An introduction to optimal control of partial differential equations, Part I

Fredi Tröltzsch

Summer School on Applied Analysis 2016

Chemnitz, 19-23 September 2016

Some monographies for PDE control

"The bible":

Lions, J. L.
Optimal Control of Systems Governed by Partial Differential Equations
Springer 1971

Basis for this course:

Tröltzsch, F.
Optimal Control of Partial Differential Equations – Theory, Methods and Applications
American Math. Society 2010

(Deutsche Version, zweite Auflage: Vieweg+Teubner 2009)

More monographies

- Neitaanmäki, P., Tiba, D.
  Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms, and Applications
  Marcel Dekker 1994
- Neitaanmäki, P., Sprekels, J., Tiba, D.
  Optimization of Elliptic Systems
  Springer 2006
- Ito, K., Kunisch, K.
  Lagrange Multiplier Approach to Variational Problems and Applications
  SIAM 2008
- Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.
  Optimization with PDE Constraints
  Springer-Verlag 2009

Introductory examples

General definition of an optimal control problem

Optimal control problem

$$\text{Minimize} \quad J(y,u)$$

subject to $\quad F(y,u) = 0, \qquad u \in U_{ad}, \qquad y \in Y_{ad}$

State equation *Control constraint* *State constraint*

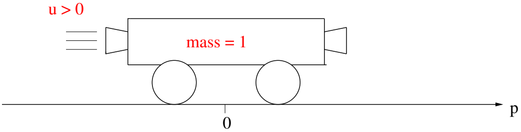
Given quantities

- Linear spaces U, V, Y
- Real-valued function $J : Y \times U \to \mathbb{R}$
- Mapping $F : Y \times U \to V$
- Non-empty subsets $U_{ad} \subset U, \quad Y_{ad} \subset Y$

We assume that, to each *control* u, there exists exactly one *state* y. Then this is really an optimal control problem, otherwise just an optimization problem.

Outline

Example – the rocket car



Move the car in shortest time from initial position p_0 and initial velocity v_0 to rest in $p = p_T$.

This is a problem of time-optimal control.

$$\min \ T \qquad \sim J(\underbrace{p, v}_{y}, u)$$

$$\left.\begin{aligned} p'(t) &= v(t) \\ v'(t) &= u(t) \\ p(0) &= p_0 \\ v(0) &= v_0 \end{aligned}\right\} \sim F(y, u) = 0$$

$$\left.\begin{aligned} p(T) &= p_T \\ v(T) &= 0 \end{aligned}\right\} \qquad \sim y \in Y_{ad}$$

$$-1 \le u(t) \le 1 \qquad \sim u \in U_{ad}$$

The optimal control is piecewise constant ± 1 with at most one switching point. It is discontinuous! This influences the choice of function spaces.

Here, the choice of function spaces is difficult, since $[0, T]$ changes $\rightarrow$ transformation to the fixed time interval $[0, 1]$ by $\tau := t/T$ with additional "control" T.

Outline

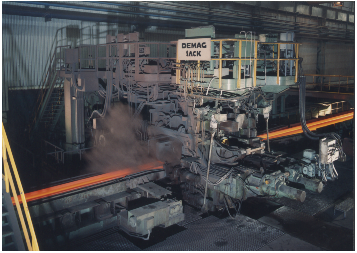
1. Optimal control, definition and introductory example
   - The rocket car
   - Optimal cooling of milled steel profiles
2. Basic ideas of elliptic optimal control problems – Optimality conditions
   - Main tools of nonlinear optimization – finite-dimensional optimal control
   - The simplest elliptic control problem
   - Analysis for a semilinear elliptic control problem

Our applied topic in Chemnitz, 1991

Optimal cooling of milled steel profiles

Cooperation with Mannesmann-Demag-Sack GmbH

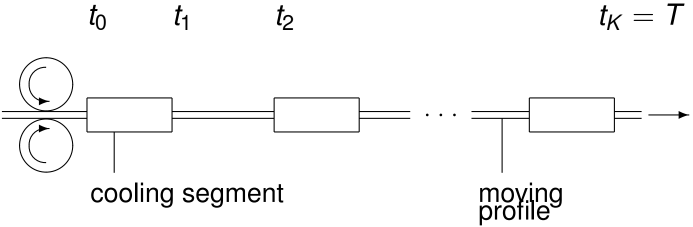
Cooling line



Cooling segment

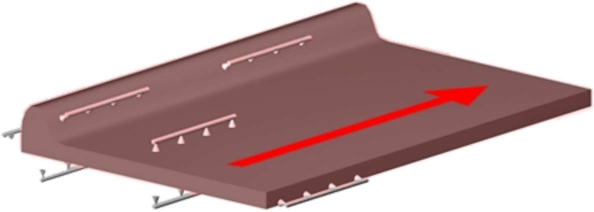
Joint work with R. Lezius, A. Unger, and K. Eppler

Scheme of a cooling line

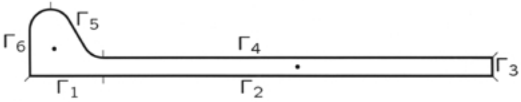


Water cooling segments are followed by air cooling segments

Moving profile and spray nozzles



Ship profile passing a cooling line



Cross section and partitioning of the boundary

Heat equation of the mathematical model

Here, the state is the temperature ϑ of the steel, the control u is the vector of intensities of all spray nozzles

$$y \sim \vartheta, \qquad u \sim (u_{ki}).$$

State equation

Temperature $\vartheta(x,t)$

$$\begin{aligned} c(\vartheta)\rho(\vartheta)\,\vartheta_t &= \operatorname{div}\,(\lambda(\vartheta)\,\operatorname{grad}\,\vartheta) && \text{in } Q, && \text{(heat eq.)} \\ \lambda(\vartheta)\,\partial_n\vartheta &= \sum_{i,k} u_{ki}\,\chi(\Sigma_{ki})\,\alpha(\cdot,\vartheta)(\vartheta_{fl}-\vartheta) && \text{in } \Sigma, && \text{(boundary cond.)} \\ \vartheta(x,0) &= \vartheta_0(x) && \text{in } \Omega && \text{(initial cond.)} \end{aligned}$$

Optimal control problem

Given: Cooling time $T > 0$, weights $a_1, \dots a_N$, finitely many points $P_i,\ R_j,\ Q_k$ in Ω, bounds c_{jk}

$$\min\ J(\vartheta) = \sum_{n=1}^{N} a_n\, \vartheta(P_n, T)$$

subject to the heat equation

$$\begin{aligned} c(\vartheta)\rho(\vartheta)\,\vartheta_t &= \operatorname{div}\,(\lambda(\vartheta)\operatorname{grad}\vartheta) && \text{in } Q,\\ \lambda(\vartheta)\,\partial_n\vartheta &= \sum_{i,k} u_{ki}\,\chi(\Sigma_{ki})\,\alpha(\cdot,\vartheta)(\vartheta_{fl}-\vartheta) && \text{in } \Sigma,\\ \vartheta(x,0) &= \vartheta_0(x) && \text{in } \Omega, \end{aligned}$$

and subject to the constraints on control and state

$$|\vartheta(R_\mu, t) - \vartheta(Q_\nu, t)| \le c_{\mu\nu},$$

$$0 \le u_{ki} \le 1.$$

Main difficulties for the mathematical analysis

$$\min \; J(\vartheta) = \sum_{n=1}^{N} a_n \, \vartheta(P_n, T)$$

subject to the heat equation

$$\begin{aligned} c(\vartheta)\rho(\vartheta)\,\vartheta_t &= \operatorname{div}\,(\lambda(\vartheta)\,\operatorname{grad}\,\vartheta) && \text{in } Q,\\ \lambda(\vartheta)\,\partial_n\vartheta &= \sum_{i,k} u_{ki}\,\chi(\Sigma_{ki})\,\alpha(\cdot,\vartheta)(\vartheta_{fl}-\vartheta) && \text{in } \Sigma,\\ \vartheta(x,0) &= \vartheta_0(x) && \text{in } \Omega, \end{aligned}$$

and subject to the constraints on control and state

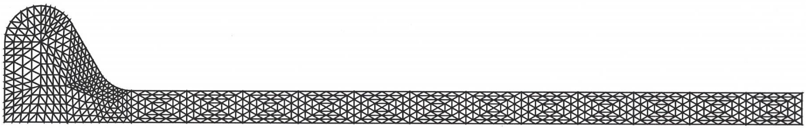
$$|\vartheta(R_\mu, t) - \vartheta(Q_\nu, t)| \le c_{\mu\nu},$$
$$0 \le u_{ki} \le 1.$$

semilinear term, state constraints, quasilinear parts

The mathematical analysis of this problem is still partially open.

Computational example

Thanks to model predictive control techniques, we were able to reduce the computing time from some days to 5 minutes. This was our contribution to real time optimization.



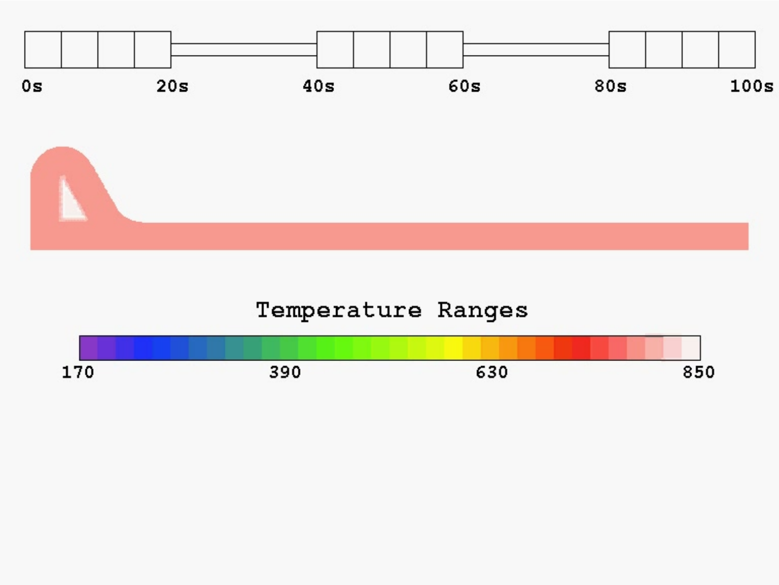
Finite element method – the grid

Numerical example – Ship profile

Initial temperature field

Final temperature fields with and without equilibration

Cooling steel – video (not playable)



Simplified finite-dimensional problems

Outline

1. Optimal control, definition and introductory example
   - The rocket car
   - Optimal cooling of milled steel profiles
2. Basic ideas of elliptic optimal control problems – Optimality conditions
   - Main tools of nonlinear optimization – finite-dimensional optimal control
   - The simplest elliptic control problem
   - Analysis for a semilinear elliptic control problem

A finite-dimensional optimal control problem

Optimal control problem in $\mathbb{R}^n$:

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

- $y_d \in \mathbb{R}^n$, $\nu > 0$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ given,
- Control $u \in \mathbb{R}^m$, state $y \in \mathbb{R}^n$,
- Let A be invertible.

Think of a discretized optimal control problem, where $A\,y = B\,u$ is the discretized PDE.

Elimination of the state:

$$y = A^{-1}\,Bu$$

Existence of an optimal control

Elimination of $y = A^{-1}Bu$ leads to the

Reduced objective function f

$$f(u) := \frac{1}{2}|A^{-1}Bu - y_d|^2 + \frac{\nu}{2}\,|u|^2$$

Theorem

There is a unique vector $\bar{u}$ that minimizes f; $\bar{u}$ is said to be the optimal control for the problem above.

Proof:

$$0 \le \inf_{u \in \mathbb{R}^m} f(u) \le f(0) = \frac{1}{2}\,|y_d|^2.$$

Hence we can concentrate on controls in the compact ball with radius $\nu^{-1}|y_d|$. The function f is continuous. Apply the Weierstraß theorem. Uniqueness follows from strict convexity of f. $\square$

Necessary optimality conditions

We have $y = A^{-1}Bu$; $S := A^{-1}B$ is the *control-to-state operator;* $y = Su$.

$$f(u) = \frac{1}{2}|A^{-1}Bu - y_d|^2 + \frac{\nu}{2}\,|u|^2 = \frac{1}{2}|S\,u - y_d|^2 + \frac{\nu}{2}\,|u|^2.$$

In the optimum, we must have $\nabla f(\bar{u}) = 0$.

Simple computation (reduced gradient)

$$\nabla f(u) = S^\top(Su - y_d) + \nu\,u = B^\top(A^{-1})^\top(y - y_d) + \nu\,u.$$

Avoid the numerical use of A^{-1}

In our control problems, A might by extremely large. Think of a dimension $n = 10^4$ in 2D or $n = 10^6$ in 3D. This prohibits the computation of A^{-1}. Moreover, in the extension to PDEs, we do not have in general an explicit representation of A^{-1} and need some substitute for analysis and numerical methods.

The adjoint state

$$0 = \nabla f(u) = B^\top \underbrace{(A^{-1})^\top (y - y_d)}_{=:p} + \nu\, u = B^\top p + \nu\, u$$

where

$$p := (A^{-1})^\top (y - y_d) = (A^\top)^{-1} (y - y_d).$$

Therefore, p solves $A^\top p = y - y_d$. $\Rightarrow$

Definition – Adjoint equation

The equation

$$A^\top p = y - y_d$$

is the adjoint equation and p is the adjoint state associated with y (or u).

$\Rightarrow$ Optimality system

$$\begin{array}{rcll} A\, y &=& B\, u & \textit{State equation} \\ A^\top p &=& y - y_d & \textit{Adjoint equation} \\ \nu\, u + B^\top p &=& 0 & \textit{Gradient equation.} \end{array}$$

A first numerical idea

Eliminating u by $u = -\nu^{-1} B^\top p$, we obtain the system of linear equations

$$\boxed{\begin{array}{rcl} A\,y & = & -\nu^{-1} B\,B^\top p \\ A^\top p & = & y - y_d. \end{array}}$$

Solve this system to obtain the optimal solution. Our problem is convex, hence the optimality system is sufficient for optimality. The solution of the optimality system is optimal.

Optimal control with control constraints

We now require in addition constraints on the control vector u.

Control constrained problem

$$\min\ J(y,u) := \frac{1}{2}|y - y_d|^2 + \frac{\nu}{2}\,|u|^2,$$
$$A\,y = B\,u, \qquad \alpha \le u \le \beta,$$

the constraints in componentwise sense

$$U_{ad} := \{u \in \mathbb{R}^m : \alpha \le u_i \le \beta \quad \forall i \in \{1,\dots,m\}\}.$$

Eliminating again y, the control problem is equivalent to

$$\min_{u \in U_{ad}} f(u).$$

Clearly, we cannot expect $\nabla f(\bar{u}) = 0$ if $\bar{u}$ is located at ∂U_{ad}!

The variational inequality

Existence of an optimal control $\bar{u}$: Trivial, since U_{ad} is compact.

Assume now that f is a general (not necessarily quadratic) functional.

Theorem

If $\bar{u}$ is optimal and f differentiable at $\bar{u}$, then the variational inequality

$$\nabla f(\bar{u}) \cdot (u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}$$

must be satisfied.

The proof is simple:

$$\begin{aligned} 0 &\leq f(\bar{u} + s\,(u - \bar{u})) - f(\bar{u}) \quad \forall s \in (0, 1) \\ 0 &\leq \lim_{s \downarrow 0} \frac{1}{s}(f(\bar{u} + s\,(u - \bar{u})) - f(\bar{u})) = \nabla f(\bar{u}) \cdot (u - \bar{u}). \qquad \square \end{aligned}$$

Discussion of the variational inequality

We re-write the variational inequality,

$$\nabla f(\bar{u}) \cdot \bar{u} \leq \nabla f(\bar{u}) \cdot u \quad \forall u \in [\alpha, \beta].$$

Important consequence

$$\bar{u}_i = \begin{cases} \alpha & \text{if} \quad \nabla f(\bar{u})_i > 0 \\ \in [\alpha, \beta] & \text{if} \quad \nabla f(\bar{u})_i = 0 \\ \beta & \text{if} \quad \nabla f(\bar{u})_i < 0. \end{cases}$$

Using the adjoint state $\bar{p}$ this means

$$\bar{u}_i = \begin{cases} \alpha & \text{if} \quad (B^\top \bar{p} + \nu \bar{u})_i > 0 \\ \in [\alpha, \beta] & \text{if} \quad (B^\top \bar{p} + \nu \bar{u})_i = 0 \\ \beta & \text{if} \quad (B^\top \bar{p} + \nu \bar{u})_i < 0. \end{cases} \quad \Rightarrow \bar{u}_i = -\nu^{-1}(B^\top \bar{p})_i$$

Use of a Lagrangian function

The Lagrangian function is a very helpful tool for expressing the optimality conditions in an easy way. Recall the optimal control problem:

$$\min J(y,u), \qquad Ay = Bu, \qquad u \in U_{ad}$$

Lagrangian function

$$\mathcal{L}(y,u,p) := J(y,u) - (Ay - Bu\,,\,p).$$

We think of "eliminating" the equality constraint by a Lagrange multiplier vector p and consider the optimality conditions for the problem with y, u decoupled

$$\min_{y\in\mathbb{R}^n,\; u\in U_{ad}} \mathcal{L}(y,u,p).$$

$$\Rightarrow \quad \underbrace{\partial_y \mathcal{L}(\bar y,\bar u,\bar p) = 0,}_{\textit{adjoint equation}} \qquad \underbrace{\partial_u \mathcal{L}(\bar y,\bar u,\bar p)(u-\bar u) \ge 0 \quad \forall u \in U_{ad}}_{\textit{variational inequality}}$$

Notice: $\nabla f(u) = B^\top p + \nu u = \partial_u \mathcal{L}(\bar y,\bar u,\bar p)$ (reduced gradient = $\partial_u \mathcal{L}$)

Nonlinear finite-dimensional optimal control problem

Let us consider the nonlinear optimal control problem

$$\min J(y,u), \quad F(y,u) = 0, \quad u \in U_{ad}.$$

We assume that $J : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $F : J : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are of class C^1. Assume that the pair $(\bar{y}, \bar{u})$ is (locally) optimal. Assume further

$$\exists\, \partial_y F(\bar{y}, \bar{u})^{-1}.$$

By an application of the implicit function theorem, we can prove the existence of a unique adjoint state $\bar{p} \in \mathbb{R}^n$ such that the variational inequality

$$(\partial_u F(\bar{y},\bar{u})^\top \bar{p} + \partial_u J(\bar{y},\bar{u})\,,\, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}$$

and the adjoint equation

$$\partial_y F(\bar{y},\bar{u})^\top \bar{p} = \partial_y J(\bar{y},\bar{u})$$

are satisfied.

Summary 1

- We have considered finite-dimensional versions of optimal control problems with state equation $Ay = Bu$ that typically might arise from the discretization of optimal control problems.
- Introducing the control-to-state operator $S : u \mapsto y$, $S = A^{-1}B$, a reduced problem was defined. The proof of existence of an optimal control followed easily by the Weierstraß theorem.
- To avoid the use of the matrix A^{-1}, an adjoint equation was introduced, one of the most important ideas in optimal control. By its solution p, the adjoint state, the first-order optimality conditions can be expressed in terms of an optimality system.
- The optimality system can also be derived by a Lagrangian function.
- In this way, the adjoint state p can be interpreted as Lagrangian multiplier to the state equation.

Control of elliptic PDEs

Outline

Outline

1. Optimal control, definition and introductory example
   - The rocket car
   - Optimal cooling of milled steel profiles

2. Basic ideas of elliptic optimal control problems – Optimality conditions
   - Main tools of nonlinear optimization – finite-dimensional optimal control
   - The simplest elliptic control problem
   - Analysis for a semilinear elliptic control problem

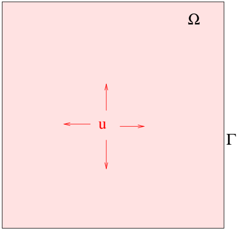
Optimally controlled heat source

"Optimal stationary heat source": Heat a domain Ω by a controlled heat source u to reach the given temperature $y_\Omega \in L^2(\Omega)$; select $\nu > 0$.

$$\min \; J(y,u) := \frac{1}{2}\int_\Omega \left|y(x) - y_\Omega(x)\right|^2 dx + \frac{\nu}{2}\int_\Omega u(x)^2\, dx$$

subject to $y \in H_0^1(\Omega),\; u \in L^2(\Omega),$

$$\begin{aligned} -\Delta y(x) &= u(x) \quad \text{in } \; \Omega \\ y(x) &= 0 \quad\quad\; \text{on } \; \Gamma \end{aligned}$$



Heat source

Linear-quadratic elliptic distributed control problem, J.L. Lions 1968.

Recall: $\Delta y = \dfrac{\partial^2 y}{\partial x_1^2} + \dfrac{\partial^2 y}{\partial x_2^2} + \dfrac{\partial^2 y}{\partial x_3^2}.$

Comparison to the finite-dimensional case

$$\min \ \frac{1}{2}\int_\Omega |y - y_\Omega|^2 \, dx + \frac{\nu}{2}\int_\Omega u^2 \, dx \qquad\qquad \min \frac{1}{2}|y - y_d|^2 + \frac{\nu}{2}|u|^2$$

$$y \in H_0^1(\Omega),\ u \in L^2(\Omega), \qquad\qquad y \in \mathbb{R}^n, \quad u \in \mathbb{R}^m$$

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

Assumptions on Ω

In all what follows, $\Omega \subset \mathbb{R}^N$, $N \geq 2$, is a bounded Lipschitz domain.

Weak solution of the elliptic PDE

We recall the notion of a weak solution and assume that the concept of weak derivatives and the space $H_0^1(\Omega)$ is known.
Multiply the strong form of the PDE by an arbitrary test function $v \in H_0^1(\Omega)$ and integrate,

$$\int_\Omega (-\Delta y(x))v(x)\,dx = \int_\Omega u(x)v(x)\,dx \quad \forall v \in H_0^1(\Omega).$$

Now integrate by parts.

Definition

A function $y \in H_0^1(\Omega)$ is said to be a weak solution of the PDE, if the variational formulation

$$\int_\Omega \nabla y(x) \cdot \nabla v(x)\,dx = \int_\Omega u(x)v(x)\,dx \quad \forall v \in H_0^1(\Omega)$$

is satisfied.

Control-to-state operator S

By the lemma of Lax and Milgram and the Friedrichs inequality, the following result is obtained:

Theorem (Well-posedness of the elliptic PDE)

For all $u \in L^2(\Omega)$, the linear elliptic PDE above has a unique weak solution $y_u \in H_0^1(\Omega)$. There is a constant $c_\Omega > 0$ that does not depend on u such that

$$\|y_u\|_{H_0^1(\Omega)} \leq c_\Omega \, \|u\|_{L^2(\Omega)}.$$

Corollary

The control-to-state mapping $G : u \mapsto y_u$ is linear and continuous from $L^2(\Omega)$ to $H_0^1(\Omega)$.

We consider now G as a mapping with range in $L^2(\Omega)$ and denote this mapping by S.

Existence of an optimal control

We follow our method from the finite-dimensional case and establish a reduced problem. Obviously, we have

$$J(y,u) = \frac{1}{2}\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} = \frac{1}{2}\|Su - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} =: f(u).$$

The reduced functional f is convex and continuous in $L^2(\Omega)$.

Theorem

The elliptic optimal control problem "optimal stationary heat source" has a unique optimal control $\bar{u}$.

Proof: Since $\nu > 0$, we can find all optimal controls in the closed ball

$$B_r(0) = \{u \in L^2(\Omega) : \|u\|_{L^2(\Omega)} \leq r\}$$

for sufficiently large $r > 0$. The space $L^2(\Omega)$ is a separable Hilbert space, the ball $B_r(0)$ is bounded, convex, and closed. Therefore, it is weakly sequentially compact. Notice that $B_r(0)$ is not compact. $\rightarrow$ weak convergence

Continuation of proof

Let $j = \inf_{u \in L^2(\Omega)} f(u)$; select an infimal sequence (u_n) such that

$$f(u_n) \to j, \quad n \to \infty.$$

By weak compactness we can select a weakly convergent subsequence – w.l.o.g (u_n) itself – such that

$$u_n \rightharpoonup \bar{u}, \quad n \to \infty.$$

Any continuous and convex functional is lower semicontinuous, hence

$$j = \lim_{n \to \infty} f(u_n) \geq f(\lim_{n \to \infty} u_n) = f(\bar{u}).$$

Therefore, $\bar{u}$ is optimal. Uniqueness follows again by the strict convexity of f. $\square$

We postpone the necessary optimality conditions to a more general semilinear control problem.

Semilinear elliptic control problem

$$\min \ \frac{1}{2}\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}$$

$$y \in H_0^1(\Omega),\ u \in L^2(\Omega),$$

$$\begin{aligned} -\Delta y + R(y) &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

$$\alpha \le u(x) \le \beta \quad \text{a.e. in } \Omega.$$

- Constant bounds $\alpha < \beta$; $R : \mathbb{R} \to \mathbb{R}$ monotone non-decreasing, differentiable with locally Lipschitz derivative;

  "Reaction term".
- Think of

  $$R(y) = y^3 \ \text{ or } \ R(y) = \exp y.$$

The constraints on u are very helpful, because they imply $u \in L^\infty(\Omega)$.

Definition: $U_{ad} = \{u \in L^2(\Omega) \mid \alpha \le u(x) \le \beta, \ \text{a.e. in } \Omega\}.$

Weak solution

Notice that, for $y \in H_0^1(\Omega)$, the function $R(y)$ is perhaps not integrable, think of $R(y) = \exp y$. Therefore, the definition of a weak solution needs some care.

Definition (Weak solution)

A function $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$ is said to be a weak solution of the semilinear elliptic pde, if

$$\int_\Omega \nabla y \cdot \nabla v \, dx + \int_\Omega R(y) \, v \, dx = \int_\Omega u \, v \, dx \quad \forall v \in H_0^1(\Omega).$$

Well-posedness of the elliptic state equation

Theorem (Existence and uniqueness)

For each $u \in L^p(\Omega)$ with $p > N/2$, $N =$ dim Ω, the semilinear elliptic state equation has exactly one solution $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$. This solution belongs even to $C(\bar{\Omega})$.

$\to$ Frèchet derivative, sinus-counterexample

Theorem (Differentiability)

- *The solution mapping $G : L^p(\Omega) \to H_0^1(\Omega) \cap C(\bar{\Omega})$, $G : u \mapsto y$, is continuously Frèchet differentiable.*
- *The derivative G' is given by $G'(u)\,v =: z$, where $z \in H_0^1(\Omega)$ is the unique solution to the linearized equation*

$$-\Delta z + R'(y)\,z = v \quad \text{in } \Omega, \qquad z = 0 \quad \text{on } \Gamma. \quad \to \text{ confirmation}$$

Reduced problem

By the solution mapping G, we can reformulate our problem: Recall that $y = G(u)$, hence

$$J(y,u) = J(G(u),u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} =: f(u).$$

Our optimal control problem is equivalent to the

Reduced problem

$$(P) \qquad \min_{u \in U_{ad}} f(u) := \frac{1}{2}\|G(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}$$

Theorem (Existence of an optimal control)

The semilinear elliptic control problem has at least one optimal control $\bar{u}$.

Proof: Is more difficult; U_{ad} is weakly compact in $L^p(\Omega)$, have $u_n \rightharpoonup \bar{u}$ in $L^p(\Omega)$. We use the boundedness of (y_n) in $L^\infty(\Omega)$ and the compactness of the embedding $H_0^1(\Omega) \subset L^2(\Omega)$. Then, a subsequence of (y_n) converges almost everywhere.

Derivation of necessary optimality conditions

Since G is Fréchet differentiable, the chain rule ensures the differentiability of f.

Variational inequality

Let $\bar u \in U_{ad}$ be a solution to (P). Then there holds

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

The proof is identical with the one from the finite-dimensional case.

This variational inequality is not yet useful. We have to work a bit.

$$\begin{aligned} 0 \leq f'(\bar u)(u - \bar u) &= \Big(\underbrace{G(\bar u)}_{\bar y} - y_\Omega \,,\, \underbrace{G'(\bar u)(u - \bar u)}_{z}\Big)_{L^2(\Omega)} + \nu\big(\bar u \,,\, u - \bar u\big)_{L^2(\Omega)} \\ &= \big(\bar y - y_\Omega \,,\, z\big)_{L^2(\Omega)} + \nu\,\big(\bar u \,,\, u - \bar u\big)_{L^2(\Omega)}, \end{aligned}$$

where $z \in H_0^1(\Omega)$ is the unique weak solution to the linearized equation

$$-\Delta z + R'(\bar y)\, z = u - \bar u.$$

Adjoint equation

Adjoint equation

The equation

$$\begin{aligned} -\Delta p + R'(\bar{y})\, p &= \bar{y} - y_\Omega \quad \text{in} \;\; \Omega \\ p &= 0 \qquad\quad\;\; \text{on} \;\; \Gamma \end{aligned}$$

is said to be the adjoint equation. Its unique solution $\bar{p}$ is called adjoint state associated with $\bar{y}$.

Theorem (Variational inequality)

If $\bar{u}$ is optimal for (P), then there exists a unique adjoint state $\bar{p} \in H_0^1(\Omega)$ such that the variational inequality

$$\int_\Omega (\bar{p}(x) + \nu\, \bar{u}(x))(u(x) - \bar{u}(x))dx \geq 0 \quad \forall u \in U_{ad}$$

is fulfilled.

Proof of the variational inequality

We recall the variational inequality above,

$$(\bar{y} - y_\Omega \, , \, z)_{L^2(\Omega)} + \nu \, (\bar{u} \, , \, u - \bar{u})_{L^2(\Omega)} \geq 0.$$

Next, we write down the variational equations for z and $\bar{p}$ with test functions $\bar{p}$ and z, respectively,

$$\begin{aligned} \int_\Omega \nabla z \cdot \nabla \bar{p} + R'(\bar{y}) z \, \bar{p} \, dx &= \int_\Omega (u - \bar{u}) \, \bar{p} \, dx \\ \int_\Omega \nabla \bar{p} \cdot \nabla z + R'(\bar{y}) \bar{p} \, z \, dx &= \int_\Omega (\bar{y} - y_\Omega) \, z \, dx. \end{aligned}$$

The left-hand sides are equal, hence the right-hand sides are equal as well. This implies

$$(\bar{p} \, , \, u - \bar{u})_{L^2(\Omega)} + \nu \, (\bar{u} \, , \, u - \bar{u})_{L^2(\Omega)} \geq 0.$$

Now simplify the two inner products. □

Application of a formal Lagrangian principle

For later use with semilinear parabolic systems, let us verify that the optimality conditions can be derived by some Lagrangian technique. It is a bit formal, since we do not discuss any function spaces. We just assume that a Lagrange multiplier p exists as a function with sufficient smoothness.

Definition (Lagrangian function)

Let $p \in H_0^1(\Omega)$, $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$.

$$\begin{aligned} \mathcal{L}(y,u,p) &= J(y,u) - \int_\Omega (-\Delta y + R(y) - u)\, p\, dx \\ &:= J(y,u) - \int_\Omega \nabla y \cdot \nabla p + R(y)\, p - u\, p\, dx \end{aligned}$$

Then one has that $\partial_y \mathcal{L}(\bar{y},\bar{u},\bar{p}) = 0$ leads to the adjoint equation while $\partial_u \mathcal{L}(\bar{y},\bar{u},\bar{p})(u-\bar{u}) \geq 0 \quad \forall u \in U_{ad}$ delivers the variational inequality (cf. the exposition on the parabolic case).P

Projection formula

With the adjoint state, the variational inequality reads

$$\int_\Omega (\bar{p}(x) + \nu\, \bar{u}(x))(u(x) - \bar{u}(x))\, dx \geq 0 \quad \forall u \in U_{ad}.$$

This can be essentially simplified by a pointwise discussion!

For a.a. $x \in \Omega$, $\bar{u}(x)$ solves $\rightarrow$ sinus-example

$$\min_{\alpha \leq u \leq \beta} \big(\bar{p}(x) + \nu \bar{u}(x)\big)\, u \qquad \text{for a.a. } x \in \Omega.$$

This is equivalent with the property that $u := \bar{u}(x)$ solves for almost all $x \in \Omega$ the simple quadratic optimization problem in $\mathbb{R}$,

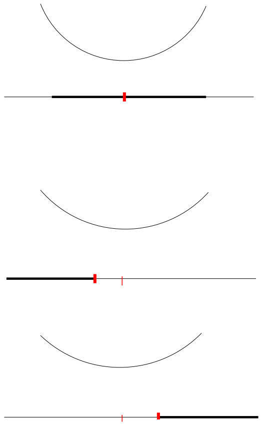
$$\min_{\alpha \leq u \leq \beta} \ \{\bar{p}(x)\, u + \frac{\nu}{2} u^2\}.$$

There is a simple explicit solution formula for this!

Projection formula

Quadratic optimization in $\mathbb{R}$

$$\min\{\frac{\nu}{2}u^2 + \bar{p}(x)\,u\}, \quad \alpha \le u \le \beta.$$



Unconstrained, the optimal value is

$$\bar{u}(x) = -\nu^{-1}\,\bar{p}(x).$$

With constraints, we have

$$\bar{u}(x) = \begin{cases} \alpha & \text{if } -\nu^{-1}\,\bar{p}(x) < \alpha \\ -\nu^{-1}\,\bar{p}(x) & \text{if } \alpha \le -\nu^{-1}\,\bar{p}(x) \le \beta \\ \beta & \text{if } -\nu^{-1}\,\bar{p}(x) > \beta \end{cases}$$

Projection formula

$$\bar{u}(x) = \mathbb{P}_{[\alpha,\beta]}\{-\nu^{-1}\bar{p}(x)\}$$

Regularity of the optimal control

A first conclusion of the projection formula is higher regularity of the optimal control.

Theorem

If $\nu > 0$, then the optimal control $\bar{u}$ belongs to $H^1(\Omega)$. If $N \leq 3$ and Ω is either convex or has a $C^{1,1}$-boundary, then $\bar{u}$ is contained in $C(\bar{\Omega})$.

Proof: The adjoint state $\bar{p}$ belongs to $H_0^1(\Omega)$. By a result from Stampacchia and Kinderlehrer, the functions $\max\{p, q\}$ and $\min\{p, q\}$ are elements of $H^1(\Omega)$ provided that $p \in H^1(\Omega)$ and $q \in H^1(\Omega)$.
The real function $v \mapsto \mathbb{P}_{[\alpha,\beta]}(v) : \mathbb{R} \to [\alpha, \beta]$ is a composition of the max- and min-function. This confirms the regularity result.

The result on continuity follows from maximal regularity of elliptic equations. The right-hand side of the adjoint equation is $\bar{y} - \bar{\Omega} \subset L^2(\Omega)$. If Ω is convex, then the solution p has the higher regularity $H^2(\Omega)$. For $N \leq 3$, we have $H^2(\Omega) \subset C(\bar{\Omega})$ (more precisely, $u \in H^2(\Omega)$ is equivalent to a function of $C(\bar{\Omega})$). The same follows for domains with $C^{1,1}$-boundary. □

The optimality system

Substituting this projection formula for u in the state equation, the optimal state $\bar{y}$ and the associated adjoint state $\bar{p}$ can be obtained by solving the

Nonsmooth optimality system

$$
\begin{aligned}
-\Delta y(x) + R(y(x)) &= \mathbb{P}_{[\alpha,\beta]}\{-\nu^{-1}p(x)\} && \text{in } \Omega \\
y(x) &= 0 && \text{on } \Gamma \\
-\Delta p(x) + R'(y(x))\,p(x) &= y(x) - y_\Omega(x) && \text{in } \Omega \\
p(x) &= 0 && \text{on } \Gamma.
\end{aligned}
$$

Brute force method: COMSOL Multiphysics is able to directly solve such coupled systems in 2D. In particular, a build-in mesh refinement deals with the two corners of $u \mapsto \mathbb{P}_{[\alpha,\beta]}(u)$.

Semismooth Newton method: This method is justified by a detailed convergence analysis in function spaces. It is locally superlinearly convergent.

The semismooth Newton method

The derivative of $u \mapsto \mathbb{P}_{[\alpha,\beta]}(u)$ is

$$\mathbb{P}'_{[\alpha,\beta]}(u) = \begin{cases} 1 & \text{if } u \in (\alpha,\beta) \\ 0 & \text{if } u \notin (\alpha,\beta). \end{cases}$$

Nondifferentiability only in $u = \alpha$ and $u = \beta$. The numerical method will most likely not see these points :-)

Rule of thumb:

Apply the Newton method formally to the non-smooth optimality system and use $\mathbb{P}'_{[\alpha,\beta]}$ as defined above. Take as $\mathbb{P}'_{[\alpha,\beta]}(u)$ any value in the points $u = \alpha,\ \beta$, say zero.

Convergence analysis, semismooth functions, Newton differentiability $\quad \rightarrow$

Ito, K., Kunisch, K.
Lagrange Multiplier Approach to Variational Problems and Applications
SIAM 2008

Second-order sufficient optimality conditions

So far, we only have considered first-order necessary optimality conditions. In the nonlinear (hence in general nonconvex) case, they are not sufficient for (local) optimality. Second order sufficient optimality conditions can be assumed to guarantee local optimality; important also for numerical approximations.

This is a very interesting but complicated issue. Many publications have been devoted to this issue since our 1996. Here is a new references, a survey paper:

Casas, E., Trö, F.
Second order optimality conditions and their role in PDE control
Jahresbericht der Deutschen Mathematiker-Vereinigung 117, 2015, 3-44.

Extensions

- Boundary control problems
- Quasilinear elliptic equation
- Pointwise constraints on the state
- More general objective functionals
- Second-order sufficient optimality conditions
- Numerical discretization and error analysis
- Numerical iteration methods and their convergence
- Sparse optimal control

Summary 2

- We considered a particular class of optimal control problems for linear and semilinear elliptic equations.
- In the linear case, the standard control space is $L^2(\Omega)$, the state space is $H_0^1(\Omega)$.
- For semilinear equations, we need continuity and differentiability properties of the nonlinear mapping $y(\cdot) \mapsto R(y(\cdot))$ that might not be fulfilled in $H_0^1(\Omega)$. Therefore, the control is taken from $L^p(\Omega)$ with sufficiently large p and y is considered in $H_0^1(\Omega) \cap C(\bar{\Omega})$.
- In both cases, an adjoint elliptic equation is introduced to set up necessary optimality conditions.
- A projection formula was derived that expresses the optimal control in terms of the associated adjoint state.
- This projection formula can be used to set up a non-smooth optimality system that can be used for computing the optimal control by a semi-smooth Newton method. Moreover, it shows that the optimal control is a function of $H^1(\Omega)$.