Contents

Contents **1**

1 Basics of Optimal Design **1**

1.1 Optimal Design Problem for a Stepped One-Dimensional Linear Elastic Body . . . 1
1.1.1 State Determination Problem . . . 3
1.1.2 An Optimal Design Problem . . . 8
1.1.3 Cross-Sectional Derivatives . . . 9
1.1.4 The Substitution Method . . . 11
1.1.5 The Direct Differentiation Method . . . 12
1.1.6 The Adjoint Variable Method . . . 14
1.1.7 Optimality Conditions . . . 19
1.1.8 Numerical Example . . . 21
1.2 Comparison of the Direct Differentiation Method and the Adjoint Variable Method . . . 22
1.2.1 The Direct Differentiation Method . . . 23
1.2.2 The Adjoint Variable Method . . . 25
1.3 An Optimal Design Problem of a Branched One-Dimensional Stokes Flow Field . . . 26
1.3.1 State Determination Problem . . . 27
1.3.2 An Optimal Design Problem . . . 29
1.3.3 Cross-Sectional Derivatives . . . 30
1.3.4 Optimality Conditions . . . 34
1.3.5 Numerical Example . . . 35
1.4 Summary . . . 36
1.5 Practice Problems . . . 37

References **41**

Chapter 1

Basics of Optimal Design

The main topic of this book is optimal design. In order to understand the mathematical structures involved in our study, we will begin by examining two simple problems. Upon finishing this book, the reader should be able to understand that even shape optimization problems of continuum structures possess the same formulation as the problems dealt with in this chapter. Moreover, even if the target continuum is changed from linear elastic body or flow field dealt in this book, the reader will recognize that their corresponding shape optimization problems maintain the fundamental structures introduced in this book.

Linear elastic solids and the Stokes flow field constitute the continuum used in our applications. We will construct optimal design problems related to one-dimensional linear elastic bodies and the one-dimensional Stokes flow field. We also show how to obtain optimality conditions for these problems. The conditions that we obtain will again be encountered in Chap. 9, where we will treat shape optimization problems of domain variation type for linear elastic bodies and the Stokes flow field, in two and three dimensions.

1.1 Optimal Design Problem for a Stepped One-Dimensional Linear Elastic Body

In order to understand the structure of optimal design problems, let us consider a mechanical system consisting of a one-dimensional linear elastic

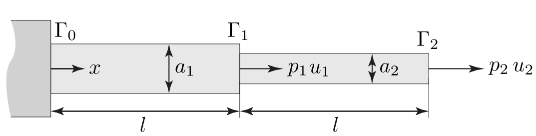

Fig. 1.1: 1D linear elastic body with two cross-sectional areas.

body with two cross-sectional areas, such as is shown in Fig. 1.1. The reason we refer to this system as one-dimensional is because, although in reality it is a three-dimensional body having a cross-sectional area and length, the x coordinate is taken in the length direction and, as is shown later, the displacement of the elastic body can be described as a function of x. In other words, it is assumed that the displacement is given as a function on a one-dimensional vector space.[1] Furthermore, as will also be shown later, the linearity of the elastic body arises from the fact that, under the assumptions that the constitutive law is given by Hooke's law and that the deformation is infinitesimal, the outer force is a linear function of displacement.

Let us now define several constants and variables in detail. In particular, l is a constant representing length, a_1 and a_2 are cross-sectional areas, and $\boldsymbol{a} = (a_1, a_2)^\top \in \mathbb{R}^2$ is a vector with two components. In this book, $\mathbb{R}$ denotes the set of all real numbers and $(\,\cdot\,)^\top$ represents the transpose. Moreover, bold lower-case Latin and Greek letters will be used in mathematical equations to represent finite-dimensional vectors. We remark that there exist positive constants a_{01} and a_{02} satisfying $a_i \geq a_{0i}$ for $i \in \{1, 2\}$. This can be expressed as $\boldsymbol{a} \geq \boldsymbol{a}_0$, where $\boldsymbol{a}_0 = (a_{01}, a_{02})^\top \in \mathbb{R}^2$. Similarly, letting p_1 and p_2 denote external forces acting on cross-sections Γ_1 and Γ_2, and u_1 and u_2 be the corresponding displacements, we write $\boldsymbol{p} = (p_1, p_2)^\top \in \mathbb{R}^2$ and $\boldsymbol{u} = (u_1, u_2)^\top \in \mathbb{R}^2$.

Now let us consider an optimal design problem, where l and $\boldsymbol{p}$ are assumed to be given. We treat $\boldsymbol{a}$ as the design variable, due to the fact that once the cross-section $\boldsymbol{a}$ is determined, the system we are attempting to design is uniquely determined. When a system is specified by determining $\boldsymbol{a}$, the variable $\boldsymbol{u}$ which satisfies the system's state equation is called a state variable. In this book, the problem of finding the state variable is referred to as the state determination problem. The state determination problem that we are currently considering will be examined in detail in Sect. 1.1.1.

When the design variable $\boldsymbol{a}$ and the state variable $\boldsymbol{u}$ are given, we define real-valued functions of $\boldsymbol{a}$ and $\boldsymbol{u}$ representing the performance of the system. Such functions are called cost functions. In Sect. 1.1.2, considering that our current system is a structure supporting an external force, a function for measuring deformation and a function for imparting a volume constraint are chosen as the cost functions. The cost functions are then used to formulate the optimal design problem through defining objective and constraint functions.

The condition which holds when an optimal solution is used in an optimal design problem constructed in this manner is called an optimality condition. An optimality condition for the current one-dimensional elastic body problem is presented in Sect. 1.1.7. For this reason, the derivative of the cost function with respect to the variation of a design variable is defined in Sect. 1.1.3, and ways to obtain them are considered from Sect. 1.1.4 to 1.1.6. These results should perhaps be presented after an explanation has been given regarding a main theorem of optimization theory, which is given in Chap. 2. Nevertheless,

[1] The finite-dimensional vector space considered here can also be called a Euclid space. Moreover, vector space is synonymous with linear space (see Definition 4.2.1).

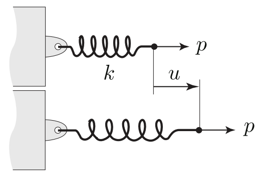

Fig. 1.2: A spring system with a single degree of freedom. The top figure represents the system's initial state, and the bottom figure illustrates the balanced state of forces.

this book sets a priority on obtaining a practical understanding of how to make use of this theorem in optimization theory.

1.1.1 State Determination Problem

Let us go through the process involved in constructing an optimal design problem. We will begin by defining a mechanical system which is the target of the design. When the design variables are specified, this system reverts to a mechanical problem constructed by standard equations of motion and boundary conditions. We refer to this problem as a state determination problem, and we examine its construction based on mechanical principles. Readers who are knowledgeable in the field of mechanics are invited to skip this section.

Before analyzing our one-dimensional elastic body, we review the fact that the equilibrium equation of forces can be obtained from minimality conditions of a potential energy [7]. The next exercise concerns the definition of potential energy.

Exercise 1.1.1 (Potential energy of a simple spring system)
Consider a spring system with a single degree of freedom, such as is shown in Fig. 1.2. Let k and p denote positive numbers representing the spring constant and an external force, respectively. Moreover, let the external force be conservative, that is a constant force generated anywhere on $\mathbb{R}$ and $u \in \mathbb{R}$ be the displacement when the spring and the external force are in balance. Assume that

$$ku - p = 0$$

holds. Find the potential energy of the spring system when $u = 0$ is set as the point of reference. □

Answer In mechanics, potential energy is defined as an amount of energy which expresses the capacity to do work. When $u = 0$ is the point of reference, the potential energy is obtained by integrating the unbalanced force $kv - p$ (v denotes an intermediate displacement) over a displacement from 0 to u:

$$\pi(u) = \int_0^u (kv - p)\,\mathrm{d}v = \frac{1}{2}ku^2 - pu. \tag{1.1.1}$$

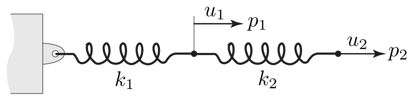

Fig. 1.3: A two-degree-of-freedom spring system.

□

The first and the second terms on the right-hand side of Eq. (1.1.1) are called the internal potential energy and the external potential energy, respectively. Notice that the internal potential energy is the part which acquires the ability to do work (potential) and is therefore positive. On the other hand, the external potential energy is the part which has already done work (the directions of the force and the displacement are the same), and so it is negative. Here we remark that, although potential energy is related to the stored energy (Hamiltonian) which appears in the law of conservation of energy (see Practice **4.3**), the two are in fact different entities.

If a potential energy of π is obtained, the force equilibrium equation is given by the stationary condition of the potential energy:

$$\frac{\mathrm{d}\pi}{\mathrm{d}u} = ku - p = 0.$$

The fact that the potential energy is minimized at this point is a consequence of the following:

$$\frac{\mathrm{d}^2\pi}{\mathrm{d}u^2} = k > 0.$$

Once the notion of potential energy is understood, one can also assess the potential energy of spring systems with two degrees of freedom. As in the following exercise, the same idea can be applied to a force equilibrium equation of a two-degree-of-freedom spring system.

Exercise 1.1.2 (Potential energy in a 2DOF spring system)
Consider a spring system consisting of two degrees of freedom, such as is shown in Fig. 1.3. Here, k_1 and k_2 are positive constants representing the spring constants, $\boldsymbol{p} = (p_1, p_2)^\top \in \mathbb{R}^2$ is a constant vector representing external forces, and $\boldsymbol{u} = (u_1, u_2)^\top \in \mathbb{R}^2$ denotes the displacement when in a balanced state with $\boldsymbol{p}$. In this case, obtain the potential energy when $\boldsymbol{u} = \mathbf{0}_{\mathbb{R}^2}$ ($\mathbf{0}_{\mathbb{R}^2}$ denotes $(0,0)^\top$ in this book) is the point of reference. Also, find the force equilibrium equation using the stationary condition of the potential energy. □

Answer The potential energy of the system can be obtained by adding together the internal and external potential energies:

$$\pi(\boldsymbol{u}) = \frac{1}{2}k_1 u_1^2 + \frac{1}{2}k_2 (u_2 - u_1)^2 - (p_1 u_1 + p_2 u_2).$$

One has the stationarity condition of potential energy:

$$\frac{\partial \pi}{\partial u_1} = k_1 u_1 - k_2 (u_2 - u_1) - p_1 = 0,$$
$$\frac{\partial \pi}{\partial u_2} = k_2 (u_2 - u_1) - p_2 = 0,$$

which can be used to obtain the force equilibrium equation. These equations can be written as

$$\begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}. \tag{1.1.2}$$

□

The fact that a vector $\boldsymbol{u}$ satisfying the stationary condition Eq. (1.1.2) minimizes the potential energy π can be shown using the approach of Exercise 2.5.8. This fact will be omitted for now.

We have confirmed that the force equilibrium equation can be obtained via minimality conditions of the potential energy, so now let us apply these conditions to the one-dimensional linear elastic body shown in Fig. 1.1. First, similar to Exercise 1.1.2, the external potential energy can be given as

$$\pi_{\mathrm{E}} (\boldsymbol{u}) = -\boldsymbol{p} \cdot \boldsymbol{u}. \tag{1.1.3}$$

In this book, $\boldsymbol{p} \cdot \boldsymbol{u} = \boldsymbol{p}^\top \boldsymbol{u}$ represents the inner product of a finite dimensional vector space.

Next, let us find the internal potential energy. Let $x \in \mathbb{R}$ be the coordinate in the length direction, where the cross-section Γ_0 in Fig. 1.1 is taken as the origin. In this case, the displacement at $x \in [0, 2l]$ is assumed to be given by

$$u(x) = \begin{cases} u_1 \dfrac{x}{l} & x \in [0, l) \\ (u_2 - u_1) \dfrac{x}{l} + 2u_1 - u_2 & x \in [l, 2l] \end{cases} \tag{1.1.4}$$

from the linearized elasticity assumption explained in the beginning of Sect. 1.1. In other words, it is assumed that three-dimensional deformations are not considered. Here, $[0, 2l]$ represents the interval $\{x \in \mathbb{R} \mid 0 \le x \le 2l\}$.

We would now like to interrupt our discussion in order to explain principles regarding the representation of sets and functions used in this book. Sets are defined in the format $\{x \in \mathbb{R} \mid 0 \le x \le 2l\}$, where the linear space (defined in Chap. 4) or an underlying set is written in the position of $\mathbb{R}$. Conditions satisfied by elements of the set are written after the $\mid$ symbol. In particular, $[0, l)$ represents the interval $\{x \in \mathbb{R} \mid 0 \le x < l\}$ and $(0, l)$ represents $\{x \in \mathbb{R} \mid 0 < x < l\}$. In Eq. (1.1.4), $u(x)$ is continuous at $x = l$ and therefore $[0, l)$ of Eq. (1.1.4) can be written as $[0, l]$ or as $(0, l)$. Hence, in this book, the domain of definition of a function is defined to be an open set (see Appendix Sect. A.1.1), and the function's boundary values are defined using properties of continuity (called the trace of the function (Theorem 4.4.2)). On the other hand, taking Eq. (1.1.4) as an example, the domain and range of the function

$u(x)$ are expressed using the notation $u : (0, 2l) \to \mathbb{R}$, where $\to$ designates that the mapping is from the domain $(0, 2l)$ into the range of real numbers $\mathbb{R}$. When specifying elements, we will sometimes write $u(x) : (0, 2l) \ni x \mapsto u \in \mathbb{R}$. Becoming too caught up with the wording of functions or variables in this book may lead to confusion, because functions themselves become variables from Chap. 4 onwards. Therefore, let us remember that the mapping notation is important in such cases.

Let us now return to our original discussion. In mechanics, equations relating variables representing phenomena such as force and displacement, or temperature and heat are called constitutive equations or constitutive laws. Hooke's law is used in the case of linear elastic bodies. Hooke's law relates the strain of a material (its rate of deformation)

$$\varepsilon(u) = \frac{\mathrm{d}u}{\mathrm{d}x} \tag{1.1.5}$$

with its stress $\sigma(u)$ (force acting per unit area) via

$$\sigma(u) = e_{\mathrm{Y}} \varepsilon(u). \tag{1.1.6}$$

Here, e_{Y} is assumed to be given by a material-specific positive constant called the modulus of longitudinal elasticity, or Young's modulus. In the one-dimensional linear elastic body of Fig. 1.1, it may be assumed that e_{Y} is given by a discontinuous function such as $e_{\mathrm{Y}} : (0, 2l) \to \mathbb{R}$, but for the sake of simplicity we shall assume that it is given by a positive real constant. Furthermore, the mechanical quantity defined using the stress and the strain:

$$w(u) = \frac{1}{2}\sigma(u)\,\varepsilon(u) \tag{1.1.7}$$

is called the strain energy density (internal potential energy density or elastic potential energy density). The fact that w is an energy per unit volume can also be confirmed from the fact that its units are $[\mathrm{Nm/m^3}]$ in the international system of units (SI). Using these definitions, the internal potential energy of the one-dimensional linear elastic body in Fig. 1.1 is given by

$$\pi_{\mathrm{I}}(\boldsymbol{u}) = \int_0^l w(u)\, a_1 \,\mathrm{d}x + \int_l^{2l} w(u)\, a_2 \,\mathrm{d}x. \tag{1.1.8}$$

Since the internal and external potential energies of the one-dimensional elastic body in Fig. 1.1 were obtained using Eq. (1.1.8) and Eq. (1.1.3), the total potential energy with $\boldsymbol{u} = \mathbf{0}_{\mathbb{R}^2}$ as a reference point is given by

$$\begin{aligned}\pi(\boldsymbol{u}) &= \pi_{\mathrm{I}}(\boldsymbol{u}) + \pi_{\mathrm{E}}(\boldsymbol{u}) \\ &= \frac{1}{2}\frac{e_{\mathrm{Y}}}{l} a_1 u_1^2 + \frac{1}{2}\frac{e_{\mathrm{Y}}}{l} a_2 (u_2 - u_1)^2 - p_1 u_1 - p_2 u_2. \end{aligned} \tag{1.1.9}$$

Therefore, the stationary condition of π is expressed as

$$\frac{\partial \pi}{\partial u_1} = \frac{e_{\mathrm{Y}}}{l} a_1 u_1 - \frac{e_{\mathrm{Y}}}{l} a_2 (u_2 - u_1) - p_1 = 0,$$

$$\frac{\partial \pi}{\partial u_2} = \frac{e_{\mathrm{Y}}}{l} a_2 (u_2 - u_1) - p_2 = 0,$$

which can also be written as

$$\frac{e_{\mathrm{Y}}}{l} \begin{pmatrix} a_1 + a_2 & -a_2 \\ -a_2 & a_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}. \tag{1.1.10}$$

This leads us to the problem of determining the displacement when external forces act on the one-dimensional linear elastic body in Fig. 1.1.

Problem 1.1.3 (Stepped 1D linear elastic body) Let $l \in \mathbb{R}$, $e_{\mathrm{Y}} \in \mathbb{R}$, $\boldsymbol{p} \in \mathbb{R}^2$ and $\boldsymbol{a} \in \mathbb{R}^2$ be given with respect to the one-dimensional linear elastic body of Fig. 1.1. Find the displacement $\boldsymbol{u} \in \mathbb{R}^2$ that satisfies

$$\boldsymbol{K}(\boldsymbol{a})\, \boldsymbol{u} = \boldsymbol{p}, \tag{1.1.11}$$

where Eq. (1.1.11) of course represents Eq. (1.1.10). □

In this book, matrices are expressed using bold capital Latin and Greek letters, such as $\boldsymbol{K}$.

Anticipating future developments, let us take a look at an alternative way of expressing Problem 1.1.3. With respect to Problem 1.1.3,

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{v} \cdot (-\boldsymbol{K}(\boldsymbol{a})\, \boldsymbol{u} + \boldsymbol{p}) \tag{1.1.12}$$

will be called a Lagrange function for a state determination problem (defined in Chap. 2). Here, $\boldsymbol{u} \in \mathbb{R}^2$ is not necessarily the solution of Problem 1.1.3 and $\boldsymbol{v} \in \mathbb{R}^2$ has been introduced as a Lagrange multiplier with respect to Eq. (1.1.11). The Lagrange multiplier with respect to a state equation is also referred to as an adjoint variable. Here, $\boldsymbol{u} \in \mathbb{R}^2$, which satisfies

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}) = 0 \tag{1.1.13}$$

for all $\boldsymbol{v} \in \mathbb{R}^2$, has the same value as the solution of Problem 1.1.3. This is because if Eq. (1.1.11) is to be satisfied, then Eq. (1.1.13) holds for all $\boldsymbol{v} \in \mathbb{R}^2$. The converse also holds.

The condition under which Eq. (1.1.13) is satisfied for all $\boldsymbol{v} \in \mathbb{R}^2$ is called the principle of virtual work. The reason for this is that the potential energy

$$\pi(\boldsymbol{u}) = \frac{1}{2} \boldsymbol{u} \cdot (\boldsymbol{K}(\boldsymbol{a})\, \boldsymbol{u}) - \boldsymbol{p} \cdot \boldsymbol{u}$$

has a stationary condition with respect to an arbitrary variation $\mathrm{d}\boldsymbol{u} \in \mathbb{R}^2$ of $\boldsymbol{u}$ (the virtual displacement), given by

$$\mathrm{d}\pi(\boldsymbol{u}) = \mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \mathrm{d}\boldsymbol{u}) = 0.$$

1.1.2 An Optimal Design Problem

Having defined the state determination problem, let us now use it to construct an optimal design problem. Let us first define the cost function. With respect to the solution $\boldsymbol{u}$ of the state determination problem (Problem 1.1.3), the following quantity will be referred to as the mean compliance:

$$f_0(\boldsymbol{u}) = \begin{pmatrix} p_1 & p_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \boldsymbol{p} \cdot \boldsymbol{u}. \tag{1.1.14}$$

Here, f_0 is equivalent to the mechanical quantity known as external work. Nevertheless, in Chaps. 8 and 9, a cost function measuring the ease of deformation (compliance) of linear elastic bodies will be given and used to define an extended notion of mean compliance. The naming in those cases is chosen in order to not imply work done by external forces. The fact that f_0 of Eq. (1.1.14) is a real-valued function representing an ease of deformation can be explained as follows. Since $\boldsymbol{u}$ is a vector representing the ease of deformation, it is not simply a real number. Here, if f_0 is thought of as a function weighted by $\boldsymbol{p}$ in order to convert $\boldsymbol{u}$ into a real number, then it can be understood that f_0 is a real-valued function expressing the ease of deformation. Relatedly,

$$f_1(\boldsymbol{a}) = l(a_1 + a_2) - c_1 = \begin{pmatrix} l & l \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} - c_1 \tag{1.1.15}$$

will be referred to as a constraint function with respect to volume. Here, c_1 is a positive constant representing an upper bound on the volume. In this section, f_0 and f_1 are defined as cost functions for an optimal design problem. Throughout this book, cost functions will be denoted by $f_0, f_1, \ldots, f_m$, where f_0 will denote the objective function, and $f_1, \ldots, f_m$ will denote constraint functions.

Let us define an optimal design problem with respect to the one-dimensional linear elastic body in Fig. 1.1 using the previous cost functions as in Problem 1.1.4. Hereinafter, linear spaces of the design variable $\boldsymbol{a}$ and the state variable $\boldsymbol{u}$ will be denoted as $X = \mathbb{R}^2$ and $U = \mathbb{R}^2$, respectively, and are called the linear space of design variables and the linear space of state variables. Moreover, with respect to a constant vector $\boldsymbol{a}_0 = (a_{01}, a_{02})^\top > \boldsymbol{0}_{\mathbb{R}^2}$, we have the admissible set of design variables:

$$\mathcal{D} = \{\boldsymbol{a} \in X \mid \boldsymbol{a} \geq \boldsymbol{a}_0\}. \tag{1.1.16}$$

In this book, capital Latin and Greek letters (including decorative scripts) are used for sets. Symbols relating to the sets X, U and $\mathcal{D}$ will be used with a unified meaning, even in the setting of optimal design problems in function space (beginning in Chap. 7).

Problem 1.1.4 (Mean compliance minimization) Let $X = \mathbb{R}^2$, $U = \mathbb{R}^2$, and $\mathcal{D}$ be given by Eq. (1.1.16). If $f_0(\boldsymbol{u})$ and $f_1(\boldsymbol{a})$ are given by Eq. (1.1.14) and Eq. (1.1.15), respectively, find $\boldsymbol{a}$ satisfying

$$\min_{(\boldsymbol{a},\boldsymbol{u}) \in \mathcal{D} \times U} \{ f_0(\boldsymbol{u}) \mid f_1(\boldsymbol{a}) \leq 0, \text{ Problem 1.1.3} \}.$$

□

We remark that problem 1.1.4 should probably be written as follows: find

$$\boldsymbol{a}^* = \arg\min_{\boldsymbol{a}\in\mathcal{D}} \{ f_0(\boldsymbol{u}) \mid f_1(\boldsymbol{a}) \leq 0,\ \boldsymbol{u} \in U,\ \text{Problem 1.1.3} \}.$$

Here, $\arg\min_{\boldsymbol{x}\in X} f(\boldsymbol{x})$ denotes a point $\boldsymbol{x}$ in the domain X where f attains its minimum value. However, to simplify the expression, it will be written as shown in Problem 1.1.4. Moreover, simple constraints such as $\boldsymbol{a} \geq \boldsymbol{a}_0$ given in Eq. (1.1.16) will sometimes be referred to as side constraints. Later on, solutions satisfying optimality conditions will be sought while disregarding side constraints. A solution is then chosen from those candidates that satisfy the side constraints. Therefore, we assume that $\boldsymbol{a}$ is an interior point of $\mathcal{D}$ ($\boldsymbol{a} \in \mathcal{D}^\circ$) and when some of the side constraints are activated, we include them in the inequality constraints (Practice **1.4**).

Problem 1.1.4 is an optimization problem with an equality and inequality constraint. Optimality conditions satisfied by solutions of this type of problem will be discussed in detail in Chap. 2 and methods for their numerical solutions will be considered in Chap. 3. An explanation regarding the details of these will be omitted at present, and we will look at how the optimality conditions are obtained by following a set of formal procedures. In the next section, a method for obtaining the derivatives of f_0 and f_1 with respect to variations of the design variable $\boldsymbol{a}$ is considered. These results are used at once with the optimality conditions in Sect. 1.1.7.

Moreover, in the numerical solutions of optimum design problems, the calculation of derivatives of cost functions with respect to an arbitrary variation of design variable becomes the pivotal ingredient. In this case, the state determination problem, which will be a boundary value problem of partial differential equations in Chaps. 8 and 9, becomes an equality constraint. Then, an understanding of how to calculate the derivatives of cost functions f_0 in the next subsection will help the reader to understand the method used to obtain the derivatives of cost functions in Chaps. 8 and 9. In order to help the reader's understanding, we use the same notation as in Chaps. 8 and 9 as much as possible.

In the next subsection, we will obtain the derivative of cost function f_0 using three methods and confirm that those results accord. However, in Chaps. 8 and 9, we will use only one of them for convenience.

1.1.3 Cross-Sectional Derivatives

We call derivatives of f_0 and f_1 with respect to the variation of the cross-sectional area $\boldsymbol{a}$ cross-sectional derivatives.

Let us start by considering the cross-sectional derivative of f_1 to which the usual definition of differentiation can be applied. Since f_1 is defined as a function of $\boldsymbol{a}$ in Eq. (1.1.15), its partial derivative with respect to $\boldsymbol{a}$ can be obtained as

$$f_{1\boldsymbol{a}} = \frac{\partial f_1}{\partial \boldsymbol{a}} = \begin{pmatrix} \partial f_1/\partial a_1 \\ \partial f_1/\partial a_2 \end{pmatrix} = \begin{pmatrix} l \\ l \end{pmatrix} = \boldsymbol{g}_1. \tag{1.1.17}$$

In this book, the partial derivative $\partial f_1/\partial \boldsymbol{a}$ will be written as $f_{1\boldsymbol{a}}$. Here, note that $f_{1\boldsymbol{a}}$ is a column vector because, although f_1 is a real number, $\boldsymbol{a}$ is a column vector. On the other hand, the Taylor expansion (Theorem 2.4.2) of f_1 at $\boldsymbol{a}$ with respect to an arbitrary $\boldsymbol{b} \in X$ is

$$\begin{aligned} f_1(\boldsymbol{a}+\boldsymbol{b}) &= f_1(\boldsymbol{a}) + f_1'(\boldsymbol{a})[\boldsymbol{b}] + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}) \\ &= f_1(\boldsymbol{a}) + \boldsymbol{g}_1 \cdot \boldsymbol{b} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}). \end{aligned} \tag{1.1.18}$$

Here $f_1'(\boldsymbol{a})[\boldsymbol{b}]$ represents the first-order variation of f_1 at $\boldsymbol{a}$ with respect to the variation $\boldsymbol{b}$. Moreover, $o(\,\cdot\,)$ denotes the Bachmann–Landau little-o symbol defined by $\lim_{\epsilon\to 0} o(\epsilon)/\epsilon = 0$, where $\|\boldsymbol{b}\|_{\mathbb{R}^2} = \sqrt{|b_1|^2 + |b_2|^2}$. We also remark that $o(\|\boldsymbol{b}\|_{\mathbb{R}^2}) = 0$ with respect to f_1. In view of Eq. (1.1.18), since $f_1'(\boldsymbol{a})[\boldsymbol{b}] = f_{1\boldsymbol{a}} \cdot \boldsymbol{b} = \boldsymbol{g}_1 \cdot \boldsymbol{b}$, we note that $f_1'(\boldsymbol{a})[\boldsymbol{b}]$ is a linear function with respect to $\boldsymbol{b}$. In other words, the corresponding vector $\boldsymbol{g}_1$ of the inner product with respect to $\boldsymbol{b}$ has been found. When the equation can be written as $f_1'(\boldsymbol{a})[\boldsymbol{b}] = \boldsymbol{g}_1 \cdot \boldsymbol{b}$ we say that it is differentiable and that $f_1'(\boldsymbol{a})[\boldsymbol{b}]$ is the cross-sectional derivative of f_1 at $\boldsymbol{a}$. We refer to $\boldsymbol{g}_1 \in \mathbb{R}^2$ as the cross-sectional-area gradient of f_1.

Next, let us consider the cross-sectional derivative of f_0 in a similar way. Although f_0 is a function of $\boldsymbol{u}$ as in Eq. (1.1.14), it is not explicitly a function of $\boldsymbol{a}$. However, $\boldsymbol{u}$ is assumed to satisfy the state equation (Problem 1.1.3) for a given $\boldsymbol{a}$, so that $\boldsymbol{u}$ varies with any variation of $\boldsymbol{a}$. In other words, $\boldsymbol{u}$ is a function of $\boldsymbol{a}$. Let us now write

$$\tilde{f}_0(\boldsymbol{a}) = \{ f_0(\boldsymbol{u}) \mid (\boldsymbol{a},\boldsymbol{u}) \in \mathcal{D} \times U,\ \text{Problem 1.1.3} \}, \tag{1.1.19}$$

and suppose that we have found a linear function $\tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}]$ with respect to $\boldsymbol{b}$ satisfying

$$\tilde{f}_0(\boldsymbol{a}+\boldsymbol{b}) = \tilde{f}_0(\boldsymbol{a}) + \tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}] + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}),$$

where we write $\tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}] = \boldsymbol{g}_0 \cdot \boldsymbol{b}$ for a certain $\boldsymbol{g}_0 \in \mathbb{R}^2$. Then f_0 is said to be differentiable with respect to $\boldsymbol{a}$, $\tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}]$ is called the cross-sectional derivative of f_0 at $\boldsymbol{a}$, and $\boldsymbol{g}_0$ is called the cross-sectional-area gradient.

Furthermore, with respect to a function $\boldsymbol{g}_0 : X \to \mathbb{R}^2$, whenever there exists $\boldsymbol{g}_0'(\boldsymbol{a})[\boldsymbol{b}_2]$ which is linear in $\boldsymbol{b}_2$ satisfying

$$\begin{aligned} \boldsymbol{g}_0(\boldsymbol{a}+\boldsymbol{b}_2)\cdot\boldsymbol{b}_1 &= \boldsymbol{g}_0(\boldsymbol{a})\cdot\boldsymbol{b}_1 + \boldsymbol{g}_0'(\boldsymbol{a})[\boldsymbol{b}_2]\cdot\boldsymbol{b}_1 + o(\|\boldsymbol{b}_2\|_{\mathbb{R}^2}) \\ &= \boldsymbol{g}_0(\boldsymbol{a})\cdot\boldsymbol{b}_1 + \tilde{f}_0''(\boldsymbol{a})[\boldsymbol{b}_1,\boldsymbol{b}_2] + o(\|\boldsymbol{b}_2\|_{\mathbb{R}^2}), \end{aligned}$$

which is expressible as $\boldsymbol{g}_0'(\boldsymbol{a})[\boldsymbol{b}_2] = \boldsymbol{H}_0\boldsymbol{b}_2$ for a certain $\boldsymbol{H}_0 \in \mathbb{R}^{2\times 2}$, then f_0 is second-order differentiable and $\boldsymbol{H}_0 \in \mathbb{R}^{2\times 2}$ is referred to as the Hesse matrix or the Hessian (Definition 2.4.1) of f_0 at $\boldsymbol{a}$. This is equivalent to the condition that $\tilde{f}_0''(\boldsymbol{a})[\boldsymbol{b}_1, \boldsymbol{b}_2]$ is a bilinear function of $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, and $\tilde{f}_0''(\boldsymbol{a})[\boldsymbol{b}_1, \boldsymbol{b}_2]$ is referred to as the second-order cross-sectional derivative. In this book, $\mathbb{R}^{m \times n}$ represents the set of all real matrices consisting of m rows and n columns.

Using these definitions, if $\tilde{f}_0(\boldsymbol{a})$ is second-order differentiable with respect to $\boldsymbol{a}$, the Taylor expansion (Theorem 2.4.2) of $\tilde{f}_0(\boldsymbol{a})$ at $\boldsymbol{a}$ can be written

$$\tilde{f}_0(\boldsymbol{a}+\boldsymbol{b}) = \tilde{f}_0(\boldsymbol{a}) + \boldsymbol{g}_0 \cdot \boldsymbol{b} + \frac{1}{2}\tilde{f}_0''(\boldsymbol{a})[\boldsymbol{b}, \boldsymbol{b}] + o\left(\|\boldsymbol{b}\|_{\mathbb{R}^2}^2\right).$$

Later on, $\boldsymbol{g}_0$ is used under the condition that f_0 takes an extreme value. Moreover, $\tilde{f}_0''(\boldsymbol{a})[\boldsymbol{b}, \boldsymbol{b}]$ (equivalently, $\boldsymbol{H}_0$) will be used in conditions to guarantee that minimum values are obtained. In the next section we will consider how to obtain $\boldsymbol{g}_0$ and $\boldsymbol{H}_0$.

1.1.4 The Substitution Method

Let us now obtain $\boldsymbol{g}_0$ and $\boldsymbol{H}_0$ directly from $\tilde{f}_0(\boldsymbol{a})$ by direct substitution of the state equation into the cost function. We remark that this method cannot be used in more complicated problems. However, we shall use it here in order to verify results obtained from the direct differentiation and adjoint variable methods (shown later).

The solution of the state equation (Eq. (1.1.11)) is obtained as

$$\boldsymbol{u} = \boldsymbol{K}^{-1}(\boldsymbol{a})\,\boldsymbol{p} = \frac{l}{e_{\mathrm{Y}}}\begin{pmatrix} \dfrac{1}{a_1} & \dfrac{1}{a_1} \\ \dfrac{1}{a_1} & \dfrac{1}{a_1} + \dfrac{1}{a_2} \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \frac{l}{e_{\mathrm{Y}}}\begin{pmatrix} \dfrac{p_1+p_2}{a_1} \\ \dfrac{p_1+p_2}{a_1} + \dfrac{p_2}{a_2} \end{pmatrix}. \tag{1.1.20}$$

Since $\tilde{f}_0(\boldsymbol{a})$ is defined by Eq. (1.1.19), the following equation can be obtained:

$$\tilde{f}_0(\boldsymbol{a}) = \boldsymbol{p} \cdot \left(\boldsymbol{K}^{-1}(\boldsymbol{a})\,\boldsymbol{p}\right) = \frac{l}{e_{\mathrm{Y}}}\left(\frac{(p_1+p_2)^2}{a_1} + \frac{p_2^2}{a_2}\right), \tag{1.1.21}$$

from which we get

$$\boldsymbol{g}_0 = \begin{pmatrix} \dfrac{\partial \tilde{f}_0}{\partial a_1} \\ \dfrac{\partial \tilde{f}_0}{\partial a_2} \end{pmatrix} = \begin{pmatrix} \boldsymbol{p} \cdot \left(\dfrac{\partial \boldsymbol{K}^{-1}}{\partial a_1}\boldsymbol{p}\right) \\ \boldsymbol{p} \cdot \left(\dfrac{\partial \boldsymbol{K}^{-1}}{\partial a_2}\boldsymbol{p}\right) \end{pmatrix} = \frac{l}{e_{\mathrm{Y}}}\begin{pmatrix} -\dfrac{(p_1+p_2)^2}{a_1^2} \\ -\dfrac{p_2^2}{a_2^2} \end{pmatrix}. \tag{1.1.22}$$

Similarly, the Hesse matrix is expressed as

$$\boldsymbol{H}_0 = \begin{pmatrix} \dfrac{\partial^2 \tilde{f}_0}{\partial a_1 \partial a_1} & \dfrac{\partial^2 \tilde{f}_0}{\partial a_1 \partial a_2} \\ \dfrac{\partial^2 \tilde{f}_0}{\partial a_2 \partial a_1} & \dfrac{\partial^2 \tilde{f}_0}{\partial a_2 \partial a_2} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \boldsymbol{g}_0}{\partial a_1} & \dfrac{\partial \boldsymbol{g}_0}{\partial a_2} \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{p} \cdot \left(\dfrac{\partial^2 \boldsymbol{K}^{-1}}{\partial a_1 \partial a_1} \boldsymbol{p} \right) & \boldsymbol{p} \cdot \left(\dfrac{\partial^2 \boldsymbol{K}^{-1}}{\partial a_1 \partial a_2} \boldsymbol{p} \right) \\ \boldsymbol{p} \cdot \left(\dfrac{\partial^2 \boldsymbol{K}^{-1}}{\partial a_1 \partial a_2} \boldsymbol{p} \right) & \boldsymbol{p} \cdot \left(\dfrac{\partial^2 \boldsymbol{K}^{-1}}{\partial a_2 \partial a_2} \boldsymbol{p} \right) \end{pmatrix}$$

$$= \frac{l}{e_{\mathrm{Y}}} \begin{pmatrix} \dfrac{2 (p_1 + p_2)^2}{a_1^3} & 0 \\ 0 & \dfrac{2 p_2^2}{a_2^3} \end{pmatrix}. \tag{1.1.23}$$

Whenever $a_1, a_2 > 0$, the eigenvalues of $\boldsymbol{H}_0$ are positive and so $\boldsymbol{H}_0$ is positive definite (Definition 2.4.5). From this property, based on Theorem 2.4.6 (shown later), $\tilde{f}_0 (\boldsymbol{a})$ can be shown to be a convex function (Definition 2.4.3). The convexity of $\tilde{f}_0 (\boldsymbol{a})$ is used as a sufficient condition for showing minimality in Sect. 1.1.7.

1.1.5 The Direct Differentiation Method

Next, let us also obtain $\boldsymbol{g}_0$ and $\boldsymbol{H}_0$ via the direct differentiation method, which utilizes the chain rule of differentiation for composite functions. The details of the direct differentiation method are presented in Sect. 2.6.5.

Note that $\boldsymbol{u}$ is determined with respect to $\boldsymbol{a}$ such that Eq. (1.1.11) is satisfied. Thus, if $\tilde{f}_0$ of Eq. (1.1.19) is Taylor expanded around $\boldsymbol{a}$ we have

$$\begin{aligned} \tilde{f}_0 (\boldsymbol{a} + \boldsymbol{b}) &= f_0 (\boldsymbol{u} (\boldsymbol{a} + \boldsymbol{b})) \\ &= f_0 (\boldsymbol{u} (\boldsymbol{a})) + \frac{\partial f_0}{\partial u_1} \left(\frac{\partial u_1}{\partial a_1} b_1 + \frac{\partial u_1}{\partial a_2} b_2 \right) \\ &\quad + \frac{\partial f_0}{\partial u_2} \left(\frac{\partial u_2}{\partial a_1} b_1 + \frac{\partial u_2}{\partial a_2} b_2 \right) + o (\| \boldsymbol{b} \|_{\mathbb{R}^2}) \\ &= f_0 (\boldsymbol{u} (\boldsymbol{a})) + \begin{pmatrix} p_1 & p_2 \end{pmatrix} \begin{pmatrix} \dfrac{\partial u_1}{\partial a_1} & \dfrac{\partial u_1}{\partial a_2} \\ \dfrac{\partial u_2}{\partial a_1} & \dfrac{\partial u_2}{\partial a_2} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + o (\| \boldsymbol{b} \|_{\mathbb{R}^2}). \end{aligned} \tag{1.1.24}$$

On the other hand, if Eq. (1.1.11) is partially differentiated with respect to a_1, then we have

$$\frac{\partial \boldsymbol{K}}{\partial a_1} \boldsymbol{u} + \boldsymbol{K} \frac{\partial \boldsymbol{u}}{\partial a_1} = \boldsymbol{0}_{\mathbb{R}^2},$$

which can be written as

$$\frac{e_{\mathrm{Y}}}{l} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \frac{e_{\mathrm{Y}}}{l} \begin{pmatrix} a_1 + a_2 & -a_2 \\ -a_2 & a_2 \end{pmatrix} \begin{pmatrix} \dfrac{\partial u_1}{\partial a_1} \\ \dfrac{\partial u_2}{\partial a_1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, it follows that

$$\frac{\partial \boldsymbol{u}}{\partial a_1} = -\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial a_1}\boldsymbol{u} = -\begin{pmatrix} \frac{1}{a_1} & \frac{1}{a_1} \\ \frac{1}{a_1} & \frac{1}{a_1}+\frac{1}{a_2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -\frac{u_1}{a_1} \\ -\frac{u_1}{a_1} \end{pmatrix}. \quad (1.1.25)$$

Similarly, partially differentiating Eq. (1.1.11) with respect to a_2 yields

$$\frac{\partial \boldsymbol{u}}{\partial a_2} = -\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial a_2}\boldsymbol{u} = -\begin{pmatrix} \frac{1}{a_1} & \frac{1}{a_1} \\ \frac{1}{a_1} & \frac{1}{a_1}+\frac{1}{a_2} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{u_2-u_1}{a_2} \end{pmatrix}. \quad (1.1.26)$$

Therefore, upon substituting Eq. (1.1.25) and Eq. (1.1.26) into Eq. (1.1.24), we obtain

$$\begin{aligned} &\tilde{f}_0(\boldsymbol{a}+\boldsymbol{b}) \\ &= f_0(\boldsymbol{u}(\boldsymbol{a})) + \begin{pmatrix} p_1 & p_2 \end{pmatrix}\begin{pmatrix} -\frac{u_1}{a_1} & 0 \\ -\frac{u_1}{a_1} & -\frac{u_2-u_1}{a_2} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}) \\ &= f_0(\boldsymbol{u}(\boldsymbol{a})) + \begin{pmatrix} -\frac{u_1}{a_1}(p_1+p_2) & -\frac{u_2-u_1}{a_2}p_2 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}) \\ &= f_0(\boldsymbol{u}(\boldsymbol{a})) + \boldsymbol{g}_0 \cdot \boldsymbol{b} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}). \end{aligned} \quad (1.1.27)$$

If the solution of the state equation (Eq. (1.1.20)) is used in the previous equation, it becomes apparent that $\boldsymbol{g}_0$ of Eq. (1.1.27) agrees with Eq. (1.1.22). Moreover, if we use the notation $\varepsilon(u_1) = u_1/l$ and $\sigma(u_1) = e_{\mathrm{Y}}\varepsilon(u_1)$ for the strain and stress, then we obtain

$$\boldsymbol{g}_0 = -\frac{e_{\mathrm{Y}}}{l}\begin{pmatrix} u_1^2 \\ (u_2-u_1)^2 \end{pmatrix} = l\begin{pmatrix} -\sigma(u_1)\varepsilon(u_1) \\ -\sigma(u_2-u_1)\varepsilon(u_2-u_1) \end{pmatrix}. \quad (1.1.28)$$

Equation (1.1.28) is an equation in which $\boldsymbol{g}_0$ is expressed as a function of the state variable $\boldsymbol{u}$.

Let us also find the second-order derivative of $\tilde{f}_0$ (the Hesse matrix of $\tilde{f}_0$) with respect to the variation of $\boldsymbol{a}$. If the chain rule of differentiation is used on $\boldsymbol{g}_0$ of Eq. (1.1.28), then we have

$$\begin{aligned} \boldsymbol{g}_0(\boldsymbol{a}+\boldsymbol{b}) &= \boldsymbol{g}_0(\boldsymbol{a}) + \frac{\partial \boldsymbol{g}_0}{\partial \boldsymbol{u}^{\top}}\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{a}^{\top}}\boldsymbol{b} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}) \\ &= \boldsymbol{g}_0(\boldsymbol{a}) + \begin{pmatrix} \frac{\partial g_{01}}{\partial u_1} & \frac{\partial g_{01}}{\partial u_2} \\ \frac{\partial g_{02}}{\partial u_1} & \frac{\partial g_{02}}{\partial u_2} \end{pmatrix}\begin{pmatrix} \frac{\partial u_1}{\partial a_1} & \frac{\partial u_1}{\partial a_2} \\ \frac{\partial u_2}{\partial a_1} & \frac{\partial u_2}{\partial a_2} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + o(\|\boldsymbol{b}\|_{\mathbb{R}^2}). \end{aligned}$$

From this, the Hesse matrix of f_0 can be computed:

$$\begin{aligned}
\boldsymbol{H}_0 &= \frac{\partial \boldsymbol{g}_0}{\partial \boldsymbol{u}^\top} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{a}^\top} \\
&= -\frac{e_{\mathrm{Y}}}{l} \begin{pmatrix} 2u_1 & 0 \\ -2(u_2 - u_1) & 2(u_2 - u_1) \end{pmatrix} \begin{pmatrix} -\dfrac{u_1}{a_1} & 0 \\ -\dfrac{u_1}{a_1} & -\dfrac{u_2 - u_1}{a_2} \end{pmatrix} \\
&= \frac{e_{\mathrm{Y}}}{l} \begin{pmatrix} \dfrac{2u_1^2}{a_1} & 0 \\ 0 & \dfrac{2(u_2 - u_1)^2}{a_2} \end{pmatrix} \\
&= l \begin{pmatrix} \dfrac{2\sigma(u_1)\varepsilon(u_1)}{a_1} & 0 \\ 0 & \dfrac{2\sigma(u_2 - u_1)\varepsilon(u_2 - u_1)}{a_2} \end{pmatrix}.
\end{aligned} \tag{1.1.29}$$

Using the solution of the state equation (Eq. (1.1.20)), it can easily be seen that the $\boldsymbol{H}_0$ of Eq. (1.1.29) agrees with that of Eq. (1.1.23). Here, although the $\boldsymbol{H}_0$ in Eq. (1.1.23) agrees with the partial derivative of $\boldsymbol{g}_0$ in Eq. (1.1.22) with respect to $\boldsymbol{a}^\top$, the $\boldsymbol{H}_0$ in Eq. (1.1.29) cannot be obtained from such a relationship. The reason for this is that the state variable is used in obtaining the $\boldsymbol{H}_0$ of Eq. (1.1.29).

1.1.6 The Adjoint Variable Method

Finally, let us find $\boldsymbol{g}_0$ through the adjoint variable method, which utilizes the Lagrange multiplier method. The details of the adjoint variable method are presented later on in Sect. 2.6.5 and so, for the moment, we shall limit ourselves to its formal application.

Let the Lagrange function for the cost function f_0 be

$$\begin{aligned}
\mathscr{L}_0(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0) &= f_0(\boldsymbol{u}) + \mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0) \\
&= \boldsymbol{p} \cdot \boldsymbol{u} - \boldsymbol{v}_0 \cdot (\boldsymbol{K}(\boldsymbol{a}) \boldsymbol{u} - \boldsymbol{p}),
\end{aligned} \tag{1.1.30}$$

where $\mathscr{L}_{\mathrm{S}}$ denotes a Lagrange function with respect to the state determination problem (Problem 1.1.3) defined by Eq. (1.1.12). Here, $\boldsymbol{v}_0 = (v_{01}, v_{02})^\top \in U = \mathbb{R}^2$ includes the subscript 0 in order to indicate that it is an adjoint variable (Lagrange multiplier) prepared for f_0. Going forward, whenever f_i is a function of the state variable $\boldsymbol{u}$, the adjoint variable will be written as $\boldsymbol{v}_i$.

The adjoint variable method is a technique for finding $\boldsymbol{g}_0$ using the stationary conditions of $\mathscr{L}_0(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)$ with respect to arbitrary variations of $\boldsymbol{u}$ and $\boldsymbol{v}_0$. The (total) derivative of $\mathscr{L}_0$ with respect to an arbitrary variation $(\boldsymbol{b}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}_0) \in X \times U \times U$ of $(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)$ is

$$\begin{aligned}
&\mathscr{L}_0'(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}_0] \\
&= \mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}] + \mathscr{L}_{0\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{u}}] + \mathscr{L}_{0\boldsymbol{v}_0}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{v}}_0].
\end{aligned} \tag{1.1.31}$$

In this book, we use the notation $(\,\cdot\,)_a$ for $\partial(\,\cdot\,)/\partial a$. The third term on the right-hand side of Eq. (1.1.31) is

$$\mathscr{L}_{0\boldsymbol{v}_0}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\hat{\boldsymbol{v}}_0]=\mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{u},\hat{\boldsymbol{v}}_0)\,. \tag{1.1.32}$$

Equation (1.1.32) is the Lagrange function with respect to the state determination problem (Problem 1.1.3) defined in Eq. (1.1.12) and, if $\boldsymbol{u}$ is a solution of the state determination problem, the third term on the right-hand side of Eq. (1.1.31) is zero.

Moreover, the second term on the right-hand side of Eq. (1.1.31) is

$$\begin{aligned}\mathscr{L}_{0\boldsymbol{u}}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\hat{\boldsymbol{u}}]&=f_{0\boldsymbol{u}}(\boldsymbol{u})[\hat{\boldsymbol{u}}]+\mathscr{L}_{\mathrm{S}\boldsymbol{u}}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\hat{\boldsymbol{u}}]\\&=\boldsymbol{p}\cdot\hat{\boldsymbol{u}}-\boldsymbol{v}_0\cdot(\boldsymbol{K}(\boldsymbol{a})\hat{\boldsymbol{u}})\\&=\hat{\boldsymbol{u}}\cdot\left(\boldsymbol{p}-\boldsymbol{K}^{\top}(\boldsymbol{a})\boldsymbol{v}_0\right).\end{aligned} \tag{1.1.33}$$

Here, if $\boldsymbol{v}_0$ can be determined so that Eq. (1.1.33) is zero for arbitrary $\hat{\boldsymbol{u}}\in U$, then the second term on the right-hand side of Eq. (1.1.31) also vanishes. The condition here is equivalent to setting $\boldsymbol{v}_0$ to be the solution of the following adjoint problem.

Problem 1.1.5 (Adjoint problem with respect to f_0) Let $\boldsymbol{K}(\boldsymbol{a})$ and $\boldsymbol{p}$ be as in Problem 1.1.3 and find $\boldsymbol{v}_0\in U$ satisfying

$$\boldsymbol{K}^{\top}(\boldsymbol{a})\boldsymbol{v}_0=\boldsymbol{p}. \tag{1.1.34}$$

□

Upon comparison of Problem 1.1.3 and Problem 1.1.5, using the fact that $\boldsymbol{K}^{\top}=\boldsymbol{K}$, we obtain

$$\boldsymbol{v}_0=\boldsymbol{u}. \tag{1.1.35}$$

As in the above equation, the relationship where the state variable is equal to the adjoint variable is called a self-adjoint relationship. In fact, the right-hand side of Eq. (1.1.34) is $\partial f_0(\boldsymbol{u})/\partial\boldsymbol{u}$. In Problem 1.1.4, $f_0(\boldsymbol{u})=\boldsymbol{p}\cdot\boldsymbol{u}$, and the self-adjoint relationship holds. That is, the self-adjoint property holds when f_0 is selected so that $\partial f_0(\boldsymbol{u})/\partial\boldsymbol{u}$ is equal to the right-hand side of the state equation (Eq. (1.1.11)). In generality, state and adjoint equations are different and, in such a case, their relationship is said to be non-self adjoint. An example of this is presented in Practice **1.1**. We would now like to consider the meaning of the adjoint equation, and we remark that the following arguments apply to other adjoint equations.

The first term on the right-hand side of Eq. (1.1.31) can be obtained:

$$\begin{aligned}&\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\boldsymbol{b}]\\&\quad=-\left\{\boldsymbol{v}_0\cdot\left(\frac{\partial\boldsymbol{K}(\boldsymbol{a})}{\partial a_1}\boldsymbol{u}\quad\frac{\partial\boldsymbol{K}(\boldsymbol{a})}{\partial a_2}\boldsymbol{u}\right)\right\}\boldsymbol{b}\end{aligned}$$

$$
\begin{aligned}
&= -\frac{e_{\mathrm{Y}}}{l}\left\{\begin{pmatrix} v_{01} & v_{02} \end{pmatrix}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right)\right\}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
&= -\frac{e_{\mathrm{Y}}}{l}\begin{pmatrix} v_{01} & v_{02} \end{pmatrix}\begin{pmatrix} u_1 & u_1 - u_2 \\ 0 & u_2 - u_1 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
&= -\frac{e_{\mathrm{Y}}}{l}\begin{pmatrix} u_1 v_{01} & (u_2 - u_1)(v_{02} - v_{01}) \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
&= l\begin{pmatrix} -\sigma(u_1)\,\varepsilon(v_{01}) & -\sigma(u_2 - u_1)\,\varepsilon(v_{02} - v_{01}) \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
&= \boldsymbol{g}_0 \cdot \boldsymbol{b}. \qquad (1.1.36)
\end{aligned}
$$

Here we remark that $\boldsymbol{g}_0$ matches the results from the direct differentiation method (Eq. (1.1.28)).

Based on the above results, if $\boldsymbol{u}$ and $\boldsymbol{v}_0$ are solutions of Problem 1.1.3 and Problem 1.1.5, respectively, the second and third terms on the right-hand side of Eq. (1.1.31) are zero, and the following equation holds:

$$
\mathscr{L}_0'(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}_0] = \mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}] = \tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}] = \boldsymbol{g}_0 \cdot \boldsymbol{b}. \qquad (1.1.37)
$$

The method of obtaining the derivative of the cost function when equality constraints (state equations) are satisfied in this way is shown in Sect. 2.6, where Eq. (1.1.36) corresponds to Eq. (2.6.25). It should, however, be noted that in Sect. 2.6, the gradient of $\tilde{f}$ is written as $\tilde{\boldsymbol{g}}$.

Let us also examine the relationship between the Lagrange function $\mathscr{L}_0$ and the Hesse matrix $\boldsymbol{H}_0$.

As explained in Sect. 2.1, if an optimal design problem is to be replaced by an optimization problem in which design variables and state variables are not distinguished, and are treated as variables, then the variable of the optimization problem becomes a combination of the state and design variables of the optimal design problem. Following this approach here, the design variable of the optimization problem is set to be $\boldsymbol{x} = \left(\boldsymbol{a}^\top, \boldsymbol{u}^\top\right)^\top \in \mathbb{R}^4$. In order to simplify the notation, $\left(\boldsymbol{a}^\top, \boldsymbol{u}^\top\right)^\top$ will be written as $(\boldsymbol{a}, \boldsymbol{u})$. The Lagrange multiplier with respect to the equality constraint will be written as $\boldsymbol{v}_0$, and the second-order derivative of the Lagrange function $\mathscr{L}_0$ with respect to arbitrary variations $(\boldsymbol{b}_2, \hat{\boldsymbol{u}}_2) \in X \times U$ and $(\boldsymbol{b}_1, \hat{\boldsymbol{u}}_1) \in X \times U$ of the design variables $(\boldsymbol{a}, \boldsymbol{u})$ will be written $\mathscr{L}_{0(\boldsymbol{a},\boldsymbol{u}),(\boldsymbol{a},\boldsymbol{u})}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[(\boldsymbol{b}_1, \hat{\boldsymbol{u}}_1), (\boldsymbol{b}_2, \hat{\boldsymbol{u}}_2)]$. In this case, we have

$$
\begin{aligned}
&\mathscr{L}_{0(\boldsymbol{a},\boldsymbol{u}),(\boldsymbol{a},\boldsymbol{u})}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[(\boldsymbol{b}_1, \hat{\boldsymbol{u}}_1), (\boldsymbol{b}_2, \hat{\boldsymbol{u}}_2)] \\
&\quad = \left(\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}_1] + \mathscr{L}_{0\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{u}}_1]\right)_{\boldsymbol{a}}[\boldsymbol{b}_2] \\
&\qquad + \left(\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}_1] + \mathscr{L}_{0\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{u}}_1]\right)_{\boldsymbol{u}}[\hat{\boldsymbol{u}}_2] \\
&\quad = \left(f_{0\boldsymbol{u}} \cdot \hat{\boldsymbol{u}}_1 + \mathscr{L}_{\mathrm{S}\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}_1] + \mathscr{L}_{\mathrm{S}\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{u}}_1]\right)_{\boldsymbol{a}}[\boldsymbol{b}_2] \\
&\qquad + \left(f_{0\boldsymbol{u}} \cdot \hat{\boldsymbol{u}}_1 + \mathscr{L}_{\mathrm{S}\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\boldsymbol{b}_1] + \mathscr{L}_{\mathrm{S}\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_0)[\hat{\boldsymbol{u}}_1]\right)_{\boldsymbol{u}}[\hat{\boldsymbol{u}}_2] \\
&\quad = \begin{pmatrix} \boldsymbol{b}_2 \\ \hat{\boldsymbol{u}}_2 \end{pmatrix} \cdot \left(\boldsymbol{H}_{\mathscr{L}_{\mathrm{S}}}\begin{pmatrix} \boldsymbol{b}_1 \\ \hat{\boldsymbol{u}}_1 \end{pmatrix}\right), \qquad (1.1.38)
\end{aligned}
$$

where

$$\boldsymbol{H}_{\mathscr{L}_{\mathrm{S}}} = \begin{pmatrix} \mathscr{L}_{\mathrm{S}\boldsymbol{a}\boldsymbol{a}} & \mathscr{L}_{\mathrm{S}\boldsymbol{a}\boldsymbol{u}} \\ \mathscr{L}_{\mathrm{S}\boldsymbol{u}\boldsymbol{a}} & \mathscr{L}_{\mathrm{S}\boldsymbol{u}\boldsymbol{u}} \end{pmatrix} = -\begin{pmatrix} \mathbf{0}_{\mathbb{R}^{2\times 2}} & \begin{pmatrix} \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_1} \\ \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_2} \end{pmatrix} \\ \begin{pmatrix} \boldsymbol{K}_{a_1}^{\top}\boldsymbol{v}_0 & \boldsymbol{K}_{a_2}^{\top}\boldsymbol{v}_0 \end{pmatrix} & \mathbf{0}_{\mathbb{R}^{2\times 2}} \end{pmatrix}. \tag{1.1.39}$$

From Eq. (1.1.39), it is apparent that the matrix $\boldsymbol{H}_{\mathscr{L}_{\mathrm{S}}}$ need not be positive definite.

Here $\boldsymbol{u}$ and $\boldsymbol{v}_0$ denote, respectively, the solutions to the state determination problem (Problem 1.1.3) and the adjoint problem (Problem 1.1.5), subject to a design variable $\boldsymbol{a}$. Furthermore, we assume that $\hat{\boldsymbol{\upsilon}}$ (the letter $\boldsymbol{\upsilon}$ is a bold Greek upsilon) denotes a variation of $\boldsymbol{u}$ under the equality constraint of the state determination problem corresponding to an arbitrary variation $\boldsymbol{b} \in X$ of $\boldsymbol{a}$. In Chap. 2, we call the set of $(\boldsymbol{b}, \hat{\boldsymbol{\upsilon}})$ the feasible direction set or the tangent plane on $X \times U$ satisfying the equality constraint of the state determination problem (see $T_V(\boldsymbol{x})$ in Eq. (2.6.2), Theorems 2.6.6 and 2.6.7). Here, the cross-sectional derivative of the Lagrange function with respect to the state determination problem is

$$\mathscr{L}_{\mathrm{S}(\boldsymbol{a},\boldsymbol{u})}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v})[\boldsymbol{b},\hat{\boldsymbol{\upsilon}}] = \boldsymbol{v}\cdot\left\{-\left(\boldsymbol{K}'(\boldsymbol{a})[\boldsymbol{b}]\right)\boldsymbol{u} - \boldsymbol{K}(\boldsymbol{a})\hat{\boldsymbol{\upsilon}}\right\} = 0. \tag{1.1.40}$$

This yields the identity

$$\hat{\boldsymbol{\upsilon}} = -\boldsymbol{K}^{-1}(\boldsymbol{a})\left(\boldsymbol{K}'(\boldsymbol{a})[\boldsymbol{b}]\right) = \begin{pmatrix} -\dfrac{u_1}{a_1} & 0 \\ -\dfrac{u_1}{a_1} & -\dfrac{u_2-u_1}{a_2} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{1.1.41}$$

Equation (1.1.41) is equivalent to the conditions expressed by Eq. (1.1.25) and Eq. (1.1.26). Moreover, using the self-adjoint relationship, Eq. (1.1.38) becomes

$$\begin{aligned} &\mathscr{L}_{0(\boldsymbol{a},\boldsymbol{u}),(\boldsymbol{a},\boldsymbol{u})}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[(\boldsymbol{b}_1,\hat{\boldsymbol{\upsilon}}_1),(\boldsymbol{b}_2,\hat{\boldsymbol{\upsilon}}_2)] \\ &\quad = \mathscr{L}_{\mathrm{S}\boldsymbol{a}\boldsymbol{u}}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\boldsymbol{b}_1,\hat{\boldsymbol{\upsilon}}_2] + \mathscr{L}_{\mathrm{S}\boldsymbol{u}\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{u},\boldsymbol{v}_0)[\boldsymbol{b}_2,\hat{\boldsymbol{\upsilon}}_1] \\ &\quad = -\begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}\cdot\left\{\begin{pmatrix} \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_1} \\ \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_2} \end{pmatrix}\begin{pmatrix} -u_1/a_1 & 0 \\ -u_1/a_1 & -(u_2-u_1)/a_2 \end{pmatrix}\right. \\ &\qquad \left. + \left(\begin{pmatrix} \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_1} \\ \boldsymbol{v}_0^{\top}\boldsymbol{K}_{a_2} \end{pmatrix}\begin{pmatrix} -u_1/a_1 & 0 \\ -u_1/a_1 & -(u_2-u_1)/a_2 \end{pmatrix}\right)^{\top}\right\}\begin{pmatrix} b_{21} \\ b_{22} \end{pmatrix} \\ &\quad = \boldsymbol{b}_1\cdot(\boldsymbol{H}_0\boldsymbol{b}_2). \end{aligned}$$

This shows that the second-order cross-sectional derivative of $\tilde{f}_0$ agrees with Eq. (1.1.29) and is expressed as

$$h_0(\boldsymbol{a})[\boldsymbol{b}_1,\boldsymbol{b}_2] = \boldsymbol{b}_1\cdot(\boldsymbol{H}_0\boldsymbol{b}_2). \tag{1.1.42}$$

The above results clarify that the Hesse matrix of the Lagrange function $\mathscr{L}_0$ with respect to an arbitrary variation of $(\boldsymbol{a},\boldsymbol{u})$ agrees with the Hesse matrix

$\boldsymbol{H}_{\mathscr{L}_\mathrm{S}}$ of $\mathscr{L}_\mathrm{S}$, and that it is not necessarily positive definite. However, when we assume that $\boldsymbol{u}$ denotes the solution of the state determination problem as the design variable is varied and that $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$ denote variations of $\boldsymbol{u}$, we showed that the Hesse matrix with respect to arbitrary variations $\boldsymbol{b}_1, \boldsymbol{b}_2 \in X$ of $\boldsymbol{a}$ from $\mathscr{L}_0$ is the same as the Hesse matrix $\boldsymbol{H}_0$ obtained via the other methods.

In the optimal design problem (Problem 1.1.4) considered in this section, we have assumed that the design variable $\boldsymbol{a} \in \mathbb{R}^2$ is a cross-sectional area. Therefore, $\boldsymbol{K}(\boldsymbol{a})$ was a linear form of $\boldsymbol{a}$ in Eq. (1.1.11) and hence, $\mathscr{L}_{\mathrm{S}\boldsymbol{a}\boldsymbol{a}} = \boldsymbol{0}_{\mathbb{R}^{2\times 2}}$ in Eq. (1.1.39) and Eq. (1.1.42). However, if the design variable $\boldsymbol{a} \in \mathbb{R}^2$ is assumed to be the length of one side of a square cross-section, then $\boldsymbol{K}(\boldsymbol{a})$ becomes $\boldsymbol{K}\left(\boldsymbol{a}^2\right)$ and $\mathscr{L}_{\mathrm{S}\boldsymbol{a}\boldsymbol{a}} \neq \boldsymbol{0}_{\mathbb{R}^{2\times 2}}$ (see Practice **1.5**). In this way, we see that $\mathscr{L}_{0\boldsymbol{a}\boldsymbol{a}}$ may not be $\boldsymbol{0}_{\mathbb{R}^{2\times 2}}$, depending on the choice of design variables or cost functions. Nevertheless, when the state determination problem is linear the condition $\mathscr{L}_{\mathrm{S}\boldsymbol{u}\boldsymbol{u}} = \boldsymbol{0}_{\mathbb{R}^{2\times 2}}$ always holds.

In the method used above, Eq. (1.1.41) was utilized to obtain the second cross-sectional derivative from the first cross-sectional derivative. Equation (1.1.41) accords with Eq. (1.1.25) and Eq. (1.1.26) from which we see that the direct differentiation method was actually applied. These results do not always hold in general, such as in those problems given in Chaps. 8 and 9. In these cases, the Lagrange multiplier method can be used to obtain the second-order derivative of cost functions. Such a method will also be described in the succeeding discussions. However, in cases where the second-order derivatives are used in solving the optimization problems, a key idea, whose detail will be presented in Chap. 3, is required.

In Chap. 4, the Fréchet derivative will be defined as a generalized derivative (Definition 4.5.4). Following the definition of the second-order derivative, here, we fix $\boldsymbol{b}_1$ and consider differentiating the first cross-sectional derivative $\tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}_1] = \boldsymbol{g}_0 \cdot \boldsymbol{b}_1$. To do this, we define the Lagrange function for $\boldsymbol{g}_0 \cdot \boldsymbol{b}_1$ by

$$\mathscr{L}_{\mathrm{I}0}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0) = \boldsymbol{g}_0(\boldsymbol{u}) \cdot \boldsymbol{b}_1 + \mathscr{L}_\mathrm{S}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0), \tag{1.1.43}$$

where $\boldsymbol{g}_0(\boldsymbol{u})$ and $\mathscr{L}_\mathrm{S}$ are given by Eq. (1.1.12) and Eq. (1.1.36), respectively. $\boldsymbol{w}_0 = (w_{01}, w_{02})^\top$ is the adjoint variable provided for $\boldsymbol{u}$ in $\boldsymbol{g}_0(\boldsymbol{u})$ satisfying the state determination problem.

With respect to arbitrary variations $(\boldsymbol{b}_2, \hat{\boldsymbol{u}}, \hat{\boldsymbol{w}}_0) \in X \times U^2$ of $(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)$, the derivative of $\mathscr{L}_{\mathrm{I}0}$ is written as

$$\begin{aligned}
&\mathscr{L}_{\mathrm{I}0}'(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\boldsymbol{b}_2, \hat{\boldsymbol{u}}, \hat{\boldsymbol{w}}_0] \\
&\quad = \mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\boldsymbol{b}_2] + \mathscr{L}_{\mathrm{I}0\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\hat{\boldsymbol{u}}] \\
&\qquad + \mathscr{L}_{\mathrm{I}0\boldsymbol{w}_0}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\hat{\boldsymbol{w}}_0].
\end{aligned} \tag{1.1.44}$$

The third term on the right-hand side of Eq. (1.1.44) vanishes if $\boldsymbol{u}$ is the solution of the state determination problem. Moreover, the second term on the right-hand side of Eq. (1.1.44) is

$$\mathscr{L}_{\mathrm{I}0\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\hat{\boldsymbol{u}}] = \boldsymbol{g}_{0\boldsymbol{u}^\top}(\boldsymbol{u})[\hat{\boldsymbol{u}}] \cdot \boldsymbol{b}_1 + \mathscr{L}_{\mathrm{S}\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0)[\hat{\boldsymbol{u}}]$$

$$= \hat{\boldsymbol{u}} \cdot \boldsymbol{q} - \boldsymbol{w}_0 \cdot (\boldsymbol{K}(\boldsymbol{a}) \hat{\boldsymbol{u}})$$
$$= \hat{\boldsymbol{u}} \cdot \left(\boldsymbol{q} - \boldsymbol{K}^\top(\boldsymbol{a}) \boldsymbol{w}_0\right), \tag{1.1.45}$$

where

$$\boldsymbol{q} = \boldsymbol{g}_{0\boldsymbol{u}}^\top(\boldsymbol{u}) \boldsymbol{b}_1 = -\frac{2e_{\mathrm{Y}}}{l} \begin{pmatrix} u_1 & u_1 - u_2 \\ 0 & u_2 - u_1 \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}. \tag{1.1.46}$$

Here, the condition that Eq. (1.1.45) is zero for arbitrary $\hat{\boldsymbol{u}} \in U$ is equivalent to setting $\boldsymbol{w}_0$ to be the solution of the following adjoint problem.

Problem 1.1.6 (Adjoint problem with respect to $\boldsymbol{g}_0(\boldsymbol{u}) \cdot \boldsymbol{b}_1$) Let $\boldsymbol{K}(\boldsymbol{a})$ be as in Problem 1.1.3, and $\boldsymbol{q}$ be given by Eq. (1.1.46). Find $\boldsymbol{w}_0 \in U$ satisfying

$$\boldsymbol{K}^\top(\boldsymbol{a}) \boldsymbol{w}_0 = \boldsymbol{q}.$$

□

The solution of Problem 1.1.6 is

$$\boldsymbol{w}_0 = \left(\boldsymbol{K}^\top(\boldsymbol{a})\right)^{-1} \boldsymbol{g}_{0\boldsymbol{u}}^\top(\boldsymbol{u}) \boldsymbol{b}_1 = -2 \begin{pmatrix} \dfrac{u_1}{a_1} & 0 \\ \dfrac{u_1}{a_1} & \dfrac{u_2 - u_1}{a_2} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}. \tag{1.1.47}$$

Here, $\boldsymbol{w}_0$ is a function of $\boldsymbol{b}_1$, and so is written as $\boldsymbol{w}_0(\boldsymbol{b}_1)$.

Finally, the first term on the right-hand side of Eq. (1.1.44) becomes

$$\mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0(\boldsymbol{b}_1))[\boldsymbol{b}_2]$$
$$= -\left\{\boldsymbol{w}_0(\boldsymbol{b}_1) \cdot \left(\frac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_1}\boldsymbol{u} \quad \frac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_2}\boldsymbol{u}\right)\right\} \boldsymbol{b}_2. \tag{1.1.48}$$

Substituting Eq. (1.1.47) into Eq. (1.1.48), we obtain

$$\mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}0, \boldsymbol{w}_0(\boldsymbol{b}_1))[\boldsymbol{b}_2]$$
$$= h_0(\boldsymbol{a})[\boldsymbol{b}_1, \boldsymbol{b}_2] = \boldsymbol{b}_1 \cdot (\boldsymbol{H}_0 \boldsymbol{b}_2) = \boldsymbol{g}_{\mathrm{H}0}(\boldsymbol{a}, \boldsymbol{b}_1) \cdot \boldsymbol{b}_2, \tag{1.1.49}$$

where

$$\boldsymbol{g}_{\mathrm{H}0}(\boldsymbol{a}, \boldsymbol{b}_1) = \mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{w}_0(\boldsymbol{b}_1)). \tag{1.1.50}$$

In this book, $\boldsymbol{g}_{\mathrm{H}0}$ is called the Hesse gradient.

1.1.7 Optimality Conditions

The previous section explored methods for calculating $\boldsymbol{g}_0$ (the gradient of $\tilde{f}_0$ with respect to the variation of the cross-section $\boldsymbol{a} \in \mathcal{D}^\circ$), the Hesse matrix $\boldsymbol{H}_0$, and the gradient $\boldsymbol{g}_1$ of f_1. We now return to Problem 1.1.4 and consider optimality conditions that the optimal cross-section satisfies.

As described in Sect. 1.1.4 the convexity of $\tilde{f}_0$ was obtained from the fact that the Hesse matrix $\boldsymbol{H}_0$ is positive definite on X (Theorem 2.4.6). We also showed that f_1 is a linear function of $\boldsymbol{a}$ and that it is therefore convex on X (Theorem 2.4.4). Problem 1.1.4 is then a convex optimization problem and, as will be shown later, a design variable $\boldsymbol{a} \in \mathcal{D}^\circ$ which satisfies the Karush–Kuhn–Tucker conditions (Theorem 2.7.5) is the minimizer of Problem 1.1.4 (Theorem 2.7.9). Let us now find the KKT conditions for Problem 1.1.4.

Let the Lagrange function for Problem 1.1.4 be

$$\mathscr{L}(\boldsymbol{a}, \lambda_1) = \tilde{f}_0(\boldsymbol{a}) + \lambda_1 f_1(\boldsymbol{a}),$$

where $\lambda_1 \in \mathbb{R}$ is a Lagrange multiplier with respect to $f_1(\boldsymbol{a}) \le 0$. Then the KKT conditions for Problem 1.1.4 are given by

$$\mathscr{L}_{\boldsymbol{a}}(\boldsymbol{a}, \lambda_1) = \boldsymbol{g}_0 + \lambda_1 \boldsymbol{g}_1 = \boldsymbol{0}_{\mathbb{R}^2}, \tag{1.1.51}$$

$$\mathscr{L}_{\lambda_1}(\boldsymbol{a}, \lambda_1) = f_1(\boldsymbol{a}) = l(a_1 + a_2) - c_1 \le 0, \tag{1.1.52}$$

$$\lambda_1 f_1(\boldsymbol{a}) = 0, \tag{1.1.53}$$

$$\lambda_1 \ge 0. \tag{1.1.54}$$

A detailed explanation regarding the meaning of the KKT conditions will be deferred until Sect. 2.7.3, but let us now take a look into their general meaning.

First of all, when the cross-section is optimal, Eq. (1.1.51) and Eq. (1.1.54) describe a trade-off relationship between the objective and the constraint functions. In fact, upon taking the inner product of $\boldsymbol{b}$ with both sides of Eq. (1.1.51), we can obtain

$$\lambda_1 = -\frac{\boldsymbol{g}_0 \cdot \boldsymbol{b}}{\boldsymbol{g}_1 \cdot \boldsymbol{b}}. \tag{1.1.55}$$

The numerator and the denominator on the right-hand side of Eq. (1.1.55) represent the amount of variation in f_1 and f_0 when the design variable is varied by $\boldsymbol{b}$. Here, $\lambda_1 > 0$ indicates the fact that the signs of the variations differ. In other words, there is a trade-off relationship between f_1 and f_0.

Finally, we remark that Eq. (1.1.52) is the original constraint condition. Also, we say that Eq. (1.1.53) is a complementarity condition. If an inequality constraint can be satisfied by an equality (referred to as active), then Eq. (1.1.53) allows $\lambda_1 > 0$. Similarly, if it can be satisfied as an inequality (referred to as inactive) then $\lambda_1 = 0$ and this acts to inactivate the constraint.

Next let us consider the physical interpretation of Eq. (1.1.51). If $\boldsymbol{g}_0$ from Eq. (1.1.36) and $\boldsymbol{g}_1$ from Eq. (1.1.17) are substituted into Eq. (1.1.51), then we obtain

$$l\begin{pmatrix} -\sigma(u_1)\,\varepsilon(u_1) \\ -\sigma(u_2 - u_1)\,\varepsilon(u_2 - u_1) \end{pmatrix} + \lambda_1 l \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This equation implies that

$$\sigma(u_1)\,\varepsilon(u_1) = \sigma(u_2 - u_1)\,\varepsilon(u_2 - u_1) = \lambda_1. \tag{1.1.56}$$

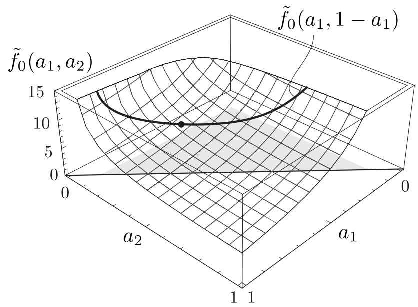

Fig. 1.4: Numerical example of the mean compliance minimization problem.

In other words, when Problem 1.1.4 is minimized, the strain energy densities (w of Eq. (1.1.7)) of the two elastic bodies agree and λ_1 is twice the strain energy density. Therefore, λ_1 is greater than zero when $\boldsymbol{p}$ generates a non-zero stress on the two one-dimensional elastic bodies, and the volume constraint is active at the minimizer.

1.1.8 Numerical Example

Let us consider a concrete example and try to obtain the minimizer.

Exercise 1.1.7 (Mean compliance minimization) Find the minimizer $\boldsymbol{a}$ in Problem 1.1.4, subject to $l = 1$, $e_Y = 1$, $c_1 = 1$, $\boldsymbol{p} = (1, 1)^\top$ and $\boldsymbol{a}_0 = (0.1, 0.1)^\top$. □

Answer Substituting $l = 1$, $e_Y = 1$ and $\boldsymbol{p} = (1, 1)^\top$ into Eq. (1.1.21) gives

$$\tilde{f}_0(\boldsymbol{a}) = \frac{4}{a_1} + \frac{1}{a_2}. \tag{1.1.57}$$

Figure 1.4 shows $\tilde{f}_0$. From Eq. (1.1.22) and Eq. (1.1.17), the cross-sectional-area derivative of $\tilde{f}_0$ and f_1 are given by

$$\boldsymbol{g}_0 = -\begin{pmatrix} 4/a_1^2 \\ 1/a_2^2 \end{pmatrix}, \quad \boldsymbol{g}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

If $\boldsymbol{a}$ is a minimizer, then from Eq. (1.1.56) we have

$$\lambda_1 = \frac{4}{a_1^2} = \frac{1}{a_2^2}.$$

If $\boldsymbol{a}$ is an element of $\mathcal{D}^\circ$ (see Eq. (1.1.16)), then λ_1 is positive and the complementarity condition allows for the inequality constraint with respect to f_1 to be satisfied with an equality. Here, if $a_2 = 1 - a_1$ is substituted into Eq. (1.1.57), then a_1 can be obtained from the stationary condition of $\tilde{f}_0$ with respect to an arbitrary variation of a_1:

$$\frac{\mathrm{d}}{\mathrm{d}a_1}\tilde{f}_0(a_1, 1 - a_1) = \frac{1}{(1 - a_1)^2} - \frac{4}{a_1^2} = 0.$$

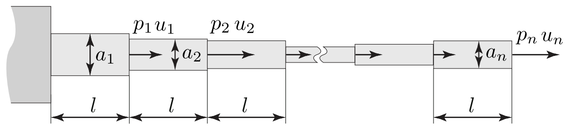

Fig. 1.5: A one-dimensional linear elastic body with n cross-sections.

We then obtain a_2 from $1 - a_1$:

$$\boldsymbol{a} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Among these values, $\boldsymbol{a} = (2/3, 1/3)^\top$ satisfies $\boldsymbol{a} \geq \boldsymbol{a}_0$. Due to the fact that $\tilde{f}_0$ and f_1 are convex functions, Problem 1.1.4 becomes a convex optimization problem, and from Theorem 2.7.9, the value of $\boldsymbol{a}$ which satisfies the KKT condition is the minimizer of the problem. □

1.2 Comparison of the Direct Differentiation Method and the Adjoint Variable Method

In Sect. 1.1, we considered how to find optimality conditions for Problem 1.1.4. The substitution method, the direct differentiation method, and the adjoint variable method were used to find the cross-sectional derivative of the cost function. Since we will later deal with optimization problems where the design variable is a function, and because the substitution method becomes quite complex in such a setting, we will exclude this method and compare the characteristics and applicable range of the direct differentiation and adjoint variable methods.

Let us consider a one-dimensional linear elastic body such as the one shown in Fig. 1.5. Here, the number of cross-sections in Problem 1.1.4 has been extended to $n \in \mathbb{N}$ (the set of all natural numbers). The linear elasticity problem in this case is as follows.

Problem 1.2.1 (Multi-stepped 1D linear elastic body) When $l \in \mathbb{R}$, $e_{\mathrm{Y}} \in \mathbb{R}$, $\boldsymbol{a} \in \mathbb{R}^n$ ($\boldsymbol{a} \geq \boldsymbol{a}_0 > \boldsymbol{0}_{\mathbb{R}^n}$) and $\boldsymbol{p} \in \mathbb{R}^n$ are given with respect to the one-dimensional linear elastic body in Fig. 1.5, find $\boldsymbol{u} \in \mathbb{R}^n$ satisfying

$$\boldsymbol{K}(\boldsymbol{a})\,\boldsymbol{u} = \boldsymbol{p}. \tag{1.2.1}$$

Here, $\boldsymbol{K}(\boldsymbol{a})$ is an extension matrix of $\boldsymbol{K}(\boldsymbol{a})$ from Problem 1.1.3 (see Practice **1.5**). □

With respect to $\boldsymbol{u} \in \mathbb{R}^n$ and the Lagrange multiplier $\boldsymbol{v} \in \mathbb{R}^n$, the Lagrange function with respect to the state determination problem (Problem 1.2.1) is

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{v} \cdot (-\boldsymbol{K}(\boldsymbol{a})\,\boldsymbol{u} + \boldsymbol{p}). \tag{1.2.2}$$

Here, Problem 1.2.1 is equivalent to finding $\boldsymbol{u} \in \mathbb{R}^n$ satisfying

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}) = 0,$$

for all $\boldsymbol{v} \in \mathbb{R}^n$.

The number of constraint functions is set to be $m \in \mathbb{N}$, and the optimal design problem is as follows. In the following problem, we let $X = \mathbb{R}^n$, $U = \mathbb{R}^n$, and for a constant vector $\boldsymbol{a}_0 > \mathbf{0}_{\mathbb{R}^n}$ we set

$$\mathcal{D} = \{\boldsymbol{a} \in X \mid \boldsymbol{a} \geq \boldsymbol{a}_0\}. \tag{1.2.3}$$

Problem 1.2.2 (Multi-design variable multi-constraint) Let $X = U = \mathbb{R}^n$, and $\mathcal{D}$ be given by Eq. (1.2.3). Also assume that a function $f_i : X \times U \to \mathbb{R}$ is given for each $i \in \{0, 1, \ldots, m\}$. Given these conditions, find $\boldsymbol{a}$ satisfying

$$\min_{(\boldsymbol{a},\boldsymbol{u}) \in \mathcal{D} \times U} \{ f_0(\boldsymbol{a}, \boldsymbol{u}) \mid f_1(\boldsymbol{a}, \boldsymbol{u}) \leq 0, \ldots, f_m(\boldsymbol{a}, \boldsymbol{u}) \leq 0, \text{ Problem 1.2.1} \}.$$

□

Let us use this problem to compare the method of direct differentiation to the adjoint variable approach, while formalizing techniques for calculating the cross-sectional-area gradients $\boldsymbol{g}_0, \ldots, \boldsymbol{g}_m$ of the cost functions $f_0, \ldots, f_m$. Hereinafter, $i \in \{0, 1, \ldots, m\}$ is the subscript of the cost function f_i, and $j \in \{1, \ldots, n\}$ is the subscript of the design variable a_j.

1.2.1 The Direct Differentiation Method

Let us first look at the method for calculating $\boldsymbol{g}_i$ using the direct differentiation method. The solution of Problem 1.2.1 corresponding to $\boldsymbol{a} + \boldsymbol{b}$ with respect to an arbitrary $\boldsymbol{b} \in \mathbb{R}^n$ will be written as $\boldsymbol{u}(\boldsymbol{a} + \boldsymbol{b})$.

From the Taylor expansion of $\tilde{f}_i(\boldsymbol{a})$ about $\boldsymbol{a}$ and the chain rule of differentiation, one can write

$$\begin{aligned}
&\tilde{f}_i(\boldsymbol{a} + \boldsymbol{b}) \\
&\quad = f_i(\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{u}(\boldsymbol{a} + \boldsymbol{b})) \\
&\quad = f_i(\boldsymbol{a}, \boldsymbol{u}(\boldsymbol{a})) + \begin{pmatrix} \dfrac{\partial f_i}{\partial a_1} & \dfrac{\partial f_i}{\partial a_2} & \cdots & \dfrac{\partial f_i}{\partial a_n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\
&\qquad + \begin{pmatrix} \dfrac{\partial f_i}{\partial u_1} & \dfrac{\partial f_i}{\partial u_2} & \cdots & \dfrac{\partial f_i}{\partial u_n} \end{pmatrix} \begin{pmatrix} \dfrac{\partial u_1}{\partial a_1} & \dfrac{\partial u_1}{\partial a_2} & \cdots & \dfrac{\partial u_1}{\partial a_n} \\ \dfrac{\partial u_2}{\partial a_1} & \dfrac{\partial u_2}{\partial a_2} & \cdots & \dfrac{\partial u_2}{\partial a_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial u_n}{\partial a_1} & \dfrac{\partial u_n}{\partial a_2} & \cdots & \dfrac{\partial u_n}{\partial a_n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
&+ o\left(\|\boldsymbol{b}\|_{\mathbb{R}^n}\right) \\
&= f_i\left(\boldsymbol{a}, \boldsymbol{u}\left(\boldsymbol{a}\right)\right) + f_{i\boldsymbol{a}} \cdot \boldsymbol{b} + f_{i\boldsymbol{u}} \cdot \left(\boldsymbol{u}_{\boldsymbol{a}^\top} \boldsymbol{b}\right) + o\left(\|\boldsymbol{b}\|_{\mathbb{R}^n}\right) \\
&= f_i\left(\boldsymbol{a}, \boldsymbol{u}\left(\boldsymbol{a}\right)\right) + \left\{ f_{i\boldsymbol{a}} + \left(\boldsymbol{u}_{\boldsymbol{a}^\top}\right)^\top f_{i\boldsymbol{u}} \right\} \cdot \boldsymbol{b} + o\left(\|\boldsymbol{b}\|_{\mathbb{R}^n}\right) \\
&= f_i\left(\boldsymbol{a}, \boldsymbol{u}\left(\boldsymbol{a}\right)\right) + \boldsymbol{g}_i \cdot \boldsymbol{b} + o\left(\|\boldsymbol{b}\|_{\mathbb{R}^n}\right).
\end{aligned} \tag{1.2.4}
$$

In this book, the matrix $\left(\partial u_i / \partial a_j\right)_{ij}$, consisting of the partial derivative of the column vector $\boldsymbol{u}$ with respect to a row vector $\boldsymbol{a}^\top$ will be written as $\boldsymbol{u}_{\boldsymbol{a}^\top}$. If $f_i\left(\boldsymbol{a}, \boldsymbol{u}\right)$ is given in Eq. (1.2.4), then $f_{i\boldsymbol{a}}$ and $f_{i\boldsymbol{u}}$ are known, so let us now consider a method for calculating $\boldsymbol{u}_{\boldsymbol{a}^\top}$ in order to find $\boldsymbol{g}_i$.

Taking the partial derivative of the state equation with respect to a_j yields

$$
\frac{\partial \boldsymbol{K}}{\partial a_j} \boldsymbol{u} + \boldsymbol{K}\left(\boldsymbol{a}\right) \frac{\partial \boldsymbol{u}}{\partial a_j} = \boldsymbol{0}_{\mathbb{R}^n},
$$

where $j \in \{1, \ldots, n\}$, or equivalently

$$
\frac{\partial \boldsymbol{u}}{\partial a_j} = -\boldsymbol{K}^{-1}\left(\boldsymbol{a}\right) \frac{\partial \boldsymbol{K}}{\partial a_j} \boldsymbol{u}. \tag{1.2.5}
$$

Arranging Eq. (1.2.5) in rows with respect to $j \in \{1, \ldots, n\}$ establishes $\boldsymbol{u}_{\boldsymbol{a}^\top}$.

The following statements can be made from the above observations.

Remark 1.2.3 (Characteristics of the direct differentiation method) Compared with the adjoint variable method, the direct differentiation method has the following properties:

(1) The direct differentiation method is effective when the number of cost functions is large, i.e., $m \gg 1$. This is because, once $\boldsymbol{u}_{\boldsymbol{a}^\top}$ has been computed from Eq. (1.2.5), $\boldsymbol{u}_{\boldsymbol{a}^\top}$ can be used for each of the cost functions $f_0, \ldots, f_m$.

(2) When the number of design variables is large ($n \gg 1$), the direct differentiation method becomes ineffective. This is because Eq. (1.2.5) must be solved n times. If the inverse matrix $\boldsymbol{K}^{-1}$ is not tracked and is recalculated for each design variable, the amount of calculation required is similar to that when the finite-difference method is used to compute the cross-sectional derivative.

□

In Remark 1.2.3 (2), the finite-difference method was used as a comparison. The method for calculating the cross-sectional derivative by the finite-difference method is described as follows. Let $\boldsymbol{g}_i = \left(g_{ij}\right)_{j \in \{1, \ldots, n\}}$. Using the solutions to state equations corresponding to $\boldsymbol{a}$ and $\boldsymbol{a} + (0, \ldots, 0, b_j, 0, \ldots, 0)^\top$, we calculate g_{ij} as follows:

$$
g_{ij} = \frac{\tilde{f}_i\left(\boldsymbol{a} + (0, \ldots, 0, b_j, 0, \ldots, 0)^\top\right) - \tilde{f}_i\left(\boldsymbol{a}\right)}{b_j}. \tag{1.2.6}
$$

Thus, finding $\boldsymbol{g}_i$ through this method requires solving the state equation n times.

1.2.2 The Adjoint Variable Method

Next, let us look at the adjoint variable method for calculating $\boldsymbol{g}_i$. Let the Lagrange function with respect to $f_i(\boldsymbol{a}, \boldsymbol{u})$ be

$$\begin{aligned}\mathscr{L}_i(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i) &= f_i(\boldsymbol{a}, \boldsymbol{u}) + \mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)\\ &= f_i(\boldsymbol{a}, \boldsymbol{u}) - \boldsymbol{v}_i \cdot (\boldsymbol{K}(\boldsymbol{a})\boldsymbol{u} - \boldsymbol{p}),\end{aligned} \tag{1.2.7}$$

where we have supposed that $\mathscr{L}_{\mathrm{S}}$ is defined by Eq. (1.2.2). Here, $\boldsymbol{v}_i \in \mathbb{R}^n$ is the adjoint variable (Lagrange multiplier) with respect to the state equation. The derivative of $\mathscr{L}_i$ with respect to an arbitrary variation $(\boldsymbol{b}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}_i) \in X \times U \times U$ of $(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)$ is expressed as

$$\begin{aligned}&\mathscr{L}_i'(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\boldsymbol{b}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}_i]\\ &\quad = \mathscr{L}_{i\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\boldsymbol{b}] + \mathscr{L}_{i\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\hat{\boldsymbol{u}}] + \mathscr{L}_{i\boldsymbol{v}_i}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\hat{\boldsymbol{v}}_i].\end{aligned} \tag{1.2.8}$$

The third term on the right-hand side of Eq. (1.2.8) satisfies

$$\mathscr{L}_{i\boldsymbol{v}_i}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\hat{\boldsymbol{v}}_i] = \mathscr{L}_{\mathrm{S}}(\boldsymbol{a}, \boldsymbol{u}, \hat{\boldsymbol{v}}_i). \tag{1.2.9}$$

Equation (1.2.9) is a Lagrange function with respect to the state determination problem (Problem 1.2.1). It is zero when $\boldsymbol{u}$ solves the state determination problem.

Moreover, computing the second term on the right-hand side of Eq. (1.2.8) yields

$$\begin{aligned}\mathscr{L}_{i\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\hat{\boldsymbol{u}}] &= f_{i\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u})[\hat{\boldsymbol{u}}] + \mathscr{L}_{\mathrm{S}\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\hat{\boldsymbol{u}}]\\ &= f_{i\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}) \cdot \hat{\boldsymbol{u}} - \boldsymbol{v}_i \cdot (\boldsymbol{K}(\boldsymbol{a})\hat{\boldsymbol{u}})\\ &= -\hat{\boldsymbol{u}} \cdot \left(\boldsymbol{K}^{\top}(\boldsymbol{a})\boldsymbol{v}_i - \frac{\partial f_i}{\partial \boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u})\right).\end{aligned} \tag{1.2.10}$$

Here, if $\boldsymbol{v}_i$ can be determined so that Eq. (1.2.10) is zero for arbitrary $\hat{\boldsymbol{u}} \in U$, then the second term on the right-hand side of Eq. (1.2.8) vanishes. This condition is equivalent to setting $\boldsymbol{v}_i$ to solve the following adjoint problem.

Problem 1.2.4 (The adjoint problem with respect to f_i) Let $\boldsymbol{K}(\boldsymbol{a})$ and f_i be as in Problem 1.2.1. Find $\boldsymbol{v}_i \in U$ satisfying

$$\boldsymbol{K}^{\top}(\boldsymbol{a})\boldsymbol{v}_i = f_{i\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{u}).$$

□

If $\boldsymbol{u}$ and $\boldsymbol{v}_i$ are solutions of Problem 1.2.1 and Problem 1.2.4, respectively, then we have

$$\begin{aligned}&\mathscr{L}_{i\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}, \boldsymbol{v}_i)[\boldsymbol{b}]\\ &\quad = f_{i\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{u}) \cdot \boldsymbol{b} - \boldsymbol{v}_i \cdot \left\{\begin{pmatrix}\dfrac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_1}\boldsymbol{u} & \cdots & \dfrac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_n}\boldsymbol{u}\end{pmatrix}\boldsymbol{b}\right\}\end{aligned}$$

$$
\begin{aligned}
&= \left\{ f_{i\boldsymbol{a}}\left(\boldsymbol{a},\boldsymbol{u}\right) - \left(\frac{\partial \boldsymbol{K}\left(\boldsymbol{a}\right)}{\partial a_1}\boldsymbol{u} \quad \cdots \quad \frac{\partial \boldsymbol{K}\left(\boldsymbol{a}\right)}{\partial a_n}\boldsymbol{u} \right)^{\top} \boldsymbol{v}_i \right\} \cdot \boldsymbol{b} \\
&= \boldsymbol{g}_i \cdot \boldsymbol{b}.
\end{aligned}
$$

This result agrees with the formula obtained from the direct differentiation method (see Sect. 2.6.5).

The discussion above leads us to the following observations regarding the adjoint variable method.

Remark 1.2.5 (Properties of the adjoint variable method)
In comparison with the method of direct differentiation, the adjoint variable method displays the following characteristics:

(1) When the number of cost functions is large ($m \gg 1$), the adjoint variable method is ineffective because the number of adjoint problems is the same as the number of cost functions, $m + 1$.

(2) When the number of design variables is large ($n \gg 1$), the adjoint variable method is effective because the number of adjoint problems, $m + 1$, does not depend on the number of design variables, n.

(3) The number of variables in an adjoint problem is the same as the number of variables in the state equation, n. This indicates the fact that the adjoint variable method is applicable even when the state variable is a function of time or defined over a domain (in such a case, the linear space for the state variable becomes infinite-dimensional).

(4) If a self-adjoint relationship is satisfied, then there is no need to explicitly solve the adjoint problem.

□

Beginning in Chap. 5, the state equation will be assumed to be a partial differential equation given by a boundary value problem. In this case, the state variable becomes a function defined in a domain of $d \in \{2, 3\}$-dimensional space. Remark 1.2.5 (3) above indicates that the adjoint variable method is indispensable when constructing cost function derivatives with respect to design variables while satisfying the state equation in shape and topology optimization problems. However, in cases wherein a relation such as Eq. (1.2.5) could be obtained by a simple partial differential equation (for example, Eq. (8.5.17) in Chap. 8) under appropriate assumptions, then the direct differentiation method is effective in shape and topology optimization problems.

1.3 An Optimal Design Problem of a Branched One-Dimensional Stokes Flow Field

So far we have looked at what optimal design problems are by considering one-dimensional linear elastic bodies. Let us now continue our investigation by

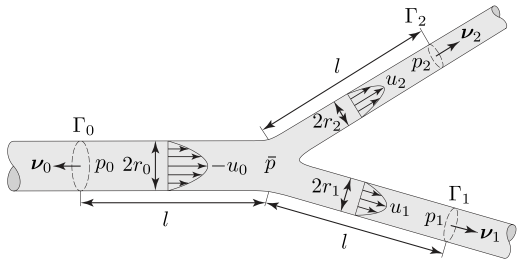

Fig. 1.6: A branched 1D Stokes flow field.

taking a look at how a similar optimal design problem can be constructed when the design target is changed to a flow field.

Consider a viscous flow field within a circular tube such as that shown in Fig. 1.6. This problem hints at Murray's law. By minimizing the sum of a blood flow transportation cost under a volume constraint, Murray showed that fluid flow through blood vessel cross-sections is proportional to the cube of the vessel radius [6]. Murray's analysis did not include a branched tube, but here we will take the cross-sectional areas as the design variables to see if the relationship still holds. Using the same cost function, Murray also derived a branch law for branch angles. The interested reader is referred to [5].

1.3.1 State Determination Problem

In Fig. 1.6, r_0, r_1, and r_2 denote the radii of their respective circular tubes, p_0, p_1, and p_2 signify the pressure at the inflow cross-section Γ_0 and the outflow cross-sections Γ_1 and Γ_2, respectively. The pressure at the branched cross-section is denoted by $\bar{p}$, l represents the length, and μ is the viscosity coefficient. The flow velocity at a radius r within a circular cross-section of radius r_i, $i \in \{0, 1, 2\}$, is assumed to be given by the Hagen-Poiseuille flow:

$$u_{\mathrm{H}i}(r) = -\frac{p_i - \bar{p}}{4\mu l}\left(r_i^2 - r^2\right), \tag{1.3.1}$$

which is derived from the solution of a stationary Stokes equation with respect to a cylindrical boundary. Note that the flow velocities within the three tubes are taken to be positive in the outward normal directions $\boldsymbol{\nu}_0$, $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ of the cross-sections Γ_0, Γ_1 and Γ_2, respectively. Meanwhile, due to the fact that $u_{\mathrm{H}0}(r)$ flows inward from Γ_0 in the opposite direction of $\boldsymbol{\nu}_0$, its flow velocity is negative. Here, let the volume of fluid flow per unit time through the tube of radius r_i be

$$u_i = \int_0^{r_i} u_{\mathrm{H}i}(r)\, 2\pi r \,\mathrm{d}r = \frac{\bar{p} - p_i}{8\pi\mu l}\left(\pi r_i^2\right)^2 = (\bar{p} - p_i)\, a_i^2, \tag{1.3.2}$$

where

$$\boldsymbol{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \frac{1}{\sqrt{8\pi\mu l}} \begin{pmatrix} \pi r_0^2 \\ \pi r_1^2 \\ \pi r_2^2 \end{pmatrix}. \tag{1.3.3}$$

According to Eq. (1.3.3), the cross-sectional area of Γ_i is $\sqrt{8\pi\mu l} a_i$. However, for the sake of simplicity a_i will be referred to as the cross-section and $\boldsymbol{a} \in X \in \mathbb{R}^3$ will be used as the design variable. Moreover, the continuity equation states that

$$u_0 + u_1 + u_2 = 0. \tag{1.3.4}$$

Substituting Eq. (1.3.2) into Eq. (1.3.4) yields

$$\bar{p} = \frac{p_0 a_0^2 + p_1 a_1^2 + p_2 a_2^2}{a_0^2 + a_1^2 + a_2^2}. \tag{1.3.5}$$

If Eq. (1.3.5) is substituted into Eq. (1.3.2), and $\bar{p}$ is eliminated, then we obtain

$$\frac{1}{a_0^2 + a_1^2 + a_2^2} \begin{pmatrix} a_0^2 \left(a_1^2 + a_2^2\right) & -a_0^2 a_1^2 & -a_0^2 a_2^2 \\ -a_0^2 a_1^2 & a_1^2 \left(a_0^2 + a_2^2\right) & -a_1^2 a_2^2 \\ -a_0^2 a_2^2 & -a_1^2 a_2^2 & a_2^2 \left(a_0^2 + a_1^2\right) \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = - \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}. \tag{1.3.6}$$

In this section, we will assume that the volume of fluid flow per unit time $\boldsymbol{u} = (u_1, u_2)^\top \in \mathbb{R}^2$ is known and that u_0 is given by Eq. (1.3.4). The values relating to the flow velocity are assumed to be known because, as will be shown in Chap. 5, the existence of solutions to the Stokes problem is guaranteed when the flow velocity is given along the entire boundary (Theorem 5.6.3). However, the matrix of coefficients in this equation is singular because the equation holds regardless of the selected values for the average pressure (uncertainty with respect to the constant term). We therefore set $p_0 = 0$ for convenience. When the flow volume per unit time $\boldsymbol{u}$ and the design variable $\boldsymbol{a}$ are given, the pressure $\boldsymbol{p} = (p_1, p_2)^\top \in P = \mathbb{R}^2$ is then uniquely determined by

$$\frac{1}{a_0^2 + a_1^2 + a_2^2} \begin{pmatrix} a_1^2 \left(a_0^2 + a_2^2\right) & -a_1^2 a_2^2 \\ -a_1^2 a_2^2 & a_2^2 \left(a_0^2 + a_1^2\right) \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{1.3.7}$$

Hence, $\boldsymbol{p} = (p_1, p_2)^\top \in P = \mathbb{R}^2$ is used as the state variable in this section and the state determination problem is defined as follows.

Problem 1.3.1 (1D branched Stokes flow field) Consider the one-dimensional Stokes flow field of Fig. 1.6. Let $\boldsymbol{a} \in \mathbb{R}^3$ and $\boldsymbol{u} \in \mathbb{R}^2$ be given. Find $\boldsymbol{p} \in \mathbb{R}^2$ satisfying

$$\boldsymbol{A}(\boldsymbol{a})\,\boldsymbol{p} = -\boldsymbol{u}. \tag{1.3.8}$$

Here, Eq. (1.3.8) expresses Eq. (1.3.7) in vector notation. □

The solution of Eq. (1.3.8) is

$$\begin{aligned}
\boldsymbol{p} &= -\boldsymbol{A}^{-1}(\boldsymbol{a})\,\boldsymbol{u} \\
&= -\begin{pmatrix} \dfrac{1}{a_0^2}+\dfrac{1}{a_1^2} & \dfrac{1}{a_0^2} \\ \dfrac{1}{a_0^2} & \dfrac{1}{a_0^2}+\dfrac{1}{a_2^2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\
&= -\begin{pmatrix} \dfrac{u_1}{a_1^2}+\dfrac{u_1+u_2}{a_0^2} \\ \dfrac{u_2}{a_2^2}+\dfrac{u_1+u_2}{a_0^2} \end{pmatrix} = -\begin{pmatrix} \dfrac{u_1}{a_1^2}-\dfrac{u_0}{a_0^2} \\ \dfrac{u_2}{a_2^2}-\dfrac{u_0}{a_0^2} \end{pmatrix}.
\end{aligned} \tag{1.3.9}$$

For later use, we define a Lagrange function with respect to Problem 1.3.1:

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{q}\cdot(\boldsymbol{A}(\boldsymbol{a})\,\boldsymbol{p}+\boldsymbol{u}), \tag{1.3.10}$$

where $\boldsymbol{p}$ is not necessarily the solution of Problem 1.3.1 and $\boldsymbol{q} = (q_1, q_2)^\top \in \mathbb{R}^2$ is introduced as a Lagrange multiplier. Comparing Eq. (1.3.10) with Eq. (1.1.12), the sign convention for the Lagrangian looks different. This sign was decided to obtain the self-adjoint relationship of Eq. (1.3.18) together with the sign of $\mathscr{L}_{\mathrm{S}}$ in $\mathscr{L}_0$ defined later in Eq. (1.3.14). If the opposite sign was used for the right-hand side of Eq. (1.3.10), then the self-adjoint relationship of Eq. (1.3.18) holds with the opposite sign. This difference comes from the difference of the objective functions. The mean compliance represents the magnitude of displacement, while the mean flow resistance represents the negative value of the magnitude of flow velocity. Problem 1.3.1 is equivalent to finding $\boldsymbol{p}$ satisfying

$$\mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}) = 0,$$

for all $\boldsymbol{q} \in \mathbb{R}^2$.

1.3.2 An Optimal Design Problem

Having defined the state determination problem, we now establish a cost function using the design variable $\boldsymbol{a}$ and the state variable $\boldsymbol{p}$.

We want to construct an objective function that measures flow resistance. According to the law of conservation of energy, the energy lost through viscosity per unit time inside the viscous flow field is equal to the negative value of the power (energy per unit time) integrated along the boundary. Now, let the objective function be

$$f_0 = -(p_0u_0 + p_1u_1 + p_2u_2) = -\boldsymbol{p}\cdot\boldsymbol{u}, \tag{1.3.11}$$

where we have used the fact that $p_0 = 0$. In Eq. (1.3.11), assuming $\boldsymbol{u}$ is given, the minimization of f_0 means the maximization of $\boldsymbol{p}$ at the out flow boundaries. It means the minimization of pressure loss. Then, this function corresponds to values often referred to as dissipation energy or power loss. However, in Chaps.

8 and 9, an extension of this definition (referred to as the mean flow resistance) will be used as a cost function to measure flow resistance in a Stokes flow field. The reason for this terminology is because, in that scenario, the quantity does not represent dissipation energy. For this reason, f_0 of Eq. (1.3.11) will also be referred to as the mean flow resistance in a one-dimensional branched Stokes flow field.

The volume constraint function is taken as

$$f_1(\boldsymbol{a}) = l(a_0 + a_1 + a_2) - c_1, \tag{1.3.12}$$

where c_1 is a positive constant.

Having defined these cost functions, the optimization problem for the one-dimensional branched Stokes flow field is defined as follows. We take $X = \mathbb{R}^3$ as the linear space for the design variable $\boldsymbol{a}$ and, with respect to a constant vector $\boldsymbol{a}_0 > \boldsymbol{0}_{\mathbb{R}^3}$, we set

$$\mathcal{D} = \{\boldsymbol{a} \in X \mid \boldsymbol{a} \geq \boldsymbol{a}_0\}. \tag{1.3.13}$$

Furthermore, $P = \mathbb{R}^2$ denotes the linear space for the respective state variables, $\boldsymbol{p}$.

Problem 1.3.2 (Mean flow resistance minimization) Let $X = \mathbb{R}^3$, $P = \mathbb{R}^2$, and $\mathcal{D}$ be given by Eq. (1.3.13). Furthermore, let $f_0(\boldsymbol{p})$ and $f_1(\boldsymbol{a})$ be given by Eq. (1.3.11) and Eq. (1.3.12), respectively. Find $\boldsymbol{a}$ satisfying

$$\min_{(\boldsymbol{a},\boldsymbol{p}) \in \mathcal{D} \times P} \{f_0(\boldsymbol{p}) \mid f_1(\boldsymbol{a}) \leq 0, \text{ Problem 1.3.1}\}.$$

□

1.3.3 Cross-Sectional Derivatives

The derivative $\tilde{f}_0'(\boldsymbol{a})[\boldsymbol{b}] = f_0'(\boldsymbol{p}(\boldsymbol{a}))[\boldsymbol{b}] = \boldsymbol{g}_0 \cdot \boldsymbol{b}$ of f_0 with respect to a variation $\boldsymbol{b}$ of $\boldsymbol{a}$ is called the cross-sectional derivative and $\boldsymbol{g}_0$ is referred to as the cross-sectional-area gradient. Let us now find $\boldsymbol{g}_0$ and the Hesse matrix $\boldsymbol{H}_0$ of f_0 using the adjoint variable method.

The Lagrange function with respect to f_0 is taken as

$$\begin{aligned}\mathscr{L}_0(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0) &= f_0(\boldsymbol{p}) - \mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)\\ &= -\boldsymbol{p}\cdot\boldsymbol{u} - \boldsymbol{q}_0 \cdot (\boldsymbol{A}(\boldsymbol{a})\boldsymbol{p} + \boldsymbol{u}),\end{aligned} \tag{1.3.14}$$

where $\boldsymbol{q}_0 \in P$ is the adjoint variable (Lagrange multiplier) with respect to the state equation (prepared for f_0). The derivative of $\mathscr{L}_0$ with respect to an arbitrary variation $(\boldsymbol{b}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}_0) \in X \times P \times P$ of $(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)$ is

$$\begin{aligned}&\mathscr{L}_0'(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\boldsymbol{b},\hat{\boldsymbol{p}},\hat{\boldsymbol{q}}_0]\\ &\quad = \mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\boldsymbol{b}] + \mathscr{L}_{0\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{p}}] + \mathscr{L}_{0\boldsymbol{q}_0}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{q}}_0].\end{aligned} \tag{1.3.15}$$

The third term on the right-hand side of Eq. (1.3.15) satisfies

$$\mathscr{L}_{0\boldsymbol{q}_0}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{q}}_0] = -\mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{p},\hat{\boldsymbol{q}}_0). \tag{1.3.16}$$

This term is zero if $\boldsymbol{p}$ solves the state determination problem.

The second term on the right-hand side of Eq. (1.3.15) can be calculated:

$$\begin{aligned}\mathscr{L}_{0\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{p}}] &= f_{0\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p})[\hat{\boldsymbol{p}}] - \mathscr{L}_{\mathrm{S}\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{p}}]\\ &= -\mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{q}_0,\hat{\boldsymbol{p}}).\end{aligned} \tag{1.3.17}$$

This term can also be made to take the value zero, provided the self-adjoint relationship holds:

$$\boldsymbol{q}_0 = \boldsymbol{p}. \tag{1.3.18}$$

Furthermore, direct calculation shows that the first term on the right-hand side of Eq. (1.3.15) can be expressed as

$$\begin{aligned}&\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\boldsymbol{b}]\\ &\quad= -\frac{1}{\left(a_0^2+a_1^2+a_2^2\right)^2}\\ &\qquad\times\begin{pmatrix} 2a_0\left(a_1^2p_1+a_2^2p_2\right)\left(a_1^2q_{01}+a_2^2q_{02}\right)\\ 2a_1\left\{a_0^2p_1+a_2^2(p_1-p_2)\right\}\left\{a_0^2q_{01}+a_2^2(q_{01}-q_{02})\right\}\\ 2a_2\left\{a_0^2p_2+a_1^2(p_2-p_1)\right\}\left\{a_0^2q_{02}+a_1^2(q_{02}-q_{01})\right\}\end{pmatrix}\cdot\begin{pmatrix}b_0\\ b_1\\ b_2\end{pmatrix}\\ &\quad= -2\begin{pmatrix}u_0^2/a_0^3\\ u_1^2/a_1^3\\ u_2^2/a_2^3\end{pmatrix}\cdot\begin{pmatrix}b_0\\ b_1\\ b_2\end{pmatrix} = \boldsymbol{g}_0\cdot\boldsymbol{b}.\end{aligned} \tag{1.3.19}$$

Here, the self-adjoint relationship has been used along with the fact that $\boldsymbol{p}$ is a solution of the state determination problem.

It can also be easily seen that

$$f_1'(\boldsymbol{a})[\boldsymbol{b}] = l\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\cdot\begin{pmatrix}b_0\\ b_1\\ b_2\end{pmatrix} = \boldsymbol{g}_1\cdot\boldsymbol{b}. \tag{1.3.20}$$

Furthermore, the Hesse matrix $\boldsymbol{H}_0$ of the mean flow resistance $\tilde{f}_0(\boldsymbol{a}) = f_0(\boldsymbol{a},\boldsymbol{p}(\boldsymbol{a}))$ can be obtained as follows. As described in Sect. 1.1.6, we use the adjoint variable method. The second-order derivative of the Lagrange function $\mathscr{L}_0$ with respect to arbitrary variations $(\boldsymbol{b}_1,\hat{\boldsymbol{p}}_1)\in X\times P$ and $(\boldsymbol{b}_2,\hat{\boldsymbol{p}}_2)\in X\times P$ of the design and state variables $(\boldsymbol{a},\boldsymbol{p})$ is computed as follows:

$$\begin{aligned}&\mathscr{L}_{0(\boldsymbol{a},\boldsymbol{p}),(\boldsymbol{a},\boldsymbol{p})}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[(\boldsymbol{b}_1,\hat{\boldsymbol{p}}_1),(\boldsymbol{b}_2,\hat{\boldsymbol{p}}_2)]\\ &\quad= \left(\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\boldsymbol{b}_1] + \mathscr{L}_{0\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{p}}_1]\right)_{\boldsymbol{a}}[\boldsymbol{b}_2]\\ &\qquad+ \left(\mathscr{L}_{0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\boldsymbol{b}_1] + \mathscr{L}_{0\boldsymbol{p}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[\hat{\boldsymbol{p}}_1]\right)_{\boldsymbol{u}}[\hat{\boldsymbol{p}}_2]\\ &\quad= \begin{pmatrix}\boldsymbol{b}_2\\ \hat{\boldsymbol{p}}_2\end{pmatrix}\cdot\left(\begin{pmatrix}\mathscr{L}_{0\boldsymbol{a}\boldsymbol{a}} & \mathscr{L}_{0\boldsymbol{a}\boldsymbol{p}}\\ \mathscr{L}_{0\boldsymbol{p}\boldsymbol{a}} & \mathscr{L}_{0\boldsymbol{p}\boldsymbol{p}}\end{pmatrix}\begin{pmatrix}\boldsymbol{b}_1\\ \hat{\boldsymbol{p}}_1\end{pmatrix}\right).\end{aligned} \tag{1.3.21}$$

Here, $\hat{\boldsymbol{p}}_1$ and $\hat{\boldsymbol{p}}_2$ are replaced by variations satisfying the state determination problem. By taking the partial derivative of Eq. (1.3.8) with respect to a_i for $i \in \{1, 2\}$, we obtain

$$\frac{\partial \boldsymbol{A}}{\partial a_i}\boldsymbol{p} + \boldsymbol{A}\frac{\partial \boldsymbol{p}}{\partial a_i} = \boldsymbol{0}_{\mathbb{R}^2}. \tag{1.3.22}$$

Solving for $\partial \boldsymbol{p}/\partial a_i$, we get

$$\frac{\partial \boldsymbol{p}}{\partial a_i} = -\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial a_i}\boldsymbol{p}, \tag{1.3.23}$$

and set

$$\hat{\boldsymbol{p}}(\boldsymbol{a})[\boldsymbol{b}] = \frac{\partial \boldsymbol{p}}{\partial \boldsymbol{a}^\top}\boldsymbol{b} = \begin{pmatrix} \partial p_1/\partial a_0 & \partial p_1/\partial a_1 & \partial p_1/\partial a_2 \\ \partial p_2/\partial a_0 & \partial p_2/\partial a_1 & \partial p_2/\partial a_2 \end{pmatrix}\begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}. \tag{1.3.24}$$

Substituting Eq. (1.3.24) into Eq. (1.3.21), we obtain

$$\mathscr{L}_{0(\boldsymbol{a},\boldsymbol{p}),(\boldsymbol{a},\boldsymbol{p})}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{q}_0)[(\boldsymbol{b}_1,\hat{\boldsymbol{p}}(\boldsymbol{a})[\boldsymbol{b}_1]),(\boldsymbol{b}_2,\hat{\boldsymbol{p}}(\boldsymbol{a})[\boldsymbol{b}_2])] = \boldsymbol{b}_1 \cdot (\boldsymbol{H}_0\boldsymbol{b}_2), \tag{1.3.25}$$

and we see that, if the self-adjoint relationship is used along with the fact that $\boldsymbol{p}$ is a solution of the state determination problem, the Hesse matrix of $\tilde{f}_0$ is expressed as

$$\begin{aligned} \boldsymbol{H}_0 &= \mathscr{L}_{0\boldsymbol{a}\boldsymbol{a}} + \mathscr{L}_{0\boldsymbol{a}\boldsymbol{p}}\frac{\partial \boldsymbol{p}}{\partial \boldsymbol{a}^\top} + \left(\mathscr{L}_{0\boldsymbol{a}\boldsymbol{p}}\frac{\partial \boldsymbol{p}}{\partial \boldsymbol{a}^\top}\right)^\top \\ &= 6\begin{pmatrix} u_0^2/a_0^4 & 0 & 0 \\ 0 & u_1^2/a_1^4 & 0 \\ 0 & 0 & u_2^2/a_2^4 \end{pmatrix}. \end{aligned} \tag{1.3.26}$$

This formula matches the result obtained through partial differentiation of $\boldsymbol{g}_0$ in Eq. (1.3.19) with respect to $\boldsymbol{a}$. This relationship holds true because the state variable $\boldsymbol{p}$ is not included in $\boldsymbol{g}_0$ of Eq. (1.3.19). In this way, we observe that $\boldsymbol{H}_0$ is positive definite.

Let us, in addition, obtain the Hesse matrix of $\tilde{f}_0$ by the Lagrange multiplier method. The Lagrange function for $\boldsymbol{g}_0 \cdot \boldsymbol{b}_1$ can be defined as

$$\begin{aligned} &\mathscr{L}_{\mathrm{I}0}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{r}_0) \\ &\quad = \boldsymbol{g}_0(\boldsymbol{p}) \cdot \boldsymbol{b}_1 - \mathscr{L}_{\mathrm{S}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{r}_0) \\ &\quad = \boldsymbol{g}_0(\boldsymbol{p}) \cdot \boldsymbol{b}_1 - \boldsymbol{r}_0 \cdot (\boldsymbol{A}(\boldsymbol{a})\boldsymbol{p} + \boldsymbol{u}), \end{aligned} \tag{1.3.27}$$

where $\boldsymbol{g}_0(\boldsymbol{p})$ is defined in the second equation of Eq. (1.3.19) substituting Eq. (1.3.18), $\mathscr{L}_{\mathrm{S}}$ in Eq. (1.3.10). $\boldsymbol{r}_0 = (r_{01}, r_{02})^\top \in P = \mathbb{R}^2$ is the adjoint variable corresponding to $\boldsymbol{p}$ in $\boldsymbol{g}_0(\boldsymbol{p})$ satisfying the state determination problem. $\boldsymbol{b}_1$ was assumed to be a constant vector in $\mathscr{L}_{\mathrm{I}0}$.

With respect to arbitrary variations $(\boldsymbol{b}_2, \hat{\boldsymbol{p}}, \hat{\boldsymbol{r}}_0) \in X \times P^2$ of $(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)$, the derivative of $\mathscr{L}_{\mathrm{I}0}$ is written as

$$\begin{aligned}&\mathscr{L}'_{\mathrm{I}0}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\boldsymbol{b}_2, \hat{\boldsymbol{p}}, \hat{\boldsymbol{r}}_0] \\ &\quad = \mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\boldsymbol{b}_2] + \mathscr{L}_{\mathrm{I}0\boldsymbol{p}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\hat{\boldsymbol{p}}] + \mathscr{L}_{\mathrm{I}0\boldsymbol{r}_0}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\hat{\boldsymbol{r}}_0].\end{aligned} \tag{1.3.28}$$

The third term on the right-hand side of Eq. (1.3.28) vanishes if $\boldsymbol{p}$ is the solution of the state determination problem.

The second term on the right-hand side of Eq. (1.3.28) can be written as

$$\begin{aligned}\mathscr{L}_{\mathrm{I}0\boldsymbol{p}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\hat{\boldsymbol{p}}] &= \boldsymbol{g}_{0\boldsymbol{p}^\top}(\boldsymbol{p})[\hat{\boldsymbol{p}}] \cdot \boldsymbol{b}_1 - \mathscr{L}_{\mathrm{S}\boldsymbol{p}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\hat{\boldsymbol{p}}] \\ &= \hat{\boldsymbol{p}} \cdot \boldsymbol{w} - \boldsymbol{r}_0 \cdot (\boldsymbol{A}(\boldsymbol{a})\hat{\boldsymbol{p}}) \\ &= \hat{\boldsymbol{p}} \cdot \left(\boldsymbol{w} - \boldsymbol{A}^\top(\boldsymbol{a})\boldsymbol{r}_0\right),\end{aligned} \tag{1.3.29}$$

where

$$\boldsymbol{w} = \boldsymbol{g}_{0\boldsymbol{p}}^\top(\boldsymbol{p})\boldsymbol{b}_1 = \begin{pmatrix} \dfrac{\partial g_{01}}{\partial p_1} & \dfrac{\partial g_{02}}{\partial p_1} & \dfrac{\partial g_{03}}{\partial p_1} \\ \dfrac{\partial g_{01}}{\partial p_2} & \dfrac{\partial g_{02}}{\partial p_2} & \dfrac{\partial g_{03}}{\partial p_2} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{12} \\ b_{13} \end{pmatrix}. \tag{1.3.30}$$

Here, the condition that Eq. (1.3.29) is zero for arbitrary $\hat{\boldsymbol{p}} \in P$ is equivalent to setting $\boldsymbol{r}_0$ to be the solution of the following adjoint problem.

Problem 1.3.3 (Adjoint problem of $\boldsymbol{r}_0$ with respect to $\boldsymbol{g}_0(\boldsymbol{p}) \cdot \boldsymbol{b}_1$) Let $\boldsymbol{A}(\boldsymbol{a})$ be as in Problem 1.3.1, and $\boldsymbol{w}$ be given by Eq. (1.3.30). Find $\boldsymbol{r}_0 \in P$ satisfying

$$\boldsymbol{A}^\top(\boldsymbol{a})\boldsymbol{r}_0 = \boldsymbol{w}.$$

□

The solution of Problem 1.3.3 is

$$\boldsymbol{r}_0 = \left(\boldsymbol{A}^\top(\boldsymbol{a})\right)^{-1} \boldsymbol{g}_{0\boldsymbol{p}}^\top(\boldsymbol{p})\boldsymbol{b}_1. \tag{1.3.31}$$

Finally, the first term on the right-hand side of Eq. (1.3.28) becomes

$$\begin{aligned}&\mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0)[\boldsymbol{b}_2] \\ &\quad = -\left\{\boldsymbol{r}_0 \cdot \left(\frac{\partial \boldsymbol{A}(\boldsymbol{a})}{\partial a_1}\boldsymbol{p} \quad \frac{\partial \boldsymbol{A}(\boldsymbol{a})}{\partial a_2}\boldsymbol{p} \quad \frac{\partial \boldsymbol{A}(\boldsymbol{a})}{\partial a_3}\boldsymbol{p}\right)\right\}\boldsymbol{b}_2.\end{aligned} \tag{1.3.32}$$

Here, substituting Eq. (1.3.31) into Eq. (1.3.32), we have the relation

$$\mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{p}, \boldsymbol{r}_0(\boldsymbol{b}_1))[\boldsymbol{b}_2] = h_0(\boldsymbol{a})[\boldsymbol{b}_1, \boldsymbol{b}_2] = \boldsymbol{g}_{\mathrm{H}0}(\boldsymbol{a}, \boldsymbol{b}_1) \cdot \boldsymbol{b}_2, \tag{1.3.33}$$

where the Hesse gradient of the mean flow resistance $\boldsymbol{g}_{\mathrm{H}0}$ is given by

$$\boldsymbol{g}_{\mathrm{H}0}(\boldsymbol{a},\boldsymbol{b}_1) = \mathscr{L}_{\mathrm{I}0\boldsymbol{a}}(\boldsymbol{a},\boldsymbol{p},\boldsymbol{r}_0(\boldsymbol{b}_1)). \tag{1.3.34}$$

The above results show that the function $\tilde{f}_0(\boldsymbol{a})$, which is obtained by rewriting the mean flow resistance $f_0(\boldsymbol{p})$ as a function of $\boldsymbol{a}$ only, is convex. As we will now show, as in the case of the mean compliance problem, $\boldsymbol{a}$ yields the minimum when it satisfies a set of KKT conditions.

1.3.4 Optimality Conditions

Let us again consider optimality using the KKT conditions. The Lagrange function with respect to the optimization problem (Problem 1.3.2) is set as

$$\mathscr{L}(\boldsymbol{a},\lambda_1) = \tilde{f}_0(\boldsymbol{a}) + \lambda_1 f_1(\boldsymbol{a}),$$

where $\lambda_1 \in \mathbb{R}$ is a Lagrange multiplier with respect to $f_1(\boldsymbol{a}) \le 0$. In this case, the KKT conditionsKarush–Kuhn–Tucker conditions of Problem 1.3.2 are given by

$$\mathscr{L}_{\boldsymbol{a}}(\boldsymbol{a},\lambda_1) = \boldsymbol{g}_0 + \lambda_1 \boldsymbol{g}_1 = \boldsymbol{0}_{\mathbb{R}^2}, \tag{1.3.35}$$

$$\mathscr{L}_{\lambda_1}(\boldsymbol{a},\lambda_1) = f_1(\boldsymbol{a}) \le 0, \tag{1.3.36}$$

$$\lambda_1 f_1(\boldsymbol{a}) = 0, \tag{1.3.37}$$

$$\lambda_1 \ge 0. \tag{1.3.38}$$

Equation (1.3.35) then becomes

$$-2\begin{pmatrix} u_0^2/a_0^3 \\ u_1^2/a_1^3 \\ u_2^2/2_1^3 \end{pmatrix} + \lambda_1 l \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

and the optimality condition with respect to the mean flow resistance minimization problem (Problem 1.3.2) is given by

$$2\frac{u_0^2}{a_0^3 l} = 2\frac{u_1^2}{a_1^3 l} = 2\frac{u_2^2}{a_2^3 l} = \lambda_1. \tag{1.3.39}$$

This optimality condition shows that Murray's law holds. In fact, using $u_{\mathrm{H}i}(r)$ from Eq. (1.3.1) shows that the shear strain velocity and shear stress on the walls can be expressed as

$$\gamma_i = \left.\frac{\mathrm{d}u_{\mathrm{H}i}}{\mathrm{d}r}\right|_{r=r_i} = -\frac{\bar{p}-p_i}{2\mu l} r_i = -\frac{u_i}{2\mu l a_i^2} r_i, \tag{1.3.40}$$

$$\tau_i = \mu\gamma_i = -\frac{u_i}{2 l a_i^2} r_i, \tag{1.3.41}$$

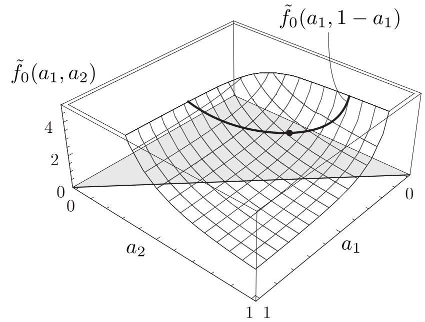

Fig. 1.7: Numerical example of mean flow resistance minimization problem.

respectively, for each $i \in \{0, 1, 2\}$. Using Eq. (1.3.3), we then obtain the following relation regarding the dissipation energy density:

$$\frac{1}{2}\tau_i \gamma_i = \frac{1}{\sqrt{8\mu l}} \frac{u_i^2}{a_i^3 l} = \frac{\sqrt{8\mu l}}{2}\lambda_1. \tag{1.3.42}$$

When the shear stresses are the same, this condition shows that the flow volume is proportional to the cube of the blood vessel radius (Murray's law).

1.3.5 Numerical Example

Let us now consider finding a minimizer through a specific exercise.

Exercise 1.3.4 (Mean flow resistance minimization problem) Let $a_0 = 1$ in Problem 1.3.2 (a_0 is not included in the design variables). Also, let $l = 1$, $c_1 = 2$, $\boldsymbol{u} = (1/3, 2/3)^\top$, and $\boldsymbol{a}_0 = (0.1, 0.1, 0.1)^\top$. Find the minimizer $\boldsymbol{a}$. □

Answer If $\boldsymbol{p}$ in Eq. (1.3.9) is substituted into f_0 (defined by Eq. (1.3.11)), then we obtain:

$$\tilde{f}_0(\boldsymbol{a}) = \frac{1}{9}\left(9 + \frac{1}{a_1} + \frac{4}{a_2}\right). \tag{1.3.43}$$

Figure 1.7 shows $\tilde{f}_0$. From Eq. (1.3.19) and Eq. (1.3.20), we obtain the cross-sectional-area derivative of $\tilde{f}_0$ and f_1:

$$\boldsymbol{g}_0 = -\begin{pmatrix} 2/9a_1^3 \\ 8/9a_2^3 \end{pmatrix}, \quad \boldsymbol{g}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Here, if $\boldsymbol{a}$ yields the minimum, then from Eq. (1.3.39) we get

$$\lambda_1 = \frac{2}{9a_1^3} = \frac{8}{9a_2^3}.$$

If $\boldsymbol{a}$ is an element of $\mathcal{D}^\circ$ (defined by Eq. (1.3.13)), then λ_1 is positive and the inequality constraint with respect to f_1 holds as an equality (due to the complementarity condition). Upon substituting $a_2 = 1 - a_1$ into Eq. (1.3.43), and using the stationary condition with respect to variations in a_1:

$$\frac{\mathrm{d}}{\mathrm{d}a_1}\tilde{f}_0\left(a_1, 1-a_1\right) = \frac{1}{\left(1-a_1\right)^2} - \frac{4}{a_1^2} = 0,$$

we obtain

$$\begin{aligned} a_1 = {} & \frac{1}{5}\left(1 + 2^{4/3} - 2^{2/3}\right),\ \frac{1}{5}\left\{1 - 2^{1/3}\left(1 - \mathrm{i}\sqrt{3}\right) + 2^{-1/3}\left(1 + \mathrm{i}\sqrt{3}\right)\right\}, \\ & \frac{1}{5}\left\{1 - 2^{1/3}\left(1 + \mathrm{i}\sqrt{3}\right) + 2^{-1/3}\left(1 - \mathrm{i}\sqrt{3}\right)\right\} \\ = {} & 0.386488,\ 0.106756 + 0.711395\mathrm{i},\ 0.106756 - 0.711395\mathrm{i}. \end{aligned}$$

Here i denotes the imaginary unit and we remark that a_2 can be obtained from $1 - a_1$. Using these results, $\boldsymbol{a} \in \mathcal{D}^{\circ}$ is then given by $\boldsymbol{a} = (0.386488, 0.613512)^{\top}$. Since $\tilde{f}_0$ and f_1 are convex-functions, Problem 1.3.2 is a convex-optimization problem and, from Theorem 2.7.9, the $\boldsymbol{a}$ which satisfies the KKT condition yields the minimum. □

1.4 Summary

In order to have a conceptional idea about how optimality conditions are sought, this chapter examined some examples of optimal design problems related to one-dimensional elastic bodies and the one-dimensional Stokes flow field. The key points are outlined below:

(1) In addition to design variables that determine the system, optimal design problems include state variables which describe the system's desired state. The state equation which determines the state variables is an equality constraint. Cost functions can be defined as functions of the design and state variables (Sect. 1.1.2, Sect. 1.3.2).

(2) When the cost function is given as a function of a state variable, its derivative with respect to a variation of the design variable needs to be obtained in a manner that satisfies the equality constraints of the state determination problem. The direct differentiation method (Sect. 1.1.5, Sect. 1.2.1), which uses the chain rule of differentiation, and the adjoint variable method (Sect. 1.1.6, Sect. 1.2.2), which is based on the Lagrange multiplier method, are both conceivable methods for obtaining the derivative.

- The direct differentiation method is advantageous for problems involving multiple constraints (Remark 1.2.3).
- The adjoint variable method is advantageous for problems with multiple design variables (Remark 1.2.5).

Later in this book, partial differential equations (boundary value problems) are assumed to be the state equation, and the state variable becomes a function (an element of an infinite-dimensional space) defined in a $d \in \{2,3\}$-dimensional domain. The adjoint variable approach is essential in such settings.

(3) In a mean compliance minimization problem (Problem 1.1.4), where the cross-sectional area of the one-dimensional elastic body was the design variable, an optimality condition stating that the strain energy density is uniform was obtained in Eq. (1.1.56) (Sect. 1.1.7).

(4) In a mean flow resistance minimization problem (Problem 1.3.2), where the cross-sectional area of the branched one-dimensional Stokes flow field is the design variable, an optimality condition stating that the dissipative energy density is uniform was derived in Eq. (1.3.42) (Sect. 1.3.4).

There is a vast amount of literature regarding optimal design problems, and here we only mention a selection [1–4, 8].

1.5 Practice Problems

1.1 Consider the problem of finding $\boldsymbol{a}$ satisfying

$$\min_{(\boldsymbol{a},\boldsymbol{u})\in\mathcal{D}\times U} \left\{ f_0(\boldsymbol{u}) = u_2^2 \;\middle|\; f_1(\boldsymbol{a}) \leq 0,\ \text{Problem 1.1.3} \right\}.$$

This is Problem 1.1.4 with f_0 changed to u_2^2. Let the adjoint variable with respect to f_0 be $\boldsymbol{v}_0 \in \mathbb{R}^2$. Derive the adjoint equation. Also, express $\boldsymbol{g}_0$ in terms of $\boldsymbol{u}$ and $\boldsymbol{v}_0$.

1.2 In Sect. 1.1.6, using the stationary conditions with respect to an arbitrary variation of $\boldsymbol{u}$ and $\boldsymbol{v}_0$ of the Lagrange function in Problem 1.1.4, we derived the cross-sectional-area gradient $\boldsymbol{g}_0$ of f_0. In this case, the self-adjoint relationship was used. In fact, if the self-adjoint relationship holds, then the cross-sectional-area gradient $\boldsymbol{g}_0$ can be obtained without using a Lagrange function. In particular, instead of using the mean compliance f_0, the potential energy

$$\pi(\boldsymbol{a},\boldsymbol{u}) = \frac{1}{2}\boldsymbol{u}\cdot(\boldsymbol{K}(\boldsymbol{a})\boldsymbol{u}) - \boldsymbol{u}\cdot\boldsymbol{p}$$

can be used to consider the problem of finding $(\boldsymbol{a},\boldsymbol{u})$ satisfying

$$\max_{\boldsymbol{a}\in\mathcal{D}} \min_{\boldsymbol{u}\in U} \pi(\boldsymbol{a},\boldsymbol{u}).$$

When $\boldsymbol{u}$ satisfies $\min_{\boldsymbol{u}\in\mathbb{R}^2}\pi$, show that the cross-sectional-area gradient of $-\pi$ is equal to $1/2$ of the $\boldsymbol{g}_0$ in Eq. (1.1.36). Note that π in this problem is the potential energy. This problem shows that the minimization of the mean compliance is equivalent to the maximization of the potential energy.

1.3 Consider Practice **1.1**, where $l = 1$, $e_{\mathrm{Y}} = 1$, $c_1 = 1$ and $\boldsymbol{p} = (1,1)^\top$. Find the minimizer $\boldsymbol{a}$.

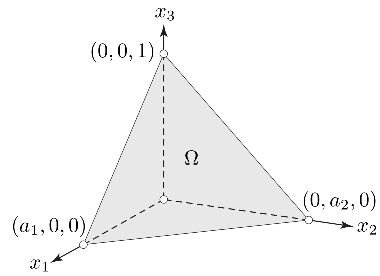

Fig. 1.8: Tetrahedron Ω.

1.4 Consider the state determination problem in Problem 1.1.4. Here, when $\boldsymbol{p} = (1, -1)^\top$, the stress of a one-dimensional linear elastic body with a cross-section of a_1 is zero. The side constraint with respect to the cross-section a_1 thus activates at the optimal solution. The optimal solution in this case is $(a_1, a_2) = (a_{01}, (c_1/l) - a_{01})$. Derive the KKT conditions in this situation and find the Lagrange multiplier. Here, assume that $\boldsymbol{g}_0$ and $\boldsymbol{g}_1$ are given by equations Eq. (1.1.28) and Eq. (1.1.17), respectively.

1.5 Consider the state determination problem of Problem 1.1.4 and assume that the design variable $\boldsymbol{a} \in \mathbb{R}^2$ is the length of one side of a square cross-section. Find the Hesse matrix $\boldsymbol{H}_0$ and the gradient $\boldsymbol{g}_0$ of f_0, with respect to variation of $\boldsymbol{a}$.

1.6 Consider the tetrahedron Ω shown in Fig. 1.8. Let the design variable be the lengths of the sides of the base $\boldsymbol{a} = (a_1, a_2)^\top \in \mathbb{R}^2$, and the cost function f be the volume of Ω. Find the Hesse matrix $\boldsymbol{H}$ and the gradient $\boldsymbol{g}$ of f, with respect to variation of $\boldsymbol{a}$.

1.7 Give a concrete expression for $\boldsymbol{K}(\boldsymbol{a})$ in Eq. (1.2.1).

1.8 The self-adjoint relationship was also established with respect to f_0 in Problem 1.3.2. Thus, in a manner similar to Practice **1.2**, the cross-sectional-area gradient $\boldsymbol{g}_0$ can be obtained without using a Lagrange function. Considering f_0, let us formally investigate the problem of finding $(\boldsymbol{a}, \boldsymbol{p})$ satisfying

$$\min_{\boldsymbol{a}\in\mathcal{D}} \max_{\boldsymbol{p}\in P} \pi(\boldsymbol{a}, \boldsymbol{p}),$$

where the potential energy of the dissipative system is set to be

$$\pi(\boldsymbol{a}, \boldsymbol{p}) = -\frac{1}{2}\boldsymbol{p} \cdot (\boldsymbol{A}(\boldsymbol{a})\,\boldsymbol{p}) - \boldsymbol{u} \cdot \boldsymbol{p}.$$

When $\boldsymbol{p}$ satisfying $\max_{\boldsymbol{p}\in\mathbb{R}^2} \pi$ is used, show that the cross-sectional-area gradient of π is the same as $1/2$ of $\boldsymbol{g}_0$ in Eq. (1.3.19). This problem shows that minimizing the mean flow resistance is the same as minimizing π when the potential energy of the dissipative system is formally set to π.

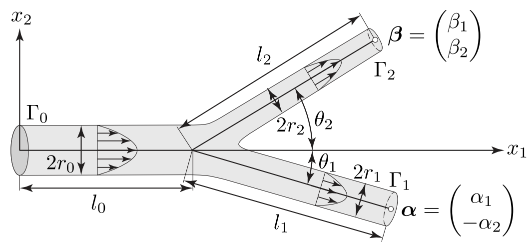

Fig. 1.9: The branch angle of a branched one-dimensional Stokes flow field.

1.9 Consider a branched one-dimensional Stokes flow field such as the one shown in Fig. 1.9. The center of the inflow boundary Γ_0 is taken to be the origin, while $\boldsymbol{\alpha} = (\alpha_1, -\alpha_2)^\top \in \mathbb{R}^2$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top \in \mathbb{R}^2$ (with respect to four positive constants α_1, α_2, β_1 and β_2) are set as the coordinates of the central position of the outward flow boundaries, Γ_1 and Γ_2. The radius of the tube is $\boldsymbol{r} = (r_0, r_1, r_2)^\top \in \mathbb{R}^3$. Assume that $\boldsymbol{r}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known and that the sum of the volumes of the three tubes is minimized. Use the branch angles θ_1 and θ_2 to show that

$$r_0^2 = r_1^2 \cos\theta_1 + r_2^2 \cos\theta_2.$$

References

[1] Arora, J. S. *Introduction to Optimum Design.* McGraw-Hill, New York; Tokyo, 1989.

[2] Choi, K. K. and Kim, N. H. *Structural Sensitivity Analysis and Optimization, 1 & 2.* Springer, New York, 2005.

[3] Haftka, R. T. and Gurdal, Z. *Elements of Structural Optimization. 3rd Rev. and Expanded Ed.* Kluwer Academic Publishers, Dordrecht, 1992.

[4] Haug, E. J., Choi, K. K., and Komkov, V. *Design Sensitivity Analysis of Structural Systems.* Academic Press, Orlando, 1986.

[5] Murray, C. D. The physiological principle of minimum work applied to the angle of branching of arteries. *The Journal of General Physiology*, 9:835–841, 1926.

[6] Murray, C. D. The physiological principle of minimum work. I. the vascular system and the cost of blood volume. *Proceedings of the National Academy of Sciences*, 12:207–214, 1926.

[7] Washizu, K. *Introduction to Energy Principles (in Japanese).* Baifukan, Tokyo, 1980.

[8] Yamakawa, H. and others, . *Optimum Design Handbook: Basics, Strategy and Application (in Japanese).* Asakura Publishing, Tokyo, 2003.