SAMSI Summer School 2016

Tutorial on

Computational Methods for PDE constrained optimization

slides by

Prof. Dr. Volker Schulz
Trier University, Germany

Contents

1 PDE constrained optimization vs. nonlinear programming

Our problem class for a moment:

$$\begin{aligned} &\min_{y,u} J(y,u)\,, && J : Y \times U \to \mathbb{R} \\ &c(y,u) = 0\,, && c : Y \times U \to Z\,, \quad \exists\, c_y^{-1} \\ &u \in U_{ad} \subset U \end{aligned}$$

Black-box-approach via implicit functions theorem:

$$y = y(u) \text{ via constraint } c(y,u) = 0\,.$$

From this, we obtain the reduced problem for the reduced objective

$$\min_u \hat{J}(u) := J(y(u),u)$$

with

$$\hat{J} : U_{ad} \to \mathbb{R}$$

Theorem 1.1. *If $U_{ad} \subset H$, for some Hilbert space H, the function $\hat{J} : U_{ad} \to \mathbb{R}$ is continuous, convex (if H is not finite dimensional) and bounded from below and the level set*

$$L_t := \{u \in U_{ad} \,|\, f(u) \leq t\}$$

is a nonempty, closed and bounded set for some $t < \infty$, then there exists an optimal solution $\hat{u} \in U_{ad}$.

Proof: See [Trö10] or other book. □

Corollary 1.2. *If $U_{ad} \subset H$ is a convex set and $\hat{J} : U_{ad} \to \mathbb{R}$ is strictly convex, then the optimal solution $\hat{u} \in U_{ad}$ is unique.*

Convexity all over $\Rightarrow$ typically only linear quadratic model problems.

Let us look at the following tracking type optimization problem

$$\min J(y,u) = \tfrac{1}{2} \int_\Omega (y(x) - z(x))^2 dx$$
$$\text{s.t.} \quad \begin{aligned} -\triangle y(x) &= u(x), \ \forall x \in \Omega \\ y(x) &= 0, \ \forall x \in \Gamma := \partial\Omega \end{aligned}$$

with the computational domain $\Omega := (0,1)^2$ and z a given function.

Discretized Problem:

- tensor grid: $x_h \in G := \{i/h \,|\, i = 1, \ldots, n-1\} \times \{j/h \,|\, j = 1, \ldots, n-1\}$, $h := 1/n$
- state discretization $y_h \in \mathbb{R}^G$
- control discretization $u_h \in \mathbb{R}^G$
- A_h as FD discretization (symmetric) of $-\triangle$ with Dirichlet b.c.

results in linear-quadratic optimization problem

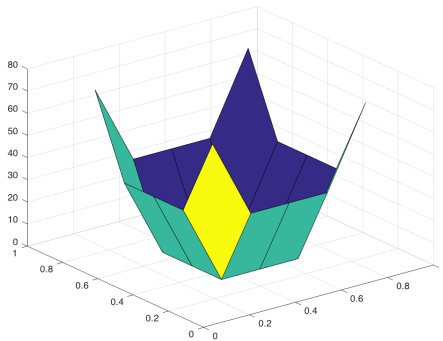
$$\min J_h(y_h, u_h) = \tfrac{1}{2}(y_h - z_h)^\top (y_h - z_h)$$

$$\text{s.t.} \qquad A_h y_h \;=\; u_h$$

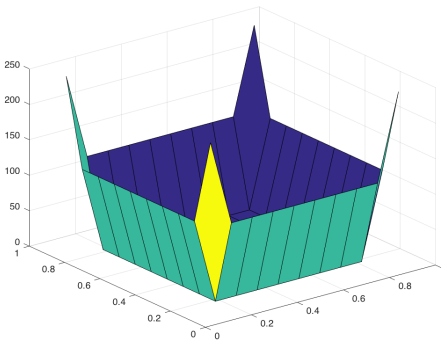
with A_h invertible, yielding the solution

$$\begin{aligned} u_h &= \left(A_h^{-1} A_h^{-1}\right)^{-1} A_h^{-1} z = A_h z_h \\ y_h &= A_h^{-1} u_h = z_h \;\text{ on } G \end{aligned}$$

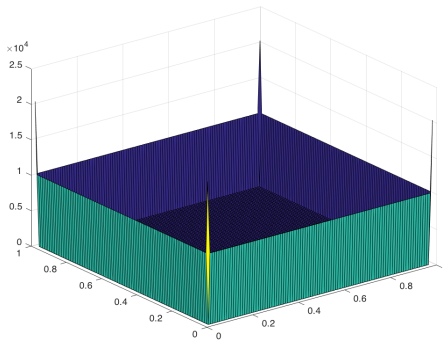
Numerical result on varying grids for $z \equiv 1$:



u_h on 5×5 grid



u_h on 10×10 grid



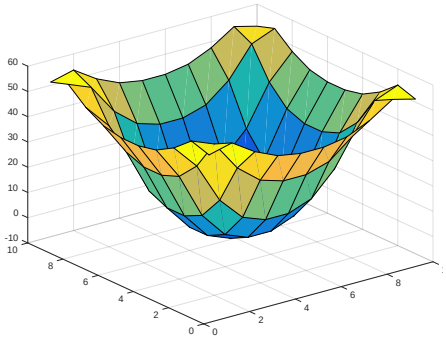
u_h on 100×100 grid

Is this really, what we have expected?

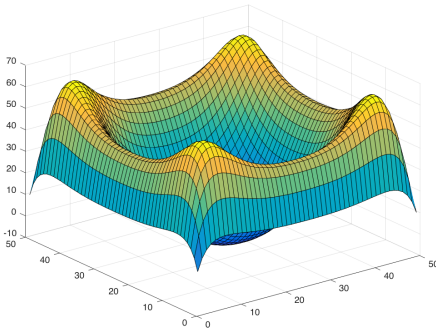
The finite dimensional problem on each mesh is well defined and has a unique solution.

The solutions on the different meshes share common feature but do not converge.

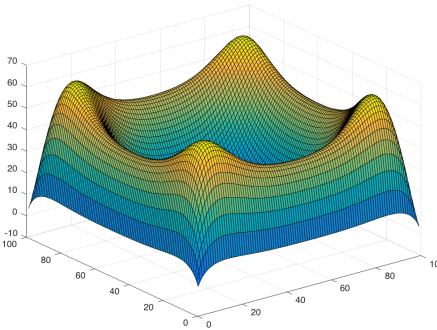
What maybe some of us have expected or hoped for:



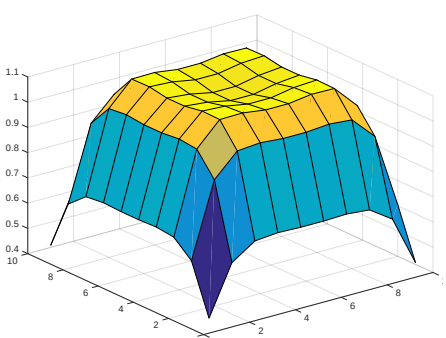
u_h on 10×10 grid



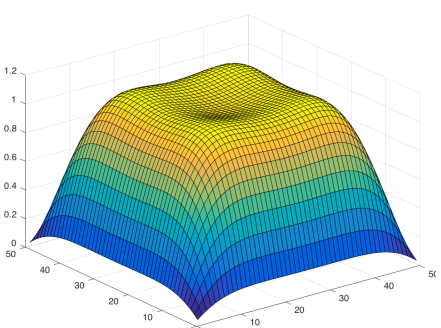
u_h on 50×50 grid



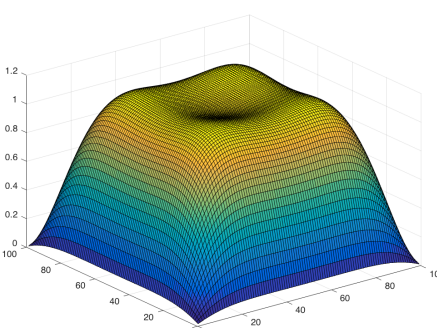
u_h on 100×100 grid



y_h on 10×10 grid



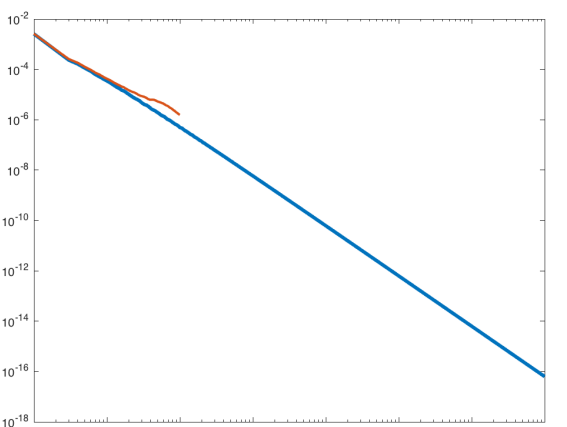
y_h on 50×50 grid



u_h on 100×100 grid

What is going on?

- We have to work in function spaces instead of $\mathbb{R}^G$ on each grid.
- Is the problem strictly convex ? (in the sense of corollary 1.2)
  - YES, in discretized form on each mesh: Hessian is A_h^{-2} with spectrum $\sigma = \{\frac{h^4}{16(\sin^2(k\pi h/2)+\sin^2(\ell\pi h/2))^2} \,|\, 1 \leq k, \ell \leq N-1\}$
  - NO, in function space: Hessian is $(-\triangle_D)^{-2}$, spectrum $\sigma = \{\frac{1}{(k^2+\ell^2)^2\pi^4} \,|\, k, \ell \in \mathbb{N}\}$



Regularized Problem: $y \in H^1(\Omega), u \in L^2(\Omega)$

$$\min J(y,u) = \tfrac{1}{2}\int_\Omega (y(x)-z(x))^2 dx + \frac{\mu}{2}\int_\Omega u(x)^2 dx\,, \qquad \mu > 0$$

$$\text{s.t.} \quad \begin{aligned} -\triangle y(x) &= u(x),\ \forall x \in \Omega \\ y(x) &= 0,\ \forall x \in \Gamma := \partial\Omega \end{aligned}$$

- Regularization term can also be interpreted as control costs.
- Hessian is now $\left((-\triangle_D)^{-2} + \mu \cdot \mathrm{id}_{L^2(\Omega)}\right)$ with spectrum bounded by μ.
- Hessian = "compact perturbation of one", more specifically, $(-\triangle_D)^{-2}$ is a (compact) Hilbert-Schmidt operator, for which BFGS update techniques work superlinearly also in Hilbert spaces (Sachs/Kelley).
- Important message: *function space framework allows for investigations into well-posedness and for construction of mesh independent solution algorithms.*

Optimality conditions: Weak problem formulation

$$\begin{aligned} \min J(y,u) &= \tfrac{1}{2}\|y-z\|_{L^2}^2 + \tfrac{\mu}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad a(y,p) &= b(u,p),\ \forall p \in H_0^1(\Omega) \end{aligned} \qquad \begin{aligned} a(y,p) &:= \int_\Omega \nabla y(x)^\top \nabla p(x) dx \\ b(u,p) &:= \int_\Omega u(x)p(x)dx \end{aligned}$$

Since $a(y(u),p) - b(u,p) = 0$ for all $p \in H_0^1(\Omega)$, we can add it to the objective

$$\hat{J}(u) = J(y(u),u) = J(y(u),u) + a(y(u),p) - b(u,p)$$

The necessary condition is

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0} \hat{J}(u + t\,w) \\ &= \frac{d}{dt}\Big|_{t=0} (J(y(u+t\,w), u+t\,w) + a(y(u+t\,w),p) - b((u+t\,w),p)) \\ &= J_u\,w + J_y\,y_u\,w + a_y\,(y_u\,w,p) - b_u(w,p) \\ &= J_u\,w + \left[J_y\,y_u\,w + a_y\,(y_u\,w,p)\right] - b_u(w,p)\,, \quad \forall\, w \in U. \end{aligned}$$

Notice that the expression in the square brackets is zero if there exists a $p \in H_0^1(\Omega)$ such that

$$J_y\,h + a_y\,(h,p) = 0\,, \quad \forall h \in H_0^1(\Omega)$$

We can now use the Lagrangian

$$\mathscr{L}(y,u,p) := J(y,u) + a(y,p) - b(u,p)$$

in order to state the necessary optimality conditions

$$\begin{array}{lrcl} \text{(adjoint eq.)} & 0 = \mathscr{L}_y & \Leftrightarrow & a_y(h,p) = -J_y h\,, \quad \forall h \in H_0^1(\Omega) \\ \text{(design eq.)} & 0 = \mathscr{L}_u & \Leftrightarrow & J_u w - b_u(w,p) = 0\,, \quad \forall w \in U \\ \text{(state eq.)} & 0 = \mathscr{L}_p & \Leftrightarrow & a(y,q) = b(u,q)\,, \quad \forall q \in H_0^1(\Omega) \end{array}$$

which is in strong differential form

$$\begin{array}{rcll} -\triangle p(x) &=& z(x) - y(x), & \forall x \in \Omega\,, \quad p(x) = 0, \ \forall x \in \partial\Omega \\ \mu u(x) - p(x) &=& 0\,, & \forall x \in \Omega \\ -\triangle y(x) &=& u(x), & \forall x \in \Omega\,, \quad y(x) = 0, \ \forall x \in \partial\Omega \end{array} \tag{1}$$

or in matrix operator notation:

$$\begin{bmatrix} \mathrm{id} & 0 & -\triangle_D \\ 0 & \mu \cdot \mathrm{id} & -\mathrm{id} \\ -\triangle_D & -\mathrm{id} & 0 \end{bmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} z \\ 0 \\ u \end{pmatrix}$$

which is the basis for "one-shot" computational approaches.

SQP methods:

Since all model equations can be written in variational form, Newton's method applied to the necessary conditions gives rise to so-called **SQP** methods. SQP methods have to solve linear-quadratic problems very similar to the problem above (in absence of inequality constraints).

Computational solution approaches: black box vs. one-shot

black box methods:

solve reduced problem
$\min_u \hat{J}(u) := J(y(u), u)$.

$\rightarrow$ advantage: easy implementation

$\rightarrow$ drawback: one full system solution per optimization iteration, resulting in slow methods

one-shot methods:

solve necessary conditions as one large coupled system.

$\rightarrow$ advantage: structure exploitation results in fast methods.

$\rightarrow$ drawback: some implementation effort required

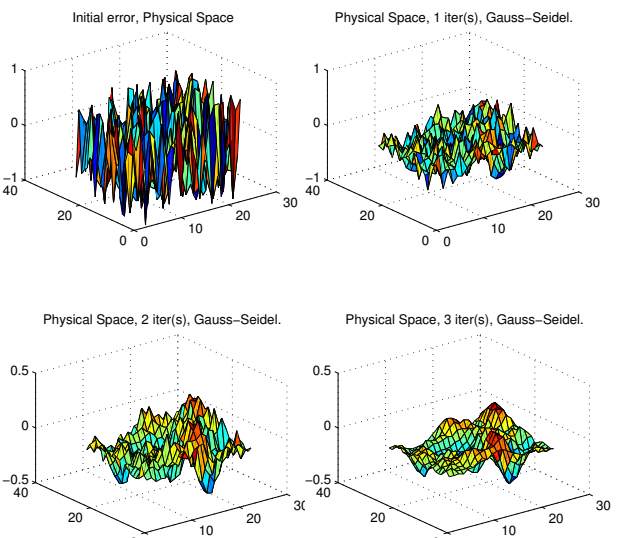
$\rightarrow$ SQP methods, optimization multigrid methods and piggy-back type methods as paradigmatic one-shot methods

Multigrid methods:

Idea based on correction schemes for equation $A_h y_h = b_h$. Starting from an initial guess y_h^0 we search a correction v_h such that the equation is solved at the next step $y_h^1 := y_h^0 + v_h$, i.e.:

$$A_h(y_h^0 + v_h) = b_h \quad \Leftrightarrow \quad v_h = -A_h^{-1}(A_h y_h^0 - b_h)$$

- A_h^{-1} is to be substituted by
  - some standard approximation: diagonal (Jacobi), lower triangle (Gauß-Seidel), etc., which smoothes the error
  - coarse grid correction $I_H^h A_H^{-1} I_h^H$, which reduces smooth errors only
- both alternate and the coarse grid correction may mean yet another application of same principle on the coarser mesh.



→ *Multigrid enables linear algorithmic complexity!*

Multigrid optimization methods: [BS12]

- iterate in multigrid manner over all variables: y, u, p
- coarse grid correction is mostly straight forward
- novel part: smoothing iteration, e.g. Gauß-Seidel:

$$\begin{bmatrix} \mathrm{I} & 0 & A_h \\ 0 & \mu \cdot I & -I \\ A_h & -I & 0 \end{bmatrix}^{-1} \text{ is approximated by } \begin{bmatrix} \mathrm{I} & 0 & L_h \\ 0 & \mu \cdot I & -I \\ L_h & -I & 0 \end{bmatrix}^{-1} \text{ with } L_h := \text{lower triangle of} A_h$$

- reordering of variables $\{y_1, y_2, \ldots, u_1, u_2, \ldots, p_1, p_2, \ldots\}$ as $\{y_1, u_1, p_1, y_2, u_2, p_2, \ldots\}$ and equations accordingly leads to collective Gauß-Seidel smoothing, where small 3×3 systems have to be solved at each grid point:

$$\begin{bmatrix} 1 & 0 & 4/h^2 \\ 0 & \mu & -1 \\ 4/h^2 & -1 & 0 \end{bmatrix} \begin{bmatrix} y_i \\ u_i \\ p_i \end{bmatrix} = \text{r.h.s.}$$

Further investigations by: A. Borzi, S. Nash, W. Zulehner, T' Asan,...

Piggy back iterations, aka approximate reduced SQP:

for ease of discussion in finite dimensions, here

constraint is assumed to be solved by a fixed-point equation (with $\rho(G_y(y,u)) < 1$):

$$y = G(y,u) \qquad \Leftrightarrow \qquad c(y,u) = 0$$

thus

$$\begin{array}{rcl} \min J(y,u) & & \min J(y,u) \\ \text{s.t. } G(y,u) - y & = & 0 \end{array} \quad \Leftrightarrow \quad \begin{array}{rcl} \min J(y,u) & & \\ \text{s.t. } c(y,u) & = & 0 \end{array}$$

new Lagrangian $L(y,u,p) := J(y,u) + p^\top (G(y,u) - y)$ gives rise to adjoint fixed-point equation (and thus iteration):

$$p = p + \nabla_y L(y,u,p)$$

which is convergent, since $\rho(G_y(y,u)^\top) = \rho(G_y(y,u)) < 1$, resulting in overall iteration:

$$\begin{aligned} u^{k+1} &= u^k - \tau B^{-1} \nabla_u \mathrm{L}(y^k, u^k, p^k) \\ y^{k+1} &= G(y^k, u^{k+1}) \\ p^{k+1} &= p^k + \nabla_y \mathrm{L}(y^{k+1}, u^{k+1}, p^k) \end{aligned}$$

B is some approximation of the reduced Hessian. The iteration can be interpreted as approximate SQP step. More details in [BGG$^+$14]

Alternatively: Krylov subspace methods for the KKT system with appropriate preconditioning (Ghattas, Sachs, Wathen,...). And POD for instationary problems.

Control and state contraints

A) Control box constraints

$$
\begin{aligned}
& \min J(y,u) = \tfrac{1}{2}\int_\Omega (y(x)-z(x))^2 dx + \frac{\mu}{2}\int_\Omega u(x)^2 dx \\
& -\triangle y(x) = u(x),\ \forall x \in \Omega \\
\text{s.t.} \quad & y(x) = 0,\ \forall x \in \Gamma := \partial\Omega \\
& u \in U_{ad} = \{\phi \in L^2(\Omega) \mid u_a(x) \le u(x) \le u_b(x)\,,\ \text{a.e.}\}
\end{aligned}
$$

problem is well posed even for $\mu = 0$

instead of the design equation in (1), we obtain the projection condition (cf. [Trö10])

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

if p is adjoint solution, u_p is a descent direction of the reduced problem which results in an overall projected gradient method.

B) General constraints [IK08, de 15]

optimality conditions together with complementarity conditions

classical active set strategies fail because of computational complexity

Remedy: primal-dual active set strategies and semi-smooth Newton methods in function spaces

adjoints to state constraints have rather weak regularity $\longrightarrow$ special regularization techniques required, e.g., Moreau-Yosida

2 PDE constrained shape optimization

- Instead of a control function $u : \Omega \to \mathbb{R}$, whole domain Ω as DOF.
- Problem: domains Ω do not constitute a linear space.
- Quick and dirty solution approach: a priori parameterization
  - less flexibility for shapes
  - numerical effort can be saved by shape calculus

Definition 2.1. *Let D be the set of all closed subsets of an open so-called "'hold-all-domain" $O \subset \mathbb{R}^d$ ($d \in \mathbb{N}$, mostly $2, 3$), i.e., $D \subset \mathcal{P}(O)$. We define for $\Omega \in D$ with $\Gamma \subset \partial\Omega$ the objective*

$$\begin{aligned} J \colon D &\to \mathbb{R} \\ \Omega &\mapsto f(\Omega) \end{aligned}$$

If constraints (in form of PDE) are present, we write

$$\begin{aligned} &\min_{(y,\Omega)} J(y,\Omega),\ y \in Y \text{ Banach space} \\ &\text{s.t. } c(y,\Omega) = 0,\ c \colon Y \times D \to Z \text{ Banach space} \end{aligned}$$

Example: Isoperimetric problem (Dido Problem)

$$\min_{\Omega} J(\Omega) = \min \int_{\Omega} -1 dx \text{ (max volume)}$$

$$\text{s.t. } c(\Omega) = \int_{\partial\Omega} 1 ds = l_0 \text{ prescribed length}$$

first solution by F. Edler 1882.

Definition 2.2. *$T_t \colon x \mapsto T_t(x)$ sufficiently smooth family of bijective mappings (e.g. perturbation of identity: $T_t(x)[V] = x + tV(x)$). We define the deformed domain Ω_t as $\Omega_t = T_t(\Omega) = \left\{ T_t(x) \middle| x \in \Omega \right\}$. For $\Omega \in D, V \in C^2\left(\mathbb{R}^d, \mathbb{R}^d\right)$, the functional $J \colon D \to \mathbb{R}$ is called shape differentiable, if the limit*

$$dJ(\Omega)[V] := \lim_{t \to 0+} \frac{J(\Omega_t) - J(\Omega)}{t}$$

exists for all V and the mapping $dJ(\Omega)[V] \colon C^2\left(\mathbb{R}^d, \mathbb{R}^d\right) \to \mathbb{R}$ is linear. Then, $dJ(\Omega)[V]$ is called the shape derivative of J in direction V.

Theorem 2.3 (Hadamard representation theorem)**.** *If $J\colon D \to \mathbb{R}$ is shape differentiable and Ω has a piecewise smooth boundary, $V \in C^k\left(\mathbb{R}^d, \mathbb{R}^d\right)$, then there is as distribution $g \in D^k(\Gamma)$ (D^k dual space to C^k), such that*

$$dJ(\Omega)[V] = \int_\Gamma \left(V, n\right) g ds$$

where $\left(V, n\right) = V^T n$, $\Gamma = \partial\Omega$, n outer normal to $\Gamma = \partial\Omega$. (proof: Sokolowski/Zolesio 1992)

Theorem 2.4. *$\Omega \subset D$ with piecewise smooth boundary, $g_1, g_2\colon D \to \mathbb{R}$, $g_3\colon D \to \mathbb{R}^d$ sufficiently smooth and* **independent** *from Ω. (D hold-on-domain)*
Then yield the shape derivatives

$$df_1(\Omega)[V] := d\left(\int_\Omega g_1(x)dx\right)[V] = \int_{\partial\Omega} \left(V, n\right) \cdot g_1(s) ds$$

$$df_2(\Omega)[V] := d\left(\int_{\partial\Omega} g_2(s)ds\right)[V] = \int_{\partial\Omega} \left(V, n\right) \cdot \left(\frac{\partial g_2}{\partial n} + \kappa g_2\right) ds$$

$$df_3(\Omega)[V] := d\left(\int_{\partial\Omega} \left(g_3(s), n(s)\right) ds\right)[V] = \int_{\partial\Omega} \left(V, n\right) \cdot \operatorname{div}(g_3) ds$$

where $\kappa = (d-1) \cdot$ mean curvature

what about PDE constraints?

- Integrands in the objective are not independent from domain!
- Each new problem needs a lenghty derivation.

Definition 2.5. *We consider a family of mappings* $\{\varphi_t : D \to \mathbb{R} \mid t \in [0, \epsilon)\}$, *which depend differentiably on the parameter* t. *Each member of the family is evaluated at* $x_t := T_t(x) \in \Omega_t$. *We call the total derivative*

$$\dot{\varphi}(x) := \frac{d}{dt}\Big|_{t=0+} \varphi_t(x_t)$$

the material derivative *with additional notation* $D_m(\varphi) := \dot{\varphi}$ *and the partial derivative the* shape derivative

$$\varphi'(x) := \frac{\partial}{\partial t}\Big|_{t=0+} \varphi_t(x)$$

Lemma 2.6. *We have* $\dot{\varphi}(x) = \varphi'(x) + \nabla\varphi^\top V$, *and*

$$\frac{\partial}{\partial t}\Big|_{t=0+} \left(\int_{\Omega_t} \varphi_t dx_t \right) = \int_\Omega \dot{\varphi} + \mathit{div}(V)\varphi dx$$

Remark: Similarly on the boundary for $z : \Gamma = \partial\Omega \to \mathbb{R}$ yields: $z' = \dot{z} - V^\top \nabla_\Gamma z$ and

$$\frac{\partial}{\partial t}\Big|_{t=0+} \left(\int_{\Gamma_t} z_t(s_t) ds_t \right) = \int_\Gamma \dot{z}(s) + z(s) \mathrm{div}_\Gamma(V(s)) ds$$

We need the following computational rules:

Theorem 2.7. *We consider material derivatives D_m of functions $\varphi, \psi : \Omega \to \mathbb{R}$ with respect to perturbation of identity. Then there holds*

(i) Product rule: $D_m(\varphi\psi) = D_m\varphi\psi + \varphi D_m\psi$

(ii) Commutation rule with gradient: $D_m(\nabla\varphi) = \nabla(D_m\varphi) - \nabla V^\top \nabla\varphi$

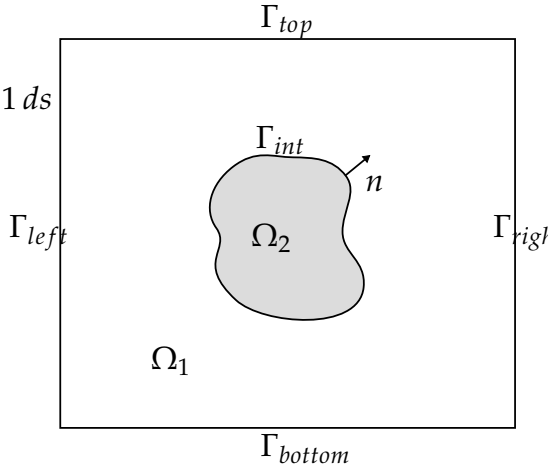
(iii) Product rule for gradients:
$D_m(\nabla\varphi^\top\nabla\psi) = \nabla D_m\varphi^\top\nabla\psi - \nabla\varphi^\top(\nabla V + \nabla V^\top)\nabla\psi + \nabla\varphi^\top\nabla D_m\psi$

Proof: [Ber10] □

Example: motivated by electrical impedance tomography

$$\min\ J(y,\Omega_2) = j(y)+j_{\text{reg}}(\Omega_2) := \frac{1}{2}\int_\Omega (y-z)^2\,dx + \mu\int_{\Gamma_{int}} 1\,ds$$

$$\text{s.t.}\ -\triangle y = f \quad \text{in } \Omega := \Omega_1 \cup \Omega_2$$
$$y = 0 \quad \text{on } \partial\Omega$$



where

$$f \equiv \begin{cases} f_1 = \text{const.} & \text{in } \Omega_1 \\ f_2 = \text{const.} & \text{in } \Omega_2 \end{cases}$$

Weak formulation

$$\min\ J(y,\Omega_2) = \frac{1}{2}\int_\Omega (y-z)^2\,dx + \mu\int_{\Gamma_{int}} 1\,ds$$
$$\text{s.t. } a(y,p) = b_{\Omega_2}(f,p) \quad \forall p \in H_0^1(\Omega)$$

gives rise to Lagrangian

$$\mathscr{L}(y,\Omega_2,p) := J(y,\Omega) + a(y,p) - b_{\Omega_2}(f,p)$$

Theorem of Correa and Seeger [DZ01] gives for the total shape derivative

$$dJ(y(\Omega_2),\Omega_2)[V] = d\mathscr{L}(y,\Omega_2,p)[V]$$

if y, p satisfy the saddle point property

$$\frac{\partial \mathscr{L}(y,\Omega_2,p)}{\partial p} = 0 = \frac{\partial \mathscr{L}(y,\Omega_2,p)}{\partial y}$$

thus, again adjoint equation:

$$\begin{aligned} -\triangle p &= z - y \text{ in } \Omega \\ p &= 0 \text{ on } \partial\Omega \end{aligned}$$

Theorem 2.8. *If the elliptic PDE and its adjoint admit a solution in $H^1(\Omega)$, the shape derivative can be expressed in volumetric form as*

$$dJ(\Omega_2)[V] =$$

$$= \int_\Omega -\nabla y^\top(\nabla V + \nabla V^\top)\nabla p - pV^\top \nabla f + \operatorname{div}(V)\left(\frac{1}{2}(y-z)^2 + \nabla y^\top \nabla p - fp\right) dx + \int_{\Gamma_{int}} \mu\kappa(V,n)$$

If $y, p \in H^2(\Omega)$, we can express this in the following boundary formulation

$$dJ(\Omega_2)[V] = \int_{\Gamma_{int}} (-[\![f]\!]p + \mu\kappa)(V,n)ds$$

Volumetric versus boundary formulation:

- The volumetric formulation requires less regularity and is more in line with finite element ansatz spaces (see below).
- The boundary formulation is preferable for analytic investigations of the solution.
- The boundary formulation seems to be more easily applicable within a boundary stepping algorithm, but this impression is misleading.

WARNING: Cea [Cea86] gives a much shorter derivation of the boundary formulation – without usage of the volumetric formulation and without the concept of a material derivative – and is surprisingly popular. By chance, this derivation gives the same expression as above. However, in general, his derivation principle leads to wrong results as illustrated in [Pan05], in particular for interface problems as below.

Parabolic interface problem

$$\min\ J(\Omega_2) = j(\Omega) + j_{\text{reg}}(\Omega_2) := \int_0^T \int_\Omega (y - z)^2 \, dx \, dt + \mu \int_{\Gamma_{int}} 1 \, ds$$

$$\begin{aligned}
\text{s.t. } \frac{\partial y}{\partial t} - \operatorname{div}(k \nabla y) &= \tilde{f} \quad \text{in } \Omega \times (0, T] \\
y &= 1 \quad \text{on } \Gamma_{top} \times (0, T] \\
\frac{\partial y}{\partial n} &= 0 \quad \text{on } (\Gamma_{bottom} \cup \Gamma_{left} \cup \Gamma_{right}) \times (0, T] \\
y &= y_0 \quad \text{in } \Omega \times \{0\}
\end{aligned}$$

where

$$k \equiv \begin{cases} k_1 = \text{const.} & \text{in } \Omega_1 \times (0, T] \\ k_2 = \text{const.} & \text{in } \Omega_2 \times (0, T] \end{cases}$$

discussed in detail in [SSW16b, SSW16a]

parabolic adjoint equation

$$\begin{aligned}
\frac{\partial p}{\partial t} - \operatorname{div}(k\nabla p) &= z - y \quad \text{in } \Omega \times [0,T) \\
p &= 0 \quad \text{on } \Gamma_{top} \times [0,T) \\
\frac{\partial p}{\partial n} &= 0 \quad \text{on } (\Gamma_{bottom} \cup \Gamma_{left} \cup \Gamma_{right}) \times [0,T) \\
p &= 0 \quad \text{in } \Omega \times \{T\}
\end{aligned}$$

Theorem 2.9. *If the parabolic PDE and its adjoint admit H^1 smoothness in each time instance, the shape derivative can be expressed in volumetric form as*

$$dJ(\Omega_2)[V] = \int_{\Gamma_{int}} \mu\kappa(V,n)ds +$$

$$\int_0^T \int_\Omega -k\nabla y^\top(\nabla V + \nabla V^\top)\nabla p - pV^\top \nabla \tilde{f} + \operatorname{div}(V)\left(\frac{1}{2}(y-z)^2 + \frac{\partial y}{\partial t}p + k\nabla y^\top \nabla p - \tilde{f}p\right) dxdt$$

If y, p admit H^2 smoothness, we can express this in the following boundary formulation

$$dJ(\Omega_2)[V] = \int_{\Gamma_{int}} \left([\![k]\!] \int_0^T \nabla y_1^\top \nabla p_2 dt + \mu\kappa \right)(V,n)dt$$

Algorithms for shape optimization

- most algorithms are based on steepest descent on the shape boundary
- boundary step needs gradient $\neq$ derivative
- gradient is representation in a scalar product $\longrightarrow$ which one?
- several choices in literature: $L^2(\Gamma), L^2(\Gamma, 1 + A\kappa), (id - A\triangle_\Gamma), ...$
- ideally it should match the shape Hessian $\longrightarrow$ in many examples an operator of order 1
- it should be compatible with the equivalence of the volumetric and boundary shape expression
- Let us use the trace of a volume deformation operator $\longrightarrow$ the inverse Poincaré-Steklov operator $(S^p)^{-1}$, where

$$\begin{aligned} S^p \colon H^{-1/2}(\Gamma_{int}) &\to H^{1/2}(\Gamma_{int}), \\ g &\mapsto (U, n) \end{aligned}$$

where $U \in H_0^1(\Omega, \mathbb{R}^d)$ solves the Neumann problem

$$a(U, V) = \int_{\Gamma_{int}} g \cdot (V, n)\, ds\,, \ \forall\ V \in H_0^1(\Omega, \mathbb{R}^d)$$

S^p based SD vs *standard scalar products S_Γ*

Initialize Γ_{int} and mesh
repeat
  solve fwd & adj prob.
  compute mesh deformation U with
    shape derivative as r.h.s.
  jointly deform mesh and Γ_{int}
until $\|U\| < TOL$
return Γ_{int}

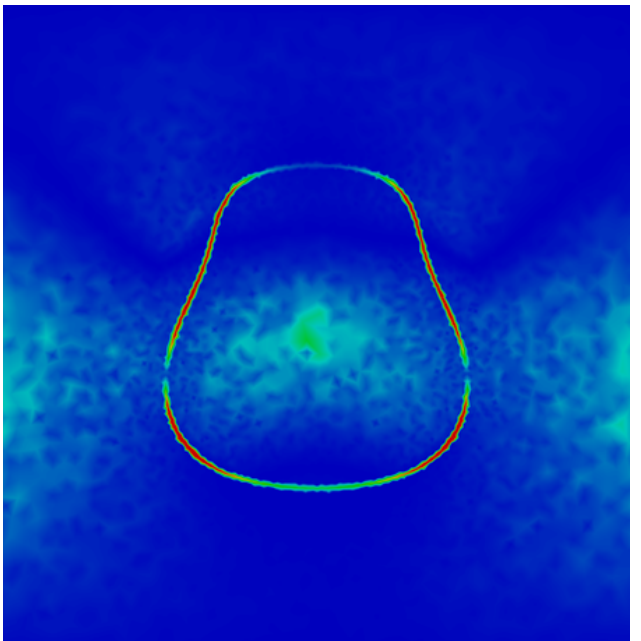
- usage of volumetric form, boundary form and mixtures consistently possible
- choice of bilinear form $a(.,.)$ not fixed
- we recommend elasticity with stiff material near optimization boundary and Lamé parameter $\lambda = 0$.
- assembly of r.h.s. in volumetric form collects noise in the domain which has to be removed

Initialize Γ_{int} and mesh
repeat
  solve fwd & adj prob.
  compute $g|_{\Gamma_{int}}$ (shape derivative)
  compute $\nabla J \leftarrow S_\Gamma^{-1} g$
  $\Gamma_{int} \leftarrow \Gamma_{int} - \tau \nabla J$
  compute mesh deformation
until $\|\nabla J\| < TOL$
return Γ_{int}

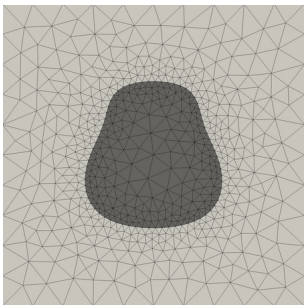
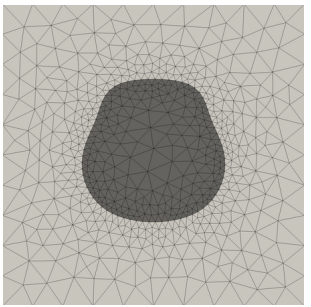
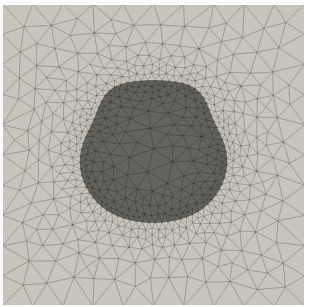
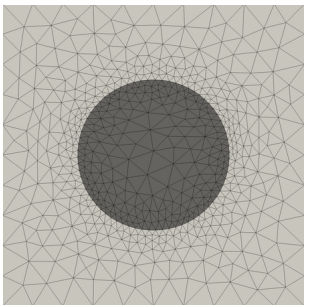
- only boundary formulation used
- mesh deformation often with Poisson or elasticity, where $-\tau \nabla J$ is Dirichlet information
- deformed mesh frequently overlaps
- tangential node movement not possible in boundary Γ_{int}

Important detail with volumetric formulation of shape derivative:

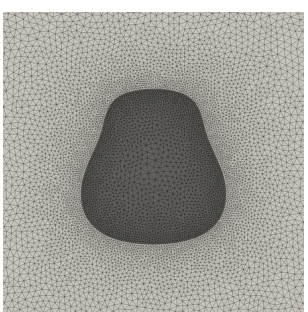
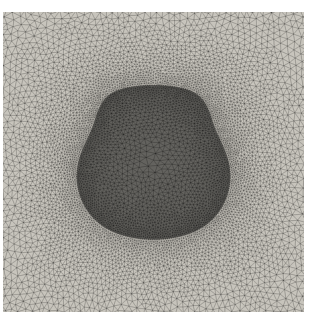
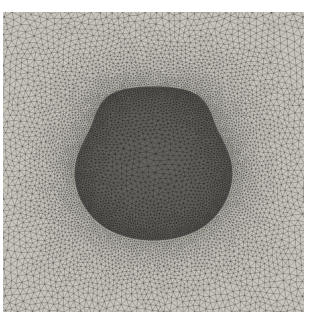
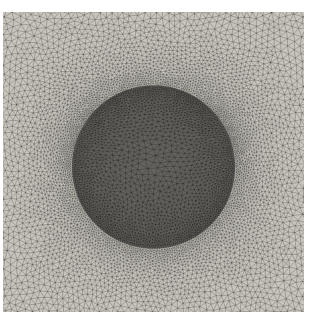
- green speckles outside of shape boundary are numerical artefacts and lead to divergence in S^p based algorithm
- should be forced to 0
- resulting algorithm converges even for coarse grid



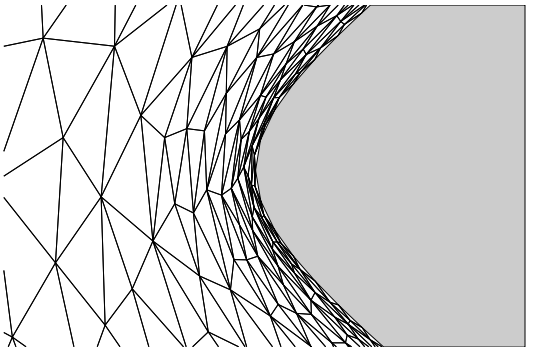
magnitude of discretized FE r.h.s.

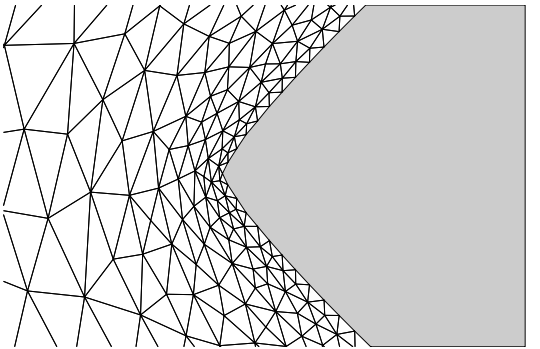
The domain-based optimization approach enables the usage of much coarser spatial discretizations than usual: approx. 1000 cells vs. 100.000 cells [SSW16a]

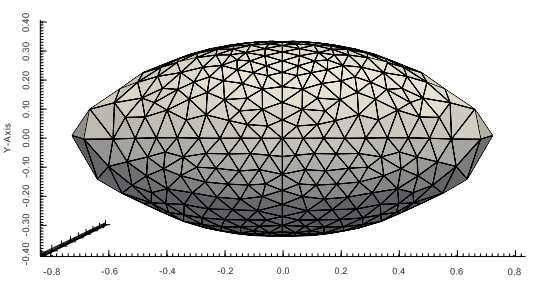
Grid comparison for shape optimization in Stokes flow [SS16]



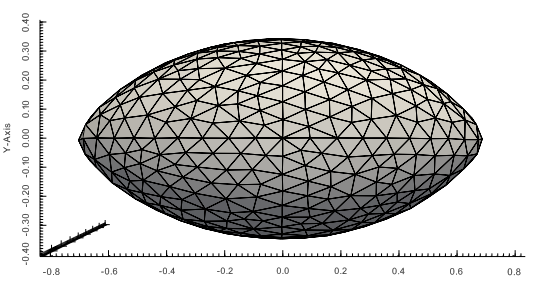
Laplace-Beltrami metric, no FEM solution possible



Steklov-Poincaré metric, converged solution



Laplace-Beltrami metric, no FEM solution possible



Steklov-Poincaré metric, converged solution

SQP for PDE constrained shape optimization

- shape space is not a vector space but can be considered a manifold (cf. [Sch14] and references therein)
- straight lines are substituted by geodesics in the following Taylor series (cf. [Sch14, AMS08])

$$J(\exp_\Omega(V_n)) = J(\Omega) + dJ(\Omega)[V_n] + \frac{1}{2}\overset{\mathcal{R}}{\mathsf{Hess}}J(\Omega)[V_n, V_n] + \mathcal{O}(\|V_n\|^3)$$

$\Rightarrow$ quadratic convergence for exact shape-Newton and superlinear convergence for shape-Quasi-Newton methods can be formulated and observed [SSW14, SSW16a, SSW16b]

$\Rightarrow$ Lagrange-Newton approach leads to linear-quadratic subproblems in a shape-SQP method in the form of an optimal control problem with Neumann boundary control [SSW14].

3 PDE constrained optimization in HPC and applications

$\longrightarrow$ see separate slides in file:

Schulz_applications.pdf

4 Hands-on playground

Task 1: Solve the optimal control problem

$$\min J(y,u) = \tfrac{1}{2}\int_\Omega (y(x)-z(x))^2 dx + \frac{\mu}{2}\int_\Omega u(x)^2 dx\,, \qquad \mu > 0$$

$$\text{s.t.} \quad \begin{aligned} -\triangle y(x) &= u(x),\ \forall x \in \Omega \\ y(x) &= 0,\ \forall x \in \Gamma := \partial\Omega \end{aligned}$$

on the unit square by two different approaches:

(1) Steepest descent method with respect to u in program `Reyes/sd.m`

(2) Multigrid optimization method in main program `MGOPT/mgopt.m`

In particular, identify the algorithmic aspects in the respective programs and called subroutines and "feel" the advantage of multigrid in terms of computing time. Feel free to play around and try out alternative options for sub-tasks in the algorithms.

Task 2: Solve the semilinear optimal control problem

$$\min J(y,u) = \tfrac{1}{2}\int_\Omega (y(x)-z(x))^2 dx + \frac{\mu}{2}\int_\Omega u(x)^2 dx\,, \qquad \mu > 0$$

$$\text{s.t.} \quad \begin{aligned} -\triangle y(x) + y(x)^3 &= u(x),\ \forall x \in \Omega \\ y(x) &= 0,\ \forall x \in \Gamma := \partial\Omega \end{aligned}$$

- Derive first order optimality conditions in weak formulation
- Formulate an SQP method (Newton method for the whole necessary conditions)
- Compare with the implementation in program `Reyes/sqp.m`
- What type of convergence do you observe [linear/superlinear/quadratic]?
- Can you generalize the programs from task 1 to this problem?

Task 3: Solve the box constrained optimal control problem

$$\min J(y,u) = \tfrac{1}{2}\int_\Omega (y(x)-z(x))^2 dx + \frac{\mu}{2}\int_\Omega u(x)^2 dx\,, \qquad \mu > 0$$

$$\begin{aligned} \text{s.t.} \qquad -\triangle y(x) &= u(x),\ \forall x \in \Omega \\ y(x) &= 0,\ \forall x \in \Gamma := \partial\Omega \\ u_a \le u(x) &\le u_b,\ \forall x \in \Omega \end{aligned}$$

- Understand and use the program `Reyes/projgrad.m`
- What is the effect of the regularization parameter μ ?
- Can you generalize the multigrid optimization method from task 1 to this problem?

Task 4: Consider the following shape optimization problems

(1) Solve $\min_{\Omega \subset \mathbb{R}^2} \int_\Omega \|x\|^2 - 1 dx$

(2)

$$\begin{aligned} \min\ J(y, \Omega_2) &:= \frac{1}{2} \int_\Omega (y - z)^2 \, dx + \mu \int_{\Gamma_{int}} 1 \, ds \\ \text{s.t. } -\triangle y + y^3 &= f \quad \text{in } \Omega := \Omega_1 \cup \Omega_2 \\ y &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

where

$$f \equiv \begin{cases} f_1 = \text{const.} & \text{in } \Omega_1 \\ f_2 = \text{const.} & \text{in } \Omega_2 \end{cases}$$

Derive the shape derivative and compare with the result in Theorem 2.8.

References

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.

[Ber10] M. Berggren. A unified discrete-continuous sensitivity analysis method for shape optimization. In W. Fitzgibbon et al., editor, *Applied and Numerical Partial Differential Equations*, volume 15 of *Computational Methods in Applied Sciences*, pages 25–39. Springer, 2010.

[BGG+14] T. Bosse, N. Gauger, A. Griewank, S. Guenther, and V. Schulz. One-shot approaches to design optimization. In *Trends in PDE Constrained Optimization*, volume 165 of *ISNM*. Birkhäuser Basel, Boston, Stuttgart, 2014.

[BS12] A. Borzì and V. Schulz. *Computational optimization of systems governed by partial differential equations*. Number 08 in SIAM book series on Computational Science and Engineering. SIAM Philadelphia, 2012.

[Cea86] J. Cea. Conception optimale ou identification de formes calcul rapide de la dérivée directionelle de la fonction coût. *RAIRO Modelisation mathématique et analyse numérique*, 20(3):371–402, 1986.

[de 15] Juan Carlos de los Reyes. *Numerical PDE-Constrained Optimization*. Springer Briefs in Optimization. Springer, 2015.

[DZ01] M. C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Analysis, Differential Calculus, and Optimization*. Advances in Design and Control. SIAM Philadelphia, 2001.

[HPUU08] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*. Number 23 in Mathematical Modelling: Theory and Applications. Springer, 2008.

[IK08] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Advances in Design and Control. SIAM, 2008.

[LBE+14] G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, and S. Ulbrich, editors. *Trends in PDE Constrained Optimization*, volume 165 of *ISNM*. Birkhäuser Basel, Boston, Stuttgart, 2014.

[LEG+12] G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich, and S. Ulbrich, editors. *Constrained Optimization and Optimal Control for Partial Differential Equations*. Number 160 in ISNM. Birkhäuser Basel, Boston, Stuttgart, 2012.

[Pan05] Olivier Pantz. Sensibilité de l'équation de la chaleur aux sauts de conductivité. *C.R. Acad. Sci. Paris, Ser. I*, 341:333–337, 2005.

[Sch14] Volker Schulz. A Riemannian view on shape optimization. *Foundations of Computational Mathematics*, 14:483–501, 2014.

[SS16] V. Schulz and M. Siebenborn. Computational comparison of surface metrics for pde constrained shape optimization. *Computational Methods in Applied Mathematics*, 2016. `http://arxiv.org/abs/1509.08601` (to appear).

[SSW14] Volker Schulz, Martin Siebenborn, and Kathrin Welker. Towards a Lagrange-Newton approach for PDE constrained shape optimization. In *Trends in PDE Constrained Optimization*, volume 165 of *International Series of Numerical Mathematics*. Springer, 2014. `http://arxiv.org/abs/1405.3266`.

[SSW16a] Volker Schulz, Martin Siebenborn, and Kathrin Welker. Efficient PDE constrained shape optimization based on Steklov-Poincaré type metrics. *SIAM Control*, 2016. `http://arxiv.org/abs/1506.02244` (in print).

[SSW16b] Volker Schulz, Martin Siebenborn, and Kathrin Welker. Structured inverse modeling in parabolic diffusion problems. *SIAM Control*, 2016. `http://arxiv.org/abs/1409.3464` (in print).

[Trö10] Fredi Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*. AMS, 2010.