On Finite Element Error Estimates for Optimal Control Problems with Elliptic PDEs

Fredi Tröltzsch

Technische Universität Berlin, Institut für Mathematik
10623 Berlin, Str. d. 17. Juni 136, Sekr. MA 4-5, Germany

Abstract. Discretizations of optimal control problems for elliptic equations by finite element methods are considered. The problems are subject to constraints on the control and may also contain pointwise state constraints. Some techniques are surveyed to estimate the distance between the exact optimal control and the associated optimal control of the discretized problem. As a particular example, an error estimate for a nonlinear optimal control problem with finitely many control values and state constraints in finitely many points of the spatial domain is derived.

1 Introduction

In this paper, we consider optimal control problems of the type

$$\min J(y,u) := \frac{1}{2}\|y - y_d\|^2 + \frac{\lambda}{2}\|u\|^2 \tag{1}$$

subject to

$$\begin{aligned} -\Delta y &= u \quad &\text{in } \Omega \\ y &= 0 \quad &\text{on } \Gamma, \end{aligned} \tag{2}$$

where also further constraints $u \in U_{ad}$ (control constraints) and $y \in Y_{ad}$ (state constraints) may be given. In this setting, $\Omega \subset \mathrm{I\!R}^2$ is a convex, bounded and polygonal domain, $y_d \in L^2(\Omega)$ a given desired state, and $\lambda > 0$ is a fixed regularization parameter. By $\|\cdot\|$, the natural norm of $L^2(\Omega)$ is denoted. In the last section, $B(u,\rho) \subset \mathrm{I\!R}^m$ is the open ball of radius ρ around u. In the paper, c is a generic constant that is more or less arbitrarily adapted.

Our main issue is to estimate the error arising from a finite element discretization of such problems. First, we consider the problem without control and state constraints. Next, we explain how the error can be estimated, if control constraints are given. We briefly survey recent results on probems with state constraints and discuss finally a problem with finite-dimensional control space and state constraints given in finitely many points of the spatial domain.

2 Optimal control problem with control constraints

2.1 The unconstrained optimal control problem

The problem Let us start the tour through error estimates by a quite simple approach for the unconstrained problem (1), (2), where no further constraint on u or y are given, i.e. $U_{ad} = L^2(\Omega)$ and $Y_{ad} = L^2(\Omega)$.

For all $u \in L^2(\Omega)$, there is exactly one state $y = y(u) \in H^2(\Omega) \cap H_0^1(\Omega)$ and the mapping $\Lambda : L^2(\Omega) \to H^2(\Omega) \cap H_0^1(\Omega)$, $\Lambda : u \mapsto y(u)$, is continuous. We consider Λ also with range in $L^2(\Omega)$ and denote this "solution operator" by S, i.e. $S = E_{H^2 \to L^2}\Lambda$, where $E_{H^2 \to L^2}$ denotes the continuous injection of $H^2(\Omega)$ in $L^2(\Omega)$.

By S, we are able to formally eliminate the PDE and to transform the problem to the control reduced quadratic optimization problem

$$(P) \qquad \min_{u \in L^2(\Omega)} f(u) := J(Su, u) = \frac{1}{2}\|Su - y_d\|^2 + \frac{\lambda}{2}\|u\|^2.$$

The existence of a unique (optimal) solution of this problem is a standard result. In all what follows, we denote by $\bar{u}$ the unique *optimal control* and by $\bar{y} = y(\bar{u})$ the associated *optimal state.*

All further results on the necessary optimality conditions for (P) and its version with additional control constraints are stated without proofs. They are discussed extensively in the forthcoming textbook [1].

Necessary and sufficient optimality condition It is clear that $f'(\bar{u}) = 0$ is necessary for the optimality of $\bar{u}$, hence

$$f'(\bar{u}) = S^*(S\bar{u} - y_d) + \lambda\bar{u} = 0$$

must hold, where S^* denotes the adjoint operator of S. It is very useful to introduce an auxiliary function p by $\bar{p} := S^*(S\bar{u} - y_d)$. This function $\bar{p}$ is called *adjoint state* associated with $\bar{u}$. Therefore, we have

$$\bar{p} + \lambda\bar{u} = 0. \tag{3}$$

The adjoint state $\bar{p}$ is the solution of the *adjoint equation*

$$\begin{aligned} -\Delta p &= \bar{y} - y_d && \text{in } \Omega \\ p &= 0 && \text{on } \Gamma. \end{aligned} \tag{4}$$

To determine the unknown triplet $(\bar{y}, \bar{u}, \bar{p})$, we have to solve the *optimality system* (2), (4), (3). Invoking (3), i.e. inserting $u = -\lambda^{-1}p$ in the state equation, the optimality system

$$\begin{aligned} -\Delta y + \lambda^{-1}p &= 0 & -\Delta p - y &= -y_d && \text{in } \Omega \\ y &= 0, & p &= 0 && \text{on } \Gamma \end{aligned} \tag{5}$$

is obtained. Having solved this, we obtain the optimal control by $\bar{u} := -\lambda^{-1}p$.

Discretized problem and error estimate We assume a regular triangulation $\mathcal{T}$ of Ω with mesh size h, triangles T_i, and piecewise linear and continuous ansatz functions Φ_i, $i = 1, \ldots, n$, which generate the finite-dimensional subspace $V_h = \text{ span } \{\Phi_1, \ldots, \Phi_n\} \subset H_0^1(\Omega)$. We do not explain the standard notion of regularity of the triangulation. Instead, we just assume that the standard finite element error estimate (7) below is satisfied.

In the discretized problem, the state y_h associated with u is determined by $y_h \in V_h$ and

$$\int_\Omega \nabla y_h \cdot \nabla \Phi_i \, dx = \int_\Omega u \; \Phi_i \, dx \qquad \forall i = 1, \ldots, n. \tag{6}$$

To each $u \in L^2(\Omega)$, there exists exactly one solution $y_h \in H_0^1(\Omega)$ of (6) denoted by $y_h(u)$. From the finite element analysis for regular grids, the error estimate

$$h \, \|y_h(u) - y(u)\|_{H^1(\Omega)} + \|y_h(u) - y(u)\| \leq c \, h^2 \|u\| \tag{7}$$

is known for all sufficiently small $h > 0$, where the constant c does not depend on u or h. Let us introduce the mapping $S_h : L^2(\Omega) \to L^2(\Omega)$, $S_h : u \mapsto y_h(u)$. In terms of S and S_h, (7) is equivalent to

$$\|S - S_h\|_{L^2(\Omega) \to L^2(\Omega)} \leq c \, h^2. \tag{8}$$

Analogously to the former section, the *discretized optimal control problem* can be formulated in control reduced form as

$$(P_h) \qquad \min f_h(u) := \frac{1}{2} \|S_h u - y_d\|^2 + \frac{\lambda}{2} \|u\|^2.$$

This discretized problem has a unique optimal control denoted by $\bar{u}_h$ with associated state $\bar{y}_h$. The reader might be surprised that the control u in (P_h) is not discretized. In fact, we do not need this here, since the optimality conditions will automatically imply $u_h \in V_h$.

It is easy to estimate the error $\|\bar{u}_h - \bar{u}\|$. We write down the necessary optimality conditions for both optimal controls,

$$\begin{aligned} S^*(S\bar{u} - y_d) + \lambda \bar{u} &= 0 \\ S_h^*(S_h \bar{u}_h - y_d) + \lambda \bar{u}_h &= 0, \end{aligned}$$

multiply them scalarly by $\bar{u} - \bar{u}_h$, subtract the results and re-order. This yields

$$\begin{aligned} \|S_h(\bar{u} - \bar{u}_h)\|^2 + \lambda \|\bar{u} - \bar{u}_h\|^2 &\leq |\, (y_d \, , \, (S - S_h)(\bar{u} - \bar{u}_h)) \,| \\ &\leq \|y_d\| \, c \, h^2 \, \|\bar{u} - \bar{u}_h\|, \end{aligned}$$

where $(\cdot \, , \, \cdot)$ denotes the natural inner product of $L^2(\Omega)$. Consequently, we have the L^2 *error estimate*

$$\|\bar{u} - \bar{u}_h\| \leq c \, \lambda^{-1} \, h^2 \, \|y_d\|. \tag{9}$$

This was easy, since we considered the unconstrained case that is not really interesting in optimization. Let us include also constraints on the control.

2.2 Constraints on the control

Optimality conditions Let $u_a < u_b$ be two real numbers. Consider the control-constrained problem

$$(PC) \qquad \min_{u \in U_{ad}} f(u) = \frac{1}{2}\|Su - y_d\|^2 + \frac{\lambda}{2}\|u\|^2$$

with

$$U_{ad} = \{u \in L^2(\Omega) \,:\, u_a \le u(x) \le u_b \text{ a.e. in } \Omega\}.$$

Again, the problem has a unique optimal control $\bar{u}$. However, due to the constraints, we cannot expect that $f'(\bar{u}) = 0$. Instead, the *variational inequality*

$$f'(\bar{u})(u - \bar{u}) \ge 0 \quad \forall u \in U_{ad}$$

is necessary and sufficient for optimality of $\bar{u}$. It expresses the intuitively clear observation that, in a minimum, the function f cannot decrease in any feasible direction. In terms of S, this means

$$(S^*(S\bar{u} - y_d) + \lambda\bar{u}\,,\, u - \bar{u}) \ge 0 \quad \forall u \in U_{ad} \tag{10}$$

or equivalently

$$\int_\Omega \big(p(x) + \lambda\bar{u}(x)\big)(u(x) - \bar{u}(x))\,dx \ge 0 \quad \forall u(\cdot) \in U_{ad}.$$

A simple pointwise discussion of this inequality reveals that almost everywhere

$$\bar{u}(x) = \begin{cases} u_a & \text{if } \; p(x) + \lambda\bar{u}(x) > 0 \\ u_b & \text{if } \; p(x) + \lambda\bar{u}(x) < 0 \end{cases} \tag{11}$$

and we have, of course, $\bar{u}(x) = -\lambda^{-1}p(x)$ if $p(x) + \lambda\bar{u}(x) = 0$. From this, one derives with some effort the well-known projection formula

$$\bar{u}(x) = \mathbb{P}_{[u_a,u_b]}\Big\{ -\frac{1}{\lambda}p(x)\Big\} \qquad \text{a.e. in } \Omega, \tag{12}$$

where $\mathbb{P}_{[u_a,u_b]} : \mathbb{R} \to [u_a, u_b]$ denotes projection onto $[a, b]$. This projection formula shows that, although we have $p \in H^2(\Omega)$, $\bar{u}$ can exhibit corners in the points, where the bounds u_a and u_b are reached. Hence in general we can only expect $u \in H^1(\Omega)$. Moreover, $(\bar{y}, \bar{u}, \bar{p})$ cannot be obtained from a smooth coupled system of PDEs as (5). Therefore, error estimates are more difficult. They also depend on the way how the control function u is discretized.

Discretization by piecewise constant controls The most common way of control discretization is working with *piecewise constant* controls. Here, the set of admissible discretized controls is defined by

$$U_{ad}^h = \{u_a \le u(\cdot) \le u_b \,:\, u \text{ is constant on each triangle } T_i\}$$

and the associated *discretized problem* is

$$(PC_h) \qquad \min_{u_h \in U_{ad}^h} f_h(u_h),$$

where f_h and y_h are defined as in the last section. The difference to (P_h) consists in the appearance of U_{ad}^h. Let $\bar{u}_h$ denote the unique optimal control of (PC_h) and let $\bar{y}_h$ be the associated (discretized) state. Then a discrete counterpart to the variational inequality (10) must be satisfied,

$$(S_h^*(S_h\bar{u}_h - y_d) + \lambda\bar{u}_h \,,\, u_h - \bar{u}_h) \ge 0 \quad \forall u_h \in U_{ad}^h. \tag{13}$$

By the discrete adjoint state $p_h := S_h^*(S_h\bar{u}_h - y_d)$, this is equivalent to

$$\bar{u}(x) = \mathbb{P}_{[u_a,u_b]}\Big\{ -\frac{1}{\lambda\,|T_i|}\int_{T_i} p_h(x)\,dx\Big\} \quad \forall\, x \in T_i,\, \forall\, i = 1,\ldots,M, \tag{14}$$

where $|T_i|$ is the area of the triangle T_i. Now we cannot derive an error estimate in the same way as before. First, $\bar{u}$ cannot be inserted in (14), as $\bar{u}_h$ is not piecewise constant. As a substitute, we use the interpolant $\Pi_h\bar{u}$ defined by

$$(\Pi_h\bar{u})(x) := -\frac{1}{|T_i|}\int_{T_i} \bar{u}(x)\,dx \quad \forall\, x \in T_i. \tag{15}$$

It holds that $\Pi_h\bar{u} \in U_{ad}^h$ and, with some $c > 0$ not depending on h,

$$\|\bar{u} - \Pi_h\bar{u}\| \le c\,h. \tag{16}$$

Now, we might insert $\bar{u}_h$ in (10), $\Pi_h\bar{u}$ in (13), add the two inequalities obtained and resolve for $\|\bar{u} - \bar{u}_h\|$ to estimate the error. This procedure only yields a non-optimal estimate of the order $\sqrt{h}$. Some further tricks avoid the square root, cf. [2] or the early papers [3], [4].

Below, we explain a *perturbation approach.* Its main idea goes back to [5] and was applied in [6] to nonlinear elliptic problems.

The function $\Pi_h\bar{u}$ will only fulfill the variational inequality (13) if it is optimal for (PC_h) by chance. However, it satisfies the variational inequality for the *perturbed control problem*

$$(PC_{\zeta_h}) \qquad \min_{u_h \in U_{ad}^h} \Big\{ f_h(u_h) + \int_\Omega \zeta_h(x)\,u_h(x)\,dx \Big\},$$

if $\zeta_h(x)$ is defined such that $\Pi_h\bar{u}$ satisfies the associated projection formula

$$\Pi_h\bar{u}(x) = \mathbb{P}_{[u_a,u_b]}\Big\{ -\frac{1}{\lambda\,|T_i|}\int_{T_i} (p_h(x) + \zeta_h(x))\,dx\Big\} \quad \forall\, x \in T_i,\, \forall\, i = 1,\ldots,M.$$

Notice that the derivative of the perturbed functional at a function u is equal to $S_h^*(S_h u - y_d) + \zeta_h + \lambda\, u = p_h + \zeta_h + \lambda\, u$ so that $p_h + \zeta_h$ plays the role of the former p_h. How the function ζ_h must be constructed? If $u_a < \varPi_h \bar u(x) < u_b$ holds in a triangle T_i, then $\lambda |T_i| \varPi_h \bar u(x) + \int_{T_i} (p_h + \zeta_h)\, dx = 0$ must hold on T_i. This follows from (14), applied to $\varPi_h \bar u$ and (PC_{ζ_h}). Therefore, we define ζ_h by

$$\zeta_h(x) \equiv -\varPi_h \bar u(x) + \frac{1}{\lambda\, |T_i|} \int_{T_i} p_h\, dx \qquad \text{on } T_i.$$

If, on T_i, $\varPi_h \bar u(x) \equiv u_a$, then $\lambda |T_i| \varPi_h \bar u(x) + \int_{T_i} (p_h + \zeta_h)\, dx \geq 0$ must hold on T_i (adapt (11) to (PC_{ζ_h})). To compensate negative values, we define ζ_h by

$$\zeta_h(x) \equiv \Big[\varPi_h \bar u(x) + \frac{1}{\lambda\, |T_i|} \int_{T_i} p_h\, dx \Big]_- \qquad \text{on } T_i,$$

where $a_- = (|a| - a)/2 \geq 0$ denotes the negative part of a real number. Analogously, we define $\zeta_h := -[\ldots]_+$ via the associated positive part to compensate a positive value, if $\varPi_h \bar u(x) \equiv u_b$. It is not difficult to show that

$$\|\zeta_h\| \leq c\, \|\bar u - \varPi_h \bar u\| \leq \tilde c\, h.$$

Now we proceed similarly as in the last section. We insert $\varPi_h \bar u$ in the variational inequality (13) for $\bar u_h$ and insert $\bar u_h$ in the perturbed variational inequality for $\varPi_h \bar u$ to obtain

$$\begin{aligned} (S_h^*(S_h \bar u_h - y_d) + \lambda \bar u_h\, ,\, \varPi_h \bar u - \bar u_h) &\geq 0, \\ (S_h^*(S_h \varPi_h \bar u - y_d) + \zeta_h + \lambda \varPi_h \bar u\, ,\, \bar u_h - \varPi_h \bar u) &\geq 0. \end{aligned}$$

Adding both inequalities, we obtain after some re-ordering and ignoring the term $\|S_h(\bar u_h - \varPi_h \bar u)\|^2$

$$\|\bar u_h - \varPi_h \bar u\|^2 \leq \lambda^{-1}\, (\zeta_h\, ,\, \bar u_h - \varPi_h \bar u) \leq \lambda^{-1}\, \|\zeta_h\|\, \|\bar u_h - \varPi_h \bar u\|.$$

In view of (16), an obvious application of the triangle inequality yields finally

$$\|\bar u_h - \bar u\| \leq c\, h \tag{17}$$

with some $c > 0$ not depending on h. This is the *optimal error estimate for piecewise constant approximations* of $\bar u$.

The same order of the error can be derived for problems with semilinear elliptic equations, both for distributed and boundary controls, [6], [7]. However, thanks to non-convexity, the situation is more delicate. Different locally optimal controls might appear. They should satisfy a second-order sufficient optimality condition to have a unique approximating locally optimal control in a certain neighborhood, cf. also the discussion in the last section.

The error analysis for *piecewise linear* controls is more difficult. We refer only to [8] for Neumann and to [9], [10], and [11] for Dirichlet boundary control problems. In the Neumann case, the order $h^{3/2}$ can be expected for the error.

Variational discretization The situation is easier, if the control functions are (formally) *not discretized*, i.e. if we consider the discretized problem

$$(P_h) \qquad \min_{u \in U_{ad}} f_h(u).$$

At first glance, this consideration seems to be useless. How should one be able to compute the optimal control without any kind of discretization? However, take a look at the finite element approximation of the optimality system

$$\begin{aligned} -\Delta y &= \mathbb{P}_{[u_a,u_b]}\big\{ -\lambda^{-1} p \big\} & -\Delta p - y &= -y_d & &\text{in } \Omega \\ y &= 0, & p &= 0 & &\text{on } \Gamma, \end{aligned} \tag{18}$$

where u is eliminated by the projection formula (12). This nonsmooth nonlinear system can be solved numerically to obtain the (discrete) state y_h and adjoint state p_h, [12]. Then $\bar{u}_h$ is found by $\bar{u}_h = \mathbb{P}_{[u_a,u_b]}\{-\lambda^{-1} p_h\}$. It is piecewise linear but does not in general belong to V_h. This approach of *variational discretization* is also useful in iterative methods of optimization, cf. [13].

For this variational discretization, an error estimate of the order h^2 is easily derived: We repeat a similar procedure as before and insert $\bar{u}_h$ in the variational inequality for $\bar{u}$, $\bar{u}$ in the variational inequality for $\bar{u}_h$ (all $u \in U_{ad}$ are now admitted!),

$$\begin{aligned} (S^*(S\bar{u} - y_d) + \lambda \bar{u} \,,\, \bar{u}_h - \bar{u}) &\geq 0, \\ (S_h^*(S_h \bar{u}_h - y_d) + \lambda \bar{u}_h \,,\, \bar{u} - \bar{u}_h) &\geq 0. \end{aligned}$$

Next, we add both inequalities and re-order as we proceeded to derive (9),

$$\|S_h(\bar{u} - \bar{u}_h)\|^2 + \lambda \|\bar{u} - \bar{u}_h\|^2 \leq |\, (y_d \,,\, (S - S_h)(\bar{u} - \bar{u}_h)) \,| \leq \|y_d\| \, c\, h^2 \, \|\bar{u} - \bar{u}_h\|.$$

Consequently, we have the optimal L^2-estimate

$$\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} \leq c\, \lambda^{-1}\, h^2\, \|y_d\|.$$

The same optimal order can be obtained under natural assumptions with piecewise constant controls by one smoothing step, cf. [14].

3 Problems with state constraints

3.1 Available error estimates

Here, we admit also state constraints. Now the error analysis is more difficult. Currently, this is a very active field of research. A detailed survey on relevant contributions would go beyond the scope of this paper. We mention first results for problems with finitely many state constraints in [15] and the convergence of discretizations for *pointwise state constraints* of the type $y(x) \leq c$ a.e. in Ω in [16]. To have a comparison with the results of the next section, we state also recent results for elliptic problems with pointwise state constraints: In [17], for $\|\bar{u} - \bar{u}_h\|$ the order $h^{1-\varepsilon}$ was derived in $\Omega \subset \mathbb{R}^2$ and $h^{1/2-\varepsilon}$ for $\Omega \subset \mathbb{R}^3$. Recently, this was improved in [18] to $h|\log h|$ for 2D and $h^{1/2}$ for 3D domains.

3.2 Control in $\mathbb{R}^m$ and state constraints in finitely many points

Let us discuss a simpler problem with semilinear elliptic equation, controls in $\mathbb{R}^m$ and state constraints in finitely many points $x_1, \ldots, x_\ell$ of Ω:

$$(PS) \qquad \begin{cases} \min\limits_{u \in U_{ad}} J(y_u, u) = \dfrac{1}{2}\|y_u - y_d\|^2 + \dfrac{\lambda}{2}|u|^2 \\ \text{subject to} \\ g_i(y_u(x_i)) = 0, \qquad \text{for all } i = 1, \ldots, k, \\ g_i(y_u(x_i)) \le 0, \qquad \text{for all } i = k+1, \ldots, \ell, \end{cases}$$

where y_u is the solution to the state equation

$$\begin{aligned} -\Delta y(x) + d(y(x), u) &= 0 \quad \text{in } \Omega \\ y(x) &= 0 \quad \text{on } \Gamma \end{aligned} \tag{19}$$

and $U_{ad} = \{u \in \mathbb{R}^m : u_a \le u \le u_b\}$ with given $u_a \le u_b$ of $\mathbb{R}^m$. We assume $l \ge 1$ and set $k = 0$, if only inequality constraints are given and $k = l$, if only equality constraints are given.

We assume for short that $d : \mathbb{R}^2 \to \mathbb{R}$ and $g_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, \ell$, are twice differentiable with locally Lipschitz second-order derivatives and that d is monotone non-decreasing with respect to y. In [19], the problem is considered in a slightly more general setting. Thanks to our assumptions, the mapping $u \mapsto y_u$ is continuous from $\mathbb{R}^m$ to $H_0^1(\Omega) \cap C(\bar{\Omega})$, hence the values $y_u(x_i)$ are well defined. Therefore, we consider $S : u \mapsto y_u$ as mapping from $\mathbb{R}^m$ to $H_0^1(\Omega) \cap C(\bar{\Omega})$.

To convert (PS) into a finite-dimensional nonlinear programming problem, we define again $f : \mathbb{R}^m \mapsto \mathbb{R}$ of class $C^{2,1}$ by $f(u) = J(y_u, u) = J(S(u), u)$. Thanks to our assumptions, in particular the Lipschitz properties of the second derivatives of d and the g_i, the mapping S has a locally Lipschitz continuous second-order derivative S''. Moreover, we define $G : \mathbb{R}^m \to \mathbb{R}^\ell$ by

$$G(u) := [g_1(y_u(x_1)), \ldots, g_\ell(y_u(x_\ell))]^\top. \tag{20}$$

To cover the equality and inequality constraints in a unified way, we introduce the convex cone $K = \{z \in \mathbb{R}^\ell : z_i = 0,\ i = 1, \ldots, k,\ z_i \le 0, i = k+1, \ldots, \ell\}$ and write $z \le_K 0$ iff $z \in K$. By these definitions, (PS) becomes equivalent to the nonlinear programming problem

$$(N) \qquad \begin{cases} \min f(u) \\ G(u) \le_K 0, \quad u \in U_{ad}. \end{cases} \tag{21}$$

The discretized optimal control problem (PS$_h$) is defined on substituting y_u by its finite-element approximation $y_{h,u}$, obtained from

$$\int_\Omega \nabla y_h \cdot \nabla v_h \, dx + \int_\Omega d(y_h, u)\, v_h \, dx = 0 \qquad \forall v_h \in V_h.$$

Introducing $G_h(u) := [g_1(y_{h,u}(x_1)), \ldots, g_\ell(y_{h,u}(x_\ell))]^\top$ we express this problem as finite-dimensional nonlinear programming problem

$$(N_h) \qquad \left\{ \begin{array}{l} \min f_h(u) \\ G_h(u) \le_K 0, \quad u \in U_{ad}. \end{array} \right. \tag{22}$$

Let Ω_0 and Ω_1 be open sets such that $\{x_1, \ldots, x_\ell\} \subset \Omega_0$ and $\bar{\Omega}_0 \subset \Omega_1 \subset \bar{\Omega}_1 \subset \Omega$. Then there exists a constant $c > 0$, independent of h and $u \in U_{ad}$, such that

$$\|y_u - y_{h,u}\|_{L^\infty(\Omega_0)} \le c \left(h^2 |\log h| \, \|y_u\|_{W^{2,\infty}(\Omega_1)} + h^2 \, \|y_u\|_{H^2(\Omega)} \right), \tag{23}$$

cf. [19]. Thanks to this estimate and to our assumptions, it holds

$$\begin{array}{l} |f(u) - f_h(u)| + |f'(u) - f_h'(u)| + |f''(u) - f_h''(u)| \le c\, h^2 \\ |G(u) - G_h(u)| + |G'(u) - G_h'(u)| + |G''(u) - G_h''(u)| \le c\, h^2 |\log h| \end{array} \tag{24}$$

for all $u \in U_{ad}$, with some constant c not depending on h and u, [19].

In all what follows, $\bar{u}$ is a locally optimal reference solution of (PS), hence also of (N). We show the existence of an associated locally optimal solution $\bar{u}_h$ of (PS$_h$), (or (N$_h$)) converging to $\bar{u}$ as $h \downarrow 0$. Our main aim is to estimate $|\bar{u} - \bar{u}_h|$. For short, we use the abbreviation $\alpha(h) = h^2 |\log h|$.

Our error analysis is based on 3 main assumptions. To formulate them, we need some standard definitions of nonlinear programming which are explained below.

We first extend the vector $G(u) \in \mathbb{R}^\ell$ to $\hat{G}(u) \in \mathbb{R}^{\ell+2m}$ by including the box constraints defining U_{ad}. We add the $2m$ further components

$$\begin{array}{ll} G_{\ell+i}(u) = u_{a,i} - u_i, & \text{for } i = 1, \ldots, m \\ G_{\ell+m+i}(u) = u_i - u_{b,i}, & \text{for } i = 1, \ldots, m, \end{array}$$

and put $\hat{G}(u) := (G_i(u))_{i=1,\ldots,\ell+2m}$. Then all constraints of the problem can be unified by $\hat{G}(u) \le_K 0$, where K is re-defined accordingly. We define the *Lagrangian function* $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^{\ell+2m}$ by

$$\mathcal{L}(u, \nu) = f(u) + \sum_{i=1}^{\ell+2m} \nu_i \, G_i(u).$$

The *index set* $\mathcal{A}(\bar{u})$ *of active constraints* at $\bar{u}$ is defined by

$$\mathcal{A}(\bar{u}) = \{ i \in \{1, \ldots, \ell + 2m\} : G_i(\bar{u}) = 0 \}.$$

We now formulate the main assumptions:

Robinson regularity condition: At $\bar{u}$, it holds that

$$0 \in \text{int } \{ G(\bar{u}) + G'(\bar{u})(U_{ad} - \bar{u}) + K \},$$

where the set in braces is defined as $\cup \{ G(\bar{u}) + G'(\bar{u})(u - \bar{u}) + k \mid u \in U_{ad}, \; k \in K \}$.

It is known that this regularity assumption is sufficient for the existence of a Lagrange multiplier $\bar{\nu}$ associated with the (locally) optimal solution $\bar{u}$.

Strong second-order sufficient optimality condition: For the pair $(\bar{u}, \bar{\nu})$, it holds

$$v^\top \frac{\partial^2 \mathcal{L}(\bar{u}, \bar{\nu})}{\partial u^2} v > 0 \qquad \forall v \in C_{\bar{u}},\ v \neq 0,$$

where $C_{\bar{u}} \subset \mathrm{I\!R}^m$ is defined by

$$C_{\bar{u}} = \{v \,|\, G_i'(\bar{u})v = 0 \ \forall i \in \{1, \ldots, k\} \cup \{i \in \{k+1, \ldots, \ell + 2m\} : \bar{\nu}_i > 0\}\}\,.$$

Linear independence condition of active gradients: This condition is satisfied if all vectors of the set $\{\nabla G_i(\bar{u}) \,|\, i \in \mathcal{A}(\bar{u})\}$ are linearly independent.

Theorem 1. *Under the three assumptions stated above, there exists a constant* $c > 0$ *not depending on* h *such that, for all sufficiently small* $h > 0$*, a unique locally optimal control* $\bar{u}_h$ *exists in a neighborhood of* $\bar{u}$ *and it holds*

$$|\bar{u} - \bar{u}_h| \leq c\, h^2\, |\log h|.$$

We do not entirely show this result here. Instead, we show an estimate of the order $h\sqrt{|\log h|}$. In this way, we prepare the proof of the full order in [19]. First, we approximate admissible vectors for (N) by admissible ones for (N$_h$) with the order $\alpha(h)$ and vice versa.

Lemma 2. *Suppose that* $\bar{u}$ *is feasible for* (N) *and satisfies the Robinson regularity condition. Then there are* $c > 0$ *independent of* h *and* $h_0 > 0$ *such that, for each* $h \in (0, h_0)$ *an admissible* u_h *for problem* (P_h) *exists with*

$$|\bar{u} - u_h| \leq c\, \alpha(h). \tag{25}$$

Proof. To be consistent with the notation of [20], we write

$$\mathcal{G}(h, u) = \begin{cases} G_h(u), & \text{if } u \in U_{ad} \text{ and } h > 0, \\ G(u), & \text{if } u \in U_{ad} \text{ and } h = 0, \\ \emptyset, & \text{if } u \notin U_{ad}. \end{cases}$$

Thanks to (24), $\mathcal{G}$ and $\partial \mathcal{G}/\partial u$ are continuous at the point $(h, u) = (0, \bar{u})$. Moreover, we have $\mathcal{G}(0, \bar{u}) \leq_K 0$. In view of the Robinson regularity condition, the assumptions of the generalized implicit function in [20] are fulfilled. We obtain the existence of neighborhoods $\mathcal{N}$ of $h = 0$ and $\mathcal{O}$ of $\bar{u}$ such that, for all $h \in \mathcal{N}$, the inequality $\mathcal{G}(h, u) \leq_K 0$ has a solution $u \in \mathcal{O}$, and it holds

$$dist[v, \varSigma(h)] \leq c\, |\mathcal{G}(h, v)_+|, \quad \forall h \in \mathcal{N}, \forall v \in \mathcal{O}, \tag{26}$$

where $\varSigma(h) = \{u \in U_{ad} \,|\, \mathcal{G}(h, u) \leq_K 0\}$ is the solution set of the inequality and $dist$ denotes the Euclidean distance of a point to a set. The value $|\mathcal{G}(h, v)_+|$ is the distance of the set $\mathcal{G}(h, v) + K$ to the origin and measures the residual of v with respect to the inequality $\mathcal{G}(h, v) \leq_K 0$, cf. [20], p. 498. Inserting $v = \bar{u}$ in (26), we deduce

$$dist[\bar{u}, \varSigma(h)] \leq c\, |\mathcal{G}(h, \bar{u})_+| \leq c(|\mathcal{G}(0, \bar{u})_+| + |\mathcal{G}(h, \bar{u})_+ - \mathcal{G}(0, \bar{u})_+|) \leq 0 + c\, \alpha(h).$$

Hence, there exists $u_h \in \varSigma(h)$ with $|\bar{u} - u_h| \leq c\, \alpha(h)$. The statement is shown.

Lemma 3. *Let the reference solution $\bar{u}$ satisfy the linear independence condition. Then, for all given $\rho > 0$ and all sufficiently small $h > 0$, the auxiliary problem*

$$(N_{h,\rho}) \qquad \begin{cases} \min f_h(u) \\ G_h(u) \leq_K 0, \\ u \in U_{ad} \cap cl\, B(\bar{u}, \rho) \end{cases} \tag{27}$$

is solvable. If $\bar{u}_h$ is any optimal solution to this problem, then an admissible element v_h for (N) exists satisfying with some $c > 0$ independent of h

$$|\bar{u}_h - v_h| \leq c\, \alpha(h). \tag{28}$$

Proof. (i) *Solvability of $(N_{h,\rho})$* : For a positive h_0 and all $h \in (0, h_0)$, the admissible set of $(N_{h,\rho})$ is not empty, because u_h constructed in Lemma 2 satisfies all constraints. The existence of an optimal $\bar{u}_h$ follows immediately. We have to find v_h in U_{ad} with $G(v_h) \leq_K 0$ and $|v_h - \bar{u}_h| \leq c\, \alpha(h)$. Below, we cover the inequality constraints of U_{ad} by the extended vector function $\hat{G}(u) : \mathbb{R}^m \to \mathbb{R}^{\ell+2m}$ introduced 2 pages before. Let us set in this proof $G := \hat{G}$ and $\ell := \ell + 2m$ to avoid an extensive use of the hat sign. Hence we have to construct v_h such that

$$G_i(v_h) = 0, \quad i = 1, \ldots, k, \qquad G_i(v_h) \leq 0, \quad i = k+1, \ldots, \ell.$$

(ii) *Construction of an equation for v_h*: Notice that $\bar{u}_h \in cl\, B(\bar{u}, \rho)$ for all $h \leq h_0$. Therefore, if ρ is taken small enough, all inactive components $G_i(\bar{u})$ are inactive for $\bar{u}_h$ as well and there exists $\varepsilon > 0$ such that

$$G_i(\bar{u}_h) \leq -\varepsilon < 0 \quad \forall i \in I, \quad \forall h \leq h_0, \tag{29}$$

where I is the set of all inactive indices i of $\bar{u}$ in $\{k+1, \ldots, \ell\}$.

Suppose that r constraints are active at $\bar{u}$, $k \leq r \leq m$. After renumbering, if necessary, we can assume that those with the numbers $1 \leq i \leq r$ are active, hence $G_1(\bar{u}) = \ldots = G_r(\bar{u}) = 0$. By the independence condition, the associated gradients $\nabla G_i(\bar{u})$ are linearly independent. If ρ is small enough, also $\nabla G_1(\bar{u}_h), \ldots, \nabla G_r(\bar{u}_h)$ are linearly independent. Consider the matrix $B_h = [\nabla G_1(\bar{u}_h), \ldots, \nabla G_r(\bar{u}_h)]^\top$. Since B_h has full rank r, we find an invertible submatrix D_h such that (after renumbering of the components of u, if necessary) $B_h = [D_h, E_h]$ holds with some matrix E_h. Define $F_h : \mathbb{R}^r \to \mathbb{R}^r$ by

$$F_{h,i}(w) := G_i(w, \bar{u}_{h,r+1}, \ldots, \bar{u}_{h,m}) - G_{h,i}(\bar{u}_h), \quad i = 1, \ldots, r.$$

To find v_h, we fix its $m - r$ last components by $v_{h,i} := \bar{u}_{h,i}$, $i = r+1, \ldots, m$, and determine the first r components as the solution w of the system

$$F_h(w) = 0, \tag{30}$$

i.e. we set $v_{h,i} := w_i$, $i = 1, \ldots, r$.

(iii) *Solvability of (30)*: In this part of the proof, we follow a technique used by Allgöwer et al. [21]. We define for convenience $\bar{w}_h := (\bar{u}_{h,1}, \ldots, \bar{u}_{h,r})^\top$, $\bar{w} := (\bar{u}_1, \ldots, \bar{u}_r)^\top$ and have

$$|F_h(\bar{w}_h)| \leq c\, \alpha(h), \tag{31}$$

since $|G_i(\bar{u}_h) - G_{h,i}(\bar{u}_h)| \le c\,\alpha(h)$ holds for all $1 \le i \le r$.

Thanks to (24) and the Lipschitz assumptions, there exist $\gamma > 0$, $\beta > 0$ with

$$\|F_h'(w_1) - F_h'(w_2)\| \le \gamma\,|w_1 - w_2| \qquad \forall w_i \in B(\bar{w},\rho),$$
$$\|(F_h'(w))^{-1}\| \le \beta \qquad \forall w \in B(\bar{w},\rho)$$

for all $0 \le h \le h_0$, if ρ is taken sufficiently small. Notice that $\partial\mathcal{G}(w)/\partial w$ is then close to $\partial\mathcal{G}(\bar{w})/\partial w$, and this matrix is invertible. Define $\eta > 0$ by $\eta := \beta|F_h(\bar{w}_h)|$. Then (31) implies $\beta\,\gamma\,\eta/2 \le 1$ for all $0 < h < h_0$, if h_0 is sufficiently small. Proceeding as in [21], the Mysovskij theorem, cf. Ortega and Rheinboldt [22], p. 412, ensures that the Newton method starting at $w_0 := \bar{w}_h$ generates a solution w of (30) in the ball $cl\,B(\bar{w}_h, c_0\,\eta)$, where c_0 is a certain constant. It follows from our construction that

$$G_i(v_h) = G_{h,i}(\bar{u}_h) \begin{cases} = 0, & i = 1,\dots,k, \\ \le 0, & i = k+1,\dots,r. \end{cases}$$

Moreover, if h is small, $G_i(v_h) < 0$ holds for $r < i \le \ell$. Therefore, we have that $G(v_h) \le_K 0$ and $v_h \in U_{ad}$. From $w \in cl\,B(\bar{w}_h, c_0\,\eta)$ it follows $|w - \bar{w}_h| \le c_0\,\eta \le c\,\alpha(h)$, hence also $|v_h - \bar{u}_h| \le c\,\alpha(h)$.

Lemma 4. *If $\rho > 0$ is taken sufficiently small and $h \in (0, h_0(\rho))$, then all solutions $\bar{u}_h$ of the auxiliary problem $(N_{h,\rho})$ belong to $B(\bar{u},\rho)$. Therefore, they are also locally optimal for the problem (N_h).*

Proof. First, we compare the solution $\bar{u}_h$ of $(N_{h,\rho})$ defined in Lemma 2 with u_h that is admissible for $(N_{h,\rho})$ and approximates $\bar{u}$ with the order $\alpha(h)$. We get

$$f_h(\bar{u}_h) \le f_h(u_h) \le |f_h(u_h) - f_h(\bar{u})| + |f_h(\bar{u}) - f(\bar{u})| + f(\bar{u}).$$

By

$$|f_h(\bar{u}) - f(\bar{u})| + |u_h - \bar{u}| + |f_h(\bar{u}_h) - f(\bar{u}_h)| \le c\,\alpha(h)$$

and by the uniform Lipschitz property of f_h, we find

$$f(\bar{u}_h) \le f(\bar{u}) + c_1\,\alpha(h). \tag{32}$$

Next, we compare $\bar{u}$ with v_h taken from Lemma 3. The assumed second-order sufficient optimality condition implies a quadratic growth condition. Inserting v_h in this condition, we obtain for small h

$$f(v_h) \ge f(\bar{u}) + \omega\,|\bar{u} - v_h|^2.$$

From $|\bar{u}_h - v_h| \le c\,\alpha(h)$ we deduce

$$f(\bar{u}_h) + c_2\alpha(h) \ge f(\bar{u}) + \omega\,|\bar{u} - \bar{u}_h|^2. \tag{33}$$

Combining the inequalities (32)–(33), it follows that

$$f(\bar{u}) + c_1\,\alpha(h) \ge f(\bar{u}) + \omega\,|\bar{u} - \bar{u}_h|^2 - c_2\alpha(h)$$

and hence we obtain the stated *auxiliary error estimate*

$$|\bar{u} - \bar{u}_h| \leq c\,\sqrt{\alpha(h)}. \tag{34}$$

For all sufficiently small h, this estimate implies $|\bar{u} - \bar{u}_h| < \rho$ so that $\bar{u}_h$ does not touch the boundary of $B(\bar{u}, \rho)$. In view of this, $\bar{u}_h$ is locally optimal for (N_h).

The error estimate (34) is not optimal. We can get rid of the square root. Moreover, we are able to show the stated local uniqueness of $\bar{u}_h$, i.e. uniqueness of local optima of (N_h) in a neighborhood of $\bar{u}$. Both tasks can be accomplished by the stability theory for optimality systems written as *generalized equations.* This would go beyond the scope of this paper and we refer the reader to the detailed presentation in [19], where we complete the proof of the optimal error estimate stated in the theorem. Moreover, we mention the recent monography [23], where the theory of generalized equations and associated applications are discussed extensively. The same estimate can also be shown for the associated Lagrange multipliers.

References

1. Tröltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods and Applications. American Math. Society, book series Graduate Studies in Mathematics, to appear 2010
2. Casas, E., Tröltzsch, F.: Error estimates for linear-quadratic elliptic control problems. In Barbu et al., V., ed.: Analysis and Optimization of Differential Systems, Boston, Kluwer Academic Publishers (2003) 89–100
3. Falk, F.: Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl. **44** (1973) 28–47
4. Geveci, T.: On the approximation of the solution of an optimal control problem problem governed by an elliptic equation. R.A.I.R.O. Analyse numérique/ Numerical Analysis **13** (1979) 313–328
5. Dontchev, A.L., Hager, W.W., Poore, A.B., Yang, B.: Optimality, stability, and convergence in nonlinear control. Applied Math. and Optimization **31** (1995) 297–326
6. Arada, N., Casas, E., Tröltzsch, F.: Error estimates for the numerical approximation of a semilinear elliptic control problem. Computational Optimization and Applications **23** (2002) 201–229
7. Casas, E., Mateos, M., Tröltzsch, F.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Computational Optimization and Applications **31** (2005) 193–220
8. Casas, E., Mateos, M.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Computational Optimization and Applications **39** (2008) 265–295
9. Casas, E., Raymond, J.P.: Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. SIAM J. Control and Optimization **45** (2006) 1586–1611
10. Vexler, B.: Finite element approximation of elliptic Dirichlet optimal control problems. Numer. Funct. Anal. Optim. **28** (2007) 957–973

11. Deckelnick, K., Günther, A., Hinze, M.: Finite element approximation of Dirichlet boundary control for elliptic PDEs on two- and three-dimensional curved domains. Preprint SPP1253-08-05, DFG-Priority Programme 1253 "Optimization with Partial Differential Equations" (2008)
12. Neitzel, I., Prüfert, U., Slawig, T.: Strategies for time-dependent PDE control with inequality constraints using an integrated modeling and simulation environment. Numerical Algorithms **50**(3) (2009) 241–269
13. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. J. Computational Optimization and Applications **30** (2005) 45–63
14. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control and Optimization **43** (2004) 970–985
15. Casas, E.: Error estimates for the numerical approximation of semilinear elliptic control problems with finitely many state constraints. ESAIM, Control, Optimisation and Calculus of Variations **8** (2002) 345–374
16. Casas, E., Mateos, M.: Uniform convergence of the FEM. Applications to state constrained control problems. J. of Computational and Applied Mathematics **21** (2002) 67–100
17. Meyer, C.: Error estimates for the finite element approximation of an elliptic control problem with pointwise constraints on the state and the control. Preprint 1159, WIAS Berlin (2006)
18. Deckelnick, K., Hinze, M.: Numerical analysis of a control and state constrained elliptic control problem with piecewise constant control approximations. In Kunisch, K., Of, G., Steinbach, O., eds.: Numerical Mathematics and Advanced Applications. Proceedings of ENUMATH 2007, the 7th European Conference on Numerical Mathematics and Advanced Applications, Graz, Austria, September 2007, Springer (2007) 597–604
19. Merino, P., Tröltzsch, F., Vexler, B.: Error estimates for the finite element approximation of a semilinear elliptic control problem with state constraints and finite dimensional control space. Accepted for publication by ESAIM, Control, Optimisation and Calculus of Variations
20. Robinson, S.M.: Stability theory for systems of inequalities, part ii: differentiable nonlinear systems. SIAM J. Numer. Analysis **13** (1976) 497–513
21. Allgower, E.L., Böhmer, K., Potra, F.A., Rheinboldt, W.C.: A mesh-independence principle for operator equations and their discretizations. SIAM Journal on Numerical Analysis **23** (1986) 160–169
22. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables. SIAM Publ., Philadelphia (2000)
23. Dontchev, A. L., Rockafellar, R. T.: Implicit functions and solution mappings. A view from variational analysis. Springer Monographs in Mathematics. Springer, New York (2009)

This article was processed using the LaTeX macro package with LLNCS style