

Research Article

Volker Schulz* and Martin Siebenborn

Computational Comparison of Surface Metrics for PDE Constrained Shape Optimization

DOI: 10.1515/cmam-2016-0009

Received December 10, 2015; revised February 5, 2016; accepted February 10, 2016

Abstract: We compare surface metrics for shape optimization problems with constraints, consisting mainly of partial differential equations (PDE), from a computational point of view. In particular, classical Laplace–Beltrami type metrics are compared with Steklov–Poincaré type metrics. The test problem is the minimization of energy dissipation of a body in a Stokes flow. We therefore set up a quasi-Newton method on appropriate shape manifolds together with an augmented Lagrangian framework, in order to enable a straightforward integration of geometric constraints for the shape. The comparison is focussed towards convergence behavior as well as effects on the mesh quality during shape optimization.

Keywords: PDE constrained shape optimization, optimization on Riemannian manifolds, Stokes problem

MSC 2010: 49Q10, 65K10, 65N30

1 Introduction

Shape optimization is a challenging field with many interesting applications. As examples, we mention aerodynamic shape optimization [14], acoustic shape optimization [20] or optimization of interfaces in transmission problems [6, 10, 11]. The general structure of such a PDE constrained shape optimization problem is of the form

$$\min_{y, \Omega} J(y, \Omega) \quad \text{subject to} \quad B_{\Omega}(y) = 0, \quad c(y, \Omega) = 0,$$

where $\Omega \subset D$ is an open subset of a hold-all domain $D \subset \mathbb{R}^d$ and J is a real valued functional. Usually only some part of the boundary $\Gamma \subset \partial\Omega$ is free for optimization. The exact description of the set from which Γ is taken is given in Section 3. The constraint $B_{\Omega}(y) = 0$ denotes a boundary value problem defined on the domain Ω given in the form of equations in appropriate function spaces, where y is the solution of the boundary value problem B_{Ω} consisting of a set of partial differential equations together with some boundary conditions. Furthermore $c(y, \Omega)$ denotes a finite number of sufficiently smooth constraints. The additional constraints may be even in the form of inequalities, but this aspect is not in the focus of this paper.

In principle, there are two major conceptual approaches: the direct parametrization approach, which a priori parameterizes the shape Γ to be optimized, e.g., within a CAD framework, and the shape calculus approach operating in shape spaces. The direct parameterization approach suffers from obvious limitations with respect to the reachable geometries, but can be embedded within a vector space framework which simplifies the numerical treatment significantly and thus enables the application of classical methods for PDE constrained optimization [2]. The shape calculus approach has received strong attention in particular concerning its theoretical framework in the form of the calculus for generating shape derivative information. Several books and publications deal with this aspect in detail, e.g., [4, 19]. In contrast to that, only relatively few publications study computational aspects of shape optimization based on the shape calculus.

*Corresponding author: Volker Schulz: Department of Mathematics, University of Trier, 54286 Trier, Germany, e-mail: volker.schulz@uni-trier.de

Martin Siebenborn: Department of Mathematics, University of Trier, 54286 Trier, Germany, e-mail: siebenborn@uni-trier.de

In particular, studies on proper shape metrics for usage within gradient-type methods and also on the efficient numerical treatment of additional constraints are missing. In most cases, steepest descent methods are applied involving representations of the shape derivative, which are based on the L^2 scalar product or a Laplace–Beltrami scalar product or combinations thereof. In [16], a novel scalar product based on the Steklov–Poincaré operator is introduced in the context of shape optimization and shown to possess the following advantageous properties: volumetric and boundary expressions of the shape derivative can be treated in a consistent manner; and the resulting shape manifold admits kinks and is complete. Furthermore, this scalar product is in line with the coercivity results for shape Hessians in [5] for elliptic problems.

Therefore, we compare in this paper the novel metric introduced in [16] with the metric used so far in many publications from a computational perspective. The test case is the computation of a shape embedded in a Stokes flow which minimizes drag and satisfies further geometric constraints. The optimal solution is well known as the so-called Haack ogive [7, 12]. The geometric constraints are taken into account within an augmented Lagrangian framework similarly to [5]. Furthermore, quasi-Newton techniques are applied in order to accelerate the convergence of the shape optimization scheme. A major issue of the comparison of metrics is the surface and volume mesh quality of the deformed computational mesh. We observe that the mesh quality only mildly deteriorates during the optimization iterations based on the Steklov–Poincaré metric, although the overall deformation is considerably large. In contrast to that, the mesh quality drastically deteriorates during iterations with standard metrics.

This paper is organized in the following way. Section 2 gives details on the specific formulation used as a benchmark for the comparison of metrics. In Section 3, we introduce the manifold point of view on shape optimization as well as the two specific metrics to be compared in later sections. Section 4 introduces an augmented Lagrangian formulation for an efficient treatment of the geometric constraints involved in the test case. Finally, the results on the computational comparison of the two types of metrics for shape optimization are presented in Section 5. Conclusions are drawn in Section 6.

2 Problem Formulation: Optimal Shapes in Stokes Flows

We consider incompressible flow which is dominated by viscous forces around an obstacle described by the Stokes equations. The aim is to shape a d -dimensional body such that the energy dissipation of the system is minimized under certain geometrical constraints. For the optimization to be reasonable, the volume and the barycenter of the body are required to be constant. Here, the dimension of the problem is fixed to $d = 2$ or $d = 3$. This situation is visualized in Figure 1, where $\Omega \subset D \subset \mathbb{R}^d$ is the obstacle and Γ its boundary, which is considered to be variable. $\Omega_{\text{ext}} \subset \mathbb{R}^d$ denotes the flow field. This is the domain for the finite elements, whereas Ω is a hole in the discretization mesh. For the geometric restrictions we need to compute the body's volume

$$\text{vol}(\Omega) = \int_{\Omega} 1 \, dx \in \mathbb{R}$$

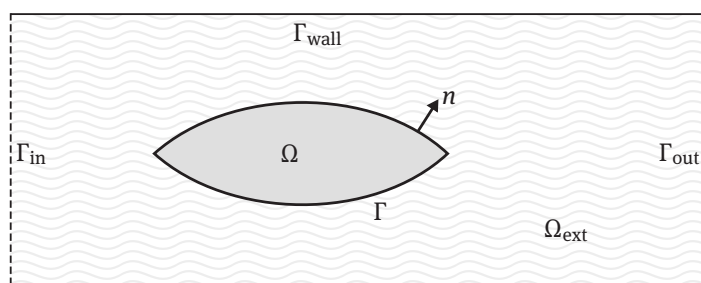


Figure 1. Schematic view of the flow field and the variable shape Ω encircled by Γ .

and its barycenter

$$\text{bc}^\Omega = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} x \, dx \in \mathbb{R}^d.$$

Finally, we end up with the PDE constraint shape optimization problem

$$\min_{(v, \Omega)} J(v, \Omega) = \int_{\Omega_{\text{ext}}} \sum_{i,j=1}^d \left(\frac{\partial v_i}{\partial x_j} \right)^2 dx \quad (2.1)$$

subject to the Stokes equations, where the viscosity is normalized to 1,

$$\begin{cases} \Delta v + \nabla p = -f & \text{in } \Omega_{\text{ext}}, \\ \text{div } v = 0 & \text{in } \Omega_{\text{ext}}, \\ v = v_{\infty} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}, \\ v = 0 & \text{on } \Gamma_{\text{wall}} \cup \Gamma, \end{cases} \quad (2.2)$$

and the geometric constraints, i.e. barycenter and volume,

$$\begin{cases} c_1(\Omega) = \text{bc}(\Omega)_1 - \text{bc}(\Omega_0)_1, \\ \vdots \\ c_d(\Omega) = \text{bc}(\Omega)_d - \text{bc}(\Omega_0)_d, \\ c_{d+1}(\Omega) = \text{vol}(\Omega) - \text{vol}(\Omega_0). \end{cases} \quad (2.3)$$

In (2.2), $v : \Omega_{\text{ext}} \rightarrow \mathbb{R}^d$ denotes the velocity and $p : \Omega_{\text{ext}} \rightarrow \mathbb{R}$. For the corresponding weak formulation of the Stokes equation we assume $v \in H^1(\Omega_{\text{ext}})^d$ and $p \in L_{2,0}(\Omega_{\text{ext}}) := \{q \in L_2(\Omega_{\text{ext}}) : \int_{\Omega_{\text{ext}}} q \, dx = 0\}$. Since we neglect body forces like gravity, we can set $f = 0$.

For a gradient based optimization we need derivatives with respect to the shape which are defined in the following way. The shape derivative in direction of a smooth vector field $V : \Omega \rightarrow \mathbb{R}^d$ is defined as

$$dJ(\Omega)[V] := \lim_{h \rightarrow 0+} \frac{J(\Omega_h) - J(\Omega)}{h},$$

where $\Omega_h = \{x + h \cdot V(x) : x \in \Omega\}$ is perturbed according to V . For the particular setting in (2.1) and (2.2) the derivative of J subject to the Stokes equation is given by (cf. [9])

$$dJ(\Omega)[V] = - \int_{\Gamma} \langle n, V \rangle \sum_{i=1}^d \left(\frac{\partial v_i}{\partial n} \right)^2 ds. \quad (2.4)$$

The derivatives of the geometric constraints can be derived by applying the calculus developed in [4] to c_1, \dots, c_{d+1} yielding

$$\begin{aligned} dc_i(\Omega)[V] &= -\frac{1}{(\text{vol}(\Omega))^2} \int_{\Gamma} \langle n, V \rangle ds \int_{\Omega} x_i \, dx + \frac{1}{\text{vol}(\Omega)} \int_{\Gamma} x_i \langle n, V \rangle ds \\ &= \frac{1}{\text{vol}(\Omega)} \int_{\Gamma} (x_i - \text{bc}(\Omega)_i) \langle n, V \rangle ds \end{aligned}$$

and

$$dc_{d+1}(\Omega)[V] = \int_{\Gamma} \langle n, V \rangle ds.$$

In order to formulate an augmented Lagrangian method for shape optimization in Section 4, derivatives of the squared constraints are also required. By applying the chain rule, we obtain

$$dc_i^2(\Omega)[V] = 2(\text{bc}(\Omega)_i - \text{bc}(\Omega_0)_i) \frac{1}{\text{vol}(\Omega)} \int_{\Gamma} (x_i - \text{bc}(\Omega)_i) \langle n, V \rangle ds$$

and

$$dc_{d+1}^2(\Omega)[V] = 2(\text{vol}(\Omega) - \text{vol}(\Omega_0)) \int_{\Gamma} \langle n, V \rangle ds.$$

3 Metrics on the Manifold of Feasible Shapes

In [15], it is pointed out that shape optimization can be viewed as optimization on Riemannian shape manifolds and resulting optimization methods can be constructed and analyzed within this framework, which combines algorithmic ideas from [1] with the differential geometric point of view established in [8]. Let us study connected and compact curves Γ as in Figure 1. Although the optimal solution of our test problem has two kinks, we first consider smooth closed curves for a better understanding of the underlying metrics.

In [8], this set of smooth closed curves is characterized by

$$B_e(S^1, \mathbb{R}^2) := \text{Emb}(S^1, \mathbb{R}^2) / \text{Diff}(S^1),$$

i.e., as the set of all equivalence classes of C^∞ embeddings of S^1 into the plane ($\text{Emb}(S^1, \mathbb{R}^2)$), where the equivalence relation is defined by the set of all C^∞ re-parameterizations, i.e., diffeomorphisms of S^1 into itself ($\text{Diff}(S^1)$). A particular point on the manifold $B_e(S^1, \mathbb{R}^2)$ is represented by a curve $\Gamma : S^1 \ni \theta \mapsto \Gamma(\theta) \in \mathbb{R}^2$. Because of the equivalence relation ($\text{Diff}(S^1)$), the tangent space is isomorphic to the set of all normal C^∞ vector fields along c , i.e.,

$$T_\Gamma B_e \cong \{h : h = \alpha n, \alpha \in C^\infty(S^1, \mathbb{R})\}$$

where n is the unit exterior normal field of the shape Γ such that $n(\theta) \perp \Gamma'(\theta)$ for all $\theta \in S^1$ and Γ' denotes the circumferential derivative as in [8]. Several intrinsic metrics are discussed in [8], among which the following Sobolev metric is used in most algorithmic approaches to shape optimization based on the shape calculus. For $A > 0$, the Sobolev metric is induced by the scalar product

$$g^1 : T_\Gamma B_e \times T_\Gamma B_e \rightarrow \mathbb{R}, \quad (h, k) \mapsto ((\text{id} - A\Delta_\Gamma)\alpha, \beta)_{L^2(\Gamma)},$$

where $h = \alpha n$ and $k = \beta n$ denote two elements from the tangent space at γ and Δ_Γ denotes the Laplace–Beltrami operator on the surface Γ . In [8], it is shown that the condition $A > 0$ guarantees that the scalar product g^1 defines a Riemannian metric on B_e and thus, geodesics can be used to measure distances.

In [16], the following scalar product g^S on the tangent space is proposed:

$$g^S : H^{1/2}(\Gamma) \times H^{1/2}(\Gamma) \rightarrow \mathbb{R}, \quad (\alpha, \beta) \mapsto \langle \alpha, (S^p)^{-1}\beta \rangle = \int_\Gamma \alpha(s) \cdot [(S^p)^{-1}\beta](s) \, ds,$$

where the symmetric and coercive operator S^p is defined by

$$S^p : H^{-1/2}(\Gamma) \rightarrow H^{1/2}(\Gamma), \quad \alpha \mapsto U^\top n$$

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

$$a(U, V) = \int_\Gamma \alpha \cdot V^\top n \, ds \quad \text{for all } V \in H_0^1(\Omega, \mathbb{R}^d), \quad (3.1)$$

where a is a coercive and symmetric bilinear form, defined for instance by the elasticity equation, and thus corresponds to an elliptic problem with fixed outer boundary and forces $\alpha \cdot n$ at the inner boundary Γ_{in} .

With the shape space B_e and a Riemannian metric on its tangent space in hand we can form the Riemannian shape gradient as a Riesz representation of a shape derivative given in the form

$$dJ(\Omega)[V] = \int_\Gamma \gamma \langle V, n \rangle \, ds.$$

In our model setting, the objective function J is given in (2.1) and its shape derivative in (2.4). The Riemannian shape gradient $\text{grad } J$ with respect to a Riemannian metric $g \in \{g^1, g^S\}$ is then obtained by

$$\text{grad } J = \gamma n \quad \text{with} \quad g(v, \alpha) = \int_\Gamma \gamma(s) \alpha(s) \, ds \quad \text{for all } \alpha n \in T_\Gamma B_e.$$

The metric g^1 , which is also used in [17] and in many other publications, necessitates a shape derivative in Hadamard form as well as efficient means to solve linear systems involving the Laplace–Beltrami operator in surfaces. All of that requires computational overhead. Furthermore, the surrounding mesh for the computation of the Stokes flow has to be deformed according to the geometry change. This mesh deformation is typically performed by the solution of an elasticity or Poisson problem, with the geometry step as a Dirichlet condition.

In contrast to that, the usage of the metric g^S corresponds to interpreting the shape derivative as a (volumetric or boundary) force to the mesh deformation process as illustrated in [16]. Thus, not only overhead is saved, but also better overall mesh properties are obtained. This paper is devoted to the detailed computational comparison of both metrics.

Furthermore, it is known that the Riemannian manifold (B_e, g^1) is not metrically complete and the solution of our test problem is not contained in it. On the other hand, according to [16], the metric g^S gives rise to the space

$$\mathcal{B}_i^{1/2}(\Gamma_0) := \mathcal{H}_i^{1/2}(\Gamma_0, \mathbb{R}^d) / \text{Homeo}(\Gamma_0), \quad (3.2)$$

where $\text{Homeo}(\Gamma_0)$ denotes the set of all homeomorphisms of the prior shape Γ_0 and

$$\begin{aligned} \mathcal{H}_i^{1/2}(\Gamma_0, D) &:= \{w : \Gamma_0 \rightarrow D \mid \text{there exists } W \in H^1(D, D), \text{ } W \text{ injective, continuous such that } W|_{\Gamma_0} = w\}, \\ D &= \Omega \cup \Gamma \cup \Omega_{\text{ext}}. \end{aligned}$$

Thus, the construction of $\mathcal{B}_i^{1/2}(\Gamma_0)$ is in complete analogy to the construction in [8]. This larger shape space (3.2) contains the optimal solution of our test problem.

4 Augmented Lagrangian Method for Shape Optimization

The general problem formulation in Section 2 includes an objective functional together with a system PDE and geometric constraints. Shape optimization problems treated by means of the shape calculus are not often set within a framework of additional constraints. In our test case, however, the geometric constraints are necessary in order to obtain nontrivial solutions: without the volume constraint c_{d+1} , the shape would shrink to a straight line as a trivial but not interesting solution, and without the barycenter constraints $c_{1,\dots,d}$, the shape would just float out of the computational domain, which again is not a desirable solution. The focus of this paper is on the comparison of the effects of metrics on shapes computed by an optimization process. Thus, we try to keep the optimization framework conceptually as simple as possible and dispense with the potential of additional algorithmic efficiency of one-shot methods [2, 14, 18], but rather stick with the conceptually simpler and otherwise widely used black-box approach. Therefore, we exploit the assumption that the model equation $B_\Omega(y) = 0$ can be solved uniquely for y , if the shape Ω is given, which itself depends uniquely on the variable boundary part Γ . Thus, we consider the system solution y as function of Γ , i.e., $y = y(\Gamma)$, which results in the following reduced problem formulation:

$$\min_{\Gamma} J(y(\Gamma)) \quad (4.1)$$

$$\text{subject to } c(\Gamma) = 0. \quad (4.2)$$

Indeed, the shape derivative in (2.4) is already the shape derivative of (4.1) with respect to Γ obeying the chain rule via the implicit function theorem. We note that in this particular self-adjoint combination of PDE and objective we do not need any adjoint equation, which is otherwise usually necessary.

In principle, the optimization problem (4.1), (4.2) can be treated by a sequential quadratic programming (SQP) approach on shape manifolds. This would result in the necessity to compute not only the Riesz representation of the shape derivative of the objective, but also of the $(d + 1)$ constraints in each optimization step, which increases the algorithmic complexity significantly, since a PDE – on the surface (g^1) or in the volume (g^S) – has to be solved for each of these Riesz representations. This problem can be circumvented by an

augmented Lagrangian approach based on the so-called augmented Lagrangian (cf. [3])

$$\mathcal{L}_A(\Gamma, \lambda) := J(y(\Gamma)) + \lambda^\top c(\Gamma) + \frac{\mu}{2} c(\Gamma)^\top c(\Gamma)$$

for some Lagrange multipliers $\lambda \in \mathbb{R}^{d+1}$ to be determined iteratively below and some $\mu > 0$ sufficiently large. The augmented Lagrangian approach has already been used in [5] in the context of shape calculus. Since the convergence for λ is very fast (cf. Figure 7), we concentrate in Section 5 on the local convergence in step 1 for given true multipliers λ .

Thus, we use the following algorithmic outline for both metrics to be compared.

0. Initialize λ^1 , $k := 1$, choose tolerance δ_f for the optimization process of the augmented Lagrangian and tolerance δ_c for the satisfaction of constraints c . Choose penalty increment factor $\mu_{\text{inc}} > 1$.
1. Solve $\Gamma_k := \operatorname{argmin}_\Gamma \mathcal{L}_A(\Gamma, \lambda)$ up to tolerance δ_f .
2. If $\|c(\Gamma_k)\| > \delta_c$
 - (a) update $\mu \leftarrow \mu_{\text{inc}} \mu$ and go to 1;
 - else
 - (b) update $\lambda^{k+1} \leftarrow \lambda^k + \mu c(\Gamma_k)$ and $k \leftarrow k + 1$.
3. If λ^k is not converged, go to 1.

The inner optimization step 1 is chosen as either steepest descent method or as limited memory BFGS-quasi-Newton method as described in [17] for the metric g^1 and in [16] for the metric g^S . We note that the discussion in Section 5 focusses on this step 1.

For the sake of completeness, we rephrase these optimization strategies in the current framework. The quasi-Newton approaches rely on the secant condition on manifolds as in [1] for a step $\Gamma_{j+1} := R_{\Gamma_j}(\eta)$ resulting from an increment $\eta_j \in T_{\Gamma_j} \mathcal{B}_i^{1/2}(\Gamma_0)$ in iteration j via a retraction R as

$$\operatorname{grad} \mathcal{L}_A(\Gamma_{j+1}) - \mathcal{T}_{\eta_j} \operatorname{grad} \mathcal{L}_A(\Gamma_j) = G_{j+1}[\mathcal{T}_{\eta_j} \eta_j],$$

where

$$\mathcal{T} : T\mathcal{B}_i^{1/2}(\Gamma_0) \oplus T\mathcal{B}_i^{1/2}(\Gamma_0) \rightarrow T\mathcal{B}_i^{1/2}(\Gamma_0), \quad (h_\Gamma, k_\Gamma) \mapsto \mathcal{T}_{h_\Gamma} k_\Gamma$$

is a vector transport associated to the retraction R and G_{j+1} is intended to approximate the Riemannian Hessian $\nabla \operatorname{grad} \mathcal{L}_A(\Gamma_{j+1})$.

In the following limited BFGS loop, we use the notation

$$s_j := \mathcal{T}_{\eta_j} \eta_j \in T_{\Gamma_{j+1}} \mathcal{B}_i^{1/2}(\Gamma_0), \quad y_j := \operatorname{grad} \mathcal{L}_A(\Gamma_{j+1}) - \mathcal{T}_{\eta_j} \operatorname{grad} \mathcal{L}_A(\Gamma_j) \in T_{\Gamma_{j+1}} \mathcal{B}_i^{1/2}(\Gamma_0).$$

In [13], superlinear convergence properties for BFGS-quasi-Newton methods on manifolds are analyzed for the case that \mathcal{T}_{η_j} is an isometry. This requirement is satisfied, e.g., if \mathcal{T} and R are the parallel transport and the exponential map. Details on the specific operators \mathcal{T}, R used here are given in Section 5.

Thus steepest descent methods and quasi-Newton methods, used here, are described jointly by Algorithm 1.

If $m = 0$ in Algorithm 1, the algorithm boils down to steepest descent methods. Note that the algorithm only describes the deformation of the shape Γ . The surrounding mesh is deformed according to an elasticity equations in the case $g = g^1$ and in the case $g = g^S$ by the usage of the elastic deformation field which is anyway available from the computation of q , such that the shape derivative is interpreted as a boundary force in the latter case.

5 Numerical Results

We are now prepared to describe a specific test case for the algorithms outlined in the previous sections. The computational domain, as depicted in Figure 1, is chosen to be $\Omega_0 \cup \Omega_{\text{ext}} = [-3, 6] \times [-2, 2]$ for the two-dimensional case. The initial body Ω_0 is a circle with barycenter $\operatorname{bc}(\Omega_0) = (0, 0)$ and radius $r = 0.5$ leading to

Algorithm 1. Details of step 1.

```

repeat
   $\rho_j \leftarrow g(y_j, s_j)^{-1}$ 
   $q \leftarrow \text{grad } \mathcal{L}_A(\Gamma_j)$ 
  for  $i = j - 1, \dots, j - m$  do
     $s_i \leftarrow \mathcal{T}_q s_i$ 
     $y_i \leftarrow \mathcal{T}_q y_i$ 
     $\alpha_i \leftarrow \rho_i g(s_i, q)$ 
     $q \leftarrow q - \alpha_i d_i$ 
  end for
   $z \leftarrow \text{grad } \mathcal{L}_A(\Gamma_j)$ 
   $q \leftarrow \frac{g(y_{j-1}, s_{j-1})}{g(y_{j-1}, y_{j-1})} \text{grad } \mathcal{L}_A(\Gamma_j)$ 
  for  $i = j - m, \dots, j - 1$  do
     $\beta_i \leftarrow \rho_i g(y_i, z)$ 
     $q \leftarrow q + (\alpha_i - \beta_i) s_i$ 
  end for
   $\Gamma_{j+1} \leftarrow R_{\Gamma_j}(q)$ 
until  $\|q\|_2 < \delta_j$ 

```

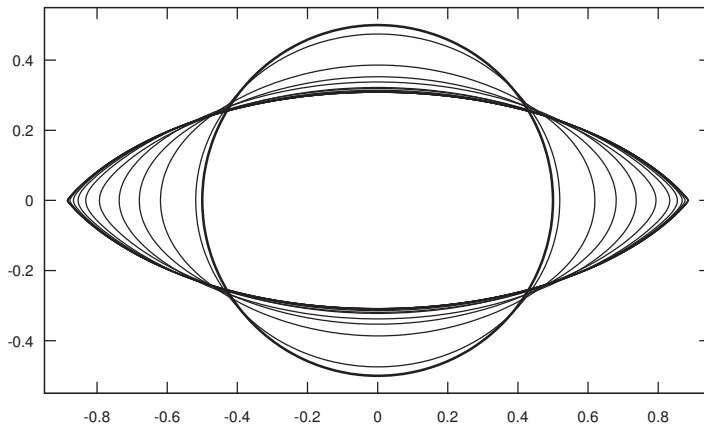


Figure 2. Iterations of the BFGS method with Steklov–Poincaré metric (initial and optimal shapes are bold).

$\text{vol}(\Omega_0) = \frac{\pi}{4}$. The computational grid consists of 10,150 triangles of which 633 form the variable surface Γ . Analogously, the three-dimensional domain is generated by rotating the two-dimensional mesh around the X axis. Here we have 27,892 tetrahedral elements with 1,206 triangles forming Γ .

Compared to the three-dimensional mesh we have chosen a much finer discretization for the two-dimensional case. This is due to the fact that we want to measure convergence of the proposed algorithm which is only practical in 2D. The coarse three-dimensional discretization shall demonstrate the ability of the shape metric g^S to also handle sharp edges in this situation. This can be seen in Figure 5. Here we again observe that the g^S metric is superior compared to g^1 with respect to the node distribution on the surface of the body. While the solution on the right-hand side is converged to the actual solution, the g^1 -algorithm on the left-hand side breaks down too early with unfeasible grids.

It should be remarked that there is no mesh inside Ω . Thus, we apply the divergence theorem to the constraints c and obtain

$$\text{vol}(\Omega) = - \int_{\Gamma} \frac{1}{d} \langle s, n \rangle ds \quad \text{and} \quad \text{bc}(\Omega) = - \frac{1}{2 \text{vol}(\Omega)} \int_{\Gamma} \langle (s_1^2, \dots, s_d^2)^T, n \rangle ds.$$

One could also constrain the volume and barycenter of Ω_{ext} to be constant which is equivalent since the outer boundaries are fixed.

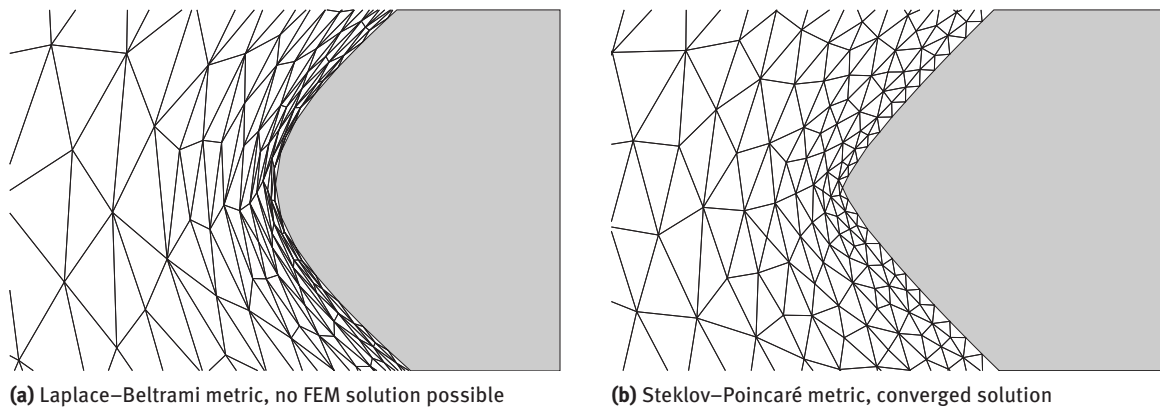


Figure 3. Visual comparison of shape metrics with respect to influence on mesh quality (tip of the body), left: 65 gradient steps, right: 16 BFGS steps.

From a computational point of view, the algorithms for the g^S metric are favorable compared to the g^1 metric. In any case, we discretize the velocity v with continuous piece-wise quadratic (P2) elements and the pressure p with continuous piece-wise linear (P1) elements in order to guarantee stability, which is the standard approach for the Stokes equation. This discretization leads to a major advantage of the g^S metric. Since v is represented in P2 functions, we end up with a discontinuous piece-wise linear representation for the term $\sum_{i=1}^d (\frac{\partial v_i}{\partial n})^2$ in (2.4). However, parts of the shape derivative coming from the geometric constraints are piece-wise linear due to the P1 shape functions of the straight-lined elements. In the g^1 case we thus have to perform L^2 -projections in order to represent all quantities in the same basis functions. Given a function $u \in V_1$ in a finite element space which has to be projected into a different space V_2 we have to solve

$$\int_{\Gamma} \bar{u} v \, ds = \int_{\Gamma} u v \, ds \quad \text{for all } v \in V_2$$

for \bar{u} . Whereas in the g^S case these terms only show up on the right-hand side of (3.1). It is thus straightforward to combine expressions which are represented in different basis functions. By solving the Neumann problem (3.1), we obtain a representation of the shape derivative with respect to the g^S metric. However, the solution U of (3.1) is not only defined on Γ but also in the entire domain Ω_{ext} . This gives us a smooth deformation field which can be applied to the finite element mesh. In contrast, using the g^1 metric, we have to solve a tangential Laplace equation yielding a representation of the shape gradient only defined at Γ . In an additional step this information is plugged into a linear elasticity problem as Dirichlet boundary condition.

Comparing the computational costs of the two algorithms, we additionally have to solve two PDEs defined at Γ in the g^1 case. This stems from the fact that gradient representation and mesh deformation are computed in one step by using the g^S metric.

Secondly, what makes the Steklov–Poincaré metrics an even more appealing concept is the mesh quality that results from these two approaches. The classical approach, which is based on the g^1 metric, only takes deformations normal to Γ into account. This reflects the Hadamard theorem stating that only the normal component of deformations affect the objective function. However, using metrics of Steklov–Poincaré type also allows surface nodes to slide along Γ which strongly influences the mesh quality. This can be seen in Figure 3. In this particular situation, a solution of the optimization problem (2.1) subject to (2.2) and (2.3) can not be achieved with the classical approach. These iterations lead to discretization meshes that do not allow further computations due to degenerated cells.

Figure 4 shows the decreasing mesh quality during the optimization iterations. Here we assume the Lagrange multipliers to be known and start the optimization with the circle geometry. The mesh quality is measured with respect to the condition of the affine mapping between reference and physical element. Values close to one indicate a good mesh quality. Here the worst element is shown. For the Steklov–Poincaré type metric we achieve mesh qualities which are within $[2.47, 3.87]$ during the entire optimization. Whereas, the

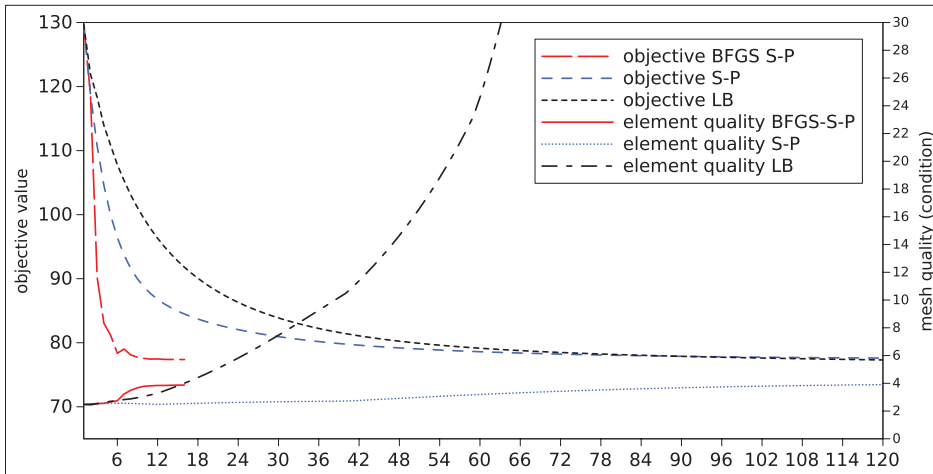


Figure 4. Objective values and mesh quality for Steklov–Poincaré and Laplace–Beltrami metric.

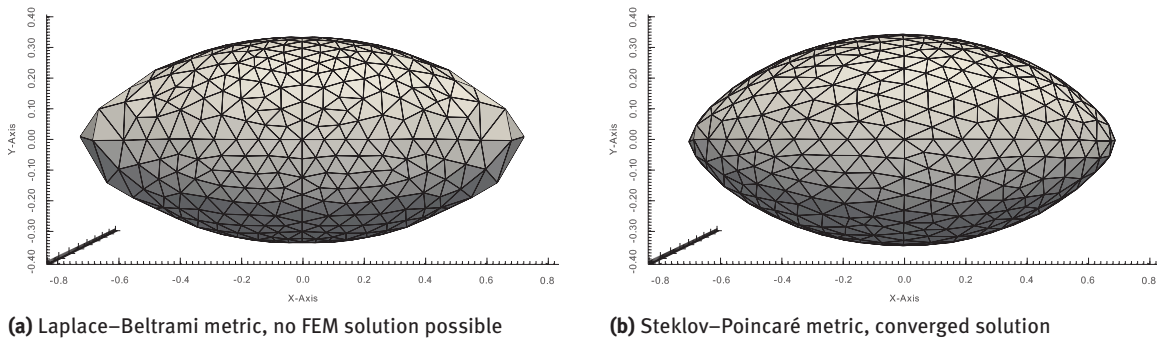


Figure 5. 3D comparison of g^1 after 62 gradient steps and g^S metric with BFGS updates after 16 iterations.

classical approach leads to degenerated elements, which can be observed by the unbound condition numbers. This behavior makes it impossible to apply BFGS updates to the Laplace–Beltrami type algorithm as the algorithm breaks down in the first iterations with unfeasible elements due to the larger step sizes. In contrast, Figure 4 also shows the speedup that can be gained by with a quasi-Newton method. In this particular case we applied a limited memory BFGS strategy with three gradients in storage. A more detailed description of quasi-Newton updates for shape optimization can be found in [16]. The corresponding 16 iterated shapes are shown in Figure 2.

The key part of the method described here is the choice of the bilinear form in (3.1). It turns out that the mesh quality heavily depends on a . The results shown in this work are obtained by choosing a as the weak form of the linear elasticity equations

$$\begin{cases} \operatorname{div}(\sigma) = 0 & \text{in } \Omega_{\text{ext}}, \\ U = 0 & \text{on } \Gamma_{\text{out}} \cup \Gamma_{\text{in}} \cup \Gamma_{\text{wall}}, \\ \frac{\partial U}{\partial n} = \gamma & \text{on } \Gamma \end{cases}$$

in terms of

$$\sigma := \lambda_{\text{elas}} \operatorname{Tr}(\epsilon) I + 2\mu_{\text{elas}} \epsilon, \quad \epsilon := \frac{1}{2}(\nabla U + \nabla U^T),$$

where σ and ϵ are the strain and stress tensor, respectively. Here λ_{elas} and μ_{elas} denote the Lamé parameters, which can be expressed in terms of Young's modulus E and Poisson's ratio ν as

$$\lambda_{\text{elas}} = \frac{\nu E}{(1 + \nu)(1 - 2\nu)}, \quad \mu_{\text{elas}} = \frac{E}{2(1 + \nu)}.$$

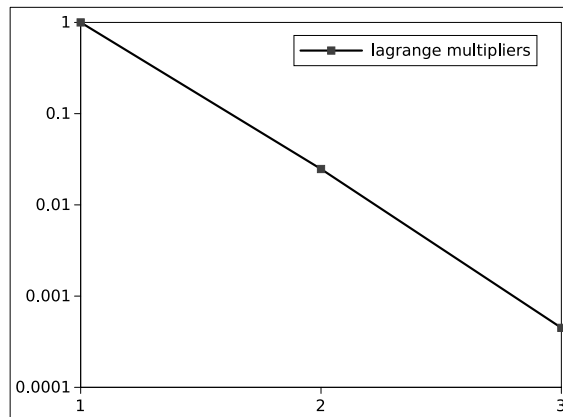


Figure 6. Convergence of the Lagrange multipliers for the geometric constraints. The figure shows the distance to the final Lagrange multiplier in L^2 -norm.

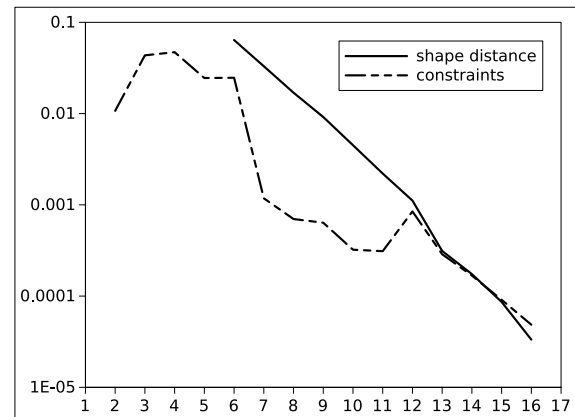


Figure 7. Shape distances to optimal solution and convergence of geometric constraints for BFGS-Steklov–Poincaré, both with fixed Lagrange multipliers.

Equation (3.1) then transforms to

$$\int_{\Omega_{\text{ext}}} \sigma(U) : \epsilon(V) \, dx = \int_{\Gamma} \gamma \langle n, V \rangle \, ds \quad \text{for all } V \in H_0^1(\Omega, \mathbb{R}^d).$$

The optimal mesh shown in Figure 3 (b) is obtained with $\lambda_{\text{elas}} = 0$ and $\mu_{\text{elas}} \in [\mu_{\min}, \mu_{\max}] = [1, 500]$ smoothly decreasing from Γ to the outer boundaries. It should be mentioned that these parameters are not meant to have a physical interpretation. They were found empirically. The choice $\lambda_{\text{elas}} = 0$ leads to a simplification of the strain tensor to $\sigma = 2\mu_{\text{elas}}\epsilon$.

In order to obtain the values for μ_{elas} we solve the Poisson equation

$$\begin{cases} \Delta \mu_{\text{elas}} = 0 & \text{in } \Omega_{\text{ext}}, \\ \mu_{\text{elas}} = \mu_{\max} & \text{on } \Gamma, \\ \mu_{\text{elas}} = \mu_{\min} & \text{on } \Gamma_{\text{out}} \cup \Gamma_{\text{in}} \cup \Gamma_{\text{wall}} \end{cases}$$

in an initial stage of the optimization. It should be remarked that the solution behavior of the Poisson equation depends on the dimension. In the three-dimensional case we therefore apply $\sqrt{\mu_{\text{elas}}}$ in contrast to μ_{elas} for the elasticity tensor of the g^S metric.

Still two questions remain open, namely the numerical realization of the operators \mathcal{T} and R . In principle, the retraction R would require yet another solution of a PDE, which is computationally too expensive. As discussed in [17], simply adding the computed deformation field U to the nodes of the finite element mesh approximates the retraction in a reasonable way. Similarly, the vector transport necessary for the limited BFGS algorithm is approximated with the identity operator. From a computational point of view, s_i and y_i are finite dimensional vectors in the memory. Thus, we use these vectors as they are, without taking into account that they approximate elements in different tangent spaces.

In order to estimate the convergence speed of the proposed method we have to use a measure for the distance between two shapes. We therefore approximate the geodesic distance of the current to the converged shape. This is done by finding the pointwise distance in normal direction between the two shapes and integrating this quantity over Γ . This can be seen in Figure 7. Here the convergence of the corresponding constraints $\|c(\Gamma^k)\|_2$ is also visualized. The underlying optimization is performed with the converged Lagrange multipliers for the geometric constraints. Note that values of the shape distance for the first iterations are missing. This is due to the fact that the underlying shape metric as described above is only reasonable for small deformations and thus does not work for the first iterations.

As mentioned in Section 4, the convergence of λ is relatively fast. This can be seen by observing $\|\lambda\|_2$ which is shown in Figure 6. For this test we have chosen the tolerance for the inner optimization loop depending on

the L^2 norm of the step U evaluated on Γ to be $\delta_f = 10^{-4}$. These computations are performed with a penalty factor $\mu = 10^2$. In principle, the augmented Lagrangian algorithm provides updates for the penalty factor, if the violation of the constraints is over a given tolerance δ_c . We yet did not encounter convergence problems in the test cases with the penalty factor $\mu = 10^2$, since, as Figure 7 suggests, a convergence of the constraints is achieved after appropriate Lagrange multipliers are found by the algorithm.

The algorithms presented within this work are entirely implemented using the GetFEM++ library for the assembly of the finite elements and the PETSc library for the solution of the linear systems. The initial grids are generated with the Gmsh mesh generator and the mesh optimization routines therein.

6 Conclusions

We compared computational aspects of the standard surface metric for shape optimization based on shape calculus with the surface metric introduced and analyzed in [16]. Advantages of the latter metric were demonstrated with regards to convergence properties, computational overall effort as well as resulting mesh quality. This holds for a two-dimensional and for a three-dimensional set-up.

Funding: This work has been partly supported by the Deutsche Forschungsgemeinschaft within the Priority program SPP 1648 “Software for Exascale Computing” under contract number Schu804/12-1. Furthermore, we acknowledge support by the Sino-German Science Center on the occasion of the Chinese-German Workshop on Computational and Applied Mathematics in Augsburg 2015.

References

- [1] P. Absil, R. Mahony and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, 2008.
- [2] A. Borzi and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, SIAM Book Ser. Comput. Sci. Eng. 8, SIAM, Philadelphia, 2012.
- [3] A. R. Conn, N. I. M. Gould and P. L. Toint, *Lancelot*, Springer, Berlin, 1992.
- [4] M. C. Delfour and J.-P. Zolésio, *Shapes and Geometries: Analysis, Differential Calculus, and Optimization*, Adv. Des. Control, SIAM, Philadelphia, 2001.
- [5] K. Eppler, H. Harbrecht and R. Schneider, On convergence in elliptic shape optimization, *SIAM J. Control Optim.* **46** (2007), no. 1, 61–83.
- [6] P. Gangl, A. Laurain, H. Meftahi and K. Sturm, Shape optimization of an electric motor subject to nonlinear magneto-statics, preprint (2015), <http://arxiv.org/abs/1501.04752>.
- [7] W. Haack, Geschoßformen kleinsten Wellenwiderstandes, *Bericht der Lilienthal-Gesellschaft* **136** (1941), no. 1, 14–28.
- [8] P. Michor and D. Mumford, Riemannian geometries on spaces of plane curves, *J. Eur. Math. Soc. (JEMS)* **8** (2006), 1–48.
- [9] B. Mohammadi and O. Pironneau, *Applied Shape Optimization for Fluids*, Num. Math. Sci. Comput., Clarendon Press, Oxford, 2001.
- [10] A. Nägel, V. Schulz, M. Siebenborn and G. Wittum, Scalable shape optimization methods for structured inverse modeling in 3D diffusive processes, *Comput. Vis. Sci.* **17** (2015), 79–88.
- [11] A. Paganini, Approximative shape gradients for interface problems, technical report 2014-12, Seminar for Applied Mathematics, ETH Zürich, 2014.
- [12] O. Pironneau, On optimum profiles in stokes flow, *J. Fluid Mech.* **59** (1973), no. 1, 117–128.
- [13] W. Ring and B. Wirth, Optimization methods on Riemannian manifolds and their application to shape space, *SIAM J. Optim.* **22** (2012), 596–627.
- [14] S. Schmidt, C. Ilıc, V. Schulz and N. Gauger, Three dimensional large scale aerodynamic shape optimization based on the shape calculus, *AIAA J.* **51** (2013), no. 11, 2615–2627.
- [15] V. Schulz, A Riemannian view on shape optimization, *Found. Comput. Math.* **14** (2014), 483–501.
- [16] V. Schulz, M. Siebenborn and K. Welker, A novel Steklov–Poincaré type metric for efficient PDE constrained optimization in shape spaces, preprint (2015), <http://arxiv.org/abs/1506.02244v4>.
- [17] V. Schulz, M. Siebenborn and K. Welker, Structured inverse modeling in parabolic diffusion problems, *SIAM J. Control Optim.* **53** (2015), no. 6, 3319–3338.

- [18] V. Schulz, M. Siebenborn and K. Welker, Towards a Lagrange–Newton approach for PDE constrained shape optimization, in: *New Trends in Shape Optimization*, Internat. Ser. Numer. Math. 166, Birkhäuser, Cham (2015), 229–249.
- [19] J. Sokolowski and J. Zolésio, *An introduction to shape optimization*, Springer Ser. Comput. Math. 16, Springer, Berlin, 1992.
- [20] R. Udawalpola and M. Berggren, Optimization of an acoustic horn with respect to efficiency and directivity, *Internat. J. Numer. Methods Engrg.* **73** (2007), no. 11, 1571–1606.