A RIEMANNIAN VIEW ON SHAPE OPTIMIZATION

VOLKER H. SCHULZ*

Abstract. Shape optimization based on the shape calculus is numerically mostly performed by means of steepest descent methods. This paper provides a novel framework to analyze shape-Newton optimization methods by exploiting a Riemannian perspective. A Riemannian shape Hessian is defined yielding often sought properties like symmetry and quadratic convergence for Newton optimization methods.

1. Introduction. Shape optimization is a vivid field of research. In particular, the usage of shape calculus for practical applications has increased steadily [2, 3, 4, 15, 28, 29, 26, 17]. Standard references for shape optimization based on the shape calculus are [9, 24, 31]. The shape Hessian is already used as a means to accelerate gradient based shape optimization methods [13, 15, 17, 27]. It is also used for characterizing the well-posedness and sufficient optimality conditions [12, 14] in particular applications and is reformulated in the framework of differential forms in [19]. A fairly general framework for descent methods for shape optimization is presented in [18].

However, a general framework for the analysis of Newton-type shape optimization algorithms is still missing. A major reason is the lack of symmetry [9, 19] of the shape Hessian, as it is commonly defined. In this paper, an attempt is made to cast shape optimization problems in the framework of optimization on Riemannian manifolds. There is a fairly large amount of publications available on the issue of optimization on Riemannian manifolds—mainly for matrix manifolds as in [1, 25].

It is proposed in this paper, to view the set of all shapes as a Riemannian manifold and follow there the ground breaking work in [22, 21, 23, 6, 5]. The resulting manifold is an abstract infinite dimensional manifold, in contrast to the finite dimensional submanifolds of $\mathbb{R}^n$ that arise in optimization on matrix manifolds. Therefore, distance concepts have to be reviewed and used with somewhat more care. The key observation of this paper is that the action of an element of the tangent space of the manifolds of shapes can be interpreted as the shape derivative of classical shape calculus. Once this link is established, the concept of Riemannian shape Hessian, shape Taylor series, shape Newton convergence and sufficient shape optimality conditions follow quite naturally.

In section 2 the notation for the manifold of shapes is introduced and for a particular example of a Riemannian geometry, the correspondence between Riemannian geometry and shape calculus is established. The key element of the co-variant derivative is rephrased in terms of the shape calculus. The Riemannian shape Hessian is defined and the Riemannian shape Taylor series formulated. Section 3 presents a generalization of the Newton convergence theory, established in [7] for linear spaces, on Riemannian manifolds. From that, convergence properties of variants of Newton's method on Riemanian manifolds follow immediately. Finally, section 4 discusses numerical experiments for shape optimization algorithms with linear and quadratic convergence properties.

2. Riemannian Shape Geometry and the Shape Calculus. The purpose of this section is to demonstrate the possibility to define a Riemannian metric on the manifold of all possible shapes with a relation to the shape calculus. Since this point of view is new, we use the established framework of differential geometry for shapes which are C^∞ embeddings of the unit sphere. Of course, this framework has to be generalized for specific applications. However, the purpose of this paper is to

---

*University of Trier, Department of Mathematics, 54296 Trier, Germany (`volker.schulz@uni-trier.de`).

convey a new point of view rather than the impact on applications. Therefore, we assume in the interest of simplicity of the presentation maximum smoothness and restrict ourself to only 2D problems.

It is assumed that the reader is familiar with the basic concepts of Riemannian geometry as they are, e.g.,: geodesics, exponential mapping, parallel transport. In [22, 21], a geometric structure of two-dimensional C^∞-shapes has been introduced and consequently generalized to shapes in higher dimension in [23, 6, 5]. Essentially, closed curves (and closed higher dimensional surfaces) are identified with mappings of the unit sphere to any shape under consideration. In two dimensions, this can be naturally motivated by the Riemannian mapping theorem. In this paper, we focus on two-dimensional shapes as subsets of $\mathbb{R}^2$ for ease of discussion, but mention the other publications above, in order to indicate that natural extensions of this paper to higher dimensional surfaces are conceivable.

Here, we mean with "shape" a singly connected and compact subset of $\Omega \subset \mathbb{R}^2$ with $\Omega \neq \emptyset$ and C^∞ boundary $\partial\Omega$. As always in shape optimization, the boundary of the shape is all that matters. Thus, we can identify the set of all shapes with the set of all those boundaries. In [22], this set is characterized by

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

i.e., as the set of all equivalence classes of C^∞ embeddings of S^1 in the plane ($\mathrm{Emb}(S^1, \mathbb{R}^2)$), where the equivalence relation is defined by the set of all C^∞ reparameterizations, i.e., diffeomorphisms of S^1 into itself ($\mathrm{Diff}(S^1)$). The set B_e is considered as a manifold in [22] and various Riemannian metrics are investigated. A particular point on the manifold $B_e(S^1, \mathbb{R}^2)$ is represented by a curve $c : S^1 \ni \theta \mapsto c(\theta) \in \mathbb{R}^2$. Because of the equivalence relation ($\mathrm{Diff}(S^1)$), the tangent space is isomorphic to the set of all normal C^∞ vector fields along c, i.e.

$$T_c B_e \cong \{h \mid h = \alpha\vec{n},\ \alpha \in C^\infty(S^1, \mathbb{R})\}$$

where $\vec{n}$ is the unit exterior normal field of the shape Ω defined by the boundary $\partial\Omega = c$ such that $\vec{n}(\theta) \perp c_\theta(\theta)$ for all $\theta \in S^1$. For our discussion, we pick among the other metrics discussed in [22] the metric family for $A \geq 0$

$$G^A : T_c B_e \times T_c B_e \to \mathbb{R}$$
$$(h, k) \mapsto \int_{S^1} \left(1 + A\kappa_c(\theta)^2\right) \langle h(\theta), k(\theta)\rangle \|c_\theta(\theta)\|\, d\theta$$

where κ_c denotes the curvature of the curve c and $\langle .,. \rangle$ and $\|.\|$ mean the standard Euclidian scalar product and norm in $\mathbb{R}^2$. This does by no means mean that this metric is uniquely defined by the purpose of establishing a connection with the shape calculus. There are probably many more Riemannian metrics available which can be used in a similar way. If $h = \alpha\vec{n}$ and $k = \beta\vec{n}$, then this scalar product on $T_c B_e$ can be expressed more simple as

$$G^A(h, k) = \int_{\partial\Omega} (1 + A\kappa_c^2)\alpha\beta ds$$

where ds is the length measure on $\partial\Omega = c$. In [22] it is shown that for $A > 0$ the scalar product G^A defines a Riemannian metric on B_e and that the resulting Riemannian manifold (B_e, G^A) is complete. Thus, geodesics can be used to measure distances and Cauchy sequences in terms of the geodesic distance converge. Unfortunately, this is not the case for the most simple member G^0 of the metric family G^A, where $A = 0$. An illustrative counter-example is given in [22]. In section 3, we use the fact that (B_e, G^A) is a Riemannian manifold extensively and

exploit the existence of the exponential map (exp) according to the usual definition in Riemannian geometry.

Now, we want to analyze the correlation of the Riemannian geometry on B_e with shape optimization. In 2D shape optimization, one searches for the solution of optimization problems of the form

$$\min_{\Omega} f(\Omega)$$

where f is a real valued shape differentiable objective function. Often, the problem formulation involves explicit constraints in the form of differential equations and additional state variables as in [2, 3, 4, 15, 28, 29, 26, 17, 13, 15, 17, 27, 9, 24, 31, 12, 14], but for the sake of ease of presentation, we can assume all those possible additional structures are contained implicitly within the mapping f. The shape derivative of f is a directional derivative in the direction of a vector C^∞ vector field $V : \mathbb{R}^2 \to \mathbb{R}^2$ which can be represented on the boundary according to the Hadamard structure theorem [31] as

$$df(\Omega)[V] = \int_{\partial\Omega} g \, \langle V, \vec{n} \rangle \, ds \,, \qquad g : \partial\Omega \to \mathbb{R}$$

If $V|_{\partial\Omega} = \alpha\vec{n}$, this can be written more concisely as

$$df(\Omega)[V] = \int_{\partial\Omega} g \, \alpha \, ds$$

In Riemannian geometry, tangential vectors are considered as directional derivatives of scalar valued functions. Since curves $c \in B_e$ can be interpreted as boundaries of domains Ω with boundary $c = \partial\Omega$, we can consider every scalar valued function $f : c = \partial\Omega \mapsto \mathbb{R}$ also as mapping $f : \Omega \mapsto \mathbb{R}$. Thus, we see that the action of a tangent vector $h \in T_c B_e$ on a scalar valued function $f : B_e \to \mathbb{R}$ can be interpreted in the shape calculus, via the unique identification of the boundary $c = \partial\Omega$ with its shape Ω, as the shape derivative of f with respect to an arbitrary C^∞ extension V of h in the whole domain Ω with $V|_{\partial\Omega} = h$. Thus, we can write

$$h(f)(c) = df(\Omega)[V] = \int_{\partial\Omega} g \, \alpha \, ds$$

if $h = \alpha\vec{n}$. Also, the gradient in terms of a Riemannian representation of the shape derivative in terms of the metric G^A can be written as

$$\mathrm{grad} f = \frac{1}{1 + A\kappa_c^2} g$$

The essential operation in Riemannian geometry is the covariant derivative $\nabla_h k$ which is a directional derivative of vector fields in terms of tangential vectors such that $\nabla_h k \in T_c B_e$, if $h, k \in T_c B_e$. Usually, the covariant derivative is written in terms of the Christoffel symbols. In [22] explicit expressions for the Christoffel symbols are derived in terms of the Riemannian metric G^A. However, in order to reveal the relation with the shape calculus, we show another representation of the covariant derivative in theorem 2.1.

THEOREM 2.1. *Let $\Omega \in \mathbb{R}^2$ be a shape and $V, W \in C^\infty(\mathbb{R}^2, \mathbb{R}^2)$ vector fields with orthogonal components $\alpha := \langle V|_{\partial\Omega}, \vec{n} \rangle$, $\beta := \langle W|_{\partial\Omega}, \vec{n} \rangle$ such that $h := \alpha\vec{n}$, $k := \beta\vec{n}$ belong to the tangent space of B_e. Then, the covariant derivative can be expressed as*

$$\nabla_V W := \nabla_h k = \frac{\partial\beta}{\partial\vec{n}}\alpha + \frac{1}{2}(\kappa_c + \frac{2A\kappa_c^3}{1 + A\kappa_c^2})\alpha\beta + A\kappa_c(\alpha\beta)_{\tau\tau}$$

$$= \langle DW\,V, \vec{n} \rangle + \frac{1}{2}(\kappa_c + \frac{2A\kappa_c^3}{1 + A\kappa_c^2}) \, \langle V, \vec{n} \rangle \, \langle W, \vec{n} \rangle + A\kappa_c(\langle V, \vec{n} \rangle \, \langle W, \vec{n} \rangle)_{\tau\tau}$$

4

where expressions "$(.)_{\tau\tau}$" mean second order derivative in unit tangential direction of the shape boundary.

  Proof. The expression in theorem 2.1 is derived from the product rule for Riemannian connections, $hG^A(k,\ell) = G^A(\nabla_h k,\ell) + G^A(k,\nabla_h\ell)$ for any $\ell \in T_c B_e$. We assume that $Z \in C^\infty(\mathbb{R}^2,\mathbb{R}^2)$ is a vector field with $\ell := Z|_{\partial\Omega} = \gamma\vec{n}$. Then

$$h(G^A(k,\ell)) = d\left(\int_{\partial\Omega}(1+A\kappa_c^2)\beta\gamma ds\right)[V] \tag{2.1}$$

$$= \int_{\partial\Omega} \frac{\partial[(1+A\kappa_c^2)\beta\gamma]}{\partial\vec{n}}\alpha + \kappa_c(1+A\kappa_c^2)\alpha\beta\gamma ds \tag{2.2}$$

$$= \int_{\partial\Omega} 2A\kappa_c(\kappa_c^2\alpha+\alpha_{\tau\tau})\beta\gamma + (1+A\kappa_c^2)\frac{\partial\beta}{\partial\vec{n}}\gamma\alpha \tag{2.3}$$

$$+ (1+A\kappa_c^2)\frac{\partial\gamma}{\partial\vec{n}}\beta\alpha + \kappa_c(1+A\kappa_c^2)\alpha\beta\gamma ds$$

$$= \int_{\partial\Omega} 2A\kappa_c^3\alpha + 2A\kappa_c\alpha(\beta\gamma)_{\tau\tau} + (1+A\kappa_c^2)\frac{\partial\beta}{\partial\vec{n}}\gamma\alpha \tag{2.4}$$

$$+ (1+A\kappa_c^2)\frac{\partial\gamma}{\partial\vec{n}}\beta\alpha + \kappa_c(1+A\kappa_c^2)\alpha\beta\gamma ds$$

where equation (2.4) is an immediate consequence of the discussion in section 2.2 of [22]. Since the remaining properties of a Riemannian connection (C^∞ linearity, the Leibniz rule and the symmetry) are obviously satisfied, the expression above is shown due to the uniqueness of the Riemannian connection. □

  The Riemannian connection now gives the means to investigate the Hessian of an objective defined on a shape, resp. its boundary. As in [1], we define the Riemannian Hessian of a function f at the point $c \in B_e$ as the linear mapping of $T_c B_e$ to itself defined by

$$\mathrm{Hess}f(c)[h] := \nabla_h \mathrm{grad} f$$

In the terminology of shape optimization and with the definition of vector fields as in theorem 2.1 and the identification of the boundary $c = \partial\Omega$ with its shape, we may identify this with a now so-called Riemannian shape Hessian.

  DEFINITION 2.2. *For the setting of theorem 2.1, we define the* Riemannian shape Hessian *as*

$$\mathrm{Hess}f(\Omega)[V] := \nabla_V \mathrm{grad} f$$

The next theorem gives a correlation of the covariant shape Hessian with the standard shape Hessian which is defined by repeated shape differentiation.

  THEOREM 2.3. *The Riemannian shape Hessian satisfies the relation*

$$G(Hessf(\Omega)[V],W) = d(df(\Omega)[W])[V] - df(\Omega)[\nabla_V W]$$

where V,W are defined as in theorem 2.1 and $d(df(\Omega)[W])[V]$ denotes the standard shape Hessian as defined in [9].

  Proof. The Riemannian product rule gives

$$h(G(\mathrm{grad}f,k)) = G(\nabla_h \mathrm{grad}f,k) + G(\mathrm{grad}f,\nabla_h k)$$

We observe

$$h(G(\mathrm{grad}f,k)) = h(k(f)) = h(df(\Omega)[k]) = d(df(\Omega)[k])[h] = d(df(\Omega)[V])[W]$$

$$G(\nabla_h \mathrm{grad}f,k) = G(\mathrm{Hess}f(c)[h],k) = G(\mathrm{Hess}f(\Omega)[V],W)$$

$$G(\mathrm{grad}f,\nabla_h k) = \nabla_h k(f) = df(\Omega)[\nabla_h k] = df(\Omega)[\nabla_V W]$$

□

COROLLARY 2.4. *The Riemannian shape Hessian is symmetric, i.e.*

$$G(Hess f(\Omega)[V], W) = G(V, Hess f(\Omega)[W])$$

Proof. The symmetry of the Riemannian connection ∇ gives with the vector field definitions of theorem 2.1

$$G(\text{Hess} f(\Omega)[V], W) = h(k(f)) - (\nabla_h k)(f) = k(h(f)) - (\nabla_k h)(f)$$
$$= G(V, \text{Hess} f(\Omega)[W])$$

□

Now we can use the Riemannian shape Hessian in order to formulate a Taylor series expansion as well as optimality conditions. Since the subsequent theorem holds in more general cases than just for B_e, we formulate it in a more general notation.

THEOREM 2.5. *Let $(\mathcal{N}, G)$ be a Riemannian manifold with metric G and norm $\|.\| := G(.,.)$. The set $U \subset \mathcal{N}$ be a convex subset of $\mathcal{N}$. We consider all $x \in \mathcal{N}$ and $\Delta x \in T_x \mathcal{N}$ with $x, \exp_x(\Delta x) \in U$ and denote the parallel transport along the geodesic $\gamma : [0,1] \to \mathcal{N}, t \mapsto \gamma(t) := \exp_x(t\Delta x)$ by $\mathcal{P}_{\alpha,\beta} : T_{\gamma(\alpha)}\mathcal{N} \to T_{\gamma(\beta)}\mathcal{N}$. We assume for the Riemannian Hessian of a function $f : U \to \mathbb{R}$ the following Lipschitz property at $x \in \mathcal{N}$:*

$$\|\mathcal{P}_{1,0} Hess\, f(\exp_x(\Delta x))\mathcal{P}_{0,1} - Hess\, f(x)\| \leq L\|\Delta x\|, \quad \forall \exp_x(\Delta x) \in U$$

with a constant $L < \infty$. Then, we achieve the estimation

$$|f(\exp_x(\Delta x)) - f(x) + G(\text{grad}\, f(x), \Delta x) + \frac{1}{2}G(Hess\, f(x)\Delta x, \Delta x)| \leq \frac{L}{6}\|\Delta x\|^3$$

Proof. Let us consider the mapping $\varphi : [0,1] \ni t \mapsto f(\exp_x(t\Delta x))$. We note that for all differentiable functions and in particular for φ yields

$$\int_0^1 \int_0^t \varphi''(s) - \varphi''(0) ds\, dt = \varphi(1) - \varphi(0) - \varphi'(0) - \frac{1}{2}\varphi''(0)$$

Since

$$\varphi(1) = f(\exp_x(\Delta x)), \ \varphi(0) = f(x), \ \varphi'(0) = G(\text{grad}\, f(x), \Delta x)$$
$$\varphi''(0) = G(Hess\, f(x)\Delta x, \Delta x)$$

we observe

$$|f(\exp_x(\Delta x)) - f(x) + G(\text{grad}\, f(x), \Delta x) + \frac{1}{2}G(Hess\, f(x)\Delta x, \Delta x)|$$

$$= \int_0^1 \int_0^t |\varphi''(s) - \varphi''(0)| ds\, dt$$

$$= \int_0^1 \int_0^t |((\mathcal{P}_{s,0} Hess\, f(\exp_x(s\Delta x))\mathcal{P}_{0,s} - Hess\, f(x))\, \Delta x, \Delta x)|\, ds\, dt$$

$$\leq \int_0^1 \int_0^t \|\mathcal{P}_{s,0} Hess\, f(\exp_x(s\Delta x))\mathcal{P}_{0,s} - Hess\, f(x)\| \, \|\Delta x\|^2 ds\, dt$$

$$\leq \int_0^1 \int_0^t sL\|\Delta x\|^3 ds\, dt = \frac{L}{6}\|\Delta x\|^3$$

□

Now, we can exploit the Taylor expansion of Theorem 2.5 for necessary and sufficient optimality conditions

THEOREM 2.6. *Under the assumptions of Theorem 2.5 we obtain:*
(a) *If $\hat{x}$ is an optimal solution, then* Hess $f(\hat{x}) \geq 0$, *i.e.* (Hess $f(\hat{x})h, h) \geq 0$, *for all $h \in T_{\hat{x}}\mathcal{N}$*
(b) *If $\hat{x}$ satisfies* grad $f(\hat{x}) = 0$, *and* Hess $f(\hat{x})$ *is coercive, i.e.* (Hess $f(\hat{x})h, h) \geq c\|h\|^2$, *for all $h \in T_{\hat{x}}\mathcal{N}$ and for some $c > 0$, then $\hat{x}$ is a local minimum, provided* Hess $f(\hat{x})$ *satisfies a Lipschitz condition as in Theorem 2.5.*

Proof. The proof is identical with the standard proof in linear spaces. □

EXAMPLE 1. Now let us study a particular example, one of the simplest but nevertheless instructive shape optimization problems

$$\min_{\Omega \subset \mathbb{R}^2} f(\Omega) := \int_{\Omega} \psi(x)dx$$

where $\psi : \mathbb{R}^2 \to \mathbb{R}$ is a sufficiently smooth scalar valued function. We use the vector field definitions of theorem 2.1. The shape derivative of this objective and the G^A-gradient are

$$df(\Omega)[V] = \int_{\partial\Omega} \psi \langle V, \vec{n}\rangle \, ds$$

$$\operatorname{grad} f(c) = \frac{\psi}{1 + A\kappa_c^2}, \qquad c = \partial\Omega$$

The shape Hessian is computed by shape differentiating the shape derivative

$$d(df(\Omega)[W][V])) = d(\int_{\partial\Omega} \psi k ds)[h] = \int_{\partial\Omega} \frac{\partial(\psi k)}{\partial\vec{n}} h + \kappa_c \psi k h ds$$

$$= \int_{\partial\Omega} \frac{\partial\psi}{\partial\vec{n}} kh + \psi \frac{\partial k}{\partial\vec{n}} h + \kappa_c \psi k h ds$$

$$= \int_{\partial\Omega} (\frac{\partial\psi}{\partial\vec{n}} + \kappa_c\psi) \langle W, \vec{n}\rangle \langle V, \vec{n}\rangle + \psi \langle DW\,V, \vec{n}\rangle \, ds$$

We observe that the shape Hessian is not symmetric and we are tempted to use the term "$\partial\psi/\partial\vec{n} + \kappa_c\psi$" within a Newton-like optimization strategy, since we do not trust the nonsymmetric part of the Hessian. In contrast to that, the Riemannian shape Hessian computed by application of theorem 2.3 is given by

$$G(\operatorname{Hess} f(\Omega)[V], W) = d(df(\Omega)[W])[V] - df(\Omega)[\nabla_V W]$$

$$= \int_{\partial\Omega} (\frac{\partial\psi}{\partial\vec{n}} + \kappa_c\psi) \langle W, \vec{n}\rangle \langle V, \vec{n}\rangle + \psi \langle DW\,V, \vec{n}\rangle \, ds - \int_{\partial\Omega} \psi \langle \nabla_V W, \vec{n}\rangle \, ds$$

$$= \int_{\partial\Omega} (\frac{\partial\psi}{\partial\vec{n}} + \kappa_c\psi) \langle W, \vec{n}\rangle \langle V, \vec{n}\rangle + \psi \langle DW\,V, \vec{n}\rangle \, ds$$

$$- \int_{\partial\Omega} \psi(\langle DW\,V, \vec{n}\rangle + \frac{\psi}{2}(\kappa_c + \frac{2A\kappa_c^3}{1 + A\kappa_c^2}) \langle V, \vec{n}\rangle \langle W, \vec{n}\rangle$$

$$+ \psi A\kappa_c(\langle V, \vec{n}\rangle \langle W, \vec{n}\rangle)_{\tau\tau} ds$$

$$= \int_{\partial\Omega} (\frac{\partial\psi}{\partial\vec{n}} + \frac{\kappa_c}{2}\psi - \frac{A\kappa_c^3}{1 + A\kappa_c^2}\psi) \langle V, \vec{n}\rangle \langle W, \vec{n}\rangle - \psi A\kappa_c(\langle V, \vec{n}\rangle \langle W, \vec{n}\rangle)_{\tau\tau} ds$$

Now, as already abstractly shown in corollary 2.4, we observe symmetry of the Riemannian shape Hessian also in this example. For small $A > 0$ the term "$\partial\psi/\partial\vec{n} + \frac{\kappa_c}{2}\psi$" is decisive rather than the putative term above. We will study this example in more specific details in section 4.

3. Convergence of Riemannian Shape Newton Methods. Now we formulate a contraction result for Newton iterations on manifolds which is in line with linear space theorems in [7, 11]. Again, since the observations hold for more general Riemannian manifolds than just B_e, we use a more general notation. We proof convergence properties for Newton-like methods on Riemannian manifolds, but have always in mind that in this paper we want to solve a particular root finding problem for the gradient of an objective. Thus, every time the Jacobian is mentioned, we can think it as the Jacobian of the gradient and thus the Hessian of an objective on a Riemannian manifold.

THEOREM 3.1. *We consider a complete Riemannian Manifold $(\mathcal{N}, G)$ with norm $\|.\| := G(.,.)$ on the tangential bundle $T\mathcal{N}$. The set $D \subset \mathcal{N}$ is assumed to be singly connected and open. We are seeking a singular point of the twice differentiable vector field $F : \mathcal{N} \to T\mathcal{N}$ by employing a Newton method on manifolds. The symbol $J := DF$ denotes the covariant derivative of F such that $DF(x)v := \nabla_v F_x$. Furthermore, we assume that for all points $x, y \in D$ with $y = \exp_x(\Delta x), \Delta x := -M(x)F(x)$ and $M(x) \in \mathrm{End}(T_x\mathcal{N}))$ and invertible and all $t \in [0, 1]$, the following Lipschitz conditions are satisfied along the geodesic $\gamma : [0, 1] \to \mathcal{N}, t \mapsto \gamma(t) := \exp_x(-tM(x)F(x))$ with parallel transport $\mathcal{P}_{\alpha,\beta} : T_{\gamma(\alpha)}\mathcal{N} \to T_{\gamma(\beta)}\mathcal{N}$.*

(1) There exists $\omega < \infty$ with

$$\|M(y)(\mathcal{P}_{t,1}J(\gamma(t))\mathcal{P}_{0,t} - \mathcal{P}_{0,1}J(x))\Delta x)\| \leq \omega t \|\Delta x\|^2$$

(2) There is a constant upper limit $\kappa < 1$ for the function $\tilde{\kappa}(x)$ in

$$\|M(y)\mathcal{P}_{0,1}(F(x) + J(x)\Delta x)\| =: \tilde{\kappa}(x) \|\Delta x\|$$

with $\tilde{\kappa}(x) \leq \kappa$.

If x_0 satisfies $\delta_0 < 1$, where $\delta_k := \kappa + \frac{\omega}{2} \|M(x_k)F(x_k)\|, k = 0, 1, \ldots$, then follows

(1) The iteration $x_{k+1} := \exp_{x_k}(-M(x_k)F(x_k))$ is well defined and stays in $D_0 := \{x \in D \mid d(x, x_0) \leq \|M(x_0)F(x_0)\| / (1 - \delta_0)\}$.

(2) There exists $\hat{x} \in D_0$ with $\lim_{k\to\infty} x^k = \hat{x}$ in the sense that $\lim_{k\to\infty} d_G(x^k, \hat{x}) = 0$, where d_G is the geodesic distance.

(3) There holds the a-priori estimation

$$d_{G^A}(x^k, \hat{x}) \leq \frac{\delta_k}{1 - \delta_k} \|\Delta x_k\| \leq \frac{\delta_0^k}{1 - \delta_0} \|\Delta x_k\|$$

where $\Delta x_k := -M(x^k)F(x^k)$.

(4) There yields the contraction property

$$\|\Delta x_{k+1}\| \leq \delta_k \|\Delta x_k\| = (\kappa + \frac{\omega}{2} \|\delta x_k\|) \|\Delta x_k\|$$

(5) If the mapping $x \mapsto M(x)$ is continuous and $M(\hat{x})$ is nonsingular, then $\hat{x}$ is not only a fixed point but rather a root of the equation $F(x) = 0$.

Proof. The proof follows the standard lines—now in manifold notation. First, we note the manifold variant of the fundamental theorem of calculus along the geodesic γ for any C^1 vector field X (cf. [16]):

$$X(\gamma(t)) = \mathcal{P}_{0,t}X(\gamma(0) + \int_0^t \mathcal{P}_{s,t}\nabla_{\dot{\gamma}(s)}X_{\gamma(s)}ds = \mathcal{P}_{0,t}X(\gamma(0) + \int_0^t \mathcal{P}_{s,t}DX(\gamma(s))\dot{\gamma}(s)ds$$

and have in mind that obviously $\dot{\gamma}(s) = \mathcal{P}_{0,s}\Delta x^k$. We show the contraction property and use the abbreviation $R(x^k) := F(x^k) - J(x^k)M(x^k)F(x^k)$

$$
\begin{aligned}
\left\|\Delta x^{k+1}\right\| &= \left\|M(x^{k+1})F(x^{k+1})\right\| \\
&= \left\|M(x^{k+1})\mathcal{P}_{0,1}R(x^k) + M(x^{k+1})\left[F(x^{k+1}) - \mathcal{P}_{0,1}R(x^k)\right]\right\| \\
&\leq \kappa\left\|\Delta x^k\right\| + \left\|M(x^{k+1})\int_0^1 \mathcal{P}_{t,1}DF(\gamma(t))\mathcal{P}_{0,t}\Delta x^k - \mathcal{P}_{0,1}J(x^k)\Delta x^k dt\right\| \\
&\leq \kappa\left\|\Delta x^k\right\| + \int_0^1 \left\|M(x^{k+1})(\mathcal{P}_{t,1}DF(\gamma(t))\mathcal{P}_{0,t} - \mathcal{P}_{0,1}J(x^k))\Delta x^k\right\| dt \\
&\leq \kappa\left\|\Delta x^k\right\| + \int_0^1 t\omega\left\|\Delta x^k\right\|^2 dt = \kappa\left\|\Delta x^k\right\| + \frac{\omega}{2}\left\|\Delta x^k\right\|^2 = \delta_k\left\|\Delta x^k\right\|
\end{aligned}
$$

Now we conclude inductively that $\delta_k < 1$ for all k and thus the series $\{\delta_k\}_{k=0}^\infty$ and $\{\|\Delta x_k\|\}_{k=0}^\infty$ are montonically decreasing. Now, we show that the series $\{x_k\}_{k=0}^\infty$ stays in D_0. We use the triangle inequality which holds because of the metric properties of the Riemannian metric.

$$
d_G(x^k, x^0) = d_G(\exp(\Delta x_{k-1}) \circ \ldots \circ \exp(\Delta x_1) \circ \exp_{x_0}(\Delta x_0), x_0) \tag{3.1}
$$

$$
\leq \sum_{j=1}^{k-1}\left\|\Delta x^j\right\| \leq \sum_{j=1}^{k-1}\delta_0^j\left\|\Delta x^0\right\| \leq \frac{\|\Delta x_0\|}{1-\delta_0} \qquad \text{(geometric series)} \tag{3.2}
$$

Analogously, we show the Cauchy property of the series $\{x_k\}_{k=0}^\infty$.

$$
d_G(x^m, x^n) = d(\exp(\Delta x_{m-1}) \circ \ldots \circ \exp(\Delta x_{n+1}) \circ \exp_{x_n}(\Delta x_n), x_n) \tag{3.3}
$$

$$
\leq \sum_{k=n}^{m-1}\left\|\Delta x^k\right\| \leq \sum_{k=n}^{m-1}\delta_n^k\left\|\Delta x^n\right\| \leq \sum_{k=n}^{m-1}\delta_0^k\left\|\Delta x^0\right\| \tag{3.4}
$$

$$
\leq \varepsilon \text{ for any } \varepsilon > 0, \text{ if } n,m \to \infty \tag{3.5}
$$

The a priori estimation is yet another application of the triangle inequality. Now, if $M(.)$ is continuous and $x^k \to \hat{x}$, we pass the defining equation

$$
x_{k+1} := \exp_{x_k}(-M(x_k)F(x_k))
$$

to the limit, having in mind that exp is continuous and $\exp_y(z) = y$ implies $z = 0$ for any y. Thus, we observe that $M(\hat{x})F(\hat{x}) = 0$. Finally, if $M(\hat{x})$ is nonsingular, we conclude that $F(\hat{x}) = 0$

□

Now, we proof quadratic convergence for the exact Newton method–again completely parallel to the discussions in [7, 11].

COROLLARY 3.2. *Together with the assumptions of theorem 3.1, we choose* $M(x) = DF(x)^{-1}$, *thus defining the exact Newton method on manifolds. The resulting iteration converges locally quadratically, i.e., there is a* $\tilde{k} \in \mathbb{N}$ *and a* $C < \infty$ *such that*

$$
d_G(x^{k+1}, \hat{x}) \leq C d_G(x^k, \hat{x})^2, \quad \forall k \geq \tilde{k}
$$

where d_G denotes again the geodesic distance.

Proof. Because of the choice $M(x) = DF(x)^{-1}$, we observe for κ in theorem 3.1 from

$$
\|M(y)\mathcal{P}_{0,1}(F(x) + J(x)\Delta x)\| = \left\|M(y)\mathcal{P}_{0,1}(F(x) - J(x)J(x)^{-1}F(x))\right\| = 0
$$

And thus $\kappa = 0$. Since $\{\delta_k\}_{k=0}^{\infty}$ is monotonically decreasing to zero, there is $\tilde{k}$ such that $\delta_k = \frac{\omega}{2} \left\| \Delta x^k \right\| \leq \frac{1}{4}$ for all $k \geq \tilde{k}$, which implies also $\delta_k/(1 - \delta_k) \leq 1/3$ for all $k \geq \tilde{k}$. From the fact that the exponential function is an isometry and from the a priori estimation in theorem 3.1, we observe

$$\left\| \Delta x^k \right\| = d_G(x^{k+1}, x^k) \leq d_G(x^{k+1}, \hat{x}) + d_G(x^k, \hat{x}) \leq \frac{\delta_k}{1 - \delta_k} \left\| \Delta x^k \right\| + d_G(x^k, \hat{x})$$

$$\leq \frac{1}{3} \left\| \Delta x^k \right\| + d_G(x^k, \hat{x})$$

and therefore $\left\| \Delta x^k \right\| \leq \frac{3}{2} d(x^k, \hat{x})$. Now, we use again the a priori estimation

$$d_G(x^{k+1}, \hat{x}) \leq \frac{\delta_{k+1}}{1 - \delta_{k+1}} \left\| \Delta x_{k+1} \right\| \leq \frac{4}{3} \frac{\omega}{2} \left\| \Delta x^k \right\|^2 \leq \frac{3}{2} \omega d_G(x^k, \hat{x})^2$$

□

In shape optimization, it is very rare that one can get hold of $J(x^k)$ (i.e., the Hessian of an objective) away from the optimal solution. However, in many cases, expressions can be derived, which deliver the exact Hessian, if evaluated at the solution. That means that often the situation occurs that an approximation M_k^{-1} of $J(x^k)$ is available with the property $M_k^{-1} \to DF(\hat{x}), k \to \infty$. We show local superlinear convergence in those cases.

COROLLARY 3.3. *Together with the assumptions of theorem 3.1, we choose* $M_k := M(x^k)$ *such that*

$$\frac{\left\| [M_k^{-1} - DF(x^k)]\Delta x^k \right\|}{\left\| \Delta x^k \right\|} \to 0, \quad k \to \infty$$

If there is a constant $C < \infty$ such that the approximation M_k is uniformly bounded, $\|M_k\| \leq C, \forall k$, then the resulting Newton iteration converges locally superlinearly.

Proof. We observe for $\kappa_k := \kappa(x^k)$ in theorem 3.1

$$\kappa_k = \frac{\left\| M_{k+1} \mathcal{P}_{0,1}(F(x^k) + J(x^k)\Delta x^k) \right\|}{\left\| \Delta x^k \right\|} = \frac{\left\| M_{k+1} \mathcal{P}_{0,1}(M_k^{-1} + J(x^k))\Delta x) \right\|}{\left\| \Delta x^k \right\|}$$

$$\leq \frac{\left\| M_{k+1} \right\| \left\| \mathcal{P}_{0,1}(M_k^{-1} + J(x^k))\Delta x) \right\|}{\left\| \Delta x^k \right\|} \to 0$$

We use again the apriori estimation in an analogous fashion as in the proof of corollary 3.2 to obtain

$$d_G(x^{k+1}, \hat{x}) \leq \frac{\delta_{k+1}}{1 - \delta_{k+1}} \frac{1 - 2\delta_k}{1 - \delta_k} d_G(x^{k+1}, \hat{x}) \leq \frac{1 - 2\delta_k}{(1 - \delta_k)^2} \delta_k d_G(x^{k+1}, \hat{x})$$

Since $0 \leq \delta_k^2$, we observe that $(1 - 2\delta_k)/(1 - \delta_k)^2 \leq 1$ which yields

$$d_G(x^{k+1}, \hat{x}) \leq \delta_k d_G(x^k, \hat{x}) \text{ with } \delta_k \to 0, \quad k \to \infty$$

□

REMARKS:
- If we can assume that the quality of the Hessian approximation satisfies even $\left\| M_k^{-1} + J(x^k) \right\| \leq C \left\| \Delta x^k \right\|$ then we observe also quadratic convergence as an immediate consequence of an obvious refinement of the proof of corollary 3.3.
- Note that the condition in corollary 3.3 is similar to the Dennis-Moré condition [10, 8], which is also applicable for quasi-Newton update techniques on Riemannian manifolds [25].

4. Numerical Experiments. Here, we study the linear and quadratic convergence properties of standard optimization algorithms applied to an example for shape optimization in $\mathbb{R}^2$ which is as simple as possible, but nevertheless reveals structures reflecting the discussion of this paper. We consider the following problem:

$$\min_{\Omega} f(\Omega) = \int_{\Omega} x_1^2 + \mu x_2^2 - 1 dx \,, \qquad \mu \geq 1 \tag{4.1}$$

The obvious solution to this optimization problem is the shape $\hat{\Omega}$ with boundary

$$\partial\hat{\Omega} = \{x \in \mathbb{R}^2 \mid x_1^2 + \mu^2 x_2^2 = 1\}$$

And from the discussion above, we observe that the Riemannian shape Hessian at this solution is a multiplication operator

$$Hess f(\hat{\Omega})V = \nu \cdot V \,, \qquad \nu(s) \in [\frac{2}{\mu}, 2], s \in \partial\hat{\Omega}$$

Therefore, the Hessian is coercive and thus the solution is locally unique and the Newton iteration should converge quadratically close to the solution.

The application of the exponential mapping within the Newton method is an expensive operation. It is recommendable to replace this step by a so-called *retraction* mapping [1, 32, 25], i.e., a smooth mapping $r_x : T_x B_e \to B_e$ with the following properties:

a) $r_x(0) = x$

b) $Dr_x(0) = id_{T_x B_e}$ (local rigidity condition [1])

with the canonical identification $T_0 T_x B_e \cong T_x B_e$. Properties a) and b) are also satisfied by the exponential mapping. An example of this kind of mapping is the following mapping that we will use in our implementation of Newton optimization methods on B_e:

$$r_x : T_x B_e \to B_e$$
$$\eta \mapsto \left(x + \eta : \begin{array}{l} S^1 \to \mathbb{R}^2 \\ \theta \mapsto x(\theta) + \eta(\theta) \end{array} \right)$$

The implementation of the retraction r_x defined above is much simpler than the implementation of the exponential map. However, one should keep in mind that this retraction does not leave B_e only for tangent vectors η which are small enough.

Now, the Newton method for optimization is generalized to

$$x^{k+1} = r_{x^k}(-M^k F(x^k)) \tag{4.2}$$

where the special case $r_x = \exp_x$ considered above falls also into this class of iterations. Since any retraction is a smooth mapping and because of the local rigidity, we observe that

$$d_G(x, r_x(\eta)) = d_G(x, \exp_x(\eta)) + cd_G(x, \exp_x(\eta))^2$$

for some constant c and for small enough tangential vectors η. Plugging this into the estimations (3.1, 3.2) and (3.3, 3.4), we conclude that locally linear, superlinear and quadratic convergence properties are not changed by using more general retractions rather than the special case of the exponential map.

The distance of a shape iterate Ω^k to the optimal shape $\hat{\Omega}$ should be measured by the length of a connecting geodesic. However, in order to reduce the numerical effort, we exploit the fact that a connecting geodesic is also the result of an exponential map, which on the other hand is an isometry. Thus, we use for shapes close enough

to the optimal solution such that $\Omega^k = \hat{\Omega} + h$ with $h = \alpha\vec{n}|_{\partial\hat{\Omega}}$ as second order accurate substitute for the geodesic distance the measure

$$\tilde{d}(\Omega^k, \hat{\Omega}) = \int_{\partial\hat{\Omega}} |\alpha| ds \qquad (4.3)$$

i.e., just the absolute value of the area between the iterates and the optimal shape.

Because of the simplicity of the objective (4.1), the integrations for the evaluations of the objective can be performed in the form of integration along the shape boundary in polar coordinates in the following way

$$\int_{\Omega} x_1^2 + \mu^2 x_2^2 - 1 dx = \frac{1}{\mu} \int_{S^1} \frac{\|c^\mu(\theta)\|^3}{3} - \|c^\mu(\theta)\| \frac{d\theta}{\|c_\theta^\mu(\theta)\|}$$

where $c^\mu := \partial \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} \Omega$ is the boundary of the domain Ω scaled by μ in the x_2-direction and the last integral can be numerically evaluated by usage of the trapezoidal rule for a piecewise linear boundary discretization. Similarly, in the line of (4.3), first order approximations of the distance of curves to the optimal solution, which differ only by a small amount are computed by

$$\bar{d}(\Omega^k, \hat{\Omega}) = \int_{\partial\hat{\Omega}} |\, \|c(s)\| - \|\hat{c}(s)\| \,| \, ds = \frac{1}{\mu} \int_{S^1} |\, \|c^\mu(s)\| - 1| \frac{d\theta}{\|c_\theta^\mu(\theta)\|}$$

Furthermore, the tangential (and thus the unit normal) vector field for the iterates is computed by central differences.

The first idea for setting μ in the numerical experiments is $\mu = 1$. However, then $Hess\, f(\hat{\Omega}) = 2 \cdot id_{T_{\partial\hat{\Omega}} B_e}$ which means that the method of steepest descent with exact line search can exactly mimic the behaviour of the Newton method—including quadratic convergence. However, we want to experience the difference between steepest descent methods and Newton's method in this example. Therefore, we choose $\mu = 2$ for performing the computations. The initial shape is defined by its boundary $\partial\Omega^0 := c^0$ as

$$c^0(s) = \frac{1}{2} \begin{pmatrix} \cos(s) - 0.15|1 - \sin(2s)| \cos(s) \\ \sin(s) - 0.15|1 - \cos(2s)| \cos(s) \end{pmatrix}, \qquad s \in [0, 2\pi]$$

just to be somewhat more interesting than a simple circle. We use a piecewise linear approximation with 100 equidistant pieces in $[0, 2\pi]$. Table 4.1 shows the performance of the shape steepest descent method versus the shape Newton method with exact line search, where we use $A = 0$ in the definition of G^A similarly as in [20]. The iterations are stopped at a distance of less than 10^{-7} to the solution. For the steepest descent method, we observe linear convergence with a factor of approximately 0.3, while for the Newton method we see clear quadratic convergence. Figure 4 shows the various shape boundaries during the respective iterations.

5. Conclusions. In this paper, a novel point of view on shape optimization is presented—the Riemannian point of view. The Riemannian shape Hessian is introduced which can serve as a much more useful notion of second shape derivative than the classical so-called shape Hessian. With this term, classical optimization results, known in linear spaces, can be formulated and proofed in the area of shape optimization. Newton optimization methods for shape optimization are analyzed. A new model problem for shape optimization is introduced which mimics the properties of a standard L^2-quadratic model problem in linear spaces—in so far as the Riemannian Hessian is just a multiplicative operator.

TABLE 4.1
Performance of shape algorithms: steepest descent (indices SD) versus Newton (indices NM), α denotes the line search parameter chosen by an exact line search applied to the objective.

| It.-No.	$f(\Omega_{SD}^k)$	$\bar{d}(\Omega_{SD}^k, \hat{\Omega})$	α_{SD}	$f(\Omega_{NM}^k)$	$\bar{d}(\Omega_{NM}^k, \hat{\Omega})$	α_{NM}
0	-1.7664	0.9222	0.51	-1.7664	0.9222	0.64
1	-2.0716	0.2142	0.31	-2.0865	0.1359	1.00
2	-2.0921	0.5882E-01	0.37	-2.0944	0.4265E-02	1.00
3	-2.0942	0.1662E-01	0.32	-2.0944	0.1037E-04	1.00
4	-2.0944	0.4459E-02	0.35	-2.0944	0.2590E-09	
5	-2.0944	0.1356E-02	0.33			
6	-2.0944	0.3951E-03	0.34			
7	-2.0944	0.1218E-03	0.33			
8	-2.0944	0.3643E-04	0.35			
9	-2.0944	0.1194E-04	0.32			
10	-2.0944	0.3421E-05	0.35			
11	-2.0944	0.1137E-05	0.32			
12	-2.0944	0.3305E-06	0.35			
13	-2.0944	0.1109E-06	0.30			
14	-2.0944	0.2894E-07	0.39			
15	-2.0944	0.1008E-07	0.36			
16	-2.0944	0.3723E-08				
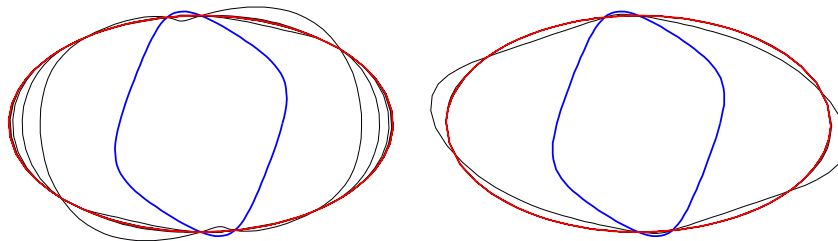
FIG. 4.1. *Visualization of the shape iterates for the steepest descent (left) and for Newton's method (right) from initial (blue) to solution (red) shape.*

Several aspects are now open for further research: This paper only deals with C^∞-boundaries of 2D shapes. Of course, for practical purposes, this is not enough. The regularity has to be reduced, one has to go from dimension 2 to 3 and one has to treat shapes which are just part of submanifolds of $\mathbb{R}^3$, rather than whole embeddings of the unit sphere. Furthermore, the implications for practical shape optimization have to be analyzed. This will all be covered in subsequent papers.

REFERENCES

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, 2008.
[2] E. Arian. Analysis of the hessian for aeroelastic optimization. Technical Report 95-84, Institute for Computer Applications in Science and Engineering (ICASE), 1995.
[3] E. Arian and S. Ta'asan. Analysis of the hessian for aerodynamic optimization: Inviscid flow. Technical Report 96-28, Institute for Computer Applications in Science and Engineering (ICASE), 1996.
[4] E. Arian and V. N. Vatsa. A preconditioning method for shape optimization governed by the

euler equations. Technical Report 98-14, Institute for Computer Applications in Science and Engineering (ICASE), 1998.

[5] M. Bauer, P. Harms, and P.W. Michor. Sobolev metrics on shape space ii: Weighted sobolev metrics and almost local metrics. arXiv:1109.0404., 2011.

[6] M. Bauer, P. Harms, and P.W. Michor. Sobolev metrics on shape space of surfaces. *Journal of Geometric Mechanics*, 3(4):389–438, 2011.

[7] H.G. Bock. Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. *Bonner Mathematische Schriften 183*, Bonn, 1987.

[8] R.H. Byrd and J. Nocedal. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *SIAM Journal on Numerical Analysis*, 26(3):727–739, 1989.

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

[10] J.E. Dennis and J.J. Moré. A characterization of superlinear convergence and its application to quasi-Newton methods. *Math. Comp.*, 28:549–560, 1974.

[11] P. Deuflhard. *Newton Methods for Nonlinear Problems, Affine Invariance and Adaptive Algorithms.* Number 35 in Springer Series in Computational Mathematics. Springer Berlin, Heidelberg, 2004.

[12] K. Eppler. Second derivatives and sufficient optimality conditions for shape functionals. *Control and Cybernetics*, 29:485–512, 2000.

[13] K. Eppler and H. Harbrecht. A regularized newton method in electrical impedance tomography using shape Hessian information. *Control and Cybernetics*, 34(1):203–225, 2005.

[14] K. Eppler and H. Harbrecht. Shape optimization for free boundary problems – analysis and numerics. In 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*, volume 160 of *International Series of Numerical Mathematics*, pages 277–288. Birkhäuser Basel, Boston, Stuttgart, 2012.

[15] K. Eppler, S. Schmidt, V. Schulz, and C. Ilic. Preconditioning the pressure tracking in fluid dynamics by shape hessian information. *Journal of Optimization Theory and Applications*, 141(3):513–531, 2009.

[16] O.P. Ferreira and B.F. Svaiter. Kantorovich's theorem on Newton's method in Riemannian manifolds. *Journal of Complexity*, 18(1):304–329, 2002.

[17] N. Gauger, C. Ilic, S. Schmidt, and V. Schulz. Non-parametric aerodynamic shape optimization. In G. Leugering, S. Engell, A. Griewank, M. Hinzea, R. Rannacher, V. Schulz, M. Ulbrich, and S. Ulbrich, editors, *Constrained Optimization and Optimal Control for Partial Differential Equations*, volume 160, pages 289–300. Birkhäuser, Basel, Boston, Berlin, 2012.

[18] M. Hintermüller. Fast level set based algorithms using shape and topological sensitivity information. *Control and Cybernetics*, 34(1):305–324, 2005.

[19] R. Hiptmair and J. Li. Shape derivatives in differential forms i: An intrinsic perspective. Technical Report 2011/42, Seminar for Applied Mathematics, ETH Zürich, 2011.

[20] S. Joshi, E. Klassen, A. Srivastava, and I. H. Jermyn. A novel representation for Riemannian analysis of elastic curves in $\mathbb{R}^n$. In *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 1–7, Minneapolis, USA, June 2007.

[21] Peter W. Michor and David Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. In *Applied and Computational Harmonic Analysis, doi:10.1016/j.acha.2006.07.004. METRIC ON SHAPE SPACE WITH EXPLICIT GEODESICS 35*, 2006.

[22] P.M. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8:1–48, 2006.

[23] P.W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documeta Math.*, 10:217–245, 2005.

[24] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids.* Numerical Mathematics and Scientific Computation. Clarendon Press Oxford, 2001.

[25] Chunhong Qi. *Numerical Optimization Methods on Riemannian Manifolds.* PhD thesis, Florida State University, 2011.

[26] C. Schillings, S. Schmidt, and V. Schulz. Efficient shape optimization for certain and uncertain aerodyamic design. *Computers and Fluids*, 46(1):78–87, 2011.

[27] S. Schmidt. *Efficient Large Scale Aerodynamic Design Based on Shape Calculus.* PhD thesis, Universit"at Trier, 2010.

[28] S. Schmidt and V. Schulz. Impulse response approximations of discrete shape hessians with application in CFD. *SIAM Journal on Control and Optimization*, 48(4):2562–2580, 2009.

[29] S. Schmidt and V. Schulz. Shape derivatives for general objective functions and the incompressible navier-stokes equations. *Control and Cybernetics*, 39(3):677–713, 2010.

[30] E. Sharon and D. Mumford. 2d-shape analysis using conformal mapping. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, volume Vol. II, pages 350–359, 2004.

14

[31] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization: Shape Sensitivity Analysis.* Springer, 1992.

[32] C. Udriste. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its Applications.* Kluwer academic publishers, 1994.