Hamburger Beiträge
zur Angewandten Mathematik

Elliptic control problems with gradient constraints - variational discrete versus piecewise constant controls

Andreas Günther and Michael Hinze

Nr. 2009-02
February 2009

Elliptic control problems with gradient constraints - variational discrete versus piecewise constant controls

Andreas Günther & Michael Hinze

Bereich Optimierung und Approximation, Universität Hamburg

Abstract: We consider an elliptic optimal control problem with pointwise bounds on the gradient of the state. To guarantee the required regularity of the state we include the L^r-norm in our cost functional with $r > d$, $(d = 2, 3)$. We investigate variational discretization of the control problem [6] as well as piecewise constant approximations of the control. In both cases we use standard piecewise linear and continuous finite elements for the discretization of the state. Pointwise bounds on the gradient of the discrete gradient are enforced element-wise. Error bounds for control and state are obtained in two and three space dimensions depending on the value of r .

Mathematics Subject Classification (2000): 49J20, 49K20, 35B37

Keywords: Elliptic optimal control problem, gradient constraints, error estimates

1 Introduction

Constraints on the gradient of the state play an important role in practical applications where cooling of melts forms a critical process. In order to accelerate such production processes it is highly desirable to speed up the cooling processes while avoiding damage of the products caused by large material stresses. Cooling processes frequently are described by systems of partial differential equations involving the temperature as a system variable, so that large (Von Mises) stresses in the optimization process can be avoided by imposing pointwise bounds on the gradient of the temperature. Pointwise bounds on the gradient in optimization in general deliver adjoint variables admitting low regularity only. This fact then necessitates the development of tailored discrete concepts which take into account the low regularity of adjoint variables and multipliers involved in the optimality conditions of the underlying optimization problem.

The present work complements the discrete approach to elliptic optimal control problems with gradient constraints presented by the authors in [2]. There, variational discretization of the controls is considered combined with the lowest order Raviart-Thomas finite element approximations of a mixed formulation of the state equation. This in particular leads to piecewise constant approximations to the state and the adjoint state, respectively. However, many existing finite element codes use finite elements based on conventional continuous piecewise polynomial Ansatz spaces. This is our motivation to provide numerical analysis for elliptic control problems with gradient constraints also for piecewise polynomial and continuous state approximations. In the present work we besides variational discretization also consider piecewise constant approximations of the controls. In both cases the state is discretized with standard piecewise linear and continuous finite elements. Our main results are stated in Theorems 2.5,2.7. It reads

$$\|y - y_h\| \leq Ch^{\frac{1}{2}(1-\frac{d}{r})}, \text{ and } \|u - u_h\|_{L^r} \leq Ch^{\frac{1}{r}(1-\frac{d}{r})},$$

and is valid for variational discretization as well as for piecewise constant control approximations. Here, y, u and y_h, u_h denote the unique solutions of the optimal control problems (1.2) and (2.9), (2.23), respectively, and $\|\cdot\|$ throughout the paper denotes the L^2-norm.
In the presence of gradient constraints variational discretization of the controls automatically leads to globally continuous approximations of the controls, if globally continuous Ansatz functions for the state are used, see relation (2.14). This is certainly a drawback of the approach, since the optimal control and the associated adjoint state may have jumps, see the numerical example in Section 3. Piecewise constant control approximations here seem to be the better choice. However, the approximation order in both cases is the same, and also the errors in the numerical experiments for both approaches are of similar size, see Tables 1,2.

The problem formulation already is presented in [2]. For the convenience of the reader it is recalled in the following. To begin with let $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) be a bounded domain with a smooth boundary $\partial\Omega$. We consider the differential operator

$$Ay := -\sum_{i,j=1}^{d} \partial_{x_j}\left(a_{ij} y_{x_i}\right) + a_0 y,$$

where for simplicity the coefficients a_{ij} and a_0 are assumed to be smooth functions on $\bar{\Omega}$. We associate with A the bilinear form

$$a(y, z) := \int_\Omega \big(\sum_{i,j=1}^{d} a_{ij}(x) y_{x_i} z_{x_j} + a_0(x) y z\big)\, dx, \qquad y, z \in H^1(\Omega)$$

and subsequently assume that $a_{ij} = a_{ji}$, $a_0 \geq 0$ in Ω and that there exists $c_0 > 0$ such that

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq c_0|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d \text{ and all } x \in \Omega.$$

From the above assumptions it follows that for a given $f \in L^r(\Omega)$ $(1 < r < \infty)$ the elliptic boundary value problem

$$\begin{aligned} Ay &= f \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{1.1}$$

has a unique solution $y \in W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega)$ which we denote by $y = \mathcal{G}(f)$. Furthermore,

$$\|y\|_{W^{2,r}} \leq C\|f\|_{L^r},$$

where $\|\cdot\|_{L^r}$ and $\|\cdot\|_{W^{k,r}}$ denote the usual Lebesgue and Sobolev norms. Moreover, for $f \in W^{-1,r}(\Omega)$ we have $\mathcal{G}(f) \in W^{1,r}(\Omega)$ (see [5] for $d = 2$, and [8] for $d = 3$) with

$$\|y\|_{W^{1,r}} \leq C\|f\|_{W^{-1,r}},$$

where the positive constant is independent of f.
Let $r > d$, $\alpha > 0$ and $y_0 \in L^2(\Omega)$ be given. We now consider the control problem

$$\begin{aligned} &\min_{u \in L^r(\Omega)} J(u) = \frac{1}{2}\int_\Omega |y - y_0|^2 + \frac{\alpha}{r}\int_\Omega |u|^r. \\ &\text{subject to } y = \mathcal{G}(u) \text{ and } \nabla y \in \mathcal{K}. \end{aligned} \tag{1.2}$$

Here,

$$\mathcal{K} = \{z \in C^0(\bar{\Omega})^d \,|\, |z(x)| \leq \delta, x \in \bar{\Omega}\}. \tag{1.3}$$

Since $r > d$ we have $y \in W^{2,r}(\Omega)$ and hence $\nabla y \in C^0(\bar{\Omega})^d$ by a well–known embedding result. We impose the following Slater condition:

$$\exists \hat{u} \in L^r(\Omega) \quad |\nabla \hat{y}(x)| < \delta,\ x \in \bar{\Omega},\ \text{where } \hat{y} \text{ solves (1.1) with } u = \hat{u}. \tag{1.4}$$

Since J is uniformly convex and the set of admissible controls and states forms a closed and convex set problem (1.2) admits a unique solution u with associated state $\mathcal{G}(u)$.
The KKT system of problem (1.2) is obtained with the help of [1, Corollary 1]. There holds

Theorem 1.1. *An element $u \in L^r(\Omega)$ is a solution of (1.2) if and only if there exist $\vec{\mu} \in \mathcal{M}(\bar{\Omega})^d$ and $p \in L^t(\Omega)$ ($t < \frac{d}{d-1}$) such that*

$$\int_\Omega p\mathcal{A}z - \int_\Omega (y - y_0)z - \int_{\bar{\Omega}} \nabla z \cdot d\vec{\mu} = 0 \qquad \forall z \in W^{2,t'}(\Omega) \cap W_0^{1,t'}(\Omega) \tag{1.5}$$

$$p + \alpha |u|^{r-2} u = 0 \qquad \text{in } \Omega \tag{1.6}$$

$$\int_{\bar{\Omega}} (\mathbf{z} - \nabla y) \cdot d\vec{\mu} \leq 0 \qquad \forall \mathbf{z} \in \mathcal{K}. \tag{1.7}$$

Here, y is the solution of (1.1) and $\frac{1}{t} + \frac{1}{t'} = 1$. Further we recall that $\mathcal{M}(\bar{\Omega})$ denotes the space of regular Borel measures.

Remark 1.2. Lemma 1 in the paper [1] of Casas and Fernandéz shows that the vector valued measure $\vec{\mu}$ appearing in Theorem 1.1 can be written in the form

$$\vec{\mu} = \frac{1}{\delta} \nabla y \, \mu,$$

where $\mu \in \mathcal{M}(\bar{\Omega})$ is a nonnegative measure that is concentrated in the set $\{x \in \bar{\Omega} \,|\, |\nabla y(x)| = \delta\}$. For an example we refer to [2].

2 Finite element discretization

We sketch an approach from [7, Section 3.3.2] which uses classical piecewise linear, continuous approximations of the states. In [2] Deckelnick, Günther and Hinze present a finite element approximation to problem (1.2) which uses mixed finite element approximations for the states. Let us recall the definition of the space of linear finite elements,

$$X_h := \{v_h \in C^0(\bar{\Omega}) \,|\, v_h \text{ is a linear polynomial on each } T \in \mathcal{T}_h\}$$

with the appropriate modification for boundary elements and let $X_{h0} := X_h \cap H_0^1(\Omega)$. Here $\mathcal{T}_h$ again denotes a quasi-uniform triangulation of Ω with maximum mesh size $h := \max_{T \in \mathcal{T}_h} \operatorname{diam}(T)$. We suppose that $\bar{\Omega}$ is the union of the elements of $\mathcal{T}_h$ so that element edges lying on the boundary are curved. Furthermore let us recall the definition of the discrete approximation of the operator $\mathcal{G}$. For a given function $v \in L^2(\Omega)$ we denote by $z_h = \mathcal{G}_h(v) \in X_{h0}$ the solution of

$$a(z_h, v_h) = \int_\Omega v v_h \quad \text{for all } v_h \in X_{h0}.$$

It is well–known that for all $v \in L^r(\Omega)$

$$\|\mathcal{G}(v) - \mathcal{G}_h(v)\|_{W^{1,\infty}} \leq C \inf_{z_h \in X_{h0}} \|\mathcal{G}(v) - z_h\|_{W^{1,\infty}} \leq C h^{1-\frac{d}{r}} \|\mathcal{G}(v)\|_{W^{2,r}} \leq C h^{1-\frac{d}{r}} \|v\|_{L^r}. \tag{2.8}$$

For each $T \in \mathcal{T}_h$ let $z_T \in \mathbb{R}^d$ denote constant vectors. We define

$$\mathcal{K}_h := \{z_h : \Omega \to \mathbb{R}^d \,|\, z_{h|T} = z_T \text{ on } T \text{ and } |z_{h|T}| \leq \delta,\ T \in \mathcal{T}_h\}.$$

Let us first consider variational discretization of problem (1.2) which reads:

$$\min_{u \in L^r(\Omega)} J_h(u) := \frac{1}{2}\int_\Omega |y_h - y_0|^2 + \frac{\alpha}{r}\int_\Omega |u|^r \\ \text{subject to } \; y_h = \mathcal{G}_h(u) \text{ and } \nabla y_h \in \mathcal{K}_h. \tag{2.9}$$

We first note that $\hat{y}_h := \mathcal{G}_h(\hat{u})$ satisfies a Slater condition similar to (1.4), since for $x_T \in T \in \mathcal{T}_h$ by (2.8)

$$|\nabla \hat{y}_h(x_T)| \le |\nabla(\hat{y}_h(x_T) - \hat{y}(x_T))| + |\nabla \hat{y}(x_T)| \le \|\nabla(\hat{y}_h - \hat{y})\|_{L^\infty} + \max_{x \in \bar{\Omega}} |\nabla \hat{y}(x)| \le \\ \le Ch^{1-\frac{d}{r}} + (1-2\epsilon)\delta \le (1-\epsilon)\delta \text{ for all } T \in \mathcal{T}_h,$$

for some $\epsilon > 0$ and $0 < h \le h_0$, so that $\left(\nabla \hat{y}_{h|T}\right)_{T \in \mathcal{T}_h} \in \mathcal{K}_h$ satisfies the Slater condition

$$|\nabla \hat{y}_h(x)| < \delta \text{ for all } x \in \bar{\Omega}. \tag{2.10}$$

This delivers

Lemma 2.1. *Problem (2.9) has a unique solution $u_h \in L^r(\Omega)$. There exist $\vec{\mu}_T \in \mathbb{R}^d, T \in \mathcal{T}_h$ and $p_h \in X_{h0}$ such that with $y_h = \mathcal{G}_h(u_h)$ we have*

$$a(v_h, p_h) = \int_\Omega (y_h - y_0)v_h + \sum_{T \in \mathcal{T}_h} |T| \, \nabla v_{h|T} \cdot \vec{\mu}_T \qquad \forall v_h \in X_{h0}, \tag{2.11}$$

$$p_h + \alpha |u_h|^{r-2} u_h = 0 \quad \text{in } \Omega, \tag{2.12}$$

$$\sum_{T \in \mathcal{T}_h} |T| \big(z_T - \nabla y_{h|T}\big) \cdot \vec{\mu}_T \le 0 \quad \forall z_h \in \mathcal{K}_h. \tag{2.13}$$

In problem (2.9) we apply variational discretization of [6]. From (2.12) we infer for the discrete optimal control

$$u_h = -\alpha^{-\frac{1}{r-1}} |p_h|^{\frac{2-r}{r-1}} p_h. \tag{2.14}$$

Further, according to Remark 1.2 we have the following representation of the discrete multipliers.

Lemma 2.2. *Let u_h denote the unique solution of (2.9) with corresponding state $y_h = \mathcal{G}_h(u_h)$ and multiplier $(\vec{\mu}_T)_{T \in \mathcal{T}_h}$. Then there holds*

$$\vec{\mu}_T = |\vec{\mu}_T| \frac{1}{\delta} \nabla y_{h|T} \text{ for all } T \in \mathcal{T}_h. \tag{2.15}$$

Proof. Fix $T \in \mathcal{T}_h$. The assertion is clear if $\vec{\mu}_T = 0$. Suppose that $\vec{\mu}_T \ne 0$ and define $z_h : \bar{\Omega} \to \mathbb{R}^d$ by

$$z_{h|\tilde{T}} := \begin{cases} \nabla y_{h|T}, & \tilde{T} \ne T, \\ \delta \frac{\vec{\mu}_T}{|\vec{\mu}_T|}, & \tilde{T} = T. \end{cases}$$

Clearly, $z_h \in \mathcal{K}_h$ so that (2.13) implies

$$\vec{\mu}_T \cdot \big(\delta \frac{\vec{\mu}_T}{|\vec{\mu}_T|} - \nabla y_{h|T}\big) \le 0,$$

and therefore, since $\left(\nabla y_{h|T}\right)_{T \in \mathcal{T}_h} \in \mathcal{K}_h$,

$$\delta |\vec{\mu}_T| \le \vec{\mu}_T \cdot \nabla y_{h|T} \le \delta |\vec{\mu}_T|.$$

Hence we obtain $\frac{\vec{\mu}_T}{|\vec{\mu}_T|} = \frac{1}{\delta}\nabla y_{h|T}$ and the lemma is proved. ∎

As a consequence of Lemma 2.2 we immediately infer that

$$|\vec{\mu}_T| = \vec{\mu}_T \cdot \frac{1}{\delta}\nabla y_{h|T} \text{ for all } T \in \mathcal{T}_h. \tag{2.16}$$

We now use (2.16) in order to derive an important a priori estimate.

Lemma 2.3. *Let $u_h \in L^r(\Omega)$ be the optimal solution of (2.9) with corresponding state $y_h \in X_{h0}$ and adjoint variables $p_h \in X_{h0}$, $\vec{\mu}_T, T \in \mathcal{T}_h$. Then there exists $h_0 > 0$ such that*

$$\|y_h\|,\ \|u_h\|_{L^r},\ \|p_h\|_{L^{\frac{r}{r-1}}},\ \sum_{T\in\mathcal{T}_h} |T|\,|\vec{\mu}_T| \le C \qquad \textit{for all } 0 < h \le h_0.$$

Proof. Combining (2.16) with (2.10) we deduce

$$\vec{\mu}_T \cdot (\nabla y_{h|T} - \nabla \hat{y}_{h|T}) \ge \delta|\vec{\mu}_T| - (1-\epsilon)\delta\,|\vec{\mu}_T| = \epsilon\delta|\vec{\mu}_T|.$$

Choosing $w_h = y_h - \hat{y}_h$ in (2.11) and using the definition of $\mathcal{G}_h$ together with (2.12) we hence obtain

$$\begin{aligned}
\epsilon\delta \sum_{T\in\mathcal{T}_h} |T||\vec{\mu}_T| &\le \sum_{T\in\mathcal{T}_h} |T|\vec{\mu}_T \cdot (\nabla y_{h|T} - \nabla \hat{y}_{h|T}) \\
&= a(y_h - \hat{y}_h, p_h) - \int_\Omega (y_h - y_0)(y_h - \hat{y}_h) \\
&= \int_\Omega (u_h - \hat{u})p_h - \int_\Omega (y_h - y_0)(y_h - \hat{y}_h) \\
&\le -\frac{\alpha}{2}\int_\Omega |u_h|^r - \frac{1}{2}\int_\Omega |y_h|^2 + C(1 + \|y_0\|^2 + \|\hat{u}\|_{L^r}^r).
\end{aligned}$$

This implies the bounds on y_h, u_h and $\vec{\mu}_T$. The bound on p_h follows from (2.12). ∎

Remark 2.4. For the measure $\vec{\mu}_h \in \mathcal{M}(\bar{\Omega})^d$ defined by

$$\int_{\bar{\Omega}} f \cdot d\vec{\mu}_h := \sum_{T\in\mathcal{T}_h} \int_T f\,dx \cdot \vec{\mu}_T \text{ for all } f \in C^0(\bar{\Omega})^d,$$

it follows immediately that

$$\|\vec{\mu}_h\|_{\mathcal{M}(\bar{\Omega})^d} \le C.$$

Now we are in the position to prove the following error estimates.

Theorem 2.5. *Let u and u_h be the solutions of (1.2) and (2.9) respectively. Then there exists $h_1 \le h_0$ such that*

$$\|y - y_h\| \le Ch^{\frac{1}{2}(1-\frac{d}{r})}, \textit{ and } \|u - u_h\|_{L^r} \le Ch^{\frac{1}{r}(1-\frac{d}{r})}$$

for all $0 < h \le h_1$.

Proof. Let us introduce $y^h := \mathcal{G}(u_h) \in W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega)$, and $\tilde{y}_h := \mathcal{G}_h(u)$. In view of Lemma 2.3 and (2.8) we have

$$\|y^h - y_h\|_{W^{1,\infty}} \le Ch^{1-\frac{d}{r}}\|u_h\|_{L^r} \le Ch^{1-\frac{d}{r}}. \tag{2.17}$$

Let us now turn to the actual error estimate. To begin, we recall that for $r \geq 2$ there exists $\theta_r > 0$ such that

$$(|a|^{r-2}a - |b|^{r-2}b)(a-b) \geq \theta_r |a-b|^r \qquad \forall a, b \in \mathbb{R}.$$

Hence, using (1.6) and (2.12),

$$\alpha\theta_r \int_\Omega |u - u_h|^r \leq \alpha \int_\Omega \big(|u|^{r-2}u - |u_h|^{r-2}u_h\big)(u - u_h) = \int_\Omega (-p + p_h)(u - u_h) =: (1) + (2).$$

Recalling (1.5) we have

$$\begin{aligned}
(1) &= \int_\Omega p\big(\mathcal{A}y^h - \mathcal{A}y\big) \\
&= \int_\Omega (y - y_0)(y^h - y) + \int_{\bar\Omega} \big(\nabla y^h - \nabla y\big) \cdot d\vec{\mu} \\
&= \int_\Omega (y - y_0)(y^h - y) + \underbrace{\int_{\bar\Omega} \big(P_\delta(\nabla y^h) - \nabla y\big) \cdot d\vec{\mu}}_{\leq 0} + \int_{\bar\Omega} \big(\nabla y^h - P_\delta(\nabla y^h)\big) \cdot d\vec{\mu}
\end{aligned}$$

where P_δ denotes the orthogonal projection onto $\bar{B}_\delta(0) = \{x \in \mathbb{R}^d \,|\, |x| \leq \delta\}$. Note that

$$|P_\delta(x) - P_\delta(\tilde{x})| \leq |x - \tilde{x}| \qquad \forall x, \tilde{x} \in \mathbb{R}^d. \tag{2.18}$$

Since $x \mapsto P_\delta(\nabla y^h(x)) \in \mathcal{K}$ we infer from (1.7)

$$(1) \leq \int_\Omega (y - y_0)(y^h - y) + \max_{x \in \bar\Omega} |\nabla y^h(x) - P_\delta(\nabla y^h(x))| \|\vec{\mu}\|_{\mathcal{M}(\bar\Omega)^d}. \tag{2.19}$$

Let $x \in \bar\Omega$, say $x \in T$ for some $T \in \mathcal{T}_h$. Since u_h is feasible for (2.9) we have that $\nabla y_{h|T} \in \bar{B}_\delta(0)$ so that (2.18) together with (2.17) implies

$$\begin{aligned}
\big|\nabla y^h(x) - P_\delta(\nabla y^h(x))\big| &\leq \big|\nabla y^h(x) - \nabla y_{h|T}\big| + \big|P_\delta(\nabla y^h(x)) - P_\delta\big(\nabla y_{h|T}\big)\big| \\
&\leq 2\big|\nabla y^h(x) - \nabla y_{h|T}\big| \leq Ch^{1-\frac{d}{r}} \|u_h\|_{L^r}.
\end{aligned} \tag{2.20}$$

Thus

$$(1) \leq \int_\Omega (y - y_0)(y^h - y) + Ch^{1-\frac{d}{r}}. \tag{2.21}$$

Similarly,

$$(2) = a(\tilde{y}_h - y_h, p_h) = \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + \sum_{T \in \mathcal{T}_h} |T| \big(\nabla \tilde{y}_{h|T} - \nabla y_{h|T}\big) \cdot \vec{\mu}_T =$$

$$= \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + \sum_{T \in \mathcal{T}_h} |T| \big(\nabla \tilde{y}_{h|T} - P_\delta(\nabla \tilde{y}_{h|T})\big) \cdot \vec{\mu}_T + \underbrace{\sum_{T \in \mathcal{T}_h} |T| \big(P_\delta(\nabla \tilde{y}_{h|T} - \nabla y_{h|T}\big) \cdot \vec{\mu}_T}_{\leq 0} \leq$$

$$\leq \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + \sum_{T \in \mathcal{T}_h} |T| \big(\nabla \tilde{y}_{h|T} - \nabla y(x_T)\big) \cdot \vec{\mu}_T + \sum_{T \in \mathcal{T}_h} |T| \big(P_\delta(\nabla y(x_T) - P_\delta(\nabla \tilde{y}_{h|T})\big) \cdot \vec{\mu}_T,$$

where $x_T \in T$, so that $(\nabla y(x_T))_{T \in \mathcal{T}_h} \in \mathcal{K}_h$. We infer from Lemma 2.3 and (2.8)

$$(2) \leq \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + 2 \max_{T \in \mathcal{T}_h} |\nabla \tilde{y}_{h|T} - \nabla y(x_T)| \sum_{T \in \mathcal{T}_h} |T| |\vec{\mu}_T| \leq$$

$$\leq \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + Ch^{1-\frac{d}{r}} \|u\|_{L^r}. \tag{2.22}$$

Combining (1) and (2) we finally obtain

$$
\begin{aligned}
\alpha\theta_r \int_\Omega |u-u_h|^r &\le \int_\Omega (y-y_0)(y^h-y) + \int_\Omega (y_h-y_0)(\tilde{y}_h-y_h) + Ch^{1-\frac{d}{r}} \\
&= -\int_\Omega |y-y_h|^2 + \int_\Omega \big((y_0-y_h)(y-\tilde{y}_h) + (y-y_0)(y^h-y_h)\big) + Ch^{1-\frac{d}{r}} \\
&\le -\int_\Omega |y-y_h|^2 + C\big(\|y-\tilde{y}_h\| + \|y^h-y_h\|\big) + Ch^{1-\frac{d}{r}} \\
&\le -\int_\Omega |y-y_h|^2 + Ch\big(\|u\| + \|u_h\|\big) + Ch^{1-\frac{d}{r}}
\end{aligned}
$$

and the result follows. ∎

Piecewise constant controls. Let us now consider the following optimal control problem with piecewise constant controls as discretization of problem (1.2);

$$
\begin{aligned}
&\min_{u_h\in U_h} J_h(u_h) := \frac{1}{2}\int_\Omega |y_h-y_0|^2 + \frac{\alpha}{r}\int_\Omega |u_h|^r \\
&\text{subject to } \; y_h = \mathcal{G}_h(u_h) \text{ and } \nabla y_h \in \mathcal{K}_h,
\end{aligned} \tag{2.23}
$$

where $U_h := \{v_h \in L^r(\Omega)\,|\, v_{h|T} \in \mathbb{R} \text{ for all } T \in \mathcal{T}_h\}$. It is not difficult to prove that this problem admits a unique solution $u_h \in U_h$. Our finite element error analysis for this problem is based on approximation properties of the orthogonal L^2-projection $Q_h : L^2(\Omega) \to U_h$ defined by

$$
(Q_h v)(x) := \fint_T v = \frac{1}{|T|}\int_T v \text{ for all } v \in L^2(\Omega), x \in T.
$$

For $v \in L^r$ and $\phi \in W^{1,r}$ we have the stability estimate

$$
\|Q_h v\|_{L^r} \le c\|v\|_{L^r} \tag{2.24}
$$

as well as the approximation property

$$
\|\phi - Q_h\phi\|_{L^r} \le Ch\|\phi\|_{W^{1,r}}, \tag{2.25}
$$

see [3, Prop. 1.135].

Let $v := \frac{1}{2}u + \frac{1}{2}\hat{u}$. Then it is not difficult to show that for $h > 0$ small enough the function $\hat{y}_h := \mathcal{G}_h(Q_h v)$ satisfies the Slater condition (2.10). For the optimal control problem (2.23) the result of Lemma 2.1 is valid if we replace (2.12) by

$$
\int_\Omega (p_h + \alpha|u_h|^{r-2}u_h)(v_h - u_h) = 0 \qquad \forall v_h \in U_h. \tag{2.26}
$$

Furthermore Lemma 2.2 holds accordingly and the analogon to Lemma 2.3 reads

Lemma 2.6. *Let $u_h \in U_h$ be the optimal solution of (2.23) with corresponding state $y_h \in X_{h0}$ and adjoint variables $p_h \in X_{h0}$, $\vec{\mu}_T, T \in \mathcal{T}_h$. Then there exists $h_0 > 0$ such that*

$$
\|y_h\|,\ \|u_h\|_{L^r},\ \sum_{T\in\mathcal{T}_h} |T|\,|\vec{\mu}_T| \le C \qquad \textit{for all } 0 < h \le h_0
$$

holds. Its proof is along the lines of the proof of Lemma 2.3 where one uses the properties of the projection Q_h.

Theorem 2.7. *Let u and u_h be the solutions of (1.2) and (2.23) respectively. Then there exists $h_1 \le h_0$ such that*

$$\|y - y_h\| \le Ch^{\frac{1}{2}(1-\frac{d}{r})}, \text{ and } \|u - u_h\|_{L^r} \le Ch^{\frac{1}{r}(1-\frac{d}{r})}$$

for all $0 < h \le h_1$.

Proof. Let us introduce $y^h := \mathcal{G}(u_h) \in W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega)$, and $\tilde{y}_h := \mathcal{G}_h(Q_h u)$. In view (2.8) we have

$$\|y^h - y_h\|_{W^{1,\infty}} \le Ch^{1-\frac{d}{r}} \|u_h\|_{L^r} \le Ch^{1-\frac{d}{r}}.$$

Let us now turn to the actual error estimate. Using (1.6) and (2.26) we have

$$\alpha\theta_r \int_\Omega |u - u_h|^r \le \alpha \int_\Omega \big(|u|^{r-2}u - |u_h|^{r-2}u_h\big)(u - u_h) =$$
$$\underbrace{\int_\Omega p(u_h - u)}_{=:(1)} + \underbrace{\int_\Omega p_h(Q_h u - u_h)}_{=:(2)} \underbrace{-\alpha \int_\Omega \underbrace{|u_h|^{r-2}u_h}_{\in U_h} \underbrace{(u - Q_h u)}_{\in U_h^\perp}}_{=0}$$

To estimate the terms (1) and (2) we follow the lines of the proof of Theorem 2.5 and obtain

$$(1) \le \int_\Omega (y - y_0)(y^h - y) + Ch^{1-\frac{d}{r}}, \tag{2.27}$$

as well as

$$(2) \le \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + 2 \max_{T\in\mathcal{T}_h} |\nabla\tilde{y}_{h|T} - \nabla y(x_T)| \sum_{T\in\mathcal{T}_h} |T||\vec{\mu}_T| \le$$
$$\le \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + C\|\nabla(\tilde{y}_h - y)\|_{L^\infty} \tag{2.28}$$

Abbreviating $v := Q_h u - u$ we estimate the last term by

$$\|\nabla(\tilde{y}_h - y)\|_{L^\infty} \le \|\nabla\mathcal{G}(v)\|_{L^\infty} + \|\mathcal{G}_h(Q_h u) - \mathcal{G}(Q_h u)\|_{W^{1,\infty}} \le$$
$$\le \|\nabla\mathcal{G}(v)\|_{L^\infty} + Ch^{1-\frac{d}{r}}\|Q_h u\|_{L^r}. \tag{2.29}$$

Furthermore

$$\|\nabla\mathcal{G}(v)\|_{L^\infty} \le C\|\nabla\mathcal{G}(v)\|_{L^r}^\beta |\nabla\mathcal{G}(v)|_{W^{1,r}}^{1-\beta} \le C\|v\|_{W^{-1,r}}^\beta \|v\|_{L^r}^{1-\beta},$$

where we have used the Lyapunov inequality ([4, Thm. 10.1]) with $0 < \beta := 1 - \frac{d}{r} < 1$. It is easy to prove

$$\|v\|_{W^{-1,r}} = \|u - Q_h u\|_{W^{-1,r}} \le ch\|u\|_{L^r},$$

so that we obtain

$$\|\nabla\mathcal{G}(v)\|_{L^\infty} \le Ch^{1-\frac{d}{r}},$$

and thus

$$(2) \le \int_\Omega (y_h - y_0)(\tilde{y}_h - y_h) + Ch^{1-\frac{d}{r}}.$$

Combining (1) and (2) we finally obtain as in the proof of Theorem 2.5

$$\alpha\theta_r \int_\Omega |u - u_h|^r + \int_\Omega |y - y_h|^2 \le Ch\big(\|u\| + \|u_h\|\big) + Ch^{1-\frac{d}{r}}$$

and the result follows. ∎

3 A numerical experiment with pointwise constraints on the gradient

We now consider the finite element approximation of problem (1.2) with the following data. We consider (1.2) with the choices $\Omega = B_2(0) \subset \mathbb{R}^2$, $\alpha = 1$,

$$\mathcal{K} = \{\mathbf{z} \in C^0(\bar{\Omega})^2 \,|\, |\mathbf{z}(x)| \le \frac{1}{2}, x \in \bar{\Omega}\}$$

as well as

$$y_0(x) := \begin{cases} \frac{1}{4} + \frac{1}{2}\log 2 - \frac{1}{4}|x|^2, & 0 \le |x| \le 1, \\ \frac{1}{2}\log 2 - \frac{1}{2}\log|x|, & 1 < |x| \le 2. \end{cases}$$

In the state equation we allow an additional right hand side f, i.e. we consider the problem

$$\begin{aligned} -\Delta y &= f + u \quad \text{in } \Omega \\ y &= 0 \qquad\quad \text{on } \partial\Omega, \end{aligned}$$

where

$$f(x) := \begin{cases} 2, & 0 \le |x| \le 1, \\ 0, & 1 < |x| \le 2. \end{cases}$$

The optimization problem then has the unique solution

$$u(x) = \begin{cases} -1, & 0 \le |x| \le 1, \\ \;\;0, & 1 < |x| \le 2 \end{cases}$$

with corresponding state $y \equiv y_0$. We note that we obtain equality in (1.6), i.e. $p = -u$. Furthermore, the action of the measure $\vec{\mu}$ applied to a vectorfield $\phi \in C^0(\bar{\Omega})^2$ is given by

$$\int_{\bar{\Omega}} \phi \cdot d\vec{\mu} = -\int_{\partial B_1(0)} x \cdot \phi \, dS.$$

Variational discretization. We solve problem (2.9), where we essentially make use of the structure of u_h in terms of equation (2.14). Figs. 1 illustrates the optimal solution u_h and corresponding adjoint state p_h on a mesh consisting of $nt = 512$ triangles. We note that due to relation (2.14) the variational control has to be a continuous function. The exact control however has a jump. We conclude that variational discretization combined with piecewise linear and continuous finite elements for the state approximation is not ideally suited to approximate control problems with gradient constraints. Here, the lowest order Raviart-Thomas finite element combined with a mixed formulation of the state equation seems to be a more appropriate choice, see [2]. However, many existing finite element codes use standard finite elements, and there exists a demand in these classical approximation approaches also in state constrained optimization of elliptic optimal control problems.
In Table 1 we investigate the experimental order of convergence for the error functionals

$$E_u^s(h) := \|u - u_h\|_{L^s(\Omega)}, \; s \in \{2, 4\}, \text{ and } E_y(h) := \|y - y_h\|_{L^2(\Omega)}.$$

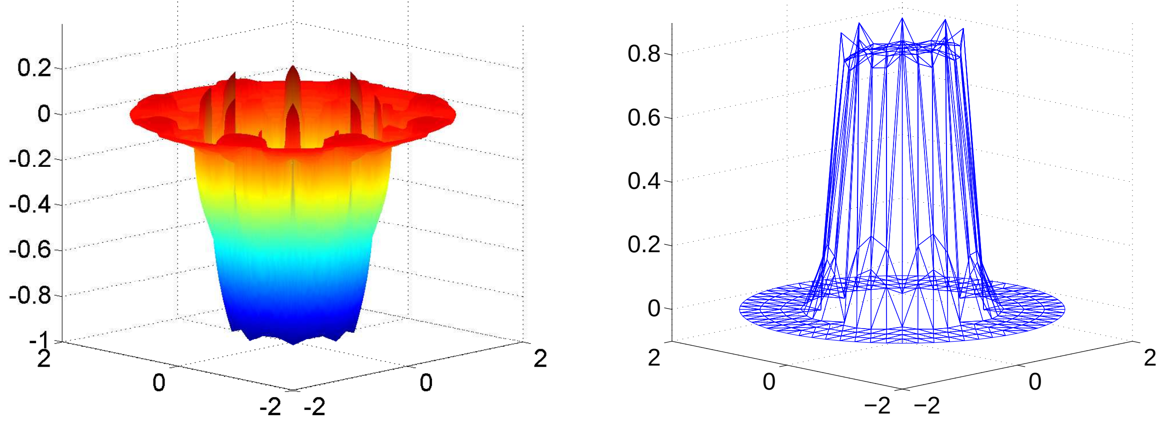

Figure 1: Control (left), and adjoint state (right) (variational discretization)

nt	$\|u-u_h\|_{L^4(\Omega)}$	$\|u-u_h\|_{L^2(\Omega)}$	$\|y-y_h\|_{L^2(\Omega)}$
32	$8.34633 \cdot 10^{-1}$	1.36003	$2.20346 \cdot 10^{-1}$
128	$5.88566 \cdot 10^{-1}$	$9.04770 \cdot 10^{-1}$	$7.97200 \cdot 10^{-2}$
512	$4.84191 \cdot 10^{-1}$	$5.82014 \cdot 10^{-1}$	$3.52102 \cdot 10^{-2}$
	0.54884	0.64041	1.59745
	0.29263	0.66136	1.22499

Table 1: Errors (top) and EOCs for the numerical example (variational discretization)

Piecewise constant controls. We use piecewise constant, discontinuous Ansatz functions for the control u_h. For the numerical solution we use the routine `fmincon` contained in the MATLAB Optimization Toolbox. The state equation is approximated with piecewise linear, continuous finite elements on quasi-uniform triangulations $\mathcal{T}_h$ of $B_2(0)$. The gradient constraints are required element-wise. The resulting discretized optimization problem then reads

$$\min_{u_h \in U_h} J_h(u_h, y_h) = \frac{1}{2}\|y_h - y_0\|^2_{L^2(\Omega)} + \frac{\alpha}{r}\|u_h\|^r_{L^r(\Omega)}$$

$$\text{subject to} \quad \begin{aligned} y_h &= \mathcal{G}_h(u_h) \\ |\nabla y_{h|T}| &\leq \delta = \tfrac{1}{2} \qquad \forall T \in \mathcal{T}_h \end{aligned}$$

In Figs. 2, 3 we present the numerical approximations u_h, y_h, and μ_h on a grid containing $nt = 8192$ triangles, where μ_h is obtained by $\vec{\mu}_h$ according to relation (2.16). Fig. 3 clearly shows that the support of μ_h is concentrated around $|x| = 1$.

In Table 2 again we document the experimental order of convergence. The controls show an approximation behaviour which is slightly better than that predicted by Theorem 2.7. However, in this example we have $\|u\|_{L^\infty}, \|u_h\|_{L^\infty} \leq C$ uniformly in h so that we could expect the convergence order .25 for the L^4-norm of the controls. The L^2-norm of the state

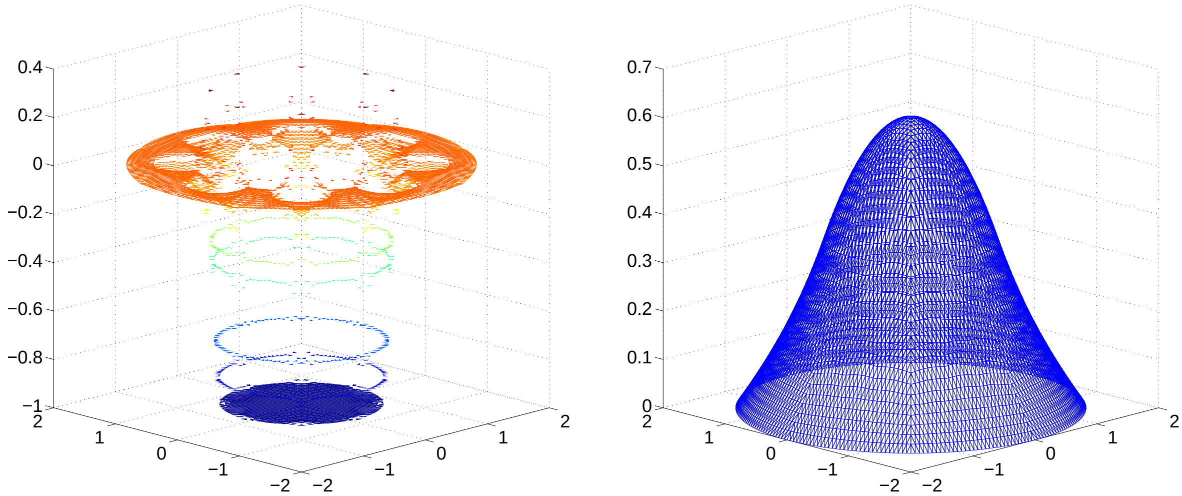

Figure 2: Control (left), and state (right) (piecewise constant controls)

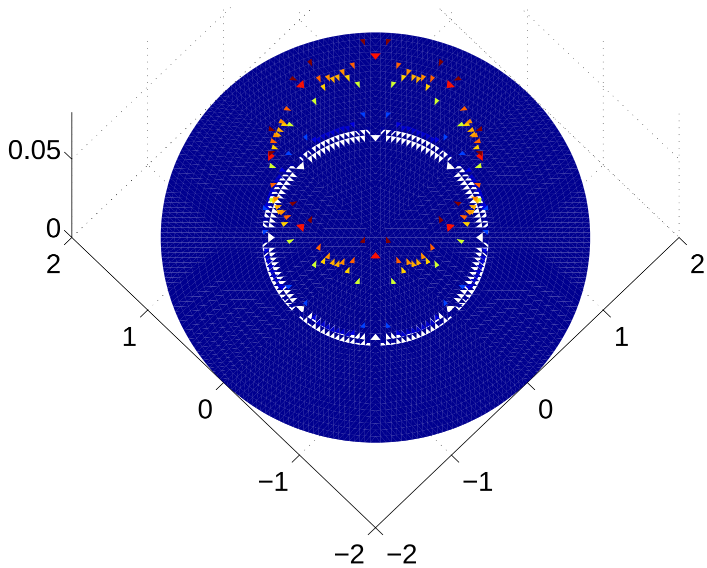

Figure 3: Discrete multiplier (piecewise constant controls)

seems to converge at least with linear order. This can be explained by the high regularity of the exact solution.

In the last column we display the values of $\sum_{T\in\mathcal{T}_h}|T||\vec{\mu}_T|$. These values are expected to converge to 2π as $h\to 0$, since this gives the value of μ applied to the function which is identically equal to 1 on $\bar{\Omega}$.

In order to explain the convergence behaviour of $\|u-u_h\|_{L^2}$ we briefly consider

nt	$\|u-u_h\|_{L^4(\Omega)}$	$\|u-u_h\|_{L^2(\Omega)}$	$\|y-y_h\|_{L^2(\Omega)}$	$\sum_{T\in\mathcal{T}_h} \|T\|\|\vec{\mu}_T\|$
32	$8.34550 \cdot 10^{-1}$	1.37619	$2.30207 \cdot 10^{-1}$	0
128	$5.41825 \cdot 10^{-1}$	$8.45567 \cdot 10^{-1}$	$8.11347 \cdot 10^{-2}$	2.497502
512	$4.57207 \cdot 10^{-1}$	$6.03292 \cdot 10^{-1}$	$3.26818 \cdot 10^{-2}$	4.216741
2048	$3.63216 \cdot 10^{-1}$	$4.11190 \cdot 10^{-1}$	$1.33259 \cdot 10^{-2}$	5.213440
8192	$2.95328 \cdot 10^{-1}$	$2.74811 \cdot 10^{-1}$	$5.27703 \cdot 10^{-3}$	5.739806
	0.67870	0.76530	1.63860	
	0.25455	0.50609	1.36307	
	0.33810	0.56318	1.31796	
	0.30116	0.58653	1.34830	

Table 2: Errors (top), EOCs and multiplier approximation for the numerical example (piecewise constant controls)

Tychonov regularization. Since $u \in L^r(\Omega)$ with $r > d \geq 2$ we may also penalize with the L^2-norm of the control. The corresponding optimal control problem reads

$$\min_{u_h\in U_h} J_h(u_h, y_h) = \frac{1}{2}\|y_h - y_0\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u_h\|^2_{L^2(\Omega)} + \frac{\alpha}{r}\|u_h\|^r_{L^r(\Omega)}$$

$$\text{subject to} \quad \begin{array}{rcl} y_h &=& \mathcal{G}_h(u_h) \\ |\nabla y_{h|T}| &\leq& \delta = \frac{1}{2} \qquad \forall T \in \mathcal{T}_h. \end{array}$$

An analytic solution can be obtained by adapting the constants in our example. Since the variational equality now reads

$$\int_\Omega (p_h + \alpha(u_h + |u_h|^{r-2}u_h))v_h = 0 \text{ for all } v_h \in U_h$$

we have a solution for the same data as before except for $\alpha = 0.5$. An analysis along the lines of Theorems 2.5,2.7 now shows that we also get

$$\|u - u_h\|_{L^2} \leq h^{\frac{1}{2}(1-d/r)},$$

so that in the case of $r = \infty$ the convergence of the L^2-norm of the control error behaves as expected. In Fig. 4 we present the numerical approximations u_h and μ_h on a grid containing $nt = 8192$ triangles. In Table 3 again we investigate the experimental order of convergence for different error functionals. Compared to the previous L^r-regularization all orders of convergence are slighly worse. The control does not oscillate that much along $\partial B_1(0)$.

Acknowledgements

The authors thank Klaus Deckelnick from the University of Magdeburg for many fruitful discussions and the collaboration on the subject of this work. The authors also greatfully acknowledge support of the DFG Priority Program 1253 through grants DFG06-381 and DFG06-382.

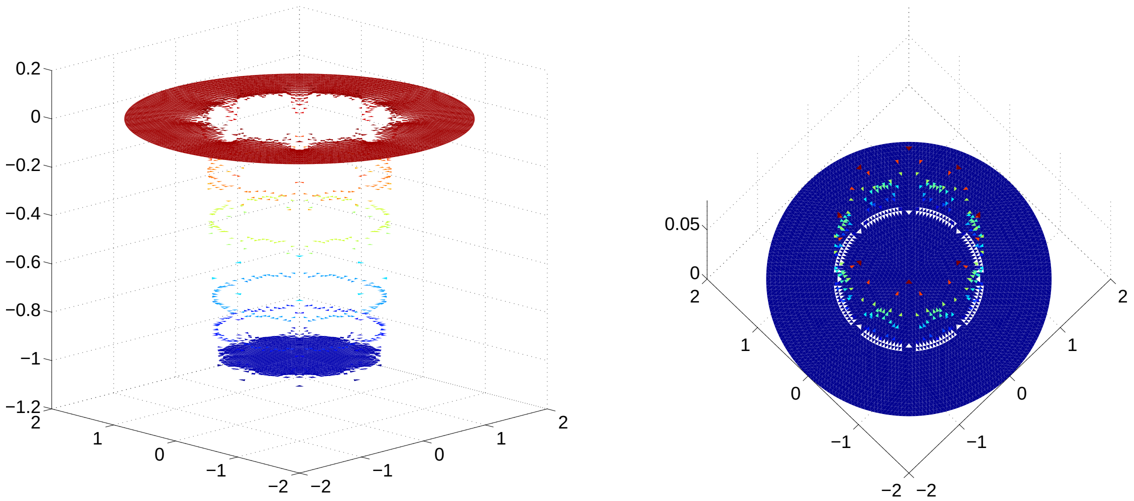

Figure 4: Control (left), and multiplier (right) (Tychonov regularization)

nt	$\|u-u_h\|_{L^4(\Omega)}$	$\|u-u_h\|_{L^2(\Omega)}$	$\|y-y_h\|_{L^2(\Omega)}$	$\sum_{T\in\mathcal{T}_h}\lvert T\rvert\lvert\vec{\mu}_T\rvert$
32	$8.63533\cdot 10^{-1}$	1.22454	$3.83556\cdot 10^{-1}$	0.923216
128	$5.30078\cdot 10^{-1}$	$7.72724\cdot 10^{-1}$	$1.14305\cdot 10^{-1}$	3.656823
512	$4.25213\cdot 10^{-1}$	$5.03372\cdot 10^{-1}$	$4.94054\cdot 10^{-2}$	4.957956
2048	$3.52524\cdot 10^{-1}$	$3.48416\cdot 10^{-1}$	$2.13540\cdot 10^{-2}$	5.602883
8192	$2.89696\cdot 10^{-1}$	$2.41345\cdot 10^{-1}$	$9.58600\cdot 10^{-3}$	5.940486
	0.76678	0.72339	1.90217	
	0.33044	0.64248	1.25741	
	0.27542	0.54054	1.23233	
	0.28570	0.53442	1.16576	

Table 3: Errors (top), EOCs and multiplier approximation for the numerical example (Tychonov regularization)

References

[1] Casas, E., Fernández, L.: *Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state*, Appl. Math. Optimization **27**, 35–56 (1993).

[2] Deckelnick, K., Günther, A., Hinze, M.: *Finite element approximation of elliptic control problems with constraints on the gradient.* Numer. Math. **111**, 335–350, (2009).

[3] Ern, A., Guermond, J.-L.: *Theory and practice of finite elements.* Appl. Math. Sci. **159**, Springer, (2004).

[4] Friedman, A.: *Partial differential equations.* Holt, Rinehart and Winston Inc, (1969).

[5] Gröger, K.: *A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations*. Math. Ann. **283**, 679–687, (1989).

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

[7] Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: *Optimization with pde constraints*. Mathematical Modelling: Theory and Applications **23**, Springer, (2009).

[8] Zanger, D. Z.: *The inhomogeneous Neumann problem in Lipschitz domains*. Comm. Part. Diff. Eqn. **25**, 1771–1808, (2000).