Second derivatives of cost functions and H^1 Newton method in shape optimization problems

Hideyuki Azegami

Abstract We derive the second-order shape derivatives (shape Hessians) of cost functions for shape optimization problems of domains in which boundary value problems of partial differential equations are defined, and propose an H^1 Newton method to solve the problems using the shape Hessians. In this paper, we formulate an abstract shape optimization problem and show the computations of the first and second order shape derivatives of cost functions under the abstract framework. Then, using the shape gradients and Hessians, we propose an H^1-Newton method to solve the given problem. As an illustration, the shape Hessians of a mean compliance and a domain measure are derived and then used for a numerical example.

1 Introduction

Optimization problems with respect to shapes of domains in which boundary value problems of partial differential equations are defined are called shape optimization problems. Among them, a problem formulated using a domain variation by the Lagrange description as a design variable is classified as a domain variation type. In the previous paper [2], we showed an evaluation method of Fréchet derivatives (shape derivatives) of cost functions with respect to domain variation and a solution to the problems based on a gradient method (H^1 gradient method). In this paper, we present an evaluation method of the second-order shape derivatives (shape Hessians) of cost functions and a solution using a Newton method (H^1 Newton method).

Notation $W^{s,p}\left(\Omega_0;\mathbb{R}^d\right)$ will be used to represent the Sobolev space for the set of functions which are defined in Ω_0, have values in $\mathbb{R}^d$, and are $s \in [0,\infty]$ times differentiable and $p \in [1,\infty]$-th order Lebesgue integrable. Moreover, $L^p\left(\Omega_0;\mathbb{R}^d\right)$ and $H^s\left(\Omega_0;\mathbb{R}^d\right)$ are denoted by $W^{0,p}\left(\Omega_0;\mathbb{R}^d\right)$ and $W^{s,2}\left(\Omega_0;\mathbb{R}^d\right)$, respectively.

Hideyuki Azegami
Graduate School of Informatics, Nagoya University, A4-2(780) Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan, e-mail: azegami@i.nagoya-u.ac.jp

2 Abstract optimum design problem

In this paper, basic results will be shown using an abstract optimum design problem as follows. Let $\phi \in \mathscr{D} \subset X$ be a design variable from an admissible set $\mathscr{D}$ in a Hilbert space X. For $\phi \in \mathscr{D}$, a state variable u in a Hilbert space U is assumed to be uniquely determined as a solution of the following problem.

Problem 2.1 (Abstract variational problem) *For $\phi \in \mathscr{D}$, let $a(\phi): U \times U \to \mathbb{R}$ be a bounded and coercive bilinear form in U, and $l(\phi) = l(\phi)(\cdot) = \langle l(\phi), \cdot \rangle \in U'$ (dual space of U). Find $u \in U$ such that*

$$a(\phi)(u,v) = l(\phi)(v) \quad \forall v \in U.$$

Problem 2.1 can be equivalently stated as follows. "Let $\tau(\phi): U \to U'$ be an isomorphism by Lax-Milgram theorem (Theorem 18.F in [3]) when $a(\phi)(\cdot,\cdot)$ is a bounded and coercive bilinear form in U. Find $u \in U$ satisfying

$$s(\phi,u) = l(\phi) - \tau(\phi)u = 0_{U'}.\text{"} \tag{1}$$

We assume that the solution u of Problem 2.1 is an element of admissible set $\mathscr{S} \subset U$ in order to assure that $\phi + \varphi \in \mathscr{D}$, where φ is a variation of the design variable ϕ obtained by the gradient or the Newton method which will be shown later. Given the pair $(\phi,u) \in \mathscr{D} \times \mathscr{S}$, we consider the following design problem.

Problem 2.2 (Abstract optimum design problem) *For $f_0, \ldots, f_m : \mathscr{D} \times \mathscr{S} \to \mathbb{R}$, find $(\phi^\star, u^\star) \in \mathscr{D} \times \mathscr{S}$ such that*

$$f_0(\phi^\star,u^\star) = \min_{(\phi,u)\in\mathscr{D}\times\mathscr{S}} \{ f_0(\phi,u) \mid f_1(\phi,u) \le 0, \ldots, f_m(\phi,u) \le 0, \text{ Problem 2.1} \}.$$

2.1 Gradient of cost function f_i with respect to ϕ

In Problem 2.2, Problem 2.1 is assumed as an equality constraint. In this section, using the following problem, we show the computations of the Fréchet derivative and Hessian of cost function f_i with respect to an arbitrary variation $\varphi \in X$ (or $\varphi \in Y \subset X$ with a linear space $Y \supset \mathscr{D}$) of the design variable $\phi \in \mathscr{D}$ subject to the given equality constraint.

Problem 2.3 (Abstract optimum design problem with equality constraint) *For $f_i : \mathscr{D} \times \mathscr{S} \to \mathbb{R}$ and $s(\phi,u)$ in (1), find $(\phi^\star, u^\star)$ such that*

$$f_i(\phi^\star,u^\star) = \min_{(\phi,u)\in\mathscr{D}\times\mathscr{S}} \{ f_i(\phi,u) \mid s(\phi,u) = 0_{U'} \}.$$

In order to show necessary and sufficient conditions for a local minimizer of Problem 2.3, we define the Lagrange function with respect to Problem 2.3 as

$$\mathscr{L}_i(\phi,u,v_i) = f_i(\phi,u) + \langle s(\phi,u), v_i\rangle = f_i(\phi,u) + \mathscr{L}_{\mathrm{S}}(\phi,u,v_i), \tag{2}$$

where $\mathscr{L}_{\mathrm{S}}(\phi,u,v_i)$ denotes the Lagrange function with respect to Problem 2.1, and u and v_i are the variables in $\mathscr{S}\subset U$ corresponding to the solution of Problem 2.1 and the Lagrange multiplier with respect to the equality constraint of Problem 2.1 for f_i respectively. In the definition of Lagrange function of (2), notice that u is not necessary to be the solution of Problem 2.1. $\mathscr{S}$ and U will be used as the admissible set and the set for test function of u and v_i in Problem 2.1 and the adjoint problem shown later in (13), respectively. With respect to an arbitrary variation $(\varphi,\hat{u},\hat{v}_i)\in X\times U^2$ of $(\phi,u,v_i)\in\mathscr{D}\times\mathscr{S}^2$, the Fréchet derivative of $\mathscr{L}_i$ can be written as

$$\begin{aligned}
&\mathscr{L}_i'(\phi,u,v_i)[\varphi,\hat{u},\hat{v}_i] = \mathscr{L}_{i\phi}(\phi,u,v_i)[\varphi] + \mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}] + \mathscr{L}_{iv_i}(\phi,u,v_i)[\hat{v}_i]\\
&= f_i'(\phi,u)[\varphi,\hat{u}] + \left\langle s'(\phi,u)[\varphi,\hat{u}], v_i\right\rangle + \langle s(\phi,u),\hat{v}_i\rangle\\
&= f_{i\phi}(\phi,u)[\varphi] + f_{iu}(\phi,u)[\hat{u}] + \left\langle s_\phi(\phi,u)[\varphi] + s_u(\phi,u)[\hat{u}], v_i\right\rangle + \langle s(\phi,u),\hat{v}_i\rangle\\
&= \left(\langle f_{i\phi}(\phi,u),\varphi\rangle + \left\langle s_\phi(\phi,u)[\varphi], v_i\right\rangle\right) + \langle f_{iu}(\phi,u) - \tau^*(\phi)v_i,\hat{u}\rangle + \langle s(\phi,u),\hat{v}_i\rangle\\
&= \langle g_i(\phi,u,v_i),\varphi\rangle + \mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}] + \mathscr{L}_{\mathrm{S}}(\phi,u,\hat{v}_i).
\end{aligned}\tag{3}$$

Here, $\tau^*(\phi): U\to U'$ is the adjoint operator of $\tau(\phi)$. Using the notations in (3), we have the following result.

Theorem 2.1 (The first-order necessary condition of a local minimizer) *Let $f_i\in C^1(\mathscr{D}\times\mathscr{S};\mathbb{R})$ and $s\in C^1(\mathscr{D}\times\mathscr{S};U')$. If (ϕ,u) is a local minimizer of Problem 2.3, there exists a $v_i\in U$ which satisfies*

$$\langle g_i(\phi,u,v_i),\varphi\rangle + \mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}] = 0 \quad \forall(\varphi,\hat{u})\in X\times U, \tag{4}$$

$$\mathscr{L}_{\mathrm{S}}(\phi,u,\hat{v}_i) = 0 \quad \forall\hat{v}_i\in U. \tag{5}$$

Proof. From the fact that $s\in C^1(\mathscr{D}\times\mathscr{S};U')$ and there is a unique solution u which satisfies $s(\phi,u)=0_{U'}$, s satisfies the following assumptions for the implicit function theorem in a neighborhood $B_X\times B_U\subset X\times U$ of $(\phi,u)\in\mathscr{D}\times\mathscr{S}$:

1. $s(\phi,u)=0_{U'}$,
2. $s\in C^0(B_X\times B_U;U')$,
3. $s(\phi,\cdot)\in C^1(B_U;U')$ with respect to an arbitrary $y=(\phi,w)\in B_X\times B_U$ and $s_u(\phi,u)=-\tau: U\to U'$ is continuous at (ϕ,u),
4. $(s_u(\phi,u))^{-1}=-\tau^{-1}: U'\to U$ is bounded and linear.

From the implicit function theorem, there exist some neighborhood $\hat{B}_X\times\hat{B}_U\subset B_X\times B_U$ and continuous mapping $\upsilon:\hat{B}_X\to\hat{B}_U$, and $s(\phi,u)=0_{U'}$ can be written as

$$u=\upsilon(\phi). \tag{6}$$

Hence, we can define $y(\phi)=(\phi,\upsilon(\phi))\in C^1(\mathscr{D};X\times U)$ and write $f_i(\phi,\upsilon(\phi))=f_i(y(\phi))$ by $\tilde{f}_i(\phi)$. Since $f_i\in C^1(\mathscr{D}\times\mathscr{S};\mathbb{R})$, then ϕ being a local minimizer implies that

$$\tilde{f}_i'(\phi)[\varphi] = y'^*(\phi) \circ g_i(\phi, \upsilon(\phi))[\varphi] = 0 \quad \forall \varphi \in X. \tag{7}$$

Here,

$$g_i(\phi, \upsilon(\phi)) = f_i'(\phi, \upsilon(\phi)) \in \mathfrak{L}\left(X; X' \times U'\right) = \mathfrak{L}(X; \mathfrak{L}(X \times U; \mathbb{R})),$$
$$y'(\phi) \in \mathfrak{L}(X; X \times U), \quad y'^*(\phi) \in \mathfrak{L}\left(X' \times U'; X'\right).$$

In this paper, $\mathfrak{L}(X;U)$ denotes the set of all bounded linear operators from X to U and $\circ$ is the composition operator. We rewrite (7) as below. Firstly, let's write the admissible set of (ϕ, u) with respect to the equality constraint as

$$S = \{(\phi, u) \in \mathscr{D} \times \mathscr{S} \mid s(\phi, u) = 0_{U'}\}. \tag{8}$$

For $y(\phi) = (\phi, u) \in S$, we denote the kernel of $s'(\phi, u) \in \mathfrak{L}(X \times U; U')$ by

$$T_S(\phi, u) = \left\{(\varphi, \hat{\upsilon}) \in X \times U \;\middle|\; s'(\phi, u)[\varphi, \hat{\upsilon}] = 0_{U'}\right\} \tag{9}$$

and the space orthogonal to $T_S(\phi, u)$ as

$$T_S'(\phi, u) = \left\{(\psi, w) \in X' \times U' \;\middle|\; \langle(\varphi, \hat{\upsilon}), (\psi, w)\rangle = 0 \; \forall (\varphi, \hat{\upsilon}) \in T_S(\phi, u)\right\}.$$

Moreover, the relationship between $T_S(\phi, u)$ and the Fréchet derivative $y'(\phi)[\varphi]$ of $y(\phi) \in S$ with respect to an arbitrary variation $\varphi \in X$ can be obtained in the following way. If we take the Fréchet derivative on both sides of $s(\phi, u) = 0_{U'}$ with respect to $\varphi \in X$, then we get

$$s'(\phi, u) \circ y'(\phi)[\varphi] = 0_{U'} \quad \forall \varphi \in X. \tag{10}$$

This relationship shows that the image space $\operatorname{Im} y'(\phi)$ of $y'(\phi)$ is actually the kernel space $\operatorname{Ker} s'(\phi, u)$ of $s'(\phi, u)$. In other words, the following relationship is established:

$$T_S(\phi, u) = \operatorname{Im} y'(\phi). \tag{11}$$

We use the relationship above to rewrite (7). When ϕ is a local minimizer, $g_i(\phi, \upsilon(\phi))$ needs to be orthogonal to an arbitrary $(\varphi, v_i) \in T_S(\phi, u)$. Hence,

$$g_i(\phi, \upsilon(\phi)) \in T_S'(\phi, u). \tag{12}$$

Now, using the relationship between the orthogonal complement space of the image space and the kernel space, we get that $T_S'(\phi, u) = \operatorname{Im} s'^*(\phi, u)$ where $s'^*(\phi, u) \in \mathfrak{L}(U; X' \times U')$ is the adjoint of $s'(\phi, u)$. Therefore, (12) means that we can find an element v_i of U such that

$$f_i'(\phi, u)[\varphi, \hat{u}] + \left\langle s'(\phi, u)[\varphi, \hat{u}], v_i\right\rangle = 0 \quad \forall (\varphi, \hat{u}) \in X \times U.$$

This established equation (4). Moreover, (5) holds if u is the solution of (1). (QED)

From Theorem 2.1, g_i can be evaluated in the following way. Let u be determined as satisfying (5). This means u is the solution of Problem 2.1. Moreover, let v_i be determined through the equation

$$\mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}] = \langle f_{iu}(\phi,u) - \tau^*(\phi)v_i, \hat{u}\rangle = 0 \quad \forall \hat{u} \in U. \tag{13}$$

This problem to determine v_i by (13) is called the adjoint problem of Problem 2.1 with respect to f_i. When we use the solutions u and v_i, $g_i \in X'$ is obtained through

$$\langle g_i(\phi,u,v_i), \varphi\rangle = \left\langle f_{i\phi}(\phi,u), \varphi\right\rangle + \left\langle s_\phi(\phi,u)[\varphi], v_i\right\rangle \quad \forall \varphi \in X. \tag{14}$$

2.2 *Hessian of cost function f_i with respect to variation of ϕ*

Furthermore, when f_i and s are the second-order Fréchet differentiable, then we may calculate the second-order partial Fréchet derivative of $\mathscr{L}_i$ with respect to arbitrary variations $(\varphi_1, \hat{\upsilon}_1), (\varphi_2, \hat{\upsilon}_2) \in T_S(\phi,u)$ of $(\phi,u) \in S$ as

$$\begin{aligned}&\mathscr{L}_{i(\phi,u),(\phi,u)}(\phi,u,v_i)[(\varphi_1,\hat{\upsilon}_1),(\varphi_2,\hat{\upsilon}_2)]\\&\quad = f_i''(\phi,u)[(\varphi_1,\hat{\upsilon}_1),(\varphi_2,\hat{\upsilon}_2)] + \left\langle s''(\phi,u)[(\varphi_1,\hat{\upsilon}_1),(\varphi_2,\hat{\upsilon}_2)], v_i\right\rangle.\end{aligned} \tag{15}$$

Using it, we have the following result.

Theorem 2.2 (The second-order necessary condition of a local minimizer) *Let f_i and s be elements of $C^2(\mathscr{D}\times\mathscr{S};\mathbb{R})$ and $C^2(\mathscr{D}\times\mathscr{S};U')$, respectively. If (ϕ,u) is a local minimizer of Problem 2.3, the following holds:*

$$\mathscr{L}_{i(\phi,u),(\phi,u)}(\phi,u,v_i)[(\varphi,\hat{\upsilon}),(\varphi,\hat{\upsilon})] \geq 0 \quad \forall(\varphi,\hat{\upsilon}) \in T_S(\phi,u). \tag{16}$$

Proof. In the proof of Theorem 2.1, the assumption for the implicit function theorem is replaced by $s(\phi,\cdot) \in C^2(B_U;U')$, and then using $\upsilon(\phi)$ in (6), $y(\phi) = (\phi,\upsilon(\phi)) \in C^2(\mathscr{D};X\times U)$ is determined. From (10), we have

$$s''(\phi,u)\left[y'(\phi)[\varphi], y'(\phi)[\varphi]\right] = 0_{U'} \tag{17}$$

with respect to $y'(\phi)[\varphi] \in T_S(\phi,u)$. Hence, if (ϕ,u) is a local minimizer of Problem 2.3,

$$\mathscr{L}_{i(\phi,u),(\phi,u)}(\phi,u,v_i)\left[y'(\phi)[\varphi], y'(\phi)[\varphi]\right] = \tilde{f}_i''(\phi)[\varphi,\varphi] \geq 0 \tag{18}$$

holds with respect to $y'(\phi)[\varphi] \in T_S(\phi,u)$. (QED)

Since the left-hand side of (18) is the Hessian of f_i with respect to an arbitrary variation $\varphi \in X$ of ϕ, we write it as $h_i(\phi,u,v_i)[\varphi,\varphi]$, and obtain the following result.

Theorem 2.3 (The second-order sufficient condition for a local minimizer) *Under the assumptions of Theorem 2.2, if* (4) *and* (5) *are satisfied at* $(\phi, u, v_i) \in \mathscr{D} \times \mathscr{S}^2$ *and* (16) *holds, then* (ϕ, u) *is a local minimizer of Problem 2.3.*

Proof. When $(\phi, u, v_i) \in \mathscr{D} \times \mathscr{S}^2$ is a stationary point of $\mathscr{L}_i$ in S, with respect to an arbitrary point $y(\phi + \varphi) = y(\phi) + z(\varphi)$ in a neighborhood $B \subset S$ of $y(\phi) = (\phi, u)$, there exists a $\theta \in (0,1)$ satisfying

$$\tilde{f}_i(\phi + \varphi) - \tilde{f}_i(\phi) = \frac{1}{2}\mathscr{L}_i''(\phi + \theta\varphi, u, v_i)[z(\varphi), z(\varphi)] \quad \forall y(\phi) + z(\varphi) \in B.$$

From the assumption, since the right-hand side is greater than or equal to 0, $\tilde{f}_i(\phi) \leq \tilde{f}_i(\phi + \varphi)$ holds. (QED)

In view of Theorems 2.2 and 2.3, h_i is calculated as

$$\begin{aligned} h_i(\phi, u, v_i)[\varphi_1, \varphi_2] &= \left(\mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi, u, v_i)[\hat{\upsilon}_1]\right)_\phi [\varphi_2] \\ &\quad + \left(\mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi, u, v_i)[\hat{\upsilon}_1]\right)_u [\hat{\upsilon}_2] \\ &= \mathscr{L}_{i\phi\phi}(\phi, u, v_i)[\varphi_1, \varphi_2] + \mathscr{L}_{iu\phi}(\phi, u, v_i)[\hat{\upsilon}_1, \varphi_2] \\ &\quad + \mathscr{L}_{i\phi u}(\phi, u, v_i)[\varphi_1, \hat{\upsilon}_2] + \mathscr{L}_{iuu}(\phi, u, v_i)[\hat{\upsilon}_1, \hat{\upsilon}_2], \end{aligned} \tag{19}$$

where, in order that $(\varphi_j, \hat{\upsilon}_j) \in T_S(\phi, u)$ for $j \in \{1,2\}$, $\hat{\upsilon}_j = \upsilon'(\phi)[\varphi_j]$ has to be determined using the equation

$$\mathscr{L}_{\mathrm{S}(\phi,u)}(\phi, u, v)[\varphi_j, \hat{\upsilon}_j] = 0 \quad \forall \varphi_j \in X. \tag{20}$$

3 Shape optimization problem of linear elastic body

Based on the results above, we illustrate how to find a solution to a shape optimization problem using Hessians. In particular, we consider a shape optimization problem of a d-dimensional linear elastic body ($d \in \{2,3\}$). In such a case of shape optimization problem of domain variation type, we set a domain variation (displacement) ϕ from an initial domain Ω_0 as a design variable, and denote the varied domain with $\Omega(\phi)$. Similarly, we denote $(\cdot)(\phi)$ as the varied. The function spaces for ϕ are defined as $X = \{\phi \in H^1(\mathbb{R}^d; \mathbb{R}^d) \mid \phi = \mathbf{0}_{\mathbb{R}^d} \text{ on } \bar{\Omega}_{\mathrm{C}0}\}$, where $\bar{\Omega}_{\mathrm{C}0}$ expresses a part of boundary or domain to fix while the domain varies, $Y = X \cap W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d)$ and $\mathscr{D} = \{\phi \in Y \mid \Omega_0 \to \Omega(\phi) \text{ is a bijection}\}$. Let $\mathbf{u} : \Omega(\phi) \to \mathbb{R}^d$ be an elastic displacement, $\mathbf{E}(\mathbf{u}) = \{\nabla\mathbf{u}^{\mathrm{T}} + (\nabla\mathbf{u}^{\mathrm{T}})^{\mathrm{T}}\}/2$ be a linear strain, $\mathbf{S}(\mathbf{u}) = \mathbf{C}(\phi)\mathbf{E}(\mathbf{u})$ be a stress, and $\mathbf{C}(\phi) : \Omega(\phi) \to \mathbb{R}^{d\times d\times d\times d}$ denote a stiffness. We use ν as the normal.

Problem 3.1 (Linear elastic problem) *For* $\phi \in \mathscr{D}$, $\mathbf{b}(\phi)$, $\mathbf{p}_{\mathrm{N}}(\phi)$, $\mathbf{u}_{\mathrm{D}}(\phi) : \mathbb{R}^d \to \mathbb{R}^d$ *and* $\mathbf{C}(\phi) : \mathbb{R}^d \to \mathbb{R}^{d\times d\times d\times d}$, *find* $\mathbf{u} : \Omega(\phi) \to \mathbb{R}^d$ *such that*

$$-\nabla^{\mathrm{T}}\mathbf{S}(\phi,\mathbf{u}) = \mathbf{b}^{\mathrm{T}}(\phi) \quad \text{in } \Omega(\phi), \quad \mathbf{S}(\phi,\mathbf{u})\,\nu = \mathbf{p}_{\mathrm{N}}(\phi) \quad \text{on } \Gamma_p(\phi),$$
$$\mathbf{S}(\phi,\mathbf{u})\,\nu = \mathbf{0}_{\mathbb{R}^d} \quad \text{on } \Gamma_{\mathrm{N}}(\phi)\setminus\bar{\Gamma}_p(\phi), \quad \mathbf{u} = \mathbf{u}_{\mathrm{D}}(\phi) \quad \text{on } \Gamma_{\mathrm{D}}(\phi).$$

When the given functions are given appropriately, $\mathbf{u}-\mathbf{u}_{\mathrm{D}}$ is found uniquely in $\mathscr{S} = U \cap W^{1,\infty}\left(\Omega(\phi);\mathbb{R}^d\right)$ with $U = \{\mathbf{u} \in H^1\left(\Omega(\phi);\mathbb{R}^d\right) \mid \mathbf{u} = \mathbf{0}_{\mathbb{R}^d} \text{ on } \Gamma_{\mathrm{D}}(\phi)\}$. Using ϕ and $\mathbf{u}$, we define an objective cost function f_0, which is called a mean compliance, and a constraint cost function f_1 for domain measure as specified below, and set a shape optimization problem in Problem 3.2.

$$f_0(\phi,\mathbf{u}) = \int_{\Omega(\phi)} \mathbf{b}\cdot\mathbf{u}\,\mathrm{d}x + \int_{\Gamma_p(\phi)} \mathbf{p}_{\mathrm{N}}\cdot\mathbf{u}\,\mathrm{d}\gamma - \int_{\Gamma_{\mathrm{D}}(\phi)} \mathbf{u}_{\mathrm{D}}\cdot(\mathbf{S}(\mathbf{u})\,\nu)\,\mathrm{d}\gamma, \tag{21}$$
$$f_1(\phi) = \int_{\Omega(\phi)} \mathrm{d}x - c_1. \tag{22}$$

Problem 3.2 (Mean compliance minimization) *Find $\Omega(\phi)$ satisfying*

$$\min_{(\phi,\mathbf{u}-\mathbf{u}_{\mathrm{D}})\in\mathscr{D}\times\mathscr{S}} \{ f_0(\phi,\mathbf{u}) \mid f_1(\phi) \le 0, \text{ Problem 3.1} \}.$$

3.1 Shape gradients of cost functions

The shape gradient $\mathbf{g}_0$ of f_0 can be obtained using the relations of (13) and (14), where the Lagrange function $\mathscr{L}_{\mathrm{S}}$ for Problem 3.1 is given as

$$\mathscr{L}_{\mathrm{S}}(\phi,\mathbf{u},\mathbf{v}_0) = l(\phi)(\mathbf{v}_0) - a(\phi)(\tilde{\mathbf{u}},\mathbf{v}_0),$$

where $\tilde{\mathbf{u}} = \mathbf{u}-\mathbf{u}_{\mathrm{D}}, \mathbf{v}_0 \in U$ and

$$a(\phi)(\tilde{\mathbf{u}},\mathbf{v}_0) = \int_{\Omega(\phi)} \mathbf{S}(\tilde{\mathbf{u}})\cdot\mathbf{E}(\mathbf{v}_0)\,\mathrm{d}x,$$
$$l(\phi)(\mathbf{v}_0) = \int_{\Omega(\phi)} \mathbf{b}\cdot\mathbf{v}_0\mathrm{d}x + \int_{\Gamma_p(\phi)} \mathbf{p}_{\mathrm{N}}\cdot\mathbf{v}_0\,\mathrm{d}\gamma - a(\phi)(\mathbf{u}_{\mathrm{D}},\mathbf{v}_0).$$

From the adjoint problem corresponding to (13), the self adjoint condition $\mathbf{v}_0 = \mathbf{u}$ is obtained. Here, we assume $\mathbf{b}(\phi)$, $\mathbf{p}_{\mathrm{N}}(\phi)$, $\mathbf{u}_{\mathrm{D}}(\phi)$ and $\mathbf{C}(\phi)$ vary with domain variation. Then, using Propositions A.1 and A.2 in [2] (Propositions 9.3.4 and 9.3.7 in [1]), we have

$$\mathscr{L}_{0\phi}(\phi,\mathbf{u},\mathbf{v}_0)[\varphi] = \langle \mathbf{g}_0,\varphi\rangle = \int_{\Omega(\phi)} \left(\mathbf{G}_{\Omega 0}\cdot\nabla\varphi^{\mathrm{T}} + g_{\Omega 0}\nabla\cdot\varphi\right)\mathrm{d}x$$
$$+ \int_{\Gamma_p(\phi)} \mathbf{g}_{p0}\cdot\varphi\,\mathrm{d}\gamma + \int_{\partial\Gamma_p(\phi)\cup\Theta(\phi)} \mathbf{g}_{\partial p0}\cdot\varphi\,\mathrm{d}\varsigma \quad \forall\varphi\in Y,$$

where

$$\mathbf{G}_{\Omega 0} = 2\mathbf{S}(\mathbf{u})\left(\nabla \mathbf{u}^{\mathrm{T}}\right)^{\mathrm{T}}, \quad g_{\Omega 0} = -\mathbf{S}(\mathbf{u}) \cdot \mathbf{E}(\mathbf{u}) + 2\mathbf{b} \cdot \mathbf{u},$$
$$\mathbf{g}_{p0} = 2\kappa(\mathbf{p}_{\mathrm{N}} \cdot \mathbf{u})\nu, \quad \mathbf{g}_{\partial p0} = 2(\mathbf{p}_{\mathrm{N}} \cdot \mathbf{u})\tau.$$

$\Theta(\phi)$ is the set of corner points in $\Gamma_p(\phi)$ for $d = 2$ and the vertices and edges for $d = 3$, τ denotes the tangent, and $\kappa = \nabla \cdot \nu$. We can confirm that $\mathbf{g}_0 \in X'$ when $\mathbf{u} \in \mathscr{S}$ (Theorem 9.7.2 in [1]).

On the other hand, the shape derivative of f_1 is obtained as

$$f_1'(\phi)[\varphi] = \langle \mathbf{g}_1, \varphi \rangle = \int_{\Omega(\phi)} \nabla \cdot \varphi \, \mathrm{d}x \quad \forall \varphi \in Y.$$

3.2 Shape Hessians of cost functions

The shape Hessian h_0 of f_0 is obtained using (19) and (20). In order to get a Hessian form, we assume thats $\mathbf{b} = \mathbf{0}_{\mathbb{R}^d}$ and Γ_{p0} is a fixed boundary, i.e. $\Gamma_{p0} \subset \bar{\Omega}_{\mathrm{C}0}$. In this case, (19) becomes

$$\begin{aligned} h_0(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_1, \varphi_2] &= \mathscr{L}_{0\phi\phi}(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_1, \varphi_2] + \mathscr{L}_{0\phi'\mathbf{u}}(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_1, \hat{\upsilon}_2] \\ &\quad + \mathscr{L}_{0\phi'\mathbf{u}}(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_2, \hat{\upsilon}_1] + \mathscr{L}_{0\mathbf{uu}}(\phi, \mathbf{u}, \mathbf{v}_0)[\hat{\upsilon}_1, \hat{\upsilon}_2]. \end{aligned} \tag{23}$$

Corresponding to (20), using Proposition A.1 and A.2 in [2], for $j \in \{1,2\}$, we get

$$\begin{aligned} &\mathscr{L}_{\mathrm{S}(\phi,\mathbf{u})}(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_j, \hat{\upsilon}_j] = \int_{\Omega(\phi)} \Big\{ \mathbf{S}(\mathbf{u}) \cdot \left(\nabla \varphi_j^{\mathrm{T}} \nabla \mathbf{v}_0^{\mathrm{T}}\right)^{\mathrm{s}} + \mathbf{S}(\mathbf{v}_0) \cdot \left(\nabla \varphi_j^{\mathrm{T}} \nabla \mathbf{u}^{\mathrm{T}}\right)^{\mathrm{s}} \\ &\quad - (\mathbf{S}(\mathbf{u}) \cdot \mathbf{E}(\mathbf{v}_0)) \nabla \cdot \varphi_j - \mathbf{S}(\hat{\upsilon}_j) \cdot \mathbf{E}(\mathbf{v}_0) \Big\} \mathrm{d}x \\ &= \int_{\Omega(\phi)} \left[\left\{ \left(\nabla \varphi_j^{\mathrm{T}}\right)^{\mathrm{T}} \mathbf{S}(\mathbf{u}) + \mathbf{C}\left(\nabla \varphi_j^{\mathrm{T}} \nabla \mathbf{u}^{\mathrm{T}}\right)^{\mathrm{s}} - \mathbf{S}(\mathbf{u}) \nabla \cdot \varphi_j - \mathbf{S}(\hat{\upsilon}_j) \right\} \left(\nabla \mathbf{v}_0^{\mathrm{T}}\right)^{\mathrm{T}} \right] \cdot \mathbf{I} \, \mathrm{d}x \\ &= 0. \end{aligned} \tag{24}$$

Here, $(\,\cdot\,)^{\mathrm{s}} = \{(\,\cdot\,) + (\,\cdot\,)^{\mathrm{T}}\}/2$, the unit matrix $\mathbf{I} \in \mathbb{R}^{d \times d}$ and the Dirichlet boundary conditions for $\mathbf{v}_0$ and $\hat{\upsilon}_j = \upsilon'(\phi)[\varphi_j]$ were used. From (24), the identity

$$\mathbf{S}(\hat{\upsilon}_j)\left(\nabla \mathbf{v}_0^{\mathrm{T}}\right)^{\mathrm{T}} = \left\{ \left(\nabla \varphi_j^{\mathrm{T}}\right)^{\mathrm{T}} \mathbf{S}(\mathbf{u}) + \mathbf{C}\left(\nabla \varphi_j^{\mathrm{T}} \nabla \mathbf{u}^{\mathrm{T}}\right)^{\mathrm{s}} - \nabla \cdot \varphi_j \mathbf{S}(\mathbf{u}) \right\} \left(\nabla \mathbf{v}_0^{\mathrm{T}}\right)^{\mathrm{T}} \tag{25}$$

is obtained. Moreover, since (24) can be written as

$$\begin{aligned} &\mathscr{L}_{\mathrm{S}(\phi,\mathbf{u})}(\phi, \mathbf{u}, \mathbf{v}_0)[\varphi_j, \hat{\upsilon}_j] \\ &= \int_{\Omega(\phi)} \left[\nabla \mathbf{v}_0^{\mathrm{T}} \mathbf{S}(\mathbf{u}) \nabla \varphi_j^{\mathrm{T}} + \mathbf{S}(\mathbf{v}_0) \left\{ \left(\nabla \mathbf{u}^{\mathrm{T}}\right)^{\mathrm{T}} \left(\left(\nabla \varphi_j^{\mathrm{T}}\right)^{\mathrm{T}} - \nabla \cdot \varphi_j \right) - \left(\nabla \hat{\upsilon}_j^{\mathrm{T}}\right)^{\mathrm{T}} \right\} \right] \cdot \mathbf{I} \, \mathrm{d}x \\ &= 0, \end{aligned}$$

we have another relation as

$$\mathbf{S}(\mathbf{v}_0)\left(\nabla\hat{v}_j^{\mathrm{T}}\right)^{\mathrm{T}} = \nabla\mathbf{v}_0^{\mathrm{T}}\mathbf{S}(\mathbf{u})\nabla\varphi_j^{\mathrm{T}} + \mathbf{S}(\mathbf{v}_0)\left(\nabla\mathbf{u}^{\mathrm{T}}\right)^{\mathrm{T}}\left\{\left(\nabla\varphi_j^{\mathrm{T}}\right)^{\mathrm{T}} - \nabla\cdot\varphi_j\right\}. \tag{26}$$

Substituting (25) and (26) into (23) and then using the self-adjoint property and the identity $h_0(\phi,\mathbf{u},\mathbf{v}_0)[\varphi_1,\varphi_2] = h_0(\phi,\mathbf{u},\mathbf{v}_0)[\varphi_2,\varphi_1]$, the shape Hessian of f_0 is obtained using Proposition A.1 and A.2 in [2] as

$$\begin{aligned} h_0(\phi,\mathbf{u},\mathbf{v}_0)[\varphi_1,\varphi_2] = \int_{\Omega(\phi)} \Big[& \mathbf{S}(\mathbf{u})\cdot\mathbf{E}(\mathbf{u})\left\{\left(\nabla\varphi_2^{\mathrm{T}}\right)^{\mathrm{T}}\cdot\nabla\varphi_1^{\mathrm{T}} + (\nabla\cdot\varphi_2)(\nabla\cdot\varphi_1)\right\} \\ & + \left(\nabla\mathbf{u}^{\mathrm{T}}\mathbf{S}(\mathbf{u})\right)\cdot\left\{\nabla\varphi_1^{\mathrm{T}}\left(\nabla\varphi_2^{\mathrm{T}}\right)^{\mathrm{T}} + \nabla\varphi_2^{\mathrm{T}}\left(\nabla\varphi_1^{\mathrm{T}}\right)^{\mathrm{T}}\right\} \\ & - 2\left(\mathbf{S}(\mathbf{u})\mathbf{E}(\mathbf{u})\right)\cdot\left\{\nabla\varphi_2^{\mathrm{T}}(\nabla\cdot\varphi_1) + \nabla\varphi_1^{\mathrm{T}}(\nabla\cdot\varphi_2)\right\}\Big]\,\mathrm{d}x \quad \forall\varphi_1,\varphi_2\in X. \end{aligned}$$

On the other hand, the shape Hessian of f_1 is given by

$$h_1(\phi)[\varphi_1,\varphi_2] = \int_{\Omega(\phi)}\left\{-\left(\nabla\varphi_2^{\mathrm{T}}\right)^{\mathrm{T}}\cdot\nabla\varphi_1^{\mathrm{T}} + (\nabla\cdot\varphi_2)(\nabla\cdot\varphi_1)\right\}\mathrm{d}x \quad \forall\varphi_1,\varphi_2\in X.$$

4 Solutions to shape optimization problems

Finally, we show solutions to shape optimization problems. In this section, we consider a shape optimization problem similar to Problem 3.2, but the number of constraints is changed to m. To distinguish it from Problem 3.2, we denote the problem as Problem 3.2 (m), and write its Lagrange function as

$$\mathscr{L}(\phi,\mathbf{u},\mathbf{v}_i) = f_0(\phi,\mathbf{u}) + \sum_{i\in\{1,\ldots,m\}}\lambda_i f_i(\phi,\mathbf{u}). \tag{27}$$

The H^1 gradient method of domain variation type was formulated by seeking $\varphi_{gi}\in X$ that decrease $f_i(\phi_k,\mathbf{u})$ with respect to iterations $k\in\{0,1,\ldots\}$ by the following methods.

Problem 4.1 (H^1 gradient method for f_i) *Let $a_X : X\times X\to\mathbb{R}$ be a bounded and coercive bilinear form in X, and c_a be a positive constant to control the magnitude of φ_{gi}. For $\mathbf{g}_i(\phi_k)\in X'$, find $\varphi_{gi}\in X$ such that*

$$c_a a_X(\varphi_{gi},\psi) = -\langle\mathbf{g}_i,\psi\rangle \quad \forall\psi\in X. \tag{28}$$

As an example, we use

$$a_X(\varphi,\psi) = \int_{\Omega(\phi)}\left\{\left(\nabla\varphi^{\mathrm{T}}\right)\cdot\left(\nabla\psi^{\mathrm{T}}\right) + c_\Omega\varphi\cdot\psi\right\}\mathrm{d}x,$$

where c_Ω is a positive constant. An algorithm to solve the generalized shape optimization problem was shown in the previous paper [2] (Section 9.9.1 in [1]).

In addition, given the $\mathbf{g}_i$ and the computed shape Hessian h_i of f_i for all $i \in \{0, 1, \ldots, m\}$, a Newton method in H^1 class function space (H^1 Newton method) can be defined as follows (Problem 9.8.8 in [1]).

Problem 4.2 (H^1 Newton method for domain variation type for f_i) *Let $\mathbf{g}_i \in X'$ and $h_i \in \mathfrak{L}^2(X \times X; \mathbb{R})$ for each $i \in \{0, 1, \ldots, m\}$ at $\phi_k \in \mathscr{D}$ be given, and*

$$h_{\mathscr{L}}(\phi_k)[\varphi_1, \varphi_2] = h_0(\phi_k)[\varphi_1, \varphi_2] + \sum_{i \in \{1,\ldots,m\}} \lambda_{ik} h_i(\phi_k)[\varphi_1, \varphi_2] \quad \forall \varphi_1, \varphi_2 \in X$$

be the shape Hessian of $\mathscr{L}$ in (27)*. Moreover, let $a_X : X \times X \to \mathbb{R}$ be a bilinear form to compensate coerciveness and regularity of $h_{\mathscr{L}}(\phi_k)$ in X, and c_a be a positive constant for adjustment. Find $\varphi_{gi} \in X$ such that*

$$h_{\mathscr{L}}(\phi_k)[\varphi_{gi}, \psi] + c_a a_X(\varphi_{gi}, \psi) = -\langle g_i(\phi_k), \psi \rangle \quad \forall \psi \in X. \tag{29}$$

A simple algorithm for solving Problem 3.2 (m) by the H^1 Newton method is shown below (Section 9.9.2 in [1]).

Algorithm 4.1 (The H^1 Newton method for shape optimization problems)
Obtain the local minimizer of Problem 3.2 (m) in the following way.

1. *Set Ω_0 and $\phi_0 = \mathbf{i}$ (the identity mapping) as $f_1(\phi_0, \mathbf{u}) \leq 0$, …, $f_m(\phi_0, \mathbf{u}) \leq 0$. Set c_a, ε_0, ε_1, …, ε_m appropriately. Set $k = 0$.*
2. *Solve the state determination problem (Problem 3.1) at ϕ_k, and compute $f_0(\phi_k, \mathbf{u})$, $f_1(\phi_k, \mathbf{u})$, …, $f_m(\phi_k, \mathbf{u})$. Set $I_{\mathrm{A}}(\phi_k) = \{ i \in \{1, \ldots, m\} \mid f_i(\phi_k, \mathbf{u}) \geq -\varepsilon_i \}$.*
3. *Do the following when the Hessian of constraint function can be calculated.*
   - *Solve the adjoint problem with respect to f_0, f_{i_1}, …, $f_{i_{|I_{\mathrm{A}}|}}$ and calculate $\mathbf{g}_0$, $\mathbf{g}_{i_1}$, …, $\mathbf{g}_{i_{|I_{\mathrm{A}}|}}$ at ϕ_k.*
   - *Use* (28) *to solve φ_{g0}, φ_{gi_1}, …, $\varphi_{gi_{|I_{\mathrm{A}}|}}$.*
   - *Solve λ_{k+1} using*

   $$\left(\langle \mathbf{g}_i, \varphi_{gj} \rangle\right)_{(i,j) \in I_{\mathrm{A}}^2(\phi_k)} \left(\lambda_{jk+1}\right)_{j \in I_{\mathrm{A}}(\phi_k)} = -\left(f_i(\phi_k) + \langle \mathbf{g}_i, \varphi_{g0} \rangle\right)_{i \in I_{\mathrm{A}}(\phi_k)}. \tag{30}$$

   If $I_{\mathrm{I}}(\phi_k) = \{ i \in I_{\mathrm{A}}(\phi_k) \mid \lambda_{ik+1} < -\varepsilon_i \}$ is not empty set, replace $I_{\mathrm{A}}(\phi_k) \setminus I_{\mathrm{I}}(\phi_k)$ with $I_{\mathrm{A}}(\phi_k)$ and solve (30) *until $I_{\mathrm{I}}(\phi_k) = \emptyset$.*
4. *Solve adjoint problems at ϕ_k, and compute $\mathbf{g}_0$, $\mathbf{g}_{i_1}$, …, $\mathbf{g}_{i_{|I_{\mathrm{A}}|}}$ and $h_0, h_{i_1}, \ldots, h_{i_{|I_{\mathrm{A}}|}}$.*
5. *Solve φ_{g0}, φ_{gi_1}, …, $\varphi_{gi_{|I_{\mathrm{A}}|}}$ using* (29)*.*
6. *Solve λ_{k+1} using* (30)*. When $I_{\mathrm{I}}(\phi_k) \neq \emptyset$, replace $I_{\mathrm{A}}(\phi_k) \setminus I_{\mathrm{I}}(\phi_k)$ with $I_{\mathrm{A}}(\phi_k)$, and resolve* (30) *until $I_{\mathrm{I}}(\phi_k) = \emptyset$.*
7. *Compute φ_g using*

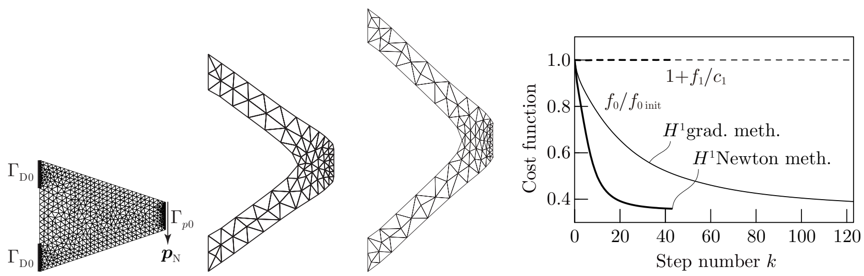
$$\varphi_g = \varphi_g(\lambda_{k+1}) = \varphi_{g0} + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{ik+1} \varphi_{gi}, \tag{31}$$

set $\phi_{k+1} = \phi_k + \varphi_g$, *and compute* $f_0(\phi_{k+1}, \mathbf{u})$, $f_1(\phi_{k+1}, \mathbf{u})$, ..., $f_m(\phi_{k+1}, \mathbf{u})$. *Set* $I_{\mathrm{A}}(\phi_{k+1}) = \{ i \in \{1, \dots, m\} \mid f_i(\phi_{k+1}, \mathbf{u}) \geq -\varepsilon_i \}$.

8. *Assess* $|f_0(\phi_{k+1}, \mathbf{u}) - f_0(\phi_k, \mathbf{u})| \leq \varepsilon_0$.
   - *If "Yes," proceed to 9.*
   - *If "No," replace* $k+1$ *with* k *and return to 4.*
9. *Stop the algorithm.*

5 Numerical example

Figure 1 shows numerical results to Problem 3.2 obtained by the program[1] written by Kento Furuki and the author with FreeFEM++[2].



(a) Initial shape (b) H^1 grad. m. (c) H^1 Newton m. (d) Iteration history of cost functions

Fig. 1 Numerical results for two dimensional linear elastic body

Acknowledgements The present study was supported by JSPS KAKENHI (16K05285).

References

1. Azegami, H.: Shape Optimization Problems (in Japanese). Morikita Publishing, Tokyo (2016)
2. Azegami, H.: Solution of Shape Optimization Problem and Its Application to Product Design, pp. 83–98. Mathematics for Industry Volume 26 2017. Springer Singapore (2016)
3. Zeidler, E.: Linear monotone operators. Springer-Verlag, New York ; Tokyo (1990)

[1] http://www.morikita.co.jp/books/mid/061461.

[2] http://www.freefem.org/ff++/