$ is a densely defined and closed linear operator with a closed range.*

Assume that there is a finite dimensional subspace $V \subset U$ of essentially bounded functions on Q , such that $R = \text{ran}A \oplus B(V)$, i.e., for each $r \in R$ there are unique $r_Y \in \text{ran}A$ and $r_V \in B(V)$ with $r = r_Y + r_V$.

Closed operators are a classical concept of functional analysis. For basic results we refer to [9, Kapitel IV.4] for more details, see [5]. In many applications A is bijective, i.e., the equation $Ay = r$ has a unique solution y for all $r \in R$. However, there are several important cases (such as pure Neumann problems), where only a Fredholm alternative holds while the corresponding optimal control problems are still well posed. Introduction of V includes these cases. If A is surjective, then $V = \{0\}$. Consider now the operator

$$\begin{aligned} T : Y \times U \supset \text{dom}A \times U &\rightarrow R \\ (y, u) &\mapsto Ay - Bu. \end{aligned} \quad (2)$$

From our assumptions it can be shown easily that T is densely defined, closed and surjective. Since T is closed, $E := \ker T$ is a closed subspace of X .

By density of $\text{dom}A$ in Y , we can define an adjoint operator A^* . For every $l \in R^*$ the mapping $y \rightarrow \langle l, Ay \rangle$ is a linear functional on $\text{dom}A$. We define $\text{dom}A^*$ as the subspace of all $l \in R^*$ for which $y \rightarrow \langle l, Ay \rangle$ is continuous on $\text{dom}A$ and can thus by density be extended uniquely to a continuous functional on Y . Hence, for all $l \in \text{dom}A^*$ there is a unique $A^*l \in Y^*$ for which $\langle l, Ay \rangle = \langle A^*l, y \rangle \forall y \in \text{dom}A$. This defines $A^* : R^* \supset \text{dom}A^* \rightarrow Y^*$.

1.2 Inequality Constraints and Convex Functionals

The inequality constraints in (1) are interpreted to hold pointwise almost everywhere and define a closed convex set of $G \subset X$. Some of the inequality constraints may not be present.

Assumption 2 *Assume that $E = \ker T$ is weakly sequentially compact. Assume that there is a strictly feasible point $\check{x} = (\check{y}, \check{u}) \in E$, which satisfies*

$$0 < d_{\min} := \text{ess inf}_{t \in \bar{\Omega}} \min \{ \check{u}(t) - \underline{u}(t), \bar{u}(t) - \check{u}(t), \check{y}(t) - \underline{y}(t), \bar{y}(t) - \check{y}(t) \}. \quad (3)$$

Assume that $J : X \rightarrow \bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous, convex, and coercive on the feasible set $E \cap G$, that J is continuous at \check{x} (cf. (3)) and that its subdifferential ∂J is uniformly bounded in X^ on bounded sets of X .*

Weak sequential compactness of E can usually shown by taking into account slightly stronger regularity properties of A . Often $\text{dom}A$ is contained in a reflexive (Sobolev)-Space.

Denote by $\chi_C(x)$ the indicator function of a set $C \subset X$, which vanishes on C and is $+\infty$ otherwise. Then we can rewrite (1) as an unconstrained minimization problem defined by the functional:

$$\begin{aligned} F : X &\rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \\ F &:= J + \chi_E + \chi_G. \end{aligned} \quad (4)$$

By our assumptions F is a lower semi-continuous, convex, and coercive functional with a non-empty domain and does thus admit a minimizer by weak compactness of E (cf. e.g. [4, Prop. II.1.2]).

Assumption 3 Assume that F is strongly convex (w.r.t. some norm $\|\cdot\|$):

$$\exists \alpha > 0 : \alpha \|x - y\|^2 \leq F(x) + F(y) - 2F\left(\frac{x+y}{2}\right) \quad \forall x, y \in \text{dom}F \quad (5)$$

Usually, optimal control problems with Tychonov regularization satisfy (5).

1.3 Example: A class of Elliptic PDEs

To illustrate our theoretical framework, we consider a class of elliptic PDEs, which was analysed by Amann [1] in an even more general framework.

Let Ω be a bounded domain in \mathbb{R}^d with a smooth boundary Γ . Let $\mathbf{a} \in C(\overline{\Omega}, \mathbb{R}^{d \times d})$, $\mathbf{b}, \mathbf{c} \in C(\overline{\Omega}, \mathbb{R}^d)$, $a_0 \in L_\infty(\Omega)$, $b_0 \in C(\Gamma)$. Assume that \mathbf{a} is symmetric positive definite, uniformly in Ω . Denote by $\gamma(\cdot) : W^{1,s} \rightarrow L_2(\Gamma)$ the boundary trace operator, which exists continuously if $s > 3/2$. For $1 < q < \infty$ and $1/q + 1/q' = 1$ consider the following continuous elliptic differential operator in the weak formulation:

$$\begin{aligned} A : W^{1,q}(\Omega) &\rightarrow (W^{1,q'}(\Omega))^* \\ \langle Ay, p \rangle &:= \int_{\Omega} \langle \mathbf{a} \nabla y + \mathbf{b} y, \nabla p \rangle + \langle \nabla y, \mathbf{c} p \rangle + a_0 y p \, dt + \int_{\Gamma} b_0 \gamma(y) \gamma(p) \, ds. \end{aligned} \quad (6)$$

Let $f \in (W^{1,q'}(\Omega))^*$. By [1, Theorem 9.2] a Fredholm alternative holds for the solvability of the equation $Ay = f$. This means that either it is uniquely solvable, or the homogenous problem has a finite dimensional space of nontrivial solutions with basis vectors $w_i \in W^{1,q}(\Omega)$. Then there is a finite number of conditions $\langle w_i, f \rangle = 0$ under which $Ay = f$ is non-uniquely solvable. This implies that A has a closed range with finite codimension and a kernel of the same dimension. In case of solvability, we have (cf. [1, 9.3(d)]):

$$\|y\|_{W^{1,q}} \leq C(\|f\|_{(W^{1,q'})^*} + \|y\|_{(W^{1,q'})^*}). \quad (7)$$

If $q > d$, then by the Sobolev embedding theorems $W^{1,q}(\Omega) \hookrightarrow C(\overline{\Omega})$ and we may redefine A as an unbounded operator

$$A : C(\overline{\Omega}) \supset W^{1,q}(\Omega) \rightarrow (W^{1,q'}(\Omega))^*.$$

Since $C^\infty(\overline{\Omega})$ is dense in $C(\overline{\Omega})$, A is densely defined, and closedness of A follows easily from (7), continuity of the embedding $W^{1,q}(\Omega) \hookrightarrow C(\overline{\Omega})$, and closedness of $\text{ran}A$. Hence, setting $Y := C(\overline{\Omega})$, $R := (W^{1,q'}(\Omega))^*$, and $\text{dom}A := W^{1,q}(\Omega)$, A fits into our framework. Its adjoint operator

$$A^* : W^{1,q'}(\Omega) \supset \text{dom}A^* \rightarrow C(\overline{\Omega})^*$$

is defined by $\langle y, A^*p \rangle = \langle Ay, p \rangle$ via the right hand side in (6). This expression is well defined for all $y \in \text{dom}A = W^{1,q}(\Omega)$, and $\text{dom}A^*$ is the set of all p , for which $\langle Ay, p \rangle$ is continuous on $\text{dom}A$ with respect to $\|y\|_\infty$ and have thus a unique continuous extension to an element of $C(\overline{\Omega})^*$.

By the choice of B we select how the control acts on the state. Two examples are distributed control

$$B_\Omega : L_2(\Omega) \rightarrow (W^{1,q'}(\Omega))^* \quad \langle B_\Omega u, p \rangle := \int_\Omega u \cdot p \, dt,$$

and Neumann or Robin boundary control

$$B_\Gamma : L_2(\Gamma) \rightarrow (W^{1,q'}(\Omega))^* \quad \langle B_\Gamma u, p \rangle := \int_\Gamma u \cdot \gamma(p) \, ds.$$

If $q' < d/(d-1)$ is chosen sufficiently large, B_Ω is continuous by the Sobolev embedding theorem for $d \leq 3$ and B_Γ is continuous by the trace theorem for $d \leq 2$. If $d = 3$, then $\gamma : W^{1,q'} \rightarrow L_2(\Gamma)$ is not continuous and thus the case $d = 3$ is not included in our framework for B_Γ . This has been a principal problem for the analysis (not only for barrier methods) of state constrained optimal control problems (cf. e.g. [3]). However, in [8] new techniques have been developed to overcome this restriction, which are likely to carry over to the analysis of barrier methods.

If $Ay = f$ is not uniquely solvable, then we have to assert that $u \in U$ can be split into $u = u_Y + u_V$, such that $\langle w_i, Bu_Y \rangle = 0$ and $u_V \in L_\infty$. Since all $w_i \in W^{1,q}$ are bounded, such an u_V can easily be constructed from these w_i in our cases $B = B_\Omega$ and $B = B_\Gamma$.

2 The Homotopy Path and its Properties

We analyse the main properties of the homotopy path of barrier regularizations. For brevity, we give only proofs here, when they differ from [7].

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

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

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

Their derivatives can be computed as $l'(z; \mu) = -\mu^q z^{-q}$. Bounds like $z \geq \bar{z}$ and $z \leq \underline{z}$, are incorporated by shifting the arguments.

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

$$\begin{aligned} b(\cdot; \mu) : Z &\rightarrow \bar{\mathbb{R}} \\ z &\mapsto \int_B l(z(t); \mu) dt. \end{aligned}$$

By $b'(z; \mu)$ we denote the formal derivative of $b(z; \mu)$, defined by

$$\langle b'(z; \mu), \delta z \rangle := \int_Q l'(z; \mu) \delta z dt,$$

if the right hand side is well defined. The following result connects these formal derivatives to the subdifferentials of convex analysis (cf. e.g. [4, Section I.5]).

Proposition 1. *Consider $b : L_p(Q) \rightarrow \bar{\mathbb{R}}$, $1 \leq p < \infty$ on a measurable set Q . Then either $\partial b(z; \mu) = \emptyset$, or $\partial b(z; \mu) = \{b'(z; \mu)\}$.*

Consider $b : C(Q) \rightarrow \bar{\mathbb{R}}$ on a compact set Q and assume $\emptyset \neq \partial b(z; \mu) \subset M(Q) \cong C(Q)^$. Then on the set of strictly feasible points $S := \{t \in Q : z(t) > 0\}$ we have*

$$m|_S = b'(z; \mu)|_S \quad \forall m \in \partial b(z; \mu). \quad (8)$$

In particular, $\partial b(z; \mu) \cap L_1(Q) = \{b'(z; \mu)\}$. Moreover,

$$\langle m, \delta z \rangle \leq \langle b'(z; \mu), \delta z \rangle \leq 0 \quad \forall 0 \leq \delta z \in C(Q) \quad (9)$$

and

$$\|b'(z; \mu)\|_{L_1(Q)} = \min_{m \in \partial b(z; \mu)} \|m\|_{M(Q)}. \quad (10)$$

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

$$\begin{aligned} F_\mu(x) &:= F(x) + b(x; \mu) = J(x) + \chi_E(x) + \chi_G(x) + b(x; \mu) \\ &= J(x) + \chi_E(x) + b(x; \mu). \end{aligned} \quad (11)$$

Our definition implies $F_0 = F$, which means that the original state constrained problem is included in our analysis.

Theorem 4 (Existence of Minimizers). *Let $F : X \rightarrow \bar{\mathbb{R}}$ be defined by (4) and suppose that Assumptions 1–2 hold. Assume that F_{μ_0} is coercive for some $\mu_0 > 0$.*

Then (11) admits a unique minimizer $x(\mu) = (u(\mu), y(\mu))$ for each $\mu \in]0, \mu_0]$. Moreover, $x(\mu)$ is strictly feasible almost everywhere in Ω and bounded in X uniformly in $\mu \in [0, \mu_0]$.

Next we study first order optimality conditions for barrier problems. For this purpose, we first have to study the subdifferential of χ_E , the characteristic function for the equality constraints $Ay - Bu = 0$, which can by (2) be written as $Tx = 0$. It is at this point, where our theory differs from [7].

Lemma 1. *If Assumption 1 holds, then there is a constant M , such that for each $u \in U$ there are $y \in Y$, $u_Y \in U$ and $u_V \in V$ with $Ay - Bu_Y = 0$ and*

$$u = u_Y + u_V \quad \|y\|_\infty + \|u_V\|_\infty \leq M \|u\|_U. \quad (12)$$

Proof. For $u \in U$ let $Bu = r$ and $r = r_Y + r_V$ as in Assumption 1. Since $\text{ran}A$ and $B(V)$ ($\dim V < \infty$) are closed, [9, Satz IV.6.3] yields a constant c independent of r , such that $\|r_Y\| + \|r_V\| \leq c \|r\| \leq c \|B\| \|u\|_U$.

By closedness of $B(V)$, the mapping $B : V \rightarrow B(V)$ is open, which yields a constant C such that for each $r_V \in B(V)$ there is $u_V \in V$ with $Bu_V = r_V$ and $\|u_V\|_U \leq C \|r_V\|$. Since all norms are equivalent on finite dimensional spaces, and V is a space of bounded functions, we even have $\|u_V\|_\infty \leq C \|r_V\|$.

Similarly, because $\text{ran}A$ is closed, $A : Y \supset \text{dom}A \rightarrow \text{ran}A$ is an open mapping by [9, Satz IV.4.4] and for each $r_Y \in \text{ran}A$ there is $y \in \text{dom}A$ with $Ay = r_Y$ and $\|y\|_\infty \leq C \|r_Y\|$.

This altogether yields $\|y\|_\infty + \|u_V\|_\infty \leq C(\|r_Y\| + \|r_V\|) \leq M \|u\|_U$ and thus (12).

Proposition 2. *Let X, R be Banach spaces and $T : X \supset \text{dom}T \rightarrow R$ a closed, densely defined, linear operator with closed range. Denote by χ_E the indicator function of $E := \ker T$. Then*

$$\partial \chi_E(x) = \text{ran} T^* \quad \forall x \in E. \quad (13)$$

Proof. Since, by definition of the subdifferential, $\partial \chi_E(x) = (\ker T)^\perp$, (13) is a consequence of the *closed range theorem* for closed operators on Banach spaces [5, Theorem IV.1.2], which asserts $(\ker T)^\perp = \text{ran} T^*$.

Theorem 5 (First Order Optimality Conditions). *Suppose that the Assumptions 1–2 hold. For $\mu \geq 0$ let x be the unique minimizer of F_μ .*

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

$$\begin{aligned} j_y + m + A^* p &= 0 \\ j_u + b'(u; \mu) - B^* p &= 0 \end{aligned} \quad (14)$$

holds. If y is strictly feasible, then $\partial b(y; \mu) = \{b'(y; \mu)\}$ and m is unique.

Proof. Let x be a minimizer of F_μ . Then $0 \in \partial F_\mu(x) = \partial(J + \chi_E + b)(x)$.

To show that (14) has a solution, we have to apply the sum-rule of convex analysis twice:

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

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

$$0 \in \text{int}(\text{dom } f - \text{dom } g). \quad (15)$$

Let now B_X be the unit ball in a normed space X . We observe that showing (15) is equivalent to showing that there is $\varepsilon > 0$ such that each $x \in \varepsilon B_X$ can be written as a difference $x_1 - x_2$ with $x_1 \in \text{dom } f$ and $x_2 \in \text{dom } g$.

By (3) there exists a strictly feasible point $\check{x} = (\check{y}, \check{u})$, which implies $\check{x} \in \text{dom}(\chi_E + b)$. Our assumptions on J include continuity at \check{x} and hence boundedness in some ball $\check{x} + \varepsilon B_X$. Thus,

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

and we conclude that (15) is fulfilled for $f = J$ and $g = \chi_E + b$. Therefore the sum-rule can be applied and yields $\partial(J + \chi_E + b) = \partial J + \partial(\chi_E + b)$.

Next we show that $\partial(\chi_E + b) = \partial\chi_E + \partial b$ by verifying (15) for b and χ_E . Here $Y = C(\overline{\Omega})$ is crucial because it guarantees that $(\check{u}, \check{y} + rB_Y) \in \text{dom } b$ for $r < d_{\min}$ via (3). By (12) there is $\delta > 0$ such that for each $u \in \delta B_U$ we find an $y \in (r/2)B_Y$ with $Ay - Bu_Y = 0$ and u_V with $\|u_V\|_\infty \leq r$, such that $u = u_Y + u_V$.

Thus $(\check{y} + y, \check{u} + u_Y) \in \text{dom } \chi_E$ and $(\check{y} + y - w, \check{u} - u_V) \in \text{dom } b$ for all $w \in (r/2)B_Y$ by (3). Consequently, for sufficiently small ε and arbitrary $(w, u) \in \varepsilon B_X$ we have

$$\begin{aligned} w &= (\check{y} + y) - (\check{y} + y - w) \\ u &= \underbrace{(\check{u} + u_Y)}_{\in \text{dom } \chi_E} - \underbrace{(\check{u} - u_V)}_{\in \text{dom } b}. \end{aligned}$$

This finally shows (15) and the sum-rule yields $0 \in \partial J + \partial\chi_E + \partial b$.

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

$$\begin{aligned} j_y + v + m &= 0 \\ j_u + \lambda + l &= 0. \end{aligned}$$

Proposition 2 applied to T as defined in (2) yields $(v, \lambda) \in \text{ran } T^*$. Hence there is $p \in \text{dom } T^*$ with $v = A^*p$ and $\lambda = B^*p$. Proposition 1 characterizes m and l in terms of barrier gradients. This yields (14). If y is strictly feasible, then $m = b'(y; \mu)$ by Proposition 1.

Once, existence of the barrier gradients is established, their uniform boundedness for $\mu \rightarrow 0$ can again be shown as in [7].

Proposition 3. *Suppose that the Assumptions 1–2 hold. Then for each $\mu_0 > 0$*

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

Just as in [7] this result allows also to derive uniform bounds on the adjoint state $p(\mu)$ in some suitable Sobolev space. The results on the analytic properties of the central path carry over literally from [7].

Theorem 6. *Suppose that the Assumptions 1–3 hold. Let $x(\mu)$ be a barrier minimizer for $\mu \geq 0$ and x_* be minimizer of F . Then*

$$F(x(\mu)) \leq F(x_*) + C\mu_0 \quad (16)$$

$$\|x(\mu) - x_*\| \leq C\sqrt{\frac{\mu}{\alpha}}. \quad (17)$$

$$\|x(\mu) - x(\tilde{\mu})\| \leq \frac{c}{\sqrt{\alpha\mu}}|\mu - \tilde{\mu}| \quad \forall \tilde{\mu} \geq 0. \quad (18)$$

Finally, we remark that the results on strict feasibility of the homotopy path, which depend on the regularity of $y(\mu)$, carry over from [7].

Acknowledgement. Supported by the DFG Research Center MATHEON “Mathematics for key technologies.”

References

- [1] Amann, H.: Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In H.J. Schmeisser and H. Triebel, eds., *Function Spaces, Differential Operators and Nonlinear Analysis*, pages 9–126. Teubner, Stuttgart, Leipzig, 1993.
- [2] Borwein, J.M., Zhu, Q.J.: *Techniques of Variational Analysis*. CMS Books in Mathematics. Springer, 2005.
- [3] Casas, E.: Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control Optim.*, 31:993–1006, 1993.
- [4] Ekeland, I., Témam, R.: *Convex Analysis and Variational Problems*. Classics in Applied Mathematics, 28. SIAM, 1999.
- [5] Goldberg, S.: *Unbounded Linear Operators*. Dover, 1966.
- [6] Hinze, M., Schiela, A.: 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, 2007.
- [7] Schiela, A.: Barrier methods for optimal control problems with state constraints. ZIB Report 07-07, Zuse Institute Berlin, 2007.
- [8] Schiela, A.: Optimality conditions for convex state constrained optimal control problems with discontinuous states. ZIB Report 07-35, Zuse Institute Berlin, 2007.
- [9] Werner, D.: *Funktionalanalysis*. Springer, 3rd ed., 2000.