Contents

Appendices

A.1 Basic Terminology

Firstly, let us define the basic terminology used in mathematics.

A.1.1 Open Sets, Closed Sets and Bounded Sets

Let X be a metric space (Definition 4.2.8), for example $\mathbb{R}^2$, and S be its subset. We define the distance between two elements in X by $d : X \times X \to \mathbb{R}$. For $\boldsymbol{x} \in X$, the open ball $B(\boldsymbol{x}, \delta) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x}, \boldsymbol{y}) < \delta\}$ of radius $\delta > 0$ is called a neighborhood. A subset S is called an open set if, for an arbitrary $\boldsymbol{x} \in S$, there exists $\delta > 0$ such that $B(\boldsymbol{x}, \delta) \subset S$ (Fig. A.1). If $X \setminus S$ is an open set, S is said to be a closed set. The smallest closed set containing S is its closure and is denoted by $\bar{S}$. For an open set S, $\partial S = \bar{S} \setminus S$ is referred to as the boundary of S. $\boldsymbol{x} \in S$ is called an inner point and $\boldsymbol{x} \in \partial S$ a boundary point. For a closed set S, the set of all inner points is called the interior and is denoted by S°. Moreover, a subset S is a bounded set if there exists an arbitrary point $\boldsymbol{x}$ of X and $\beta > 0$ such that for every $\boldsymbol{y} \in S$, $d(\boldsymbol{x}, \boldsymbol{y}) \le \beta$ holds.

A.1.2 Continuity of Functions

The continuity of a function on a metric space is defined as follows. Let X and Y be metric spaces with corresponding distances $d_X : X \times X \to \mathbb{R}$ and $d_Y : Y \times Y \to \mathbb{R}$. A function $f : X \to Y$ is said to be continuous on $\boldsymbol{x} \in X$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$d_Y(f(\boldsymbol{y}), f(\boldsymbol{x})) < \epsilon$$

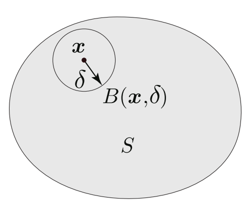


Fig. A.1: Open set S.

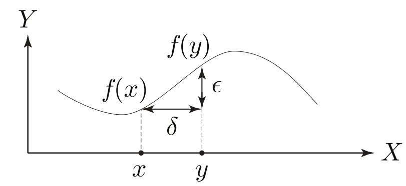


Fig. A.2: Continuity of a function.

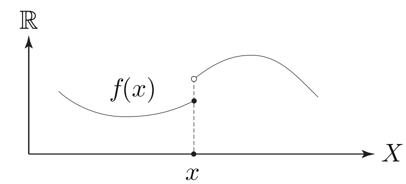


Fig. A.3: Lower semi-continuity of a function.

holds for any $\boldsymbol{y} \in X$ that satisfies $d_X(\boldsymbol{x}, \boldsymbol{y}) < \delta$ (Fig. A.2). Moreover, if it is continuous on all $\boldsymbol{x} \in X$, it is called uniformly continuous.

From this definition, a uniformly continuous function is continuous but its converse is not necessarily true. Take, for instance, $f(x) = x^2$. This function is clearly continuous on $\mathbb{R}$. However, as $|x| \to \infty$, the gradient quickly becomes large. In this case, $\delta \to 0$ for any given $\epsilon > 0$, contradicting the requirement that we can find a $\delta > 0$. Nevertheless, a continuous function is also uniformly continuous and bounded if its domain is a compact (bounded and closed) set.

Moreover, an extended real-valued function $f : X \to \mathbb{R} \cup \{-\infty, \infty\}$ is said to be lower semi-continuous on $\boldsymbol{x} \in X$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) - \epsilon$$

holds for any $\boldsymbol{y} \in X$ that satisfies $d_X(\boldsymbol{x}, \boldsymbol{y}) < \delta$ (Fig. A.3).

A.2 Positive Definiteness of Real Symmetric Matrix

In Chaps. 2 and 3, the positive definite real symmetric matrix plays an important role. Here, let us state a fundamental theorem used to check the positive definiteness of a matrix based on the notion of eigenvalues.

Theorem A.2.1 (Positive definiteness of a real symmetric matrix)
Let $\boldsymbol{A} \in \mathbb{R}^{d \times d}$ be a real symmetric matrix. The necessary and sufficient conditions for $\boldsymbol{A}$ to be positive definite is that all of its eigenvalues $\lambda_1, \cdots, \lambda_d \in \mathbb{R}$ are positive. □

Proof Let us show the necessity part. For each $i \in \{1, \ldots, d\}$, let $\boldsymbol{x}_i \in \mathbb{R}^d$ be the eigenvector corresponding to the eigenvalue λ_i. If $\boldsymbol{A}$ is positive definite, the condition

$$\boldsymbol{x}_i \cdot A\boldsymbol{x}_i = \boldsymbol{x}_i \cdot (\lambda_i \boldsymbol{x}_i) = \lambda_i \|\boldsymbol{x}_i\|^2 > 0$$

holds for every $i \in \{1, \ldots, d\}$. It follows that $\lambda_i > 0$ for all $i \in \{1, \ldots, d\}$. For the sufficiency part, it can be shown that $\boldsymbol{x}_i$, for $i \in \{1, \ldots, d\}$, are mutually orthogonal. An arbitrary vector $\boldsymbol{x} \in \mathbb{R}^d$ is given by a linear combination of d independent vectors. That is, we can write

$$\boldsymbol{x} = \sum_{i \in \{1,\ldots,d\}} \boldsymbol{x}_i \xi_i$$

and from this, we have

$$\boldsymbol{x} \cdot A\boldsymbol{x} = \sum_{i \in \{1,\ldots,d\}} \lambda_i \|\boldsymbol{x}_i\|^2 \xi_i^2 > 0$$

for all $\boldsymbol{x} \in \mathbb{R}^d \setminus \{\boldsymbol{0}_{\mathbb{R}^d}\}$. □

A basis for evaluating positive definiteness based on minors is known as Sylvester's criterion (see, e.g., [3]).

Theorem A.2.2 (Sylvester's criterion) Let $\boldsymbol{A} = (A_{ij})_{ij} \in \mathbb{R}^{d \times d}$ be a real symmetric matrix. The necessary and sufficient condition for $\boldsymbol{A}$ to be positive definite is that all minors with respect to $i \in \{1, \ldots, d\}$ satisfy

$$|\boldsymbol{A}_i| = \begin{vmatrix} A_{11} & \cdots & A_{1i} \\ \vdots & \ddots & \vdots \\ A_{i1} & \cdots & A_{ii} \end{vmatrix} > 0.$$

□

A.3 Null Space, Image space and Farkas's Lemma

In Chap. 2, the relationship between null space and image space as well as Farkas's lemma are used in the proofs of some important theorems. Here, let us summarize these auxiliary results.

Let m and n be natural numbers and $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{pmatrix} \in \mathbb{R}^{n \times m}$. The set

$$\operatorname{Ker} \boldsymbol{A} = \{\, \boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{A}\boldsymbol{y} = \boldsymbol{0}_{\mathbb{R}^n} \} \tag{A.3.1}$$

is called the null space or the kernel space of $\boldsymbol{A}$. Moreover,

$$\operatorname{Im} \boldsymbol{A} = \{\, \boldsymbol{A}\boldsymbol{z} \in \mathbb{R}^n \mid \boldsymbol{z} \in \mathbb{R}^m \} \tag{A.3.2}$$

is called the image space or the range space of $\boldsymbol{A}$. On the other hand, when a linear subspace V of $\mathbb{R}^n$ is given, the linear space constructed of all vectors orthogonal to V is referred to as the orthogonal complement space of V and is written as $V^{\perp}$. The spaces $\operatorname{Ker} \boldsymbol{A}$ and $\operatorname{Im} \boldsymbol{A}$ are related as follows.

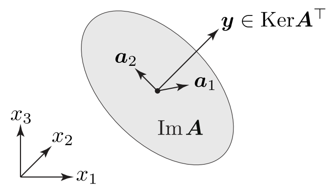


Fig. A.4: Null space and image space.

Lemma A.3.1 (Orthogonal complement of null space and image space) For a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, the following relations hold:

$$\mathrm{Im}\,\boldsymbol{A} = \left(\mathrm{Ker}\,\boldsymbol{A}^\top\right)^\perp, \quad \left(\mathrm{Im}\,\boldsymbol{A}^\top\right)^\perp = \mathrm{Ker}\,\boldsymbol{A}.$$

□

Proof By definition, we have

$$\mathrm{Ker}\,\boldsymbol{A}^\top = \left\{\boldsymbol{y} \in \mathbb{R}^n \;\middle|\; \boldsymbol{A}^\top \boldsymbol{y} = \boldsymbol{0}_{\mathbb{R}^m}\right\},$$
$$\left(\mathrm{Ker}\,\boldsymbol{A}^\top\right)^\perp = \left\{\boldsymbol{w} \in \mathbb{R}^n \;\middle|\; \boldsymbol{w}\cdot\boldsymbol{y} = 0,\ \boldsymbol{y} \in \mathbb{R}^n,\ \boldsymbol{A}^\top \boldsymbol{y} = \boldsymbol{0}_{\mathbb{R}^m}\right\}.$$

Let us choose a $\boldsymbol{z} \in \mathbb{R}^m$ such that, for any $\boldsymbol{w} \in \left(\mathrm{Ker}\,\boldsymbol{A}^\top\right)^\perp$, $\boldsymbol{w} = \boldsymbol{A}\boldsymbol{z}$. Note, however, that for any $\boldsymbol{y} \in \mathrm{Ker}\,\boldsymbol{A}^\top$, one can arbitrarily pick a $\boldsymbol{z} \in \mathbb{R}^m$ such that

$$\boldsymbol{w}\cdot\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{z})\cdot\boldsymbol{y} = \boldsymbol{z}\cdot\left(\boldsymbol{A}^\top\boldsymbol{y}\right) = 0.$$

Hence, form the definition of Eq. (A.3.2), the relation $\mathrm{Im}\,\boldsymbol{A} = \left(\mathrm{Ker}\,\boldsymbol{A}^\top\right)^\perp$ immediately follows. The same can be shown for $\left(\mathrm{Im}\,\boldsymbol{A}^\top\right)^\perp = \mathrm{Ker}\,\boldsymbol{A}$. □

For $n = 3$ and $m = 2$, the relationship between the null space and the image space can be geometrically illustrated as in Fig. A.4.

Next, let us think about the set when the equality constraint is replaced by an inequality condition. In relation to a null space, the set

$$\mathrm{Kco}\,\boldsymbol{A} = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{A}\boldsymbol{y} \le \boldsymbol{0}_{\mathbb{R}^n}\}$$

is called a non-positive cone. Moreover, with respect to an image space, the set

$$\mathrm{Ico}\,\boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{z} \in \mathbb{R}^n \mid \boldsymbol{z} \ge \boldsymbol{0}_{\mathbb{R}^m}\}$$

is called an image cone. Furthermore, when $C \subset \mathbb{R}^n$ is a cone,

$$C' = \{\boldsymbol{z} \in \mathbb{R}^n \mid \boldsymbol{y}\cdot\boldsymbol{z} \le 0 \text{ for all } \boldsymbol{y} \in C\}$$

is called the dual cone of C. The next result is a rephrased version of Farkas's lemma (see, e.g., [8, Lemma 2.1, p. 27]).

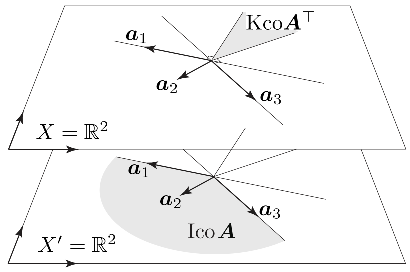


Fig. A.5: Farkas's lemma (when $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3)$).

Lemma A.3.2 (Farkas) For a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, the following relations hold:

$$\mathrm{Ico}\,\boldsymbol{A} = \left(\mathrm{Kco}\,\boldsymbol{A}^\top\right)', \quad (\mathrm{Ico}\,\boldsymbol{A})' = \mathrm{Kco}\,\boldsymbol{A}^\top.$$

□

For $n = 2$ and $m = 3$, a geometric interpretation of Farkas's lemma is shown in Fig. A.5. In this diagram,

$$\mathrm{Ico}\,\boldsymbol{A} = \left\{ \begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{pmatrix} \boldsymbol{z} \in \mathbb{R}^2 \;\middle|\; \boldsymbol{z} \geq \boldsymbol{0}_{\mathbb{R}^3} \right\}$$

represents the vector domain (cone) in which all of $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ are in the positive direction. On the other hand,

$$\mathrm{Kco}\,\boldsymbol{A}^\top = \left\{ \boldsymbol{y} \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix}\boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \boldsymbol{a}_3^\top\end{pmatrix} \boldsymbol{y} \leq \boldsymbol{0}_{\mathbb{R}^3} \right\}$$

represents a vector domain (cone) in which all of the inner vectors with $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ are non-positive. These mutually have the relationship of cone and dual cone.

A.4 Implicit Function Theorem

In the proof shown in Chaps. 2 and 7 of the Lagrange multiplier method (adjoint variable method), a key result known as the implicit function theorem was applied. Here, let us recall a version of this important theorem for implicit functions on finite-dimensional vector spaces (see, e.g., [10, Section 1.37, p. 30]).

Theorem A.4.1 (Implicit function theorem) Let m, n and k be natural numbers and suppose that $\boldsymbol{h} : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies the following conditions in a neighborhood $B_{\mathbb{R}^m} \times B_{\mathbb{R}^n}$ of $(\boldsymbol{x}_0, \boldsymbol{y}_0) \in \mathbb{R}^m \times \mathbb{R}^n$:

(1) $\boldsymbol{h}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}_{\mathbb{R}^n}$,

(2) $\boldsymbol{h} \in C^k(B_{\mathbb{R}^m} \times B_{\mathbb{R}^n}; \mathbb{R}^n)$,

(3) for every $\boldsymbol{x} \in B_{\mathbb{R}^m}$, $\boldsymbol{h}(\boldsymbol{x}, \cdot) \in C^1(B_{\mathbb{R}^n}; \mathbb{R}^n)$, and $\boldsymbol{h}_{\boldsymbol{y}^\top}(\boldsymbol{x}_0, \boldsymbol{y}_0) = (\partial h_i / \partial y_j (\boldsymbol{x}_0, \boldsymbol{y}_0))_{ij} \in \mathbb{R}^{n \times n}$ is regular.

In this case, there exists a neighborhood $U_{\mathbb{R}^m} \times U_{\mathbb{R}^n} \subset B_{\mathbb{R}^m} \times B_{\mathbb{R}^n}$ of $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ and a function $\boldsymbol{v} \in C^k(U_{\mathbb{R}^m}; U_{\mathbb{R}^n})$ such that $\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y}) = 0$ is equivalent to

$$\boldsymbol{y} = \boldsymbol{v}(\boldsymbol{x})$$

for all $(\boldsymbol{x}, \boldsymbol{y}) \in U_{\mathbb{R}^m} \times U_{\mathbb{R}^n}$. □

The next result is another version of implicit function theorem but for functions on Banach spaces (see, e.g., [1, Section 3.1.10, p. 115]).

Theorem A.4.2 (Implicit function theorem on Banach space) Let X, Y and Z be real Banach spaces. Suppose $h : X \times Y \to Z$ satisfies the following conditions in a neighborhood $B_X \times B_Y$ of $(\boldsymbol{x}_0, \boldsymbol{y}_0) \in X \times Y$:

(1) $h(\boldsymbol{x}_0, \boldsymbol{y}_0) = \mathbf{0}_Z$,

(2) $h \in C^k(B_X \times B_Y; Z)$,

(3) with respect to an arbitrary $\boldsymbol{x} \in B_X$, $h(\boldsymbol{x}, \cdot) \in C^1(B_Y; Z)$, and $(h_{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}))^{-1} : Z \to Y$ is bounded and linear.

In this case, there exists a neighborhood $U_X \times U_Y \subset B_X \times B_Y$ of $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ and $\boldsymbol{v} \in C^k(U_X; U_Y)$ such that $\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y}) = 0$ is equivalent to

$$\boldsymbol{y} = \boldsymbol{v}(\boldsymbol{x})$$

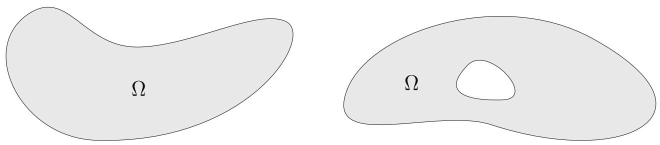
for all $(\boldsymbol{x}, \boldsymbol{y}) \in U_X \times U_Y$. □

A.5 Lipschitz Domain

After Chap. 5, boundary value problems of partial differential equations are taken up. In defining these problems, a definition regarding the smoothness with respect to boundary of domain is used. Let us summarize these notions of regularities of domains in this section (see, e.g., [4, Definition 1.2.1.1, p. 5], [7, Definition 6.28, p. 146] for references).

The characteristic that open sets cannot be split into two open sets (joined to be one) is called connected. For a natural number d, a connected open subset Ω in $\mathbb{R}^d$ is called a domain. In particular, if an arbitrarily chosen closed curve in the domain can be continuously constricted to a point in the domain, as in Fig. A.6 (a), then the domain is called simply connected. Otherwise, the domain is called multiply connected. Figure A.6 (b) shows an example of a doubly connected domain.

On a given domain, in order for a boundary value problem to be defined, it is not sufficient that the boundary is continuous. In fact, the additional assumption that the domain is Lipschitz is also required.



(a) Simply connected domain (b) Doubly connected domain

Fig. A.6: Domains.

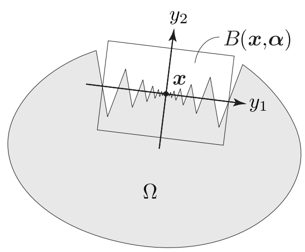


Fig. A.7: Lipschitz domain.

Definition A.5.1 (Lipschitz domain) Let Ω be a $d \in \{2, 3, \ldots\}$-dimensional bounded domain with boundary denoted by $\partial\Omega$. $\partial\Omega$ is called a Lipschitz boundary if, for all $\boldsymbol{x} \in \partial\Omega$, there exists a $\boldsymbol{\alpha} = (\alpha_i)_i \in \mathbb{R}^d$ ($\boldsymbol{\alpha} > \mathbf{0}_{\mathbb{R}^d}$) and a function (graph) φ belonging to $C^{0,1}\left(\mathbb{R}^{d-1}; \mathbb{R}\right)$ (Definition 4.3.1) such that using a new coordinate system $\boldsymbol{y} = (y_1, \cdots, y_d)^\top = \left(\boldsymbol{y}'^\top, y_d\right)^\top \in \mathbb{R}^d$ in the neighborhood

$$B(\boldsymbol{x}, \boldsymbol{\alpha}) = \left\{ \boldsymbol{x} + (y_1, \cdots, y_d)^\top \mid -\alpha_i < y_i < \alpha_i,\ i \in \{1, \ldots, d\} \right\}$$

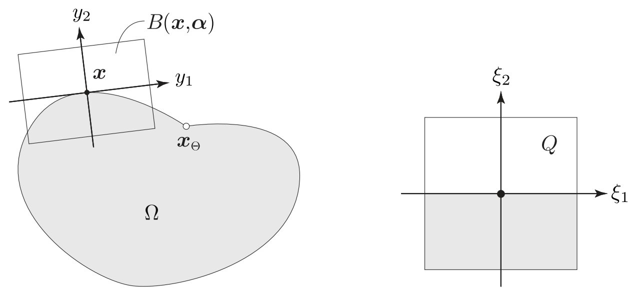
of $\boldsymbol{x}$, the condition

$$\Omega \cap B(\boldsymbol{x}, \boldsymbol{\alpha}) = \left\{ \boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \left(\boldsymbol{y}'^\top, y_d\right)^\top \in B(\boldsymbol{x}, \boldsymbol{\alpha}) \mid y_d < \varphi(\boldsymbol{y}') \right\}$$

holds. Moreover, Ω is called a Lipschitz domain. □

An extreme case of a Lipschitz domain is shown in Fig. A.7. In this figure, if the gradient of the graph φ is bounded below by some value and converges to zero as it nears $\boldsymbol{x}$, the boundary is still Lipschitz even if it consists of regions having rapid oscillations. If the derivative of the graph is not continuous, the normal, as described later, can be defined almost everywhere on the graph except on points of discontinuity. However, for a curvature to be defined, an even smoother boundary is needed. In such situations, a mapping from the neighborhood $B(\boldsymbol{x}, \boldsymbol{\alpha})$ of a point $\boldsymbol{x}$ in Fig. A.8 (a) to a standard domain Q in Fig. A.8 (b) can be constructed.

Definition A.5.2 (C^k-class domain) Let Ω be $d \in \{2, 3, \ldots\}$-dimensional bounded domain with boundary $\partial\Omega$ and $B(\boldsymbol{x}, \boldsymbol{\alpha})$, as defined in Definition A.5.1,



(a) Neighborhood $B(\boldsymbol{x}, \boldsymbol{\alpha})$ of $\boldsymbol{x} \in \partial\Omega$ (b) Standard domain Q

Fig. A.8: Piecewise C^1-class boundary.

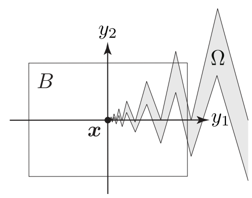


Fig. A.9: A case where a continuous boundary cannot be defined.

be a neighborhood of a point $\boldsymbol{x} \in \partial\Omega$. For a $k \in \mathbb{N}$, $\partial\Omega$ is called a C^k-class boundary if there exists a function $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_d)^\top : B(\boldsymbol{x}, \boldsymbol{\alpha}) \to Q = (0,1)^d$ such that

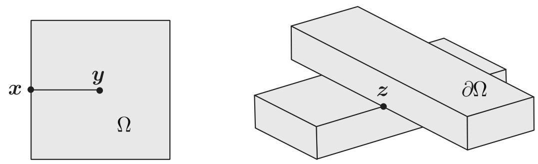
$$\Omega \cap B(\boldsymbol{x}, \boldsymbol{\alpha}) = \{\boldsymbol{x} + \boldsymbol{y} \in B(\boldsymbol{x}, \boldsymbol{\alpha}) \mid \phi_d(\boldsymbol{y}) < 0\}$$

holds for all $\boldsymbol{x} \in \partial\Omega$. The function $\boldsymbol{\phi}$ and its inverse map $\boldsymbol{\phi}^{-1}$ are uniformly bijective (one-to-one and onto mapping) and both belong to $C^k\left(B(\boldsymbol{x}, \boldsymbol{\alpha}); \mathbb{R}^d\right)$. Moreover, Ω is called a C^k-class domain. □

In Definition A.5.2, $k = 0$ is excluded. If a Lipschitz continuous function is chosen for $\boldsymbol{\phi}$ in Definition A.5.2, even a domain such as that in Fig. A.9 becomes permissible [4, Section 1.2, p. 4]. Hence, an important point to note about Definition A.5.1 is that it assumes the existence of a Lipschitz continuous graph φ. Two examples of simple domains with boundaries which do not satisfy the condition of being Lipschitz are shown in Fig. A.10; refer to the neighborhood of points $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ (see, e.g., [6, Figure 2, p. 91]).

Moreover, if for all $\boldsymbol{x} \in \Gamma$, where Γ is an open subset of $\partial\Omega$, the condition in Definition A.5.2 holds, then Γ is referred to as a C^k-class boundary. Furthermore, let the set of measure zero on the boundary where smoothness changes (see $\boldsymbol{x}_\Theta$ in Fig. A.8) be written as Θ. If the condition in Definition A.5.2 holds for all $\boldsymbol{x} \in \partial\Omega \setminus \Theta$ (points on $\partial\Omega$ excluding Θ), then $\partial\Omega$ is called a piecewise C^k-class boundary.

These definitions and notations are used to define the tangent, normal, and curvature of a point $\boldsymbol{x}$ on $\partial\Omega$ as follows.



(a) Domain with a crack (b) Connected hexahedra

Fig. A.10: Cases where there is not a Lipschitz boundary.

Definition A.5.3 (Tangent) Let $\partial\Omega$ be a Lipschitz boundary. For every $\boldsymbol{x} \in \partial\Omega \setminus \Theta$, let the neighborhood $B(\boldsymbol{x}, \boldsymbol{\alpha})$ and the function $\varphi \in C^{0,1}(\partial\Omega \cap B(\boldsymbol{x}, \boldsymbol{\alpha}); \mathbb{R})$ be as defined in Definition A.5.1. In this case, the vector

$$\begin{aligned}\begin{pmatrix}\boldsymbol{\tau}_1 & \cdots & \boldsymbol{\tau}_{d-1}\end{pmatrix}^\top &= \begin{pmatrix}\dfrac{\partial \varphi}{\partial y_1}(\boldsymbol{x})\,\boldsymbol{e}_1 & \cdots & \dfrac{\partial \varphi}{\partial y_{d-1}}(\boldsymbol{x})\,\boldsymbol{e}_{d-1}\end{pmatrix}^\top \\ &\in \left(L^\infty\left(\partial\Omega \cap B(\boldsymbol{x}, \boldsymbol{\alpha}); \mathbb{R}^d\right)\right)^{d-1}\end{aligned}$$

is the tangent at $\boldsymbol{x}$. Here, $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{d-1}$ represent the unit vectors of the coordinate system $(y_1, \cdots, y_{d-1})$ defined by Definition A.5.1. □

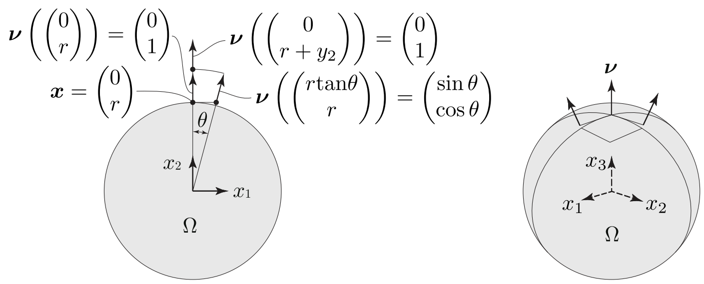
Definition A.5.4 (Normal) Let $\partial\Omega$ be a Lipschitz boundary. Suppose $\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_{d-1}$ are tangents (Definition A.5.3) at $\boldsymbol{x} \in \partial\Omega \setminus \Theta$. The unit vector $\boldsymbol{\nu}$ that is orthogonal to $\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_{d-1}$ and in the direction going from an inner point of Ω to the boundary is called the outer unit normal or normal at $\boldsymbol{x}$. □

Definition A.5.5 (Mean curvature) Suppose $\partial\Omega$ is a piecewise $C^{1,1}$-class ($C^{1,1}$-class on $\partial\Omega \setminus \Theta$) boundary. Given a point $\boldsymbol{x} \in \partial\Omega \setminus \Theta$, let $B(\boldsymbol{x}, \boldsymbol{\alpha})$ and the function $\boldsymbol{\phi} \in C^{1,1}\left(B(\boldsymbol{x}, \boldsymbol{\alpha}); \mathbb{R}^d\right)$ be defined as in Definition A.5.2 and let $\boldsymbol{\nu} = \boldsymbol{\nabla}\phi_d(\boldsymbol{x}) / \|\boldsymbol{\nabla}\phi_d(\boldsymbol{x})\|_{\mathbb{R}^d}$ and $\kappa = \boldsymbol{\nabla} \cdot \boldsymbol{\nu}$. Here, $\kappa/(d-1)$ is called the mean curvature of $\partial\Omega$. □

In what follows, we apply the definition of κ to a circle and a sphere and look at its relationship to the radius of curvature. Let us think about a circular domain with radius r centered at the origin such as that shown in Fig. A.11 (a). Let us calculate κ at a point $\boldsymbol{x} = (0, r)^\top$ on the boundary. In Fig. A.11 (a), the normal at the point moving from $\boldsymbol{x}$ in the x_1 direction by $\mathrm{d}x_1 = r\tan\theta$ is given by $(\sin\theta, \cos\theta)^\top$ and the normal at a point moving from $\boldsymbol{x}$ in the x_2 direction by $\mathrm{d}x_2 = y_2$ is $(0, 1)^\top$. Here,

$$\kappa(\boldsymbol{x}) = \boldsymbol{\nabla} \cdot \boldsymbol{\nu} = \frac{\partial \nu_1}{\partial x_1} + \frac{\partial \nu_2}{\partial x_2} = \lim_{\theta \to 0} \frac{\sin\theta}{r\tan\theta} = \frac{1}{r}$$

is obtained, where r is called the radius of curvature. Moreover, at the point $\boldsymbol{x} = (0, 0, r)^\top$ on a sphere with a radius r centered at the origin such as that shown in Fig. A.11 (b), the mean curvature κ becomes twice $(d-1=2)$ the



(a) $\Omega \subset \mathbb{R}^2$ (b) $\Omega \subset \mathbb{R}^3$

Fig. A.11: Mean curvature of circle and sphere.

value of the inverse of the radius of curvature; that is,

$$\kappa(\boldsymbol{x}) = \boldsymbol{\nabla} \cdot \boldsymbol{\nu} = \frac{\partial \nu_1}{\partial x_1} + \frac{\partial \nu_2}{\partial x_2} + \frac{\partial \nu_3}{\partial x_3} = \frac{2}{r}.$$

It should be noted that, besides the notion of mean curvature, there is also the definition for the so-called total curvature or Gauss curvature. In order to avoid confusion, let us point out the main differences between the mean and the total curvature. A smooth curved surface in a $d = 3$-dimensional space is generally a curved surface which has a spherical surface, as depicted in Fig. A.11 (b), replaced by an elliptical surface. At a point $\boldsymbol{x}$ on the surface of such an ellipsoid, the curvatures (whose sign is positive if it is concave or negative otherwise) at $\boldsymbol{x}$ of curves created on all the planes that contain the normal crossing the elliptical surface are called normal curvatures. The set of the maximum value κ_1 and minimum value κ_2 of the normal curvatures is called principal curvature. Here, the total curvature is defined by the product $\kappa_1 \kappa_2$ while the mean curvature is defined by $(\kappa_1 + \kappa_2)/2$.

A.6 Heat Conduction Problem

From Chap. 5, a Poisson problem is used as a prototype problem of a boundary value problem of a partial differential equation. A Poisson problem is used as a mathematical model representing various phenomena occurring in static and balanced situations. Here, let us look at a heat conduction phenomenon as an example in order to see how a Poisson problem represents such a state of a steady heat balance. Firstly, let us consider a time-dependent heat conduction problem of a one-dimensional continuous body. Afterwards, we extend it to a time-dependent heat conduction problem of a $d \in \{2, 3\}$-dimensional continuous body.

A.6.1 One-Dimensional Problem

Let us consider a one-dimensional continuous body such as in Fig. A.12. Let $(0, t_{\top})$ be a time domain and $(0, l)$ be the spatial domain for a one-dimensional

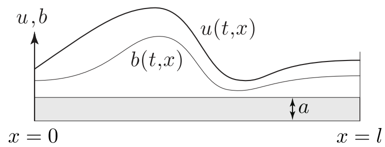


Fig. A.12: One-dimensional heat conduction problem.

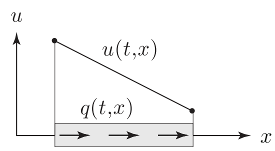


Fig. A.13: Fourier's law.

continuous body. Let a be a positive real constant representing the cross sectional area. Let $b : (0, t_{\mathrm{T}}) \times (0, l) \to \mathbb{R}$ be the heat released internally per unit volume and unit time, and $u : (0, t_{\mathrm{T}}) \times (0, l) \to \mathbb{R}$ be the temperature distribution. Now, let us look at the steps on how to derive the equation of heat conduction for seeking u with respect to b.

Firstly, let us assume that heat and temperature are related in a certain way.

Definition A.6.1 (Constitutive equation of heat and temperature) Let $u(t, x)$ be the temperature distribution at $(t, x) \in (0, t_{\mathrm{T}}) \times (0, l)$. Then, the amount of heat per unit volume of an object to which heat conducts to is given by

$$w(t, x) = c_{\mathrm{V}}(x)\, u(t, x), \tag{A.6.1}$$

where $c_{\mathrm{V}} : (0, l) \to \mathbb{R}$ is a positive-valued function representing the volume heat capacity. □

Next, we assume that the transfer of heat follows Fourier's law of heat conduction which is given as follows (see Fig. A.13).

Definition A.6.2 (Fourier's law of heat conduction) Let $u(t, x)$ be the temperature distribution at $(t, x) \in (0, t_{\mathrm{T}}) \times (0, l)$. The amount of heat (heat flux) passing through the cross-sectional region at x per unit time and area is given by

$$q(t, x) = -\lambda(x) \frac{\partial u}{\partial x}(t, x), \tag{A.6.2}$$

where $\lambda : (0, l) \to \mathbb{R}$ is a positive-valued function representing the heat conduction rate. □

Here, the change in the amount of heat in an infinitesimal volume and time $a\mathrm{d}x\mathrm{d}t$ (Fig. A.14) with respect to an arbitrary $(t, x) \in (0, t_{\mathrm{T}}) \times (0, l)$ becomes

$$(w(t + \mathrm{d}t, x) - w(t, x))\, a\mathrm{d}x = (b(t, x)\, \mathrm{d}x - q(t, x + \mathrm{d}x) + q(t, x))\, a\mathrm{d}t.$$

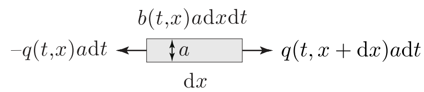


Fig. A.14: Heat balance.

In this case, if the limit of $\mathrm{d}x \to 0$, $\mathrm{d}t \to 0$ is taken,

$$\frac{\partial w}{\partial t}(t,x) = b(t,x) - \frac{\partial q}{\partial x}(t,x)$$

holds. Furthermore, if Eq. (A.6.1) and Eq. (A.6.2) are used,

$$c_{\mathrm{V}}\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\lambda\frac{\partial u}{\partial x}\right) = b$$

is obtained. This equation is called an equation of heat conduction.

The heat conduction equation is a second-order differential equation with respect to space and a first-order differential equation with respect to time. In order to uniquely determine u, two boundary conditions and one initial condition are required. For example, the following conditions could be considered:

(1) Let $u_{\mathrm{D}} : (0, t_{\top}) \to \mathbb{R}$ be a known temperature and suppose

$$u(t,0) = u_{\mathrm{D}}(t)$$

is satisfied at $x = 0$. This sort of condition specifying u is called a fundamental boundary condition or first-type boundary condition, and is commonly known in the literature as a Dirichlet condition.

(2) Suppose $p_{\mathrm{N}} : (0, t_{\top}) \to \mathbb{R}$ is a known heat flux (Definition A.6.2) and that

$$\lambda\frac{\partial u}{\partial x}(t,l) = p_{\mathrm{N}}(t)$$

is satisfied at $x = l$. A condition which specifies such a derivative of u is called a natural boundary condition or second-type boundary condition, and is commonly known in the literature as a Neumann condition.

(3) Suppose $u_0 : (0, l) \to \mathbb{R}$ is a known temperature such that

$$u(0,x) = u_0(x)$$

is satisfied at $t = 0$. This sort of condition specifying u at a certain time is called an initial condition.

Initial conditions can be viewed as boundary conditions of the time domain. Hence, a problem seeking u as the solution of a partial differential equation with specified boundary conditions and initial conditions is called a boundary value problem of a partial differential equation. Heat conduction is classified

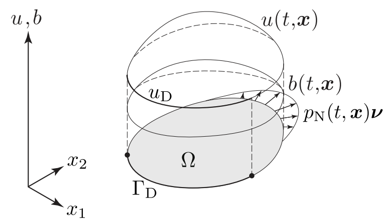


Fig. A.15: Two-dimensional heat conduction problem.

as a linear second-order partial differential equation. Specifically, the heat conduction equation can be classified as a parabolic partial differential equation (Sect. A.7). In a steady-state case, $b(t, x) = b(x)$ and $u(t, x) = u(x)$ and we get

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda \frac{\mathrm{d}u}{\mathrm{d}x}\right) = b. \tag{A.6.3}$$

Equation (A.6.3) is called a steady-state heat conduction equation. If this equation is extended to a d-dimensional domain, it becomes a partial differential equation such as shown later in Eq. (A.6.5). This resulting equation is classified as an elliptic partial differential equation (Sect. A.7).

A.6.2 d-Dimensional Problem

Next, let us think about a heat conduction phenomenon in a $d \in \{2, 3\}$-dimensional object. Figure A.15 shows a two-dimensional case. Let Ω be a Lipschitz domain on $\mathbb{R}^d$, Γ_{D} be a subset of its boundary $\partial\Omega$, and $\Gamma_{\mathrm{N}} = \partial\Omega \setminus \bar{\Gamma}_{\mathrm{D}}$. Also, let $b : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}$ be the amount of heat emitted internally per unit time and per unit volume, and $u : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}$ denote the temperature distribution on the domain at a given time. In this case, Fourier's law of heat conduction can be stated as follows.

Definition A.6.3 (Fourier's law of heat conduction in $\mathbb{R}^d$)
Let $u : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}$ denote the temperature distribution on the domain at a given time. The amount of heat conducted in an object per unit time and unit area (heat flux) $\boldsymbol{q} : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}^d$ satisfies

$$\boldsymbol{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_d \end{pmatrix} = -\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1d} \\ \vdots & \ddots & \vdots \\ \lambda_{d1} & \cdots & \lambda_{dd} \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_d} \end{pmatrix} u = -\boldsymbol{\Lambda}\boldsymbol{\nabla} u.$$

Here, $\boldsymbol{\Lambda} = (\lambda_{ij})_{ij} : \Omega \to \mathbb{R}^{d \times d}$ is a function taking the values of a positive definite real symmetric matrix (Definition 2.4.5) which represents the heat conduction

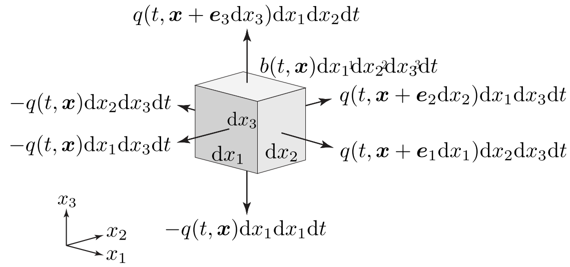


Fig. A.16: Three-dimensional heat balance.

rate. If the heat conduction rate is isotropic, a function taking positive real numbers $\lambda : \Omega \to \mathbb{R}$ can be used to write $\boldsymbol{\Lambda} = \lambda \boldsymbol{I}$ ($\boldsymbol{I}$ is a unit matrix). In this case,

$$\boldsymbol{q} = -\lambda \boldsymbol{\nabla} u. \tag{A.6.4}$$

□

With respect to an arbitrary $(t, \boldsymbol{x}) \in (0, t_{\top}) \times \Omega$, if the change in the amount of heat in an infinitesimal $\mathrm{d}x_1 \cdots \mathrm{d}x_d \mathrm{d}t$ is considered, we get

$$\begin{aligned}
&(w(t + \mathrm{d}t, \boldsymbol{x}) - w(t, \boldsymbol{x}))\, \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_d \\
&\quad = \left\{ b(t, \boldsymbol{x}) - \sum_{i \in \{1, \dots, d\}} (q_i(t, \boldsymbol{x} + \boldsymbol{e}_i \mathrm{d}x_i) - q_i(t, \boldsymbol{x})) \right\} \mathrm{d}t
\end{aligned}$$

where $\boldsymbol{e}_i$ denotes the unit vector in the x_i-axis direction (Fig. A.16). Here, if we take the limit of $\mathrm{d}x_1, \cdots, \mathrm{d}x_d \to 0$ and $\mathrm{d}t \to 0$, the following equation holds:

$$\frac{\partial w}{\partial t} = b - \sum_{i \in \{1, \dots, d\}} \frac{\partial q_i}{\partial x_i}.$$

If Eq. (A.6.1) and Eq. (A.6.4) are used,

$$\begin{aligned}
&c_{\mathrm{V}} \frac{\partial u}{\partial t} - \begin{pmatrix} \dfrac{\partial}{\partial x_1} & \cdots & \dfrac{\partial}{\partial x_d} \end{pmatrix} \left(\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1d} \\ \vdots & \ddots & \vdots \\ \lambda_{d1} & \cdots & \lambda_{dd} \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_3} \end{pmatrix} \right) u \\
&\quad = c_{\mathrm{V}} \frac{\partial u}{\partial t} - \boldsymbol{\nabla} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) = b
\end{aligned}$$

is obtained. This equation is called a d-dimensional heat conduction equation. If the heat conduction rate is isotropic, we get

$$c_{\mathrm{V}} \frac{\partial u}{\partial t} - \boldsymbol{\nabla} \cdot (\lambda \boldsymbol{\nabla} u) = b.$$

Furthermore, if λ is a real constant, we have

$$c_{\mathrm{V}} \frac{\partial u}{\partial t} - \lambda \Delta u = b.$$

Here, $\Delta = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$ is called the Laplace operator or harmonic operator. It can be written as ∇^2, but in this book Δ will be used.

In order to uniquely determine the solution u which satisfies the d-dimensional heat conduction equation, certain boundary conditions have to be imposed such as the following:

(1) Suppose $u_{\mathrm{D}} : (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{D}} \to \mathbb{R}$ is a known temperature and that

$$u = u_{\mathrm{D}} \quad \text{on } \ (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{D}}$$

is satisfied (fundamental boundary condition).

(2) Suppose $p_{\mathrm{N}} : (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{N}} \to \mathbb{R}$ is a known heat flux and that

$$\boldsymbol{\nu} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) = p_{\mathrm{N}} \quad \text{on } \ (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{N}}$$

is satisfied (natural boundary condition). If the heat conduction rate is isotropic, we get

$$\lambda \partial_{\nu} u = p_{\mathrm{N}} \quad \text{on } \ (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{N}},$$

where $\partial_{\nu}(\,\cdot\,)$ represents $(\partial(\,\cdot\,)/\partial \boldsymbol{x}) \cdot \boldsymbol{\nu}$.

(3) Let $u_0 : \Omega \to \mathbb{R}$ be a known temperature and that

$$u\,(t_0, \boldsymbol{x}) = u_0\,(\boldsymbol{x}) \quad \text{in } \boldsymbol{x} \in \Omega$$

is satisfied with respect to $t_0 \in (0, t_{\mathrm{T}})$ (initial condition).

In a steady-state case, we have $b\,(t, \boldsymbol{x}) = b\,(\boldsymbol{x})$ and $u\,(t, \boldsymbol{x}) = u\,(\boldsymbol{x})$ and the heat conduction equation becomes

$$-\boldsymbol{\nabla} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) = b. \tag{A.6.5}$$

Moreover, when a natural boundary condition is assumed across the entirety of $\partial\Omega$, there remains a constant indefiniteness. In order to uniquely determine u, the condition $|\Gamma_{\mathrm{D}}| > 0$ is required (Exercise 5.2.6).

Summarizing the above, the heat conduction problem can be defined as follows. Let Ω be a $d \in \{2, 3\}$-dimensional Lipschitz domain. Moreover, $\Gamma_{\mathrm{D}} \subset \partial\Omega$ and $\Gamma_{\mathrm{N}} = \partial\Omega \setminus \bar{\Gamma}_{\mathrm{D}}$. Also, let $c_{\mathrm{V}} : \Omega \to \mathbb{R}$ be a positive-valued function and $\boldsymbol{\Lambda} : \Omega \to \mathbb{R}^{d \times d}$ be a function taking positive definite real symmetric matrix values.

Problem A.6.4 (Heat conduction problem) Let $b : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}$, $p_{\mathrm{N}} : (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{N}} \to \mathbb{R}$, $u_{\mathrm{D}} : (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{D}} \to \mathbb{R}$, $u_0 : \Omega \to \mathbb{R}$ be given. Find $u : (0, t_{\mathrm{T}}) \times \Omega \to \mathbb{R}$ which satisfies

$$\begin{aligned} c_{\mathrm{V}} \frac{\partial u}{\partial t} - \boldsymbol{\nabla} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) &= b \quad \text{in } (0, t_{\mathrm{T}}) \times \Omega, \\ \boldsymbol{\nu} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) &= p_{\mathrm{N}} \quad \text{on } (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{N}}, \\ u &= u_{\mathrm{D}} \quad \text{on } (0, t_{\mathrm{T}}) \times \Gamma_{\mathrm{D}}, \\ u &= u_0 \quad \text{in } \Omega \text{ at } t = 0. \end{aligned}$$

□

In a steady state case, the heat conduction problem can be stated as follows.

Problem A.6.5 (Steady-state heat conduction problem) Let $b : \Omega \to \mathbb{R}$, $p_{\mathrm{N}} : \Gamma_{\mathrm{N}} \to \mathbb{R}$, $u_{\mathrm{D}} : \Gamma_{\mathrm{D}} \to \mathbb{R}$ be given. Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\boldsymbol{\nabla} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) &= b \quad \text{in } \Omega, \\ \boldsymbol{\nu} \cdot (\boldsymbol{\Lambda} \boldsymbol{\nabla} u) &= p_{\mathrm{N}} \quad \text{on } \Gamma_{\mathrm{N}}, \\ u &= u_{\mathrm{D}} \quad \text{on } \Gamma_{\mathrm{D}}. \end{aligned}$$

□

In Problem A.6.5, if we let $\boldsymbol{\Lambda} = \boldsymbol{I}$, the problem becomes a Poisson problem.

A.7 Classification of Second-Order Partial Differential Equations

As seen in Sect. A.6, a heat conduction problem is classified as parabolic if the problem is time-dependent and is classified as elliptic when viewed as a steady-state problem. Here, let us summarize the ways of classifying a system of partial differential equations (PDE) based on the standard form of a (linear) second-order partial differential equation with constant coefficients.

Definition A.7.1 (Classification of linear second-order PDE) Suppose the partial differential operator $\partial/\partial x_i$ ($i \in \{1, \dots, d\}$) is expressed as ξ_i and that the characteristic equation of the term (the primary term) whose sum of orders is a maximum is $f(\xi_1, \xi_2, \cdots, \xi_d) = 0$. Here, it will be classified depending on the following condition:

(1) When the characteristic equation does not have a real solution other than $(\xi_1, \dots, \xi_d) = (0, \dots, 0)$, it is called an elliptic partial differential equation.

(2) When the characteristic equation always has two different real solutions of $(\xi_1, \dots, \xi_d) \neq (0, \dots, 0)$, it is called a hyperbolic partial differential equation.

(3) When the characteristic equation $f(\xi_1, \xi_2, \cdots, \xi_d) = 0$ can be written as $\xi_1 - f_1(\xi_2, \cdots, \xi_d) = 0$, $f_1(\xi_2, \cdots, \xi_d) = 0$ has no real solution other than $(\xi_2, \ldots, \xi_d) = (0, \ldots, 0)$, and it is called a parabolic partial differential equation.

□

The typical form of an elliptic partial differential equation is the Laplace equation:

$$\Delta u = \left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_d^2} \right) u = 0.$$

Apparently, in this case, we have

$$f(\xi_1, \cdots, \xi_d) = \xi_1^2 + \cdots + \xi_d^2 = 0$$

and there are no real solutions other than $(x_1, \ldots, x_d) = (0, \ldots, 0)$. Apart from Laplace equations, Poisson equation $\Delta u = b$ or Helmholz equation $\Delta u + \omega^2 u = 0$, etc., where b and ω are real numbers, are also classified as elliptic partial differential equations. These characteristics are such that they:

- are balanced-type,
- and require closed boundary conditions.

Here, "closed boundary conditions" means that at all of the points on the boundary of the domain in which a partial differential equation is defined, either a type 1 boundary condition (Dirichlet condition), type 2 boundary condition (Neumann condition) or type 3 boundary condition (Robin condition) is given. Examples include steady heat conduction (temperature), steady electric field (electric potential), steady linear elastic body (displacement), flow field of ideal gases (potential) and Stokes flow field (flow speed and pressure), etc.

On the other hand, the typical form of a hyperbolic partial differential equation is the wave motion equation:

$$\ddot{u} - c^2 \Delta u = \frac{\partial^2 u}{\partial t^2} - c^2 \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) u = 0,$$

where c is a positive real number and is called the wave's speed. Consequently, in this case, the equation

$$f(\xi_1, \cdots, \xi_d) = \xi_1^2 - c^2 \left(\xi_2^2 + \cdots + \xi_d^2 \right) = 0$$

always has two different real solutions of $(x_1, \ldots, x_d) \neq (0, \ldots, 0)$. The characteristics of a hyperbolic partial differential equation are that:

- it is time-dependent,
- and requires a closed boundary condition and two initial conditions.

Furthermore, the typical form of a parabolic partial differential equation is the diffusion equation:

$$\dot{u} - a\Delta u = \frac{\partial u}{\partial t} - a\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right) u = 0,$$

where a is a positive real number and is called the diffusion coefficient. Accordingly, in this case, we have

$$f\left(\xi_1, \cdots, \xi_d\right) = \xi_1 - a\left(\xi_2^2 + \cdots + \xi_d^2\right) = 0$$

The characteristics of a parabolic partial differential equation are that:

- it is time-dependent,
- and requires closed boundary conditions and one initial condition.

A.8 Divergence Theorems

Beyond Chap. 5, integral formulae based on divergence theorems are used frequently. Here, we recall Gauss's divergence theorem and the Gauss–Green theorem.

Theorem A.8.1 (Gauss's divergence theorem) Let Ω be a d-dimensional, $d \in \{2, 3, \ldots\}$, Lipschitz domain and $\boldsymbol{\nu}$ be an outward unit normal (Definition A.5.4) on boundary $\partial\Omega$. In this case, the identity

$$\int_\Omega \boldsymbol{\nabla} f \,\mathrm{d}x = \int_{\partial\Omega} f \boldsymbol{\nu} \,\mathrm{d}\gamma$$

holds for any function $f \in C^1(\Omega; \mathbb{R})$. □

Proof We only outline the proof of the theorem. For more details, we refer the readers to [6, Theorem 3.34, p. 97]. Suppose $\Omega \subset \mathbb{R}^2$ is convex. The integral of $\partial f / \partial x_1$ on Ω can change the result of integrating with a very small domain in a band form of height $\mathrm{d}x_2$ such as in Fig. A.17 into one, which is integrated over the interval x_2. Here, we have

$$\int_\Omega \frac{\partial f}{\partial x_1} \,\mathrm{d}x = \int_{x_{\mathrm{B}2}}^{x_{\mathrm{T}2}} \left(f_{\mathrm{R}} - f_{\mathrm{L}}\right) \mathrm{d}x_2 = \int_{\partial\Omega} f \nu_1 \,\mathrm{d}\gamma.$$

The same results can be obtained with respect to $\partial f / \partial x_2$. Moreover, if $\Omega \subset \mathbb{R}^2$ is not convex, it can be split into convex partial domains and similar results can be obtained in the split domains. The same results can be established even when Ω is extended to $d \in \{3, 4, \ldots\}$-dimensions. □

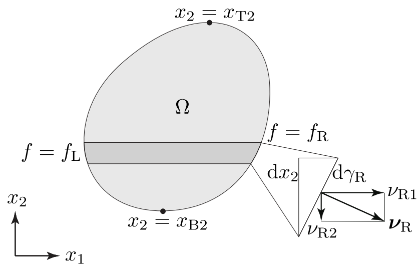


Fig. A.17: Gauss's divergence theorem.

Theorem A.8.2 (Gauss–Green theorem) Let Ω be a d-dimensional, $d \in \{2, 3, \ldots\}$, Lipschitz domain and $\boldsymbol{\nu}$ be an outward unit normal on the boundary $\partial\Omega$. Also, let $p, q \in (1, \infty)$ be such that $1/p + 1/q = 1$, and assume $f \in W^{1,p}(\Omega; \mathbb{R})$ and $g \in W^{1,q}(\Omega; \mathbb{R})$. In this case, the following identity holds:

$$\int_\Omega \boldsymbol{\nabla} f g \,\mathrm{d}x = \int_{\partial\Omega} f g \boldsymbol{\nu} \,\mathrm{d}\gamma - \int_\Omega f \boldsymbol{\nabla} g \,\mathrm{d}x.$$

□

Proof We only give the main working equation where the identity follows. For a more detailed proof, we refer the readers to [4, Theorem 1.5.3.1, p. 52]. From Leibnitz's law and Gauss's divergence theorem, the following equation holds:

$$\int_\Omega \boldsymbol{\nabla} f g \mathrm{d}x = \int_\Omega \boldsymbol{\nabla} (f g) \,\mathrm{d}x - \int_\Omega f \boldsymbol{\nabla} g \mathrm{d}x = \int_{\partial\Omega} f g \boldsymbol{\nu} \,\mathrm{d}\gamma - \int_\Omega f \boldsymbol{\nabla} g \mathrm{d}x.$$

□

A.9 Inequalities

Theorems with respect to several inequalities are used in the proof of theorems and propositions beyond Chap. 4. Here, we shall summarize the inequalities used this book.

We first state the well-known Hölder's inequality (see, e.g., [5, Theorem 2.8, p. 40]) which we have used in proving that the Lebesgue space $L^p(\Omega; \mathbb{R})$ (Definition 4.3.3) is a Banach space in Chap. 4, and also in investigating the regularity of solutions of the state determination problems in Chaps. 8 and 9.

Theorem A.9.1 (Hölder's inequality) Let Ω be a measurable set on $\mathbb{R}^d$, $d \in \mathbb{N}$, and $p, q \in (1, \infty)$ be such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then, for any pair of functions $f \in L^p(\Omega; \mathbb{R})$ and $g \in L^q(\Omega; \mathbb{R})$, the following inequality holds:

$$\|fg\|_{L^1(\Omega;\mathbb{R})} \leq \|f\|_{L^p(\Omega;\mathbb{R})} \|g\|_{L^q(\Omega;\mathbb{R})}.$$

□

In Theorem A.9.1, when $p = q = 2$, it is called Schwarz's inequality.

Furthermore, in showing that a Lebesgue space $L^p(\Omega;\mathbb{R})$ is a linear space in Chap. 4, we have used Minkowski's inequality which can be stated as follows (see, e.g., [5, Theorem 2.9, p. 41]).

Theorem A.9.2 (Minkowski's inequality) Let Ω be a measurable set on $\mathbb{R}^d$, $d \in \mathbb{N}$, and suppose $p \in [1,\infty)$. Then, for any pair of functions f, $g \in L^p(\Omega;\mathbb{R})$, the following inequality holds:

$$\|f + g\|_{L^p(\Omega;\mathbb{R})} \leq \|f\|_{L^p(\Omega;\mathbb{R})} + \|g\|_{L^p(\Omega;\mathbb{R})}.$$

□

Theorem A.9.2 is equivalent to triangle inequality at $L^p(\Omega;\mathbb{R})$.

Next, we state the widely known Poincaré's inequality (see, e.g, [4, Theorem 1.4.3.4, p. 26], [2, Theorem 6.1-2 (b), p. 276]) which is a key result used in Chap. 5 when showing the existence of a unique solution to the Poisson problem (see Exercise 5.2.5). The said inequality was also used in error analyses in Chaps. 8 and 9.

Theorem A.9.3 (Poincaré inequality) Let Ω be a measurable set on $\mathbb{R}^d$, $d \in \mathbb{N}$, and suppose $p \in [1,\infty)$. Then, for a function $f \in W^{1,p}(\Omega;\mathbb{R})$, there exists a positive constant c depending only on Ω and p, such that the inequality

$$\|f - f_0\|_{L^p(\Omega;\mathbb{R})} \leq c\,\|\boldsymbol{\nabla} f\|_{L^p(\Omega;\mathbb{R}^d)}, \quad f_0 = \frac{1}{|\Omega|}\int_\Omega f\,\mathrm{d}x$$

holds, where $|\Omega| = \displaystyle\int_\Omega \mathrm{d}x$. Moreover, when $f = 0$ on $\Gamma_{\mathrm{D}} \subset \partial\Omega$ such that $|\Gamma_{\mathrm{D}}| > 0$,

$$\|f\|_{L^p(\Omega;\mathbb{R})} \leq c\,\|\boldsymbol{\nabla} f\|_{L^p(\Omega;\mathbb{R}^d)}$$

holds. □

Corollary A.9.4 (Poincaré inequality) Let Ω be a measurable set on $\mathbb{R}^d$, $d \in \mathbb{N}$. Also, let $k \in \mathbb{N}$ and $p \in [1,\infty)$. If $f = 0$ on $\Gamma_{\mathrm{D}} \subset \partial\Omega$ such that $|\Gamma_{\mathrm{D}}| > 0$, then there exists a positive constant c which depends only on Ω and p such that the inequality

$$|f|_{W^{k,p}(\Omega;\mathbb{R})} \leq \|f\|_{W^{k,p}(\Omega;\mathbb{R})} \leq c\,|f|_{W^{k,p}(\Omega;\mathbb{R})}$$

holds, where $|\cdot|_{W^{k,p}(\Omega;\mathbb{R}^d)}$ represents

$$|f|_{W^{k,p}(\Omega;\mathbb{R})} = \begin{cases} \left(\displaystyle\sum_{|\boldsymbol{\beta}|=k} \left\|\nabla^{\boldsymbol{\beta}} f\right\|^p_{L^p(\Omega;\mathbb{R})} \right)^{1/p} & \text{for } p \in [0,\infty), \\ \displaystyle\max_{|\boldsymbol{\beta}|=k} \left\|\nabla^{\boldsymbol{\beta}} f\right\|_{L^\infty(\Omega;\mathbb{R})} & \text{for } p = \infty. \end{cases}$$

□

Furthermore, with respect to a function of $\Omega \to \mathbb{R}^d$, such as that in the case of a linear elastic problem, Korn's inequality (for example, [9, Lemma 5.4.21, p. 312]) and Korn's second inequality shown next are used (for example, [9, Lemma 5.4.18, p. 312], where $\Gamma_{\mathrm{D}} = \partial\Omega$ is assumed and $c = 2$ is obtained).

Theorem A.9.5 (Korn's inequality) Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. For a function $\boldsymbol{f} = (f_i)_i \in H^1\left(\Omega; \mathbb{R}^d\right)$, let

$$\boldsymbol{E}(\boldsymbol{f}) = (e_{ij}(\boldsymbol{f}))_{ij} = \frac{1}{2}\left(\frac{\partial f_i}{\partial x_j} + \frac{\partial f_j}{\partial x_i}\right)_{ij},$$

$$\|\boldsymbol{E}(\boldsymbol{f})\|^2_{L^2(\Omega;\mathbb{R}^{d\times d})} = \int_\Omega \sum_{(i,j)\in\{1,\dots,d\}^2} e_{ij}(\boldsymbol{f})\, e_{ij}(\boldsymbol{f})\, \mathrm{d}x.$$

Then, there exist positive constants c_1 and c_2 which are only dependent on Ω such that the inequality

$$\|\boldsymbol{E}(\boldsymbol{f})\|^2_{L^2(\Omega;\mathbb{R}^{d\times d})} \geq c_1 \|\boldsymbol{f}\|^2_{H^1(\Omega;\mathbb{R}^d)} - c_2 \|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)}$$

holds. □

Theorem A.9.6 (Korn's second inequality) In Theorem A.9.5, when $\boldsymbol{f} = \boldsymbol{0}_{\mathbb{R}^d}$ on $\Gamma_{\mathrm{D}} \subset \partial\Omega$ such that $|\Gamma_{\mathrm{D}}| > 0$, there exists a positive constant c depending only on Ω such that

$$\|\boldsymbol{f}\|^2_{H^1(\Omega;\mathbb{R}^d)} \leq c\, \|\boldsymbol{E}(\boldsymbol{f})\|^2_{L^2(\Omega;\mathbb{R}^{d\times d})}$$

holds. □

A.10 Ascoli–Arzelà Theorem

In this book, optimization problems in which continuous functions are chosen as design variables are considered in Chaps. 8 and 9. In these cases, admissible sets for design variables are required to be compact in the linear spaces where the design variables were defined. Regarding the compactness of set of continuous functions, the following theorem is known (cf. [11, Section III.3, p. 85]).

Theorem A.10.1 (Ascoli–Arzelà theorem) Let X be a complete metric space and V be a compact subset of X. If a sequence $\{f_n\}_{n\in\mathbb{N}}$ of continuous functions on V is uniformly bounded and equicontinuous on V, then there exists a subsequence of $\{f_n\}_{n\in\mathbb{N}}$ that converges uniformly. □

This theorem assures that a bounded subset of a continuous function's set $C^k(\Omega; \mathbb{R})$ defined on a bounded domain $\Omega \subset \mathbb{R}^d$ has a subsequence converging uniformly (compact) with $\|\cdot\|_{C^k(\Omega;\mathbb{R})}$. This result shows that when considering an optimization problem constructed with continuous functions as design variables, if a bounded sequence of functions included in a bounded subset of $C^k(\Omega; \mathbb{R})$ could be found by an optimization algorithm and is confirmed to be converged, then the convergent point is in the bounded subset.

References

[1] Berger, M. S. *Nonlinearity and Functional Analysis: Lectures on Nonlinear Problems in Mathematical Analysis.* Academic Press, New York, 1977.

[2] Ciarlet, P. G. *Three-Dimensional Elasticity.* North-Holland, Amsterdam; Tokyo, 1988.

[3] Gilbert, G. T. Positive definite matrices and Sylvester's criterion. *The American Mathematical Monthly*, 98(1):44–46, 1991. doi:10.2307/2324036.

[4] Grisvard, P. *Elliptic Problems in Nonsmooth Domains.* Pitman Advanced Pub. Program, Boston, 1985.

[5] Kurodai, S. *Functional Analysis (in Japanese).* Kyoritsu Shuppan, Tokyo, 1980.

[6] McLean, W. *Strongly Elliptic Systems and Boundary Iintegral Equations.* Cambridge University Press, Cambridge, 2000.

[7] Miyajima, S. *Basics and Applications of Sobolev Spaces (in Japanese).* Kyoritsu Shuppan, Tokyo, 2006.

[8] Tamura, A. and Matsumura, M. *Optimization Method (in Japanese).* Kyoritsu Shuppan, Tokyo, 2002.

[9] Trangenstein, J. A. *Numerical Solution of Elliptic and Parabolic Partial Differential Equations.* Cambridge University Press, Cambridge, 2013.

[10] Warner, F. *Foundations of Differentiable Manifolds and Lie Groups.* Springer, New York, 1983.

[11] Yosida, K. *Functional Analysis, 6th Ed.* Springer, Berlin; New York, 1980.