# Tutorial

# Introduction to the theory and numerical solution of PDE constrained optimization problems


Matthias Heinkenschloss
Rice University

Fredi Tröltzsch
Technische Universität Berlin


**Workshop on PDE Constrained Optimization, Tomar 2005**

1. Examples of elliptic control problems
2. Existence of optimal solutions
3. Optimality conditions
4. Discretization strategies
5. Optimization algorithms

# 1. Examples of elliptic control problems

## 1.1 Convex problems

## 1.2 Nonconvex problems

## 1.3 Linear elliptic PDEs

### 1.3.1 Spaces

### 1.3.2 Poisson equation

### 1.3.3 Robin boundary condition

## 1.1 Convex problems

**Optimal stationary boundary temperature:** Heating of a body $\Omega$ by a controlled boundary temperature $u$ to reach a target temperature $y_\Omega$.

$$\min \ J(y,u) := \frac{1}{2}\int_\Omega \big(y(x) - y_\Omega(x)\big)^2\,dx + \frac{\alpha}{2}\int_\Gamma u(x)^2\,ds(x)$$

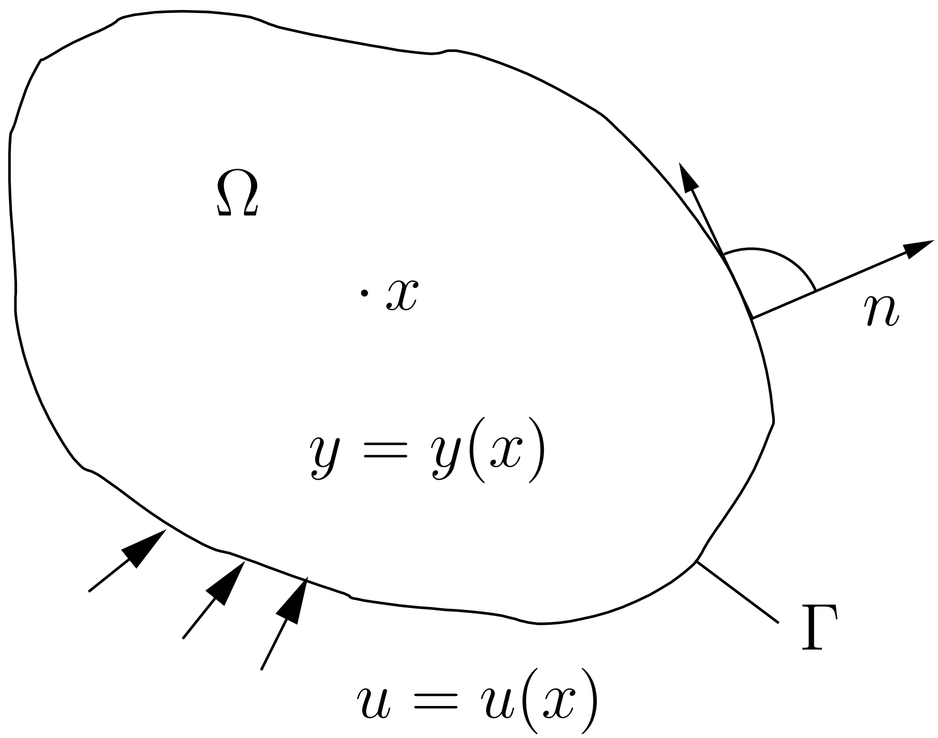
subject to the *state equation* (state $y$)

$$\begin{aligned} -\Delta y &= 0 && \text{in } \Omega \\ \frac{\partial y}{\partial n} &= \sigma\,(u - y) && \text{on } \Gamma \end{aligned}$$

and to the *control-constraints* (control $u$)

$$u_a(x) \le u(x) \le u_b(x) \quad \text{on } \Gamma.$$

This is a linear-quadratic elliptic *boundary control problem*.




*Boundary control*

**Optimal stationary heat source:** Heating of a body $\Omega$ by a controlled heat source $u$ (say electromagnetic induction or microwaves) to reach the target $y_\Omega$.

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

subject to

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

$$u_a(x) \le u(x) \le u_b(x) \quad \text{in } \Omega.$$

**Optimal stationary heat source:** Heating of a body $\Omega$ by a controlled heat source $u$ (say electromagnetic induction or microwaves) to reach the target $y_\Omega$.

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

subject to

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

$$u_a(x) \le u(x) \le u_b(x) \quad \text{in} \;\; \Omega.$$

This is a linear-quadratic elliptic *distributed control problem*.

**Optimal stationary heat source:** Heating of a body $\Omega$ by a controlled heat source $u$ (say electromagnetic induction or microwaves) to reach the target $y_\Omega$.

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

subject to

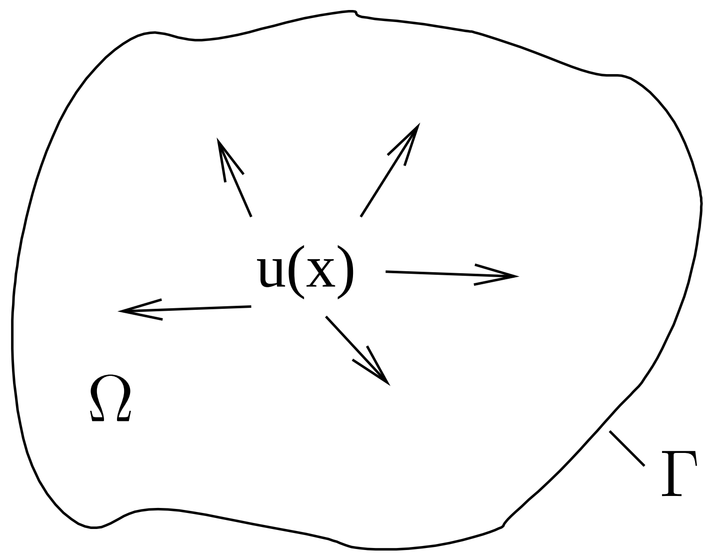
$$\begin{aligned} -\Delta y &= u \quad \text{in} \quad \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

$$u_a(x) \le u(x) \le u_b(x) \quad \text{in} \;\; \Omega.$$

This is a linear-quadratic elliptic *distributed control problem*.

It might be important to include also *pointwise state constraints*

$$y_a \le y(x) \le y_b.$$




*Controlled heat source*

Consider the same objective functional as before, but with *semilinear* state equation:

**Stefan-Boltzmann radiation condition:** Sachs, 1978 (parabolic case)

$$\begin{aligned} -\Delta y &= 0 && \text{in } \Omega \\ \frac{\partial y}{\partial n} &= \sigma\left(u^4 - y^4\right) && \text{on } \Gamma. \end{aligned}$$

Consider the same objective functional as before, but with *semilinear* state equation:

**Stefan-Boltzmann radiation condition:** Sachs, 1978 (parabolic case)

$$\begin{aligned} -\Delta y &= 0 && \text{in } \Omega \\ \frac{\partial y}{\partial n} &= \sigma\,(u^4 - y^4) && \text{on } \Gamma. \end{aligned}$$

**Simplified equation in super conductivity:**

The following Ginzburg-Landau model has been discussed by Ito and Kunisch, 1996:

$$\begin{aligned} -\Delta y - y + y^3 &= u && \text{in } \Omega \\ y &= 0 && \text{on } \Gamma. \end{aligned}$$

Consider the same objective functional as before, but with *semilinear* state equation:

**Stefan-Boltzmann radiation condition:** Sachs, 1978 (parabolic case)

$$\begin{aligned} -\Delta y &= 0 && \text{in } \Omega \\ \frac{\partial y}{\partial n} &= \sigma\,(u^4 - y^4) && \text{on } \Gamma. \end{aligned}$$

**Simplified equation in super conductivity:**

The following Ginzburg-Landau model has been discussed by Ito and Kunisch, 1996:

$$\begin{aligned} -\Delta y - y + y^3 &= u && \text{in } \Omega \\ y &= 0 && \text{on } \Gamma. \end{aligned}$$

Then nonconvex optimal control problems are obtained. Nonconvex problems require different spaces and a different analysis. Associated numerical algorithms are more complex.

**Stationary Navier-Stokes-equations:** The stationary fluid flow in a domain $\Omega$ can be modelled by

$$
\begin{aligned}
-\frac{1}{Re}\,\Delta y + (y\cdot\nabla)\,y + \nabla p &= u \quad \text{in} \quad \Omega \\
y &= 0 \quad \text{on } \Gamma \\
\operatorname{div} y &= 0 \quad \text{in} \quad \Omega,
\end{aligned}
$$

where $y = y(x) \in \mathbb{R}^3$ is the velocity vector of a particle located at $x \in \Omega$, $p = p(x)$ is the pressure and $u = u(x)$ the controlled density of volume forces. $Re$ is the *Reynolds number*.

**Notation:**

$$
(y\cdot\nabla)\,y = y_1\,D_1 y + y_2 D_2 y + y_3 D_3 y = \sum_{i=1}^{3} y_i \begin{bmatrix} D_i y_1 \\ D_i y_2 \\ D_i y_3 \end{bmatrix}
$$

**Target:** For instance, a desired velocity field $y_d$.

## 1.3 Linear elliptic PDEs

**Sobolev spaces:** We shall use the Sobolev space

$$H^1(\Omega) = \{y \in L^2(\Omega) \,:\, D_i\, y \in L^2(\Omega),\ i = 1, \ldots, N\},$$

endowed with the norm

$$\|y\|_{H^1(\Omega)} = \Big(\int\limits_{\Omega} \big(y^2 + |\nabla y|^2\big)\, dx\Big)^{1/2}$$

(where $|\nabla y|^2 = (D_1\, y)^2 + \ldots (D_N\, y)^2$). With the inner product

$$\big(u\,,\, v\big)_{H^1(\Omega)} = \int\limits_{\Omega} u\, v\, dx + \int\limits_{\Omega} \nabla u \cdot \nabla v\, dx,$$

$H^1(\Omega)$ is a Hilbert space. Moreover, we need $H^1_0(\Omega)$, the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$ (functions of $H^1(\Omega)$ with boundary value (=trace) zero).

### 1.3.1 Poisson equation

Consider for given $f \in L^2(\Omega)$

$$\boxed{\begin{array}{rcll} -\Delta y & = & f & \text{in } \Omega \\ y & = & 0 & \text{on } \Gamma. \end{array}}$$

**Weak formulation:** Multiply the PDE by an arbitray but fixed *test function* $v \in H_0^1(\Omega)$ and integrate on $\Omega$. Then

$$-\int_\Omega v\,\Delta y\,dx = \int_\Omega f\,v\;dx$$

and integrating by parts,

$$-\int_\Gamma v\,\partial_n y\,ds + \int_\Omega \nabla y \cdot \nabla v\;dx = \int_\Omega f\,v\;dx.$$

$\partial_n := \partial/\partial n$; By $v|_\Gamma = 0$:

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx.$$

**Definition:** *A function $y \in H_0^1(\Omega)$ is a* weak solution *of the Poisson equation, if the* weak formulation

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx \qquad \forall \; v \in H_0^1(\Omega)$$

*is satisfied.*

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx.$$

**Definition:** *A function $y \in H_0^1(\Omega)$ is a* weak solution *of the Poisson equation, if the* weak formulation

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx \qquad \forall \; v \in H_0^1(\Omega)$$

*is satisfied.*

**Theorem:** *In any bounded Lipschitz domain $\Omega$, for each $f \in L^2(\Omega)$, the Poisson equation admits a unique weak solution $y \in H_0^1(\Omega)$. There exists a constant $c_P$, independent of $f$, such that*

$$\|y\|_{H^1(\Omega)} \leq c_P \, \|f\|_{L^2(\Omega)}.$$

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx.$$

**Definition:** *A function $y \in H_0^1(\Omega)$ is a* weak solution *of the Poisson equation, if the* weak formulation

$$\int_\Omega \nabla y \cdot \nabla v \; dx = \int_\Omega f \, v \; dx \qquad \forall \; v \in H_0^1(\Omega)$$

*is satisfied.*

**Theorem:** *In any bounded Lipschitz domain $\Omega$, for each $f \in L^2(\Omega)$, the Poisson equation admits a unique weak solution $y \in H_0^1(\Omega)$. There exists a constant $c_P$, independent of $f$, such that*

$$\|y\|_{H^1(\Omega)} \leq c_P \, \|f\|_{L^2(\Omega)}.$$

**Conclusion:** *The mapping $G : f \mapsto y$ is continuous from $L^2(\Omega)$ to $H_0^1(\Omega)$.*

**1.3.2 Robin boundary condition** . We proceed similarly for

$$\boxed{\begin{array}{rclll} -\Delta y + c_0\, y & = & f & \text{in} & \Omega \\ \partial_n y + \sigma\, y & = & g & \text{on} & \Gamma \end{array}}$$

where are given: $f \in L^2(\Omega)$, $g \in L^2(\Gamma)$, $c_0 \in L^\infty(\Omega)$, $\sigma \in L^\infty(\Gamma)$.

Multiplying by a test function $v \in H^1(\Omega)$,

$$-\int_\Gamma v\, \partial_n y\, ds + \int_\Omega \nabla y \cdot \nabla v\ dx + \int_\Omega c_0\, y\, v\ dx = \int_\Omega f\, v\ dx.$$

Inserting the boundary condition $\partial_n y = g - \sigma\, y$

$$\int_\Omega \nabla y \cdot \nabla v\ dx + \int_\Omega c_0\, y\, v\ dx + \int_\Gamma \sigma\, y\, v\, ds = \int_\Omega f\, v\ dx + \int_\Gamma g\, v\ ds$$

for all $v \in H^1(\Omega)$.

**Definition:** *An* $y \in H^1(\Omega)$ *is a* weak solution *of the Robin problem, if the weak formulation*

$$\int_\Omega \nabla y \cdot \nabla v \; dx + \int_\Omega c_0 \, y \, v \; dx + \int_\Gamma \sigma \, y \, v \, ds = \int_\Omega f \, v \; dx + \int_\Gamma g \, v \; ds$$

*is satisfied for all* $v \in H^1(\Omega)$.

**Definition:** *An* $y \in H^1(\Omega)$ *is a* weak solution *of the Robin problem, if the weak formulation*

$$\int_\Omega \nabla y \cdot \nabla v \; dx + \int_\Omega c_0 \, y \, v \; dx + \int_\Gamma \sigma \, y \, v \, ds = \int_\Omega f \, v \; dx + \int_\Gamma g \, v \; ds$$

*is satisfied for all* $v \in H^1(\Omega)$.

**Theorem:** *Let a bounded Lipschitz domain* $\Omega$ *and non-negative* $c_0 \in L^\infty(\Omega)$ *und* $\sigma \in L^\infty(\Gamma)$ *be given such that*

$$\int_\Omega c_0(x)^2 \, dx + \int_\Gamma \sigma(x)^2 \, ds(x) > 0.$$

*Then the Robin problem admits for each pair* $f \in L^2(\Omega)$, $g \in L^2(\Gamma)$ *a unique weak solution* $y \in H^1(\Omega)$. *There is a constant* $c_R$, *independent of* $f$ *und* $g$ *such that*

$$\|y\|_{H^1(\Omega)} \leq c_R \left( \|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Gamma)} \right).$$

We have introduced several examples of linear and semilinear elliptic equations that form the state equation of optimal control problems.

For linear equations, the spaces $H_0^1(\Omega)$ and $H^1(\Omega)$ are adequate. Associated optimal control problems fall into the class of convex optimization problems.

If semilinear state equations are given, then the problems are in general nonconvex optimization problems, even if the objective functional to be minimized is convex.

# 2. Existence of optimal controls

## 2.1 Distributed control – optimal heat source

### Main assumptions I:

- $\Omega$ bounded Lipschitz domain
- Bounds $u_a \leq u_b$ bounded and measurable (or just missing)
- $c_0$ and $\sigma$ nonnegative with $\|c_0\|_\infty + \|\sigma\|_\infty \neq 0$
- target $y_\Omega \in L^\infty(\Omega)$
- $\alpha \geq 0$

## 2.1 Distributed control – optimal heat source

$$\min\ J(y,u) := \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

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

$$u_a(x) \le u(x) \le u_b(x) \qquad \text{a.e. in } \Omega.$$

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

$U_{ad}$ is non-empty, closed and bounded in $L^2(\Omega)$. The functions of $U_{ad}$ are the *feasible controls*.

**Control-to-state mapping:** $G:$ $u \mapsto y$, $G : L^2(\Omega) \to H_0^1(\Omega)$.

**Solution operator:** $S:$ $u \mapsto y$, $S : L^2(\Omega) \to L^2(\Omega)$.

**Definition:** *A control* $\bar{u} \in U_{ad}$ *is said to be* optimal *with associated optimal state* $\bar{y} = y(\bar{u})$, *if*

$$J(\bar{y}, \bar{u}) \leq J(y(u), u) \qquad \forall\, u \in U_{ad}.$$

**Definition:** *A control* $\bar{u} \in U_{ad}$ *is said to be* optimal *with associated optimal state* $\bar{y} = y(\bar{u})$*, if*

$$J(\bar{y}, \bar{u}) \leq J(y(u), u) \qquad \forall\, u \in U_{ad}.$$

**Transformation:** We formally eliminate the PDE by

$$\frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)} = \frac{1}{2}\,\|S\,u - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)} =: \widehat{J}(u)$$

**Definition:** *A control $\bar{u} \in U_{ad}$ is said to be* optimal *with associated optimal state $\bar{y} = y(\bar{u})$, if*

$$J(\bar{y}, \bar{u}) \leq J(y(u), u) \qquad \forall\, u \in U_{ad}.$$

**Transformation:** We formally eliminate the PDE by

$$\frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)} = \frac{1}{2}\,\|S\,u - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)} =: \widehat{J}(u)$$

Thus the optimal control problem admits the form of a quadratic optimal control problem in the Hilbert space $U = L^2(\Omega)$:

$$\min_{u \in U_{ad}} \widehat{J}(u) := \frac{1}{2}\,\|S\,u - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}.$$

**Theorem:** *Let real Hilbert spaces $U$ and $H$, a nonempty, closed, bounded and convex set $U_{ad} \subset U$, $y_d \in U$ and $\alpha \geq 0$ be given. Let $S : U \to H$ be a linear and continuous operator. Then the quadratic optimization problem*

$$\min_{u \in U_{ad}} \widehat{J}(u) := \frac{1}{2} \|S\,u - y_d\|_H^2 + \frac{\alpha}{2} \|u\|_U^2$$

*admits an optimal solution $\bar{u}$. It is unique for $\alpha > 0$ .*

The proof uses the fact that, under the assumptions above, the set $U_{ad}$ is weakly sequentially compact. Moreover, by continuity and convexity, $f$ is weakly lower semicontinuous. By standard arguments, this permits to prove this known result.

**Theorem:** *Let real Hilbert spaces $U$ and $H$, a nonempty, closed, bounded and convex set $U_{ad} \subset U$, $y_d \in U$ and $\alpha \geq 0$ be given. Let $S : U \to H$ be a linear and continuous operator. Then the quadratic optimization problem*

$$\min_{u \in U_{ad}} \widehat{J}(u) := \frac{1}{2} \|S\,u - y_d\|_H^2 + \frac{\alpha}{2} \|u\|_U^2$$

*admits an optimal solution $\bar{u}$. It is unique for $\alpha > 0$ .*

The proof uses the fact that, under the assumptions above, the set $U_{ad}$ is weakly sequentially compact. Moreover, by continuity and convexity, $f$ is weakly lower semicontinuous. By standard arguments, this permits to prove this known result.

**Corollary:** *For $\alpha > 0$, the problem of optimal heat source admits a unique optimal control $\bar{u}$.*

**Theorem:** *Let real Hilbert spaces $U$ and $H$, a nonempty, closed, bounded and convex set $U_{ad} \subset U$, $y_d \in U$ and $\alpha \geq 0$ be given. Let $S : U \to H$ be a linear and continuous operator. Then the quadratic optimization problem*

$$\min_{u \in U_{ad}} \widehat{J}(u) := \frac{1}{2} \|S\,u - y_d\|_H^2 + \frac{\alpha}{2} \|u\|_U^2$$

*admits an optimal solution $\bar{u}$. It is unique for $\alpha > 0$ .*

The proof uses the fact that, under the assumptions above, the set $U_{ad}$ is weakly sequentially compact. Moreover, by continuity and convexity, $f$ is weakly lower semicontinuous. By standard arguments, this permits to prove this known result.

**Corollary:** *For $\alpha > 0$, the problem of optimal heat source admits a unique optimal control $\bar{u}$.*

**Case of optimal boundary control:** Since $G : L^2(\Gamma) \to H^1(\Omega)$ is continuous, the same result can be derived for the problem of optimal stationary boundary temperature.

### 2.2.1 Existence for semilinear equations

In the semilinear elliptic case, we consider the optimal control of the following

**Model problem:**

$$\boxed{\begin{array}{rcll} -\Delta y + c_0(x)\, y + d(y) & = & f & \text{in} \quad \Omega \\ \partial_n y + \sigma(x)\, y + b(y) & = & g & \text{on} \quad \Gamma. \end{array}}$$

The functions $c_0$ and $\sigma$ fulfill the same assumptions as before and $d$, $b : \mathbb{R} \to \mathbb{R}$ are monotone non-decreasing, differentiable with locally Lipschitz first derivative.

## 2.2.1 Existence for semilinear equations

In the semilinear elliptic case, we consider the optimal control of the following

**Model problem:**

$$\boxed{\begin{array}{rcll} -\Delta y + c_0(x)\, y + d(y) & = & f & \text{in} \quad \Omega \\ \partial_n y + \sigma(x)\, y + b(y) & = & g & \text{on} \quad \Gamma. \end{array}}$$

The functions $c_0$ and $\sigma$ fulfill the same assumptions as before and $d,\, b : \mathbb{R} \to \mathbb{R}$ are monotone non-decreasing, differentiable with locally Lipschitz first derivative.

**Choice of the state space:** *Now, $H^1(\Omega)$ is not in general the suitable space of solutions $y$, since $y \in H^1(\Omega)$ does not guarantee that $d(y) \in L^2(\Omega)$ and even not $d(y) \in L^1(\Omega)$.*

We define the solution in $Y = H^1(\Omega) \cap C(\bar{\Omega})$. The space $C(\bar{\Omega})$ is also important to deal with state-constrained problems later.

If $b$ and $d$ are in addition uniformly bounded on $\mathbb{R}$ and vanish at $0$, then the theorem on monotone operators by Browder and Minty can be applied to show that

$$\begin{aligned} -\Delta y + c_0(x)\, y + d(y) &= f \qquad \text{in} \quad \Omega \\ \partial_n y + \sigma(x)\, y + b(y) &= g \qquad \text{on} \quad \Gamma \end{aligned}$$

has for all pairs $(f, g) \in L^2(\Omega) \times L^2(\Gamma)$ a unique weak solution $y \in H^1(\Omega)$ that satisfies the estimate

$$\|y\|_{H^1(\Omega)} \le c_M \left(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Gamma)}\right).$$

Here, the constant does not depend on $f$ and $g$ and even not on $b$, $d$.

If $b$ and $d$ are in addition uniformly bounded on IR and vanish at $0$, then the theorem on monotone operators by Browder and Minty can be applied to show that

$$\begin{aligned} -\Delta y + c_0(x)\, y + d(y) &= f \qquad \text{in} \quad \Omega \\ \partial_n y + \sigma(x)\, y + b(y) &= g \qquad \text{on} \quad \Gamma \end{aligned}$$

has for all pairs $(f, g) \in L^2(\Omega) \times L^2(\Gamma)$ a unique weak solution $y \in H^1(\Omega)$ that satisfies the estimate

$$\|y\|_{H^1(\Omega)} \leq c_M \left(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Gamma)}\right).$$

Here, the constant does not depend on $f$ and $g$ and even not on $b$, $d$.

**Weak solution:**

$$\int_\Omega \nabla y \cdot \nabla v\, dx + \int_\Omega (c_0\, y + d(y))\, v\, dx + \int_\Gamma (\sigma\, y + b(y))\, v\, ds = \int_\Omega f\, v\, dx + \int_\Gamma g\, v\, ds$$

for all $v \in H^1(\Omega)$.

If the degree of integrability of $f$ and $g$ is sufficiently high, then the solution $y$ of the semilinear equation is bounded and even continuous. Therefore, it is quite natural that the assumption on uniform boundedness of $b$ and $d$ is not needed:

**Assumption:** *Let $c_0$ and $\sigma$ be as before, $d$, $b$ monotone non-decreasing and continuous.*

If the degree of integrability of $f$ and $g$ is sufficiently high, then the solution $y$ of the semilinear equation is bounded and even continuous. Therefore, it is quite natural that the assumption on uniform boundedness of $b$ and $d$ is not needed:

**Assumption:** *Let $c_0$ and $\sigma$ be as before, $d$, $b$ monotone non-decreasing and continuous.*

**Theorem:** *Let the assumption above be satisfied, $r > N/2$, $s > N - 1$. Then, for each pair $f \in L^r(\Omega)$, $g \in L^s(\Gamma)$, the semilinear model equation has a unique weak solution $y \in H^1(\Omega) \cap C(\bar{\Omega})$. If, in addition, $b(0) = d(0) = 0$ holds, then*

$$\|y\|_{H^1(\Omega)} + \|y\|_{C(\bar{\Omega})} \leq c_\infty \left( \|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)} \right)$$

*holds with a constant $c_\infty$ that does not depend on $d$, $b$, $f$ and $g$.*

Casas 1993, Alibert and Raymond 1997; Stampacchia method

Without the assumption $b(0) = d(0) = 0$, the estimate

$$\|y\|_{H^1(\Omega)} + \|y\|_{C(\bar{\Omega})} \leq c_\infty \big(\|f\|_{L^r(\Omega)} + \|g\|_{L^s(\Gamma)} + 1\big)$$

holds true.

**Dimensions:** In the following cases, data from $L^2$ are sufficient to have continuous solutions:

Distributed control: $r = 2 > N/2 \Leftrightarrow N < 4$

Boundary control: $s = 2 > N - 1 \Leftrightarrow N < 3$

### 2.2.2 Control problem and existence of optimal controls

We already have discussed the solvability of the equations in $H^1(\Omega) \cap C(\bar{\Omega})$.

Now we consider the associated control problem

$$\min \, J(y,u) :=:= \frac{1}{2} \, \|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \, \|u\|^2_{L^2(\Omega)}$$

subject to

$$\boxed{\begin{array}{rcll} -\Delta y + y + d(y) & = & u & \text{in} \quad \Omega \\ \partial_n y & = & 0 & \text{auf} \;\; \Gamma \end{array}}$$

and

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

**Theorem:** *Under the assumptions posed above, the distributed optimal problem for the semilinear elliptic model equation admits at least one optimal control.*

**Remarks:**

– There can be more than one optimal control, even infinitely many different ones might exist.

– Locally optimal controls are of interest as well.

– We need necessary / sufficient conditons to find them.

**Theorem:** *Under the assumptions posed above, the distributed optimal problem for the semilinear elliptic model equation admits at least one optimal control.*

**Remarks:**

– There can be more than one optimal control, even infinitely many different ones might exist.

– Locally optimal controls are of interest as well.

– We need necessary / sufficient conditons to find them.

Let us start with the necessary ones...

### 2.2.3 Derivatives

Let $U$, $V$ be real Banach spaces and $F : U \to V$ mapping.

**Definition:** *$F : U \to V$ is said to be Fréchet-differentiable at $u$, if there exist a linear and continuous operator $A : U \to V$ and a mapping $r :\ U \times U \to V$ with the following properties: For all $h \in U$,*

$$F(u+h) = F(u) + A\,h + r(u,h)$$

*and the remainder term $r$ satisfies*

$$\frac{\|r(u,h)\|_V}{\|h\|_U} \to 0 \qquad \text{as} \qquad \|h\|_U \to 0.$$

*$A$ is the Fréchet derivative of $F$ at $u$, $A = F'(u)$.*

**Example 1:**

Let $U$ and $H$ real Hilbert spaces, $z \in H$, $S : U \to H$ linear and continuous. Then

$$F(u) := \|S\,u - z\|_H^2$$

is Fréchet-differentiable on $U$ and

$$F'(u)\,h = 2\left(S^*(S\,u - z)\,,\,h\right)_U.$$

This is the expression for the derivative. The element $2\,S^*(S\,u - z)$ is said to be the gradient of $F$,

$$F'(u) = 2\,S^*(S\,u - z).$$

[Riesz representation]

**Example 2:** The mapping $\Phi : y(\cdot) \mapsto \sin(y(\cdot))$,

$$(\Phi(y))(x) := \sin(y(x))$$

is called a superposition operator or Nemytskij operator.

It is

- Lipschitz continuous from $L^p(\Omega)$ to $L^p(\Omega)$ for all $1 \leq p \leq \infty$.

**Example 2:** The mapping $\Phi : y(\cdot) \mapsto \sin(y(\cdot))$,

$$(\Phi(y))(x) := \sin(y(x))$$

is called a superposition operator or Nemytskij operator.

It is

- Lipschitz continuous from $L^p(\Omega)$ to $L^p(\Omega)$ for all $1 \le p \le \infty$.
- not F-differentiable in any $L^p$-space with $p < \infty$.

**Example 2:** The mapping $\Phi : y(\cdot) \mapsto \sin(y(\cdot))$,

$$(\Phi(y))(x) := \sin(y(x))$$

is called a superposition operator or Nemytskij operator.

It is

- Lipschitz continuous from $L^p(\Omega)$ to $L^p(\Omega)$ for all $1 \le p \le \infty$.
- not F-differentiable in any $L^p$-space with $p < \infty$.
- F-differentiable in $L^\infty(\Omega)$ and in $C(\bar{\Omega})$. Here, the derivative is

$$(\Phi'(y)\, h)(x) := \cos(y(x))\, h(x).$$

**Example 2:** The mapping $\Phi : y(\cdot) \mapsto \sin(y(\cdot))$,

$$(\Phi(y))(x) := \sin(y(x))$$

is called a superposition operator or Nemytskij operator.

It is

- Lipschitz continuous from $L^p(\Omega)$ to $L^p(\Omega)$ for all $1 \le p \le \infty$.
- not F-differentiable in any $L^p$-space with $p < \infty$.
- F-differentiable in $L^\infty(\Omega)$ and in $C(\bar{\Omega})$. Here, the derivative is

$$(\Phi'(y)\,h)(x) := \cos(y(x))\,h(x).$$

- It is F-differentiable from $L^p(\Omega)$ to $L^q(\Omega)$ with $q < p$...

In the linear-quadratic case, the convexity of the problem and boundedness of the feasible set $U_{ad}$ guarantee the existence of at least one optimal control that is unique for $\alpha > 0$. We are justified to work in the Hilbert spaces $H^1(\Omega)$ and $L^2(\Omega)$.

In the case of semilinear equations, the theory is more difficult. To have existence for the equation, we need in general the space $H^1(\Omega) \cap C(\bar{\Omega})$. Then existence can be shown by the theory of monotone operators and the Stampacchia method.

Existence of optimal controls can be shown as well.

# 3. First order necessary optimality conditions

## 3.1 Quadratic optimization in Hilbert space

## 3.2 Distributed control

### 3.2.1 Adjoint equation

### 3.2.2 Projection formula

### 3.2.3 Test examples

### 3.2.4 Lagrange multipliers

### 3.2.5 Karush-Kuhn-Tucker system

### 3.2.6 The reduced gradient

3.3 Boundary control

3.4 The formal Lagrange principle

3.5 Control of semilinear equations

3.6 Second order sufficient conditions

3.7 Pointwise state constraints

# 3. First order necessary optimality conditions

## 3.1 Quadratic optimization in Hilbert space

We have transformed our elliptic optimal control problems to the following optimization problem in Hilbert space:

$$(P) \qquad \min_{u \in U_{ad}} \qquad \widehat{J}(u) := \frac{1}{2}\|Su - y_d\|_H^2 + \frac{\alpha}{2}\|u\|_U^2.$$

**Lemma:** *Let $U$ be a real Banach space, $C \subset U$ a convex set and $f: \ C \to I\!R$ F-differentiable on $C$. Let $\bar{u} \in C$ be a solution of (P). Then the following variational inequality is satisfied:*

$$\widehat{J}'(\bar{u})(u - \bar{u}) \geq 0 \quad \forall u \in C.$$

Application to $\widehat{J}$ above...

**Application to (P):** If $\bar{u}$ is a solution to (P), then

$$\widehat{J}'(\bar{u})(u - \bar{u}) \quad = (S^*(S\,\bar{u} - y_d) + \alpha\bar{u}\,,\, u - \bar{u})_H \geq 0 \quad \forall u \in C.$$

**Application to (P):** If $\bar{u}$ is a solution to (P), then

$$\widehat{J}'(\bar{u})(u-\bar{u}) \quad = (S^*(S\,\bar{u}-y_d)+\alpha\bar{u}\,,\,u-\bar{u})_H \geq 0 \quad \forall u \in C.$$

## 3.2. Distributed control

### 3.2.1 Adjoint equation

**Recall:** In the distributed elliptic problem, we have $S: u \mapsto y$, $S: L^2(\Omega) \to L^2(\Omega)$. The problem was

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

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

$$u_a(x) \leq u(x) \leq u_b(x) \qquad \text{a.e. in } \Omega.$$

To apply the variational inequality above, we need $S^*$.

**Lemma:** *In the case of the Poisson equation, the adjoint operator $S^* : L^2(\Omega) \to L^2(\Omega)$ is given by*

$$S^* z := p,$$

*where $p \in H_0^1(\Omega)$ is the weak solution of the following Poisson equation:*

$$\begin{aligned} -\Delta p &= z \quad \textit{in} \quad \Omega \\ p &= 0 \quad \textit{auf } \Gamma. \end{aligned}$$

**Application:** $S^*(S\,\bar{u} - y_\Omega) = S^*\,(\bar{y} - y_\Omega) = p$, where

$$\begin{aligned} -\Delta p &= \bar{y} - y_\Omega \quad \text{in} \quad \Omega \\ p &= 0 \quad \text{on} \quad \Gamma \end{aligned}$$

**Definition:** *The weak solution $p \in H_0^1(\Omega)$ of the adjoint equation*

$$\boxed{\begin{array}{rcll} -\Delta p & = & \bar{y} - y_\Omega & \textit{in} \quad \Omega \\ p & = & 0 & \textit{auf } \Gamma \end{array}}$$

*is called adjoint state associated with $\bar{y}$.*

We get $S^*(S\,\bar{u} - y_\Omega) + \alpha\bar{u} = p + \alpha\bar{u}$ hence

$$\big(p + \alpha\bar{u}\,,\, u - \bar{u}\big)_{L^2(\Omega)} \geq 0 \quad \forall\, u \in U_{ad}.$$

**Theorem:** *If $\bar{u}$ is optimal for the distributed problem above and $\bar{y}$ the associated state, then there exists a unique adjoint state $p$ defined by the adjoint equation and the following variational inequality is satisfied:*

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

Summarizing up, we have the

**Optimality system:**

$$\boxed{\begin{array}{c} \begin{array}{rclcrcl} -\Delta y &=& u & \qquad & -\Delta p &=& y - y_\Omega \\ y|_\Gamma &=& 0 & \qquad & p|_\Gamma &=& 0 \end{array} \\ u \in U_{ad} \\ \big(p + \alpha\, u\,,\, v - u\big)_{L^2(\Omega)} \geq 0 \quad \forall v \in U_{ad}. \end{array}}$$

Each $u \in U_{ad}$ that satisfies together with $y$ and the adjoint state $p$ the optimality system, is optimal. This follows from the convexity of the problem.

Each $u \in U_{ad}$ that satisfies together with $y$ and the adjoint state $p$ the optimality system, is optimal. This follows from the convexity of the problem.

### 3.2.2 Projection formula

**Discussion of the variational inequality**

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

$$\int_\Omega (p + \alpha\,\bar{u})\,\bar{u}\ dx \leq \int_\Omega (p + \alpha\,\bar{u})\,u\ dx \qquad \forall u \in U_{ad}$$

$$\int_\Omega (p + \alpha\,\bar{u})\,\bar{u}\ dx = \min_{u \in U_{ad}} \int_\Omega (\,p + \alpha\,\bar{u})\,u\ dx.$$

Each $u \in U_{ad}$ that satisfies together with $y$ and the adjoint state $p$ the optimality system, is optimal. This follows from the convexity of the problem.

### 3.2.2 Projection formula

**Discussion of the variational inequality**

$$\int\limits_{\Omega} \big(p(x) + \alpha\,\bar{u}(x)\big)\big(u(x) - \bar{u}(x)\big)\; dx \geq 0 \qquad \forall\, u \in U_{ad}$$

$$\int\limits_{\Omega} (p + \alpha\,\bar{u})\,\bar{u}\; dx \leq \int\limits_{\Omega} (p + \alpha\,\bar{u})\, u\; dx \qquad \forall u \in U_{ad}$$

$$\int\limits_{\Omega} (p + \alpha\,\bar{u})\,\bar{u}\; dx = \min_{u \in U_{ad}} \int\limits_{\Omega} (\,p + \alpha\,\bar{u})\, u\; dx.$$

$$\Rightarrow \quad \big(p(x) + \alpha\,\bar{u}(x)\big)\,\bar{u}(x) = \min_{v \in [u_a(x), u_b(x)]} \big(p(x) + \alpha\,\bar{u}(x)\big)\, v \quad \text{a.e. on } \Omega.$$

We obtain immediately

$$\bar{u}(x) = \begin{cases} u_a(x), & \text{where} \quad p(x) + \alpha\,\bar{u}(x) > 0 \\ \in [u_a(x), u_b(x)], & \text{where} \quad p(x) + \alpha\,\bar{u}(x) = 0 \\ u_b(x), & \text{where} \quad p(x) + \alpha\,\bar{u}(x) < 0. \end{cases}$$

$\alpha = \mathbf{0}$ :

$$\bar{u}(x) = \begin{cases} u_a(x) & \text{where} \quad p(x) > 0 \\ u_b(x) & \text{where} \quad p(x) < 0. \end{cases}$$

We obtain immediately

$$\bar{u}(x) = \begin{cases} u_a(x), & \text{where} \quad p(x) + \alpha\,\bar{u}(x) > 0 \\ \in [u_a(x), u_b(x)], & \text{where} \quad p(x) + \alpha\,\bar{u}(x) = 0 \\ u_b(x), & \text{where} \quad p(x) + \alpha\,\bar{u}(x) < 0. \end{cases}$$

$\alpha = \mathbf{0}$ :

$$\bar{u}(x) = \begin{cases} u_a(x) & \text{where} \quad p(x) > 0 \\ u_b(x) & \text{where} \quad p(x) < 0. \end{cases}$$

**Theorem:** *For $\alpha > 0$, $\bar{u}$ is optimal iff the projection formula*

$$\bar{u}(x) = P_{[u_a(x), u_b(x)]}\left\{-\frac{1}{\alpha}p(x)\right\}$$

*is satisfied for a.a. $x \in \Omega$ with the associated $p$. Here, $P_{[a,b]}$ denotes the projection from $\mathbb{R}$ to $[a, b]$.*

**$\alpha > 0$ and $\mathbf{U_{ad}} = \mathbf{L^2(\Omega)}$:** We get

$$\bar{u} = -\frac{1}{\alpha}\, p.$$

Inserting this in the state equation gives the optimality system

$$\boxed{\begin{array}{rclcrcl} -\Delta y &=& -\alpha^{-1}\, p & \quad & -\Delta p &=& y - y_\Omega \\ y|_\Gamma &=& 0 & & p|_\Gamma &=& 0, \end{array}}$$

a coupled system of two elliptic boundary value problems to find $y = \bar{y}$ and $p$. After having computed $p$, we find $\bar{u}$ by $\bar{u} = -p/\alpha$.

### 3.2.3 Test examples

**Bang-Bang control:** We consider the following problem

$$\min \int_{\Omega} (y - y_\Omega)^2 \, dx$$

$$\begin{aligned} -\Delta y &= u + e_\Omega \\ y|_\Gamma &= 0 \end{aligned}$$

$$-1 \leq u(x) \leq 1.$$

The term $e_\Omega$ does not change the optimality conditions. We take the unit square $\Omega = (0,1)^2$. To obtain a "checkerboard function" as optimal control, we proceed as follows: $\Omega$ is partitioned into $8 \times 8 = 64$ subdomains, where the control admits the values $1$ and $-1$.

Free to adapt the problem: $y_\Omega$, $e_\Omega$

## Optimality system:

$$
\begin{aligned}
-\Delta y &= u + e_\Omega & \qquad -\Delta p &= y - y_\Omega \\
y|_\Gamma &= 0 & p|_\Gamma &= 0
\end{aligned}
$$

$$u(x) = -\ \text{sign}\ p(x)$$

## Optimality system:

$$\begin{aligned} -\Delta y &= u + e_\Omega & -\Delta p &= y - y_\Omega \\ y|_\Gamma &= 0 & p|_\Gamma &= 0 \end{aligned}$$

$$u(x) = -\ \text{sign}\ p(x)$$

We just define the optimal state and adjoint state by

$$\begin{aligned} \bar{y}(x) &= y(x_1, x_2) = \sin(\pi\, x_1)\sin(\pi\, x_2), \\ p(x) &= p(x_1, x_2) = -\tfrac{1}{128\,\pi^2}\sin(8\,\pi\, x_1)\sin(8\,\pi\, x_2) \end{aligned}$$

## Optimality system:

$$\begin{aligned} -\Delta y &= u + e_\Omega & \qquad -\Delta p &= y - y_\Omega \\ y|_\Gamma &= 0 & p|_\Gamma &= 0 \end{aligned}$$

$$u(x) = -\ \text{sign}\ p(x)$$

We just define the optimal state and adjoint state by

$$\begin{aligned} \bar{y}(x) &= y(x_1, x_2) = \sin(\pi\, x_1)\sin(\pi\, x_2), \\ p(x) &= p(x_1, x_2) = -\tfrac{1}{128\,\pi^2}\sin(8\,\pi\, x_1)\sin(8\,\pi\, x_2) \end{aligned}$$

To satisfy both equations, we adapt $y_\Omega$ and $e_\Omega$ by

$$\begin{aligned} e_\Omega &= -\Delta\bar{y} - \bar{u} = 2\pi^2\ \sin(\pi\, x_1)\sin(\pi\, x_2) + \text{sign}\left(-\sin(8\,\pi\, x_1)\sin(8\,\pi\, x_2)\right) \\ y_\Omega &= \bar{y} + \Delta p = y_\Omega(x) = \sin(\pi\, x_1)\sin(\pi\, x_2) + \sin(8\,\pi\, x_1)\sin(8\,\pi\, x_2). \end{aligned}$$

The checkerboard function $\bar{u}$ satisfies these conditions, hence (convexity!) it is optimal.

**Problem with Neumann boundary condition:**

$$\min\ J(y,u) := \frac{1}{2}\int_\Omega (y - y_\Omega)^2\,dx + \int_\Gamma e_\Gamma\,y\,ds + \frac{1}{2}\int_\Omega u^2\,dx$$

$$\begin{aligned} -\Delta y + y &= u + e_\Omega, \qquad 0 \le u(x) \le 1 \\ \partial_n y &= 0. \end{aligned}$$

Again, $\Omega = (0,1)^2$ with midpoint $\hat{x} = (0.5, 0.5)^\top$.

**Problem with Neumann boundary condition:**

$$\min\ J(y,u) := \frac{1}{2}\int_{\Omega} (y - y_\Omega)^2\,dx + \int_{\Gamma} e_\Gamma\, y\,ds + \frac{1}{2}\int_{\Omega} u^2\,dx$$

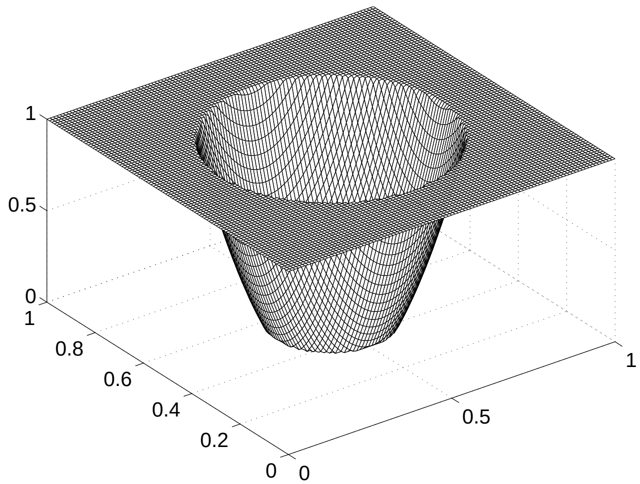
$$\begin{aligned} -\Delta y + y &= u + e_\Omega, \qquad 0 \le u(x) \le 1 \\ \partial_n y &= 0. \end{aligned}$$

Again, $\Omega = (0,1)^2$ with midpoint $\hat{x} = (0.5, 0.5)^\top$. We play with $y_\Omega$, $e_\Omega$, $e_\Gamma$.

$$r := |x - \hat{x}| = \sqrt{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}$$

**Desired optimal control:**

$$\bar{u}(x) = \begin{cases} 1 & \text{for} \quad r > \frac{1}{3} \\ 12\,r^2 - \frac{1}{3} & \text{for} \quad r \in [\frac{1}{6}, \frac{1}{3}] \\ 0 & \text{for} \quad r < \frac{1}{6} \end{cases}$$



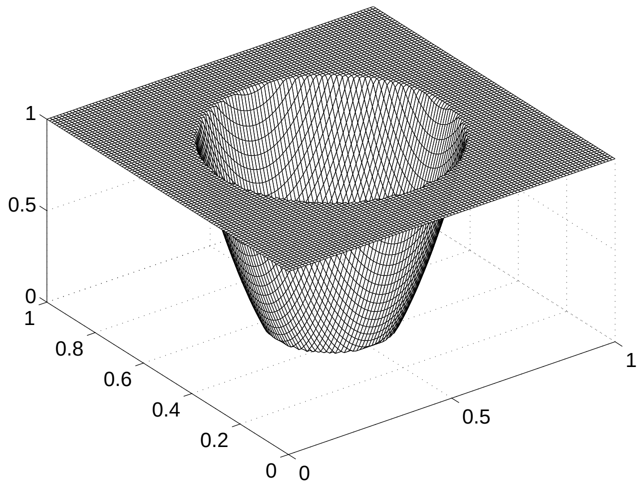

*Optimal control*

**Adjoint equation:**

$$\begin{aligned} -\Delta p + p &= \bar{y} - y_\Omega \\ \partial_n p &= e_\Gamma \end{aligned}$$

**Projection formula:**

$$\bar{u}(x) = \mathsf{P}_{[0,1]}\{-p(x)\}$$




*Optimal control*

## Adjoint equation:

$$
\begin{aligned}
-\Delta p + p &= \bar{y} - y_\Omega \\
\partial_n p &= e_\Gamma
\end{aligned}
$$

## Projection formula:

$$\bar{u}(x) = \mathsf{P}_{[0,1]}\{-p(x)\}$$

$$
\begin{aligned}
p(x) &:= -12\,|x-\hat{x}|^2 + \tfrac{1}{3} = -12\,r^2 + \tfrac{1}{3} \\
y(x) &:= 1 \\
e_\Omega &= 1 - \min\{1, \max\{0,\ 12\,r^2 - \tfrac{1}{3}\}\} \\
y_\Omega(x) &= 1 - 48 - \tfrac{1}{3} + 12\,|x-\hat{x}|^2 = -\tfrac{142}{3} + 12\,r^2 \\
e_\Gamma &= \partial_n p = -12.
\end{aligned}
$$

### 3.2.4 Lagrange multipliers

**Theorem:** *The variational inequality* $(\alpha\bar{u} + p\,,\, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}$ *is equivalent to the existence of a.e. nonnegative functions* $\mu_a$, $\mu_a \in L^2(\Omega)$ *such that the equation*

$$p + \alpha\bar{u} - \mu_a + \mu_b = 0$$

*and the* complementarity conditions

$$\mu_a(x)\left(u_a(x) - \bar{u}(x)\right) = \mu_b(x)\left(\bar{u}(x) - u_b(x)\right) = 0$$

*are satisfied a.e. in* $\Omega$.

## 3.2.4 Lagrange multipliers

**Theorem:** *The variational inequality* $(\alpha\bar{u} + p\,,\, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}$ *is equivalent to the existence of a.e. nonnegative functions* $\mu_a$, $\mu_a \in L^2(\Omega)$ *such that the equation*

$$p + \alpha\bar{u} - \mu_a + \mu_b = 0$$

*and the* complementarity conditions

$$\mu_a(x)\,\big(u_a(x) - \bar{u}(x)\big) = \mu_b(x)\,\big(\bar{u}(x) - u_b(x)\big) = 0$$

*are satisfied a.e. in* $\Omega$*.*

**Proof:** We define

$$\begin{aligned} \mu_a(x) \quad &:= \quad \big(p(x) + \alpha u(x)\big)^+ \\ \mu_b(x) \quad &:= \quad \big(p(x) + \alpha u(x)\big)^-, \end{aligned}$$

By definition, we have $\mu_a \geq 0$, $\mu_b \geq 0$ und $p + \alpha u = \mu_a - \mu_b$.

Moreover, we know the following implications:

$$(\,p + \alpha\,\bar{u})(x) > 0 \Rightarrow \bar{u}(x) = u_a(x)$$

$$(\,p + \alpha\,\bar{u})(x) < 0 \Rightarrow \bar{u}(x) = u_b(x)$$

$$u_a(x) < \bar{u}(x) < u_b(x) \Rightarrow (\,p + \alpha\,\bar{u})(x) = 0.$$

Moreover, we know the following implications:

$$(p + \alpha\,\bar{u})(x) > 0 \Rightarrow \bar{u}(x) = u_a(x)$$

$$(p + \alpha\,\bar{u})(x) < 0 \Rightarrow \bar{u}(x) = u_b(x)$$

$$u_a(x) < \bar{u}(x) < u_b(x) \Rightarrow (p + \alpha\,\bar{u})(x) = 0.$$

This gives the complementarity conditions, since always one of the two factors is zero. For instance,

$$\mu_a(x) > 0 \Rightarrow \mu_b(x) = 0, \quad \text{and} \quad (p + \alpha\,\bar{u})(x) = \mu_a > 0$$

and thus $\bar{u}(x) - u_a(x) = 0$. Therefore,

$$(u(x) - u_a(x))\,\mu_a(x) = 0.$$

$\square$

## 3.2.5 Optimality system (KKT system)

$$
\boxed{
\begin{array}{c}
\begin{array}{rclcrcl}
-\Delta y & = & u & \qquad & -\Delta p & = & y - y_\Omega \\
y|_\Gamma & = & 0 & & p|_\Gamma & = & 0
\end{array} \\
p + \alpha u - \mu_a + \mu_b = 0 \\
u_a \le u \le u_b, \quad \mu_a \ge 0, \quad \mu_b \ge 0, \\
\mu_a(x)\left(u_a(x) - \bar{u}(x)\right) = \mu_b(x)\left(\bar{u}(x) - u_b(x)\right) = 0.
\end{array}
}
$$

The nondifferential equations are satisfied a.e. in $\Omega$.

### 3.2.6 The gradient of the objective functional

The adjoint state permits a simple expression for the gradient of $\widehat{J}(u) = J\big(y(u), u\big)$.

**Lemma:** *The gradient of the functional*

$$\widehat{J}(u) = J\big(y(u), u\big) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

*is given by*

$$\nabla \widehat{J}(u) = p + \alpha\, u,$$

*where $p \in H^1_0(\Omega)$ is the weak solution of the adjoint equation*

$$\begin{aligned} -\Delta p &= y - y_\Omega \quad \text{in} \quad \Omega \\ p &= 0 \quad\quad\quad\ \ \text{on} \ \Gamma \end{aligned}$$

*and $y = y(u)$ is the state associated with $u$.*

### 3.2.6 The gradient of the objective functional

The adjoint state permits a simple expression for the gradient of $\widehat{J}(u) = J\big(y(u), u\big)$.

**Lemma:** *The gradient of the functional*

$$\widehat{J}(u) = J\big(y(u), u\big) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

*is given by*

$$\nabla \widehat{J}(u) = p + \alpha\, u,$$

*where $p \in H_0^1(\Omega)$ is the weak solution of the adjoint equation*

$$\begin{aligned} -\Delta p &= y - y_\Omega \quad \textit{in} \quad \Omega \\ p &= 0 \qquad\quad\ \textit{on}\ \Gamma \end{aligned}$$

*and $y = y(u)$ is the state associated with $u$.*

This follows from

$$\widehat{J}'(u)\,h = \big(S^*(S\,u - y_\Omega) + \alpha\,u\,,\,h\big)_{L^2(\Omega)} = \big(p + \alpha\,u\,,\,h\big)_{L^2(\Omega)}$$

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

**S1** Compute $y_k$ associated with $u_k$ (state equation)

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

**S1** Compute $y_k$ associated with $u_k$ (state equation)

**S2** Compute the associated adjoint state $p_k$ from

$$\begin{aligned} -\Delta p &= y_k - y_\Omega \quad \text{in} \quad \Omega \\ p &= 0 \quad\quad\quad\quad \text{on} \ \Gamma. \end{aligned}$$

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

**S1** Compute $y_k$ associated with $u_k$ (state equation)

**S2** Compute the associated adjoint state $p_k$ from

$$\begin{aligned} -\Delta p &= y_k - y_\Omega \quad \text{in} \quad \Omega \\ p &= 0 \qquad\qquad \text{on} \quad \Gamma. \end{aligned}$$

Direction of descent: $-\nabla \widehat{J}(u_k) = -(\alpha u_k + p_k)$ (antigradient)

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

**S1** Compute $y_k$ associated with $u_k$ (state equation)

**S2** Compute the associated adjoint state $p_k$ from

$$\begin{array}{rcll} -\Delta p &=& y_k - y_\Omega & \text{in } \Omega \\ p &=& 0 & \text{on } \Gamma. \end{array}$$

Direction of descent: $-\nabla \widehat{J}(u_k) = -(\alpha u_k + p_k)$ (antigradient)

**S3** $u_{k+1} = u_k - \tau_k \nabla \widehat{J}(u_k)$

where the optimal stepsize $\tau_k$ solves the quadratic problem $\min\limits_{\tau \geq 0} \widehat{J}(u_k - \tau \nabla \widehat{J}(u_k))$.

Can be done analytically...

We consider for simplicity the case without control constraints and assume $\alpha > 0$ for existence.

**Gradient method:** Let $u_1, \cdots u_k$ already have been computed.

**S1** Compute $y_k$ associated with $u_k$ (state equation)

**S2** Compute the associated adjoint state $p_k$ from

$$\begin{aligned} -\Delta p &= y_k - y_\Omega \quad \text{in} \quad \Omega \\ p &= 0 \quad \text{on} \quad \Gamma. \end{aligned}$$

Direction of descent: $-\nabla \widehat{J}(u_k) = -(\alpha u_k + p_k)$ (antigradient)

**S3** $u_{k+1} = u_k - \tau_k \nabla \widehat{J}(u_k)$

where the optimal stepsize $\tau_k$ solves the quadratic problem $\min\limits_{\tau \geq 0} \widehat{J}(u_k - \tau \nabla \widehat{J}(u_k))$.
Can be done analytically...

Continue by S1, if the descent is sufficiently large, otherwise stop.

### 3.2.7 Primal-dual active set strategy

(Bergounioux, Ito and Kunisch)

**Optimality system:**

$$\boxed{\begin{array}{c} \begin{array}{rclcrcl} -\Delta y & = & u & \qquad & -\Delta p & = & y - y_\Omega \\ y|_\Gamma & = & 0 & & p|_\Gamma & = & 0 \end{array} \\ u \in U_{ad} \\ \big( p + \alpha\, u\,,\, v - u \big)_{L^2(\Omega)} \geq 0 \quad \forall v \in U_{ad}. \end{array}}$$

The variational inequality is equivalent to

$$\bar{u}(x) = \mathsf{P}_{[u_a(x), u_b(x)]}\big\{ -\alpha^{-1}\, p(x) \big\}.$$

$$\bar{u}(x) = \mathsf{P}_{[u_a(x), u_b(x)]}\left\{ -\alpha^{-1} p(x) \right\}.$$

Define

$$\mu = -(\alpha^{-1} p + \bar{u}) = -\alpha^{-1} \nabla \widehat{J}(\bar{u}).$$

$$\bar{u}(x) = \mathsf{P}_{[u_a(x), u_b(x)]}\big\{ - \alpha^{-1}\, p(x) \big\}.$$

Define

$$\mu = -(\alpha^{-1}\, p + \bar{u}) = -\alpha^{-1}\, \nabla \widehat{J}(\bar{u}).$$

We find

$$\bar{u}(x) = \begin{cases} u_a(x) & \text{if} \quad -\alpha^{-1}\, p(x) < u_a(x) & (\Leftrightarrow \mu(x) < 0) \\ -\alpha^{-1}\, p(x) & \text{if} \quad -\alpha^{-1}\, p(x) \in [u_a(x), u_b(x)] & (\Leftrightarrow \mu(x) = 0) \\ u_b(x) & \text{if} \quad -\alpha^{-1}\, p(x) > u_b(x) & (\Leftrightarrow \mu(x) > 0). \end{cases}$$

$$\bar{u}(x) = \mathsf{P}_{[u_a(x), u_b(x)]}\{-\alpha^{-1} p(x)\}.$$

Define

$$\mu = -(\alpha^{-1} p + \bar{u}) = -\alpha^{-1} \nabla \widehat{J}(\bar{u}).$$

We find

$$\bar{u}(x) = \begin{cases} u_a(x) & \text{if} \quad -\alpha^{-1} p(x) < u_a(x) & (\Leftrightarrow \mu(x) < 0) \\ -\alpha^{-1} p(x) & \text{if} \quad -\alpha^{-1} p(x) \in [u_a(x), u_b(x)] & (\Leftrightarrow \mu(x) = 0) \\ u_b(x) & \text{if} \quad -\alpha^{-1} p(x) > u_b(x) & (\Leftrightarrow \mu(x) > 0). \end{cases}$$

Take, for instance, the upper case. Then, by definition of $\mu$ and by $\bar{u} = u_a$, we have $\mu(x) < 0$ and hence $\bar{u}(x) + \mu(x) < u_a(x)$

$$u(x) = \begin{cases} u_a(x) & \text{if} \quad u(x) + \mu(x) < u_a(x) \\ -\alpha^{-1} p(x) & \text{if} \quad u(x) + \mu(x) \in [u_a(x), u_b(x)] \\ u_b(x) & \text{if} \quad u(x) + \mu(x) > u_b(x). \end{cases}$$

Fix initial functions $u_0$, $\mu_0$ in $L^2(\Omega)$. Current iterate: $u_{k-1}$ und $\mu_{k-1}$. Next:

**S1** *(New active and inactive sets)*

$$\begin{aligned} A_k^+ &= \{x : u_{k-1}(x) + \mu_{k-1}(x) > u_b(x)\} \\ A_k^- &= \{x : u_{k-1}(x) + \mu_{k-1}(x) < u_a(x)\} \\ I_k &= \Omega \setminus (A_k^+ \cup A_k^-). \end{aligned}$$

If $A_k^+ = A_{k-1}^+$ and $A_k^- = A_{k-1}^-$, then terminate because of optimality. Otherwise continue.

**S2** *(New control)* Solve

$$\begin{aligned} -\Delta y &= u \\ -\Delta p &= y - y_\Omega \end{aligned}$$

$$u = \begin{cases} u_a & \text{on} \quad A_k^- \\ -\alpha^{-1}\, p & \text{on} \quad I_k \\ u_b & \text{on} \quad A_k^+ \end{cases}$$

with $y,\, p \in H_0^1(\Omega)$.

Define $u_k := u,\ p_k := p,\ \mu_k := -(\alpha^{-1}\, p_k + u_k),\ k := k+1,$

goto $\mathbf{S1}$.

## 3.3 Boundary control

Here, $G : u \mapsto y$ is linear and continuous from $L^2(\Gamma)$ to $H^1(\Omega)$. We consider it as operator $S$ with range in $L^2(\Omega)$. With $S$, the objective functional reads

$$\widehat{J}(u) = \frac{1}{2} \|Su - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u\|^2_{L^2(\Gamma)}.$$

Let $\bar{u} \in U_{ad}$ be optimal with state $\bar{y}$. The adjoint state is defined by

$$\boxed{\begin{aligned} -\Delta p &= \bar{y} - y_\Omega && \text{in } \Omega \\ \partial_n p + \sigma\, p &= 0 && \text{on } \Gamma. \end{aligned}}$$

**Theorem:** *If $\bar{u}$ is optimal with adjoint state $p$, then for a.a. $x \in \Gamma$, the minimum*

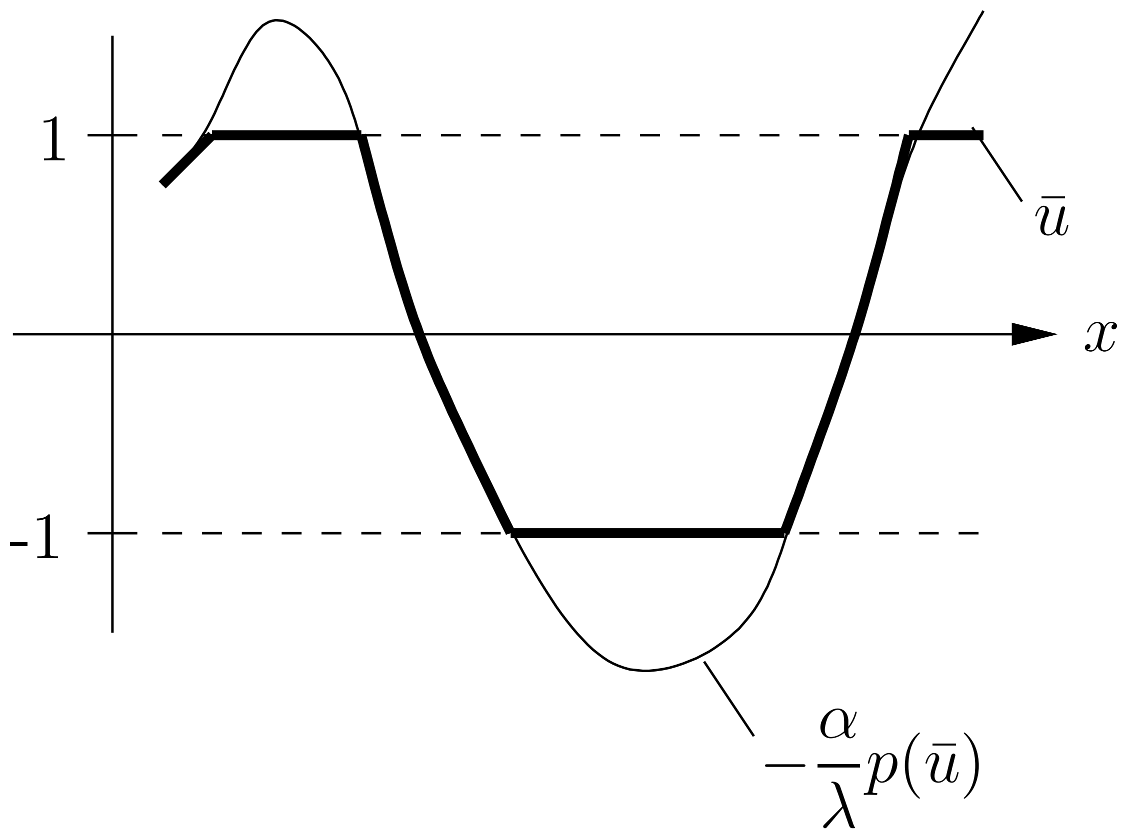
$$\min_{u_a(x) \leq v \leq u_b(x)} \left\{ \sigma(x)\, p(x)\, v + \frac{\alpha}{2} v^2 \right\}$$

*is attained by $v = \bar{u}(x)$. Therefore, for $\alpha > 0$ the projection formula*

$$\bar{u}(x) = P_{[u_a(x), u_b(x)]} \left\{ -\frac{1}{\alpha} \sigma(x)\, p(x) \right\}$$

*is fulfilled for a.a. $x \in \Gamma$.*

**Case $\alpha > 0$, $|u| \leq 1$:**




*Optimal control for $\alpha > 0$*

## 3.4 The formal Lagrange principle

The necessary optimality conditions (variational inequality, adjoint equation) can be easily obtained by a Lagrange function. We consider again the boundary control,

$$\min J(y,u) := \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Gamma)}$$

$$\begin{aligned} -\Delta y &= 0 && \text{in } \Omega \\ \partial_n y + \sigma y &= \sigma\, u && \text{on } \Gamma \end{aligned}$$

$$u_a(x) \le u(x) \le u_b(x) \quad \text{a.e. on } \Gamma.$$

We eliminate the differential constraints (PDE, boundary condition) by Lagrange multipliers $p_1$, $p_2$. Later, we shall see $p_2 = p_1|_\Gamma$. Therefore we put $p := p_1$ and $p_2 := p|_\Gamma$.

**Definition :**

$$\mathcal{L} = \mathcal{L}(y,u,p) = J(y,u) - \int_{\Omega} (-\Delta y)\, p \; dx - \int_{\Gamma} \big(\partial_n y - \sigma(u-y)\big)\, p \; ds.$$

*with Lagrange multiplier $p$.*

**Definition :**

$$\mathcal{L} = \mathcal{L}(y,u,p) = J(y,u) - \int_{\Omega} (-\Delta y)\, p\, dx - \int_{\Gamma} \big(\partial_n y - \sigma(u-y)\big)\, p\, ds.$$

*with Lagrange multiplier $p$.*

This is a little bit too formal, since $-\Delta\, y$ is not in general a function. Integrating by parts:

**Definition:** *The Lagrange function $\mathcal{L}: \; H^1(\Omega) \times L^2(\Gamma) \times H^1(\Omega) \to I\!R$ for the boundary control problem is*

$$\boxed{\mathcal{L}(y,u,p) := J(y,u) - \int_{\Omega} \nabla y \cdot \nabla p\, dx + \int_{\Gamma} \sigma\,(u-y)\, p\, ds.}$$

The multiplier $p$ is identical with the adjoint state.

It is easy to verify that

- $D_y\mathcal{L}(\bar{y},\bar{u},p)\,h = 0 \quad \forall\, h \in H^1(\Omega)$

  is the weak formulation of the adjoint equation.

- $D_u\mathcal{L}(\bar{y},\bar{u},p)\,(u-\bar{u}) \geq 0 \quad \forall\, u \in U_{ad}$

  gives the variational inequality.

- The gradient of the reduced functional $\widehat{J}(u) = J(y(u),u)$ is obtained by

$$\widehat{J}'(u) = D_u\mathcal{L}(y,u,p),$$

  with $y = y(u)$ und $p = p(u)$.

## Extension

In the same way, the box constraints on the control can be eliminated by Lagrange multipliers $\mu_a$, $\mu_b$.

## Extended Lagrange function:

$$\begin{aligned}\mathcal{L}(y,u,p,\mu_a,\mu_b) \quad := \quad & J(y,u) - \int_\Omega \nabla y \cdot \nabla p \, dx + \int_\Gamma \sigma\,(u-y)\,p\,ds \\ & + \int_\Omega \big(\mu_a(u_a - u) + \mu_b(u - u_b)\big)\,dx.\end{aligned}$$

Again, $D_y\mathcal{L} = 0$ gives the adjoint equation, while $D_u\mathcal{L} = 0$ yields the gradient equation.

## 3.5. Control of semilinear equations

We already have discussed the solvability of the equations in $H^1(\Omega) \cap C(\bar{\Omega})$.

Now we consider the control problem

$$\min \; J(y,u) :=:= \frac{1}{2}\, \|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\, \|u\|^2_{L^2(\Omega)}$$

subject to

$$\boxed{\begin{array}{rcll} -\Delta y + y + d(y) & = & u & \text{in} \quad \Omega \\ \partial_n y & = & 0 & \text{auf} \;\; \Gamma \end{array}} \tag{1}$$

and

$$u_a(x) \leq u(x) \leq u_b(x) \quad \text{a.e. in } \Omega. \tag{2}$$

**Definition:**

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

We repeat the assumptions for convenience.

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $y_\Omega \in L^\infty(\Omega)$, $\alpha \ge 0$, $u_a$, $u_b \in L^\infty(\Omega)$ with $u_a(x) \le u_b(x)$.*

**Definition:** *A control $\bar{u} \in U_{ad}$ is locally optimal in the sense of $L^r(\Omega)$, if there exists $\varepsilon > 0$ such that*

$$J\big(y(\bar{u}), \bar{u}\big) \le J\big(y(u), u\big)$$

*holds for all $u \in U_{ad}$ with $\|u - \bar{u}\|_{L^r(\Omega)} \le \varepsilon$.*

Our state equation is

$$\boxed{\begin{array}{rcll} -\Delta y + y + d(y) & = & u & \text{in} \quad \Omega \\ \partial_n y & = & 0 & \text{on} \quad \Gamma. \end{array}}$$

To each $u \in U := L^r(\Omega)$, $r > N/2$, there exists exactly one state $y \in Y = H^1(\Omega) \cap C(\bar{\Omega})$. We denote the associated control-to-state mapping by $G : U \to Y$, $G(u) = y$.

$G$ is twice continuously differentiable:

**Theorem:** *$G$ is twice continuously Fréchet-differentiable from $L^r(\Omega)$ to $H^1(\Omega) \cap C(\bar{\Omega})$, $r > N/2$. It holds $G'(\bar{u})\, u = y$, where $y$ is the solution of the linearized problem*

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\, y &= u \quad \text{in} \quad \Omega \\ \partial_n y &= 0 \quad \text{on} \quad \Gamma. \end{aligned}$$

**Theorem:** *$G$ is twice continuously Fréchet-differentiable from $L^r(\Omega)$ to $H^1(\Omega) \cap C(\bar{\Omega})$, $r > N/2$. It holds $G'(\bar{u})\, u = y$, where $y$ is the solution of the linearized problem*

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\, y &= u \quad \text{in} \quad \Omega \\ \partial_n y &= 0 \quad \text{on} \quad \Gamma. \end{aligned}$$

*The second derivative is given by $G''(\bar{u})[u_1, u_2] = z$, where $z$ solves*

$$\begin{aligned} -\Delta z + z + d'(\bar{y})\, z &= -d''(\bar{y})\, y_1\, y_2 \\ \partial_n z &= 0 \end{aligned}$$

*and $y_i \in H^1(\Omega)$ are defined by $y_i = G'(u)\, u_i$.*

**Theorem:** *$G$ is twice continuously Fréchet-differentiable from $L^r(\Omega)$ to $H^1(\Omega) \cap C(\bar{\Omega})$, $r > N/2$. It holds $G'(\bar{u})\,u = y$, where $y$ is the solution of the linearized problem*

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\,y &= u \quad \textit{in} \quad \Omega \\ \partial_n y &= 0 \quad \textit{on} \quad \Gamma. \end{aligned}$$

*The second derivative is given by $G''(\bar{u})[u_1, u_2] = z$, where $z$ solves*

$$\begin{aligned} -\Delta z + z + d'(\bar{y})\,z &= -d''(\bar{y})\,y_1\,y_2 \\ \partial_n\,z &= 0 \end{aligned}$$

*and $y_i \in H^1(\Omega)$ are defined by $y_i = G'(u)\,u_i$.*

This is the basis to derive optimality conditions. However, we just rely on the formal Lagrange technique.

**Lagrange function:** *The Lagrange function $\mathcal{L}: H^1(\Omega) \times L^2(\Gamma) \times H^1(\Omega) \longrightarrow I\!R$ for the semilinear distributed control problem is*

$$\boxed{\mathcal{L}(y,u,p) := J(y,u) - \int_\Omega \nabla y \cdot \nabla p \, dx - \int_\Omega (y + d(y) - u)\, p \, dx.}$$

**Lagrange function:** *The Lagrange function* $\mathcal{L}: H^1(\Omega)\times L^2(\Gamma)\times H^1(\Omega)\to I\!R$ *for the semilinear distributed control problem is*

$$\boxed{\mathcal{L}(y,u,p) := J(y,u) - \int_\Omega \nabla y\cdot\nabla p\; dx - \int_\Omega (y + d(y) - u)\, p\; dx.}$$

We proceed as before. The adjoint equation is obtained by $D_y\mathcal{L}y = 0$ for all $y$:

$$(\bar y - y_\Omega\,,\, y) - (\nabla y\,,\,\nabla p) - (p + d'(\bar y)p\,,\, y) = 0 \quad \forall\, y\in H^1(\Omega).$$

This is the weak formulation of the adjoint equation

$$\begin{aligned} -\Delta p + p + d'(\bar y)\,p &= \bar y - y_\Omega && \text{in } \Omega\\ \partial_n p &= 0 && \text{on } \Gamma. \end{aligned}$$

The variational inequality is obtained from $D_u\mathcal{L}(u-\bar{u}) \geq 0$ for all $u \in U_{ad}$. We obtain

$$(\alpha\bar{u} + p\,,\, u - \bar{u}) \geq 0.$$

Consequence for $\alpha > 0$:

$$\boxed{\bar{u}(x) = \mathsf{P}_{[u_a(x),u_b(x)]}\big\{ -\frac{1}{\alpha}\, p(x)\big\}.}$$

**Example:** "Superconductivity"

$$\min\; J(y,u) := \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

subject to $-2 \leq u(x) \leq 2$ and

$$\begin{aligned} -\Delta y + y + y^3 &= u \\ \partial_n y &= 0. \end{aligned}$$

## Optimality system:

$$
\begin{aligned}
-\Delta y + y + y^3 &= u \qquad & -\Delta p + p + 3\,\bar{y}^2\,p &= \bar{y} - y_\Omega \\
\partial_n y &= 0 & \partial_n p &= 0
\end{aligned}
$$

$$
\bar{u}(x) = P_{[u_a(x), u_b(x)]}\left\{ -\frac{1}{\alpha} p(x) \right\},
$$

if $\alpha > 0$. Therefore $\bar{u} \in H^1(\Omega) \cap C(\bar{\Omega})$, if $u_a$, $u_b$ in $H^1(\Omega) \cap C(\bar{\Omega})$. ◇

**Test example:** For $\alpha = 1$ and $y_\Omega \equiv 9$, $\bar{u}(x) \equiv 2$ satisfies the first order conditions.

Test it...

Is it locally optimal? ◇

## 3.6. Second-order sufficient optimality conditions

To check for local optimality, we need second-order sufficient optimality conditions. In infinite dimensions, the theory of second-order conditions is more difficult than in finite-dimensional spaces. We only formally state them without discussing the main difficulties behind.

**Critical cone:**

$$C(\bar{u}) = \big\{ u \in L^\infty(\Omega) \,:\, u(x) \geq 0, \text{ where } \bar{u}(x) = u_a(x), \quad u(x) \leq 0, \text{ where } \bar{u}(x) = u_b(x) \big\}$$

**SSC:** *There exists* $\delta > 0$ *such that*

$$\widehat{J}''(\bar{u})\, u^2 \geq \delta\, \|u\|^2_{L^2(\Omega)} \quad \forall u \in C(\bar{u}).$$

**(SSC)** is equivalent with

$$\boxed{\int\limits_\Omega \left\{ \big(1 - p\, d''(\bar{y})\big)\, y^2 + \alpha\, u^2 \right\} dx \ge \delta\, \|u\|^2_{L^2(\Omega)}}$$

for all $u \in C(\bar{u})$ and all $y \in H^1(\Omega)$ that satisfy

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\, y &= u \\ \partial_n y &= 0. \end{aligned}$$

In terms of the Lagrange function, the second-order sufficient condition can be expressed as follows:

$$\mathcal{L}''(\bar{u}, \bar{y}, p)[y, u]^2 \geq \delta \|u\|^2_{L^2(\Omega)}$$

for all $u \in C(\bar{u})$ and all $y$ that solve

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\, y &= u \\ \partial_n y &= 0. \end{aligned}$$

In terms of the Lagrange function, the second-order sufficient condition can be expressed as follows:

$$\mathcal{L}''(\bar{u}, \bar{y}, p)[y, u]^2 \geq \delta \|u\|^2_{L^2(\Omega)}$$

for all $u \in C(\bar{u})$ and all $y$ that solve

$$\begin{aligned} -\Delta y + y + d'(\bar{y})\, y &= u \\ \partial_n y &= 0. \end{aligned}$$

**Theorem:** *Let $\bar{u} \in U_{ad}$, $\bar{y} = G(\bar{u})$ and $p$ satisfy together the first-order necessary and second-order sufficient conditions. Then there exist constants $\varepsilon > 0$ and $\sigma > 0$ such that the quadratic growth condition*

$$J(y, u) \geq J(\bar{y}, \bar{u}) + \sigma \|u - \bar{u}\|^2_{L^2(\Omega)}$$

*holds for all $u \in U_{ad}$ with $\|u - \bar{u}\|_{L^\infty(\Omega)} \leq \varepsilon$ and $y = G(u)$. Therefore, $\bar{u}$ is locally optimal in the sense of $L^\infty(\Omega)$.*

We show that the solution of our test example "Superconductivity" is locally optimal. Recall the problem:

$$\min\ J(y,u) := \frac{1}{2}\,\|y-9\|^2_{L^2(\Omega)} + \frac{1}{2}\,\|u\|^2_{L^2(\Omega)}$$

subject to $-2 \le u(x) \le 2$ and

$$\begin{aligned} -\Delta y + y + y^3 &= u \\ \partial_n y &= 0. \end{aligned}$$

Our candidate was $u = 2,\ y = 1,\ p = -2.$ We get

$$\begin{aligned} \mathcal{L}''(\bar u,\bar y,p)[y,u]^2 &= \int_\Omega \Big\{ \big(1 - p\, d''(\bar y)\big)\, y^2 + u^2 \Big\}\, dx \\ &= \int_\Omega \Big\{ \big(1 + 2\cdot 6\,\big)\, y^2 + u^2 \Big\}\, dx \ge 1\cdot\ \|u\|^2_{L^2(\Omega)}. \end{aligned}$$

**Conclusion:** *The control $\bar u \equiv 2$ is locally optimal in $L^\infty(\Omega)$.*

## 3.7. Pointwise state constraints

Often, in addition to the control constraints, bounds on the state $y$ are given which have to be satisfied in the whole domain $\Omega$. Such pointwise state constraints are difficult in theory and numerics.

**State constrained problem:**

$$\min\ J(y,u) :=:= \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

$$\boxed{\begin{array}{rclll} -\Delta y + y + d(y) & = & u & \text{in} & \Omega \\ \partial_n y & = & 0 & \text{on} & \Gamma \end{array}}$$

$$\begin{array}{lll} u_a(x) \le u(x) \le u_b(x) & \text{a.e. in } \Omega & \text{(control constraints)} \\ y_a(x) \le y(x) \le y_b(x) & \forall x \in \bar{\Omega} & \text{(state constraints)} \end{array}$$

**Control space:** $U = L^\infty(\Omega)$

**State space:** $Y = C(\bar\Omega)$

**Control-to-state operator:** $G : u \mapsto y,\ G : L^\infty(\Omega) \to (H^1(\Omega)\cap)\ C(\bar\Omega)$

**Abstract formulation:**

$$\min\ \widehat{J}(u) := J(G(u), u), \quad u_a \le u \le u_b, \quad y_a \le G(u) \le y_b$$

**Lagrange function 1:**

$$\begin{aligned} L(u, \mu_a, \mu_b, \nu_a, \nu_b) \quad := \quad & \widehat{J}(u) + \int_\Omega (u_a - u)\mu_a\, dx + \int_\Omega (u - u_b)\mu_b\, dx \\ & + \int_{\bar\Omega} (y_a - G(u)) d\nu_a + \int_{\bar\Omega} (G(u) - y_b) d\nu_b \end{aligned}$$

Here, the Lagrange multipliers $\mu_a$, $\mu_b$ are functions from certain spaces $L^p(\Omega)$, while $\nu_a$, $\nu_b$ are regular Borel measures (elements of $C(\bar\Omega)^*$.)

## Lagrange function 2:

$$\begin{aligned}\mathcal{L}(y,u,p,\mu_a,\mu_b,\nu_a,\nu_b) \quad := \quad & J(y,u) - \int_\Omega \nabla y \cdot \nabla p \, dx - \int_\Omega (y + d(y) - u)\, p \, dx \\ & + \int_\Omega \big((u_a - u)\mu_a + (u - u_b)\mu_b\big)\, dx \\ & + \int_{\bar{\Omega}} \big((y_a - y) d\nu_a + (y - y_b) d\nu_b\big)\end{aligned}$$

## Lagrange function 2:

$$
\begin{aligned}
\mathcal{L}(y,u,p,\mu_a,\mu_b,\nu_a,\nu_b) \quad := \quad & J(y,u) - \int_\Omega \nabla y \cdot \nabla p \, dx - \int_\Omega (y + d(y) - u)\, p \, dx \\
& + \int_\Omega \big( (u_a - u)\mu_a + (u - u_b)\mu_b \big) \, dx \\
& + \int_{\bar\Omega} \big( (y_a - y) d\nu_a + (y - y_b) d\nu_b \big)
\end{aligned}
$$

To guarantee existence of Lagrange multipliers, a constraint qualification is needed. For a locally optimal $\bar u$, we assume the

**Linearized Slater condition:** There exist $\varepsilon > 0$ and $\tilde u \in L^\infty(\Omega)$ such that

$$
\begin{aligned}
& u_a(x) + \varepsilon \le \tilde u(x) \le u_b(x) - \varepsilon && \text{a.e. on } \Omega \\
& y_a(x) < \bar y(x) + G'(\bar y)(\tilde u - \bar u)(x) < y_a(x) && \forall x \in \bar\Omega.
\end{aligned}
$$

**Theorem:** *If $\bar{u}$ is locally optimal and the constraint qualification is satisfied, then there exist non-negative Lagrange multipliers $\mu_a,\ \mu_b, \nu_a, \nu_b$ and an adjoint state $p \in W^{1,s}(\Omega)$ for all $s \in [1, N/(N-1))$ such that*

$$\boxed{\begin{aligned} &D_y\mathcal{L}(\bar{y}, \bar{u}, p, \mu_a,\ \mu_b, \nu_a, \nu_b) = 0 \\ &D_u\mathcal{L}(\bar{y}, \bar{u}, p, \mu_a,\ \mu_b, \nu_a, \nu_b) = 0 \\ &(\bar{u} - u_a\ ,\ \mu_a)_{L^2(\Omega)} = (\bar{u} - u_b\ ,\ \mu_b)_{L^2(\Omega)} = 0 \\ &(\bar{y} - y_a\ ,\ \nu_a)_{C(\bar{\Omega}),M(\bar{\Omega})} = (\bar{y} - y_b\ ,\ \nu_b)_{C(\bar{\Omega}),M(\bar{\Omega})} = 0 \end{aligned}}$$

The first equation leads to an adjoint equations with measures in the right hand side discussed e.g. by Casas and also by Alibert and Raymond. The second is the so-called gradient equation. The last conditions are the complementarity conditions.

In some cases (mixed control state constraints, pure state constraints), a Lavrentiev type regularization may help to avoid measures $\nu_a, \nu_b$ and to obtain instead functions (Meyer, Rösch, Tröltzsch 2004, Meyer,Tröltzsch 2005).

In this case, the following constraints are considered:

**Regularized constraints:**

$$u_a(x) \leq u(x) \leq u_b(x), \qquad y_a(x) \leq \rho\, u(x) + y(x) \leq y_b(x)$$

Then numerical methods can be set up in function space and next discretized.

We consider the following example that has been solved by a primal-dual active set strategy and by a primal-dual interior point method as well, Meyer, Prüfert, Tröltzsch 2005.

$$
(E) \begin{cases} \text{minimize} & J(y,u) := \dfrac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \dfrac{\kappa}{2}\|u - u_d\|^2_{L^2(\Omega)} \\ \text{subject to} & -\Delta\, y(x) + y(x) = u(x) \quad \text{in } \Omega \\ & \partial_n y(x) = 0 \qquad \text{on } \Gamma \\ \text{and} & y(x) \le y_b(x) \quad \text{a.e. in } \Omega \end{cases}
$$

in $\Omega = B(0,1)$, with

$$
\text{(E)} \begin{cases} \text{minimize} & J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\|u - u_d\|^2_{L^2(\Omega)} \\ \text{subject to} & -\Delta\, y(x) + y(x) = u(x) \quad \text{in } \Omega \\ & \partial_n y(x) = 0 \qquad \text{on } \Gamma \\ \text{and} & y(x) \le y_b(x) \quad \text{a.e. in } \Omega \end{cases}
$$

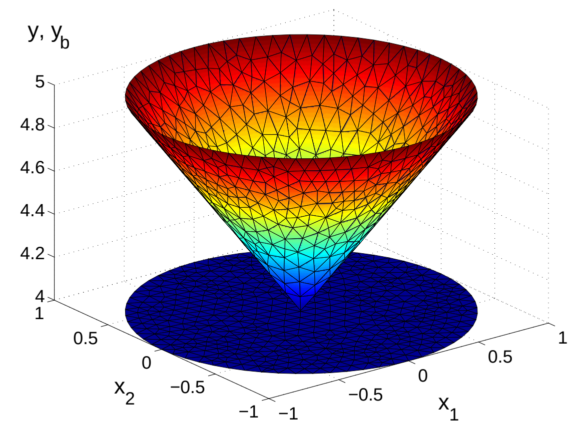
in $\Omega = B(0,1)$, with

$$
\begin{aligned}
y_d(r,\varphi) &= 4 + \tfrac{1}{\pi} - \tfrac{1}{4\pi} r^2 + \tfrac{1}{2\pi} \log(r), \\
u_d(r,\varphi) &= 4 + \tfrac{1}{4\pi\kappa} r^2 - \tfrac{1}{2\pi\kappa} \log(r), \\
y_b(r,\varphi) &= r + 4,\ \kappa = 0.5 \cdot 10^{-5}
\end{aligned}
$$

$\bar{y}(r,\varphi) \equiv 4 \quad \leq y_b(r,\varphi) = r + 4$

$\Rightarrow$ feasible, constraint only active at $r = 0$

$\Rightarrow \quad \bar{u} = -\Delta\bar{y} + \bar{y} = 4$

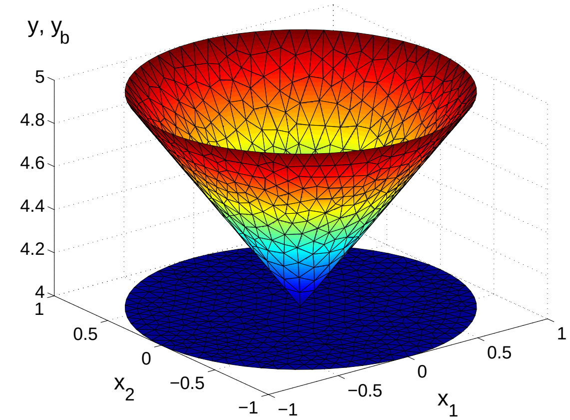



$$\mu_b = \delta_0 \in C(\bar{\Omega})^*$$

$\bar{y}(r,\varphi) \equiv 4 \quad \leq y_b(r,\varphi) = r + 4$

$\Rightarrow$ feasible, constraint only active at $r = 0$

$\Rightarrow \quad \bar{u} = -\Delta\bar{y} + \bar{y} = 4$



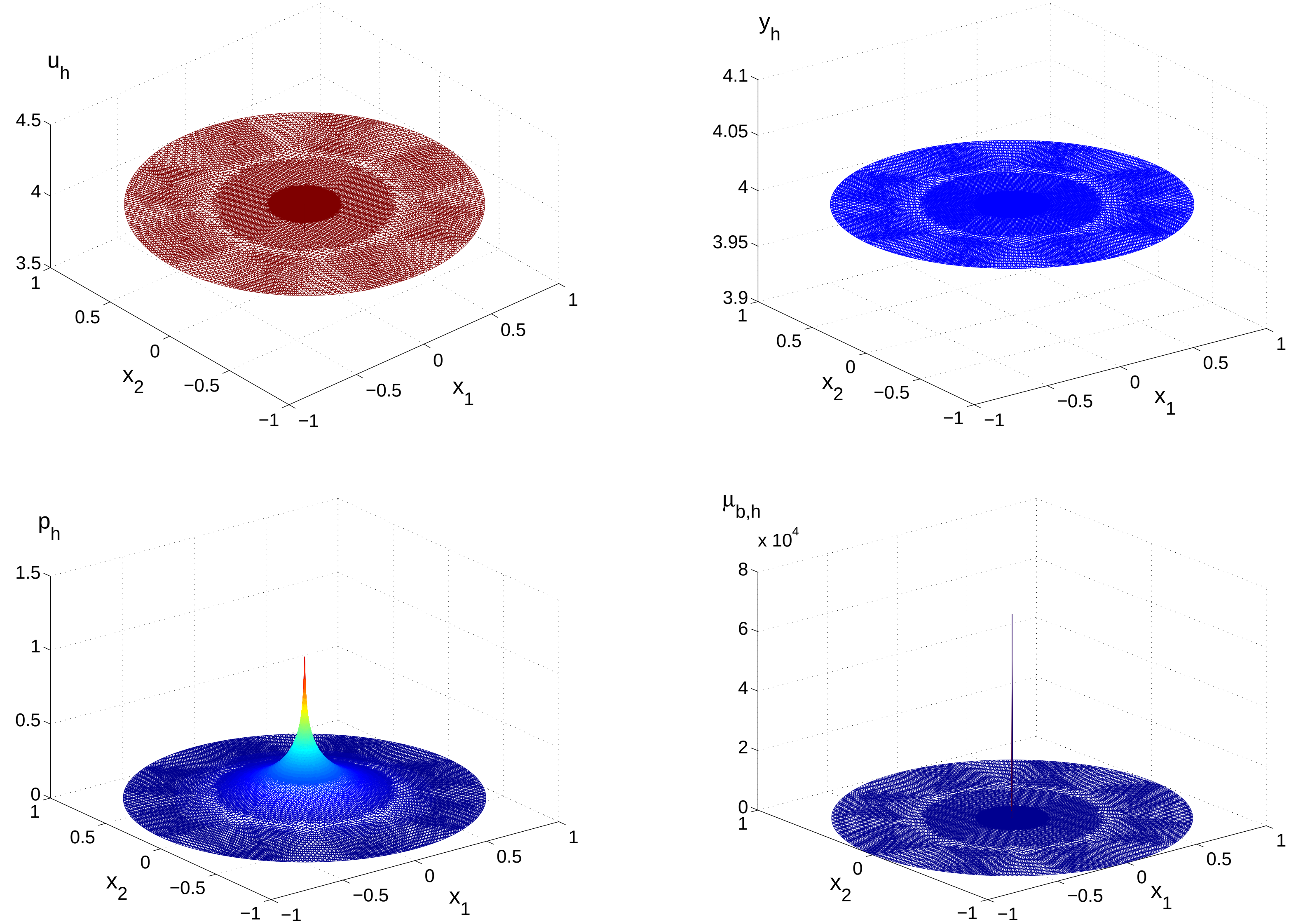

$p(r,\varphi) = \frac{1}{4\pi} r^2 + \Phi(r,\varphi)$

with: $\Phi = -\frac{1}{2\pi}\log(r)$ - fundamental solution of the Poisson eq. in $\mathbb{R}^2$

$\Rightarrow -\Delta\Phi = \delta_0$: Dirac measure

$$\mu_b = \delta_0 \in C(\bar{\Omega})^*$$

# Result for $\rho = 0.75 \cdot 10^{-4}$



*We have derived first order necessary optimality conditions for linear-quadratic and semilinear elliptic optimal control problems. They can be used to check for optimality and to construct test examples.*

*We have derived first order necessary optimality conditions for linear-quadratic and semilinear elliptic optimal control problems. They can be used to check for optimality and to construct test examples.*

*Based on the expression of the gradient, we have set up a gradient method for the numerical solution. Moreover, the optimality conditions were discussed in detail to obtain a projection formula. This formula was applied in a primal-dual active set strategy for the numerical solution of linear-quadratic elliptic problems.*

*We have derived first order necessary optimality conditions for linear-quadratic and semilinear elliptic optimal control problems. They can be used to check for optimality and to construct test examples.*

*Based on the expression of the gradient, we have set up a gradient method for the numerical solution. Moreover, the optimality conditions were discussed in detail to obtain a projection formula. This formula was applied in a primal-dual active set strategy for the numerical solution of linear-quadratic elliptic problems.*

*In the semilinear case, second-order sufficient optimality conditions have been studied. They require positive definiteness of the Lagrange function on a certain critical cone. They are important for the justification of numerical methods.*

## Classical reference:

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

## New book on optimal control of PDEs:

Tröltzsch, F., *Optimale Steuerung partieller Differentialgleichungen*, Vieweg Verlag, 2005.

ISBN 3-528-03224-3

A big part of sections 1–3 can be found in detail there.

# 4. Discretization

- Convergence analysis for the optimal control problem is more than convergence analysis for the governing PDE.
- Properties of the optimization problem (second order sufficient optimality conditions) are important.
- Two approaches: Optimize–then–discretize and discretize–then–optimize.
- May not always lead to the same result.

$$\text{Minimize } \frac{1}{2}\int_\Omega (y(x)-\widehat{y}(x))^2 dx + \frac{\alpha}{2}\int_\Omega u^2(x)dx$$

subject to

$$\begin{aligned} -\Delta y(x) &= f(x)+u(x), \quad x\in\Omega,\\ y(x) &= 0, \qquad\qquad\quad x\in\Gamma, \end{aligned}$$

where $f, \widehat{y} \in L^2(\Omega)$, $\alpha > 0$,

Define state space $Y = H_0^1(\Omega)$ and control space $U = L^2(\Omega)$.

- Weak form of the state equation:

$$\underbrace{\int_\Omega \nabla y(x)\nabla v(x)dx}_{a(y,v)} - \underbrace{\int_\Omega u(x)v(x)dx}_{b(u,v)} = \underbrace{\int_\Omega f(x)v(x)dx}_{l(v)} \quad \forall v \in Y.$$

- Weak form of the state equation:

$$\underbrace{\int_\Omega \nabla y(x)\nabla v(x)dx}_{a(y,v)} - \underbrace{\int_\Omega u(x)v(x)dx}_{b(u,v)} = \underbrace{\int_\Omega f(x)v(x)dx}_{l(v)} \quad \forall v \in Y.$$

- Optimal Control Problem

$$\min \frac{1}{2}\|y-\widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2,$$
$$\text{s.t. } a(y,v) + b(u,v) = l(v) \quad \forall v \in Y.$$

- Weak form of the state equation:

$$\underbrace{\int_\Omega \nabla y(x)\nabla v(x)dx}_{a(y,v)} - \underbrace{\int_\Omega u(x)v(x)dx}_{b(u,v)} = \underbrace{\int_\Omega f(x)v(x)dx}_{l(v)} \quad \forall v \in Y.$$

- Optimal Control Problem

$$\min \frac{1}{2}\|y-\widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2,$$
$$\text{s.t. } a(y,v)+b(u,v)=l(v) \quad \forall v \in Y.$$

- Lagrangian

$$L(y,u,p) = \frac{1}{2}\|y-\widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 - a(y,p) - b(u,p) + l(p).$$

Lagrangian

$$L(y,u,p) = \frac{1}{2}\|y-\widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 - a(y,p) - b(u,p) + l(p).$$

Necessary and sufficient optimality conditions

$$\begin{aligned}
a(v,p) - \langle y, v\rangle_{L^2} &= -\langle \widehat{y}, v\rangle_{L^2} && \forall v \in Y,\\
-b(w,p) + \alpha\langle u, w\rangle_{L^2} &= 0 && \forall w \in U,\\
a(y,v) + b(u,v) &= l(v) && \forall v \in Y.
\end{aligned}$$

Necessary and sufficient optimality conditions are the weak forms of

$$
\begin{aligned}
-\Delta p(x) &= (y(x) - \widehat{y}(x)), && x \in \Omega, \\
p(x) &= 0, && x \in \Gamma, \\
p(x) + \alpha u(x) &= 0 && x \in \Omega, \\
-\Delta y(x) &= f(x) + u(x), && x \in \Omega, \\
y(x) &= 0, && x \in \Gamma.
\end{aligned}
$$

- Let

$$Y^h = \text{span}\{\varphi_1, \dots, \varphi_n\} \subset Y,$$
$$U^h = \text{span}\{\psi_1, \dots, \psi_m\} \subset U$$

be finite dimensional subspaces of the state and control space with bases $\varphi_1, \dots, \varphi_n$ and $\psi_1, \dots, \psi_m$, respectively.

- Replace $y$ by $y^h = \sum_{i=1}^n y_i \varphi_i$, replace $u$ by $u^h = \sum_{i=1}^m u_i \psi_i$, and require that $y^h, u^h$ satisfy the state equation for all $v = \varphi_i$, $i = 1, \dots, n$.

- Discretized optimal control problem:

$$\min \frac{1}{2}\|y^h - \widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u^h\|_{L^2}^2,$$
$$\text{s.t. } a(y^h, \varphi_i) + b(u^h, \varphi_i) = l(\varphi_i) \quad i = 1, \dots, n.$$

- Discretized optimal control problem:

$$\min \frac{1}{2}\|y^h - \widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u^h\|_{L^2}^2,$$
$$\text{s.t. } a(y^h, \varphi_i) + b(u^h, \varphi_i) = l(\varphi_i) \quad i = 1, \dots, n.$$

- Necessary and sufficient optimality conditions for the discretized optimal control problem:

$$\begin{aligned} a(\varphi_i, p^h) - \langle y^h, \varphi_i \rangle_{L^2} &= -\langle \widehat{y}, \varphi_i \rangle_{L^2} && i = 1, \dots, n, \\ -b(\psi_i, p^h) + \alpha \langle u^h, \psi \rangle_{L^2} &= 0 && i = 1, \dots, m, \\ a(y^h, \varphi_i) + b(u^h, \varphi_i) &= l(\varphi_i) && i = 1, \dots, n. \end{aligned}$$

- Discretized optimal control problem:

$$\min \frac{1}{2}\|y^h - \widehat{y}\|^2_{L^2} + \frac{\alpha}{2}\|u^h\|^2_{L^2},$$
$$\text{s.t. } a(y^h, \varphi_i) + b(u^h, \varphi_i) = l(\varphi_i) \quad i = 1, \dots, n.$$

- Necessary and sufficient optimality conditions for the discretized optimal control problem:

$$\begin{aligned} a(\varphi_i, p^h) - \langle y^h, \varphi_i\rangle_{L^2} &= -\langle \widehat{y}, \varphi_i\rangle_{L^2} && i = 1, \dots, n, \\ -b(\psi_i, p^h) + \alpha\langle u^h, \psi\rangle_{L^2} &= 0 && i = 1, \dots, m, \\ a(y^h, \varphi_i) + b(u^h, \varphi_i) &= l(\varphi_i) && i = 1, \dots, n. \end{aligned}$$

- The necessary and sufficient optimality conditions for the discretized optimal control problem are identical to the conditions that arise if we discretize the optimality conditions directly, i.e, replace $y$ by $y^h = \sum_{i=1}^n y_i\varphi_i$, replace $u$ by $u^h = \sum_{i=1}^m u_i\psi_i$, replace $p$ by $p^h = \sum_{i=1}^n p_i\varphi_i$, and require that $y^h, u^h, p^h$ satisfy the optimality conditions for all $v = \varphi_i$, $i = 1, \dots, n$, $w = \psi_i$, $i = 1, \dots, m$.

If we define $\vec{y} = (y_1, \ldots, y_n)^T$, $\vec{u} = (u_1, \ldots, u_m)^T$, $\vec{p} = (p_1, \ldots, p_n)^T$, matrices $\mathbb{A} \in \mathbb{R}^{n \times n}$, $\mathbb{B} \in \mathbb{R}^{n \times m}$, $\mathbb{A} \in \mathbb{R}^{n \times n}$, with entires

$$\mathbb{A}_{ij} = a(\varphi_j, \varphi_i), \quad \mathbb{B}_{ij} = b(\psi_j, \varphi_i), \quad \text{etc.},$$

and vectors $\vec{c} = (\langle \varphi_1, \widehat{y} \rangle_{L^2}, \ldots, \langle \varphi_n, \widehat{y} \rangle_{L^2})^T$, $\vec{f} = (l(\varphi_1), \ldots, l(\varphi_n))^T$, then the discretized optimal control problem can be written as

$$\begin{aligned} &\min \frac{1}{2} \vec{y}^T \mathbb{M} \vec{y} - \vec{y}^T \vec{c} + \frac{\alpha}{2} \vec{u}^T \mathbb{Q} \vec{u} \\ &\text{s.t. } \mathbb{A}\vec{y} + \mathbb{B}\vec{u} = \vec{f}. \end{aligned}$$

If we define $\vec{y} = (y_1, \ldots, y_n)^T$, $\vec{u} = (u_1, \ldots, u_m)^T$, $\vec{p} = (p_1, \ldots, p_n)^T$, matrices $\mathbb{A} \in \mathbb{R}^{n\times n}$, $\mathbb{B} \in \mathbb{R}^{n\times m}$, $\mathbb{A} \in \mathbb{R}^{n\times n}$, with entires

$$\mathbb{A}_{ij} = a(\varphi_j, \varphi_i), \quad \mathbb{B}_{ij} = b(\psi_j, \varphi_i), \quad \text{etc.},$$

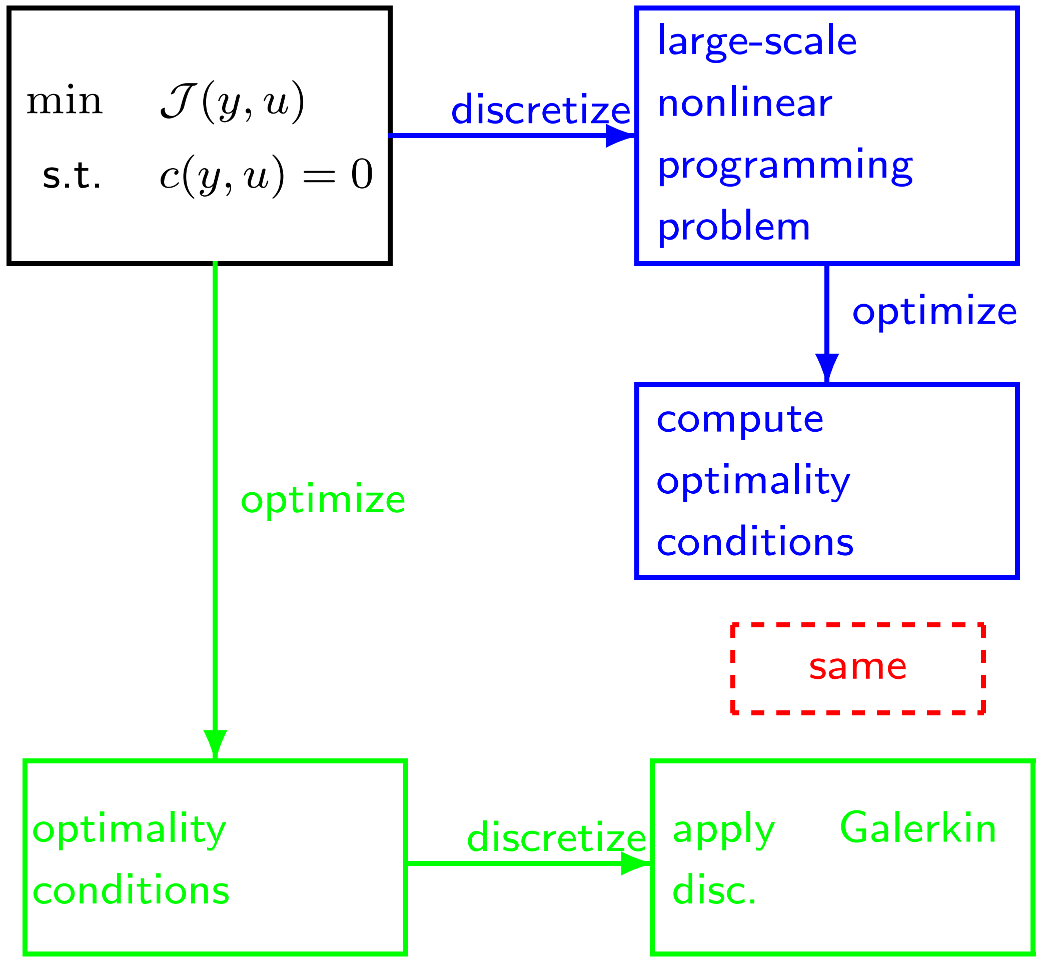
and vectors $\vec{c} = (\langle \varphi_1, \widehat{y}\rangle_{L^2}, \ldots, \langle \varphi_n, \widehat{y}\rangle_{L^2})^T$, $\vec{f} = (l(\varphi_1), \ldots, l(\varphi_n))^T$, then the discretized optimal control problem can be written as

$$\begin{aligned} &\min \frac{1}{2}\vec{y}^T\mathbb{M}\vec{y} - \vec{y}^T\vec{c} + \frac{\alpha}{2}\vec{u}^T\mathbb{Q}\vec{u} \\ &\text{s.t. } \mathbb{A}\vec{y} + \mathbb{B}\vec{u} = \vec{f}. \end{aligned}$$

The necessary and sufficient optimality conditions are given by

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}^T \\ 0 & \mathbb{Q} & \mathbb{B}^T \\ \mathbb{A} & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ -\vec{p} \end{pmatrix} = \begin{pmatrix} \vec{c} \\ 0 \\ \vec{f} \end{pmatrix}.$$

This is just the matrix version of the optimality system on the previous slide.
Systems of this type are called KKT systems (Karush-Kuhn-Tucker systems).


min $\mathcal{J}(y,u)$
s.t. $c(y,u) = 0$
discretize
large-scale nonlinear programming problem
optimize
compute optimality conditions
same
optimize
optimality conditions
discretize
apply Galerkin disc.

$$
\begin{aligned}
a(v,p) - \langle y, v\rangle_{L^2} &= -\langle \widehat{y}, v\rangle_{L^2} && \forall v \in Y,\\
-b(w,p) + \alpha\langle u, w\rangle_{L^2} &= 0 && \forall w \in U,\\
a(y,v) + b(u,v) &= l(v) && \forall v \in Y
\end{aligned}
$$

can be written as

$$\mathbf{K}\mathbf{x} = \mathbf{r} \quad \text{in } \mathcal{X}'$$

where $\mathcal{X} = Y \times U \times Y$ and $\mathbf{K} \in \mathcal{L}(\mathcal{X}, \mathcal{X}')$.

$$
\begin{aligned}
a(v,p) - \langle y, v\rangle_{L^2} &= -\langle \widehat{y}, v\rangle_{L^2} && \forall v \in Y,\\
-b(w,p) + \alpha\langle u, w\rangle_{L^2} &= 0 && \forall w \in U,\\
a(y,v) + b(u,v) &= l(v) && \forall v \in Y
\end{aligned}
$$

can be written as

$$\mathbf{K}\mathbf{x} = \mathbf{r} \quad \text{in } \mathcal{X}'$$

where $\mathcal{X} = Y \times U \times Y$ and $\mathbf{K} \in \mathcal{L}(\mathcal{X}, \mathcal{X}')$.

$$
\begin{aligned}
a(\varphi_i, p^h) - \langle y^h, \varphi_i\rangle_{L^2} &= -\langle \widehat{y}, \varphi_i\rangle_{L^2} && i = 1, \dots, n,\\
-b(\psi_i, p^h) + \alpha\langle u^h, \psi\rangle_{L^2} &= 0 && i = 1, \dots, m,\\
a(y^h, \varphi_i) + b(u^h, \varphi_i) &= l(\varphi_i) && i = 1, \dots, n,
\end{aligned}
$$

can be written as

$$\mathbf{K}_h\mathbf{x}_h = \mathbf{r}_h \quad \text{in } \mathcal{X}_h'$$

where $\mathcal{X}_h = Y^h \times U^h \times Y^h$ and $\mathbf{K}_h \in \mathcal{L}(\mathcal{X}_h, \mathcal{X}_h')$.

- Consider

$$\mathbf{K}\mathbf{x} = \mathbf{r}$$

and

$$\mathbf{K}_h\mathbf{x}_h = \mathbf{r}_h,$$

- Let $\mathbf{R}_h : \mathcal{X} \to \mathcal{X}_h$ be a restriction operator.

- Subtract $\mathbf{K}_h\mathbf{R}_h(\mathbf{x})$ from $\mathbf{K}_h\mathbf{x}_h = \mathbf{r}_h$,

$$\mathbf{K}_h(\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})) = \mathbf{r}_h - \mathbf{K}_h\mathbf{R}_h(\mathbf{x}),$$

to obtain the estimate

$$\|\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})\|_h \leq \|\mathbf{K}_h^{-1}\|_h \|\mathbf{r}_h - \mathbf{K}_h\mathbf{R}_h(\mathbf{x})\|_h.$$

Stability

$$\|\mathbf{K}_h^{-1}\|_h \leq \kappa \quad \text{for all } h$$

and consistency

$$\|\mathbf{r}_h - \mathbf{K}_h \mathbf{R}_h(\mathbf{x})\|_h = O(h^p)$$

imply

$$\|\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})\|_h \leq O(h^p).$$

Hence

$$\|\mathbf{x}_h - \mathbf{x}\|_h \leq \|\mathbf{x} - \mathbf{R}_h(\mathbf{x})\|_h + \underbrace{\|\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})\|_h}_{= O(h^p)}$$

If there exist $\alpha_1, \alpha_2, \beta, \gamma > 0$ with

- $a(y,y) \geq \alpha_1 \|y\|_Y^2$, $a(y,v) \leq \alpha_2 \|y\|_Y \|v\|_Y$, (true for model problem)
- $b(u,v) \leq \beta \|u\|_U \|v\|_Y$, (true for model problem)
- $\frac{1}{2}\|Su - \widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \geq \gamma \|u^h\|_{L^2}^2$, (of course true for model problem)

then there exists $\kappa > 0$ independent of $h$ such that

$$\|\mathbf{K}_h^{-1}\|_h \leq \kappa \quad \text{for all } h.$$

- Finite Elements

$$\begin{aligned} Y_h &= \{y_h \in Y \,:\, y_h|_T \in P_k(T) \text{ for all } T \in \mathcal{T}_h\}\,, \\ U_h &= \{u_h \in U \,:\, u_h|_T \in P_k(T) \text{ for all } T \in \mathcal{T}_h\}\,. \end{aligned}$$

- $\mathbf{R}_h(x) \in Y_h \times U_h \times Y_h$ interpolation of $x = (y, u, p)$.
- Consistency

$$\|\mathbf{r}_h - \mathbf{K}_h \mathbf{R}_h(\mathbf{x})\|_h \le C h^k (|y|_{k+1} + |p|_{k+1}).$$

- Convergence

$$\|y - y_h\|_{H^1} + \|u - u_h\|_{L^2} + \|p - p_h\|_{H^1} \le C h^k (|y|_{k+1} + |p|_{k+1}).$$

| $h$ | $\|y - y_h\|_{H^1}$ | | $\|u - u_h\|_{L^2}$ | | $\|p - p_h\|_{H^1}$ | |
|---|---|---|---|---|---|---|
| 1.00e-01 | 3.08e-02 | | 1.89e-02 | | 1.50e-02 | |
| 5.00e-02 | 1.55e-02 | 0.99 | 4.73e-03 | 2.00 | 7.52e-03 | 1.00 |
| 2.50e-02 | 7.76e-03 | 1.00 | 1.18e-03 | 2.00 | 3.76e-03 | 1.00 |
| 1.25e-02 | 3.88e-03 | 1.00 | 2.96e-04 | 2.00 | 1.88e-03 | 1.00 |
| 6.25e-03 | 1.94e-03 | 1.00 | 7.40e-05 | 2.00 | 9.40e-04 | 1.00 |

| $h$ | $\|y - y_h\|_{L^2}$ | | $\|u - u_h\|_{L^2}$ | | $\|p - p_h\|_{L^2}$ | |
|---|---|---|---|---|---|---|
| 1.00e-01 | 9.03e-04 | | 1.89e-02 | | 1.89e-04 | |
| 5.00e-02 | 2.30e-04 | 1.97 | 4.73e-03 | 2.00 | 4.73e-05 | 2.00 |
| 2.50e-02 | 5.78e-05 | 1.99 | 1.18e-03 | 2.00 | 1.18e-05 | 2.00 |
| 1.25e-02 | 1.45e-05 | 2.00 | 2.96e-04 | 2.00 | 2.96e-06 | 2.00 |
| 6.25e-03 | 3.62e-06 | 2.00 | 7.40e-05 | 2.00 | 7.40e-07 | 2.00 |

Higher convergence order can be established similar to the PDE case.

$$\min \frac{1}{2}\int_\Omega (y(x)-\widehat{y}(x))^2 dx + \frac{\alpha}{2}\int_\Omega u^2(x)dx$$

subject to

$$\begin{array}{rcll} -\epsilon\Delta y(x) + c(x)\cdot\nabla y(x) + r(x)y(x) & = & f(x)+u(x), & x\in\Omega, \\ y(x) & = & d(x), & x\in\Gamma_d, \\ \epsilon\dfrac{\partial}{\partial n}y(x) & = & g(x), & x\in\Gamma_n, \end{array}$$

where $f,\widehat{y}\in L^2(\Omega)$, $\alpha>0$,

$$\epsilon>0, \quad c\in\left(W^{1,\infty}(\Omega)\right)^2, \quad r\in L^\infty(\Omega),$$

$$r(x)-\tfrac{1}{2}\nabla\cdot c(x)\geq r_0>0 \text{ a.e. in } \Omega, \quad n\cdot c(x)\geq 0 \text{ on } \Gamma_n.$$

We are interested in the case $\epsilon \ll \|c(x)\|$.

$$-\epsilon \Delta y(x) + c(x) \cdot \nabla y(x) + r(x)y(x) = f(x) + u(x) \quad x \in \Omega,$$
$$y(x) = d(x) \quad x \in \Gamma_d, \qquad \epsilon \frac{\partial}{\partial n} y(x) = 0 \quad x \in \Gamma_n.$$

Weak form:

$$Y = \left\{ y \in H^1(\Omega) \ : \ y = d \text{ on } \Gamma_d \right\}, \quad V = \left\{ v \in H^1(\Omega) \ : \ y = 0 \text{ on } \Gamma_d \right\}.$$

Find $y \in Y$ such that

$$a(y, v) + b(u, v) = \langle f, v \rangle \quad \forall v \in V,$$

where

$$\begin{aligned} a(y, v) &= \int_\Omega \epsilon \nabla y(x) \cdot \nabla v(x) + c(x) \cdot \nabla y(x) v(x) + r(x) y(x) v(x) dx, \\ &= \int_\Omega \epsilon \nabla y \cdot \nabla v + \tfrac{1}{2} c \cdot \nabla y v - \tfrac{1}{2} c \cdot \nabla v y + (r - \tfrac{1}{2} \nabla \cdot c) y v dx + \int_{\Gamma_n} \tfrac{1}{2} (n \cdot c) y v dx, \end{aligned}$$

$$b(u, v) = -\int_\Omega u(x) v(x) dx, \qquad \langle f, v \rangle = \int_\Omega f(x) v(x) dx.$$

If

$$\epsilon > 0, \quad c \in \left(W^{1,\infty}(\Omega)\right)^2, \quad r \in L^\infty(\Omega),$$

$$r(x) - \tfrac{1}{2}\nabla \cdot c(x) \ge r_0 > 0 \text{ a.e. in } \Omega, \quad n \cdot c(x) \ge 0 \text{ on } \Gamma_n,$$

then $a$ is continuous on $V \times V$ and $V$–elliptic.

Thus, for given control $u \in L^2(\Omega)$, the state equation has a unique solution $y \in Y$.

Standard Galerkin Method:

$$
\begin{aligned}
Y_h &= \{y_h \in Y \;:\; y_h|_T \in P_k(T) \text{ for all } T \in \mathcal{T}_h\}\,, \\
V_h &= \{v_h \in V \;:\; v_h|_T \in P_k(T) \text{ for all } T \in \mathcal{T}_h\}\,.
\end{aligned}
$$

Find $y_h \in Y_h$ such that

$$
a(y_h, v_h) + b(u, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h. \qquad (*)
$$

The discretized state equation (*) has a unique solution.

If $h \gtrsim \epsilon / \|c\|_\infty$, solution often exhibits spurious oscillations.

Add stabilization term

$$
\begin{aligned}
& a(y_h, v_h) + b(u, v_h) + \sum_{T_e \in \mathcal{T}_h} \tau_e \langle -\epsilon \Delta y_h + c \cdot \nabla y_h + r y_h - u, \sigma(v_h) \rangle_{T_e} \\
= \quad & \langle f, v_h \rangle_{L^2} + \sum_{T_e \in \mathcal{T}_h} \tau_e \langle f, \sigma(v_h) \rangle_{T_e},
\end{aligned}
$$

where $\sigma(v_h) = c \cdot \nabla v_h$.

- For the solution $y$ of the state equation,

$$
\sum_{T_e \in \mathcal{T}_h} \tau_e \langle -\epsilon \Delta y + c \cdot \nabla y + r y - u, \sigma(v_h) \rangle_{T_e} = \sum_{T_e \in \mathcal{T}_h} \tau_e \langle f, \sigma(v_h) \rangle_{T_e}.
$$

- If $y_h = v_h$,

$$
\sum_{T_e \in \mathcal{T}_h} \tau_e \underbrace{\langle c \cdot \nabla v_h, c \cdot \nabla v_h \rangle_{T_e}}_{\text{diffusion}} + \langle -\epsilon \Delta v_h + r v_h, c \cdot \nabla v_h \rangle_{T_e}
$$

is added to the Galerkin bilinear form $a$.

Streamline upwind Petrov-Galerkin (SUPG) method (Brooks/Hughes 1979).

Stabilized weak form

$$a_h(y_h, v_h) + b_h(u_h, v_h) = \langle f, v_h \rangle_h \quad \forall v_h \in V_h,$$

where

$$\begin{aligned} a_h(y, v_h) &= a(y, v_h) + \sum_{T_e \in \mathcal{T}_h} \tau_e \langle -\epsilon \Delta y + c \cdot \nabla y + ry, \sigma(v_h) \rangle_{T_e}, \\ b_h(u, v_h) &= -\langle u, v_h \rangle_{L^2} - \sum_{T_e \in \mathcal{T}_h} \tau_e \langle u, \sigma(v_h) \rangle_{T_e}, \\ \langle f, v_h \rangle_h &= \langle f, v_h \rangle_{L^2} + \sum_{T_e \in \mathcal{T}_h} \tau_e \langle f, \sigma(v_h) \rangle_{T_e} \end{aligned}$$

and

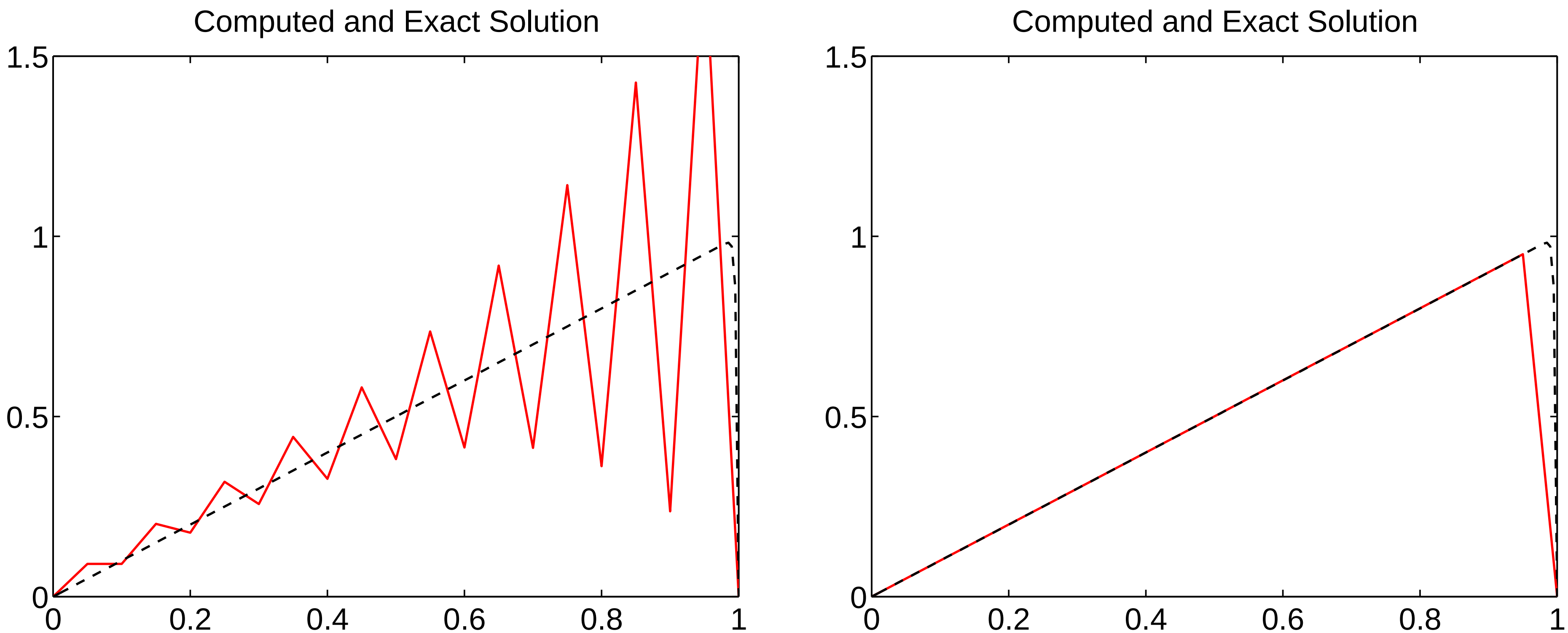
$$\sigma(v_h) = c \cdot \nabla v_h.$$

Choice of stabilization parameter $\tau_e$ depends on the mesh Péclet number

$$\mathrm{Pe}_e = \frac{\|c\|_{\infty, T_e} h_e}{2\epsilon}$$

## 1D example

$$-0.0025y''(x) + y'(x) = 1 \text{ on } (0,1), \qquad y(0) = y(1) = 0.$$

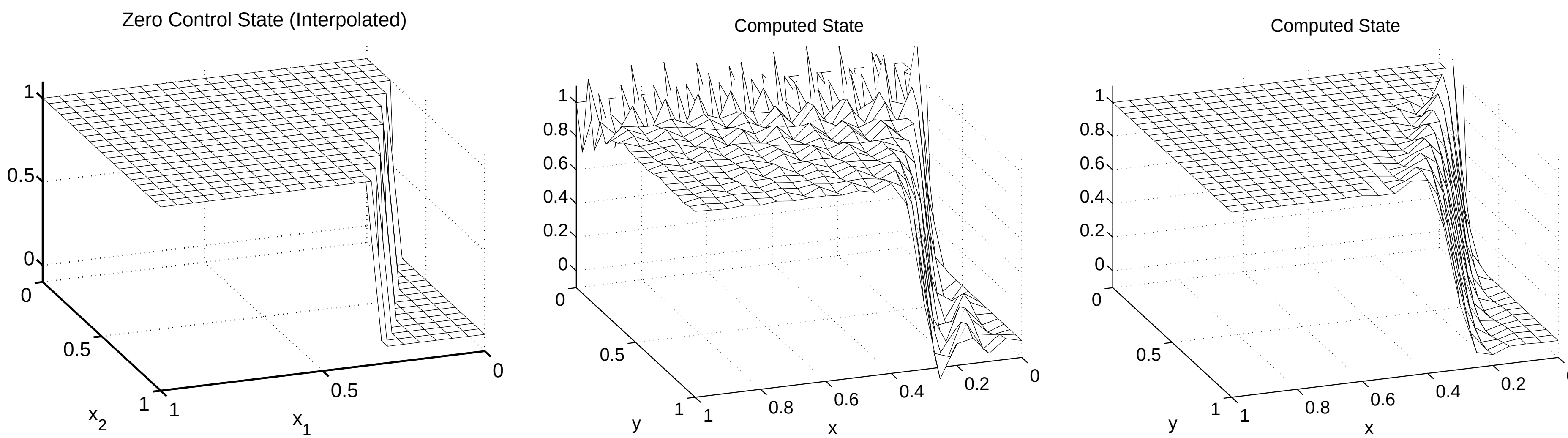



Galerkin $n_x = 20$ SUPG $n_x = 20$

2D example

$$
\begin{aligned}
-\epsilon\Delta y(x) + \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}\cdot\nabla y(x) &= 0 \quad \text{in } \Omega = (0,1)^2,\\
y(x) &= 1 \quad \text{on } \Gamma_1 = \big((0,1)\times\{0\}\big)\cup\big(\{0\}\times(0,0.2)\big),\\
y(x) &= 0 \quad \text{on } \Gamma_2 = \{0\}\times(0.2,1),\\
\frac{\partial}{\partial n} y(x) &= 0 \quad \text{on } \Gamma_3 = \partial\Omega\setminus(\Gamma_1\cup\Gamma_2),
\end{aligned}
$$

where $\epsilon = 10^{-6}$ and $\theta = 67.5^0$.



**Theorem.** If $\tau_e$ satisfies and

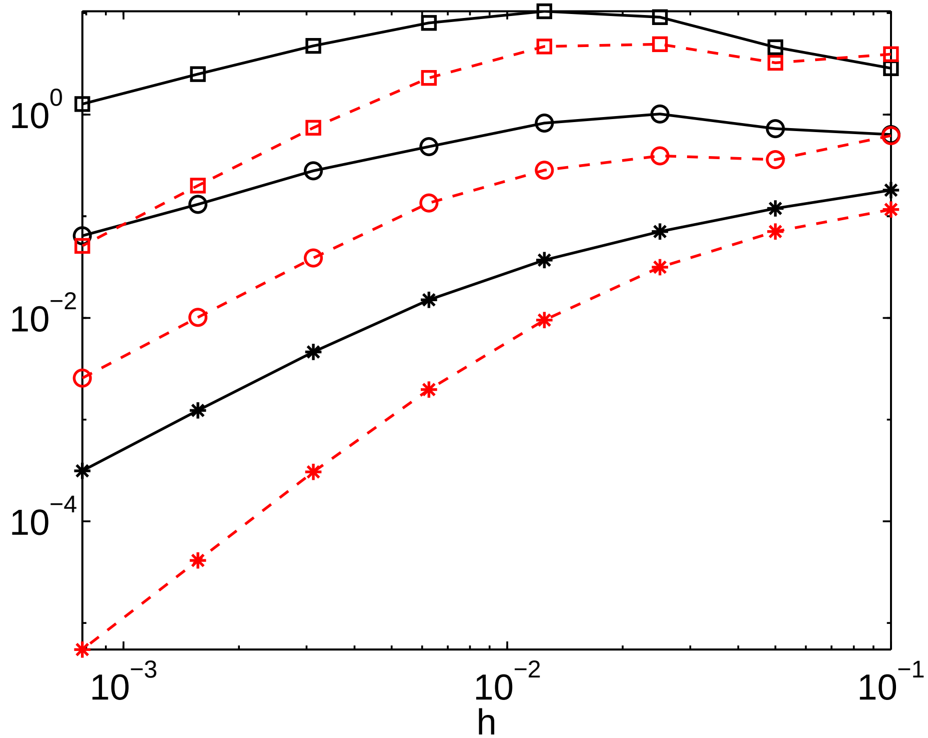
$$0 < \tau_e \le \min\left\{ \frac{h_e^2}{\epsilon \mu_{\mathrm{inv}}^2}, \frac{r_0}{\|r\|_{\infty,T_e}} \right\} \quad \text{and} \quad \tau_e = \begin{cases} \tau_1 \frac{h_e^2}{\epsilon}, & \mathrm{Pe}_e \le 1, \\ \tau_2 h_e, & \mathrm{Pe}_e > 1, \end{cases}$$

then the solution $y_h$ of the discretized state equation obeys

$$\|y - y_h\|_{SD} \le C h^k (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) |y|_{k+1},$$

where $\|v\|_{SD}^2 \stackrel{\mathrm{def}}{=} \epsilon |v|_1^2 + r\|v\|^2 + \sum_{T_e \in \mathcal{T}_h} \tau_e \|c \cdot \nabla v\|_{T_e}^2$.

Error in PDE solution, example 1 ($\mathrm{Pe}_e = 1$ for $h = 5 * 10^{-3}$).




Linear elements —-, quadratic elements - - -

$* = \|y_h - y_{\mathrm{ex}}\|_{L^2}$, $\square = \|y_h - y_{\mathrm{ex}}\|_{H^1}$, $\circ = \|y_h - y_{\mathrm{ex}}\|_{SD}$.

$$\min \frac{1}{2}\int_\Omega (y(x)-\widehat{y}(x))^2 dx + \frac{\alpha}{2}\int_\Omega u^2(x)dx$$

subject to

$$\begin{aligned}
-\epsilon\Delta y(x) + c(x)\cdot\nabla y(x) + r(x)y(x) &= f(x)+u(x), & x\in\Omega,\\
y(x) &= d(x), & x\in\Gamma_d,\\
\epsilon\frac{\partial}{\partial n}y(x) &= 0, & x\in\Gamma_n,
\end{aligned}$$

where $f, \widehat{y} \in L^2(\Omega)$, $\alpha > 0$,

$$\epsilon > 0, \quad c \in \left(W^{1,\infty}(\Omega)\right)^2, \quad r \in L^\infty(\Omega),$$

$$r(x) - \tfrac{1}{2}\nabla\cdot c(x) \geq r > 0 \text{ a.e. in } \Omega, \quad n\cdot c(x) \geq 0 \text{ on } \Gamma_n.$$

Optimal Control Problem

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|y-\widehat{y}\|^2+\frac{\alpha}{2}\|u\|^2, \\ \text{subject to} \quad & a(y,v)+b(u,v)=\langle f,v\rangle_{L^2} \quad \forall v\in V. \end{aligned}$$

Lagrangian

$$L(y,u,p)=\frac{1}{2}\|y-\widehat{y}\|^2+\frac{\alpha}{2}\|u\|^2-a(y,p)-b(u,p)+\langle f,p\rangle.$$

Optimality Conditions:

Adjoint equation: $a(\psi,p)=\langle y-\hat{y},\psi\rangle_{L^2} \quad \forall\psi\in V$.

Gradient equation: $-b(w,p)+\alpha\langle u,w\rangle_{L^2}=0 \quad \forall w\in U$.

State equation: $a(y,v)+b(u,v)=\langle f,v\rangle \quad \forall v\in V$.

Adjoint equation: $a(\psi, p) = \langle y - \hat{y}, \psi \rangle_{L^2} \quad \forall \psi \in V$.

Gradient equation: $-b(w, p) + \alpha \langle u, w \rangle_{L^2} = 0 \quad \forall w \in U$.

State equation: $a(y, v) + b(u, v) = \langle f, v \rangle \quad \forall v \in V$.

Adjoint equation:

$$
\begin{aligned}
-\epsilon \Delta p(x) - c(x) \cdot \nabla p(x) + (r(x) - \nabla \cdot c(x)) p(x) \\
= y(x) - \hat{y}(x), \quad x \in \Omega,
\end{aligned}
$$

$$
p(x) = 0 \quad x \in \Gamma_d, \qquad \epsilon \frac{\partial}{\partial n} p(x) = -c(x) \cdot n\, p(x) \quad x \in \Gamma_n.
$$

Gradient equation: $\quad p(x) + \alpha u(x) = 0,\ x \in \Omega$.

State equation:

$$
\begin{aligned}
-\epsilon \Delta y(x) + c(x) \cdot \nabla y(x) + r(x) y(x) = f(x) + u(x), \quad x \in \Omega, \\
y(x) = d(x), \quad x \in \Gamma_d, \qquad \epsilon \frac{\partial}{\partial n} y(x) = 0, \quad x \in \Gamma_n.
\end{aligned}
$$

- Discretize–then–optimize.
  Discretize the optimal control problem, then apply finite dimensional optimization.

- Optimize–then–discretize.
  Formulate the optimality conditions on the PDE level, then discretize the PDEs in the optimality conditions individually.

Discretized optimal control problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|y_h - \widehat{y}\|^2 + \frac{\alpha}{2}\|u_h\|^2, \\ \text{subject to} & a_h^{\mathsf{s}}(y_h, v_h) + b_h^{\mathsf{s}}(u_h, v_h) = \langle f, v_h \rangle_h^{\mathsf{s}} \quad \forall v_h \in V_h, \end{array}$$

where

$$\begin{aligned} a_h^{\mathsf{s}}(y, v_h) &= a(y, v_h) + \sum_{T_e \in \mathcal{T}_h} \tau_e^{\mathsf{s}} \langle -\epsilon \Delta y + c \cdot \nabla y + ry, \sigma^{\mathsf{s}}(v_h) \rangle_{T_e}, \\ b_h^{\mathsf{s}}(u, v_h) &= -\langle u, v_h \rangle_{L^2} - \sum_{T_e \in \mathcal{T}_h} \tau_e^{\mathsf{s}} \langle u, \sigma^{\mathsf{s}}(v_h) \rangle_{T_e}, \\ \langle f, v_h \rangle_h^{\mathsf{s}} &= \langle f, v_h \rangle_{L^2} + \sum_{T_e \in \mathcal{T}_h} \tau_e^{\mathsf{s}} \langle f, \sigma^{\mathsf{s}}(v_h) \rangle_{T_e} \end{aligned}$$

and

$$\sigma^{\mathsf{s}}(v_h) = c \cdot \nabla v_h.$$

Optimality Conditions for Discretized System

Discrete adjoint equation: $a_h^{\mathsf{s}}(v_h, p_h) = \langle y_h - \hat{y}, v_h \rangle_{L^2} \quad \forall v_h \in V_h.$

Discrete gradient equations: $-b_h^{\mathsf{s}}(w_h, p_h) + \alpha \langle u_h, w_h \rangle_{L^2} = 0 \qquad \forall w_h \in U_h.$

Discretized state equations: $a_h^{\mathsf{s}}(y_h, v_h) + b_h^{\mathsf{s}}(u_h, v_h) = \langle f, v_h \rangle_h^{\mathsf{s}} \quad \forall v_h \in V_h.$

Recall discrete adjoint equations

$$a_h^{\mathsf{s}}(v_h, p_h) = \langle y_h - \hat{y}, v_h \rangle_{L^2} \quad \forall v_h \in V_h.$$

Compare stabilization in the discrete adjoint equation

$$\begin{aligned} a(\psi_h, p_h) + \sum_{T_e \in \mathcal{T}_h} \tau_e \langle -\epsilon\Delta\psi_h + c \cdot \nabla\psi_h + r\psi_h, c \cdot \nabla p_h \rangle_{T_e} \\ = \quad \langle y_h - \hat{y}, \psi_h \rangle_{L^2} + \sum_{T_e \in \mathcal{T}_h} 0 \end{aligned}$$

with the adjoint equation

$$-\epsilon\Delta p(x) - c(x) \cdot \nabla p(x) + (r(x) - \nabla \cdot c(x))p(x) = y(x) - \hat{y}(x), \; x \in \Omega.$$

Stabilization adds 'the right' amount of diffusion, but is not strongly consistent.

Adjoint equation:

$$-\epsilon\Delta p(x) - c(x)\cdot\nabla p(x) + (r(x) - \nabla\cdot c(x))p(x) = y(x) - \hat{y}(x),\ x\in\Omega,$$
$$p(x) = 0 \quad x\in\Gamma_d, \qquad \epsilon\frac{\partial}{\partial n}p(x) = -c(x)\cdot n\, p(x) \quad x\in\Gamma_n$$

Apply SUPG method to adjoint equation

$$a_h^{\mathsf{a}}(\psi_h, p) = \langle y - \hat{y}, \psi_h\rangle_h^{\mathsf{a}} \quad \forall\psi_h\in V_h,$$

where

$$\begin{aligned} a_h^{\mathsf{a}}(\psi_h, p) &= a(\psi_h, p) + \sum_{T_e\in\mathcal{T}_h}\tau_e^{\mathsf{a}}\langle -\epsilon\Delta p - c\cdot\nabla p + (r - \nabla\cdot c)p, \sigma^{\mathsf{a}}(\psi_h)\rangle_{T_e}, \\ \langle y - \hat{y}, \psi_h\rangle_h^{\mathsf{a}} &= \langle y - \hat{y}, \psi_h\rangle_{L^2} + \sum_{T_e\in\mathcal{T}_h}\tau_e^{\mathsf{a}}\langle y - \hat{y}, \sigma^{\mathsf{a}}(\psi_h)\rangle_{T_e} \end{aligned}$$

and $\sigma^{\mathsf{a}}(\psi_h) = -c\cdot\nabla\psi_h$.

Gradient equation:

$$p(x) + \alpha u(x) = 0, \quad x \in \Omega.$$

Discretized gradient equation

$$-b(w_h, p_h) + \alpha \langle u_h, w_h \rangle_{L^2} = 0 \quad \forall w_h \in U_h,$$

where (as before)

$$b(w_h, p_h) = -\langle w_h, p_h \rangle_{L^2}.$$

State equation:

Apply SUPG as before.

Discretized adjoint equations: $a_h^{\mathsf{a}}(\psi_h, p_h) = \langle y_h - \hat{y}, \psi_h \rangle_h^{\mathsf{a}} \quad \forall \psi_h \in V_h$.

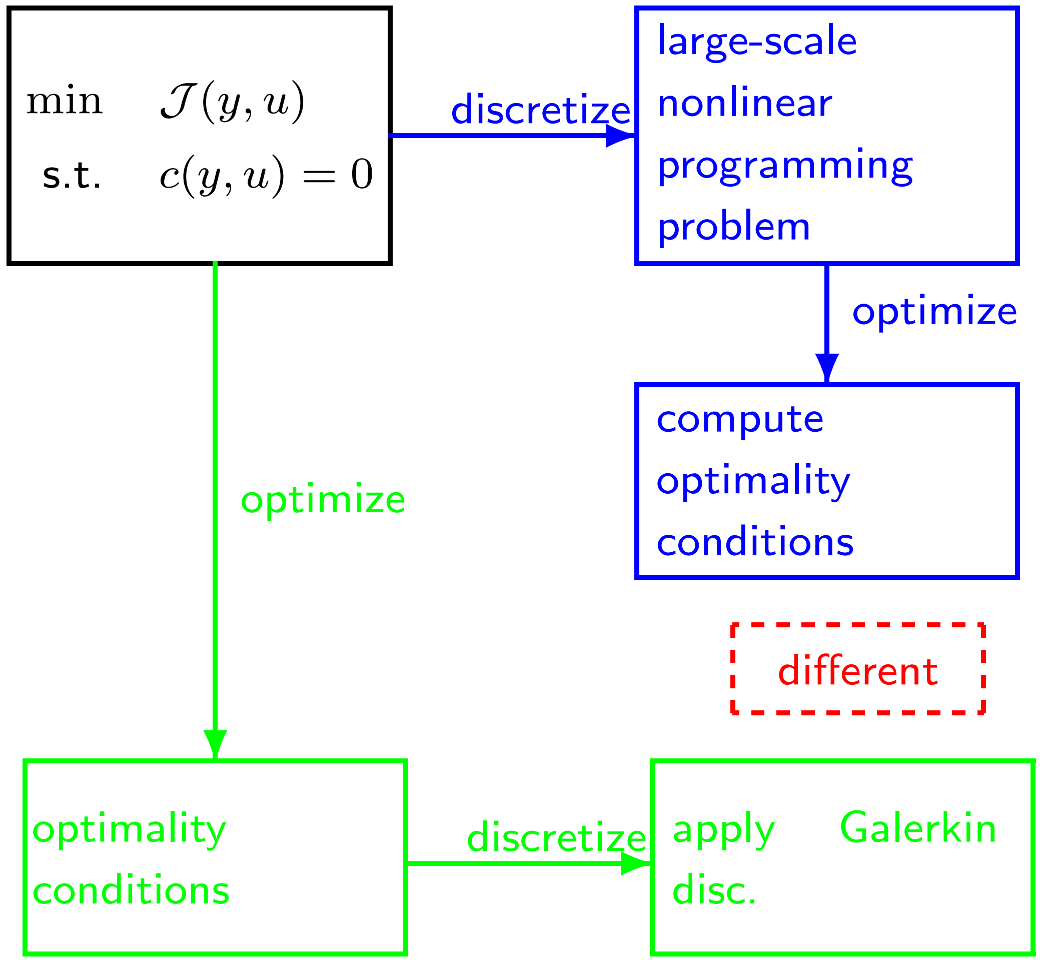
Discretized gradient equation: $-b(w_h, p_h) + \alpha \langle u_h, w_h \rangle_{L^2} = 0 \quad \forall w_h \in U_h$.

Discretized state equation: $a_h^{\mathsf{s}}(y_h, v_h) + b_h^{\mathsf{s}}(u_h, v_h) = \langle f, v_h \rangle_h^{\mathsf{s}} \quad \forall v_h \in V_h$.

Note

$$\begin{aligned} a_h^{\mathsf{s}}(y_h, p_h) &\neq a_h^{\mathsf{a}}(y_h, p_h), \\ b_h^{\mathsf{s}}(u_h, v_h) &\neq b(u_h, v_h), \\ \langle y_h, \psi_h \rangle_h^{\mathsf{a}} &\neq \langle \psi_h, y_h \rangle_h^{\mathsf{a}}. \end{aligned}$$

The discretized optimality system leads to a nonsymmetric linear system.
Non-symmetry is the smaller the smaller the stabilization parameters $\tau_e^{\mathsf{s}}$ and $\tau_e^{\mathsf{a}}$.


min $\mathcal{J}(y,u)$
s.t. $c(y,u) = 0$
discretize
large-scale nonlinear programming problem
optimize
compute optimality conditions
different
optimize
optimality conditions
discretize
apply Galerkin disc.

We proceed exactly as before:

- Write the optimality conditions as

$$\mathbf{K}\mathbf{x} = \mathbf{r},$$
$$\mathbf{K}_h\mathbf{x}_h = \mathbf{r}_h,$$

- Choose restriction operator $\mathbf{R}_h : \mathcal{X} \to \mathcal{X}_h$.
- Subtract $\mathbf{K}_h\mathbf{R}_h(\mathbf{x})$ from $\mathbf{K}_h\mathbf{x}_h = \mathbf{r}_h$,

$$\mathbf{K}_h(\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})) = \mathbf{r}_h - \mathbf{K}_h\mathbf{R}_h(\mathbf{x}),$$

  to obtain the estimate

$$\begin{aligned} \|\mathbf{x}_h - \mathbf{x}\|_h &\le \|\mathbf{x} - \mathbf{R}_h(\mathbf{x})\|_h + \|\mathbf{x}_h - \mathbf{R}_h(\mathbf{x})\|_h \\ &\le \underbrace{\|\mathbf{x} - \mathbf{R}_h(\mathbf{x})\|_h}_{= O(h^q)?} + \underbrace{\|\mathbf{K}_h^{-1}\|_h}_{\le \kappa?} \underbrace{\|\mathbf{r}_h - \mathbf{K}_h\mathbf{R}_h(\mathbf{x})\|_h}_{= O(h^q)?} \end{aligned}$$

Suppose there exist $\alpha_1, \alpha_2, \beta, \gamma > 0$ with

- $a(y,y) \geq \alpha_1 \|y\|_Y^2$, $a(y,v) \leq \alpha_2 \|y\|_Y \|v\|_Y$, (true for model problem)
- $b(u,v) \leq \beta \|u\|_U \|v\|_Y$, (true for model problem)
- $\frac{1}{2}\|Su - \widehat{y}\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \geq \gamma \|u\|_{L^2}^2$ (of course true for model problem).

then there exists $\kappa > 0$ independent of $h$ such that

$$\|\mathbf{K}_h^{-1}\|_h \leq \kappa \quad \text{for all } h.$$

**Discretize–Then–Optimize:**
**There exists $\kappa$ such that $\|\mathbf{K}_h^{-1}\|_h < \kappa$ for all $h$.**

**Optimize–Then–Discretize:**
**There exist $\kappa$, $h_0$ $(= h_0(\alpha, c, r))$ such that $\|\mathbf{K}_h^{-1}\|_h < \kappa$ for all $h \leq h_0$.**

Let

$$0 < \tau_e^{\mathsf{s,a}} \le \min\left\{ \frac{h_e^2}{\epsilon \mu_{\text{inv}}^2}, \frac{r_0}{\|r\|_{\infty,T_e}}, \frac{r_0}{\|r - \nabla \cdot c\|_{\infty,T_e}} \right\} \text{ and } \tau_e^{\mathsf{s,a}} = \begin{cases} \tau_1 \frac{h_e^2}{\epsilon}, & \mathrm{Pe}_e \le 1, \\ \tau_2 h_e, & \mathrm{Pe}_e > 1, \end{cases}$$

## Discretize–then-optimize

$$\begin{aligned} &\|\mathbf{r}_h - \mathbf{K}_h \mathbf{R}_h(\mathbf{x})\|_h \\ &\le C \begin{cases} (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) h^k |y|_{k+1} + h\epsilon^{-\frac{1}{2}} \|\nabla p^I\| + h^{k+1}|y|_{k+1}, & \mathrm{Pe}_e \le 1, \\ (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) h^k |y|_{k+1} + (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) \|\nabla p^I\| + h^{k+1}|y|_{k+1}, & \mathrm{Pe}_e > 1, \end{cases} \end{aligned}$$

## Optimize–then–discretize

$$\|\mathbf{r}_h - \mathbf{K}_h \mathbf{R}_h(\mathbf{x})\|_h \le C h^k (\epsilon^{1/2} + h^{1/2})(|y|_{k+1} + |p|_{k+1}).$$

Let

$$0 < \tau_e^{\mathsf{s,a}} \le \min\left\{ \frac{h_e^2}{\epsilon \mu_{\text{inv}}^2}, \frac{r_0}{\|r\|_{\infty,T_e}}, \frac{r_0}{\|r - \nabla \cdot c\|_{\infty,T_e}} \right\} \text{ and } \tau_e^{\mathsf{s,a}} = \begin{cases} \tau_1 \frac{h_e^2}{\epsilon}, & \mathrm{Pe}_e \le 1, \\ \tau_2 h_e, & \mathrm{Pe}_e > 1, \end{cases}$$

Discretize–then-optimize

$$\begin{aligned} &\|y - y_h\|_{SD} + \|u - u_h\|_{L^2} + \|p - p_h\|_{SD} \\ &\le C \begin{cases} (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) h^k |y|_{k+1} + h\epsilon^{-\frac{1}{2}} \|\nabla p^I\| + h^{k+1} |y|_{k+1}, & \mathrm{Pe}_e \le 1, \\ (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) h^k |y|_{k+1} + (\epsilon^{\frac{1}{2}} + h^{\frac{1}{2}}) \|\nabla p^I\| + h^{k+1} |y|_{k+1}, & \mathrm{Pe}_e > 1, \end{cases} \end{aligned}$$

Optimize–then–discretize

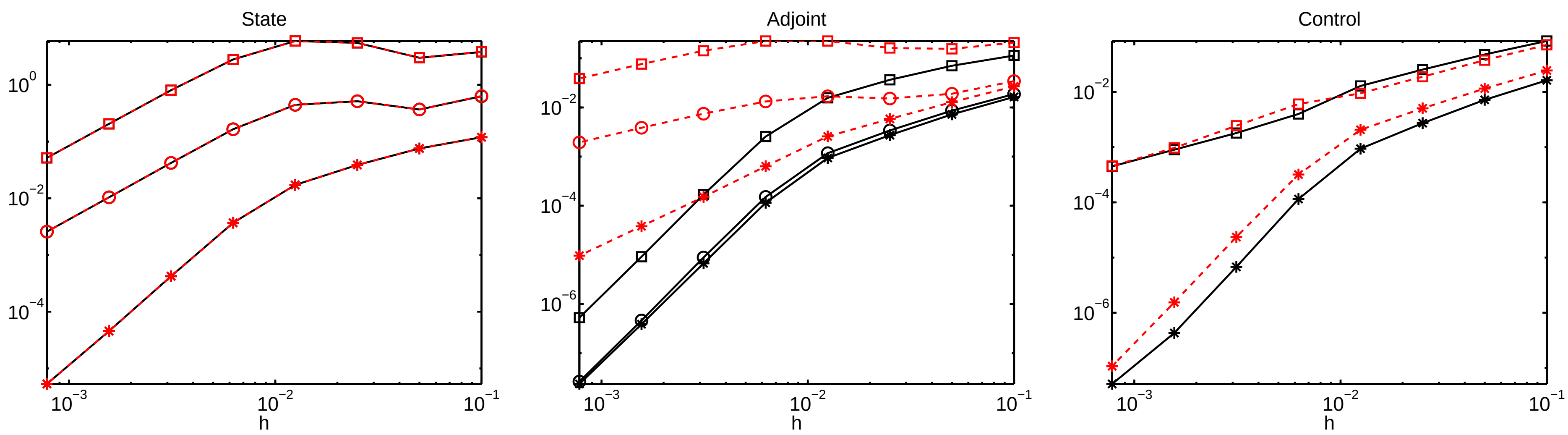
$$\|y - y_h\|_{SD} + \|u - u_h\|_{L^2} + \|p - p_h\|_{SD} \le C h^k (\epsilon^{1/2} + h^{1/2})(|y|_{k+1} + |p|_{k+1}).$$

State equation is

$$-0.0025y''(x) + y'(x) = f(x) + u(x) \text{ on } (0,1), \quad y(0) = y(1) = 0.$$

We use $\alpha = 1$. The solution to the optimal control problem is

$$y_{\text{ex}}(x) = -\frac{\exp(\frac{x-1}{\epsilon}) - \exp(-\frac{1}{\epsilon})}{1 - \exp(-\frac{1}{\epsilon})}, \quad u_{\text{ex}}(x) = x(x-1), \quad p_{\text{ex}}(x) = \alpha u_{\text{ex}}(x).$$




Discretize–then–optimize - - -, optimize–then–discretize ——

$$* = \|\cdot\|_{L^2}, \quad \square = \|\cdot\|_{H^1}, \quad \circ = \|\cdot\|_{SD}.$$

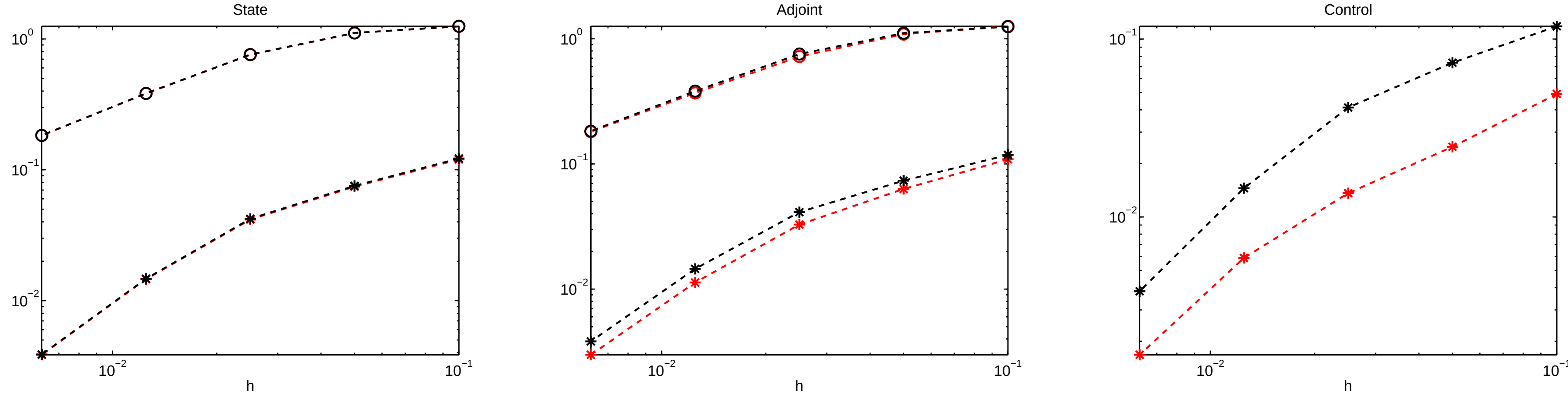
State equation

$$-\epsilon\Delta y(x) + \mathbf{c}(x)\cdot\nabla y(x) = u(x) \text{ in } \Omega = (0,1)^2,$$
$$y(x) = y_{\text{ex}}(x) \text{ on } \partial\Omega$$

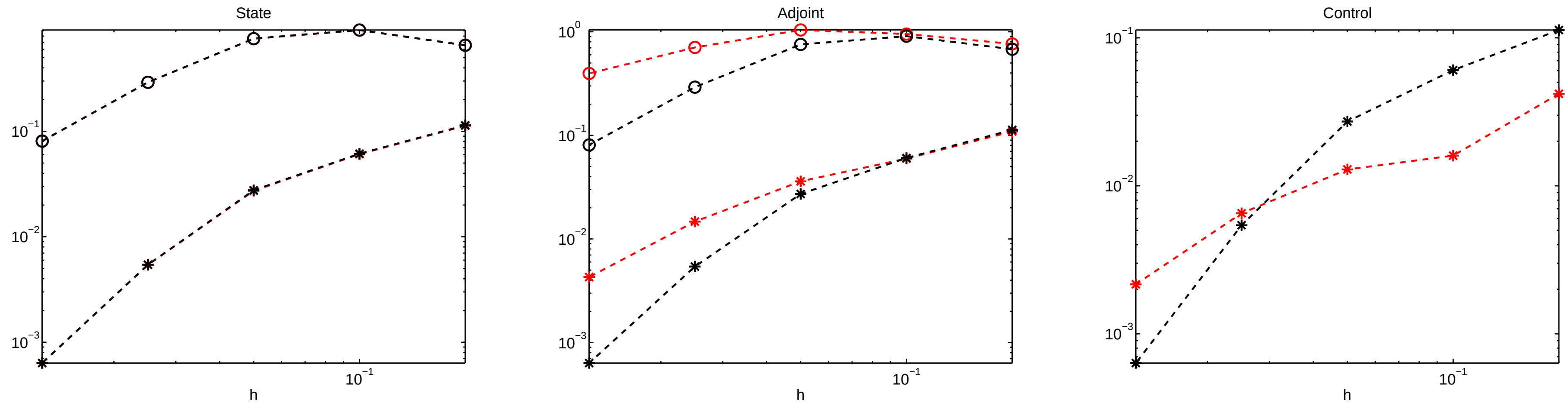
where $\epsilon = 10^{-2}$, $\theta = 45^0$.

Regularization parameter $\alpha = 1$.

## $k = \ell = m = 1$

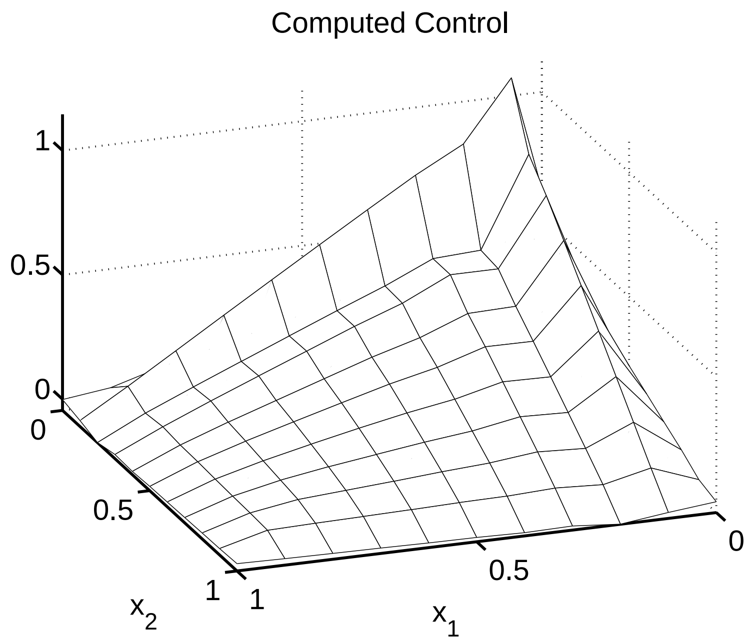



## $k = \ell = m = 2$

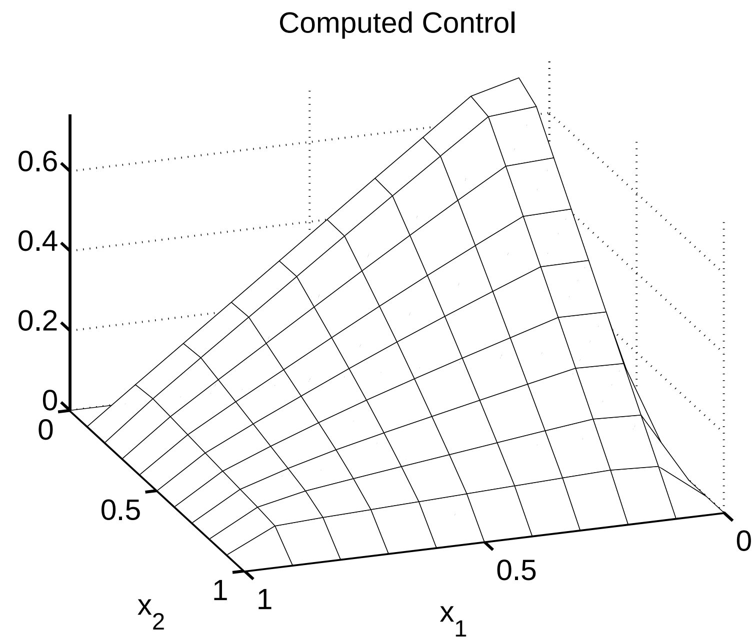



Discretize–optimize - - -, optimize–discretize ——, $* = \|\cdot\|_{L^2}$, $\circ = \|\cdot\|_{SD}$.

## Linear Finite Elements

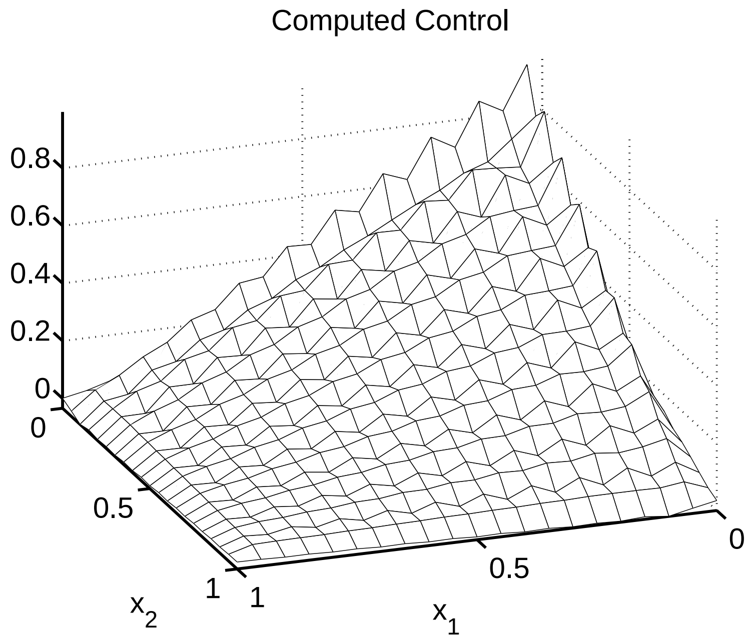



discretize-then-optimize

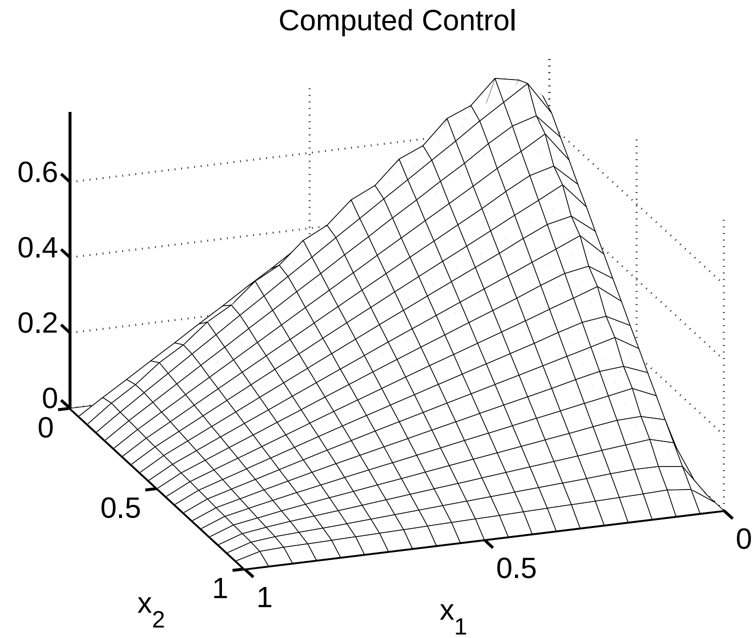



discretize-then-optimize

## Quadratic Finite Elements




discretize-then-optimize




optimize-then-discretize

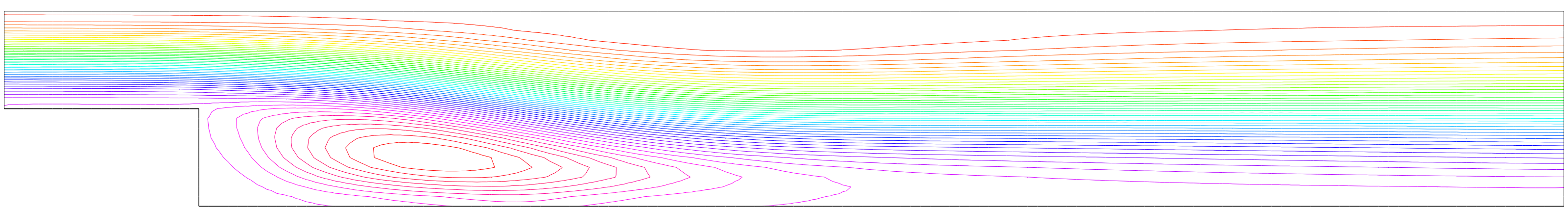
$$\min \frac{1}{2}\int_{\Omega_{\text{obs}}} |\nabla\times\mathbf{u}|^2 d\Omega + \frac{\alpha}{2}\int_{\Gamma_c} |\mathbf{g}|^2 d\Gamma_c,$$

subject to

$$\begin{aligned}(\mathbf{a}\cdot\nabla)\mathbf{u} - \nabla\cdot[-p\mathbf{I} + \mu(\nabla\mathbf{u} + \nabla\mathbf{u}^T)] &= 0 \quad \text{in} \quad \Omega,\\ \nabla\cdot\mathbf{u} &= 0 \quad \text{in} \quad \Omega,\end{aligned}$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } \Gamma_c, \qquad \mathbf{u} = \mathbf{u}_{\text{in}} \quad \text{on } \Gamma_{\text{in}}, \qquad \mathbf{u} = \mathbf{0} \quad \text{on } \partial\Omega \setminus (\Gamma_c \cup \Gamma_{\text{in}} \cup \Gamma_{\text{out}}),$$

$$\mathbf{n}\cdot[-p\mathbf{I} + \mu(\nabla\mathbf{u} + \nabla\mathbf{u}^T)] = \mathbf{0} \quad \text{on } \Gamma_{\text{out}}.$$



**Note change of notation:** $\mathbf{u}, p$ (velocities, pressure) states, $\mathbf{g}$ control, $\boldsymbol{\lambda}, \theta$ adjoint variables.

Adjoint equation:

$$-(\mathbf{a}\cdot\boldsymbol{\nabla})\boldsymbol{\lambda}_{\mathbf{u}} - (\boldsymbol{\nabla}\cdot\mathbf{a})\boldsymbol{\lambda} - \boldsymbol{\nabla}[-\theta\mathbf{I} + \mu(\boldsymbol{\nabla}\boldsymbol{\lambda} + \boldsymbol{\nabla}\boldsymbol{\lambda}^T)] = (\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{u})|_{\Omega_{\text{obs}}} \quad \text{in}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\lambda} = 0, \quad \text{in}$$

$$\boldsymbol{\lambda} = \mathbf{0} \quad \text{on } \Gamma_c, \qquad \boldsymbol{\lambda} = \mathbf{0} \quad \text{on } \Gamma_{\text{in}}, \qquad \boldsymbol{\lambda} = \mathbf{0} \quad \text{on } \partial\Omega \setminus (\Gamma_c \cup \Gamma_{\text{in}} \cup \Gamma_{\text{out}}),$$

$$\mathbf{n}\cdot[-\theta\mathbf{I} + \mu(\boldsymbol{\nabla}\boldsymbol{\lambda} + \boldsymbol{\nabla}\boldsymbol{\lambda}^T)] = \mathbf{0} \quad \text{on } \Gamma_{\text{out}}.$$

Gradient equation:

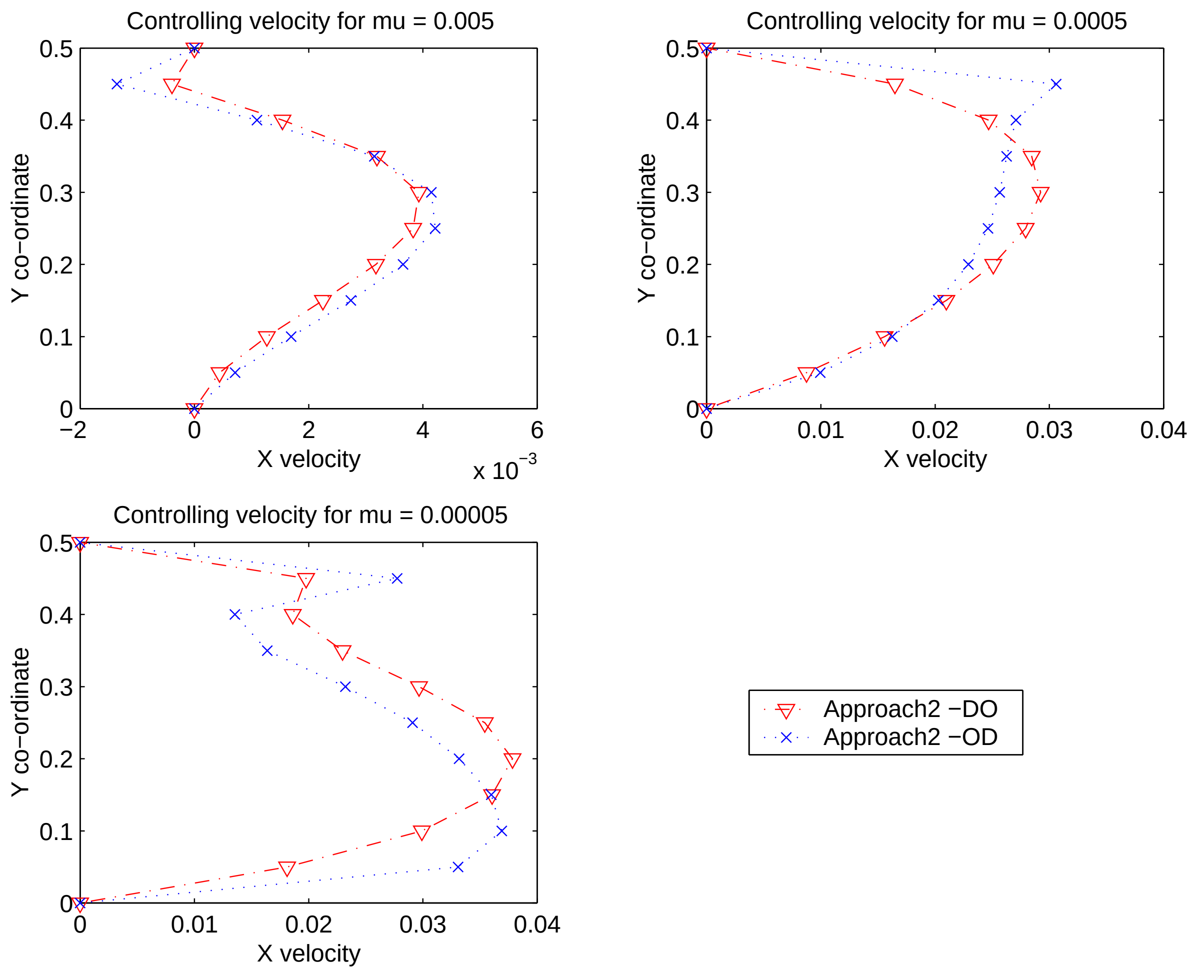
$$\text{“}\boldsymbol{\lambda}\cdot\mathbf{n} + \alpha\mathbf{g} = 0'' \quad \text{on } \Gamma_c.$$

State equation: As before.

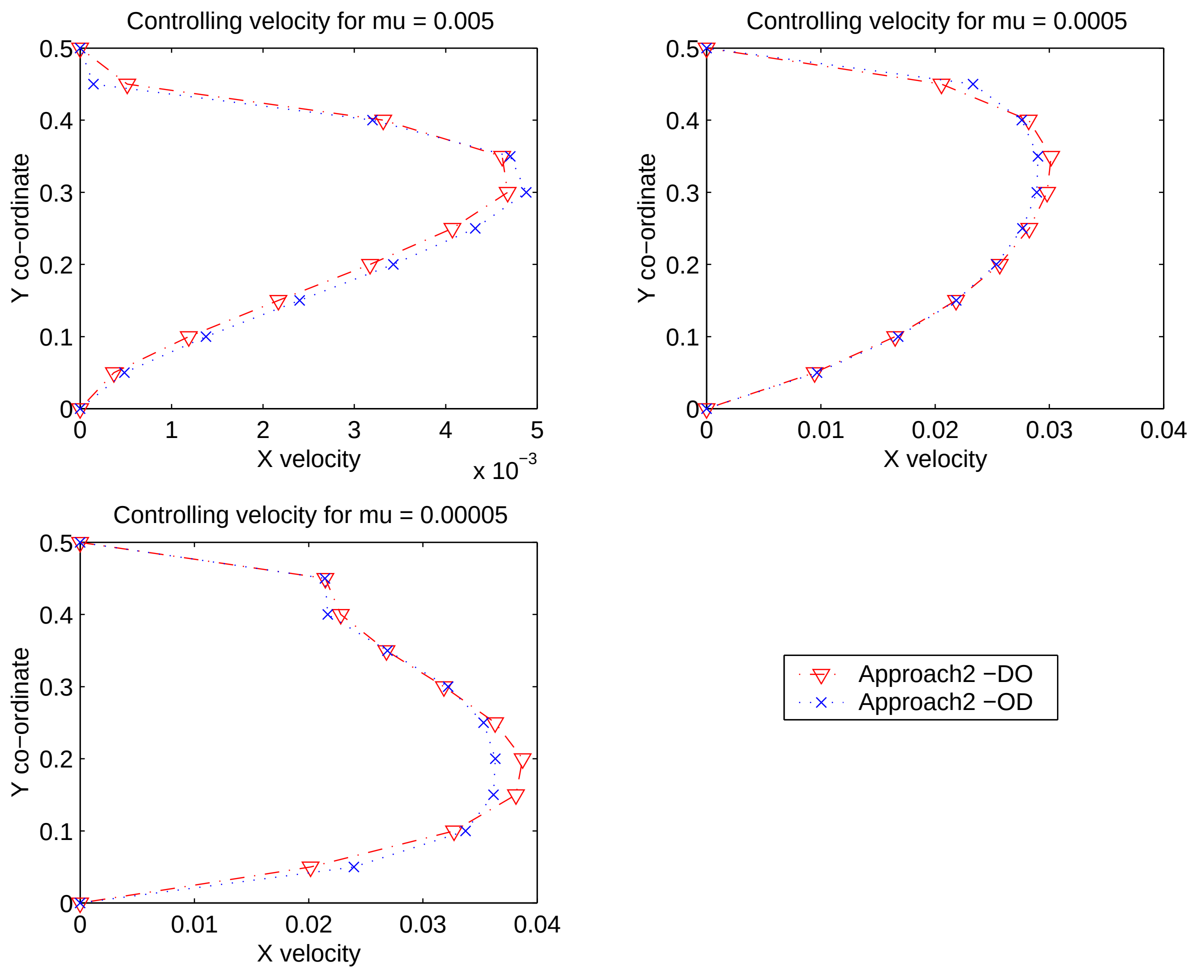
- Galerkin/Least Squares (GaLS) stabilization.
- Linear finite elements.
- Replace Dirichlet boundary condition $\mathbf{u} = \mathbf{g}$ on $\Gamma_c$ by

$$\mathbf{n}\cdot[-p\mathbf{I} + \mu(\nabla\mathbf{u} + \nabla\mathbf{u}^T)] + 10^5\mathbf{u} = 10^5\mathbf{g} \text{ on } \Gamma_c.$$

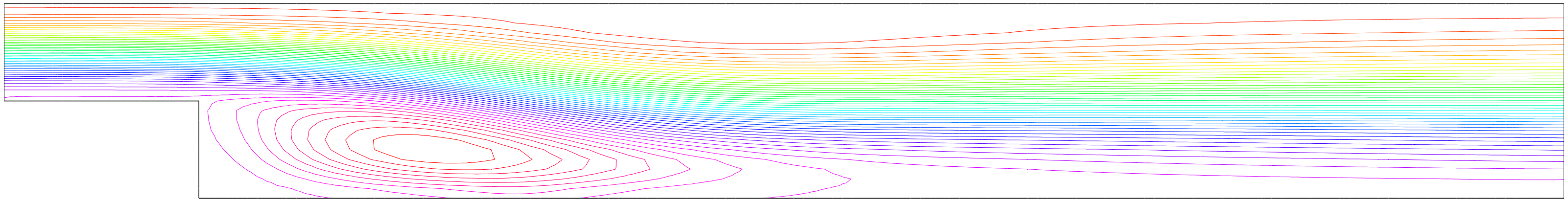
Hou/Ravindran (1998)



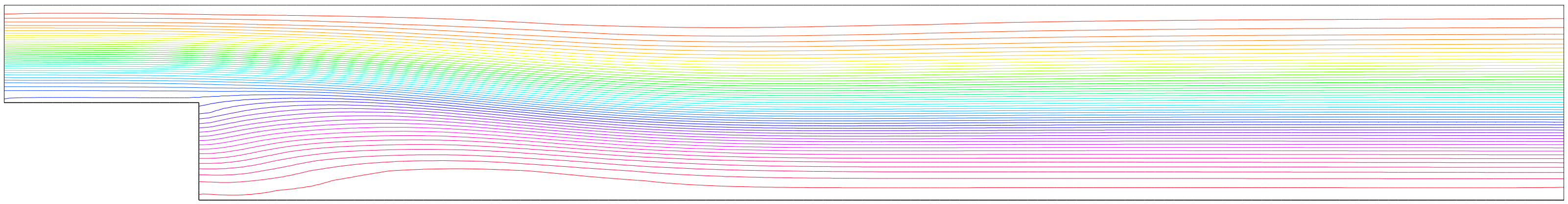

Optimal controls obtained using DO and OD, coarse discretization, $\alpha = 10^{-5}$.




Optimal controls obtained using DO and OD, fine discretization, $\alpha = 10^{-5}$.



Uncontrolled flow



Controlled flow

$\mu = 5 * 10^{-4}$, $\alpha = 10^{-5}$

The discretization of the optimal control problem implies a discretization for the adjoint differential equation (discretize-then-optimize). This implied discretization scheme of the adjoint equation may not have the same convergence properties as the discretization scheme for the state equation.

The discretization of the optimal control problem (discretize-then-optimize) and the discretization of the optimality conditions (optimize-then-discretize) may lead to systems whose solution better approximates the solution of the optimal control problem. However, the discretized optimality systems may be nonsymmetric.

Both approaches, discretize-then-optimize and optimize-then-discretize, may offer advantages and disadvantages. It is important to look at both.

# 5. Optimization Algorithms

The infinite dimensional optimization problem strongly influences the convergence behavior of the optimization algorithm applied to the discretized problem.

Therefore, it is important to study optimization algorithms in function spaces.

$$\begin{aligned} \min \quad & J(y,u) \\ \text{s.t.} \quad & c(y,u) = 0, \\ & g(y,u) = 0, \\ & h(y,u) \in K \end{aligned}$$

where

$$J : \mathcal{Y} \times \mathcal{U} \to \mathbb{R}, \qquad c : \mathcal{Y} \times \mathcal{U} \to \mathcal{C},$$
$$g : \mathcal{Y} \times \mathcal{U} \to \mathcal{G}, \qquad h : \mathcal{Y} \times \mathcal{U} \to \mathcal{H},$$

$\mathcal{Y}, \mathcal{U}, \mathcal{C}, \mathcal{G}, \mathcal{H}$ are Banach spaces, and $K \subset \mathcal{H}$ is a cone.

Notation:
$y$: states, $\mathcal{Y}$: state space, $u$: controls, $\mathcal{U}$: control space,
$c(y,u) = 0$ state equation.

$$\begin{aligned} \min \quad & J(y,u) \\ \text{s.t.} \quad & c(y,u) = 0, \\ & g(y,u) = 0, \\ & h(y,u) \in K \end{aligned}$$

$$\Downarrow$$

$y(u)$ is the unique solution of $c(y,u) = 0$

$$\Downarrow$$

$$\left.\begin{aligned} \min \quad & \widehat{J}(u) \\ \text{s.t.} \quad & \widehat{g}(u) = 0, \\ & \widehat{h}(u) \in K, \end{aligned}\right\} \text{reduced problem}$$

where $\widehat{J}(u) \stackrel{\text{def}}{=} J(y(u),u)$, $\widehat{g}(u) \stackrel{\text{def}}{=} g(y(u),u)$, $\widehat{h}(u) \stackrel{\text{def}}{=} h(y(u),u)$.

We want to solve

$$\min \widehat{J}(u)$$

where $\widehat{J}(u) \stackrel{\text{def}}{=} J(y(u), u)$ using gradient based methods.

We want to solve

$$\min \widehat{J}(u)$$

where $\widehat{J}(u) \stackrel{\text{def}}{=} J(y(u), u)$ using gradient based methods.

Gradient type methods:

$$u_{k+1} = u_k - \tau_k \nabla \widehat{J}(u_k),$$

with step size $\tau_k \in (0, 1]$.

We want to solve

$$\min \widehat{J}(u)$$

where $\widehat{J}(u) \stackrel{\text{def}}{=} J(y(u), u)$ using gradient based methods.

Gradient type methods:

$$u_{k+1} = u_k - \tau_k \nabla \widehat{J}(u_k),$$

with step size $\tau_k \in (0, 1]$.

Newton type methods

$$\nabla^2 \widehat{J}(u_k) s = -\nabla \widehat{J}(u_k),$$
$$u_{k+1} = u_k + \tau_k s,$$

with step size $\tau_k \in (0, 1]$.

We want to solve

$$\min \widehat{J}(u)$$

where $\widehat{J}(u) \stackrel{\text{def}}{=} J(y(u), u)$ using gradient based methods.

Gradient type methods:

$$u_{k+1} = u_k - \tau_k \nabla \widehat{J}(u_k),$$

with step size $\tau_k \in (0, 1]$.

Newton type methods

$$\nabla^2 \widehat{J}(u_k)s = -\nabla \widehat{J}(u_k),$$
$$u_{k+1} = u_k + \tau_k s,$$

with step size $\tau_k \in (0, 1]$.

Computation of $\nabla \widehat{J}(u_k)$? Computation of $\nabla^2 \widehat{J}(u_k)$?

Consider

$$\widehat{J}(u) = J(y(u), u),$$

$u \in \mathcal{U}$ Hilbert space, where $y = y(u)$ is the unique solution of $c(y, u) = 0$.

Consider

$$\widehat{J}(u) = J(y(u), u),$$

$u \in \mathcal{U}$ Hilbert space, where $y = y(u)$ is the unique solution of $c(y, u) = 0$.

F-derivative of $\widehat{J}(u)$ applied to $u'$:

$$\begin{aligned} D\widehat{J}(u)u' &= D_u J(y,u)u' + \langle D_y J(y,u), y_u(u)u' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \\ &= D_u J(y,u)u' + \langle D_y J(y,u), [-c_y(y,u)^{-1} c_u(y,u)]u' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \\ &= D_u J(y,u)u' + \langle \underbrace{-c_y(y,u)^{-*} D_y J(y,u)}_{p}, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}} \end{aligned}$$

Adjoint equation: Compute $p$ such that

$$\langle c_y(y,u)^* p, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} = \langle D_y J(y,u), y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \quad \forall y' \in \mathcal{Y}.$$

Consider

$$\widehat{J}(u) = J(y(u), u),$$

$u \in \mathcal{U}$ Hilbert space, where $y = y(u)$ is the unique solution of $c(y, u) = 0$.

F-derivative of $\widehat{J}(u)$ applied to $u'$:

$$\begin{aligned} D\widehat{J}(u)u' &= D_u J(y,u)u' + \langle D_y J(y,u), y_u(u)u' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \\ &= D_u J(y,u)u' + \langle D_y J(y,u), [-c_y(y,u)^{-1} c_u(y,u)]u' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \\ &= D_u J(y,u)u' + \langle -\underbrace{c_y(y,u)^{-*} D_y J(y,u)}_{p}, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}} \end{aligned}$$

Adjoint equation: Compute $p$ such that

$$\langle c_y(y,u)^* p, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} = \langle D_y J(y,u), y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \quad \forall y' \in \mathcal{Y}.$$

Gradient: Riesz representation of $D\widehat{J}(u)$. Find $\nabla \widehat{J}(u) \in \mathcal{U}$ such that

$$\langle \nabla \widehat{J}(u), u' \rangle_{\mathcal{U}} = \langle \nabla_u J(y,u), u' \rangle_{\mathcal{U}} - \langle p, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}} \quad \forall u' \in \mathcal{U}.$$

Let $p$ solve the adjoint equation

$$\langle c_y(y,u)^* p, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} = \langle D_y J(y,u), y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \quad \forall y' \in \mathcal{Y}.$$

Define the Lagrangian

$$L(y,u,p) = J(y,u) - \langle p, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}}.$$

Observe

$$\begin{aligned} D\widehat{J}(u)u' &= D_u J(y,u)u' - \langle p, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}}, \\ &= D_u L(y,u,p)u' \end{aligned}$$

Gradient: Riesz representation of $D\widehat{J}(u)$. Find $\nabla \widehat{J}(u) \in \mathcal{U}$ such that

$$\langle \nabla \widehat{J}(u), u' \rangle_{\mathcal{U}} = D_u L(y,u,p)u' \quad \forall u' \in \mathcal{U}.$$

Consider

$$\widehat{J}(u) := \frac{1}{2}\,\|y(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

where $y(u)$ solves

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

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $y_\Omega \in L^\infty(\Omega)$, $\alpha \geq 0$.*

Consider

$$\widehat{J}(u) := \frac{1}{2}\,\|y(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)}$$

where $y(u)$ solves

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

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $y_\Omega \in L^\infty(\Omega)$, $\alpha \geq 0$.*

The problem

$$\min \widehat{J}(u)$$

is well posed if the control space satisfied Let $\mathcal{U} \subset L^r(\Omega)$, $r > N/2$.
For $N = 1, 2$, or $3$ choose $\mathcal{U} = L^2(\Omega)$.

Lagrangian

$$L(y,u,p) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{L^2(\Omega)} - \int_\Omega \nabla y \nabla p + yp + d(y)p - up\,dx$$

Solve adjoint equation

$$\int_\Omega \nabla v \nabla p + vp + d'(y)vp\,dx = \int_\Omega (y - y_\Omega)v\,dx \quad \forall v \in H^1(\Omega).$$

Compute $g = \nabla \widehat{J}(u) \in L^2(\Omega)$ such that

$$\int_\Omega g(x)u'(x)dx = \alpha \int_\Omega u(x)u'(x)dx + \int_\Omega u'(x)p(x)dx \quad \forall u' \in L^2(\Omega).$$

Hence

$$\nabla \widehat{J}(u) = \alpha u + p.$$

Consider

$$\widehat{J}(u) := \frac{1}{2} \|y(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u\|^2_{\mathcal{U}}$$

where $y(u)$ solves

$$\begin{aligned} -\Delta y + y + d(y) &= f && \text{in} \quad \Omega \\ \partial_n y &= u && \text{on } \Gamma_c \\ \partial_n y &= 0 && \text{on } \Gamma \setminus \Gamma_c \end{aligned}$$

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $f \in L^2(\Omega)$, $y_\Omega \in L^\infty(\Omega)$, $\alpha \geq 0$.*

Consider

$$\widehat{J}(u) := \frac{1}{2}\,\|y(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{\mathcal{U}}$$

where $y(u)$ solves

$$\begin{aligned} -\Delta y + y + d(y) &= f && \text{in } \Omega \\ \partial_n y &= u && \text{on } \Gamma_c \\ \partial_n y &= 0 && \text{on } \Gamma \setminus \Gamma_c \end{aligned}$$

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $f \in L^2(\Omega)$, $y_\Omega \in L^\infty(\Omega)$, $\alpha \geq 0$.*

The problem

$$\min \widehat{J}(u)$$

is well posed if the control space satisfied Let $\mathcal{U} \subset L^s(\Gamma_c)$, $s > N - 1$.

Consider

$$\widehat{J}(u) := \frac{1}{2}\,\|y(u) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\,\|u\|^2_{\mathcal{U}}$$

where $y(u)$ solves

$$\begin{aligned} -\Delta y + y + d(y) &= f && \text{in } \Omega \\ \partial_n y &= u && \text{on } \Gamma_c \\ \partial_n y &= 0 && \text{on } \Gamma \setminus \Gamma_c \end{aligned}$$

**Assumptions:** *$\Omega \subset I\!R^N$ is a bounded Lipschitz domain. The function $d : I\!R \to I\!R$ is monotone non-decreasing, twice differentiable with locally Lipschitz second derivative. Moreover, $f \in L^2(\Omega)$, $y_\Omega \in L^\infty(\Omega)$, $\alpha \geq 0$.*

The problem

$$\min\ \widehat{J}(u)$$

is well posed if the control space satisfied Let $\mathcal{U} \subset L^s(\Gamma_c)$, $s > N-1$.

For $N = 1,2$ choose $\mathcal{U} = L^2(\Omega)$. For $N = 3$ choose $\mathcal{U} = H^1(\Omega)$.

Lagrangian

$$L(y,u,p) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\int_{\Gamma_c} |\nabla_s u|^2 + u^2 dx - \int_\Omega \nabla y \nabla p + yp + d(y)p dx + \int_{\Gamma_c} up dx$$

Solve adjoint equation

$$\int_\Omega \nabla v \nabla p + vp + d'(y)vp dx = \int_\Omega (y - y_\Omega) v dx \quad \forall v \in H^1(\Omega).$$

Compute $g = \nabla \widehat{J}(u) \in H^1(\Gamma_c)$ such that

$$\int_{\Gamma_c} \nabla_s g \nabla_s u' + gu' dx = \alpha \int_{\Gamma_c} \nabla_s u \nabla_s u' + uu' dx + \int_{\Gamma_c} u' p dx \quad \forall u' \in H^1(\Gamma_c).$$

Lagrangian

$$L(y,u,p) = \frac{1}{2}\,\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\int_{\Gamma_c} |\nabla_s u|^2 + u^2 dx - \int_\Omega \nabla y \nabla p + yp + d(y)p dx + \int_{\Gamma_c} up dx$$

Solve adjoint equation

$$\int_\Omega \nabla v \nabla p + vp + d'(y)vp dx = \int_\Omega (y - y_\Omega) v dx \quad \forall v \in H^1(\Omega).$$

Compute $g = \nabla \widehat{J}(u) \in H^1(\Gamma_c)$ such that

$$\int_{\Gamma_c} \nabla_s g \nabla_s u' + gu' dx = \alpha \int_{\Gamma_c} \nabla_s u \nabla_s u' + uu' dx + \int_{\Gamma_c} u' p dx \quad \forall u' \in H^1(\Gamma_c).$$

Note:

$$\nabla \widehat{J}(u) \neq \alpha u + p|_{\Gamma_c}.$$

Computation of $\nabla^2 \widehat{J}(u)\delta u \in \mathcal{U}$

- Given $u \in \mathcal{U}$.
- Compute solution $y \in \mathcal{Y}$ of the state equation $c(y,u) = 0$.
- Compute solution $p$ of the adjoint equation
  $\langle c_y(y,u)^* p, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} = \langle D_y J(y,u), y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \quad \forall y' \in \mathcal{Y}$.
- Compute solution $\delta y \in \mathcal{Y}$ of the linearized state equation
  $c_y(y(u),u)\delta y = -c_u(y(u),u)\delta u$.
- Compute solution $\eta$ of

$$\langle c_y(y,u)^* \eta, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} = \langle D^2_{yy} L(y,u,p)\delta y D^2_{yu} L(y,u,p)\delta u, y' \rangle_{\mathcal{Y}^* \times \mathcal{Y}} \quad \forall y' \in \mathcal{Y}$$

- Find $z = \nabla^2 \widehat{J}(u)\delta u \in \mathcal{U}$ such that

$$\begin{aligned} \langle \nabla^2 \widehat{J}(u)\delta u, u' \rangle_{\mathcal{U}} =& \langle D^2_{uy} L(y,u,p)\delta y + D^2_{uu} L(y,u,p)\delta u, u' \rangle_{\mathcal{U}^* \times \mathcal{U}} \\ & - \langle \eta, c_u(y,u)u' \rangle_{\mathcal{C}^* \times \mathcal{C}} \quad \forall u' \in \mathcal{U}. \end{aligned}$$

Hessian $\nabla^2 \widehat{J}(u)$ usually not available in matrix form. Only operator-vector multiplications can be computed.

Newton step

$$\nabla^2 \widehat{J}(u_k) s = -\nabla \widehat{J}(u_k), \quad (*)$$

has to be computed using iterative methods (conjugate gradient method).

Hessian $\nabla^2 \widehat{J}(u)$ usually not available in matrix form. Only operator-vector multiplications can be computed.

Newton step

$$\nabla^2 \widehat{J}(u_k) s = -\nabla \widehat{J}(u_k), \quad (*)$$

has to be computed using iterative methods (conjugate gradient method).

Solution $s$ of (*) is the second component of the solution vector $s_y, s_u$ of

$$\min \left\langle \begin{bmatrix} D_y L \\ D_u L \end{bmatrix}, \begin{bmatrix} s_y \\ s_u \end{bmatrix} \right\rangle + \frac{1}{2} \left\langle \begin{bmatrix} D^2_{yy} L & D^2_{yu} L \\ D^2_{uy} L & D^2_{uu} L \end{bmatrix} \begin{bmatrix} s_y \\ s_u \end{bmatrix}, \begin{bmatrix} s_y \\ s_u \end{bmatrix} \right\rangle,$$

$$\text{s.t. } c_y(y,u) s_y + c_u(y,u) s_u = 0,$$

where $L = L(y,u,p)$, $y$ is the solution of the state equation and $p$ is the solution of the adjoint equation

Want to solve

$$\min_{u \in \mathcal{U}} \widehat{J}(u)$$

Newton method

- Given $u_k$.
- Solve $\nabla^2 \widehat{J}(u_k)s = -\nabla \widehat{J}(u_k)$.
- Set $u_{k+1} = u_k + s$.

Convergence: Let $\widehat{J}(u)$ be twice continuously blue F-differentiable, let $\nabla \widehat{J}(u_*) = 0$ and $\langle \nabla^2 \widehat{J}(u_*)v, v\rangle_{\mathcal{U}} \geq \delta \|v\|_{\mathcal{U}}^2$ for some $\delta > 0$ (second order sufficient optimality conditions). There exists $\epsilon > 0$ and $c > 0$ such that if $\|u_0 - u_*\| \leq \epsilon$ then $\lim u_k = 0$ and

$$\|u_{k+1} - u_*\|_{\mathcal{U}} \leq c\|u_k - u_*\|_{\mathcal{U}}^2 \quad \forall k.$$

Want to solve

$$\min_{u \in \mathcal{U}} \widehat{J}(u) \qquad\qquad \min_{u^h \in \mathcal{U}^h} \widehat{J}_h(u^h)$$

Mesh independence (basic version)

- $\mathcal{U}^h \subset \mathcal{U}$.
- $\|\nabla \widehat{J}(u) - \nabla \widehat{J}_h(u^h)\| \leq \hat{c}h^q \quad \forall u^h \in \mathcal{U}^h$.
- $\|\nabla^2 \widehat{J}(u) - \nabla^2 \widehat{J}_h(u^h)\| \leq \hat{c}h^q \quad \forall u^h \in \mathcal{U}^h$.

Let the assumptions for the convergence of Newton's method in $\mathcal{U}$ hold. There exist $h_0$, $\epsilon$ and $c > 0$ such that for all $h \leq h_0$ and all $u_0^h \in \mathcal{U}^h$ with $\|u_0^h - u_*\|_{\mathcal{U}} \leq \epsilon$,

- $\lim u_k^h = u_*^h$,
- $\|u_{k+1}^h - u_*^h\|_{\mathcal{U}} \leq c\|u_k^h - u_*^h\|_{\mathcal{U}}^2 \quad \forall k$.
- $\|u_*^h - u_*\|_{\mathcal{U}} \leq ch^q$.
- If $\|u_0^h - u_0\|_{\mathcal{U}} \leq ch^q$, then $\|u_k^h - u_k\|_{\mathcal{U}} \leq ch^q \quad \forall k$.

| **Number of Iterations** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Example 1** | | | | | | | | | | | |
| | TOL $= 10^{-8}$ | | | | | | TOL $= 10^{-6}$ | | | | | |
| $\alpha \quad h^{-1}$ | 12 | 24 | 48 | 96 | 192 | 384 | 12 | 24 | 48 | 96 | 192 | 384 |
| $10^{-6}$ | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| $10^{-4}$ | 8 | 8 | 8 | 8 | 8 | 8 | 7 | 7 | 7 | 7 | 7 | 7 |
| $10^{-2}$ | 10 | 10 | 10 | 10 | 10 | 10 | 8 | 8 | 8 | 8 | 8 | 8 |
| | **Example 2** | | | | | | | | | | | |
| $10^{-6}$ | 7 | 7 | 7 | 7 | 7 | 7 | 6 | 6 | 6 | 6 | 6 | 6 |
| $10^{-4}$ | 9 | 9 | 9 | 9 | 9 | 9 | 7 | 7 | 7 | 7 | 7 | 7 |
| $10^{-2}$ | 9 | 9 | 9 | 9 | 9 | 9 | 7 | 7 | 7 | 7 | 7 | 7 |

Want to solve

$$\min_{u \in \mathcal{U}} \widehat{J}(u)$$

BFGS method

- Given $u_k$, $H_k$.
- Solve $H_k s = -\nabla \widehat{J}(u_k)$.
- Set $u_{k+1} = u_k + s$, $v = \nabla \widehat{J}(u_{k+1}) - \nabla \widehat{J}(u_k)$ and

$$H_{k+1} = H_k + \frac{v \otimes v}{\langle v, s \rangle_{\mathcal{U}}} - \frac{(H_k s) \otimes (H_k s)}{\langle s, H_k s \rangle_{\mathcal{U}}}.$$

  Here $(v \otimes w)x = \langle w, x \rangle_{\mathcal{U}} v$.

Convergence: Let $\widehat{J}(u)$ be twice continuously blue F-differentiable, let $\langle \nabla \widehat{J}(u_*) = 0$ and $\langle \nabla^2 \widehat{J}(u_*)v, v \rangle_{\mathcal{U}} \geq \delta \|v\|^2_{\mathcal{U}}$ for some $\delta > 0$. There exists $\epsilon > 0$ such that if $\|u_0 - u_*\| \leq \epsilon$, $\|H_0 - \nabla^2 \widehat{J}(u_*)\| \leq \epsilon$, and if $H_0 - \nabla^2 \widehat{J}(u_*)$ is compact, then there exists $c_k \geq 0$, $\lim c_k = 0$, $\lim u_k = 0$ and

$$\|u_{k+1} - u_*\|_{\mathcal{U}} \leq c_k \|u_k - u_*\|_{\mathcal{U}} \quad \forall k.$$

$$\min J(\mathbf{u}, \mathbf{g}) = \frac{1}{2} \int_{\{x_2=0.4\}} |u_2(x)|^2 \, dx + \frac{\gamma}{2} \, \|\mathbf{g}\|^2_{H^1(\Gamma_c)}$$

subject to

$$\begin{aligned} -\tfrac{1}{Re}\, \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega = (0,1)^2, \\ \text{div } \mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{b} && \text{on } \Gamma_u, \\ \mathbf{u} &= \mathbf{g} && \text{on } (0,1) \times \{1\}. \end{aligned}$$

**Note change of notation:** $\mathbf{u}, p$ (velocities, pressure) states, $\mathbf{g}$ control.

$$\min J(\mathbf{u}, \mathbf{g}) = \frac{1}{2} \int_{\{x_2=0.4\}} |u_2(x)|^2 \, dx + \frac{\gamma}{2} \, \|\mathbf{g}\|^2_{H^1(\Gamma_c)}$$

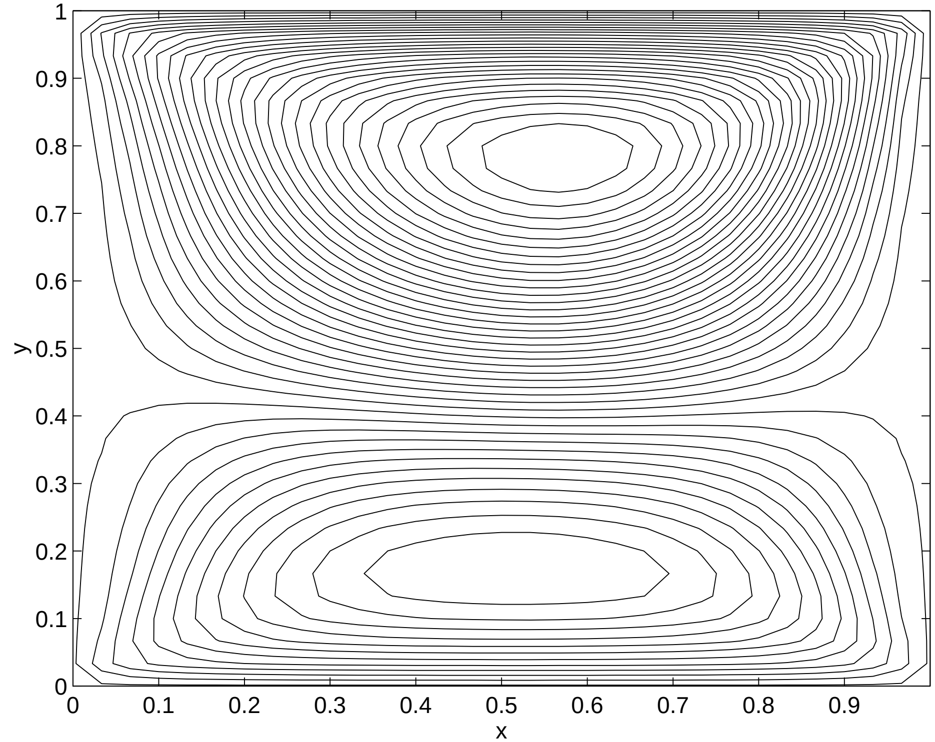
subject to

$$\begin{aligned} -\tfrac{1}{Re} \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega = (0,1)^2, \\ \text{div } \mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{b} && \text{on } \Gamma_u, \\ \mathbf{u} &= \mathbf{g} && \text{on } (0,1) \times \{1\}. \end{aligned}$$

**Note change of notation:** $\mathbf{u}, p$ (velocities, pressure) states, $\mathbf{g}$ control.

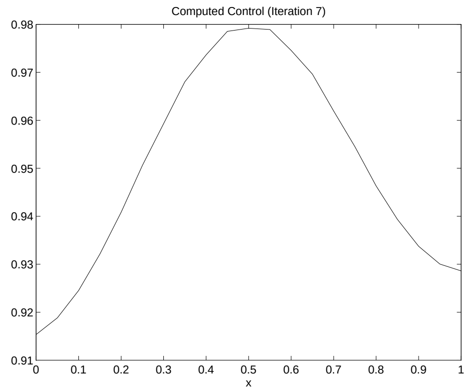
Control space is $H^1(\Gamma_c)$. Gradient computation analogous to the procedure described for boundary control of semilinear elliptic equation.

$\longrightarrow$ Velocity $\mathbf{g}$ (control) $\longrightarrow$

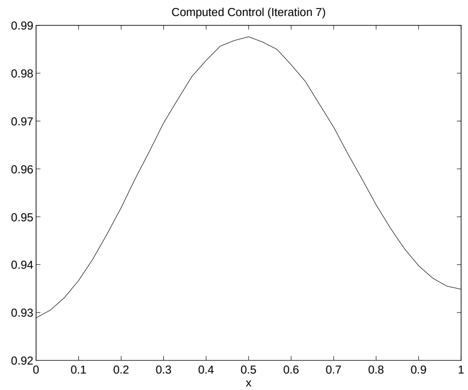



$\longrightarrow$ Velocity $\mathbf{b}$ (given $0.5$) $\longrightarrow$

Infinite dimensional approach.

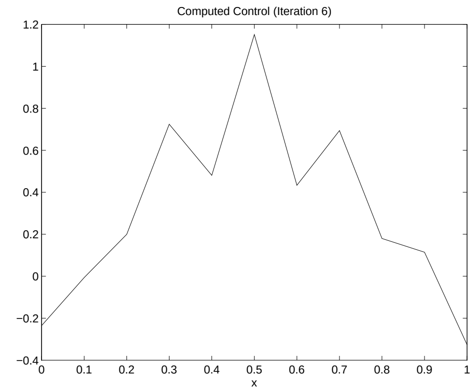



(Grid size $h = 1/10$).

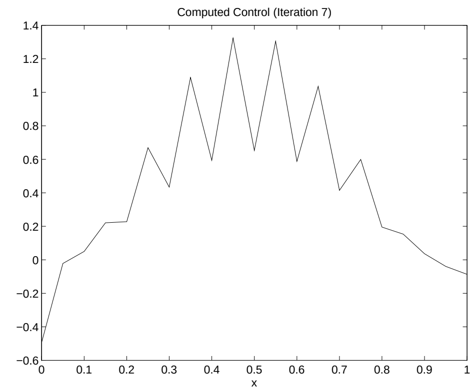



(Grid size $h = 1/15$).

Finite dimensional approaach.




(Grid size $h = 1/5$).




(Grid size $h = 1/10$).

Consider

$$\min_{\mathcal{U}} \widehat{J}(u),$$

where $\mathcal{U}$ Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$.

Consider

$$\min_{\mathcal{U}} \widehat{J}(u),$$

where $\mathcal{U}$ Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$.

After a discretization, this leads to

$$\min_{\mathcal{U}^h} \widehat{J}(u_h)$$

for some finite dimensional subspace $\mathcal{U}^h \subset \mathcal{U}$.

Consider

$$\min_{\mathcal{U}} \widehat{J}(u),$$

where $\mathcal{U}$ Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$.

After a discretization, this leads to

$$\min_{\mathcal{U}^h} \widehat{J}(u_h)$$

for some finite dimensional subspace $\mathcal{U}^h \subset \mathcal{U}$.

We can identify $\mathcal{U}^h$ with $\mathbb{R}^n$, but the inner product leads to a weighted Euclidean product

$$\langle u_1, u_2 \rangle_{\mathcal{U}} = \mathbf{u}_1^\top \mathbf{T} \mathbf{u}_2$$

for some positive definite $\mathbf{T} \in \mathbb{R}^{n \times n}$.

The discretized problem can be viewed as problem in $\mathbb{R}^n$

$$\min_{\mathbb{R}^n} \widehat{J}(\mathbf{u}),$$

but $\mathbb{R}^n$ is equipped with the weighted Euclidean product

$$\mathbf{u}_1^\top \mathbf{T} \mathbf{u}_2$$

not with $\mathbf{u}_1^\top \mathbf{u}_2$.

The discretized problem can be viewed as problem in $\mathbb{R}^n$

$$\min_{\mathbb{R}^n} \widehat{J}(\mathbf{u}),$$

but $\mathbb{R}^n$ is equipped with the weighted Euclidean product

$$\mathbf{u}_1^\top \mathbf{T} \mathbf{u}_2$$

not with $\mathbf{u}_1^\top \mathbf{u}_2$.

Let's see when and why this matters.

## Gradient Computation

Let $\widehat{J} : \mathbb{R}^n \to \mathbb{R}$. Denote the derivative of $\widehat{J}$ by $D\widehat{J}$.

The gradient $\nabla \widehat{J}(\mathbf{u})$ is defined to be the vector that satisfies

$$\langle \nabla \widehat{J}(\mathbf{u}), \mathbf{u}' \rangle = D\widehat{J}(\mathbf{u})\mathbf{u}' \quad \forall \mathbf{u}'$$

(Riesz representation). Thus $\nabla \widehat{J}(\mathbf{u})$ depends on the inner product.

If we use

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \mathbf{u}_1^\top \mathbf{u}_2,$$

then

$$\nabla \widehat{J}(\mathbf{u}) = \nabla_1 \widehat{J}(\mathbf{u}) := \left( \frac{\partial}{\partial \mathbf{u}_j} \widehat{J}(\mathbf{u}) \right)_{j=1,\ldots,n},$$

i.e., $\nabla \widehat{J}(\mathbf{u})$ is the vector of partial derivatives.

## Gradient Computation

Let $\widehat{J} : \mathbb{R}^n \to \mathbb{R}$. Denote the derivative of $\widehat{J}$ by $D\widehat{J}$.

The gradient $\nabla \widehat{J}(\mathbf{u})$ is defined to be the vector that satisfies

$$\langle \nabla \widehat{J}(\mathbf{u}), \mathbf{u}' \rangle = D\widehat{J}(\mathbf{u})\mathbf{u}' \quad \forall \mathbf{u}'$$

(Riesz representation). Thus $\nabla \widehat{J}(\mathbf{u})$ depends on the inner product.

If we use

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \mathbf{u}_1^\top \mathbf{T} \mathbf{u}_2,$$

then

$$D\widehat{J}(\mathbf{u})\mathbf{u}' = \nabla_1 \widehat{J}(\mathbf{u})^\top \mathbf{u}' = \left( \mathbf{T}^{-1} \nabla_1 \widehat{J}(\mathbf{u}) \right)^\top \mathbf{T}\mathbf{u}'$$

i.e., $\nabla \widehat{J}(\mathbf{u}) = \mathbf{T}^{-1} \nabla_1 \widehat{J}(\mathbf{u})$.

Same result as scaling of the $\mathbf{u}$-variable by $\mathbf{T}^{1/2}$.

## Hessian Computation

Let $\widehat{J} : \mathbb{R}^n \rightarrow \mathbb{R}$. Denote the second derivative of $\widehat{J}$ by $D^2\widehat{J}$.

The Hessian $\nabla^2\widehat{J}(\mathbf{u})$ is defined to be the matrix that satisfies

$$\langle \nabla^2\widehat{J}(\mathbf{u})\mathbf{u}_1, \mathbf{u}_2\rangle = D^2\widehat{J}(\mathbf{u})[\mathbf{u}_1, \mathbf{u}_2] \quad \forall \mathbf{u}_1, \mathbf{u}_2.$$

Thus $\nabla^2\widehat{J}(\mathbf{u})$ depends on the inner product.

If we use

$$\langle \mathbf{u}_1, \mathbf{u}_2\rangle = \mathbf{u}_1^\top \mathbf{u}_2,$$

then

$$\nabla^2\widehat{J}(\mathbf{u}) = \nabla_1^2\widehat{J}(\mathbf{u}) := \left(\frac{\partial^2}{\partial \mathbf{u}_i \partial \mathbf{u}_j}\widehat{J}(\mathbf{u})\right)_{i,j=1,\ldots,n},$$

i.e., $\nabla^2\widehat{J}(\mathbf{u})$ is the matrix of second partial derivatives.

## Hessian Computation

Let $\widehat{J} : \mathbb{R}^n \to \mathbb{R}$. Denote the second derivative of $\widehat{J}$ by $D^2\widehat{J}$.

The Hessian $\nabla^2\widehat{J}(\mathbf{u})$ is defined to be the matrix that satisfies

$$\langle \nabla^2\widehat{J}(\mathbf{u})\mathbf{u}_1, \mathbf{u}_2 \rangle = D^2\widehat{J}(\mathbf{u})[\mathbf{u}_1, \mathbf{u}_2] \quad \forall \mathbf{u}_1, \mathbf{u}_2.$$

Thus $\nabla^2\widehat{J}(\mathbf{u})$ depends on the inner product.

If we use

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \mathbf{u}_1^\top \mathbf{T}\mathbf{u}_2,$$

then

$$D^2\widehat{J}(\mathbf{u})[\mathbf{u}_1, \mathbf{u}_2] = \left(\nabla_1^2\widehat{J}(\mathbf{u})\mathbf{u}_1\right)^\top \mathbf{u}_2 = \left(\mathbf{T}^{-1}\nabla_1^2\widehat{J}(\mathbf{u})\mathbf{u}_1\right)^\top \mathbf{T}\mathbf{u}_2$$

i.e., $\nabla^2\widehat{J}(\mathbf{u}) = \mathbf{T}^{-1}\nabla_1^2\widehat{J}(\mathbf{u})$.

Same result as scaling of the $\mathbf{u}$-variable by $\mathbf{T}^{1/2}$.

## Quasi Newton: BFGS Update

$$\mathbf{H}_{k+1} = \mathbf{H} + \frac{\mathbf{v} \otimes \mathbf{v}}{\langle \mathbf{v}, \mathbf{s} \rangle} - \frac{(\mathbf{Hs}) \otimes (\mathbf{Hs})}{\langle \mathbf{s}, \mathbf{Hs} \rangle}.$$

where

$$(\mathbf{x} \otimes \mathbf{v})\mathbf{w} = \langle \mathbf{v}, \mathbf{w} \rangle \, \mathbf{x}.$$

If $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^\top \mathbf{w}$, then we obtain the standard BFGS update.

If $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^\top \mathbf{T} \mathbf{w}$, then

$$\mathbf{H}_{k+1} = \mathbf{H} + \frac{\mathbf{v}(\mathbf{Tv})^\top}{\mathbf{v}^\top \mathbf{Ts}} - \frac{(\mathbf{Hs})(\mathbf{THs})^\top}{\mathbf{s}^\top \mathbf{THs}}.$$

This is the BFGS update resulting from a scaling of the independent variables $\mathbf{u}$ by $\mathbf{T}^{1/2}$.

- If we discretize the optimal control and solve the discretized problem as a nonlinear problem in $\mathbb{R}^n$ with standard Euclidean inner product, the convergence of
  - gradient
  - quasi-Newton
  - conjugate gradient (CG)
  - Newton CG
  - ...

  methods depend on the mesh size.
- Often, the finer the mesh size, the more poorly scaled the discretized nonlinear programming problems become.

The problem we want to solve

$$\begin{array}{ll} \min & J_h(y_h, u_h) \\ \text{s.t.} & c_h(y_h, u_h) = 0 \end{array}$$

$$\Downarrow$$

$$\min \widehat{J}_h(u_h) \stackrel{\text{def}}{=} J_h(y_h(u_h), u_h)$$

- Solution of $c_h(y_h, u_h) = 0$ is determined iteratively
  $\longrightarrow$ only $\tilde{y}_h(u_h) \approx y_h(u_h)$ is known
  $\longrightarrow$ only $\widehat{J}_h(u_h), \nabla \widehat{J}_h(u_h)$ are not known
- $\|\tilde{y}_h(u_h) - y_h(u_h)\|$ can be controlled, but often only asymptotic estimates are known.

Optimization problem

$$\begin{aligned} \min \quad & J(y,u) \\ \text{s.t.} \quad & c(y,u) = 0 \end{aligned}$$

Optimality conditions

$$D_y^2 L(y,u,p) = 0, \quad D_u^2 L(y,u,p) = 0, \quad c(y,u) = 0.$$

Newton's Method

$$\begin{bmatrix} D_{yy}^2 L(y,u,p) & D_{yu}^2 L(y,u,p) & c_y(y,u)^* \\ D_{uy}^2 L(y,u,p) & D_{uu}^2 L(y,u,p) & c_u(y,u)^* \\ c_y(y,u) & c_u(y,u) & 0 \end{bmatrix} \begin{bmatrix} s_y \\ s_p \\ s_u \end{bmatrix} = - \begin{bmatrix} D_y L(y,u,p) \\ D_u L(y,u,p) \\ c(y,u) \end{bmatrix}$$

Quadratic problem

$$\min \left\langle \begin{bmatrix} D_y L \\ D_u L \end{bmatrix}, \begin{bmatrix} s_y \\ s_u \end{bmatrix} \right\rangle + \frac{1}{2} \left\langle \begin{bmatrix} D^2_{yy} L & D^2_{yu} L \\ D^2_{uy} L & D^2_{uu} L \end{bmatrix} \begin{bmatrix} s_y \\ s_u \end{bmatrix}, \begin{bmatrix} s_y \\ s_u \end{bmatrix} \right\rangle,$$

s.t. $c_y(y,u)s_y + c_u(y,u)s_u = -c(y,u)$, The problem we want to solve

$$\begin{aligned} \min \quad & J(y,u) \\ \text{s.t.} \quad & c(y,u) = 0, g(y,u) = 0, h(y,u) \in K. \end{aligned} \qquad (P)$$

The problem we can solve

$$\begin{aligned} \min \quad & J_h(y_h,u_h) \\ \text{s.t.} \quad & c_h(y_h,u_h) = 0, g_h(y_h,u_h) = 0, h_h(y_h,u_h) \in K_h. \end{aligned} \qquad (P_h)$$

- The infinite dimensional optimization problem $(P)$ strongly influences the convergence behavior of the optimization algorithm applied to the discretized problem $(P_h)$.
  - Mesh independece principles.
  - Convergence of quasi-Newton methods.
  - Development of new opt. algorithms for infinite dim. problems.

- There is not one, but a sequence of optimization problems $(P_h)$. The better $(P_h)$ approximates $(P)$, the larger $(P_h)$ becomes. Want to use inexpensive problems as long as possible.
- Efficient solution of optimization subproblems at fixed level $(P_h)$.

- Optimization algorithms for problems with control and state constraints Talks by Kunisch, Hintermüller, de los Reyes, Schiela,Wehrstedt.
- Preconditioning of KKT systems Talk by Sachs.
- Optimization and discretization Talks by Rannacher, Hoppe.
- Reduced order Models: Gunzburger, Patera.