

TOWARDS A LAGRANGE–NEWTON APPROACH FOR PDE CONSTRAINED SHAPE OPTIMIZATION

VOLKER H. SCHULZ, MARTIN SIEBENBORN, AND KATHRIN WELKER*

Abstract. The novel Riemannian view on shape optimization developed in [27] is extended to a Lagrange–Newton approach for PDE constrained shape optimization problems. The extension is based on optimization on Riemannian vector space bundles and exemplified for a simple numerical example.

1. Introduction. Shape optimization problems arise frequently in technological processes, which are modeled in the form of partial differential equations as in [2, 3, 4, 13, 14, 24, 25, 26]. In many practical circumstances, the shape under investigation is parameterized by finitely many parameters, which on the one hand allows the application of standard optimization approaches, but on the other hand limits the space of reachable shapes unnecessarily. Shape calculus, which has been the subject of several monographs [12, 22, 29] presents a way out of that dilemma. However, so far it is mainly applied in the form of gradient descent methods, which can be shown to converge. The major difference between shape optimization and the standard PDE constrained optimization framework is the lack of the linear space structure in shape spaces. If one cannot use a linear space structure, then the next best structure is the Riemannian manifold structure as discussed for shape spaces in [5, 6, 19, 20, 21]. The publication [27] makes a link between shape calculus and shape manifolds and thus enables the usage of optimization techniques on manifolds in the context of shape optimization.

PDE constrained shape optimization however, is confronted with function spaces defined on varying domains. The current paper presents a vector bundle framework based on the Riemannian framework established in [27], which enables the discussion of Lagrange–Newton methods within the shape calculus framework for PDE constrained shape optimization.

The paper first presents the novel Riemannian vector bundle framework on section 2, discusses this approach for a specific academic example in section 3 and presents numerical results in section 4 .

2. Constrained Riemannian shape optimization. The typical set-up of an equality constrained optimization problem is

$$\begin{aligned} \min_{y,u} J(y,u), \quad J: Y \times U \rightarrow \mathbb{R} \\ \text{s.t. } c(y,u) = 0, \quad c: Y \times U \rightarrow Z \end{aligned}$$

where U, Y, Z are linear spaces and c, J sufficiently smooth nonlinear functions [8]. In some situations the constraint c allows to apply the implicit function theorem in order to define a unique control to state mapping $y(u)$ and thus the constrained problem maybe reduced to an unconstrained one of the form

$$\min_u J(y(u), u).$$

However, the constrained formulation is often computationally advantageous, because it allows the usage of pre-existing solver technology for the constraint and it is geared towards an efficient SAND (simultaneous analysis and design) or one-shot approach based on linear KKT systems. So far, shape optimization methods

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

based on the shape calculus, have been mainly considered with the reduced black-box framework above via the implicit function theorem – mainly because the set of all admissible shapes is typically not a linear space – unlike the space U above. The publication [27] has developed a Riemannian framework for shape optimization in the reduced unconstrained paradigm, which enables Newton-like iteration techniques and convergence results. This publication aims at generalizing those results to the constrained perspective – in particular for the case that the constraint is of the form of a set of partial differential equations (PDE).

Within that framework, the space Y for the state variable is a linear (function) space depending explicitly on $u \in U$, e.g., $H^1(\Omega(u))$, where $\Omega(u)$ is the interior of a shape u . This line of thinking leads to vector bundles of function spaces as discussed in detail in [18]. Thus, we now consider a Riemannian manifold (\mathcal{N}, G) of class C^q ($q \geq 0$), where G is a smooth mapping assigning any point $p \in \mathcal{N}$ an inner product $G_p(\cdot, \cdot)$ on the tangential bundle $T\mathcal{N}$. For each $u \in \mathcal{N}$, there is given a Hilbert space $H(u)$ such that the set

$$E := \{(H(u), u) \mid u \in \mathcal{N}\}$$

is the total space of a vector bundle (E, π, \mathcal{N}) . In particular, there is a bundle-projection $\pi : E \rightarrow \mathcal{N}$ and for an open covering $\{U_i\}$ of \mathcal{N} a local C^q isomorphism

$$\tau_i : \pi^{-1}(U_i) \rightarrow H_0 \times U_i$$

where H_0 is a Hilbert space. In particular, we have an isomorphism on each fiber

$$\tau_i(u) : \pi^{-1}(x) = H(u) \rightarrow H_0$$

and for $u \in U_i \cap U_j$, the mapping $\tau_i(u) \circ \tau_j(u)^{-1} : H_0 \rightarrow H_0$ is a linear isomorphism. The total space E of the vector bundle (E, π, \mathcal{N}) is by itself a Riemannian manifold, where the tangential bundle TE satisfies

$$T_{(y,u)}E \cong H(u) \times T_y\mathcal{N}.$$

In Riemannian geometry, tangential vectors are considered as first order differential operators acting on germs of scalar valued functions (e.g. [10]). Such a differential operator will be notated by $h(J)(e)$, if $J : E \rightarrow \mathbb{R}$ is a differentiable function and $e \in E$. We will have to deal with derivatives, where we will always use directional derivatives of scalar valued functions only, but notate them in the usual fashion. Let the derivative of J at e in direction $h \in TE$ be denoted by $DJ(e)h$. Then, we define in this setting

$$DJ(e)h := h(J)(e), \quad h \in TE.$$

In particular, we denote

$$\begin{aligned} \frac{\partial}{\partial y} J(y, u) h_y &:= h_1(J)(y, u), \quad h_1 := (h_y, 0) \in TE \\ \frac{\partial}{\partial u} J(y, u) h_u &:= h_2(J)(y, u), \quad h_2 := (0, h_u) \in TE \end{aligned}$$

where $h_y \in H(u)$ and $h_u \in T_y\mathcal{N}$, if $h_1, h_2 \in T_{(y,u)}E$.

We consider now the following constrained optimization problem

$$\min_{(y,u) \in E} J(y, u), \quad J : E \rightarrow \mathbb{R} \tag{2.1}$$

$$\text{s.t. } a_u(y, p) = b_u(p), \quad \forall p \in H \tag{2.2}$$

where $a_u(\cdot, \cdot)$ is a bilinear form and $b_u(\cdot)$ a linear form defined on the fiber H which are C^q with respect to u . The scalar valued function J is assumed to be C^q . Intentionally, the weak formulation of the PDE is chosen for ease of presentation. Now, it will be necessary to define the Lagrangian \mathcal{L} in order to formulate the adjoint and design equation to the constrained optimization problem (2.1–2.2).

DEFINITION 2.1. *We define the Lagrangian in the setting above for $(y, u, p) \in F$ as*

$$\mathcal{L}(y, u, p) := J(y, u) + a_u(y, p) - b_u(p)$$

where $F := \{(H(u), u, H(u)) \mid u \in \mathcal{N}\}$ with $T_{(y, u, p)}F \cong H(u) \times T_y\mathcal{N} \times H(u)$.

Let $(\hat{y}, \hat{u}) \in E$ solves the optimization problem (2.1–2.2). Then, the (adjoint) variational problem which we get by differentiating \mathcal{L} with respect to y is given by

$$a_{\hat{u}}(z, p) = -\frac{\partial}{\partial y} J(\hat{y}, \hat{u})z, \quad \forall z \in H(\hat{u}) \quad (2.3)$$

and the design problem which we get by differentiating \mathcal{L} with respect to u is given by

$$\left. \frac{\partial}{\partial u} \right|_{u=\hat{u}} [J(\hat{y}, u) + a_u(\hat{y}, \hat{p}) - b_u(\hat{p})] w = 0, \quad \forall w \in T_{\hat{u}}\mathcal{N} \quad (2.4)$$

where $\hat{p} \in H$ solves (2.3). If we differentiate \mathcal{L} with respect to p , we get the state equation (2.2). These (KKT) conditions (2.2–2.4) could be collected in the following condition:

$$D\mathcal{L}(\hat{y}, \hat{u}, \hat{p})h = 0, \quad \forall h \in T_{(y, u, p)}F. \quad (2.5)$$

REMARK 1. *In a vector space setting, the existence of a solution $p \in H$ of the (adjoint) variational problem (2.3) is typically guaranteed by so-called constraint qualifications. From this point of view, here, the existence itself can be interpreted as formulation of a constraint qualification.*

By using a Riemannian metric G on $T\mathcal{N}$ and a smoothly varying scalar product $\langle \cdot, \cdot \rangle_u$ on the Hilbert space $H(u)$, we can envision $T_{(y, u, p)}F$ as a Hilbert space with a canonical scalar product

$$\left\langle \begin{pmatrix} z_1 \\ w_1 \\ q_1 \end{pmatrix}, \begin{pmatrix} z_2 \\ w_2 \\ q_2 \end{pmatrix} \right\rangle_{T_{(y, u, p)}F} := \langle z_1, z_2 \rangle_u + G_u(w_1, w_2) + \langle q_1, q_2 \rangle_u \quad (2.6)$$

and thus also $(F, \langle \cdot, \cdot \rangle_{TF})$ as Riemannian manifold. This scalar product can be used to apply the Riesz representation theorem in order to define the gradient of the Lagrangian $\text{grad}\mathcal{L} \in TF$ by the condition

$$\langle \text{grad}\mathcal{L}, h \rangle_{T_{(y, u, p)}F} := D\mathcal{L}(y, u, p)h, \quad \forall h \in T_{(y, u, p)}F.$$

Now, similar to standard nonlinear programming we can solve the problem of finding $(y, u, p) \in F$ with

$$\text{grad}\mathcal{L}(y, u, p) = 0 \quad (2.7)$$

as a means to find solutions to the optimization problem (2.1–2.2). The nonlinear problem (2.7) has exactly the form of the root finding problems discussed in [27]. Exploiting the Riemannian structure on TF , we can formulate a Newton iteration involving the Riemannian Hessian which is based on the resulting Riemannian connection:

k -th iteration:

(1) compute increment $\Delta\xi$ as solution of

$$\text{Hess}\mathcal{L}(\xi^k)\Delta\xi = -\text{grad}\mathcal{L}(\xi^k) \quad (2.8)$$

(2) increment $\xi^{k+1} := \exp_{\xi^k}(\alpha^k \cdot \Delta\xi)$ for some steplength α^k

This iteration will be detailed out below. However, before that, we have to specify the scalar product on the Hilbert space involved. Since we will use the exponential map based on the Riemannian metric on F , we would like to choose a metric that is in the Hilbert space parts as simple as possible. Therefore we use the metric defined on the Hilbert space $(H_0, \langle \cdot, \cdot \rangle_0)$ and transfer that canonically to the Hilbert spaces $H(u)$. Thus, we assume now that in the sequel we only have to deal with one particular chart (U_i, τ_i) from the covering $\{U_i\}$ and define there

$$\langle z_1, z_2 \rangle_u := \langle \tau_i(u)z_1, \tau_i(u)z_2 \rangle_0, \quad \forall u \in U_i.$$

Now, geodesics in the Hilbert space parts of F are represented just by straight lines in H_0 and the exponential map can be expressed in the form

$$\begin{aligned} & \exp_{(y,u,p)}(z, w, q) \\ &= \left(\tau_i \left(\exp_u^{\mathcal{N}}(w) \right)^{-1} \circ \tau_i(u)(y + z), \exp_u^{\mathcal{N}}(w), \tau_i \left(\exp_u^{\mathcal{N}}(w) \right)^{-1} \circ \tau_i(u)(p + q) \right) \end{aligned}$$

where $\exp^{\mathcal{N}}$ denotes the exponential map on the manifold \mathcal{N} .

Within iteration (2.8), the Hessian has to be discussed. It is based on the Riemannian connection ∇ on F at $u \in \mathcal{N}$. The expression $\nabla_u^{\mathcal{N}}$ may denote the Riemannian covariant derivative on $T_u\mathcal{N}$. Since the scalar product in H is completely independent from the location $u \in \mathcal{N}$, we observe that mixed covariant derivatives of vectors from H with respect to tangential vectors in $T\mathcal{N}$ are reduced to simple directional derivatives – which is the case for derivatives in linear spaces anyway. Thus:

$$\begin{aligned} \nabla_{(h_y, h_u, h_p)}: T_{(y,u,p)}F &\rightarrow T_{(y,u,p)}F \\ \begin{pmatrix} z \\ w \\ q \end{pmatrix} &\mapsto \begin{pmatrix} \frac{\partial}{\partial y} z[h_y] + \frac{\partial}{\partial u} z[h_u] + \frac{\partial}{\partial p} z[h_p] \\ \frac{\partial}{\partial y} w[h_y] + \nabla_u^{\mathcal{N}} w[h_u] + \frac{\partial}{\partial p} w[h_p] \\ \frac{\partial}{\partial y} q[h_y] + \frac{\partial}{\partial u} q[h_u] + \frac{\partial}{\partial p} q[h_p] \end{pmatrix} \end{aligned}$$

From the definition of the Hessian as $\text{Hess}\mathcal{L}[h] := \nabla_h \text{grad}\mathcal{L}$ we conclude the following block structure of the Hessian:

$$\text{Hess}\mathcal{L} = \begin{pmatrix} D_y \text{grad}_y \mathcal{L} & D_u \text{grad}_y \mathcal{L} & D_p \text{grad}_y \mathcal{L} \\ D_y \text{grad}_u \mathcal{L} & \nabla_u^{\mathcal{N}} \text{grad}_u \mathcal{L} & D_p \text{grad}_u \mathcal{L} \\ D_y \text{grad}_p \mathcal{L} & D_u \text{grad}_p \mathcal{L} & 0 \end{pmatrix} \quad (2.9)$$

From a practical point of view, it may be advantageous to solve equation (2.8) in a weak formulation as

$$\nabla(D\mathcal{L}(y, u, p)h) \begin{pmatrix} z \\ w \\ q \end{pmatrix} = -D\mathcal{L}(y, u, p)h, \quad \forall h \in T_{(y,u,p)}F \quad (2.10)$$

i.e., in detail, the following equations have to be satisfied for all $h := (\bar{z}, \bar{w}, \bar{q})^T \in T_{(y,u,p)}F$:

$$H_{11}(z, \bar{z}) + H_{12}(w, \bar{z}) + H_{13}(q, \bar{z}) = -a_u(\bar{z}, p) - \frac{\partial}{\partial y} J(y, u) \bar{z} \quad (2.11)$$

$$H_{21}(z, \bar{w}) + H_{22}(w, \bar{w}) + H_{23}(q, \bar{w}) = -\frac{\partial}{\partial u} [J(y, u) + a_u(y, p) - b_u(p)] \bar{w} \quad (2.12)$$

$$H_{31}(z, \bar{q}) + H_{32}(w, \bar{q}) + H_{33}(q, \bar{q}) = -a_u(y, \bar{q}) + b_u(\bar{q}) \quad (2.13)$$

where

$$\begin{aligned}
H_{11}(z, \bar{z}) &= \frac{\partial^2}{\partial y^2} J(y, u) z \bar{z} \\
H_{12}(w, \bar{z}) &= \frac{\partial}{\partial u} \left[a_u(\bar{z}, p) + \frac{\partial}{\partial y} J(y, u) \bar{z} \right] w \\
H_{13}(q, \bar{z}) &= a_u(\bar{z}, q) \\
H_{21}(z, \bar{w}) &= \frac{\partial}{\partial y} \frac{\partial}{\partial u} ([J(y, u) + a_u(y, p)] \bar{w}) z \\
H_{22}(w, \bar{w}) &= G(\text{Hess}^{\mathcal{N}}(J(y, u) + a_u(y, p) - b_u(p)) w, \bar{w}) \\
H_{23}(q, \bar{w}) &= \frac{\partial}{\partial u} [a_u(y, q) - b_u(q)] \bar{w} \\
H_{31}(z, \bar{q}) &= a_u(z, \bar{q}) \\
H_{32}(w, \bar{q}) &= \frac{\partial}{\partial u} [a_u(y, \bar{q}) - b_u(\bar{q})] w \\
H_{33}(q, \bar{q}) &= 0
\end{aligned}$$

One should note that the covariant derivative ∇ reveals natural symmetry properties and thus obvious symmetries can be observed in the components above not involving second shape derivatives. A key observation in [27] is that even the expression $H_{22}(w, \bar{w})$ is symmetric in the solution of the shape optimization problem. This motivates a shape-SQP method as outlined below, where away from the solution only expressions in $H_{22}(w, \bar{w})$ are used which are nonzero at the solution. Its basis is the following observation:

If the term $H_{22}(w, \bar{w})$ is replaced by an approximation $\hat{H}_{22}(w, \bar{w})$, which omits all terms in $H_{22}(w, \bar{w})$, which are zero at the solution and if the reduced Hessian of (2.9) built with this approximation is coercive, equation (2.10) is equivalent to the linear-quadratic problem

$$\min_{(z, w)} \frac{1}{2} \left(H_{11}(z, z) + 2H_{12}(w, z) + \hat{H}_{22}(w, w) \right) + \frac{\partial}{\partial y} J(y, u) z + \frac{\partial}{\partial u} J(y, u) w \quad (2.14)$$

$$\text{s.t. } a_u(z, \bar{q}) + \frac{\partial}{\partial u} [a_u(y, \bar{q}) - b_u(\bar{q})] w = -a_u(y, \bar{q}) + b_u(\bar{q}), \quad \forall \bar{q} \in H(u) \quad (2.15)$$

where the adjoint variable to the constraint (2.15) is just $p + q$. In the next sections, we also omit terms in H_{11} and H_{12} , which are zero, when evaluated at the solution of the optimization problems. Nevertheless, quadratic convergence of the resulting SQP method is to be expected and indeed observed in section 4.

3. Discussion for a Poisson-type model problem. In this section, we apply the theoretical discussion of section 2 to a PDE constrained shape optimization problem, which is inspired by the standard tracking-type elliptic optimal control problem and motivated by electrical impedance tomography. It is very close to the model problem of example 2 in [9] and the inverse interface problem in [17].

Let the domain $\Omega := (0, 1)^2 \subset \mathbb{R}^2$ split into the two subdomains $\Omega_1, \Omega_2 \subset \Omega$ such that $\Omega_1 \cup \Gamma \cup \Omega_2 = \Omega$ and $\partial\Omega_1 \cap \partial\Omega_2 = \Gamma$. The interface Γ is replaced by u and an element of the following manifold

$$B_e^0([0, 1], \mathbb{R}^2) := \text{Emb}^0([0, 1], \mathbb{R}^2) / \text{Diff}^0([0, 1])$$

i.e., an element of the set of all equivalence classes of the set of embeddings

$$\text{Emb}^0([0, 1], \mathbb{R}^2) := \{ \phi \in C^\infty([0, 1], \mathbb{R}^2) \mid \phi(0) = (0.5, 0), \phi(1) = (0.5, 1), \\ \phi \text{ injective immersion} \}$$

where the equivalence relation is defined by the set of all C^∞ re-parameterizations, i.e., by the set of all diffeomorphisms

$$\text{Diff}^0([0, 1], \mathbb{R}^2) := \{\phi: [0, 1] \rightarrow [0, 1] \mid \phi(0) = (0.5, 0), \phi(1) = (0.5, 1), \\ \phi \text{ diffeomorphism}\}.$$

In Figure 3 the construction of the domain Ω from the interface $u \in B_e^0([0, 1], \mathbb{R}^2)$ is illustrated. Now, we consider Ω dependent on u . Therefore, we denote it by $\Omega(u) = \Omega_1(u) \cup u \cup \Omega_2(u)$.

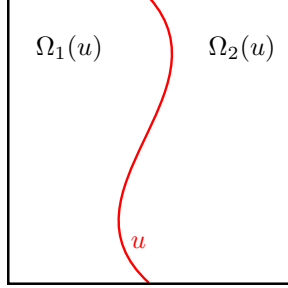


FIG. 3.1. Example of a domain $\Omega(u) = \Omega_1(u) \cup u \cup \Omega_2(u)$

REMARK 2. The manifold $B_e^0([0, 1], \mathbb{R}^2)$ is constructed in analogy to the manifold $B_e(S^1, \mathbb{R}^2)$ in [20] as a set of equivalence classes in a set of embeddings with respect to a equivalence relation which is given by a set of diffeomorphisms. Moreover, a particular point on the manifold $B_e^0([0, 1], \mathbb{R}^2)$ is represented by a curve $c: [0, 1] \rightarrow \mathbb{R}^2$, $\theta \mapsto c(\theta)$. Because of the equivalence relation $\text{Diff}([0, 1])$, the tangent space is isomorphic to the set of all smooth vector fields along c , i.e.,

$$T_c B_e^0([0, 1], \mathbb{R}^2) \cong \{h \mid h = \alpha n, \alpha \in C^\infty([0, 1], \mathbb{R})\}$$

where n is the unit outer normal to $\Omega_1(u)$ at u . Thus, all considerations of [27] carry easily over to our manifold $B_e^0([0, 1], \mathbb{R}^2)$.

The PDE constrained shape optimization problem is given in strong form by

$$\min_u J(y, u) \equiv \frac{1}{2} \int_{\Omega(u)} (y - \bar{y})^2 dx + \mu \int_u 1 ds \quad (3.1)$$

$$\text{s.t. } -\Delta y = f \quad \text{in } \Omega(u) \quad (3.2)$$

$$y = 0 \quad \text{on } \partial\Omega(u) \quad (3.3)$$

where

$$f \equiv \begin{cases} f_1 = \text{const.} & \text{in } \Omega_1(u) \\ f_2 = \text{const.} & \text{in } \Omega_2(u) \end{cases}. \quad (3.4)$$

The perimeter regularization with $\mu > 0$ in the objective (3.1) is a frequently used means to overcome ill-posedness of the optimization problem (e.g. [1]). Let n be the unit outer normal to $\Omega_1(u)$ at u . We observe that the unit outer normal to $\Omega_2(u)$ at u is equal to $-n$, which enables us to use only one normal n for the subsequent discussions. Furthermore, we have interface conditions at the interface u . We formulate explicitly the continuity of the state and of the flux at the boundary u as

$$\llbracket y \rrbracket = 0, \quad \left[\left[\frac{\partial y}{\partial n} \right] \right] = 0 \quad \text{on } u \quad (3.5)$$

where the jump symbol $[[\cdot]]$ denotes the discontinuity across the interface u and is defined by $[[v]] := v_1 - v_2$ where $v_1 := v|_{\Omega_1}$ and $v_2 := v|_{\Omega_2}$.

The boundary value problem (3.2-3.5) is written in weak form as

$$a_u(y, p) = b_u(p), \quad \forall p \in H_0^1(\Omega(u)) \quad (3.6)$$

where

$$a_u(y, p) := \int_{\Omega(u)} \nabla y^T \nabla p \, dx - \int_u \left[\left[\frac{\partial y}{\partial n} \right] p \right] ds \quad (3.7)$$

$$b_u(p) := \int_{\Omega(u)} f p \, dx. \quad (3.8)$$

Now, F from definition 2.1 takes the specific form

$$F := \{ (H_0^1(\Omega(u)), u, H_0^1(\Omega(u))) \mid u \in B_\varepsilon^0([0, 1], \mathbb{R}^2) \}.$$

The metric in the vector space parts is constructed by employing a "mesh deformation". Mesh deformations are often used to deform a computational mesh smoothly in accordance with a deformation of the boundary of the computational domain. Here, we use this in the form of a deformation of the computational domain rather than of the mesh only and assume that there is a bijective C^∞ -mapping

$$\Phi_u : [0, 1]^2 \rightarrow \Omega(u),$$

e.g., Φ_u is the deformation given by the solution of a linear elasticity problem. Thus, we can construct the necessary bijective identification

$$\tau(u) : H_0^1(\Omega(u)) \rightarrow H_0^1((0, 1)^2), \quad g \mapsto g \circ \Phi_u.$$

We have to detail the expressions in equation (2.8) or respectively (2.10). First, the Lagrangian is defined for $(y, u, p) \in F$ as

$$\mathcal{L}(y, u, p) := J(y, u) + a_u(y, p) - b_u(p)$$

where $J(y, u)$ is defined in (3.1) and a_u, b_u are defined in (3.7, 3.8). Now, we focus on the shape derivative of \mathcal{L} in direction of a continuous vector field V . It is defined by

$$\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] := \lim_{t \rightarrow 0^+} \frac{\mathcal{L}(y, u_t, p) - \mathcal{L}(y, u, p)}{t} \quad (3.9)$$

if for all directions V this limit exists and the mapping $V \mapsto \frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V]$ is linear and continuous. The perturbed boundaries u_t in (3.9) are defined by

$$u_t := F_t(u) = \{F_t(x) : x \in u\} \text{ with } u_0 = u \quad (3.10)$$

where $F_t(x) := x + tV(x)$ denotes the perturbation of identity and $t \in [0, T]$ with $T > 0$.

REMARK 3. One should note that we get perturbed domains Ω_t given by

$$\Omega_t := F_t(\Omega(u)) = \{F_t(x) : x \in \Omega(u)\} \text{ with } \Omega_0 = \Omega(u) \quad (3.11)$$

due to the perturbed boundaries u_t .

REMARK 4. The perturbation of u or respectively $\Omega(u)$ could also be described by the velocity method, i.e., as the flow $F_t(x) := \xi(t, x)$ determined by the initial value problem

$$\begin{aligned} \frac{d\xi(t, x)}{dt} &= V(\xi(t, x)) \\ \xi(0, x) &= x \end{aligned} \quad (3.12)$$

instead of the perturbation of identity.

We first consider the objective J in (3.1) without perimeter regularization. Then the shape derivative $\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V]$ can be expressed as an integral over the domain $\Omega(u)$ as well as an integral over the interface u which is better suited for a finite element implementation as already mentioned for example in [12, remark 2.3, p. 531]. An important point to note here is that the shape derivative of our \mathcal{L} evaluated in its saddle-point is equal to the one of J due to the theorem of Correa and Seeger [11, theorem 2.1]. Such a saddle-point is given by

$$\frac{\partial \mathcal{L}(\Omega, y, p)}{\partial y} = \frac{\partial \mathcal{L}(\Omega, y, p)}{\partial p} = 0 \quad (3.13)$$

which leads to the adjoint equation

$$-\Delta p = -(y - \bar{y}) \quad \text{in } \Omega(u) \quad (3.14)$$

$$p = 0 \quad \text{on } \partial\Omega(u) \quad (3.15)$$

$$[[p]] = 0 \quad \text{on } u \quad (3.16)$$

$$\left[\left[\frac{\partial p}{\partial n} \right] \right] = 0 \quad \text{on } u \quad (3.17)$$

and to the state equation

$$-\Delta y = f \quad \text{in } \Omega(u). \quad (3.18)$$

Like in [28] we first deduce a representation of the shape derivative expressed as a domain integral which will later allow us to calculate the boundary expression of the shape derivative by means of integration by parts on the interface u . One should note however, that by the Hadamard structure theorem [29, theorem 2.27] only the normal part of the continuous vector field has an impact on its value. Applying the following common rule for differentiating domain integrals

$$\frac{d^+}{dt} \left(\int_{\Omega_t} \eta(t) \right) \Big|_{t=0} = \int_{\Omega} (D_m \eta + \operatorname{div}(V) \eta) \quad (3.19)$$

which was proved in [15, lemma 3.3] yields

$$\begin{aligned} \frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \lim_{t \rightarrow 0^+} \frac{\mathcal{L}(y, u_t, p) - \mathcal{L}(y, u, p)}{t} = \frac{d^+}{dt} \mathcal{L}(y, u_t, p) \Big|_{t=0} \\ &= \int_{\Omega(u)} D_m \left(\frac{1}{2} (y - \bar{y})^2 \right) + D_m (\nabla y^T \nabla p) - D_m (f p) \\ &\quad + \operatorname{div}(V) \left(\frac{1}{2} (y - \bar{y})^2 + \nabla y^T \nabla p - f p \right) dx \\ &\quad - \int_u D_m \left(\left[\left[\frac{\partial y}{\partial n} p \right] \right] \right) + \operatorname{div}_u(V) \left[\left[\frac{\partial y}{\partial n} p \right] \right] ds \end{aligned} \quad (3.20)$$

where D_m denotes the material derivative with respect to $F_t = id + tV$ which is defined by

$$D_m(j(x)) := \lim_{t \rightarrow 0^+} \frac{(j \circ F_t)(x) - j(x)}{t} = \frac{d^+}{dt} (j \circ F_t)(x) \Big|_{t=0} \quad (3.21)$$

for a generic function $j: \Omega_t \rightarrow \mathbb{R}$. For the material derivative the product rule holds. Moreover, the following equality was proved in [7]

$$D_m(\nabla j) = \nabla(D_m(j)) - \nabla V^T \nabla j. \quad (3.22)$$

Combining (3.20), the product rule and (3.22) we obtain

$$\begin{aligned}
\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \int_{\Omega(u)} (y - \bar{y}) D_m(y) + \nabla(D_m(y))^T \nabla p + \nabla y^T \nabla(D_m(p)) \\
&\quad - \nabla y^T (\nabla V + \nabla V^T) \nabla p - D_m(f)p - f D_m(p) \\
&\quad + \operatorname{div}(V) \left(\frac{1}{2} (y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) dx \\
&\quad - \int_u \left[D_m \left(\frac{\partial y}{\partial n} \right) p + \frac{\partial y}{\partial n} D_m(p) \right] + \operatorname{div}_u(V) \left[\frac{\partial y}{\partial n} p \right] ds.
\end{aligned} \tag{3.23}$$

From this we get

$$\begin{aligned}
\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \int_{\Omega(u)} ((y - \bar{y}) - \triangle p) D_m(y) + (-\triangle y - f) D_m(p) \\
&\quad - \nabla y^T (\nabla V + \nabla V^T) \nabla p - D_m(f)p \\
&\quad + \operatorname{div}(V) \left(\frac{1}{2} (y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) dx \\
&\quad + \int_u \left[\frac{\partial p}{\partial n} D_m(y) - D_m \left(\frac{\partial y}{\partial n} \right) p \right] + \operatorname{div}_u(V) \left[\frac{\partial y}{\partial n} p \right] ds.
\end{aligned} \tag{3.24}$$

To deal with the term $D_m(f)p$, we note that the shape derivative of a generic function $j: \Omega_t \rightarrow \mathbb{R}$ with respect to the vector field V is given by

$$Dj[V] := D_m j - V^T j. \tag{3.25}$$

Therefore $D_m(f)p$ is equal to $pV^T \nabla f$ in the both subdomains $\Omega_1(u)$, $\Omega_2(u)$. Due to the continuity of the state and of the flux (3.5) their material derivative is continuous. Thus, we get

$$\left[\frac{\partial p}{\partial n} D_m(y) \right] = D_m(y) \left[\frac{\partial p}{\partial n} \right] \stackrel{(3.17)}{=} 0 \quad \text{on } u \tag{3.26}$$

$$\left[D_m \left(\frac{\partial y}{\partial n} \right) p \right] = D_m \left(\frac{\partial y}{\partial n} \right) [p] \stackrel{(3.16)}{=} 0 \quad \text{on } u. \tag{3.27}$$

That

$$\left[\frac{\partial y}{\partial n} p \right] = 0 \quad \text{on } u \tag{3.28}$$

follows from (3.5), (3.16) and the identity

$$[ab] = [a] b_1 + a_2 [b] = a_1 [b] + [a] b_2 \tag{3.29}$$

which implies

$$[ab] = 0 \text{ if } [a] = 0 \wedge [b] = 0. \tag{3.30}$$

By combining (3.14), (3.18) and (3.24–3.28), we obtain

$$\boxed{
\begin{aligned}
\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \int_{\Omega(u)} - \nabla y^T (\nabla V + \nabla V^T) \nabla p - p V^T \nabla f \\
&\quad + \operatorname{div}(V) \left(\frac{1}{2} (y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) dx
\end{aligned}
} \tag{3.31}$$

i.e., the shape derivative of \mathcal{L} expressed as domain integral which is equal to the one of J due to the theorem of Correa and Seeger. Now, we convert this domain integral into a boundary integral as mentioned above. Integration by parts in (3.31) yields

$$\begin{aligned}
& \int_{\Omega(u)} \operatorname{div}(V) \left(\frac{1}{2}(y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) dx \\
&= - \int_{\Omega(u)} V^T ((y - \bar{y}) \nabla y + \nabla (\nabla y^T \nabla p) - \nabla(fp)) dx \\
&\quad + \int_u \left\| \left(\frac{1}{2}(y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right\| ds \\
&\quad + \int_{\partial\Omega(u)} \left(\frac{1}{2}(y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle ds.
\end{aligned} \tag{3.32}$$

Since the outer boundary $\partial\Omega$ is not variable, we can choose the deformation vector field V equals zero in small neighbourhoods of $\partial\Omega(u)$. Therefore, the outer integral in (3.32) disappears. Combining (3.31), (3.32) and the vector calculus identity

$$\nabla y^T (\nabla V + \nabla V^T) \nabla p + V^T \nabla (\nabla y^T \nabla p) = \nabla p^T \nabla (V^T \nabla y) + \nabla y^T \nabla (V^T \nabla p)$$

which was proved in [7] gives

$$\begin{aligned}
\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \int_{\Omega(u)} -\nabla p^T \nabla (V^T \nabla y) - \nabla y^T \nabla (V^T \nabla p) \\
&\quad - (y - \bar{y}) V^T \nabla y + f V^T \nabla p dx \\
&\quad + \int_u \left\| \left(\frac{1}{2}(y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right\| ds.
\end{aligned} \tag{3.33}$$

Then, applying integration by parts in (3.33) we get

$$\begin{aligned}
& \int_{\Omega(u)} \nabla y^T \nabla (V^T \nabla p) dx \\
&= - \int_{\Omega(u)} \Delta y V^T \nabla p dx + \int_u \left\| \frac{\partial y}{\partial n} V^T \nabla p \right\| ds + \int_{\partial\Omega(u)} \frac{\partial y}{\partial n} V^T \nabla p ds
\end{aligned} \tag{3.34}$$

and analogously

$$\begin{aligned}
& \int_{\Omega(u)} \nabla p^T \nabla (V^T \nabla y) dx \\
&= - \int_{\Omega(u)} \Delta p V^T \nabla y dx + \int_u \left\| \frac{\partial p}{\partial n} V^T \nabla y \right\| ds + \int_{\partial\Omega(u)} \frac{\partial p}{\partial n} V^T \nabla y ds.
\end{aligned} \tag{3.35}$$

Like in (3.32) the outer integral in (3.34) as well as in (3.35) vanishes due to the fixed outer boundary $\partial\Omega(u)$. Thus, it follows that

$$\begin{aligned}
\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] &= \int_{\Omega(u)} V^T \nabla p (\Delta y + f) + V^T \nabla y (\Delta p - (y - \bar{y})) dx \\
&\quad + \int_u \left\| \left(\frac{1}{2}(y - \bar{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right\| \\
&\quad - \left\| \frac{\partial y}{\partial n} V^T \nabla p \right\| - \left\| \frac{\partial p}{\partial n} V^T \nabla y \right\| ds.
\end{aligned} \tag{3.36}$$

The domain integral in (3.36) vanishes due to (3.14) and (3.18). Moreover, the term $\llbracket \frac{1}{2}(y - \bar{y})^2 \langle V, n \rangle \rrbracket$ disappears because of (3.5) and the term $\llbracket \nabla y^T \nabla p \langle V, n \rangle \rrbracket$ because of the continuity of ∇y and ∇p . That

$$\left\llbracket \frac{\partial y}{\partial n} V^T \nabla p \right\rrbracket = \left\llbracket \frac{\partial p}{\partial n} V^T \nabla y \right\rrbracket = \langle V, n \rangle \left\llbracket \frac{\partial y}{\partial n} \frac{\partial p}{\partial n} \right\rrbracket = 0 \quad (3.37)$$

follows from (3.5), (3.17) and (3.30). Thus, we obtain the shape derivative of \mathcal{L} expressed as interface integral:

$$\boxed{\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] = - \int_u \llbracket f \rrbracket p \langle V, n \rangle ds} \quad (3.38)$$

Now, we consider the objective J in (3.1) with perimeter regularization. Combining (3.38) with proposition 5.1 in [23] we get

$$\boxed{\frac{\partial}{\partial u} \mathcal{L}(y, u, p)[V] = \int_u (-\llbracket f \rrbracket p + \mu \kappa) \langle V, n \rangle ds} \quad (3.39)$$

where κ denotes the curvature corresponding to the normal n .

REMARK 5. Note that (3.38) is equal to $\frac{\partial}{\partial u} J(y, u)[V]$ without perimeter regularization and (3.39) is equal to $\frac{\partial}{\partial u} J(y, u)[V]$ with perimeter regularization due to the theorem of Correa and Seeger as mentioned above.

We focus now on the weak formulation (2.11-2.13) and observe the following for the right hand sides in the case of (3.1-3.3):

$$-a_u(\bar{z}, p) - \frac{\partial}{\partial y} J(y, u) \bar{z} = - \int_{\Omega(u)} \nabla \bar{z}^T \nabla p + (y - \bar{y}) \bar{z} dx \quad (3.40)$$

$$-\frac{\partial}{\partial u} [J(y, u) + a_u(y, p) - b_u(p)] \bar{w} = \int_u (\llbracket f \rrbracket p - \mu \kappa) \langle \bar{w}, n \rangle ds \quad (3.41)$$

$$-a_u(y, \bar{q}) + b_u(\bar{q}) = \int_{\Omega(u)} -\nabla y^T \nabla \bar{q} + f \bar{q} dx \quad (3.42)$$

These expressions are set to zero, in order to define the necessary conditions of optimality.

Now, we discuss more details about the Hessian operators in the left hand sides of (2.11-2.13). We first consider them without the term H_{22} which requires special care. These are at the solution $(y, u, p) \in F$ of the optimization problem (3.1-3.3) for all $h := (\bar{z}, \bar{w}, \bar{q})^T \in T_{(y, u, p)} F$ as follows:

$$\begin{aligned} H_{11}(z, \bar{z}) &= \frac{\partial^2}{\partial y^2} J(y, u) z \bar{z} = \int_{\Omega(u)} z \bar{z} dx \\ H_{12}(w, \bar{z}) &= \frac{\partial}{\partial u} \left[a_u(\bar{z}, p) + \frac{\partial}{\partial y} J(y, u) \bar{z} \right] w = 0 \\ H_{13}(q, \bar{z}) &= a_u(\bar{z}, q) = \int_{\Omega(u)} \nabla \bar{z}^T \nabla q dx \\ H_{21}(z, \bar{w}) &= \frac{\partial}{\partial y} \frac{\partial}{\partial u} ([J(y, u) + a_u(y, p)] \bar{w}) z = 0 \\ H_{23}(q, \bar{w}) &= \frac{\partial}{\partial u} [a_u(y, q) - b_u(q)] \bar{w} = - \int_u \llbracket f \rrbracket q \langle \bar{w}, n \rangle ds \\ H_{31}(z, \bar{q}) &= a_u(z, \bar{q}) = \int_{\Omega(u)} \nabla z^T \nabla \bar{q} dx \\ H_{32}(w, \bar{q}) &= \frac{\partial}{\partial u} [a_u(y, \bar{q}) - b_u(\bar{q})] w = - \int_u \llbracket f \rrbracket \bar{q} \langle w, n \rangle ds \\ H_{33}(q, \bar{q}) &= 0 \end{aligned}$$

We compute now the term H_{22} . It will be evaluated at the solution of the optimization problem which means that it consists only of the second shape derivative. In section 4 this solution will be a straight line connection of the points $(0.5, 0)$ and $(0.5, 1)$, i.e., the curvature is equal to zero. Combining proposition 5.1 in [23] with the following rule for differentiating boundary integrals

$$\frac{d^+}{dt} \left(\int_{\Gamma_t} \eta(t) \right) \Big|_{t=0} = \int_{\Gamma} \left(D\eta[V] + \left(\frac{\partial \eta}{\partial n} + \eta \kappa \right) \langle V, n \rangle \right) \quad (3.43)$$

which was proved in [16] yields

$$\begin{aligned} H_{22}(w, \bar{w}) &= G(\text{Hess}^{\mathcal{N}}(J(y, u) + a_u(y, p) - b_u(p)) w, \bar{w}) \\ &= \int_u -D(\llbracket f \rrbracket p)[\bar{w}] \langle w, n \rangle - \llbracket f \rrbracket \left(\kappa p + \frac{\partial p}{\partial n} \right) \langle \bar{w}, n \rangle \langle w, n \rangle \\ &\quad + \mu \frac{\partial w}{\partial \tau} \frac{\partial \bar{w}}{\partial \tau} \langle \bar{w}, n \rangle \langle w, n \rangle ds \end{aligned} \quad (3.44)$$

where $\partial/\partial\tau$ denotes the derivative tangential to u . We have to evaluate the shape derivative $D(\llbracket f \rrbracket p)[\bar{w}]$ in (3.44). We observe in our special case

$$p = 0 \quad \text{on } u \quad (3.45)$$

because of the necessary optimality condition (3.41). Thus, it holds that

$$Dp[\bar{w}] = -\bar{w}^T \nabla p = -\bar{w}^T \frac{\partial p}{\partial n} n \quad \text{on } u \quad (3.46)$$

due to (3.25). Applying the product rule for shape derivatives yields

$$\begin{aligned} D(\llbracket f \rrbracket p)[\bar{w}] &= \llbracket Df[\bar{w}] p \rrbracket + \llbracket f Dp[\bar{w}] \rrbracket \stackrel{(3.5)}{=} \llbracket Df[\bar{w}] \rrbracket p + \llbracket f \rrbracket Dp[\bar{w}] \\ &\stackrel{(3.45)}{=} \stackrel{(3.46)}{=} -\llbracket f \rrbracket \frac{\partial p}{\partial n} \langle \bar{w}, n \rangle \quad \text{on } u. \end{aligned} \quad (3.47)$$

Thus, the Hessian operator H_{22} reduces to

$$\hat{H}_{22}(w, \bar{w}) = \int_u \left(\mu \frac{\partial w}{\partial \tau} \frac{\partial \bar{w}}{\partial \tau} - \llbracket f \rrbracket \kappa p \right) \langle w, n \rangle \langle \bar{w}, n \rangle ds. \quad (3.48)$$

By using the expressions above, we can formulate the QP (2.14, 2.15) at the solution in the following form:

$$\min_{(z, w)} F(z, w, y, p) \quad (3.49)$$

$$\begin{aligned} \text{s.t. } & \int_{\Omega(u)} \nabla z^T \nabla \bar{q} dx - \int_u \llbracket f \rrbracket \bar{q} w ds \\ &= - \int_{\Omega(u)} \nabla y^T \nabla \bar{q} dx + \int_{\Omega(u)} f \bar{q} dx, \quad \forall \bar{q} \in H_0^1(\Omega(u)) \end{aligned} \quad (3.50)$$

where the objective function F is given by

$$\begin{aligned} F(z, w, y, p) &= \int_{\Omega(u)} \frac{z^2}{2} + (y - \bar{y})z dx + \int_u \mu \kappa w - \llbracket f \rrbracket p w ds \\ &\quad + \frac{1}{2} \int_u \mu \left(\frac{\partial w}{\partial \tau} \right)^2 - \llbracket f \rrbracket \kappa p w^2 ds. \end{aligned} \quad (3.51)$$

This QP in weak formulation can be rewritten in the more intelligible strong form of an optimal control problem:

$$\min_{(z,w)} F(z, w, y, p) \quad (3.52)$$

$$\text{s.t. } -\Delta z = \Delta y + f_1 \quad \text{in } \Omega_1(u) \quad (3.53)$$

$$-\Delta z = \Delta y + f_2 \quad \text{in } \Omega_2(u) \quad (3.54)$$

$$\frac{\partial z}{\partial n} = f_1 w \quad \text{on } u \quad (3.55)$$

$$-\frac{\partial z}{\partial n} = f_2 w \quad \text{on } u \quad (3.56)$$

$$z = 0 \quad \text{on } \partial\Omega(u) \quad (3.57)$$

The adjoint problem to this optimal control problem is the boundary value problem:

$$-\Delta q = -z - (y - \bar{y}) \quad \text{in } \Omega(u) \quad (3.58)$$

$$q = 0 \quad \text{on } \partial\Omega(u) \quad (3.59)$$

The resulting design equation for the optimal control problem (3.52–3.57) is

$$0 = -\llbracket f \rrbracket (p + \kappa p w + q) + \mu \kappa - \mu \frac{\partial^2 w}{\partial \tau^2} \quad \text{on } u. \quad (3.60)$$

4. Numerical Results. In this section, we use the QP (3.49, 3.50) away from the optimal solution as a means to determine the step in the shape normal direction and thus create an iterative solution technique very similar to SQP techniques known from linear spaces. We solve the optimal control problem (3.52–3.57) by employing a CG-iteration for the reduced problem (3.60). I.e., we iterate over the variable w and each time the CG-iteration needs a residual of equation (3.60) from w^k , we compute the state variable z^k from (3.53–3.57) and then the adjoint variable q^k from (3.58, 3.59), which enables the evaluation of the residual

$$r^k := -\llbracket f \rrbracket (p + \kappa p w^k + q^k) + \mu \kappa - \mu \frac{\partial^2 w^k}{\partial \tau^2}. \quad (4.1)$$

The particular values for the parameters are chosen as $f_1 = 1000$, $f_2 = 1$ and $\mu = 10$. The data \bar{y} are generated from a solution of the state equation (3.2, 3.2) with u being the straight line connection of the points (0.5, 0) and (0.5, 1). The starting point of our iterations is described by a B-spline defined by the two control points (0.6, 0.7) and (0.4, 0.3). We build a coarse unstructured tetrahedral grid Ω_h^1 with roughly 6000 triangles as shown in the leftmost picture of figure 4. We also perform computations on uniformly refined grids Ω_h^2, Ω_h^3 with roughly 24000 and 98000 triangles. In figure 4 are also shown the next two iterations on the coarsest grid, where table 4 gives the distances of each shape to the solution approximated by

$$\text{dist}(u^k, u^*) := \int_{u^*} \left| \langle u^k, e_1 \rangle - \frac{1}{2} \right| ds$$

where u^* denotes the solution shape and $e_1 = (1, 0)$ is the first unit vector. Similar to [27], the retraction chosen for the shape is just the addition of the $q^k n_1$ to the current shape. In each iteration, the volume mesh is deformed according to the elasticity equation. Table [27] demonstrates that indeed quadratic convergence can

be observed on the finest mesh, but also that the mesh resolution has a strong influence on the convergence properties revealed.

The major advantage of the Newton method over a standard shape calculus steepest method based on the (reduced) shape derivative

$$dJ(y, u)[V] = - \int_u (\llbracket f \rrbracket p - \mu \kappa) \langle V, n \rangle ds$$

is the natural scaling of the step, which is just 1 near to the solution. When first experimenting with a steepest descent method, we found by trial and error, that one needs a scaling around 10 000 in order to obtain sufficient progress.

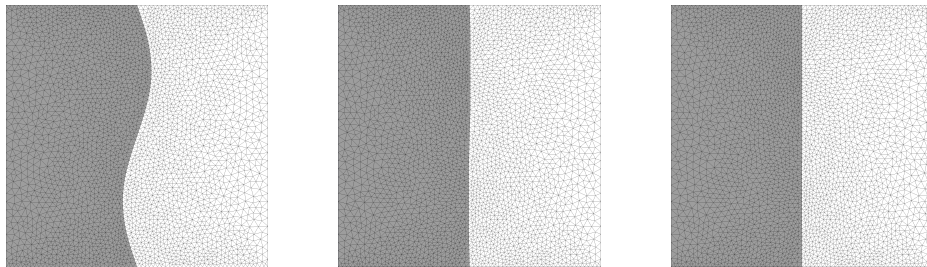


FIG. 4.1. Iterations 0, 1 and 2 (left to right) together with deformations of mesh Ω_h^1

TABLE 4.1

Performance of shape Lagrange–Newton algorithms: distances $\text{dist}(u^k, u^*)$ from the optimal solution on meshes with varying refinement. Quadratic convergence on the finest grid can be observed.

It.–No.	Ω_h^1	Ω_h^2	Ω_h^3
0	0.0705945	0.070637	0.0706476
1	0.0043115	0.004104	0.0040465
2	0.0003941	0.000104	0.0000645

5. Conclusions. This paper presents a generalization of the Riemannian shape calculus framework in [27] to Lagrange–Newton approaches for PDE constrained shape optimization problems. It is based on the idea that Riemannian shape Hessians do not differ from classical shape Hessians in the solution of a shape optimization problem and that Newton methods still converge locally quadratically, if Hessian terms are neglected which are zero at the solution anyway. It is shown that this approach is viable and leads to computational methods with superior convergence properties, when compared to only linearly converging standard steepest descent methods. Nevertheless, several issues have to be addressed in future investigations, like:

- More refined retractions have to be developed for large shape deformations.
- As observed during the computations, the shape deformation sometimes leads to shapes, where normal vectors can no longer be reliably evaluated. Provisions for those cases have to be developed.
- Full Lagrange–Newton methods may turn out being not very computationally efficient. However, this paper lays the foundation for the construction of appropriate preconditioners for the reduced optimization problem in many practical cases.
- The Riemannian shape space properties including quadratic convergence of the Lagrange–Newton approach seem to materialize only on very fine grids.

A logical next development is then to use locally adapted meshes near the shape front to be optimized.

Acknowledgment. This research has been partly funded by the DFG within the collaborative project EXASOLVERS as part of the DFG priority program SPP 1648 SPPEXA. Furthermore, the authors are very grateful to an unknown referee for insightful comments, which helped significantly to improve the paper from an earlier version.

REFERENCES

- [1] H. B. Ameur, M. Burger, and B. Hackl. Level set methods for geometric inverse problems in linear elasticity. *Inverse Problems*, 20:673–696, 2004.
- [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] 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.
- [8] A. Borzi and V. H. Schulz. *Computational optimization of systems governed by partial differential equations*. Number 08 in SIAM book series on Computational Science and Engineering. SIAM Philadelphia, 2012.
- [9] J. Céa. Conception optimale ou identification de formes calcul rapide de la dérivée directionnelle de la fonction coût. *RAIRO Modelisation mathématique et analyse numérique*, 20(3):371–402, 1986.
- [10] L. Conlon. *Differentiable Manifolds*. Birkhäuser advanced texts: Basler Lehrbücher. Birkhäuser, 2001.
- [11] R. Correa and A. Seeger. Directional derivative of a minmax function. *Nonlinear Anal.*, 9(1):13–22, 1985.
- [12] M. C. Delfour and J.–P. Zolésio. *Shapes and Geometries: Analysis, Differential Calculus, and Optimization*. Advances in Design and Control. SIAM Philadelphia, 2001.
- [13] K. Eppler, S. Schmidt, V. H. 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.
- [14] N. Gauger, C. Ilic, S. Schmidt, and V. H. Schulz. Non-parametric aerodynamic shape optimization. In G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. H. 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.
- [15] J. Haslinger and M. A. E. Mäkinen. *Introduction to Shape Optimization: Theory, Approximation, and Computation*. Advances in Design and Control. SIAM Philadelphia, 2008.
- [16] E. J. Haug, K. K. Choi, and V. Komkov. *Design Sensitivity Analysis of Structural Systems*. Academic Press, Orlando, FL, 1986.
- [17] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Advances in Design and Control. SIAM, 2008.
- [18] S. Lang. *Fundamentals of Differential Geometry*. Springer Verlag, 2001.
- [19] P. W. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23:74–113, 2007.
- [20] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8:1–48, 2006.
- [21] P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Math.*, 10:217–245, 2005.
- [22] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids*. Numerical Mathematics and Scientific Computation. Clarendon Press Oxford, 2001.

- [23] A. Novruzi and M. Pierre. Structure of shape derivatives. *Journal of Evolution Equations*, 2:365–382, 2002.
- [24] C. Schillings, S. Schmidt, and V. H. Schulz. Efficient shape optimization for certain and uncertain aerodynamic design. *Computers and Fluids*, 46(1):78–87, 2011.
- [25] S. Schmidt and V. H. Schulz. Impulse response approximations of discrete shape Hessians with application in CFD. *SIAM Journal on Control and Optimization*, 48(4):2562–2580, 2009.
- [26] S. Schmidt and V. H. Schulz. Shape derivatives for general objective functions and the incompressible Navier–Stokes equations. *Control and Cybernetics*, 39(3):677–713, 2010.
- [27] V. H. Schulz. A Riemannian view on shape optimization. *Foundations of Computational Mathematics*, 14:483–501, 2014.
- [28] V. H. Schulz, M. Siebenborn, and K. Welker. *Structured inverse modeling in parabolic diffusion problems*. submitted to SICON, arXiv:1409.3464, 2014.
- [29] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization: Shape Sensitivity Analysis*. Springer, 1992.