Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARTIN WEISER ANTON SCHIELA

Function space interior point methods for PDE constrained optimization

ZIB-Report 04-27 (June 2004)

Function space interior point methods for PDE constrained optimization*

Martin Weiser Anton Schiela

June 24, 2004

Abstract

A primal-dual interior point method for optimal control problems with PDE constraints is considered. The algorithm is directly applied to the infinite dimensional problem. Existence and convergence of the central path are analyzed. Numerical results from an inexact continuation method applied to a model problem are shown.

AMS MSC 2000: 49M15, 90C48, 90C51

Keywords: interior point methods in function space, optimal control, complementarity functions

1 Introduction

This paper extends [6] from ODE constrained optimal control problems to PDE constraints and adds numerical examples. For a detailed introduction we refer to [6, 7] and the references therein. In order to ease the presentation, we will restrict the discussion to a simple model problem. The extension to more general, possibly nonlinear, elliptic PDE constraints is straightforward along the lines of [6]. On the Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, we consider the optimal control problem

$$\min J(u,y) = \frac{1}{2}\|y-\tilde{y}\|^2_{L_2} + \frac{\alpha}{2}\|u\|^2_{L_2} \quad \text{subject to} \quad \begin{aligned} c(u,y) &= 0 \quad \text{a.e.} \\ g(u) &\geq 0 \quad \text{a.e.} \end{aligned} \tag{1}$$

with $x = (u,y) \in L_\infty(\Omega) \times (H^1_0(\Omega) \cap L_\infty(\Omega))$, $c(x) = \Delta y + u$, $g(u) = (u - \underline{u}, \overline{u} - u)^T$, $\tilde{y}, \underline{u}, \overline{u} \in L_\infty$, and $\alpha > 0$. For the whole paper we will simplify the notation by omitting Ω from the function spaces.

*Supported by the DFG Research Center "Mathematics for key technologies" (FZT 86) in Berlin.

The first order necessary conditions for (1) state the existence of Lagrange multipliers λ and η, such that

$$J'(u,y) - c'(u,y)^*\lambda - g'(u)^*\eta = 0 \tag{2}$$
$$c(u,y) = 0 \tag{3}$$
$$g(u) \geq 0\,, \quad \eta \geq 0\,, \quad \langle \eta, g(u)\rangle = 0 \tag{4}$$

holds for the solution point (u,y). The function space interior point method discussed here replaces the complementarity condition (4) by a pointwise application of the Fischer-Burmeister function $\psi(a,b;\mu) = a + b - \sqrt{a^2+b^2+2\mu}$:

$$\Psi(g(u),\eta;\mu) = 0\,. \tag{5}$$

As will be shown lateron, performing a homotopy $\mu \to 0$ leads to a Kuhn-Tucker point satisfying the first order necessary conditions. For (5) to be continuously differentiable we have to *assume* $\eta \in L_\infty$ and $\lambda \in H_0^1 \cap L_\infty$.

We define the Lagrangian as $L(u,y,\lambda,\eta) = J(u,y) - \langle \lambda, c(u,y)\rangle - \langle \eta, g(u)\rangle$ and the homotopy in terms of

$$F(u,y,\lambda,\eta;\mu) = \begin{bmatrix} \partial_x L(x,\lambda,\eta) \\ -c(x) \\ \Psi(\eta, g(x);\mu) \end{bmatrix} = \begin{bmatrix} (y-\tilde{y}) - \Delta\lambda \\ \alpha u - \lambda - G\eta \\ -\Delta y - u \\ \psi(g(u),\eta;\mu) \end{bmatrix}\,. \tag{6}$$

2 The central path

We adapt the ODE related Theorems 3.2, 3.8, and 3.9 to the current linear PDE setting. The main differences are the choice of appropriate function spaces, the way of establishing the smoothing property of the state equation solution operator Δ^{-1} from $L_2 \to L_\infty$, and specializing the assumptions to the linear case.

Theorem 2.1. *Define $Y = H_0^1 \cap L_\infty$, $V = Y \times L_\infty \times Y \times L_\infty$, $R = \Delta(Y) \subset H^{-1}$ as the image of the Laplace operator applied to Y, and $Z = R \times L_\infty \times R \times L_\infty$. Then the complementarity formulation* (6) *is a continuously differentiable mapping from $V \times \mathbb{R}_+$ to Z which satisfies the Lipschitz condition*

$$\|\partial_v F(v+\delta v;\mu) - \partial_v F(v;\mu)\|_{V\to Z} \leq c(1+\mu^{-1/2})\,\|\delta v\|_V\,. \tag{7}$$

Proof. The image spaces of the adjoint equation w.r.t. u, the state equation and the complementarity equation are immediately clear. As for the adjoint equation w.r.t. y we first notice that $-\Delta y = f, y|_{\partial\Omega} = 0$ with $f \in L_\infty$ implies $y \in L_\infty$ (see [3, Thm. 8.16]). Thus we infer that $L_\infty \subset R$ and hence $y - \tilde{y} \in R$.

As for the Lipschitz condition (7), only the complementarity function Ψ contributes to the difference. The particular value of the Lipschitz constant is due to the Fischer-Burmeister function, see Theorem 3.2 in [6] for details. □

Next we establish bounded invertibilty of the derivative. Therefore it is necessary to introduce a splitting of the domain into nearly active and nearly inactive regions.

Definition 2.2. *For some $\rho > 0$ and functions $u, \eta \in L_\infty$, define the characteristic function $\chi^A = \chi^A(\xi; u, \eta)$ of the nearly active set vector Ω^A componentwise as*

$$\chi_i^A(\xi) = \begin{cases} 1, & \tilde{g}_i(u(\xi)) \le \rho\eta_i(\xi) \\ 0, & \textit{otherwise.} \end{cases}$$

The corresponding characteristic function χ^I of the nearly inactive set vector Ω^I is defined as $\mathbf{1} - \chi^A$, where $\mathbf{1} \in L_\infty$ is the constant function with value 1.

Note that pointwise multiplication with χ^A defines an orthogonal projector onto the corresponding L_∞ space over the nearly active set vector Ω^A.

Theorem 2.3. *Suppose there is a constant $\beta > 0$, such that $\overline{u} - \underline{u} \ge \beta$ almost everywhere. Then there is a constant $c > 0$ such that $\partial_v F(v; \mu)$ has an inverse which is bounded uniformly in the neighborhood $U(c)$ of the central path:*

$$U(c) = \{(v, \mu) \in V \times \mathbb{R} : \|v - v(\mu)\|_V \le c\sqrt{\mu}, 0 \le \mu \le \beta^2/(4\rho)\}$$

Proof. The proof is completely analogous to the one of Theorem 3.8 in [6]. Note that the essential assumptions — controllability of the state equation, strengthened Legendre-Clebsch condition, and positive definiteness of the Hessian on the nullspace of the state equation — are trivially satisfied in the current setting. On the central path $v(\mu)$ we consider the system

$$\partial_v F(v(\mu); \mu)\delta v = r$$

and derive a bound on $\delta v = (\delta u, \delta y, \delta\lambda, \delta\eta)^T$ in terms of $r = (a, b, c, d)^T$. Block elimination of δy and $\delta\lambda$ leads to

$$\begin{bmatrix} \alpha I + \Delta^{-2} & -G^* \\ \partial_g \Psi G & \partial_\eta \Psi \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\eta \end{bmatrix} = \begin{bmatrix} a - \Delta^{-1}(b + \Delta^{-1}c) \\ d \end{bmatrix}$$

with $G = g'(u)$. Multiplication by $\partial_g \Psi^{-1}$ and elimination of only the *inactive* part of η allows to use bounds on the derivatives of Ψ which are independent of μ. The resulting system reads

$$\begin{bmatrix} \alpha I + \Delta^{-2} + G^* D^I G & G^*\chi^A \\ \chi^A G & -D^A \end{bmatrix} \begin{bmatrix} \delta u \\ \chi^A \delta\eta \end{bmatrix} = \begin{bmatrix} a - \Delta^{-1}(b + \Delta^{-1}c) + G^*\chi^I \partial_\eta \Psi^{-1} d \\ \chi^A \partial_g \Psi^{-1} d \end{bmatrix},$$

where both $D^I = \chi^I \partial_\eta \Psi^{-1} \partial_g \Psi = \chi^I \frac{\eta}{g(u)}$ and $D^A = \chi^A \partial_g \Psi^{-1} \partial_\eta \Psi = \chi^A \frac{g(u)}{\eta}$ are nonnegative and bounded independently of μ. Note that due to $4\rho\mu < \beta^2$ in each point ξ at most one component of the inequality constraints is active, such that the following inf-sup-condition is satisfied:

$$\inf_{\xi \in L_2} \sup_{u \in L_2} \frac{\langle \chi^A \xi, Gu \rangle}{\|\chi^A \xi\|_{L_2} \, \|u\|_{L_2}} \geq \hat{\beta} > 0$$

Using the saddle point lemma from [1], an L_2-bound on δu and $\chi^A \delta\eta$ can be obtained. Using the smoothing property $\Delta^{-2} : L_2 \to L_\infty$ (see [3, Thm. 8.16]) to move $\Delta^{-2}\delta u$ to the right hand side and pointwise application of the saddle point lemma then provides an L_∞-bound for δu and $\chi^A \delta\eta$. Tracing the elimination chain back finally yields a constant γ independent of μ such that $\|\partial_v F(v(\mu), \mu)^{-1}\|_{Z \to V} \leq \gamma$.

Using Theorem 2.1 we estimate for $v \neq v(\mu)$

$$\begin{aligned}
\|\partial_v F(v,\mu)^{-1}\| &\leq \|\partial_v F(v,\mu)^{-1}(\partial_v F(v(\mu),\mu) - \partial_v F(v,\mu))\partial_v F(v(\mu),\mu)^{-1}\| \\
&\quad + \|\partial_v F(v(\mu),\mu)\| \\
&\leq \|\partial_v F(v,\mu)^{-1}\| \frac{C}{\sqrt{\mu}} \|v - v(\mu)\| \|\partial_v F(v(\mu),\mu)^{-1}\| + \|\partial_v F(v(\mu),\mu)\|
\end{aligned}$$

and hence

$$\|\partial_v F(v,\mu)^{-1}\| \leq \frac{\|\partial_v F(v(\mu),\mu)\|}{1 - \frac{C}{\sqrt{\mu}}\|v - v(\mu)\| \|\partial_v F(v(\mu),\mu)^{-1}\|} \leq \frac{C_1}{1 - \frac{C}{\sqrt{\mu}} c\sqrt{\mu} C_2} \leq C_3$$

for sufficiently small $c > 0$ independently of μ. □

The fact that the inverse of $\partial_v F$ can be bounded independently of μ limits the length of the central path and thus ensures convergence.

Theorem 2.4. *Suppose the assumptions of Theorem 2.3 hold. Then the central path* $v(\mu)$ *exists for all* $0 < \mu < \beta^2/(4\rho)$ *and converges to a Kuhn-Tucker point* $v(0)$*:*

$$\|v(\mu) - v(0)\|_V \leq \text{const}\,\sqrt{\mu}$$

Proof. Via an implicit function theorem, the bounds given by Theorems 2.1 and 2.3 provide local existence of the central path on the interval $[\sigma\mu, \mu/\sigma]$ around μ with a constant $\sigma < 1$. Thus the central path $v(\mu)$ can be continued up to $\mu > 0$. Its derivative $v'(\mu) = -\partial_v F(v;\mu)^{-1}\partial_\mu F(v;\mu)$ is bounded by $c\mu^{-1/2}$ for some constant c. Integrating the derivative gives a bound of $c\sqrt{\mu}$ for the length of the path, such that the path converges to some limit point $v(0)$. By continuity of F, $v(0)$ satisfies the first order necessary conditions (2)–(4). □

3 Numerical example

In [6], linear convergence of an exact short step pathfollowing method in function space has been shown. The result extends naturally to the current PDE setting. In [5, 7], an *inexact* continuation method has been developed, which can actually be implemented (see Fig. 1). It relies on adaptive mesh refinement when solving linear KKT operator equations in order to meet the accuracy requirements imposed by the inexact Newton corrector. This algorithm has been extended to elliptic PDE constrained optimal control problems and implemented in the FEM code KASKADE using piecewise linear elements for state and Lagrange multiplier and piecewise constant elements for the control.

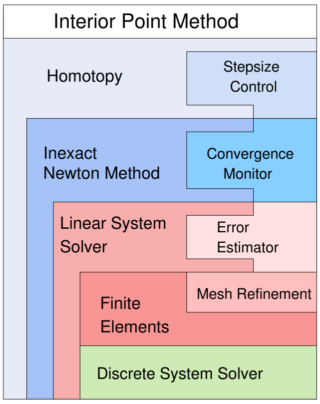

Figure 1: Building blocks of the function space oriented interior point algorithm.

As a numerical example, we present the application of the algorithm to a simplified problem from regional hyperthermia treatment planning. Hyperthermia is a cancer therapy which aims at heating the tumor by microwave radiation and thus making it more susceptible to an accompanying radio- or chemotherapy [2]. The governing PDE is the stationary bio-heat-transfer equation [4]

$$
\begin{aligned}
-\nabla(\kappa \nabla y) + (y - 37)w = u \quad & \text{in } \Omega \\
\beta y + \partial_n y = g \quad & \text{on } \partial\Omega
\end{aligned}
$$

for the temperature y on the relevant part of the human body. The control u, assumed to be freely adjustable within the bounds $0 \leq u \leq u_{\max}$, is the energy absorption of the tissue and is directly related to the amplitude of the time harmonic electric field generated by the microwave generator. The thermal effect of perfusion w with arterial blood of 37° C from different body regions is accounted for by the Helmholtz term. Heat flow through the skin is modeled by the Robin boundary conditions. The aim is to achieve a desired therapeutical temperature distribution

$$
\tilde{y} = \begin{cases} 45 & \text{in } \Omega_t \\ 37 & \text{in } \Omega \backslash \Omega_t \end{cases}
$$

that affects only the tumor tissue $\Omega_t \subset \Omega$ (see Fig. 2). For this example, the regularization parameter α has been set to 10^{-12}.

As can be expected, the solution shown in Fig. 3 just deposits almost all the energy into the tumor region and almost nothing outside. The narrow band of very steep increase in the control is due to the small regularization parameter. Similar results are obtained for the case of Dirichlet boundary conditions, which is completely covered by the theory.

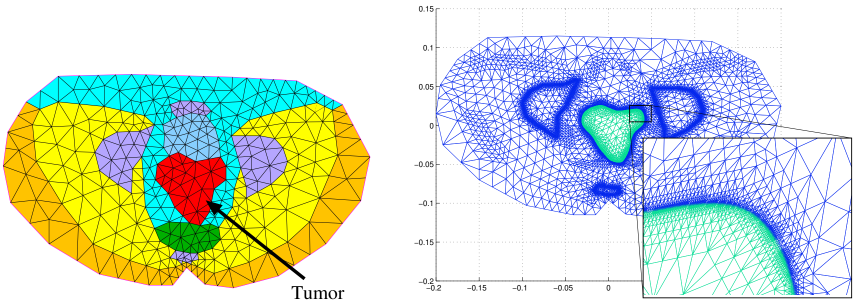

Figure 2: Cross-section Ω of the pelvic region with different tissue types (left) and adaptively refined mesh (right).

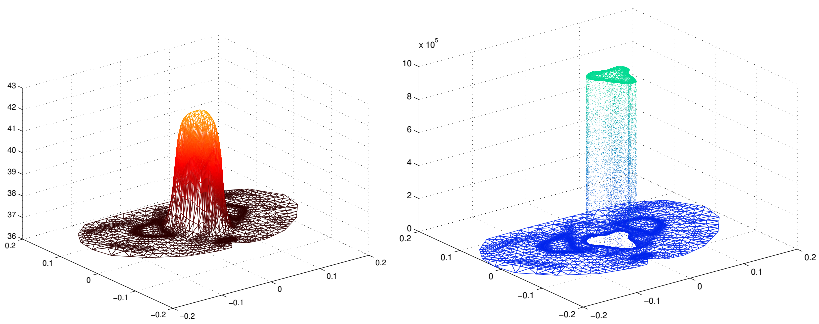

Figure 3: Resulting temperature profile (left) and control (right) for $\mu = 10^{-6}$.

References

[1] D. Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.* **57**, 747–761 (1990).

[2] P. Deuflhard, M. Seebaß, D. Stalling, R. Beck, and H.-C. Hege. Hyperthermia Treatment Planning in Clinical Cancer Therapy: Modelling, Simulation and Visualization. In A. Sydow, editor, *Proc. of the 15th IMACS World Congress 1997 on Scientific Computation: Modelling and Applied Mathematics* **3**, 9–17. Wissenschaft und Technik Verlag (1997).

[3] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* (Springer, 1977).

[4] H.H. Pennes. Analysis of tissue and arterial blood temperatures in the resting human forearm. *J. Appl. Phys.* **1**, 93–122 (1948).

[5] M. Weiser. *Function Space Complementarity Methods for Optimal Control Problems.* PhD thesis, Free University of Berlin (2001).

[6] M. Weiser. Interior point methods in function space. ZIB Report 03-35, Zuse Institute Berlin (2003).

[7] M. Weiser and P. Deuflhard. The central path towards the numerical solution of optimal control problems. ZIB Report 01-12, Zuse Institute Berlin (2001).