Contents

Chapter 7

Abstract Optimum Design Problem

We have seen in Chaps. 5 and 6 how boundary problems of partial differential equations are constructed and how they are solved. If we compare them to the optimum design problems looked at in Chap. 1, they correspond to state determination problems. From this chapter, we finally start thinking about optimum design problems targeting the shape or topology of a domain in which the boundary value problem is defined. In this chapter, we construct an abstract problem common to both problems, and explore the ways to solve them.

In Chap. 1, the basics of the optimum design problems were looked at. In those problems considered in the chapter, the linear spaces for design variable and the state variable were both finite-dimensional vector spaces. In this chapter, such vector spaces will be extended to function spaces. Moreover, a state determination problem is replaced by an abstract variation problem. In this case the following points should be noted. The dual space of a finite-dimensional vector space was the same finite-dimensional vector space. Therefore, the derivative of the cost function with respect to a variation of the design variable is an element of the same finite-dimensional vector space as that of the design variable. However, if a function space is selected as a vector space, its dual space generally becomes a different vector space. In this chapter, there is a need to be careful of this. However, other than this, the same results as the conditions satisfied by the optimal solution in Chap. 1 should be attainable.

Furthermore, with respect to numerical solutions (algorithms), the abstract gradient method and abstract Newton method are defined by expanding the gradient method and the Newton method shown in Chap. 3. Using their solutions, the gradient method and Newton method with respect to constrained problems can be considered similarly to those shown in Chap. 3. In these cases, the numerical solutions of an abstract optimum design problem can be constructed using the algorithms shown in Chap. 3 just by replacing the corresponding terms.

Hence, if the content of this chapter is understood, the remaining issue in the shape or topology optimization problem is specifying the admissible set or function spaces with respect to these problems within the framework of an abstract optimum design problem. In this case, clarifying the method of calculation of the Fréchet derivative with respect to a variation of design variable, and confirming that the solution with the abstract gradient method or abstract Newton method using them is included in the admissible set of the design variable become the focal points. These will be looked at for each problem in Chaps. 8 and 9.

7.1 Linear Spaces of Design Variables

Think about making the optimum design problem abstract by remembering the optimum design problem seen in Chap. 1. Here, let us use the mean compliance minimization problem (Problem 1.1.4) of the stepped one-dimensional linear elastic body in order to look at its correspondence with the abstract optimum design problem.

In Problem 1.1.4, the linear space of design variables was set to be $X = \mathbb{R}^2$ and the linear space of state variables to be $U = \mathbb{R}^2$. In this chapter, X and U can be also be function spaces. In this case, if the element of X is given, we assume that a state determination problem can be constructed, the state variables can be determined as an element of U, and cost function (functional) defined on $X \times U$ can also be calculated. Here, if solutions using the gradient or Hessian such as those looked at in Chap. 3 are considered, there is a need for X and U to be a function space in which the Fréchet derivative can be defined. Furthermore, if solutions via the abstract gradient method (Problem 7.6.1) or abstract Newton method (Problem 7.6.4) to be shown later are to be considered, X needs to be a real Hilbert space. Hence, in this chapter, we will assume X in the following way.

In the optimization problems shown in Chaps. 8 and 9, the domains over which the boundary problems of partial differential equations are defined are the scope of the design. In this case, as shown in Chap. 5, in order for the boundary problem of a partial differential equation to be defined, the domain needs to have at least a Lipschitz boundary (Section A.5). In Chaps. 8 and 9, the functions representing the density or domain variation are chosen to be the design variables. In this case, if the Lipschitz boundary is to be defined using these functions, the function space Y for the functions needs to be of class $C^{0,1}$. Moreover, to show the existence of the optimum solution in Sect. 7.4, Y needs to be compactly embedded in X ($\mathcal{D} \Subset X$). In fact, in Chap. 8, for the $d \in \{2, 3\}$-dimensional bounded domain D, the function spaces X and Y will be chosen as $H^1(D; \mathbb{R})$ and $H^2(D; \mathbb{R}) \cap C^{0,1}(D; \mathbb{R})$, respectively. Similarly, in Chap. 9, for a bounded domain $D \subset \mathbb{R}^d$, X and Y will be defined as $H^1(D; \mathbb{R}^d)$ and $H^2(D; \mathbb{R}^d) \cap C^{0,1}(D; \mathbb{R}^d)$, respectively. In relation to this, $Y \Subset X$ is guaranteed by the Rellich–Kondrachov compact embedding theorem (Theorem 4.4.15). Furthermore, the admissible set $\mathcal{D}$ of the design variables will be defined as sets satisfying additional conditions.

Hence, in this chapter, we shall denote the design variable as ϕ, which is an element of $\mathcal{D} \subset Y \Subset X$. Furthermore, since the bounded conditions correspond to the side constraint in Chap. 1, after this chapter, we assume that ϕ is an interior point of $\mathcal{D}$ ($\phi \in \mathcal{D}^\circ$) when considering a gradient method or Newton method. When some of the side constraints are activated, we include them in the inequality constraints. Moreover, we will assume a variation of the design variable as $\varphi \in X$ ($\varphi \in Y$ in Chap. 9) and define the Fréchet derivatives of functions or functionals with respect to an arbitrary $\varphi \in X$ as an element in the dual space X' of X (Definition 4.4.5).

7.2 State Determination Problem

In the optimum design problem of Problem 1.1.4, the state variable was defined by $\boldsymbol{u}$ and constructed so that it can be uniquely determined as a solution of the state determination problem (Problem 1.1.3) when $\boldsymbol{a} \in \mathcal{D}$ is given. The linear space containing $\boldsymbol{u}$ was $U = \mathbb{R}^2$.

In this chapter, the state variable is written as u and uniquely determined as a solution to the state determination problem given by an abstract variational problem as shown later when the design variable $\phi \in \mathcal{D}$ is given. This problem is the same as Problem 5.2.3 but from the fact that the bilinear form a and the linear form l depend on ϕ, they are rewritten as $a(\phi)$ and $l(\phi)$ respectively. U is a real Hilbert space as per Problem 5.2.3.

Problem 7.2.1 (Abstract variational problem for ϕ) Let $\phi \in \mathcal{D}$ and define $a(\phi) : U \times U \to \mathbb{R}$ as a bounded and coercive bilinear form on U and $l(\phi) = l(\phi)(\,\cdot\,) = \langle l(\phi), \,\cdot\, \rangle \in U'$. In this case, find $u \in U$ such that

$$a(\phi)(u, v) = l(\phi)(v)$$

for every $v \in U$. □

Let us write Problem 7.2.1 in the following way. Let $\tau(\phi) : U \to U'$ be the isomorphism given by the Lax–Milgram theorem (Theorem 5.2.4) for a given bounded and coercive bilinear form $a(\phi)(\,\cdot\,,\,\cdot\,)$ and known term $l(\phi) \in U'$. In this case, find $u \in U$ which satisfies

$$s(\phi, u) = l(\phi) - \tau(\phi) u = 0_{U'}. \tag{7.2.1}$$

Moreover, as shown in Exercise 5.2.5 in Chap. 5, a non-homogeneous Dirichlet problem is contained in an abstract variational problem by replacing with $u \in U$ in Eq. (7.2.1) by $\tilde{u} = u - u_{\mathrm{D}} \in U$. Here, $l(\phi)$ can be replaced by $\hat{l}(\phi) = l(\phi) - \tau(\phi) u_{\mathrm{D}}$ and becomes

$$s(\phi, \tilde{u}) = \hat{l}(\phi) - \tau(\phi) \tilde{u} = 0_{U'}. \tag{7.2.2}$$

For simplicity, in this chapter, we use Eq. (7.2.1).

Moreover, as shown in Remark 7.6.3 later, in order to define the Fréchet derivative of cost function with respect to a variation of the design variable, the solution u of Problem 7.2.1 needs to be an element of the admissible set of state variables $\mathcal{S} \subset U$. In order for this to be satisfied, the known term $l(\phi)$ or the regularity of domain need to be appropriately set. Their conditions will be shown in Chaps. 8 and 9 depending on the specific optimum design problems. Here, the design variable u is assumed to be obtained as an element of $\mathcal{S}$.

Under this type of setting, in a similar manner to the state determination problem (Problem 1.1.3) in Chap. 1, $v \in U$ is taken to be an adjoint variable (or a Lagrange multiplier) and

$$\mathscr{L}_{\mathrm{S}}(\phi, u, v) = -a(\phi)(u, v) + l(\phi)(v) \tag{7.2.3}$$

is referred to as the Lagrange function for the state determination problem. Here, u is not necessarily the solution of Problem 7.2.1. However, the element $u \in U$ which satisfies

$$\mathscr{L}_{\mathrm{S}}(\phi, u, v) = 0 \tag{7.2.4}$$

with respect to an arbitrary $v \in U$ is equivalent to the weak-form solution of Problem 7.2.1.

7.3 Abstract Optimum Design Problem

In Problem 1.1.4, the cost functions f_0 and f_1 were defined as a function of the design variable and the state variable. Here, the functionals $f_0, \ldots, f_m$ defined on the admissible set $\mathcal{D} \subset X$ of the design variables defined in Sect. 7.1 and the admissible set $\mathcal{S} \subset U$ of the state variables defined in Sect. 7.2 are set to be cost functions and used in an abstract optimum design problem as follows.

Problem 7.3.1 (Abstract optimum design problem) For $(\phi, u) \in \mathcal{D} \times \mathcal{S}$, if $f_0, \ldots, f_m : \mathcal{D} \times \mathcal{S} \to \mathbb{R}$ is given, obtain ϕ which satisfies

$$\min_{(\phi,u)\in\mathcal{D}\times\mathcal{S}} \{ f_0(\phi, u) \mid f_1(\phi, u) \le 0, \ldots, f_m(\phi, u) \le 0, \text{ Problem 7.2.1} \}.$$

□

Problem 7.3.1 can be thought of in the following way by using Figs. 2.1.1 and 2.1.3. Even if X becomes a real Hilbert space, there is no need to change the image of the plane within the diagrams. Furthermore, if $\mathcal{D}$ only imposes constraint conditions such as smoothness on an element in X with no constraint conditions imposed using the norm of X directly, it again becomes a similar plane image as X. However, the plane in this case can be thought to be a plane made of only elements satisfying constraint conditions such as smoothness, just like the set of rational numbers in the real numbers. Moreover, the set S of

Eq. (2.1.1) called the admissible set of design variables in Chap. 2 can be replaced by

$$S = \{(\phi, u(\phi)) \in \mathcal{D} \times \mathcal{S} \mid f_1(\phi, u(\phi)) \le 0, \cdots, f_m(\phi, u(\phi)) \le 0, \\ s(\phi, u) = 0_{U'}\} \tag{7.3.1}$$

in this chapter. This set is an image of the sets on a plane satisfying $f_1 \le 0$ and $f_2 \le 0$ in Figs. 2.1.1 and 2.1.3.

We shall now look at the Fréchet derivatives of the cost functions and KKT conditions with respect to Problem 7.3.1. In this case, the notation of the Lagrange function is used in several ways. Here, in order to avoid confusion, let us summarize these relationships. Let the Lagrange function with respect to Problem 7.3.1 be

$$\mathscr{L}(\phi, u, v_0, v_1, \ldots, v_m) = \mathscr{L}_0(\phi, u, v_0) + \sum_{i \in \{1,\ldots,m\}} \lambda_i \mathscr{L}_i(\phi, u, v_i). \tag{7.3.2}$$

Here, $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_m\}^\top$ is a Lagrange multiplier with respect to $f_1 \le 0, \ldots,$ $f_m \le 0$. Furthermore, if the cost function f_i is given as a functional of the solution u of a state determination problem (Problem 7.2.1),

$$\mathscr{L}_i(\phi, u, v_i) = f_i(\phi, u) + \mathscr{L}_{\mathrm{S}}(\phi, u, v_i) \tag{7.3.3}$$

is referred to as the Lagrange function with respect to $f_i(\phi, u)$. Here, $\mathscr{L}_{\mathrm{S}}$ is the Lagrange function with respect to Problem 7.2.1 defined by Eq. (7.2.3). Moreover, v_i is a Lagrange multiplier defined with respect to f_i. If f_i contains boundary integrals on the Dirichlet boundary, for example $\int_{\Gamma_{\mathrm{D}}} v_{\mathrm{D}i} \partial_\nu u \, \mathrm{d}\gamma$ with respect to a Poisson problem, $\tilde{v}_i = v_i - v_{\mathrm{D}i}$ is assumed to be an element of U. Details are shown in Chaps. 8 and 9. Hereinafter, boundary integrals on the Dirichlet boundary are not included in f_i and v_i is an element of U.

7.4 Existence of an Optimum Solution

The abstract optimum design problem was defined in Problem 7.3.1. In this section, we will confirm the existence of an optimum solution. To do this, Weierstrass's theorem (Theorem 2.3.2) shown in Chap. 2 becomes a basic principle. Here, we will consider a corresponding theorem for the abstract optimum design problem. The concept used here are explained precisely in [2, Section 2.3, p. 38, Section 2.4, p. 45].

In the optimization problems considered in Chap. 2, the cost functions were defined only for the design variables. In contrast, in the case of optimum design problems, the cost functions are defined as functions of the design variable ϕ and state variable $u(\phi)$ which is determined with ϕ. Hence, the assumption of Theorem 2.3.2 in Chap. 2 that an admissible set of design variables is a bounded closed subset is replaced with that the admissible set for $(\phi, u(\phi))$, or its subset S defined in Eq. (7.3.1), is compact on $X \times U$ in the case of abstract optimum design problems. Since $u(\phi)$ is continuously determined from $\phi \in \mathcal{D}$, the admissible set of $(\phi, u(\phi))$ is given by the graph of $u(\phi)$ with respect to ϕ defined as

$$\mathcal{F} = \{(\phi, u(\phi)) \in \mathcal{D} \times \mathcal{S} \mid \text{Problem 7.2.1}\}. \tag{7.4.1}$$

Then, we need to show that $\mathcal{F}$ is compact on $X \times U$ in the abstract optimum design problem (Problem 7.3.1). To do it, we form the following assumption in addition to the hypothesis shown in the abstract variational problem (Problem 7.2.1).

Hypothesis 7.4.1 (Continuity of $a(\phi)$ and $l(\phi)$) Let $a(\phi)$ and $l(\phi)$ defined in Problem 7.2.1 be continuous with respect to $\phi \in \mathcal{D}$, that is, $a(\phi_n) \to a(\phi)$ and $l(\phi_n) \to l(\phi)$ hold with respect to an arbitrary Cauchy sequence $\phi_n \to \phi$ on X which is uniformly convergent in $\mathcal{D}$. □

The compactness of $\mathcal{F}$ can be shown as follows [2, Lemma 2.1, p. 14, Lemma 2.12, p. 39].

Lemma 7.4.2 (Compactness of $\mathcal{F}$) In addition to the hypothesis in Problem 7.2.1, let Hypothesis 7.4.1 be satisfied. With respect to an arbitrary Cauchy sequence $\phi_n \to \phi$ on X which is uniformly convergent in $\mathcal{D}$ and the solutions $u_n = u(\phi_n) \in U$ $(n \to \infty)$ of Problem 7.2.1, the convergence

$$u_n \to u \quad \text{strongly in } U$$

holds, and $u = u(\phi) \in U$ solves Problem 7.2.1. □

Proof With respect to the solution u_n of Problem 7.2.1 for ϕ_n,

$$\alpha_n \|u_n\|_U^2 \le a(\phi_n)(u_n, u_n) = l(\phi_n)(u_n) \le \|l(\phi_n)\|_{U'} \|u_n\|_U$$

holds. Here, α_n is a positive constant used in the definition of coerciveness for $a(\phi_n)$. When $\phi_n \to \phi$ is uniformly convergent in $\mathcal{D}$, α_n can be replaced by a positive constant α not depending on n. From the equation, it can be confirmed that $\{u_n\}_{n\in\mathbb{N}}$ is bounded. Then, there exists a subsequence such that $u_n \to u$ weakly in U.

Next, we will show that u is the solution of Problem 7.2.1 for ϕ. From the definition of Problem 7.2.1,

$$\lim_{n\to\infty} a(\phi_n)(u_n, v) = \lim_{n\to\infty} l(\phi_n)(v) \tag{7.4.2}$$

holds with respect to an arbitrary $v \in U$. Using Hypothesis 7.4.1, the right-hand side of Eq. (7.4.2) becomes

$$\lim_{n\to\infty} l(\phi_n)(v) = l(\phi)(v). \tag{7.4.3}$$

The left-hand side of Eq. (7.4.2) becomes

$$\lim_{n\to\infty} a(\phi_n)(u_n, v) = a(\phi)(u, v). \tag{7.4.4}$$

Indeed, we have Eq. (7.4.4) from

$$\begin{aligned}
&|a(\phi_n)(u_n,v) - a(\phi)(u,v)| \\
&\quad \le |a(\phi_n)(u_n,v) - a(\phi)(u_n,v)| + |a(\phi)(u_n,v) - a(\phi)(u,v)| \\
&\quad \le \|a(\phi_n) - a(\phi)\|_{\mathcal{L}(U\times U,\mathbb{R})} \|u_n\|_U \|v\|_U + |a(\phi)(u_n - u, v)|
\end{aligned}$$

and by using Hypothesis 7.4.1 and $u_n \to u$ weakly in U. Substituting Eq. (7.4.3) and Eq. (7.4.4) in Eq. (7.4.2), it is confirmed that u is the solution of Problem 7.2.1 for ϕ.

Since the weak convergence was shown, then to prove the strong convergence of $\{u_n\}_{n\in\mathbb{N}}$ to u, it is sufficient to show that

$$\|u_n\|_U \to \|u\|_U \quad (n \to \infty). \tag{7.4.5}$$

Indeed, when using

$$|||v||| = \langle \tau(\phi) v, v\rangle$$

as a norm on U, we have

$$\begin{aligned}
|||u_n||| &= \langle \tau(\phi) u_n, u_n\rangle = \langle (\tau(\phi) - \tau(\phi_n)) u_n, u_n\rangle + \langle \tau(\phi_n) u_n, u_n\rangle \\
&= \langle (\tau(\phi) - \tau(\phi_n)) u_n, u_n\rangle + l(\phi_n)(u_n) \\
&\to l(\phi)(u) = |||u||| \quad (n \to \infty).
\end{aligned} \tag{7.4.6}$$

Then, $u_n \to u$ strongly in U is proved. □

We assume that $u(\phi)$ belongs to $\mathcal{S}$ is guaranteed in the setting of Problem 7.2.1.

On the other hand, we form the following hypothesis for the objective function.

Hypothesis 7.4.3 (Continuity of f_0) Let f_0 be lower semi-continuous on S defined in Eq. (7.3.1). That is, with respect to an arbitrary Cauchy sequence $\phi_n \to \phi$ on X which is uniformly convergent in $\mathcal{D}$, by which we determine a Cauchy sequence $u(\phi_n) \to u(\phi)$ ($(\phi_n, u(\phi_n)), (\phi, u(\phi)) \in S$), it holds that

$$\liminf_{n\to\infty} f_0(\phi_n, u(\phi_n)) \ge f_0(\phi, u(\phi)).$$

□

Using the hypotheses and the previous lemma above, we have the following result for the existence of a solution to the abstract optimum design problem (Problem 7.3.1) [2, Theorem 2.1, p. 16, Theorem 2.8, p. 41].

Theorem 7.4.4 (Existence of an optimum solution) In addition to the hypothesis in Problem 7.2.1, suppose Hypothesis 7.4.1 is satisfied. Let S in Eq. (7.3.1) not be empty and compact in $X \times U$. Moreover, f_0 is lower semi-continuous (Hypothesis 7.4.3) on S. Then, there exists a minimum point in Problem 7.3.1. □

Proof Let $\{\phi_n\}_{n\in\mathbb{N}}$ ($\phi_n \in \mathcal{D}$) be a minimizing sequence in Problem 7.3.1, and

$$q = \inf_{(\phi,u(\phi))\in S} f_0(\phi, u(\phi)) = \lim_{n\to\infty} f_0(\phi_n, u(\phi_n)). \tag{7.4.7}$$

Since $\mathcal{D}$ is compact, there exists a subsequence which we still denote by $\{\phi_n\}_{n\in\mathbb{N}}$ and $\phi^* \in \mathcal{D}$ such that

$$\phi_n \to \phi^* \quad \text{strongly in } X. \tag{7.4.8}$$

From Lemma 7.4.2, we have

$$u(\phi_n) \to u(\phi^*) \quad \text{strongly in } U, \tag{7.4.9}$$

where $u(\phi_n)$ and $u(\phi^*)$ are the solutions of Problem 7.2.1 with respect to ϕ_n and ϕ^*, respectively. Using Eq. (7.4.8), Eq. (7.4.9), Eq. (7.4.7) and the lower semi-continuity of f_0 on S, we conclude that the limit

$$q = \liminf_{n\to\infty} f_0(\phi_n, u(\phi_n)) = f_0(\phi^*, u(\phi^*))$$

holds. It means that $(\phi^*, u(\phi^*)) \in S$ is a minimum point in Problem 7.3.1. □

7.5 Derivatives of Cost Functions

From this point onward, assuming that the conditions for the existence of a solution of the abstract optimum design problem (Problem 7.3.1) are satisfied, we shall examine a solution to an optimization problem with equality constraints. In this book, we focus on an approach based on the gradient method, so next let us think about the way to seek the Fréchet derivative of a cost function f_i with respect to a variation of ϕ on X. Here, there is a need to seek the Fréchet derivative on X when the equality constraints of the abstract variational problem (Problem 7.2.1) are satisfied. With respect to the equality-constrained problems on a finite-dimensional vector space, the Lagrange multiple method described in Section 2.6.2 (or adjoint variable method described in Section 2.6.5) was used. This principle is based on Theorem 2.6.4. Here, let us think about expanding this into the function space.

Let us expand Problem 2.6.1 to a problem defined on $X \times U$. Let us consider an optimization problem with equality constraint such as the following. Here, let f_i be a cost function for $i \in \{1, \ldots, m\}$.

Problem 7.5.1 (Optimization problem with equality constraint) Let $(\phi, u) \in X \times U$. If $f_i : X \times U \to \mathbb{R}$ is given, find (ϕ, u) which satisfies

$$\min_{(\phi,u)\in X\times U} \{ f_i(\phi, u) \mid s(\phi, u) = 0_{U'} \},$$

where $s(\phi, u)$ is defined by Eq. (7.2.1). □

In this chapter, an arbitrary variation of (ϕ, u) will be denoted by $(\varphi, w) \in X \times U$ and the Fréchet derivative of s and f_i will be written as

$$\begin{aligned} f_i'(\phi, u)[\varphi, w] &= f_{i\phi}(\phi, u)[\varphi] + f_{iu}(\phi, u)[w] \\ &= \langle g_{f_i}, \varphi \rangle + f_{iu}(\phi, u)[w], \end{aligned} \tag{7.5.1}$$

$$\begin{aligned} s'(\phi, u)[\varphi, w] &= s_\phi(\phi, u)[\varphi] + s_u(\phi, u)[w] \\ &= g_h[\varphi] - \tau(\phi) w, \end{aligned} \tag{7.5.2}$$

respectively. We shall use these notations to show the result of Theorem 2.6.4 being expanded.

Theorem 7.5.2 (1st necessary condition for a minimizer) Let f_i and s of Problem 7.5.1 be elements of $C^1(X \times U; \mathbb{R})$ and $C^1(X \times U; U')$, respectively. Let the Fréchet derivatives of f_i and s with respect to an arbitrary $\varphi \in X$ be given by Eq. (7.5.1) and Eq. (7.5.2), respectively. In this case, if (ϕ, u) is the minimal point of Problem 7.5.1, there exists a $v_i \in U$ which satisfies

$$\langle g_{f_i}, \varphi \rangle + \langle g_h[\varphi], v_i \rangle + \langle f_{iu}(\phi, u) - \tau^*(\phi) v_i, w \rangle = 0, \tag{7.5.3}$$

$$\langle l(\phi) - \tau(\phi) u, w \rangle = 0 \tag{7.5.4}$$

for an arbitrary $(\varphi, w) \in X \times U$. Here, $\tau^*(\phi) : U \to \hat{u}$ is the adjoint operator of $\tau(\phi)$. □

Proof From the fact that we assume $s \in C^1(X \times U; U')$ and there is a unique solution u which satisfies $s(\phi, u) = 0_{U'}$, s satisfies the following assumptions for the implicit function theorem (Theorem A.4.2) in neighborhood $B_X \times B_U \subset X \times U$ of $(\phi, u) \in X \times U$:

(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 $y = (\varphi, w) \in B_X \times B_U$ 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.

Hence, from the implicit function theorem, there exist some neighborhood $U_X \times U_U \subset B_X \times B_U$ and continuous mapping $\upsilon : U_X \to U_U$ (υ is the Greek letter upsilon), and $s(\phi, u) = 0_{U'}$ can be written as

$$u = \upsilon(\phi). \tag{7.5.5}$$

Therefore, $y(\phi) = (\phi, \upsilon(\phi)) \in C^1(\mathcal{D}; X \times U)$ can be defined.

Hence, write $\tilde{f}_i(\phi) = f_i(\phi, \upsilon(\phi)) = f_i(y(\phi))$. Since $f_i \in C^1(X \times U; \mathbb{R})$, when ϕ is a local minimizer,

$$\tilde{f}_i'(\phi)[\varphi] = y'^*(\phi) \circ g_i(\phi, \upsilon(\phi))[\varphi] = 0 \tag{7.5.6}$$

holds with respect to an arbitrary $\varphi \in X$. Here,

$$g_i(\phi, \upsilon(\phi)) = f_i'(\phi, \upsilon(\phi)) \in \mathcal{L}\left(X; X' \times U'\right) = \mathcal{L}(X; \mathcal{L}(X \times U; \mathbb{R})),$$

$$y'(\phi) \in \mathcal{L}(X; X \times U), \quad y'^{*}(\phi) \in \mathcal{L}\left(X' \times U'; X'\right).$$

$\mathcal{L}(X; U)$ represents the bounded linear operator $X \to U$. $\circ$ represents a composition operator. We rewrite the relationship of Eq. (7.5.6) as follows.

Firstly, let us write the admissible set of (ϕ, u) as

$$S = \{(\phi, u) \in X \times U \mid s(\phi, u) = 0_{U'}\}. \tag{7.5.7}$$

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

$$T_S(\phi, u) = \left\{(\varphi, \hat{v}) \in X \times U \;\middle|\; s'(\phi, u)[\varphi, \hat{v}] = 0_{U'}\right\} \tag{7.5.8}$$

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

$$\begin{aligned} & T_S'(\phi, u) \\ & \quad = \left\{ (\psi, w) \in X' \times U' \mid \langle (\varphi, \hat{v}), (\psi, w) \rangle = 0 \text{ for all } (\varphi, \hat{v}) \in T_S(\phi, u) \right\}. \end{aligned} \tag{7.5.9}$$

Moreover, the relationship between $T_S(\phi, u)$ and the Fréchet derivative $y'(\phi)$ of $y(\phi)$ 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 an arbitrary $\varphi \in X$, we get

$$s'(\phi, u) \circ y'(\phi)[\varphi] = 0_{U'}. \tag{7.5.10}$$

Here, the invertibility of $\tau(\phi)$ was used. 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 is established:

$$T_S(\phi, u) = \operatorname{Im} y'(\phi). \tag{7.5.11}$$

We use the relationship above to rewrite Eq. (7.5.6). 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,

$$g_i(\phi, \upsilon(\phi)) \in T_S'(\phi, u). \tag{7.5.12}$$

Here, from the theorem relating to the orthogonal complement space of the image space and the kernel space and Eq. (7.5.11),

$$T_S'(\phi, u) = (T_S(\phi, u))^{\perp} = \left(\operatorname{Ker} s'(\phi, u)\right)^{\perp} = \operatorname{Im} s'^{*}(\phi, u)$$

is established. Here, $s'^{*}(\phi, u) \in \mathcal{L}(U; X' \times U')$. Therefore, Eq. (7.5.12) is equivalent that there exists some $v_i \in U$ and

$$\begin{aligned} & f_{i\phi}(\phi, u)[\varphi] + f_{iu}(\phi, u)[w] + \langle s_{\phi}(\phi, u)[\varphi], v_i \rangle + \langle s_u(\phi, u)[w], v_i \rangle \\ & \quad = \langle g_{f_i}, \varphi \rangle + \langle g_h[\varphi], v_i \rangle + \langle f_{iu}(\phi, u) - \tau^{*}(\phi) v_i, w \rangle = 0 \end{aligned}$$

holds with respect to an arbitrary $(\varphi, w) \in X \times U$. In other words, Eq. (7.5.3) is established. Moreover, Eq. (7.5.4) holds if u is the solution of Eq. (7.2.1). $\square$

7.5.1 Adjoint Variable Method

Let us define the adjoint variable method in the following way based on Theorem 7.5.2. $v_i \in U$ is called an adjoint problem with respect to f_i, and is determined so that the second term on the left-hand side of Eq. (7.5.3) becomes zero. In other words, let it be the solution of the following problem.

Problem 7.5.3 (Adjoint problem with respect to f_i) When $\phi \in X$ and the solution $u \in U$ of Eq. (7.2.1) in this case as well as $f_{iu}(\phi, u) \in U'$ are given, obtain a function $v_i \in U$ which satisfies

$$f_{iu}(\phi, u) - \tau^*(\phi)\, v_i = 0_{U'}. \tag{7.5.13}$$

Here, $\tau(\phi)$ is the same as Eq. (7.2.1). □

If the solution v_i of Problem 7.5.3 is used, Eq. (7.5.3) can be written as

$$\langle g_{f_i}, \varphi \rangle + \langle g_h[\varphi], v_i \rangle = \langle g_i, \varphi \rangle = 0. \tag{7.5.14}$$

When u is the solution of Eq. (7.2.1) and (ϕ, u) is the minimal point of Problem 7.5.1, Eq. (7.5.14) holds by Theorem 7.5.2.

The g_i in this case is the gradient of the Fréchet derivative of f_i with respect to $\varphi \in X$ when u continues to be the solution of the state determination problem (Problem 7.2.1), even when the design variable varies with an arbitrary $\varphi \in X$. Here, if $\upsilon(\phi)$ of Eq. (7.5.5) defined in the proof of Theorem 7.5.2 is used, the following can be written with respect to $\tilde{f}_i(\phi) = f_i(\phi, \upsilon(\phi))$:

$$\tilde{f}_i'(\phi)[\varphi] = \langle g_i, \varphi \rangle. \tag{7.5.15}$$

7.5.2 Lagrange Multiplier Method

The gradient g_i of the Fréchet derivative of the cost function f_i with respect to an arbitrary variation $\varphi \in X$ of design variable can also be obtained from the Lagrange multiplier method shown next. This method is used in Chaps. 8 and 9 because the process is explicit.

As defined in Problem 2.6.5, the Lagrange multiplier method is a method for finding candidates for solutions by replacing optimization problems with equality constraints with stationary conditions of Lagrange functions. Hence, set the Lagrange function of Problem 7.5.1 to be

$$\mathscr{L}_i(\phi, u, v_i) = f_i(\phi, u) + \langle s(\phi, u), v_i \rangle = f_i(\phi, u) + \mathscr{L}_{\mathrm{S}}(\phi, u, v_i). \tag{7.5.16}$$

Here, $\mathscr{L}_{\mathrm{S}}(\phi, u, v_i)$ is a Lagrange function of the state determination problem (Problem 7.2.1). The function u is not necessary for the solution of Problem 7.2.1. v_i is the Lagrange multiplier with respect to the state determination problem prepared for f_i and assumed to be an element of U, similarly to Theorem 7.5.2. In this case, the Fréchet derivative of $\mathscr{L}_i(\phi, u, v_i)$ with respect to an arbitrary variation $(\varphi, \hat{u}, \hat{v}_i) \in X \times U \times U$ of (ϕ, u, v_i) becomes

$$\mathscr{L}_i'(\phi, u, v_i)[\varphi, \hat{u}, \hat{v}_i]$$

$$= \mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi] + \mathscr{L}_{iu}(\phi, u, v_i)[\hat{u}] + \mathscr{L}_{iv_i}(\phi, u, v_i)[\hat{v}_i]. \quad (7.5.17)$$

The following can be obtained with respect to the third term on the right-hand side of Eq. (7.5.17):

$$\mathscr{L}_{iv_i}(\phi, u, v_i)[\hat{v}_i] = \mathscr{L}_{\mathrm{S}}(\phi, u, \hat{v}_i). \quad (7.5.18)$$

The right-hand side of Eq. (7.5.18) is the Lagrange function of the state determination problem (Problem 7.2.1). In this case, if u is a solution of the state determination problem, the third term on the right-hand side of Eq. (7.5.17) becomes zero. Moreover, the following equation holds:

$$\begin{aligned}\mathscr{L}_{iu}(\phi, u, v_i)[\hat{u}] &= f_{iu}(\phi, u)[\hat{u}] + \mathscr{L}_{\mathrm{S}u}(\phi, u, v_i)[\hat{u}] \\ &= \langle f_{iu}(\phi, u) - \tau^*(\phi) v_i, \hat{u}\rangle. \end{aligned} \quad (7.5.19)$$

The conditions under which Eq. (7.5.19) becomes zero with respect to an arbitrary $\hat{u} \in U$ is the same as the weak-form equation of the adjoint problem (Problem 7.5.3). Hence, if the weak solution of the adjoint problem is set to be v_i, the second term on the right-hand side of Eq. (7.5.17) becomes zero.

Furthermore, the first term on the right-hand side of Eq. (7.5.17) becomes

$$\mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi] = \langle g_{f_i}, \varphi\rangle + \langle g_h[\varphi], v_i\rangle = \langle g_i, \varphi\rangle. \quad (7.5.20)$$

The g_i of Eq. (7.5.20) matches the g_i of Eq. (7.5.14).

This relationship is an expression which is an abstract format of Eq. (1.1.37) in Chap. 1. In Chaps. 8 and 9, the stationary conditions of a Lagrange function with respect to f_i will be used to seek g_i as shown here.

7.5.3 Second-Order Fréchet Derivatives of Cost functions

Furthermore, let us think about the second-order derivatives of the cost functions with respect to a variation of design variable based on the definition of a Fréchet derivative (Definition 4.5.4).

In Section 1.1.6, the second-order derivative of mean compliance with respect to a variation of design variables when a stepped one-dimensional linear elastic problem is set to be a state determination problem was sought by using the second-order derivative of the Lagrange function $\mathscr{L}_0$. In this regard, it was crucial to substitute for the variation $\hat{\boldsymbol{u}}$ of the state variable $\boldsymbol{u}$ with the variation $\hat{\boldsymbol{u}}$ which satisfies the equality constraints of the state determination problem based on Theorems 2.6.6 and 2.6.7. Here, let us think about something similar with respect to an abstract optimum design problem (Problem 7.5.1) with equality constraints.

With respect to $\mathscr{L}_i$ defined by Eq. (7.5.16), (ϕ, u) is thought of as a design variable based on the definitions in Chap. 2. In this case, the second-order Fréchet derivative of $\mathscr{L}_i$ with respect to arbitrary variations $(\varphi_1, \hat{u}_1)$ and $(\varphi_2, \hat{u}_2) \in T_S(\phi, u)$ of $(\phi, u) \in S$ becomes

$$\mathscr{L}_{i(\phi,u)(\phi,u)}(\phi, u, v_i)[(\varphi_1, \hat{u}_1), (\varphi_2, \hat{u}_2)]$$

$$
\begin{aligned}
&= \left(\mathscr{L}_{i\phi}(\phi,u,v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}_1]\right)_\phi [\varphi_2] \\
&\quad + \left(\mathscr{L}_{i\phi}(\phi,u,v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi,u,v_i)[\hat{u}_1]\right)_u [\hat{u}_2] \\
&= \mathscr{L}_{i\phi\phi}(\phi,u,v_i)[\varphi_1,\varphi_2] + \mathscr{L}_{iu\phi}(\phi,u,v_i)[\hat{u}_1,\varphi_2] \\
&\quad + \mathscr{L}_{i\phi u}(\phi,u,v_i)[\varphi_1,\hat{u}_2] + \mathscr{L}_{iuu}(\phi,u,v_i)[\hat{u}_1,\hat{u}_2]\,. \qquad (7.5.21)
\end{aligned}
$$

Using it, we have the following result corresponding to Theorem 2.6.6.

Theorem 7.5.4 (The second-order necessary condition) Let f_i and s be elements of $C^2(X \times U;\mathbb{R})$ and $C^2(X \times U;U')$, respectively. If (ϕ,u) is a local minimizer of Problem 7.5.1,

$$
\mathscr{L}_{i(\phi,u)(\phi,u)}(\phi,u,v_i)[(\varphi,\hat{v}),(\varphi,\hat{v})] \geq 0 \qquad (7.5.22)
$$

holds with respect to an arbitrary $(\varphi,\hat{v}) \in T_S(\phi,u)$. □

Proof In the proof of Theorem 7.5.2, the assumption $s(\phi,\,\cdot\,) \in C^1(B_U;U')$ for the implicit function theorem is replaced by $s(\phi,\,\cdot\,) \in C^2(B_U;U')$, and then using $\upsilon(\phi)$ in Eq. (7.5.5), $y(\phi) = (\phi,\upsilon(\phi)) \in C^2(\mathcal{D};X \times U)$ is determined. From Eq. (7.5.10), we have

$$
s''(\phi,u)\left[y'(\phi)[\varphi],y'(\phi)[\varphi]\right] = 0_{U'} \qquad (7.5.23)
$$

with respect to $y'(\phi)[\varphi] \in T_S(\phi,u)$. Hence, if (ϕ,u) is a local minimizer of Problem 7.5.1,

$$
\mathscr{L}_{i(\phi,u)(\phi,u)}(\phi,u,v_i)\left[y'(\phi)[\varphi],y'(\phi)[\varphi]\right] = \tilde{f}_i''(\phi)[\varphi,\varphi] \geq 0 \qquad (7.5.24)
$$

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

Moreover, corresponding to Theorem 2.6.7, we obtain the following result.

Theorem 7.5.5 (The second-order sufficient condition) Under the assumptions of Theorem 7.5.4, if Eq. (7.5.3) and Eq. (7.5.4) are satisfied at $(\phi,u,v_i) \in X \times U^2$ and Eq. (7.5.22) replacing $\geq$ with $>$ holds, then (ϕ,u) is a local minimizer of Problem 7.5.1. □

Proof When $(\phi,u,v_i) \in X \times U^2$ 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

$$
\tilde{f}_i(\phi+\varphi) - \tilde{f}_i(\phi) = \frac{1}{2}\mathscr{L}_{i(\phi,u)(\phi,u)}(\phi+\theta\varphi,u(\phi+\theta\varphi),v_i)[z(\varphi),z(\varphi)]
$$

for all $y(\phi)+z(\varphi) \in B$. From the assumption, since the right-hand side is greater than or equal to zero, $\tilde{f}_i(\phi) \leq \tilde{f}_i(\phi+\varphi)$ holds. □

In view of Theorems 7.5.4 and 7.5.5, since the left-hand side of Eq. (7.5.24) is the Hessian of $\tilde{f}_i$ with respect to an arbitrary variation $\varphi \in X$ of ϕ, we write it as $h_i(\phi,u,v_i) \in \mathcal{L}^2(X \times X;\mathbb{R})$ (Definition 4.5.4). h_i is calculated as

$$
h_i(\phi,u,v_i)[\varphi_1,\varphi_2]
$$

$$
\begin{aligned}
&= \left(\mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi, u, v_i)[\hat{v}_1]\right)_{\phi}[\varphi_2] \\
&\quad + \left(\mathscr{L}_{i\phi}(\phi, u, v_i)[\varphi_1] + \mathscr{L}_{iu}(\phi, u, v_i)[\hat{v}_1]\right)_{u}[\hat{v}_2] \\
&= \mathscr{L}_{i\phi\phi}(\phi, u, v_i)[\varphi_1, \varphi_2] + \mathscr{L}_{iu\phi}(\phi, u, v_i)[\hat{v}_1, \varphi_2] \\
&\quad + \mathscr{L}_{i\phi u}(\phi, u, v_i)[\varphi_1, \hat{v}_2] + \mathscr{L}_{iuu}(\phi, u, v_i)[\hat{v}_1, \hat{v}_2], \qquad (7.5.25)
\end{aligned}
$$

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

$$
\mathscr{L}_{\mathrm{S}\phi u}(\phi, u, v)[\varphi_j, \hat{v}_j] = 0 \qquad (7.5.26)
$$

for all $\varphi_j \in X$. These specific results of h_i are shown in Chaps. 8 and 9.

7.5.4 Second-Order Fréchet Derivative of Cost Function Using Lagrange Multiplier Method

The application of the Lagrange multiplier method in obtaining the second-order Fréchet derivative of a cost function is described as follows. Recalling the definition of the second-order Fréchet derivative (Definition 4.5.4), and that u and v_i are the solutions of the state determination and adjoint problems, respectively, we fix φ_1 and define the Lagrange function with respect to $\langle g_i, \varphi_1\rangle$ in Eq. (7.5.20) by

$$
\mathscr{L}_{\mathrm{I}i}(\phi, u, v_i, w_i, z_i) = \langle g_i, \varphi_1\rangle + \mathscr{L}_{\mathrm{S}}(\phi, u, w_i) + \mathscr{L}_{\mathrm{A}i}(\phi, v_i, z_i), \qquad (7.5.27)
$$

where $\mathscr{L}_{\mathrm{S}}$ is defined by Eq. (7.2.3). $\mathscr{L}_{\mathrm{A}i}$ is the Lagrange function of the adjoint problem (Problem 7.5.3) with respect to f_i defined by

$$
\mathscr{L}_{\mathrm{A}i}(\phi, v_i, z_i) = \mathscr{L}_{iu}(\phi, u, v_i)[z_i] = \langle f_{iu}(\phi, u) - \tau^*(\phi) v_i, z_i\rangle, \qquad (7.5.28)
$$

where $\mathscr{L}_{iu}(\phi, u, v_i)[z_i]$ is given by Eq. (7.5.19). $w_i \in U$ and $z_i \in U$ are the adjoint variables provided for u and v_i in g_i.

With respect to arbitrary variations $(\varphi_2, \hat{u}, \hat{v}_i, \hat{w}_i, \hat{z}_i) \in X \times U^4$ of (ϕ, u, v_i, w_i, z_i), the derivative of $\mathscr{L}_{\mathrm{I}i}$ is written as

$$
\begin{aligned}
&\mathscr{L}'_{\mathrm{I}i}(\phi, u, v_i, w_i, z_i)[\varphi_2, \hat{u}, \hat{v}_i, \hat{w}_i, \hat{z}_i] \\
&\quad = \mathscr{L}_{\mathrm{I}i\phi}(\phi, u, v_i, w_i, z_i)[\varphi_2] + \mathscr{L}_{\mathrm{I}iu}(\phi, u, v_i, w_i, z_i)[\hat{u}] \\
&\qquad + \mathscr{L}_{\mathrm{I}iv_i}(\phi, u, v_i, w_i, z_i)[\hat{v}_i] + \mathscr{L}_{\mathrm{I}iw_i}(\phi, u, v_i, w_i, z_i)[\hat{w}_i] \\
&\qquad + \mathscr{L}_{\mathrm{I}iz_i}(\phi, u, v_i, w_i, z_i)[\hat{z}_i]. \qquad (7.5.29)
\end{aligned}
$$

The fourth term on the right-hand side of Eq. (7.5.29) vanishes if u is the solution of the state determination problem. If v_i can be determined as the solution of the adjoint problem, the fifth term of Eq. (7.5.29) also vanishes.

The condition that the second term on the right-hand side of Eq. (7.5.29) satisfies

$$
\mathscr{L}_{\mathrm{I}iu}(\phi, u, v_i, w_i, z_i)[\hat{u}] = 0 \qquad (7.5.30)
$$

with respect to arbitrary variation of $\hat{u} \in U$ gives the adjoint problem with respect to $\langle g_i, \varphi_1 \rangle$ to determine w_i. The condition that the third term on the right-hand side of Eq. (7.5.29) satisfies

$$\mathscr{L}_{\mathrm{I}iv_i}(\phi, u, v_i, w_i, z_i)[\hat{v}_i] = 0 \tag{7.5.31}$$

with respect to arbitrary variation of $\hat{v}_i \in U$ gives the adjoint problem with respect to $\langle g_i, \varphi_1 \rangle$ to determine z_i.

Here, u, v_i, $w_i(\varphi_1)$ and $z_i(\varphi_1)$ are assumed to be the weak solutions of Problems 7.2.1, 7.5.3, Eq. (7.5.30) and Eq. (7.5.31), respectively. If we denote $f_i(\phi, u)$ by $\tilde{f}_i(\phi)$, then we can write

$$\begin{aligned}\mathscr{L}_{\mathrm{I}i\phi}(\phi, u, v_i, w_i(\varphi_1), z_i(\varphi_1))[\varphi_2] &= \tilde{f}_i''(\phi)[\varphi_1, \varphi_2] \\ &= g_{\mathrm{H}i}(\phi, \varphi_1)[\varphi_2].\end{aligned} \tag{7.5.32}$$

In this book, $g_{\mathrm{H}i}(\phi, \varphi_1)[\varphi_2]$ is called the Hesse gradient.

7.6 Descent Directions of Cost Functions

In Sect. 7.5, it became apparent that the first and second-order Fréchet derivative of cost functions $\tilde{f}_0, \ldots, \tilde{f}_m$ with respect to a variation of design variable can be obtained. Hence, let us think about making the solution to the optimization problem shown in Chap. 3 abstract.

7.6.1 Abstract Gradient Method

Let us make the gradient method abstract. From now on, let us go with the notation in Chap. 3 to write $\tilde{f}_i(\phi)$ as $f_i(\phi)$. Here, let us assume that the Fréchet derivative of $f_i(\phi)$ can be calculated and think about obtaining the minimum point of $f_i(\phi)$.

In the gradient method on a finite-dimensional vector space seen in Section 3.3 (Problem 3.3.1), when $X = \mathbb{R}^d$, the bilinear form $a_X(\cdot, \cdot) = (\cdot) \cdot (\boldsymbol{A}(\cdot))$ using the positive definite symmetric matrix $\boldsymbol{A}$ was an operator $X \times X \to \mathbb{R}$ which is coercive (Definition 5.2.1), bounded and symmetric. Focusing on these characteristics, an abstract gradient method such as the one below can be thought of.

Problem 7.6.1 (Abstract gradient method) Let $X \Supset \mathcal{D}$ be a real Hilbert space. Let $a_X : X \times X \to \mathbb{R}$ be a coercive and bounded bilinear form on X. In other words, with respect to an arbitrary $\varphi, \psi \in X$, there exists some $\alpha, \beta > 0$ and

$$a_X(\varphi, \varphi) \geq \alpha \|\varphi\|_X^2, \quad |a_X(\varphi, \psi)| \leq \beta \|\varphi\|_X \|\psi\|_X \tag{7.6.1}$$

is taken to form. With respect to $f_i \in C^1(X; \mathbb{R})$ (Definition 4.5.4), let $g_i(\phi_k) \in X'$ be the Fréchet derivative at $\phi_k \in \mathcal{D}^\circ$ which is not a local minimizer. In this case, obtain a $\varphi_{gi} \in X$ which satisfies

$$a_X(\varphi_{gi}, \varphi) = -\langle g_i(\phi_k), \varphi \rangle \tag{7.6.2}$$

with respect to an arbitrary $\varphi \in X$. □

In Problem 7.6.1, the symmetry $a_X(\varphi, \psi) = a_X(\psi, \varphi)$ of a_X was not assumed. This is because the desired result can be obtained later in Theorem 7.6.2 without using symmetry. In reality, it is possible to think of a non-symmetric example among coercive and bounded linear forms in a real Hilbert space. For example, with respect to an arbitrary $u, v \in X$ defined on $X = H_0^1(\Omega; \mathbb{R})$,

$$a(u, v) = \int_\Omega \left(\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v + \frac{\partial u}{\partial x_1} v \right) \mathrm{d}x$$

is a coercive and bounded bilinear form even though it is asymmetric.[1] However, when considering the numerical solution, it is desirable to assume the symmetry of a_X. Specific provisions are shown in response to the problems in Chaps. 8 and 9.

The following result is obtained with respect to Problem 7.6.1.

Theorem 7.6.2 (Abstract gradient method) The solution φ_{gi} of Problem 7.6.1 exists uniquely in X and the inequality

$$\|\varphi_{gi}\|_X \le \frac{1}{\alpha} \|g_i(\phi_k)\|_{X'} \tag{7.6.3}$$

is established. Here, α is a positive constant used in Eq. (7.6.1). Furthermore, φ_{gi} is the descent direction of f_i at ϕ. □

Proof The unique existence and Eq. (7.6.3) can be seen from the Lax–Milgram theorem. Furthermore, φ_{gi} satisfies Eq. (7.6.2), hence

$$\begin{aligned} f_i(\phi + \bar{\epsilon}\varphi_{gi}) - f_i(\phi) &= \bar{\epsilon}\langle g_i, \varphi_{gi}\rangle + o(|\bar{\epsilon}|) = -\bar{\epsilon} a_X(\varphi_{gi}, \varphi_{gi}) + o(|\bar{\epsilon}|) \\ &\le -\bar{\epsilon}\alpha \|\varphi_{gi}\|_X^2 + o(|\bar{\epsilon}|) \end{aligned}$$

holds with respect to a positive constant $\bar{\epsilon}$. □

Even among the abstract gradient methods, let us refer to the case when a function space of H^1-class is chosen in X as the H^1 gradient method.

Theorem 7.6.2 shows that the solution of the abstract gradient method (Problem 7.6.1) φ_{gi} is in X. However, there is no guarantee that φ_{gi} is in $\mathcal{D}$. Hence, there is a need to note the following.

Remark 7.6.3 (Solution of abstract gradient method) In order to use the solution φ_{gi} of the abstract gradient method (Problem 7.6.1) in the solution for the abstract optimum design problem (Problem 7.3.1), φ_{gi} should be obtained as an element of $\mathcal{D}$. The following needs to be noted to satisfy the condition:

[1] This a appears in the weak form of the problem in which a convective term is added to the homogeneous Poisson problem [F. Kikuchi, personal communication].

(1) In order for the solution φ_{gi} of the abstract gradient method (Problem 7.6.1) to be included in $\mathcal{D}$, g_i needs to be included within a set of functions with appropriate regularity. For this result, the known term $l(\phi)$ and boundary regularities need to be set appropriately in order for the solution u of the state determination problem (Problem 7.2.1) to be included in an appropriate set $\mathcal{S}$ of functions. Furthermore, $f_{iu}(\phi, u)$ needs to be appropriately set so that the solution v_i of the adjoint problem (Problem 7.5.3) is in the appropriate function set $\mathcal{S}$. Details of these are shown in Chaps. 8 and 9.

(2) When this is not possible (for example when there are special constraint conditions imposed on $\mathcal{D}$), in seeking φ_{gi}, or after it is been sought, there is a need to add extra procedures to satisfy the necessary constraint conditions under $\mathcal{D}$.

□

7.6.2 Abstract Newton Method

Next let us make the Newton method abstract. Here, let us assume that the Fréchet derivative $\langle g_i(\phi_k), \varphi\rangle$ of $f_i(\phi)$ and second-order Fréchet derivative $h_i(\phi_k)[\varphi_1, \varphi_2]$ can be calculated, and think about obtaining the minimum point of $f_i(\phi)$.

As seen in Section 3.5, the bilinear form $a_X(\,\cdot\,,\,\cdot\,) = (\,\cdot\,)\cdot(\boldsymbol{A}(\,\cdot\,))$ used in the gradient method was replaced by $h(\boldsymbol{x}_k)[\,\cdot\,,\,\cdot\,] = (\,\cdot\,)\cdot(\boldsymbol{H}(\boldsymbol{x}_k)(\,\cdot\,))$ using the Hessian matrix $\boldsymbol{H}$ in the Newton method (Problem 3.5.1). In a real Hilbert space X, the abstract Newton method such as the following can be thought of.

Problem 7.6.4 (Abstract Newton method) Let $X \Supset \mathcal{D}$ be a real Hilbert space. With respect to $f_i \in C^2(X;\mathbb{R})$ (Definition 4.5.4), the gradient of the Fréchet derivative and Hessian of f_i at a non-local minimum point $\phi_k \in \mathcal{D}^\circ$ are denoted as $g_i(\phi_k) \in X'$ and $h_i(\phi_k) \in \mathcal{L}^2(X \times X;\mathbb{R})$, respectively. Moreover, $a_X : X \times X \to \mathbb{R}$ is a coercive and bounded bilinear form on X. Here, obtain a $\varphi_{gi} \in X$ which satisfies

$$h_i(\phi_k)[\varphi_{gi}, \varphi] + a_X(\varphi_{gi}, \varphi) = -\langle g_i(\phi_k), \varphi\rangle \tag{7.6.4}$$

with respect to an arbitrary $\varphi \in X$. □

In Problem 7.6.4, the a_X was introduced in order to compensate for the lack of coerciveness and boundedness of the left-hand side of Eq. (7.6.4) and to ensure the regularity of φ_{gi}. Even among the abstract Newton methods, the case when function space of H^1-class is chosen in X is called the H^1 Newton method. In Problem 7.6.4, as in Theorem 3.5.2, when ϕ_k is sufficiently close to the local minimum, it is hoped that the point sequence generated by the abstract Newton method would have quadratic convergence to the local minimum point. Moreover, Remark 7.6.3 with respect to the solution to the abstract gradient method is valid here too.

Furthermore, in the case of the abstract Newton method when the second-order Fréchet derivative of $f_i(\phi)$ is given by the Hesse gradient, Problem 7.6.4 is replaced with the following problem.

Problem 7.6.5 (Abstract Newton method using Hesse gradient) Under the assumption of Problem 7.6.4, the gradient of the Fréchet derivative of f_i, search vector and Hesse gradient of f_i at a non-local minimum point $\phi_k \in \mathcal{D}^\circ$ are denoted by $g_i(\phi_k) \in X'$, $\bar{\varphi}_{gi} \in X$ and $g_{\mathrm{H}i}(\phi_k, \bar{\varphi}_{gi}) \in X'$, respectively. Given a coercive and bounded bilinear form $a_X : X \times X \to \mathbb{R}$ on X, find a $\varphi_{gi} \in X$ which satisfies

$$a_X(\varphi_{gi}, \varphi) = -\langle (g_i(\phi_k) + g_{\mathrm{H}i}(\phi_k, \bar{\varphi}_{gi})), \varphi \rangle \tag{7.6.5}$$

with respect to an arbitrary $\varphi \in X$. □

The solution φ_{gi} of Problem 7.6.5 accords with the solution of the abstract Newton method if $\bar{\varphi}_{gi} = \varphi_{gi}$.

7.7 Solution of Abstract Optimum Design Problem

Now that the abstract gradient method and abstract Newton method have been defined, let us think about the solution of the abstract optimum design problem (Problem 7.3.1) here.

7.7.1 Gradient Method for Constrained Problems

Firstly, let us bear in mind what was learned in Section 3.7 and think about the gradient method with respect to a constrained problem. Here, the gradients $g_0, \ldots, g_m \in X'$ of the Fréchet derivatives of the cost functions $f_0, \ldots, f_m$ are assumed to be calculable using the method shown in Sect. 7.5.

Here, let us show the KKT conditions with respect to Problem 7.3.1. The content shown here is an expansion of the KKT conditions Eq. (1.1.51) to Eq. (1.1.54) with respect to Problem 1.1.4 in Chap. 1. In Problem 1.1.4, $X = \mathbb{R}^2$ and $U = \mathbb{R}^2$. In contrast, in Problem 7.3.1, X and U were assumed to be real Hilbert spaces. The Fréchet derivatives of cost functions with respect to an arbitrary variation of design variable are included in the dual space X' of X. If this relationship is remembered, the following result can be obtained.

Let the Lagrange function with respect to Problem 7.3.1 be

$$\mathscr{L}(\phi, \boldsymbol{\lambda}) = f_0(\phi) + \sum_{i \in \{1,\ldots,m\}} \lambda_i f_i(\phi). \tag{7.7.1}$$

Here, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}^m$ is a Lagrange multiplier with respect to $f_1(\phi) \le 0, \ldots, f_m(\phi) \le 0$.

In this case, the KKT conditions with respect to Problem 7.3.1 are given by

$$g_0(\phi) + \sum_{i \in \{1,\ldots,m\}} \lambda_i g_i(\phi) = 0_{X'}, \tag{7.7.2}$$

$$f_i(\phi) \leq 0 \quad \text{for } i \in \{1,\ldots,m\}, \tag{7.7.3}$$

$$\lambda_i f_i(\phi) = 0 \quad \text{for } i \in \{1,\ldots,m\}, \tag{7.7.4}$$

$$\lambda_i \geq 0 \quad \text{for } i \in \{1,\ldots,m\}. \tag{7.7.5}$$

Let us think about the solution to the abstract optimum design problem (Problem 7.3.1) based on this condition following the gradient method with respect to the constrained problem shown in Section 3.7. With respect to $k \in \{0,1,2,\ldots\}$, the trial point ϕ_k is assumed to be an element of admissible set S defined by Eq. (7.3.1). Denote the set of suffixes with respect to active constraints with respect to ϕ_k as

$$I_{\mathrm{A}}(\phi_k) = \{ i \in \{1,\ldots,m\} \mid f_i(\phi_k) \geq 0\} = \left\{ i_1, \ldots, i_{|I_{\mathrm{A}}|} \right\}. \tag{7.7.6}$$

If there is no confusion, denote $I_{\mathrm{A}}(\phi_k)$ as I_{A}. Moreover, the size of the search vector (step size) is adjusted by the size of a positive constant c_a. In this case, the problem seeking the search vector $\varphi_g \in X$ satisfying the inequalities constraint with respect to the cost functions around ϕ_k is constructed in the following way.

Problem 7.7.1 (Gradient method for constrained problems) Suppose that for a trial point $\phi_k \in \mathcal{D}$ of Problem 7.3.1 satisfying the inequality constraints, $f_0(\phi_k)$, $f_{i_1}(\phi_k) = 0$, $\ldots$, $f_{i_{|I_{\mathrm{A}}|}}(\phi_k) = 0$ and $g_0(\phi_k)$, $g_{i_1}(\phi_k)$, $\ldots$, $g_{i_{|I_{\mathrm{A}}|}}(\phi_k) \in X'$ are given. Let $a_X : X \times X \to \mathbb{R}$ be a coercive and bounded bilinear form on X. Moreover, c_a is taken to be a positive constant. Obtain $\phi_{k+1} = \phi_k + \varphi_g$ which satisfies

$$q(\varphi_g) = \min_{\varphi \in X} \left\{ q(\varphi) = \frac{c_a}{2} a_X(\varphi, \varphi) + \langle g_0(\phi_k), \varphi \rangle \;\middle|\; f_i(\phi_k) + \langle g_i(\phi_k), \varphi \rangle \leq 0 \text{ for } i \in I_{\mathrm{A}}(\phi_k) \right\}.$$

□

Similarly to Problem 3.7.1, Problem 7.7.1 is a convex optimization problem. In this regard, φ_g satisfying the KKT conditions becomes the local minimizer of Problem 7.7.1. Let us consider the solution for Problem 7.3.1 by focusing on this. The method below is an abstract version of the method shown in Section 3.7.

Let the Lagrange function of Problem 7.7.1 be

$$\mathscr{L}_{\mathrm{Q}}(\varphi_g, \boldsymbol{\lambda}) = q(\varphi_g) + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_i \left(f_i(\phi_k) + \langle g_i(\phi_k), \varphi_g \rangle \right).$$

Here, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}^m$ is a Lagrange multiplier with respect to the inequality constraint conditions. The KKT conditions with respect to the minimum point φ_g of Problem 7.7.1 are that the following hold with respect to an arbitrary $\psi \in X$:

$$c_a a_X (\varphi_g, \psi) + \langle g_0 (\phi_k), \psi \rangle + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_i \langle g_i (\phi_k), \psi \rangle = 0, \tag{7.7.7}$$

$$f_i (\phi_k) + \langle g_i (\phi_k), \varphi_g \rangle \leq 0 \quad \text{for } i \in I_{\mathrm{A}} (\phi_k) \tag{7.7.8}$$

$$\lambda_{k+1\, i} (f_i (\phi_k) + \langle g_i (\phi_k), \psi \rangle) = 0 \quad \text{for } i \in I_{\mathrm{A}} (\phi_k), \tag{7.7.9}$$

$$\lambda_{k+1\, i} \geq 0 \quad \text{for } i \in I_{\mathrm{A}} (\phi_k). \tag{7.7.10}$$

$(\varphi_g, \boldsymbol{\lambda}_{k+1}) \in X \times \mathbb{R}^{|I_{\mathrm{A}}|}$ satisfying these can be obtained as follows.

Let φ_{g0}, φ_{i_1}, ... , $\varphi_{i_{|I_{\mathrm{A}}|}}$ be the solution of the abstract gradient method (Problem 7.6.1). Here, Eq. (7.6.2) is changed to

$$c_a a_X (\varphi_{gi}, \psi) = - \langle g_i, \psi \rangle \tag{7.7.11}$$

with respect to an arbitrary $\psi \in X$. In this case,

$$\varphi_g = \varphi_g (\lambda_{k+1\, i}) = \varphi_{g0} + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{k+1\, i}\, \varphi_{gi} \tag{7.7.12}$$

satisfies Eq. (7.7.7). On the other hand, Eq. (7.7.8) becomes

$$\begin{pmatrix} \langle g_{i_1}, \varphi_{g i_1} \rangle & \cdots & \left\langle g_{i_1}, \varphi_{g i_{|I_{\mathrm{A}}|}} \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle g_{i_{|I_{\mathrm{A}}|}}, \varphi_{g i_1} \right\rangle & \cdots & \left\langle g_{i_{|I_{\mathrm{A}}|}}, \varphi_{g i_{|I_{\mathrm{A}}|}} \right\rangle \end{pmatrix} \begin{pmatrix} \lambda_{k+1\, i_1} \\ \vdots \\ \lambda_{k+1\, i_{|I_{\mathrm{A}}|}} \end{pmatrix} = - \begin{pmatrix} f_{i_1} + \langle g_{i_1}, \varphi_{g0} \rangle \\ \vdots \\ f_{i_{|I_{\mathrm{A}}|}} + \left\langle g_{i_{|I_{\mathrm{A}}|}}, \varphi_{g0} \right\rangle \end{pmatrix}.$$

This equation is written as

$$(\langle g_i, \varphi_{gj} \rangle)_{(i,j) \in I_{\mathrm{A}}^2} (\lambda_{k+1\, j})_{j \in I_{\mathrm{A}}} = - (f_i + \langle g_i, \varphi_{g0} \rangle)_{i \in I_{\mathrm{A}}}. \tag{7.7.13}$$

In Eq. (7.7.13), the matrix $(\langle g_i, \varphi_{gj} \rangle)_{(i,j) \in I_{\mathrm{A}}^2}$ is symmetric because $\langle g_i, \varphi_{gj} \rangle = a_X (\varphi_{gi}, \varphi_{gj})$. If $g_1, \ldots, g_m$ are linearly independent, Eq. (7.7.13) is solvable about $\boldsymbol{\lambda}_{k+1}$. Moreover, if the active constraint functions f_{i_1}, ... , $f_{i_{|I_{\mathrm{A}}|}}$ all have the value zero, from the fact that they hold even when an arbitrary real number is multiplied all of $\varphi_{g i_1}$, ... , $\varphi_{g i_{|I_{\mathrm{A}}|}}$, it becomes possible to obtain $\boldsymbol{\lambda}_{k+1}$ even when the step size $\|\varphi_g\|_X$ is not appropriately set. In addition, as adapted in Chap. 3, Eq. (7.7.13) is solved possibly several times, removing each time the constraints where the associated Lagrange multiplier is negative (active set method), although if we set appropriate constraint functions which have trade-off property with respect to an objective function, the Lagrange multipliers become always positive.

Using the definitions so far, a simple Algorithm 3.7.2 shown in Section 3.7.1 can be applied. In this case, the following changes can be made:

(1) Replace the design variable $\boldsymbol{x}$ and its variation $\boldsymbol{y}$ as ϕ and φ respectively.

(2) Replace Eq. (3.7.10) with Eq. (7.7.11).

(3) Replace Eq. (3.7.11) with Eq. (7.7.12).

(4) Replace Eq. (3.7.12) with Eq. (7.7.13).

Furthermore, when considering a complicated Algorithm 3.7.6 with parameter adjustments shown in Section 3.7.2, the following needs to be taken into consideration:

(i) Functionality for determining c_a, so that the initial step size becomes $\|\varphi_g\| = \epsilon_g$ with a given value ϵ_g.

(ii) When the design variable is updated to ϕ_{k+1}, the functionality of amending $\boldsymbol{\lambda}_{k+1} = (\lambda_{k+1\,i})_{i \in I_{\mathrm{A}}(\phi_{k+1})}$ so that $|f_i(\phi_{k+1})| \leq \epsilon_i$ and $\lambda_{k+1\,i} > 0$ with respect to $i \in I_{\mathrm{A}}(\phi_{k+1})$ are satisfied.

(iii) The functionality to make suitable the admissible values $\epsilon_1, \ldots, \epsilon_m$ of the constraint functions $f_1, \ldots, f_m$ with respect to the convergence determination value ϵ_0 of the objective function f_0.

(iv) Functionality for adjusting the step size $\|\varphi_g\|$ so that global convergence is guaranteed.

With respect to the aforementioned (i), the content shown in Section 3.7.2 will hold as it is by replacing $\boldsymbol{y}$ by φ.

The same is true for (ii) above. In other words, Algorithm 3.7.6 can be used exactly by replacing the update of $\boldsymbol{\lambda}_{k+1}$ using Eq. (3.7.21) of the Newton–Raphson method by

$$(\delta\lambda_j)_{j \in I_{\mathrm{A}}} = -\left(\langle g_i(\boldsymbol{\lambda}_{k+1\,l}), \varphi_{gj}(\boldsymbol{\lambda}_{k+1\,l})\rangle\right)^{-1}_{(i,j)\in I_{\mathrm{A}}^2} \left(f_i(\boldsymbol{\lambda}_{k+1\,l})\right)_{i \in I_{\mathrm{A}}}. \tag{7.7.14}$$

Moreover, with respect to (iii) above, the method for replacing ϵ_i so that Eq. (3.7.25) is satisfied has already been incorporated into Algorithm 3.7.6.

Furthermore, with respect to (iv) above, the following type of replacement would allow Algorithm 3.7.6 to be used as it is. The Lagrange function with respect to the abstract optimum design problem (Problem 7.3.1) is given by $\mathscr{L}(\phi, \boldsymbol{\lambda})$ of Eq. (7.7.1). In this case, the Armijo criterion becomes

$$\mathscr{L}(\phi_k + \varphi_g, \boldsymbol{\lambda}_{k+1}) - \mathscr{L}(\phi_k, \boldsymbol{\lambda}_k)$$

$$\leq \xi \left\langle g_0\left(\phi_k\right) + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{ki} g_i\left(\phi_k\right), \varphi_g \right\rangle \tag{7.7.15}$$

with respect to a $\xi \in (0,1)$. The Wolfe criterion is given with respect to a μ $(0 < \xi < \mu < 1)$ by

$$\begin{aligned} &\mu \left\langle g_0\left(\phi_k\right) + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{ki} g_i\left(\phi_k\right), \varphi_g \right\rangle \\ &\quad \leq \left\langle g_0\left(\phi_k + \varphi_g\right) + \sum_{i \in I_{\mathrm{A}}(\phi_{k+1})} \lambda_{k+1\, i} g_i\left(\phi_k + \varphi_g\right), \varphi_g \right\rangle. \end{aligned} \tag{7.7.16}$$

Using these replacements, Algorithm 3.7.6 can be applied with respect to the abstract optimum design problem (Problem 7.3.1). In this case, the following changes are made in addition to (1) to (4) above:

(5) Replace the Armijo criterion Eq. (3.7.26) with Eq. (7.7.15).

(6) Replace the Wolfe criterion Eq. (3.7.27) with Eq. (7.7.16).

(7) Replace Eq. (3.7.21) by Eq. (7.7.14) to update $\boldsymbol{\lambda}$ by the Newton–Raphson method.

In order for this algorithm to function well, there is a need for the points made in Remark 7.6.3 to be satisfied. If these are not satisfied, there is a possibility of numerical instability arising. In order to prevent these situations, there is a need to ensure that the new design variable is always included within the admissible set $\mathcal{D}$ by adding an appropriate process after the design variable is updated.

7.7.2 Newton Method for Constrained Problems

If the second-order Fréchet derivatives of the cost functions can be obtained, it is possible to change the gradient method with respect to a constrained problem to a Newton method with respect to a constrained problem. Here, let us use the abstract Newton method (Problem 7.6.4) in order to make Problem 3.8.1 in Chap. 3 abstract.

Problem 7.7.2 (Newton method for constrained problems) At a trial point $\phi_k \in \mathcal{D}$ of Problem 7.3.1 satisfying the inequality constraints, the Lagrange multiplier $\boldsymbol{\lambda}_k \in \mathbb{R}^{|I_{\mathrm{A}}|}$ is assumed to satisfy Eq. (7.7.8) to Eq. (7.7.10) (where $k+1$ is viewed as k). Moreover, $f_0\left(\phi_k\right)$, $f_{i_1}\left(\phi_k\right) = 0$, $\ldots$, $f_{i_{|I_{\mathrm{A}}|}}\left(\phi_k\right) = 0$ and $g_0\left(\phi_k\right)$, $g_{i_1}\left(\phi_k\right)$, $\ldots$, $g_{i_{|I_{\mathrm{A}}|}}\left(\phi_k\right) \in X'$ as well as $h_0\left(\phi_k\right)$, $h_{i_1}\left(\phi_k\right)$, $\ldots$, $h_{i_{|I_{\mathrm{A}}|}}\left(\phi_k\right) \in \mathcal{L}^2\left(X \times X; \mathbb{R}\right)$ are taken to be known and

$$h_{\mathscr{L}}\left(\phi_k\right) = h_0\left(\phi_k\right) + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{ik} h_i\left(\phi_k\right). \tag{7.7.17}$$

Moreover, let $a_X : X \times X \to \mathbb{R}$ be a coercive and bounded bilinear form on X. In this case, obtain $\phi_{k+1} = \phi_k + \varphi_g$ which satisfies

$$q(\varphi_g) = \min_{\varphi \in X} \left\{ q(\varphi) = \frac{1}{2}\left(h_{\mathscr{L}}(\phi_k)[\varphi, \varphi] + a_X(\varphi, \varphi)\right) + \langle g_0(\phi_k), \varphi \rangle \right.$$
$$\left. + f_0(\phi_k) \;\middle|\; f_i(\phi_k) + \langle g_i(\phi_k), \varphi \rangle \leq 0 \text{ for } i \in I_{\mathrm{A}}(\phi_k) \right\}.$$

□

In Problem 7.7.2, the a_X was introduced in order to compensate for the lack of coerciveness and boundedness of $h_{\mathscr{L}}(\phi_k)$ on X and to ensure the regularity of φ_{gi}.

Problem 7.7.2 is classified to be a second-order optimization problem. When $h_{\mathscr{L}}(\phi_k)[\varphi, \varphi] + a_X(\varphi, \varphi)$ is a coercive and bounded bilinear form on X, Problem 7.7.2 becomes a convex optimization problem. It is not necessarily the case. However, a φ_g satisfying the KKT conditions shown next is a candidate for the minimum point with respect to Problem 7.7.2. Focusing on this, let us look at what has been learned in Section 3.8 in order to think of the solution to Problem 7.7.2.

It is assumed that KKT conditions at the minimum point φ_g of Problem 7.7.2 hold. In other words, the following holds with respect to an arbitrary $\psi \in X$:

$$h_{\mathscr{L}}(\phi_k)[\varphi, \psi] + a_X(\varphi, \psi) + \langle g_0(\phi_k), \psi \rangle + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{k+1\,i} \langle g_i(\phi_k), \psi \rangle = 0, \tag{7.7.18}$$

$$f_i(\phi_{k+1}) = f_i(\phi_k) + \langle g_i(\phi_k), \varphi_g \rangle \leq 0 \quad \text{for } i \in I_{\mathrm{A}}(\phi_k), \tag{7.7.19}$$

$$\lambda_{k+1\,i}\left(f_i(\phi_k) + \langle g_i(\phi_k), \varphi_g \rangle\right) = 0 \quad \text{for } i \in I_{\mathrm{A}}(\phi_k), \tag{7.7.20}$$

$$\lambda_{k+1\,i} \geq 0 \quad \text{for } i \in I_{\mathrm{A}}(\phi_k). \tag{7.7.21}$$

$(\varphi_g, \boldsymbol{\lambda}_{k+1}) \in X \times \mathbb{R}^{|I_{\mathrm{A}}|}$ satisfying these can be obtained as follows.

In the gradient method (Sect. 7.7.1) with constraints, φ_{g0}, φ_{i_1}, …, $\varphi_{i_{|I_{\mathrm{A}}|}}$ were taken to be the solution for the abstract gradient method. Here, these are replaced by the solution of the abstract Newton method. Problem 7.6.4 is rewritten as follows. "Let the known functions in Problem 7.7.2 be given. Find $\varphi_{gi} \in X$ which satisfy the following with respect to an arbitrary $\varphi \in X$:

$$h_{\mathscr{L}}(\phi_k)[\varphi_{gi}, \varphi] + a_X(\varphi_{gi}, \varphi) = -\langle g_i(\phi_k), \varphi \rangle .\text{"} \tag{7.7.22}$$

Here, φ_g defined by Eq. (7.7.12) satisfies Eq. (7.7.18). On the other hand, Eq. (7.7.19) becomes Eq. (7.7.13). Hence, if Eq. (7.7.13) is used to obtain $\boldsymbol{\lambda}_{k+1}$, Eq. (7.7.19) is established and the KKT conditions at the minimal point φ_g of Problem 7.7.2 hold. In this case, Eq. (7.7.20) and Eq. (7.7.21) are satisfied by choosing $I_{\mathrm{A}}(\phi_{k+1})$ appropriately in the algorithm using the active set method.

Using the definitions so far, Algorithm 3.8.4 shown in Section 3.8.1 can be applied. In this case, it will be changed in the following way:

(1) Replace the design variable $\boldsymbol{x}$ and its fluctuation $\boldsymbol{y}$ by ϕ and φ respectively.

(2) Replace Eq. (3.8.9) with Eq. (7.7.22).

(3) Replace Eq. (3.7.11) with Eq. (7.7.12).

(4) Replace Eq. (3.7.12) with Eq. (7.7.13).

Furthermore, the abstract Newton method when the second-order Fréchet derivative of $f_i(\phi)$ is obtained as a Hesse gradient can be illustrated as follows. Equations (7.7.17) and (7.7.22) are replaced with

$$g_{\mathrm{H}\mathscr{L}}(\phi_k, \bar{\varphi}_g) = g_{\mathrm{H}0}(\phi_k, \bar{\varphi}_g) + \sum_{i \in I_{\mathrm{A}}(\phi_k)} \lambda_{ik} g_{\mathrm{H}i}(\phi_k, \bar{\varphi}_g), \tag{7.7.23}$$

$$a_X(\varphi_{gi}, \varphi) = -\langle (g_i(\phi_k) + g_{\mathrm{H}\mathscr{L}}(\phi_k, \bar{\varphi}_g)), \varphi \rangle, \tag{7.7.24}$$

respectively. Using the definitions, the following step is added:

(5) Replace Eq. (3.8.11) with Eq. (7.7.24).

In this way, the difference between the gradient method with respect to a constrained problem and the Newton method is just that $a_X(\cdot, \cdot)$ of the abstract gradient method is replaced with $h_i(\phi_k)[\cdot, \cdot] + a_X(\cdot, \cdot)$ or $g_i(\phi_k)$ is replaced with $g_i(\phi_k) + g_{\mathrm{H}i}(\phi_k, \bar{\varphi}_g)$. However, with this method, a second-order derivative of a cost function is used. Hence, it is hoped that the characteristics of the Newton method mentioned in Remark 3.5.4 will hold. However, as explained in Remark 3.8.2, because the constraint condition is approximated to be up to first-order derivative, there is a need to be careful with the step size when the non-linearity of the constraint functions is strong.

Furthermore, with respect to the methods for achieving coerciveness of the bilinear form and adding the functionality for adjusting the step size, the explanation provided in Section 3.8.2 is still valid here.

Whether such a Newton method can be used or not depends on whether the calculation $h_i(\phi_k)[\cdot, \cdot]$ or $g_{\mathrm{H}i}(\phi_k, \bar{\varphi}_g)$ is possible or not. Let us look at the specific calculation methods of these in Chaps. 8 and 9.

7.8 Summary

In Chap. 7, abstract problems, which may be common to the optimum design problems targeting the topology and shape with respect to domain of a boundary value problem of partial differential equation such as shown in Chaps. 8 and 9, were constructed and their solutions were looked at. The following are the key points:

(1) Real Hilbert spaces are chosen for the linear spaces of a design variable and state variable (Sect. 7.1). This is because the abstract gradient method is defined on a Hilbert space.

(2) An abstract optimum design problem was defined with an abstract variational problem as a state determination problem (Sect. 7.2) and using cost functions defined via functionals of the design variable (Sect. 7.3) and the state variable (solution of state determination problem).

(3) With respect to an abstract optimum design problem, the derivative of a cost function can be obtained via the adjoint variable method (Sect. 7.5.1) or the Lagrange multiplier method (Sect. 7.5.2). Moreover, the second-order derivative of a cost function can be obtained by substituting the derivative of the solution of the state determination problem when the equality constraints of the state determination problem are satisfied in the second-order derivative of the Lagrange function (Sect. 7.5.3).

(4) The abstract gradient method is defined on a real Hilbert space (Sect. 7.6.1). The unique existence of the solution of the abstract gradient method is shown by the Lax–Milgram theorem. Moreover, the solution is on the downward slope of the cost function (Theorem 7.6.2). Furthermore, the abstract Newton method is defined as a method in which the bilinear form in the abstract gradient method is replaced by the sum of a second-order derivative of the cost function and a bilinear form which compensates for the coerciveness and boundedness of the second-order derivative (Sect. 7.6.2).

(5) The solution to the abstract optimum design problem is constructed with the same framework as the gradient method and Newton method with respect to constrained problems shown in Chap. 3 (Sect. 7.7.1 and Sect. 7.7.2).

Let us mention a few books which are useful references for this chapter. Chap. 5 of [3] is useful in relation to the Lagrange multiplier method on a function space. The paper [1] and Section 4.4 of [4] are useful with respect to the gradient method on function spaces.

References

[1] Cea, J. Numerical methods of shape optimal design. In Haug, E. J. and Cea, J., editors, *Optimization of Distributed Parameter Structures*, volume 2, pages 1049–1088. Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, 1981.

[2] Haslinger, J. and Mäkinen, R. A. E. *Introduction to Shape Optimization: Theory, Approximation, and Computation.* SIAM, Philadelphia, 2003.

[3] Jahn, J. *Introduction to the Theory of Nonlinear Optimization. 2nd Rev. Ed.* Springer, Berlin; New York, 1996.

[4] Pironneau, O. *Optimal Shape Design for Elliptic Systems.* Springer, New York, 1984.