J Optim Theory Appl
DOI 10.1007/s10957-013-0520-4

On Computation of the Shape Hessian of the Cost Functional Without Shape Sensitivity of the State Variable

H. Kasumba · K. Kunisch

Received: 15 May 2012 / Accepted: 30 December 2013

Abstract A framework for calculating the shape Hessian for the domain optimization problem, with a partial differential equation as the constraint, is presented. First and second order approximations of the cost with respect to geometry perturbations are arranged in an efficient manner that allows the computation of the shape derivative and Hessian of the cost without the necessity to involve the shape derivative of the state variable. In doing so, the state and adjoint variables are only required to be Hölder continuous with respect to geometry perturbations.

Keywords Shape optimization · Shape derivative · Shape Hessian · Cost functional

Mathematics Subject Classification (2000) 49J20 · 49Q10 · 35Q93 · 49Q12

1 Introduction

Many important questions arise in the study of shape optimization problems. For instance, the existence and stability of optimal domains [1], the analysis of convergence of fixed point methods for free boundary problems [2], and speeding up of gradient type methods in shape optimization problems, where the topological struc-

Communicated by Günter Leugering.

H. Kasumba (✉)
Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences, Altenbergerstraße 69, 4040 Linz, Austria
e-mail: henry.kasumba@oeaw.ac.at

K. Kunisch
Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens
Universität Graz, Heinrichstraße 36, 8010 Graz, Austria
e-mail: karl.kunisch@uni-graz.at

Published online: 28 January 2014

ture of the shape changes during iteration [3]. All these questions are obviously linked to the second order information of the shape functional. Since the pioneering work of Fujii [4], the computation of second order shape derivatives has received a growing amount of attention, see, for instance, [1,2,5–7]. In some of these contributions, the characterization of the shape Hessian is given only in a formal manner. The approach taken by Fujii [4] and Simon [5] involves differentiation of the state equation with respect to the domain. The state variable varies in a Hilbert space which depends on the geometry with respect to which optimization is carried out. To obtain sensitivity information of the reduced cost functional, a chain rule approach involving the shape derivative of the state variable with respect to the domain is chosen. The rigorous analysis of this intermediate step is a non-trial task and as shown in Ito et al [8], there are cases where the assumptions of this paper are applicable, while shape differentiability of the state does not hold. Other techniques that bypass the computation of the derivative of the state with respect to the domain are presented by Delfour and Zolesio in [6] and [7]. In [6,7], they use function space parametrization and function space embedding methods, respectively, to characterize the shape Hessian of the cost functional. However, these techniques depend strongly on sophisticated differentiability properties of saddle point problems.

Concerning monographs, which are devoted to shape sensitivity analysis, we refer to [9–11], and recently [12]. In this paper, we present a computation of the shape Hessian of J under minimal regularity assumptions. The technique we employ was first suggested in [13] for computing the first order information, and allows to compute the shape derivative of the reduced cost functional without using the shape derivative of the state variable with respect to the geometry. The method and the associated computation, we present, are general and are applicable to a large class of boundary value problems. However, to make the exposition more transparent, we present the results on a simple example rather than give a general exhaustive theory. The remainder of this paper is organized as follows. Section 2 describes the setting of the optimization problem, useful notations, and definitions. The computation of the shape derivative via re-arrangement of the cost is given in Sect. 3. In Sect. 4, the computation of the shape Hessian via re-arrangement of the cost is presented and the conclusions of this work are drawn.

2 Problem Setting, Notations and Definitions

We describe an approach to compute the shape Hessian without recourse to the shape derivative of the state variable by means of the following shape optimization problem:

$$\min J(\Omega) = \frac{1}{2} \int_{\Omega} |y - y_d|^2 \, d\Omega \tag{1}$$

subject to

$$-\Delta y = f \text{ in } \Omega, \quad y = 0 \text{ on } \Gamma. \tag{2}$$

Here, Ω is a domain with boundary $\Gamma := \partial\Omega$ of class $C^{2,1}$. We shall denote by $H^m(\mathscr{S})$, $m \in \mathbb{R}$, the standard Sobolev space of order m defined by

$$H^m(\mathscr{S}) := \left\{ u \in L^2(\mathscr{S}) \mid D^\alpha u \in L^2(\mathscr{S}), \quad \text{for } 0 \le |\alpha| \le m \right\},$$

where D^α is the weak (or distributional) partial derivative, and α is a multi-index. Here, $\mathscr{S}$ is either the domain Ω, or its boundary Γ, or part of its boundary. The norm $||\cdot||_{H^m(\mathscr{S})}$ associated with $H^m(\mathscr{S})$ is given by

$$||u||^2_{H^m(\mathscr{S})} = \sum_{|\alpha| \le m} \int_{\mathscr{S}} |D^\alpha u|^2 \, \mathrm{d}\mathscr{S}.$$

Note that $H^0(\mathscr{S}) = L^2(\mathscr{S})$ and $||\cdot||_{H^0(\mathscr{S})} = ||\cdot||_{L^2(\mathscr{S})}$.

The data f are assumed to belong to the space $H^s(\mathscr{D})$, where $\mathscr{D}$ is a bounded hold all domain, and $s \ge 0$ will be specified later on. We shall say that y is the state and (2) is the state equation. The desired state y_d in (1) is assumed to be in $H^1(\mathscr{D})$. Our objective is to compute the first and second order derivatives of the cost J with respect to Ω without the necessity of involving the derivative of the state y with respect to Ω.

2.1 Shape Derivative

Shapes are difficult entities to be dealt with directly, so we manipulate them by means of transformations. If Ω is the initial admissible shape, and Ω_t is the shape at time t, one considers transformations $T_t : \Omega \mapsto \Omega_t$. Such transformations can be constructed, for instance, by perturbation of the identity [9]. To construct an admissible class of these transformations, let $\Omega \subset \bar{\mathscr{D}}$ be a bounded domain, and let

$$\mathscr{T}_{\text{ad}} = \{\mathbf{V} \in C^{2,1}(\bar{\mathscr{D}}) : \mathbf{V}|_{\partial\mathscr{D}} = 0\}$$

be the space of deformation fields. The fields $\mathbf{V} \in \mathscr{T}_{\text{ad}}$ define for $t > 0$, a perturbation of Ω by means of

$$\begin{aligned} T_t : \Omega &\mapsto \Omega_t(\mathbf{V}), \\ x &\mapsto T_t(x) = x + t\mathbf{V}(x). \end{aligned}$$

For each $\mathbf{V} \in \mathscr{T}_{\text{ad}}$, there exists $\tau > 0$ such that $T_t(\mathscr{D}) = \mathscr{D}$, and $\{T_t\}$ is a family of $C^{2,1}$-diffeomorphisms for $|t| < \tau$ [9]. For each $t \in \mathbb{R}$ with $|t| < \tau$, we set $\Omega_t = T_t(\Omega), \quad \Gamma_t = T_t(\Gamma)$. Thus, $\Omega_0 = \Omega,\ \Gamma_0 = \Gamma,\ \Omega_t \subset \mathscr{D}$.

Let $\mathscr{J} = [0, \tau]$ with τ sufficiently small. Then, the following regularity properties of the transformation T_t can be shown; see, for example, ([8,11], [9, Chap. 7]):

$$\begin{aligned}
&T_0 = id, && t \mapsto T_t \in C^1(\mathscr{J}, C^1(\bar{\mathscr{D}}; \mathbb{R}^d)), \\
&t \mapsto T_t^{-1} \in C(\mathscr{J}, C^1(\bar{\mathscr{D}}; \mathbb{R}^d)), && t \mapsto I_t \in C^1(\mathscr{J}, C(\bar{\mathscr{D}})), \\
&t \mapsto (DT_t)^{-T} \in C^1(\mathscr{J}, C(\bar{\mathscr{D}}; \mathbb{R}^{d\times d})), && \tfrac{d}{\mathrm{d}t} T_t|_{t=0} = \mathbf{V}, \\
&\tfrac{d}{\mathrm{d}t} T_t^{-1}|_{t=0} = -\mathbf{V}, && \tfrac{d}{\mathrm{d}t} DT_t|_{t=0} = D\mathbf{V}, \\
&\tfrac{d}{\mathrm{d}t} DT_t^{-1}|_{t=0} = -D\mathbf{V}, && \tfrac{d}{\mathrm{d}t} I_t|_{t=0} = \operatorname{div} \mathbf{V}, \\
&I_t|_{t=0} = 1, && I_t^{-1}|_{t=0} = 1, \\
&\tfrac{d}{\mathrm{d}t} A(t)|_{t=0} = \operatorname{div}\mathbf{V} - (D\mathbf{V} + (D\mathbf{V})^T),
\end{aligned} \tag{3}$$

where

$$I_t := \det DT_t, \quad A(t) := I_t(DT_t)^{-1}(DT_t)^{-T}, \tag{4}$$

and the limits defining the derivatives at $t = 0$ exist uniformly in $x \in \bar{\mathscr{D}}$.

Definition 2.1 For given $\mathbf{V} \in \mathscr{T}_{\text{ad}}$, the Eulerian derivative of J at Ω in the direction $\mathbf{V}$ is defined as

$$dJ(\Omega)\mathbf{V} := \lim_{t\downarrow 0^+} \frac{J(\Omega_t(\mathbf{V})) - J(\Omega)}{t}. \tag{5}$$

The functional J is said to be shape differentiable at Ω iff $dJ(\Omega)\mathbf{V}$ exists for all $\mathbf{V} \in \mathscr{T}_{\text{ad}}$ and the mapping $\mathbf{V} \mapsto dJ(\Omega)\mathbf{V}$ is linear and continuous on $\mathscr{T}_{\text{ad}}$. If J is shape differentiable, then, there exists a distribution $\mathscr{G}$ in $\mathscr{T}^*_{\text{ad}}$ such that

$$dJ(\Omega)\mathbf{V} = \langle \mathscr{G}, \mathbf{V}\rangle_{\mathscr{T}^*_{\text{ad}} \times \mathscr{T}_{\text{ad}}}. \tag{6}$$

The distribution $\mathscr{G}$, that is uniquely defined by (6), is called the shape gradient of J at Ω.

2.2 Shape Hessian

Let $\mathbf{V}$ and $\mathbf{W}$ be given vector fields in $\mathscr{T}_{\text{ad}}$. As in the previous subsection, we associate with $\mathbf{V}$ and $\mathbf{W}$ the transformed domains $\Omega_t(\mathbf{V})$ and $\Omega_t(\mathbf{W})$, respectively.

Definition 2.2 [9] Assume that $dJ(\Omega_t(\mathbf{W}))\mathbf{V}$ exists for all $t \in [0, \tau]$. Then, the functional J is said to have a second order Eulerian semi-derivative at Ω in directions $(\mathbf{V}, \mathbf{W})$ iff the following limit exists

$$d^2J(\Omega)(\mathbf{V}, \mathbf{W}) := \lim_{t\downarrow 0^+} \frac{dJ(\Omega_t(\mathbf{W}))\mathbf{V} - dJ(\Omega)\mathbf{V}}{t}. \tag{7}$$

The functional J is said to be twice shape differentiable at Ω iff, for all $\mathbf{V}, \mathbf{W} \in \mathscr{T}_{\text{ad}}$, the second Eulerian semi-derivative $d^2J(\Omega)(\mathbf{V}, \mathbf{W})$ exists and the mapping

$$(\mathbf{V}, \mathbf{W}) \mapsto d^2J(\Omega)(\mathbf{V}, \mathbf{W}) : \ \mathscr{T}_{\text{ad}} \times \mathscr{T}_{\text{ad}} \mapsto \mathbb{R} \tag{8}$$

is bilinear and continuous.

The distribution associated with the mapping in (8) is called the shape Hessian. It will be shown that the shape Hessian has its support on the boundary of Ω and that it is dependent on both the normal and the tangential components of $\mathbf{V}$ and $\mathbf{W}$ on the boundary. Furthermore, it will be shown that the shape Hessian can be decomposed into symmetric and non-symmetric parts, where the non-symmetric part involves the shape gradient multiplied by the derivative of the first vector field in the direction of the second vector field. Hence the shape Hessian is typically not symmetric.

3 Shape Derivative via Re-arrangement of the Cost

In this section, we compute the shape derivative of J in (1) by re-arranging the first perturbation of the cost with respect to the geometry. This result can be obtained using the general theory developed in Ito et al. [8]. However, we felt compelled to repeat some essential steps to give a basis for computing the second order information.

Using Definition 2.1, the first derivative can be expressed as

$$dJ(\Omega)\mathbf{V} = \lim_{t\downarrow 0^+} \frac{1}{2t} \left\{ \int\limits_{\Omega_t(\mathbf{V})} |y_t - y_d|^2 \, d\Omega_t(\mathbf{V}) - \int\limits_{\Omega} |y - y_d|^2 \, d\Omega \right\}, \tag{9}$$

where y_t satisfies

$$-\Delta y_t = f_t \text{ in } \Omega_t, \quad y_t = 0 \text{ on } \Gamma_t. \tag{10}$$

We assume that $f \in H^1(\mathscr{D})$ and $y_d \in H^1(\mathscr{D})$. The variational form of (10) is given by:

Find $y_t \in X_t := H_0^1(\Omega_t)$ such that

$$\langle \mathscr{E}(y_t^t, \Omega_t), \psi_t \rangle_{X_t^* \times X_t} := (\nabla y_t, \nabla \psi_t)_{\Omega_t} - (f_t, \psi_t)_{\Omega_t} = 0 \tag{11}$$

holds for all $\psi_t \in X_t$.

Proposition 3.1 *[14] There exists a unique solution y_t to (10). Moreover, the domain is assumed to be of class $C^{2,1}$, we have $y_t \in H^3(\Omega_t) \cap H_0^1(\Omega_t)$.*

Observe that at $t = 0$, $y_t|_{t=0} = y \in X$, $\Omega_t|_{t=0} = \Omega$, $X_t|_{t=0} = X$, and (11) becomes

$$\langle \mathscr{E}(y, \Omega), \psi \rangle_{X^* \times X} := (\nabla y, \nabla \psi)_\Omega - (f, \psi)_\Omega = 0, \tag{12}$$

which is the weak formulation of the state Eq. (2) with homogeneous Dirichlet boundary conditions. The functions y_t and y in (9) are defined on different domains. Therefore, to compute (9), one needs to transport y_t back to Ω. Any function $y_t : \Omega_t \mapsto \mathbb{R}^2$, can be mapped back to the reference domain by

$$y^t = y_t \circ T_t : \Omega \mapsto \mathbb{R}^2, \tag{13}$$

where $(y_t \circ T_t)(x) = y_t(T_t(x))$. Furthermore, the chain rule guarantees that the gradients of y_t and y^t are related by

$$(\nabla y_t) \circ T_t = B_t \nabla y^t, \tag{14}$$

(see [11] Prop. 2.29) where $B_t := (DT_t)^{-T}$. Consequently, the transformation of (11) back to Ω is obtained as follows: Find $y^t \in X$ such that

$$\langle \tilde{\mathscr{E}}(y^t, t), \psi \rangle_{X^* \times X} := (A(t)\nabla y^t, \nabla \psi)_\Omega - (f^t I_t, \psi)_\Omega = 0, \quad \text{for all} \quad \psi \in X. \tag{15}$$

It is shown in [15] that, for any $0 < \alpha < 1$, the following result

$$(A(t)\xi, \xi) \geq \frac{\alpha}{2}|\xi|^2,$$

holds for $(\xi, x) \in \mathbb{R}^2 \times \Omega$ and τ sufficiently small. Thus, the bilinear form in (15) is elliptic uniformly in $t \in \mathscr{J}$. The adjoint state $p \in H_0^1(\Omega)$, is defined as the solution to

$$\langle \mathscr{E}_y(y, \Omega)\psi, p\rangle_{X^* \times X} = (y - y_d, \psi)_\Omega, \quad \text{for all} \ \ \psi \in H_0^1(\Omega), \tag{16}$$

where

$$\langle \mathscr{E}_y(y, \Omega)\psi, p\rangle_{X^* \times X} := (\nabla\psi, \nabla p)_\Omega. \tag{17}$$

Integrating the term $(\nabla\psi, \nabla p)_\Omega$, on the right-hand side of (17) by parts, one obtains the strong form of the adjoint equation in (16), that we express as

$$-\Delta p = y - y_d \quad \text{in} \quad \Omega, \quad p = 0 \quad \text{on} \quad \Gamma. \tag{18}$$

The adjoint Eq. (18) possesses a unique solution $p \in H_0^1(\Omega)$. Moreover, the domain is assumed to be of class $C^{2,1}$, we have that $p \in H^3(\Omega) \cap H_0^1(\Omega)$.

The existence of the primal and adjoint states allows the formulation of first order optimality conditions for the optimization problem (1)–(2). The following lemmas shall be utilized.

Lemma 3.1 *[8]*

(1) Let $g \in C(\mathscr{J}, W^{1,1}(\mathscr{D}))$, and assume that $\frac{\partial g}{\partial t}(0)$ exists in $L^1(\mathscr{D})$. Then

$$\frac{d}{\mathrm{d}t}\int_{\Omega_t} g(t, x)\, d\Omega_t|_{t=0} = \int_{\Omega} \frac{\partial g}{\partial t}(0, x)\, d\Omega + \int_{\Gamma} g(0, x)\mathbf{V}\cdot\mathbf{n}\, d\Gamma.$$

(2) Let $g \in C(\mathscr{J}, W^{2,1}(\mathscr{D}))$, and assume that $\frac{\partial g}{\partial t}(0)$ exists in $W^{1,1}(\mathscr{D})$. Then

$$\frac{d}{\mathrm{d}t}\int_{\Gamma_t} g(t, x)\, d\Omega_t|_{t=0} = \int_{\Gamma} \frac{\partial g}{\partial t}(0, x)\, d\Gamma + \int_{\Gamma} \Big(\frac{\partial g(0, x)}{\partial \mathbf{n}} + \kappa g(0, x)\Big)\mathbf{V}\cdot\mathbf{n}\, d\Gamma,$$

where κ stands for the mean curvature of Γ.

The assumptions of Lemma 3.1 can be verified using the following result.

Lemma 3.2 *[11, Chap. 2]*

(1) If $y \in L^p(\mathscr{D})$, then $t \mapsto y \circ T_t^{-1} \in C(\mathscr{J}, L^p(\mathscr{D}))$, $1 \leq p < \infty$.
(2) If $y \in H^2(\mathscr{D})$, then $t \mapsto y \circ T_t^{-1} \in C(\mathscr{J}, H^2(\mathscr{D}))$.
(3) If $y \in H^2(\mathscr{D})$, then $\frac{d}{\mathrm{d}t}(y \circ T_t^{-1})|_{t=0}$ exists in $H^1(\mathscr{D})$ and is given by

$$\frac{d}{\mathrm{d}t}(y \circ T_t^{-1})|_{t=0} = -(Dy)\mathbf{V}.$$

Remark 3.1 As a consequence of Lemma 3.2, we note that $\frac{d}{dt}\nabla(y \circ T_t^{-1})\Big|_{t=0}$ exists in $L^2(\mathscr{D})$ and is given by

$$\frac{d}{dt}\nabla(y \circ T_t^{-1})\Big|_{t=0} = -\nabla(Dy\mathbf{V}).$$

For the transformation of domain integrals, the following well-known fact will be used repeatedly.

Lemma 3.3 *Let $\phi_t \in L^1(\Omega_t)$, then $\phi_t \circ T_t \in L^1(\Omega)$ and*

$$\int_{\Omega_t} \phi_t \, d\Omega_t = \int_{\Omega} \phi_t \circ T_t \; I_t \, d\Omega.$$

Lemma 3.4 *[13] For any $f \in L^p(\mathscr{D})$, $p \geq 1$, we have $\lim_{t\downarrow 0^+} f \circ T_t = f$ in $L^p(\mathscr{D})$.*

3.1 Preliminary Results

Lemma 3.5 *There exists a unique solution $y^t \in H_0^1(\Omega) \cap H^2(\Omega)$ to (15) for $t > 0$ sufficiently small. Moreover,*

$$\lim_{t\downarrow 0^+} \frac{||y^t - y||_{H^2(\Omega)}}{|t|^{\frac{1}{2}}} = 0 \tag{19}$$

holds, where y is the weak solution of (2).

Proof The existence and uniqueness of a solution to (15) are established in [9, p. 396]. Therefore, it suffices to show the second statement. Subtracting (15) from (12), one obtains

$$\begin{aligned}(A(t)\nabla(y^t - y), \nabla\psi)_\Omega &= -((A(t) - I)\nabla y, \nabla\psi)_\Omega + (f^t I_t - f, \psi)_\Omega \\ &= (\mathrm{div}((A(t) - I)\nabla y), \psi)_\Omega + (f^t I_t - f, \psi)_\Omega.\end{aligned}$$

We have that $y^t - y$ belongs to $H_0^1(\Omega)$. By standard elliptic regularity theory [16, p. 317], we obtain that $y^t - y \in H_0^1(\Omega) \cap H^2(\Omega)$ and

$$||\frac{1}{t}(y^t - y)||_{H^2(\Omega)} \leq C\wp(t), \tag{20}$$

where $\wp(t) := \left\{||\mathrm{div}(\frac{1}{t}(A(t) - I)\nabla y)||_{L^2(\Omega)} + ||\frac{1}{t}(f^t I_t - f)||_{L^2(\Omega)}\right\}$ and C is some appropriate constant which may be chosen independently of t. Following [15], we can show that the right-hand side in (20) is bounded uniformly in t. Consequently,

$$\lim_{t\downarrow 0^+} \frac{1}{\sqrt{t}}||y^t - y||_{H^2(\Omega)} \leq C \lim_{t\downarrow 0^+} \wp(t)\sqrt{t} = 0,$$

which implies (19). □

By linearity of $\mathscr{E}$ in y, we have

$$\langle \mathscr{E}(v,\Omega) - \mathscr{E}(y,\Omega) - \mathscr{E}_y(y,\Omega)(v-y), \psi\rangle_{X^*\times X} = 0, \text{ for all } v,\ \psi \in X. \tag{21}$$

Lemma 3.6 *The following result holds*

$$\lim_{t\downarrow 0^+} \frac{1}{t}\left\langle \tilde{\mathscr{E}}(y^t,t) - \tilde{\mathscr{E}}(y,t) - \big(\mathscr{E}(y^t,\Omega) - \mathscr{E}(y,\Omega)\big), \psi\right\rangle_{X^*\times X} = 0,$$

for every $\psi \in X$.

Proof Let

$$\mathscr{G}(t) := \left\langle \tilde{\mathscr{E}}(y^t,t) - \tilde{\mathscr{E}}(y,t) - \big(\mathscr{E}(y^t,\Omega) - \mathscr{E}(y,\Omega)\big), \psi\right\rangle_{X^*\times X}.$$

Then $\mathscr{G}(t) = \Big[(I_t B_t \nabla(y^t - y), B_t\psi)_\Omega - (\nabla(y^t-y),\psi)_\Omega\Big]$, and $\lim_{t\downarrow 0^+}\frac{1}{t}\mathscr{G}(t) = 0$ follows from Lemma 3.5 and the differentiability of the mappings $t \mapsto I_t,\ t\mapsto B_t$. □

Theorem 3.1 *Let y and p be the solutions to* (2) *and* (18)*, respectively. Then*

$$\frac{d}{dt}\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X}|_{t=0} = -\int_\Gamma \left[\frac{\partial y}{\partial \mathbf{n}}\frac{\partial p}{\partial \mathbf{n}}\right]\mathbf{V}\cdot\mathbf{n}\, d\Gamma - \int_\Omega (y-y_d)\nabla y \cdot \mathbf{V}\, d\Omega. \tag{22}$$

Proof The proof is a minor modification of a similar result in [8]. Since $\Omega \in C^{2,1}$, y, and p can be extended to functions in $H^2(\mathscr{D})$ which we again denote by the same symbol. Furthermore, observe that

$$\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X} := (A(t)\nabla y, \nabla p)_\Omega - (f^t I_t, p)_\Omega$$

can be mapped back to Ω_t to obtain

$$\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X} = (\nabla(y\circ T_t^{-1}), \nabla(p\circ T_t^{-1}))_{\Omega_t} - (f, (p\circ T_t^{-1}))_{\Omega_t}.$$

Thus, Lemma 3.1(1) and Lemma 3.2 imply that

$$\frac{d}{dt}\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X}|_{t=0} = \int_\Omega \Big(\nabla\dot{y}\cdot\nabla p + \nabla y\cdot\nabla\dot{p} - f\dot{p}\Big)\, d\Omega + \int_\Gamma (\nabla y\cdot\nabla p)\mathbf{V}\cdot\mathbf{n}\, d\Gamma,$$

where $\dot{y} := -\nabla y\cdot\mathbf{V}$, $\quad \dot{p} := -\nabla p\cdot\mathbf{V}$. Note that $\dot{y}$, as well as $\dot{p}$, do not belong to $H_0^1(\Omega)$, but they are elements of $H^1(\Omega)$. Applying Greens theorem implies

$$\begin{aligned}&\frac{d}{dt}\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X}|_{t=0}\\ &\quad = \int_\Omega (-\Delta y - f)\dot{p}\, d\Omega + \int_\Gamma \left(\frac{\partial y}{\partial \mathbf{n}}\dot{p} + \dot{y}\frac{\partial p}{\partial \mathbf{n}} + \nabla y\cdot\nabla p\, \mathbf{V}\cdot\mathbf{n}\right) d\Gamma + \int_\Omega (-\Delta p)\dot{y}\, d\Omega.\end{aligned}$$

Since $y, p \in \mathbf{H}_0^1(\Omega)$, we have

$$\int_\Gamma \left(\frac{\partial y}{\partial \mathbf{n}}\dot{p} + \dot{y}\frac{\partial p}{\partial \mathbf{n}} + \nabla y \cdot \nabla p \, \mathbf{V}\cdot\mathbf{n}\right) d\Gamma = -\int_\Gamma \left[\frac{\partial y}{\partial \mathbf{n}}\frac{\partial p}{\partial \mathbf{n}}\right] \mathbf{V}\cdot\mathbf{n}\, d\Gamma.$$

The strong forms of the state (2) and adjoint (18) imply

$$\frac{d}{\mathrm{d}t}\langle\tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X}|_{t=0} = -\int_\Gamma \left[\frac{\partial y}{\partial \mathbf{n}}\frac{\partial p}{\partial \mathbf{n}}\right] \mathbf{V}\cdot\mathbf{n}\, d\Gamma - (y - y_d, \nabla y \cdot \mathbf{V})_\Omega.$$

□

3.2 Shape Derivative

In this subsection, we establish the expression for the shape derivative for the cost functional J. Since $y - y_d \in L^2(\Omega)$, the cost functional $J(\Omega)$ is well defined. The associated adjoint state $p \in X$ is given as a solution to (18).

Theorem 3.2 *The shape derivative of $J(y,\Omega)$ exists and it is given by the expression*

$$dJ(\Omega)\mathbf{V} = \int_\Gamma \left[\frac{\partial y}{\partial \mathbf{n}}\frac{\partial p}{\partial \mathbf{n}} + \frac{1}{2}(y - y_d)^2\right] \mathbf{V}\cdot\mathbf{n}\, d\Gamma. \tag{23}$$

Proof The general result in [8] can be utilized to derive the expression in (23). However, we provide the proof here in a more elegant way than earlier in [8]. Let $\boldsymbol{\Delta}_1 := J(y_t, \Omega_t) - J(y,\Omega)$. Then

$$\boldsymbol{\Delta}_1 = \int_\Omega \left(I_t \frac{1}{2}(y^t - y_d)^2 - \frac{1}{2}(y - y_d)^2\right) d\Omega.$$

We can express $\boldsymbol{\Delta}_1$ as $\boldsymbol{\Delta}_1 = \boldsymbol{\Delta}_{1,1}(t) + \boldsymbol{\Delta}_{1,2}(t) + \boldsymbol{\Delta}_{1,3}(t) + \boldsymbol{\Delta}_{1,4}(t)$, where

$$\begin{aligned}
\boldsymbol{\Delta}_{1,1}(t) &= \int_\Omega I_t\left[\frac{1}{2}(y^t - y_d)^2 - \frac{1}{2}(y - y_d)^2 - (y - y_d, y^t - y)\right] d\Omega,\\
\boldsymbol{\Delta}_{1,2}(t) &= \int_\Omega (I_t - 1)(y - y_d, y^t - y)\, d\Omega, \quad \boldsymbol{\Delta}_{1,3}(t) = \int_\Omega (y - y_d, y^t - y)\, d\Omega,\\
\boldsymbol{\Delta}_{1,4}(t) &= \int_\Omega (I_t - 1)\frac{1}{2}(y - y_d)^2\, d\Omega.
\end{aligned} \tag{24}$$

From (24) and the embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$, it follows that

$$|\boldsymbol{\Delta}_{1,1}(t)| \le K||y^t - y||^2_{H^1(\Omega)},$$

where $K > 0$ does not depend on $t \in \mathscr{J}$. Using Lemma 3.5, we have that

$$\lim_{t\downarrow 0^+} \frac{1}{t}\boldsymbol{\Delta}_{1,1}(t) = 0.$$

Next, observe that

$$\left|\frac{\boldsymbol{\Delta}_{1,2}(t)}{t}\right| \leq K \left\|\frac{(I_t - I)}{t}\right\|_{L^\infty} ||y - y_d||_{H^1(\Omega)}||y^t - y||_{H^1(\Omega)}.$$

Therefore, by Lemma 3.5 and (3), one obtains $\lim\limits_{t\downarrow 0^+} \left|\frac{1}{t}\boldsymbol{\Delta}_{1,2}(t)\right| = 0$. Using the adjoint Eq. (16) with $\psi = y^t - y \in X$, we have that

$$\begin{aligned}
\boldsymbol{\Delta}_{1,3}(t) &= \langle \mathscr{E}_y(y,\Omega)(y^t - y), p\rangle_{X^*\times X}\\
&= -\langle \tilde{\mathscr{E}}(y,t) - \tilde{\mathscr{E}}(y,0), p\rangle_{X^*\times X}\\
&\quad -\langle \mathscr{E}(y^t,\Omega) - \mathscr{E}(y,\Omega) - \mathscr{E}_y(y,\Omega)(y^t - y), p\rangle_{X^*\times X}\\
&\quad -\langle \tilde{\mathscr{E}}(y^t,t) - \tilde{\mathscr{E}}(y,t) - \mathscr{E}(y^t,\Omega) + \mathscr{E}(y,\Omega), p\rangle_{X^*\times X}.
\end{aligned}$$

By Lemma 3.5, (21), and Lemma 3.6, we find

$$\begin{aligned}
\lim_{t\downarrow 0^+} \frac{\boldsymbol{\Delta}_{1,3}(t)}{t} &= -\frac{d}{\mathrm{d}t}\langle \tilde{\mathscr{E}}(y,t), p\rangle_{X^*\times X}|_{t=0},\\
&= \int_\Gamma \frac{\partial y}{\partial \mathbf{n}}\frac{\partial p}{\partial \mathbf{n}}\mathbf{V}\cdot\mathbf{n}\, d\Gamma + \int_\Omega (y - y_d)\nabla y\cdot\mathbf{V}\, d\Omega,
\end{aligned} \tag{25}$$

where we use (22). Since $y \in H^2(\Omega)$, it follows that $\frac{1}{2}(y - y_d)^2 \in W^{1,1}(\Omega)$. This implies that $\frac{d}{\mathrm{d}t}\left[\frac{1}{2}(y\circ T_t^{-1} - y_d)^2\right]_{t=0}$ exists in $L^1(\Omega)$, [11, p. 65]. Hence, $\boldsymbol{\Delta}_{1,4}(t) = \int_{\Omega_t}\frac{1}{2}(y\circ T_t^{-1} - y_d)^2\, d\Omega_t - \int_\Omega \frac{1}{2}(y - y_d)^2\, d\Omega$, we have by Lemma 3.1(1), that

$$\begin{aligned}
\lim_{t\downarrow 0^+} \frac{\boldsymbol{\Delta}_{1,4}(t)}{t} &= \frac{d}{\mathrm{d}t}\int_{\Omega_t} \frac{1}{2}(y\circ T_t^{-1} - y_d)^2\Big|_{t=0} d\Omega_t,\\
&= \int_\Omega \frac{d}{\mathrm{d}t}\left[\frac{1}{2}(y\circ T_t^{-1} - y_d)^2\right]_{t=0} d\Omega + \int_\Gamma \frac{1}{2}(y - y_d)^2\mathbf{V}\cdot\mathbf{n}\, d\Gamma,\\
&= -\int_\Omega (y - y_d)\nabla y\cdot\mathbf{V}\, d\Omega + \int_\Gamma \frac{1}{2}(y - y_d)^2\mathbf{V}\cdot\mathbf{n}\, d\Gamma.
\end{aligned} \tag{26}$$

Hence, (25) and (26) yield the desired expression for the shape derivative. □

4 Shape Hessian via Re-arrangement of the Cost

In this section, we compute the shape Hessian of J in (1) by re-arranging the second order perturbation of the cost with respect to the geometry. Using the divergence theorem, we can express the first derivative derived in the previous section as follows

$$dJ(\Omega)\mathbf{V} = \int_{\Omega} \operatorname{div}\left(\left[\frac{1}{2}(y_d)^2 + \nabla y \cdot \nabla p\right]\mathbf{V}\right) d\Omega.$$

Using Definition 2.2, the second order Eulerian semi-derivative of J at Ω in direction $(\mathbf{V}, \mathbf{W})$ can be expressed as

$$d^2 J(\Omega)(\mathbf{V}, \mathbf{W}) = \lim_{t \downarrow 0^+} \frac{dJ(\Omega_t(\mathbf{W}))\mathbf{V} - dJ(\Omega)\mathbf{V}}{t},$$

where

$$dJ(\Omega_t(\mathbf{W}))\mathbf{V} = \int_{\Omega_t(\mathbf{W})} \operatorname{div}\left(\left[\frac{1}{2}y_d^2 + \nabla y_t \cdot \nabla p_t\right]\mathbf{V}\right) d\Omega_t(\mathbf{W}), \tag{27}$$

and y_t, p_t satisfy

$$-\Delta y_t = f_t \quad \text{in} \quad \Omega_t(\mathbf{W}), \quad y_t = 0 \quad \text{on} \quad \Gamma_t(\mathbf{W}), \tag{28}$$

$$-\Delta p_t = y_t - y_d \quad \text{in} \quad \Omega_t(\mathbf{W}), \quad p_t = 0 \quad \text{on} \quad \Gamma_t(\mathbf{W}). \tag{29}$$

The weak form of (29) is given by
Find $p_t \in H_0^1(\Omega_t(\mathbf{W}))$ such that

$$(\nabla p_t, \nabla \psi_t)_{\Omega_t(\mathbf{W})} = (y_t - y_d, \psi_t)_{\Omega_t(\mathbf{W})}, \quad \text{for all} \;\; \psi_t \in H_0^1(\Omega_t(\mathbf{W})). \tag{30}$$

The transformation of (30) back to Ω leads to the problem
Find $p^t \in X$ such that

$$(A(t)\nabla p^t, \nabla \psi)_\Omega - ((y^t - y_d)I_t, \psi)_\Omega = 0, \quad \text{for all} \quad \psi \in X. \tag{31}$$

4.1 Preliminary Results

Lemma 4.1 *There exists a unique solution $p^t \in H_0^1(\Omega) \cap H^2(\Omega)$ to (31), for $t > 0$ sufficiently small. Moreover,*

$$\lim_{t \downarrow 0^+} \frac{||p^t - p||_{H^2(\Omega)}}{|t|^{\frac{1}{2}}} = 0 \tag{32}$$

holds, where p is the weak solution of (18).

Proof The existence and uniqueness of a solution to (31) can be established in an analogous way as in Lemma 3.5. Subtracting (31) from (30) at $t=0$, one obtains for $\psi \in H_0^1(\Omega)$

$$\begin{aligned}(A(t)\nabla(p^t-p),\nabla\psi)_\Omega &= -((A(t)-I)\nabla p,\nabla\psi)_\Omega+((y^t-y_d)I_t-(y-y_d),\psi)_\Omega\\ &= (\operatorname{div}((A(t)-I)\nabla p),\psi)_\Omega+((y^t-y_d)I_t-(y-y_d),\psi)_\Omega.\end{aligned}$$

We have that p^t-p belongs to $H_0^1(\Omega)$. By standard elliptic regularity theory [16, p. 317], we obtain that $p^t-p\in H_0^1(\Omega)\cap H^2(\Omega)$ and

$$\left\|\frac{1}{t}(p^t-p)\right\|_{H^2(\Omega)}\le C\wp(t),$$

where $\wp(t):=\left\{||\operatorname{div}(\frac{1}{t}(A(t)-I)\nabla p)||_{L^2(\Omega)}+||\frac{1}{t}((y^t-y_d)I_t-(y-y_d)||_{L^2(\Omega)}\right\}$ and C is some appropriate constant which may be chosen independently of t. Following [15], we can show that the term $||\operatorname{div}(\frac{1}{t}(A(t)-I)\nabla p)||_{L^2(\Omega)}$ is bounded uniformly in t. Furthermore, Lemma 3.5 also suggests that $||\frac{1}{t}((y^t-y_d)I_t-(y-y_d)||_{L^2(\Omega)}$ is bounded uniformly in t. Consequently,

$$\lim_{t\downarrow 0^+}\frac{1}{\sqrt{t}}||p^t-p||_{H^2(\Omega)}\le C\lim_{t\downarrow 0^+}\wp(t)\sqrt{t}=0,$$

which implies (32). □

For the computation of the second order derivative, the presence of y and p as states suggests the introduction of two adjoint states $(\hat{\Sigma},\hat{P})\in H^2(\Omega)\times H^2(\Omega)$. Following [7], we introduce $(\hat{\Sigma},\hat{P})$ as solutions to

$$\begin{aligned}&\int_\Omega \operatorname{div}\Big([\nabla y\cdot\nabla\tilde{p}]\mathbf{V}\Big)+\Delta\tilde{p}\hat{\Sigma}+\tilde{p}\Delta\hat{\Sigma}+\nabla\tilde{p}\cdot\nabla\hat{\Sigma}\,d\Omega=0,\quad \text{for all}\quad \tilde{p}\in H^2(\Omega),\\ &\int_\Omega \operatorname{div}\Big([\nabla\tilde{y}\cdot\nabla p]\mathbf{V}\Big)+\Delta\tilde{y}\hat{P}+\tilde{y}\Delta\hat{P}+\nabla\tilde{y}\cdot\nabla\hat{P}+\tilde{y}\hat{\Sigma}\,d\Omega=0,\quad \text{for all}\quad \tilde{y}\in H^2(\Omega).\end{aligned}$$

By integrating the second order terms in the expressions above, we obtain

$$\begin{aligned}&\int_\Omega \operatorname{div}\Big([\nabla y\cdot\nabla\tilde{p}]\mathbf{V}\Big)-\nabla\tilde{p}\cdot\nabla\hat{\Sigma}\,d\Omega+\int_\Gamma\frac{\partial\tilde{p}}{\partial\mathbf{n}}\hat{\Sigma}+\frac{\partial\hat{\Sigma}}{\partial\mathbf{n}}\tilde{p}\,d\Gamma=0,\quad \text{for all}\quad \tilde{p}\in\mathscr{V},\\ &\int_\Omega \operatorname{div}\Big([\nabla\tilde{y}\cdot\nabla p]\mathbf{V}\Big)-\nabla\tilde{y}\cdot\nabla\hat{P}+\tilde{y}\hat{\Sigma}\,d\Omega+\int_\Gamma\frac{\partial\tilde{y}}{\partial\mathbf{n}}\hat{P}+\frac{\partial\hat{P}}{\partial\mathbf{n}}\tilde{y}\,d\Gamma=0,\quad \text{for all}\quad \tilde{y}\in\mathscr{V},\end{aligned} \tag{33}$$

where $\mathscr{V} = H^2(\Omega)$. Note that multiplying (2) and (18) by $\hat{P} \in \mathscr{V}$ and $\hat{\Sigma} \in \mathscr{V}$, respectively, one obtains

$$\begin{aligned}\langle E(y,\Omega), \hat{P}\rangle_\Upsilon = 0, \quad \text{for all} \quad \hat{P} \in \mathscr{V},\\ \langle E(p,y,\Omega), \hat{\Sigma}\rangle_\Upsilon = 0, \quad \text{for all} \quad \hat{\Sigma} \in \mathscr{V},\end{aligned}$$

where $\Upsilon := \mathscr{V}^* \times \mathscr{V}$ and

$$\langle E(y,\Omega), \hat{P}\rangle_\Upsilon := \int_\Omega (f\hat{P} - \nabla y \cdot \nabla \hat{P})\, d\Omega + \int_\Gamma \left(\frac{\partial y}{\partial \mathbf{n}}\hat{P} + y\frac{\partial \hat{P}}{\partial \mathbf{n}}\right) d\Gamma, \tag{34}$$

$$\langle E(p,y,\Omega), \hat{\Sigma}\rangle_\Upsilon := \int_\Omega \Big((y-y_d)\hat{\Sigma} - \nabla p \cdot \nabla \hat{\Sigma}\Big)\, d\Omega + \int_\Gamma \left(\frac{\partial p}{\partial \mathbf{n}}\hat{\Sigma} + \frac{\partial \hat{\Sigma}}{\partial \mathbf{n}} p\right) d\Gamma. \tag{35}$$

Using the divergence theorem on the boundary terms in (34) and (35), one obtains the equations on $\Omega_t(\mathbf{W})$ as follows

$$\begin{aligned}\langle E(y_t,\Omega_t(\mathbf{W})), \hat{P}_t\rangle_{\Upsilon_t} = 0, \quad \text{for all} \quad \hat{P}_t \in \mathscr{V}_t,\\ \langle E(p_t,y_t,\Omega_t(\mathbf{W})), \hat{\Sigma}_t\rangle_{\Upsilon_t} = 0, \quad \text{for all} \quad \hat{\Sigma}_t \in \mathscr{V}_t,\end{aligned} \tag{36}$$

where $\Upsilon_t := \mathscr{V}_t^* \times \mathscr{V}_t$ and

$$\langle E(y_t,\Omega_t(\mathbf{W})), \hat{P}_t\rangle_{\Upsilon_t} := \int_{\Omega_t(\mathbf{W})} f_t\hat{P}_t - \nabla y_t \cdot \nabla \hat{P}_t + \operatorname{div}(\nabla y_t\ \hat{P}_t + \nabla \hat{P}_t\ y_t)\, d\Omega_t(\mathbf{W}),$$

$$\begin{aligned}\langle E(p_t,y_t,\Omega_t(\mathbf{W})), \hat{\Sigma}_t\rangle_{\Upsilon_t} := &\int_{\Omega_t(\mathbf{W})} (y_t - y_d)\hat{\Sigma}_t - \nabla p_t \cdot \nabla \hat{\Sigma}_t\, d\Omega_t(\mathbf{W})\\ &+ \int_{\Omega_t(\mathbf{W})} \operatorname{div}(\nabla p_t\ \hat{\Sigma}_t + \nabla \hat{\Sigma}_t\ p_t)\, d\Omega_t(\mathbf{W}).\end{aligned}$$

Transforming (36) back to Ω, one obtains

$$\begin{aligned}\langle \tilde{E}(y^t,t), \hat{P}^t\rangle_\Upsilon &:= \int_\Omega (I_t f^t \hat{P}^t - A(t)\nabla y^t \cdot \nabla \hat{P}^t + I_t(B_t)_k \nabla\theta_k^t)\, d\Omega = 0,\\ \langle \tilde{E}(p^t,y^t,t), \hat{\Sigma}^t\rangle_\Upsilon &:= \int_\Omega (I_t(y^t - y_d)\hat{\Sigma}^t - A(t)\nabla p^t \cdot \nabla \hat{\Sigma}^t + I_t(B_t)_k \nabla\vartheta_k^t)\, d\Omega = 0,\end{aligned} \tag{37}$$

where

$$\theta^t := B_t \nabla y^t \hat{P}^t + B_t \nabla \hat{P}^t y^t, \quad \vartheta^t := B_t \nabla p^t \hat{\Sigma}^t + B_t \nabla \hat{\Sigma}^t p^t, \quad \theta := B_t \nabla y \hat{P} + B_t \nabla \hat{P} y,$$
$$\theta_0 := \nabla y \hat{P} + \nabla \hat{P} y, \quad \theta^r = B_t \nabla y^t \hat{P} + B_t \nabla \hat{P} y^t.$$

Lemma 4.2 *Let*

$$\mathscr{W}(t) := \left\langle \tilde{E}(y^t, t) - \tilde{E}(y, t) - \big(E(y^t, \Omega) - E(y, \Omega)\big), \hat{P} \right\rangle_{\Upsilon},$$

and

$$\mathscr{Y}(t) := \left\langle \tilde{E}(p^t, y^t, t) - \tilde{E}(p, y, t) - \big(E(p^t, y^t, \Omega) - E(p, y, \Omega)\big), \hat{\Sigma} \right\rangle_{\Upsilon}.$$

Then

$$\lim_{t \downarrow 0^+} \frac{1}{t} \mathscr{W}(t) = 0, \quad \lim_{t \downarrow 0^+} \frac{1}{t} \mathscr{Y}(t) = 0.$$

Proof It suffices to prove that $\lim_{t \downarrow 0^+} \frac{1}{t} \mathscr{W}(t) = 0$, since the second expression follows in an analogous way. Using (37) with $\hat{P}^t$ replaced by $\hat{P}$ and $(y^t, \hat{P}^t)$ replaced by $(y, \hat{P})$, we obtain

$$\langle \tilde{E}(y^t, t), \hat{P} \rangle_{\Upsilon} = \int_{\Omega} I_t f^t \hat{P} - A(t) \nabla y^t \cdot \nabla \hat{P} \, d\Omega + \int_{\Omega} I_t (B_t)_k \nabla \theta_k^r \, d\Omega, \tag{38}$$

and

$$\langle \tilde{E}(y, t), \hat{P} \rangle_{\Upsilon} = \int_{\Omega} I_t f^t \hat{P} - A(t) \nabla y \cdot \nabla \hat{P} \, d\Omega + \int_{\Omega} I_t (B_t)_k \nabla \theta_k \, d\Omega, \tag{39}$$

respectively. Subtracting (39) from (38), one obtains

$$\langle \tilde{E}(y^t, t), \hat{P} \rangle_{\Upsilon} - \langle \tilde{E}(y, t), \hat{P} \rangle_{\Upsilon} = \int_{\Omega} -A(t) \nabla \delta y \cdot \nabla \hat{P} + I_t (B_t)_k \nabla (\theta_k^r - \theta_k) \, d\Omega,$$

where $\delta y := y^t - y$ and $\delta p := p^t - p$. Replacing y by y^t in (34) and subtracting the result from (34), we obtain

$$\langle E(y, \Omega) - E(y^t, \Omega), \hat{\Sigma} \rangle_{\Upsilon} = \int_{\Omega} \nabla \delta y \cdot \nabla \hat{P} \, d\Omega - \int_{\Gamma} \left(\frac{\partial(\delta y)}{\partial \mathbf{n}} \hat{P} + \frac{\partial \hat{P}}{\partial \mathbf{n}} \delta y \right) d\Gamma.$$

Furthermore, the divergence theorem implies that

$$\int_{\Gamma} \left(\frac{\partial(\delta y)}{\partial \mathbf{n}} \hat{P} + \frac{\partial \hat{P}}{\partial \mathbf{n}} \delta y \right) d\Gamma = \int_{\Omega} \mathrm{div}\gamma \; d\Omega,$$

where $\gamma = (\nabla \delta y)\hat{P} + (\nabla \hat{P})\delta y$. Therefore, $\mathscr{W}(t) = \mathscr{W}_1(t) + \mathscr{W}_2(t)$, where

$$\mathscr{W}_1(t) := \int_{\Omega} \Big((I - A(t)) \nabla \delta y \nabla \hat{P} \Big) \, d\Omega, \quad \mathscr{W}_2(t) := \int_{\Omega} I_t (B_t)_k \nabla(\theta_k^r - \theta_k) - e_k \nabla \Gamma_k \; d\Omega.$$

Observe that $\mathscr{W}_2(t)$ can be expressed as

$$\begin{aligned}
\mathscr{W}_2(t) &= \int_{\Omega} (I_t(B_t)_k - e_k)\nabla(\theta_k^r - \theta_k) + e_k \nabla(\theta_k^r - \theta_k - \gamma_k) \; d\Omega \\
&= \int_{\Omega} (I_t(B_t)_k - e_k)\nabla(\theta_k^r - \theta_k) + \mathrm{div}(\theta^r - \theta - \gamma) \; d\Omega \\
&= \int_{\Omega} (I_t(B_t)_k - e_k)\nabla(\theta_k^r - \theta_k) \; d\Omega + \int_{\Gamma} (\theta^r - \theta - \gamma) \cdot \mathbf{n} \; d\Gamma,
\end{aligned}$$

where

$$(\theta^r - \theta - \gamma) \cdot \mathbf{n} = (B_t - I) \frac{\partial(\delta y)}{\partial \mathbf{n}} \hat{P} + (B_t - I) \frac{\partial \hat{P}}{\partial \mathbf{n}} \delta y,$$

and

$$\theta^r - \theta = (B_t \nabla \delta y)\hat{P} + (B_t \nabla \hat{P})\delta y.$$

Let $\mathscr{W}_{2,1}(t) := \int_{\Omega} (I_t(B_t)_k - e_k)\nabla(\theta_k^r - \theta_k) \, d\Omega$, and $\mathscr{W}_{2,2}(t) := \int_{\Gamma} (\theta^r - \theta - \gamma) \cdot \mathbf{n} \, d\Gamma$. Then, the following estimates hold,

$$\begin{aligned}
|\mathscr{W}_{2,1}(t)| &\leq ||(I_t(B_t)_k - e_k)||_\infty ||\theta_k^r - \theta_k||_{H^1} \\
&\leq ||(I_t(B_t)_k - e_k)||_\infty \Big(||B_t||_\infty ||\delta y||_{H^2} ||\hat{P}||_{H^1} + ||B_t||_\infty ||\hat{P}||_{H^2} ||\delta y||_{H^2} \Big), \\
|\mathscr{W}_{2,2}(t)| &\leq 2||(B_t) - I)||_\infty \Big(||\delta y||_{H^2} ||\hat{P}||_{H^2} \Big), \\
|\mathscr{W}_1(t)| &\leq ||I - A(t))||_\infty \Big(||\delta y||_{H^2} ||\hat{P}||_{H^2} \Big).
\end{aligned}$$

Analogously, we can find the estimates for $\mathscr{Y}(t)$. The result follows from Lemmas 3.5, 4.1, and the differentiability of the mappings $t \mapsto A(t)$, $t \mapsto I_t$, and $t \mapsto B_t$. □

The following lemma will become important in what follows.

Lemma 4.3 *Let* $p_\phi := (p \circ T_t^{-1})$, $\hat{\Sigma}_\phi := (\hat{\Sigma} \circ T_t^{-1})$, $y_\phi := y \circ T_t^{-1}$, $\hat{P}_\phi := (\hat{P} \circ T_t^{-1})$, *and*

$$\begin{aligned}
&\tilde{G}(\Omega_t(\mathbf{W}), y, p, \hat{P}, \hat{\Sigma}) \\
&\quad := \int\limits_{\Omega_t(\mathbf{W})} f\hat{P}_\phi - \nabla y_\phi \cdot \nabla \hat{P}_\phi + div(\nabla y_\phi\ \hat{P}_\phi + \nabla \hat{P}_\phi\ y_\phi)\, d\Omega_t(\mathbf{W}) \\
&\qquad + \int\limits_{\Omega_t(\mathbf{W})} (y_\phi - y_d)\hat{\Sigma}_\phi - \nabla p_\phi \cdot \nabla \hat{\Sigma}_\phi + div(\nabla p_\phi\ \hat{\Sigma}_\phi + \nabla \hat{\Sigma}_\phi\ p_\phi)\, d\Omega_t(\mathbf{W}).
\end{aligned}$$

Then, the partial derivative of $\tilde{G}(\Omega_t(\mathbf{W}), y, p, \hat{P}, \hat{\Sigma})$ *with respect to* t *is given by*

$$\begin{aligned}
\partial_t \tilde{G}|_{t=0} = &-\int\limits_{\Omega} div\Big([\nabla \dot{y} \cdot \nabla p + \nabla y \cdot \nabla \dot{p}]\mathbf{V}\Big)\, d\Omega \\
&+ \int\limits_{\Gamma} (\nabla y \cdot \nabla \hat{P} + \nabla p \cdot \nabla \hat{\Sigma})\mathbf{W} \cdot \mathbf{n}\, d\Gamma,
\end{aligned}$$

where $\dot{y} := -\nabla y \cdot \mathbf{W} \in H^2(\Omega)$, $\dot{p} := -\nabla p \cdot \mathbf{W} \in H^2(\Omega)$, $\dot{\hat{P}} := -\nabla \hat{P} \cdot \mathbf{W} \in H^1(\Omega)$, *and* $\dot{\hat{\Sigma}} := -\nabla \hat{\Sigma} \cdot \mathbf{W} \in H^1(\Omega)$.

Proof Since $(y,\ p) \in H^3(\Omega) \times H^3(\Omega)$ and $(\hat{\Sigma},\ \hat{P}) \in H^2(\Omega) \times H^2(\Omega)$, by Lemma 3.2(2), $(y_\phi,\ p_\phi)$, and $(\hat{\Sigma}_\phi,\ \hat{P}_\phi)$ also belong to $H^3(\Omega) \times H^3(\Omega)$ and $H^2(\Omega) \times H^2(\Omega)$, respectively. Furthermore, the derivatives of $(y_\phi,\ p_\phi)$ and $(\hat{\Sigma}_\phi,\ \hat{P}_\phi)$, with respect to t at $t = 0$ exist in $H^2(\Omega) \times H^2(\Omega)$ and $H^1(\Omega) \times H^1(\Omega)$, respectively, by Lemma 3.2(3) and are given by $(\dot{y}, \dot{p})$ and $(\dot{\hat{\Sigma}}, \dot{\hat{P}})$, respectively.

Taking the derivative of $\tilde{G}$ is complicated by the fact that $(\hat{\Sigma},\ \hat{P})$ belongs $H^2(\Omega) \times H^2(\Omega)$ only. To overcome this difficulty, we follow [15] and choose sequences $(y_k)_{k=1}^\infty \subset C^\infty(\bar{\Omega})$, $(p_k)_{k=1}^\infty \subset C^\infty(\bar{\Omega})$ which approximate y in $H^3(\Omega)$ and p in $H^3(\Omega)$, respectively. We denote by $\hat{\Sigma}_k,\ \hat{P}_k$ the solutions to (33) with $-\frac{\partial y}{\partial n}\mathbf{V} \cdot \mathbf{n}$ and $-\frac{\partial p}{\partial n}\mathbf{V} \cdot \mathbf{n}$ replaced by $-\frac{\partial y_k}{\partial n}\mathbf{V} \cdot \mathbf{n}$ and $-\frac{\partial p_k}{\partial n}\mathbf{V} \cdot \mathbf{n}$, respectively. Then $\hat{\Sigma}_k,\ \hat{P}_k \in H^3(\Omega) \times H^3(\Omega)$ and the limits $\lim_{k\mapsto\infty} \hat{\Sigma}_k = \hat{\Sigma}$, $\lim_{k\mapsto\infty} \hat{P}_k = \hat{P}$ hold in $H^2(\Omega)$. Extending $y_k,\ p_k,\ \hat{\Sigma}_k$, and $\hat{P}_k$ to elements in $H^3(\mathscr{D})$, the derivative of $\tilde{G}$ with respect to t can be calculated using Lemma 3.1 (1) and Lemma 3.2. Denoting $g(x,t)$ in Lemma 3.1 (1) by

$$\begin{aligned}
g(x,t) := &f\hat{P}_{k,\phi} - \nabla y_{k,\phi} \cdot \nabla \hat{P}_{k,\phi} + \mathrm{div}(\nabla y_{k,\phi}\ \hat{P}_{k,\phi} + \nabla \hat{P}_{k,\phi}\ y_{k,\phi}) \\
&+ (y_{k,\phi} - y_d)\hat{\Sigma}_{k,\phi} - \nabla P_{k,\phi} \cdot \nabla \hat{\Sigma}_{k,\phi} + \mathrm{div}(\nabla P_{k,\phi}\ \hat{\Sigma}_{k,\phi} + \nabla \hat{\Sigma}_{k,\phi}\ P_{k,\phi}),
\end{aligned}$$

where $p_{k,\phi} := (p_k \circ T_t^{-1})$, $\hat{\Sigma}_{k,\phi} := (\hat{\Sigma}_k \circ T_t^{-1})$, $y_{k,\phi} := y_k \circ T_t^{-1}$, $\hat{P}_{k,\phi} := (\hat{P}_k \circ T_t^{-1})$, we obtain

$$\begin{aligned}
\partial_t \tilde{G}|_{t=0} = \lim_{k\to\infty} \Big[&\int_\Gamma \Big(f\hat{P}_k - \nabla y_k \cdot \nabla \hat{P}_k + \operatorname{div}(\nabla y_k\ \hat{P}_k + \nabla \hat{P}_k\ y_k)\Big) \mathbf{W}\cdot \mathbf{n}\, d\Gamma \\
&+ \int_\Gamma \Big((y_k - y_d)\hat{\Sigma}_k - \nabla p_k \cdot \nabla \hat{\Sigma}_k + \operatorname{div}(\nabla p_k\ \hat{\Sigma}_k + \nabla \hat{\Sigma}_k\ p_k)\Big) \mathbf{W}\cdot \mathbf{n}\, d\Gamma \\
&+ \int_\Omega \Big(f\dot{\hat{P}}_k - \nabla y_k \cdot \nabla \dot{\hat{P}}_k + \operatorname{div}(\nabla y_k\ \dot{\hat{P}}_k + \nabla \dot{\hat{P}}_k\ y_k)\Big)\, d\Omega \\
&+ \int_\Omega \Big((y_k - y_d)\dot{\hat{\Sigma}}_k - \nabla p_k \cdot \nabla \dot{\hat{\Sigma}}_k + \operatorname{div}(\nabla p_k\ \dot{\hat{\Sigma}}_k + \nabla \dot{\hat{\Sigma}}_k\ p_k)\Big)\, d\Omega \\
&+ \int_\Omega \Big(\dot{y}_k\hat{\Sigma}_k - \nabla \dot{y} \cdot \nabla \hat{P}_k + \operatorname{div}(\nabla \dot{y}_k\ \hat{P}_k + \nabla \hat{P}_k\ \dot{y}_k)\Big)\, d\Omega \\
&+ \int_\Omega \Big(- \nabla \dot{p}_k \cdot \nabla \hat{\Sigma}_k + \operatorname{div}(\nabla \dot{p}_k\ \hat{\Sigma}_k + \nabla \hat{\Sigma}_k\ \dot{p}_k)\Big)\, d\Omega \Big].
\end{aligned} \tag{40}$$

Following arguments in [15], it can be shown that the above limit exists. Note that (34), (35) imply that the third and forth integrals in (40) vanish in the limit, i.e.,

$$\langle E(y,\Omega), \dot{\hat{P}}\rangle_\Upsilon = \lim_{k\to\infty} \int_\Omega \Big(f\dot{\hat{P}}_k - \nabla y_k \cdot \nabla \dot{\hat{P}}_k + \operatorname{div}(\nabla y_k\ \dot{\hat{P}}_k + \nabla \dot{\hat{P}}_k\ y_k)\Big)\, d\Omega = 0,$$

$$\langle E(p,y,\Omega), \dot{\hat{\Sigma}}\rangle_\Upsilon = \lim_{k\to\infty} \int_\Omega \Big((y_k - y_d)\dot{\hat{\Sigma}}_k - \nabla p_k \cdot \nabla \dot{\hat{\Sigma}}_k + \operatorname{div}(\nabla p_k\ \dot{\hat{\Sigma}}_k + \nabla \dot{\hat{\Sigma}}_k\ p_k)\Big)\, d\Omega = 0,$$

respectively. In addition, utilizing (33) with $\tilde{p} = \dot{p}$ and $\tilde{y} = \dot{y}$, one obtains in the limit, for the last two integrals in (40)

$$\int_\Omega \Big(\dot{y}\hat{\Sigma} - \nabla \dot{y} \cdot \nabla \hat{P} + \operatorname{div}(\nabla \dot{y}\ \hat{P} + \nabla \hat{P}\ \dot{y})\Big)\, d\Omega = -\int_\Omega \operatorname{div}\Big([\nabla \dot{y} \cdot \nabla p]\mathbf{V}\Big)\, d\Omega,$$

$$\int_\Omega \Big(- \nabla \dot{p} \cdot \nabla \hat{\Sigma} + \operatorname{div}(\nabla \dot{p}\ \hat{\Sigma} + \nabla \hat{\Sigma}\ \dot{p})\Big)\, d\Omega = -\int_\Omega \operatorname{div}\Big([\nabla y \cdot \nabla \dot{p}]\mathbf{V}\Big)\, d\Omega.$$

Furthermore, observe that

$$\int_\Gamma \operatorname{div}(\nabla y_k\ \hat{P}_k + \nabla \hat{P}_k\ y_k)\, d\Gamma = \int_\Gamma 2\nabla y_k \cdot \nabla \hat{P}_k + \Delta y_k\ \hat{P}_k + \Delta \hat{P}_k\ y_k\, d\Gamma.$$

Since $\lim_{k\to\infty} \Delta y_k + f = 0$ in $L^2(\Gamma)$, and $y_k = 0$ on Γ, we have, in the limit, for the first integral in (40)

$$\lim_{k\to\infty}\int_\Gamma \Big(f\hat{P}_k - \nabla y_k\cdot\nabla\hat{P}_k + \operatorname{div}(\nabla y_k\ \hat{P}_k + \nabla\hat{P}_k\ y_k)\Big)\mathbf{W}\cdot\mathbf{n}\,d\Gamma$$
$$= \int_\Gamma \nabla y\cdot\nabla\hat{P}\ \mathbf{W}\cdot\mathbf{n}\,d\Gamma.$$

Analogously, for the second integral,

$$\lim_{k\to\infty}\int_\Gamma \Big((y_k-y_d)\hat{\Sigma}_k - \nabla p_k\cdot\nabla\hat{\Sigma}_k + \operatorname{div}(\nabla p_k\ \hat{\Sigma}_k + \nabla\hat{\Sigma}_k\ p_k)\Big)\mathbf{W}\cdot\mathbf{n}\,d\Gamma$$
$$= \int_\Gamma \nabla p\cdot\nabla\hat{\Sigma}\ \mathbf{W}\cdot\mathbf{n}\,d\Gamma.$$

Consequently,

$$\partial_t\tilde{G}|_{t=0} = -\int_\Omega \operatorname{div}\Big([\nabla\dot{y}\cdot\nabla p+\nabla y\cdot\nabla\dot{p}]\mathbf{V}\Big)\,d\Omega + \int_\Gamma (\nabla y\cdot\nabla\hat{P}+\nabla p\cdot\nabla\hat{\Sigma})\mathbf{W}\cdot\mathbf{n}\,d\Gamma.$$

□

4.2 Shape Hessian

In this subsection, we establish the expression for the shape Hessian for the cost functional J. In what follows, we shall make use of the summation convention. For instance, where necessary, the divergence of a vectorial function $\boldsymbol{\psi}$ shall be expressed as

$$(\operatorname{div}\boldsymbol{\psi}) = e_i\nabla\psi_i, \tag{41}$$

where e_i stands for the i-th canonical basis vector in $\mathbb{R}^d$. Furthermore, we shall make use of the transformation of the divergence from $\Omega_t(\mathbf{W})$ to Ω. Using (14), and the summation convection in (41), one can express this transformation as

$$(\operatorname{div}\boldsymbol{\psi}_t)\circ T_t = D\psi_i^t B_t^T e_i = (B_t)_i\nabla\psi_i^t, \tag{42}$$

where $(B_t)_i$ denotes the i-th row of $B_t := (\mathrm{d}t_t)^{-T}$.

Theorem 4.1 *The shape Hessian of $J(\Omega)$ exists and it is given by the expression*

$$d^2J(\Omega)(\mathbf{V},\mathbf{W}) = \int_\Gamma \Big(\nabla y\cdot\nabla\hat{P} + \nabla p\cdot\nabla\hat{\Sigma} + div\left(\left[\tfrac{1}{2}y_d^2 + \nabla y\cdot\nabla p\right]\mathbf{V}\right)\Big)\mathbf{W}\cdot\mathbf{n}\,d\Gamma.$$

Proof Let $\Delta := dJ(\Omega_t(\mathbf{W}))\mathbf{V} - dJ(\Omega)\mathbf{V}$, $\Theta := [\frac{1}{2}y_d^2 + (\nabla y,\nabla p)]\mathbf{V}$, and $\Theta^t := [\frac{1}{2}y_d^2 + (B_t\nabla y^t, B_t\nabla p^t)]\mathbf{V}$. Then, using Lemma 3.3, (14) and (42), we can express Δ as

$$\Delta = \int\limits_{\Omega} I_t (B_t)_k \nabla \Theta_k^t - e_k \nabla \Theta_k \, d\Omega. \tag{43}$$

We re-write the right-hand side of (43) such that $\Delta = S(t) + R(t)$, where

$$S(t) := \int\limits_{\Omega} I_t ((B_t)_k \nabla \Theta_k^t - e_k \nabla \Theta_k^r) \, d\Omega, \tag{44}$$

$$R(t) := \int\limits_{\Omega} (I_t e_k \nabla \Theta_k^r - e_k \nabla \Theta_k) \, d\Omega, \tag{45}$$

and

$$\Theta^r := \left[\frac{1}{2} y_d^2 + (\nabla y^t, \nabla p^t) \right] \mathbf{V}.$$

The task now is to evaluate $\lim\limits_{t \downarrow 0^+} |R(t)|/t + |S(t)|/t$. We will do this in several steps. To this end, the terms on the right-hand side of (45) and (44) are rearranged to obtain $R(t) = R_1(t) + R_2(t) + R_3(t)$ and $S(t) = S_1(t) + S_2(t)$, respectively, where

$$\begin{aligned}
R_1(t) &:= \int\limits_{\Omega} (I_t - 1) e_k \nabla (\Theta_k^r - \Theta_k) \, d\Omega, \quad R_2(t) := \int\limits_{\Omega} e_k \nabla (\Theta_k^r - \Theta_k) \, d\Omega, \\
R_3(t) &:= \int\limits_{\Omega} (I_t - 1) e_k \nabla \Theta_k \, d\Omega, \quad S_1(t) \\
&:= \int\limits_{\Omega} I_t ((B_t)_k \nabla (\Theta_k^t - \Theta_k^s) - e_k \nabla (\Theta_k^r - \Theta_k)) \, d\Omega, \\
S_2(t) &:= \int\limits_{\Omega} I_t ((B_t)_k \nabla \Theta_k^s - e_k \nabla \Theta_k) \, d\Omega,
\end{aligned} \tag{46}$$

and

$$\Theta^s := \left[\frac{1}{2} y_d^2 + (B_t \nabla y, B_t \nabla p) \right] \mathbf{V}.$$

We now evaluate $\lim\limits_{t \downarrow 0^+} |R_1(t)|$. Note that by using the relation $ab - cd = (a - c)(b - d) + (a - c)d + c(b - d)$, with $a = \nabla y^t$, $b = \nabla p^t$, $c = \nabla y$, and $d = \nabla p$, we can express $\Theta^r - \Theta$ as

$$\begin{aligned}
\Theta^r - \Theta &= [(\nabla y^t, \nabla p^t) - (\nabla y, \nabla p)] \mathbf{V}, \\
&= [(\nabla \delta y, \nabla \delta p)] \mathbf{V} + [(\nabla \delta y, \nabla p) + (\nabla y, \nabla \delta p)] \mathbf{V},
\end{aligned} \tag{47}$$

where $\delta y = y^t - y$ and $\delta p = p^t - p$.

Using (47), the divergence theorem, the trace theorem, and the fact that $\delta y \in H_0^1(\Omega) \cap H^2(\Omega)$, $\delta p \in H_0^1(\Omega) \cap H^2(\Omega)$, we can estimate $\frac{1}{t}|R_1(t)|$ as

$$\begin{aligned}\frac{1}{t}|R_1(t)| &\leq \frac{1}{t}||I_t - 1||_{L^\infty} \int_\Gamma \left|\frac{\partial(\delta y)}{\partial \mathbf{n}} \frac{\partial(\delta p)}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n}\right| d\Gamma \\ &\quad + \frac{1}{t}||I_t - 1||_{L^\infty} \int_\Gamma \left|\left(\frac{\partial(\delta y)}{\partial \mathbf{n}} \frac{\partial p}{\partial \mathbf{n}} + \frac{\partial y}{\partial \mathbf{n}} \frac{\partial(\delta p)}{\partial \mathbf{n}}\right) \mathbf{V} \cdot \mathbf{n}\right| d\Gamma \\ &\leq \frac{1}{t}||I_t - 1||_{L^\infty} ||\delta y||_{H^2(\Omega)} ||\delta p||_{H^2(\Omega)} ||\mathbf{V} \cdot \mathbf{n}||_{L^\infty(\Gamma)} \\ &\quad + \frac{1}{t}||I_t - 1||_{L^\infty} ||\delta y||_{H^2(\Omega)} ||p||_{H^2(\Omega)} ||\mathbf{V} \cdot \mathbf{n}||_{L^\infty(\Gamma)} \\ &\quad + \frac{1}{t}||I_t - 1||_{L^\infty} ||y||_{H^2(\Omega)} ||\delta p||_{H^2(\Omega)} ||\mathbf{V} \cdot \mathbf{n}||_{L^\infty(\Gamma)}.\end{aligned}$$

Using Lemma 3.5, Lemma 4.1, and the differentiability of the mapping $t \mapsto I_t$, it follows that $\lim_{t \downarrow 0^+} \frac{1}{t}|R_1(t)| = 0$.

Next, we evaluate $\lim_{t \downarrow 0^+} \frac{R_2(t)}{t}$. Observe that by using (41) and (47), one can express $R_2(t)$ in (46) as

$$R_2(t) = \mathscr{X}_1(t) + \mathscr{X}_2(t),$$

where $\mathscr{X}_1(t) := \int_\Omega \operatorname{div}\Big([(\nabla \delta y, \nabla \delta p)]\mathbf{V}\Big)\, d\Omega$, $\mathscr{X}_2(t) := \int_\Omega \operatorname{div}\Big([(\nabla \delta y, \nabla p) + (\nabla y, \nabla \delta p)]\mathbf{V}\Big)\, d\Omega$.

Using the divergence theorem and the fact that $\delta y, \delta p \in H_0^1(\Omega) \cap H^2(\Omega)$, we can express $\mathscr{X}_1(t)$ as

$$\mathscr{X}_1(t) = \int_\Gamma \frac{\partial(\delta y)}{\partial \mathbf{n}} \frac{\partial(\delta p)}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n}\, d\Gamma,$$

and the following estimate holds

$$\begin{aligned}|\mathscr{X}_1(t)| \leq \int_\Gamma \left|\frac{\partial(\delta y)}{\partial \mathbf{n}} \frac{\partial(\delta p)}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n}\right| d\Gamma &\leq \left|\left|\frac{\partial(\delta y)}{\partial \mathbf{n}}\right|\right|_{L^2(\Gamma)} \left|\left|\frac{\partial(\delta p)}{\partial \mathbf{n}}\right|\right|_{L^2(\Gamma)} ||\mathbf{V} \cdot \mathbf{n}||_{L^\infty(\Gamma)} \\ &\leq \left|\left|\delta y\right|\right|_{H^2(\Omega)} \left|\left|\delta p\right|\right|_{H^2(\Omega)} ||\mathbf{V} \cdot \mathbf{n}||_{L^\infty(\Gamma)},\end{aligned}$$

by the trace theorem.

Therefore, by Lemma 3.5 and Lemma 4.1 we have

$$\lim_{t\downarrow 0^+}\frac{1}{t}|\mathscr{X}_1(t)| \leq \lim_{t\downarrow 0^+}\frac{1}{\sqrt{t}}||\delta y||_{H^2(\Omega)}\frac{1}{\sqrt{t}}||\delta p||_{H^2(\Omega)}||\mathbf{V}\cdot\mathbf{n}||_{L^\infty(\Gamma)} = 0.$$

Next, we need to evaluate $\lim\limits_{t\downarrow 0^+}\dfrac{\mathscr{X}_2(t)}{t}$. Using (33), we can express $\mathscr{X}_2(t)$ as

$$\begin{aligned}\mathscr{X}_2(t) = &\int\limits_{\Omega}\left(\nabla\delta y\cdot\nabla\hat{P} - \delta y\hat{\Sigma}\right)d\Omega - \int\limits_{\Gamma}\left(\frac{\partial\delta y}{\partial\mathbf{n}}\hat{P} + \frac{\partial\hat{P}}{\partial\mathbf{n}}\delta y\right)d\Gamma\\ &+\int\limits_{\Omega}(\nabla\delta p\cdot\nabla\hat{\Sigma})\,d\Omega - \int\limits_{\Gamma}\left(\frac{\partial(\delta p)}{\partial\mathbf{n}}\hat{\Sigma} + \frac{\partial\hat{\Sigma}}{\partial\mathbf{n}}(\delta p)\right)d\Gamma,\end{aligned}$$

where $(y, p) \in H^2(\Omega)\times H^2(\Omega)$ satisfy (2) and (18). Observe that $\mathscr{X}_2(t)$ can be expressed as

$$\begin{aligned}\mathscr{X}_2(t) = &\left\langle\tilde{E}(y^t, t) - \tilde{E}(y, t) - \left(E(y^t, \Omega) - E(y, \Omega)\right), \hat{P}\right\rangle_\Upsilon\\ &+\left\langle\tilde{E}(p^t, y^t, t) - \tilde{E}(p, y, t) - \left(E(p^t, y^t, \Omega) - E(p, y, \Omega)\right), \hat{\Sigma}\right\rangle_\Upsilon\\ &+\langle\tilde{E}(p, y, t), \hat{\Sigma}\rangle_\Upsilon - \langle\tilde{E}(p, y, 0), \hat{\Sigma}\rangle_\Upsilon + \langle\tilde{E}(y, t), \hat{P}\rangle_\Upsilon - \langle\tilde{E}(y, 0), \hat{P}\rangle_\Upsilon,\end{aligned}$$

where the extra terms $\langle\tilde{E}(y^t, t), \hat{P}\rangle_\Upsilon$, $\langle\tilde{E}(y, 0), \hat{P}\rangle_\Upsilon$, $\langle\tilde{E}(p^t, y^t, t), \hat{\Sigma}\rangle_\Upsilon$, and $\langle\tilde{E}(p, y, 0), \hat{\Sigma}\rangle_\Upsilon$ introduced, vanish by (36), (37), and the fact that $\langle\tilde{E}(y, 0), \hat{P}\rangle_\Upsilon = \langle\tilde{E}(y, \Omega), \hat{P}\rangle_\Upsilon$, and $\langle\tilde{E}(p, y, 0), \hat{\Sigma}\rangle_\Upsilon = \langle E(p, y, \Omega), \hat{\Sigma}\rangle_\Upsilon$. Utilizing the notation used in Lemma 4.2, we obtain

$$\begin{aligned}\mathscr{X}_2(t) = &\mathscr{W}(t) + \mathscr{Y}(t) + \langle\tilde{E}(p, y, t), \hat{\Sigma}\rangle_\Upsilon - \langle\tilde{E}(p, y, 0), \hat{\Sigma}\rangle_\Upsilon + \langle\tilde{E}(y, t), \hat{P}\rangle_\Upsilon\\ &-\langle\tilde{E}(y, 0), \hat{P}\rangle_\Upsilon.\end{aligned}$$

Using Lemma 4.2, we obtain

$$\lim_{t\downarrow 0^+}\frac{1}{t}\mathscr{X}_2(t) = \frac{d}{\mathrm{d}t}\tilde{E}(y, \hat{P}; t)|_{t=0} + \frac{d}{\mathrm{d}t}\tilde{E}(p, y, \hat{\Sigma}; t)|_{t=0}. \tag{48}$$

The expression on the right-hand side of (48) can be computed by transforming $y, \hat{P}, p, \hat{\Sigma}$ defined in Ω back to $\Omega_t(\mathbf{W})$ via $y \mapsto y\circ T_t^{-1}$, $p \mapsto p\circ T_t^{-1}$, $\hat{P} \mapsto \hat{P}\circ T_t^{-1}$, and $\hat{\Sigma} \mapsto \hat{\Sigma}\circ T_t^{-1}$. Utilizing Lemma 4.3, one obtains

$$\begin{aligned}\lim_{t\downarrow 0^+}\frac{1}{t}\mathscr{X}_2(t) = \partial_t\tilde{G}|_{t=0} = &-\int\limits_{\Omega}\mathrm{div}\Big([\nabla\dot{y}\cdot\nabla p + \nabla y\cdot\nabla\dot{p}]\mathbf{V}\Big)\,d\Omega\\ &+\int\limits_{\Gamma}(\nabla y\cdot\nabla\hat{P} + \nabla p\cdot\nabla\hat{\Sigma})\mathbf{W}\cdot\mathbf{n}\,d\Gamma.\end{aligned} \tag{49}$$

We shall revisit $R_3(t)$ later on. Let us now provide an estimate for $|S_1(t)|/t$. Observe that

$$S_1(t) = \int_\Omega I_t(((B_t)_k - e_k)\nabla(\Theta_k^t - \Theta_k^s) + e_k \nabla \mathscr{Q}_k)\, d\Omega$$
$$= \int_\Omega I_t(\mathrm{div}((B_t - I)(\Theta^t - \Theta^s)) + \mathrm{div}\mathscr{Q})\, d\Omega,$$

where

$$\mathscr{Q} := \Theta^t - \Theta^s - \Theta^r + \Theta = [((B_t - I)\nabla\delta y, B_t\nabla p^t) + (\nabla\delta y, (B_t - I)\nabla p^t) + (B_t\nabla y, (B_t - I)\nabla\delta p) + ((B_t - I)\nabla y, (B_t - I)\nabla\delta p)]\mathbf{V},$$

and

$$\Theta^t - \Theta^s = [(B_t\nabla\delta y, B_t\nabla p^t) + (B_t\nabla y, B_t\nabla\delta p)]\mathbf{V}.$$

Hence, we have

$$\left|\frac{1}{t}S_1(t)\right| \leq ||I_t||_{L^\infty(\Omega)} \int_\Omega \left|\mathrm{div}\left(\frac{1}{t}(B_t - I)(\Theta^t - \Theta^s) + \mathscr{Q}\right)\right| d\Omega.$$

Using the divergence theorem and the trace theorem, we obtain the following estimate

$$\left|\frac{1}{t}S_1(t)\right| \leq ||I_t||_{L^\infty(\Omega)} ||\mathbf{V}\cdot\mathbf{n}||_{L^\infty(\Gamma)}(\mathscr{P}(t) + \mathscr{R}(t)),$$

where

$$\mathscr{P}(t) := \left\|\frac{1}{t}(B_t - I)\right\|_{L^\infty} ||B_t||^2_{L^\infty}(||\delta y||_{H^2(\Omega)}||p^t||_{H^2(\Omega)} + ||y||_{H^2(\Omega)}||\delta p||_{H^2(\Omega)}),$$

and

$$\mathscr{R}(t) := \left\|\frac{1}{t}(B_t - I)\right\|_{L^\infty} ||B_t||_{L^\infty}\Big(||\delta y||_{H^2(\Omega)}||p^t||_{H^2(\Omega)} + ||y||_{H^2(\Omega)}||\delta p||_{H^2(\Omega)}\Big)$$
$$+ ||(B_t - I)||_{L^\infty}||\delta y||_{H^2(\Omega)}||p^t||_{H^2(\Omega)}.$$

Using Lemmas 3.5, 4.1, and the differentiability of the mapping $t \mapsto B_t$, one obtains $\lim_{t\downarrow 0^+} \left|\frac{1}{t}S_1(t)\right| = 0$.

We combine the remaining expressions, i.e., $R_3(t)$ and $S_2(t)$ into $Q(t)$ such that $Q(t) := S_2(t) + R_3(t)$. Then, $Q(t)$ can be expressed as

$$Q(t) = \int_{\Omega} I_t (B_t)_k \nabla\Theta_k^s - e_k \nabla\Theta_k \, d\Omega.$$

Next, we evaluate $\lim_{t\downarrow 0^+} \dfrac{Q(t)}{t}$. Using Lemma 3.3, (41), (42), and (13), we obtain

$$\begin{aligned} Q(t) &= \int_{\Omega_t(\mathbf{W})} (B_t)_k \nabla\Theta_k^s \circ T_t^{-1} \, d\Omega_t(\mathbf{W}) - \int_{\Omega} e_k \nabla\Theta_k \, d\Omega \\ &= \int_{\Omega_t(\mathbf{W})} \operatorname{div}\,(\Theta^s \circ T_t^{-1}) \, d\Omega_t(\mathbf{W}) - \int_{\Omega} \operatorname{div}\,\Theta \, d\Omega. \end{aligned} \tag{50}$$

Dividing the expression in (50) by t and making use of (14), the definition of the shape derivative, Lemma 3.1(1), and Lemma 3.2(3), we obtain

$$\begin{aligned} \lim_{t\downarrow 0^+} \frac{Q(t)}{t} &= \frac{d}{dt} \int_{\Omega_t(\mathbf{W})} \operatorname{div}\Big([\frac{1}{2}y_d^2 + \nabla y_\phi \cdot \nabla p_\phi]\mathbf{V}\Big) \, d\Omega_t(\mathbf{W})\Big|_{t=0} \\ &= \int_{\Omega} \operatorname{div}\Big(\Big[\nabla\dot{y} \cdot \nabla p + \nabla y \cdot \nabla\dot{p}\Big]\mathbf{V}\Big) \, d\Omega \\ &\quad + \int_{\Gamma} \operatorname{div}\left(\left[\frac{1}{2}y_d^2 + \nabla y \cdot \nabla p\right]\mathbf{V}\right) \mathbf{W}\cdot\mathbf{n} \, d\Gamma. \end{aligned} \tag{51}$$

Collecting the results in (49) and (51), we obtain

$$\begin{aligned} d^2 J(\Omega)(\mathbf{V},\mathbf{W}) &= \lim_{t\downarrow 0^+} \frac{Q(t)}{t} + \lim_{t\downarrow 0^+} \frac{\mathscr{X}_2(t)}{t} \\ &= \int_{\Gamma} \Big(\nabla y \cdot \nabla\hat{P} + \nabla p \cdot \nabla\hat{\Sigma} + \operatorname{div}([\frac{1}{2}y_d^2 + \nabla y \cdot \nabla p]\mathbf{V})\Big)\mathbf{W}\cdot\mathbf{n} \, d\Gamma. \end{aligned} \tag{52}$$

□

Remark 4.1 In shape calculus, the shape Hessian of a cost functional is typically not symmetric and can be decomposed into symmetric and non-symmetric parts, see, e.g., [17, pp. 286–287], [9, pp. 384–387]. As we shall show, this also holds true for the expression of the shape Hessian of J in (52).

To this end, we analyze each term in (52) for symmetry. In the first two addends, $\hat{P}$ and $\hat{\Sigma}$ depend linearly on $\mathbf{V}$. We denote them by $\hat{P}_V$ and $\hat{\Sigma}_V$ and note that $(\hat{\Sigma}_V, \hat{P}_V) \in H^2(\Omega) \times H^2(\Omega)$ satisfy

$$-\Delta\hat{\Sigma}_V = 0 \quad \text{in} \quad \Omega, \quad \hat{\Sigma}_V = -\frac{\partial y}{\partial \mathbf{n}}\mathbf{V}\cdot\mathbf{n} \quad \text{on} \quad \Gamma, \tag{53}$$

$$-\Delta \hat{P}_V = \hat{\Sigma}_V \quad \text{in} \quad \Omega, \quad \hat{P}_V = -\frac{\partial p}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n} \quad \text{on} \quad \Gamma. \tag{54}$$

Utilizing (53)–(54) and following the arguments in [18], we find

$$\int_\Gamma (\nabla y \cdot \nabla \hat{P}_V + \nabla p \cdot \nabla \hat{\Sigma}_V)\, \mathbf{W} \cdot \mathbf{n}\, d\Gamma = \int_\Gamma (\nabla y \cdot \nabla \hat{P}_W + \nabla p \cdot \nabla \hat{\Sigma}_W)\, \mathbf{V} \cdot \mathbf{n}\, d\Gamma,$$

where $(\hat{\Sigma}_W, \hat{P}_W) \in H^2(\Omega) \times H^2(\Omega)$ satisfy (53)–(54) with $\mathbf{V}$ replaced by $\mathbf{W}$. This expression is symmetric in $(\mathbf{V}, \mathbf{W})$.

Now let us turn to the third additive term in (52) and set $r := \frac{1}{2} y_d^2 + \nabla y \cdot \nabla p$. From [9, p. 374], it follows that

$$\begin{aligned}
&\int_\Gamma \operatorname{div}(r\, \mathbf{V})\, \mathbf{W} \cdot \mathbf{n}\, d\Gamma \\
&\quad = \int_\Gamma \Big\{ \Big(\frac{\partial r}{\partial \mathbf{n}} + \kappa\, r \Big) v_n w_n + r (D_\Gamma \mathbf{n} \mathbf{V}_\Gamma \cdot \mathbf{W}_\Gamma - \mathbf{V}_\Gamma \cdot \nabla_\Gamma w_n - \mathbf{W}_\Gamma \cdot \nabla_\Gamma v_n) \Big\}\, d\Gamma \\
&\qquad + \int_\Gamma r\, D\mathbf{V}\mathbf{W} \cdot \mathbf{n}\, d\Gamma,
\end{aligned} \tag{55}$$

where $v_n = \mathbf{V} \cdot \mathbf{n}$, $w_n := \mathbf{W} \cdot \mathbf{n}$, $\mathbf{V}_\Gamma := \mathbf{V} - v_n \mathbf{n}$, $\mathbf{W}_\Gamma := \mathbf{W} - w_n \mathbf{n}$, κ is the mean curvature of Γ, $\nabla_\Gamma v_n := \nabla v_n|_\Gamma - (\nabla v_n \cdot \mathbf{n})\mathbf{n}$, and $\nabla_\Gamma w_n := \nabla w_n|_\Gamma - (\nabla w_n \cdot \mathbf{n})\mathbf{n}$. The first term in (55) is symmetric and only depends on the trace of r on Γ and a group of terms involving $\mathbf{V}_\Gamma$, $\mathbf{W}_\Gamma$ and the tangential derivatives of v_n and w_n on Γ. The second term in (55) is non-symmetric. It involves the shape gradient multiplied by the derivative of the first vector field in the direction of the second vector field.

Remark 4.2 The proof of Theorem 4.1 reveals that the essential ingredients for establishing the expression for the shape Hessian of cost functional J are the properties of Lemma 3.5 and Lemma 4.1.

Remark 4.3 The method and the associated computation we presented can be extended to more rough cost functionals, e.g., the volume functional

$$J(\Omega) = \int_\Omega |\nabla(y - y_d)|^2\, d\Omega$$

subject to the constraint (2). If $f \in H^1(\mathscr{D})$, $y_d \in H^3(\mathscr{D})$, and $\Omega \subset \mathbb{R}^3$ is of class $C^{2,1}$, then, it is shown in [8] that $J(\Omega)$ has a shape derivative given by

$$dJ(\Omega)\mathbf{h} = \int_\Gamma \Big(\frac{\partial y}{\partial \mathbf{n}} \frac{\partial p}{\partial \mathbf{n}} - \frac{\partial (y - y_d)}{\partial \mathbf{n}} \frac{\partial (y + y_d)}{\partial \mathbf{n}} \Big)^2\, d\Gamma. \tag{56}$$

By careful revision of the steps in Sect. 4.2, we can further show that, under the above regularity assumption on the data, $J(\Omega)$ has a second order shape derivative given by

$$d^2 J(\Omega)(\mathbf{V},\mathbf{W}) = \int_\Gamma \Big(\nabla y \cdot \nabla \hat{P} + \nabla p \cdot \nabla \hat{\Sigma} - 2\nabla y \nabla \hat{\Sigma} + \operatorname{div}([\nabla y \cdot \nabla p - \nabla(y - y_d) \cdot \nabla(y + y_d)]\mathbf{V})\Big)\mathbf{W} \cdot \mathbf{n}\, d\Gamma, \quad (57)$$

where $(\hat{P}, \hat{\Sigma}) \in H^2(\Omega) \times H^2(\Omega)$ satisfy

$$\Delta \hat{P} = 2\Delta \hat{\Sigma} \quad \text{in} \quad \Omega, \quad \hat{P} = -\frac{\partial p}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n} \quad \text{on} \quad \Gamma, \quad (58)$$

$$\Delta \hat{\Sigma} = 0 \quad \text{in} \quad \Omega, \quad \hat{\Sigma} = -\frac{\partial y}{\partial \mathbf{n}} \mathbf{V} \cdot \mathbf{n} \quad \text{on} \quad \Gamma. \quad (59)$$

Remark 4.4 The regularity assumptions on the data required for the derivation of (56) and (57) can be reduced. Specifically, if $f \in W^{1,q}(\mathscr{D})$ and $y_d \in W^{3,q}(\mathscr{D})$ with $q \in (1, \frac{6}{5})$, then using the re-arrangement of the first order perturbation of the cost, it is shown in [8] that $J(\Omega)$ has a shape derivative given by (56).

On the other hand, it was shown in [8] that not even the first order shape derivative of the state y in (2) exists in $H^1(\Omega)$. Hence the shape derivative of $J(\Omega)$ cannot be obtained using the chain rule approach.

By careful revision of the above steps, we can further show that, under the above regularity assumption on the data, $J(\Omega)$ has a second order shape derivative given by (57) where $(\hat{P}, \hat{\Sigma}) \in W^{2,q} \times W^{2,q}$ satisfy (58), (59).

5 Conclusions

It was demonstrated that, for a certain class of shape optimization problems, the first and second order shape derivatives can be obtained without recourse to the shape derivative of the state variable. The methodology is more general than considered here. This can be exploited in future work.

Acknowledgments The second author was supported in part by the Austrian Science Fund Fond zur Forderung der Wissenschaftlichen Forschung (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences").

References

1. Henrot, A., Pierre, M.: About critical points of the energy in electromagnetic shaping problem. In: Zolesio J.-P. (ed.) Boundary Control and Boundary Variation. Lecture notes in Control and Information Sciences, vol. 178, pp. 238–252. Springer, Berlin (1991)
2. Tiihonen, T.: Shape optimization and trial methods for free boundary problems. ESAIM Math. Model. Numer. Anal. **31**, 805–825 (1997)
3. Burger, M.: Levenberg–Marquardt level set methods for inverse obstacle problems. Inverse Problems **20**, 259–282 (2004)

4. Fujii, N.: Second variation and its application in domain optimization problem. in control of distributed parameter systems. In: Proceedings of the 4th IFAC Symposium. Pergamon Press, Oxford (1987)
5. Simon, J.: Second variations for domain optimization problems. In: Kappel F. et al. (eds.) Control and Estimation of Distributed Parameter Systems. International Series of Numerical Mathematics vol. 91, pp. 361–378 (1989)
6. Delfour, M., Zolesio, J.P.: Anatomy of the shape Hessian. Ann. Mat. Pura Appl. **519**, 315–339 (1991)
7. Delfour, M., Zolesio, J.P.: Velocity method and Lagrangian formulation for the computation of the shape Hessian. SIAM J. Control Optim. **29**(6), 1414–1442 (1991)
8. Ito, K., Kunisch, K., Peichl, G.: Variational approach to shape derivatives. ESAIM Control Optim. Calc. Var. **14**, 517–539 (2008)
9. Delfour, M.C., Zolésio, J.P.: Shapes and geometries: analysis, differential calculus, and optimization. Society for Industrial and Applied Mathematics, Philadelphia (2001)
10. Haslinger, J., Mäkinen, R.A.E.: Introduction to Shape Optimization: Theory, Approximation, and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2003)
11. Sokolowski, J., Zolésio, J.P.: Introduction to Shape Optimization. Shape Sensitivity Analysis. Springer-Verlag, Berlin (1992)
12. Plotnikov, P., Sokolowski, J.: Compressible Navier–Stokes Equations: Theory and Shape Optimization. Monografie Matematyczne, vol. 73. Birkhäuser, Basel (2012)
13. Ito, K., Kunisch, K., Peichl, G.: Variational approach to shape derivatives for a class of Bernoulli problems. J. Math. Anal. Appl. **314**(1), 126–149 (2006)
14. Girault, V., Raviart, P.A.: Finite Element Methods for Navier–Stokes. Springer-Verlag, Berlin (1986)
15. Haslinger, J., Ito, K., Kozubek, T., Kunisch, K., Peichl, G.: On the shape derivative for problems of Bernoulli type. Interfaces Free Bound. **11**(2), 317–330 (2009)
16. Evans, L.C.: Partial Differential Equations. American Mathematical Society, Providence (1998)
17. Khludnev, A.M., Sokolowski, J.: Modelling and Control in Solid Mechanics. Birkhäuser, Basel (1997)
18. Bucur, D., Zolesio, J.P.: Anatomy of the shape Hessian via lie brackets. Ann. Mat. Pura. Appl. **173**(4), 127–143 (1997)