Springer Monographs in Mathematics

Kazuaki Taira

Semigroups, Boundary Value Problems and Markov Processes

Second Edition

Springer

S*pringer* **M***onographs* in **M***athematics*

For further volumes:
http://www.springer.com/series/3733

Kazuaki Taira

Semigroups, Boundary Value Problems and Markov Processes

Second Edition

Springer

Kazuaki Taira
Institute of Mathematics
University of Tsukuba
Tsukuba Ibaraki
Japan

ISSN 1439-7382 ISSN 2196-9922 (electronic)
Springer Monographs in Mathematics
ISBN 978-3-662-43695-0 ISBN 978-3-662-43696-7 (eBook)
DOI 10.1007/978-3-662-43696-7
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014946082

Mathematics Subject Classification: 35J25, 47D07, 47D05, 60J35, 60J60

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my mother
Yasue Taira (1918–)
and to the memory of my father
Yasunori Taira (1915–1990)

Preface to the Second Edition

In this monograph we study a general class of elliptic boundary value problems for second-order integro-differential operators in partial differential equations, and prove that this class of elliptic boundary value problems provides a general class of Feller semigroups in functional analysis. As an application, we construct a general class of Markov processes in probability in which a Markovian particle moves chaotically in the state space, incessantly changing its direction of motion until it "dies" at the time when it reaches the set where the particle is definitely absorbed.

In the early 1950s, W. Feller completely characterized the analytic structure of one-dimensional diffusion processes; he gave an intrinsic representation of the infinitesimal generator of a one-dimensional diffusion process and determined all possible boundary conditions which describe the domain of the infinitesimal generator. The probabilistic meaning of Feller's work was clarified by E.B. Dynkin, K. Itô, H.P. McKean, Jr., D. Ray and others. One-dimensional diffusion processes are fully understood, both from the analytic and the probabilistic viewpoints.

The main purpose of the present monograph is to generalize Feller's work to the multi-dimensional case. In 1959 A.D., Ventcel' (Wentzell) studied the problem of determining all possible boundary conditions for multi-dimensional diffusion processes. This monograph is devoted to a careful and accessible exposition of functional analytic methods for the problem of constructing Markov processes with Ventcel' boundary conditions in probability. Analytically, a Markovian particle in a domain of Euclidean space is governed by an integro-differential operator, called a Waldenfels integro-differential operator, in the interior of the domain, and it obeys a boundary condition, called a Ventcel' boundary condition, on the boundary of the domain. Probabilistically, a Markovian particle moves both by jumps and continuously in the state space and it obeys the Ventcel' boundary condition which consists of six terms corresponding to the diffusion along the boundary, the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon, the jump phenomenon on the boundary and the inward jump phenomenon from the boundary. In this monograph we prove existence theorems for Feller semigroups with Ventcel' boundary conditions for second-order elliptic Waldenfels integro-differential operators.

Many people found the first edition of my book *Semigroups, Boundary Value Problems and Markov Processes* (2004) useful. However, in the 10 years since the first edition appeared, there have been a number of much more comprehensive treatments of the material. The theory has reached a state of completion that makes it ideal for presentation in book form. This augmented second edition is amply illustrated. In order to keep the book up-to-date, additional references have been included in the bibliography. All chapters are rounded off with Notes and Comments where, primarily, bibliographical references are discussed. These Notes and Comments are intended to supplement the text and place it in a better perspective. The errors in the first edition have been corrected thanks to the kind input of many friends.

Our approach to the problem of constructing Markov processes with Ventcel' boundary conditions is distinguished by the extensive use of ideas and techniques characteristic of recent developments in the theory of partial differential equations. In particular, the theory of pseudo-differential operators – a modern theory of potentials – continues to be one of the most influential topics in the modern history of analysis, and is a very refined mathematical tool whose full power is yet to be exploited. Several recent developments in the theory of pseudo-differential operators have made possible further progress in the study of elliptic boundary value problems and hence in the study of Markov processes. The presentation of these new results is the main purpose of this monograph.

The idea behind our approach is as follows:

(1) We reduce the problem of existence of Feller semigroups to the unique solvability of boundary value problems for Waldenfels integro-differential operators with Ventcel' boundary conditions, and then prove existence theorems for Feller semigroups. To do this, we consider the Dirichlet problem for Waldenfels integro-differential operators, and prove an existence and uniqueness theorem in the framework of Hölder spaces. Then, by using the Green and harmonic operators for the Dirichlet problem, we can reduce the study of the boundary value problems to that of the Fredholm pseudo-differential equations on the boundary.

(2) The crucial point in the proof is how to calculate the complete symbol of pseudo-differential operators on the boundary. Furthermore, we make use of an existence and uniqueness theorem for a class of pseudo-differential operators on the boundary. The proof of this theorem is carried out by using a method of *elliptic regularizations* essentially due to Oleĭnik and Radkevič and developed for second-order differential operators with non-negative characteristic form.

(3) In this way, by using the Hölder space theory of pseudo-differential operators, we can prove that if the Ventcel' boundary conditions are *transversal* on the boundary, then we can verify all the conditions of the generation theorems of Feller semigroups.

(4) Moreover, we consider the general *non-transversal* case, and prove the unique solvability of boundary value problems for Waldenfels integro-differential operators with Ventcel' boundary conditions. In other words, we construct Feller

semigroups corresponding to the diffusion phenomenon where a Markovian particle moves both by jumps and continuously in the state space until it “dies” at the time when it reaches the set where the particle is definitely absorbed.

The present monograph is an expanded and revised version of a set of lecture notes for the graduate courses given at Sophia University, Hokkaido University, Tôhoku University, Tokyo Metropolitan University, Hiroshima University, Tokyo Institute of Technology and the University of Tsukuba. These lectures were addressed to the advanced undergraduates and beginning-graduate students with interest in functional analysis, partial differential equations and probability.

Unlike many other books on Markov processes, this monograph focuses on the relationship between four subjects in analysis: semigroups, pseudo-differential operators, elliptic boundary value problems and Markov processes. For graduate students, it may serve as an effective introduction to these four interrelated fields of analysis. For the graduate students about to major in the subject and for the mathematicians in the field looking for a coherent overview, it will provide a method for the analysis of elliptic boundary value problems in the framework of Sobolev and Besov spaces. Filling a mathematical gap between textbooks on Markov processes and recent developments in analysis, this monograph describes a powerful method capable of extensive further development. This monograph would be suitable as a textbook in a 1-year, advanced graduate course on functional analysis and partial differential equations, with emphasis on their strong interrelations with probability theory. In the introductory Chap. 1, I have attempted to state our problems and results in such a fashion that a broad spectrum of readers will be able to understand.

The second edition of *Boundary Value Problems and Markov Processes* (2009), which was published in the *Springer Lecture Notes in Mathematics* series, may be considered as a short introduction to the present more advanced monograph.

I began this work at the Ecole Normale Supérieure d’Ulm and Université de Paris-Sud (1976–1978) with the financial support of the French Government while I was on leave from the Tokyo Institute of Technology, and pursued it at the Institute for Advanced Study (1980–1981) with the financial support of the National Science Foundation. A major part of the work was done at the University of Tsukuba (1981–2009) and Hiroshima University (1995–1998) with the aid of Grant-in-Aid for General Scientific Research, Ministry of Education, Culture, Sports, Science and Technology, Japan. I take this opportunity to express my sincere gratitude to all these institutions.

In preparing this monograph, I am indebted to many people. I would like to express my hearty thanks to Professors Francesco Altomare, Wolfgang Arendt, Angelo Favini, Yasushi Ishikawa, Yuji Kasahara, Tamotu Kinoshita, Akihiko Miyachi, Silvia Romanelli, Elmar Schrohe, Seiichiro Wakabayashi and Atsushi Yagi. It is a pleasure to thank Schrohe and Ishikawa who have read and commented on portions of various preliminary drafts of this monograph. Especially, Schrohe helped me to learn the material that is presented in this monograph. Appendix B is based on his paper entitled “A short introduction to Boutet de Monvel’s calculus”.

I am deeply indebted to the late Professor Kiyosi Itô for his constant interest in my work.

I would like to thank the editorial staff of Springer-Verlag Heidelberg for their cooperation during the production of this augmented second edition.

Last but not least, I owe a great debt of gratitude to my family who gave me moral support during the preparation of this monograph.

Tsukuba, Japan
January 2014

Kazuaki Taira

Preface to the First Edition

The purpose of this book is to provide a careful and accessible account along modern lines of the subject of the title, and to discuss problems of current interest in the field. Unlike many other books on Markov processes, this book focuses on the relationship between Markov processes and elliptic boundary value problems, with an emphasis on the study of analytic semigroups. More precisely, this book is devoted to the functional analytic approach to a class of *degenerate* boundary value problems for second-order elliptic integro-differential operators, called Waldenfels operators, which includes as particular cases the Dirichlet and Robin problems. We prove that this class of boundary value problems provides a new example of analytic semigroups both in the L^p topology and in the topology of uniform convergence. As an application, we construct a strong Markov process corresponding to the physical phenomenon where a Markovian particle moves both by jumps and continuously in the state space until it "dies" at the time when it reaches the set where the particle is definitely absorbed.

The approach here is distinguished by the extensive use of techniques characteristic of recent developments in the theory of partial differential equations. The main technique used is the calculus of pseudo-differential operators which may be considered as a modern theory of potentials. Several recent developments in the theory of partial differential equations have made possible further progress in the study of boundary value problems and hence the study of Markov processes. The presentation of these new results is the main purpose of this book. We have confined ourselves to the simple but important boundary condition. This makes it possible to develop our basic machinery with a minimum of bother and the principal ideas can be presented concretely and explicitly.

This monograph is an expanded and revised version of a set of lecture notes for graduate courses given by the author at Hiroshima University in 1995–1997 and at the University of Tsukuba in 1998–2000. In 1988–1990 I gave a course in functional analysis and partial differential equations at the University of Tsukuba, the notes for which were published in the Springer Lecture Notes in Mathematics series under the title *Boundary Value Problems and Markov Processes* in 1991. These notes were found useful by a number of people, but they went out of print after a few years.

Moreover, in the 10 years since the lecture notes appeared, there have been a number of much more comprehensive treatments of the material.

Out of all this has emerged the present book. This new edition has been revised to streamline some of the analysis and to give better coverage of important examples and applications. It is addressed to advanced undergraduates or beginning-graduate students and also mathematicians with an interest in functional analysis, partial differential equations and probability. For the former, it may serve as an effective introduction to these three interrelated fields of mathematics. For the latter, it provides a method for the analysis of elliptic boundary value problems, a powerful method clearly capable of extensive further development. I have revised and updated the bibliography, but I have preferred to give references to expository books and articles rather than to research papers.

It is possible to regard the present book as the first volume of a three-volume series by the author, the others being *Analytic Semigroups and Semilinear Initial Boundary Value Problems* (Cambridge University Press, 1995) and *Brownian Motion and Index Formulas for the de Rham Complex* (Wiley-VCH, 1998). I hope that this monograph will lead to a better insight into the study of three interrelated subjects in analysis: semigroups, elliptic boundary value problems and Markov processes.

I would like to express my hearty thanks to the two referees whose comments and corrigenda concerning portions of various preliminary drafts of this book have resulted in a number of improvements. I am grateful to my colleagues and students at Tsukuba and Hiroshima who have provided the stimulating environment in which ideas germinate and flourish.

My special thanks go to the editorial staff of Springer-Verlag Tokyo and Heidelberg for their unfailing helpfulness and cooperation during the production of the book.

This research was partially supported by Grant-in-Aid for General Scientific Research (No. 13440041), Ministry of Education, Culture, Sports, Science and Technology, Japan.

Last, but not least, I owe a great debt of gratitude to my wife, Naomi, who gave me moral support during the preparation of this book.

Tsukuba, Japan Kazuaki Taira
April 2003

Contents

Part II Elements of Partial Differential Equations

Chapter 1
Introduction and Main Results

In this monograph we solve the problem of the existence of Feller semigroups associated with strong Markov processes. More precisely, we prove the unique solvability of boundary value problems for Waldenfels integro-differential operators with general Ventcel' (Wentzell) boundary conditions, and construct Feller semigroups corresponding to the diffusion phenomenon where a Markovian particle moves chaotically in the state space, incessantly changing its direction of motion until it "dies" at the time when it reaches the set where the particle is definitely absorbed. This monograph provides a careful and accessible exposition of the functional analytic approach to the problem of constructing strong Markov processes with Ventcel' boundary conditions in probability. Our approach here is distinguished by the extensive use of ideas and techniques characteristic of recent developments in the theory of partial differential equations.

The following diagram gives a bird's eye view of Markov processes, Feller semigroups and elliptic boundary value problems and how these relate to each other:

Probability	**Functional Analysis**	**Boundary Value Problems**
Strong Markov process $\mathcal{X} = (x_t)$	Feller semigroup $\{T_t\}$	Infinitesimal generator $\mathfrak{W}$
Markov transition function $p_t(\cdot, dy)$	$T_t f(\cdot) = \int p_t(\cdot, dy) f(y)$	$T_t = \exp[t\mathfrak{W}]$
Chapman–Kolmogorov equation	Semigroup property $T_{t+s} = T_t \cdot T_s$	Waldenfels integro-differential operator $W = P + S$
Various diffusion phenomena	Function spaces $C(\overline{D})$, $C_0(\overline{D} \setminus M)$	Ventcel' boundary condition L

In the early 1950s W. Feller [Fe1, Fe2] completely characterized the analytic structure of one-dimensional diffusion processes; he gave an intrinsic representation

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_1,

of the infinitesimal generator $\mathfrak{A}$ of a one-dimensional diffusion process and determined all possible boundary conditions which describe the domain $D(\mathfrak{A})$ of definition of the infinitesimal generator $\mathfrak{A}$. The probabilistic meaning of Feller's work was clarified by E.B. Dynkin [Dy1, Dy2], K. Itô and H.P. McKean, Jr. [IM], D. Ray [Ra] and others. One-dimensional diffusion processes are fully understood both from the analytic and the probabilistic points of view. The main purpose of this monograph is to generalize Feller's work to the multi-dimensional case.

Now we take a close look at Feller's work. Let $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ be a one-dimensional diffusion process with state space K. A point x of K is called a right (resp. left) singular point if $x_t(\omega) \geq x$ (resp. $x_t(\omega) \leq x$) for all $t \in [0, \zeta(\omega))$ with P_x-measure one, where the random variable ζ is the lifetime of the process $\mathcal{X}$. A right and left singular point is called a trap. A point which is neither right nor left singular is called a regular point.

For simplicity, we assume that the state space K is the half-line

$$K = [0, \infty),$$

and all its interior points are regular. Feller proved that there exist a strictly increasing, continuous function s on $(0, \infty)$ and Borel measures m and k on $(0, \infty)$ such that the infinitesimal generator $\mathfrak{A}$ of the process $\mathcal{X}$ can be expressed as follows:

$$\mathfrak{A} f(x) = \lim_{y \downarrow x} \frac{f^+(y) - f^+(x) - \int_{(x,y]} f(z)\, dk(z)}{m((x,y])}. \tag{1.1}$$

Here:

(1) $f^+(x) = \lim_{\varepsilon \downarrow 0} \frac{f(x+\varepsilon)-f(x)}{s(x+\varepsilon)-s(x)}$, the right-derivative of f at x with respect to s.
(2) The measure m is positive for non-empty open subsets, and is finite for compact sets.
(3) The measure k is finite for compact subsets.

The function s is called a canonical scale, and the measures m and k are called a canonical measure (or speed measure) and a killing measure for the process $\mathcal{X}$, respectively. They determine the behavior of a Markovian particle in the interior of the state space K.

We remark that the right-hand side of (1.1) is a generalization of the second-order differential operator

$$a(x) f'' + b(x) f' + c(x) f,$$

where $a(x) > 0$ and $c(x) \leq 0$ on K. For example, the formula

$$\mathfrak{A} f = a(x) f'' + b(x) f'$$

can be written in the form (1.1), if we take

$$s(x) = \int_0^x \exp\left[-\int_0^y \frac{b(z)}{a(z)}\,dz\right] dy,$$

$$dm(x) = \frac{1}{a(x)} \exp\left[\int_0^x \frac{b(y)}{a(y)}\,dy\right] dx,$$

$$dk(x) = 0.$$

The boundary point 0 is called a regular boundary if we have, for an arbitrary point $r \in (0, \infty)$,

$$\int_{(0,r)} [s(r) - s(x)][dm(x) + dk(x)] < \infty,$$

$$\int_{(0,r)} [m((x,r)) + k((x,r))]ds(x) < \infty.$$

Intuitively, the regularity of the boundary point means that a Markovian particle approaches the boundary in finite time with positive probability, and also enters the interior from the boundary.

The behavior of a Markovian particle at the boundary point is characterized by boundary conditions. In the case of regular boundary points, Feller determined all possible boundary conditions which are satisfied by the functions $f(x)$ in the domain $D(\mathfrak{A})$ of the infinitesimal generator $\mathfrak{A}$ of the process $\mathcal{X}$. A general boundary condition is of the form

$$\gamma f(0) - \delta \mathfrak{A} f(0) + \mu f^+(0) = 0,$$

where γ, δ and μ are constants such that $\gamma \leq 0$, $\delta \geq 0$, $\mu \geq 0$, $\mu + \delta > 0$. If we admit jumps from the boundary into the interior, then a general boundary condition takes the form

$$\gamma f(0) - \delta \mathfrak{A} f(0) + \mu f^+(0) + \int_{(0,\infty)} [f(x) - f(0)]\, d\nu(x) = 0, \tag{1.2}$$

where ν is a Borel measure with respect to which the function $\min(1, s(x) - s(+0))$ is integrable.

First, we introduce a class of (temporally homogeneous) Markov processes. A Markov process is said to be one-dimensional or multi-dimensional according as the state space is a subset of $\mathbf{R}$ or $\mathbf{R}^N$ $(N \geq 2)$. Intuitively, the Markov property is the principle of the lack of any "memory" in the system. Markov processes are an abstraction of the idea of Brownian motion. From the point of view of analysis, however, the transition function of a Markov process is something more convenient than the Markov process itself. Secondly, we give the precise definition of a Markov transition function adapted to the Hille–Yosida theory of semigroups. Moreover, we introduce a class of semigroups associated with Markov processes, called Feller semigroups, and prove generation theorems for Feller semigroups which form a functional analytic background for the construction of Markov processes. Following Ventcel' (Wentzell) [We], we can describe analytically the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}$ in the case where the state space is the closure $\overline{D}$ of a bounded domain D in Euclidean space $\mathbf{R}^N$. Analytically, a Markovian particle in $\overline{D}$ is governed by an integro-differential operator W, called a Waldenfels operator, in the interior D of the domain, and it obeys a boundary condition L, called a Ventcel' boundary condition, on the boundary ∂D of the domain. Probabilistically, a Markovian particle moves both by jumps and continuously in the state space and it obeys the Ventcel' boundary condition which consists of six terms corresponding to the diffusion along the boundary, the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary. Therefore, we are reduced to the study of the boundary value problem for Waldenfels integro-differential operators W with Ventcel' boundary conditions L in the theory of partial differential equations.

1.1 Formulation of the Problem

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$ with smooth boundary ∂D; its closure $\overline{D} = D \cup \partial D$ is an N-dimensional, compact smooth manifold with boundary (see Fig. 1.1).

Let $C(\overline{D})$ be the space of real-valued, continuous functions on $\overline{D}$. We equip the space $C(\overline{D})$ with the topology of uniform convergence on the whole $\overline{D}$; hence it is a Banach space with the maximum norm

$$\|f\|_\infty = \max_{x \in \overline{D}} |f(x)|.$$

A strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on the space $C(\overline{D})$ is called a *Feller semigroup* on $\overline{D}$ if it is non-negative and contractive on $C(\overline{D})$:

$$f \in C(\overline{D}),\ 0 \leq f \leq 1 \quad \text{on } \overline{D} \implies 0 \leq T_t f \leq 1 \quad \text{on } \overline{D}.$$

It is known (see [Dy1, Dy2, Ta2]) that if T_t is a Feller semigroup on $\overline{D}$, then there exists a unique Markov transition function p_t on $\overline{D}$ such that

$$T_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y) \quad \text{for all } f \in C(\overline{D}).$$

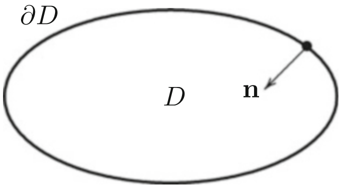


Fig. 1.1 The bounded domain D with smooth boundary ∂D

It can be shown that the function p_t is the transition function of some *strong Markov process* $\mathcal{X}$; hence the value $p_t(x, E)$ expresses the transition probability that a Markovian particle starting at position x will be found in the set E at time t.

On the other hand, it is known (see [BCP, SU, Ta2, Wa, We]) that the infinitesimal generator of a Feller semigroup $\{T_t\}_{t\geq 0}$ is described analytically by a Waldenfels integro-differential operator W (formula (1.3)) and a Ventcel' boundary condition L (formula (1.5)), which we shall formulate precisely (see Sects. 9.4 and 9.5).

Let W be a second-order *elliptic* integro-differential operator with real smooth coefficients such that

$$\begin{aligned} & W_D u(x) \\ = \; & Pu(x) + S_D u(x) \\ := \; & \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\ & + \sum_{j=1}^{N} a_\sigma^j(x) \frac{\partial u}{\partial x_j}(x) + a_\sigma(x)u(x) \\ & + \int_D s(x,y) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^{N} (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy. \end{aligned} \tag{1.3}$$

Here:

(1) $a^{ij} \in C^\infty(\mathbf{R}^N)$, $a^{ij}(x) = a^{ji}(x)$ for all $x \in \mathbf{R}^N$ and $1 \leq i, j \leq N$, and there exists a constant $a_0 > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i \xi_j \geq a_0 |\xi|^2 \quad \text{for all } (x, \xi) \in \mathbf{R}^N \times \mathbf{R}^N.$$

(2) $b^i \in C^\infty(\mathbf{R}^N)$ for all $1 \leq i \leq N$.
(3) $c \in C^\infty(\mathbf{R}^N)$ and $P1(x) = c(x) \leq 0$ in D.
(4) The integral kernel $s(x, y)$ is the distribution kernel of a properly supported, pseudo-differential operator $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$, $\kappa > 0$, which has the *transmission property* with respect to ∂D due to Boutet de Monvel [Bo] (see Sect. 7.6), and

$s(x, y) \geq 0$ off the *diagonal* $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$. The measure dy is the Lebesgue measure on $\mathbf{R}^N$.

(5) The function $\sigma(x, y)$ is a local unity function on $\overline{D}$, that is, $\sigma(x, y)$ is a smooth function on $\overline{D} \times \overline{D}$ such that $\sigma(x, y) = 1$ in a neighborhood of the *diagonal* $\Delta_{\overline{D}} = \{(x, x) : x \in \overline{D}\}$ in $\overline{D} \times \overline{D}$. The function $\sigma(x, y)$ depends on the shape of the domain D. More precisely, it depends on a family of local charts on D in each of which the Taylor expansion is valid for functions u. For example, if D is *convex*, we may take $\sigma(x, y) \equiv 1$ on $\overline{D} \times \overline{D}$.

(6) $a_\sigma^j(x) = (S_D \sigma_x^j)(x)$ where $\sigma_x^j(y) = \sigma(x, y)(y_j - x_j)$ for all $1 \leq j \leq N$.

(7) $a_\sigma(x) = (S_D \sigma_x)(x)$ where $\sigma_x(y) = \sigma(x, y)$.

Under these conditions, we show that the operator S_D can be *formally* written in the simple form

$$(S_D u)(x) = S(u^0)\big|_D = \int_D s(x, y) u(y)\, dy, \quad x \in D,$$

where u^0 is the extension of v to $\mathbf{R}^N$ by zero outside of $\overline{D}$

$$u^0(x) = \begin{cases} u(x) & \text{for } x \in \overline{D}, \\ 0 & \text{for } x \in \mathbf{R}^N \setminus \overline{D}. \end{cases}$$

In fact, we have

$$\begin{aligned} &S_D u(x) \\ &:= \sum_{j=1}^N a_\sigma^j(x) \frac{\partial u}{\partial x_j}(x) + a_\sigma(x) u(x) \\ &\quad + \int_D s(x, y) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy \\ &= \sum_{j=1}^N \int_D s(x, y)\, \sigma(x, y)(y_j - x_j)\, dy \cdot \frac{\partial u}{\partial x_j}(x) + \int_D s(x, y)\, \sigma(x, y)\, dy \cdot u(x) \\ &\quad + \int_D s(x, y) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy \\ &= \int_D s(x, y) u(y)\, dy, \quad x \in D. \end{aligned}$$

Due to the non-local character of the pseudo-differential operator S, it is natural to use the zero-extension of functions in the interior D outside of the closure $\overline{D} = D \cup \partial D$. This extension has a probabilistic interpretation. Namely,

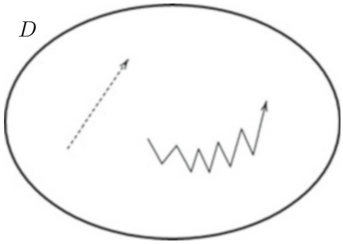


Fig. 1.2 A Markovian particle moves by jumps and continuously

it corresponds to stopping the diffusion process with jumps in the whole space $\mathbf{R}^N$ at the first exit time of the closure $\overline{D}$. However, the zero-extension produces a singularity of solutions at the boundary ∂D. The transmission property guarantees that if v is smooth up to the boundary ∂D, then so is $S_D v$.

The operator S_D can be visualized as follows:

$$S_D : C^\infty(\overline{D}) \longrightarrow \mathcal{E}'(\mathbf{R}^N) \overset{S}{\longrightarrow} \mathcal{E}'(\mathbf{R}^N) \longrightarrow C^\infty(D),$$

where the first arrow is the zero extension to $\mathbf{R}^N$ and the last one is the restriction to D.

(8) The operator W_D satisfies the condition

$$\begin{aligned} W_D 1(x) &= P1(x) + S_D 1(x) \\ &= c(x) + a_\sigma(x) + \int_D s(x,y)\,[1-\sigma(x,y)]\,dy \leq 0 \quad \text{in } D. \end{aligned} \tag{1.4}$$

The operator W is called a second-order *Waldenfels integro-differential operator* or simply a *Waldenfels operator* (cf. [Wa]). The differential operator P is called a diffusion operator which describes analytically a strong Markov process with continuous paths (diffusion process) in the interior D, and the functions $a^{ij}(x)$, $b^i(x)$ and $c(x)$ are called the diffusion coefficients, the drift coefficients and the termination coefficient, respectively. The integro-differential operator

$$S_r u = \int_D s(x,y)\left[u(y) - \sigma(x,y)\left(u(x) + \sum_{j=1}^N (y_j - x_j)\frac{\partial u}{\partial x_j}(x)\right)\right] dy$$

is called a second-order Lévy operator which corresponds to the jump phenomenon in the interior D (see [St]). That is, a Markovian particle moves by jumps to a random point, chosen with kernel $s(x,y)$, in the interior D. Therefore, the Waldenfels integro-differential operator $W = P + S_r$ corresponds to the diffusion phenomenon where a Markovian particle moves both by jumps and continuously in the state space D (see Fig. 1.2).

In order to remove a singularity in the interior D, we introduce condition (1.4) on the structure of jumps for the Waldenfels integro-differential operator W. The intuitive meaning of condition (1.4) is that the jump phenomenon from a point $x \in D$ to the outside of a neighborhood of x in the interior D is "dominated" by the absorption phenomenon at x. In particular, if $P1(x) = c(x) \equiv 0$ in D, then condition (1.4) implies that any Markovian particle does not move by jumps from $x \in D$ to the outside of a neighborhood $V(x)$ of x in the interior D. Indeed, we have

$$\int_D s(x,y)\,[1-\sigma(x,y)]\,dy = 0,$$

and so, by conditions (4) and (5),

$$s(x,y) = 0 \quad \text{for all } y \in D \setminus V(x).$$

Probabilistically, this is a condition on the support of the Lévy measure $s(\cdot, y)\,dy$ associated with the pseudo-differential operator S.

It should be emphasized that the integro-differential operator S_r is a "regularization" of S (see inequalities (10.2) and (10.3)).

Let L be a second-order boundary condition such that, in terms of local coordinates $(x_1, x_2, \ldots, x_{N-1})$ of ∂D,

$$\begin{aligned}
&Lu(x') \qquad (1.5)\\
&= Qu(x') + \mu(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x')W_D u(x') + \Gamma u(x')\\
&:= \left(\sum_{i,j=1}^{N-1} \alpha^{ij}(x')\frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x')\frac{\partial u}{\partial x_i}(x') + \gamma(x')u(x')\right)\\
&\quad + \mu(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x')W_D u(x')\\
&\quad + \sum_{j=1}^{N-1} \eta_\tau^j(x')\frac{\partial u}{\partial x_j}(x') + \eta_\tau(x')u(x')\\
&\quad + \int_{\partial D} r(x',y')\left[u(y') - \tau(x',y')\left(u(x') + \sum_{j=1}^{N-1}(y_j - x_j)\frac{\partial u}{\partial x_j}(x')\right)\right] dy'\\
&\quad + \int_D t(x',y)\left[u(y) - u(x')\right] dy.
\end{aligned}$$

Here:

(1) The operator Q is a second-order degenerate elliptic differential operator on ∂D with non-positive principal symbol. In other words, the α^{ij} are the

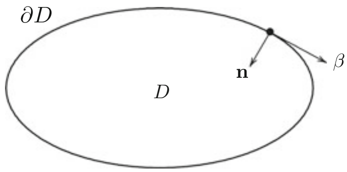


Fig. 1.3 The vector field β and the unit interior normal $\mathbf{n}$ to ∂D

components of a smooth symmetric contravariant tensor of type $\binom{2}{0}$ on ∂D satisfying the condition

$$\sum_{i,j=1}^{N-1} \alpha^{ij}(x')\xi_i\xi_j \geq 0 \quad \text{for all } x' \in \partial D \text{ and } \xi' = \sum_{j=1}^{N-1} \xi_j \, dx_j \in T^*_{x'}(\partial D),$$

where $T^*_{x'}(\partial D)$ is the cotangent space of ∂D at x'.

(2) $\beta(x') = \sum_{i=1}^{N-1} \beta^i(x')\partial/\partial x_i$ is a smooth vector field on ∂D (see Fig. 1.3).
(3) $Q1 = \gamma \in C^\infty(\partial D)$ and $\gamma(x') \leq 0$ on ∂D.
(4) $\mu \in C^\infty(\partial D)$ and $\mu(x') \geq 0$ on ∂D.
(5) $\delta \in C^\infty(\partial D)$ and $\delta(x') \geq 0$ on ∂D.
(6) $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ is the unit interior normal to the boundary ∂D (see Fig. 1.3).
(7) The integral kernel $r(x', y')$ is the distribution kernel of a pseudo-differential operator $R \in L^{2-\kappa_1}_{1,0}(\partial D)$, $\kappa_1 > 0$, and $r(x', y') \geq 0$ off the diagonal $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$ in $\partial D \times \partial D$. The density dy' is a strictly positive density on ∂D.
(8) The integral kernel $t(x, y)$ is the distribution kernel of a properly supported, pseudo-differential operator $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, $\kappa_2 > 0$, which has the transmission property with respect to the boundary ∂D, and $t(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$.
(9) The function $\tau(x, y)$ is a local unity function on $\overline{D}$; more precisely, $\tau(x, y)$ is a smooth function on $\overline{D} \times \overline{D}$, with compact support in a neighborhood of the diagonal $\Delta_{\partial D}$, such that, at each point x' of ∂D, $\tau(x', y) = 1$ for y in a neighborhood of x' in $\overline{D}$. The function $\tau(x, y)$ depends on the shape of the boundary ∂D.
(10) $\eta^j_\tau(x') = R(\tau^j_{x'})(x')$ where $\tau^j_{x'}(y') = \tau(x', y')(y_j - x_j)$ for all $1 \leq j \leq N - 1$.
(11) $\eta_\tau(x') = R(\tau_{x'})(x')$ where $\tau_{x'}(y') = \tau(x', y')$.

Under these conditions, we show that the boundary operator Γ can be *formally* written in the simple form

$$\Gamma u(x') = \int_{\partial D} r(x', y')u(y')\, dy' + \int_D t(x', y)\left[u(y) - u(x')\right] dy, \quad x' \in \partial D.$$

In fact, we have

$$
\begin{aligned}
&\sum_{j=1}^{N-1} \eta_\tau^j(x') \frac{\partial u}{\partial x_j}(x') + \eta_\tau(x') u(x') \\
&+ \int_{\partial D} r(x',y') \left[u(y') - \tau(x',y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\
&+ \int_D t(x',y) \left[u(y) - u(x') \right] dy \\
= &\sum_{j=1}^{N-1} \int_{\partial D} r(x',y')\, \tau(x',y')(y_j - x_j)\, dy' \cdot \frac{\partial u}{\partial x_j}(x) \\
&+ \int_{\partial D} r(x',y')\, \tau(x',y')\, dy' \cdot u(x') \\
&+ \int_{\partial D} r(x',y') \left[u(y') - \tau(x',y') \left(u(x') + \sum_{j=1}^{N=1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\
&+ \int_D t(x',y) \left[u(y) - u(x') \right] dy \\
= &\int_{\partial D} r(x',y') u(y')\, dy' + \int_D t(x',y) \left[u(y) - u(x') \right] dy, \quad x' \in \partial D.
\end{aligned}
$$

(12) The boundary operator Γ is of order $2 - \kappa_1$, and satisfies the condition

$$
\begin{aligned}
&Q1(x') + \Gamma 1(x') \\
&= \gamma(x') + \eta_\tau(x') + \int_{\partial D} r(x',y') \left[1 - \tau(x',y') \right] dy' \leq 0 \quad \text{on } \partial D.
\end{aligned}
\tag{1.6}
$$

The boundary condition L is called a second-order *Ventcel' boundary condition* (cf. [We]). It should be noted that L is a generalization of the boundary condition (1.2) to the multi-dimensional case. The six terms of L

$$
\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') d \frac{\partial u}{\partial x_i}(x'),
$$

$$
\gamma(x') u(x'), \quad \mu(x') \frac{\partial u}{\partial \mathbf{n}}(x'), \quad \delta(x') W_D u(x'),
$$

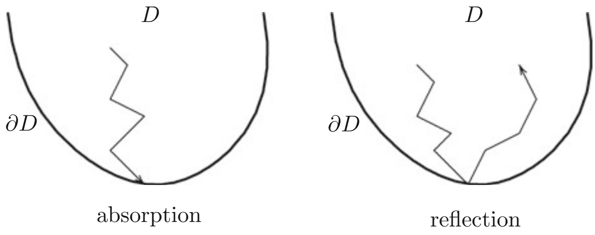


Fig. 1.4 The absorption phenomenon and the reflection phenomenon

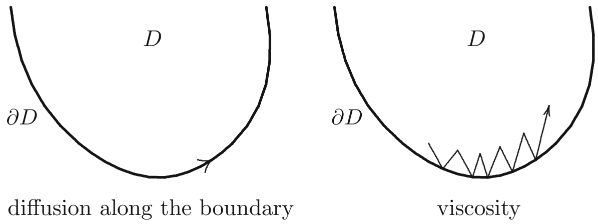


Fig. 1.5 The diffusion along $\partial\Omega$ and the viscosity phenomenon

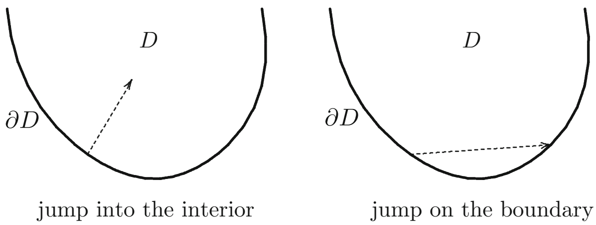


Fig. 1.6 The inward jump phenomenon from ∂ and the jump phenomenon on ∂D

$$\int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') - \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy',$$

$$\int_D t(x', y) \left[u(y) - u(x') \right] dy$$

correspond to the diffusion along the boundary, the absorption phenomenon, the reflection phenomenon, the viscosity phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary, respectively (see Figs. 1.4–1.6).

For the probabilistic meanings of Ventcel' boundary conditions, the reader might refer to Dynkin–Yushkevich [DY].

For the Ventcel' boundary condition L, we are forced to impose condition (1.6) in order to remove a singularity of solutions at the boundary ∂D. The intuitive meaning of condition (1.6) is that the jump phenomenon from a point $x' \in \partial D$ to the outside of a neighborhood of x' on the boundary ∂D is "dominated" by the absorption phenomenon at x'. In particular, if $\gamma(x') \equiv 0$ on ∂D, then condition (1.6) implies that any Markovian particle does not move by jumps from $x' \in \partial D$ to the outside of a neighborhood $V(x')$ of x' on the boundary ∂D. Indeed, we have

$$\int_{\partial D} r(x', y') \left[1 - \tau(x', y')\right] dy' = 0,$$

and so, by conditions (7) and (9),

$$r(x', y') = 0 \quad \text{for all } y' \in \partial D \setminus V(x').$$

Probabilistically, this is a condition on the support of the Lévy measure $r(\cdot, y')\, dy'$ associated with the pseudo-differential operator R.

It should be noted that the integro-differential operator

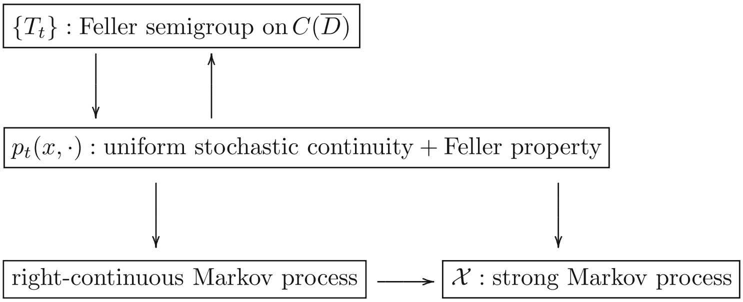
$$\begin{aligned} &\Gamma_r u(x') \\ &= \int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\ &\quad + \int_D t(x', y) \left[u(y) - u(x')\right] dy, \quad x' \in \partial D, \end{aligned}$$

is a "regularization" of $R \in L^{2-\kappa_1}_{1,0}(\partial D)$ and $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$ (see inequalities (10.5) and (10.6)).

This monograph is devoted to functional analytic methods for the problem of the existence of strong Markov processes in probability theory. More precisely, we consider the following problem:

Problem 1.1. Conversely, given analytic data (W, L), can we construct a Feller semigroup $\{T_t\}_{t \geq 0}$ whose infinitesimal generator is characterized by (W, L) ?

It is known that any right-continuous Markov process whose transition function has the Feller property is a strong Markov process. Our functional analytic method for the problem of the existence of strong Markov processes in probability theory may be visualized as follows:



1.2 Statement of Main Results

Now we say that the boundary condition L is *transversal* on the boundary ∂D if it satisfies the condition

$$\int_D t(x', y)\, dy = +\infty \quad \text{if } \mu(x') = \delta(x') = 0. \tag{1.7}$$

The intuitive meaning of condition (1.7) is that a Markovian particle jumps away "instantaneously" from the points $x' \in \partial D$ where neither reflection nor viscosity phenomenon occurs (which is similar to the reflection phenomenon). Probabilistically, this means that every Markov process on the boundary ∂D is the "trace" on ∂D of trajectories of some Markov process on the closure $\overline{D} = D \cup \partial D$.

The next theorem asserts that there exists a Feller semigroup on $\overline{D}$ corresponding to the diffusion phenomenon where one of the reflection phenomenon, the viscosity phenomenon and the inward jump phenomenon from the boundary occurs at each point of the boundary ∂D (see Fig. 1.7):

Theorem 1.2. *We define a linear operator $\mathfrak{A}$ from the space $C(\overline{D})$ into itself as follows:*

(a) The domain of definition $D(\mathfrak{A})$ of $\mathfrak{A}$ is the set

$$D(\mathfrak{A}) = \left\{u \in C(\overline{D}) : W_D u \in C(\overline{D}),\ Lu = 0\right\}. \tag{1.8}$$

(b) $\mathfrak{A}u = W_D u$ for every $u \in D(\mathfrak{A})$.

Here $W_D u$ and Lu are taken in the sense of distributions.

Assume that the boundary condition L is transversal *on the boundary ∂D. Then the operator $\mathfrak{A}$ generates a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D}$.*

Next we generalize Theorem 1.2 to the *non-transversal* case. To do this, we assume the following condition (H):

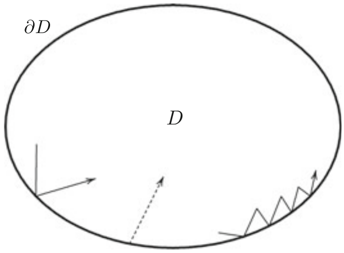


Fig. 1.7 The intuitive meaning of Theorem 1.2

(H) There exists a second-order Ventcel' boundary condition L_ν such that

$$Lu = m(x')\, L_\nu u + \gamma(x')\, u \quad \text{on } \partial D,$$

where

(3′) $m \in C^\infty(\partial D)$ and $m(x') \geq 0$ on ∂D,

and L_ν is given, in terms of local coordinates $(x_1, x_2, \ldots, x_{N-1})$ of ∂D, by the formula

$$\begin{aligned} & L_\nu u(x') \\ = \; & \overline{Q}u(x') + \overline{\mu}(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \overline{\delta}(x') W_D u(x') + \overline{\Gamma}u(x') \\ := \; & \left(\sum_{i,j=1}^{N-1} \overline{\alpha}^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \overline{\beta}^i(x') \frac{\partial u}{\partial x_i}(x') + \overline{\gamma}(x') \right) \\ & + \overline{\mu}(x') \frac{\partial u}{\partial \mathbf{n}}(x') - \overline{\delta}(x')\, W_D u(x') \\ & + \sum_{j=1}^{N-1} \overline{\eta}_\tau^j(x') \frac{\partial u}{\partial x_j}(x') + \overline{\eta}_\tau(x') u(x') \\ & + \int_{\partial D} \overline{r}(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} \left(y_j - x_j\right) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\ & + \int_D \overline{t}(x', y) \left[u(y) - u(x') \right] dy, \end{aligned}$$

and the operator $\overline{\Gamma}$ is a boundary condition of order $2-\kappa_1$, and satisfies the condition

$$\begin{aligned} \overline{Q}1(x') + \overline{\Gamma}1(x') &= \overline{\gamma}(x') + \overline{\eta}_\tau(x') + \int_{\partial D} \overline{r}(x', y') \left[1 - \tau(x', y')\right] dy' \\ &\leq 0 \quad \text{on } \partial D. \end{aligned} \tag{1.9}$$

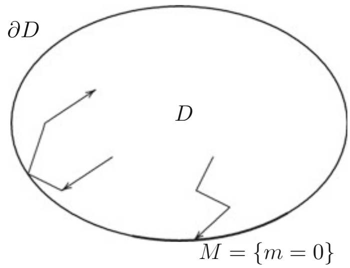


Fig. 1.8 A Markovian particle dies when it reaches the set M

Furthermore, we assume that L_ν satisfies the *transversality* condition

$$\int_D \overline{t}(x', y)\,dy = +\infty \quad \text{if } \overline{\mu}(x') = \overline{\delta}(x') = 0. \tag{1.10}$$

We let

$$M = \left\{ x' \in \partial D : \mu(x') = \delta(x') = 0,\ \int_D t(x', y)\,dy < \infty \right\}.$$

Then, by condition (1.10) it follows that

$$M = \{x' \in \partial D : m(x') = 0\},$$

since we have

$$\begin{aligned} \mu(x') &= m(x')\,\overline{\mu}(x'), \\ \delta(x') &= m(x')\,\overline{\delta}(x'), \\ t(x', y) &= m(x')\,\overline{t}(x', y). \end{aligned}$$

Hence we find that the boundary condition L is *not* transversal on ∂D.

Furthermore, we assume the following condition (A):

(A) $m(x') - \gamma(x') > 0$ on ∂D.

The intuitive meaning of conditions (H) and (A) is that a Markovian particle does not stay on ∂D for any period of time until it "dies" at the time when it reaches the set M where the particle is definitely absorbed (see Fig. 1.8).

We introduce a subspace of $C(\overline{D})$ which is associated with the non-transversal boundary condition L.

By condition (A), we find that the boundary condition

$$Lu = m(x')\,L_\nu u + \gamma(x')\,u = 0 \quad \text{on } \partial D$$

includes the condition

$$u = 0 \quad \text{on } M.$$

With this fact in mind, we let

$$C_0(\overline{D} \setminus M) = \{u \in C(\overline{D}) : u = 0 \text{ on } M\}.$$

The space $C_0(\overline{D} \setminus M)$ is a closed subspace of $C(\overline{D})$; hence it is a Banach space.

A strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on the space $C_0(\overline{D} \setminus M)$ is called a *Feller semigroup* on $\overline{D} \setminus M$ if it is non-negative and contractive on $C_0(\overline{D} \setminus M)$:

$$f \in C_0(\overline{D} \setminus M),\, 0 \leq f \leq 1 \text{ on } \overline{D} \setminus M \implies 0 \leq T_t f \leq 1 \text{ on } \overline{D} \setminus M.$$

We define a linear operator $\mathfrak{W}$ from $C_0(\overline{D} \setminus M)$ into itself as follows:

(a) The domain of definition $D(\mathfrak{W})$ of $\mathfrak{W}$ is the set

$$D(\mathfrak{W}) = \left\{u \in C_0(\overline{D} \setminus M) : W_D u \in C_0(\overline{D} \setminus M),\ Lu = 0\right\}. \tag{1.11}$$

(b) $\mathfrak{W}u = W_D u$ for every $u \in D(\mathfrak{W})$.

Here $W_D u$ and Lu are taken in the sense of distributions.

The next theorem is a generalization of Theorem 1.2 to the non-transversal case:

Theorem 1.3. *Assume that conditions (H) and (A) are satisfied. Then the operator* $\mathfrak{W}$, *defined by formula* (1.11), *generates a Feller semigroup* $\{T_t\}_{t \geq 0}$ *on* $\overline{D} \setminus M$.

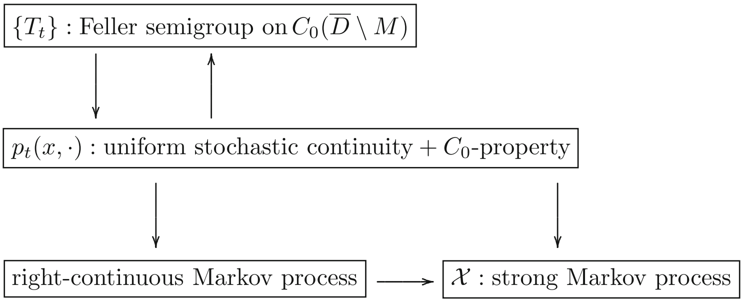
If T_t is a Feller semigroup on $\overline{D} \setminus M$, then there exists a unique Markov transition function p_t on $\overline{D} \setminus M$ such that

$$T_t f(x) = \int_{\overline{D} \setminus M} p_t(x, dy) f(y) \quad \text{for all } f \in C_0(\overline{D} \setminus M),$$

and further that p_t is the transition function of some strong Markov process $\mathcal{X}$.

On the other hand, the intuitive meaning of conditions (A) and (H) is that the absorption phenomenon occurs at each point of the set $M = \{x' \in \partial D : m(x') = 0\}$. Therefore, Theorem 1.3 asserts that there exists a Feller semigroup on $\overline{D} \setminus M$ corresponding to the diffusion phenomenon where a Markovian particle moves both by jumps and continuously in the state space $\overline{D} \setminus M$ until it "dies" at the time when it reaches the set M.

It is known (see [Dy1, Theorem 5.10]) that any right-continuous Markov process whose transition function has the C_0-property is a strong Markov process. Our functional analytic method for the problem of the existence of strong Markov processes in probability theory may be visualized as follows (see Sect. 9.2):



Finally, we consider the case where all the operators S, T and R are pseudo-differential operators of order *less than one*. Then we can take $\sigma(x, y) \equiv 1$ on $\overline{D} \times \overline{D}$, and write the Waldenfels integro-differential operator W in the following form:

$$\begin{aligned} W_D u(x) &= Pu(x) + S_D u(x) \qquad (1.12) \\ &:= \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\ &\quad + a(x)u(x) + \int_D s(x, y) \, [u(y) - u(x)] \, dy, \end{aligned}$$

where:

(4′) The integral kernel $s(x, y)$ is the distribution kernel of a properly supported, pseudo-differential operator $S \in L^{1-\kappa}_{1,0}(\mathbf{R}^N)$, $\kappa > 0$, which has the transmission property with respect to the boundary ∂D, and $s(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$.
(7′) $a(x) = (S_D 1)(x)$.
(8′) $W_D 1(x) = P1(x) + S_D 1(x) = c(x) + a(x) \leq 0$ in D.

Similarly, the boundary condition L can be written in the following form:

$$\begin{aligned} & Lu(x') \qquad (1.13) \\ &= Qu(x') + \mu(x') \frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x') W_D u(x') + \Gamma u(x') \\ &:= \left(\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x') + \gamma(x')u(x') \right) \\ &\quad + \mu(x') \frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x') W_D u(x') + \eta_\tau(x')u(x') \\ &\quad + \int_{\partial D} r(x', y') \left[u(y') - u(x') \right] dy' + \int_D t(x', y) \left[u(y) - u(x') \right] dy. \end{aligned}$$

Here:

(3′) $Q1(x') = \gamma(x') \leq 0$ on ∂D.

(7′) The integral kernel $r(x', y')$ is the distribution kernel of a pseudo-differential operator $R \in L^{1-\kappa_1}_{1,0}(\partial D)$, $\kappa_1 > 0$, and $r(x', y') \geq 0$ off the diagonal $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$ in $\partial D \times \partial D$.

(8′) The integral kernel $t(x, y)$ is the distribution kernel of a properly supported, pseudo-differential operator $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, $\kappa_2 > 0$, which has the transmission property with respect to the boundary ∂D, and $t(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$.

(12′) The boundary operator Γ satisfies the condition

$$Q1(x') + \Gamma 1(x') = \gamma(x') + \eta_\tau(x') \leq 0 \quad \text{on } \partial D.$$

Then Theorems 1.2 and 1.3 may be simplified as follows:

Theorem 1.4. *Assume that the Waldenfels operator W and the Ventcel' boundary condition L are of the forms* (1.12) *and* (1.13)*, respectively. If the boundary condition L is* transversal *on the boundary ∂D, then the operator $\mathfrak{W}$, defined by formula* (1.8)*, generates a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D}$.*

Theorem 1.5. *Assume that the Waldenfels operator W and the Ventcel' boundary condition L are of the forms* (1.12) *and* (1.13)*, respectively, and further that conditions (H) and (A) are satisfied:*

(H) There exists a second-order transversal *Ventcel' boundary condition L_ν such that*

$$Lu = m(x')\, L_\nu u + \gamma(x')\, u \quad \text{on } \partial D.$$

(A) $m(x') - \gamma(x') > 0$ on ∂D.

Then the operator $\mathfrak{W}$, defined by formula (1.11)*, generates a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D} \setminus M$.*

In this monograph we do not prove Theorems 1.4 and 1.5, since their proofs are essentially the same as those of Theorems 1.2 and 1.3, respectively. Theorems 1.2–1.5 solve the problem of the existence of Feller semigroups with Ventcel' boundary conditions for *elliptic* Waldenfels integro-differential operators from the viewpoint of functional analysis.

1.3 Summary of the Contents

This introductory Chap. 1 is intended as a brief introduction to our problem and results in such a fashion that it could be understood by a broad spectrum of readers. The contents of the monograph are divided into Parts I–III.

Part I (Chaps. 2–4) provides elements of measure theory, real analysis, probability theory, functional analysis and semigroup theory. The material in these preparatory chapters is given for completeness, to minimize the necessity of consulting too many outside references. This makes the book fairly self-contained.

Chapter 2 is intended as a brief introduction to probability theory. Section 2.1 serves to illustrate some results of measure theory, since measure spaces are the natural setting for the study of probability. We study measurable spaces and measurable functions. In particular, we prove the monotone class theorem (Theorem 2.4) and the Dynkin class theorem (Corollary 2.5) which will be useful for the study of the measurability of functions in Chap. 9. In Sect. 2.2 we introduce probability spaces and in Sect. 2.3 we consider random variables and their expectations. One of the most important concepts in probability theory is that of independence. It is the concept of independence more than anything else which gives probability theory a life of its own, distinct from other branches of analysis. In Sect. 2.4 we study independent events, independent random variables and independent algebras. In Sect. 2.5, as an application of the Radon–Nikodým theorem we introduce conditional probabilities and conditional expectations. Section 2.6 is devoted to the general theory of conditional expectations which will play a vital role in the study of Markov processes in Chap. 9.

Chapter 3 is devoted to a review of standard topics from functional analysis such as quasinormed and normed linear spaces and closed, compact and Fredholm linear operators in Banach spaces. These topics form a necessary background for what follows. In Sects. 3.1–3.3 we study linear operators and functionals, quasinormed and normed linear spaces. In a normed linear space we consider continuous linear functionals as generalized coordinates of the space. The existence of non-trivial, continuous linear functionals is based on the Hahn–Banach extension theorem (Theorem 3.21). In particular, Mazur's theorem (Theorem 3.25) asserts that there exists a non-trivial, continuous linear functional which separates a point and a closed convex, balanced set. In Sect. 3.3.2, by using Mazur's theorem we prove that a closed convex, balanced subset is weakly closed (Corollary 3.26). This corollary plays an important role in the proof of an existence and uniqueness theorem for a class of second-order classical pseudo-differential operators in the framework of Hölder spaces (Theorem 10.23) in Chap. 10. In Sect. 3.4 we prove the Riesz–Markov representation theorem (Theorem 3.41) which describes an intimate relationship between Radon measures and non-negative linear functionals on the spaces of continuous functions. This fact constitutes an essential link between measure theory and functional analysis, providing a powerful tool for constructing Markov transition functions in Chap. 9. Section 3.5 is devoted to closed operators and Sect. 3.6 is devoted to complemented subspaces in a normed linear space, respectively. Section 3.7 is devoted to the Riesz–Schauder theory for compact operators. More precisely, for a compact operator T on a Banach space, the eigenvalue problem can be treated fairly completely in the sense that the classical theory of Fredholm integral equations may be extended to the linear functional equation $Tx - \lambda x = y$ with a complex parameter λ (Theorem 3.61). In Sect. 3.8 we state important properties of Fredholm operators (Theorems 3.64–3.66). Roughly speaking, the Fredholm property of T conveys that the operator T is both "almost" injective and "almost" surjective, that is, it is "almost" an isomorphism. Moreover, the index ind T indicates how far the operator T is from being "invertible". Namely, the further ind T is from zero, the more "non-invertible" T is. The stability theorem for indices of Fredholm operators (Theorem 3.66) plays an essential role in the proof of Theorem 1.2 in Chap. 10.

Chapter 4 is devoted to the general theory of semigroups. In Sects. 4.1–4.3 we study Banach space valued functions, operator valued functions and exponential functions, generalizing the numerical case. Section 4.4 is devoted to the theory of contraction semigroups. A typical example of a contraction semigroup is the semigroup associated with the heat kernel (Example 4.11). We consider when a linear operator is the infinitesimal generator of some contraction semigroup. This question is answered by the Hille–Yosida theorem (Theorem 4.10). In Sect. 4.5 we consider when a linear operator is the infinitesimal generator of some (C_0) semigroup (Theorem 4.28), generalizing the theory of contraction semigroups developed in Sect. 4.4. Moreover, we study an initial-value problem associated with a (C_0) semigroup, and prove an existence and uniqueness theorem for the initial-value problem (Theorem 4.30). These topics form the necessary background for the proof of Theorems 13.3, 13.4, 13.9 and 13.10 in Chap. 13.

Part II (Chaps. 5–8) provides elements of distributions, Sobolev and Besov spaces, pseudo-differential operators and maximum principles for second-order elliptic Waldenfels operators which play a crucial role throughout the book. Each chapter has its own focus.

Chapter 5 is a summary of the basic definitions and results about the theory of distributions or generalized functions which will be used in subsequent chapters. Distribution theory has become a convenient tool in the study of partial differential equations. Many problems in partial differential equations can be formulated in terms of abstract operators acting between suitable spaces of distributions, and these operators are then analyzed by the methods of functional analysis. The virtue of this approach is that a given problem is stripped of extraneous data, so that the analytic core of the problem is revealed. Section 5.1 serves to settle questions of notation and such. In Sect. 5.2 we study L^p spaces, the spaces of C^k functions and test functions, and Hölder spaces on an open subset of Euclidean space. Moreover, we introduce Friedrichs' mollifiers and show how mollifiers can be used to approximate a function by smooth functions (Theorem 5.4). In Sect. 5.3 we study differential operators and state that differential operators are local operators (Peetre's theorem 5.7). In Sect. 5.4 we present a brief description of the basic concepts and results of distributions. In particular, the importance of tempered distributions lies in the fact that they have Fourier transforms. In Sect. 5.4.10 we calculate the Fourier transform of a tempered distribution which is closely related to the stationary phase theorem (Example 5.29). In Sect. 5.5 we prove the Schwartz kernel theorem (Theorem 5.36) which characterizes continuous linear operators in terms of distributions. In Sect. 5.6 we describe the classical single and double layer potentials arising in the Dirichlet and Neumann problems for the Laplacian Δ in the case of the half-space $\mathbf{R}^n_+$ (formulas (5.14) and (5.15)). Moreover, we prove Green's representation formula (5.16). This formula will be formulated in terms of pseudo-differential operators in Chap. 7. Some results in Sects. 5.3–5.5 can be extended to distributions, differential operators, and operators and kernels on a manifold in Sect. 5.7. Manifolds are an abstraction of the idea of a surface in Euclidean space. The virtue of manifold theory is that it provides a geometric insight into the study of partial differential equations, and intrinsic properties of partial differential equations may be revealed. In Sect. 5.8 we introduce the notion of domains of class C^r from the viewpoint of manifold theory.

Chapter 6 is devoted to the precise definitions and statements of Sobolev and Besov spaces of L^p type with some detailed proofs. One of the most useful ways of measuring differentiability properties of functions on $\mathbf{R}^n$ is in terms of L^p norms, and is provided by the Sobolev spaces on $\mathbf{R}^n$. The great advantage of this approach lies in the fact that the Fourier transform works very well in $L^p(\mathbf{R}^n)$. The function spaces we shall treat are the following:

(i) The generalized Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$ of L^p type on an open subset Ω of $\mathbf{R}^n$, which will be used in subsequent chapters. When Ω is a Lipschitz domain, these spaces coincide with each other.
(ii) The Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$ on $\mathbf{R}^{n-1}$ are function spaces defined in terms of the L^p modulus of continuity. The Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ of a Lipschitz domain Ω are defined to be locally the Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity.

In studying boundary value problems in the domain Ω, we need to make sense of the restriction $u|_{\partial\Omega}$ as an element of a function space on the boundary $\partial\Omega$ when u belongs to a Sobolev space of L^p type on Ω. In Sect. 6.1 we prove Hardy's inequality on the interval $(0, \infty)$ (Theorem 6.2) which is used systematically in the proof of a trace theorem (Theorem 6.6). In Sect. 6.2 we present some basic definitions and results of the Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$. In Sect. 6.3 we give the precise definition of the Besov space $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$. It should be emphasized that the Besov spaces $B^{s,p}(\partial\Omega)$ enter naturally in connection with boundary value problems in the framework of function spaces of L^p type. In Sect. 6.4 we prove a trace theorem (Theorem 6.6) which plays an important role in the study of boundary value problems in Chap. 7.

In Chap. 7 we present a brief description of the basic concepts and results of the theory of pseudo-differential operators – a modern theory of potentials – which is used in the subsequent chapters. In recent years there has been a trend in the theory of partial differential equations towards constructive methods. The development of the theory of pseudo-differential operators has made possible such an approach to the study of elliptic boundary value problems.

The purpose of Sect. 7.1 is to summarize the basic facts about manifolds with boundary and the double of a manifold which are most frequently used in the theory of partial differential equations. We formulate two fundamental theorems on smooth manifolds with boundary. The first theorem states that if Ω is a bounded domain of Euclidean space $\mathbf{R}^n$ with smooth boundary $\partial\Omega$, then $\partial\Omega$ has an open neighborhood in Ω which is diffeomorphic to $\partial\Omega \times [0, 1)$ (the product neighborhood theorem). The second theorem states that Ω is a submanifold of some n-dimensional, smooth manifold M without boundary. This manifold M is called the double of Ω.

In Sect. 7.2 we define the generalized Sobolev spaces $H^{s,p}(M)$ and the Besov spaces $B^{s,p}(\partial\Omega)$ where $M = \hat{\Omega}$ is the double of Ω. In Sect. 7.3 we introduce the Fourier integral distribution

$$K(x) = \int_{\mathbf{R}^N} e^{i\varphi(x,\theta)} a(x, \theta)\, d\theta$$

associated with the phase function $\varphi(x, \theta)$ and the amplitude $a(x, \theta)$. The operator A is called the Fourier integral operator associated with the phase function $\varphi(x, y, \theta)$ and the amplitude $a(x, y, \theta)$ if its distribution kernel $K_A(x, y)$ is given by the Fourier integral distribution

$$K_A(x, y) = \int_{\mathbf{R}^N} e^{i\varphi(x,y,\theta)} a(x, y, \theta)\, d\theta.$$

In Sect. 7.4 we define pseudo-differential operators. A pseudo-differential operator of order m is a Fourier integral operator associated with the phase function $\varphi(x, y, \xi) = (x - y) \cdot \xi$ and some amplitude $a(x, y, \xi) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^n)$. In this section we study their basic properties such as the behavior of transposes, adjoints and compositions of such operators, and the effect of a change of coordinates on such operators. It should be emphasized that Theorem 7.18 contains all the machinery necessary for the theory of pseudo-differential operators, and its proof is based on Example 5.29 and the stationary phase theorem. By using the multiplier theorem of Marcinkiewicz just as in Coifman–Meyer [CM], Bourdaud [Bd] proved an L^p boundedness theorem for pseudo-differential operators (Theorem 7.24) which plays a fundamental role throughout the book. A *global version* of Theorem 7.24 is proved in Appendix A, due to its length. This appendix is a refinement of Appendix A of the first edition of the present monograph.

The calculus of pseudo-differential operators is applied to elliptic boundary value problems in Chaps. 9 and 10. In Sect. 7.5 we describe the classical surface and volume potentials arising in boundary value problems for elliptic differential operators in terms of pseudo-differential operators. One of the important questions in the theory of elliptic boundary value problems is that of the smoothness of a solution near the boundary. In Sect. 7.6, following Boutet de Monvel [Bo], we introduce a condition concerning symbols in the normal direction at the boundary (the transmission property) in order to ensure the boundary regularity property. It should be noted that the notion of transmission property is invariant under a change of coordinates which preserves the boundary. Hence this notion can be transferred to manifolds with boundary. Section 7.7 is devoted to the Boutet de Monvel calculus. Elliptic boundary value problems cannot be treated directly by pseudo-differential operator methods. It was Boutet de Monvel who introduced the operator-algebraic aspect with his calculus in 1971. He constructed a relatively small "algebra" which contains the boundary value problems for elliptic differential operators as well as their parametrices. The operators in the Boutet de Monvel calculus may be regarded as operator-valued pseudo-differential operators. This point of view, going back to an idea of Schulze, was first sketched by Schrohe–Schulze [SS1]. In Appendix B, following Schrohe [Sr5], we follow the pseudo-differential spirit of Boutet de Monvel's construction more closely than the older descriptions. We modify Schrohe's paper [Sr5] in such a fashion that a broad spectrum of readers could understand the Boutet de Monvel calculus. In Sect. 7.8 we prove that the distribution kernel $s(x, y)$ of a pseudo-differential operator $S \in L^m_{1,0}(\mathbf{R}^n)$ satisfies the estimate

$$|s(x, y)| \leq \frac{C}{|x - y|^{m+n}}, \quad x, y \in \mathbf{R}, \quad x \neq y.$$

In Chap. 8, following Bony–Courrège–Priouret [BCP], we prove various maximum principles for second-order elliptic Waldenfels operators which play an essential role throughout the book. In Sect. 8.1 we give complete characterizations of linear operators which satisfy the positive maximum principle related to condition (β') in Theorem 9.50 in Chap. 9 (Theorems 8.2, 8.4 and 8.8). In Sect. 8.2 we prove the weak and strong maximum principles and Hopf's boundary point lemma for second-order elliptic Waldenfels operators (Theorems 8.11, 8.13 and 8.15). It should be emphasized that these characterizations give, as a byproduct, characterizations of distributions of order 2 which are non-negative outside the origin. Chapter 8 is an expanded and revised version of Sect. 3.4 and Appendix C of the first edition of the present monograph.

Part III (Chaps. 9–13) is the heart of the subject, and is devoted to the functional analytic approach to the problem of constructing Markov processes with Ventcel' boundary condition for second-order elliptic Waldenfels operators with smooth coefficients. We describe how the problem can be solved, using the mathematics presented in Parts I and II.

In Chap. 9 we introduce a class of (temporally homogeneous) Markov processes which we will deal with in this monograph. Intuitively, the (temporally homogeneous) Markov property is that the prediction of subsequent motion of a physical particle, knowing its position at time t, depends neither on the value of t nor on what has been observed during the time interval $[0, t)$; that is, a physical particle "starts afresh". From the point of view of analysis, however, the transition function of a Markov process is something more convenient than the Markov process itself. In fact, it can be shown that the transition functions of Markov processes generate solutions of certain parabolic partial differential equations such as the classical diffusion equation; and, conversely, these partial differential equations can be used to construct and study the transition functions and the Markov processes themselves. In Sect. 9.1 we give the precise definition of a (temporally homogeneous) Markov transition function adapted to the theory of semigroups (Definition 9.4). A Markov process is called a strong Markov process if the "starting afresh" property holds not only for every fixed moment but also for suitable random times. In Sect. 9.1.7

we formulate precisely this "strong" Markov property (Definition 9.25), and give a useful criterion for the strong Markov property (Theorem 9.26). In Sect. 9.1.8 we introduce the notion of uniform stochastic continuity of transition functions (Definition 9.27), and give simple criteria for the strong Markov property in terms of transition functions (Theorems 9.28 and 9.29). In Sect. 9.2 we introduce a class of semigroups associated with Markov processes (Definition 9.30), called Feller semigroups, and we give a characterization of Feller semigroups in terms of Markov transition functions (Theorems 9.33 and 9.34). Section 9.3 is devoted to a version of the Hille–Yosida theorem (Theorem 3.10) adapted to the present context. We prove generation theorems for Feller semigroups (Theorems 9.35 and 9.50) which form a functional analytic background for the proof of Theorem 1.2 in Chap. 10. In particular, Theorem 9.50 and Corollary 9.51 give useful criteria in terms of maximum principles. In Sects. 9.4 and 9.5, following Ventcel' [We], we study the problem of determining all possible boundary conditions for multi-dimensional diffusion processes. More precisely, we describe analytically the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}$ in the case where the state space is the closure $\overline{D}$ of a bounded domain D in Euclidean space $\mathbf{R}^N$ (Theorems 9.52 and 9.53). Theorems 9.52 and 9.53 are essentially due to Ventcel' [We]. Our proof of these theorems follows Bony–Courrège–Priouret [BCP], where the infinitesimal generators of Feller semigroups are studied in great detail in terms of the maximum principle (see Chap. 8). Analytically, a Markovian particle in $\overline{D}$ is governed by an integro-differential operator W, called a Waldenfels operator, in the interior D of the domain, and it obeys a boundary condition L, called a Ventcel' boundary condition, on the boundary ∂D of the domain. Probabilistically, a Markovian particle moves both by jumps and continuously in the state space and it obeys the Ventcel' boundary condition which consists of six terms corresponding to the diffusion along the boundary, the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary.

In this way, we can reduce the problem of the existence of Feller semigroups to the unique solvability of the boundary value problem for Waldenfels integro-differential operators W with Ventcel' boundary conditions L in the theory of partial differential equations.

Chapter 10 is devoted to the proof of Theorem 1.2 which is a refinement of [Ta6, Theorem 1]. In Sects. 10.1 and 10.2 we formulate our problem and Theorem 1.2, generalizing Feller's work to the multi-dimensional case. Our functional analytic approach to the problem of constructing Markov processes with Ventcel' boundary conditions is adapted from Bony–Courrège–Priouret [BCP], Cancelier [Cn], Sato–Ueno [SU] and Taira [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9].

The idea of our approach is as follows (cf. [BCP, SU, Ta5]): First, in Sect. 10.3 we consider the following *Dirichlet problem* for the Waldenfels integro-differential operator $W = P + S$:

$$\begin{cases} (\alpha - W_D)v = f & \text{in } D, \\ \gamma_0 v = \varphi & \text{on } \partial D, \end{cases} \tag{D}$$

where α is a positive parameter. We show that if $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$ has the *transmission property* with respect to ∂D, then Dirichlet problem (D) is uniquely solvable in the framework of Hölder spaces (Theorem 10.4). In the proof, we estimate the integro-differential operator S in terms of Hölder norms, and show that the pseudo-differential operator case $(W, \gamma_0) = (P + S, \gamma_0)$ may be considered as a perturbation of a *compact operator* to the differential operator case (P, γ_0) in the framework of Hölder spaces.

We let

$$v = G^0_\alpha f.$$

The operator G^0_α is called the Green operator for Dirichlet problem (D). We remark that the operator G^0_α is a generalization of the classical *Green representation formula* for Dirichlet problem (D).

In Sect. 10.4 we reduce the problem of the existence of Feller semigroups to the *unique solvability* of the boundary value problem for the Waldenfels integro-differential operator $W = P + S$ (Theorem 10.19)

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ (\lambda - L)u = \varphi & \text{on } \partial D, \end{cases}$$

and then prove existence theorems for Feller semigroups (Theorem 10.2). Here α is a positive parameter and λ is a non-negative constant.

However, we find that a function u is a solution of the problem

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ Lu = 0 & \text{on } \partial D \end{cases} \tag{$*$}$$

if and only if the function $w = u - v$ is a solution of the problem

$$\begin{cases} (\alpha - W_D)w = 0 & \text{in } D, \\ Lw = -Lv = LG^0_\alpha f & \text{on } \partial D. \end{cases}$$

On the other hand, we show that every solution w of the equation

$$(\alpha - W_D)w = 0 \quad \text{in } D$$

can be expressed by means of a single layer potential as follows:

$$w = H_\alpha \psi.$$

The operator H_α is called the harmonic operator for Dirichlet problem (D). We remark that the operator H_α is a generalization of the classical *Poisson integral formula* for Dirichlet problem (D).

Therefore, by using the Green operator G_α^0 and harmonic operator H_α we can reduce the study of problem $(*)$ to that of the equation

$$LH_\alpha\psi = -LG_\alpha^0 f \quad \text{on } \partial D.$$

This is a generalization of the classical *Fredholm integral equation* on the boundary.

Section 10.5 is devoted to the proof of Theorem 1.2. The first essential step in the proof is to show that if $T \in L_{1,0}^{1-\kappa_2}(\mathbf{R}^N)$ has the *transmission property* with respect to ∂D, then the operator LH_α is the sum of a second-order degenerate elliptic differential operator P_α and a classical pseudo-differential operator S_α with *non-negative distribution kernel* on the boundary ∂D. The second essential step in the proof is to calculate the *complete symbol* of the pseudo-differential operator LH_α. In particular, we calculate the complete symbol of the first-order classical pseudo-differential operator Π_α defined by the formula

$$\Pi_\alpha\varphi = \frac{\partial}{\partial \mathbf{n}}(H_\alpha\varphi)\bigg|_{\partial D}.$$

In the special case where P is the usual Laplacian Δ, we can write down concretely the complete symbol $p(x', \xi'; \alpha)$ of Π_α as follows (cf. formula (10.55)):

$$\begin{aligned} &p(x', \xi'; \alpha) \\ &= -|\xi'| - \frac{1}{2}\left(\frac{\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1)M(x')\right) + \sqrt{-1}\,\frac{1}{2}\operatorname{div}\delta_{(\xi')}(x') \\ &\quad + \text{terms of order} \leq -1/2 \text{ depending on } \alpha. \end{aligned}$$

Here:

(a) $|\xi'|$ is the length of ξ' with respect to the Riemannian metric $(g_{ij}(x'))$ of ∂D induced by the natural metric of $\mathbf{R}^N$.
(b) $M(x')$ is the *mean curvature* of ∂D at x'.
(c) $\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})$ is the *second fundamental form* of ∂D at x', while $\widehat{\xi'} \in T_{x'}(\partial D)$ is the tangent vector corresponding to the cotangent vector $\xi' \in T_{x'}^*(\partial D)$ by the duality between $T_{x'}(\partial D)$ and $T_{x'}^*(\partial D)$ with respect to the Riemannian metric $(g_{ij}(x'))$ of ∂D.
(d) $\operatorname{div}\delta_{(\xi')}$ is the *divergence* of a real smooth vector field $\delta_{(\xi')}$ on ∂D defined (in local coordinates) by the formula

$$\delta_{(\xi')} = \sum_{j=1}^{N-1} \frac{\partial|\xi'|}{\partial \xi_j}\frac{\partial}{\partial x_j} \quad \text{for } \xi' \neq 0.$$

This is carried out in Sect. 10.7 due to its length.

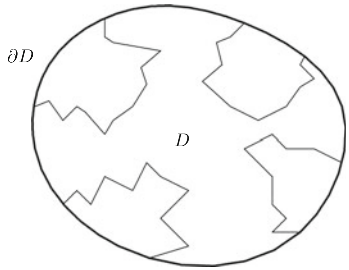


Fig. 1.9 A Markov process on ∂D pieced together with a W-diffusion in D

In the third essential step in the proof we formulate an existence and uniqueness theorem for a class of second-order classical pseudo-differential operators in the framework of Hölder spaces (Theorem 10.23) which enters naturally into the study of the pseudo-differential operator LH_α. Theorem 10.23 is proved in Sect. 10.6, by using a method of *elliptic regularizations* essentially due to Oleĭnik–Radkevič [OR] (Theorem 10.27). In order to prove the fundamental estimate (10.28), we need an interpolation argument. Moreover, we remark that Corollary 3.26 to Mazur's theorem (Theorem 3.25) in Chap. 3 plays an important role in the proof of estimate (10.28). Section 10.6 is a refinement of Appendix B of the first edition of the present monograph.

In this way, by using the Hölder space theory of pseudo-differential operators we can show that if the boundary condition L is *transversal* on the boundary ∂D, then the operator LH_α is *bijective* in the framework of Hölder spaces. More precisely, we find that a unique solution u of problem (∗) can be expressed in the form

$$u = G_\alpha^0 f - H_\alpha \left(LH_\alpha^{-1} \left(LG_\alpha^0 f \right) \right).$$

This formula allows us to verify all the conditions of the generation theorems of Feller semigroups (Theorems 9.35 and 9.50) discussed in Sect. 9.3.

Intuitively, if the boundary condition L is transversal on the boundary ∂D, then we can "piece together" a strong Markov process on the boundary ∂D with W-diffusion in the interior D to construct a strong Markov process on the closure $\overline{D} = D \cup \partial D$. It seems that our method of construction of Feller semigroups is, in spirit, not far removed from the probabilistic method of construction of diffusion processes by means of Poisson point processes of Brownian excursions used by Watanabe [Wb]. The situation may be represented schematically by Fig. 1.9.

In Chap. 11 we prove Theorem 1.3 (Theorem 11.1), generalizing Theorem 1.2 to the non-transversal case. The following idea of proof can be traced back to the work of Taira [Ta9] and [Ta6]. In fact, Theorem 1.3 is a refinement of [Ta6, Theorem 2]. In Sect. 11.1 we consider a one-point compactification $K_\partial = K \cup \{\partial\}$ of the space $K = \overline{D} \setminus M$, where

$$M = \{x' \in \partial D : m(x') = 0\},$$

and introduce a closed subspace of $C(K_\partial)$ by

$$C_0(K) = \{u \in C(K_\partial) : u(\partial) = 0\}.$$

Then we have the isomorphism

$$C_0(K) \cong C_0(\overline{D} \setminus M) = \left\{u \in C(\overline{D}) :\ u = 0 \text{ on } M\right\}.$$

In Sect. 11.2 we apply part (ii) of Theorem 9.35 to the operator $\mathfrak{W}$ defined by formula (1.11).

Our functional analytic approach may be described as follows (see Sect. 9.1): First, we note that if condition (H) is satisfied, then the boundary condition L can be written in the form

$$Lu = m(x')\, L_\nu u + \gamma(x')\, u \quad \text{on } \partial D,$$

where the boundary condition L_ν is *transversal* on ∂D. Hence, by applying Theorem 1.2 to the boundary condition L_ν we can solve uniquely the following boundary value problem:

$$\begin{cases} (\alpha - W_D)v = f & \text{in } D, \\ L_\nu v = 0 & \text{on } \partial D. \end{cases}$$

We let

$$v = G_\alpha^\nu f.$$

The operator G_α^ν is called the Green operator for the boundary condition L_ν. Then it follows that a function u is a solution of the problem

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ Lu = 0 & \text{on } \partial D \end{cases} \tag{$**$}$$

if and only if the function

$$w = u - v$$

is a solution of the problem

$$\begin{cases} (\alpha - W_D)w = 0 & \text{in } D, \\ Lw = -Lv = -\gamma(x')\, v & \text{on } \partial D. \end{cases}$$

Thus, just as in the proof of Theorem 1.2, we can reduce the study of problem ($**$) to that of the equation

$$LH_\alpha \psi = -LG_\alpha^\nu f = -\gamma(x')\, G_\alpha^\nu f \quad \text{on } \partial D.$$

By using the Hölder space theory of pseudo-differential operators as in the proof of Theorem 1.2, we can show that if condition (A) is satisfied, then the operator LH_α is *bijective* in the framework of Hölder spaces.

Therefore, we find that a unique solution u of problem ($**$) can be expressed as follows:

$$u = G_\alpha^\nu f - H_\alpha \left(LH_\alpha^{-1} \left(LG_\alpha^\nu f \right) \right).$$

This formula allows us to verify all the conditions of the generation theorems of Feller semigroups (Theorem 10.28), especially the *density* of the domain $D(\mathfrak{W})$ in $C_0(\overline{D} \setminus M)$.

It is worth pointing out that if we had used instead of G_α^ν the Green operator G_α^0 for Dirichlet problem (D) as in the proof of Theorem 1.2, then our proof would break down.

In this monograph we study mainly Markov transition functions with only informal references to the random variables which actually form the Markov processes themselves. In Chap. 12 we study this neglected side of our subject. The discussion will have a more measure-theoretical flavor than hitherto. Section 12.1 is devoted to a review of the basic definitions and properties of Markov processes. In Sect. 12.2 we consider when the paths of a Markov process are actually continuous. In Sect. 12.3 we give a useful criterion for path-continuity of a Markov process $\{x_t\}$ in terms of the infinitesimal generator $\mathfrak{A}$ of the associated Feller semigroup $\{T_t\}$. Section 12.4 is devoted to examples of multi-dimensional diffusion processes. More precisely, we prove that (1) reflecting barrier Brownian motion, (2) reflecting and absorbing barrier Brownian motion, (3) reflecting, absorbing and drifting barrier Brownian motion, are typical examples of multi-dimensional diffusion processes, that is, examples of continuous strong Markov processes on a bounded domain. It should be emphasized that these three Brownian motions correspond to the Neumann boundary value problem, the Robin boundary value problem and the oblique derivative boundary value problem for the Laplacian in terms of elliptic boundary value problems, respectively.

In Chap. 13 we summarize the contents of the first edition of the present monograph "Semigroups, boundary value problems and Markov processes" which was published in 2004. In Sect. 13.1 we study a class of *degenerate* boundary value problems for second-order elliptic differential operators which includes as particular cases the Dirichlet and Robin problems. We state existence and uniqueness theorems for this class of degenerate elliptic boundary value problems (Theorems 13.1 and 13.2). The crucial point is how to define modified boundary spaces $B_{L_0}^{1-1/p,p}(\partial D)$ and $C_{L_0}^{1+\theta}(\partial D)$ in which our boundary value problems are uniquely solvable. The purpose of Sect. 13.2 is to study our degenerate elliptic boundary value problems from the viewpoint of the theory of analytic semigroups, and to generalize generation theorems for *analytic semigroups* both in the L^p topology and in the topology of uniform convergence (Theorems 13.3 and 13.4). As an application, we state generation theorems for Feller semigroups corresponding to the diffusion phenomenon where a Markovian particle moves continuously until it "dies" at the time when it reaches the set where the particle is definitely absorbed (Theorem 13.5). In Sect. 13.3 we assume that the domain D is *convex*, and extend the existence and uniqueness theorems for degenerate elliptic boundary value problems in Sect. 13.1 (Theorems 13.6) and the generation theorems for analytic and Feller semigroups in Sect. 13.2 to the *integro-differential operator* case (Theorems 13.7, 13.9 and 13.10). Due to the non-local character of integro-differential operators, we are forced to impose various conditions on the structure of jumps of Markovian particles such as the *moment condition*. Moreover, in order to remove a singularity of solutions at the boundary ∂D, we impose the condition that no jumps outside the closure $\overline{D}$ are allowed.

As concluding remarks, we give an overview of general results on generation theorems for Feller semigroups proved mainly by the author [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9, Ta10, Ta11] using the theory of pseudo-differential operators [Ho1, Se1, Se2] and the Calderón–Zygmund theory of singular integral operators [CZ].

Bibliographical references are discussed primarily in the Notes and Comments at the end of the chapters. These notes are intended to supplement the text and place it in a better perspective.

1.4 Notes and Comments

Our functional analytic approach to the problem of constructing Markov processes with Ventcel' boundary conditions in probability is adapted from Bony–Courrège–Priouret [BCP], Cancelier [Cn], Sato–Ueno [SU] and Taira [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9]. The approach here is distinguished by the extensive use of ideas and techniques characteristic of recent developments in the theory of pseudo-differential operators. The theory of pseudo-differential operators continues to be one of the most influential topics in the modern history of analysis, and is a very refined mathematical tool whose full power is yet to be exploited (see [Ho4]). Several recent developments in the theory of pseudo-differential operators have made possible further progress in the study of elliptic boundary value problems and hence in the study of Markov processes. The presentation of these new results is the main purpose of this monograph.

This monograph is an expanded and revised version of the previous works Taira [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9], and has been revised to streamline some of the analysis and to give better coverage of important examples and applications. The theory has reached a state of completion that makes it ideal for presentation in book form. More precisely, Theorems 1.2 and 1.3 are an expanded and revised version of [Ta6, Theorem 1] and [Ta6, Theorem 2], respectively.

This book fills a mathematical gap between textbooks on Markov processes such as Bass [Ba1, Ba2], Garroni–Menaldi [GM2], Skubachevskii [Sk] and recent developments in analysis.

Skubachevskii [Sk] studies elliptic boundary value problems containing Non-local terms with support inside the domain and small non-local terms with support near the boundary. Based on the classical works of Agmon–Douglis–Nirenberg [ADN] and Agranovich–Vishik [AV], he proves the Fredholm property and the stability of the index of elliptic boundary value problems. As an application of these results, he gives sufficient conditions for the existence of Feller semigroups with Ventcel' boundary conditions. However, his sufficient conditions in the main theorems are formulated in terms of general integro-differential operators acting on Hölder spaces. He does not give a probabilistic interpretation of his sufficient conditions for the existence of Feller semigroups with Ventcel' boundary conditions.

On the other hand, Jacob [Ja] discusses the martingale problem for a large class of pseudo-differential operators in the whole space $\mathbf{R}^N$. In particular, he studies the martingale problem for generators of Lévy type. However, the objectives of his works are different from this monograph, which focuses on general Ventcel' boundary conditions. Due to the non-local character of the Waldenfels Integro-differential operator W, we find more difficulties in the bounded domain D than in the whole space $\mathbf{R}^N$.

For detailed studies of stochastic differential equations with jumps, the reader might refer to Applebaum [Ap] and Bichteler [Bi].

Part I
Elements of Analysis

Chapter 2
Elements of Probability Theory

This chapter is intended as a brief introduction to probability theory.

First, we present a brief dictionary of the Probabilists' dialect due to Folland [Fo2, Chapter 10]:

Real Analysis	**Probability**
Measure space $(X, \mathcal{M}, \mu)$	Probability space $(\Omega, \mathcal{F}, P)$
σ-algebra	σ-field
Measurable set	Event
Real-valued measurable function $f(x)$	Random variable $X(\omega)$
Integral $\int_X f(x)\, d\mu$	Expectation (Mean) $E(X) = \int_\Omega X(\omega)\, dP$
L^p-norm $\int_X \lvert f\rvert^p\, d\mu$	p-th moment $E(\lvert X\rvert^p)$
Convergence in measure	Convergence in probability
almost everywhere (a.e.)	almost surely (a.s.)
Borel probability measure on $\mathbf{R}$	Distribution
Fourier transform of a measure	Characteristic function of a distribution
Characteristic function	Indicator function

Section 2.1 serves to illustrate some results of measure theory, since measure spaces are the natural setting for the study of probability. In particular, we prove the monotone class theorem (Theorem 2.4) and the Dynkin class theorem

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_2,

(Corollary 2.5) which will be useful for the study of measurability of functions in Chap. 9. In Sect. 2.2 we introduce probability spaces and in Sect. 2.3 we consider random variables and their expectations. One of the most important concepts in probability theory is that of independence. It is the concept of independence more than anything else which gives probability theory a life of its own, distinct from other branches of analysis. In Sect. 2.4 we study independent events, independent random variables and independent algebras. In Sect. 2.5, as an application of the Radon–Nikodým theorem, we introduce conditional probabilities and conditional expectations (Definitions 2.23, 2.26 and 2.32). Section 2.6 is devoted to the general theory of conditional expectations which will play a vital role in the study of Markov processes in Chap. 9.

2.1 Measurable Spaces and Functions

This section serves to illustrate some results of measure theory, since measure spaces are the natural setting for the study of probability. We study measurable spaces and measurable functions. In particular, we prove the monotone class theorem (Theorem 2.4) and the Dynkin class theorem (Corollary 2.5) which will be useful for the study of measurability of functions in Chap. 9.

2.1.1 Measurable Spaces

Let X be a non-empty set. An *algebra* of sets on X is a non-empty collection $\mathcal{A}$ of subsets of X which is closed under finite unions and complements, that is, if it has the following two properties (F1) and (F2):

(F1) If $E \in \mathcal{A}$, then its complement $E^c = X \setminus E$ belongs to $\mathcal{A}$.
(F2) If $\{E_j\}_{j=1}^n$ is an arbitrary finite collection of members of $\mathcal{A}$, then the union $\bigcup_{j=1}^n E_j$ belongs to $\mathcal{A}$.

Two subsets E and F of X are said to be *disjoint* if $E \cap F = \emptyset$, that is, if there are no elements common to E and F. A *disjoint union* is a union of sets that are mutually disjoint.

A collection $\mathcal{E}$ of subsets of X is called an *elementary family* on X if it has the following three properties (EF1)–(EF3):

(EF1) The empty set $\emptyset$ belongs to $\mathcal{E}$.
(EF2) If $E, F \in \mathcal{E}$, then their intersection $E \cap F$ belongs to $\mathcal{E}$.
(EF3) If $E \in \mathcal{E}$, then the complement $E^c = X \setminus E$ is a finite disjoint union of members of $\mathcal{E}$.

It should be noted that if $\mathcal{E}$ is an elementary family, then the collection of finite disjoint unions of members of $\mathcal{E}$ is an algebra.

A *σ-algebra* of sets on X is an algebra which is closed under countable unions and complements. More precisely, a non-empty collection $\mathcal{M}$ of subsets of X is called a σ-algebra if it has the following three properties (S1)–(S3):

(S1) The empty set $\emptyset$ belongs to $\mathcal{M}$.
(S2) If $A \in \mathcal{M}$, then its complement $A^c = X \setminus A$ belongs to $\mathcal{M}$.
(S3) If $\{A_n\}_{n=1}^{\infty}$ is an arbitrary countable collection of members of $\mathcal{M}$, then the union $\bigcup_{n=1}^{\infty} A_n$ belongs to $\mathcal{M}$.

The pair $(X, \mathcal{M})$ is called a *measurable space* and the members of $\mathcal{M}$ are called *measurable sets* in X.

It is easy to see that the intersection of any family of σ-algebras on X is a σ-algebra. Therefore, for any collection $\mathcal{F}$ of subsets of X we can find a unique smallest σ-algebra $\sigma(\mathcal{F})$ on X which contains $\mathcal{F}$, that is, the intersection of all σ-algebras containing $\mathcal{F}$. This $\sigma(\mathcal{F})$ is sometimes called the *σ-algebra generated by $\mathcal{F}$*.

If X is a topological space, then the σ-algebra $\mathcal{B}(X)$ generated by the family $\mathcal{O}_X$ of open sets in X is called the *Borel σ-algebra* on X. In other words, we have

$$\mathcal{B}(X) = \sigma(\mathcal{O}_X).$$

The members of $\mathcal{B}(X)$ are called *Borel sets* in X.

We sometimes consider measurability on subsets of X. If Ω is a non-empty Borel set of X, then the collection

$$\mathcal{B}(\Omega) = \{\Omega \cap A : A \in \mathcal{B}(X)\}$$

is a σ-algebra on Ω.

The next proposition asserts that the Borel σ-algebra $\mathcal{B}(\mathbf{R})$ on $\mathbf{R}$ can be generated in a number of different ways [Fo2, Proposition 1.2]:

Proposition 2.1. *The Borel σ-algebra $\mathcal{B}(\mathbf{R})$ is generated by each of the following five collections (a)–(e):*

(a) The open rays: $\mathcal{E}_1 = \{(a, \infty) : -\infty < a < \infty\}$ *or* $\mathcal{E}_2 = \{(-\infty, b) : -\infty < b < \infty\}$.
(b) The closed rays: $\mathcal{E}_3 = \{[a, \infty) : -\infty < b < \infty\}$ *or* $\mathcal{E}_4 = \{(-\infty, b] : -\infty < b < \infty\}$.
(c) The open intervals: $\mathcal{E}_5 = \{(a, b) : -\infty < a < b < \infty\}$.
(d) The half-open intervals: $\mathcal{E}_6 = \{(a, b] : -\infty < a < b < \infty\}$ *or* $\mathcal{E}_7 = \{[a, b) : -\infty < a < b < \infty\}$.
(e) The closed intervals: $\mathcal{E}_8 = \{[a, b] : -\infty < a < b < \infty\}$.

Moreover, the next proposition asserts that the Borel σ-algebra $\mathcal{B}(\mathbf{R}^n)$ generated by the family $\mathcal{O}_n$ of open sets in $\mathbf{R}^n$ can be generated in a number of different ways:

Proposition 2.2. *The Borel σ-algebra $\mathcal{B}(\mathbf{R}^n)$ on $\mathbf{R}^n$ is generated by each of the following three collections (a)–(c):*

(a) The open intervals: $\mathcal{E}_1 = \{(a_1, b_1) \times \cdots \times (a_n, b_n) : -\infty < a_i < b_i < \infty \ (1 \leq i \leq n)\}$.
(b) The half-open intervals: $\mathcal{E}_2 = \{(a_1, b_1] \times \cdots \times (a_n, b_n] : -\infty < a_i < b_i < \infty \ (1 \leq i \leq n)\}$ *or* $\mathcal{E}_3 = \{[a_1, b_1) \times \cdots \times [a_n, b_n) : -\infty < a_i < b_i < \infty \ (1 \leq i \leq n)\}$.
(c) The product of Borel sets of $\mathbf{R}$*:* $\mathcal{E}_4 = \mathcal{B}(\mathbf{R}) \times \cdots \times \mathcal{B}(\mathbf{R}) = \{A_1 \times \cdots \times A_n : A_i \in \mathcal{B}(\mathbf{R}) \ (1 \leq i \leq n)\}$.

Now we introduce a new class of subsets of X which is closely related to σ-algebras:

Definition 2.3. Let $\mathcal{F}$ be a collection of subsets of X.

(i) $\mathcal{F}$ is called a π-*system* in X if it is closed under finite intersections.
(ii) $\mathcal{F}$ is called a d-*system* in X if it has the following three properties (a)–(c):

  (a) The set X itself belongs to $\mathcal{F}$.
  (b) If A, $B \in \mathcal{F}$ and $A \subset B$, then the difference $B \setminus A$ belongs to $\mathcal{F}$.
  (c) If $\{A_n\}_{n=1}^{\infty}$ is an increasing sequence of members of $\mathcal{F}$, then the union $\bigcup_{n=1}^{\infty} A_n$ belongs to $\mathcal{F}$.

It should be emphasized that a collection $\mathcal{F}$ is a σ-algebra if and only if it is both a π-system and a d-system. For any collection $\mathcal{F}$ of subsets of X, there exists a smallest d-system $d(\mathcal{F})$ which contains $\mathcal{F}$. Indeed, it suffices to note that the intersection of an arbitrary number of d-systems is again a d-system.

The next theorem gives a useful criterion for the d-system $d(\mathcal{F})$ to be a σ-algebra:

Theorem 2.4 (the monotone class theorem). *If a collection $\mathcal{F}$ of subsets of X is a π-system, then it follows that*

$$d(\mathcal{F}) = \sigma(\mathcal{F}).$$

Proof. Since we have

$$d(\mathcal{F}) \subset \sigma(\mathcal{F}),$$

we have only to show that $d(\mathcal{F})$ is a σ-algebra. To do this, it suffices to show that $d(\mathcal{F})$ is a π-system. The proof is divided into two steps.

Step 1: First, we let

$$\mathcal{D}_1 := \{B \in d(\mathcal{F}) : B \cap A \in d(\mathcal{F}) \quad \text{for all } A \in \mathcal{F}\}.$$

Then it is easy to verify that $\mathcal{D}_1$ is a d-system and further that

$$\mathcal{F} \subset \mathcal{D}_1,$$

since $\mathcal{F}$ is a π-system. Hence we have

$$d(\mathcal{F}) \subset \mathcal{D}_1,$$

and so

$$\mathcal{D}_1 = d(\mathcal{F}).$$

Step 2: Secondly, we let

$$\mathcal{D}_2 := \{B \in d(\mathcal{F}) : B \cap A \in d(\mathcal{F}) \quad \text{for all } A \in d(\mathcal{F})\}.$$

Again, it is easy to verify that $\mathcal{D}_2$ is a d-system. Moreover, if A is an arbitrary element of $\mathcal{F}$, then we have, for all $B \in \mathcal{D}_1 = d(\mathcal{F})$,

$$B \cap A \in d(\mathcal{F}).$$

This proves that

$$\mathcal{F} \subset \mathcal{D}_2.$$

Hence we have

$$d(\mathcal{F}) \subset \mathcal{D}_1,$$

and so

$$\mathcal{D}_2 = d(\mathcal{F}).$$

This implies that $d(\mathcal{F})$ is closed under finite intersections, that is, $d(\mathcal{F})$ is a π-system.

The proof of Theorem 2.4 is complete.

The next version of the monotone class theorem will be useful for the study of measurability of functions in Chap. 9:

Corollary 2.5 (the Dynkin class theorem). *Let $\mathcal{F}$ be a π-system. If $\mathcal{D}$ is a d-system which contains $\mathcal{F}$, then it follows that*

$$\sigma(\mathcal{F}) \subset \mathcal{D}.$$

Indeed, it follows from an application of Theorem 2.4 that

$$\sigma(\mathcal{F}) = d(\mathcal{F}) \subset \mathcal{D},$$

since $\mathcal{D}$ is a d-system which contains $\mathcal{F}$.

2.1.2 Measures

Let X be a non-empty set equipped with a collection $\mathcal{D}$ of subsets of X. An extended real-valued *set function* μ is a function defined on $\mathcal{D}$ taking extended real numbers. The collection $\mathcal{D}$ is called the *domain of definition* of μ.

Let μ be a set function defined on an algebra $\mathcal{A}$ on X. The set function μ is said to be *finitely additive* if we have, for any $m \in \mathbf{N}$,

$$\mu\left(\sum_{j=1}^{m} A_j\right) = \sum_{j=1}^{m} \mu(A_j)$$

provided that the A_j are mutually disjoint sets of $\mathcal{A}$. We say that μ is *countably additive* if we have the equality

$$\mu\left(\sum_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

provided that the A_j are mutually disjoint sets of $\mathcal{A}$ such that $\sum_{j=1}^{\infty} A_j$.

Let $(X, \mathcal{M})$ be a measurable space. An extended real-valued function μ defined on the σ-algebra $\mathcal{M}$ is called a *non-negative measure* or simply a *measure* if it has the following three properties (M1)–(M3):

(M1) $0 \leq \mu(E) \leq \infty$ for all $E \in \mathcal{M}$.
(M2) $\mu(\emptyset) = 0$.
(M3) The function μ is countably additive, that is,

$$\mu\left(\sum_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$$

for any disjoint countable collection $\{E_j\}_{j=1}^{\infty}$ of members of $\mathcal{M}$.

The triplet $(X, \mathcal{M}, \mu)$ is called a *measure space*. In other words, a measure space is a measurable space which has a non-negative measure defined on the σ-algebra of its measurable sets.

If $\mu(X) < \infty$, then the measure μ is called a *finite measure* and the space $(X, \mathcal{M}, \mu)$ is called a *finite measure space*. If X is a countable union of sets of finite measure, then the measure μ is said to be σ*-finite* on X. We also say that the measure space $(X, \mathcal{M}, \mu)$ is σ*-finite*.

Some basic properties of measures are summarized in the following four properties (a)–(d):

(a) (Monotonicity) If $E, F \in \mathcal{M}$ and $E \subset F$, then $\mu(E) \leq \mu(F)$.
(b) (Subadditivity) If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$, then we have the inequality

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu(E_j).$$

(c) (Continuity from below) If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ and if $E_1 \subset E_2 \subset \ldots$, then we have the equality

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \lim_{j \to \infty} \mu(E_j).$$

(d) (Continuity from above) If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ and if $E_1 \supset E_2 \supset \ldots$ and $\mu(E_1) < \infty$, then we have the equality

$$\mu\left(\bigcap_{j=1}^{\infty} E_j\right) = \lim_{j \to \infty} \mu(E_j).$$

Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $\mathcal{A}$ be an algebra in $\mathcal{M}$. The next approximation theorem asserts that every set in the σ-algebra $\sigma(\mathcal{A})$ can be approximated by sets in the algebra $\mathcal{A}$:

Theorem 2.6 (the approximation theorem). *If $\Lambda \in \sigma(\mathcal{A})$ and $\mu(\Lambda) < \infty$, then there exists a sequence $\{A_n\}_{n=1}^{\infty}$ in $\mathcal{A}$ such that*

$$\lim_{n \to \infty} \mu(\Lambda \triangle A_n) = 0,$$

where

$$A \triangle B = (A \setminus B) \cup (B \setminus A) \tag{2.1}$$

is the symmetric difference of A and B.

Proof. We let

$$\mathcal{C} = \{\Lambda \in \mathcal{M} : \text{ condition (2.1) holds true}\}.$$

We have only to show that $\mathcal{C}$ is a σ-algebra which contains $\mathcal{A}$. Indeed, we then have the assertion

$$\sigma(\mathcal{A}) \subset \mathcal{C}.$$

In other words, the desired condition (2.1) holds true for all $\Lambda \in \sigma(\mathcal{A})$.

(a) First, it follows that $\mathcal{A} \subset \mathcal{C}$. Indeed, it suffices to take $A_n := \Lambda$ if $\Lambda \in \mathcal{A}$.
(b) Secondly, if $\Lambda \in \mathcal{C}$, then it is easy to see that

$$\Lambda^c \triangle A_n^c = \Lambda \triangle A_n,$$

so that

$$\lim_{n\to\infty} \mu\left(\Lambda^c \triangle A_n^c\right) = \lim_{n\to\infty} \mu(\Lambda \triangle A_n) = 0.$$

This proves that $\Lambda^c \in \mathcal{C}$, since $A_n^c \in \mathcal{A}$.
(c) Thirdly, if $\{\Lambda_n\}_{n=1}^{\infty} \subset \mathcal{C}$, then it follows that $\Lambda = \cup_{n=1}^{\infty} \Lambda_n \in \mathcal{C}$. Indeed, since $\mu(\Lambda) < \infty$, it follows from the continuity from above of the measure μ that, for any given $\varepsilon > 0$, there exists a positive number $N = N(\varepsilon)$ such that

$$\mu\left(\Lambda \setminus \bigcup_{n=1}^{N} \Lambda_n\right) < \frac{\varepsilon}{2}. \tag{2.2}$$

On the other hand, for each $\Lambda_n \in \mathcal{C}$ there exists a set $A_n \in \mathcal{A}$ such that

$$\mu\left(\Lambda_n \triangle A_n\right) < \frac{\varepsilon}{2^{n+1}}, \quad n = 1, 2, \ldots, N. \tag{2.3}$$

However, it is easy to see that

$$\left(\bigcup_{n=1}^{N} A_n\right) \triangle \Lambda \subset \bigcup_{n=1}^{N} (A_n \triangle \Lambda_n) \bigcup \left(\Lambda \setminus \bigcup_{n=1}^{N} \Lambda_n\right). \tag{2.4}$$

Therefore, by combining inequalities (2.2) and (2.3) we obtain from assertion (2.4) that

$$\begin{aligned}\mu\left(\Lambda \setminus \bigcup_{n=1}^{N} \Lambda_n\right) &\leq \sum_{n=1}^{N} \mu\left(A_n \triangle \Lambda_n\right) + \mu\left(\Lambda \setminus \bigcup_{n=1}^{N} \Lambda_n\right) \\ &< \frac{\varepsilon}{2}\left(\frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^N}\right) + \frac{\varepsilon}{2} < \varepsilon, \quad \bigcup_{n=1}^{N} A_n \in \mathcal{A}.\end{aligned}$$

This proves that $\Lambda = \cup_{n=1}^{\infty} \Lambda_n \in \mathcal{C}$.

Summing up, we have proved that $\mathcal{C}$ is a σ-algebra which contains $\mathcal{A}$.

The proof of Theorem 2.6 is complete.

2.1.3 Measurable Functions

We let

$$\overline{\mathbf{R}} = \{-\infty\} \cup \mathbf{R} \cup \{\infty\}$$

with the obvious ordering, where $\infty = +\infty$. The topology on $\overline{\mathbf{R}}$ is defined by declaring that the open sets in $\overline{\mathbf{R}}$ are those which are unions of segments of the types

$$(a, b), \quad [-\infty, a), \quad (a, \infty].$$

The elements of $\overline{\mathbf{R}}$ are called *extended real numbers*.

If we define a mapping

$$\phi : \mathbf{R} \longrightarrow (-1, 1)$$

by the formula

$$\phi(x) = \frac{x}{1 + |x|} \quad \text{for each } x \in \mathbf{R},$$

then it is easy to verify that the space $\overline{\mathbf{R}}$ is topologically isomorphic to the closed interval $[-1, 1]$.

Let $(X, \mathcal{M})$ be a measurable space. An extended real-valued function $f(x)$ defined on a set $A \in \mathcal{M}$ is said to be $\mathcal{M}$-*measurable* or simply *measurable* if the set

$$f^{-1}((a, \infty)) = \{x \in A : f(x) > a\}$$

is in $\mathcal{M}$ for every $a \in \mathbf{R}$. It is easy to verify that an extended real-valued function $f(x)$ is $\mathcal{M}$-measurable if and only if it satisfies the following two conditions (1) and (2):

(1) $f^{-1}(B) \in \mathcal{M}$ for every $B \in \mathcal{B}(\mathbf{R})$.
(2) $f^{-1}(\infty) \in \mathcal{M}$ and $f^{-1}(-\infty) \in \mathcal{M}$.

It should be emphasized that the sets $f^{-1}(\infty)$ and $f^{-1}(-\infty)$ are measurable for extended real-valued functions.

If $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ is the Borel measurable space, we say that $f(x)$ is *Borel measurable*. It is easy to see that continuous functions on $\mathbf{R}^n$ are Borel measurable.

Some basic properties of measurable functions are summarized in the following three properties (MF1)–(MF3):

(MF1) If f and g are measurable functions, then $f \pm g$ and fg are measurable functions.

(MF2) If $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions, then the functions

$$\sup_{n \geq 1} f_n, \quad \inf_{n \geq 1} f_n,$$

$$\limsup_{n \to \infty} f_n, \quad \liminf_{n \to \infty} f_n$$

are all measurable.

(MF3) If a sequence $\{f_n\}_{n=1}^{\infty}$ of measurable functions converges to a function g, then the limit function g is measurable.

If A is a subset of X, we let

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{for } x \notin A. \end{cases}$$

The function χ_A is called the *characteristic function* of A.

A real-valued function $f(x)$ on X is called a *simple function* if it takes on only a finite number of values. Thus, if $a_1, a_2, \ldots, a_m$ are the distinct values of $f(x)$, then $f(x)$ can be written as follows (see Fig. 2.1):

$$f(x) = \sum_{j=1}^{m} a_j \, \chi_{A_j}(x),$$

where

$$A_j = \left\{ x \in X : f(x) = a_j \right\}.$$

We call this formula the standard representation of $f(x)$. It expresses $f(x)$ as a linear combination, with distinct coefficients a_j, of characteristic functions χ_{A_j} of disjoint sets A_j whose union is X. We note that the function $f(x)$ is measurable if and only if each set A_j is measurable.

The next theorem asserts that arbitrary measurable functions can be approximated in a nice way by simple functions [Fo2, Theorem 2.10]:

Theorem 2.7. *An extended real-valued function defined on a measurable set is measurable if and only if it is a pointwise limit of a sequence of measurable simple functions. Furthermore, every non-negative measurable function is a pointwise limit of an increasing sequence of non-negative measurable simple functions.*

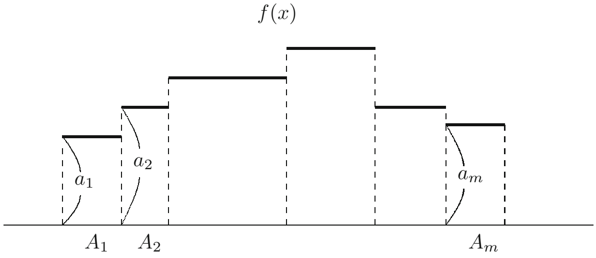


Fig. 2.1 The simple function $f(x)$

Let $(X, \mathcal{M})$ be a measurable space. If $E \in \mathcal{M}$, we let

$$\mathcal{M}(E) = \{E \cap F : F \in \mathcal{M}\}.$$

It is easy to see that $\mathcal{M}(E)$ is a σ-algebra on E. An extended real-valued function f defined on X is said to be *measurable* on E if the restriction $f|_E$ of f to E is $\mathcal{M}(E)$-measurable. In other words, f is measurable on E if and only if it satisfies the condition

$$E \cap f^{-1}(B) \in \mathcal{M} \quad \text{for every } B \in \mathcal{B}(\mathbf{R}).$$

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. A mapping $f : X \to Y$ is said to be $\mathcal{M}/\mathcal{N}$*-measurable* or simply *measurable* if it satisfies the condition

$$f^{-1}(\mathcal{N}) \subset \mathcal{M}.$$

The next proposition is often useful:

Proposition 2.8. *Assume that the σ-algebra $\mathcal{N}$ is generated by a non-empty collection $\mathcal{E}$ of subsets of Y: $\sigma(\mathcal{E}) = \mathcal{N}$. Then the mapping $f : X \to Y$ is $\mathcal{M}/\mathcal{N}$-measurable if and only if $f^{-1}(\mathcal{E}) \subset \mathcal{M}$. Moreover, we have*

$$\sigma\left(f^{-1}(\mathcal{E})\right) = f^{-1}\left(\sigma(\mathcal{E})\right) = f^{-1}(\mathcal{N}).$$

The next theorem is a version of the monotone class theorem (Theorem 2.4), and will be useful for the study of measurability of functions in Chap. 9:

Theorem 2.9. *Let $\mathcal{F}$ be a π-system in X and let $\mathcal{H}$ be a linear space of real-valued functions on X. Assume that the following two conditions (i) and (ii) are satisfied:*

(i) $1 \in \mathcal{H}$ *and* $\chi_A \in \mathcal{H}$ *for all* $A \in \mathcal{F}$.
(ii) *If* $\{f_n\}_{n=1}^{\infty}$ *is an increasing sequence of non-negative functions in* $\mathcal{H}$ *such that* $f = \sup_{n \geq 1} f_n$ *is bounded, then it follows that* $f \in \mathcal{H}$.

Then the linear space $\mathcal{H}$ contains all real-valued, bounded functions on X which are $\sigma(\mathcal{F})$-measurable.

Proof. The proof is divided into three steps.

Step 1: We let

$$\mathcal{D} := \{A \subset X : \chi_A \in \mathcal{H}\},$$

and prove that

$$\sigma(\mathcal{F}) \subset \mathcal{D}.$$

To do this, we show that $\mathcal{D}$ is a d-system containing $\mathcal{F}$.

(a) First, by condition (i) it follows that $X \in \mathcal{D}$ and that

$$\mathcal{F} \subset \mathcal{D}.$$

(b) Secondly, since $\mathcal{H}$ is a linear space, it follows that we have, for all $A_1, A_2 \in \mathcal{D}$ with $A_1 \subset A_2$,

$$\chi_{A_2 \setminus A_2} = \chi_{A_2} - \chi_{A_1} \in \mathcal{H}.$$

This implies that $A_2 \setminus A_1 \in \mathcal{D}$.

(c) Finally, if $\{A_n\}_{n=1}^{\infty}$ is an arbitrary sequence of sets in $\mathcal{D}$, then it follows from condition (ii) that

$$\chi_{\cup_{n=1}^{\infty} A_n} = \sup_{n \geq 1} \chi_{A_n} \in \mathcal{H}.$$

This implies that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Therefore, we have proved that $\mathcal{D}$ is a d-system containing $\mathcal{F}$, so that

$$d(\mathcal{F}) \subset \mathcal{D}.$$

On the other hand, since $\mathcal{F}$ is a π-system, it follows from an application of the monotone class theorem (Theorem 2.4) that

$$d(\mathcal{F}) = \sigma(\mathcal{F}).$$

Summing up, we obtain that

$$\sigma(\mathcal{F}) = d(\mathcal{F}) \subset \mathcal{D}.$$

Step 2: If $f(x)$ is an arbitrary real-valued, bounded $\sigma(\mathcal{F})$-measurable function on X, then we can write it in the form

$$f(x) = f^+(x) - f^-(x),$$

where

$$f^+(x) = \max\{f(x), 0\},$$
$$f^-(x) = \max\{-f(x), 0\}.$$

It should be noted that both $f^+(x)$ and $f^-(x)$ are non-negative and $\sigma(\mathcal{F})$-measurable functions on X.

Moreover, by applying Theorem 2.7 we obtain that every non-negative, $\sigma(\mathcal{F})$-measurable function $g(x)$ is a pointwise limit of an increasing sequence $\{g_n(x)\}_{n=1}^{\infty}$ of non-negative, $\sigma(\mathcal{F})$-measurable simple functions.

However, since $\sigma(\mathcal{F}) \subset \mathcal{D}$ and since $\mathcal{H}$ is a linear space, it follows that the simple functions $g_n(x)$ are in $\mathcal{H}$ and further from condition (ii) that

$$g(x) = \sup_{n \geq 1} g(x) \in \mathcal{H}.$$

Step 3: Since $f^{\pm}(x)$ are non-negative and $\sigma(\mathcal{F})$-measurable, we obtain from Step 2 that

$$f^+(x), \ f^-(x) \in \mathcal{H},$$

so that

$$f(x) = f^+(x) - f^-(x) \in \mathcal{H}.$$

The proof of Theorem 2.9 is complete.

2.2 Probability Spaces

Let Ω be a non-empty set and let $\mathcal{F}$ be a σ-algebra of subsets of Ω, that is, $\mathcal{F}$ is a collection of subsets which contains the empty set $\emptyset$ and is closed under the formation of complements and of the union of countably many of members. A function P defined on $\mathcal{F}$ is called a *probability measure* if it satisfies the following three conditions (P1)–(P3):

(P1) $P(A) \geq 0$ for all $A \in \mathcal{F}$.
(P2) $P(\Omega) = 1$.

(P3) The function P is countably additive, that is,

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

for any disjoint countable collection $\{A_n\}_{n=1}^{\infty}$ of members of $\mathcal{F}$.

The triplet $(\Omega, \mathcal{F}, P)$ is called a *probability space*. The elements of Ω are known as sample points, those of $\mathcal{F}$ as events and the values $P(A)$, $A \in \mathcal{F}$, are their probabilities.

We remark that conditions (P1)–(P3) imply the following *continuity condition* (P4) of the probability P:

(P4) If $A_n \in \mathcal{F}$, and $A_{n+1} \subset A_n$ and $\cup_{n=1}^{\infty} = \emptyset$, then it follows that

$$\lim_{n\to\infty} P(A_n) = 0.$$

2.3 Random Variables and Expectations

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A real-valued, $\mathcal{F}$-measurable function X defined on Ω is called a *random variable*. In other words, a real-valued function X on Ω is a random variable if and only if it satisfies the condition

$$X^{-1}(A) \in \mathcal{F} \quad \text{for every } A \in \mathcal{B}(\mathbf{R}). \tag{2.5}$$

However, it is easy to verify (see Proposition 2.1) that X satisfies condition (2.5) if and only if $X^{-1}((-\infty, r)) \in \mathcal{F}$ for every $r \in \mathbf{Q}$. Indeed, it suffices to note that we have, for every $b \in \mathbf{R}$,

$$(-\infty, b] = \bigcap_{\substack{r_n \in \mathbf{Q} \\ r_n > b}} (-\infty, r_n).$$

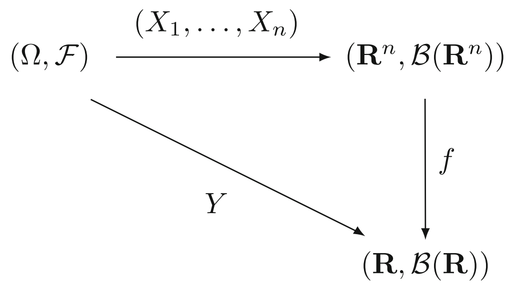
Moreover, we have the following theorem:

Theorem 2.10. *Let X_1, X_2, ..., X_n be random variables. If f is a real-valued, $\mathcal{B}(\mathbf{R}^n)$-measurable function on $\mathbf{R}^n$, then the composite function*

$$Y(\omega) = f(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)), \quad \omega \in \Omega,$$

is $\mathcal{F}$-measurable.

The situation can be visualized in the following diagram:



Proof. (1) First, we show that

$$(X_1, X_2, \ldots, X_n)^{-1}(A) \in \mathcal{F} \quad \text{for all } A \in \mathcal{B}(\mathbf{R}^n). \tag{2.6}$$

To do this, we let

$$\mathcal{B} = \left\{ A \in \mathcal{B}(\mathbf{R}^n) : (X_1, X_2, \ldots, X_n)^{-1}(A) \in \mathcal{F} \right\}.$$

We remark that the Borel σ-algebra $\mathcal{B}(\mathbf{R}^n)$ is generated by the collection $\mathcal{C}$ of all Borel cylinder sets in $\mathbf{R}^n$:

$$\mathcal{C} = \{A_1 \times A_2 \times \cdots \times A_n : A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbf{R})\},$$
$$\sigma(\mathcal{C}) = \mathcal{B}(\mathbf{R}^n).$$

Since X_1, X_2, ..., X_n are $\mathcal{F}$-measurable, it follows that

$$(X_1, X_2, \ldots, X_n)^{-1}(A_1 \times A_2 \times \cdots \times A_n) = \bigcap_{i=1}^{n} X_i^{-1}(A_i) \in \mathcal{F}.$$

This proves that

$$\mathcal{C} \subset \mathcal{B}.$$

However, we find that $\mathcal{B}$ is a σ-algebra. Indeed, it suffices to note that the mapping $(X_1, X_2, \ldots, X_n)^{-1}$ preserves unions, intersections and complements.

Hence it follows that

$$\mathcal{B}(\mathbf{R}^n) = \sigma(\mathcal{C}) \subset \mathcal{B} \subset \mathcal{B}(\mathbf{R}^n),$$

so that

$$\mathcal{B} = \mathcal{B}(\mathbf{R}^n).$$

This proves the desired assertion (2.6).

(2) Secondly, since f is $\mathcal{B}(\mathbf{R}^n)$-measurable, it follows that

$$f^{-1}(B) \in \mathcal{B}(\mathbf{R}^n) \quad \text{for every } B \in \mathcal{B}(\mathbf{R}).$$

Therefore, by applying assertion (2.6) with

$$A := f^{-1}(B),$$

we obtain that

$$Y^{-1}(B) = (X_1, X_2, \ldots, X_n)^{-1}\left(f^{-1}(B)\right) \in \mathcal{F} \quad \text{for every } B \in \mathcal{B}(\mathbf{R}).$$

This proves that Y is $\mathcal{F}$-measurable.

The proof of Theorem 2.10 is complete.

If X_1, X_2, …, X_n are random variables on Ω, then an n-dimensional, vector-valued function X on Ω, defined by the formula

$$X(\omega) = (X_1(\omega), X_2(\omega), \ldots, X_n(\omega)), \quad \omega \in \Omega,$$

is called an *n-dimensional random variable* on Ω. Then we have the following theorem:

Theorem 2.11. *A vector-valued function $X = (X_1, X_2, \ldots, X_n)$ is a random variable on Ω if and only if it satisfies the condition*

$$X^{-1}(A) \in \mathcal{F} \quad \textit{for every } A \in \mathcal{B}(\mathbf{R}^n). \tag{2.7}$$

Proof. (1) The "if" part: For any $B \in \mathcal{B}(\mathbf{R})$, we let

$$A = \mathbf{R} \times \cdots \times \overset{(i)}{\to} B \times \cdots \times \mathbf{R} \in \mathcal{B}(\mathbf{R}^n).$$

Then it follows from condition (2.7) that

$$X_i^{-1}(B) = (X_1, X_2, \ldots, X_n)^{-1}(A) = X^{-1}(A) \in \mathcal{F}.$$

This proves that each X_i is $\mathcal{F}$-measurable.

(3) The "only if" part: If we let

$$\mathcal{B} = \left\{A \in \mathcal{B}(\mathbf{R}^n) : X^{-1}(A) \in \mathcal{F}\right\},$$

then we find that condition (2.7) is equivalent to the following condition:

$$\mathcal{B} = \mathcal{B}(\mathbf{R}^n).$$

To do this, we recall that the Borel σ-algebra $\mathcal{B}(\mathbf{R}^n)$ is generated by the collection $\mathcal{C}$ of all Borel cylinder sets in $\mathbf{R}^n$:

$$\begin{aligned} &\mathcal{C} = \{A_1 \times A_2 \times \cdots \times A_n : A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbf{R})\}, \\ &\sigma(\mathcal{C}) = \mathcal{B}(\mathbf{R}^n). \end{aligned}$$

First, we find that $\mathcal{B}$ is a σ-algebra. Indeed, it suffices to note that the mapping X^{-1} preserves unions, intersections and complements. Moreover, since X_1, X_2, ..., X_n are $\mathcal{F}$-measurable, it follows that

$$\begin{aligned} X^{-1}\left(A_1 \times A_2 \times \cdots \times A_n\right) &= (X_1, X_2, \ldots, X_n)^{-1}\left(A_1 \times A_2 \times \cdots \times A_n\right) \\ &= \bigcap_{i=1}^{n} X_i^{-1}(A_i) \in \mathcal{F}. \end{aligned}$$

This implies that

$$\mathcal{C} \subset \mathcal{B}.$$

Hence we obtain that

$$\mathcal{B}(\mathbf{R}^n) = \sigma(\mathcal{C}) \subset \mathcal{B} \subset \mathcal{B}(\mathbf{R}^n),$$

so that

$$\mathcal{B} = \mathcal{B}(\mathbf{R}^n).$$

Therefore, we have proved that condition (2.7) is satisfied.

The proof of Theorem 2.11 is complete.

If X_1, X_2, ..., X_n are real-valued functions on Ω, we define a σ-algebra $\sigma(X_1, X_2, \ldots, X_n)$ by the formula

$$\sigma\left(X_1, X_2, \ldots, X_n\right) = \left\{(X_1, X_2, \ldots, X_n)^{-1}\left(A\right) : A \in \mathcal{B}(\mathbf{R}^n)\right\}.$$

Then we have the following proposition:

Proposition 2.12. *The σ-algebra $\mathcal{B} = \sigma\left(X_1, X_2, \ldots, X_n\right)$ is the smallest σ-algebra with respect to which all the variables X_1, X_2, ..., X_n are measurable.*

Proof. (1) First, it is easy to see that $\mathcal{B}$ is a σ-algebra. Indeed, it suffices to note that the mapping $(X_1, X_2, \ldots, X_n)^{-1}$ preserves unions, intersections and complements.

(2) Secondly, we show that each X_i is $\mathcal{B}$-measurable. For any $B \in \mathcal{B}(\mathbf{R})$, we let

$$A = \mathbf{R} \times \cdots \times \overset{(i)}{\to} B \times \cdots \times \mathbf{R} \in \mathcal{B}(\mathbf{R}^n).$$

Then it follows that

$$X_i^{-1}(B) = (X_1, X_2, \ldots, X_n)^{-1}(A) \in \sigma(X_1, X_2, \ldots, X_n) = \mathcal{B}.$$

This proves that each X_i is $\mathcal{B}$-measurable.

(3) Finally, we show that $\mathcal{B}$ is the smallest σ-algebra with respect to which all the variables X_1, X_2, ..., X_n are measurable. To do this, we assume that $\tilde{\mathcal{B}}$ is a σ-algebra with respect to which all the variables $X_1, X_2, \ldots, X_n$ are measurable. Let $\mathcal{C}$ be the collection of all Borel cylinder sets in $\mathbf{R}^n$:

$$\mathcal{C} = \{A_1 \times A_2 \times \cdots \times A_n : A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbf{R})\}.$$

We recall that the Borel σ-algebra $\mathcal{B}(\mathbf{R}^n)$ is generated by $\mathcal{C}$:

$$\sigma(\mathcal{C}) = \mathcal{B}(\mathbf{R}^n).$$

Since $X_1, X_2, \ldots, X_n$ are $\tilde{\mathcal{B}}$-measurable, it follows that

$$(X_1, X_2, \ldots, X_n)^{-1}(A_1 \times A_2 \times \cdots \times A_n) = \bigcap_{i=1}^{n} X_i^{-1}(A_i) \in \tilde{\mathcal{B}}.$$

This implies that

$$(X_1, X_2, \ldots, X_n)^{-1}(\mathcal{C}) \subset \tilde{\mathcal{B}}.$$

Hence we have

$$\begin{aligned}\mathcal{B} &= (X_1, X_2, \ldots, X_n)^{-1}(\mathcal{B}(\mathbf{R}^n)) \\ &= (X_1, X_2, \ldots, X_n)^{-1}(\sigma(\mathcal{C})) = \sigma\left((X_1, X_2, \ldots, X_n)^{-1}(\mathcal{C})\right) \\ &\subset \tilde{\mathcal{B}}.\end{aligned}$$

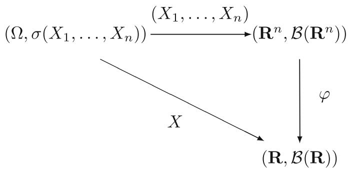
Therefore, we have proved that $\mathcal{B}$ is the smallest σ-algebra with respect to which all the variables $X_1, X_2, \ldots, X_n$ are measurable.

The proof of Proposition 2.12 is complete.

Theorem 2.13. *Assume that a function* $X : \Omega \to \mathbf{R}$ *is* $\sigma(X_1, X_2, \ldots, X_n)$*-measurable. Then we can find a Borel measurable function* $\varphi(x_1, x_2, \ldots, x_n)$ *on* $\mathbf{R}^n$ *such that*

$$X(\omega) = \varphi(X_1(\omega), X_2(\omega), X_n(\omega)) \quad \textit{for every } \omega \in \Omega. \tag{2.8}$$

The situation can be visualized in the following diagram:



Proof. The proof is divided into two steps.

Step 1: Assume that X is a non-negative, $\sigma(X_1, X_2, \ldots, X_n)$-measurable function. Then we let

$$\Lambda_i^{(k)} = X^{-1}\left(\left[\frac{i}{2^k}, \frac{i+1}{2^k}\right)\right) \in \sigma(X_1, X_2, \ldots, X_n), \quad i = 0, 1, \ldots, 2^{2k}.$$

Hence we can find disjoint subsets $A_i^{(k)} \in \mathcal{B}(\mathbf{R}^n)$ such that

$$\Lambda_i^{(k)} = (X_1, X_2, \ldots, X_n)^{-1}\left(A_i^{(k)}\right), \quad i = 0, 1, \ldots, 2^{2k}.$$

Moreover, without loss of generality, we may assume that

$$\begin{aligned}
A_i^{(k)} &= A_{2i}^{(k+1)} + A_{2i+1}^{(k+1)}, \\
\Lambda_{2i}^{(k+1)} &= X^{-1}\left(\left[\frac{2i}{2^{k+1}}, \frac{2i+1}{2^{k+1}}\right)\right) = (X_1, X_2, \ldots, X_n)^{-1}\left(A_{2i}^{(k+1)}\right), \\
\Lambda_{2i+1}^{(k+1)} &= X^{-1}\left(\left[\frac{2i+1}{2^{k+1}}, \frac{2i+2}{2^{k+1}}\right)\right) = (X_1, X_2, \ldots, X_n)^{-1}\left(A_{2i+1}^{(k+1)}\right).
\end{aligned}$$

Now we define simple functions $X^{(k)}(\omega)$ on Ω and $\varphi_k(x_1, x_2, \ldots, x_n)$ on $\mathbf{R}^n$ respectively as follows:

$$X^{(k)}(\omega) = \sum_{i=0}^{2^{2k}} \frac{i}{2^k} \chi_{\Lambda_i^{(k)}}(\omega), \quad \Lambda_i^{(k)} = X^{-1}\left(\left[\frac{i}{2^k}, \frac{i+1}{2^k}\right)\right),$$

and

$$\varphi_k(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{2^{2k}} \frac{i}{2^k} \chi_{A_i^{(k)}}(x_1, x_2, \ldots, x_n),$$

$$\Lambda_i^{(k)} = (X_1, X_2, \ldots, X_n)^{-1}\left(A_i^{(k)}\right), \quad A_i^{(k)} \in \mathcal{B}(\mathbf{R}^n).$$

Here it should be noted that

$$\varphi_k(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)) = \sum_{i=0}^{2^{2k}} \frac{i}{2^k} \chi_{\Lambda_i^{(k)}}(\omega) = X^{(k)}(\omega), \quad \omega \in \Omega.$$

Therefore, we obtain the following two assertions (a) and (b):

(a) The $\varphi_k(x_1, x_2, \ldots, x_n)$ are Borel measurable functions on $\mathbf{R}^n$.
(b) $\varphi_k\,(X_1, X_2, \ldots, X_n) = X^{(k)} \uparrow X$ in Ω.

If we introduce a Borel measurable function $\varphi(x_1, x_2, \ldots, x_n)$ on $\mathbf{R}^n$ by the formula

$$\begin{aligned} &\varphi(x_1, x_2, \ldots, x_n) \\ &= \begin{cases} \sup_k \varphi_k(x_1, x_2, \ldots, x_n) & \text{if } \sup_k \varphi_k(x_1, x_2, \ldots, x_n) < \infty, \\ 0 & \text{if } \sup_k \varphi_k(x_1, x_2, \ldots, x_n) = \infty, \end{cases} \end{aligned}$$

then we have the equality

$$X(\omega) = \varphi(X_1(\omega), X_2(\omega), X_n(\omega)) \quad \text{for every } \omega \in \Omega.$$

This proves the desired equality (2.8) for the non-negative case.

Step 2: If X is a $\sigma(X_1, X_2, \ldots, X_n)$-measurable function, it can be decomposed into the positive and negative parts:

$$X(\omega) = X^+(\omega) - X^-(\omega), \quad \omega \in \Omega,$$

where

$$X^+(\omega) = \max\{X(\omega), 0\},$$

$$X^-(\omega) = \max\{-X(\omega), 0\}.$$

By Step 1, we can find two Borel measurable functions $\varphi^+(x_1, x_2, \ldots, x_n)$ and $\varphi_-(x_1, x_2, \ldots, x_n)$ such that

$$X^+(\omega) = \varphi^+(X_1(\omega), X_2(\omega), X_n(\omega)) \quad \text{for every } \omega \in \Omega,$$
$$X^-(\omega) = \varphi^-(X_1(\omega), X_2(\omega), X_n(\omega)) \quad \text{for every } \omega \in \Omega.$$

If we let

$$\varphi(x_1, x_2, \ldots, x_n) = \varphi^+(x_1, x_2, \ldots, x_n) - \varphi^-(x_1, x_2, \ldots, x_n),$$

then we obtain that the Borel measurable function $\varphi(x_1, x_2, \ldots, x_n)$ satisfies the condition

$$\begin{aligned} X(\omega) &= X^+(\omega) - X^-(\omega) \\ &= \varphi^+(X_1(\omega), X_2(\omega), X_n(\omega)) - \varphi^-(X_1(\omega), X_2(\omega), X_n(\omega)) \\ &= \varphi(X_1(\omega), X_2(\omega), X_n(\omega)) \quad \text{for every } \omega \in \Omega. \end{aligned}$$

This proves the desired equality (2.8) for the general case.

Now the proof of Theorem 2.13 is complete.

Let $\mathcal{B}(\mathbf{R})$ be the σ-algebra of all Borel sets in $\mathbf{R}$. For any $A \in \mathbf{B}(\mathbf{R})$, we let

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\},$$

and define

$$P(X \in A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}).$$

Then it is easy to see that a real-valued function P_X on $\mathcal{B}(\mathbf{R})$, defined by the formula

$$P_X(A) = P(X \in A) \quad \text{for every } A \in \mathcal{B}(\mathbf{R}),$$

is a probability measure on $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$. The measure P_X is called the *distribution* of the random variable X. It should be emphasized that if $X : \Omega \to \mathbf{R}$ is a random variable, then the probability space $(\Omega, \mathcal{F}, P)$ can be transferred to the probability space $(\mathbf{R}, \mathcal{B}(\mathbf{R}), P_X)$.

Moreover, we can introduce a real-valued function $F_X(x)$ on $\mathbf{R}$ by the formula

$$F_X(x) = \mu(-\infty, x] = P(X \leq x) = P(\omega \in \Omega : X(\omega) \leq x) \quad \text{for every } x \in \mathbf{R}.$$

The function $F_X(x)$ is called the *distribution function* of the random variable X. It is easy to verify that the distribution function $F_X(x)$ enjoys the following four properties (F1)–(F4):

(F1) $x \leq y \implies F_X(x) \leq F_X(y)$ (monotonicity).
(F2) $x \downarrow c \implies F_X(x) \downarrow F_X(c)$ (right-continuity).
(F3) $\lim_{x \to +\infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0$.

(F4) $F_X(x)$ has a jump discontinuity of magnitude $\delta > 0$ at $x = a$ if and only if $P(X = a) = \delta$. In particular, $F_X(x)$ is continuous at $x = a$ if and only if $P(X = a) = 0$.

It is a general principle that all properties of random variables which are relevant to probability theory can be formulated in terms of their distributions.

The integral

$$\int_\Omega X(\omega)\, dP$$

is called the *expectation* or *mean* of X, and is denoted by $E(X)$. When we speak of $E(X)$, it is understood that the integral of $|X(\omega)|$ is finite. If $\Lambda \in \mathcal{F}$, we let

$$E(X; \Lambda) = E(X\chi_\Lambda) = \int_\Omega X(\omega)\, \chi_\Lambda(\omega)\, dP = \int_\Lambda X(\omega)\, dP.$$

Some basic properties of the expectations are summarized in the following seven properties (E1)–(E7):

(E1) $E(X)$ exists if and only if $E(|X|)$ exists.
(E2) If either $E(|X|) < \infty$ or $E(|Y|) < \infty$, then we have, for all $a, b \in \mathbf{R}$,

$$E(aX + bY) = aE(X) + bE(Y).$$

(E3) If $X = c$ for some constant c almost everywhere in Ω, then $E(X) = c$.
(E4) If $X = Y$ almost everywhere in Ω, then $E(X) = E(Y)$.
(E5) If $X \leq Y$ almost everywhere in Ω, then $E(X) \leq E(Y)$.
(E6) If $X \geq 0$ almost everywhere in Ω, then $E(X) \geq 0$.
(E7) $|E(X)| \leq E(|X|)$.

We remark that expectations can be computed using P_X or F_X instead of the integral over Ω as follows:

$$E(X) = \int_\Omega X(\omega)\, dP = \int_{\mathbf{R}} x\, dP_X = \int_{-\infty}^{\infty} x\, dF_X(x), \tag{2.9}$$

where the last expression is interpreted as an improper Riemann–Stieltjes integral and the third one is interpreted as a Lebesgue–Stieltjes integral.

The second equality in formula (2.9) is a special case of the following measure-theoretic construction: Let $(\Omega', \mathcal{F}')$ be another measurable space, and assume that a mapping $\phi : \Omega \to \Omega'$ is measurable in the sense that $\phi^{-1}(A') \in \mathcal{F}$ for every $A' \in \mathcal{F}'$. Then the measure P induces an *image measure* P_ϕ on $(\Omega', \mathcal{F}')$ by the formula

$$P_\phi(A') = P(\phi^{-1}(A')) \quad \text{for every } A' \in \mathcal{F}'. \tag{2.10}$$

Indeed, it suffices to note that the mapping ϕ^{-1} preserves unions and intersections.

Then we have the following theorem:

Theorem 2.14. *If X' is a measurable function from $(\Omega', \mathcal{F}')$ into $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, then the composite function $X(\omega) = X'(\phi(\omega))$ is a random variable on $(\Omega, \mathcal{F}, P)$, and we have the equality*

$$E(X) = \int_{\Omega} X(\omega)\, dP = \int_{\Omega'} X'(\omega')\, dP_{\phi}, \tag{2.11}$$

where the existence of either side implies that of the other.

Proof. First, it is clear that $X = X' \circ \phi$ is $\mathcal{F}'$-measurable. The proof of (2.11) is divided into three steps.

Step 1: If X' is a characteristic function of a set $A' \in \mathcal{F}'$, then it follows that X is also a characteristic function of the set $\phi^{-1}(A')$:

$$X = X' \circ \phi = \chi_{A'} \circ \phi = \chi_{\phi^{-1}(A')}.$$

Hence we have, by definition (2.10),

$$E(X) = P(\phi^{-1}(A')) = P_{\phi}(A') = \int_{\Omega'} \chi_{A'}(\omega')\, dP_{\phi} = \int_{\Omega'} X'(\omega')\, dP_{\phi}.$$

This proves the desired equality (2.11) for characteristic functions. The extension to simple functions follows by taking finite linear combinations, since the mapping ϕ^{-1} preserves unions and intersections.

Step 2: If X' is a non-negative, $\mathcal{F}'$-measurable function, then we can find an increasing sequence of simple functions $\{X'_n\}$ which converges to X'. Hence it follows from an application of the monotone convergence theorem [Fo2, Theorem 2.14] that

$$\lim_{n\to\infty} \int_{\Omega'} X'_n(\omega')\, dP_{\phi} = \int_{\Omega'} X'(\omega')\, dP_{\phi}. \tag{2.12}$$

However, we note that the composite functions $X_n = X'_n \circ \phi$ are an increasing sequence of simple functions which converges to X. By applying again the monotone convergence theorem, we obtain that

$$E(X) = \int_{\Omega} X(\omega)\, dP = \lim_{n\to\infty} \int_{\Omega} X_n(\omega)\, dP = \lim_{n\to\infty} E(X_n). \tag{2.13}$$

Since (2.11) holds true for the simple functions X_n and X'_n, it follows from (2.13) and (2.12) that

$$E(X) = \lim_{n\to\infty} E(X_n) = \lim_{n\to\infty} \int_{\Omega'} X'_n(\omega')\, dP_\phi = \int_{\Omega'} X'(\omega')\, dP_\phi.$$

This proves the desired equality (2.11) for non-negative, $\mathcal{F}'$-measurable functions.

Step 3: Finally, the general case of (2.11) follows by applying separately to the positive and negative parts of X':

$$X'(\omega') = X'^{+}(\omega') - X'^{-}(\omega'), \quad \omega' \in \Omega',$$

where

$$X'^{+}(\omega') = \max\{X'(\omega'), 0\},$$
$$X'^{-}(\omega') = \max\{-X'(\omega'), 0\}.$$

Indeed, since we have

$$X'^{+}(\omega') = X^{+}(\phi(\omega)) = \max\{X(\phi(\omega)), 0\},$$
$$X'^{-}(\omega') = X^{-}(\phi(\omega)) = \max\{-X(\phi(\omega)), 0\},$$

we obtain from Step 2 that

$$\begin{aligned}\int_{\Omega'} X'(\omega')\, dP_\phi &= \int_{\Omega'} X'^{+}(\omega')\, dP_\phi - \int_{\Omega'} X'^{-}(\omega')\, dP_\phi \\ &= \int_{\Omega} X^{+}(\omega)\, dP - \int_{\Omega} X^{-}(\omega)\, dP = \int_{\Omega} X(\omega)\, dP \\ &= E(X).\end{aligned}$$

This proves the desired equality (2.11) for general $\mathcal{F}'$-measurable functions.

Moreover, it is easily seen that $E(X)$ exists if and only if X' is integrable with respect to P_ϕ.

The proof of Theorem 2.14 is complete.

Corollary 2.15. *Let $g(x)$ be a Borel measurable function from $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ into $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$. If a vector-valued function $X = (X_1, X_2, \ldots, X_n)$ is a random variable on Ω and if the expectation $E(g(X))$ exists, then we have*

$$\begin{aligned}E(g(X)) &= \int_{\Omega} g(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))\, dP \qquad (2.14) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_n)\, dP_X.\end{aligned}$$

Indeed, Corollary 2.15 follows from an application of Theorem 2.14 with

$$\Omega' := \mathbf{R}^n, \quad \mathcal{F}' = \mathcal{B}(\mathbf{R}^n),$$
$$\phi := X, \quad X' = g, \quad X := g(X).$$

Now we let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces, and let $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ be the Cartesian product of the measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$. We recall that a rectangle is a set of the form $A_1 \times A_2$ where $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$, and further that the collection $\mathcal{A}$ of finite disjoint unions of rectangles forms an algebra. Moreover, we have

$$\sigma(\mathcal{A}) = \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2.$$

If $A \in \mathcal{F}$, then the ω_1-section A_{ω_1} of A is defined by the formula

$$A_{\omega_1} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\}, \quad \omega_1 \in \Omega_1.$$

If X is an $\mathcal{F}$-measurable function on Ω, then the ω_1-section X_{ω_1} of X is defined by the formula

$$X_{\omega_1}(\omega_2) = X(\omega_1, \omega_2), \quad \omega_2 \in \Omega_2.$$

The next theorem will be useful for the study of measurability of functions in Chap. 9:

Theorem 2.16. *Let $P(\omega_1, A_2)$ be a function defined on $\Omega_1 \times \mathcal{F}_2$. Assume that the following two conditions (i) and (ii) are satisfied:*

(i) For each $\omega_1 \in \Omega_1$, $P(\omega_1, \cdot)$ is a probability measure on $(\Omega_2, \mathcal{F}_2)$.
(ii) For each $A_2 \in \mathcal{F}_2$, $P(\cdot, A_2)$ is an $\mathcal{F}_1$-measurable function on Ω_1.

If h is a bounded, $\mathcal{F}$-measurable function on Ω, we let

$$H(\omega_1) = \int_{\Omega_2} h(\omega_1, \omega_2)\, P(\omega_1, d\omega_2).$$

Then H is a bounded, $\mathcal{F}_1$-measurable function on Ω_1.

Proof. The proof is divided into two steps.

Step 1: We prove the boundedness of H on Ω_1.
First, since $h(\omega_1, \cdot)$ is $\mathcal{F}_2$-measurable for all ω_1, it follows that the function H is well-defined. Moreover, we have the inequalit

$$\begin{aligned}|H(\omega_1)| &= \left|\int_{\Omega_2} h(\omega_1,\omega_2)\, P(\omega_1, d\omega_2)\right| \le \int_{\Omega_2} |h(\omega_1,\omega_2)|\, P(\omega_1, d\omega_2) \\ &\le \sup_{\Omega} |h| \int_{\Omega_2} P(\omega_1, d\omega_2) = \sup_{\Omega} |h|\, P(\omega_1, \Omega_2) \\ &= \sup_{\Omega} |h|.\end{aligned}$$

This proves that H is bounded on Ω_1.

Step 2: Secondly, we prove the $\mathcal{F}_1$-measurability of H.

Step 2-1: If $h = \chi_{A_1 \times A_2}$ is a characteristic function with $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$, then it follows that

$$H(\omega_1) = \int_{\Omega_2} \chi_{A_1}(\omega_1)\, \chi_{A_2}(\omega_2)\, P(\omega_1, d\omega_2) = \chi_{A_1}\,(\omega_1\; P(\omega_1, A_2))\,.$$

This proves that H is $\mathcal{F}_1$-measurable, since χ_{A_1} and $P(\cdot, A_2)$ are $\mathcal{F}_1$-measurable.

Step 2-2: Let $\mathcal{A}$ be the collection of finite disjoint unions of rectangles in Ω. If $h = \chi_A$ with $A \in \mathcal{A}$, then it follows that h is a simple function of the form

$$h = \sum_{j=1}^{k} a_j\, \chi_{A_j^{(1)} \times A_j^{(2)}}, \quad A_j^{(1)} \in \mathcal{F}_1,\; A_j^{(2)} \in \mathcal{F}_2.$$

Hence we have

$$H(\omega_1) = \sum_{j=1}^{k} a_j\, \chi_{A_j^{(1)}}(\omega_1\; P(\omega_1, A_j^{(2)})).$$

By Step 2-1, this proves that H is $\mathcal{F}_1$-measurable.

Step 2-3: We let

$$\mathcal{M} = \left\{ A \in \mathcal{F} : \int_{\Omega_2} \chi_A(\omega_1,\omega_2)\, P(\omega_1, d\omega_2) \text{ is } \mathcal{F}_1\text{-measurable} \right\}.$$

By Step 2-2, it follows that

$$\mathcal{A} \subset \mathcal{M}.$$

Moreover, by applying the monotone convergence theorem [Fo2, Theorem 2.14] we obtain that $\mathcal{M}$ is a d-system. Therefore, it follows from an application of the Dynkin class theorem (Corollary 2.5) that

$$\mathcal{F} = \sigma(\mathcal{A}) \subset \mathcal{M} \subset \mathcal{F},$$

so that

$$\mathcal{M} = \mathcal{F}.$$

This proves that the function

$$\int_{\Omega_2} \chi_A(\omega_1, \omega_2)\, P(\omega_1, d\omega_2)$$

is $\mathcal{F}_1$-measurable for every $A \in \mathcal{F}$.

Step 2-4: If h is a general simple function of the form

$$h = \sum_{j=1}^{k} a_j\, \chi_{A_j}, \quad A_j \in \mathcal{F},$$

then it follows that

$$H(\omega_1) = \sum_{j=1}^{k} a_j \int_{\Omega_2} \chi_{A_j}(\omega_1, \omega_2)\, P(\omega_1, d\omega_2).$$

By Step 2-3, this proves that H is $\mathcal{F}_1$-measurable.

Step 2-5: If h is a bounded, $\mathcal{F}$-measurable function on Ω, it can be decomposed into the positive and negative parts:

$$h(\omega_1, \omega_2) = h^+(\omega_1, \omega_2) - h^-(\omega_1, \omega_2), \quad (\omega_1, \omega_2) \in \Omega = \Omega_1 \times \Omega_2,$$

where

$$\begin{aligned} h^+(\omega_1, \omega_2) &= \max\{h(\omega_1, \omega_2), 0\}, \\ h^-(\omega_1, \omega_2) &= \max\{-h(\omega_1, \omega_2), 0\}. \end{aligned}$$

However, we know that the function h^+ is a pointwise limit of an increasing sequence $\{h_n^+\}_{n=1}^{\infty}$ of non-negative, $\mathcal{F}$-measurable simple functions and that the function h^- is a pointwise limit of an increasing sequence $\{h_n^-\}_{n=1}^{\infty}$ of non-negative, $\mathcal{F}$-measurable simple functions. Hence it follows from an application of the monotone convergence theorem [Fo2, Theorem 2.14] that

$$\int_{\Omega_2} h^{\pm}(\omega_1, \omega_2)\, P(\omega_1, d\omega_2) = \lim_{n\to\infty} \int_{\Omega_2} h_n^{\pm}(\omega_1, \omega_2)\, P(\omega_1, d\omega_2).$$

By Step 2-4, we find that the functions

$$\int_{\Omega_2} h^{\pm}(\omega_1, \omega_2)\, P(\omega_1, d\omega_2)$$

are $\mathcal{F}_1$-measurable.

Summing up, we have proved that the function

$$\begin{aligned} H(\omega_1) &= \int_{\Omega_2} h(\omega_1, \omega_2)\, P(\omega_1, d\omega_2) \\ &= \int_{\Omega_2} h^+(\omega_1, \omega_2)\, P(\omega_1, d\omega_2) - \int_{\Omega_2} h^-(\omega_1, \omega_2)\, P(\omega_1, d\omega_2) \end{aligned}$$

is $\mathcal{F}_1$-measurable.

The proof of Theorem 2.16 is complete.

2.4 Independence

One of the most important concepts in probability theory is that of independence. It is the concept of independence more than anything else which gives probability theory a life of its own, distinct from other branches of analysis.

2.4.1 Independent Events

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Two events A and B in $\mathcal{F}$ are said to be *independent* if the following product rule holds true:

$$P(A \cap B) = P(A)P(B).$$

A collection $\{E_1, E_2, \ldots, E_n\}$ of events in $\mathcal{F}$ is said to be *independent* if the product rule holds true for every subcollection of them, that is, if every subcollection $\{E_{i_1}, E_{i_2}, \ldots, E_{i_k}\}$ satisfies the condition

$$P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}) = P(E_{i_1})P(E_{i_2}) \ldots P(E_{i_k}). \tag{2.15}$$

A collection $\mathcal{A} = \{E_i : i \in I\}$ of events in $\mathcal{F}$, where I is a finite or infinite index set, is said to be *independent* if every finite subcollection of $\mathcal{A}$ is independent, that is, if condition (2.15) holds true for all $k \in \mathbf{N}$ and all distinct $i_1, i_2, \ldots, i_k \in I$.

We note that if $\mathcal{A} = \{E_i : i \in I\}$ is an independent class, then so is the class $\mathcal{A}'$ obtained by replacing the E_i in any subclass of $\mathcal{A}$ by either $\emptyset$, Ω or E_i^c. For example, if A and B are independent, then the following product rules hold true:

$$\begin{aligned} &P(A \cap B^c) = P(A)P(B^c), \;\; P(A^c \cap B) = P(A^c)P(B), \\ &P(A^c \cap B^c) = P(A^c)P(B^c). \end{aligned}$$

2.4.2 Independent Random Variables

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A collection $\{X_1, X_2, \ldots, X_n\}$ of random variables on Ω is said to be *independent* if the events $E_1 = X^{-1}(B_1)$, $E_2 = X^{-1}(B_2)$, …, $E_n = X^{-1}(B_n)$ satisfy condition (2.15) for every choice of Borel sets B_1, B_2, …, $B_n \in \mathcal{B}(\mathbf{R})$:

$$P\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \ldots, X_{i_k} \in B_{i_k}\right) = \prod_{j=1}^{k} P\left(X_{i_j} \in B_{i_j}\right). \tag{2.16}$$

A collection $\mathcal{C} = \{X_i : i \in I\}$ of random variables on Ω, where I is a finite or infinite index set, is said to be *independent* if every finite subcollection of $\mathcal{C}$ is independent, that is, if condition (2.16) holds true for all $k \in \mathbf{N}$ and all distinct i_1, i_2, …, $i_k \in I$.

For any finite sequence $\{X_1, X_2, \ldots, X_n\}$ of random variables on Ω, we consider $(X_1, X_2, \ldots, X_n)$ as a map of Ω into $\mathbf{R}^n$

$$(X_1, X_2, \ldots, X_n) : \Omega \longrightarrow \mathbf{R}^n,$$

and define the image measure $P_{(X_1,X_2,\ldots,X_n)}$ on $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ by the formula

$$P_{(X_1,X_2,\ldots,X_n)}(B) = P\left((X_1, X_2, \ldots, X_n)^{-1}(B)\right) \quad \text{for every Borel set } B \in \mathcal{B}(\mathbf{R}^n).$$

The probability measure $P_{(X_1,X_2,\ldots,X_n)}$ on $\mathbf{R}^n$ is called the *joint distribution* of $(X_1, X_2, \ldots, X_n)$.

The next theorem gives a characterization of independent random variables in terms of their joint distributions:

Theorem 2.17. *A collection $\{X_i : i \in I\}$ of random variables on Ω is independent if and only if the joint distribution $P_{(X_{\alpha_1},X_{\alpha_2},\ldots,X_{\alpha_n})}$ of any finite set $\{X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n}\}$ is the product of their individual distributions:*

$$P_{(X_{\alpha_1},X_{\alpha_2},\ldots,X_{\alpha_n})} = \prod_{j=1}^{n} P_{X_{\alpha_j}} \quad \text{on } \mathcal{B}(\mathbf{R}^n). \tag{2.17}$$

Proof. First, we have, for all Borel sets B_1, B_2, …, $B_n \in \mathcal{B}(\mathbf{R})$,

$$\begin{aligned} & P_{(X_{\alpha_1},X_{\alpha_2},\ldots,X_{\alpha_n})}(B_1 \times B_2 \ldots \times B_n) \\ &= P\left((X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})^{-1}(B_1 \times B_2 \ldots \times B_n)\right) \\ &= P\left(X_{\alpha_1}^{-1}(B_1) \cap X_{\alpha_2}^{-1}(B_2) \cap \ldots \cap X_{\alpha_n}^{-1}(B_n)\right). \end{aligned} \tag{2.18}$$

(1) The "only if" part: If we define a measurable rectangle (Borel cylinder set) to be a set of the form

$$B_1 \times B_2 \times \ldots \times B_n, \quad B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathbf{R}),$$

then it follows that the collection $\mathcal{A}$ of finite disjoint unions of rectangles forms an algebra and further that $\mathcal{A}$ generates the σ-algebra $\mathcal{B}(\mathbf{R}^n)$: $\sigma(\mathcal{A}) = \mathcal{B}(\mathbf{R}^n)$.

If $\{X_i\}$ is independent, then it follows from (2.18) that

$$\begin{aligned} & P_{(X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})}(B_1 \times B_2 \ldots \times B_n) \\ &= \prod_{j=1}^{n} P\left(X_{\alpha_j}^{-1}(B_j)\right) = \prod_{j=1}^{n} P_{X_{\alpha_j}}(B_j) \\ &= \left(\prod_{j=1}^{n} P_{X_{\alpha_j}}\right)(B_1 \times B_2 \times \ldots \times B_n). \end{aligned} \tag{2.19}$$

We let

$$\mathcal{M} = \left\{ A \in \mathcal{B}(\mathbf{R}^n) : P_{(X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})}(A) = \prod_{j=1}^{n} P_{X_{\alpha_j}}(A) \right\}.$$

Then we find from (2.19) that $\mathcal{A} \subset \mathcal{M}$. Moreover, it is easy to see that $\mathcal{M}$ is a d-system, since $P_{(X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})}$ and $\prod_{j=1}^{n} P_{X_{\alpha_j}}$ are measures on $\mathcal{B}(\mathbf{R}^n)$. Therefore, by applying the Dynkin class theorem (Corollary 2.5) we obtain that

$$\sigma(\mathcal{A}) \subset \mathcal{M},$$

so that

$$\mathcal{M} = \mathcal{B}(\mathbf{R}^n).$$

This proves the desired equality (2.17).

(2) The "if" part: If (2.17) holds true, then we have, by (2.18) and (2.19),

$$\begin{aligned} & P\left(X_{\alpha_1}^{-1}(B_1) \cap X_{\alpha_2}^{-1}(B_2) \cap \ldots \cap X_{\alpha_n}^{-1}(B_n)\right) \\ &= P\left((X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})^{-1}(B_1 \times B_2 \ldots \times B_n)\right) \\ &= P_{(X_{\alpha_1}, X_{\alpha_2}, \ldots, X_{\alpha_n})}(B_1 \times B_2 \ldots \times B_n) \end{aligned}$$

$$= \left(\prod_{j=1}^{n} P_{X_{\alpha_j}} \right) (B_1 \times B_2 \times \ldots \times B_n)$$

$$= \prod_{j=1}^{n} P_{X_{\alpha_j}}(B_j) = \prod_{j=1}^{n} P\left(X_{\alpha_j}^{-1}(B_j)\right),$$

or equivalently,

$$P\left(X_{\alpha_1} \in B_1, X_{\alpha_2} \in B_2, \ldots, X_{\alpha_n} \in B_n\right) = \prod_{j=1}^{n} P\left(X_{\alpha_j} \in B_j\right).$$

This proves that $\{X_i\}$ is independent.

The proof of Theorem 2.17 is complete.

2.4.3 Independent Algebras

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A collection $\mathfrak{A} = \{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ of subalgebras of $\mathcal{F}$ is said to be *independent* if we have, for any event $E_i \in \mathcal{A}_i$,

$$P(E_1 \cap E_2 \cap \ldots \cap E_n) = P(E_1)P(E_2)\ldots P(E_n). \tag{2.20}$$

A collection $\mathfrak{A} = \{\mathcal{A}_i : i \in I\}$ of subalgebras of $\mathcal{F}$, where I is an infinite index set, is said to be *independent* if every finite subcollection of $\mathfrak{A}$ is independent, that is, if condition (2.20) holds true for all $n \in \mathbf{N}$ and all distinct $i_1, i_2, \ldots, i_n \in I$:

$$P\left(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n}\right) = P(E_{i_1})P(E_{i_2})\ldots P(E_{i_n}), \quad E_{i_k} \in \mathcal{A}_{i_k}.$$

If A is an event of Ω, we define the σ-algebra $\sigma(A)$ as follows:

$$\sigma(A) = \{\emptyset, A, A^c, \Omega\}.$$

Then it is easy to see that a collection $\{A_1, A_2, \ldots, A_n\}$ of events is independent if and only if the collection $\{\sigma(A_1), \sigma(A_2), \ldots, \sigma(A_n)\}$ of σ-algebras is independent.

We recall that a collection $\{X_1, X_2, \ldots, X_n\}$ of random variables on Ω is independent if the events $E_1 = X_1^{-1}(B_1)$, $E_2 = X_2^{-1}(B_2)$, ..., $E_n = X_n^{-1}(B_n)$ satisfy condition (2.20) for every choice of Borel sets $B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathbf{R})$:

$$P(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n) = \prod_{j=1}^{n} P\left(X_j \in B_j\right). \tag{2.21}$$

If X is a random variable on Ω, we define the σ-algebra $\sigma(X)$ by the formula

$$\sigma(X) = \left\{X^{-1}(A) : A \in \mathcal{B}(\mathbf{R})\right\}.$$

Then we have the following theorem:

Theorem 2.18. *Let $\mathcal{C} = \{X_i : i \in I\}$ be a collection of random variables on Ω, where I is a finite or infinite index set. Then the collection $\mathcal{C}$ is independent if and only if the collection $\mathfrak{A} = \{\sigma(X_i) : i \in I\}$ of σ-algebras is independent.*

Proof. The proof is divided into two steps.

Step 1: The "only if" part: Let $\{X_{i_1}, X_{i_2}, \ldots, X_{i_n}\}$ be an arbitrary finite subcollection of $\mathfrak{A}$. If $E_{i_k} = X_{i_k}^{-1}(B_{i_k})$, $B_{i_k} \in \mathcal{B}(\mathbf{R})$, is an element of $\sigma\left(X_{i_k}\right)$, we obtain from (2.21) that

$$\begin{aligned} P\left(\cap_{k=1}^n E_{i_k}\right) &= P\left(\cap_{k=1}^n X_{i_k}^{-1}(B_{i_k})\right) \\ &= P\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \ldots, X_{i_n} \in B_{i_n}\right) \\ &= \prod_{k=1}^n P\left(X_{i_k} \in B_{i_k}\right) = \prod_{k=1}^n P\left(X_{i_k}^{-1}(B_{i_k})\right) \\ &= \prod_{k=1}^n P\left(E_{i_k}\right). \end{aligned}$$

This proves the independence of the collection $\{\sigma(X_{i_1}), \sigma(X_{i_2}), \ldots, \sigma(X_{i_n})\}$ of σ-algebras.

Step 2: The "if" part: Let $\{X_{i_1}, X_{i_2}, \ldots, X_{i_n}\}$ be an arbitrary finite subcollection of $\mathfrak{A}$. For any Borel set $B_{i_k} \in \mathcal{B}(\mathbf{R})$, it follows that

$$E_{i_k} = X_{i_k}^{-1}(B_{i_k}) \in \sigma\left(X_{i_k}\right).$$

Hence we have, by the independence of $\{\sigma(X_{i_1}), \sigma(X_{i_2}), \ldots, \sigma(X_{i_n})\}$,

$$\begin{aligned} &P\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \ldots, X_{i_n} \in B_{i_n}\right) \\ &= P\left(\cap_{k=1}^n X_{i_k}^{-1}(B_{i_k})\right) = P\left(\cap_{k=1}^n E_{i_k}\right) = \prod_{k=1}^n P\left(E_{i_k}\right) \\ &= \prod_{k=1}^n P\left(X_{i_k} \in B_{i_k}\right). \end{aligned}$$

This proves the independence of the collection $\{X_{i_1}, X_{i_2}, \ldots, X_{i_n}\}$ of random variables.

The proof of Theorem 2.18 is complete.

The next theorem asserts that functions of independent random variables are independent:

Theorem 2.19. *Let $\varphi_{i,j} : \mathbf{R}^{k_i} \to \mathbf{R}$ be Borel measurable functions for $1 \le i \le n$, $1 \le j \le \ell_i$, and let $\mathbf{X}_i : \Omega \to \mathbf{R}^{k_i}$ be random variables for $1 \le i \le n$. If the random variables $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ are independent, then the random variables $\mathbf{Y}_i$, defined by the formula*

$$\mathbf{Y}_i = (\varphi_{i,1}(X_i), \varphi_{i,2}(X_i), \ldots, \varphi_{i,\ell_i}(X_i)) \quad \textit{for every } 1 \le i \le n,$$

are independent.

Proof. First, we have, for any set $B \in \mathcal{B}(\mathbf{R})$,

$$\varphi_{i,j}^{-1}(B) \in \mathcal{B}(\mathbf{R}^{k_i}), \quad 1 \le i \le n, \ 1 \le j \le \ell_i,$$
$$\mathbf{Y}_i^{-1}(B) = \mathbf{X}_i^{-1}\left(\varphi_{i,j}^{-1}(B)\right), \quad 1 \le i \le n,$$

and so

$$\sigma(\mathbf{Y}_i) = \left\{\mathbf{Y}_i^{-1}(B) : B \in \mathcal{B}(\mathbf{R})\right\} \subset \sigma(\mathbf{X}_i) = \mathbf{X}_i^{-1}(A) : A \in \mathcal{B}(\mathbf{R}^{k_i}).$$

This proves that $\sigma(\mathbf{Y}_i)$ is a sub-σ-algebra of $\sigma(\mathbf{X}_i)$ for $1 \le i \le n$.

Now we assume that the random variables $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ are independent. Then it follows from an application of Theorem 2.18 that the collection

$$\{\sigma(\mathbf{X}_1), \sigma(\mathbf{X}_2), \ldots, \sigma(\mathbf{X}_n)\}$$

is independent. Hence we find that the collection $\{\sigma(\mathbf{Y}_1), \sigma(\mathbf{Y}_2), \ldots, \sigma(\mathbf{Y}_n)\}$ of σ-algebras is independent, since $\sigma(\mathbf{Y}_i) \subset \sigma(\mathbf{X}_i)$ for $1 \le i \le n$. Therefore, by applying again Theorem 2.18, we obtain that the random variables $\{\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n\}$ are independent.

The proof of Theorem 2.19 is complete.

We give examples of operations which preserve the independence of algebras:

Theorem 2.20. *Let $\mathfrak{A} = \{\mathcal{A}_i : i \in I\}$ be a collection of sub-algebras of $\mathcal{F}$, where I is a finite or infinite index set. If the collection $\mathfrak{A}$ is independent, then the collection $\mathfrak{B} = \{\sigma(\mathcal{A}_i) : i \in I\}$ of σ-algebras is independent. Here $\sigma(\mathcal{A}_i)$ is the σ-algebra generated by the algebra $\mathcal{A}_i$.*

Proof. We have only to prove that if every finite subcollection

$$\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$$

of $\mathfrak{A}$ is independent, then the collection $\{\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \ldots, \sigma(\mathcal{A}_n)\}$ of σ-algebras is independent.

Let Λ_i be an arbitrary element of $\sigma(\mathcal{A}_i)$ with $1 \leq i \leq n$. By Theorem 2.6, for any positive $\varepsilon > 0$ we can find a subset $A_i \in \mathcal{A}_i$ such that

$$P\left(\Lambda_i \bigtriangleup A_i\right) < \frac{\varepsilon}{2n} \quad \text{for every } 1 \leq i \leq n,$$

where

$$A \bigtriangleup B = (A \setminus B) \cup (B \setminus A)$$

is the symmetric difference of A and B. Then we have the inequalities

$$\left|P\left(\cap_{i=1}^n \Lambda_i\right) - P\left(\cap_{i=1}^n A_i\right)\right| \leq P\left(\left(\cap_{i=1}^n \Lambda_i\right) \bigtriangleup \left(\cap_{i=1}^n A_i\right)\right) \leq \sum_{i=1}^n P\left(\Lambda_i \bigtriangleup A_i\right) < \frac{\varepsilon}{2}, \tag{2.22}$$

and

$$\left|\prod_{i=1}^n P\left(\Lambda_i\right) - \prod_{i=1}^n P\left(A_i\right)\right| \leq \sum_{i=1}^n \left|P\left(\Lambda_i\right) - P\left(A_i\right)\right| \leq \sum_{i=1}^n P\left(\Lambda_i \bigtriangleup A_i\right) < \frac{\varepsilon}{2n} \times n = \frac{\varepsilon}{2}. \tag{2.23}$$

However, we have

$$P\left(\cap_{i=1}^n A_i\right) = \prod_{i=1}^n P\left(A_i\right) \quad \text{for all } A_i \in \mathcal{A}_i, \tag{2.24}$$

since the collection $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ is independent.

Therefore, we obtain from inequalities (2.22), (2.23) and formula (2.24) that

$$\begin{aligned}
&\left|P\left(\cap_{k=1}^n \Lambda_i\right) - P\left(\cap_{k=1}^n \Lambda_i\right)\right| \\
&\leq \left|P\left(\cap_{k=1}^n \Lambda_i\right) - P\left(\cap_{k=1}^n A_i\right)\right| + \left|P\left(\cap_{k=1}^n A_i\right) - \prod_{k=1}^n P\left(\Lambda_i\right)\right| \\
&= \left|P\left(\cap_{k=1}^n \Lambda_i\right) - P\left(\cap_{k=1}^n A_i\right)\right| + \left|\prod_{k=1}^n P\left(A_i\right) - \prod_{k=1}^n P\left(\Lambda_i\right)\right| \\
&< \varepsilon.
\end{aligned}$$

Since ε is arbitrary, we have, for each $\Lambda_i \in \sigma(\mathcal{A}_i)$ with $1 \leq i \leq n$,

$$P\left(\cap_{i=1}^{n} \Lambda_i\right) = P\left(\cap_{i=1}^{n} \Lambda_i\right).$$

This proves the independence of the collection $\{\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \ldots, \sigma(\mathcal{A}_n)\}$.

The proof of Theorem 2.20 is complete.

Theorem 2.21. *Let $\mathfrak{A} = \{\mathcal{A}_i : i \in I\}$ be a collection of subalgebras of $\mathcal{F}$, where I is a finite or infinite index set. If the collection $\mathfrak{A}$ is independent and if J_1, J_2, …, J_p are disjoint subsets of I, then the collection*

$$\left\{\sigma\left(\cup_{j\in J_1}\mathcal{A}_j\right), \sigma\left(\cup_{j\in J_2}\mathcal{A}_j\right), \ldots, \sigma\left(\cup_{j\in J_p}\mathcal{A}_j\right)\right\}$$

of σ-algebras is independent.

Proof. We consider the case where

$$J_1 = \{i_1, i_2, \ldots\}.$$

We let

$$\tilde{\mathcal{A}} = \left\{\cap_{j=1}^{k} A_{i_j} : A_{i_j} \in \mathcal{A}_{i_i},\ k = 1, 2, \ldots\right\},$$

and define an algebra $\mathcal{A}$ as follows:

$$\mathcal{A} = \text{the collection of finite unions of members of } \tilde{\mathcal{A}}.$$

Then it is easy to see that

$$\mathcal{A}_j \subset \mathcal{A} \quad \text{for all } j \in J_1,$$
$$\sigma\left(\cup_{j\in J_1}\mathcal{A}_j\right) = \sigma(\mathcal{A}).$$

First, we show that if $\tilde{J} = \{\ell_1, \ell_2, \ldots, \ell_n\}$ such that $\tilde{J} \cap J_1 = \emptyset$, then the collection

$$\{\sigma(\mathcal{A}),\ \mathcal{A}_{\ell_1},\ \mathcal{A}_{\ell_2}, \ldots,\ \mathcal{A}_{\ell_n}\}$$

is independent. By Theorem 2.20, it suffices to show that the collection

$$\{\mathcal{A}, \mathcal{A}_{\ell_1}, \mathcal{A}_{\ell_2}, \ldots, \mathcal{A}_{\ell_n}\}$$

of subalgebras is independent.

We remark that every element of $\mathcal{A}$ can be expressed as a finite disjoint union of members of $\tilde{\mathcal{A}}$. Now let A be an arbitrary element of $\mathcal{A}$ such that

$$A = B_1 + B_2 + \ldots + B_m,$$

where

$$B_i = \cap_{j=1}^{k} A_{i_j} \in \tilde{\mathcal{A}}, \quad A_{i_j} \in \mathcal{A}_{i_j}.$$

Then we have, for each $A_{\ell_k} \in \mathcal{A}_{\ell_k}$ with $\ell_k \in \tilde{J}$,

$$P\left(B_i \cap \cap_{i=1}^{n} A_{\ell_i}\right) = P\left(\cap_{i \in J_1 \cup \tilde{J}} A_i\right) = \prod_{i \in J_1 \cup \tilde{J}} P\left(A_i\right) \tag{2.25}$$

$$= P(B_i) \cdot \prod_{i=1}^{n} P\left(A_{\ell_i}\right),$$

since the collection $\{\mathcal{A}_i : i \in J_1 \cup \tilde{J}\}$ is independent.

Therefore, we obtain from formula (2.25) that

$$\begin{aligned} P\left(A \cap \cap_{i=1}^{n} A_{\ell_i}\right) &= \sum_{j=1}^{m} P\left(B_j \cap \cap_{i=1}^{n} A_{\ell_i}\right) \\ &= \sum_{j=1}^{m} P(B_j) \cdot \prod_{i=1}^{n} P\left(A_{\ell_i}\right) \\ &= P(A) \prod_{i=1}^{n} P\left(A_{\ell_i}\right) \quad \text{for all } A \in \mathcal{A} \text{ and } A_{\ell_k} \in \mathcal{A}_{\ell_k}. \end{aligned}$$

This proves the independence of the collection $\{\mathcal{A}, \mathcal{A}_{\ell_1}, \mathcal{A}_{\ell_2}, \ldots, \mathcal{A}_{\ell_n}\}$.

By repeating this process, we can prove that the collection

$$\left\{\sigma\left(\cup_{j \in J_1} \mathcal{A}_j\right), \sigma\left(\cup_{j \in J_2} \mathcal{A}_j\right), \ldots, \sigma\left(\cup_{j \in J_p} \mathcal{A}_j\right)\right\}$$

is independent.

The proof of Theorem 2.21 is complete.

The next example asserts that vector-valued functions of independent random variables are independent:

Example 2.22. Let $\{X_1, X_2, \ldots, X_n\}$ be a collection of independent random variables on Ω. If we let

$$\begin{aligned} \mathbf{Y}_0 &= (X_1, \ldots, X_{k_1}), \\ \mathbf{Y}_1 &= (X_{k_1+1}, \ldots, X_{k_2+k_2}), \\ &\vdots \\ \mathbf{Y}_\ell &= \left(X_{k_1+\ldots+k_\ell+1}, \ldots, X_n\right), \end{aligned}$$

then the collection $\{\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_n\}$ is independent.

Proof. If we let

$$\mathcal{A}_i = \sigma(X_i) = \left\{X_i^{-1}(A) : A \in \mathcal{B}(\mathbf{R})\right\} \quad \text{for every } 1 \leq i \leq n,$$

it follows from an application of Theorem 2.18 that the collection

$$\{X_1, X_2, \dots, X_n\}$$

of random variables is independent if and only if the collection

$$\{\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n\}$$

of σ-algebras is independent. Moreover, by applying Theorem 2.21 to our situation we find that the collection

$$\left\{\sigma\left(\mathcal{A}_1, \dots, \mathcal{A}_{k_1}\right), \sigma\left(\mathcal{A}_{k_1+1}, \dots, \mathcal{A}_{k_2+k_2}\right), \dots, \sigma\left(\mathcal{A}_{k_1+\dots+k_\ell+1}, \dots, \mathcal{A}_n\right)\right\}$$

of σ-algebras is independent. However, it is easy to see that

$$\begin{aligned}
\sigma(\mathbf{Y}_0) &= \sigma\left(\mathcal{A}_1, \dots, \mathcal{A}_{k_1}\right), \\
\sigma(\mathbf{Y}_1) &= \sigma\left(\mathcal{A}_{k_1+1}, \dots, \mathcal{A}_{k_2+k_2}\right), \\
&\vdots \\
\sigma(\mathbf{Y}_\ell) &= \sigma\left(\mathcal{A}_{k_1+\dots+k_\ell+1}, \dots, \mathcal{A}_n\right).
\end{aligned}$$

Therefore, by applying again Theorem 2.18, we obtain that the collection

$$\{\mathbf{Y}_0, \mathbf{Y}_1, \dots, \mathbf{Y}_n\}$$

is independent.

The proof of Example 2.22 is complete.

2.5 Conditional Probabilities

In this section we introduce conditional probabilities which play a crucial role in the study of Markov processes in Chap. 9. We begin with the definition of a conditional probability:

Definition 2.23. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{F}$ and $\Lambda \in \mathcal{F}$. A random variable Y on Ω is called a version of the *conditional probability* of Λ for given $\mathcal{B}$ if it satisfies the following two conditions (CP1) and (CP2):

(CP1) The function Y is $\mathcal{B}$-measurable.
(CP2) $P(\Lambda \cap \Theta) = E(Y; \Theta)$ for every $\Theta \in \mathcal{B}$. That is, we have, for every $\Theta \in \mathcal{B}$,

$$P(\Lambda \cap \Theta) = \int_{\Theta} Y(\omega)\, dP. \tag{2.26}$$

We shall write

$$Y = P(\Lambda \mid \mathcal{B}).$$

The existence of $P(\Lambda \mid \mathcal{B})$ is based on the Radon–Nikodým theorem [Fo2, Theorem 3.8]. Indeed, by applying the Radon–Nikodým Theorem we can find a non-negative, $\mathcal{B}$-measurable function $f(\omega)$ on Ω such that

$$P(\Lambda \cap \Theta) = \int_{\Theta} f(\omega)\, dP \quad \text{for all } \Theta \in \mathcal{B}.$$

Moreover, if $g(\omega)$ is another $\mathcal{B}$-measurable function on Ω such that

$$P(\Lambda \cap \Theta) = \int_{\Theta} g(\omega)\, dP \quad \text{for all } \Theta \in \mathcal{B},$$

then it follows that $f(\omega) = g(\omega)$ almost everywhere in Ω. That is, the conditional probability $P(\Lambda \mid \mathcal{B})$ can be determined up to a set in $\mathcal{B}$ of measure zero.

Now let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. If we let

$$\Lambda = X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\} = \{X \leq x\} \quad \text{for every } x \in \mathbf{R},$$

and consider the conditional probability

$$P(\Lambda \mid \mathcal{B}) = P(X \leq x \mid \mathcal{B}),$$

then we have the following lemma:

Lemma 2.24. *The conditional probabilities*

$$P(X \leq x \mid \mathcal{B}), \quad x \in \mathbf{R},$$

enjoy the following three properties (CD1)–(CD3):

(CD1) $x < y \implies P(X \leq x \mid \mathcal{B})(\omega) \leq P(X \leq y \mid \mathcal{B})(\omega)$ *for almost all* $\omega \in \Omega$ *(monotonicity).*
(CD2) $\lim_{k \to +\infty} P(X \leq k \mid \mathcal{B})(\omega) = 1$ *for almost all* $\omega \in \Omega$.
(CD3) $\lim_{k \to -\infty} P(X \leq k \mid \mathcal{B})(\omega) = 0$ *for almost all* $\omega \in \Omega$.

Proof. (i) The proof of property (CD1): If we let

$$\Theta = \{\omega \in \Omega : P(X \leq x \mid \mathcal{B})(\omega) > P(X \leq y \mid \mathcal{B})(\omega)\},$$

it follows that $\Theta \in \mathcal{B}$. Assume, to the contrary, that

$$P(\Theta) > 0.$$

Then we obtain from condition (2.26) that

$$\begin{aligned} P(\{X \leq x\} \cap \Theta) &= \int_{\Theta} P(X \leq x \mid \mathcal{B})(\omega)\, dP \\ &> \int_{\Theta} P(X \leq y \mid \mathcal{B})(\omega)\, dP = P(\{X \leq y\} \cap \Theta), \end{aligned}$$

so that

$$P(\{X \leq x\} \cap \Theta) > P(\{X \leq y\} \cap \Theta).$$

This is a contradiction, since we have

$$x < y \implies \{X \leq x\} \cap \Theta \subset \{X \leq y\} \cap \Theta.$$

(ii) The proof of property (CD2): By property (CD1), it follows that the sequence $\{P(X \leq k \mid \mathcal{B})(\omega)\}$ is increasing with respect to k, for almost all $\omega \in \Omega$. Hence we obtain that the limit

$$\lim_{k \to \infty} P(X \leq k \mid \mathcal{B})(\omega)$$

exists for almost all $\omega \in \Omega$. For any given $\varepsilon > 0$, we let

$$\Theta_\varepsilon = \left\{\omega \in \Omega : \lim_{k \to \infty} P(X \leq x \mid \mathcal{B})(\omega) \leq 1 - \varepsilon\right\}.$$

Then it follows that $\Theta_\varepsilon \in \mathcal{B}$, and further from condition (2.26) that

$$\begin{aligned} P(\Theta_\varepsilon) &= \lim_{k \to \infty} P(\{X \leq k\} \cap \Theta_\varepsilon) = \lim_{k \to \infty} E(P(X \leq k \mid \mathcal{B}); \Theta_\varepsilon) \\ &= \int_{\Theta_\varepsilon} \lim_{k \to \infty} P(X \leq k \mid \mathcal{B})(\omega)\, dP \leq (1 - \varepsilon) P(\Theta_\varepsilon). \end{aligned}$$

This proves that $P(\Theta_\varepsilon) = 0$.

Since $\varepsilon > 0$ is arbitrary, we find that

$$\lim_{k \to \infty} P(X \leq k \mid \mathcal{B})(\omega) = 1$$

for almost all $\omega \in \Omega$.

(iii) The proof of property (CD3): Similarly, by letting

$$\Lambda_\varepsilon = \left\{ \omega \in \Omega : \lim_{k \to -\infty} P(X \leq x \mid \mathcal{B})(\omega) \geq \varepsilon \right\},$$

we can prove that

$$\lim_{k \to -\infty} P(X \leq k \mid \mathcal{B})(\omega) = 0$$

for almost all $\omega \in \Omega$.

The proof of Lemma 2.24 is complete.

Moreover, we can prove the following theorem:

Theorem 2.25. *Let X be a random variable on $(\Omega, \mathcal{F}, P)$. If $\mathcal{B}$ is a sub-σ-algebra of $\mathcal{F}$, then there exists a function $\mu_{\mathcal{B}}$ on $\Omega \times \mathcal{B}(\mathbf{R})$ which satisfies the following two conditions (1) and (2):*

(1) $\mu_{\mathcal{B}}(\omega, \cdot)$ is a probability measure on $\mathcal{B}(\mathbf{R})$ for almost all $\omega \in \Omega$.
(2) For every $A \in \mathcal{B}(\mathbf{R})$, $\mu_{\mathcal{B}}(\cdot, A)$ is a version of the conditional probability $P(X \in A \mid \mathcal{B})$. In particular, we have

$$P\left((X \in A) \cap \Theta\right) = \int_\Theta \mu_{\mathcal{B}}(\omega, A)\, dP \quad \textit{for all } \Theta \in \mathcal{B}.$$

Moreover, the function $\mu_{\mathcal{B}}$ is uniquely determined in the sense that any two such functions are equal with respect to P. More precisely, if a function $\tilde{\mu}_{\mathcal{B}}$ on $\Omega \times \mathcal{B}(\mathbf{R})$ satisfies conditions (1) and (2), then we have, for almost all $\omega \in \Omega$,

$$\tilde{\mu}_{\mathcal{B}}(\omega, A) = \mu_{\mathcal{B}}(\omega, A), \quad A \in \mathcal{B}(\mathbf{R}). \tag{2.27}$$

Proof. The proof is divided into four steps.

Step 1: First, we prove the uniqueness of the function $\mu_{\mathcal{B}}$, that is, we prove equality (2.27) for almost all $\omega \in \Omega$.

To do this, we let

$$\mathcal{M} = \{A \in \mathcal{B}(\mathbf{R}) : \tilde{\mu}_{\mathcal{B}}(\omega, A) = \mu_{\mathcal{B}}(\omega, A) \quad \text{for almost all } \omega \in \Omega\}.$$

It suffices to show that $\mathcal{M} = \mathcal{B}(\mathbf{R})$.

If we let

$$\mathcal{E} = \text{the collection of sets of the form } (x, y] \text{ or } (x, \infty) \text{ or } \emptyset \tag{2.28}$$
$$\text{where } -\infty \leq x < y < \infty,$$

then it is easy to see that $\mathcal{E}$ is an elementary family. Hence we find that the collection $\mathcal{A}$ of finite disjoint unions of members in $\mathcal{E}$ is an algebra, and further that the σ-algebra $\sigma(\mathcal{A})$ generated by $\mathcal{A}$ is $\mathcal{B}(\mathbf{R})$.
For every rational number $r \in \mathbf{Q}$, it follows that $(-\infty, r] \in \mathcal{B}(\mathbf{R})$. Hence we have, for almost all $\omega \in \Omega$,

$$\tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, r]) = \mu_{\mathcal{B}}(\omega, (-\infty, r]). \tag{2.29}$$

By the right-continuity of measures, we obtain from (2.29) that, for all $x \in \mathbf{R}$,

$$\begin{aligned}\tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, x]) &= \lim_{\substack{r_n \in \mathbf{Q} \\ r_n \downarrow x}} \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, r_n]) = \lim_{\substack{r_n \in \mathbf{Q} \\ r_n \downarrow x}} \mu_{\mathcal{B}}(\omega, (-\infty, r_n] \\ &= \mu_{\mathcal{B}}(\omega, (-\infty, x]).\end{aligned} \tag{2.30}$$

Moreover, we have, for all $(x, \infty) = \mathbf{R} \setminus (-\infty, x]$ with $x \in \mathbf{R}$,

$$\begin{aligned}\tilde{\mu}_{\mathcal{B}}(\omega, (x, \infty)) = \tilde{\mu}_{\mathcal{B}}(\omega, \mathbf{R} \setminus (-\infty, x]) &= \tilde{\mu}_{\mathcal{B}}(\omega, \mathbf{R}) - \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, x]) \\ = 1 - \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, x]) &= 1 - \mu_{\mathcal{B}}(\omega, (-\infty, x]) \\ = \mu_{\mathcal{B}}(\omega, \mathbf{R} \setminus (-\infty, x]) &= \tilde{\mu}_{\mathcal{B}}(\omega, (x, \infty)),\end{aligned} \tag{2.31}$$

and, for all $(x, y] = (-\infty, y] \setminus (-\infty, x]$ with $y > x$,

$$\begin{aligned}\tilde{\mu}_{\mathcal{B}}(\omega, (x, y]) &= \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, y] \setminus (-\infty, x]) \\ &= \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, y]) - \tilde{\mu}_{\mathcal{B}}(\omega, (-\infty, x]) \\ &= \mu_{\mathcal{B}}(\omega, (-\infty, y]) - \mu_{\mathcal{B}}(\omega, (-\infty, x]) \\ &= \mu_{\mathcal{B}}(\omega, (x, y]).\end{aligned} \tag{2.32}$$

Hence, we obtain from (2.30) to (2.32) that $\mathcal{A} \subset \mathcal{M}$.
However, it is easy to see that $\mathcal{M}$ is a d-system, since $\tilde{\mu}_{\mathcal{B}}$ and $\mu_{\mathcal{B}}$ are probability measures on $\mathcal{B}(\mathbf{R})$. Therefore, by applying the Dynkin class theorem (Corollary 2.5) we obtain that

$$\mathcal{B}(\mathbf{R}) = \sigma(\mathcal{A}) \subset \mathcal{M} \subset \mathcal{B}(\mathbf{R}),$$

so that

$$\mathcal{M} = \mathcal{B}(\mathbf{R}).$$

This proves the desired equality (2.27).

Step 2: We prove the existence of the function $\mu_{\mathcal{B}}$. The proof is divided into two steps.

Step 2-1: For every $r \in \mathbf{Q}$, we define a $\mathcal{B}$-measurable function $G_{\mathcal{B}}(r)$ by the formula

$$G_{\mathcal{B}}(r)(\omega) = P\left(X \leq r \mid \mathcal{B}\right)(\omega) \quad \text{for almost all } \omega \in \Omega.$$

For each integer $n \in \mathbf{N}$, we let

$$\Omega_n = \left\{\omega \in \Omega : G_{\mathcal{B}}\left(\frac{k}{n}\right)(\omega) \leq G_{\mathcal{B}}\left(\frac{k+1}{n}\right)(\omega) \quad \text{for all } k \in \mathbf{Z}, \right.$$
$$\left. \lim_{k\to\infty} G_{\mathcal{B}}\left(\frac{k}{n}\right)(\omega) = 1, \lim_{k\to-\infty} G_{\mathcal{B}}\left(\frac{k}{n}\right)(\omega) = 0\right\}.$$

Then it follows from an application of Lemma 2.24 that

$$\begin{cases} \Omega_n \in \mathcal{B} & \text{for all } n \in \mathbf{N}, \\ P\left(\Omega_n\right) = 1 & \text{for all } n \in \mathbf{N}, \end{cases}$$

so that

$$\begin{cases} \tilde{\Omega} = \bigcap_{n=1}^{\infty} \Omega_n \in \mathcal{B}, \\ P(\tilde{\Omega}) = 1. \end{cases}$$

Hence we can define a function $F_{\mathcal{B}}$ on $\Omega \times \mathbf{R}$ as follows:

$$F_{\mathcal{B}}(\omega, x) = \begin{cases} \lim_{\substack{r\in\mathbf{Q} \\ r\downarrow x}} G_{\mathcal{B}}(r)(\omega) & \text{if } \omega \in \tilde{\Omega}, \\ 0 & \text{if } \omega \notin \tilde{\Omega}. \end{cases}$$

Then it is easy to see that $F_{\mathcal{B}}(\omega, \cdot)$ is a distribution function for each $\omega \in \tilde{\Omega}$ and that $F_{\mathcal{B}}(\cdot, x)$ is $\mathcal{B}$-measurable for each $x \in \mathbf{R}$. Moreover, it follows from an application of the monotone convergence theorem [Fo2, Theorem 2.14] that

$$\begin{aligned} P(\{X \leq x\} \cap \Theta) &= \lim_{\substack{r\in\mathbf{Q} \\ r\downarrow x}} P(\{X \leq r\} \cap \Theta) = \lim_{\substack{r\in\mathbf{Q} \\ r\downarrow x}} \int_{\Theta} G_{\mathcal{B}}(r)(\omega)\, dP \\ &= \int_{\Theta} F_{\mathcal{B}}(\omega, x)\, dP \quad \text{for every } \Theta \in \mathcal{B}. \end{aligned}$$

This proves that $F_{\mathcal{B}}(\cdot, x)$ is a version of the conditional probability $P(X \leq x \mid \mathcal{B})$.

Step 2-2: We note that the distribution function $F_{\mathcal{B}}(\omega, \cdot)$ determines a probability measure $\mu_{\mathcal{B}}(\omega)$ on $\mathbf{R}$ for each $\omega \in \tilde{\Omega}$. In particular, we have

$$F_{\mathcal{B}}(\omega, x) = \mu_{\mathcal{B}}(\omega)\left((-\infty, x]\right), \quad x \in \mathbf{R}.$$

Therefore, for every $A \in \mathcal{B}(\mathbf{R})$ we can define a function $\mu_{\mathcal{B}}(\cdot, A)$ on Ω by the formula

$$\mu_{\mathcal{B}}(\omega, A) = \begin{cases} \mu_{\mathcal{B}}(\omega)(A) & \text{if } \omega \in \tilde{\Omega}, \\ 0 & \text{if } \omega \notin \tilde{\Omega}. \end{cases} \tag{2.33}$$

We show that the function $\mu_{\mathcal{B}}(\cdot, A)$ is $\mathcal{B}$-measurable for every $A \in \mathcal{B}(\mathbf{R})$. To do this, we let

$$\mathcal{L} = \{A \in \mathcal{B}(\mathbf{R}) : \ \mu_{\mathcal{B}}(\cdot, A) \text{ is } \mathcal{B}\text{-measurable}\}.$$

It suffices to show that $\mathcal{L} = \mathcal{B}(\mathbf{R})$.

(a) First, it follows that $(-\infty, y] \in \mathcal{L}$ for all $y \in \mathbf{R}$, since the function

$$\mu_{\mathcal{B}}(\omega, (-\infty, y]) = \begin{cases} \mu_{\mathcal{B}}(\omega)\left((-\infty, y]\right) = F_{\mathcal{B}}(\omega, y) & \text{if } \omega \in \tilde{\Omega}, \\ 0 & \text{if } \omega \notin \tilde{\Omega} \end{cases}$$

is $\mathcal{B}$-measurable.

(b) Secondly, it follows that $(x, y] \in \mathcal{L}$ for all $y > x$, since we have, for $(x, y] = (-\infty, y] \setminus (-\infty, x]$,

$$\mu_{\mathcal{B}}(\omega, (x, y]) = \mu_{\mathcal{B}}(\omega, (-\infty, y]) - \mu_{\mathcal{B}}(\omega, (-\infty, x]).$$

(c) Thirdly, it follows that $(x, \infty) \in \mathcal{L}$ for all $x \in \mathbf{R}$, since we have

$$\mu_{\mathcal{B}}(\omega, (x, \infty)) = \sum_{j=0}^{\infty} \mu_{\mathcal{B}}(\omega, (x + j, x + j + 1]).$$

Therefore, we find that the elementary family $\mathcal{E}$ defined by formula (2.28) is contained in $\mathcal{L}$ and further that the collection $\mathcal{A}$ of finite disjoint unions of members in $\mathcal{E}$ is an algebra contained in $\mathcal{L}$.

(d) However, it is easy to see that $\mathcal{L}$ is a d-system, since $\mu_{\mathcal{B}}(\omega, \cdot)$ are probability measures on $\mathcal{B}(\mathbf{R})$. Therefore, by applying the Dynkin class theorem (Corollary 2.5) we obtain that

$$\mathcal{B}(\mathbf{R}) = \sigma(\mathcal{A}) \subset \mathcal{L} \subset \mathcal{B}(\mathbf{R}),$$

so that

$$\mathcal{L} = \mathcal{B}(\mathbf{R}).$$

Step 2-3: Finally, we prove that $\mu_{\mathcal{B}}(\cdot, A)$ is a version of the conditional probability $P(X \in A \mid \mathcal{B})$. It remains to show that we have, for every $\Theta \in \mathcal{B}$,

$$P((X \in A) \cap \Theta) = \int_{\Theta} \mu_{\mathcal{B}}(\omega, A)\, dP. \tag{2.34}$$

To do this, we let

$$\mathcal{N} = \{A \in \mathcal{B}(\mathbf{R}) : \text{ Equation (2.34) holds true for all } \Theta \in \mathcal{B}\}.$$

It suffices to show that $\mathcal{N} = \mathcal{B}(\mathbf{R})$.
By arguing just as in Step 2-1, we obtain that $\mathcal{E} \subset \mathcal{A} \subset \mathcal{N}$ and further that $\mathcal{N}$ is a d-system. Indeed, if $\{A_n\}_{n=1}^{\infty}$ is an increasing sequence of members of $\mathcal{N}$, then it follows from an application of the monotone convergence theorem [Fo2, Theorem 2.14] that

$$\begin{aligned} P\left((X \in \cup_{n=1}^{\infty} A_n) \cap \Theta\right) &= \lim_{n\to\infty} P\left((X \in A_n) \cap \Theta\right) = \lim_{n\to\infty} \int_{\Theta} \mu(\omega, A_n)\, dP \\ &= \int_{\Theta} \mu_{\mathcal{B}}(\omega, \cup_{n=1}^{\infty} A_n)\, dP. \end{aligned}$$

This proves that the union $\cup_{n=1}^{\infty} A_n$ belongs to $\mathcal{N}$. Therefore, by applying the Dynkin class theorem (Corollary 2.5) we obtain that

$$\mathcal{B}(\mathbf{R}) = \sigma(\mathcal{A}) \subset \mathcal{N} \subset \mathcal{B}(\mathbf{R}),$$

so that

$$\mathcal{N} = \mathcal{B}(\mathbf{R}).$$

Summing up, we have proved that $\mu_{\mathcal{B}}(\omega, A) = P(X \in A \mid \mathcal{B})(\omega)$ for every $A \in \mathcal{B}(\mathbf{R})$.

Now the proof of Theorem 2.25 is complete.

Definition 2.26. The function $\mu_{\mathcal{B}}$ on $\Omega \times \mathcal{B}(\mathbf{R})$ is called a version of the *conditional probability of X with respect to $\mathcal{B}$*, and will be denoted by $P(X \in \cdot \mid \mathcal{B})$. If $\mathcal{B}$ is generated by the random variables X_1, X_2, ..., X_n, that is, if $\mathcal{B} = \sigma(X_1, X_2, \ldots, X_n)$, then $P(X \in \cdot \mid \mathcal{B})$ will be denoted as follows:

$$P\,(X \in \cdot \mid X_1, X_2, \ldots, X_n) = P(X \in \cdot \mid \sigma(X_1, X_2, \ldots, X_n)).$$

The next theorem is an $\mathbf{R}^n$-version of Theorem 2.25:

Theorem 2.27. *Let X be a random variable and let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a vector-valued random variable on the probability space $(\Omega, \mathcal{F}, P)$. Then there exists a function $\psi(x, A)$ on $\mathbf{R}^n \times \mathcal{B}(\mathbf{R})$ which satisfies the following two conditions (i) and (ii):*

(i) $\psi(x, \cdot)$ *is a probability measure on* $\mathcal{B}(\mathbf{R})$ *for* $\mu_{\mathbf{X}}$*-almost all* $x \in \mathbf{R}^n$*. Here* $\mu_{\mathbf{X}}$ *is the joint distribution of* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$.

(ii) For every $A \in \mathcal{B}(\mathbf{R})$, $\psi(\cdot, A)$ *is a Borel measurable function on* $\mathbf{R}^n$ *and we have, for almost all* $\omega \in \Omega$,

$$\begin{aligned} &\psi(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), \cdot) \\ &= P\left(X \in \cdot \mid X_1, X_2, \ldots, X_n\right)(\omega). \end{aligned} \tag{2.35}$$

Moreover, the function ψ *is uniquely determined in the sense that any two of them are equal with respect to* $\mu_{\mathbf{X}}$*. More precisely, if a function* $\tilde{\psi}$ *on* $\mathbf{R}^n \times \mathcal{B}(\mathbf{R})$ *satisfies conditions (i) and (ii), then we have, for* $\mu_{\mathbf{X}}$*-almost all* $x \in \mathbf{R}^n$,

$$\tilde{\psi}(x, A) = \psi(x, A), \quad A \in \mathcal{B}(\mathbf{R}). \tag{2.36}$$

Proof. The proof is divided into three steps.

Step 1: First, we construct a function ψ by using the function $\mu_{\mathcal{B}}$ in Theorem 2.25.

Since $\mathcal{B} = \sigma(X_1, X_2, \ldots, X_n)$ and $\mu_{\mathcal{B}}(\cdot, A)$ is $\mathcal{B}$-measurable for every $A \in \mathcal{B}(\mathbf{R})$, it follows from an application of Theorem 2.13 that there exists a Borel measurable function $\Phi(\cdot, A)$ on $\mathbf{R}^n$ such that

$$\begin{aligned} &\Phi\left(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), A\right) \\ &= \mu_{\mathcal{B}}(\omega, A) = \begin{cases} \mu_{\mathcal{B}}(\omega)(A) & \text{if } \omega \in \tilde{\Omega}, \\ 0 & \text{if } \omega \notin \tilde{\Omega}. \end{cases} \end{aligned} \tag{2.37}$$

Here we recall that $\mu_{\mathcal{B}}(\omega)$ is a probability measure on $\mathbf{R}$ for every $\omega \in \tilde{\Omega}$.

If we let

$$G(x, y) = \Phi(x, (-\infty, y]), \quad x \in \mathbf{R}^n, \ y \in \mathbf{R},$$

it follows that the function $G(x, y) = \Phi(x, (-\infty, y])$ is a distribution function of y for every $x \in \mathbf{X}(\tilde{\Omega})$. Hence, if we let

$$\begin{aligned} \Gamma = \Big\{ &x \in \mathbf{R}^n : G(x, r) \leq G(x, r') \quad \text{for all } r < r' \text{ with } r, r' \in \mathbf{Q}, \\ &\lim_{\substack{r \in \mathbf{Q} \\ r \to -\infty}} G(x, r) = 0, \ \lim_{\substack{r' \in \mathbf{Q} \\ r' \to \infty}} G(x, r') = 1 \Big\}, \end{aligned}$$

then we have

$$\Gamma \in \mathcal{B}(\mathbf{R}^n),$$
$$\mathbf{X}(\tilde{\Omega}) \subset \Gamma.$$

If we define a function $F(x, y)$ on $\mathbf{R}^n \times \mathbf{R}$ by the formula

$$F(x, y) = \begin{cases} \lim_{\substack{r \in \mathbf{Q} \\ r \downarrow y}} G(x, r) & \text{if } x \in \Gamma, \\ 0 & \text{if } x \notin \Gamma, \end{cases}$$

we have the following four assertions (a)–(d):

(a) For each $x \in \mathbf{R}^n$, the function

$$y \longmapsto F(x, y)$$

is right-continuous on $\mathbf{R}$.

(b) For each $y \in \mathbf{R}$, the function

$$x \longmapsto F(x, y)$$

is Borel measurable on $\mathbf{R}^n$.

(c) For each $x \in \Gamma$, the function $F(x, y)$ is a distribution function of y, and $F(x, y) = G(x, y)$ for all $x \in \mathbf{X}(\tilde{\Omega})$.

(d) For each $x \in \Gamma$, the function $F(x, \cdot)$ determines a probability measure $\psi(x, \cdot)$ on $\mathbf{R}$. In particular, we have, for all $x \in \mathbf{X}(\tilde{\Omega})$,

$$\psi(x, \cdot) = \Phi(x, \cdot).$$

Moreover, if we define a function $\psi(x, A)$ by the formula

$$\psi(x, A) = \begin{cases} \Phi(x, A) & \text{if } x \in \Gamma, \\ 0 & \text{if } x \notin \Gamma, \end{cases}$$

then we have

$$\psi(\mathbf{X}(\omega), A) = \Phi(\mathbf{X}(\omega), A) \quad \text{for all } \omega \in \Omega.$$

Indeed, it suffices to note that

$$\psi(x, A) = \Phi(x, A) = 0 \quad \text{for all } x \in \mathbf{X}(\tilde{\Omega}^c) \subset \Gamma^c.$$

Therefore, we obtain from formula (2.37) that, for almost all $\omega \in \Omega$,

$$\begin{aligned}\psi(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), A) &= \Phi(\mathbf{X}(\omega), A) = \mu_{\mathcal{B}}(\omega, A) \\ &= P\left(X \in A \mid X_1, X_2, \ldots, X_n\right)(\omega).\end{aligned}$$

This proves the desired equality (2.35).

Step 2: Secondly, we show that $\psi(x, A)$ is a Borel measurable function of $x \in \mathbf{R}^n$ for every $A \in \mathcal{B}(\mathbf{R})$. To do this, we let

$$\mathcal{M} = \{A \in \mathcal{B}(\mathbf{R}) : \psi(\cdot, A) \text{ is Borel measurable on } \mathbf{R}^n\}.$$

Then, by arguing just as in the proof of Theorem 2.25 we can prove that

$$\mathcal{M} = \mathcal{B}(\mathbf{R}).$$

Moreover, since $\psi(x, \cdot)$ is a probability measure on $\mathbf{R}$ for each $x \in \Gamma$ and since $\mathbf{X}(\tilde{\Omega}) \subset \Gamma$, it follows that

$$1 = P(\tilde{\Omega}) \leq P\left(\mathbf{X}^{-1}(\Gamma)\right) = \mu_{\mathbf{X}}(\Gamma) \leq 1.$$

Therefore, we have proved that $\psi(x, \cdot)$ is a probability measure on $\mathcal{B}(\mathbf{R})$ for $\mu_{\mathbf{X}}$-almost all $x \in \mathbf{R}^n$.

Step 3: Finally, we prove the uniqueness of the function ψ.

Assume that we have, for every $A \in \mathcal{B}(\mathbf{R})$,

$$\begin{aligned}\psi(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), A) &= P\left(X \in A \mid X_1, X_2, \ldots, X_n\right)(\omega) \\ &= \tilde{\psi}(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), A).\end{aligned} \tag{2.38}$$

If we let

$$\begin{aligned}F(x, y) &= \psi\left(x, (-\infty, y]\right), \quad x \in \mathbf{R}^n, \ y \in \mathbf{R}, \\ \tilde{F}(x, y) &= \tilde{\psi}\left(x, (-\infty, y]\right), \quad x \in \mathbf{R}^n, \ y \in \mathbf{R},\end{aligned}$$

then we obtain that the functions $F(x, y)$ and $\tilde{F}(x, y)$ are distribution functions of y for $\mu_{\mathbf{X}}$-almost all $x \in \mathbf{R}^n$.

For each $y \in \mathbf{Q}$, we let

$$B_1 = \left\{x \in \mathbf{R}^n : F(x, y) > \tilde{F}(x, y)\right\}.$$

Then it follows that $B_1 \in \mathcal{B}(\mathbf{R}^n)$. Moreover, we have, by assertion (2.38) with $x := \mathbf{X}(\omega)$ and $A := (-\infty, y]$,

$$\begin{aligned}\int_{B_1} \tilde{F}(x,y)\,d\mu_{\mathbf{X}}(x) &= \int_{(\mathbf{X}\in B_1)} \tilde{\psi}(\mathbf{X}(\omega),(-\infty,y])\,dP\\ &= P\left((X\in(-\infty,y])\cap(\mathbf{X}\in B_1)\right)\\ &= \int_{(\mathbf{X}\in B_1)} \psi(\mathbf{X}(\omega),(-\infty,y])\,dP = \int_{B_1} F(x,y)\,d\mu_{\mathbf{X}}(x),\end{aligned}$$

so that

$$0 = \int_{B_1} \left(F(x,y) - \tilde{F}(x,y)\right) d\mu_{\mathbf{X}}(x).$$

This proves that

$$\mu_{\mathbf{X}}(B_1) = 0,$$

since the integrand is positive on B_1.

Similarly, if we let

$$B_2 = \left\{x \in \mathbf{R}^n : \tilde{F}(x,y) > F(x,y)\right\},$$

it follows that

$$\mu_{\mathbf{X}}(B_2) = 0.$$

Therefore, we obtain that

$$\mu_{\mathbf{X}}\left(\left\{x \in \mathbf{R}^n : F(x,y) \neq \tilde{F}(x,y)\right\}\right) = \mu_{\mathbf{X}}(B_1 \cup B_2) = 0 \quad \text{for all } y \in \mathbf{Q}.$$

Hence we have

$$\begin{aligned}&\mu_{\mathbf{X}}\left(\cup_{r\in\mathbf{Q}}\left\{x \in \mathbf{R}^n : F(x,r) \neq \tilde{F}(x,r)\right\}\right)\\ &\leq \sum_{r\in\mathbf{Q}} \mu_{\mathbf{X}}\left(\left\{x \in \mathbf{R}^n : F(x,r) \neq \tilde{F}(x,r)\right\}\right) = 0.\end{aligned}$$

Namely, we have

$$\begin{aligned}&\mu_{\mathbf{X}}\left(\left\{x \in \mathbf{R}^n : F(x,r) = \tilde{F}(x,r) \quad \text{for all } r \in \mathbf{Q}\right\}\right)\\ &= \mu_{\mathbf{X}}\left(\cap_{r\in\mathbf{Q}}\left\{x \in \mathbf{R}^n : F(x,r) = \tilde{F}(x,r)\right\}\right)\\ &= \mu_{\mathbf{X}}\left(\left(\cup_{r\in\mathbf{Q}}\left\{x \in \mathbf{R}^n : F(x,r) \neq \tilde{F}(x,r)\right\}\right)^c\right) = 1.\end{aligned}$$

Moreover, since $F(x, y)$ and $\tilde{F}(x, y)$ are right-continuous functions of y, we have

$$F(x, y) = \lim_{\substack{r \in \mathbf{Q} \\ r \downarrow y}} F(x, r) = \lim_{\substack{r \in \mathbf{Q} \\ r \downarrow y}} \tilde{F}(x, r) = \tilde{F}(x, y) \quad \text{for every } y \in \mathbf{R}.$$

This proves that

$$F(x, \cdot) = \tilde{F}(x, \cdot) \quad \text{for } \mu_{\mathbf{X}}\text{-almost all } x \in \mathbf{R}^n.$$

Summing up, we have proved the desired assertion (2.36).
The proof of Theorem 2.27 is complete.

Definition 2.28. The function ψ on $\mathbf{R}^n \times \mathcal{B}(\mathbf{R})$ is called a *conditional distribution of X with respect to* $(X_1, X_2, \ldots, X_n)$. We shall write

$$P\,(X \in A \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \psi(x_1, x_2, \ldots, x_n, A),$$
$$A \in \mathcal{B}(\mathbf{R}).$$

Example 2.29. Let Y be a random variable and let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a vector-valued random variable on the probability space $(\Omega, \mathcal{F}, P)$. Then Y and $\mathbf{X}$ are independent if and only if we have, for $\mu_{\mathbf{X}}$-almost all $x \in \mathbf{R}^n$,

$$P\,(Y \in A \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = P(Y \in A), \tag{2.39}$$
$$A \in \mathcal{B}(\mathbf{R}).$$

Proof. (i) The "if" part: Since we have

$$\mathcal{B} = \sigma\,(\mathbf{X}) = \left\{\mathbf{X}^{-1}(B) : B \in \mathcal{B}(\mathbf{R}^n)\right\},$$

it follows from formula (2.35) and condition (2.39) that

$$\psi(x, A) = P\,(Y \in A \mid \mathbf{X} = x) = P(Y \in A), \quad A \in \mathcal{B}(\mathbf{R}).$$

Hence we have, for every $B \in \mathcal{B}(\mathbf{R}^n)$,

$$\begin{aligned} P\,((Y \in A) \cap (\mathbf{X} \in B)) &= \int_{(\mathbf{X} \in B)} P\,(Y \in A \mid \mathbf{X})\,(\omega)\,dP \\ &= \int_B \psi(x, A)\,d\mu_{\mathbf{X}} = \int_B P(Y \in A)\,d\mu_{\mathbf{X}} \\ &= P(Y \in A)\,\mu_{\mathbf{X}}(B) = P(Y \in A)\,P(\mathbf{X} \in B). \end{aligned}$$

This proves that the random variables Y and $\mathbf{X}$ are independent.

(ii) The "only if" part: If Y and $\mathbf{X}$ are independent variables, it follows from an application of Theorem 2.18 that the σ-algebras $\sigma(Y)$ and $\mathcal{B} = \sigma(\mathbf{X})$ are independent. Hence we have, for every $B \in \mathcal{B}(\mathbf{R}^n)$,

$$P((Y \in A) \cap (\mathbf{X} \in B)) = P(Y \in A)\, P(\mathbf{X} \in B) = \int_{(\mathbf{X} \in B)} P(Y \in A)\, dP.$$

This proves that

$$P(Y \in A \mid \mathbf{X})(\omega) = P(X \in A),$$

or equivalently,

$$P(Y \in A \mid \mathbf{X} = x) = P(Y \in A) \quad \text{for } \mu_{\mathbf{X}}\text{-almost all } x \in \mathbf{R}^n.$$

This proves condition (2.39).

Example 2.30. Let X and Y be random variables such that the joint distribution of (X, Y) has a density $f(x, y)$. Then a version of the conditional distribution of Y with respect to X has a density function

$$\frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y)\, dy}. \tag{2.40}$$

Proof. We have, for all Borel sets A, $B \in \mathcal{B}(\mathbf{R})$,

$$P(X \in B, Y \in A) = \int_B \int_A f(x, y)\, dx\, dy.$$

By taking $A = \mathbf{R}$, we obtain that the distribution of X is given by the formula

$$P(X \in B) = \int_B \left(\int_{-\infty}^{\infty} f(x, y)\, dy \right) dx.$$

That is, the distribution $\mu = P \circ X^{-1}$ of X has a density

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy,$$

since we have

$$\mu(B) = P \circ X^{-1}(B) = P(X \in B) = \int_B g(x)\, dx \quad \text{for every } B \in \mathcal{B}(\mathbf{R}).$$

If we let

$$\Delta = \{x \in \mathbf{R} : g(x) = 0\},$$

then we have

$$\mu(\Delta) = \int_{\Delta} g(x)\, dx = 0.$$

Therefore, we can define a function $\psi(x, A)$ on $\mathbf{R}^n \times \mathcal{B}(\mathbf{R})$ by the formula

$$\psi(x, A) = \begin{cases} \dfrac{1}{g(x)} \int_A f(x, y)\, dy & \text{if } x \notin \Delta, \\ 0 & \text{if } x \in \Delta. \end{cases}$$

Then it follows that $\psi(x, \cdot)$ is a probability measure on $\mathbf{R}$ if $x \notin \Delta$. Moreover, we obtain that ψ is a version of the conditional distribution of Y with respect to X. Indeed, it suffices to note that

$$\begin{aligned} &P\left((Y \in A) \cap (X \in B)\right) \\ &= \int_B \left(\int_A f(x, y)\, dy \right) dx \\ &= \int_{B \setminus \Delta} g(x)\psi(x, A)\, dx + \int_{B \cap \Delta} \left(\int_A f(x, y)\, dy \right) dx \\ &= \int_{B \setminus \Delta} g(x)\psi(x, A)\, dx = \int_B g(x)\psi(x, A)\, dx = \int_B \psi(x, A)\, d\mu \\ &= \int_{(X \in B)} \psi(X(\omega), A)\, dP \quad \text{for every } B \in \mathcal{B}(\mathbf{R}). \end{aligned}$$

Summing up, we have proved that $\psi(x, \cdot) = P(Y \in \cdot \mid X = x)$ has the density function (2.40).

The proof of Example 2.30 is complete.

2.6 Conditional Expectations

The general theory of conditional expectations plays a vital role in the study of Markov processes in Chap. 9.

Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. If $\mathcal{B}$ is a sub-σ-algebra of $\mathcal{F}$, then it follows from an application of Theorem 2.25 that there exists a conditional probability $P(X \in \cdot \mid \mathcal{B})$ of X with respect to $\mathcal{B}$.

The next theorem will play a crucial role in the study of Markov processes in Sect. 10.1:

Theorem 2.31. *Assume that $E(|X|) < \infty$. Then the integral*

$$Y(\omega) = \int_{-\infty}^{\infty} x \, P(X \in dx \mid \mathcal{B})(\omega) \tag{2.41}$$

exists for almost all $\omega \in \Omega$, and satisfies the following two conditions (CE1) and (CE2):

(CE1) The function Y is $\mathcal{B}$-measurable.
(CE2) $E(X; \Lambda) = E(Y; \Lambda)$ for every $\Lambda \in \mathcal{B}$. That is, we have, for every $\Lambda \in \mathcal{B}$,

$$\int_{\Lambda} X(\omega)\, dP = \int_{\Lambda} Y(\omega)\, dP.$$

Moreover, the function Y is uniquely determined in the sense that any two such functions are equal with respect to P.

Proof. The proof is divided into two steps.

Step 1: First, we show that the function $Y(\omega)$ is a real-valued, $\mathcal{B}$-measurable random variable, that is, it satisfies condition (CE1).
To do this, we let

$$A_{n,k} = (-(k+1)/2^n, -k/2^n] \cup [k/2^n, (k+1)/2^n), \quad n, k \in \mathbf{N},$$
$$\Lambda_k^{(n)} = X^{-1}(A_{n,k}) = \{\omega \in \Omega : X(\omega) \in A_{n,k}\}, \quad n, k \in \mathbf{N},$$

and

$$Z_n(\omega) = \sum_{k=0}^{\infty} \frac{k}{2^n} \chi_{\Lambda_k^{(n)}}(\omega) = \sum_{k=0}^{\infty} \frac{k}{2^n} \chi_{A_{n,k}}(X(\omega)), \quad n \in \mathbf{N}.$$

Then it is easy to see the following two assertions (a) and (b) (see Fig. 2.2):

(a) The Z_n are $\mathcal{B}$-measurable functions.
(b) $Z_n \uparrow |X|$ almost everywhere in Ω.

Hence, by applying the monotone convergence theorem [Fo2, Theorem 2.14] we obtain that

$$\begin{aligned}\int_{\Omega} |X(\omega)|\, dP &= \lim_{n\to\infty} \int_{\Omega} Z_n(\omega)\, dP \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \int_{\Omega} \chi_{\Lambda_k^{(n)}}(\omega)\, dP \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \int_{\Omega} \chi_{A_{n,k}}(X(\omega))\, dP\end{aligned} \tag{2.42}$$

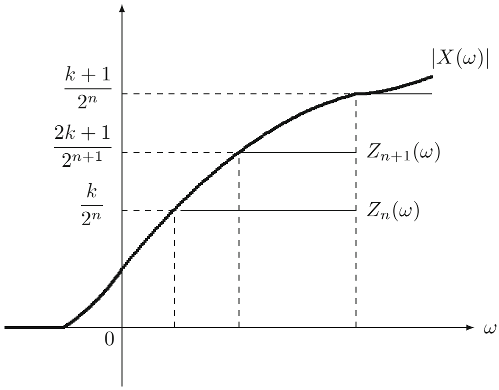


Fig. 2.2 The approximations $Z_n(\omega)$ to $|X(\omega)|$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} P\left(X \in A_{n,k}\right)$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} P\left(|X| \in [k/2^n, (k+1)/2^n)\right).$$

However, we have

$$\int_{-\infty}^{\infty} |x| \; P(X \in dx \mid \mathcal{B})(\omega) = \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} P\left(X \in A_{n,k} \mid \mathcal{B}\right)(\omega), \tag{2.43a}$$

$$\int_{\Omega} \chi_{\Lambda_k^{(n)}}(\omega)\, dP = P\left(\Lambda_k^{(n)}\right) = \int_{\Omega} P(\Lambda_k^{(n)} \mid \mathcal{B})(\omega)\, dP \tag{2.43b}$$

$$= \int_{\Omega} P(X \in A_{n,k} \mid \mathcal{B})(\omega)\, dP.$$

Hence, by using the monotone convergence theorem [Fo2, Theorem 2.14] we obtain from formula (2.42), (2.43a) and (2.43b) that

$$\int_{\Omega} \left(\int_{-\infty}^{\infty} |x| \; P(X \in dx \mid \mathcal{B})(\omega) \right) dP$$

$$= \int_{\Omega} \left(\lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} P\left(X \in A_{n,k} \mid \mathcal{B}\right)(\omega) \right) dP$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \int_{\Omega} P\left(X \in A_{n,k} \mid \mathcal{B}\right)(\omega)\, dP$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \int_{\Omega} \chi_{A_{n,k}}(X(\omega))\, dP = E\left(|X|\right) < \infty.$$

This proves that the $\mathcal{B}$-measurable function

$$\int_{-\infty}^{\infty} |x| \; P(X \in dx \mid \mathcal{B})(\omega)$$

is finite for almost all $\omega \in \Omega$.

Therefore, by letting

$$Y_n(\omega) = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega),$$

we obtain from assertion (2.41) that the series

$$\sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega)$$

converges absolutely for almost all $\omega \in \Omega$, and further that the limit function

$$\begin{aligned}\lim_{n\to\infty} Y_n(\omega) &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega) \\ &= \int_{-\infty}^{\infty} x \, P(X \in dx \mid \mathcal{B})(\omega) = Y(\omega)\end{aligned}$$

is finite for almost all $\omega \in \Omega$.

Step 2: Secondly, we let

$$X_n(\omega) = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \chi_{[k/2^n,(k+1)/2^n)}(X(\omega)), \quad n \in \mathbf{N}.$$

Then it is easy to see the following three assertions (a)–(c):

(a) The X_n are $\mathcal{B}$-measurable functions.
(b) $|X_n| \leq |X|$ almost everywhere in Ω.
(c) $X_n \to X$ almost everywhere in Ω.

Therefore, by using the dominated convergence theorem we obtain from formulas (2.41) and condition (2.43a) that

$$\begin{aligned}E(Y; \Lambda) &= \int_\Lambda Y(\omega)\, dP = \int_\Lambda \left(\int_{-\infty}^{\infty} x\, P(X \in dx \mid \mathcal{B})(\omega)\right) dP \\ &= \int_\Lambda \left(\lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n}\, P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega)\right) dP\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int_\Lambda P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega)\, dP \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(P\left(X \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right); \Lambda\right) \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left((X \in [k/2^n, (k+1)/2^n)) \cap \Lambda\right) \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int_\Lambda \chi_{[k/2^n, (k+1)/2^n)}(X(\omega))\, dP \\
&= \lim_{n\to\infty} \int_\Lambda X_n(\omega)\, dP = \int_\Lambda X(\omega)\, dP \\
&= E(X; \Lambda) \quad \text{for every } \Lambda \in \mathcal{B}.
\end{aligned}
$$

This proves that Y satisfies condition (CE2).

Step 3: Finally, we assume that a function $\hat{Y}$ satisfies conditions (CE1) and (CE2). If we let

$$\Lambda_1 = \{\omega \in \Omega : Y(\omega) > \hat{Y}(\omega)\},$$

it follows that $\Lambda_1 \in \mathcal{B}$. Moreover, since we have

$$E(X; \Lambda_1) = E(Y; \Lambda_1) = E(\hat{Y}; \Lambda_1),$$

it follows that

$$\int_{\Lambda_1} \left(Y(\omega) - \hat{Y}(\omega)\right) dP = E(Y - \hat{Y}; \Lambda_1) = 0.$$

This proves that $P(\Lambda_1) = 0$, since the integrand $Y - \hat{Y}$ is positive on the set Λ_1.

Similarly, if we let

$$\Lambda_2 = \{\omega \in \Omega : \hat{Y}(\omega) > Y(\omega)\},$$

it follows that $P(\Lambda_2) = 0$.

Summing up, we have proved that

$$P\left(\{\omega \in \Omega : Y(\omega) \neq \hat{Y}(\omega)\}\right) = P\left(\Lambda_1 \cup \Lambda_2\right) = 0,$$

so that Y and $\hat{Y}$ are equal with respect to P.

Now the proof of Theorem 2.31 is complete.

Definition 2.32. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{F}$. An integrable random variable Y is called a version of the *conditional expectation* of X for the given $\mathcal{B}$ if it satisfies conditions (CE1) and (CE2). We shall write

$$Y = E(X \mid \mathcal{B}) = \int_{-\infty}^{\infty} x\, P(X \in dx \mid \mathcal{B}).$$

Theorem 2.33. *Assume that a Borel function $f(x)$ on $\mathbf{R}$ satisfies the condition*

$$E(|f(X)|) = \int_{\Omega} |f(X(\omega))|\, dP < \infty.$$

Then we have

$$E(f(X) \mid \mathcal{B})(\omega) = \int_{-\infty}^{\infty} f(x)\, P(X \in dx \mid \mathcal{B})(\omega) \tag{2.44}$$

$$\text{for almost all } \omega \in \Omega.$$

Proof. First, by arguing just as in Theorem 2.31, we obtain that the right-hand side of (2.44)

$$\begin{aligned} &\int_{-\infty}^{\infty} f(x)\, P(X \in dx \mid \mathcal{B})(\omega) \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(f(X) \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega) \end{aligned}$$

is $\mathcal{B}$-measurable and finite for almost all $\omega \in \Omega$. Indeed, it suffices to note that

$$\begin{aligned} &\int_{\Omega} \left(\int_{-\infty}^{\infty} |f(x)|\ P(X \in dx \mid \mathcal{B})(\omega)\right) dP \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{2^n} \int_{\Omega} P\left(f(X) \in A_{n,k} \mid \mathcal{B}\right)(\omega)\, dP = E\left(|f(X)|\right) < \infty. \end{aligned}$$

Moreover, by using the dominated convergence theorem we have

$$\begin{aligned} &E(f(X); \Lambda) \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int_{\Lambda} P\left(f(X) \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right)(\omega)\, dP \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(\left(X \in f^{-1}[k/2^n, (k+1)/2^n)\right) \cap \Lambda\right) \end{aligned}$$

$$= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(P\left(X \in f^{-1}[k/2^n, (k+1)/2^n) \mid \mathcal{B}\right); \Lambda\right)$$

$$= E\left(\lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left(f(X) \in [k/2^n, (k+1)/2^n) \mid \mathcal{B}\right); \Lambda\right)$$

$$= E\left(\int_{-\infty}^{\infty} f(x)\, P(X \in dx \mid \mathcal{B}); \Lambda\right) \quad \text{for every } \Lambda \in \mathcal{B}.$$

This proves that the right-hand side of (2.44) satisfies condition (CE2).
The proof of Theorem 2.33 is complete.

When X is the characteristic function χ_A of a set $A \in \mathcal{F}$, then we have

$$E(\chi_A \mid \mathcal{B}) = P(A \mid \mathcal{B}).$$

Indeed, it suffices to note that the conditional probability $P(A \mid \mathcal{B})$ is a $\mathcal{B}$-measurable random variable which satisfies the condition

$$\int_\Lambda P(A \mid \mathcal{B})(\omega)\, dP = P(A \cap \Lambda) = \int_\Lambda \chi_A(\omega)\, dP \quad \text{for all } \Lambda \in \mathcal{B}.$$

The next theorem summarizes the basic properties of the conditional expectation:

Theorem 2.34. *Assume that $E(|X|) < \infty$. Then we have the following seven assertions (i)–(vii):*

(i) If X is $\mathcal{B}$-measurable, then it follows that $X(\omega) = E(X \mid \mathcal{B})$ for almost all $\omega \in \Omega$. In particular, we have $E(X \mid \mathcal{B})(\omega) = E(X)$ for almost all $\omega \in \Omega$ if $\mathcal{B} = \{\emptyset, \Omega\}$.

(ii) Conditional expectation is linear in X. That is, we have, for all $a_1, a_2 \in \mathbf{R}$,

$$\begin{aligned} &E(a_1 X_1 + a_2 X_2 \mid \mathcal{B})(\omega) \\ &= a_1 E(X_1 \mid \mathcal{B})(\omega) + a_2 E(X_2 \mid \mathcal{B})(\omega) \quad \text{for almost all } \omega \in \Omega. \end{aligned} \tag{2.45}$$

(iii) If $X_1(\omega) \le X_2(\omega)$ for almost all $\omega \in \Omega$, then it follows that $E(X_1 \mid \mathcal{B})(\omega) \le E(X_2 \mid \mathcal{B})(\omega)$ for almost all $\omega \in \Omega$.

(iv) If $X(\omega) \ge 0$ for almost all $\omega \in \Omega$, then it follows that $E(X \mid \mathcal{B})(\omega) \ge 0$ for almost all $\omega \in \Omega$. More precisely, we have the inequality

$$|E(X \mid \mathcal{B})(\omega)| \le E(|X| \mid \mathcal{B})(\omega) \quad \text{for almost all } \omega \in \Omega. \tag{2.46}$$

(v) If Y is $\mathcal{B}$-measurable and if $E(|XY|) < \infty$, then it follows that $E(XY \mid \mathcal{B})(\omega) = Y\, E(X \mid \mathcal{B})(\omega)$ for almost all $\omega \in \Omega$.

(vi) If $X_n(\omega) \uparrow X(\omega)$ for almost all $\omega \in \Omega$, then it follows that $E(X_n \mid \mathcal{B})(\omega) \uparrow E(X \mid \mathcal{B})(\omega)$ for almost all $\omega \in \Omega$.

(vii) If the σ-algebras $\sigma(X)$ and $\mathcal{B}$ are independent, then we have

$$E(X \mid \mathcal{B})(\omega) = E(X) \quad \textit{for almost all } \omega \in \Omega. \tag{2.47}$$

Proof. (i) This is trivial, since the function X itself satisfies conditions (CE1) and (CE2). Moreover, if $\mathcal{B} = \{\emptyset, \Omega\}$, it follows that $E(X \mid \mathcal{B})(\omega) = E(X)$ for almost all $\omega \in \Omega$. Indeed, we have

$$\int_\Omega E(X)\, dP = E(X) = \int_\Omega X(\omega)\, dP.$$

(ii) First, it follows that the function

$$a_1\, E(X_1 \mid \mathcal{B}) + a_2\, E(X_2 \mid \mathcal{B})$$

is $\mathcal{B}$-measurable. Moreover, we have, by assertion (i),

$$\begin{aligned}
&E\left(a_1 E(X_1 \mid \mathcal{B}) + a_2 E(X_2 \mid \mathcal{B}); \Lambda\right) \\
&= \int_\Lambda \left(a_1\, E(X_1 \mid \mathcal{B})(\omega) + a_2\, E(X_2 \mid \mathcal{B})(\omega)\right) dP \\
&= a_1 \int_\Lambda E(X_1 \mid \mathcal{B})(\omega)\, dP + a_2 \int_\Lambda E(X_2 \mid \mathcal{B})(\omega)\, dP \\
&= a_1 \int_\Lambda X_1(\omega)\, dP + a_2 \int_\Lambda X_2(\omega)\, dP \\
&= \int_\Lambda \left(a_1 X_1(\omega) + a_2 X_2(\omega)\right) dP \\
&= E\left(a_1 X_1 + a_2 X_2; \Lambda\right) \quad \text{for every } \Lambda \in \mathcal{B}.
\end{aligned}$$

This proves the desired equality (2.45).

(iii) If we let

$$\Lambda = \{\omega \in \Omega : E(X_2 \mid \mathcal{B})(\omega) < E(X_1 \mid \mathcal{B})(\omega)\},$$

then it follows that $\Lambda \in \mathcal{B}$. However, we have, by assertions (i) and (ii),

$$\begin{aligned}
0 &\le \int_\Lambda \left(E(X_1 \mid \mathcal{B})(\omega) - E(X_2 \mid \mathcal{B})(\omega)\right) dP \\
&= \int_\Lambda E(X_1 - X_2 \mid \mathcal{B})(\omega)\, dP = \int_\Lambda (X_1 - X_2)(\omega)\, dP \le 0,
\end{aligned}$$

so that

$$\int_{\Lambda} (E(X_1 \mid \mathcal{B})(\omega) - E(X_2 \mid \mathcal{B})(\omega)) \, dP = 0.$$

This proves that $P(\Lambda) = 0$, since the integrand $E(X_1 \mid \mathcal{B}) - E(X_2 \mid \mathcal{B})$ is positive on Λ.

(iv) Since we have the inequality

$$-|X(\omega)| \leq X(\omega) \leq |X(\omega)| \quad \text{for almost all } \omega \in \Omega,$$

it follows from an application of assertion (iii) that

$$-E(|X| \mid \mathcal{B})(\omega) \leq E(X \mid \mathcal{B})(\omega) \leq E(|X| \mid \mathcal{B})(\omega) \quad \text{for almost all } \omega \in \Omega.$$

This proves the desired inequality (2.46).

(v) We let

$$Y_n(\omega) = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \chi_{[k/2^n,(k+1)/2^n)} (Y(\omega)) , \quad n \in \mathbf{N}.$$

Then it is easy to see the following two assertions (a) and (b):

(a) The Y_n are $\mathcal{B}$-measurable functions.
(b) $Y_n(\omega) \to Y(\omega)$ for almost all $\omega \in \Omega$.

Moreover, since $|XY_n| \leq |XY|$ in Ω and $E(|XY|) < \infty$, by using the dominated convergence theorem we obtain that

$$\begin{aligned} & E(XY; \Lambda) \qquad (2.48) \\ &= \int_{\Lambda} X(\omega)Y(\omega) \, dP = \lim_{n\to\infty} \int_{\Lambda} X(\omega)Y_n(\omega) \, dP \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \int_{\Lambda} X(\omega) \frac{k}{2^n} \chi_{[k/2^n,(k+1)/2^n)}(Y(\omega)) \, dP \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} E\left(X \frac{k}{2^n} \chi_{[k/2^n,(k+1)/2^n)}(Y); \Lambda \right) \\ &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(X; \Lambda \cap Y^{-1}[k/2^n, (k+1)/2^n)\right) \quad \text{for every } \Lambda \in \mathcal{B}. \end{aligned}$$

However, since Y is $\mathcal{B}$-measurable, it follows that

$$\Lambda \cap Y^{-1}[k/2^n, (k+1)/2^n) \in \mathcal{B}.$$

Hence we obtain from condition (CE2) and the dominated convergence theorem that

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(X; \Lambda \cap Y^{-1}[k/2^n, (k+1)/2^n)\right) \qquad (2.49)\\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(E(X \mid \mathcal{B}); \Lambda \cap Y^{-1}[k/2^n, (k+1)/2^n)\right)\\
&= \lim_{n\to\infty} \int_\Lambda \sum_{k=-\infty}^{\infty} E(X \mid \mathcal{B})(\omega) \frac{k}{2^n} \chi_{[k/2^n,(k+1)/2^n)}(Y(\omega))\, dP\\
&= \lim_{n\to\infty} \int_\Lambda Y_n(\omega) E(X \mid \mathcal{B})(\omega)\, dP\\
&= E\left(Y\, E(X \mid \mathcal{B}); \Lambda\right) \quad \text{for every } \Lambda \in \mathcal{B}.
\end{aligned}$$

By combining formulas (2.48) and (2.49), we have proved that

$$\begin{aligned}
E(XY; \Lambda) &= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} E\left(X; \Lambda \cap Y^{-1}[k/2^n, (k+1)/2^n)\right)\\
&= E\left(Y\, E(X \mid \mathcal{B}); \Lambda\right) \quad \text{for every } \Lambda \in \mathcal{B}.
\end{aligned}$$

This proves that $E(XY \mid \mathcal{B})(\omega) = Y\, E(X \mid \mathcal{B})(\omega)$ for almost all $\omega \in \Omega$.

(vi) Since $X_n(\omega) \uparrow X(\omega)$ for almost all $\omega \in \Omega$, by applying assertion (iii) we find that $E(X_n \mid \mathcal{B})(\omega)$ is increasing in n. Hence, if we let

$$Y(\omega) = \min\left(\limsup_{n\to\infty} E(X_n \mid \mathcal{B})(\omega),\ E(X \mid \mathcal{B})(\omega)\right),$$

then it follows that Y is $\mathcal{B}$-measurable and further that $E(X_n \mid \mathcal{B})(\omega) \uparrow Y(\omega)$ for almost all $\omega \in \Omega$.

Therefore, we obtain from the monotone convergence theorem [Fo2, Theorem 2.14] and condition (CE2) that

$$\begin{aligned}
E(X; \Lambda) = \lim_{n\to\infty} E(X_n; \Lambda) &= \lim_{n\to\infty} E(E(X_n \mid \mathcal{B}); \Lambda)\\
&= E(Y; \Lambda) \quad \text{for every } \Lambda \in \mathcal{B}.
\end{aligned}$$

This proves that $Y(\omega) = E(X \mid \mathcal{B})(\omega)$ for almost all $\omega \in \Omega$.

(vii) Since X is independent of every set B in $\mathcal{B}$, it follows that

$$P\,(B \cap (X \in A)) = P(B)P\,(X \in A) \quad \text{for all } A \in \mathcal{B}(\mathbf{R}).$$

Therefore, we have

$$\begin{aligned}
&\int_B E(X \mid \mathcal{B})(\omega)\,dP \\
&= \int_B X(\omega)\,dP = \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int_B \chi_{[k/2^n,(k+1)/2^n)}(X(\omega))\,dP \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int_{B\cap X^{-1}[k/2^n,(k+1)/2^n)} dP \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\,(B \cap (X \in [k/2^n, (k+1)/2^n)) \\
&= \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\,(X \in [k/2^n, (k+1)/2^n)) \cdot P(B) = \int_\Omega X(\omega)\,dP \cdot P(B) \\
&= \int_B E(X)\,dP \quad \text{for every } B \in \mathcal{B}.
\end{aligned}$$

This proves the desired formula (2.47).

Now the proof of Theorem 2.34 is complete.

Theorem 2.35. *Assume that $E\,(|X|) < \infty$. If $\mathcal{B}_1 \subset \mathcal{B}_2$, then we have*

$$E(X \mid \mathcal{B}_1)(\omega) = E\,(E(X \mid \mathcal{B}_2) \mid \mathcal{B}_1)\,(\omega) = E\,(E(X \mid \mathcal{B}_1) \mid \mathcal{B}_2)\,(\omega) \tag{2.50}$$

$$\text{for almost all } \omega \in \Omega.$$

Proof. First, we have, for every $\Lambda \in \mathcal{B}_1 \subset \mathcal{B}_2$,

$$E(X; \Lambda) = \int_\Lambda X(\omega)\,dP = \int_\Lambda E(X \mid \mathcal{B}_2)(\omega)\,dP = E\,(E(X \mid \mathcal{B}_2); \Lambda)\,. \tag{2.51}$$

However, it follows from an application of assertion (iv) of Theorem 2.34 that

$$|E(X \mid \mathcal{B}_2)(\omega)| \leq E(|X| \mid \mathcal{B}_2)(\omega) \quad \text{for almost all } \omega \in \Omega,$$

so that

$$\begin{aligned}
E\,(|E(X \mid \mathcal{B}_2)|) \leq E\,(E(|X| \mid \mathcal{B}_2)) &= \int_\Omega E(|X| \mid \mathcal{B}_2)(\omega)\,dP \\
&= \int_\Omega |X|(\omega)\,dP = E\,(|X|) < \infty.
\end{aligned}$$

Hence, by taking the conditional expectation of $Z := E(X \mid \mathcal{B}_2)$ with respect to $\mathcal{B}_1$ we obtain that, for every $\Lambda \in \mathcal{B}_1$,

$$
\begin{aligned}
E\left(E(X \mid \mathcal{B}_2); \Lambda\right) &= E(Z; \Lambda) = E\left(E(Z \mid \mathcal{B}_1); \Lambda\right) \\
&= E\left(E(E(X \mid \mathcal{B}_2) \mid \mathcal{B}_1); \Lambda\right).
\end{aligned} \tag{2.52}
$$

Therefore, it follows from (2.51) and (2.52) that

$$
E(X; \Lambda) = \int_{\Lambda} E\left(E(X \mid \mathcal{B}_2) \mid \mathcal{B}_1\right)(\omega)\, dP \quad \text{for every } \Lambda \in \mathcal{B}_1.
$$

This proves the first equality in (2.50).

Moreover, since $E(X \mid \mathcal{B}_1)$ is $\mathcal{B}_2$-measurable, by applying assertion (i) of Theorem 2.34 with

$$
X := E(X \mid \mathcal{B}_1), \quad \mathcal{B} := \mathcal{B}_1,
$$

we obtain that

$$
E\left(E(X \mid \mathcal{B}_1) \mid \mathcal{B}_2\right)(\omega) = E(X \mid \mathcal{B}_1)(\omega) \quad \text{for almost all } \omega \in \Omega.
$$

This proves the second equality in (2.50).

Now the proof of Theorem 2.35 is complete.

Example 2.36. Let X be a random variable and let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a vector-valued random variables on the probability space $(\Omega, \mathcal{F}, P)$. We recall (see Definition 2.28) that the conditional distribution ψ of X with respect to the random variable $(X_1, X_2, \ldots, X_n)$ is given by the formula

$$
\begin{aligned}
&\psi(x_1, x_2, \ldots, x_n, A) \\
&= P\left(X \in A \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right) \\
&\qquad \text{for } \mu_{\mathbf{X}}\text{-almost all } (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n \text{ and } A \in \mathcal{B}(\mathbf{R}).
\end{aligned}
$$

Assume that a Borel measurable function $g(z, x_1, \ldots, x_n)$ on $\mathbf{R}^{n+1}$ satisfies the condition

$$
\begin{aligned}
&E\left(|g(X, X_1, X_2, \ldots, X_n)|\right) \\
&= \int_{\Omega} |g(X(\omega), X_1(\omega), X_2(\omega) \ldots, X_n(\omega))|\, dP < \infty.
\end{aligned}
$$

Then we have

$$E\left(g(X, X_1, \ldots, X_n) \mid X_1, X_2, \ldots, X_n\right)(\omega) \tag{2.53}$$
$$= \int_{-\infty}^{\infty} g(x, X_1(\omega), \ldots, X_n(\omega))\psi\left(X_1(\omega), X_2(\omega), \ldots, X_n(\omega), dx\right)$$
$$\text{for almost all } \omega \in \Omega.$$

If we define a Borel measurable function $h(x_1, x_2, \ldots, x_n)$ on $\mathbf{R}^n$ by the formula

$$h(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{\infty} g(x, x_1, x_2, \ldots, x_n)\psi\left(x_1, x_2, \ldots, x_n, dx\right),$$

then we obtain from (2.53) that

$$E\left(g(X, X_1, \ldots, X_n) \mid X_1, X_2, \ldots, X_n\right)(\omega) = h(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$$
$$\text{for almost all } \omega \in \Omega.$$

We shall write

$$h(x_1, x_2, \ldots, x_n)$$
$$= E\left(g(X, X_1, X_2, \ldots, X_n) \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right).$$

Example 2.37. Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ and $\mathbf{Y} = (X_{n+1}, X_{n+2}, \ldots, X_{n+\ell})$ be vector-valued random variables on the probability space $(\Omega, \mathcal{F}, P)$. Then there exists a conditional distribution Φ of $\mathbf{Y}$ with respect to the random variable $(X_1, X_2, \ldots, X_n)$:

$$\Phi(x_1, x_2, \ldots, x_n, A)$$
$$= P\left((X_{n+1}, X_{n+2}, \ldots, X_{n+\ell}) \in A \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right)$$
$$\text{for } \mu_{\mathbf{X}}\text{-almost all } (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n \text{ and } A \in \mathcal{B}(\mathbf{R}^\ell).$$

Proof. Let ψ_k be the conditional distribution ψ of X_{k+1} with respect to the random variable $(X_1, X_2, \ldots, X_k)$:

$$\psi_k(x_1, x_2, \ldots, x_k, dx_{k+1})$$
$$= P\left(X_{k+1} \in dx_{k+1} \mid X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k\right).$$

We let

$$\begin{aligned}
&\Phi(x_1, x_2, \ldots, x_n, A) \\
&= \int_{-\infty}^{\infty} \psi_n(x_1, x_2, \ldots, x_n, dx_{n+1}) \int_{-\infty}^{\infty} \psi_{n+1}(x_1, \ldots x_{n+1}, dx_{n+2}) \cdots \\
&\quad \times \int_{-\infty}^{\infty} \chi_A(x_{n+1}, \ldots, x_{n+\ell})\, \psi_{n+\ell-1}\,(x_1, \ldots, x_{n+\ell-1}, dx_{n+\ell})\,, \quad A \in \mathcal{B}(\mathbf{R}^\ell).
\end{aligned}$$

Then we can prove the following two assertions (1) and (2):

(1) For $\mu_{\mathbf{X}}$-almost all $x \in \mathbf{R}^n$, $\Phi(x_1, x_2, \ldots, x_n, \cdot)$ is a probability measure on $\mathbf{R}^\ell$.
(2) For every $A \in \mathcal{B}(\mathbf{R}^\ell)$, $\Phi(\cdot, A)$ is a Borel measurable function on $\mathbf{R}^n$.

Moreover, by applying Theorem 2.35 we have

$$\begin{aligned}
&\Phi(X_1(\omega), \ldots, X_n(\omega), A) \\
&= E\Big(E \cdots E\Big(E\,(\chi_A\,(X_{n+1}, \ldots, X_{n+\ell}) \mid X_1, \ldots, X_{n+\ell-1}) \\
&\qquad\qquad \mid X_1, \ldots, X_{n+\ell-2}\Big) \cdots \mid X_1, \ldots, X_n\Big)(\omega) \\
&= E\,(\chi_A\,(X_{n+1}, \ldots, X_{n+\ell}) \mid X_1, \ldots, X_n)\,(\omega) \\
&= P\,(X_{n+1}, \ldots, X_{n+\ell} \in A \mid X_1, \ldots, X_n)\,(\omega) \quad \text{for almost all } \omega \in \Omega.
\end{aligned}$$

This proves that

$$\begin{aligned}
&\Phi(x_1, x_2, \ldots, x_n, A) \\
&= P\,((X_{n+1}, X_{n+2}, \ldots, X_{n+\ell}) \in A \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)\,.
\end{aligned}$$

The proof of Example 2.37 is complete.

2.7 Notes and Comments

The results discussed here are based on Blumenthal–Getoor [BG], Lamperti [La], Nishio [Ni] and Folland [Fo2].

Section 2.1: The monotone class (Theorem 2.4) and the Dynkin class theorem (Corollary 2.5) were first proved by Dynkin [Dy1]. Our proof is due to Blumenthal–Getoor [BG, Chapter 0]. The approximation theorem (Theorem 2.6) is taken from Nishio [Ni, Chapter 2, Section 3, Theorem 5].

Sections 2.2–2.4: The material in these sections is taken from Nishio [Ni].

Section 2.5: Theorems 2.25 and 2.27 are adapted from Nishio [Ni, Chapter 7, Section 1].

Section 2.6: Theorems 2.31, 2.33 and 2.34 are adapted from Nishio [Ni, Chapter 7, Section 2] and Lamperti [La, Appendix 2].

Chapter 3
Elements of Functional Analysis

This chapter is devoted to a review of standard topics from functional analysis such as quasinormed and normed linear spaces and closed, compact and Fredholm linear operators on Banach spaces. These topics form a necessary background for what follows. In Sects. 3.1–3.3 we study linear operators and functionals, quasinormed and normed linear spaces. In a normed linear space we consider continuous linear functionals as generalized coordinates of the space. The existence of non-trivial, continuous linear functionals is based on the Hahn–Banach extension theorem (Theorem 3.21). In particular, Mazur's theorem (Theorem 3.25) asserts that there exists a non-trivial, continuous linear functional which separates a point and a closed convex, balanced set. In Sect. 3.3.2, by using Mazur's theorem we prove that a closed convex, balanced subset is weakly closed (Corollary 3.26). This corollary plays an important role in the proof of an existence and uniqueness theorem for a class of pseudo-differential operators in the framework of Hölder spaces (Theorem 10.23) in Chap. 10. In Sect. 3.4 we prove the Riesz–Markov representation theorem (Theorem 3.41) which describes an intimate relationship between Radon measures and non-negative linear functionals on the spaces of continuous functions. This fact constitutes an essential link between measure theory and functional analysis, providing a powerful tool for constructing Markov transition functions in Chap. 9. Section 3.5 is devoted to closed operators and Sect. 3.6 is devoted to complemented subspaces in a normed linear space, respectively. Section 3.7 is devoted to the Riesz–Schauder theory for compact operators. More precisely, for a compact operator T on a Banach space, the eigenvalue problem can be treated fairly completely in the sense that the classical theory of Fredholm integral equations may be extended to the linear functional equation $Tx - \lambda x = y$ with a complex parameter λ (Theorem 3.61). In Sect. 3.8 we state important properties of Fredholm operators (Theorems 3.64–3.66). Roughly speaking, the Fredholm property of T conveys that the operator T is both "almost" injective and "almost" surjective, that is, it is "almost" an isomorphism. Moreover, the index ind T indicates how far the operator

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_3,

T is from being bijective. Namely, the further ind T is from zero, the more bijective T is. The stability theorem for indices of Fredholm operators (Theorem 3.66) plays an essential role in the proof of Theorem 1.2 in Chap. 10.

3.1 Linear Operators and Functionals

Let X, Y be linear spaces over the same scalar field $\mathbf{K}$. A mapping T defined on a linear subspace $\mathcal{D}$ of X and taking values in Y is said to be *linear* if it preserves the operations of addition and scalar multiplication:

$$T(x_1 + x_2) = Tx_1 + Tx_2 \quad \text{for all } x_1, x_2 \in \mathcal{D}. \tag{L1}$$

$$T(\alpha x) = \alpha Tx \quad \text{for all } x \in \mathcal{D} \text{ and } \alpha \in \mathbf{K}. \tag{L2}$$

We often write Tx, rather than $T(x)$, if T is linear. We let

$$D(T) = \mathcal{D},$$
$$R(T) = \{Tx : x \in D(T)\},$$
$$N(T) = \{x \in D(T) : Tx = 0\},$$

and call them the *domain*, the *range* and the *null space* of T, respectively. The mapping T is called a *linear operator* from $D(T) \subset X$ into Y. We also say that T is a linear operator from X into Y with domain $D(T)$. In the particular case when $Y = \mathbf{K}$, the mapping T is called a *linear functional* on $D(T)$. In other words, a linear functional is a $\mathbf{K}$-valued function on $D(T)$ which satisfies conditions (L1) and (L2).

If a linear operator T is a one-to-one map of $D(T)$ onto $R(T)$, then it is easy to see that the inverse mapping T^{-1} is a linear operator on $R(T)$ onto $D(T)$. The mapping T^{-1} is called the *inverse operator* or simply the *inverse* of T. A linear operator T admits an inverse if and only if $Tx = 0$ implies that $x = 0$.

Let T_1 and T_2 be two linear operators from a linear space X into a linear space Y with domains $D(T_1)$ and $D(T_2)$, respectively. Then we say that $T_1 = T_2$ if and only if $D(T_1) = D(T_2)$ and $T_1 x = T_2 x$ for every $x \in D(T_1) = D(T_2)$. If $D(T_1) \subset D(T_2)$ and $T_1 x = T_2 x$ for every $x \in D(T_1)$, then we say that T_2 is an *extension* of T_1 and also that T_1 is a *restriction* of T_2, and we write $T_1 \subset T_2$.

3.2 Quasinormed Linear Spaces

Let X be a linear space over the real or complex number field $\mathbf{K}$. A real-valued function p defined on X is called a *seminorm* on X if it satisfies the following three conditions (S1)–(S3):

$$0 \leq p(x) < \infty \quad \text{for all } x \in X. \tag{S1}$$

$$p(\alpha x) = |\alpha| p(x) \quad \text{for all } \alpha \in \mathbf{K} \text{ and } x \in X. \tag{S2}$$

$$p(x+y) \leq p(x) + p(y) \quad \text{for all } x, y \in X. \tag{S3}$$

Let $\{p_i\}$ be a countable family of seminorms on X such that

$$p_1(x) \leq p_2(x) \leq \cdots \leq p_i(x) \leq \cdots \quad \text{for every } x \in X, \tag{3.1}$$

and define

$$V_{ij} = \left\{ x \in X : p_i(x) < \frac{1}{j} \right\}, \quad i, j = 1, 2, \ldots.$$

Then it is easy to verify that a countable family of the sets

$$x + V_{ij} = \{x + y : y \in V_{ij}\}$$

satisfies the axioms of a fundamental neighborhood system of x; hence X is a topological space which satisfies the first axiom of countability.

Furthermore, we have the following theorem:

Theorem 3.1. *Let $\{p_i\}$ be a countable family of seminorms on a linear space X which satisfies condition (3.1). Assume that*

For every non-zero element $x \in X$, there exists a seminorm p_i such that $p_i(x) > 0$. (3.2)

Then the space X is metrizable *by the metric*

$$\rho(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{p_i(x)}{1 + p_i(x)} \quad \textit{for all } x, y \in X.$$

If we let

$$|x| = \rho(x, 0) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{p_i(x)}{1 + p_i(x)} \quad \textit{for every } x \in X, \tag{3.3}$$

then the quantity $|x|$ enjoys the following four properties (Q1)–(Q4):

(Q1) $|x| \geq 0$; $|x| = 0$ *if and only if* $x = 0$.
(Q2) $|x + y| \leq |x| + |y|$ *(the triangle inequality).*
(Q3) $\alpha_n \to 0$ *in* $\mathbf{K} \Longrightarrow |\alpha_n x| \to 0$ *for every* $x \in X$.
(Q4) $|x_n| \to 0 \Longrightarrow |\alpha x_n| \to 0$ *for every* $\alpha \in \mathbf{K}$.

This quantity $|x|$ is called a *quasinorm* of x, and the space X is called a *quasinormed linear space*.

Theorem 3.1 may be restated as follows:

Theorem 3.2. *A linear space X, topologized by a countable family $\{p_i\}_{i=1}^{\infty}$ of seminorms satisfying conditions (3.1) and (3.2), is a quasinormed linear space with respect to the quasinorm $|x|$ defined by formula (3.3).*

Let X be a quasinormed linear space. The convergence

$$\lim_{n\to\infty} |x_n - x| = 0$$

in X is denoted by $s - \lim_{n\to\infty} x_n = x$ or simply by $x_n \to x$, and we say that the sequence $\{x_n\}_{n=1}^{\infty}$ *converges strongly* to x. A sequence $\{x_n\}_{n=1}^{\infty}$ is called a *Cauchy sequence* if it satisfies Cauchy's (convergence) condition

$$\lim_{m,n\to\infty} |x_m - x_n| = 0.$$

A quasinormed linear space X is called a *Fréchet space* if it is complete, that is, if every Cauchy sequence in X converges strongly to a point in X. If a quasinormed linear space X is topologized by a countable family $\{p_i\}$ of seminorms which satisfies conditions (3.1) and (3.2), then the above definitions may be reformulated in terms of seminorms as follows.

(i) A sequence $\{x_n\}_{n=1}^{\infty}$ in X converges strongly to a point x in X if and only if, for every seminorm p_i and every positive ε, there exists a positive integer $N = N(i, \varepsilon)$ such that

$$n \geq N \Longrightarrow p_i(x_n - x) < \varepsilon.$$

(ii) A sequence $\{x_n\}_{n=1}^{\infty}$ in X is a Cauchy sequence if and only if, for every seminorm p_i and every positive ε, there exists a positive integer $N = N(i, \varepsilon)$ such that

$$m, n \geq N \Longrightarrow p_i(x_m - x_n) < \varepsilon.$$

Let X be a quasinormed linear space. A linear subspace of X is called a *closed subspace* if it is a closed subset of X. For example, the closure $\overline{M}$ of a linear subspace M is a closed subspace. Indeed, the elements of $\overline{M}$ are limits of sequences in M; thus, if $x = \lim_{n\to\infty} x_n$ with $x_n \in M$ and $y = \lim_{n\to\infty} y_n$ with $y_n \in M$, then it follows that

$$x + y = \lim_{n\to\infty}(x_n + y_n), \quad x_n + y_n \in M,$$

$$\alpha x = \lim_{n\to\infty} \alpha x_n, \quad \alpha x_n \in M.$$

This proves that $x + y \in \overline{M}$ and $\alpha x \in \overline{M}$ for every $\alpha \in \mathbf{K}$.

3.2.1 Bounded Sets

Let X be a quasinormed linear space, topologized by a countable family $\{p_i\}$ of seminorms which satisfies conditions (3.1) and (3.2). A set B in X is said to be *bounded* if we have, for every seminorm p_i,

$$\sup_{x \in B} p_i(x) < +\infty.$$

We remark that every compact set is bounded.

Throughout the rest of this section, let X and Y be quasinormed linear spaces over the same scalar field, topologized respectively by countable families $\{p_i\}_{i=1}^{\infty}$ and $\{q_i\}_{i=1}^{\infty}$ of seminorms which satisfy conditions (3.1) and (3.2).

3.2.2 Continuity of Linear Operators

Let T be a linear operator from X into Y with domain $D(T)$. By virtue of the linearity of T, it follows that T is continuous everywhere on $D(T)$ if and only if it is continuous at one point of $D(T)$. Furthermore, we have the following theorem:

Theorem 3.3. *A linear operator T from X into Y with domain $D(T)$ is continuous everywhere on $D(T)$ if and only if, for every seminorm q_j on Y, there exist a seminorm p_i on X and a positive constant C such that*

$$q_j(Tx) \leq Cp_i(x) \quad \text{for all } x \in D(T).$$

3.2.3 Topologies of Linear Operators

We let

$$\mathcal{L}(X, Y) = \text{the collection of continuous linear operators on } X \text{ into } Y.$$

We define in the set $\mathcal{L}(X, Y)$ addition and scalar multiplication of operators in the usual way:

$$(T + S)x = Tx + Sx \quad \text{for all } x \in X,$$
$$(\alpha T)x = \alpha(Tx) \quad \text{for all } \alpha \in \mathbf{K} \text{ and } x \in X.$$

Then it follows that $\mathcal{L}(X, Y)$ is a linear space.

We introduce three different topologies on the space $\mathcal{L}(X, Y)$.

(1) *Simple convergence topology*: This is the topology of convergence at each point of X; a sequence $\{T_n\}_{n=1}^{\infty}$ in $\mathcal{L}(X, Y)$ converges to an element T of $\mathcal{L}(X, Y)$ in the simple convergence topology if and only if $T_n x \to Tx$ in Y as $n \to \infty$, for each $x \in X$.
(2) *Compact convergence topology*: This is the topology of uniform convergence on compact sets in X; $T_n \to T$ in the compact convergence topology if and only if $T_n x \to Tx$ in Y as $n \to \infty$, uniformly for x ranging over compact sets in X.
(3) *Bounded convergence topology*: This is the topology of uniform convergence on bounded sets in X; $T_n \to T$ in the bounded convergence topology if and only if $T_n x \to Tx$ in Y as $n \to \infty$, uniformly for x ranging over bounded sets in X.

The simple convergence topology is weaker than the compact convergence topology, and the compact convergence topology is weaker than the bounded convergence topology.

3.2.4 The Banach–Steinhaus Theorem

We introduce three different definitions of boundedness for sets in the space $\mathcal{L}(X, Y)$:

(1) A set H in $\mathcal{L}(X, Y)$ is said to be bounded in the simple convergence topology if, for each $x \in X$, the set $\{Tx : T \in H\}$ is bounded in Y.
(2) A set H in $\mathcal{L}(X, Y)$ is said to be bounded in the compact convergence topology if, for every compact set K in X, the set $\bigcup_{T \in H} T(K)$ is bounded in Y.
(3) A set H in $\mathcal{L}(X, Y)$ is said to be bounded in the bounded convergence topology if, for every bounded set B in X, the set $\bigcup_{T \in H} T(B)$ is bounded in Y.

Furthermore, a set H in $\mathcal{L}(X, Y)$ is said to be *equicontinuous* if, for every seminorm q_j on Y, there exist a seminorm p_i on X and a positive constant C such that

$$\sup_{T \in H} q_j(Tx) \leq C p_i(x) \quad \text{for all } x \in X.$$

The next theorem states one of the fundamental properties of Fréchet spaces:

Theorem 3.4 (Banach–Steinhaus). *Let X be a Fréchet space and Y a quasi-normed linear space. Then the following four conditions (i)–(iv) are equivalent:*

(i) A set H in $\mathcal{L}(X, Y)$ is bounded in the simple convergence topology.
(ii) A set H in $\mathcal{L}(X, Y)$ is bounded in the compact convergence topology.

(iii) A set H in $\mathcal{L}(X, Y)$ is bounded in the bounded convergence topology.
(iv) A set H in $\mathcal{L}(X, Y)$ is equicontinuous.

3.2.5 Product Spaces

Let X and Y be quasinormed linear spaces over the same scalar field $\mathbf{K}$. Then the Cartesian product $X \times Y$ becomes a linear space over $\mathbf{K}$ if we define the algebraic operations coordinatewise

$$\{x_1, y_1\} + \{x_2, y_2\} = \{x_1 + x_2, y_1 + y_2\} \text{ for all } x_1, x_2 \in X \text{ and } y_1, y_2 \in Y,$$

$$\alpha\{x, y\} = \{\alpha x, \alpha y\} \text{ for all } \alpha \in \mathbf{K} \text{ and } x \in X, y \in Y.$$

It is easy to verify that the quantity

$$|\{x, y\}| = \left(|x|^2 + |y|^2\right)^{1/2}, \quad \{x, y\} \in X \times Y, \tag{3.4}$$

satisfies axioms (Q1)–(Q4) of a quasinorm; hence the product space $X \times Y$ is a quasinormed linear space with respect to the quasinorm defined by formula (3.4). Furthermore, if X and Y are Fréchet spaces, then so is $X \times Y$. In other words, the completeness is inherited by the product space.

3.3 Normed Linear Spaces

A quasinormed linear space is called a *normed linear space* if it is topologized by just one seminorm which satisfies condition (3.2). We give the precise definition of a normed linear space.

Let X be a linear space over the real or complex number field $\mathbf{K}$. A real-valued function $\|\cdot\|$ defined on X is called a *norm* on X if it satisfies the following three conditions (N1)–(N3):

(N1) $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$.
(N2) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbf{K}$ and $x \in X$.
(N3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (the triangle inequality).

A linear space X equipped with a norm $\|\cdot\|$ is called a normed linear space. The topology on X is defined by the metric

$$\rho(x, y) = \|x - y\| \quad \text{for all } x, y \in X.$$

The convergence

$$\lim_{n \to \infty} \|x_n - x\| = 0$$

in X is denoted by $s - \lim_{n\to\infty} x_n = x$ or simply $x_n \to x$, and we say that the sequence $\{x_n\}_{n=1}^{\infty}$ converges strongly to x. A sequence $\{x_n\}_{n=1}^{\infty}$ in X is called a *Cauchy sequence* if it satisfies the condition

$$\lim_{n,m \to \infty} \|x_n - x_m\| = 0.$$

A normed linear space X is called a *Banach space* if it is complete, that is, if every Cauchy sequence in X converges strongly to a point in X.

Two norms $\| \cdot \|_1$ and $\| \cdot \|_2$ defined on the same linear space X are said to be *equivalent* if there exist positive constants c and C such that

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \text{for all } x \in X.$$

Equivalent norms induce the same topology.

If X and Y are normed linear spaces over the same scalar field, then the product space $X \times Y$ is a normed linear space with the norm

$$\|\{x, y\}\| = \left(\|x\|_X^2 + \|y\|_Y^2\right)^{1/2} \quad \text{for all } \{x, y\} \in X \times Y.$$

If X and Y are Banach spaces, then so is $X \times Y$.

Let X be a normed linear space. If Y is a closed linear subspace of X, then the factor space X/Y is a normed linear space by the norm

$$\|\tilde{x}\| = \inf_{z \in \tilde{x}} \|z\| \quad \text{for every } \tilde{x} \in X/Y. \tag{3.5}$$

If X is a Banach space, then so is X/Y. The space X/Y, normed by (3.5), is called a *normed factor space*.

Throughout the rest of this section, the letters X, Y, Z denote normed linear spaces over the same scalar field.

The next theorem is a normed linear space version of Theorem 3.3:

Theorem 3.5. *Let T be a linear operator from X into Y with domain $D(T)$. Then T is continuous everywhere on $D(T)$ if and only if there exists a positive constant C such that*

$$\|Tx\| \leq C\|x\| \quad \textit{for all } x \in D(T). \tag{3.6}$$

Remark 3.6. In inequality (3.6), the quantity $\|x\|$ is the norm of x in X and the quantity $\|Tx\|$ is the norm of Tx in Y. Frequently several norms appear together, but it is clear from the context which is which.

One of the consequences of Theorem 3.5 is the following extension theorem for a continuous linear operator:

Theorem 3.7. *If T is a continuous linear operator from X into Y with domain $D(T)$ and if Y is a Banach space, then T has a unique continuous extension $\tilde{T}$ whose domain is the closure $\overline{D(T)}$ of $D(T)$.*

As another consequence of Theorem 3.5, we give a necessary and sufficient condition for the existence of a continuous inverse of a linear operator:

Theorem 3.8. *Let T be a linear operator from X into Y with domain $D(T)$. Then T admits a continuous inverse T^{-1} if and only if there exists a positive constant c such that*

$$\|Tx\| \geq c\|x\| \quad \text{for all } x \in D(T).$$

A linear operator T from X into Y with domain $D(T)$ is called an *isometry* if it is norm-preserving, that is, if we have

$$\|Tx\| = \|x\| \quad \text{for every } x \in D(T).$$

It is clear that if T is an isometry, then it is injective and both T and T^{-1} are continuous.

If T is a continuous, one-to-one linear mapping from X onto Y and if its inverse T^{-1} is also a continuous mapping, then it is called an *isomorphism* of X onto Y. Two normed linear spaces are said to be *isomorphic* if there is an isomorphism between them.

By combining Theorems 3.5 and 3.8, we obtain the following theorem:

Theorem 3.9. *Let T be a linear operator on X onto Y. Then T is an isomorphism if and only if there exist positive constants c and C such that*

$$c\|x\| \leq \|Tx\| \leq C\|x\| \quad \text{for all } x \in X.$$

If T is a continuous linear operator from X into Y with domain $D(T)$, we let

$$\|T\| = \inf\{C : \|Tx\| \leq C\|x\|, x \in D(T)\}.$$

Then, in view of the linearity of T we have

$$\|T\| = \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in D(T) \\ \|x\|=1}} \|Tx\| = \sup_{\substack{x \in D(T) \\ \|x\| \leq 1}} \|Tx\|. \tag{3.7}$$

This proves that $\|T\|$ is the smallest non-negative number such that

$$\|Tx\| \leq \|T\| \cdot \|x\| \quad \text{for all } x \in D(T). \tag{3.8}$$

Theorem 3.5 asserts that a linear operator T from X into Y is continuous if and only if it maps bounded sets in X into bounded sets in Y. Thus a continuous linear operator from X into Y is usually called a *bounded linear operator* from X into Y. We let

$$\mathcal{L}(X, Y) = \text{the space of bounded (continuous) linear operators from } X \text{ into } Y.$$

In the case of normed linear spaces, the simple convergence topology on $\mathcal{L}(X, Y)$ is usually called the *strong topology* of operators, and the bounded convergence topology on $\mathcal{L}(X, Y)$ is called the *uniform topology* of operators. In view of formulas (3.7) and (3.8), it follows that the quantity $\|T\|$ satisfies axioms (N1)–(N3) of a norm; hence the space $\mathcal{L}(X, Y)$ is a normed linear space with the norm $\|T\|$ given by formula (3.7). The topology on $\mathcal{L}(X, Y)$ induced by the operator norm $\|T\|$ is just the uniform topology of operators.

We give a sufficient condition for the space $\mathcal{L}(X, Y)$ to be complete:

Theorem 3.10. *If Y is a Banach space, then so is $\mathcal{L}(X, Y)$.*

If T is a linear operator from X into Y with domain $D(T)$ and S is a linear operator from Y into Z with domain $D(S)$, then we define the product ST as follows:

(a) $D(ST) = \{x \in D(T) : Tx \in D(S)\}$.
(b) $(ST)(x) = S(Tx)$ for every $x \in D(ST)$.

As for the product of linear operators, we have the following proposition:

Proposition 3.11. *If $T \in \mathcal{L}(X, Y)$ and $S \in \mathcal{L}(Y, Z)$, then it follows that $ST \in \mathcal{L}(X, Z)$. Moreover, we have the inequality*

$$\|ST\| \leq \|S\| \cdot \|T\|.$$

We often make use of the following theorem in constructing the bounded inverse of a bounded linear operator:

Theorem 3.12. *If T is a bounded linear operator on a Banach space X into itself and satisfies $\|T\| < 1$, then the operator $I - T$ has a unique bounded linear inverse $(I - T)^{-1}$ which is given by* C. Neumann's series

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n = I + T + T^2 + \ldots + T^n + \ldots.$$

Here I is the identity operator: $Ix = x$ for every $x \in X$, and $T^0 = I$.

As an important application of Theorem 3.12, we give the functional analytic background for the method of continuity in nonlinear analysis:

Theorem 3.13. *Let X be a Banach space and let Y be a normed linear space. If T_0 and T_1 are two bounded linear operators from X into Y, we define a family of bounded linear operators*

$$T_t = (1-t)T_0 + t\,T_1 : X \longrightarrow Y \quad \text{for every } t \in [0,1].$$

Assume that there exists a positive constant C, independent of t, such that

$$\|x\|_X \leq C\,\|T_t x\|_Y \quad \text{for all } x \in X. \tag{3.9}$$

Then the operator T_1 maps X onto Y if and only if the operator T_0 maps X onto Y.

The next theorem is a normed linear space version of the Banach–Steinhaus theorem (Theorem 3.4):

Theorem 3.14 (The principle of uniform boundedness). *Let X be a Banach space and let Y be a normed linear space. If H is a subset of $\mathcal{L}(X,Y)$, then the boundedness of the set $\{\|Tx\| : T \in H\}$ at each $x \in X$ implies the boundedness of the set $\{\|T\| : T \in H\}$.*

Corollary 3.15. *Let X be a Banach space, Y a normed linear space and $\{T_n\}_{n=1}^{\infty}$ a sequence in $\mathcal{L}(X,Y)$. If the limit*

$$s - \lim_{n\to\infty} T_n x = Tx \tag{3.10}$$

exists for every $x \in X$, then it follows that $T \in \mathcal{L}(X,Y)$ and we have the inequality

$$\|T\| \leq \liminf_{n\to\infty} \|T_n\|.$$

The operator T obtained above is called the *strong limit* of $\{T_n\}_{n=1}^{\infty}$, since the convergence (3.19) is in the strong topology of operators. We then write

$$T = s - \lim_{n\to\infty} T_n.$$

3.3.1 Finite Dimensional Spaces

The next theorem asserts that there is no point in studying abstract finite dimensional normed linear spaces:

Theorem 3.16. *All n-dimensional normed linear spaces over the same scalar field $\mathbf{K}$ are isomorphic to $\mathbf{K}^n$ with the maximum norm*

$$\|\alpha\| = \max_{1\leq i\leq n} |\alpha_i| \quad \text{for every } \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbf{K}^n.$$

The topological properties of the space $\mathbf{K}^n$ apply to all finite dimensional normed linear spaces.

Corollary 3.17. *All finite dimensional normed linear spaces are complete.*

Corollary 3.18. *Every finite dimensional linear subspace of a normed linear space is closed.*

Corollary 3.19. *A subset of a finite dimensional normed linear space is compact if and only if it is closed and bounded.*

By Corollary 3.19, it follows that the closed unit ball in a finite dimensional normed linear space is compact. Conversely, this property characterizes finite dimensional spaces:

Theorem 3.20. *If the closed unit ball in a normed linear space X is compact, then X is finite dimensional.*

3.3.2 The Hahn–Banach Extension Theorem

In a normed linear space we consider continuous linear functionals as generalized coordinates of the space. The existence of non-trivial, continuous linear functionals is based on the following Hahn–Banach extension theorem:

Theorem 3.21 (The Hahn–Banach extension theorem). *Let X be a linear space over the real or complex number field $\mathbf{K}$ and let p be a seminorm defined on X. If M is a linear subspace of X and if f is a linear functional defined on M such that $|f(x)| \leq p(x)$ on M, then there exists a linear functional F on X such that F is an extension of f and further that $|F(x)| \leq p(x)$ on X.*

The next theorem is a version of the Hahn–Banach extension theorem for normed linear spaces:

Theorem 3.22. *Let X be a normed linear space over the real or complex number field $\mathbf{K}$, and let M be a linear subspace of X. If f is a continuous linear functional defined on M, then it can be extended to a continuous linear functional f_0 on X so that*

$$\|f_0\| = \|f\|. \tag{3.11}$$

Proof. We let

$$p(x) = \|f\| \cdot \|x\| \quad \text{for every } x \in X.$$

Then it follows that p is a continuous seminorm on X such that

$$|f(x)| \leq \|f\| \cdot \|x\| = p(x) \quad \text{on } M.$$

By applying Theorem 3.21 to our situation, we can find a linear functional f_0 on X such that f_0 is an extension of f and that $|f_0(x)| \leq p(x)$ on X. Hence we obtain that

$$\|f_0\| \leq p(x) = \|f\| \cdot \|x\| \quad \text{for every } x \in X,$$

so that

$$\|f_0\| \leq \|f\|.$$

This proves the desired assertion (3.11), since f_0 is an extension of f.

The proof of Theorem 3.22 is complete.

Now we can prove the existence of non-trivial, continuous linear functionals:

Theorem 3.23. *Let X be a normed linear space over the real or complex number field $\mathbf{K}$, and let x_0 be a point of X. If p is a continuous seminorm defined on X, then there exists a continuous linear functional F on X such that $F(x_0) = p(x_0)$ and further that $|F(x)| \leq p(x)$ on X.*

Proof. We let

$$M = \{\alpha x_0 : \alpha \in \mathbf{K}\},$$

and define a functional f on M by the formula

$$f(\alpha x_0) = \alpha\, p(x_0) \quad \text{for every } \alpha \in \mathbf{K}.$$

Then it is easy to see that f is linear and continuous on M and that

$$|f(\alpha x_0)| = |\alpha|\, p(x_0) = p(\alpha x_0) \quad \text{for every } \alpha \in \mathbf{K}.$$

By applying Theorem 3.21, we can find a linear functional F on X such that F is an extension of f and that

$$|F(x)| \leq p(x) \quad \text{on } X.$$

This proves that $F(x_0) = f(x_0) = p(x_0)$ and further that $F(x)$ is continuous at $x = 0$ with $p(x)$. By the linearity of F, it follows that $F(x)$ is continuous at every point of X.

The proof of Theorem 3.23 is complete.

Corollary 3.24. *Let X be a normed linear space. For each non-zero element x_0 of X, there exists a continuous linear functional f_0 on X such that*

$$\begin{cases} f_0(x_0) = \|x_0\|, \\ \|f_0\| = 1. \end{cases}$$

Proof. In the proof of Theorem 3.23, we take

$$p(x) = \|x\| \quad \text{for every } x \in X.$$

Then it follows from an application of Theorem 3.23 that there exists a continuous linear functional f_0 on X such that $f_0(x_0) = p(x_0) = \|x_0\|$ and further that

$$|f_0(x)| \leq p(x) = \|x\| \quad \text{for every } x \in X.$$

This proves that $\|f_0\| = 1$.

The proof of Corollary 3.24 is complete.

A continuous linear functional on X is usually called a *bounded linear functional* on X.

A closed subset M of a normed linear space X is said to be *balanced* if it satisfies the condition

$$x \in M, \ |\alpha| \leq 1 \Longrightarrow \alpha x \in M.$$

The next theorem asserts that there exists a non-trivial, continuous linear functional which separates a point and a closed convex, balanced set:

Theorem 3.25 (Mazur). *Let X be a real or complex, normed linear space and let M be a closed convex, balanced subset of X. Then, for any point $x_0 \notin M$ there exists a continuous linear functional f_0 on X such that $f_0(x_0) > 1$ and $|f_0(x)| \leq 1$ on M.*

Proof. Let

$$\text{dist}\,(x_0, M) = \inf_{z \in M} \|x_0 - z\|$$

be the distance from x_0 to the set M. Since M is closed and $x_0 \notin M$, it follows that

$$\text{dist}\,(x_0, M) > 0.$$

If $0 < d < \text{dist}\,(x_0, M)$, we let

$$B\left(0, \frac{d}{2}\right) = \left\{x \in X : \|x\| \leq \frac{d}{2}\right\},$$

$$B\left(x_0, \frac{d}{2}\right) = x_0 + B\left(0, \frac{d}{2}\right) = \left\{x \in X : \|x - x_0\| \leq \frac{d}{2}\right\},$$

$$U = \left\{x \in X : \text{dist}\,(x, M) \leq \frac{d}{2}\right\}.$$

Then we have

$$U \cap B\left(x_0, \frac{d}{2}\right) = \emptyset, \tag{3.12a}$$

$$B\left(0, \frac{d}{2}\right) \subset U, \tag{3.12b}$$

since $0 \in M$.

Moreover, since M is convex and balanced, it is easy to verify the following three assertions (a)–(c):

(a) U is convex.
(b) U is balanced.
(c) U is absorbing, that is, for any $x \in X$, there exists a constant $\alpha > 0$ such that $\alpha^{-1}x \in U$.

Indeed, assertion (c) follows from assertion (3.12b).

Therefore, we can define the *Minkowski functional* p_U of U by the formula

$$p_U(x) = \inf_{\substack{\alpha>0 \\ \alpha^{-1}x \in U}} \alpha \quad \text{for every } x \in X.$$

By virtue of assertion (3.12b), we find that

$$p_U(x) \leq \frac{2}{d}\|x\| \quad \text{for every } x \in X.$$

This proves that p_U is a continuous seminorm on X. Since U is closed, it is easy to verify the following assertions (3.13):

$$p_U(x) > 1 \quad \text{if } x \notin U, \tag{3.13a}$$

$$p_U(x) \leq 1 \quad \text{if } x \in U. \tag{3.13b}$$

Therefore, by applying Theorem 3.23 with

$$p := p_U,$$

we can find a continuous linear functional f_0 on X such that

$$\begin{cases} f_0(x_0) = p_U(x_0) > 1 & \text{for } x_0 \notin U, \\ |f_0(x)| \leq p_U(x) & \text{on } X. \end{cases}$$

In particular, we have, by assertion (3.13b),

$$|f_0(x)| \leq p_U(x) \leq 1 \quad \text{on } M,$$

since $M \subset U$.

The proof of Theorem 3.25 is complete.

The next corollary will play an important role in the proof of Theorem 10.23 in Chap. 10:

Corollary 3.26. *Let X be a real or complex, normed linear space. If M is a closed convex, balanced subset of X, then it is closed in the weak topology of X.*

Proof. Assume, to the contrary, that M is not weakly closed in the weak topology of X. Then there exists a point $x_0 \notin M$ such that x_0 is an accumulation point of M in the weak topology of X. That is, there exists a sequence $\{x_n\}$ of M such that $\{x_n\}$ converges weakly to x_0. However, by using Mazur's theorem (Theorem 3.25) we can find a continuous linear functional f_0 on X such that

$$\begin{cases} f_0(x_0) > 1, \\ |f_0(x)| \leq 1 \quad \text{on } M. \end{cases}$$

Hence we have

$$1 < |f_0(x_0)| = \lim_{n \to \infty} |f_0(x_n)| \leq 1.$$

This contradiction proves Corollary 3.26.

3.3.3 Dual Spaces

Let X be a normed linear space over the real or complex number field $\mathbf{K}$. Then the space $\mathcal{L}(X, \mathbf{K})$ of all bounded linear functionals on X is called the *dual space* of X, and is denoted by X'. The bounded (resp. simple) convergence topology on X' is called the *strong* (resp. *weak**) *topology* on X' and the dual space X' equipped with this topology is called the *strong* (resp. *weak**) *dual space* of X.

It follows from an application of Theorem 3.10 with $Y := \mathbf{K}$ that the strong dual space X' is a Banach space with the norm

$$\|f\| = \sup_{\substack{x \in X \\ \|x\| \leq 1}} |f(x)|.$$

Corollary 3.24 asserts that the dual space X' separates points of X, that is, for arbitrary two distinct points x_1, x_2 of X, there exists a functional $f \in X'$ such that $f(x_1) \neq f(x_2)$.

Example 3.27. Let $(\Omega, \mathcal{M}, \mu)$ be a measure space. If $1 \leq p < \infty$, we let

$$L^p(\Omega) = \text{the space of equivalence classes of measurable functions } f(x) \text{ on } \Omega \text{ such that } |f|^p \text{ is integrable on } \Omega.$$

We remark that $L^p(\Omega)$ is a Banach space with the norm

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \, d\mu \right)^{1/p}.$$

If $1 < p < \infty$, we let $q = p/(p-1)$, so that $1 < q < \infty$ and

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

Assume that $(\Omega, \mathcal{M}, \mu)$ is σ-finite. Then we have the following two assertions (i) and (ii):

(i) If $v \in L^q(\Omega)$, we define a linear functional f_v by the formula

$$f_v(u) = \int_\Omega u(x)v(x) \, d\mu \quad \text{for every } u \in L^p(\Omega),$$

then we have

$$\begin{cases} f_v \in (L^p(\Omega))', \\ \|f_v\| = \|v\|_{L^q(\Omega)}. \end{cases}$$

(ii) Conversely, every element $f \in (L^p(\Omega))'$ can be expressed as $f = f_v$ for some function $v \in L^q(\Omega)$.

3.3.4 Annihilators

Let A be a subset of a normed linear space X. An element f of the dual space X' is called an *annihilator* of A if it satisfies the condition

$$f(x) = 0 \quad \text{for all } x \in A.$$

We let

$$A^0 = \{f \in X' : f(x) = 0 \text{ for all } x \in A\}$$

be the set of all annihilators of A. This is not a one way proposition. If B is a subset of X', we let

$$^0B = \{x \in X : f(x) = 0 \text{ for all } f \in B\}$$

be the set of all annihilators of B.

Here are some basic properties of annihilators:

(i) The sets A^0 and ${}^0B$ are closed linear subspaces of X and X', respectively.
(ii) If M is a closed linear subspace of X, then ${}^0(M^0) = M$.
(iii) If A is a subset of X and M is the closure of the subspace spanned by A, then $M^0 = A^0$ and $M = {}^0(A^0)$.

3.3.5 Dual Spaces of Normed Factor Spaces

Let M be a closed linear subspace of a normed linear space X. Then each element f of M^0 defines a bounded linear functional $\tilde{f}$ on the normed factor space X/M by the formula

$$\tilde{f}(\tilde{x}) = f(x) \quad \text{for all } \tilde{x} \in X/M.$$

Indeed, the value $f(x)$ on the right-hand side does not depend on the choice of a representative x of the equivalence class $\tilde{x}$, and we have

$$\|\tilde{f}\| = \|f\|.$$

Furthermore, it is easy to see that the mapping

$$\pi : f \longmapsto \tilde{f}$$

of M^0 into $(X/M)'$ is linear and surjective; hence we have the following theorem:

Theorem 3.28. *The strong dual space $(X/M)'$ of the factor space X/M can be identified with the space M^0 of all annihilators of M by the linear isometry π.*

3.3.6 Bidual Spaces

Each element x of a normed linear space X defines a bounded linear functional Jx on the strong dual space X' by the formula

$$Jx(f) = f(x) \quad \text{for every } f \in X'. \tag{3.14}$$

Then Corollary 3.24 asserts that

$$\|Jx\| = \sup_{\substack{f \in X' \\ \|f\| \leq 1}} |Jx(f)| = \|x\|,$$

so that the mapping J is a linear isometry of X into the strong dual space $(X')'$ of X'. The space $(X')'$ is called the *strong bidual space* or *strong second dual space* of X.

Summing up, we have the following theorem:

Theorem 3.29. *A normed linear space X can be embedded into its strong bidual space $(X')'$ by the linear isometry J defined by formula (3.14).*

If the mapping J is surjective, that is, if $X = (X')'$, then we say that X is *reflexive*.

3.3.7 Weak Convergence

A sequence $\{x_n\}_{n=1}^{\infty}$ in a normed linear space X is said to be *weakly convergent* if a finite $\lim_{n\to\infty} f(x_n)$ exists for every f in the dual space X' of X. A sequence $\{x_n\}_{n=1}^{\infty}$ in X is said to *converge weakly* to an element x of X if $\lim_{n\to\infty} f(x_n) = f(x)$ for every $f \in X'$; we then write $w - \lim_{n\to\infty} x_n = x$ or simply $x_n \to x$ weakly. Since the space X' separates points of X, the limit x is uniquely determined. Theorem 3.29 asserts that X may be considered as a linear subspace of its bidual space $(X')'$; hence the weak topology on X is just the simple convergence topology on the bidual space $(X')' = \mathcal{L}(X', \mathbf{K})$.

For weakly convergent sequences, we have the following theorem:

Theorem 3.30. *(i) $s - \lim_{n\to\infty} x_n = x$ implies that $w - \lim_{n\to\infty} x_n = x$.*
(ii) A weakly convergent sequence $\{x_n\}_{n=1}^{\infty}$ is bounded:

$$\sup_{n\geq 1} \|x_n\| < \infty.$$

Furthermore, if $w - \lim_{n\to\infty} x_n = x$, then the sequence $\{x_n\}_{n=1}^{\infty}$ is bounded and we have the inequality

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Part (ii) of Theorem 3.30 has a converse:

Theorem 3.31. *A sequence $\{x_n\}_{n=1}^{\infty}$ in X converges weakly to an element x of X if the following two conditions (a) and (b) are satisfied:*

(a) The sequence $\{x_n\}_{n=1}^{\infty}$ is bounded.
(b) $\lim_{n\to\infty} f(x_n) = f(x)$ for every f in some strongly dense subset of X'.

3.3.8 Weak* Convergence

A sequence $\{f_n\}_{n=1}^{\infty}$ in the dual space X' is said to be *weakly* convergent* if a finite $\lim_{n\to\infty} f_n(x)$ exists for every $x \in X$. A sequence $\{f_n\}_{n=1}^{\infty}$ in X' is said to *converge* *weakly** to an element f of X' if $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in X$; we then write $w*-\lim_{n\to\infty} f_n = f$ or simply $f_n \to f$ weakly*. The weak* topology on X' is just the simple topology on the dual space $X' = \mathcal{L}(X, \mathbf{K})$.

We have the following analogue of Theorem 3.30:

Theorem 3.32. *(i)* $s-\lim_{n\to\infty} f_n = f$ *implies that* $w*-\lim_{n\to\infty} f_n = f$.
(ii) If X *is a Banach space, then a weakly* convergent sequence* $\{f_n\}_{n=1}^{\infty}$ *in* X' *converges weakly* to an element f of* X' *and we have the inequality*

$$\|f\| \leq \liminf_{n\to\infty} \|f_n\|.$$

One of the important consequences of Theorem 3.32 is the *sequential weak* compactness* of bounded sets:

Theorem 3.33. *Let* X *be a separable Banach space. Then every bounded sequence in the strong dual space* X' *has a subsequence which converges weakly* to an element of* X'.

3.3.9 Transposes

Let T be a linear operator from X into Y with domain $D(T)$ dense in X. Such operators are called *densely defined operators*. Each element g of the dual space Y' of Y defines a linear functional G on $D(T)$ by the formula

$$G(x) = g(Tx) \quad \text{for every } x \in D(T).$$

If this functional G is continuous everywhere on $D(T)$, it follows from an application of Theorem 3.7 that G can be extended uniquely to a continuous linear functional g' on the closure

$$\overline{D(T)} = X,$$

that is, there exists a unique element g' of the dual space X' of X which is an extension of G. So we let

$$D(T') = \text{the totality of those } g \in Y' \text{ such that the mapping}$$
$$x \longmapsto g(Tx)$$
$$\text{is continuous everywhere on } D(T),$$

and define

$$T'g = g'.$$

In other words, the mapping T' is a linear operator from Y' into X' with domain $D(T')$ such that

$$g(Tx) = (T'g)(x) \quad \text{for all } x \in D(T) \text{ and } g \in D(T'). \tag{3.15}$$

The operator T' is called the *transpose* of T.

Frequently we write $\langle f, x \rangle$ or $\langle x, f \rangle$ for the value $f(x)$ of a functional f at a point x. For example, we write formula (3.15) as follows:

$$\langle Tx, g \rangle = \langle x, T'g \rangle \quad \text{for all } x \in D(T) \text{ and } g \in D(T').$$

The next theorem states that the continuity of operators is inherited by the transposes:

Theorem 3.34. *Let X, Y be normed linear spaces and X', Y' be their strong dual spaces, respectively. If T is a bounded linear operator from X into Y, then its transpose T' is a bounded linear operator from Y' into X', and we have*

$$\|T'\| = \|T\|.$$

3.4 Continuous Functions and Measures

One of the fundamental theorems in analysis is the Riesz–Markov representation theorem which describes an intimate relationship between measures and linear functionals.

3.4.1 Spaces of Continuous Functions

A topological space is said to be *locally compact* if every point has a compact neighborhood. Let (X, ρ) be a locally compact metric space. Let $C(X)$ be the collection of real-valued, continuous functions on X. We define on the set $C(X)$ addition and scalar multiplication of functions in the usual way:

$$(f + g)(x) = f(x) + g(x) \quad \text{for all } x \in X.$$
$$(\alpha f)x = \alpha f(x) \quad \text{for all } \alpha \in \mathbf{K} \text{ and } x \in X.$$

Then $C(X)$ is a real linear space.

The next two results show that locally compact metric spaces have a rich supply of continuous functions that vanish outside compact sets (cf. Dugundji [Dg]):

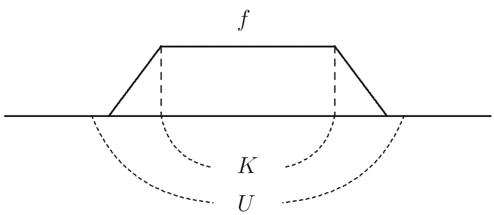


Fig. 3.1 The function f in Lemma 3.35

Lemma 3.35 (Urysohn). *Let (X, ρ) be a locally compact metric space. If $K \subset U \subset X$ where K is compact and U is open, then there exists a real-valued, continuous function $f \in C(X)$ such that (see Fig. 3.1)*

$$\begin{cases} 0 \leq f(x) \leq 1 & \text{on } X, \\ f(x) = 1 & \text{on } K, \\ f(x) = 0 & \text{outside a compact subset of } U. \end{cases}$$

Theorem 3.36 (Tietze's extension theorem). *Let (X, ρ) be a locally compact metric space and K a compact subset of X. If $f \in C(K)$, then there exists a real-valued, continuous function $F \in C(X)$ such that $F = f$ on K. Moreover, the function $F(x)$ may be taken to vanish outside a compact set.*

If $f \in C(X)$, the *support* of f, denoted by $\operatorname{supp} f$, is the smallest closed set outside of which $f(x)$ vanishes, that is, the closure of the set $\{x \in X : f(x) \neq 0\}$. If $\operatorname{supp} f$ is compact, we say that f is *compactly supported*. We define a subspace of $C(X)$ as follows:

$$C_c(X) = \{f \in C(X) : \operatorname{supp} f \text{ is compact}\}.$$

In other words, $C_c(X)$ is the space of compactly supported, continuous functions on X. The notation 'C_c' is used only for the moment in this section (later we will generally work in 'C_0').

If (X, ρ) is a non-compact, locally compact metric space, then we can make X into a compact space by adding a single "point at infinity" in such a way that the functions in $C_0(X)$ are precisely those continuous functions f such that $f(x) \to 0$ as x approaches the point at infinity.

More precisely, let ∂ denote a point that is not an element of X and let $X_\partial = X \cup \{\partial\}$. Then we have the following proposition:

Proposition 3.37. *Let (X, ρ) be a locally compact metric space and let $\mathcal{T}$ be the collection of all subsets U of $X_\partial = X \cup \{\partial\}$ such that either (i) U is an open subset of X or (ii) $\partial \in U$ and $U^c = X_\partial \setminus U$ is a compact subset of X. Then the space $(X_\partial, \mathcal{T})$ is a compact space and the inclusion map $i : X \to X_\partial$ is an embedding. Furthermore, if $f \in C(X)$, then $f(x)$ extends continuously to X_∂ if and only if $f(x) = g(x) + c$ where $g \in C_0(X)$ and $c \in \mathbf{R}$, in which case the continuous extension is given by $f(\partial) = c$ (see Fig. 3.2).*

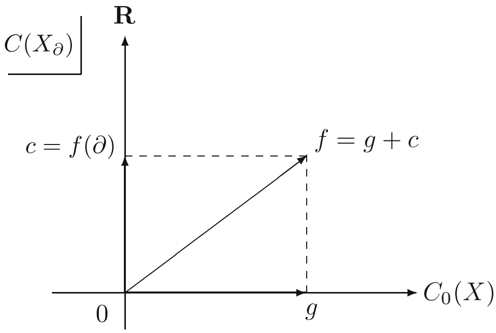


Fig. 3.2 The function $f = g + c$ in Proposition 3.37

The space $X_\partial = X \cup \{\partial\}$ is called the *one-point compactification* of X and the point ∂ is called the *point at infinity*.

Furthermore, if $f \in C(X)$, we say that f is *vanishes at infinity* if the set $\{x \in X : |f(x)| \geq \varepsilon\}$ is compact for every $\varepsilon > 0$, and we write

$$\lim_{x \to \partial} f(x) = 0.$$

We define a subspace $C_0(X)$ of $C(X)$ as follows:

$$C_0(X) = \left\{ f \in C(X) : \lim_{x \to \partial} f(x) = 0 \right\}.$$

It is easy to see that $C_0(X)$ is a Banach space with the supremum (maximum) norm

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

The next proposition asserts that $C_0(X)$ is the uniform closure of $C_c(X)$:

Proposition 3.38. *Let (X, ρ) be a locally compact metric space. The space $C_0(X)$ is the closure of $C_c(X)$ in the topology of uniform convergence.*

Proof. Assume that $\{f_n\}$ is a sequence in $C_c(X)$ which converges uniformly to some function $f \in C(X)$. Then, for any given $\varepsilon > 0$ there exists a number $n \in \mathbf{N}$ such that

$$\|f - f_n\|_\infty < \varepsilon.$$

Hence we have

$$|f(x)| < \varepsilon \quad \text{if } x \in X \setminus \operatorname{supp} f_n.$$

This proves that the set $\{x \in X : |f(x)| \geq \varepsilon\}$ is compact for every $\varepsilon > 0$, so that $f \in C_0(X)$.

Conversely, if $f \in C_0(X)$, we let

$$K_n = \left\{ x \in X : |f(x)| \geq \frac{1}{n} \right\}, \quad n \in \mathbf{N}.$$

Since K_n is compact, by applying Urysohn's lemma (Lemma 3.35) we can find a function $g_n \in C_c(X)$ such that $0 \leq g_n \leq 1$ and $g_n = 1$ on K_n. Then it follows that $f_n = g_n f \in C_c(X)$ and that

$$\|f - f_n\|_\infty = \|(1 - g_n) f\|_\infty \leq \frac{1}{n}.$$

This proves that $\{f_n\}$ converges uniformly to $f \in C(X)$.

The proof of Proposition 3.38 is complete.

3.4.2 Space of Signed Measures

Let $(X, \mathcal{M})$ be a measurable space. A real-valued function μ defined on the σ-algebra $\mathcal{M}$ is called a *signed measure* or *real measure* if it is countably additive, that is,

$$\mu\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for any disjoint countable collection $\{A_i\}_{i=1}^{\infty}$ of members of $\mathcal{M}$. It should be noted that every rearrangement of the series $\sum_i \mu(A_i)$ also converges, since the disjoint union $\sum_{i=1}^{\infty} A_i$ is not changed if the subscripts are permuted. A signed measure takes its values in $(-\infty, \infty)$, but a non-negative measure may take ∞; hence the non-negative measures do not form a subclass of the signed measures.

If μ and λ are signed measures on $\mathcal{M}$, we define the sum $\mu + \lambda$ and the scalar multiple $\alpha\mu$ ($\alpha \in \mathbf{R}$) as follows:

$$(\mu + \lambda)(A) = \mu(A) + \lambda(A) \quad \text{for all } A \in \mathcal{M},$$
$$(\alpha\mu)(A) = \alpha\mu(A) \quad \text{for all } \alpha \in \mathbf{K} \text{ and } A \in \mathcal{M}.$$

Then it is clear that $\mu + \lambda$ and $\alpha\mu$ are signed measures.

If μ is a signed measure, we define a function $|\mu|$ on $\mathcal{M}$ by

$$|\mu|(A) = \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| \right\} \quad \text{for all } A \in \mathcal{M},$$

where the supremum is taken over all finite partitions $\{A_i\}$ of A into members of $\mathcal{M}$. Then the function $|\mu|$ is a finite non-negative measure on $\mathcal{M}$. The measure $|\mu|$ is called the *total variation measure* of μ, and the quantity $|\mu|(X)$ is called the *total variation* of μ. We observe that

$$|\mu(A)| \leq |\mu|(A) \leq |\mu|(X) \quad \text{for all } A \in \mathcal{M}. \tag{3.16}$$

Furthermore, we can verify that the quantity $|\mu|(X)$ satisfies axioms (N1)–(N3) of a norm. Thus the totality of signed measures on $\mathcal{M}$ is a normed linear space with the norm $\|\mu\| := |\mu|(X)$.

If we define two functions μ^+ and μ^- on $\mathcal{M}$ by the formulas

$$\mu^+ = \frac{1}{2}(|\mu| + \mu),$$
$$\mu^- = \frac{1}{2}(|\mu| - \mu),$$

then it follows from inequalities (3.16) that both μ^+ and μ^- are finite non-negative measures on $\mathcal{M}$. It should be emphasized that the measures μ^+ and μ^- are the positive and negative variation measures of μ, respectively. We also have the *Jordan decomposition* of μ:

$$\mu = \mu^+ - \mu^-.$$

3.4.3 The Riesz–Markov Representation Theorem

The object of this section is to show that non-negative linear functionals on the spaces of continuous functions are given by integration against Radon measures. This fact constitutes an essential link between measure theory and functional analysis, providing a powerful tool for constructing measures.

Let (X, ρ) be a locally compact metric space and K a compact subset of X, and

$$C_c(X) = \{f \in C(X) : \text{ supp } f \text{ is compact}\}.$$
$$C_0(X) = \{f \in C(X) : f \text{ vanishes at infinity}\}.$$

First, we characterize the non-negative linear functionals on $C_c(X)$. A linear functional I on $C_c(X)$ is said to be *non-negative* if $I(f) \geq 0$ whenever $f \geq 0$, that is, if it satisfies the condition

$$f \in C_c(X), \ f(x) \geq 0 \text{ on } X \implies I(f) \geq 0.$$

The next proposition asserts that non-negativity implies a rather strong continuity property:

Proposition 3.39. *If I is a non-negative linear functional on $C_c(X)$, then, for every compact set $K \subset X$ there exists a positive constant C_K such that*

$$|I(f)| \leq C_K \|f\|_\infty \quad \textit{for all } f \in C_c(X) \textit{ with } \operatorname{supp} f \subset K.$$

Here

$$\|f\|_\infty = \sup_{x \in X} |f(x)| .$$

If μ is Borel measure on X such that $\mu(K) < \infty$ for every compact set $K \subset X$, then it follows that

$$C_c(X) \subset L^1(X, \mu).$$

Hence the map

$$I_\mu : f \longmapsto \int_X f(x)\, d\mu(x)$$

is a non-negative linear functional on the space $C_c(X)$. The purpose of this subsection is to prove that every non-negative linear functional on $C_c(X)$ arises in this fashion. In doing this, we impose some additional conditions on μ, subject to which μ is uniquely determined.

A *Radon measure* on X is a Borel measure that is finite on all compact sets in X, and is outer regular on all Borel sets in X and inner regular on all open sets in X.

If U is a open set in X and $f \in C_c(X)$, then we write

$$f \prec U$$

to mean that (see Fig. 3.3)

$$\begin{cases} 0 \leq f(x) \leq 1 & \text{on } X, \\ \operatorname{supp} f \subset U. \end{cases}$$

On the other hand, if K is a subset of X, we let

$$\chi_K(x) = \begin{cases} 1 & \text{if } x \in K, \\ 0 & \text{if } x \in X \setminus K, \end{cases}$$

Fig. 3.3

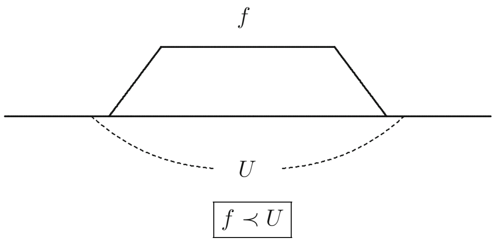
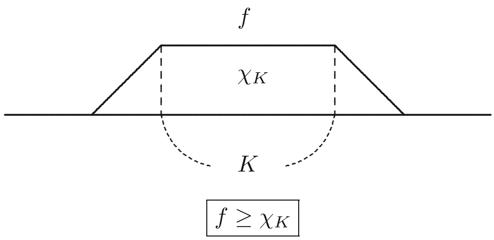


Fig. 3.4



and we write

$$f \geq \chi_K$$

to mean that (see Fig. 3.4)

$$f(x) \geq \chi_K(x) \quad \text{on } X.$$

The next theorem asserts that non-negative linear functionals on the space $C_c(X)$ are given by integration against Radon measures:

Theorem 3.40. *Let (X, ρ) be a locally compact metric space. If I is a non-negative linear functional on the space $C_c(X)$, then there is a unique Radon measure μ on X such that*

$$I(f) = \int_X f(x)\, d\mu(x) \quad \textit{for all } f \in C_c(X). \tag{3.17}$$

Furthermore, the Radon measure μ enjoys the following two properties (3.18) *and* (3.19)*:*

$$\mu(U) = \sup\{I(f) : f \in C_c(X),\ f \prec U\} \textit{ for every open set } U \subset X. \tag{3.18}$$

$$\mu(K) = \inf\{I(f) : f \in C_c(X),\ f \geq \chi_K\} \textit{ for every compact set } K \subset X. \tag{3.19}$$

Proof. The proof is divided into two steps.

Step 1: First, we prove the uniqueness of the Radon measure. More precisely, we show that a Radon measure μ is determined by I on all Borel subsets of X.

Assume that μ is a Radon measure such that

$$I(f) = \int_X f(x)\, d\mu \quad \text{for all } f \in C_c(X).$$

Let U be an arbitrary open subset of X. Then we have, for every function $f \prec U$,

$$I(f) = \int_X f(x)\, d\mu \leq \int_U d\mu = \mu(U).$$

On the other hand, if K is a compact subset of U, then it follows from an application of Urysohn's lemma (Lemma 3.35) that there exists a function $f \in C_c(X)$ such that $f \prec U$ and $f = 1$ on K. Hence we have the inequality

$$\mu(K) = \int_X \chi_K(x)\, d\mu \leq \int_X f(x)\, d\mu = I(f).$$

However, since μ is inner regular, we obtain that

$$\begin{aligned}\mu(U) &= \sup\{\mu(K) : K \subset U,\ K \text{ is compact}\} \\ &\leq \sup\{I(f) : f \in C_c(X),\ f \prec U\} \leq \mu(U).\end{aligned}$$

Therefore, we have, for every open set $U \subset X$,

$$\mu(U) = \sup\{I(f) : f \in C_c(X),\ f \prec U\}.$$

This proves that the Radon measure μ is determined by I on open subsets U of X, and hence by I on all Borel subsets of X, since it is outer regular on all Borel sets.

Step 2: The proof of the uniqueness suggests how to construct a Radon measure μ. More precisely, we begin by defining $\mu(U)$ for an arbitrary open set $U \subset X$ by

$$\mu(U) = \sup\{I(f) : f \in C_c(X),\ f \prec U\},$$

and then define $\mu^*(E)$ for an arbitrary set $E \subset X$ by the formula

$$\mu^*(E) = \inf\{\mu(U) : U \supset X,\ U \text{ is open}\}.$$

It should be noted that $\mu^*(U) = \mu(U)$ if U is open, since we have $\mu(U) \leq \mu(V)$ for $U \subset V$.

The idea of the proof may be explained as follows:

(i) First, we prove that μ^* is an outer measure:

$$\mu^*(\emptyset) = 0.$$
$$\mu^*(E) \leq \mu^*(F) \quad \text{if } E \subset F.$$
$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu^*(E_j).$$

(ii) Secondly, we prove that every open subset U of X is μ^*-measurable:

$$\mu^*(E) = \mu^*(E \cap U) + \mu^*(E \setminus U) \quad \text{for all } E \subset X \text{ such that } \mu^*(E) < \infty.$$

At this point, it follows from an application of Carathéodory's theorem [Fo2, Theorem 1.11] that every Borel set is μ^*-measurable and further that the restriction μ of μ^* to the σ-algebra $\mathcal{B}_X$ is a Borel measure. It should be emphasized that $\mu^*(U) = \mu(U)$ if U is open and that the measure μ is outer regular and satisfies condition (3.18).

(iii) Thirdly, we prove that the measure μ satisfies the condition (3.18).

This implies that μ is finite on compact subsets of X and is *inner regular* on open subsets U of X:

$$\mu(U) = \sup\{\mu(K) : K \subset U, \ K \text{ is compact}\}.$$

Indeed, if U is open and if α is an arbitrary number satisfying the condition $\alpha < \mu(U)$, then we can choose a function $f \in C_c(X)$ such that $f \prec U$ and that $I(f) > \alpha$. We let

$$K = \text{supp } f.$$

If g is a function in $C_c(X)$ satisfying the condition $g \geq \chi_K$, then it follows that

$$g - f \geq 0$$

so that, by the positivity of I,

$$I(g) \geq I(f) > \alpha.$$

However, we have, by formula (3.19),

$$\mu(K) > \alpha.$$

This proves that μ is inner regular on open sets, since α is an arbitrary number satisfying the condition $\alpha < \mu(U)$.

(iv) Finally, we prove (3.17).

Proof of Assertion (i): It suffices to show that if $\{U_j\}$ is a sequence of open sets in X and if $U = \cup_{j=1}^{\infty} U_j$, then we have the inequality

$$\mu(U) \le \sum_{j=1}^{\infty} \mu(U_j).$$

Indeed, it follows from this inequality that we have, for any subset $E \subset X$,

$$\mu^*(E) = \inf\left\{ \sum_{j=1}^{\infty} \mu(U_j) : E \subset \bigcup_{j=1}^{\infty} U_j,\ U_j \text{ is open} \right\},$$

and further [Fo2, Proposition 1.10] that the expression on the right-hand side defines an outer measure.

If $U = \cup_{j=1}^{\infty} U_j$ and if $f \in C_c(X)$ such that $f \prec U$, then we let

$$K = \text{supp } f.$$

Since K is compact, it follows that, for some finite n,

$$K \subset \bigcup_{j=1}^{n} U_j.$$

Moreover, we can find functions $g_1, g_2, \ldots, g_n \in C_c(X)$ such that $g_j \prec U_j$ and $\sum_{j=1}^{n} g_j = 1$ on K (a partition of unity subordinate to the covering $\{U_j\}$). However, since we have

$$f = \sum_{j=1}^{n} f\, g_j, \quad f\, g_j \prec U_j,$$

we obtain that, for any function $f \prec U$,

$$I(f) = \sum_{j=1}^{n} I(f g_j) \le \sum_{j=1}^{n} \mu(U_j) \le \sum_{j=1}^{\infty} \mu(U_j),$$

so that

$$\mu(U) = \sup\{I(f) : f \in C_c(X),\ f \prec U\} \le \sum_{j=1}^{\infty} \mu(U_j).$$

Proof of Assertion (ii): It suffices to show that

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U) \qquad (3.20)$$
$$\text{for all } E \subset X \text{ such that } \mu^*(E) < \infty.$$

First, we consider the case where E is open. Then, for any given $\varepsilon > 0$ we can find a function $f \in C_c(X)$ such that

$$\begin{cases} f \prec E \cap U, \\ I(f) > \mu(E \cap U) - \varepsilon. \end{cases}$$

Moreover, since the set $E \setminus \text{supp}\, f$ is also open, we can find a function $g \in C_c(X)$ such that

$$\begin{cases} g \prec E \setminus \text{supp}\, f, \\ I(g) > \mu(E \setminus \text{supp}\, f) - \varepsilon. \end{cases}$$

However, we have

$$f + g \prec E, \quad \text{supp}\, f \subset U,$$

and so

$$\begin{aligned} \mu(E) &\geq I(f) + I(g) > \mu(E \cap U) + \mu(E \setminus \text{supp}\, f) - 2\varepsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U) - 2\varepsilon. \end{aligned}$$

Therefore, by letting $\varepsilon \downarrow 0$ in this inequality we obtain the desired inequality (3.20).

Secondly, we consider the general case where $\mu^*(E) < \infty$. Then, for any given $\varepsilon > 0$ we can find an open subset $V \supset E$ such that

$$\mu(V) < \mu^*(E) + \varepsilon.$$

Hence it follows that

$$\begin{aligned} \mu^*(E) + \varepsilon > \mu(V) &\geq \mu^*(V \cap U) + \mu^*(V \setminus U) \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U). \end{aligned}$$

Therefore, by letting $\varepsilon \downarrow 0$ in this inequality we obtain the desired inequality (3.20).

Proof of Assertion (iii): Let K be an arbitrary compact subset of X, and let $f \in C_c(X)$ such that $f \geq \chi_K$. If ε is an arbitrary positive number, we define an open set U_ε as follows:

$$U_\varepsilon = \{x \in X : f(x) > 1 - \varepsilon\}.$$

Then it follows that we have, for any function $g \prec U_\varepsilon$,

$$\frac{1}{1-\varepsilon} f - g \geq 0,$$

and so, by the positivity of I,

$$I(g) \leq \frac{1}{1-\varepsilon} I(f).$$

Hence we have the inequality

$$\mu(K) \leq \mu(U_\varepsilon) = \sup\{I(g) : g \in C_c(X),\ g \prec U_\varepsilon\} \leq \frac{1}{1-\varepsilon} I(f).$$

Therefore, by letting $\varepsilon \downarrow 0$ in this inequality we obtain that

$$\mu(K) \leq I(f).$$

This proves that we have, for every compact set $K \subset X$,

$$\mu(K) \leq \inf\{I(f) : f \in C_c(X),\ f \geq \chi_K\}. \tag{3.21}$$

On the other hand, for any open set $U \supset K$, by using Urysohn's lemma (Lemma 3.35) we can find a function $h \in C_c(X)$ such that $h \geq \chi_K$ and $h \prec U$. Hence we have the inequality

$$I(h) \leq \mu(U) = \sup\{I(f) : f \in C_c(X),\ f \prec U\}.$$

However, since μ is outer regular on K, it follows that

$$\mu(K) = \inf\{\mu(U) : U \supset K,\ U \text{ is open}\}.$$

Hence we have proved that

$$I(h) \leq \mu(K).$$

This proves that

$$\inf\{I(f) : f \in C_c(X),\ f \geq \chi_K\} \leq I(h) \leq \mu(K). \tag{3.22}$$

Therefore, the desired formula (3.19) follows by combining inequalities (3.21) and (3.22).

Proof of Assertion (iv): To do this, we have only to show that

$$I(f) = \int_X f(x)\,d\mu \quad \text{for all } f \in C_c(X, [0, 1]).$$

Indeed, it suffices to note that the space $C_c(X)$ is the linear span of functions in the space $C_c(X, [0, 1])$.

For any positive integer $N \in \mathbf{N}$, we let

$$K_j = \left\{ x \in X : f(x) \geq \frac{j}{N} \right\}, \quad 1 \leq j \leq N,$$

and

$$K_0 = \operatorname{supp} f.$$

Moreover, we define functions $f_1, f_2, \ldots, f_N \in C_c(X, [0, 1])$ by the formulas

$$f_j(x) = \min \left\{ \max \left\{ f(x) - \frac{j-1}{N}, 0 \right\}, \frac{1}{N} \right\}, \quad 1 \leq j \leq N.$$

Here it should be noted that

$$f_j(x) = \begin{cases} 0 & \text{if } x \notin K_{j-1}, \\ f(x) - \frac{j-1}{N} & \text{if } x \in K_{j-1} \setminus K_j, \\ \frac{1}{N} & \text{if } x \in K_j. \end{cases}$$

Then it follows that

$$\frac{1}{N} \chi_{K_j} \leq f_j \leq \frac{1}{N} \chi_{K_{j-1}},$$

so that

$$\begin{aligned} \frac{1}{N} \mu(K_j) &= \frac{1}{N} \int_X \chi_{K_j}(x)\, d\mu \leq \int_X f_j(x)\, d\mu \\ &\leq \frac{1}{N} \int_X \chi_{K_{j-1}}(x)\, d\mu = \frac{1}{N} \mu(K_{j-1}). \end{aligned} \tag{3.23}$$

Also, if U is an open set containing K_{j-1}, then we have the condition

$$N f_j \prec U,$$

and the inequality

$$I(f_j) \leq \frac{\mu(U)}{N}.$$

Hence, by formula (3.19) and outer regularity of μ it follows that

$$\begin{aligned}\frac{1}{N}\mu(K_j) &= \frac{1}{N}\inf\left\{I(f) : f \in C_c(X),\ f \geq \chi_{K_j}\right\} \leq I(f_j) \\ &\leq \frac{1}{N}\inf\left\{\mu(U) : U \supset K_{j-1},\ U \text{ is open}\right\} = \frac{1}{N}\mu(K_{j-1}).\end{aligned} \tag{3.24}$$

However, since we have

$$f = \sum_{j=1}^{N} f_j,$$

it follows from inequalities (3.23) and (3.24) that

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq \sum_{j=1}^{N}\int_X f_j(x)\,d\mu = \int_X f(x)d\mu \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j),$$

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq \sum_{j=1}^{N} I(f_j) = I(f) \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j).$$

Hence we have the inequalities

$$\left|I(f) - \int_X f(x)\,d\mu\right| \leq \frac{\mu(K_0) - \mu(K_N)}{N} \leq \frac{\mu(\operatorname{supp} f)}{N}.$$

Therefore, by letting $N \to \infty$ in this inequality we obtain the desired formula (3.17), since $\mu(\operatorname{supp} f) < \infty$.

Now the proof of Theorem 3.40 is complete.

We recall (Proposition 3.38) that $C_0(X)$ is the uniform closure of $C_c(X)$. Hence we find that if μ is a Radon measure on X, then the linear functional

$$I_\mu : f \longmapsto \int_X f(x)\,d\mu(x)$$

extends continuously to $C_0(X)$ if and only if it is bounded with respect to the uniform norm. This happens only when

$$\mu(X) = \sup\{I(f) : f \in C_c(X),\ 0 \leq f \leq 1 \text{ on } X\} < \infty,$$

in which case $\mu(X)$ is the operator norm $\|I\|$ of F.

Therefore, we have the following locally compact version of the Riesz–Markov representation theorem:

Theorem 3.41 (Riesz–Markov). *Let (X, ρ) be a locally compact metric space. If F is a non-negative linear functional on the space $C_0(X)$, then there exists a unique Radon measure μ on X such that*

$$F(f) = \int_X f(x)\, d\mu(x) \quad \text{for all } f \in C_0(X),$$

and we have

$$\mu(X) = \sup\left\{ \int_X f(x)\, d\mu(x) : f \in C_0(X),\ 0 \leq f \leq 1 \text{ on } X \right\} = \|F\|.$$

Corollary 3.42. *Let (K, ρ) be a* compact *metric space. Then we have the following two assertions (i) and (ii):*

(i) *To each non-negative linear functional T on $C(K)$, there corresponds a unique Radon measure μ on K such that*

$$T(f) = \int_K f(x)\, d\mu(x) \quad \text{for all } f \in C(K), \tag{3.25}$$

and we have

$$\|T\| = \mu(K). \tag{3.26}$$

(ii) *Conversely, every finite Radon measure μ on K defines a non-negative linear functional T on $C(K)$ through formula (3.25), and relation (3.26) holds true.*

Remark 3.43. It is easy to see that every open set in a compact metric space is a σ-compact. Thus we find that every finite Radon measure μ is regular.

Now we can characterize the space of all bounded linear functionals on $C(K)$, that is, the dual space $C(K)'$ of $C(K)$. Recall (see formula (3.7)) that the dual space $C(K)'$ is a Banach space with the operator norm

$$\|T\| = \sup_{\substack{f \in C(K) \\ \|f\| \leq 1}} |Tf|.$$

The compact version of the Riesz–Markov representation theorem reads as follows:

Theorem 3.44 (Riesz–Markov). *Let (K, ρ) be a* compact *metric space. Then we have the following two assertions (i) and (ii):*

(i) *To each $T \in C(K)'$, there corresponds a unique real Borel measure μ on K such that (3.25) holds true for all $f \in C(K)$, and we have*

$$\|T\| = \text{the total variation } |\mu|(K) \text{ of } \mu. \tag{3.27}$$

(ii) Conversely, every real Borel measure μ on K defines a bounded linear functional $T \in C(K)'$ through formula (3.25), and relation (3.27) holds true.

Remark 3.45. The positive and negative variation measures μ^+, μ^- of a real Borel measure μ are both regular.

We recall that the space of all real Borel measures μ on K is a normed linear space with the norm

$$\|\mu\| = \text{the total variation } |\mu|(K) \text{ of } \mu. \tag{3.28}$$

Therefore, we can restate Theorem 3.44 as follows:

Theorem 3.46. *The dual space $C(K)'$ of $C(K)$ can be identified with the space of all real Borel measures on K normed by (3.28).*

3.4.4 Weak Convergence of Measures

Let K be a compact metric space and let $C(K)$ be the Banach space of real-valued continuous functions on K with the supremum (maximum) norm

$$\|f\|_\infty = \sup_{x \in K} |f(x)|.$$

A sequence $\{\mu_n\}_{n=1}^{\infty}$ of real Borel measures on K is said to *converge weakly* to a real Borel measure μ on K if it satisfies the condition

$$\lim_{n\to\infty} \int_K f(x)\, d\mu_n(x) = \int_K f(x)\, d\mu(x) \quad \text{for every } f \in C(K). \tag{3.29}$$

Theorem 3.46 asserts that the space of all real Borel measures on K normed by formula (3.28) can be identified with the strong dual space $C(K)'$ of $C(K)$. Thus the weak convergence (3.29) of real Borel measures is just the weak* convergence of $C(K)'$.

One more result is important when studying the weak convergence of measures:

Theorem 3.47. *The Banach space $C(K)$ is separable, that is, it contains a countable, dense subset.*

The next theorem is one of the fundamental theorems in measure theory:

Theorem 3.48. *Every sequence $\{\mu_n\}_{n=1}^{\infty}$ of real Borel measures on K satisfying the condition*

$$\sup_{n\geq 1} |\mu_n|(K) < +\infty \tag{3.30}$$

has a subsequence which converges weakly to a real Borel measure μ *on* K. *Furthermore, if the measures* μ_n *are all non-negative, then the measure* μ *is also non-negative.*

Proof. By virtue of Theorem 3.46, we can apply Theorem 3.33 with $X := C(K)$ to obtain the first assertion, since condition (3.30) implies the boundedness of the Borel measures μ_n. The second assertion is an immediate consequence of the first assertion of Corollary 3.42.

3.5 Closed Operators

Let X and Y be normed linear spaces over the same scalar field. Let T be a linear operator from X into Y with domain $D(T)$. The *graph* $G(T)$ of T is the set

$$G(T) = \{\{x, Tx\} : x \in D(T)\}$$

in the product space $X \times Y$. Note that $G(T)$ is a linear subspace of $X \times Y$. We say that T is *closed* if its graph $G(T)$ is closed in $X \times Y$. This is equivalent to saying that

$$\begin{aligned} &\{x_n\} \subset D(T),\ x_n \longrightarrow x \text{ in } X,\ Tx_n \longrightarrow y \text{ in } Y \\ &\Longrightarrow\ x \in D(T),\ Tx = y. \end{aligned}$$

In particular, if T is continuous and its domain $D(T)$ is closed in X, then T is a closed linear operator.

We remark that if T is a closed linear operator which is also injective, then its inverse T^{-1} is a closed linear operator. Indeed, this follows from the fact that the mapping $\{x, y\} \longmapsto \{y, x\}$ is a homeomorphism of $X \times Y$ onto $Y \times X$.

A linear operator T is said to be *closable* if the closure $\overline{G(T)}$ in $X \times Y$ of $G(T)$ is the graph of a linear operator, say, $\overline{T}$, that is, $\overline{G(T)} = G(\overline{T})$.

A linear operator is called a *closed extension* of T if it is a closed linear operator which is also an extension of T. It is easy to see that if T is closable, then every closed extension of T is an extension of $\overline{T}$. Thus the operator $\bar{T}$ is called the minimal closed extension of T.

The next theorem gives a necessary and sufficient condition for a linear operator to be closable:

Theorem 3.49. *A linear operator* T *from* X *into* Y *with domain* $D(T)$ *is closable if and only if the following condition is satisfied:*

$$\{x_n\} \subset D(T),\ x_n \longrightarrow 0 \textit{ in } X,\ Tx_n \longrightarrow y \textit{ in } Y \ \Longrightarrow\ y = 0.$$

Now we state two important theorems concerning closed linear operators:

Theorem 3.50 (Banach's open mapping theorem). *Let X and Y be Banach spaces. Then every continuous linear operator from X onto Y is open, that is, it maps every open set in X onto an open set in Y.*

Theorem 3.51 (Banach's closed graph theorem). *Let X and Y be Banach spaces. Then every closed linear operator from X into Y is continuous.*

Corollary 3.52. *Let X and Y be Banach spaces. If T is a continuous, one-to-one linear operator from X onto Y, then its inverse T^{-1} is also continuous; hence T is an isomorphism.*

Indeed, the inverse T^{-1} is a closed linear operator, so that Theorem 3.51 applies.

We give useful characterizations of closed linear operators with closed range:

Theorem 3.53. *Let X and Y be Banach spaces and T a closed linear operator from X into Y with domain $D(T)$. Then the range $R(T)$ of T is closed in Y if and only if there exists a positive constant C such that*

$$\operatorname{dist}(x, N(T)) \leq C \|Tx\| \quad \text{for all } x \in D(T).$$

Here

$$\operatorname{dist}(x, N(T)) = \inf_{z \in N(T)} \|x - z\|$$

is the distance from x to the null space $N(T)$ of T.

Theorem 3.54 (Banach's closed range theorem). *Let X and Y be Banach spaces and T a densely defined, closed linear operator from X into Y. Then the following four conditions (i)–(iv) are equivalent:*

(i) *The range $R(T)$ of T is closed in Y.*
(ii) *The range $R(T')$ of the transpose T' is closed in X'.*
(iii) $R(T) = {}^0N(T') = \{x \in X : \langle x, x' \rangle = 0 \text{ for all } x' \in N(T')\}$.
(iv) $R(T') = {}^0N(T) = \{x' \in X' : \langle x', x \rangle = 0 \text{ for all } x \in N(T)\}$.

3.6 Complemented Subspaces

Let X be a linear space. Two linear subspaces M and N of X are said to be *algebraic complements* in X if X is the direct sum of M and N, that is, if $X = M \dot{+} N$. Algebraic complements M and N in a normed linear space X are said to be *topological complements* in X if the addition mapping

$$\{y, z\} \longmapsto y + z$$

is an isomorphism of $M \times N$ onto X. We then write

$$X = M \oplus N.$$

As an application of Corollary 3.52, we obtain the following theorem:

Theorem 3.55. *Let X be a Banach space. If M and N are closed algebraic complements in X, then they are topological complements.*

A closed linear subspace of a normed linear space X is said to be *complemented* in X if it has a topological complement. By Theorem 3.55, this is equivalent in Banach spaces to the existence of a closed algebraic complement.

The next theorem gives two criteria for a closed subspace to be complemented:

Theorem 3.56. *Let X be a Banach space and M a closed subspace of X. If M has either finite dimension or finite codimension, that is, if either (i)* $\dim M < \infty$ *or (ii)* $\operatorname{codim} M = \dim X/M < \infty$*, then it is complemented in X.*

3.7 Compact Operators

For a compact operator T on a Banach space, the eigenvalue problem can be treated fairly completely in the sense that the classical theory of Fredholm integral equations may be extended to the linear functional equation $Tx - \lambda x = y$ with a complex parameter λ. This result is known as the Riesz–Schauder theory.

3.7.1 Definition and Basic Properties of Compact Operators

Let X and Y be normed linear spaces over the same scalar field $\mathbf{K}$. A linear operator T from X into Y is said to be *compact* or *completely continuous* if it maps every bounded subset of X onto a relatively compact subset of Y, that is, if the closure of $T(B)$ is compact in Y for every bounded subset B of X. This is equivalent to saying that, for every bounded sequence $\{x_n\}_{n=1}^{\infty}$ in X, the sequence $\{Tx_n\}_{n=1}^{\infty}$ has a subsequence which converges in Y.

We list some basic facts which follow at once:

(i) Every compact operator is bounded.
Indeed, a compact operator maps the unit sphere onto a bounded set.

(ii) Every bounded linear operator with finite dimensional range is compact.
This is an immediate consequence of Corollary 3.19.

(iii) No isomorphism between infinite dimensional spaces is compact.
This follows from an application of Theorem 3.20.

(iv) A linear combination of compact operators is compact.
(v) The product of a compact operator with a bounded operator is compact.

The next theorem states that if Y is a Banach space, then the compact operators form a closed subspace of the space $\mathcal{L}(X, Y)$ of bounded linear operators:

Theorem 3.57. *Let X be a normed linear space and Y a Banach space. If $\{T_n\}_{n=1}^{\infty}$ is a sequence of compact linear operators which converges to an operator T in the space $\mathcal{L}(X, Y)$ with the uniform topology, then the limit operator T is compact.*

As for the transposes of compact operators, we have the following theorem:

Theorem 3.58. *Let X and Y be normed linear spaces. If T is a compact linear operator from X into Y, then its transpose T' is a compact linear operator from Y' into X'.*

3.7.2 The Riesz–Schauder Theory

Now we state the most interesting results on compact linear operators, which are essentially due to F. Riesz in the Hilbert space setting. The results were extended to Banach spaces by Schauder:

Theorem 3.59. *Let X be a Banach space and T a compact linear operator from X into itself. Set*

$$S = I - T.$$

Then we have the following three assertions (i)–(iii):

(i) The null space $N(S)$ of S is finite dimensional and the range $R(S)$ of S is closed in X.
(ii) The null space $N(S')$ of the transpose S' is finite dimensional and the range $R(S')$ of S' is closed in X'.
(iii) $\dim N(S) = \dim N(S')$.

The next result is an extension of the theory of linear mappings for finite dimensional linear spaces:

Corollary 3.60 (The Fredholm alternative). *Let T be a compact linear operator from a Banach space X into itself. If $S = I - T$ is either one-to-one or onto, then it is an isomorphism of X onto itself.*

Let T be a bounded linear operator from X into itself. The *resolvent set* of T, denoted $\rho(T)$, is defined to be the set of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is an isomorphism of X onto itself. In this case, the inverse $(\lambda I - T)^{-1}$ is called the *resolvent* of T. The complement of $\rho(T)$, that is, the set of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is not an isomorphism of X onto itself is called the *spectrum* of T, and is denoted by $\sigma(T)$.

The set $\sigma_p(T)$ of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is not one-to-one forms a subset of $\sigma(T)$, and is called the *point spectrum* of T. A scalar $\lambda \in \mathbf{K}$ belongs to $\sigma_p(T)$ if and only if there exists a non-zero element $x \in X$ such that $Tx = \lambda x$. In this case, λ is called an *eigenvalue* of T and x an *eigenvector* of T corresponding to λ, respectively. Also the null space $N(\lambda I - T)$ of $\lambda I - T$ is called the *eigenspace* of T corresponding to λ, and the dimension of $N(\lambda I - T)$ is called the *multiplicity* of λ.

By using C. Neumann's series (Theorem 3.12), we find that the resolvent set $\rho(T)$ is open in $\mathbf{K}$ and that

$$\{\lambda \in \mathbf{K} : |\lambda| > \|T\|\} \subset \rho(T).$$

Hence the spectrum $\sigma(T) = \mathbf{K} \setminus \rho(T)$ is closed and bounded in $\mathbf{K}$.

If T is a compact operator and λ is a non-zero element of $\sigma(T)$, then, by applying Corollary 3.60 to the operator $\lambda^{-1}T$ we obtain that $\lambda I - T$ is not one-to-one, that is, $\lambda \in \sigma_p(T)$. Also note that if X is infinite dimensional, then T is not an isomorphism of X onto itself; hence $0 \in \sigma(T)$. Therefore, the scalar field $\mathbf{K}$ can be decomposed as follows:

$$\mathbf{K} = \big(\sigma_p(T) \cup \{0\}\big) \cup \rho(T).$$

We can say rather more about the spectrum $\sigma(T)$:

Theorem 3.61 (Riesz–Schauder). *Let T be a compact linear operator from a Banach space X into itself. Then we have the following three assertions (i)–(iii):*

(i) *The spectrum $\sigma(T)$ of T is either a finite set or a countable set accumulating only at 0; and every non-zero element of $\sigma(T)$ is an eigenvalue of T.*
(ii) $\dim N(\lambda I - T) = \dim N(\overline{\lambda} - T') < \infty$ *for all* $\lambda \neq 0$.
(iii) *Let $\lambda \neq 0$. The equation*

$$(\lambda I - T)x = y$$

has a solution if and only if y is orthogonal to the space $N(\overline{\lambda} - T')$. Similarly, the equation

$$(\overline{\lambda} I - T')z = w$$

has a solution if and only if w is orthogonal to the space $N(\lambda I - T)$. Moreover, the operator $\lambda I - T$ is onto if and only if it is one-to-one.

3.8 Fredholm Operators

Throughout this section, the letters X, Y, Z denote Banach spaces over the same scalar field. The Fredholm property of $T : X \to Y$ conveys that the operator T is both "almost" injective and "almost" surjective, that is, it is "almost" an isomorphism.

3.8.1 Definition and Basic Properties of Fredholm Operators

A linear operator T from X into Y is called a *Fredholm operator* if the following five conditions (i)–(v) are satisfied:

(i) The domain $D(T)$ of T is dense in X.
(ii) T is a closed operator.
(iii) The null space $N(T) = \{x \in D(T) : Tx = 0\}$ of T has finite dimension, that is, $\dim N(T) < \infty$.
(iv) The range $R(T) = \{Tx : x \in D(T)\}$ of T is closed in Y.
(v) The range $R(T)$ of T has finite codimension, that is,

$$\operatorname{codim} R(T) = \dim Y/R(T) < \infty.$$

Then the *index* of T is defined by the formula

$$\operatorname{ind} T := \dim N(T) - \operatorname{codim} R(T).$$

Roughly speaking, the index $\operatorname{ind} T$ indicates how far the operator $T : X \to Y$ is from being bijective. Namely, the further $\operatorname{ind} T$ is from zero, the more bijective T is.

First, we give a characterization of Fredholm operators:

Theorem 3.62. *If T is a Fredholm operator from X into Y with domain $D(T)$, then there exist a bounded linear operator $S : Y \to X$ and compact linear operators $P : X \to X$, $Q : Y \to Y$ such that*

(a) $ST = I - P$ on $D(T)$.
(b) $TS = I - Q$ on Y.

Furthermore, we have

$$R(P) = N(T),$$
$$\dim R(Q) = \operatorname{codim} R(T) = \dim Y/R(T).$$

Theorem 3.62 has a converse:

Theorem 3.63. *Let T be a closed linear operator from X into Y with domain $D(T)$ dense in X. Assume that there exist bounded linear operators $S_1 : Y \to X$, $S_2 : Y \to X$ and compact linear operators $K_1 : X \to X$, $K_2 : Y \to Y$ such that*

(a) $S_1 T = I - K_1$ *on* $D(T)$.
(b) $TS_2 = I - K_2$ *on* Y.

Then T is a Fredholm operator.

Secondly, we state an important property of the product of Fredholm operators:

Theorem 3.64. *If T is a Fredholm operator from X into Y and if S is a Fredholm operator from Y into Z, then the product ST is a Fredholm operator from X into Z. Moreover, we have*

$$\operatorname{ind}(ST) = \operatorname{ind} S + \operatorname{ind} T.$$

As for the transposes of Fredholm operators, we have the following theorem:

Theorem 3.65. *If T is a Fredholm operator from X into Y and if Y is reflexive, then the transpose T' of T is a Fredholm operator from Y' into X'. Moreover, we have*

$$\operatorname{ind} T' = -\operatorname{ind} T.$$

3.8.2 Stability Theorem for Indices of Fredholm Operators

By combining Theorems 3.59 and 3.54 (or Theorem 3.61), we obtain that if $X = Y$ and T is compact, then the operator $I - T$ is a Fredholm operator and ind $(I - T) = 0$.

More precisely, the next theorem asserts that the index $\operatorname{ind} T$ is stable under compact perturbations and small perturbations in the space $\mathcal{L}(X, Y)$ of bounded linear operators on X into Y:

Theorem 3.66. *Let T be a Fredholm operator from X into Y. Then we have the following two assertions (i) and (ii):*

(i) *If K is a compact linear operator on X into Y, then the sum $T + K$ is a Fredholm operator, and we have*

$$\operatorname{ind}(T + K) = \operatorname{ind} T.$$

(ii) *There exists a constant $\varepsilon > 0$ such that if an operator $S \in \mathcal{L}(X, Y)$ satisfies the condition $\|S\| < \varepsilon$, then the sum $T + S$ is a Fredholm operator, and we have*

$$\operatorname{ind}(T + S) = \operatorname{ind} T.$$

3.9 Notes and Comments

For more thorough treatments of functional analysis, the reader might be referred to Friedman [Fr2], Rudin [Ru], Schechter [Sh] and Yosida [Yo].

Sections 3.1 and 3.2: For the theory of linear topological spaces, see Treves [Tv].

Section 3.3: The method of continuity (Theorem 3.13) is adapted from Gilberg–Trudinger [GT, Theorem 5.2]. The Hahn–Banach extension theorem (Theorem 3.21) is taken from Yosida [Yo, Chapter IV, Section 1, Theorem; Section 4, Theorem].

Section 3.4: Proposition 3.38 is taken from Folland [Fo2, Proposition 4.35] and the proof of Theorem 3.40 is adapted from Folland [Fo2, Theorem 7.2]. The locally compact version of the Riesz–Markov representation theorem (Theorem 3.41) is taken from Folland [Fo2, Theorem 7.17] and the compact version of the Riesz–Markov representation theorem (Theorem 3.44) is taken from Folland [Fo2, Corollary 7.18], respectively.

Sections 3.5–3.7: The material in these sections is adapted from Friedman [Fr2] and Yosida [Yo].

Section 3.8: For further material on Fredholm operators, see Gohberg–Kreĭn [GK] and Palais [Pl].

The following diagram gives a bird's eye view of Linear Algebra, Integral Equations and Functional Analysis and how these relate to each other:

Linear Algebra	**Integral Equations**	**Functional Analysis**
Finite-dimensional vector spaces	Spaces of continuous functions	Banach spaces (of infinite-dimension)
Vectors $\mathbf{x}$	Functions $f(x)$	Vectors x
Matrices $A = (a_{ij})$	Integral kernels $K(x, y)$	Linear operators T
Linear systems of equations $A\mathbf{x} = \mathbf{y}$	Integral equations $\int K(x, y)f(y)dy = g(x)$	Linear equations $Tx = y$
Identity matrices $E = (\delta_{ij})$	Dirac measures $\delta(x - y)$	Identity operators I
Inverse matrices $AA^{-1} = A^{-1}A = E$	Green functions $\int K(x, z)G(z, y)dz = \delta(x - y)$	Inverse operators $TT^{-1} = T^{-1}T = I$

Chapter 4
Theory of Semigroups

This chapter is devoted to the general theory of semigroups. These topics form the necessary background for the proof of Theorems 1.2 and 1.3. In Sects. 4.1–4.3 we study Banach space valued functions, operator valued functions and exponential functions, generalizing the numerical case. Section 4.4 is devoted to the theory of contraction semigroups. A typical example of contraction semigroups is the semigroup associated with the heat kernel (Example 4.11). We consider when a linear operator is the infinitesimal generator of some contraction semigroup. This question is answered by the Hille–Yosida theorem (Theorem 4.10). In Sect. 4.5 we consider when a linear operator is the infinitesimal generator of some (C_0) semigroup (Theorem 4.28), generalizing the theory of contraction semigroups developed in Sect. 4.4. Moreover, we study an initial-value problem associated with a (C_0) semigroup, and prove an existence and uniqueness theorem (Theorem 4.30).

4.1 Banach Space Valued Functions

Let E be a Banach space over the real or complex number field, equipped with a norm $\|\cdot\|$. A function $u(t)$ defined on an interval I with values in E is said to be *strongly continuous* at a point t_0 of I if it satisfies the condition

$$\lim_{t \to t_0} \|u(t) - u(t_0)\| = 0.$$

If $u(t)$ is strongly continuous at every point of I, then it is said to be strongly continuous on I. If $u(t)$ is strongly continuous on I, then the function $\|u(t)\|$ is continuous on I and also, for any f in the dual space E' of E, the function $f(u(t))$ is continuous on I.

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_4,

As in the case of scalar valued functions, the following two results (1) and (2) hold true:

(1) If $u(t)$ is strongly continuous on a bounded closed interval I, then it is uniformly strongly continuous on I.
(2) If a sequence $\{u_n(t)\}$ of strongly continuous functions on I converges uniformly strongly to a function $u(t)$ on I, then the limit function $u(t)$ is strongly continuous on I.

If $u(t)$ is a strongly continuous function on I such that

$$\int_I \|u(t)\| \, dt < \infty, \tag{4.1}$$

then the Riemann integral

$$\int_I u(t) \, dt$$

can be defined just as in the case of scalar valued functions. Then we say that the function $u(t)$ is *strongly integrable* on I. By the triangle inequality, we have the inequality

$$\left\| \int_I u(t) \, dt \right\| \leq \int_I \|u(t)\| \, dt.$$

Furthermore, we easily obtain the following theorem (cf. [Tn, Chapter 7, Theorem 7-1]):

Theorem 4.1. *Let $u(t)$ be a strongly continuous function defined on an interval I which satisfies condition (4.1), and let T be a bounded linear operator on E into itself. Then the function $Tu(t)$ is strongly integrable on I, and we have*

$$T\left(\int_I u(t) \, dt\right) = \int_I Tu(t) \, dt.$$

Similarly, we have, for any $f \in E'$,

$$f\left(\int_I u(t) \, dt\right) = \int_I f(u(t)) \, dt.$$

As in the case of scalar valued functions, the following two results (3) and (4) hold true:

(3) If a sequence $\{u_n(t)\}$ of strongly continuous functions on a bounded closed interval I converges uniformly strongly to a function $u(t)$ on I, then the limit function $u(t)$ is strongly integrable on I, and we have

$$\int_I u(t)\,dt = \lim_{n\to\infty} \int_I u_n(t)\,dt.$$

(4) If $u(t)$ is strongly continuous in a neighborhood of a point t_0 of I, then we have

$$\lim_{h\to 0} \left\| \frac{1}{h} \int_{t_0}^{t_0+h} u(t)\,dt - u(t_0) \right\| = 0.$$

A function $u(t)$ defined on an open interval I is said to be *strongly differentiable* at a point t_0 of I if the limit

$$\lim_{h\to 0} \frac{u(t_0+h) - u(t_0)}{h} \tag{4.2}$$

exists in E. The value of (4.2) is denoted by

$$\frac{du}{dt}(t_0) \text{ or } u'(t_0).$$

If $u(t)$ is strongly differentiable at every point of I, then it is said to be strongly differentiable on I. A strongly differentiable function is strongly continuous.

As in the case of scalar valued functions, the following two results (5) and (6) hold true:

(5) If $u(t)$ is strongly differentiable on I and $u'(t)$ is strongly continuous on I, then we have, for any $a, b \in I$,

$$u(b) - u(a) = \int_a^b u'(t)\,dt.$$

(6) If $u(t)$ is strongly continuous on I, then, for each $c \in I$, the integral $\int_c^t u(s)\,ds$ is strongly differentiable on I, and we have

$$\frac{d}{dt}\left(\int_c^t u(s)\,ds\right) = u(t).$$

4.2 Operator Valued Functions

Let $\mathcal{L}(E, E)$ be the space of all bounded linear operators on a Banach space E into itself. The space $\mathcal{L}(E, E)$ is a Banach space with the operator norm

$$\|T\| = \sup_{\substack{x\in E\\ x\neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x\in E\\ \|x\|\leq 1}} \|Tx\|.$$

A function $T(t)$ defined on an interval I with values in the space $\mathcal{L}(E, E)$ is said to be *strongly continuous* at a point t_0 of I if it satisfies the condition

$$\lim_{t \to t_0} \|T(t)x - T(t_0)x\| = 0 \quad \text{for every } x \in E.$$

We say that $T(t)$ is *norm continuous* at t_0 if it satisfies the condition

$$\lim_{t \to t_0} \|T(t) - T(t_0)\| = 0.$$

If $T(t)$ is strongly (resp. norm) continuous at every point of I, then it is said to be strongly (resp. norm) continuous on I. A norm continuous function is strongly continuous.

The next theorem is an immediate consequence of the principle of uniform boundedness (see Theorem 3.14):

Theorem 4.2. *If $T(t)$ is strongly continuous on I, then the function $\|T(t)\|$ is bounded uniformly in t over bounded closed intervals contained in I.*

A function $T(t)$ defined on an open interval I is said to be *strongly differentiable* at a point t_0 of I if there exists an operator $S(t_0)$ in $\mathcal{L}(E, E)$ such that

$$\lim_{h \to 0} \left\| \left(\frac{T(t_0 + h) - T(t_0)}{h} \right) x - S(t_0)x \right\| = 0 \quad \text{for every } x \in E.$$

We say that $T(t)$ is *norm differentiable* at t_0 if it satisfies the condition

$$\lim_{h \to 0} \left\| \frac{T(t_0 + h) - T(t_0)}{h} - S(t_0) \right\| = 0.$$

The operator $S(t_0)$ is denoted by

$$\frac{dT}{dt}(t_0) \text{ or } T'(t_0).$$

If $T(t)$ is strongly (resp. norm) differentiable at every point of I, then it is said to be strongly (resp. norm) differentiable on I. A norm differentiable function is strongly differentiable.

It should be emphasized that the *Leibniz formula* can be extended to strongly or norm differentiable functions:

Theorem 4.3. *(i) If $u(t)$ and $T(t)$ are both strongly continuous (resp. differentiable) on I, then the function $T(t)u(t)$ is also strongly continuous (resp. differentiable) on I. In the differentiable case, we have the formula*

$$\frac{d}{dt}(T(t)u(t)) = \frac{dT}{dt}(t)u(t) + T(t)\frac{du}{dt}(t).$$

(ii) If $T(t)$ and $S(t)$ are both norm (resp. strongly) differentiable on I, then the function $S(t)T(t)$ is also norm (resp. strongly) differentiable on I, and we have the formula

$$\frac{d}{dt}(S(t)T(t)) = \frac{dS}{dt}(t)T(t) + S(t)\frac{dT}{dt}(t).$$

4.3 Exponential Functions

Let E be a Banach space and $\mathcal{L}(E, E)$ the space of all bounded linear operators on E into itself. Just as in the case of numerical series, we have the following theorem:

Theorem 4.4. *If $A \in \mathcal{L}(E, E)$, we let*

$$e^{tA} = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m \quad \text{for every } t \in \mathbf{R}. \tag{4.3}$$

Then it follows that the right-hand side converges in the Banach space $\mathcal{L}(E, E)$, and enjoys the following three properties (a)–(c):

(a) $\|e^{tA}\| \leq e^{|t|\|A\|}$ *for all* $t \in \mathbf{R}$.
(b) $e^{tA}e^{sA} = e^{(t+s)A}$ *for all* $t, s \in \mathbf{R}$.
(c) The exponential function e^{tA} is norm differentiable on $\mathbf{R}$, and satisfies the formula

$$\frac{d}{dt}(e^{tA}) = Ae^{tA} = e^{tA}A \quad \text{for all } t \in \mathbf{R}. \tag{4.4}$$

Proof. (a) Since we have, for any $m \in \mathbf{N}$,

$$\|A^m\| \leq \|A\|^m,$$

it follows that

$$\begin{aligned}\|e^{tA}\| &= \|I + tA + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \dots\| \\ &\leq \|I\| + |t|\|A\| + \frac{|t|^2\|A\|^2}{2!} + \frac{|t|^3\|A\|^3}{3!} + \dots \\ &= e^{|t|\|A\|} \quad \text{for all } t \in \mathbf{R}.\end{aligned}$$

This proves that the series (4.3) converges in the space $\mathcal{L}(E, E)$ for all $t \in \mathbf{R}$, and enjoys property (a).

(b) Just as in the case of numerical series, we can rearrange the series

$$e^{(t+s)A} = \sum_{m=0}^{\infty} \frac{(t+s)^m}{m!} A^m$$

to obtain that

$$\left(\sum_{m=0}^{\infty} \frac{t^m}{m!} A^m\right)\left(\sum_{m=0}^{\infty} \frac{s^m}{m!} A^m\right).$$

(c) We remark that the series

$$Ae^{tA} = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^{m+1} = e^{tA} A$$

converges in $\mathcal{L}(E, E)$ uniformly in t over bounded intervals of $\mathbf{R}$. Hence we have, by termwise integration,

$$\int_0^t Ae^{sA}\, ds = \sum_{m=0}^{\infty} \frac{t^{m+1}}{(m+1)!} A^{m+1} = e^{tA} - I. \tag{4.5}$$

Therefore, we find that the left-hand side of formula (4.5) and hence the function e^{tA} is norm differentiable on $\mathbf{R}$, and that the desired formula (4.4) holds true.

The proof of Theorem 4.4 is complete.

Theorem 4.5. *If A and B are bounded linear operators on a Banach space E into itself and if A and B commute, then we have*

$$e^{A+B} = e^A\, e^B = e^B\, e^A. \tag{4.6}$$

Proof. Since $AB = BA$, it follows that

$$Ae^{tB} = \sum_{m=0}^{\infty} \frac{t^m}{m!} AB^m = \sum_{m=0}^{\infty} \frac{t^m}{m!} B^m A = e^{tB} A,$$

just as in the numerical case.

However, we have, by formula (4.4),

$$\frac{d}{dt}\left(e^{(1-t)(A+B)}\right) = -(A+B)\, e^{(1-t)(A+B)} \quad \text{for all } t \in \mathbf{R}.$$

Hence, if we let

$$K(t) = e^{tA}\, e^{tB}\, e^{(1-t)(A+B)} \quad \text{for every } t \in \mathbf{R},$$

we find from the Leibniz formula that

$$\begin{aligned}\frac{d}{dt}(K(t)) &= e^{tA} A e^{tB} e^{(1-t)(A+B)} + e^{tA} e^{tB} B e^{(1-t)(A+B)} \\ &\quad - e^{tA} e^{tB} (A+B) e^{(1-t)(A+B)} \\ &= e^{tA} e^{tB} A e^{(1-t)(A+B)} + e^{tA} e^{tB} B e^{(1-t)(A+B)} \\ &\quad - e^{tA} e^{tB} (A+B) e^{(1-t)(A+B)} \\ &= 0 \quad \text{for all } t \in \mathbf{R},\end{aligned}$$

so that $K(t)$ is a constant function. In particular, we have

$$e^{A+B} = K(0) = K(1) = e^A e^B.$$

This proves the desired formula (4.6).

The proof of Theorem 4.5 is complete.

4.4 Contraction Semigroups

Let E be a Banach space. A one-parameter family $\{T_t\}_{t \geq 0}$ of bounded linear operators on E into itself is called a *contraction semigroup of class* (C_0) or simply a *contraction semigroup* if it satisfies the following three conditions (i)–(iii):

(i) $T_{t+s} = T_t \cdot T_s$ for all $t, s \geq 0$.
(ii) $\lim_{t \downarrow 0} \|T_t x - x\| = 0$ for every $x \in E$.
(iii) $\|T_t\| \leq 1$ for all $t \geq 0$.

Condition (i) is called the *semigroup property*.

Remark 4.6. In view of conditions (i) and (ii), it follows that $T_0 = I$. Hence condition (ii) is equivalent to the strong continuity of $\{T_t\}_{t \geq 0}$ at $t = 0$. Moreover, it is easy to verify that a contraction semigroup $\{T_t\}_{t \geq 0}$ is strongly continuous on the interval $[0, \infty)$.

4.4.1 The Hille–Yosida Theory of Contraction Semigroups

If $\{T_t\}_{t \geq 0}$ is a contraction semigroup of class (C_0), then we let

$$\mathcal{D} = \text{the set of all } x \in E \text{ such that the limit} \quad \lim_{h \downarrow 0} \frac{T_h x - x}{h} \quad \text{exists in } E.$$

Then we define a linear operator $\mathfrak{A}$ from E into itself as follows:

(a) The domain $D(\mathfrak{A})$ of $\mathfrak{A}$ is the set $\mathcal{D}$.
(b) $\mathfrak{A}x = \lim_{h \downarrow 0} \frac{T_h x - x}{h}$ for every $x \in D(\mathfrak{A})$.

The operator $\mathfrak{A}$ is called the *infinitesimal generator* of $\{T_t\}_{t \geq 0}$.

First, we derive a differential equation associated with a contraction semigroup of class (C_0) in terms of its infinitesimal generator:

Proposition 4.7. *Let $\mathfrak{A}$ be the infinitesimal generator of a contraction semigroup $\{T_t\}_{t \geq 0}$. If $x \in D(\mathfrak{A})$, then we have $T_t x \in D(\mathfrak{A})$ for all $t > 0$, and the function $T_t x$ is strongly differentiable on the interval $(0, \infty)$ and satisfies the equation*

$$\frac{d}{dt}(T_t x) = \mathfrak{A}(T_t x) = T_t(\mathfrak{A}x) \quad \textit{for all } t > 0. \tag{4.7}$$

Proof. Let $h > 0$. Then we have, by the semigroup property,

$$\frac{T_h(T_t x) - T_t x}{h} = T_t \left(\frac{T_h - I}{h} x \right).$$

However, since T_t is bounded and $x \in D(\mathfrak{A})$, it follows that

$$T_t \left(\frac{T_h - I}{h} x \right) \longrightarrow T_t(\mathfrak{A}x) \quad \text{as } h \downarrow 0.$$

This implies that

$$\begin{cases} T_t x \in D(\mathfrak{A}), \\ \mathfrak{A}(T_t x) = T_t(\mathfrak{A}x). \end{cases}$$

Therefore, we find that $T_t x$ is strongly right-differentiable on $(0, \infty)$ and satisfies the equation

$$\frac{d}{dt}^{+} (T_t x) = \mathfrak{A}(T_t x) = T_t(\mathfrak{A}x) \quad \text{for all } t > 0.$$

Similarly, we have, for each $0 < h < t$,

$$\frac{T_{t-h}x - T_t x}{-h} - T_t(\mathfrak{A}x) = T_{t-h} \left(\frac{T_h - I}{h} x - T_h(\mathfrak{A}x) \right).$$

However, since we have the inequality

$$\|T_{t-h}\| \leq 1$$

and, as $h \downarrow 0$,

$$T_h(\mathfrak{A}x) \longrightarrow \mathfrak{A}x,$$

we obtain that

$$\frac{T_{t-h}x - T_t x}{-h} \longrightarrow T_t(\mathfrak{A}x) \quad \text{as } h \downarrow 0.$$

This proves that $T_t x$ is strongly left-differentiable on $(0, \infty)$ and satisfies the equation

$$\frac{d^-}{dt}(T_t x) = \mathfrak{A}(T_t x) = T_t(\mathfrak{A}x) \quad \text{for all } t > 0.$$

Summing up, we have proved that $T_t x$ is strongly differentiable on the interval $(0, \infty)$ and satisfies Eq. (4.7).

The proof of Proposition 4.7 is complete.

The next proposition characterizes the infinitesimal generator $\mathfrak{A}$:

Proposition 4.8. *Let $\mathfrak{A}$ be the infinitesimal generator of a contraction semigroup $\{T_t\}_{t \geq 0}$. Then $\mathfrak{A}$ is a densely defined, closed linear operator in E.*

Proof. The proof is divided into two steps.

Step 1: First, we show that the operator $\mathfrak{A}$ is *closed*.

To do this, we assume that

$$x_n \in D(\mathfrak{A}), \ x_n \to x_0 \text{ and } \mathfrak{A}x_n \to y_0 \text{ in } E.$$

Then it follows from an application of Eq. (4.7) that

$$T_t x_n - x_n = \int_0^t \frac{d}{ds}(T_s x_n)\, ds = \int_0^t T_s(\mathfrak{A}x_n)\, ds. \tag{4.8}$$

However, we have, as $n \to \infty$,

$$T_t x_n - x_n \longrightarrow T_t x_0 - x_0$$

and also

$$\begin{aligned}\left\| \int_0^t T_s(\mathfrak{A}x_n)\, ds - \int_0^t T_s y_0\, ds \right\| &= \left\| \int_0^t T_s(\mathfrak{A}x_n - y_0)\, ds \right\| \\ &\leq \int_0^t \|T_s(\mathfrak{A}x_n - y_0)\|\, ds \\ &\leq t\, \|\mathfrak{A}x_n - y_0\| \longrightarrow 0.\end{aligned}$$

Hence, by letting $n \to \infty$ in (4.8) we obtain that

$$T_t x_0 - x_0 = \int_0^t T_s y_0 \, ds. \tag{4.9}$$

Furthermore, it follows that, as $t \downarrow 0$,

$$\frac{1}{t} \int_0^t T_s y_0 \, ds \longrightarrow T_0 y_0 = y_0,$$

since the integrand $T_s y_0$ is strongly continuous.
Therefore, we find from (4.9) that, as $t \downarrow 0$,

$$\frac{T_t x_0 - x_0}{t} = \frac{1}{t} \int_0^t T_s y_0 \, ds \longrightarrow y_0.$$

This proves that

$$\begin{cases} x_0 \in D(\mathfrak{A}), \\ \mathfrak{A} x_0 = y_0. \end{cases}$$

Therefore, we have proved that the operator $\mathfrak{A}$ is a closed operator.
Step 2: Secondly, we show the *density* of the domain $D(\mathfrak{A})$ in E.
Let x be an arbitrary element of E. For each $\delta > 0$, we let

$$x_\delta = \frac{1}{\delta} \int_0^\delta T_s x \, ds.$$

Then we have, for any $0 < h < \delta$,

$$T_h(x_\delta) = \frac{1}{\delta} \int_0^\delta T_h(T_s x) \, ds = \frac{1}{\delta} \int_0^\delta T_{h+s} x \, ds = \frac{1}{\delta} \int_h^{\delta+h} T_s x \, ds.$$

Hence it follows that

$$\begin{aligned} \left(\frac{T_h - I}{h} \right) x_\delta &= \frac{1}{\delta} \left(\frac{1}{h} \int_h^{\delta+h} T_s x \, ds - \frac{1}{h} \int_0^\delta T_s x \, ds \right) \\ &= \frac{1}{\delta} \left(\frac{1}{h} \int_\delta^{\delta+h} T_s x \, ds - \frac{1}{h} \int_0^h T_s x \, ds \right). \end{aligned} \tag{4.10}$$

However, it follows that, as $h \downarrow 0$,

$$\frac{1}{h}\int_{\delta}^{\delta+h} T_s x\, ds \longrightarrow T_\delta x,$$
$$\frac{1}{h}\int_{0}^{h} T_s x\, ds \longrightarrow T_0 x = x,$$

since the integrand $T_s x$ is strongly continuous.
Therefore, we find from (4.10) that, as $h \downarrow 0$,

$$\left(\frac{T_h - I}{h}\right) x_\delta \longrightarrow \frac{1}{\delta}(T_\delta x - x).$$

This proves that

$$\begin{cases} x_\delta \in D(\mathfrak{A}), \\ \mathfrak{A} x_\delta = \dfrac{1}{\delta}(T_\delta x - x). \end{cases}$$

Moreover, it follows that, as $\delta \downarrow 0$,

$$x_\delta = \frac{1}{\delta}\int_0^\delta T_s x\, ds \longrightarrow T_0 x = x.$$

Summing up, we have proved that $D(\mathfrak{A})$ is dense in E.
The proof of Proposition 4.8 is complete.

Let $\{T_t\}_{t\geq 0}$ be a contraction semigroup. Then the integral

$$\int_0^s e^{-\alpha t} T_t x\, dt, \quad x \in E, \tag{4.11}$$

is strongly integrable for all $s > 0$, since the integrand is strongly continuous on the interval $[0, \infty)$. Moreover, if $\alpha > 0$, then the limit $G_\alpha x$ of the integral (4.11) exists in E as $s \to \infty$:

$$G_\alpha x := \int_0^\infty e^{-\alpha t} T_t x\, dt = \lim_{s\to\infty} \int_0^s e^{-\alpha t} T_t x\, dt, \quad x \in E,\ \alpha > 0.$$

Thus $G_\alpha x$ is defined for all $x \in E$ if $\alpha > 0$. It is easy to see that the operator G_α is a bounded linear operator from E into itself with norm $1/\alpha$:

$$\|G_\alpha x\| \leq \frac{1}{\alpha}\|x\| \quad \text{for all } x \in E. \tag{4.12}$$

The family $\{G_\alpha\}_{\alpha>0}$ of bounded linear operators is called the *resolvent* of the semigroup $\{T_t\}_{t\geq 0}$.

The next proposition characterizes the resolvent G_α:

Proposition 4.9. *Let $\{T_t\}_{t\geq 0}$ be a contraction semigroup defined on a Banach space E and $\mathfrak{A}$ the infinitesimal generator of $\{T_t\}$. For each $\alpha > 0$, the operator $(\alpha I - \mathfrak{A})$ is a* bijection *of $D(\mathfrak{A})$ onto E, and its inverse $(\alpha I - \mathfrak{A})^{-1}$ is the resolvent G_α:*

$$(\alpha I - \mathfrak{A})^{-1} x = G_\alpha x = \int_0^\infty e^{-\alpha t} T_t x \, dt \quad \textit{for every } x \in E. \tag{4.13}$$

Proof. The proof is divided into three steps.

Step 1: First, we show that $(\alpha I - \mathfrak{A})$ is *surjective* for each $\alpha > 0$.

Let x be an arbitrary element of E. Then we have, for each $h > 0$,

$$\begin{aligned} T_h(G_\alpha x) &= \int_0^\infty e^{-\alpha t} T_h(T_t x) \, dt = \int_0^\infty e^{-\alpha t} T_{t+h} x \, dt \\ &= e^{\alpha h} \int_h^\infty e^{-\alpha t} T_t x \, dt. \end{aligned}$$

Hence it follows that

$$\begin{aligned} T_h(G_\alpha x) - G_\alpha x &= e^{\alpha h} \int_h^\infty e^{-\alpha t} T_t x \, dt - \int_0^\infty e^{-\alpha t} T_t x \, dt \\ &= \left(e^{\alpha h} - 1\right) \int_h^\infty e^{-\alpha t} T_t x \, dt - \int_0^h e^{-\alpha t} T_t x \, dt, \end{aligned}$$

so that

$$\begin{aligned} &\frac{T_h(G_\alpha x) - G_\alpha x}{h} \\ &= \left(\frac{e^{\alpha h} - 1}{h}\right) \int_h^\infty e^{-\alpha t} T_t x \, dt - \frac{1}{h} \int_0^h e^{-\alpha t} T_t x \, dt. \end{aligned} \tag{4.14}$$

However, we obtain that, as $h \downarrow 0$,

$$\begin{aligned} &\left(\frac{e^{\alpha h} - 1}{h}\right) \longrightarrow \alpha, \\ &\int_h^\infty e^{-\alpha t} T_t x \, dt \longrightarrow \int_0^\infty e^{-\alpha t} T_t x \, dt = G_\alpha x, \\ &\frac{1}{h} \int_0^h e^{-\alpha t} T_t x \, dt \longrightarrow e^{-\alpha t} T_t x|_{t=0} = x. \end{aligned}$$

By letting $h \downarrow 0$ in (4.14), we have proved that

$$\frac{T_h(G_\alpha x) - G_\alpha x}{h} \longrightarrow \alpha G_\alpha x - x \quad \text{as } h \downarrow 0.$$

This implies that

$$\begin{cases} G_\alpha x \in D(\mathfrak{A}), \\ \mathfrak{A}(G_\alpha x) = \alpha G_\alpha x - x, \end{cases}$$

or equivalently,

$$(\alpha I - \mathfrak{A}) G_\alpha x = x \quad \text{for every } x \in E.$$

Therefore, we have proved that the operator $(\alpha I - \mathfrak{A})$ is surjective for each $\alpha > 0$.

Step 2: Secondly, we show that $(\alpha I - \mathfrak{A})$ is *injective* for each $\alpha > 0$.

Now we assume that

$$x \in D(\mathfrak{A}), \ (\alpha I - \mathfrak{A})x = 0.$$

If we introduce a function $u(t)$ by the formula

$$u(t) = e^{-\alpha t} T_t x \quad \text{for all } t > 0,$$

then it follows from an application of Proposition 4.7 that

$$\begin{aligned} \frac{d}{dt}(u(t)) = -\alpha e^{-\alpha t} T_t x + e^{-\alpha t} T_t \mathfrak{A} x &= -e^{-\alpha t} T_t (\alpha I - \mathfrak{A}) x \\ &= 0, \end{aligned}$$

so that

$$u(t) = \text{a constant}, \quad t > 0.$$

However, we have, by letting $t \downarrow 0$,

$$u(t) = u(0) = e^{-\alpha t} T_t x\big|_{t=0} = x.$$

On the other hand, we have, by letting $t \uparrow +\infty$,

$$u(t) = u(+\infty) = \lim_{t \uparrow +\infty} u(t) = 0.$$

Indeed, it suffices to note that

$$\|u(t)\| = e^{-\alpha t} \|T_t x\| \leq e^{-\alpha t} \|x\| \quad \text{for all } t > 0.$$

Hence it follows that $x = 0$. This proves that the operator $(\alpha I - \mathfrak{A})$ is injective for each $\alpha > 0$.

Step 3: Summing up, we have proved that $(\alpha I - \mathfrak{A})$ is a *bijection* of $D(\mathfrak{A})$ onto E and that $(\alpha I - \mathfrak{A})^{-1} = G_\alpha$.

The proof of Proposition 4.9 is complete.

Now we consider when a linear operator is the infinitesimal generator of some contraction semigroup. This question is answered by the following Hille–Yosida theorem [CP, Chapitre 6, Théorème 6.12]:

Theorem 4.10 (Hille–Yosida). *Let $\mathfrak{A}$ be a linear operator from a Banach space E into itself with domain $D(\mathfrak{A})$. In order that $\mathfrak{A}$ is the infinitesimal generator of some contraction semigroup, it is necessary and sufficient that $\mathfrak{A}$ satisfies the following three conditions:*

(i) The operator $\mathfrak{A}$ is closed and its domain $D(\mathfrak{A})$ is dense in E.
(ii) For every $\alpha > 0$ the equation

$$(\alpha I - \mathfrak{A})x = y$$

has a unique solution $x \in D(\mathfrak{A})$ for any $y \in E$; we then write

$$x = (\alpha I - \mathfrak{A})^{-1} y.$$

(iii) For any $\alpha > 0$, we have the inequality

$$\left\| (\alpha I - \mathfrak{A})^{-1} \right\| \leq \frac{1}{\alpha}. \tag{4.15}$$

Proof. The necessity of conditions (i)–(iii) follows from Propositions 4.8 and 4.9 and inequality (4.12).

The sufficiency is proved in six steps.

Step 1: For each $\alpha > 0$, we define linear operators

$$J_\alpha = \alpha(\alpha I - \mathfrak{A})^{-1},$$

and

$$\mathfrak{A}_\alpha = \mathfrak{A} J_\alpha.$$

Then we can prove the following two assertions (4.16) and (4.17):

$$\| J_\alpha \| \leq 1, \tag{4.16a}$$

$$\lim_{\alpha \to +\infty} J_\alpha x = x \quad \text{for every } x \in E, \tag{4.16b}$$

and

$$\|\mathfrak{A}_\alpha\| \leq 2\alpha, \tag{4.17a}$$

$$\lim_{\alpha \to +\infty} \mathfrak{A}_\alpha x = \mathfrak{A}x \quad \text{for every } x \in D(\mathfrak{A}). \tag{4.17b}$$

The operators $\mathfrak{A}_\alpha$ are called the *Yosida approximations* to $\mathfrak{A}$.
First, we note that assertion (4.16a) is an immediate consequence of inequality (4.15). Furthermore, we have, for all $x \in D(\mathfrak{A})$,

$$\begin{aligned} J_\alpha x - x &= \alpha(\alpha I - \mathfrak{A})^{-1}x - (\alpha I - \mathfrak{A})^{-1}(\alpha I - \mathfrak{A})x \\ &= (\alpha I - \mathfrak{A})^{-1}(\alpha x - \alpha x + \mathfrak{A}x) = (\alpha I - \mathfrak{A})^{-1}(\mathfrak{A}x). \end{aligned}$$

Hence it follows from inequality (4.15) that, as $\alpha \to +\infty$,

$$\|J_\alpha x - x\| \leq \left\|(\alpha I - \mathfrak{A})^{-1}\right\| \|\mathfrak{A}x\| \leq \frac{1}{\alpha} \|\mathfrak{A}x\| \longrightarrow 0.$$

This proves assertion (4.16b), since $\|J_\alpha\| \leq 1$ and $D(\mathfrak{A})$ is dense in E.
Assertion (4.17b) follows from assertion (4.16b). Indeed, we have, as $\alpha \to +\infty$,

$$\mathfrak{A}_\alpha x = \mathfrak{A}J_\alpha x = J_\alpha(\mathfrak{A}x) \longrightarrow \mathfrak{A}x \quad \text{for every } x \in D(\mathfrak{A}).$$

On the other hand, it follows that

$$\mathfrak{A}_\alpha = -\alpha I + \alpha J_\alpha,$$

so that

$$\|\mathfrak{A}_\alpha\| \leq \alpha + \alpha \|J_\alpha\| \leq 2\alpha.$$

This proves assertion (4.17a).

Step 2: We define a linear operator

$$T_t(\alpha) = \exp[t\mathfrak{A}_\alpha] \quad \text{for every } \alpha > 0.$$

Since we have

$$\mathfrak{A}_\alpha = -\alpha I + \alpha J_\alpha,$$

it follows from an application of Theorem 4.5 that the operators

$$T_t(\alpha) = e^{-\alpha t} \exp[\alpha t J_\alpha], \quad t \geq 0, \tag{4.18}$$

form a contraction semigroup for each $\alpha > 0$. Indeed, it suffices to note that

$$\begin{aligned}\|T_t(\alpha)\| &= e^{-\alpha t} \|\exp[\alpha t J_\alpha]\| \leq e^{-\alpha t} \left(\sum_{n=0}^{\infty} \frac{(\alpha t)^n}{n!} \|J_\alpha^n\|\right) \\ &\leq e^{-\alpha t} \left(\sum_{n=0}^{\infty} \frac{(\alpha t)^n}{n!}\right) = e^{-\alpha t} e^{\alpha t} = 1.\end{aligned}$$

Step 3: We show that the operator $T_t(\alpha)$ has a strong limit T_t as $\alpha \to +\infty$:

$$T_t x = \lim_{\alpha \to +\infty} T_t(\alpha) x \quad \text{for every } x \in E.$$

Moreover, this convergence is uniform in t over bounded intervals $[0, t_0]$ for all $t_0 > 0$.
If x is an arbitrary element of $D(\mathfrak{A})$, then it follows from an application of Proposition 4.7 that

$$\begin{aligned}&T_t(\alpha)x - T_t(\beta)x \\ &= \int_0^t \frac{d}{ds} \left(T_{t-s}(\beta) T_s(\alpha) x\right) ds \\ &= \int_0^t \left(\frac{d}{ds}(T_{t-s}(\beta)) \cdot T_s(\alpha)x + T_{t-s}(\beta) \cdot \frac{d}{ds}(T_s(\alpha)x)\right) ds \\ &= \int_0^t \left(T_{t-s}(\beta)(-\mathfrak{A}_\beta) T_s(\alpha)x + T_{t-s}(\beta) T_s(\alpha)(\mathfrak{A}_\alpha x)\right) ds \\ &= \int_0^t T_{t-s}(\beta) T_s(\alpha) \left(\mathfrak{A}_\alpha x - \mathfrak{A}_\beta x\right) ds.\end{aligned}$$

Hence we have the inequality

$$\begin{aligned}&\|T_t(\alpha)x - T_t(\beta)x\| \\ &= \left\|\int_0^t T_{t-s}(\beta) T_s(\alpha) \left(\mathfrak{A}_\alpha x - \mathfrak{A}_\beta x\right) ds\right\| \\ &\leq \int_0^t \|T_{t-s}(\beta)\| \, |T_s(\alpha)\| \, ds \cdot \left\|\mathfrak{A}_\alpha x - \mathfrak{A}_\beta x\right\| \leq t \left\|\mathfrak{A}_\alpha x - \mathfrak{A}_\beta x\right\| \\ &\leq t_0 \left\|\mathfrak{A}_\alpha x - \mathfrak{A}_\beta x\right\| \quad \text{for all } t \in [0, t_0],\end{aligned}$$

since $\|T_{t-s}(\beta)\| \leq 1$ and $\|T_s(\alpha)\| \leq 1$. However, we recall (see assertion (4.17b) that, as $\alpha \to +\infty$,

$$\mathfrak{A}_\alpha x \longrightarrow \mathfrak{A} x \quad \text{for every } x \in D(\mathfrak{A}).$$

Therefore, we obtain that, as $\alpha, \beta \to +\infty$,

$$\|T_t(\alpha)x - T_t(\beta)x\| \longrightarrow 0,$$

and that this convergence is uniform in t over the interval $[0, t_0]$.
We can define a linear operator T_t by the formula

$$T_t x = \lim_{\alpha \to +\infty} T_t(\alpha)x \quad \text{for every } x \in D(\mathfrak{A}).$$

Furthermore, since $\|T_t(\alpha)\| \leq 1$ and $D(\mathfrak{A})$ is dense in E, it follows that the operator $T_t(\alpha)$ has a strong limit T_t as $\alpha \to +\infty$:

$$T_t x = \lim_{\alpha \to +\infty} T_t(\alpha)x \quad \text{for every } x \in E, \tag{4.19}$$

and further that the convergence is uniform in t over bounded intervals $[0, t_0]$ for each $t_0 > 0$.

Step 4: We show that the family $\{T_t\}_{t \geq 0}$ forms a contraction semigroup of class (C_0).
First, it follows from an application of the principle of uniform boundedness (Theorem 3.14) that the operator T_t is bounded and satisfies the condition

$$\|T_t\| \leq \liminf_{\alpha \to +\infty} \|T_t(\alpha)\| \leq 1 \quad \text{for all } t \geq 0.$$

Secondly, the semigroup property of $\{T_t\}$

$$T_t(T_s x) = T_{t+s}x, \quad x \in E$$

follows from that of $\{T_t(\alpha)\}$. Indeed, we have, as $\alpha \to +\infty$,

$$\begin{aligned}
&\|T_t(T_s x) - T_t(\alpha)(T_s(\alpha)x)\| \\
&\leq \|(T_t - T_t(\alpha))T_s x\| + \|T_t(\alpha)(T_s x - T_s(\alpha)x)\| \\
&\leq \|(T_t - T_t(\alpha))T_s x\| + \|(T_s - T_s(\alpha))x\| \longrightarrow 0,
\end{aligned}$$

so that

$$\begin{aligned}
T_t(T_s x) &= \lim_{\alpha \to +\infty} T_t(\alpha)(T_s(\alpha)x) = \lim_{\alpha \to +\infty} T_{t+s}(\alpha)x \\
&= T_{t+s}x \quad \text{for every } x \in E.
\end{aligned}$$

Furthermore, since the convergence of (4.19) is uniform in t over bounded subintervals of the interval $[0, \infty)$, it follows that the function $T_t x$, $x \in E$, is strongly continuous on the interval $[0, \infty)$. Consequently, the family $\{T_t\}_{t \geq 0}$ forms a contraction semigroup.

Step 5: We show that the infinitesimal generator of the semigroup $\{T_t\}_{t\geq 0}$ thus obtained is precisely the operator $\mathfrak{A}$.
Let $\mathfrak{A}_0$ be the infinitesimal generator of $\{T_t\}_{t\geq 0}$ with domain $D(\mathfrak{A}_0)$. If x is an arbitrary element of the domain $D(\mathfrak{A})$, it follows from an application of Proposition 4.7 that

$$e^{t\mathfrak{A}_\alpha}x - x = \int_0^t \frac{d}{ds}\left(e^{s\mathfrak{A}_\alpha}x\right)\,ds = \int_0^t e^{s\mathfrak{A}_\alpha}(\mathfrak{A}_\alpha x)\,ds. \tag{4.20}$$

However, we have, as $\alpha \to +\infty$,

$$e^{t\mathfrak{A}_\alpha}x - x = T_t(\alpha)x - x \longrightarrow T_t x - x \quad \text{for every } x \in D(\mathfrak{A}),$$

and also

$$\int_0^t e^{s\mathfrak{A}_\alpha}(\mathfrak{A}_\alpha x)\,ds \longrightarrow \int_0^t T_s(\mathfrak{A}x)\,ds \quad \text{for every } x \in D(\mathfrak{A}).$$

Indeed, it suffices to note that, as $\alpha \to +\infty$,

$$\begin{aligned}
&\left\| \int_0^t e^{s\mathfrak{A}_\alpha}(\mathfrak{A}_\alpha x)\,ds - \int_0^t T_s(\mathfrak{A}x)\,ds \right\| \\
&\leq \left\| \int_0^t e^{s\mathfrak{A}_\alpha}(\mathfrak{A}_\alpha x - \mathfrak{A}x)\,ds \right\| + \left\| \int_0^t (e^{s\mathfrak{A}_\alpha} - T_s)(\mathfrak{A}x)\,ds \right\| \\
&\leq \int_0^t \left\| e^{s\mathfrak{A}_\alpha} \right\|\,ds\,\|\mathfrak{A}_\alpha x - \mathfrak{A}x\| + \int_0^t \left\| e^{s\mathfrak{A}_\alpha}(\mathfrak{A}x) - T_s(\mathfrak{A}x) \right\|\,ds \\
&\leq t\,\|\mathfrak{A}_\alpha x - \mathfrak{A}x\| + \int_0^t \|T_s(\alpha)(\mathfrak{A}x) - T_s(\mathfrak{A}x)\|\,ds \longrightarrow 0.
\end{aligned}$$

Hence, by letting $\alpha \to +\infty$ in (4.20) we have, for all $x \in D(\mathfrak{A})$,

$$T_t x - x = \int_0^t T_s(\mathfrak{A}x)\,ds.$$

Moreover, it follows that, as $t \downarrow 0$,

$$\frac{T_t x - x}{t} = \frac{1}{t}\int_0^t T_s(\mathfrak{A}x)\,ds \longrightarrow T_0(\mathfrak{A}x) = \mathfrak{A}x \quad \text{for every } x \in D(\mathfrak{A}),$$

since the integrand $T_s(\mathfrak{A}x)$ is strongly continuous.
Summing up, we have proved that

$$\begin{cases} x \in D(\mathfrak{A}_0), \\ \mathfrak{A}_0 x = \mathfrak{A}x. \end{cases}$$

This implies that

$$\mathfrak{A} \subset \mathfrak{A}_0.$$

It remains to show that

$$D(\mathfrak{A}) = D(\mathfrak{A}_0).$$

If y is an arbitrary element of $D(\mathfrak{A}_0)$, we let

$$x = (I - \mathfrak{A})^{-1}(I - \mathfrak{A}_0)y.$$

Then we have

$$\begin{cases} x \in D(\mathfrak{A}) \subset D(\mathfrak{A}_0), \\ (I - \mathfrak{A})x = (I - \mathfrak{A}_0)y, \end{cases}$$

and so

$$(I - \mathfrak{A}_0)x = (I - \mathfrak{A}_0)y,$$

This implies that

$$y = x \in D(\mathfrak{A}),$$

since the operator $(I - \mathfrak{A}_0)$ is bijective.

Step 6: Finally, we show the uniqueness of the semigroup.

Let $\{U_t\}_{t\geq 0}$ be another contraction semigroup which has $\mathfrak{A}$ as its infinitesimal generator. For each $x \in D(\mathfrak{A})$ and each $t > 0$, we introduce a function $w(s)$ as follows:

$$w(s) = T_{t-s}(U_s x), \quad 0 \leq s \leq t.$$

Then it follows from an application of Proposition 4.7 that

$$\begin{aligned} \frac{dw}{ds} &= \left(\frac{d}{ds}T_{t-s}\right)U_s x + T_{t-s}\left(\frac{d}{ds}U_s x\right) \\ &= -\mathfrak{A}T_{t-s}(U_s x) + T_{t-s}(\mathfrak{A}U_s x) = -T_{t-s}(\mathfrak{A}U_s x) + T_{t-s}(\mathfrak{A}U_s x) \\ &= 0, \quad 0 < s < t, \end{aligned}$$

so that

$$w(s) = \text{a constant}, \quad 0 \leq s \leq t.$$

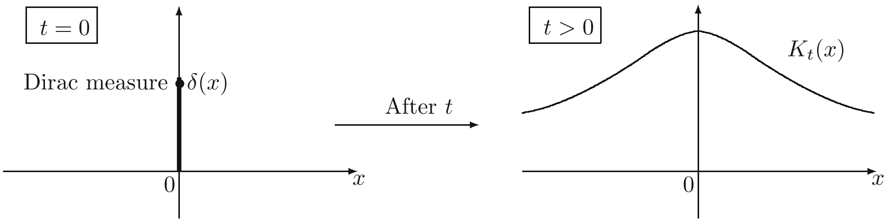


Fig. 4.1 The intuitive meaning of the heat kernel $K_t(x)$

In particular, we obtain that $w(0) = w(t)$, that is,

$$T_t x = U_t x \quad \text{for all } x \in D(\mathfrak{A}).$$

This implies that $T_t = U_t$ for all $t \geq 0$, since T_t and U_t are both bounded and since $D(\mathfrak{A})$ is dense in E.

Now the proof of Theorem 4.10 is complete.

4.4.2 The Contraction Semigroup Associated with the Heat Kernel

In this subsection we study the semigroup associated with the heat kernel

$$K_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}, \quad t > 0, \quad x \in \mathbf{R}^n.$$

Physically, the heat kernel $K_t(x)$ expresses a thermal distribution of position x at time t in a homogeneous isotropic medium $\mathbf{R}^n$ with unit coefficient of thermal diffusivity, given that the initial thermal distribution is the Dirac measure $\delta(x)$ (see Fig. 4.1):

Now we consider the heat kernel $K_t(x)$ on the function space

$$C_0(\mathbf{R}^n) = \{u \in C(\mathbf{R}^n) : \lim_{x \to \infty} u(x) = 0\}.$$

We recall (see Sect. 4.4.1) that a function $u \in C(\mathbf{R}^n)$ is said to *vanish at infinity* if the set

$$\{x \in \mathbf{R}^n : |u(x)| \geq \varepsilon\}$$

is compact for every $\varepsilon > 0$, and we write

$$\lim_{x \to \infty} u(x) = 0.$$

It is easy to see that the function space $C_0(\mathbf{R}^n)$ is a Banach space with the supremum (maximum) norm

$$\|u\|_\infty = \sup_{x \in \mathbf{R}^n} |u(x)|.$$

The purpose of this subsection is to prove the following example [Tn, Chapter 12, Example 3]:

Example 4.11. A one-parameter family $\{T_t\}_{t \geq 0}$ of bounded linear operators, defined by the formula

$$T_t u(x) = \begin{cases} u(x) & \text{for } t = 0, \\ \int_{\mathbf{R}^n} K_t(x-y)u(y)\,dy & \text{for } t > 0, \end{cases}$$

forms a *contraction semigroup* of class (C_0) on the Banach space $C_0(\mathbf{R}^n)$.

Proof. The proof of Example 4.11 is given by a series of several claims. In the following we shall write $T(t)$ for T_t.

Step 1: First, the next lemma proves that the operators $T(t)$ map $C_0(\mathbf{R}^n)$ into itself:

Lemma 4.12. *We have, for all $t > 0$,*

$$T(t) : C_0(\mathbf{R}^n) \longrightarrow C_0(\mathbf{R}^n).$$

Proof. Let $u(x)$ be an arbitrary function in $C_0(\mathbf{R}^n)$. Then it follows that

$$\begin{aligned} T(t)u(x) &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u(y)\,dy \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|z|^2}{4t}} u(x-z)\,dz. \end{aligned}$$

(1) First, we show that $T(t)u \in C(\mathbf{R}^n)$: Since $u \in C_0(\mathbf{R}^n)$ is uniformly continuous on $\mathbf{R}^n$, for any given number $\varepsilon > 0$ we can find a constant $\delta = \delta(\varepsilon) > 0$ such that

$$|x_1 - x_2| < \delta \implies |u(x_1) - u(x_2)| < \varepsilon.$$

In particular, we have

$$|x - y| < \delta \implies |u(x-z) - u(y-z)| < \varepsilon.$$

Therefore, we obtain that

$$\begin{aligned}
&|T(t)u(x) - T(t)u(y)| \\
&= \left| \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|z|^2}{4t}} u(x-z)\, dz - \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|z|^2}{4t}} u(y-z)\, dy \right| \\
&\leq \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|z|^2}{4t}} |u(x-z) - u(y-z)|\, dz \\
&\leq \frac{\varepsilon}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|z|^2}{4t}}\, dz = \varepsilon.
\end{aligned}$$

This proves that the function $T(t)u$ is uniformly continuous on $\mathbf{R}^n$.

(2) Secondly, we show that $T(t)u \in C_0(\mathbf{R}^n)$, that is,

$$\lim_{x\to\infty} T(t)u(x) = 0. \tag{4.21}$$

Since we have the condition

$$\lim_{x\to\infty} u(x) = 0,$$

for any given number $\varepsilon > 0$ we can find a positive integer $N = N(\varepsilon) \in \mathbf{N}$ such that

$$|u(y)| < \varepsilon \quad \text{for all } |y| > N. \tag{4.22}$$

Then we decompose the integral $T(t)u(x)$ into the two terms:

$$\begin{aligned}
&T(t)u(x) \\
&= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u(y)\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \int_{|y|\leq N} e^{-\frac{|x-y|^2}{4t}} u(y)\, dy + \frac{1}{(4\pi t)^{n/2}} \int_{|y|>N} e^{-\frac{|x-y|^2}{4t}} u(y)\, dy \\
&:= I_1(x) + I_2(x).
\end{aligned}$$

However, by condition (4.22) we can estimate the term $I_2(x)$ as follows:

$$\begin{aligned}
|I_2(x)| &\leq \frac{1}{(4\pi t)^{n/2}} \int_{|y|>N} e^{-\frac{|x-y|^2}{4t}} |u(y)|\, dy \qquad (4.23) \\
&< \frac{\varepsilon}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u(y)\, dy = \varepsilon.
\end{aligned}$$

The term $I_1(x)$ may be estimated as follows:

$$\begin{aligned}|I_1(x)| &\leq \frac{1}{(4\pi t)^{n/2}} \int_{|y|\leq N} e^{-\frac{|x-y|^2}{4t}} |u(y)|\, dy \\ &\leq \frac{1}{(4\pi t)^{n/2}} \int_{|y|\leq N} e^{-\frac{(|x|-N)^2}{4t}} |u(y)|\, dy \\ &= \frac{1}{(4\pi t)^{n/2}} e^{-\frac{(|x|-N)^2}{4t}} \int_{|y|\leq N} |u(y)|\, dy \\ &\leq \frac{1}{(4\pi t)^{n/2}} e^{-\frac{(|x|-N)^2}{4t}} \int_{\mathbf{R}^n} |u(y)|\, dy.\end{aligned}$$

Moreover, we have, as $x \to \infty$,

$$e^{-\frac{(|x|-N)^2}{4t}} \longrightarrow 0,$$

and hence

$$I_1(x) \longrightarrow 0 \quad \text{as } x \to \infty. \tag{4.24}$$

Summing up, we obtain from assertions (4.23) and (4.24) that

$$\limsup_{x\to\infty} |T(t)u(x)| \leq \limsup_{x\to\infty} (|I_1(x)| + |I_2(x)|) \leq \varepsilon.$$

This proves the desired assertion (4.21), since ε is arbitrary.
The proof of Lemma 4.12 is complete.

Moreover, we find that the operators $\{T(t)\}_{t>0}$ are bounded on the space $C_0(\mathbf{R}^n)$. Indeed, it suffices to note that we have, for all $x \in \mathbf{R}^n$,

$$\begin{aligned}&|T(t)u(x)| \\ &\leq \int_{\mathbf{R}^n} K_t(x-y)|u(y)|\, dy \\ &\leq \|u\|_\infty \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}}\, dy = \|u\|_\infty \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|y|^2}{4t}}\, dy \\ &= \|u\|_\infty,\end{aligned}$$

and hence

$$\|T(t)u\|_\infty \leq \|u\|_\infty \quad \text{for all } u \in C_0(\mathbf{R}^n).$$

This proves that $\|T(t)\| \leq 1$ for all $t > 0$.

Step 2: Secondly, we show that the family $\{T(t)\}_{t \geq 0}$ forms a semigroup on the space $C_0(\mathbf{R}^n)$:

Step 2-1: To do this, we need the following Chapman–Kolmogorov equation (4.25) for the heat kernel (cf. formula (9.4)):

Lemma 4.13 (the Chapman–Kolmogorov equation). *For all $t, s > 0$, we have the equation*

$$K_{t+s}(x) = \int_{\mathbf{R}^n} K_t(x-y)\, K_s(y)\, dy. \tag{4.25}$$

Proof. (1) The proof is based on the following elementary formula (4.26):

$$\frac{|x-y|^2}{4t} + \frac{|y|^2}{4s} = \frac{|x|^2}{4(t+s)} + \frac{t+s}{4ts}\left|y - \frac{s}{t+s}x\right|^2. \tag{4.26}$$

Indeed, the right-hand side is calculated as follows:

$$\begin{aligned}
&\frac{|x|^2}{4(t+s)} + \frac{t+s}{4ts}\left|y - \frac{s}{t+s}x\right|^2 \qquad (4.27)\\
&= \frac{|x|^2}{4(t+s)} + \frac{t+s}{4ts}\left(y - \frac{s}{t+s}x, y - \frac{s}{t+s}x\right)\\
&= \frac{|x|^2}{4(t+s)} + \frac{t+s}{4ts}\left(|y|^2 - \frac{2s}{t+s}(x,y) + \left(\frac{s}{t+s}\right)^2 |x|^2\right)\\
&= \frac{|x|^2}{4(t+s)} + \frac{t+s}{4ts}|y|^2 - \frac{1}{2t}(x,y) + \frac{s}{4t(t+s)}|x|^2\\
&= \frac{t+s}{4t(t+s)}|x|^2 + \frac{t+s}{4ts}|y|^2 - \frac{1}{2t}(x,y)\\
&= \frac{1}{4t}|x|^2 - \frac{1}{2t}(x,y) + \frac{t+s}{4ts}|y|^2.
\end{aligned}$$

Similarly, the left-hand side is calculated as follows:

$$\begin{aligned}
\frac{|x-y|^2}{4t} + \frac{|y|^2}{4s} &= \frac{1}{4t}(x-y, x-y) + \frac{1}{4s}|y|^2 \qquad (4.28)\\
&= \frac{1}{4t}|x|^2 - \frac{1}{2t}(x,y) + \frac{1}{4t}|y|^2 + \frac{1}{4s}|y|^2\\
&= \frac{1}{4t}|x|^2 - \frac{1}{2t}(x,y) + \frac{t+s}{4ts}|y|^2.
\end{aligned}$$

Therefore, the desired formula (4.26) follows from formulas (4.27) and (4.28).

(2) By using formula (4.26), we can prove the Chapman–Kolmogorov equation (4.25) as follows:

$$\begin{aligned}
&\int_{\mathbf{R}^n} K_t(x-y)\, K_s(y)\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} \cdot e^{-\frac{|y|^2}{4s}}\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} \int_{\mathbf{R}^n} e^{-\left(\frac{|x-y|^2}{4t} + \frac{|y|^2}{4s}\right)}\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x|^2}{4(t+s)} - \frac{t+s}{4ts}|y - \frac{s}{t+s}x|^2}\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} e^{-\frac{|x|^2}{4(t+s)}} \int_{\mathbf{R}^n} e^{-\frac{t+s}{4ts}|y - \frac{s}{t+s}x|^2}\, dy \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} e^{-\frac{|x|^2}{4(t+s)}} \int_{\mathbf{R}^n} e^{-\frac{t+s}{4ts}|z|^2}\, dz \\
&= \frac{1}{(4\pi t)^{n/2}} \frac{1}{(4\pi s)^{n/2}} e^{-\frac{|x|^2}{4(t+s)}} \left(\frac{t+s}{4ts}\right)^{-n/2} \int_{\mathbf{R}^n} e^{-|w|^2}\, dw \\
&= \frac{1}{(4\pi)^n} \frac{4^{n/2}}{(t+s)^{n/2}} e^{-\frac{|x|^2}{4(t+s)}} (\sqrt{\pi})^n = \frac{1}{(4\pi(t+s))^{n/2}} e^{-\frac{|x|^2}{4(t+s)}} \\
&= K_{t+s}(x).
\end{aligned}$$

The proof of Lemma 4.13 is complete.

Step 2-2: The next lemma proves that the family $\{T(t)\}_{t \geq 0}$ forms a semigroup:

Lemma 4.14. *For all $t, s > 0$, we have*

$$T(t)(T(s)u)(x) = T(t+s)u(x), \quad u \in C_0(\mathbf{R}^n). \tag{4.29}$$

Proof. By using the Chapman–Kolmogorov equation (4.25), we obtain that

$$\begin{aligned}
T(t)(T(s)u)(x) &= \int_{\mathbf{R}^n} K_t(x-y) T(s)u(y)\, dy \\
&= \int_{\mathbf{R}^n} K_t(x-y) \left(\int_{\mathbf{R}^n} K_s(y-z)u(z)\, dz\right) dy \\
&= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} K_t(x-y)\, K_s(y-z)\, dy\right) u(z)\, dz
\end{aligned}$$

$$= \int_{\mathbf{R}^n} K_{t+s}(x-z)u(z)\,dz$$
$$= T(t+s)u(x) \quad \text{for every } u \in C_0(\mathbf{R}^n).$$

The proof of Lemma 4.14 is complete.

Step 3: Finally, the next lemma proves that the operators $\{T(t)\}_{t>0}$ converge strongly to the identity operator I:

Lemma 4.15. *We have, for all $u \in C_0(\mathbf{R}^n)$,*

$$\lim_{t \downarrow 0} \|T(t)u - u\|_\infty = 0. \tag{4.30}$$

Proof. Since we have

$$\frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|y-y|^2}{4t}}\,dy = \frac{1}{(4\pi t)^{n/2}} \left(\frac{1}{4t}\right)^{-n/2} \int_{\mathbf{R}^n} e^{-|w|^2}\,dw = 1,$$

it follows that

$$\begin{aligned} T(t)u(x) - u(x) &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u(y)\,dy - u(x) \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} (u(y) - u(x))\,dy. \end{aligned}$$

However, since $u \in C_0(\mathbf{R}^n)$ is uniformly continuous on $\mathbf{R}^n$, for any given number $\varepsilon > 0$ we can find a constant $\delta = \delta(\varepsilon) > 0$ such that

$$|x-y| < \delta \implies |u(x) - u(y)| < \varepsilon. \tag{4.31}$$

Then we decompose the term $T(t)u(x) - u(x)$ into the two terms:

$$\begin{aligned} T(t)u(x) - u(x) &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} (u(y) - u(x))\,dy \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{|x-y|<\delta} e^{-\frac{|x-y|^2}{4t}} (u(y) - u(x))\,dy \\ &\quad + \frac{1}{(4\pi t)^{n/2}} \int_{|x-y|\geq\delta} e^{-\frac{|x-y|^2}{4t}} (u(y) - u(x))\,dy \\ &:= I_3(x) + I_4(x). \end{aligned}$$

By condition (4.31), we can estimate the term $I_3(x)$ as follows:

$$|I_3(x)| \leq \frac{1}{(4\pi t)^{n/2}} \int_{|x-y|<\delta} e^{-\frac{|x-y|^2}{4t}} |u(y) - u(x)|\, dy \tag{4.32}$$
$$< \frac{\varepsilon}{(4\pi t)^{n/2}} \int_{|x-y|<\delta} e^{-\frac{|x-y|^2}{4t}}\, dy$$
$$< \frac{\varepsilon}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}}\, dy = \varepsilon \quad \text{for all } x \in \mathbf{R}^n.$$

The term $I_4(x)$ may be estimated as follows:

$$|I_4(x)| \leq \frac{1}{(4\pi t)^{n/2}} \int_{|x-y|\geq\delta} e^{-\frac{|x-y|^2}{4t}} |u(y) - u(x)|\, dy$$

However, since $u \in C_0(\mathbf{R}^n)$, we can find a positive constant M such that

$$|u(y) - u(x)| \leq |u(y)| + |u(x)| \leq 2M.$$

Hence we have the inequality

$$|I_4(x)| \leq \frac{2M}{(4\pi t)^{n/2}} \int_{|x-y|\geq\delta} e^{-\frac{|x-y|^2}{4t}}\, dy$$
$$= \frac{2M}{(4\pi t)^{n/2}} \int_{|z|\geq\delta} e^{-\frac{|z|^2}{4t}}\, dz \quad (z = 2\sqrt{t}w)$$
$$= \frac{2M}{(4\pi t)^{n/2}} \int_{2\sqrt{t}|w|\geq\delta} e^{-|w^2|} (2\sqrt{t})^n\, dz$$
$$= \frac{2M}{\pi^{n/2}} \int_{|w|\geq\delta/(2\sqrt{t})} e^{-|w^2|} dz.$$

This proves that

$$I_4(x) \longrightarrow 0 \quad \text{for all } t \downarrow 0. \tag{4.33}$$

It should be noted that this convergence is uniform in $x \in \mathbf{R}^n$.
Summing up, we obtain from assertions (4.32) and (4.33) that

$$\limsup_{t\downarrow 0} |T(t)u(x) - u(x)| \leq \limsup_{t\downarrow 0} (|I_3(x)| + |I_4(x)|) \leq \varepsilon \quad \text{for all } x \in \mathbf{R}^n.$$

This proves the desired assertion (4.30), since ε is arbitrary.
The proof of Lemma 4.15 is complete.

Now the proof of Example 4.11 is complete.

Remark 4.16. We can prove that the function

$$w(x,t) = T(t)u(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u(y)\, dy$$

is a solution of the following initial-value problem for the heat equation:

$$\begin{cases} \dfrac{\partial w}{\partial t} - \Delta w = 0 & \text{in } \mathbf{R}^n \times (0,\infty), \\ w(\cdot, 0) = u & \text{on } \mathbf{R}^n. \end{cases}$$

Roughly speaking, we may write the function $w(x,t)$ in the form

$$w(x,t) = T(t)u(x) = e^{t\,\Delta} u(x).$$

Physically, the function $w(x,t)$ represents the temperature at position x and time t in a homogeneous isotropic medium $\mathbf{R}^n$ with unit coefficient of thermal diffusivity, given that the temperature at position x and time 0 is $u(x)$.

4.5 The Hille–Yosida Theory of (C_0) Semigroups

This section is devoted to the Hille–Yosida theory of (C_0) semigroups, generalizing the theory of contraction semigroups developed in Sect. 4.4 (Theorem 4.28). Moreover, we study an initial-value problem associated with a (C_0) semigroup, and prove an existence and uniqueness theorem (Theorem 4.30).

4.5.1 Semigroups and Their Infinitesimal Generators

Let E be a Banach space and let $\mathcal{L}(E,E)$ be the space of all bounded linear operators on E into itself. The space $\mathcal{L}(E,E)$ is a Banach space with the operator norm

$$\|T\| = \sup_{\substack{x\in E \\ x\neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x\in E \\ \|x\|\leq 1}} \|Tx\|.$$

Definition 4.17. Let $\{T_t\}_{t\geq 0}$ be a family of linear operators on E. The family $\{T_t\}_{t\geq 0}$ is called a *semigroup* if it satisfies the following three conditions (i)–(iii):

(i) $T_t \in \mathcal{L}(E,E)$ for all $t \geq 0$.
(ii) $T_0 = I$ (the identity operator).
(iii) $T_t \cdot T_s = T_{t+s}$ for all $t, s \geq 0$ (the semigroup property).

Definition 4.18. Let $\{T_t\}_{t\geq 0}$ be a semigroup on E.

(i) The semigroup $\{T_t\}_{t\geq 0}$ is said to be *quasibounded* if there exist a positive constant M and a real number β such that

$$\|T_t\| \leq M\, e^{\beta t} \quad \text{for all } t \geq 0.$$

(ii) The semigroup $\{T_t\}_{t\geq 0}$ is said to be *bounded* if there exists a positive constant M such that

$$\|T_t\| \leq M \quad \text{for all } t \geq 0.$$

In particular, if $M = 1$, that is, if $\|T_t\| \leq 1$ for all $t \geq 0$, it is called a *contraction semigroup*.

Definition 4.19. If a semigroup $\{T_t\}_{t\geq 0}$ is strongly continuous for all $t \geq 0$, that is, if it satisfies the condition

$$\lim_{s\downarrow 0} \|T_{t+s}x - T_t x\| = 0 \quad \text{for every } x \in E,$$

then it is called a *semigroup of class* (C_0) or simply a (C_0) *semigroup*.

First, we prove the following fundamental result [Tn, Chapter 12, Theorem 12-1]:

Proposition 4.20. *Every (C_0) semigroup $\{T_t\}_{t\geq 0}$ is a quasi-bounded semigroup. More precisely, there exist a constant $M \geq 1$ and a real number ω such that*

$$\|T_t\| \leq M\, e^{\omega t} \quad \textit{for all } t \geq 0. \tag{4.34}$$

Proof. In this proof, we shall write

$$T(t) = T_t \quad \text{for every } t \geq 0.$$

Since $\{T(t)\}_{t\geq 0}$ is strongly continuous on the interval $[0, 1]$, it follows from an application of the principle of uniform boundedness (Theorem 3.14) that

$$M_1 \equiv \sup_{t\in[0,1]} \|T(t)\| < \infty.$$

If we denote by $[t]$ the integral part of t, that is, if we express $t = [t] + \tau$ with $0 \leq \tau < 1$, then we obtain from the semigroup property of $\{T(t)\}_{t\geq 0}$ that

$$\|T_t\| = \|T(1)^{[t]}T(\tau)\| \leq \|T(1)\|^{[t]}\|T(\tau)\| \leq M_1\, e^{[t]\log\|T(1)\|}.$$

(i) The case where $\|T(1)\| \geq 1$: Since $t > [\tau]$, we have the inequality

$$M_1\, e^{[t]\log\|T(1)\|} \leq M_1\, e^{t\log\|T(1)\|}$$

(ii) The case where $\|T(1)\| < 1$: Since $\log \|T(1)\| < 0$ and since $[t] = t-\tau \geq t-1$, it follows that

$$[t] \log \|T(1)\| \leq (t-1) \log \|T(1)\|.$$

Hence we have the inequality

$$M_1\, e^{[t] \log \|T(1)\|} \leq M_1\, e^{(t-1) \log \|T(1)\|} = M_1\, e^{-\log \|T(1)\|} e^{t \log \|T(1)\|}.$$

Summing up, we have proved that

$$\|T(t)\| \leq \begin{cases} M_1\, e^{t \log \|T(1)\|} & \text{if } \|T(1)\| \geq 1, \\ M_1\, e^{(t-1) \log \|T(1)\|} = M_1\, e^{-\log \|T(1)\|}\, e^{t \log \|T(1)\|} & \text{if } \|T(1)\| < 1. \end{cases}$$

Therefore, the desired inequality follows by taking

$$\begin{cases} M = M_1 & \text{if } \omega = \log \|T(1)\| \geq 0, \\ M = M_1\, e^{-\log \|T(1)\|} & \text{if } \omega = \log \|T(1)\| < 0. \end{cases}$$

Finally, we observe that

$$1 = \|T(0)\| \leq \lim_{t \downarrow 0} M\, e^{\omega t} = M.$$

The proof of Proposition 4.20 is complete.

Proposition 4.20 has a converse:

Proposition 4.21. *If a quasi-bounded semigroup $\{T_t\}_{t \geq 0}$ is strongly right-continuous at $t = 0$, then it is a (C_0) semigroup.*

Proof. For a quasi-bounded semigroup $\{T_t\}$, we let

$$S_t = e^{-\beta t} T_t, \quad t \geq 0,$$

then it follows that $\{S_t\}$ is a bounded semigroup. That is, we may assume that $\{T_t\}$ is a bounded semigroup:

$$\|T_t\| \leq M \quad \text{for all } t \geq 0.$$

If t_0 is an arbitrary positive number, it suffices to show that $\{T_t\}$ is strongly continuous at $t = t_0$. First, we have, for each $x \in E$,

$$T_{t_0+h}x - T_{t_0}x = (T_h - I)(T_{t_0}x) \longrightarrow 0 \quad \text{as } h \downarrow 0.$$

This proves that $\{T_t\}_{t \geq 0}$ is strongly right-continuous at $t = t_0$.

Secondly, we have, as $h \downarrow 0$,

$$\|T_{t_0-h}x - T_{t_0}x\| = \|T_{t_0-h}(I - T_h)x\| \leq M\|(T_h - I)x\| \longrightarrow 0.$$

This proves that $\{T_t\}_{t\geq 0}$ is strongly left-continuous at $t = t_0$.

Summing up, we have proved that $\{T_t\}_{t\geq 0}$ is strongly continuous at $t = t_0$.

The proof of Proposition 4.21 is complete.

Definition 4.22. If $\{T_t\}_{t\geq 0}$ is a (C_0) semigroup, then we let

$$\mathcal{D} = \left\{ x \in E : \text{the limit} \lim_{h\downarrow 0} \frac{T_h x - x}{h} \text{ exists} \right\},$$

and introduce a linear operator $\mathfrak{A}$ by the formulas:

(a) $D(\mathfrak{A}) = \mathcal{D}$.

(b) $\mathfrak{A}x = \lim_{h\downarrow 0} \dfrac{T_h x - x}{h}$ for every $x \in D(\mathfrak{A})$.

The operator $\mathfrak{A}$ is called the *infinitesimal generator* of the semigroup $\{T_t\}_{t\geq 0}$.

First, we derive a differential equation associated with a (C_0) semigroup in terms of its infinitesimal generator [Tn, Chapter 12, Theorem 12-3]:

Lemma 4.23. *If $\mathfrak{A}$ is the infinitesimal generator of a (C_0) semigroup $\{T_t\}_{t\geq 0}$, then we have the following two assertions (i) and (ii):*

(i) If $x \in D(\mathfrak{A})$, then it follows that $T_t x \in D(\mathfrak{A})$ for all $t > 0$ and further that

$$\mathfrak{A}(T_t x) = T_t(\mathfrak{A}x) \quad \text{for all } t > 0.$$

Moreover, the function $T_t x$, $x \in D(\mathfrak{A})$, is continuously differentiable for all $t > 0$, and satisfies the equation

$$\frac{d}{dt}(T_t x) = T_t(\mathfrak{A}x) = \mathfrak{A}(T_t x) \quad \text{for all } t > 0. \tag{4.35}$$

(ii) The operator $\mathfrak{A}$ is a closed linear operator.

Proof. (i) If $x \in D(\mathfrak{A})$, we have, for all $h > 0$,

$$\left(\frac{T_h - I}{h}\right) T_t x = T_t \left(\frac{T_h - I}{h} x\right).$$

However, since $x \in D(\mathfrak{A})$, it follows that

$$T_t \left(\frac{T_h x - x}{h}\right) \longrightarrow T_t(\mathfrak{A}x) \quad \text{as } h \downarrow 0.$$

Therefore, we obtain that

$$\left(\frac{T_h - I}{h}\right) T_t x = T_t \left(\frac{T_h - I}{h} x\right) \longrightarrow T_t(\mathfrak{A}x) \quad \text{as } h \downarrow 0.$$

This proves that $T_t \mathfrak{A} x \in D(\mathfrak{A})$ and that $\mathfrak{A}(T_t x) = T_t(\mathfrak{A}x)$.

Moreover, it follows that $T_t x$ is strongly right-differentiable:

$$\begin{aligned} \frac{d}{dt}^{+}(T_t x) &= \lim_{h \downarrow 0} \frac{T_{t+h}x - T_t x}{h} = \lim_{h \downarrow 0} T_t \left(\frac{T_h x - x}{h}\right) \\ &= T_t(\mathfrak{A}x). \end{aligned}$$

On the other hand, we have, for all sufficiently small $h > 0$ with $t - h > 0$,

$$\frac{T_{t-h}x - T_t x}{-h} - T_t(\mathfrak{A}x) = T_{t-h} \left(\frac{x - T_h x}{-h} - T_h(\mathfrak{A}x)\right). \tag{4.36}$$

However, it follows that

$$\begin{aligned} &\frac{x - T_h x}{-h} = \frac{T_h x - x}{h} \longrightarrow \mathfrak{A}x \quad \text{as } h \downarrow 0, \\ &T_h(\mathfrak{A}x) \longrightarrow \mathfrak{A}x \quad \text{as } h \downarrow 0. \end{aligned}$$

Since we have the inequality (see inequality (4.34))

$$\|T_{t-h}\| \leq M e^{\omega(t-h)},$$

we obtain from (4.36) that

$$\frac{T_{t-h}x - T_t x}{-h} - T_t(\mathfrak{A}x) = T_{t-h} \left(\frac{x - T_h x}{-h} - T_h(\mathfrak{A}x)\right) \longrightarrow 0 \quad \text{as } h \downarrow 0.$$

This proves that $T_t x$ is strongly left-differentiable:

$$\frac{d}{dt}^{-}(T_t x) = \lim_{h \downarrow 0} \frac{T_{t-h}x - T_t x}{-h} = T_t(\mathfrak{A}x).$$

Therefore, we have proved that $T_t x$ is strongly differentiable, and that Eq. (4.35) holds true. Moreover, we find that the derivative

$$\frac{d}{dx}(T_t x) = T_t(\mathfrak{A}\mathfrak{x}), \quad x \in D(\mathfrak{A}),$$

is strongly continuous.

(ii) Let $\{u_n\}$ be an arbitrary sequence in the domain $D(\mathfrak{A})$ such that $u_n \to u$ and $\mathfrak{A}u_n \to v$. We show that $u \in D(\mathfrak{A})$ and $\mathfrak{A}u = v$.

Since we have, for all $x \in D(\mathfrak{A})$,

$$\frac{d}{dt}(T_t x) = T_t(\mathfrak{A}x),$$

it follows that

$$T_t x - x = \int_0^t \frac{d}{ds}(T_s x)\, ds = \int_0^t T_s(\mathfrak{A}x)\, ds.$$

In particular, we have

$$T_t u_n - u_n = \int_0^t T_s(\mathfrak{A}u_n)\, ds. \tag{4.37}$$

However, we note that

$$T_t u_n - u_n \longrightarrow T_t u - u \quad \text{as } n \to \infty,$$

and that

$$\begin{aligned}\left\| \int_0^t (T_s(\mathfrak{A}u_n) - T_s v)\, ds \right\| &\le \int_0^t \|T_s(\mathfrak{A}u_n) - T_s v\| \\ &\le t \cdot \max_{s\in[0,t]} \|T_s\| \cdot \|\mathfrak{A}u_n - v\| \longrightarrow 0 \quad \text{as } n \to \infty.\end{aligned}$$

Therefore, by letting $n \to \infty$ in (4.37) we obtain that

$$T_t u - u = \int_0^t T_s v\, ds.$$

Since the function $T_s v$ is strongly continuous, we have

$$\lim_{t\downarrow 0} \frac{T_t u - u}{t} = \lim_{t\downarrow 0} \frac{1}{t} \int_0^t T_s v\, ds = T_s v|_{s=0} = v.$$

This proves that $u \in D(\mathfrak{A})$ and $\mathfrak{A}u = v$, so that the operator $\mathfrak{A}$ is closed.

The proof of Lemma 4.23 is complete.

Lemma 4.24. *The domain $D(\mathfrak{A})$ of the infinitesimal generator $\mathfrak{A}$ of a (C_0) semigroup $\{T_t\}_{t\ge 0}$ is* dense *in the space E.*

Proof. First, we choose a real-valued function $\varphi \in C_0^\infty(\mathbf{R})$ such that

$$\operatorname{supp} \varphi \subset \mathbf{R}^+ = (0, \infty).$$

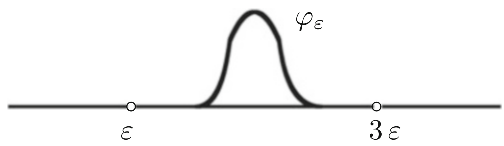


Fig. 4.2 The function φ_ε

If x is an arbitrary element of E, we let

$$u = \int_0^\infty \varphi(t) T_t x \, dt.$$

Then we have, for all sufficiently small $h > 0$,

$$\begin{aligned} T_h u &= \int_0^\infty \varphi(t)\, T_h(T_t x)\, dt = \int_0^\infty \varphi(t)\, T_{t+h} x \, dt = \int_h^\infty \varphi(s-h)\, T_s x \, ds \\ &= \int_0^\infty \varphi(s-h)\, T_s x \, ds, \end{aligned}$$

and hence

$$\frac{T_h u - u}{h} = -\int_0^\infty \frac{\varphi(s-h) - \varphi(s)}{-h} T_s x \, ds.$$

By letting $h \downarrow 0$ in this formula, we obtain from the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] that

$$\lim_{h \downarrow 0} \frac{T_h u - u}{h} = -\int_0^\infty \varphi'(t) T_t x \, dt. \tag{4.38}$$

This proves that $u \in D(\mathfrak{A})$ and that

$$\mathfrak{A} u = -\int_0^\infty \varphi'(t) T_t x \, dt.$$

For any given number $\varepsilon > 0$, we choose a real-valued function $\varphi_\varepsilon \in C_0^\infty(\mathbf{R})$ such that (see Fig. 4.2)

$$\operatorname{supp} \varphi_\varepsilon \subset [\varepsilon, 3\,\varepsilon]\,,$$
$$\int_0^\infty \varphi_\varepsilon(t)\, dt = 1.$$

If we let

$$u_\varepsilon = \int_0^\infty \varphi_\varepsilon(t) T_t x \, dt,$$

it follows from an application of formula (4.38) with $u := u_\varepsilon$ that $u_\varepsilon \in D(\mathfrak{A})$.

Moreover, since we have

$$x = \int_0^\infty \varphi_\varepsilon(t) x \, dt,$$

we obtain that

$$\begin{aligned} \|u_\varepsilon - x\| &\leq \int_0^\infty \varphi_\varepsilon(t) \|T_t x - x\| \, dt = \int_\varepsilon^{3\varepsilon} \varphi_\varepsilon(t) \|T_t x - x\| \, dt \\ &= \sup_{t \in [\varepsilon, 3\varepsilon]} \|T_t x - x\| \longrightarrow 0 \quad \text{as } \varepsilon \downarrow 0. \end{aligned}$$

This proves the density of $D(\mathfrak{A})$ in E.

The proof of Lemma 4.24 is complete.

Corollary 4.25. *Every (C_0) semigroup $\{T_t\}_{t \geq 0}$ is uniquely determined by its infinitesimal generator $\mathfrak{A}$.*

Proof. Assume that two (C_0) semigroups $\{T_t\}$ and $\{S_t\}$ have a closed linear operator $\mathfrak{A}$ as their infinitesimal generator. For any positive time t_0, it suffices to show that

$$T_{t_0} = S_{t_0}.$$

If x is an arbitrary element of the domain $D(\mathfrak{A})$, we let

$$W(t) = T_{t_0 - t} \cdot S_t \cdot x, \quad 0 \leq t \leq t_0.$$

Then it follows from an application of formula (4.22) that

$$\begin{aligned} \frac{d}{dt}(W(t)) &= \left(\frac{d}{dt}(T_{t_0 - t})\right) S_t x + T_{t_0 - t} \left(\frac{d}{dt}(S_t x)\right) \\ &= T_{t_0 - t}(-\mathfrak{A})(S_t x) + T_{t_0 - t}(\mathfrak{A} S_t x) = 0, \end{aligned}$$

so that

$$\frac{dW}{dt} \equiv 0 \quad \text{for all } t \in [0, t_0].$$

This implies that

$$T_{t_0} x = W(0) = W(t_0) = S_{t_0} x \quad \text{for all } x \in D(\mathfrak{A}). \tag{4.39}$$

However, we know from Lemma 4.24 that the domain $D(\mathfrak{A})$ is dense in E. Hence we have, by assertion (4.39),

$$T_{t_0} = S_{t_0},$$

since the operators T_{t_0} and S_{t_0} are bounded.

The proof of Corollary 4.25 is complete.

If $\{T_t\}_{t\geq 0}$ is a (C_0) semigroup, we shall write formally

$$T_t = e^{t\mathfrak{A}} = \exp(t\mathfrak{A}),$$

by using its infinitesimal generator $\mathfrak{A}$.

4.5.2 Infinitesimal Generators and Their Resolvents

Let E be a Banach space and let $A : E \to E$ be a closed linear operator with domain $D(A)$. We recall the following definitions:

(1) The *resolvent set* of A, denoted by $\rho(A)$, is defined to be the set of scalars $\lambda \in \mathbf{C}$ such that $\lambda I - A$ is injective and that $(\lambda I - A)^{-1} \in \mathcal{L}(E, E)$.
(2) If $\lambda \in \rho(A)$, the inverse operator $(\lambda I - A)^{-1}$ is called the *resolvent* of A, and is denoted by $R(\lambda; A)$:

$$R(\lambda; A) = (\lambda I - A)^{-1}, \quad \lambda \in \rho(A).$$

(3) The complement of $\rho(A)$ is called the *spectrum* of A, and is denoted by $\sigma(A)$: $\sigma(A) = \mathbf{C} \setminus \rho(A)$. The set $\sigma_p(A)$ of scalars $\lambda \in \mathbf{C}$ such that the operator $\lambda I - A$ is not one-to-one forms a subset of $\sigma(A)$, and is called the *point spectrum* of A. A scalar $\lambda \in \mathbf{C}$ belongs to $\sigma_p(A)$ if and only if there exists a non-zero element $x \in E$ such that $Ax = \lambda x$. In this case, λ is called an *eigenvalue* of A and x an *eigenvector* of A corresponding to λ. Also the null space $N(\lambda I - A)$ of $\lambda I - A$ is called the *eigenspace* of A corresponding to λ, and the dimension of $N(\lambda I - A)$ is called the *geometric multiplicity* of λ.

First, we have the following results [Tn, Chapter 8, Theorem 8-2]:

Lemma 4.26. *Let $A : E \to E$ be a closed linear operator with domain $D(A)$. If $\lambda, \mu \in \rho(A)$, we have the following two formulas (i) and (ii):*

(i) $(\lambda - A)^{-1} - (\mu - A)^{-1} = (\mu - \lambda)(\lambda - A)^{-1}(\mu - A)^{-1}$.
(ii) $(\lambda - A)^{-1}(\mu - A)^{-1} = (\mu - A)^{-1}(\lambda - A)^{-1}$.

The formula (i) is called the resolvent equation.

Proof. (i) The first formula may be proved as follows:

$$\begin{aligned}(\lambda - A)^{-1} - (\mu - A)^{-1} &= (\lambda - A)^{-1}(\mu - A)(\mu - A)^{-1} \\ &\quad -(\lambda - A)^{-1}(\lambda - A)(\mu - A)^{-1} \\ &= (\lambda - A)^{-1}\{(\mu - A) - (\lambda - A)\}(\mu - A)^{-1} \\ &= (\lambda - A)^{-1}(\mu - \lambda)(\mu - A)^{-1} \\ &= (\mu - \lambda)(\lambda - A)^{-1}(\mu - A)^{-1}.\end{aligned}$$

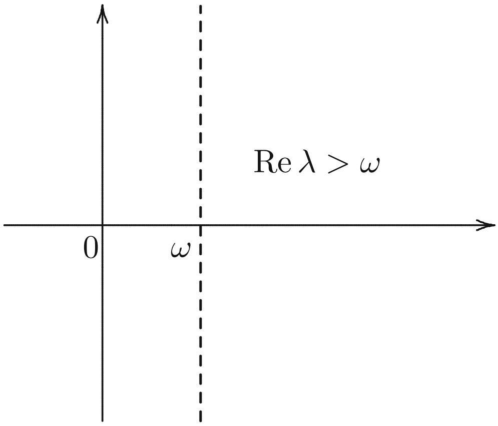


Fig. 4.3 The half-space $\{\mathrm{Re}\,\lambda > \omega\}$

(ii) Moreover, if we interchange λ and μ in formula (i), it follows that

$$(\mu - A)^{-1} - (\lambda - A)^{-1} = (\lambda - \mu)(\mu - A)^{-1}(\lambda - A)^{-1}.$$

Therefore, by combining this formula with formula (i) we obtain that

$$\begin{aligned}(\mu - A)^{-1}(\lambda - A)^{-1} &= \frac{1}{\lambda - \mu}\left((\mu - A)^{-1} - (\lambda - A)^{-1}\right) \\ &= \frac{1}{\lambda - \mu}\left((\lambda - \mu)(\lambda - A)^{-1}(\mu - A)^{-1}\right) \\ &= (\lambda - A)^{-1}(\mu - A)^{-1}.\end{aligned}$$

The proof of Lemma 4.26 is complete.

Let $\{T(t)\}_{t\geq 0}$ be a (C_0) semigroup and let A be its infinitesimal generator. The next theorem characterizes the resolvent set $\rho(A)$ and the resolvent $R(\lambda; A) = (\lambda I - A)^{-1}$ [Tn, Chapter 12, Theorem 12-4]:

Theorem 4.27. *Let $\{T(t)\}_{t\geq 0}$ be a (C_0) semigroup that satisfies the inequality*

$$\|T_t\| \leq M\, e^{\omega t} \quad \text{for all } t \geq 0. \tag{4.40}$$

Then we have the following two assertions (i) and (ii):

(i) The infinitesimal generator A of $\{T_t\}_{t\geq 0}$ is a closed operator and its domain $D(A)$ is dense in E.
(ii) The resolvent set $\rho(A)$ of A contains the half-plane $\{\lambda \in \mathbf{C} : \mathrm{Re}\,\lambda > \omega\}$ (see Fig. 4.3) and the resolvent $R(\lambda; A) = (\lambda I - A)^{-1}$ is expressed in the integral form

$$R(\lambda; A)u = \int_0^\infty e^{-\lambda t} T(t)u\, dt, \quad \operatorname{Re}\lambda > \omega, \quad u \in E. \tag{4.41}$$

Moreover, we have the inequalities for the powers of $R(\lambda; A)$

$$\|R(\lambda; A)^m\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^m}, \quad \operatorname{Re}\lambda > \omega, \ \ m = 1, 2, \ldots. \tag{4.42}$$

Proof. The proof is divided into six steps.

Step 1: First, if λ is a complex number such that $\operatorname{Re}\lambda > \omega$, then we have, by inequality (4.40),

$$\begin{aligned}
\int_0^\infty |e^{-\lambda t}| \|T(t)u\|\, dt &\leq \int_0^\infty |e^{-\lambda t}| \|T(t)\| \|u\|\, dt \\
&\leq M\|u\| \int_0^\infty e^{-(\operatorname{Re}\lambda - \omega)t}\, dt = M\|u\| \left[-\frac{e^{-(\operatorname{Re}\lambda-\omega)t}}{\operatorname{Re}\lambda - \omega} \right]_0^\infty \\
&= \frac{M\|u\|}{\operatorname{Re}\lambda - \omega} \quad \text{for all } u \in E.
\end{aligned}$$

Hence, if we let

$$\tilde{R}(\lambda)u = \int_0^\infty e^{-\lambda t} T(t)u\, dt \quad \text{for every } u \in E, \tag{4.43}$$

it follows that $\tilde{R}(\lambda) \in \mathcal{L}(x)$ and that

$$\|\tilde{R}(\lambda)\| \leq \frac{M}{\operatorname{Re}\lambda - \omega}, \quad \operatorname{Re}\lambda > \omega. \tag{4.44}$$

Step 2: Secondly, we show that, for all $u \in E$,

$$\begin{cases} \tilde{R}(\lambda)u \in D(A), \\ (\lambda I - A)\tilde{R}(\lambda)u = u. \end{cases} \tag{4.45}$$

For all sufficiently small $h > 0$, it follows that

$$\begin{aligned}
&\frac{T(h)(\tilde{R}(\lambda)u) - \tilde{R}(\lambda)u}{h} \\
&= \frac{1}{h}\int_0^\infty e^{-\lambda t} T(h)T(t)u\, dt - \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)u\, dt \\
&= \frac{1}{h}\int_0^\infty e^{-\lambda t} \{T(h+t) - T(t)\}u\, dt
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t+h)u\,dt - \frac{1}{h}\int_0^\infty T(t)u\,dt \\
&= \frac{1}{h}\int_h^\infty e^{-\lambda(s-h)} T(s)u\,ds - \frac{1}{h}\int_0^\infty T(t)u\,dt \\
&= \frac{1}{h}\left\{\int_0^\infty e^{-\lambda(t-h)} T(t)u\,dt - \int_0^h e^{-\lambda(t-h)} T(t)u\,dt\right\} \\
&\quad -\frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)u\,dt \\
&= \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t} T(t)u\,dt - \frac{1}{h}\int_0^h e^{-\lambda h(1-s)} T(hs)u\,ds \\
&:= I_1 + I_2.
\end{aligned}$$

However, we have, as $h \downarrow 0$,

$$I_1 = \frac{e^{\lambda h}-1}{h}\tilde{R}(\lambda)u = \lambda \frac{e^{\lambda h}-1}{\lambda h}\tilde{R}(\lambda)u \longrightarrow \lambda \tilde{R}(\lambda)u,$$

and also

$$I_2 = -\frac{1}{h}\int_0^h e^{-\lambda h(1-s)} T(hs)u\,ds \longrightarrow -T(0)u = -u.$$

Therefore, we have proved that

$$\begin{aligned}
&\lim_{h\downarrow 0} \frac{T(h)(\tilde{R}(\lambda)u) - (\tilde{R}(\lambda)u)}{h} \qquad (4.46)\\
&= \lim_{h\downarrow 0} \frac{1}{h}\int_0^\infty e^{-\lambda t}\{T(h+t) - T(t)\}u\,dt = \lambda \tilde{R}(\lambda)u - u.
\end{aligned}$$

This proves that

$$\begin{cases} \tilde{R}(\lambda)u \in D(A), \\ A(\tilde{R}(\lambda)u) = \lambda \tilde{R}(\lambda)u - u, \end{cases}$$

or equivalently,

$$(\lambda I - A)\tilde{R}(\lambda)u = u \quad \text{for all } u \in E.$$

We remark that the operator $\lambda I - A$ is surjective for all $\operatorname{Re}\lambda > \omega$.

Step 3: Thirdly, we show that

$$\tilde{R}(\lambda)(\lambda I - A)u = u \quad \text{for all } u \in D(A). \tag{4.47}$$

We have, by formula (4.43) and assertion (4.46),

$$\begin{aligned}\tilde{R}(\lambda)(Au) &= \lim_{h\downarrow 0} \tilde{R}(\lambda)\left(\frac{T(h)u-u}{h}\right)\\ &= \lim_{h\downarrow 0}\frac{1}{h}\left\{\int_0^\infty e^{-\lambda t}T(h)T(t)u\,dt - \int_0^\infty e^{-\lambda t}T(t)u\,dt\right\}\\ &= \lim_{h\downarrow 0}\frac{1}{h}\left\{\int_0^\infty e^{-\lambda t}(T(h+t)-T(t))u\,dt\right\}\\ &= \lambda\tilde{R}(\lambda)u - u.\end{aligned}$$

This proves the desired assertion (4.47). We remark that the operator $\lambda I - A$ is injective for all $\operatorname{Re}\lambda > \omega$.

Step 4: Fourthly, we show that

$$\lim_{\lambda\downarrow\infty} \lambda\tilde{R}(\lambda)u = u \quad \text{for every } u \in E. \tag{4.48}$$

Since we have, for all $\lambda > \omega$,

$$\int_0^\infty \lambda e^{-\lambda t}\,dt = 1,$$

it follows that

$$\begin{aligned}\left\|u - \lambda\tilde{R}(\lambda)u\right\| &= \left\|\int_0^\infty \lambda e^{-\lambda t}u\,dt - \int_0^\infty \lambda e^{-\lambda t}T(t)u\,dt\right\|\\ &\leq \int_0^\infty \lambda e^{-\lambda t}\|u - T(t)u\|\,dt.\end{aligned}$$

However, for any given number $\varepsilon > 0$ we can find a number $\delta = \delta(\varepsilon) > 0$ such that

$$0 \leq t < \delta \Longrightarrow \|u - T(t)u\| < \varepsilon.$$

Then we decompose the integral $\int_0^\infty \lambda e^{-\lambda t}\|u - T(t)u\|\,dt$ into the two terms I_3 and I_4:

$$\begin{aligned}
&\int_0^\infty \lambda e^{-\lambda t} \|u - T(t)u\| \, dt \\
&= \int_0^\delta \lambda e^{-\lambda t} \|u - T(t)u\| \, dt + \int_\delta^\infty \lambda e^{-\lambda t} \|u - T(t)u\| \, dt \\
&:= I_3 + I_4.
\end{aligned}$$

The term I_3 may be estimated as follows:

$$I_3 \leq \varepsilon \int_0^\delta \lambda e^{-\lambda t} \, dt = \varepsilon \left[-e^{-\lambda t} \right]_0^\delta = \varepsilon (1 - e^{-\lambda \delta}) < \varepsilon.$$

The term I_4 may be estimated as follows:

$$\begin{aligned}
I_4 &\leq \int_\delta^\infty \lambda e^{-\lambda t} (1 + \|T(t)\|) \, \|u\| \, dt \\
&\leq \|u\| \int_\delta^\infty \lambda e^{-\lambda t} (1 + M e^{\omega t}) \, dt \\
&= \|u\| \int_\delta^\infty (\lambda e^{-\lambda t} + M \lambda e^{-(\lambda - \omega) t}) \, dt = \|u\| \left[-e^{-\lambda t} - \frac{M \lambda e^{-(\lambda - \omega) t}}{\lambda - \omega} \right]_\delta^\infty \\
&= \|u\| \left(e^{-\lambda \delta} + \frac{M \lambda e^{-(\lambda - \omega) \delta}}{\lambda - \omega} \right).
\end{aligned}$$

Therefore, we obtain that

$$\limsup_{\lambda \to \infty} \|u - \lambda \tilde{R}(\lambda) u\| \leq \lim_{\lambda \downarrow \infty} (I_3 + I_4) \leq \varepsilon.$$

This proves the desired assertion (4.48), since $\varepsilon > 0$ is arbitrary.

Step 5: By combining assertions (4.45) and (4.48), we obtain that the domain $D(A)$ is dense in the space E. Moreover, it follows from assertions (4.45) and (4.47) and inequality (4.44) that $\tilde{R}(\lambda) = (\lambda I - A)^{-1}$ is the resolvent of A, that is,

$$R(\lambda; A)u = \tilde{R}(\lambda)u = \int_0^\infty e^{-\lambda t} T(t) u \, dt, \quad \operatorname{Re} \lambda > \omega, \quad u \in E,$$

and further that $A = \lambda I - \tilde{R}(\lambda)^{-1}$ is a closed linear operator.

Step 6: Finally, we prove inequality (4.42). If we differentiate formula (4.41) with respect to λ, it follows that

$$\frac{d}{d\lambda} (R(\lambda; A)u) = -\int_0^\infty t e^{-\lambda t} T(t) u \, dt.$$

On the other hand, by using the resolvent equation in the Banach space $\mathcal{L}(E, E)$ we obtain that

$$\frac{d}{d\lambda}\left((\lambda I - A)^{-1}\right) = -(\lambda I - A)^{-2} = -R(\lambda; A)^2.$$

Hence we have

$$R(\lambda; A)^2 = \int_0^\infty t\, e^{-\lambda t} T(t) u\, dt.$$

Similarly, if we differentiate this formula with respect to λ, we obtain that

$$\frac{d}{d\lambda}\left(R(\lambda; A)^2 u\right) = -\int_0^\infty t^2 e^{-\lambda t} T(t) u\, dt,$$

and that

$$\frac{d}{d\lambda}(\lambda I - A)^{-2} = -2(\lambda I - A)^{-3} = -2R(\lambda; A)^3.$$

Hence we have

$$2R(\lambda; A)^3 u = \int_0^\infty t^2 e^{-\lambda t} T(t) u\, dt,$$

or equivalently,

$$R(\lambda; A)^3 u = \frac{1}{2}\int_0^\infty t^2 e^{-\lambda t} T(t) u\, dt.$$

Continuing this process, we have, after $m - 1$ steps,

$$R(\lambda; A)^m u = \frac{1}{(m-1)!}\int_0^\infty t^{m-1} e^{-\lambda t} T(t) u\, dt.$$

Moreover, by using inequality (4.40) we obtain that

$$\|R(\lambda; A)^m u\| \leq \frac{M\,\|u\|}{(m-1)!}\int_0^\infty t^{m-1} e^{-(\operatorname{Re}\lambda - \omega)t}\, dt.$$

However, the integral on the right-hand side can be calculated as follows:

$$\begin{aligned}
&\int_0^\infty t^{m-1} e^{-(\operatorname{Re}\lambda - \omega)t}\, dt \\
&= \left[-\frac{t^{m-1} e^{-(\operatorname{Re}\lambda - \omega)t}}{\operatorname{Re}\lambda - \omega}\right]_{t=0}^{t=\infty} + \int_0^\infty \frac{(m-1)t^{m-2} e^{-(\operatorname{Re}\lambda - \omega)t}}{\operatorname{Re}\lambda - \omega}\, dt
\end{aligned}$$

$$= \int_0^\infty \frac{(m-1)t^{m-2}e^{-(\operatorname{Re}\lambda-\omega)t}}{\operatorname{Re}\lambda-\omega}\,dt$$

$$\vdots$$

$$= \frac{(m-1)!}{(\operatorname{Re}\lambda-\omega)^{m-1}}\int_0^\infty e^{-(\operatorname{Re}\lambda-\omega)t}\,dt = \frac{(m-1)!}{(\operatorname{Re}\lambda-\omega)^{m-1}}\left[-\frac{e^{-(\operatorname{Re}\lambda-\omega)t}}{\operatorname{Re}\lambda-\omega}\right]_{t=0}^{t=\infty}$$

$$= \frac{(m-1)!}{(\operatorname{Re}\lambda-\omega)^{m}}.$$

Therefore, we have proved that

$$\|R(\lambda;A)^m u\| \leq \frac{M\,\|u\|}{(m-1)!}\cdot\frac{(m-1)!}{(\operatorname{Re}\lambda-\omega)^m} = \frac{M\,\|u\|}{(\operatorname{Re}\lambda-\omega)^m} \quad \text{for all } u \in E.$$

This proves the desired inequality (4.42).

Now the proof of Theorem 4.27 is complete.

4.5.3 The Hille–Yosida Theorem

Now we consider when a linear operator is the infinitesimal generator of some (C_0) semigroup. This question is answered by the following Hille–Yosida theorem [Tn, Chapter 12, Theorem 12-5]:

Theorem 4.28 (Hille–Yosida). *Let E be a Banach space, and let $A : E \to E$ be a closed linear operator with domain $D(A)$. The operator A is the infinitesimal generator of some (C_0) semigroup $\{T_t\}_{t\geq 0}$ if and only if it satisfies the following two conditions (i) and (ii):*

(i) The operator A is a densely defined, closed linear operator.
(ii) There exists a real number ω such that the half-line (ω,∞) is contained in the resolvent set $\rho(A)$ of A, and the resolvent $R(\lambda;A) = (\lambda I - A)^{-1}$ satisfies the inequality

$$\|R(\lambda;A)^m\| \leq \frac{M}{(\lambda-\omega)^m}, \quad \lambda > \omega, \quad m = 1,2,\ldots. \tag{4.49}$$

Proof. (I) The "only if" part follows immediately from Theorem 4.27.

(II) The proof of the "if" part is divided into four steps. In the following we shall write $T(t)$ for T_t.

Step 1: If n is a positive integer such that $n > \omega$, we let

$$J_n := \left(I - \frac{1}{n} A \right)^{-1} = n(nI - A)^{-1} = nR(n; A) \in \mathcal{L}(E, E),$$

$$A_n := AJ_n = nA(nI - A)^{-1} = \{-n(nI - A)^{-1} + n^2 I\}(nI - A)^{-1}$$
$$= -nI + n^2 R(n; A) \in \mathcal{L}(E, E).$$

The operators A_n are called *Yosida approximations*.

First, we show that

$$s - \lim_{n\to\infty} J_n = I, \tag{4.50}$$

$$\lim_{n\to\infty} A_n u = Au \quad \text{for every } u \in D(A). \tag{4.51}$$

Since we have the inequality

$$\|R(\lambda; A)^m\| \leq \frac{M}{(\lambda - \omega)^m}, \quad \lambda > \omega, \quad m = 1, 2, \ldots,$$

it follows that

$$\|R(n; A)\| = \left\| \frac{J_n}{n} \right\| \leq \frac{M}{n - \omega}, \tag{4.52}$$

so that

$$\|J_n\| \leq \frac{nM}{n - \omega}.$$

This proves that the operators J_n are uniformly bounded in the space $\mathcal{L}(E, E)$.

On the other hand, we have, for all $u \in D(A)$,

$$J_n u = nR(n; A)u$$
$$= R(n; A)(nu - Au + Au) = R(n; A)\{(nI - A)u - Au\}$$
$$= u + R(n; A)(Au).$$

However, it follows from inequality (4.52) that

$$\|R(n; A)(Au)\| \leq \|R(n; A)\| \cdot \|Au\| \leq \frac{M\|Au\|}{n - \omega} \longrightarrow 0 \quad \text{as } n \to \infty.$$

Hence we have, as $n \to \infty$,

$$J_n u \longrightarrow u \quad \text{for every } u \in D(A).$$

Since $D(A)$ is dense in E and since J_n are uniformly bounded, we obtain that

$$J_n u \longrightarrow u \quad \text{for every } u \in E \text{ as } n \to \infty.$$

This proves the desired assertion (4.50).

Moreover, since we have

$$\frac{1}{n} J_n A = (nI - A)^{-1} A = (nI - A)^{-1}\{nI - (nI - A)\} = n(nI - A)^{-1} - I,$$
$$\frac{1}{n} A J_n = A(nI - A)^{-1} = \{nI - (nI - A)\}(nI - A)^{-1} = n(nI - A)^{-1} - I,$$

it follows that the operators A and J_n are commutative on $D(A)$. Hence we have, by assertion (4.50),

$$A_n u = A J_n u = J_n(Au) \longrightarrow Au \quad \text{for every } u \in D(A) \text{ as } n \to \infty.$$

This proves the desired assertion (4.51).

Step 2: For any positive integer $n > \omega$, we let

$$T_n(t) = e^{tA_n} = e^{tAJ_n},$$

and show that

$$\|T_n(t)\| \le M e^{\frac{n\omega t}{n-\omega}}. \tag{4.53}$$

Since we have

$$T_n(t) = e^{t(-n + n^2 R(n;A))} = e^{-nt} \cdot e^{n^2 t R(n;A)},$$

it follows that

$$\begin{aligned}
\|T_n(t)\| &\le e^{-nt} \sum_{m=0}^{\infty} \frac{n^{2m} t^m}{m!} \|R(n;A)^m\| \\
&\le M\, e^{-nt} \sum_{m=0}^{\infty} \frac{n^{2m} t^m}{m!} \cdot \frac{1}{(n-\omega)^m} = M\, e^{-nt} \cdot e^{\frac{n^2 t}{n-\omega}} \\
&= M\, e^{\frac{n\omega t}{n-\omega}} \quad \text{for all positive integer } n > \omega.
\end{aligned}$$

This proves the desired inequality (4.53).

Step 3: Now we show the following two assertions (i) and (ii):

(i) The strong limit $T(t) := s - \lim T_n(t)$ exists in the Banach space $\mathcal{L}(E, E)$. More precisely, the function $T_n(t)u$, $u \in E$, converges to $T(t)u$ uniformly in t on bounded intervals of $[0, \infty)$.

(ii) The operators $\{T(t)\}_{t\geq 0}$ form a (C_0) semigroup, and satisfies the inequality $\|T(t)\| \leq M\, e^{\omega t}$ for all $t \geq 0$.

If we let

$$\omega_0 := \max\{0, \omega\},$$

then it follows that

$$\frac{n\omega t}{n-\omega} \leq 2\omega_0 t \quad \text{for all positive integer } n > 2\omega_0.$$

Let $\tau > 0$ be an arbitrary positive number. Then we have, by Step 2,

$$\|T_n(t)\| \leq M\, e^{\frac{n\omega t}{n-\omega}} \leq M\, e^{2\omega_0 t} \leq M\, e^{2\omega_0 \tau}, \quad t \in [0, \tau], \quad n > 2\omega_0. \tag{4.54}$$

On the other hand, by Theorem 4.4 it follows that

$$\frac{d}{dt}(T_n(t)) = \frac{d}{dt}\left(e^{tA_n}\right) = T_n(t)A_n.$$

Since A_m and A_n are commutative and since A_m and $T_n(s)$ are commutative, we have

$$\begin{aligned}
T_n(t) - T_m(t) &= [T_m(t-s)T_n(s)]_{s=0}^{s=t} \\
&= \int_0^t \frac{d}{ds}\{T_m(t-s)T_n(s)\}\, ds \\
&= \int_0^t \left\{\frac{d}{ds}T_m(t-s)\cdot T_n(s) + T_m(t-s)\cdot \frac{d}{ds}T_n(s)\right\}\, ds \\
&= \int_0^t \{-T_m(t-s)A_m T_n(s) + T_m(t-s)T_n(s)A_n\}\, ds \\
&= \int_0^t \{T_m(t-s)T_n(s)A_n - T_m(t-s)T_n(s)A_m\}\, ds \\
&= \int_0^t T_m(t-s)T_n(s)(A_n - A_m)\, ds \quad \text{for all } t \in [0, \tau].
\end{aligned}$$

In view of inequality (4.54), this proves that

$$\|T_n(t)u - T_m(t)u\| \leq M^2\, e^{4\omega_0 \tau}\tau \|A_n u - A_m u\|, \tag{4.55}$$
$$\text{for all } t \in [0, \tau] \text{ and all } n, m > 2\omega_0.$$

However, it follows from assertion (4.38) that we have, for all $u \in D(A)$,

$$A_n u \longrightarrow Au \quad \text{as } n \to \infty.$$

Therefore, by letting $n, m \to \infty$ in inequality (4.42) we find that the function $T_n(t)u$, $u \in D(A)$, converges uniformly in $t \in [0, \tau]$, for each $\tau > 0$.

We consider the general case where $u \in E$: For any given number $\varepsilon > 0$, we can find an element $v \in D(A)$ such that $\|u - v\| < \varepsilon$. Then we have the inequality

$$\begin{aligned} &\|T_n(t)u - T_m(t)u\| \\ &\leq \|T_n(t)(u-v)\| + \|T_n(t)v - T_m(t)v\| + \|T_m(t)(u-v)\| \\ &\leq 2M\,\varepsilon\, e^{2\omega_0\tau} + \|T_n(t)v - T_m(t)v\|, \quad t \in [0,\tau], \quad n, m > 2\omega_0. \end{aligned}$$

Since the function $T_n(t)v$, $v \in D(A)$, converges uniformly in $t \in [0, \tau]$, it follows that

$$\limsup_{n,m\to\infty} \|T_n(t)u - T_m(t)u\| \leq 2M\,\varepsilon\, e^{2\omega_0\tau} \quad \text{for all } t \in [0,\tau].$$

This proves that the function $T_n(t)u$, $u \in E$, also converges uniformly in $t \in [0, \tau]$, for each $\tau > 0$.

Therefore, we have proved that the function $T_n(t)u$, $u \in E$, converges uniformly in t over bounded intervals of $[0, \infty)$.

We can define a family $\{T(t)\}_{t\geq 0}$ of linear operators on E by the formula

$$T(t)u := \lim_{n\to\infty} T_n(t)u \quad \text{for every } u \in E.$$

First, we note that the function $T(t)u$, $u \in E$, is strongly continuous for all $t \geq 0$, since this convergence is uniform in t over bounded intervals of $[0, \infty)$. Secondly, it follows from an application of the principle of uniform boundedness (Theorem 3.14) that

$$\|T(t)\| \leq \liminf_{n\to\infty} \|T_n(t)\| \leq \liminf_{n\to\infty} M\, e^{\frac{n\omega t}{n-\omega}} = M\, e^{\omega t} \quad \text{for all } t \geq 0,$$

so that

$$T(t) \in \mathcal{L}(E, E) \quad \text{for all } t \geq 0.$$

Thirdly, since we have the group property for the operators

$$T_n(t+s) = T_n(t)T_n(s) \quad \text{for all } t, s \in \mathbf{R},$$

by passing to the limit we obtain the semigroup property for the operators

$$T(t+s) = T(t)T(s) \quad \text{for all } t, s \geq 0.$$

Summing up, we have proved that $\{T(t)\}_{t\geq 0}$ form a (C_0) semigroup, and satisfies the inequality $\|T(t)\| \leq M\, e^{\omega t}$ for all $t \geq 0$.

Step 4: Finally, we show that the infinitesimal generator $\mathfrak{A}$ of $\{T(t)\}_{t\geq 0}$ coincides with the operator A.

Since we have, for all $n > \omega_0$,

$$\frac{d}{dt}\left(T_n(t)u\right) = \frac{d}{dt}\left(e^{tA_n}u\right) = A_n T_n(t)u = T_n(t)A_n u, \quad u \in D(A),$$

it follows that

$$T_n(h)u - u = \int_0^h \frac{d}{dt}\left(T_n(t)u\right)\, dt = \int_0^h T_n(t)A_n u\, dt. \tag{4.56}$$

Hence we have, by inequality (4.40),

$$\begin{aligned}\|T_n(t)(A_n u) - T(t)(Au)\| &\leq \|T_n(t)\|\,\|A_n u - Au\| + \|(T_n(t) - T(t))Au\| \\ &\leq M\, e^{2\omega h}\|A_n u - Au\| + \|(T_n(t) - T(t))(Au)\|.\end{aligned}$$

It should be noted that the convergence

$$T_n(t)(A_n u) \longrightarrow T(t)(Au)$$

is uniform in $t \in [0, \tau]$ as $n \to \infty$.

By letting $n \to \infty$ in (4.56), we obtain that

$$T(h)u - u = \int_0^h T(t)Au\, dt, \quad u \in D(A).$$

Hence it follows that

$$\lim_{h\downarrow 0} \frac{T(h)u - u}{h} = \lim_{h\downarrow 0} \frac{1}{h}\int_0^h T(t)Au\, dt = T(0)Au = Au.$$

This proves that

$$u \in D(\mathfrak{A}),$$
$$\mathfrak{A}u = Au,$$

so that $A \subset \mathfrak{A}$.

In order to prove that $A = \mathfrak{A}$, it suffices to show that $D(\mathfrak{A}) \subset D(A)$.

Let u be an arbitrary element of $D(\mathfrak{A})$. Since the operator

$$\lambda I - A : D(A) \longrightarrow E$$

is bijective for $\lambda > \omega$, we can find a unique element $v \in D(A)$ such that

$$(\lambda I - A)v = (\lambda I - \mathfrak{A})u.$$

However, since $A \subset \mathfrak{A}$, it follows that

$$(\lambda I - A)v = (\lambda I - \mathfrak{A})v,$$

so that

$$(\lambda I - \mathfrak{A})(u - v) = 0.$$

This proves that $u = v \in D(A)$, that is, $D(\mathfrak{A}) \subset D(A)$, since the operator

$$\lambda I - \mathfrak{A} : D(\mathfrak{A}) \longrightarrow E$$

is bijective for $\lambda > \omega$ (see assertion (ii) of Theorem 4.27).

Now the proof of Theorem 4.28 is complete.

The next corollary gives a simple necessary and sufficient condition for contraction (C_0) semigroups (cf. Theorem 4.10):

Corollary 4.29. *Let E be a Banach space, and let $A : E \to E$ be a densely defined, closed linear operator with domain $D(A)$. The operator A is the infinitesimal generator of some contraction (C_0) semigroup $\{T_t\}_{t \geq 0}$ if and only if the half-line $(0, \infty)$ is contained in the resolvent set $\rho(A)$ of A and the resolvent $R(\lambda; A) = (\lambda I - A)^{-1}$ satisfies the inequality*

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda} \quad \text{for all } \lambda > 0.$$

Proof. The "only if" part follows from Theorem 4.27 with $\omega := 0$ and $M := 1$.

The "if" part may be proved as follows: Since we have, for every integer $m \in \mathbf{N}$,

$$\|R(\lambda; A)^m\| \leq \|R(\lambda; A)\|^m \leq \frac{1}{\lambda^m} \quad \text{for all } \lambda > 0,$$

it follows that condition (ii) of Theorem 4.28 is satisfied with $\omega := 0$ and $M := 1$. Therefore, by applying Theorem 4.28 to our situation we obtain that A is the infinitesimal generator of some contraction (C_0) semigroup.

The proof of Corollary 4.29 is complete.

4.5.4 (C_0) Semigroups and Initial-Value Problems

Finally, we consider an initial-value problem associated with a (C_0) semigroup. More precisely, we prove the following existence and uniqueness theorem for an initial-value problem associated with a (C_0) semigroup ([CP, Chapitre 6, Théorème 6.9]; [Kr, Chapter I, Theorem 3.3]):

Theorem 4.30. *Let $\{T_t\}_{t\geq 0}$ be a (C_0) semigroup with infinitesimal generator A. If $x \in D(A)$, then the function $u(t) = T_t x$ is a unique solution of the initial-value problem*

$$\begin{cases} \dfrac{du}{dt} = Au & \text{for all } t > 0, \\ u(0) = x \end{cases} \tag{$*$}$$

which satisfies the following three conditions (a)–(c):

(a) The function $u(t)$ is continuously differentiable for all $t > 0$.
(b) $\|u(t)\| \leq M\, e^{\beta t}$ for all $t \geq 0$.
(c) $u(t) \to x$ as $t \downarrow 0$.

In other words, the initial-value problem $(*)$ is well-posed.

The proof of Theorem 4.30 is based on the following result on the Laplace transform:

Lemma 4.31. *Let $u(t)$ be an E-valued, bounded continuous function defined on the open interval $\mathbf{R}^+ = (0, \infty)$. If we have, for all $\lambda > 0$,*

$$\int_0^\infty e^{-\lambda t} u(t)\, dt = 0, \tag{4.57}$$

then it follows that

$$u(t) = 0 \quad \text{for all } t \geq 0.$$

Proof. (1) If f is an arbitrary element of the dual space E' of E, then it follows that the function

$$\mathbf{R}^+ \ni t \longrightarrow f(u(t))$$

is bounded and continuous. Moreover, it is easy to verify the formula

$$\int_0^\infty e^{-\lambda t} f(u(t))\, dt = f\left(\int_0^\infty e^{-\lambda t} u(t)\, dt\right), \tag{4.58}$$

since the integrals can be approximated by Riemann sums. Indeed, let

$$\Delta = \{t_0 = 0, t_1, \dots, t_n = M\}$$

be the division of the interval $[0, M]$ for any $M > 0$. Then we have, for the corresponding Riemann sums,

$$\sum_{i=0}^{n} e^{-\lambda t_i} f(u(t_i)) \cdot |\Delta| = f\left(\sum_{i=0}^{n} e^{-\lambda t_i} u(t_i) \cdot |\Delta|\right), \tag{4.59}$$

where

$$|\Delta| = \max_{1 \le i \le n} |t_i - t_{i-1}|.$$

However, by letting $|\Delta| \to 0$ in (4.59), we obtain that

$$\sum_{i=0}^{n} e^{-\lambda t_i} f(u(t_i)) \cdot |\Delta| \longrightarrow \int_0^M e^{-\lambda t} f(u(t))\, dt,$$

$$f\left(\sum_{i=0}^{n} e^{-\lambda t_i} u(t_i) \cdot |\Delta|\right) \longrightarrow f\left(\int_0^M e^{-\lambda t} u(t)\, dt\right).$$

Hence we have, for all $M > 0$,

$$\int_0^M e^{-\lambda t} f(u(t))\, dt = f\left(\int_0^M e^{-\lambda t} u(t)\, dt\right). \tag{4.60}$$

The desired formula (4.58) follows by letting $M \to \infty$ in (4.60).

(2) By combining condition (4.57) and formula (4.58), we obtain that

$$\int_0^\infty e^{-\lambda t} f(u(t))\, dt = 0 \quad \text{for all } \lambda > 0.$$

Hence we have, by the fundamental property of the Laplace transform,

$$f(u(t)) = 0 \quad \text{for all } t \ge 0.$$

This proves that $u(t) = 0$ for all $t \ge 0$, since f is an arbitrary element of the dual space E'.

The proof of Lemma 4.31 is complete.

Corollary 4.32. *Let $u(t)$ be an E-valued, bounded continuous function defined on the open interval $\mathbf{R}^+ = (0, \infty)$. Assume that there exist a constant $M \ge 1$ and a real number β such that*

$$\|u(t)\| \leq M\, e^{\beta t} \quad \textit{for all } t \geq 0. \tag{4.61}$$

If we have, for all $\lambda > \beta$*,*

$$\int_0^\infty e^{-\lambda t} u(t)\, dt = 0,$$

then it follows that

$$u(t) = 0 \quad \textit{for all } t \geq 0.$$

Proof. If we let

$$v(t) = u(t) e^{-\beta t},$$
$$\xi = \lambda - \beta,$$

then it follows that the function $v(t)$ satisfies the bounded condition

$$\|v(t)\| \leq M \quad \text{for all } t \geq 0,$$

and the condition

$$\int_0^\infty e^{-\xi t} v(t)\, dt = \int_0^\infty e^{-(\lambda-\beta)t} v(t)\, dt = \int_0^\infty e^{-\lambda t} u(t)\, dt$$
$$= 0 \quad \text{for all } \xi = \lambda - \beta > 0.$$

Therefore, by applying Lemma 4.31 we obtain that

$$v(t) = u(t)\, e^{-\beta t} = 0 \quad \text{for all } t \geq 0,$$

so that

$$u(t) = 0 \quad \text{for all } t \geq 0.$$

The proof of Corollary 4.32 is complete.

Proof of Theorem 4.30

(1) By Lemma 4.23 with $\mathfrak{A} := A$, it follows that the function $u(t) = T_t x$, $x \in D(A)$, is a solution of problem $(*)$.
(2) We have only to show the *uniqueness* of the solution. Assume that u_1 and u_2 are two solutions of problem $(*)$ which satisfy conditions (a)–(c):

$$\|u_1(t)\| \leq M\, e^{\beta_1 t} \quad \text{for all } t \geq 0,$$
$$\|u_2(t)\| \leq M\, e^{\beta_2 t} \quad \text{for all } t \geq 0.$$

Then it follows that the function

$$v(t) = u_1(t) - u_2(t)$$

is a solution of the initial-value problem

$$\begin{cases} \dfrac{dv}{dt} = Av, \\ v(0) = 0 \end{cases}$$

which satisfies the condition

$$\|v(t)\| \leq 2M\, e^{\beta t} \quad \text{for all } t \geq 0, \tag{4.62}$$

where

$$\beta = \max\{\beta_1, \beta_2\}.$$

Now we take an arbitrary real number λ such that $\lambda > \beta$, and let

$$W(t) = e^{-\lambda t} v(t) = e^{-\lambda t}\left(u_1(t) - u_2(t)\right).$$

Then it follows that

$$\begin{aligned} \frac{d}{dt}(W(t)) &= \frac{d}{dt}\left(e^{-\lambda t} v(t)\right) = -\lambda W(t) + e^{-\lambda t} \frac{d}{dt} v(t) \\ &= -\lambda W(t) + e^{-\lambda t} A v(t) = -(\lambda I - A) W(t), \end{aligned}$$

so that

$$\begin{aligned} \int_0^s W(t)\, dt &= -(\lambda I - A)^{-1} \int_0^s \frac{dW(t)}{dt}\, dt \\ &= -(\lambda I - A)^{-1} W(s), \end{aligned}$$

since $W(0) = v(0) = 0$. Hence we have, by inequality (4.62),

$$\begin{aligned} \left\| \int_0^s W(t)\, dt \right\| &= \left\| \int_0^s e^{-\lambda t} v(t)\, dt \right\| = \left\| -(\lambda I - A)^{-1} W(s) \right\| \\ &= \left\| -(\lambda I - A)^{-1} \left(e^{-\lambda s} v(s)\right) \right\| \\ &\leq \frac{1}{\lambda - \beta}\, e^{-\lambda s}\, 2M\, e^{\beta s} = \frac{2M}{\lambda - \beta} e^{-(\lambda - \beta) s}. \end{aligned}$$

Since $\lambda > \beta$, it follows that

$$\left\| \int_0^s W(t)\, dt \right\| \leq \frac{2M}{\lambda - \beta} e^{-(\lambda-\beta)s} \longrightarrow 0 \quad \text{as } s \to \infty.$$

Therefore, we have proved that

$$\int_0^\infty e^{-\lambda t} v(t)\, dt = \int_0^\infty W(t)\, dt = 0 \quad \text{for all } \lambda > \beta.$$

In view of condition (4.62), by applying Corollary 4.32 to the function $v(t)$ we obtain that

$$v(t) = u_1(t) - u_2(t) = 0 \quad \text{for all } t \geq 0,$$

so that

$$u_1(t) = u_2(t) \quad \text{for all } t \geq 0.$$

This proves the uniqueness theorem for problem $(*)$.

Now the proof of Theorem 4.30 is complete. □

4.6 Notes and Comments

Hille–Phillips [HP] and Yosida [Yo] are the classics for semigroup theory. The material in this chapter is adapted from Chazarain–Piriou [CP], Friedman [Fr1], J. Goldstein [Gj], S.G. Kreĭn [Kr], Pazy [Pa], Tanabe [Tn] and also part of Taira [Ta7]. For more leisurely treatments of semigroups, the reader is referred to Engel–Nagel [EN].

Many problems in partial differential equations can be formulated in terms of abstract operators acting between suitable Banach spaces of distributions, and these operators are then analyzed by the methods of semigroup theory. The virtue of this approach is that a given problem is stripped of extraneous data, so that the analytic core of the problem is revealed.

For example, Taira [Ta12] is devoted to a semigroup approach to an initial-boundary value problem of *linear elastodynamics* in the case where the boundary condition is a regularization of the genuine mixed displacement-traction boundary condition (see [MH, Chapter 6, Section 6.3]). More precisely, let Ω be an open, connected subset of Euclidean space $\mathbf{R}^n$, $n \geq 2$, with smooth boundary $\partial\Omega$. We think of its closure $\overline{\Omega} = \Omega \cup \partial\Omega$ as representing the volume occupied by an undeformed body. In [Ta12] we study the following initial-boundary value problem of linear elastodynamics: For given $\mathbf{R}^n$-valued functions $\boldsymbol{f}(x) = (f_i(x))$,

$\boldsymbol{u}_0(x) = (u_{0,i}(x))$ and $\boldsymbol{u}_1(x) = (u_{1,i}(x))$ defined in Ω, find an $\mathbf{R}^n$-valued function $\boldsymbol{u}(x) = (u_i(x))$ in Ω such that

$$\begin{cases} \dfrac{\partial^2 \boldsymbol{u}}{\partial t^2} - \operatorname{div}(\mathbf{a}(x) \cdot \nabla \boldsymbol{u}) = \boldsymbol{f} & \text{in } \Omega \times (0, \infty), \\ \boldsymbol{u}|_{t=0} = \boldsymbol{u}_0 & \text{in } \Omega, \\ \frac{\partial \boldsymbol{u}}{\partial t}|_{t=0} = \boldsymbol{u}_1 & \text{in } \Omega, \\ \alpha(x)\,(\mathbf{a}(x) \cdot \nabla \boldsymbol{u} \cdot \mathbf{n}) + (1 - \alpha(x))\boldsymbol{u} = \mathbf{0} & \text{on } \partial\Omega \times (0, \infty). \end{cases} \tag{4.63}$$

Here:

(1) $\mathbf{a}(x) = (a_{ij\ell m}(x))$ is a smooth elasticity tensor.
(2) $\alpha(x)$ is a smooth real-valued function on $\partial\Omega$ such that $0 \le \alpha(x) \le 1$ on $\partial\Omega$.
(3) $\mathbf{n} = (n_i)$ is the outward unit normal to $\partial\Omega$.

It is worth pointing out that, componentwise, the initial-boundary value problem (4.63) can be written in the form

$$\begin{cases} \dfrac{\partial^2 u_i}{\partial t^2} - \sum_{j=1}^n \dfrac{\partial}{\partial x_j}\left(\sum_{\ell,m=1}^n a_{ij\ell m}(x)\dfrac{\partial u_\ell}{\partial x_m}\right) = f_i(x), \\ u_i|_{t=0} = u_{0,i}(x), \\ \frac{\partial u_i}{\partial t}|_{t=0} = u_{1,i}(x), \\ \alpha(x)\sum_{j=1}^n \left(\sum_{\ell,m=1}^n a_{ij\ell m}(x)\dfrac{\partial u_\ell}{\partial x_m}\right) n_j(x) + (1 - \alpha(x))u_i(x) = 0. \end{cases}$$

It should be emphasized that our boundary condition

$$\boldsymbol{B}_\alpha \boldsymbol{u} = \alpha(x)\,(\mathbf{a}(x) \cdot \nabla \boldsymbol{u} \cdot \mathbf{n}) + (1 - \alpha(x))\boldsymbol{u} \tag{4.64}$$

is a smooth linear combination of displacement and traction boundary conditions. It is easy to see that $\boldsymbol{B}_\alpha$ is non-degenerate (or coercive) if and only if either $\alpha(x) > 0$ on $\partial\Omega$ (the Robin case) or $\alpha(x) \equiv 0$ on $\partial\Omega$ (the Dirichlet case). However, our boundary condition (4.64) is *degenerate* from an analytical point of view. This is due to the fact that the so-called Shapiro–Lopatinskii complementary condition is violated at the points $x \in \partial\Omega$ where $\alpha(x) = 0$ (cf. [Ho4]).

The crucial point in our semigroup approach is to generalize the classical variational approach to the degenerate case, by using the theory of fractional powers of analytic semigroups [Ta7].

Finally, we give two simple but important examples of the initial-boundary value problem (4.63):

Example 4.33. If we take

$$\mathbf{a}(x) = (a_{ij\ell m}(x)) = \left(\delta_{i\ell}\delta_{jm}\right),$$

then our problem (4.63) becomes the mixed displacement-traction problem for the wave equation

$$\begin{cases} \dfrac{\partial^2 \boldsymbol{u}}{\partial t^2} - \Delta \boldsymbol{u} = \boldsymbol{f} & \text{in } \Omega \times (0, \infty), \\ \boldsymbol{u}|_{t=0} = \boldsymbol{u}_0 & \text{in } \Omega, \\ \frac{\partial \boldsymbol{u}}{\partial t}|_{t=0} = \boldsymbol{u}_1 & \text{in } \Omega, \\ \alpha(x) \frac{\partial \boldsymbol{u}}{\partial \mathbf{n}} + (1 - \alpha(x))\boldsymbol{u} = \mathbf{0} & \text{on } \partial\Omega \times (0, \infty). \end{cases}$$

Example 4.34. If we take

$$\mathbf{a}(x) = (a_{ij\ell m}(x)) = \left(\lambda \delta_{ij} \delta_{\ell m} + \mu \delta_{i\ell} \delta_{jm}\right),$$

where λ, μ are Lamé moduli, then our problem (4.63) becomes the mixed displacement-traction problem for the elastodynamic wave equation

$$\begin{cases} \dfrac{\partial^2 \boldsymbol{u}}{\partial t^2} - (\mu \Delta \boldsymbol{u} + (\lambda + \mu) \operatorname{grad}(\operatorname{div} \boldsymbol{u})) = \boldsymbol{f} & \text{in } \Omega \times (0, \infty), \\ \boldsymbol{u}|_{t=0} = \boldsymbol{u}_0 & \text{in } \Omega, \\ \frac{\partial \boldsymbol{u}}{\partial t}|_{t=0} = \boldsymbol{u}_1 & \text{in } \Omega, \\ \alpha(x) \left(\boldsymbol{\tau}(\boldsymbol{u}) \cdot \mathbf{n}\right) + (1 - \alpha(x))\boldsymbol{u} = \mathbf{0} & \text{on } \partial\Omega \times (0, \infty). \end{cases}$$

Here we recall that

$$\boldsymbol{e}(\boldsymbol{u}) = (e_{ij}(x)) = \left(\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right)$$

is the linearized strain tensor and

$$\boldsymbol{\tau}(\boldsymbol{u}) = (\tau_{ij}(x)) = \left(\lambda \sum_{k=1}^{n} e_{kk}(x) \delta_{ij} + 2\mu e_{ij}(x)\right)$$

is the linearized stress tensor.

Part II
Elements of Partial Differential Equations

Chapter 5
Theory of Distributions

This chapter is a summary of the basic definitions and results from the theory of distributions or generalized functions which will be used in subsequent chapters. Distribution theory has become a convenient tool in the study of partial differential equations. Many problems in partial differential equations can be formulated in terms of abstract operators acting between suitable spaces of distributions, and these operators are then analyzed by the methods of functional analysis. The virtue of this approach is that a given problem is stripped of extraneous data, so that the analytic core of the problem is revealed.

Section 5.1 serves to settle questions of notation and such. In Sect. 5.2 we study L^p spaces, the spaces of C^k functions and test functions, and also Hölder spaces on an open subset of Euclidean space. Moreover, we introduce Friedrichs' mollifiers and show how Friedrichs' mollifiers can be used to approximate a function by smooth functions (Theorem 5.4). In Sect. 5.3 we study differential operators and state that differential operators are local operators (Peetre's theorem 5.7). In Sect. 5.4 we present a brief description of the basic concepts and results in the theory of distributions. In particular, the importance of tempered distributions lies in the fact that they have Fourier transforms. In Sect. 5.4.10 we calculate the Fourier transform of a tempered distribution which is closely related to the stationary phase theorem (Example 5.29). In Sect. 5.5 we prove the Schwartz kernel theorem (Theorem 5.36) which characterizes continuous linear operators in terms of distributions. In Sect. 5.6 we describe the classical single and double layer potentials arising in the Dirichlet and Neumann problems for the Laplacian Δ in the case of the half-space $\mathbf{R}^n_+$ (formulas (5.73) and (5.74)). Moreover, we prove the Green representation formula (5.75). This formula will be formulated in terms of pseudo-differential operators in Chap. 7 (Sect. 7.5). Some results in Sects. 5.3–5.5 can be extended to distributions, differential operators, and operators and kernels on a manifold in Sect. 5.7. The virtue of manifold theory is that it provides a geometric insight into the study of partial differential equations, and intrinsic properties of partial differential equations may be revealed. In Sect. 5.8 we introduce the notion of domains of class C^r from the viewpoint of manifold theory.

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_5,

5.1 Notation

5.1.1 Points in Euclidean Spaces

Let $\mathbf{R}^n$ be the n-dimensional Euclidean space. We use the conventional notation

$$x = (x_1, x_2, \ldots, x_n).$$

If $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are points in $\mathbf{R}^n$, we set

$$x \cdot y = \sum_{j=1}^{n} x_j y_j,$$

$$|x| = \left(\sum_{j=1}^{n} x_j^2 \right)^{1/2}.$$

5.1.2 Multi-indices and Derivations

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an n-tuple of non-negative integers. Such an n-tuple α is called a *multi-index*. We let

$$|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n,$$
$$\alpha! = \alpha_1! \alpha_2! \cdots \alpha_n!.$$

If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ are multi-indices, we define

$$\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n).$$

The notation $\alpha \leq \beta$ means that $\alpha_j \leq \beta_j$ for each $1 \leq j \leq n$. Then we let

$$\binom{\beta}{\alpha} = \binom{\beta_1}{\alpha_1}\binom{\beta_2}{\alpha_2}\cdots\binom{\beta_n}{\alpha_n}.$$

We use the shorthand

$$\partial_j = \frac{\partial}{\partial x_j},$$

$$D_j = \frac{1}{i}\frac{\partial}{\partial x_j} \qquad (i = \sqrt{-1})$$

for derivatives on $\mathbf{R}^n$. Higher-order derivatives are expressed by multi-indices as follows:

$$\partial^\alpha = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n},$$
$$D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n}.$$

Similarly, if $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, we write

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

5.2 Function Spaces

5.2.1 L^p Spaces

Let Ω be an open subset of $\mathbf{R}^n$. Two Lebesgue measurable functions f, g on Ω are said to be *equivalent* if they are equal almost everywhere in Ω with respect to the Lebesgue measure dx, that is, if $f(x) = g(x)$ for every x outside a set of Lebesgue measure zero. This is obviously an equivalence relation.

If $1 \leq p < \infty$, we let

$$L^p(\Omega) = \text{the space of equivalence classes of Lebesgue measurable functions } f(x) \text{ on } \Omega \text{ such that } |f(x)|^p \text{ is integrable on } \Omega.$$

The space $L^p(\Omega)$ is a Banach space with the norm

$$\|f\|_p = \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p}.$$

Furthermore, the space $L^2(\Omega)$ is a Hilbert space with the inner product

$$(f, g) = \int_\Omega f(x) \overline{g(x)} \, dx.$$

A Lebesgue measurable function $f(x)$ on Ω is said to be *essentially bounded* if there exists a constant $C > 0$ such that $|f(x)| \leq C$ almost everywhere (a.e.) in Ω. We define

$$\operatorname{ess\,sup}_{x \in \Omega} |f(x)| = \inf\{C : |f(x)| \leq C \quad \text{a.e. in } \Omega\}.$$

For $p = \infty$, we let

$$L^\infty(\Omega) = \text{the space of equivalence classes of essentially bounded,}$$
$$\text{Lebesgue measurable functions on } \Omega.$$

The space $L^\infty(\Omega)$ is a Banach space with the norm

$$\|f\|_\infty = \text{ess sup}_{x\in\Omega}|f(x)|.$$

If $1 < p < \infty$, we let

$$p' = \frac{p}{p-1},$$

so that $1 < p' < \infty$ and

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

The number p' is called the *exponent conjugate* to p.

We recall that the most basic inequality for L^p-functions is the following:

Theorem 5.1 (Hölder's inequality). *If $1 < p < \infty$ and $f \in L^p(\Omega)$, $g \in L^{p'}(\Omega)$, then the product $f(x)g(x)$ is in $L^1(\Omega)$ and we have the inequality*

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}. \tag{5.1}$$

It should be noted that inequality (5.1) holds true for the two cases $p = 1$, $p' = \infty$ and $p = \infty$, $p' = 1$. Inequality (5.1) in the case $p = p' = 2$ is referred to as *Schwarz's inequality*.

5.2.2 Convolutions

We give a general theorem about integral operators on a measure space [Fo2, Theorem 6.18]:

Theorem 5.2 (Schur's lemma). *Let $(X, \mathcal{M}, \mu)$ be a measure space. Assume that $K(x, y)$ is a measurable function on the product space $X \times X$ such that*

$$\sup_{x\in X} \int_X |K(x, y)|\, d\mu(y) \leq C$$

and

$$\sup_{y\in X} \int_X |K(x, y)|\, d\mu(x) \leq C$$

where C is a positive constant. If $f \in L^p(X)$ with $1 \le p \le \infty$, then the function $Tf(x)$, defined by the formula

$$Tf(x) = \int_X K(x, y) f(y)\, d\mu(y),$$

is well-defined for almost all $x \in X$, and is in $L^p(X)$.

Furthermore, we have the inequality

$$\|Tf\|_p \le C \|f\|_p.$$

Corollary 5.3 (Young's inequality). *If $f \in L^1(\mathbf{R}^n)$ and $g \in L^p(\mathbf{R}^n)$ with $1 \le p \le \infty$, then the convolution $(f * g)(x)$, defined by the formula*

$$(f * g)(x) = \int_{\mathbf{R}^n} f(x - y) g(y)\, dy,$$

is well-defined for almost all $x \in \mathbf{R}^n$, and is in $L^p(\mathbf{R}^n)$.

Furthermore, we have the inequality

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

5.2.3 Spaces of C^k Functions

Let Ω be an open subset of $\mathbf{R}^n$. We let

$$C(\Omega) = \text{the space of continuous functions on } \Omega.$$

If K is a compact subset of Ω, we define a seminorm p_K on $C(\Omega)$ by the formula

$$C(\Omega) \ni \varphi \longmapsto p_K(\varphi) = \sup_{x \in K} |\varphi(x)|.$$

We equip the space $C(\Omega)$ with the topology defined by the family $\{p_K\}$ of seminorms where K ranges over all compact subsets of Ω.

If k is a positive integer, we let

$$C^k(\Omega) = \text{the space of } C^k \text{ functions on } \Omega.$$

We define a seminorm $p_{K,k}$ on $C^k(\Omega)$ by the formula

$$C^k(\Omega) \ni \varphi \longmapsto p_{K,k}(\varphi) = \sup_{\substack{x \in K \\ |\alpha| \le k}} |\partial^\alpha \varphi(x)|. \tag{5.2}$$

We equip the space $C^k(\Omega)$ with the topology defined by the family $\{p_{K,k}\}$ of seminorms where K ranges over all compact subsets of Ω. This is the topology of uniform convergence on compact subsets of Ω of the functions and their derivatives of order $\leq k$.

We set

$$C^\infty(\Omega) = \bigcap_{k=1}^{\infty} C^k(\Omega),$$

and

$$C^0(\Omega) = C(\Omega).$$

Let m be a non-negative integer or $m = \infty$. Let $\{K_\ell\}$ be a sequence of compact subsets of Ω such that K_ℓ is contained in the interior of $K_{\ell+1}$ for each ℓ and that

$$\Omega = \bigcup_{\ell=1}^{\infty} K_\ell.$$

For example, we may take

$$K_\ell = \left\{ x \in \Omega : |x| \leq \ell, \ \ \mathrm{dist}(x, \partial\Omega) \geq \frac{1}{\ell} \right\}.$$

Such a sequence $\{K_\ell\}$ is called an exhaustive sequence of compact subsets of Ω. It is easy to see that the countable family

$$\{p_{K_\ell, j}\}_{\substack{\ell=1,2,\ldots \\ 0\leq j\leq m}}$$

of seminorms suffices to define the topology on $C^m(\Omega)$ and further that $C^m(\Omega)$ is complete. Hence the space $C^m(\Omega)$ is a Fréchet space.

Furthermore, we let

$$C(\overline{\Omega}) = \text{the space of functions in } C(\Omega) \text{ having continuous extensions to the closure } \overline{\Omega} \text{ of } \Omega.$$

If k is a positive integer, we let

$$C^k(\overline{\Omega}) = \text{the space of functions in } C^k(\Omega) \text{ all of whose derivatives of order} \leq k \text{ have continuous extensions to } \overline{\Omega}.$$

We set

$$C^{\infty}(\overline{\Omega}) = \bigcap_{k=1}^{\infty} C^k(\overline{\Omega}),$$

and

$$C^0(\overline{\Omega}) = C(\overline{\Omega}).$$

Let m be a non-negative integer or $m = \infty$. We equip the space $C^m(\overline{\Omega})$ with the topology defined by the family $\{p_{K,j}\}$ of seminorms where K ranges over all compact subsets of $\overline{\Omega}$ and $0 \leq j \leq m$.

Let $\{F_\ell\}$ be an increasing sequence of compact subsets of $\overline{\Omega}$ such that

$$\bigcup_{\ell=1}^{\infty} F_\ell = \overline{\Omega}.$$

For example, we may take

$$F_\ell = \{x \in \overline{\Omega} : |x| \leq \ell\}.$$

Such a sequence $\{F_\ell\}$ is called an exhaustive sequence of compact subsets of $\overline{\Omega}$. It is easy to see that the countable family

$$\{p_{F_\ell,j}\}_{\substack{\ell=1,2,\ldots \\ 0 \leq j \leq m}}$$

of seminorms suffices to define the topology on $C^m(\overline{\Omega})$ and further that $C^m(\overline{\Omega})$ is complete. Hence the space $C^m(\overline{\Omega})$ is a Fréchet space.

If Ω is bounded and $0 \leq m < \infty$, then the space $C^m(\overline{\Omega})$ is a Banach space with the norm

$$\|\varphi\|_{C^m(\overline{\Omega})} = \sup_{\substack{x \in \overline{\Omega} \\ |\alpha| \leq m}} |\partial^\alpha \varphi(x)|.$$

5.2.4 *The Space of Test Functions*

Let Ω be an open subset of $\mathbf{R}^n$ and let $u(x)$ be a continuous function on Ω. The *support* of u, denoted $\operatorname{supp} u$, is the closure in Ω of the set $\{x \in \Omega : u(x) \neq 0\}$. In other words, the support of u is the smallest closed subset of Ω outside of which u vanishes.

Let m be a non-negative integer or $m = \infty$. If K is a compact subset of Ω, we let

$$C_K^m(\Omega) = \text{the space of functions in } C^m(\Omega) \text{ with support in } K.$$

The space $C^m_K(\Omega)$ is a closed subspace of $C^m(\Omega)$. Furthermore, we let

$$C^m_0(\Omega) = \bigcup_{K\subset\Omega} C^m_K(\Omega),$$

where K ranges over all compact subsets of Ω, so that $C^m_0(\Omega)$ is the space of functions in $C^m(\Omega)$ with compact support in Ω. It should be emphasized that the space $C^m_0(\Omega)$ can be identified with the space of functions in $C^m_0(\mathbf{R}^n)$ with support in Ω. If $\{K_\ell\}$ is an exhaustive sequence of compact subsets of Ω, we equip the space $C^m_0(\Omega)$ with the *inductive limit topology* of the spaces $C^m_{K_\ell}(\Omega)$, that is, the strongest locally convex linear space topology such that each injection

$$C^m_{K_\ell}(\Omega) \longrightarrow C^m_0(\Omega)$$

is continuous. We can verify that this topology on $C^m_0(\Omega)$ is independent of the sequence $\{K_\ell\}$ used.

We list some basic properties of the topology on $C^m_0(\Omega)$:

(1) A sequence $\{\varphi_j\}$ in $C^m_0(\Omega)$ converges to an element φ in $C^m_0(\Omega)$ if and only if the functions φ_j and φ are supported in a *common* compact subset K of Ω and $\varphi_j \to \varphi$ in $C^m_K(\Omega)$.
(2) A subset of $C^m_0(\Omega)$ is bounded if and only if it is bounded in $C^m_K(\Omega)$ for some compact $K \subset \Omega$.
(3) A linear mapping from $C^m_0(\Omega)$ into a linear topological space is continuous if and only if its restriction to $C^m_K(\Omega)$ for every compact $K \subset \Omega$ is continuous.

The elements of $C^\infty_0(\Omega)$ are often called *test functions*.

5.2.5 Hölder Spaces

Let D be a subset of $\mathbf{R}^n$ and let $0 < \theta < 1$. A function φ defined on D is said to be *Hölder continuous* with exponent θ if the quantity

$$[\varphi]_{\theta;D} = \sup_{\substack{x,y\in D\\ x\neq y}} \frac{|\varphi(x) - \varphi(y)|}{|x-y|^\theta}$$

is finite. We say that φ is *locally Hölder continuous* with exponent θ if it is Hölder continuous with exponent θ on compact subsets of D. Hölder continuity may be viewed as a fractional differentiability.

Let Ω be an open subset of $\mathbf{R}^n$ and $0 < \theta < 1$. We let

$$C^\theta(\Omega) = \text{the space of functions in } C(\Omega) \text{ which are locally Hölder continuous with exponent } \theta \text{ on } \Omega.$$

If k is a positive integer, we let

$$C^{k+\theta}(\Omega) = \text{the space of functions in } C^k(\Omega) \text{ all of whose } k\text{-th order derivatives are locally Hölder continuous with exponent } \theta \text{ on } \Omega.$$

If K is a compact subset of Ω, we define a seminorm $q_{K,k}$ on $C^{k+\theta}(\Omega)$ by the formula

$$C^{k+\theta}(\Omega) \ni \varphi \longmapsto q_{K,k}(\varphi) = \sup_{\substack{x\in K\\|\alpha|\le k}} |\partial^\alpha \varphi(x)| + \sup_{|\alpha|=k} [\partial^\alpha \varphi]_{\theta;K}.$$

It is easy to see that the Hölder space $C^{k+\theta}(\Omega)$ is a Fréchet space.

Furthermore, we let

$$C^{\theta}(\overline{\Omega}) = \text{the space of functions in } C(\overline{\Omega}) \text{ which are Hölder continuous with exponent } \theta \text{ on } \overline{\Omega}.$$

If k is a positive integer, we let

$$C^{k+\theta}(\overline{\Omega}) = \text{the space of functions in } C^k(\overline{\Omega}) \text{ all of whose } k\text{-th order derivatives are Hölder continuous with exponent } \theta \text{ on } \overline{\Omega}.$$

Let m be a non-negative integer. We equip the space $C^{m+\theta}(\overline{\Omega})$ with the topology defined by the family $\{q_{K,k}\}$ of seminorms where K ranges over all compact subsets of $\overline{\Omega}$. It is easy to see that the Hölder space $C^{m+\theta}(\overline{\Omega})$ is a Fréchet space.

If Ω is bounded, then $C^{m+\theta}(\overline{\Omega})$ is a Banach space with the norm

$$\begin{aligned}\|\varphi\|_{C^{m+\theta}(\overline{\Omega})} &= \|\varphi\|_{C^m(\overline{\Omega})} + \sup_{|\alpha|=m} [\partial^\alpha \varphi]_{\theta;\overline{\Omega}} \\ &= \sup_{\substack{x\in\overline{\Omega}\\|\alpha|\le m}} |\partial^\alpha \varphi(x)| + \sup_{|\alpha|=m} [\partial^\alpha \varphi]_{\theta;\overline{\Omega}}.\end{aligned}$$

5.2.6 Friedrichs' Mollifiers

Let $\rho(x)$ be a non-negative, C^∞ function on $\mathbf{R}^n$ satisfying the following conditions (5.3):

$$\operatorname{supp}\rho = \{x \in \mathbf{R}^n : |x| \le 1\}. \tag{5.3a}$$

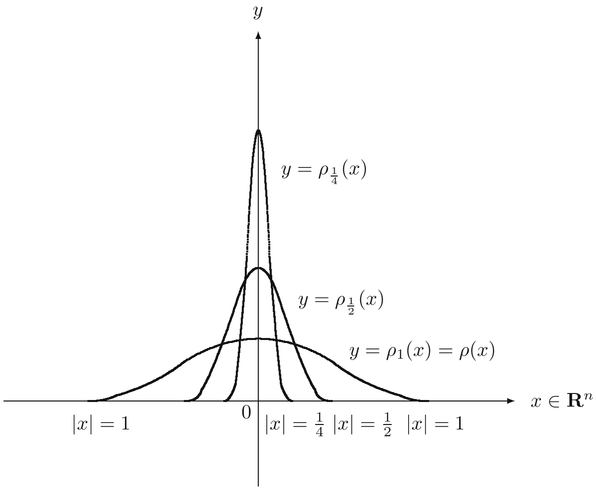


Fig. 5.1 Friedrichs' mollifiers $\{\rho_\varepsilon\}$

$$\int_{\mathbf{R}^n} \rho(x)\,dx = 1. \tag{5.3b}$$

For example, we may take

$$\rho(x) = \begin{cases} k \exp[-1/(1-|x|^2)] & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where the constant factor k is so chosen that condition (5.3b) is satisfied.

For each $\varepsilon > 0$, we define

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^n}\rho\left(\frac{x}{\varepsilon}\right),$$

then $\rho_\varepsilon(x)$ is a non-negative, C^∞ function on $\mathbf{R}^n$, and satisfies the conditions

$$\operatorname{supp} \rho_\varepsilon = \{x \in \mathbf{R}^n : |x| \leq \varepsilon\}; \tag{5.4a}$$

$$\int_{\mathbf{R}^n} \rho_\varepsilon(x)\,dx = 1. \tag{5.4b}$$

The functions $\{\rho_\varepsilon\}$ are called *Friedrichs' mollifiers* (see Fig. 5.1).

The next theorem shows how Friedrichs' mollifiers can be used to approximate a function by smooth functions:

Theorem 5.4. *Let Ω be an open subset of $\mathbf{R}^n$. Then we have the following two assertions (i) and (ii):*

(i) If $u \in L^p(\Omega)$ with $1 \leq p < \infty$ and vanishes outside a compact subset K of Ω, then it follows that $\rho_\varepsilon * u \in C_0^\infty(\Omega)$ provided that $\varepsilon < \operatorname{dist}(K, \partial\Omega)$, and further that $\rho_\varepsilon * u \to u$ in $L^p(\Omega)$ as $\varepsilon \downarrow 0$.
(ii) If $u \in C_0^m(\Omega)$ with $0 \leq m < \infty$, then it follows that $\rho_\varepsilon * u \in C_0^\infty(\Omega)$ provided that $\varepsilon < \operatorname{dist}(\operatorname{supp} u, \partial\Omega)$, and further that $\rho_\varepsilon * u \to u$ in $C_0^m(\Omega)$ as $\varepsilon \downarrow 0$.

Here

$$\operatorname{dist}(K, \partial\Omega) = \inf\{|x - y| : x \in K, \ y \in \partial\Omega\}.$$

The functions $\rho_\varepsilon * u$ are called *regularizations* of the function u.

Corollary 5.5. *The space $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ for each $1 \leq p < \infty$.*

Proof. Corollary 5.5 is an immediate consequence of part (i) of Theorem 5.4, since L^p functions with compact support are dense in $L^p(\Omega)$.

The next result gives another useful construction of smooth functions that vanish outside compact sets:

Corollary 5.6. *Let K be a compact subset of $\mathbf{R}^n$. If Ω is an open subset of $\mathbf{R}^n$ such that $K \subset \Omega$, then there exists a function $f \in C_0^\infty(\Omega)$ such that*

$$0 \leq f(x) \leq 1 \quad \text{in } \Omega,$$
$$f(x) = 1 \quad \text{on } K.$$

Proof. Let

$$\delta = \operatorname{dist}(K, \partial\Omega),$$

and define a relatively compact subset U of Ω, containing K, as follows:

$$U = \left\{x \in \Omega : |x - y| < \frac{\delta}{2} \text{ for some } y \in K\right\}.$$

Then it is easy to verify that the function

$$f(x) = \rho_\varepsilon * \chi_U(x) = \frac{1}{\varepsilon^n} \int_U \rho\left(\frac{x - y}{\varepsilon}\right) dy, \quad 0 < \varepsilon < \frac{\delta}{2},$$

satisfies all the conditions.

5.3 Differential Operators

Let Ω be an open subset of $\mathbf{R}^n$. If m is a non-negative integer, we let

$$P(x, D) = \sum_{|\alpha| \le m} a_\alpha(x) D^\alpha, \quad a_\alpha \in C^\infty(\Omega).$$

It is clear that $P(x, D)$ is a continuous linear mapping from $C^\infty(\Omega)$ into itself. Such mappings are called *differential operators* of order m on Ω. We see that $P = P(x, D)$ satisfies the condition

$$\operatorname{supp} Pu \subset \operatorname{supp} u \quad \text{for every } u \in C^\infty(\Omega), \tag{5.5}$$

since differentiation is a purely local process. We express this fact by saying that differential operators are *local*.

The next theorem states that the converse is also true (Peetre [Pe]):

Theorem 5.7 (Peetre). *Assume that P is a linear mapping from $C^\infty(\Omega)$ into itself which satisfies conditions* (5.4a) *and* (5.4b)*. Then, for every relatively compact subset Ω' of Ω, there exist a non-negative integer m and C^∞ functions $a_\alpha(x)$ in Ω such that*

$$P(x, D)u(x) = \sum_{|\alpha| \le m} a_\alpha(x) D^\alpha u(x), \quad u \in C^\infty(\Omega'), \ x \in \Omega'.$$

5.4 Distributions and the Fourier Transform

In this section we present a brief description of the basic concepts and results from the theory distributions.

5.4.1 Definitions and Basic Properties of Distributions

Let Ω be an open subset of $\mathbf{R}^n$. A *distribution* on Ω is a continuous linear functional on $C_0^\infty(\Omega)$. The space of distributions on Ω is denoted by $\mathcal{D}'(\Omega)$. In other words, the space $\mathcal{D}'(\Omega)$ is the dual space $\left(C_0^\infty(\Omega)\right)' = \mathcal{L}(C_0^\infty(\Omega), \mathbf{C})$. If $u \in \mathcal{D}'(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$, we denote the action of u on φ by $\langle u, \varphi \rangle$ or sometimes by $\langle \varphi, u \rangle$.

We give some useful characterizations of distributions:

Theorem 5.8. *Let u be a linear functional on $C_0^\infty(\Omega)$. Then the following three conditions (i), (ii) and (iii) are equivalent:*

(i) The functional u is a distribution.

(ii) For any compact subset K of Ω, there exist a constant $C > 0$ and a non-negative integer m such that

$$|\langle u, \varphi \rangle| \leq C p_{K,m}(\varphi) \quad \text{for all } \varphi \in C_K^\infty(\Omega),$$

where

$$p_{K,m}(\varphi) = \sup_{\substack{x \in K \\ |\alpha| \leq m}} |\partial^\alpha \varphi(x)|.$$

(iii) $\langle u, \varphi_j \rangle \to 0$ *whenever* $\varphi_j \to 0$ *in* $C_0^\infty(\Omega)$.

Part (ii) of Theorem 5.4 tells us that the space $C_0^\infty(\Omega)$ is a dense subspace of $C_0^m(\Omega)$ for $0 \leq m < \infty$. Also it is clear that the injection of $C_0^\infty(\Omega)$ into $C_0^m(\Omega)$ is continuous. Hence the dual space $\mathcal{D}^m(\Omega)' = \mathcal{L}(C_0^m(\Omega), \mathbf{C})$ can be identified with a linear subspace of $\mathcal{D}'(\Omega)$, by the identification of a continuous linear functional on $C_0^m(\Omega)$ with its restriction to $C_0^\infty(\Omega)$. The elements of $\mathcal{D}^m(\Omega)'$ are called distributions of *order* $\leq m$ on Ω. In other words, the distributions of order $\leq m$ on Ω are precisely those distributions on Ω that have continuous extensions to $C_0^m(\Omega)$.

Now we give some important examples of distributions.

Example 5.9. We let

$$L^1_{\mathrm{loc}}(\Omega) = \text{the space of equivalence classes of Lebesgue measurable functions on } \Omega \text{ which are integrable on every compact subset of } \Omega.$$

The elements of $L^1_{\mathrm{loc}}(\Omega)$ are called *locally integrable functions* on Ω. For example ($n = 1$), it is easy to verify the following two assertions (a) and (b):

(a) $\log|x| \in L^1_{\mathrm{loc}}(\mathbf{R})$.
(b) $Y(x) \in L^1_{\mathrm{loc}}(\mathbf{R})$.

Here $Y(x)$ is the Heaviside step function defined by the formula

$$Y(x) = \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0. \end{cases}$$

Every element f of $L^1_{\mathrm{loc}}(\Omega)$ defines a distribution T_f of order zero on Ω by the formula

$$\langle T_f, \varphi \rangle = \int_\Omega f(x)\varphi(x)\, dx \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$

Indeed, we have, for all $\varphi \in C_K^\infty(\Omega)$,

$$|\langle T_f, \varphi \rangle| \le \left(\int_K |f(x)|\, dx \right) p_{K,0}(\varphi).$$

Moreover, we can prove that the mapping

$$f \longmapsto T_f$$

induces an *injection* of $L^1_{\mathrm{loc}}(\Omega)$ into $\mathcal{D}'(\Omega)$. Indeed, we can prove the following lemma of Du Bois Raymond:

Lemma 5.10 (Du Bois Raymond). *Assume that $f \in L^1_{\mathrm{loc}}(\Omega)$ satisfies the condition*

$$\int_\Omega f(x)\varphi(x)\, dx = 0 \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{5.6}$$

Then it follows that

$$f(x) = 0 \quad \text{almost everywhere in } \Omega.$$

Proof. It suffices to show that we have, for any compact subset K of Ω,

$$f(x) = 0 \quad \text{almost everywhere in } K.$$

Now we take a function $\chi \in C_0^\infty(\Omega)$ such that $\chi(x) = 1$ on K, and let

$$f_\chi(x) = \chi(x) f(x), \quad x \in \mathbf{R}^n.$$

Then we see that

$$f_\chi \in L^1(\mathbf{R}^n).$$

Hence it follows from an application of part (i) of Theorem 5.4 with $p := 1$ that

$$\rho_\varepsilon * f_\chi \longrightarrow f_\chi \quad \text{in } L^1(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0. \tag{5.7}$$

However, we have

$$\rho_\varepsilon * f_\chi(x) = \int_{\mathbf{R}^n} \rho_\varepsilon(x-y) f_\chi(y)\, dy = \int_{\mathbf{R}^n} f(y) \left(\chi(y)\rho_\varepsilon(x-y)\right) dy,$$

and, for all sufficiently small $\varepsilon > 0$,

$$\chi(\cdot)\rho_\varepsilon(x - \cdot) \in C_0^\infty(\Omega) \quad \text{for all } x \in \mathbf{R}^n.$$

Therefore, by applying condition (5.6) to our situation we obtain that

$$\rho_\varepsilon * f_\chi(x) = 0 \quad \text{for all } x \in \mathbf{R}^n.$$

Hence we have, by assertion (5.7),

$$\|f_\chi\|_{L^1} = \lim_{\varepsilon \downarrow 0} \|\rho_\varepsilon * f_\chi\|_{L^1} = 0.$$

This proves that

$$f_\chi(x) = \chi(x) f(x) = 0 \quad \text{for almost all } x \in \mathbf{R}^n,$$

so that

$$f(x) = 0 \quad \text{for almost all } x \in K.$$

The proof of Lemma 5.10 is complete.

By virtue of Lemma 5.10, we can regard locally integrable functions as distributions. We say that such distributions "are" functions. In particular, the functions in $C^m(\Omega)$ $(0 \le m \le \infty)$ and in $L^p(\Omega)$ $(1 \le p \le \infty)$ are distributions on Ω.

Example 5.11. More generally, every complex Borel measure μ on Ω defines a distribution of order zero on Ω by the formula

$$\langle \mu, \varphi \rangle = \int_\Omega \varphi(x)\, d\mu(x) \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$

In particular, if we take μ to be the point mass at a point x_0 of Ω, we obtain the *Dirac measure* δ_{x_0} defined by the formula

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0) \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$

In other words, the Dirac measure δ_{x_0} is the point evaluation functional for $x_0 \in \Omega$. We denote δ_0 just by δ in the case $x = 0$.

Example 5.12. Let $f(x)$ be a continuous function on $\mathbf{R}^n \setminus \{0\}$ which is positively homogeneous of degree $-n$ and has mean zero on the unit sphere Σ_n:

$$f(\lambda x) = \lambda^{-n} f(x), \quad x \in \mathbf{R}^n,\ \lambda > 0, \tag{5.8a}$$

$$\int_{\Sigma_n} f(\sigma)\, d\sigma = 0. \tag{5.8b}$$

Here σ is the surface measure on Σ_n.

Then the formula

$$\langle \text{v. p.}\, f(x), \varphi \rangle = \lim_{\varepsilon \downarrow 0} \int_{|x|>\varepsilon} f(x)\varphi(x)\, dx, \quad \varphi \in C_0^\infty(\mathbf{R}^n),$$

defines a distribution on $\mathbf{R}^n$. Here "v.p." stands for Cauchy's "valeur principale" in French.

For example ($n = 1$), the distribution v. p.$(1/x)$ is defined by the formula

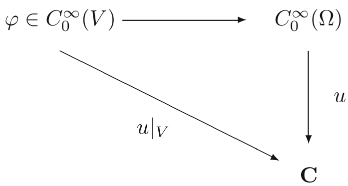
$$\begin{aligned}\left\langle \text{v. p.}\, \frac{1}{x}, \varphi \right\rangle &= \lim_{\varepsilon \downarrow 0} \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\, dx \\ &= \int_0^\infty \frac{\varphi(x) - \varphi(-x)}{x}\, dx \quad \text{for every } \varphi \in C_0^\infty(\mathbf{R}).\end{aligned}$$

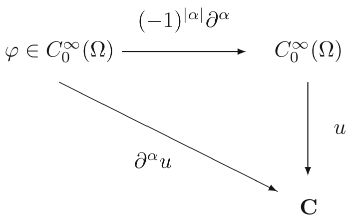
We define various operations on distributions.

(a) *Restriction*: If $u \in \mathcal{D}'(\Omega)$ and V is an open subset of Ω, we define the restriction $u|_V$ to V of u by the formula

$$\langle u|_V, \varphi \rangle = \langle u, \varphi \rangle \quad \text{for every } \varphi \in C_0^\infty(V).$$

Then it follows that $u|_V \in \mathcal{D}'(V)$.





(b) *Differentiation*: The derivative $\partial^\alpha u$ of a distribution $u \in \mathcal{D}'(\Omega)$ is the distribution on Ω defined by the formula

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$

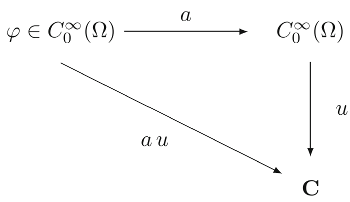
For example ($n = 1$), we have

(1) $Y(x)' = \delta(x)$.

(2) $(\log |x|)' = \text{v. p.}\, \dfrac{1}{x}$.

(c) *Multiplication by functions*: The product au of a function $a \in C^\infty(\Omega)$ and a distribution $u \in \mathcal{D}'(\Omega)$ is the distribution on Ω defined by the formula

$$\langle au, \varphi \rangle = \langle u, a\varphi \rangle \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$



For example ($n = 1$), we have

(1) $x\,\delta(x) = 0$.
(2) $x\left(\text{v. p.}\,\dfrac{1}{x}\right) = 1$.

The *Leibniz formula* for the differentiation of a product remains valid:

$$D^\beta(au) = \sum_{\alpha \le \beta} \binom{\beta}{\alpha} D^{\beta-\alpha} a \cdot D^\alpha u.$$

(d) We can combine operations (b) and (c). We let

$$P(x, D) = \sum_{|\alpha| \le m} a_\alpha(x) D^\alpha, \quad a_\alpha \in C^\infty(\Omega)$$

be a differential operator of order m on Ω. If $u \in \mathcal{D}'(\Omega)$, we define $P(x, D)u$ by the formula

$$\langle P(x, D)u, \varphi \rangle = \left\langle u, \sum_{|\alpha| \le m} (-1)^{|\alpha|} D^\alpha(a_\alpha \varphi) \right\rangle \quad \text{for every } \varphi \in C_0^\infty(\Omega).$$

Then it follows that $P(x, D)u \in \mathcal{D}'(\Omega)$.

(e) *Conjugation*: The conjugate $\overline{u}$ of a distribution $u \in \mathcal{D}'(\Omega)$ is the distribution on Ω defined by the formula

$$\langle \overline{u}, \varphi \rangle = \overline{\langle u, \overline{\varphi} \rangle} \quad \text{for every } \varphi \in C_0^\infty(\Omega),$$

where $\bar{\cdot}$ denotes complex conjugation.

5.4.2 Topologies on $\mathcal{D}'(\Omega)$

Let Ω be an open subset of $\mathbf{R}^n$. There are two natural topologies on the space $\mathcal{D}'(\Omega)$ of distributions on Ω:

(1) *Weak* topology* τ_s: This is the topology of convergence at each element of $C_0^\infty(\Omega)$. The space $\mathcal{D}'(\Omega)$ endowed with this topology is denoted by $\mathcal{D}'(\Omega)_s$. A sequence $\{u_j\}$ of distributions converges to a distribution u in $\mathcal{D}'(\Omega)_s$ if and only if the sequence $\{\langle u_j, \varphi\rangle\}$ converges to $\langle u, \varphi\rangle$ for every $\varphi \in C_0^\infty(\Omega)$.
(2) *Strong topology* τ_b: This is the topology of uniform convergence on all bounded subsets of $C_0^\infty(\Omega)$. The space $\mathcal{D}'(\Omega)$ endowed with this topology is denoted by $\mathcal{D}'(\Omega)_b$. A sequence $\{u_j\}$ of distributions converges to a distribution u in $\mathcal{D}'(\Omega)_b$ if and only if the sequence $\{\langle u_j, \varphi\rangle\}$ converges to $\langle u, \varphi\rangle$ uniformly in φ over all bounded subsets of $C_0^\infty(\Omega)$.

We list some basic topological properties of $\mathcal{D}'(\Omega)$:

(I) In the case of a sequence of distributions, the two notions of convergence coincide, that is, $u_j \to u$ in $\mathcal{D}'(\Omega)_s$ if and only if $u_j \to u$ in $\mathcal{D}'(\Omega)_b$. Let Ω_1 and Ω_2 be open subsets of $\mathbf{R}^{n_1}$ and $\mathbf{R}^{n_2}$, respectively and let A be a linear operator on $C_0^\infty(\Omega_2)$ into $\mathcal{D}'(\Omega_1)$. Then the continuity of A does not depend on the topology (τ_s or τ_b) on $\mathcal{D}'(\Omega_1)$. Indeed, $A : C_0^\infty(\Omega_2) \mapsto \mathcal{D}'(\Omega_1)$ is continuous if and only if its restriction to $C_{K_2}^\infty(\Omega_2)$ for every compact $K_2 \subset \Omega_2$ is continuous; so it suffices to base our reasoning on sequences.
(II) If $\{u_j\}$ is a sequence in $\mathcal{D}'(\Omega)$ and the limit

$$\langle u, \varphi\rangle = \lim_{j\to\infty} \langle u_j, \varphi\rangle$$

exists for every $\varphi \in C_0^\infty(\Omega)$, then it follows that $u \in \mathcal{D}'(\Omega)$. Thus we have $u_j \to u$ in $\mathcal{D}'(\Omega)_s$ and hence in $\mathcal{D}'(\Omega)_b$. This is one of the important consequences of the Banach–Steinhaus theorem.
(III) The strong dual space of $\mathcal{D}'(\Omega)_b$ can be identified with $C_0^\infty(\Omega)$. This fact is referred to as the *reflexivity* of $C_0^\infty(\Omega)$.

5.4.3 The Support of a Distribution

Let Ω be an open subset of $\mathbf{R}^n$. Two distributions u_1 and u_2 on Ω are said to be *equal* in an open subset V of Ω if the restrictions $u_1|_V$ and $u_2|_V$ are equal. In particular, we have $u = 0$ in V if and only if $\langle u, \varphi\rangle = 0$ for all $\varphi \in C_0^\infty(V)$.

The local behavior of a distribution determines it completely. More precisely, we have the following theorem:

Theorem 5.13. *The space $\mathcal{D}'(\Omega)$ has the* sheaf property*; this means the following two properties (S1) and (S2) hold:*

(S1) *If* $\{U_\lambda\}_{\lambda\in\Lambda}$ *is an open covering of* Ω *and if a distribution* $u \in \mathcal{D}'(\Omega)$ *is zero in each* U_λ, *then* $u = 0$ *in* Ω.

(S2) *Given an open covering* $\{U_\lambda\}_{\lambda\in\Lambda}$ *of* Ω *and a family of distributions* $u_\lambda \in \mathcal{D}'(U_\lambda)$ *such that* $u_j = u_k$ *in every* $U_\lambda \cap U_\mu$, *there exists a distribution* $u \in \mathcal{D}'(\Omega)$ *such that* $u = u_\lambda$ *in each* U_λ.

Proof. Let $\{\varphi_\lambda\}_{\lambda\in\Lambda}$ be a partition of unity subordinate to the open covering $\{U_\lambda\}_{\lambda\in\Lambda}$ of Ω (see Sect. 5.7.2). That is, the family $\{\varphi_\lambda\}_{\lambda\in\Lambda}$ in $C^\infty(\Omega)$ satisfies the following three conditions (a), (b) and (c):

(a) $0 \leq \varphi_\lambda(x) \leq 1$ for all $x \in \Omega$ and $\lambda \in \Lambda$.
(b) $\operatorname{supp}\varphi_\lambda \subset U_\lambda$ for each $\lambda \in \Lambda$.
(c) The collection $\{\operatorname{supp}\varphi_\lambda\}_{\lambda\in\Lambda}$ is locally finite and

$$\sum_{\lambda\in\Lambda} \varphi_\lambda(x) = 1 \quad \text{for every } x \in \Omega.$$

Here $\operatorname{supp}\varphi_\lambda$ is the *support* of φ_λ, i.e. the closure in Ω of the set $\{x \in \Omega : \varphi_\lambda(x) \neq 0\}$.

(1) For any given $\varphi \in C_0^\infty(\Omega)$, it follows that

$$\begin{cases} \varphi = \sum_{\lambda\in\Lambda} \varphi_\lambda\,\varphi, \\ \varphi_\lambda\,\varphi \in C_0^\infty(U_\lambda) \quad \text{for each } \lambda \in \Lambda. \end{cases}$$

Here it should be emphasized that the summation $\sum_{\lambda\in\Lambda}$ is finite, since $\operatorname{supp}\varphi$ is compact. Therefore, since u is zero in each U_λ, we have

$$\langle u, \varphi\rangle = \left\langle u, \sum_{\lambda\in\Lambda} \varphi_\lambda\,\varphi \right\rangle = \sum_{\lambda\in\Lambda} \langle u, \varphi_\lambda\,\varphi\rangle = 0.$$

This proves that $u = 0$ in Ω.

(2) If we define a distribution $u \in \mathcal{D}'(\Omega)$ by the formula

$$\langle u, \varphi\rangle = \sum_{\lambda\in\Lambda} \langle u_\lambda, \varphi_\lambda\,\varphi\rangle \quad \text{for every } \varphi \in C_0^\infty(\Omega),$$

then it is easy to verify that $u = u_\lambda$ in each U_λ.

The proof of Theorem 5.13 is complete.

If $u \in \mathcal{D}'(\Omega)$, the *support* of u is the smallest closed subset of Ω outside of which u is zero. The support of u is denoted by $\operatorname{supp} u$. We observe that if $\varphi \in C_0^\infty(\Omega)$ such that $\operatorname{supp}\varphi \cap \operatorname{supp} u = \emptyset$, then we have $\langle u, \varphi\rangle = 0$. It should be emphasized that the present definition of support coincides with the previous one if u is a continuous function on Ω.

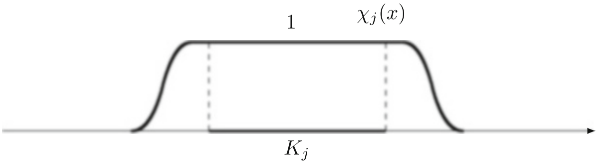


Fig. 5.2 The approximate function $\chi_j(x)$

Example 5.14. In the case where $n = 1$, it is easy to verify the following three assertions (1), (2) and (3):

(1) $\operatorname{supp} \delta_{x_0} = \{x_0\}$.
(2) $\operatorname{supp} Y(x) = [0, \infty)$.
(3) $\operatorname{supp} \text{v. p.} \frac{1}{x} = (-\infty, \infty)$.

5.4.4 The Dual Space of $C^\infty(\Omega)$

Let Ω be an open subset of $\mathbf{R}^n$. The injection of $C_0^\infty(\Omega)$ into $C^\infty(\Omega)$ is continuous and the space $C_0^\infty(\Omega)$ is a dense subspace of $C^\infty(\Omega)$.
Indeed, if $\{K_j\}$ is an exhaustive sequence of compact subsets of Ω, by using Corollary 5.6 we can construct a sequence $\{\chi_j\}$ of functions in $C_0^\infty(\Omega)$ such that $\chi_j(x) = 1$ on K_j (see Fig. 5.2). For any given function $\varphi \in C^\infty(\Omega)$, it is easy to verify that

$$\chi_j \varphi \longrightarrow \varphi \quad \text{in } C^\infty(\Omega) \text{ as } j \to \infty.$$

Hence the dual space $\mathcal{E}'(\Omega) = \mathcal{L}(C^\infty(\Omega), \mathbf{C})$ can be identified with a linear subspace of $\mathcal{D}'(\Omega)$, by the identification of a continuous linear functional on $C^\infty(\Omega)$ with its restriction to $C_0^\infty(\Omega)$. In other words, the elements of $\mathcal{E}'(\Omega)$ are precisely those distributions that have continuous extensions to $C^\infty(\Omega)$. The situation can be visualized in the following diagram:

$$\begin{array}{ccc} C^\infty(\Omega) & \longrightarrow & \mathcal{D}'(\Omega) = (C_0^\infty(\Omega))' \\ \uparrow & & \uparrow \\ C_0^\infty(\Omega) & \longrightarrow & \mathcal{E}'(\Omega) = (C^\infty(\Omega))' \end{array}$$

More precisely, we have the following theorem:

Theorem 5.15. *(i) The dual space $\mathcal{E}'(\Omega)$ of $C^\infty(\Omega)$ consists of those elements of $\mathcal{D}'(\Omega)$ with compact support.*
(ii) The dual space $\mathcal{E}'^m(\Omega)$ of $C^m(\Omega)$, $0 \leq m < \infty$, consists of those elements of $\mathcal{D}'^m(\Omega)$ with compact support, and $\mathcal{E}'(\Omega) = \cup_{m=0}^\infty \mathcal{E}'^m(\Omega)$.

As in the case of $\mathcal{D}'(\Omega)$, we equip the space $\mathcal{E}'(\Omega)$ with two natural topologies τ_s and τ_b, and denote $(\mathcal{E}'(\Omega), \tau_s)$ and $(\mathcal{E}'(\Omega), \tau_b)$ by $\mathcal{E}'(\Omega)_s$ and $\mathcal{E}'(\Omega)_b$ respectively. We have the same topological properties of $\mathcal{E}'(\Omega)$ as those of $\mathcal{D}'(\Omega)$.

5.4.5 Tensor Products of Distributions

Let X and Y be open subsets of $\mathbf{R}^n$ and $\mathbf{R}^p$, respectively. If $\varphi \in C_0^\infty(X)$ and $\psi \in C_0^\infty(Y)$, we define the *tensor product* $\varphi \otimes \psi$ of φ and ψ by the formula

$$(\varphi \otimes \psi)(x, y) = \varphi(x)\, \psi(y).$$

It is clear that $\varphi \otimes \psi \in C_0^\infty(X \times Y)$. We let

$$\begin{aligned} C_0^\infty(X) \otimes C_0^\infty(Y) = {} & \text{the space of finite combinations of the form } \varphi \otimes \psi \\ & \text{where } \varphi \in C_0^\infty(X) \text{ and } \psi \in C_0^\infty(Y). \end{aligned}$$

The space $C_0^\infty(X) \otimes C_0^\infty(Y)$ is a linear subspace of $C_0^\infty(X \times Y)$. Furthermore, it is sequentially dense in $C_0^\infty(X \times Y)$; that is, for every $\Phi \in C_0^\infty(X \times Y)$, there exists a sequence $\{\Phi_j\}$ in $C_0^\infty(X) \otimes C_0^\infty(Y)$ such that $\Phi_j \to \Phi$ in $C_0^\infty(X \times Y)$.

The next lemma asserts that the space $C_0^\infty(X) \otimes C_0^\infty(Y)$ is sequentially dense in $C_0^\infty(X \times Y)$:

Lemma 5.16. *The space $C_0^\infty(X) \otimes C_0^\infty(Y)$ is* sequentially dense *in $C_0^\infty(X \times Y)$. That is, for every function $\Phi \in C_0^\infty(X \times Y)$ there exists a sequence $\{\Phi_j\}$ in $C_0^\infty(X) \otimes C_0^\infty(Y)$ such that $\Phi_j \to \Phi$ in $C_0^\infty(X \times Y)$.*

Proof. Let $\Phi(x, y)$ be an arbitrary function in $C_0^\infty(X \times Y)$. We choose a closed cube K of side length T such that $\operatorname{supp} \Phi$ is contained in the interior of K (see Fig. 5.3), and we extend the function Φ to the periodic function $\tilde{\Phi} \in C^\infty(\mathbf{R}^n \times \mathbf{R}^p)$ with period T. Moreover, we choose two functions $\theta(x) \in C^\infty(\mathbf{R}^n)$ and $\zeta(y) \in C^\infty(\mathbf{R}^p)$ such that

$$\begin{aligned} & \operatorname{supp}\, (\theta \otimes \zeta) \subset K, \\ & \theta(x) \otimes \zeta(y) = 1 \quad \text{on } \operatorname{supp} \Phi. \end{aligned}$$

Then we have

$$\Phi(x, y) = (\theta(x) \otimes \zeta(y))\, \tilde{\Phi}(x, y), \tag{5.9}$$

and the Fourier expansion of $\tilde{\Phi}$

$$\tilde{\Phi}(x, y) = \sum_{\substack{\alpha \in \mathbf{N}^n \\ \beta \in \mathbf{N}^p}} c_{\alpha,\beta}\, e^{\frac{2\pi i}{T} \alpha \cdot x}\, e^{\frac{2\pi i}{T} \beta \cdot y}, \quad (x, y) \in \mathbf{R}^n \times \mathbf{R}^p. \tag{5.10}$$

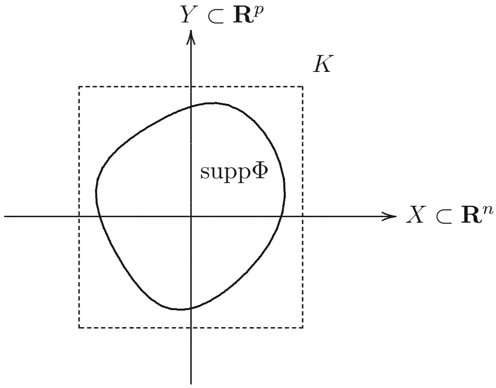


Fig. 5.3 The interior of the closed cube K

Here the Fourier coefficients $c_{\alpha,\beta}$ of $\tilde{\Phi}$ are given by the formula

$$c_{\alpha,\beta} = \frac{1}{T^{n+p}} \int_{\mathbf{R}^n} \int_{\mathbf{R}^p} e^{-\frac{2\pi i}{T}\alpha\cdot x}\, e^{-\frac{2\pi i}{T}\beta\cdot y}\, \Phi(x, y)\, dx\, dy.$$

However, we have, by integration by parts,

$$\begin{aligned} &\alpha^\gamma\, \beta^\delta\, c_{\alpha,\beta} \\ &= \frac{1}{T^{n+p}} \left(\frac{T}{2\pi}\right)^{|\gamma+\delta|} \int_{\mathbf{R}^n} \int_{\mathbf{R}^p} e^{-\frac{2\pi i}{T}\alpha\cdot x}\, e^{-\frac{2\pi i}{T}\beta\cdot y}\, D_x^\gamma\, D_y^\delta \Phi(x, y)\, dx\, dy \\ &= \frac{1}{T^{n+p}} \left(\frac{T}{2\pi}\right)^{|\gamma+\delta|} \int_X \int_Y e^{-\frac{2\pi i}{T}\alpha\cdot x}\, e^{-\frac{2\pi i}{T}\beta\cdot y}\, D_x^\gamma\, D_y^\delta \Phi(x, y)\, dx\, dy \end{aligned}$$

for all multi-indices γ and δ.

Hence, for any positive integer N we can find a positive constant C_N such that

$$(1 + |\alpha| + |\beta|)^N \left|c_{\alpha,\beta}\right| \le C_N \quad \text{for all } (\alpha, \beta) \in \mathbf{N}^n \times \mathbf{N}^p. \tag{5.11}$$

Therefore, we obtain from formulas (5.9), (5.10) and inequality (5.11) that the series

$$\begin{aligned} \Phi(x, y) &= (\theta(x) \otimes \zeta(y))\, \tilde{\Phi}(x, y) \\ &= \sum_{\substack{\alpha\in\mathbf{N}^n \\ \beta\in\mathbf{N}^p}} c_{\alpha,\beta} \left(\theta(x)\, e^{\frac{2\pi i}{T}\alpha\cdot x}\right) \left(\zeta(y)\, e^{\frac{2\pi i}{T}\beta\cdot y}\right) \end{aligned}$$

converges in the space $C_0^\infty(X \times Y)$.

Now the proof of Lemma 5.16 is complete.

The sequential density of $C_0^\infty(X) \otimes C_0^\infty(Y)$ in $C_0^\infty(X \times Y)$ allows us to obtain the following theorem:

Theorem 5.17. *If $u \in \mathcal{D}'(X)$ and $v \in \mathcal{D}'(Y)$, there exists a unique distribution $u \otimes v \in \mathcal{D}'(X \times Y)$ such that*

$$\langle u \otimes v, \Phi \rangle = \langle u, \varphi \rangle = \langle v, \psi \rangle \quad \textit{for all } \Phi \in C_0^\infty(X \times Y),$$

where $\varphi(x_1) = \langle v, \Phi(x, \cdot) \rangle$ and $\psi(y) = \langle u, \Phi(\cdot, y) \rangle$.

The distribution $u \otimes v$ is called the *tensor product* of u and v.

We list some basic properties of the tensor product:

(1) $\langle u \otimes v, \varphi \otimes \psi \rangle = \langle u, \varphi \rangle \langle v, \psi \rangle$ for all $\varphi \in C_0^\infty(X)$ and $\psi \in C_0^\infty(Y)$.
(2) $\operatorname{supp}(u \otimes v) = \operatorname{supp} u \times \operatorname{supp} v$.
(3) $D_x^\alpha D_y^\beta (u \otimes v) = D_x^\alpha u \otimes D_y^\beta v$.

5.4.6 Convolutions of Distributions

The Young inequality (Corollary 5.3) tells us that if $u \in L^1(\mathbf{R}^n)$ and $v \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$, then the convolution

$$(u * v)(x) = \int_{\mathbf{R}^n} u(x - y) v(y)\, dy$$

is well-defined for almost all $x \in \mathbf{R}^n$, and is in $L^p(\mathbf{R}^n)$. Furthermore, it follows from Fubini's theorem [Fo2, Theorem 2.37] that

$$\langle u * v, \varphi \rangle = \int\int_{\mathbf{R}^n \times \mathbf{R}^n} u(x) v(y) \varphi(x + y)\, dx\, dy \quad \text{for all } \varphi \in C_0^\infty(\mathbf{R}^n).$$

We use this formula to extend the definition of convolution to the case of distributions.

Let $u, v \in \mathcal{D}'(\mathbf{R}^n)$ and assume that one of them has compact support. If $\varphi \in C_0^\infty(\mathbf{R}^n)$, then the support of the function

$$\tilde{\varphi} : (x, y) \longmapsto \varphi(x + y)$$

is contained in the strip

$$\{(x, y) \in \mathbf{R}^n \times \mathbf{R}^n : x + y \in \operatorname{supp} \varphi\}.$$

Thus it is easy to see that the intersection

$$\operatorname{supp}(u \otimes v) \cap \operatorname{supp} \tilde{\varphi}$$

is a compact subset of $\mathbf{R}^n \times \mathbf{R}^n$. We choose a function θ in $C_0^\infty(\mathbf{R}^n \times \mathbf{R}^n)$ such that $\theta = 1$ in a neighborhood of $\operatorname{supp}(u \otimes v) \cap \operatorname{supp} \tilde{\varphi}$, and define

$$\langle u \otimes v, \tilde{\varphi} \rangle = \langle u \otimes v, \theta \tilde{\varphi} \rangle .$$

Observe that $\langle u \otimes v, \theta \tilde{\varphi} \rangle$ is independent of the function θ chosen, and further that the mapping

$$C_0^\infty(\mathbf{R}^n) \ni \varphi \longmapsto \langle u \otimes v, \tilde{\varphi} \rangle$$

is continuous. This discussion justifies the following definition:

Definition 5.18. The *convolution* $u * v$ of two distributions u and v in $\mathcal{D}'(\mathbf{R}^n)$, one of which has compact support, is a distribution on $\mathbf{R}^n$ defined by the formula

$$\langle u * v, \varphi \rangle = \langle u \otimes v, \tilde{\varphi} \rangle \quad \text{for every } \varphi \in C_0^\infty(\mathbf{R}^n).$$

We state some basic facts concerning the convolution product:

(1) $u * v = v * u$.
(2) $\operatorname{supp}(u * v) \subset \operatorname{supp} u + \operatorname{supp} v = \{x + y : x \in \operatorname{supp} u,\ y \in \operatorname{supp} v\}$.
(3) $D^\alpha(u * v) = (D^\alpha u) * v = u * (D^\alpha v)$.
(4) If either $u \in \mathcal{D}'(\mathbf{R}^n)$, $v \in C_0^\infty(\mathbf{R}^n)$ or $u \in \mathcal{E}'(\mathbf{R}^n)$, $v \in C^\infty(\mathbf{R}^n)$, then we have

$$u * v \in C^\infty(\mathbf{R}^n),$$
$$(u * v)(x) = \big\langle u_y, v(x - y) \big\rangle ,$$

where u_y means that the distribution u operates on $v(x - y)$ as a function of y with x fixed.
(5) Let $\rho(x)$ be a non-negative, C^∞ function on $\mathbf{R}^n$ such that

$$\rho(-x) = \rho(x) \quad \text{for all } x \in \mathbf{R}^n,$$
$$\operatorname{supp} \rho = \{x \in \mathbf{R}^n : |x| \leq 1\},$$
$$\int_{\mathbf{R}^n} \rho(x)\, dx = 1,$$

and define a function $\rho_\varepsilon(x)$ by the formula (see Fig. 5.1)

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^n} \rho\left(\frac{x}{\varepsilon}\right) \quad \text{for every } \varepsilon > 0.$$

If $u \in \mathcal{D}'(\mathbf{R}^n)$ (resp. $u \in \mathcal{E}'(\mathbf{R}^n)$), then it follows that the convolutions $u * \rho_\varepsilon$ are in $C^\infty(\mathbf{R}^n)$ and further that we have, for every $\varphi \in C_0^\infty(\mathbf{R}^n)$ (resp. $\varphi \in C^\infty(\mathbf{R}^n)$),

$$\langle u * \rho_\varepsilon, \varphi \rangle = \langle u, \rho_\varepsilon * \varphi \rangle \longrightarrow \langle u, \varphi \rangle \quad \text{as } \varepsilon \downarrow 0.$$

This proves that

$$u * \rho_\varepsilon \longrightarrow u \quad \text{in } \mathcal{D}'(\mathbf{R}^n) \text{ (resp. in } \mathcal{E}'(\mathbf{R}^n)) \text{ as } \varepsilon \downarrow 0.$$

Rephrased, distributions can be approximated in the weak* topology of distributions by smooth functions. The functions $u * \rho_\varepsilon$ are called *regularizations* of the distribution u.

5.4.7 The Jump Formula

If $x = (x_1, x_2, \ldots, x_n)$ is a point of $\mathbf{R}^n$, we write

$$x = (x', x_n), \quad x' = (x_1, x_2, \ldots, x_{n-1}).$$

If $u \in C^\infty(\overline{\mathbf{R}^n_+})$, we define its extension u^0 to the whole space $\mathbf{R}^n$ by the formula

$$u^0(x', x_n) = \begin{cases} u(x', x_n) & \text{for } x_n \geq 0, \\ 0 & \text{for } x_n < 0. \end{cases}$$

Then it follows that u^0 is a distribution on $\mathbf{R}^n$ and further that its j-th derivative $\partial_n^j(u^0)$ with respect to the normal variable x_n is expressed as follows:

$$\frac{\partial^j(u^0)}{\partial x_n^j} = \left(\frac{\partial^j u}{\partial x_n^j}\right)^0 + \sum_{k=0}^{j-1} \gamma_{j-k-1} u \otimes \delta^{(k)}(x_n),$$

where $\gamma_k u$ is a C^∞ function on $\mathbf{R}^{n-1}_{x'}$ defined by the formula

$$(\gamma_k u)(x') = \frac{\partial^k u}{\partial x_n^k}(x', 0) \quad \text{for every } x' \in \mathbf{R}^{n-1},$$

and $\delta(x_n)$ is the Dirac measure at 0 on $\mathbf{R}_{x_n}$.

Furthermore, if Δ is the usual Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2},$$

then we have the following formula (5.12):

$$\Delta(u^0) = (\Delta u)^0 + \gamma_1 u \otimes \delta(x_n) + \gamma_0 u \otimes \delta^{(\prime)}(x_n) \quad (5.12)$$
$$= (\Delta u)^0 + \frac{\partial u}{\partial x_n}(x', 0) \otimes \delta(x_n) + u(x', 0) \otimes \delta^{(\prime)}(x_n).$$

More generally, if the operator

$$P(x, D_x) = \sum_{j=0}^{m} P_j(x, D_{x'}) D_n^j$$

is a differential operator of order m with C^∞ coefficients on $\mathbf{R}^n$, then we have

$$P(u^0) = (Pu)^0 + \frac{1}{\sqrt{-1}} \sum_{\ell+k+1 \leq m} P_{\ell+k+1}(x, D_{x'}) \gamma_\ell u \otimes D_n^k \delta(x_n). \quad (5.13)$$

Here $P_j(x, D_{x'})$ is a differential operator of order $m - j$ with respect to x'. Formula (5.13) will be referred to as the *jump formula*.

5.4.8 Regular Distributions with Respect to One Variable

If $x = (x_1, x_2, \ldots, x_n)$ is the variable in $\mathbf{R}^n$, we write

$$x = (x', x_n), \quad x' = (x_1, x_2, \ldots, x_{n-1}),$$

so x' is the variable in $\mathbf{R}^{n-1}$.

A function $U(x_n)$ defined on $\mathbf{R}$ with values in $\mathcal{D}'(\mathbf{R}^{n-1}_{x'})$ is said to be *continuous* if, for every $\phi \in C_0^\infty(\mathbf{R}^{n-1}_{x'})$, the function $\langle U(x_n), \phi \rangle$ is continuous on $\mathbf{R}$.

We let

$$C(\mathbf{R}; \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$$
$$= \text{the space of } \mathcal{D}'(\mathbf{R}^{n-1}_{x'})\text{-valued continuous functions on } \mathbf{R}.$$

If $U \in C(\mathbf{R}; \mathcal{D}'(\mathbf{R}^{n-1}))$, we can associate injectively a distribution $u \in \mathcal{D}'(\mathbf{R}^n)$ by the formula

$$\langle u, \varphi \rangle = \int_{\mathbf{R}} \langle U(x_n), \varphi(\cdot, x_n) \rangle \, dx_n \quad \text{for every } \varphi \in C_0^\infty(\mathbf{R}^n).$$

Such a distribution u is said to be *continuous with respect to* x_n with values in $\mathcal{D}'(\mathbf{R}^{n-1}_{x'})$. We let

$$\gamma_0 u = U(0) \in \mathcal{D}'(\mathbf{R}^{n-1}_{x'}).$$

The distribution $\gamma_0 u$ is called the *sectional trace of order zero* on the hyperplane $\{x_n = 0\}$ of u.

Let k be a positive integer. A function $U(x_n)$, defined on $\mathbf{R}$ with values in $\mathcal{D}'(\mathbf{R}^{n-1}_{x'})$, is said to be *of class* C^k if, for every $\phi \in C_0^\infty(\mathbf{R}^{n-1}_{x'})$, the function $\langle U(x_n), \phi \rangle$ is of class C^k on $\mathbf{R}$.

We let

$$C^k(\mathbf{R}; \mathcal{D}'(\mathbf{R}^{n-1}_{x'})) = \text{the space of } \mathcal{D}'(\mathbf{R}^{n-1}_{x'})\text{-valued } C^k \text{ functions on } \mathbf{R}.$$

If $U \in C^k(\mathbf{R}; \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$, we have, for $0 \leq j \leq k$,

$$\left\langle \partial_n^j u, \varphi \right\rangle = \int_{\mathbf{R}} \left\langle U^{(j)}(x_n), \varphi(\cdot, x_n) \right\rangle dx_n, \quad \varphi \in C_0^\infty(\mathbf{R}^n).$$

This shows that the distribution $\partial_n^j u$ on $\mathbf{R}^n$ is the distribution associated with $U^{(j)} \in C(\mathbf{R}; \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$. We say that u is *of class* C^k *with respect to* x_n with values in $\mathcal{D}'(\mathbf{R}^{n-1}_{x'})$. We define the *sectional trace* $\gamma_j u$ of order j on the hyperplane $\{x_n = 0\}$ of u by the formula

$$\gamma_j u = D_n^j U(0) \in \mathcal{D}'(\mathbf{R}^{n-1}), \quad 0 \leq j \leq k.$$

We make no distinction between U and u for notational convenience.

It is obvious what we mean by $C^m([0, \infty); \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$, $0 \leq m \leq \infty$. If $u \in C([0, \infty); \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$, we define a distribution $u^0 \in \mathcal{D}'(\mathbf{R}^n)$ by the formula

$$\left\langle u^0, \varphi \right\rangle = \int_0^\infty \langle u(x_n), \varphi(\cdot, x_n) \rangle \, dx_n \quad \text{for every } \varphi \in C_0^\infty(\mathbf{R}^n).$$

The distribution u^0 is an extension to the whole space $\mathbf{R}^n$ of u which is equal to zero for $x_n < 0$.

If $u \in C^m([0, \infty); \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$, we define its sectional traces $\gamma_j u$, $0 \leq j \leq m$, on the hyperplane $\{x_n = 0\}$ by the formula

$$\gamma_j u = \lim_{x_n \downarrow 0} D_n^j u(\cdot, x_n) \quad \text{in } \mathcal{D}'(\mathbf{R}^{n-1}).$$

Then it is easy to verify that the jump formula (5.13) can be extended to the space $C^m([0, \infty); \mathcal{D}'(\mathbf{R}^{n-1}_{x'}))$.

5.4.9 The Fourier Transform

If $f \in L^1(\mathbf{R}^n)$, we define its (direct) *Fourier transform* $\hat{f}$ by the formula

$$\hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-ix\cdot\xi} f(x)\, dx, \quad \xi = (\xi_1, \xi_2, \ldots, \xi_n), \tag{5.14}$$

where $x \cdot \xi = x_1\xi_1 + x_2\xi_2 + \ldots + x_n\xi_n$. It follows from an application of the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] that the function $\hat{f}(\xi)$ is continuous on $\mathbf{R}^n$, and further we have the inequality

$$\|\hat{f}\|_\infty = \sup_{\mathbf{R}^n} |\hat{f}(\xi)| \le \|f\|_1.$$

We also denote $\hat{f}$ by $\mathcal{F} f$.

Example 5.19. If $f, g \in L^1(\mathbf{R}^n)$, then the Fourier transform $\widehat{f * g}$ of the convolution $f * g$ is given by the formula

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\, \hat{g}(\xi), \quad \xi \in \mathbf{R}^n.$$

Indeed, we have, by Fubini's theorem,

$$\begin{aligned}\widehat{f * g}(\xi) &= \int_{\mathbf{R}^n} e^{-ix\cdot\xi} \left(\int_{\mathbf{R}^n} f(x-y) g(y)\, dy \right) dx \\ &= \int_{\mathbf{R}^n} g(y) e^{-iy\cdot\xi}\, dy \cdot \int_{\mathbf{R}^n} f(x-y) e^{-i\xi\cdot(x-y)}\, dy \\ &= \hat{f}(\xi)\, \hat{g}(\xi).\end{aligned}$$

Similarly, if $g \in L^1(\mathbf{R}^n)$, we define the function $\check{g}(x)$ by the formula

$$\check{g}(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} g(\xi)\, d\xi.$$

The function $\check{g}(x)$ is called the *inverse Fourier transform* of g. We also denote $\check{g}$ by $\mathcal{F}^* g$.

Now we introduce a subspace of $L^1(\mathbf{R}^n)$ which is invariant under the Fourier transform. We let

$$\mathcal{S}(\mathbf{R}^n) = \text{the space of } C^\infty \text{ functions } \varphi(x) \text{ on } \mathbf{R}^n \text{ such that,}$$

$$\text{for any non-negative integer } j\text{, the quantity}$$

$$p_j(\varphi) = \sup_{\substack{x \in \mathbf{R}^n \\ |\alpha| \leq j}} \left\{ (1 + |x|^2)^{j/2} |\partial^\alpha \varphi(x)| \right\}$$

is finite.

The space $\mathcal{S}(\mathbf{R}^n)$ is called the *Schwartz space* or *space of C^∞ functions on* $\mathbf{R}^n$ *rapidly decreasing at infinity*. We equip the space $\mathcal{S}(\mathbf{R}^n)$ with the topology defined by the countable family $\{p_j\}$ of seminorms. It is easy to verify that $\mathcal{S}(\mathbf{R}^n)$ is complete; so it is a Fréchet space.

Now we give typical examples of functions in $\mathcal{S}(\mathbf{R}^n)$:

Example 5.20. (1) For every $a > 0$, it follows that

$$\varphi(x) = e^{-a|x|^2} \in \mathcal{S}(\mathbf{R}^n).$$

The Fourier transform $\hat{\varphi}(\xi)$ of $\varphi(x)$ is given by the formula

$$\hat{\varphi}(\xi) = \int_{\mathbf{R}^n} e^{-ix\cdot\xi} e^{-a|x|^2} \, dx = \left(\frac{\pi}{a}\right)^{n/2} e^{-\frac{|\xi|^2}{4a}}, \quad \xi \in \mathbf{R}^n. \tag{5.15}$$

(2) The Fourier transform $\widehat{K_t}(\xi)$ of the heat kernel

$$K_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}, \quad x \in \mathbf{R}^n, \ t > 0,$$

is given by the formula

$$\widehat{K_t}(\xi) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-ix\cdot\xi} e^{-\frac{|x|^2}{4t}} \, dx = e^{-t|\xi|^2}, \quad \xi \in \mathbf{R}^n, \ t > 0. \tag{5.16}$$

Proof. (1) We have only to prove (5.15) for $n = 1$, since we have

$$\exp\left[-a|x|^2\right] = \exp\left[-a\left(\sum_{j=1}^n x_j^2\right)\right] = \prod_{j=1}^n \exp\left[-ax_j^2\right].$$

The proof is divided into three steps.

Step 1: If $\xi = 0$, then formula (5.15) is reduced to the well known formula

$$\int_{-\infty}^{\infty} e^{-ax^2} \, dx = \sqrt{\frac{\pi}{a}}.$$

Step 2: Now we consider the case where $\xi < 0$. Since the function

$$\mathbf{C} \ni z \longmapsto e^{-az^2} e^{-iz\xi}$$

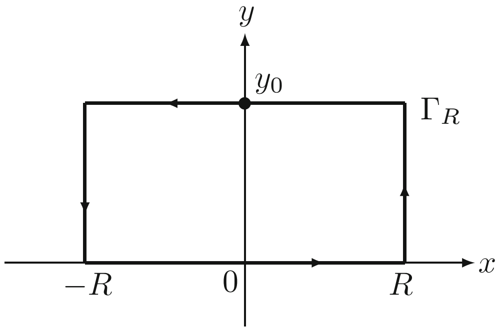


Fig. 5.4 The integral path Γ_R consisting of the rectangle

is an entire function of $z = x + iy$, it follows from an application of Cauchy's theorem that

$$\begin{aligned} 0 &= \int_{\Gamma_R} e^{-az^2} e^{-iz\xi}\, dz \qquad (5.17)\\ &= \int_{-R}^{R} e^{-ax^2-ix\xi}\, dx + \int_0^{y_0} e^{-a(R+iy)^2-i(R+iy)\xi} i\, dy \\ &\quad + \int_R^{-R} e^{-a(x+iy_0)^2} e^{-i(x+iy_0)\xi}\, dx + \int_{y_0}^{0} e^{-a(-R+iy)^2} e^{-i(-R+iy)\xi} i\, dy \\ &:= I + II + III + IV. \end{aligned}$$

Here Γ_R is a path consisting of the rectangle as in Fig. 5.4:

(a) Since we have, for $y\xi \leq 0$,

$$\left| e^{-a(\pm R+iy)^2-i(\pm R+iy)\xi} \right| = e^{-aR^2+ay^2+y\xi} \leq e^{-aR^2+ay^2},$$

we can estimate the second term II and the fourth term IV as follows:

$$\begin{aligned} |II|\,,\ |IV| &= \left| \int_0^{y_0} e^{-a(\pm R+iy)^2-i(\pm R+iy)\xi}\, dy \right| \\ &\leq e^{-aR^2} \int_0^{y_0} e^{-ay^2}\, dy \longrightarrow 0 \quad \text{as } R \to \infty. \end{aligned}$$

(b) In order to estimate the third term III as $R \to \infty$, we note that

$$\int_R^{-R} e^{-a(x+iy_0)^2} e^{-i(x+iy_0)\xi}\, dx = \int_R^{-R} e^{-ax^2+ay_0{}^2+y_0\xi-i(2ay_0+\xi)x}\, dx.$$

If we take

$$y_0 = -\frac{\xi}{2a},$$

then it follows that

$$a{y_0}^2 + y_0\xi = a\left(-\frac{\xi}{2a}\right)^2 + \left(-\frac{\xi}{2a}\right)\xi = -\frac{\xi^2}{4a}.$$

Hence the third term III can be estimated as follows:

$$\begin{aligned} III &= \int_R^{-R} e^{-ax^2+a{y_0}^2+y_0\xi-i(2ay_0+\xi)x}\,dx \\ &= e^{a{y_0}^2+y_0\xi}\int_R^{-R} e^{-ax^2}\,dx \\ &\longrightarrow e^{-\frac{\xi^2}{4a}}\int_\infty^{-\infty} e^{-ax^2}\,dx = -e^{-\frac{\xi^2}{4a}}\sqrt{\frac{\pi}{a}} \quad \text{as } R \to \infty. \end{aligned}$$

Therefore, by letting $R \to \infty$ in (5.17) we obtain the desired formula (5.15) for $\xi < 0$:

$$\int_{-\infty}^{\infty} e^{-ix\xi}\,e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}\,e^{-\frac{\xi^2}{4a}}.$$

Step 3: The case where $\xi > 0$ can be treated similarly.

(2) Formula (5.16) follows by applying (5.15) with $a := 1/(4t)$ to the heat kernel $K_t(x)$.

The proof of Example 5.20 is complete.

The next theorem summarizes the basic properties of the Fourier transform:

Theorem 5.21. *(i) The Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ map $\mathcal{S}(\mathbf{R}^n)$ continuously into itself. Furthermore, we have, for all multi-indices α and β,*

$$\begin{aligned} \widehat{D^\alpha\varphi}(\xi) &= \xi^\alpha\hat{\varphi}(\xi), \quad \varphi \in \mathcal{S}(\mathbf{R}^n), \\ D^\beta\hat{\varphi}(\xi) &= \widehat{(-x)^\beta\varphi}(\xi), \quad \varphi \in \mathcal{S}(\mathbf{R}^n). \end{aligned}$$

(ii) The Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ are isomorphisms of $\mathcal{S}(\mathbf{R}^n)$ onto itself; more precisely, $\mathcal{F}\mathcal{F}^* = \mathcal{F}^*\mathcal{F} = I$ on $\mathcal{S}(\mathbf{R}^n)$. In particular, we have

$$\varphi(x) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{ix\cdot\xi}\hat{\varphi}(\xi)\,d\xi \quad \textit{for every } \varphi \in \mathcal{S}(\mathbf{R}^n). \tag{5.18}$$

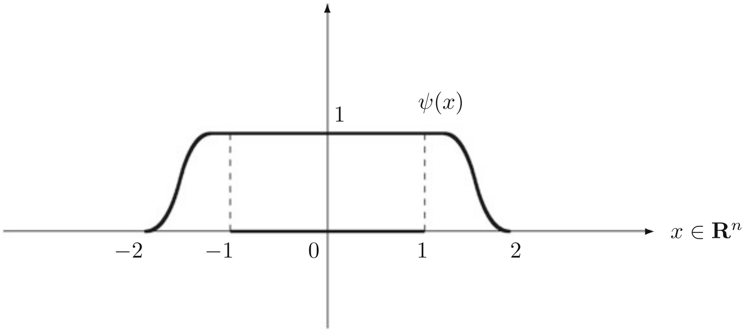


Fig. 5.5 The function $\psi(x)$

(iii) If φ, $\psi \in \mathcal{S}(\mathbf{R}^n)$*, we have*

$$\int_{\mathbf{R}^n} \varphi(x)\hat{\psi}(x)\,dx = \int_{\mathbf{R}^n} \hat{\varphi}(\xi)\psi(\xi)\,d\xi, \tag{5.19a}$$

$$\int_{\mathbf{R}^n} \varphi(x)\psi(x)\,dx = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} \varphi(\xi)\psi(\xi)\,d\xi. \tag{5.19b}$$

Formula (5.18) is called the *Fourier inversion formula* and formulas (5.19a) and (5.19b) are called the *Parseval formulas*.

5.4.10 Tempered Distributions

For the three spaces $C_0^\infty(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{R}^n)$ and $C^\infty(\mathbf{R}^n)$, we have the following two inclusions (i) and (ii):

(i) The injection of $C_0^\infty(\mathbf{R}^n)$ into $\mathcal{S}(\mathbf{R}^n)$ is continuous and the space $C_0^\infty(\mathbf{R}^n)$ is dense in $\mathcal{S}(\mathbf{R}^n)$.
(ii) The injection of $\mathcal{S}(\mathbf{R}^n)$ into $C^\infty(\mathbf{R}^n)$ is continuous and the space $\mathcal{S}(\mathbf{R}^n)$ is dense in $C^\infty(\mathbf{R}^n)$.

Indeed, we take a function $\psi \in C_0^\infty(\mathbf{R}^n)$ such that (see Fig. 5.5)

$$\psi(x) = \begin{cases} 1 & \text{if } |x| < 1, \\ 0 & \text{if } |x| > 2, \end{cases}$$

and let

$$\psi_j(x) = \psi\left(\frac{x}{j}\right) \quad \text{for every integer } j \geq 1.$$

For any given function $\varphi \in \mathcal{S}(\mathbf{R}^n)$ (resp. $\varphi \in C^\infty(\mathbf{R}^n)$), it is easy to verify that

$$\psi_j \, \varphi \longrightarrow \varphi \quad \text{in } \mathcal{S}(\mathbf{R}^n) \text{ (resp. in } C^\infty(\mathbf{R}^n)) \text{ as } j \to \infty.$$

Hence the dual space $\mathcal{S}'(\mathbf{R}^n) = \mathcal{L}(\mathcal{S}(\mathbf{R}^n), \mathbf{C})$ can be identified with a linear subspace of $\mathcal{D}'(\mathbf{R}^n)$ containing $\mathcal{E}'(\mathbf{R}^n)$, by the identification of a continuous linear functional on $\mathcal{S}(\mathbf{R}^n)$ with its restriction to $C_0^\infty(\mathbf{R}^n)$. Thus, we have the inclusions

$$\mathcal{E}'(\mathbf{R}^n) \subset \mathcal{S}'(\mathbf{R}^n) \subset \mathcal{D}'(\mathbf{R}^n).$$

The elements of $\mathcal{S}'(\mathbf{R}^n)$ are called *tempered distributions* on $\mathbf{R}^n$. In other words, the tempered distributions are precisely those distributions on $\mathbf{R}^n$ that have continuous extensions to $\mathcal{S}(\mathbf{R}^n)$.

Roughly speaking, the tempered distributions are those which grow at most polynomially at infinity, since the functions in $\mathcal{S}(\mathbf{R}^n)$ die out faster than any power of x at infinity. In fact, we have the following examples (1)–(4) of tempered distributions:

(1) The functions in $L^p(\mathbf{R}^n)$ $(1 \leq p \leq \infty)$ are tempered distributions.
(2) A locally integrable function on $\mathbf{R}^n$ is a tempered distribution if it grows at most polynomially at infinity.
(3) If $u \in \mathcal{S}'(\mathbf{R}^n)$ and $f(x)$ is a C^∞ function on $\mathbf{R}^n$ all of whose derivatives grow at most polynomially at infinity, then the product $f u$ is a tempered distribution.
(4) Any derivative of a tempered distribution is also a tempered distribution.

More precisely, we can prove the following structure theorem for tempered distributions:

Theorem 5.22 (the structure theorem). *Let $u \in \mathcal{S}'(\mathbf{R}^n)$. Then there exist a non-negative integer m and functions $\{f_\alpha\}_{|\alpha| \leq m}$ in $L^\infty(\mathbf{R}^n)$ such that*

$$u = \sum_{|\alpha| \leq m} \partial_{x_1} \cdots \partial_{x_n} \partial_x^\alpha \left[(1 + |x|^2)^{m/2} f_\alpha \right]. \tag{5.20}$$

Proof. The proof is divided into three steps.

Step 1: Since $u : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ is continuous, we can find a seminorm $p_m(\cdot)$ of $\mathcal{S}(\mathbf{R}^n)$ and a constant $\delta > 0$ such that

$$\psi \in \mathcal{S}(\mathbf{R}^n), \;\; p_m(\psi) < \delta \Longrightarrow |\langle u, \psi \rangle| < 1. \tag{5.21}$$

Here we recall that

$$p_m(\psi) = \sup_{\substack{x \in \mathbf{R}^n \\ |\alpha| \leq m}} \left\{ (1 + |x|^2)^{m/2} |\partial^\alpha \psi(x)| \right\}.$$

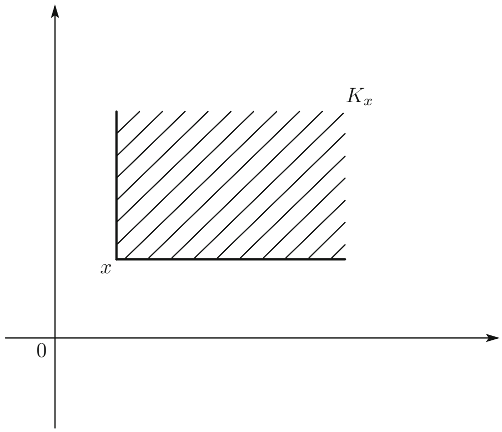


Fig. 5.6 The set K_x

Let φ be an arbitrary non-zero function in $\mathcal{S}(\mathbf{R}^n)$. By letting

$$\psi = \frac{\delta}{2} \frac{\varphi}{p_m(\varphi)},$$

we obtain from assertion (5.21) that

$$|\langle u, \varphi \rangle| < \frac{2}{\delta}\, p_m(\varphi).$$

This proves that

$$|\langle u, \varphi \rangle| \leq C\; p_m(\varphi) \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}^n), \tag{5.22}$$

where

$$C := \frac{2}{\delta}.$$

Step 2: For each point $x \in \mathbf{R}^n$, we let (see Fig. 5.6)

$$K_x = \{y \in \mathbf{R}^n : |x_j| \leq |y_j|,\; x_j\, y_j \geq 0\},$$

where we use the convention

$$x_j\, y_j \geq 0 \text{ and } x_j = 0 \implies y_j \geq 0.$$

Since every function $\varphi \in \mathcal{S}(\mathbf{R}^n)$ satisfies the estimate

$$|\partial_x^\alpha \varphi(x)| \leq \int_{K_x} \left| \partial_{y_1} \cdots \partial_{y_n} \partial_y^\alpha \varphi(y) \right| dy,$$

we have, for all $x \in \mathbf{R}^n$ and $|\alpha| \leq m$,

$$
\begin{aligned}
(1+|x|^2)^{m/2}|\partial_x^\alpha \varphi(x)| &\leq \int_{K_x} (1+|x|^2)^{m/2}|\partial_{y_1}\cdots\partial_{y_n}\partial_y^\alpha\varphi(y)|\,dy \\
&\leq \int_{K_x} (1+|y|^2)^{m/2}|\partial_{y_1}\cdots\partial_{y_n}\partial_y^\alpha\varphi(y)|\,dy \\
&\leq \int_{\mathbf{R}^n} (1+|y|^2)^{m/2}|\partial_{y_1}\cdots\partial_{y_n}\partial_y^\alpha\varphi(y)|\,dy.
\end{aligned}
$$

This proves that

$$
\begin{aligned}
p_m(\varphi) &\leq \sum_{|\alpha|\leq m} \int_{\mathbf{R}^n} \left(1+|x|^2\right)^{m/2} |\partial_{x_1}\cdots\partial_{x_n}\partial_x^\alpha\varphi(x)|\,dx \\
&= \sum_{|\alpha|\leq m} \|\psi_\alpha\|_{L^1(\mathbf{R}^n)},
\end{aligned}
\tag{5.23}
$$

where the function

$$
\psi_\alpha(x) := \left(1+|x|^2\right)^{m/2} \partial_{x_1}\cdots\partial_{x_n}\partial_x^\alpha\varphi(x), \quad |\alpha| \leq m,
$$

belongs to the space $L^1(\mathbf{R}^n)$.
By combining inequalities (5.22) and (5.23), we obtain that

$$
|\langle u, \varphi\rangle| \leq C \sum_{|\alpha|\leq m} \|\psi_\alpha\|_{L^1(\mathbf{R}^n)} \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}^n). \tag{5.24}
$$

Step 3: Now we let

$$
N = \text{the number of multi-indices } \alpha \text{ such that } |\alpha| \leq m,
$$

and introduce a subspace Δ of the product space $L^1(\mathbf{R}^n)^N$ as follows:

$$
\Delta = \left\{(\psi_\alpha)_{|\alpha|\leq m} : \varphi \in \mathcal{S}(\mathbf{R}^n)\right\}.
$$

Then it follows that the mapping

$$
\mathcal{S}(\mathbf{R}^n) \ni \varphi \longmapsto (\psi_\alpha)_{|\alpha|\leq m} \in \Delta
$$

is bijective. Indeed, it suffices to note that

$$
\varphi \in \mathcal{S}(\mathbf{R}^n) \text{ and } \partial_{x_1}\cdots\partial_{x_n}\varphi = 0 \implies \varphi = 0.
$$

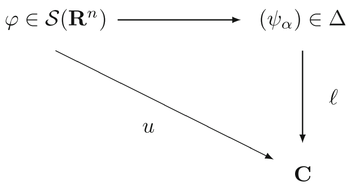
On the other hand, we introduce a norm $\|\cdot\|$ on the product space $L^1(\mathbf{R}^n)^N$ by the formula

$$\|(f_1, \ldots, f_N)\| = \sum_{j=1}^{N} \|f_j\|_{L^1} \quad \text{for every } (f_1, \ldots, f_N) \in L^1(\mathbf{R}^n)^N,$$

and define a linear functional ℓ on the subspace Δ by the formula

$$\ell\left((\psi_\alpha)_{|\alpha| \le m}\right) = \langle u, \varphi \rangle \quad \text{for every } (\psi_\alpha)_{|\alpha| \le m} \in \Delta.$$

By virtue of inequality (5.24), it follows that the linear functional ℓ is continuous. The situation can be visualized in the following diagram:



By applying the Hahn–Banach Theorem (Theorem 3.22) to our situation, we obtain that the linear functional ℓ can be extended a continuous linear functional $\tilde{\ell}$ on the whole space $L^1(\mathbf{R}^n)^N$. However, we recall that the dual space $\left(L^1(\mathbf{R}^n)\right)'$ can be identified with the space $L^\infty(\mathbf{R}^n)$. Hence we have

$$\tilde{\ell} \in \left(L^1(\mathbf{R}^n)^N\right)' = L^\infty(\mathbf{R}^n)^N.$$

Therefore, we can find functions $\{g_\alpha\}_{|\alpha| \le m}$ of $L^\infty(\mathbf{R}^n)$ such that

$$\tilde{\ell}\left((\psi_\alpha)_{|\alpha| \le m}\right) = \sum_{|\alpha| \le m} \langle g_\alpha, \psi_\alpha \rangle \quad \text{for every } (\psi_\alpha)_{|\alpha| \le m} \in L^1(\mathbf{R}^n)^N.$$

Summing up, we have proved that

$$\begin{aligned} \langle u, \varphi \rangle &= \ell\left((\psi_\alpha)_{|\alpha| \le m}\right) = \sum_{|\alpha| \le m} \langle g_\alpha, \psi_\alpha \rangle \\ &= \sum_{|\alpha| \le m} \left\langle g_\alpha, (1 + |x|^2)^{m/2} \partial_{x_1} \cdots \partial_{x_n} \partial_x^\alpha \varphi \right\rangle \\ &= \sum_{|\alpha| \le m} (-1)^{n+|\alpha|} \left\langle \partial_{x_1} \cdots \partial_{x_n} \partial_x^\alpha \left[(1 + |x|^2)^{m/2} g_\alpha\right], \varphi \right\rangle. \end{aligned}$$

If we let

$$f_\alpha(x) = (-1)^{n+|\alpha|} g_\alpha(x) \in L^\infty(\mathbf{R}^n),$$

we obtain that

$$\langle u, \varphi \rangle = \left\langle \sum_{|\alpha| \le m} \partial_{x_1} \cdots \partial_{x_n} \partial_x^\alpha \left[(1 + |x|^2)^{m/2} f_\alpha \right], \varphi \right\rangle \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}^n).$$

This proves the desired formula (5.20).

The proof of Theorem 5.22 is complete.

Furthermore, we can prove a simplified version of formula (5.20):

Corollary 5.23. *Every distribution $u \in \mathcal{S}'(\mathbf{R}^n)$ can be expressed in the form*

$$u = \partial_{x_1}^{m+3} \cdots \partial_{x_n}^{m+3} G, \tag{5.25}$$

where $G(x)$ is a tempered continuous function on $\mathbf{R}^n$.

Proof. If we let

$$h_\alpha(x) = \int_0^{x_1} \cdots \int_0^{x_n} (1 + |y|^2)^{m/2} f_\alpha(y)\, dy_1 \cdots dy_n, \tag{5.26}$$

we obtain that the function $h_\alpha(x)$ is continuous and satisfies the inequality

$$\begin{aligned} |h_\alpha(x)| &\le (1 + |x|^2)^{m/2} \int_0^{x_1} \cdots \int_0^{x_n} |f_\alpha(y)|\, dy_1 \cdots dy_n \\ &\le \| f_\alpha \|_{L^\infty(\mathbf{R}^n)} \left(1 + |x|^2\right)^{(m+n)/2}. \end{aligned} \tag{5.27}$$

This proves that $h_\alpha(x)$ is a tempered function on $\mathbf{R}^n$.

On the other hand, we have the following elementary formula:

$$h(t) = \left(\frac{d}{dt} \right)^k \int_0^t \frac{(t-\tau)^{k-1}}{(k-1)!} h(\tau)\, d\tau, \quad k \in \mathbf{N}.$$

By making use of this formula, we obtain that

$$\begin{aligned} & h_\alpha(x) \\ &= \partial_{x_1}^{m+1} \cdots \partial_{x_n}^{m+1} \left(\int_0^{x_1} \cdots \int_0^{x_n} \frac{(x_1 - y_1)^m \cdots (x_n - y_n)^m}{m! \cdots m!}\, h_\alpha(y)\, dy_1 \cdots dy_n \right) \\ &= \partial_{x_1}^{m+1} \cdots \partial_{x_n}^{m+1} G_\alpha(x), \end{aligned} \tag{5.28}$$

where

$$G_\alpha(x) := \int_0^{x_1} \cdots \int_0^{x_n} \frac{(x_1 - y_1)^m \cdots (x_n - y_n)^m}{m! \cdots m!}\, h_\alpha(y)\, dy_1 \cdots dy_n.$$

By inequality (5.27), it follows that $G_\alpha(x)$ is a tempered continuous function on $\mathbf{R}^n$.

Therefore, if we define a tempered continuous function $G(x)$ by the formula

$$G(x) = \sum_{|\alpha| \leq m} G_\alpha(x),$$

then we obtain from formulas (5.20), (5.26) and (5.28) that

$$\begin{aligned} u &= \sum_{|\alpha| \leq m} \partial_{x_1} \cdots \partial_{x_n} \partial_x^\alpha \left[(1 + |x|^2)^{m/2} f_\alpha\right] \\ &= \sum_{|\alpha| \leq m} \partial_{x_1}^2 \cdots \partial_{x_n}^2 \partial_x^\alpha h_\alpha = \sum_{|\alpha| \leq m} \partial_{x_1}^{m+3} \cdots \partial_{x_n}^{m+3} G_\alpha \\ &= \partial_{x_1}^{m+3} \cdots \partial_{x_n}^{m+3} G. \end{aligned}$$

This proves the desired formula (5.25).

The proof of Corollary 5.23 is complete.

By combining Theorem 5.15 and Corollaries 5.6 and 5.23, we can obtain the following structure theorem for distributions with compact support:

Theorem 5.24 (the structure theorem). *Let $u \in \mathcal{E}'(\mathbf{R}^n)$. For any given compact neighborhood ω of* supp u*, there exist a non-negative integer m and a continuous function $H(x)$ on $\mathbf{R}^n$ with support in ω such that*

$$u = \partial_{x_1}^{m+3} \cdots \partial_{x_n}^{m+3} H. \tag{5.29}$$

Proof. By applying Corollary 5.6 to our situation, we can construct a function $\chi \in C_0^\infty(\omega)$ such that

$$\begin{aligned} 0 \leq \chi(x) \leq 1 &\quad \text{on } \omega, \\ \chi(x) = 1 &\quad \text{on } \operatorname{supp} u. \end{aligned}$$

The desired formula (5.29) follows from formula (5.25) by replacing $G(x)$ by $H(x) = \chi(x)\, G(x)$.

The proof of Theorem 5.24 is complete.

Now we give some concrete and important examples of tempered distributions:

Example 5.25. (a) Dirac measures: $\delta(x)$.

(b) Riesz potentials:

$$R_\alpha(x) = \frac{\Gamma((n-\alpha)/2)}{2^\alpha \pi^{n/2} \Gamma(\alpha/2)} \frac{1}{|x|^{n-\alpha}}, \quad 0 < \alpha < n.$$

(c) Newtonian potentials:

$$N(x) = \frac{\Gamma((n-2)/2)}{4\pi^{n/2}} \frac{1}{|x|^{n-2}}, \quad n \geq 3.$$

(d) Bessel potentials:

$$G_\alpha(x) = \frac{1}{\Gamma(\alpha/2)} \frac{1}{(4\pi)^{n/2}} \int_0^\infty e^{-t-\frac{|x|^2}{4t}} t^{\frac{\alpha-n}{2}} \frac{dt}{t}, \quad \alpha > 0.$$

It is known (see Aronszajn–Smith [AS]) that the function $G_\alpha(x)$ is represented as follows:

$$G_\alpha(x) = \frac{1}{2^{(n+\alpha-2)/2}\pi^{n/2}\Gamma(\alpha/2)} K_{(n-\alpha)/2}(|x|)|x|^{\frac{\alpha-n}{2}},$$

where $K_{(n-\alpha)/2}(z)$ is the modified Bessel function of the third kind (cf. Watson [Wt]).

(e) Riesz kernels:

$$R_j(x) = -\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \text{v. p.} \frac{x_j}{|x|^{n+1}}, \quad 1 \leq j \leq n.$$

The distribution

$$\text{v. p.} \frac{x_j}{|x|^{n+1}}$$

is an extension of v. p. $(1/x)$ to the multi-dimensional case (see Example 5.12).

(f) In the case $n = 1$, we have

(1) $Y(x) \in \mathcal{S}'(\mathbf{R})$.
(2) $\text{v. p.} \dfrac{1}{x} \in \mathcal{S}'(\mathbf{R})$.

The importance of tempered distributions lies in the fact that they have Fourier transforms.

If $u \in \mathcal{S}'(\mathbf{R}^n)$, we define its (direct) *Fourier transform* $\mathcal{F}u = \hat{u}$ by the formula

$$\langle \mathcal{F}u, \varphi \rangle = \langle u, \mathcal{F}\varphi \rangle \quad \text{for all } \varphi \in \mathcal{S}(\mathbf{R}^n). \tag{5.30}$$

Then we have $\mathcal{F}u \in \mathcal{S}'(\mathbf{R}^n)$, since the Fourier transform

$$\mathcal{F} : \mathcal{S}(\mathbf{R}^n) \longrightarrow \mathcal{S}(\mathbf{R}^n)$$

is an isomorphism. Furthermore, in view of formulas (5.19a) and (5.19b) it follows that the above definition (5.30) agrees with definition (5.14) if $u \in \mathcal{S}(\mathbf{R}^n)$. We also denote $\mathcal{F}u$ by $\hat{u}$.

Similarly, if $v \in \mathcal{S}'(\mathbf{R}^n)$, we define its *inverse Fourier transform* $\mathcal{F}^* v = \check{v}$ by the formula

$$\langle \mathcal{F}^* v, \psi \rangle = \langle v, \mathcal{F}^* \psi \rangle \quad \text{for all } \psi \in \mathcal{S}(\mathbf{R}^n).$$

The next theorem, which is a consequence of Theorem 5.21, summarizes the basic properties of Fourier transforms in the space $\mathcal{S}'(\mathbf{R}^n)$:

Theorem 5.26. *(i) The Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ map $\mathcal{S}'(\mathbf{R}^n)$ continuously into itself. Furthermore, we have, for all multi-indices α and β,*

$$\mathcal{F}(D^\alpha u)(\xi) = \xi^\alpha \mathcal{F}u(\xi), \quad u \in \mathcal{S}'(\mathbf{R}^n),$$
$$D_\xi^\beta(\mathcal{F}u(\xi)) = \mathcal{F}((-x)^\beta u)(\xi), \quad u \in \mathcal{S}'(\mathbf{R}^n).$$

(ii) The Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ are isomorphisms of $\mathcal{S}'(\mathbf{R}^n)$ onto itself; more precisely, $\mathcal{F}\mathcal{F}^* = \mathcal{F}^*\mathcal{F} = I$ on $\mathcal{S}'(\mathbf{R}^n)$.

(iii) The transforms $\mathcal{F}$ and $\mathcal{F}^*$ are norm-preserving operators on $L^2(\mathbf{R}^n)$ and $\mathcal{F}\mathcal{F}^* = \mathcal{F}^*\mathcal{F} = I$ on $L^2(\mathbf{R}^n)$.

Assertion (iii) is referred to as the *Plancherel theorem*.

We remark that Theorems 5.21 and 5.26 can be visualized as follows:

$$\begin{array}{ccc}
\mathcal{S}'(\mathbf{R}^n) & \xrightarrow{\mathcal{F}} & \mathcal{S}'(\mathbf{R}^n) \\
\uparrow & & \uparrow \\
L^2(\mathbf{R}^n) & \xrightarrow{\mathcal{F}} & L^2(\mathbf{R}^n) \\
\uparrow & & \uparrow \\
\mathcal{S}(\mathbf{R}^n) & \xrightarrow[\mathcal{F}]{} & \mathcal{S}(\mathbf{R}^n)
\end{array}$$

$$\begin{array}{ccc}
\mathcal{S}'(\mathbf{R}^n) & \xleftarrow{\mathcal{F}^*} & \mathcal{S}'(\mathbf{R}^n) \\
\uparrow & & \uparrow \\
L^2(\mathbf{R}^n) & \xleftarrow{\mathcal{F}^*} & L^2(\mathbf{R}^n) \\
\uparrow & & \uparrow \\
\mathcal{S}(\mathbf{R}^n) & \xleftarrow[\mathcal{F}^*]{} & \mathcal{S}(\mathbf{R}^n)
\end{array}$$

5.4.11 The Fourier Transform of Tempered Distributions

In this subsection we calculate explicitly the Fourier transform of some important examples of tempered distributions. First, we consider the distributions v. p. $\frac{1}{x}$ and $Y(x)$ in the case $n = 1$:

Example 5.27. (1) For the distribution v. p. $\frac{1}{x}$, we have

$$\mathcal{F}\left(\text{v. p. }\frac{1}{x}\right)(\xi) = -\pi\, i \operatorname{sgn}\xi = \begin{cases} -\pi i & \text{for } \xi > 0, \\ \pi i & \text{for } \xi < 0. \end{cases} \tag{5.31}$$

(2) For the Heaviside function $Y(x)$, we have

$$(\mathcal{F}Y)(\xi) = \hat{Y}(\xi) = \frac{1}{i}\ \text{v. p. }\frac{1}{\xi} + \pi\delta(\xi). \tag{5.32}$$

Proof. (1) We calculate the Fourier transform of the distribution

$$h(x) = \frac{1}{\pi}\ \text{v. p. }\frac{1}{x}.$$

For $0 < \varepsilon < \mu$, we let

$$h_{\varepsilon,\mu}(x) = \begin{cases} \frac{1}{\pi x} & \text{if } \varepsilon < |x| < \mu, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$h_\varepsilon(x) = \begin{cases} \frac{1}{\pi x} & \text{if } |x| > \varepsilon, \\ 0 & \text{if } |x| \leq \varepsilon. \end{cases}$$

Then it follows that

$$\begin{aligned}(\mathcal{F}h_{\varepsilon,\mu})(\xi) &= \int_{\varepsilon<|x|<\mu} \frac{e^{-ix\cdot\xi}}{\pi x}\,dx = -\frac{2i}{\pi}\int_\varepsilon^\mu \frac{\sin(x\cdot\xi)}{x}\,dx \\ &= \begin{cases} -\frac{2i}{\pi}\int_{\varepsilon\xi}^{\mu\xi} \frac{\sin x}{x}\,dx & \text{if } \xi > 0, \\ \frac{2i}{\pi}\int_{-\varepsilon\xi}^{-\mu\xi} \frac{\sin x}{x}\,dx & \text{if } \xi < 0. \end{cases}\end{aligned}$$

Hence we have, as $\mu \uparrow \infty$,

$$(\mathcal{F}h_\varepsilon)(\xi) = -\frac{2i}{\pi}\int_{\varepsilon|\xi|}^\infty \frac{\sin x}{x}\,dx \cdot \operatorname{sign}\xi. \tag{5.33}$$

Moreover, by letting $\varepsilon \downarrow 0$ in formula (5.33) we obtain that

$$\hat{h}(\xi) = (\mathcal{F}h)(\xi) = -i \ \operatorname{sign} \xi.$$

This implies that

$$\mathcal{F}\left(\text{v. p.} \frac{1}{x}\right)(\xi) = \pi\, \hat{h}(\xi) = -\pi\, i \ \operatorname{sgn} \xi.$$

The proof of the desired formula (5.31) is complete.

(2) Similarly, we can calculate the inverse Fourier transform of v. p. $\frac{1}{x}$ as follows:

$$\mathcal{F}^*\left(\text{v. p.} \frac{1}{\xi}\right)(x) = \frac{i}{2} \operatorname{sgn} x. \tag{5.34}$$

By formula (5.34), it follows that

$$\mathcal{F}^*\left(\frac{2}{i} \ \text{v. p.} \frac{1}{\xi} + 2\pi\delta(\xi)\right) = \operatorname{sgn} x + 1 = \begin{cases} 2 & \text{for } x > 0, \\ 0 & \text{for } x < 0 \end{cases} \tag{5.35}$$
$$= 2Y(x).$$

Therefore, by applying the Fourier transform $\mathcal{F}$ to both sides of formula (5.35) we obtain from part (ii) of Theorem 5.26 that

$$\hat{Y}(\xi) = (\mathcal{F}Y)(\xi) = \frac{1}{2}\mathcal{F}\mathcal{F}^*\left(\frac{2}{i} \ \text{v. p.} \frac{1}{\xi} + 2\pi\delta(\xi)\right) = \frac{1}{i} \ \text{v. p.} \frac{1}{\xi} + \pi\delta(\xi).$$

The proof of the desired formula (5.32) is complete.

The proof of Example 5.27 is now complete.

Secondly, we calculate the inverse Fourier transform of some homogeneous functions:

Example 5.28. (i) For $\lambda \neq -n - 2k$ with $k = 0, 1, 2, \ldots$, we let

$$\sigma(\xi) = |\xi|^\lambda, \quad \xi \neq 0.$$

Then its inverse Fourier transform

$$k(x) = (\mathcal{F}^*\sigma)(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \, |\xi|^\lambda \, d\xi$$

is given by the formula

$$k(x) = \frac{2^\lambda}{\pi^{n/2}} \frac{\Gamma\left(\frac{n+\lambda}{2}\right)}{\Gamma\left(-\frac{\lambda}{2}\right)} \frac{1}{|x|^{n+\lambda}}, \quad x \neq 0.$$

(ii) The Fourier transforms $\widehat{R_\alpha}(\xi)$ and $\hat{N}(\xi)$ are obtained from part (i) by taking $\lambda := \alpha - n$ and $\lambda := 2 - n$, respectively. More precisely, we have the following two formulas (a) and (b):

(a) Riesz potentials:

$$R_\alpha(x) = \frac{\Gamma((n-\alpha)/2)}{2^\alpha \pi^{n/2} \Gamma(\alpha/2)} \frac{1}{|x|^{n-\alpha}}, \quad 0 < \alpha < n,$$

$$\widehat{R_\alpha}(\xi) = \frac{1}{|\xi|^\alpha}.$$

(b) Newtonian potentials:

$$N(x) = \frac{\Gamma((n-2)/2)}{4\pi^{n/2}} \frac{1}{|x|^{n-2}}, \quad n \geq 3,$$

$$\hat{N}(\xi) = \frac{1}{|\xi|^2}.$$

Proof. The proof of part (i) is divided into two steps.

Step 1: We let

$$\Phi(x) = \int_{\mathbf{R}^n} e^{ix\cdot\xi} \, |\xi|^\lambda \, d\xi,$$

and show that there exists a constant C such that

$$\Phi(x) = \frac{C}{|x|^{n+\lambda}} \quad \text{for every } x \neq 0. \tag{5.36}$$

(1) First, we show that the function $\Phi(x)$ is invariant under rotations, that is, we have, for every orthogonal $n \times n$ matrix A,

$$\Phi(Ax) = \Phi(x), \quad x \neq 0. \tag{5.37}$$

Indeed, since we have

$$Ax \cdot \xi = \langle Ax, \xi \rangle = \left\langle x, {}^t A\xi \right\rangle = \left\langle x, A^{-1}\xi \right\rangle = x \cdot A^{-1}\xi, \quad \xi \in \mathbf{R}^n,$$

it follows that

$$\begin{aligned}\Phi(Ax) &= \int_{\mathbf{R}^n} e^{iAx\cdot\xi}\,|A\xi|^\lambda\,d\xi = \int_{\mathbf{R}^n} e^{ix\cdot A^{-1}\xi}\,|A\xi|^\lambda\,d\xi \\ &= \int_{\mathbf{R}^n} e^{ix\cdot\eta}\,|\eta|^\lambda\,|\det A|\,d\eta \qquad (\xi = A\eta) \\ &= \int_{\mathbf{R}^n} e^{ix\cdot\eta}\,|\eta|^\lambda\,d\eta = \Phi(x) \quad \text{for every } x \neq 0.\end{aligned}$$

(2) Secondly, the function $\Phi(x)$ is positively homogeneous of degree $-\lambda - n$ in x, that is, we have, for all $r > 0$,

$$\Phi(rx) = r^{-\lambda-n}\,\Phi(x), \quad x \neq 0. \tag{5.38}$$

Indeed, it is easy to verify that

$$\begin{aligned}\Phi(rx) &= \int_{\mathbf{R}^n} e^{irx\cdot\xi}\,|\xi|^\lambda\,d\xi = \int_{\mathbf{R}^n} e^{ix\cdot r\xi}\,|\xi|^\lambda\,d\xi \\ &= \int_{\mathbf{R}^n} e^{ix\cdot\eta}\,|\eta|^\lambda\,r^{-\lambda}r^{-n}d\eta \quad (\xi = \eta/r) \\ &= r^{-\lambda-n}\int_{\mathbf{R}^n} e^{ix\cdot\eta}\,|\eta|^\lambda\,d\eta = r^{-\lambda-n}\Phi(x), \quad x \neq 0.\end{aligned}$$

(3) Therefore, we find from assertions (5.37) and (5.38) that the function $|x|^{n+\lambda}\,\Phi(x)$ is a constant. This proves the desired formula (5.36).

Step 2: Now we calculate the inverse Fourier transform $k(x)$ of the function $\sigma(\xi)$

$$k(x) = (\mathcal{F}^*\sigma)(x) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{ix\cdot\xi}\,|\xi|^\lambda\,d\xi, \quad x \neq 0.$$

By formula (5.36), we can write the function $k(x)$ in the form

$$k(x) = \frac{\alpha}{|x|^{n+\lambda}}, \quad x \neq 0, \tag{5.39}$$

where

$$\alpha := \frac{C}{(2\pi)^n}.$$

We have only to show that

$$\alpha = \frac{C}{(2\pi)^n} = \frac{2^\lambda}{\pi^{n/2}}\,\frac{\Gamma\left(\frac{n+\lambda}{2}\right)}{\Gamma\left(-\frac{\lambda}{2}\right)}. \tag{5.40}$$

To do this, by using formula (5.19a) and formula (5.39) we have, for all $\varphi \in \mathcal{S}(\mathbf{R}^n)$,

$$\int_{\mathbf{R}^n} |\xi|^\lambda \, (\mathcal{F}^*\varphi)(\xi)\, d\xi = \int_{\mathbf{R}^n} \mathcal{F}^*\left(|\xi|^\lambda\right)(x)\varphi(x)\, dx \qquad (5.41)$$
$$= \alpha \int_{\mathbf{R}^n} \frac{1}{|x|^{n+\lambda}} \varphi(x)\, dx.$$

However, if we take

$$\varphi(x) = e^{-|x|^2} \quad \text{for every } x \in \mathbf{R}^n,$$

then it follows from formula (5.15) with $a := 1$ that

$$\mathcal{F}^*\varphi(\xi) = \left(\frac{1}{4\pi}\right)^{n/2} e^{-\frac{|\xi|^2}{4}} \quad \text{for every } \xi \in \mathbf{R}^n.$$

Therefore, we obtain from formula (5.41) that

$$\left(\frac{1}{4\pi}\right)^{n/2} \int_{\mathbf{R}^n} |\xi|^\lambda \, e^{-\frac{|\xi|^2}{4}}\, d\xi = \alpha \int_{\mathbf{R}^n} \frac{1}{|x|^{n+\lambda}} e^{-|x|^2}\, dx. \qquad (5.42)$$

(1) We calculate the right-hand side of formula (5.41): By using the polar coordinates, we have

$$\alpha \int_{\mathbf{R}^n} \frac{1}{|x|^{n+\lambda}} e^{-|x|^2}\, dx \qquad (5.43)$$
$$= \alpha \int_{\Sigma_n} \int_0^\infty \frac{1}{r^{n+\lambda}} e^{-r^2} r^{n-1}\, dr\, d\sigma \quad (x = r\sigma,\ \sigma \in \Sigma_n)$$
$$= \alpha\, |\Sigma_n| \int_0^\infty e^{-r^2}\, r^{-\lambda-1}\, dr = \frac{\alpha}{2}\, \omega_n \int_0^\infty e^{-s}\, s^{-(\lambda/2)-1}\, dr \quad (s = r^2)$$
$$= \frac{1}{2}\alpha\, \omega_n\, \Gamma\left(-\frac{\lambda}{2}\right).$$

Here

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad n \geq 2,$$

is the surface area of the unit sphere Σ_n.

(2) We calculate the left-hand side of formula (5.41): Similarly, by using the polar coordinates we have

$$\begin{aligned}
&\left(\frac{1}{4\pi}\right)^{n/2} \int_{\mathbf{R}^n} |\xi|^\lambda \, e^{-\frac{|\xi|^2}{4}} \, d\xi \qquad (5.44)\\
&= \left(\frac{1}{4\pi}\right)^{n/2} \int_{\Sigma_n} \int_0^\infty r^\lambda e^{-\frac{r^2}{4}} \, r^{n-1} \, dr \, d\sigma \qquad (\xi = r\sigma, \ \sigma \in \Sigma_n)\\
&= \frac{1}{(4\pi)^{n/2}} \, |\Sigma_n| \int_0^\infty r^{n+\lambda-1} \, e^{-\frac{r^2}{4}} \, dr\\
&= \frac{2^\lambda}{\pi^{n/2}} \, \omega_n \int_0^\infty s^{n+\lambda-1} \, e^{-s^2} \, ds \qquad (r = 2s)\\
&= \frac{2^\lambda}{\pi^{n/2}} \frac{\omega_n}{2} \int_0^\infty t^{(n+\lambda)/2-1} \, e^{-t} \, dt \qquad (t = s^2)\\
&= \frac{2^\lambda}{\pi^{n/2}} \frac{\omega_n}{2} \, \Gamma\left(\frac{n+\lambda}{2}\right).
\end{aligned}$$

Therefore, the desired formula (5.40) follows by combining formulas (5.43) and (5.44):

$$\alpha \, \frac{\omega_n}{2} \, \Gamma\left(-\frac{\lambda}{2}\right) = \frac{2^\lambda}{\pi^{n/2}} \frac{\omega_n}{2} \, \Gamma\left(\frac{n+\lambda}{2}\right).$$

The proof of Example 5.28 is now complete.

Thirdly, we calculate the Fourier transform of a tempered distribution which is closely related to the stationary phase theorem [CP, Chapitre III, Théorème 9.3]:

Example 5.29. (1) For any $\lambda \in \mathbf{R} \setminus \{0\}$, we consider the function

$$f(x) = \exp\left[\frac{i\lambda}{2} x^2\right], \quad x \in \mathbf{R}.$$

Then its Fourier transform $(\mathcal{F} f)(\xi) = \hat{f}(\xi)$ is given by the formula

$$\hat{f}(\xi) = \frac{\sqrt{2\pi}}{\sqrt{|\lambda|}} \exp\left[\frac{i\pi}{4} \frac{\lambda}{|\lambda|}\right] \exp\left[-\frac{i}{2\lambda} \xi^2\right], \quad \xi \in \mathbf{R}. \qquad (5.45)$$

(2) For any symmetric, non-singular $q \times q$ matrix Q, we consider the function

$$G(y) = \exp\left[\frac{i}{2} \langle Qy, y \rangle\right], \quad y \in \mathbf{R}^q.$$

Then its Fourier transform $(\mathcal{F} G)(\eta) = \hat{G}(\eta)$ is given by the formula

$$\hat{G}(\eta) = \frac{(2\pi)^{q/2}}{\sqrt{|\det Q|}} \exp\left[\frac{i\pi}{4}\operatorname{sign} Q\right] \exp\left[-\frac{i}{2}\left\langle Q^{-1}\eta, \eta\right\rangle\right], \quad \eta \in \mathbf{R}^q. \quad (5.46)$$

Here the signature sign Q of Q is the number α of plus ones minus the number β of the minus ones in the diagonalized $q \times q$ matrix: sign $Q = \alpha - \beta$.

Proof. (1) The proof of formula (5.45) is divided into two steps.

Step 1: For any complex number $\alpha \in \mathbf{C}$, we let

$$T_\alpha(x) = \exp\left[-\frac{\alpha}{2}x^2\right], \quad x \in \mathbf{R},$$

and define a distribution $T_\alpha \in \mathcal{S}'(\mathbf{R})$ by the formula

$$\langle T_\alpha, \varphi\rangle = \int_{-\infty}^{\infty} \exp\left[-\frac{\alpha}{2}x^2\right] \varphi(x)\, dx \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}).$$

(a) First, we show that the function

$$\mathbf{C} \ni \alpha \longmapsto T_\alpha \in \mathcal{S}'(\mathbf{R})$$

is *continuous* for $\alpha = \xi + i\eta$ in the closed half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$. Let α_0 be an arbitrary point of $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$. Then we have, for any complex number $\alpha = \xi + i\eta$ near $\alpha_0 = \xi_0 + i\eta_0$,

$$\left|\exp\left[-\frac{\alpha}{2}x^2\right]\varphi(x)\right| \leq \exp\left[-\frac{\xi}{2}x^2\right]|\varphi(x)| \leq |\varphi(x)|.$$

Therefore, by applying the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] we obtain that

$$\begin{aligned}\lim_{\alpha\to\alpha_0} \langle T_\alpha, \varphi\rangle &= \lim_{\alpha\to\alpha_0} \int_{-\infty}^{\infty} \exp\left[-\frac{\alpha}{2}x^2\right]\varphi(x)\, dx \\ &= \int_{-\infty}^{\infty} \exp\left[-\frac{\alpha_0}{2}x^2\right]\varphi(x)\, dx \\ &= \langle T_{\alpha_0}, \varphi\rangle \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}).\end{aligned}$$

This proves that the function $\alpha \longmapsto T_\alpha$ is continuous for α in the closed half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$.

(b) Secondly, we show that the function

$$\mathbf{C} \ni \alpha \longmapsto T_\alpha \in \mathcal{S}'(\mathbf{R})$$

is *holomorphic* for α in the open half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$.

Let $\alpha_0 = \xi_0 + i\eta_0$ be an arbitrary point of $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$. For any complex number $\alpha = \xi + i\eta$ near $\alpha_0 = \xi_0 + i\eta_0$, we consider a function $F(\xi, \eta)$ defined by the formula

$$F(\xi, \eta) = \langle T_\alpha, \varphi \rangle = \int_{-\infty}^{\infty} \exp\left[-\frac{\alpha}{2}x^2\right] \varphi(x)\, dx$$
$$= \int_{-\infty}^{\infty} \exp\left[-\frac{\xi}{2}x^2\right] \exp\left[-\frac{i\eta}{2}x^2\right] \varphi(x)\, dx.$$

Then it is easy to see that the partial derivative

$$\frac{\partial F}{\partial \xi}(\xi_0, \eta_0) = -\int_{-\infty}^{\infty} \frac{x^2}{2} \exp\left[-\frac{\xi_0}{2}x^2\right] \exp\left[-\frac{i\eta_0}{2}x^2\right] \varphi(x)\, dx$$

exists. Moreover, by arguing just as in Step (a) we find that the function

$$\frac{\partial F}{\partial \xi}(\xi, \eta) = -\int_{-\infty}^{\infty} \frac{x^2}{2} \exp\left[-\frac{\xi}{2}x^2\right] \exp\left[-\frac{i\eta}{2}x^2\right] \varphi(x)\, dx$$

is a continuous function of $\alpha = \xi + i\eta$ in $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$.
Similarly, it is easy to see that the partial derivative

$$\frac{\partial F}{\partial \eta}(\xi, \eta) = -i\int_{-\infty}^{\infty} \frac{x^2}{2} \exp\left[-\frac{\xi}{2}x^2\right] \exp\left[-\frac{i\eta}{2}x^2\right] \varphi(x)\, dx$$

exists, and is a continuous function of α in $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$.
Summing up, we have proved that the function

$$F(\xi, \eta) = \langle T_\alpha, \varphi \rangle = \int_{-\infty}^{\infty} \exp\left[-\frac{\xi}{2}x^2\right] \exp\left[-\frac{i\eta}{2}x^2\right] \varphi(x)\, dx$$

is of class C^1 in the closed half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$, and satisfies the Cauchy–Riemann equation

$$\frac{\partial F}{\partial \xi} + i\frac{\partial F}{\partial \eta} = 0 \quad \text{in the open half-plane } \{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}.$$

This proves that the function $\alpha \longmapsto T_\alpha$ is holomorphic in the open half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$.

Step 2: We note that

$$T_\alpha(x) = \exp\left[-\frac{\alpha}{2}x^2\right] \in \mathcal{S}(\mathbf{R}) \quad \text{for every } \alpha > 0,$$

and further (see formula (5.15) for $n := 1$) that

$$\widehat{T_\alpha}(\xi) = \frac{\sqrt{2\pi}}{\sqrt{\alpha}} \exp\left[-\frac{1}{2\alpha}\xi^2\right] \in \mathcal{S}(\mathbf{R}) \quad \text{for every } \alpha > 0. \tag{5.47}$$

Hence it follows from assertions (a) and (b) that the function

$$\mathbf{C} \ni \alpha \longmapsto \left\langle \widehat{T_\alpha}, \varphi \right\rangle = \langle T_\alpha, \hat{\varphi} \rangle \qquad (\varphi \in \mathcal{S}(\mathbf{R}))$$

is continuous in the closed half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$ and is holomorphic in the open half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$. This proves that the Fourier transform $\widehat{T_\alpha}$ of T_α is holomorphic in $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$, and is continuous in $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha \geq 0\}$.
On the other hand, it is easy to see that the function

$$\frac{1}{\sqrt{\alpha}} = \exp\left[-\frac{1}{2}\log\alpha\right] = \frac{1}{\sqrt{|\alpha|}} \exp\left[-\frac{1}{2} i \arg\alpha\right]$$

is holomorphic in the open half-plane $\{\alpha \in \mathbf{C} : \operatorname{Re}\alpha > 0\}$ and is continuous in the half-plane $\{\alpha \in \mathbf{C} : \alpha \neq 0,\ \operatorname{Re}\alpha \geq 0\}$.
By virtue of the unicity theorem, it follows from formula (5.47) that

$$\widehat{T_\alpha}(\xi) = \frac{\sqrt{2\pi}}{\sqrt{\alpha}} \exp\left[-\frac{1}{2\alpha}\xi^2\right], \quad \operatorname{Re}\alpha > 0. \tag{5.48}$$

Furthermore, by passing to the limit in formula (5.48) we obtain that

$$\widehat{T_\alpha}(\xi) = \frac{\sqrt{2\pi}}{\sqrt{\alpha}} \exp\left[-\frac{1}{2\alpha}\xi^2\right], \quad \operatorname{Re}\alpha \geq 0, \quad \alpha \neq 0.$$

In particular, we have

$$\hat{f}(\xi) = \frac{\sqrt{2\pi}}{\sqrt{-i\lambda}} \exp\left[-\frac{i}{2\lambda}\xi^2\right], \quad \lambda \in \mathbf{R} \setminus \{0\}, \tag{5.49}$$

by taking

$$\alpha = -i\lambda.$$

However, we note that

$$\frac{1}{\sqrt{-i\lambda}} = \begin{cases} \frac{1}{\sqrt{|\lambda|}} \exp\left[\frac{\pi}{4} i\right] & \text{if } \lambda > 0, \\ \frac{1}{\sqrt{|\lambda|}} \exp\left[-\frac{\pi}{4} i\right] & \text{if } \lambda < 0. \end{cases} \tag{5.50}$$

Therefore, the desired formula (5.45) follows by combining formulas (5.49) and (5.50).

(2) In order to prove (5.46), we take an orthogonal $q \times q$ matrix M such that

$$ {}^{t}MQM = D = \begin{pmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & \lambda_q \end{pmatrix} $$

and let

$$ y = Mx, \quad x \in \mathbf{R}^q. $$

Then it follows that

$$ \langle Qy, y \rangle = \langle QMx, Mx \rangle = \left\langle {}^{t}MQMx, x \right\rangle = \langle Dx, x \rangle = \sum_{j=1}^{q} \lambda_j^2 x_j^2. $$

Hence we have

$$ \begin{aligned} \hat{G}(\eta) &= \int_{\mathbf{R}^q} \exp\left[-iy \cdot \eta\right] \exp\left[\frac{i}{2} \langle Qy, y \rangle\right] dy \\ &= \int_{\mathbf{R}^q} \exp\left[-iMx \cdot \eta\right] \exp\left[\frac{i}{2} \sum_{j=1}^{q} \lambda_j^2 x_j^2\right] |\det M| \, dx \\ &= \int_{\mathbf{R}^q} \exp\left[-ix \cdot M^{-1}\eta\right] \exp\left[\frac{i}{2} \sum_{j=1}^{q} \lambda_j^2 x_j^2\right] dx. \end{aligned} $$

Moreover, by replacing η by $M\eta$ we have

$$ \begin{aligned} \hat{G}(M\eta) &= \int_{\mathbf{R}^q} \exp\left[-ix \cdot \eta\right] \exp\left[\frac{i}{2} \sum_{j=1}^{q} \lambda_j^2 x_j^2\right] dx \\ &= \prod_{j=1}^{q} \int_{\mathbf{R}} \exp\left[-ix_j \eta_j\right] \exp\left[\frac{i}{2} \lambda_j^2 x_j^2\right] dx_j. \end{aligned} $$

Therefore, by using formula (5.45) we obtain that

$$\begin{aligned}\hat{G}(M\eta) &= \prod_{j=1}^{q} \int_{\mathbf{R}} \exp\left[-ix_j\eta_j\right] \exp\left[\frac{i}{2}\lambda_j^2 x_j^2\right] dx_j \qquad (5.51)\\ &= \prod_{j=1}^{q} \int_{\mathbf{R}} \sqrt{2\pi}\frac{1}{\sqrt{|\lambda_j|}} \exp\left[i\frac{\pi}{4}\frac{\lambda_j}{|\lambda_j|}\right] \exp\left[-\frac{i}{2}\frac{\eta_j^2}{\lambda_j}\right]\\ &= \frac{(2\pi)^{q/2}}{\sqrt{|\det Q|}} \exp\left[\frac{i\pi}{4}\operatorname{sign} Q\right] \exp\left[-\frac{i}{2}\left\langle Q^{-1}M\eta, M\eta\right\rangle\right].\end{aligned}$$

Indeed, it suffices to note the following three formulas (a), (b) and (c):

(a) $|\det Q| = |\det D| = \left|\lambda_1\lambda_2\cdots\lambda_q\right|$.
(b) $\sum_{j=1}^{q} \frac{\lambda_j}{|\lambda_j|} = \operatorname{sign} Q$.
(c) $\sum_{j=1}^{q} \frac{\eta_j^2}{\lambda_j} = \left\langle D^{-1}\eta, \eta\right\rangle = \left\langle {}^tMQ^{-1}M\eta, \eta\right\rangle = \left\langle Q^{-1}M\eta, M\eta\right\rangle$.

The desired formula (5.46) follows from (5.51) if we replace $M\eta$ by η.
The proof of Example 5.29 is now complete.

Furthermore, we can calculate explicitly the Fourier transforms of Bessel potentials and Riesz potentials in Examples 5.25 as follows:

Example 5.30. (a) Bessel potentials (see Theorem 6.1):

$$\widehat{G_\alpha}(\xi) = \frac{1}{(1+|\xi|^2)^{\alpha/2}}, \quad \alpha > 0.$$

(b) Riesz kernels:

$$\widehat{R_j}(\xi) = i\,\frac{\xi_j}{|\xi|}, \quad 1 \le j \le n.$$

Finally, as for distributions with compact support we have the following theorem:

Theorem 5.31. *Let $u \in \mathcal{E}'(\mathbf{R}^n)$. Then we have the following two assertions (i) and (ii):*

(i) Its Fourier transform $(\mathcal{F}u)(\xi) = \hat{u}(\xi)$ is a C^∞ function on $\mathbf{R}^n$ given by the formula

$$(\mathcal{F}u)(\xi) = \left\langle u, e^{-ix\cdot\xi}\right\rangle \quad \text{for every } \xi \in \mathbf{R}^n. \qquad (5.52)$$

(ii) The function $\mathcal{F}u(\xi)$ is slowly increasing, that is, there exist constants $C > 0$ and $\mu \in \mathbf{R}$ such that

$$|(\mathcal{F}u)(\xi)| \le C(1+|\xi|)^\mu \quad \text{for all } \xi \in \mathbf{R}^n. \qquad (5.53)$$

Proof. The proof is divided into three steps.

Step 1: For every $\xi \in \mathbf{R}^n$, we let

$$\phi(\xi) = \langle u, e^{-ix\cdot\xi} \rangle .$$

Since $u \in \mathcal{E}'(\mathbf{R}^n)$, by considering difference quotients of ϕ we obtain that $\phi(\xi) \in C^\infty(\mathbf{R}^n)$ with derivatives given by the formula

$$(\partial^\alpha \phi)(\xi) = (-i)^{|\alpha|} \langle u, x^\alpha \, e^{-ix\cdot\xi} \rangle \quad \text{for every multi-index } \alpha.$$

Step 2: Now we prove that $\phi(\xi) = (\mathcal{F}u)(\xi)$ for all $\xi \in \mathbf{R}^n$, that is, formula (5.52). If $\varphi \in C_0^\infty(\mathbf{R}^n)$, then it follows that $u * \varphi \in C_0^\infty(\mathbf{R}^n)$ and further that we have

$$\begin{aligned}
\widehat{u * \varphi}(\xi) &= \int_{\mathbf{R}^n} e^{-ix\cdot\xi} \, u * \varphi(x) \, dx \\
&= \langle u * \varphi, e^{-ix\cdot\xi} \rangle = \langle u_x \otimes \varphi_y, e^{-i(x+y)\cdot\xi} \rangle = \langle u, e^{-ix\cdot\xi} \rangle \langle \varphi, e^{-iy\cdot\xi} \rangle \\
&= \langle u, e^{-ix\cdot\xi} \rangle \, \hat{\varphi}(\xi).
\end{aligned}$$

In particular, by taking

$$\varphi(x) := \rho_\varepsilon(x) = \frac{1}{\varepsilon^n} \rho\left(\frac{x}{\varepsilon}\right), \quad \varepsilon > 0,$$

we obtain that

$$\widehat{u * \rho_\varepsilon}(\xi) = \langle u, e^{-ix\cdot\xi} \rangle \, \widehat{\rho_\varepsilon}(\xi). \tag{5.54}$$

(a) First, by letting $\varepsilon \downarrow 0$ in the left-hand side of (5.54) we obtain that

$$u * \rho_\varepsilon \longrightarrow u \quad \text{in } \mathcal{S}'(\mathbf{R}^n).$$

This proves that

$$\widehat{u * \rho_\varepsilon} \longrightarrow \mathcal{F}u \quad \text{in } \mathcal{S}'(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0, \tag{5.55}$$

since $\mathcal{F} : \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ is continuous.

(b) On the other hand, it follows from an application of the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] that

$$\begin{aligned}
&\widehat{\rho_\varepsilon}(\xi) = \int_{|y|\le 1} \rho(y) e^{-i\varepsilon y\cdot\xi} \, dy \qquad (x = \varepsilon y) \\
&\longrightarrow \int_{|y|\le 1} \rho(y) \, dy = 1 \text{ uniformly in } \xi \text{ over compact subsets of } \mathbf{R}^n \text{ as } \varepsilon \downarrow 0.
\end{aligned}$$

Hence, by letting $\varepsilon \downarrow 0$ in the right-hand side of (5.54) we obtain that

$$\left\langle u, e^{-ix\cdot\xi} \right\rangle \widehat{\rho_\varepsilon}(\xi) \longrightarrow \left\langle u, e^{-ix\cdot\xi} \right\rangle \quad \text{in } \mathcal{D}'(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0. \tag{5.56}$$

Therefore, the desired formula (5.52) follows by combining assertions (5.55) and (5.56).

Step 3: Finally, we prove that the function $(\mathcal{F}u)(\xi)$ is slowly increasing.
Since $u \in \mathcal{E}'(\mathbf{R}^n)$, we can find a compact set K, a non-negative integer μ and a positive constant γ such that

$$p_{K,\mu}(\psi) = \sup_{\substack{x \in K \\ |\alpha| \leq \mu}} |\partial^\alpha \psi(x)| < \gamma \Longrightarrow |\langle u, \psi \rangle| < 1.$$

For all $\varphi \in C^\infty(\mathbf{R}^n)$ and $\lambda > 0$, by letting

$$\psi(x) = \frac{\gamma\, \varphi(x)}{p_{K,\mu}(\varphi) + \lambda}, \quad x \in \mathbf{R}^n,$$

we obtain that

$$p_{K,\mu}(\psi) = \gamma \frac{p_{K,\mu}(\varphi)}{p_{K,\mu}(\varphi) + \lambda} < \gamma,$$

so that

$$|\langle u, \psi \rangle| = \gamma \frac{|\langle u, \varphi \rangle|}{p_{K,\mu}(\varphi) + \lambda} < 1.$$

This proves that

$$|\langle u, \varphi \rangle| \leq \frac{1}{\gamma}\, p_{K,\mu}(\varphi) \quad \text{for all } \varphi \in C^\infty(\mathbf{R}^n), \tag{5.57}$$

since $\lambda > 0$ is arbitrary.

Therefore, the desired inequality (5.53) follows by taking $\varphi(x) = e^{-i\cdot x\xi}$ in (5.57).

The proof of Theorem 5.31 is now complete.

Example 5.32. If δ_{x_0} is Dirac measure at a point x_0 of $\mathbf{R}^n$, then it follows that

$$\widehat{\delta_{x_0}}(\xi) = \left\langle \delta_{x_0}, e^{-ix\cdot\xi} \right\rangle = e^{-ix_0\cdot\xi} \quad \text{for all } \xi \in \mathbf{R}^n,$$

and further that

$$\left|\widehat{\delta_{x_0}}(\xi)\right| = 1 \quad \text{for all } \xi \in \mathbf{R}^n.$$

5.5 Operators and Kernels

Let X and Y be open subset of $\mathbf{R}^n$ and $\mathbf{R}^p$, respectively. In this section we characterize continuous linear operators from $C_0^\infty(Y)$ into $\mathcal{D}'(X)$ in terms of distributions.

Example 5.33. Let $K \in C^\infty(X \times Y)$. If we define a linear operator

$$A : C_0^\infty(Y) \longrightarrow C^\infty(X)$$

by the formula

$$A\psi(x) = \int_Y K(x,y)\psi(y)\,dy \quad \text{for every } \psi \in C_0^\infty(Y),$$

then it follows that A is continuous.

Furthermore, the operator A can be extended to a continuous linear operator

$$\tilde{A} : \mathcal{E}'(Y)_b \longrightarrow C^\infty(X).$$

Proof. (1) Let M be an arbitrary compact subset of Y. If $\psi \in C_0^\infty(Y)$ with $\operatorname{supp}\psi \subset M$, then we have, for any compact subset L of X and any non-negative integer j,

$$\begin{aligned} p_{L,j}(A\psi) &= \sup_{\substack{x\in L\\ |\alpha|\le j}} \left| \int_Y \partial_x^\alpha K(x,y)\psi(y)\,dy \right| \\ &\le \sup_{\substack{x\in L\\ |\alpha|\le j}} \int_M |\partial_x^\alpha K(x,y)|\,dy \cdot \sup_{y\in M} |\psi(y)| \\ &= \sup_{\substack{x\in L\\ |\alpha|\le j}} \int_M |\partial_x^\alpha K(x,y)|\,dy \cdot p_{M,0}(\psi)\,. \end{aligned}$$

This proves the continuity of $A : C_0^\infty(Y) \to C^\infty(X)$.

(2) Furthermore, we can extend A to a continuous linear operator

$$\tilde{A} : \mathcal{E}'(Y)_b \longrightarrow C^\infty(X)$$

as follows: If we let

$$\tilde{A}v(x) = \langle v, K(x,\cdot)\rangle \quad \text{for every } v \in \mathcal{E}'(Y),$$

then it follows that $\tilde{A}v \in C^\infty(X)$, since $\operatorname{supp} v$ is compact. More precisely, we have, for any compact subset H of X and any non-negative integer m,

$$p_{H,m}(\tilde{A}v) = \sup_{\substack{x\in H \\ |\alpha|\le m}} \left|\partial^{\alpha}\left(\tilde{A}v\right)\right| = \sup_{\substack{x\in H \\ |\alpha|\le m}} |\langle v, \partial_x^{\alpha} K(x,\cdot)\rangle| ,$$

where the functions

$$\{\partial_x^{\alpha} K(x,\cdot)\}_{\substack{x\in H \\ |\alpha|\le m}}$$

form a bounded subset of $C^{\infty}(Y)$. However, we recall that a sequence $\{v_j\}$ of distributions converges to a distribution v in the space $\mathcal{E}'(Y)_b$ if and only if the sequence $\{\langle v_j, \psi\rangle\}$ converges to $\langle v, \psi\rangle$ uniformly in ψ over all bounded subsets of $C_0^{\infty}(Y)$.

Therefore, we obtain the continuity of $\tilde{A} : \mathcal{E}'(Y)_b \to C^{\infty}(X)$.

The proof of Example 5.33 is complete.

More generally, we have the following example:

Example 5.34. If K is a distribution in $\mathcal{D}'(X \times Y)$, we can define a continuous linear operator

$$A \in \mathcal{L}\left(C_0^{\infty}(Y), \mathcal{D}'(X)\right)$$

by the formula

$$\langle A\psi, \varphi\rangle = \langle K, \varphi \otimes \psi\rangle \quad \text{for all } \varphi \in C_0^{\infty}(X) \text{ and } \psi \in C_0^{\infty}(Y).$$

We then write $A = \mathrm{Op}(K)$.

Proof. (1) First, we show that $A\psi \in \mathcal{D}'(X)$. Assume that $\varphi_j \to 0$ in $C_0^{\infty}(X)$ as $j \to \infty$. Then it follows that $\varphi_j \otimes \psi \to 0$ in $C_0^{\infty}(X\times Y)$ for every $\psi \in C_0^{\infty}(Y)$. Hence we have

$$\left\langle A\psi, \varphi_j\right\rangle = \left\langle K, \varphi_j \otimes \psi\right\rangle \longrightarrow 0 \quad \text{as } j \to \infty.$$

This proves that $A\psi \in \mathcal{D}'(X)$.

(2) Secondly, we show that $A : C_0^{\infty}(Y) \to \mathcal{D}'(X)$ is continuous. Assume that $\psi_j \to 0$ in $C_0^{\infty}(Y)$ as $j \to \infty$. Then it follows that $\varphi \otimes \psi_j \to 0$ in $C_0^{\infty}(X\times Y)$ for every $\varphi \in C_0^{\infty}(X)$. Therefore, we have

$$\left\langle A\psi_j, \varphi\right\rangle = \left\langle K, \varphi \otimes \psi_j\right\rangle \longrightarrow 0 \quad \text{as } j \to \infty,$$

and hence

$$A\psi_j \longrightarrow 0 \quad \text{in } \mathcal{D}'(X).$$

This proves the continuity of $A : C_0^{\infty}(Y) \to \mathcal{D}'(X)$.

The proof of Example 5.34 is complete.

We give a simple example of an operator $\mathrm{Op}\,(K)$:

Example 5.35. Let $D = \{(x, x) : x \in X\}$ be the diagonal in the product space $X \times X$, and define a distribution $\delta_D \in \mathcal{D}'(X \times X)$ by the formula

$$\langle \delta_D, \Phi \rangle = \int_X \Phi(x, x)\, dx \quad \text{for every } \Phi \in C_0^\infty(X \times X).$$

Then it follows that $\mathrm{Op}\,(\delta_D) = I$.

Proof. Indeed, it suffices to note that we have, for all $\varphi, \psi \in C_0^\infty(X)$,

$$\langle \mathrm{Op}\,(\delta_D)\psi, \varphi \rangle = \langle \delta_D, \varphi \otimes \psi \rangle = \int_X \varphi(x)\psi(x)\, dx = \langle \psi, \varphi \rangle\,.$$

This proves that $\mathrm{Op}\,(\delta_D)\psi = \psi$ for all $\psi \in C_0^\infty(X)$, that is, $\mathrm{Op}\,(\delta_D) = I$.

The proof of Example 5.35 is complete.

5.5.1 Schwartz's Kernel Theorem

By Lemma 5.16, we know that the space $C_0^\infty(X) \otimes C_0^\infty(Y)$ is sequentially dense in $C_0^\infty(X \times Y)$. Hence it follows that the mapping

$$\mathcal{D}'(X \times Y) \ni K \longmapsto \mathrm{Op}(K) \in \mathcal{L}(C_0^\infty(Y), \mathcal{D}'(X))$$

is injective. The next theorem asserts that it is also *surjective*:

Theorem 5.36 (Schwartz's kernel theorem). *If A is a continuous linear operator from $C_0^\infty(Y)$ into $\mathcal{D}'(Y)$, then there exists a unique distribution $K \in \mathcal{D}'(X \times Y)$ such that* $A = \mathrm{Op}(K)$.

Proof. The proof is divided into three steps.

Step 1: The bilinear form

$$C_0^\infty(X) \times C_0^\infty(Y) \ni (\varphi, \psi) \longrightarrow \langle A\psi, \varphi \rangle = {}_{\mathcal{D}'(X)}\langle A\psi, \varphi \rangle_{C_0^\infty(X)} \in \mathbf{C}$$

induces a linear form K on the algebraic tensor product $C_0^\infty(X) \otimes C_0^\infty(Y)$. More precisely, if $\Phi \in C_0^\infty(X \times Y)$ is of the form

$$\Phi = \sum_{j=1}^{N} \varphi_j \otimes \psi_j, \quad \varphi_j \in C_0^\infty(X),\ \psi_j \in C_0^\infty(Y), \tag{5.58}$$

then we define the linear form $\langle K, \Phi \rangle$ by the formula

$$\langle K, \Phi \rangle = \sum_{j=1}^{N} \langle A\psi_j, \varphi_j \rangle. \tag{5.59}$$

It should be emphasized that the right-hand side of (5.59) does not depend on the decomposition (5.58) of Φ.

Step 2: By virtue of Lemma 5.16, it suffices to show that K is a linear continuous form on $C_0^\infty(X) \otimes C_0^\infty(Y)$ for the topology induced by $C_0^\infty(X \times Y)$. In order to show that $K \in \mathcal{D}'(X \times Y)$, we need the following elementary lemma from functional analysis:

Lemma 5.37. *Let E, F be Fréchet spaces, and let $B : E \times F \to \mathbf{C}$ be a bilinear form. Assume that the linear functional $E \ni x \mapsto B(x, y)$ is continuous for each $y \in F$ and further that the linear functional $F \ni y \mapsto B(x, y)$ is continuous for each $x \in E$. Then it follows that the bilinear form $B : E \times F \to \mathbf{C}$ is continuous.*

Proof. (i) First, we show that the bilinear form $B : E \times F \to \mathbf{C}$ is *separately continuous*, that is, we prove the following two assertions (a) and (b):

(a) If $x_j \to 0$ in E, then $B(x_j, y) \to 0$ uniformly for y in any bounded set of F.
(b) If $y_k \to 0$ in F, then $B(x, y_k) \to 0$ uniformly for x in any bounded set of E.

Proof of Assertion (a): Since the linear functional $E \ni x \mapsto B(x, y)$ is continuous for each $y \in F$, it follows that $E \ni x \mapsto |B(x, y)|$ is a continuous seminorm for each $y \in F$. Similarly, it follows that $F \ni y \mapsto |B(x, y)|$ is a continuous seminorm for each $x \in E$. Hence, if M is a bounded subset of F, we have, for each $x \in E$,

$$\sup_{y \in M} |B(x, y)| < \infty. \tag{5.60}$$

This implies that the set $\{B(\cdot, y)\}_{y \in M}$ is bounded in the dual space $E' = \mathcal{L}(E, \mathbf{C})$ of E with respect to the simple topology. By applying the Banach–Steinhaus theorem (Theorem 3.4) with

$$X := E, \quad Y := \mathbf{C}, \quad H := \{B(\cdot, y)\}_{y \in M},$$

we obtain that the set $\{B(\cdot, y)\}_{y \in M}$ is equicontinuous in E'. Hence there exists a neighborhood U of 0 such that

$$\sup_{\substack{y \in M \\ x \in U}} |B(x, y)| \leq 1. \tag{5.61}$$

Let A be an arbitrary compact subset of E. Then we can find a finite set of points $\{x_1, x_2, \ldots, x_p\}$ in A such that

$$A \subset \bigcup_{i=1}^{p} \{x_i + U\}.$$

Thus we have, by assertions (5.60) and (5.61),

$$\sup_{y \in M} |B(x, y)| \leq \max_{1 \leq i \leq p} \left(\sup_{y \in M} |B(x_i, y)| \right) + 1 \quad \text{for all } x \in A.$$

This implies that the seminorm

$$\rho(x) = \sup_{y \in M} |B(x, y)|, \quad x \in E,$$

is bounded in the compact convergence topology of E. By applying again the Banach–Steinhaus theorem, we obtain that the seminorm $\rho(x)$ is continuous on E. Therefore, we have the desired assertion

$$x_j \longrightarrow 0 \quad \text{in } E \Longrightarrow \rho(x_j) = \sup_{y \in M} |B(x_j, y)| \longrightarrow 0.$$

Assertion (b) can be proved similarly.

(ii) Now we assume that

$$x_j \longrightarrow x \quad \text{in } E, \quad y_k \longrightarrow y \quad \text{in } F.$$

Then it follows that the sequence $\{y_k\}$ is bounded in F. Hence we have, by assertion (b),

$$B(x_j, y_k) - B(x, y_k) = B(x_j - x, y_k) \longrightarrow 0,$$

and also

$$B(x, y_k) - B(x, y) = B(x, y_k - y) \longrightarrow 0.$$

Therefore, we obtain that

$$\left|B(x_j, y_k) - B(x, y)\right| \leq \left|B(x_j, y_k) - B(x, y_k)\right| + |B(x, y_k) - B(x, y)| \longrightarrow 0.$$

This proves the continuity of B.

The proof of Lemma 5.37 is complete.

Step 3: We show that, for given compact subsets $H \subset X$ and $L \subset Y$, respectively, there exist a non-negative integer m and a positive constant C such that we have, for all $\Phi \in C_0^\infty(X) \otimes C_0^\infty(Y)$ with $\operatorname{supp} \Phi \subset H \times L$,

$$|\langle K, \Phi \rangle| \leq C \, p_{H \times L, m}(\Phi), \tag{5.62}$$

where

$$p_{H \times L, m}(\Phi) = \sup_{\substack{(x,y) \in H \times L \\ |\alpha| \leq m}} \left| \partial_{x,y}^\alpha \Phi(x, y) \right|.$$

(a) If $\Phi \in C_0^\infty(X) \otimes C_0^\infty(Y)$ with $\operatorname{supp} \Phi \subset H \times L$, we can find compact neighborhoods H_1 of H and L_1 of L such that

$$\Phi = \sum_{j=1}^N \varphi_j \otimes \psi_j, \quad \varphi_j \in \mathcal{D}_{H_1}(X), \ \ \psi_j \in \mathcal{D}_{L_1}(Y),$$

where

$$\mathcal{D}_{H_1}(X) = \{\varphi \in C_0^\infty(X) : \operatorname{supp} \varphi \subset H_1\},$$
$$\mathcal{D}_{L_1}(Y) = \{\psi \in C_0^\infty(Y) : \operatorname{supp} \psi \subset L_1\}.$$

Furthermore, we take functions $f \in C_0^\infty(X)$ and $g \in C_0^\infty(Y)$ such that

$$f = 1 \quad \text{near } H_1, \quad \operatorname{supp} f = H_2 \subset X,$$
$$g = 1 \quad \text{near } L_1, \quad \operatorname{supp} g = L_2 \subset Y.$$

Then we have, by the Fourier inversion formula,

$$\varphi_j(x) = f(x)\varphi_j(y) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \left(f(x) e^{ix \cdot \xi} \right) \hat{\varphi}_j(\xi) \, d\xi,$$
$$\psi_j(y) = g(y)\psi_j(y) = \frac{1}{(2\pi)^p} \int_{\mathbf{R}^p} \left(g(y) e^{iy \cdot \eta} \right) \hat{\psi}_j(\eta) \, d\eta.$$

We remark that the second integral converges in the topology of $\mathcal{D}_{L_2}(Y)$. Since $A : C_0^\infty(Y) \to \mathcal{D}'(X)$ is linear and continuous, it follows that

$$A\psi_j = \frac{1}{(2\pi)^p} \int_{\mathbf{R}^p} A \left(g(y) e^{iy \cdot \eta} \right) \hat{\psi}_j(\eta) \, d\eta \quad \text{in } \mathcal{D}'(X).$$

Therefore, we obtain the formula

$$\begin{aligned}
&\langle A\psi_j, \varphi_j\rangle \qquad (5.63)\\
&= \left\langle \frac{1}{(2\pi)^p} \int_{\mathbf{R}^n} A\left(g(y)e^{iy\cdot\eta}\right) \hat{\psi}_j(\eta)\, d\eta, \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \left(f(x)e^{ix\cdot\xi}\right) \hat{\varphi}_j(\xi)\, d\xi \right\rangle\\
&= \frac{1}{(2\pi)^p} \frac{1}{(2\pi)^n} \int_{\mathbf{R}^p} \left(\int_{\mathbf{R}^n} F(\xi,\eta)\, \hat{\varphi}_j(\xi)\, d\xi\right) \hat{\psi}_j(\eta)\, d\eta,
\end{aligned}$$

where

$$F(\xi,\eta) = {}_{\mathcal{D}'(X)} < A\left(g(y)e^{iy\cdot\eta}\right),\ f(x)e^{ix\cdot\xi} >_{C_0^\infty(X)} .$$

(b) We show that $F(\xi,\eta) \in C\left(\mathbf{R}^n \times \mathbf{R}^p\right)$ and further that there exist a positive constant C' and non-negative integers k and ℓ such that

$$|F(\xi,\eta)| \leq C' (1 + |\xi|)^k (1 + |\eta|)^\ell \quad \text{for all } (\xi,\eta) \in \mathbf{R}^n \times \mathbf{R}^p. \tag{5.64}$$

Since $A : C_0^\infty(Y) \to \mathcal{D}'(X)$ is continuous, we find that the bilinear form

$$B : \mathcal{D}_{H_2}(X) \times \mathcal{D}_{L_2}(Y) \ni (\varphi,\psi) \longrightarrow \langle A\psi, \varphi\rangle = {}_{\mathcal{D}'(X)} \langle A\psi, \varphi\rangle_{C_0^\infty(X)} \in \mathbf{C}$$

is continuous for each φ and for each ψ, that is, we have the following two assertions (1) and (2):

(1) If $\varphi_j \to 0$ in $\mathcal{D}_{H_2}(X)$, then it follows that $B(\varphi_j, \psi) \to 0$ for each $\psi \in \mathcal{D}_{L_2}(Y)$.
(2) If $\psi_k \to 0$ in $\mathcal{D}_{L_2}(Y)$, then it follows that $B(\varphi, \psi_k) \to 0$ for each $\varphi \in \mathcal{D}_{H_2}(X)$.

Hence, by applying Lemma 5.37 to our situation we obtain that the bilinear form $B : \mathcal{D}_{H_2}(X) \times \mathcal{D}_{L_2}(Y) \to \mathbf{C}$ is continuous. In particular, it follows that $F(\xi,\eta)$ is a continuous function of $(\xi,\eta) \in \mathbf{R}^n \times \mathbf{R}^p$ if we take

$$\varphi(x) := f(x)e^{ix\cdot\xi}, \quad \psi(y) := g(y)e^{iy\cdot\eta}.$$

Furthermore, since the mapping $(\varphi,\psi) \mapsto \langle A\psi, \varphi\rangle$ is continuous, there exist non-negative integers k and ℓ and a positive constant C such that

$$|\langle A\psi, \varphi\rangle| \leq C\, p_{H_2,k}(\varphi)\, p_{L_2,\ell}(\psi) \quad \text{for all } (\varphi,\psi) \in \mathcal{D}_{H_2}(X) \times \mathcal{D}_{L_2}(Y). \tag{5.65}$$

Therefore, the desired inequality (5.64) follows from inequality (5.65) by taking

$$\varphi(x) := f(x)e^{ix\cdot\xi}, \quad \psi(y) := g(y)e^{iy\cdot\eta}.$$

(c) By combining (5.59) and (5.63), we obtain that

$$\begin{aligned}\langle K, \Phi\rangle &= \sum_{j=1}^{N} \langle A\psi_j, \varphi_j\rangle \\ &= \frac{1}{(2\pi)^p}\frac{1}{(2\pi)^n}\sum_{j=1}^{N}\int_{\mathbf{R}^p}\left(\int_{\mathbf{R}^n} F(\xi,\eta)\,\hat{\varphi}_j(\xi)\,d\xi\right)\hat{\psi}_j(\eta)\,d\eta.\end{aligned}$$

Hence, by virtue of Fubini's theorem it follows from (5.58) that

$$\begin{aligned}\langle K, \Phi\rangle &= \frac{1}{(2\pi)^p}\frac{1}{(2\pi)^n}\sum_{j=1}^{N}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} F(\xi,\eta)\,\hat{\varphi}_j(\xi)\hat{\psi}_j(\eta)\,d\xi\,d\eta \qquad (5.66)\\ &= \frac{1}{(2\pi)^p}\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} F(\xi,\eta)\left(\sum_{j=1}^{N}\hat{\varphi}_j(\xi)\hat{\psi}_j(\eta)\right)d\xi\,d\eta \\ &= \frac{1}{(2\pi)^p}\frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} F(\xi,\eta)\,\hat{\Phi}(\xi,\eta)\,d\xi\,d\eta.\end{aligned}$$

However, if $\Phi \in C_0^\infty(X) \otimes C_0^\infty(Y)$ with $\operatorname{supp}\Phi \subset H \times L$, then we have, by integration by parts,

$$\begin{aligned}&\xi^\alpha\,\eta^\beta\,\hat{\Phi}(\xi,\eta)\\ &= \int_H\int_L e^{-ix\xi}\,e^{-iy\eta}\left(D_x^\alpha D_y^\beta \Phi(x,y)\right)dx\,dy \quad \text{for all multi-indices } \alpha \text{ and } \beta.\end{aligned}$$

Therefore, for any positive integer m we can find a positive constant C'' such that

$$(1+|\xi|+|\eta|)^m\left|\hat{\Phi}(\xi,\eta)\right| \leq C''\,p_{H\times L,m}(\Phi) \quad \text{for all } (\xi,\eta)\in \mathbf{R}^n\times\mathbf{R}^p. \tag{5.67}$$

Now, if we choose positive numbers m_1 and m_2 such that

$$m_1 > k+n, \quad m_2 > \ell + p,$$

and let

$$m := m_1 + m_2$$

in inequality (5.67), then we obtain from formula (5.66) and inequality (5.64) that

$$
\begin{aligned}
&|\langle K, \Phi\rangle| \\
&\leq \frac{1}{(2\pi)^p} \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} |F(\xi, \eta)| \left|\hat{\Phi}(\xi, \eta)\right| d\xi \, d\eta \\
&\leq \frac{C'}{(2\pi)^p} \frac{C''}{(2\pi)^n} \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} \frac{(1+|\xi|)^k (1+|\eta|)^\ell}{(1+|\xi|+|\eta|)^m} d\xi \, d\eta \cdot p_{H\times L,m}(\Phi) \\
&\leq \frac{C'}{(2\pi)^p} \frac{C''}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{(1+|\xi|)^k}{(1+|\xi|)^{m_1}} d\xi \cdot \int_{\mathbf{R}^n} \frac{(1+|\eta|)^\ell}{(1+|\eta|)^{m_2}} d\eta \cdot p_{H\times L,m}(\Phi) \\
&= \frac{C'}{(2\pi)^p} \frac{C''}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{1}{(1+|\xi|)^{m_1-k}} d\xi \cdot \int_{\mathbf{R}^n} \frac{1}{(1+|\eta|)^{m_2-\ell}} d\eta \cdot p_{H\times L,m}(\Phi).
\end{aligned}
$$

This proves the desired inequality (5.66) with

$$
C := \frac{C'}{(2\pi)^p} \frac{C''}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{1}{(1+|\xi|)^{m_1-k}} d\xi \cdot \int_{\mathbf{R}^n} \frac{1}{(1+|\eta|)^{m_2-\ell}} d\eta.
$$

Now the proof of Theorem 5.36 is complete.

The distribution K is called the *kernel* of A. Formally, we have

$$
(A\psi)(x) = \int_Y K(x, y)\psi(y)\, dy \quad \text{for all } \psi \in C_0^\infty(Y).
$$

Now we give some important examples of distribution kernels:

Example 5.38. (a) Riesz potentials: $X = Y = \mathbf{R}^n$, $0 < \alpha < n$.

$$
\begin{aligned}
(-\Delta)^{-\alpha/2} u(x) &= R_\alpha * u(x) \\
&= \frac{\Gamma((n-\alpha)/2)}{2^\alpha \pi^{n/2} \Gamma(\alpha/2)} \int_{\mathbf{R}^n} \frac{1}{|x-y|^{n-\alpha}} u(y)\, dy, \quad u \in C_0^\infty(\mathbf{R}^n).
\end{aligned}
$$

(b) Newtonian potentials: $X = Y = \mathbf{R}^n$, $n \geq 3$.

$$
\begin{aligned}
(-\Delta)^{-1} u(x) &= N * u(x) \\
&= \frac{\Gamma((n-2)/2)}{4\pi^{n/2}} \int_{\mathbf{R}^n} \frac{1}{|x-y|^{n-2}} u(y)\, dy, \quad u \in C_0^\infty(\mathbf{R}^n).
\end{aligned}
$$

(c) Bessel potentials: $X = Y = \mathbf{R}^n$, $\alpha > 0$.

$$
(I-\Delta)^{-\alpha/2} u(x) = G_\alpha * u(x) = \int_{\mathbf{R}^n} G_\alpha(x-y)\, u(y)\, dy, \quad u \in C_0^\infty(\mathbf{R}^n).
$$

(d) Riesz operators: $X = Y = \mathbf{R}^n$, $1 \leq j \leq n$.

$$\begin{aligned} Y_j u(x) &= R_j * u(x) \\ &= i\, \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \text{v. p.} \int_{\mathbf{R}^n} \frac{x_j - y_j}{|x-y|^{n+1}}\, u(y)\, dy, \quad u \in C_0^\infty(\mathbf{R}^n). \end{aligned}$$

(e) The Calderón–Zygmund integro-differential operator: $X = Y = \mathbf{R}^n$.

$$\begin{aligned} (-\Delta)^{1/2} u(x) &= \frac{1}{\sqrt{-1}} \sum_{j=1}^n Y_j \left(\frac{\partial u}{\partial x_j} \right)(x) \\ &= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \sum_{j=1}^n \text{v. p.} \int_{\mathbf{R}^n} \frac{x_j - y_j}{|x-y|^{n+1}} \frac{\partial u}{\partial y_j}(y)\, dy, \\ &\qquad u \in C_0^\infty(\mathbf{R}^n). \end{aligned}$$

If $A : C_0^\infty(Y) \to \mathcal{D}'(X)$ is a continuous linear operator, we define its *transpose* A' by the formula

$$\langle A'\varphi, \psi \rangle = \langle \varphi, A\psi \rangle \quad \text{for all } \varphi \in C_0^\infty(X) \text{ and } \psi \in C_0^\infty(Y).$$

Then the transpose A' is a continuous linear operator on $C_0^\infty(X)$ into $\mathcal{D}'(Y)$. The distribution kernel of A' is obtained from the distribution kernel $K(x, y)$ of A by interchanging the roles of x and y; formally this means that

$$(A'\varphi)(y) = \int_X K(y, x) \varphi(x)\, dx \quad \text{for all } \varphi \in C_0^\infty(X).$$

Also we have $(A')' = A$.

Similarly, we define the *adjoint* A^* of A by the formula

$$\langle A^*\varphi, \overline{\psi} \rangle = \langle \varphi, \overline{A\psi} \rangle \quad \text{for all } \varphi \in C_0^\infty(X) \text{ and } \psi \in C_0^\infty(Y).$$

Then the adjoint A^* is a continuous linear operator on $C_0^\infty(X)$ into $\mathcal{D}'(Y)$. The distribution kernel of A^* is obtained from the distribution kernel $\overline{K(x, y)}$ by interchanging the roles of x and y; formally this means that

$$(A^*\varphi)(y) = \int_X \overline{K(y, x)}\, \varphi(x)\, dx \quad \text{for all } \varphi \in C_0^\infty(X).$$

We also have $(A^*)^* = \text{A}$.

Example 5.39. If $X = Y$ is an open subset Ω of $\mathbf{R}^n$ and if

$$A = \sum_{|\alpha| \le m} a_\alpha(x)\, D^\alpha$$

is a differential operator of order m with C^∞ coefficients in Ω, then we have

$$A' = \sum_{|\alpha|\le m} (-1)^{|\alpha|} D^\alpha \left(a_\alpha(x)\cdot\right),$$

$$A^* = \sum_{|\alpha|\le m} (-1)^{|\alpha|} D^\alpha \left(\overline{a_\alpha(x)\cdot}\right).$$

This shows that both A' and A^* are differential operators of order m with C^∞ coefficients in Ω.

5.5.2 Regularizers

Let X and Y be open subsets of $\mathbf{R}^n$ and $\mathbf{R}^p$, respectively. A continuous linear operator $A : C_0^\infty(Y) \to \mathcal{D}'(X)$ is called a *regularizer* if it extends to a continuous linear operator from $\mathcal{E}'(Y)$ into $C^\infty(X)$.

The next theorem gives a characterization of regularizers in terms of distribution kernels:

Theorem 5.40. *A continuous linear operator $A : C_0^\infty(Y) \to \mathcal{D}'(X)$ is a regularizer if and only if its distribution kernel $K(x, y)$ is in $C^\infty(X \times Y)$.*

Proof. The proof is divided into two steps.

Step 1: First, we prove the "if" part. If $A = \operatorname{Op}(K)$ with $K \in C^\infty(X \times Y)$, then it follows from Example 5.33 that A extends to a continuous linear operator $\tilde{A} : \mathcal{E}'(Y) \to C^\infty(X)$.

Step 2: Secondly, we prove the "only if" part. The proof of Step 2 is divided into four steps.

Step 2-a: We assume that $A = \operatorname{Op}(K)$ extends to a continuous linear operator $\tilde{A} : \mathcal{E}'(Y) \to C^\infty(X)$. First, by letting

$$a(x, y) = \left(\tilde{A}\delta_y\right)(x) \quad \text{for every } (x, y) \in X \times Y,$$

we show that $a \in C(X \times Y)$. Here we observe that $\tilde{A}\delta_y \in C^\infty(X)$. Now let (x_0, y_0) be an arbitrary point of $X \times Y$. Then it follows that

$$\begin{aligned} &|a(x, y) - a(x_0, y_0)| \\ &= \left|\left(\tilde{A}\delta_y\right)(x) - \left(\tilde{A}\delta_{y_0}\right)(x_0)\right| \\ &\le \left|\left(\tilde{A}\delta_y\right)(x) - \left(\tilde{A}\delta_{y_0}\right)(x)\right| + \left|\left(\tilde{A}\delta_{y_0}\right)(x) - \left(\tilde{A}\delta_{y_0}\right)(x_0)\right|. \end{aligned} \tag{5.68}$$

However, since $A\delta_{y_0} \in C^\infty(X)$, for any given $\varepsilon > 0$ we can find a constant $\eta' = \eta'(x_0, \varepsilon) > 0$ such that

$$|x - x_0| \leq \eta' \Longrightarrow \left|\left(\tilde{A}\delta_{y_0}\right)(x) - \left(\tilde{A}\delta_{y_0}\right)(x_0)\right| \leq \frac{\varepsilon}{2}. \tag{5.69}$$

Moreover, since $A : \mathcal{E}'(Y) \to C^\infty(X)$ is continuous, for any given $\varepsilon > 0$ we can find a constant $\eta'' = \eta''(x_0, \varepsilon, \eta') > 0$ such that

$$\begin{aligned} |y - y_0| \leq \eta'' \quad \Longrightarrow \quad & \sup_{|x-x_0|\leq\eta'} \left|\left(\tilde{A}\delta_y\right)(x) - \left(\tilde{A}\delta_{y_0}\right)(x)\right| \qquad (5.70) \\ & = p_{\{|x-x_0|\leq\eta'\},0}\left(\tilde{A}\delta_y - \tilde{A}\delta_{y_0}\right) \leq \frac{\varepsilon}{2}. \end{aligned}$$

Indeed, it suffices to note that $y \to y_0$ in Y implies that $\delta_y \to \delta_{y_0}$ in $\mathcal{E}'(Y)$. Therefore, by combining inequalities (5.68)–(5.70) we obtain that

$$\sup_{\substack{|x-x_0|\leq\eta' \\ |y-y_0|\leq\eta''}} |a(x, y) - a(x_0, y_0)| \leq \varepsilon.$$

This proves that $a(x, y) = \left(\tilde{A}\delta_y\right)(x) \in C(X \times Y)$.

Step 2-b: Secondly, we show that $a \in C^\infty(X \times Y)$. Since $\tilde{A} : \mathcal{E}'(Y) \to C^\infty(X)$ is continuous, it is easy to verify that we have, for all multi-indices α and β,

$$\partial_x^\alpha \partial_y^\beta a(x, y) = (-1)^{|\beta|} \partial_x^\alpha \left(\tilde{A}\delta_y^{(\beta)}\right)(x) \quad \text{in } \mathcal{D}'(X \times Y). \tag{5.71}$$

However, by arguing just as in Step 2-a we find that

$$\partial_x^\alpha \partial_y^\beta a(x, y) = (-1)^{|\beta|} \partial_x^\alpha \left(\tilde{A}\delta_y^{(\beta)}\right)(x) \in C(X \times Y).$$

This proves that $a(x, y) = \left(\tilde{A}\delta_y\right)(x) \in C^\infty(X \times Y)$.

Step 2-c: Finally, we show that $a(x, y) = K(x, y)$. To do this, we need the following lemma:

Lemma 5.41. *The linear combinations of distributions of the form* $\delta_y^{(\beta)}$ *are* dense *in the space* $\mathcal{E}'(Y)$.

Proof. Let v be an arbitrary distribution of $\mathcal{E}'(Y)$. By Theorem 5.24, we can find a compact neighborhood V of $\operatorname{supp} v$ and a non-negative integer m such that

$$v = \partial^{m+2} G,$$

where

$$G \in C(Y), \quad \operatorname{supp} G \subset V.$$

For every function $\psi \in C^\infty(Y)$, we have

$$\begin{aligned}\langle v, \psi\rangle &= \left\langle \partial^{m+2} G, \psi\right\rangle = (-1)^{m+2}\left\langle G, \partial^{m+2}\psi\right\rangle \\ &= (-1)^{m+2}\int_Y G(y)\partial^{m+2}\psi(y)\,dy.\end{aligned}$$

The integrand $G(y)\partial^{m+2}\psi(y)$ is continuous and supported in the compact subset V, so the integral can be approximated by Riemann sums. More precisely, for each large number $N \in \mathbf{N}$ we can approximate $\operatorname{supp} G$ by a union of cubes of side length 2^{-N} and volume 2^{-nN} centered at points y_1^N, $y_2^N, \ldots, y_{k(N)}^N \in \operatorname{supp} G$. Then we find that the corresponding Riemann sums

$$S^N = \frac{(-1)^{m+2}}{2^{nN}}\sum_{j=1}^{k(N)} G(y_j^N)\partial^{m+2}\psi(y_j^N)$$

are supported in the common compact subset V, and converge uniformly to $\langle v, \psi\rangle$ as $N \to \infty$. Hence we have, for every $\psi \in C^\infty(Y)$,

$$\langle v, \psi\rangle = \lim_{N\to\infty}\left\langle S^N, \psi\right\rangle = \lim_{N\to\infty}\left\langle \frac{1}{2^{nN}}\sum_{j=1}^{k(N)} G(y_j^N)\delta_{y_j^N}^{(m+2)}, \psi\right\rangle.$$

This proves that

$$\frac{1}{2^{nN}}\sum_{j=1}^{k(N)} G(y_j^N)\delta_{y_j^N}^{(m+2)} \longrightarrow v \quad \text{in } \mathcal{E}'(Y) \text{ as } N \to \infty.$$

The proof of Lemma 5.41 is complete.

Step 2-d: If $A_1 = \operatorname{Op}(a)$, it follows from Step 1 that $A_1 : C_0^\infty(Y) \to \mathcal{D}'(X)$ is a regularizer. That is, it extends to a continuous linear operator $\tilde{A}_1 : \mathcal{E}'(Y) \to C^\infty(X)$. Hence we have, by formula (5.71),

$$\tilde{A}_1\left(\delta_y^{(\beta)}\right) = \left\langle a(\cdot, y), \delta_y^{(\beta)}\right\rangle = (-1)^{|\beta|}\partial_y^\beta a(\cdot, y) = \tilde{A}\left(\delta_y^{(\beta)}\right).$$

However, by Lemma 5.41 it follows that the linear combinations of distributions of type $\delta_y^{(\beta)}$ are dense in the space $\mathcal{E}'(Y)$.
Therefore, we obtain from the continuity of $\tilde{A}_1$ and $\tilde{A}$ that

$$\operatorname{Op}(a) = \tilde{A}_1 = \tilde{A} = \operatorname{Op}(K) \quad \text{on } \mathcal{E}'(Y).$$

By the uniqueness of kernels, this implies that $a(x, y) = K(x, y)$.

Summing up, we have proved that $A = \operatorname{Op}(K)$ with $K \in C^\infty(X \times Y)$.
Now the proof of Theorem 5.40 is complete.

5.6 Layer Potentials

The purpose of this section is to describe the classical layer potentials arising in the Dirichlet and Neumann problems for the Laplacian Δ in the case of the half-space $\mathbf{R}^n_+$.

5.6.1 Single and Double Layer Potentials

Recall that the Newtonian potential is defined by the formula (see Example 5.25)

$$\begin{aligned}(-\Delta)^{-1} f(x) &= N * f(x) \\ &= \frac{\Gamma((n-2)/2)}{4\pi^{n/2}} \int_{\mathbf{R}^n} \frac{1}{|x-y|^{n-2}} f(y)\, dy \\ &= \frac{1}{(n-2)\omega_n} \int_{\mathbf{R}^n} \frac{1}{|x-y|^{n-2}} f(y)\, dy, \quad f \in C_0^\infty(\mathbf{R}^n).\end{aligned} \tag{5.72}$$

Here

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad n \geq 2,$$

is the surface area of the unit sphere Σ_n. In the case $n = 3$, we have the classical Newtonian potential

$$u(x) = \frac{1}{4\pi} \int_{\mathbf{R}^3} \frac{f(y)}{|x-y|}\, dy.$$

Up to an appropriate constant of proportionality, the Newtonian potential

$$\frac{1}{4\pi} \frac{1}{|x-y|}$$

is the gravitational potential at position x due to a unit point mass at position y, and so the function $u(x)$ is the gravitational potential due to a continuous mass distribution with density $f(x)$. In terms of electrostatics, the function $u(x)$ describes the electrostatic potential due to a charge distribution with density $f(x)$.

We define a *single layer potential* with density φ by the formula

$$\begin{aligned}&N * (\varphi(x') \otimes \delta(x_n)) \\ &= \frac{1}{(n-2)\omega_n} \int_{\mathbf{R}^{n-1}} \frac{\varphi(y')}{(|x'-y'|^2 + x_n^2)^{(n-2)/2}}\, dy', \quad \varphi \in C_0^\infty(\mathbf{R}^{n-1}).\end{aligned} \tag{5.73}$$

In the case $n = 3$, the function $N * (\varphi \otimes \delta)$ is related to the distribution of electric charge on a conductor Ω. In equilibrium, mutual repulsion causes all the charge to reside on the surface $\partial\Omega$ of the conducting body with density φ, and $\partial\Omega$ is an equipotential surface.

We define a *double layer potential* with density ψ by the formula

$$\begin{aligned} & N * (\psi(x') \otimes \delta^{(\prime)}(x_n)) \\ &= \frac{1}{\omega_n} \int_{\mathbf{R}^{n-1}} \frac{x_n \psi(y')}{(|x' - y'|^2 + x_n^2)^{n/2}} \, dy', \quad \psi \in C_0^\infty(\mathbf{R}^{n-1}). \end{aligned} \tag{5.74}$$

In the case $n = 3$, the function $N*(\psi\otimes\delta^{(\prime)})$ is the potential induced by a distribution of dipoles on $\mathbf{R}^2$ with density $\psi(y')$, the axes of the dipoles being normal to $\mathbf{R}^2$.

5.6.2 *The Green Representation Formula*

By applying the Newtonian potential to both sides of the jump formula (5.13), we obtain that

$$\begin{aligned} u^0 &= (N * (-\Delta))(u^0) \\ &= N * ((-\Delta u)^0) - N * (\gamma_1 u \otimes \delta(x_n)) - N * (\gamma_0 u \otimes \delta^{(\prime)}(x_n)) \\ &= - \int_{\mathbf{R}^n} N(x-y)\Delta u(y)\, dy - \int_{\mathbf{R}^{n-1}} N(x'-y', x_n) \frac{\partial u}{\partial y_n}(y', 0)\, dy' \\ &\quad + \int_{\mathbf{R}^{n-1}} \frac{\partial N}{\partial y_n}(x'-y', x_n) u(y', 0)\, dy. \end{aligned}$$

Hence we arrive at the *Green representation formula*:

$$\begin{aligned} u(x) &= \frac{1}{(2-n)\omega_n} \int_{\mathbf{R}^n_+} \frac{1}{|x-y|^{n-2}} \Delta u(y)\, dy \\ &\quad + \frac{1}{(2-n)\omega_n} \int_{\mathbf{R}^{n-1}} \frac{1}{(|x'-y'|^2 + x_n^2)^{(n-2)/2}} \frac{\partial u}{\partial y_n}(y', 0)\, dy' \\ &\quad + \frac{1}{\omega_n} \int_{\mathbf{R}^{n-1}} \frac{x_n}{(|x'-y'|^2 + x_n^2)^{n/2}} u(y', 0)\, dy', \quad x \in \mathbf{R}^n_+. \end{aligned} \tag{5.75}$$

By formulas (5.72)–(5.74), we find that the first term is the Newtonian potential and the second and third terms are the single and double layer potentials, respectively.

On the other hand, it is easy to verify that if $\varphi(x')$ is bounded and continuous on $\mathbf{R}^{n-1}$, then the function

$$u(x', x_n) = \frac{2}{\omega_n} \int_{\mathbf{R}^{n-1}} \frac{x_n}{(|x' - y'|^2 + x_n^2)^{n/2}} \varphi(y')\, dy' \tag{5.76}$$

is well-defined for $(x', x_n) \in \mathbf{R}^n_+$, and is a (unique) solution of the homogeneous Dirichlet problem for the Laplacian

$$\begin{cases} \Delta u = 0 & \text{in } \mathbf{R}^n_+, \\ u = \varphi & \text{on } \mathbf{R}^{n-1}. \end{cases}$$

Formula (5.76) is called the *Poisson integral formula* for the solution of the Dirichlet problem.

Furthermore, by using the Fourier transform we can express formula (5.76) for $\varphi \in \mathcal{S}(\mathbf{R}^{n-1})$ as follows:

$$u(x', x_n) = \frac{1}{(2\pi)^{n-1}} \int_{\mathbf{R}^{n-1}} e^{ix'\cdot\xi'}\, e^{-x_n|\xi'|}\, \hat{\varphi}(\xi')\, d\xi'. \tag{5.77}$$

To do this, we need the following two elementary formulas (5.78) and (5.79):

Lemma 5.42. *(i) For any $a > 0$, we have*

$$\int_{-\infty}^{\infty} e^{i\alpha x}\, e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}\, e^{-\frac{\alpha^2}{4a}}. \tag{5.78}$$

(ii) For any $\beta > 0$, we have

$$e^{-\beta} = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-s}}{\sqrt{s}}\, e^{-\frac{\beta^2}{4s}}\, ds. \tag{5.79}$$

Proof. (i) Formula (5.78) is a special case of formula (5.15) for $n := 1$.
(ii) First, we prove

$$e^{-\beta} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{i\beta x}}{1 + x^2}\, dx. \tag{5.80}$$

To do this, we observe that the function

$$\mathbf{C} \ni z \longmapsto f(z) = \frac{e^{i\beta z}}{1 + z^2}$$

has a pole at $z = i$ in the closed half-plane $\{z \in \mathrm{C} : \operatorname{Im} z \geq 0\}$, and further that its residue is given by the formula

$$\operatorname{Res}\,[f(z)]_{z=i} = \lim_{z \to i} (z - i)\, f(z) = \frac{e^{-\beta}}{2i} = -\frac{i}{2} e^{-\beta}.$$

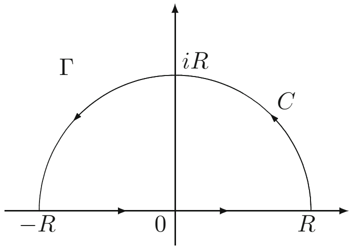


Fig. 5.7 The integral path Γ consisting of the semicircle

Hence we have, by the residue theorem,

$$\int_\Gamma f(z)\,dz = 2\pi i\left(-\frac{i}{2}e^{-\beta}\right) = \pi\, e^{-\beta}. \tag{5.81}$$

Here Γ is a path consisting of the semicircle as in Fig. 5.7:
Then we can rewrite (5.81) as follows:

$$\pi\, e^{-\beta} = \int_\Gamma f(z)\,dz = \int_C f(z)\,dz + \int_{-R}^{R} f(x)\,dx := I + II. \tag{5.82}$$

However, since we have, for all $z = x + iy \in C$,

$$\left|e^{i\beta z}\right| = \left|e^{i\beta x}e^{-\beta y}\right| = \left|e^{-\beta y}\right| \leq 1,$$

we can estimate the first term I as follows:

$$\begin{aligned}\left|\int_C f(z)\,dz\right| &= \left|\int_0^\pi \frac{e^{i\beta Re^{i\theta}}}{1+R^2e^{2i\theta}}Rie^{i\theta}\,d\theta\right| \\ &\leq \int_0^\pi \frac{R}{R^2-1}\,d\theta = \frac{\pi R}{R^2-1} \longrightarrow 0 \quad \text{as } R \to \infty.\end{aligned}$$

Therefore, (5.80) follows by letting $R \to \infty$ in (5.82).

Now, by using Fubini's theorem we obtain from (5.78) with $\alpha := \beta$ and $a := s$ and (5.80) that

$$\begin{aligned} &e^{-\beta} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{i\beta x}}{1+x^2}\,dx \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{i\beta x}\left(\int_0^\infty e^{-(1+x^2)s}\,ds\right)dx = \frac{1}{\pi}\int_0^\infty e^{-s}\left(\int_{-\infty}^{\infty} e^{i\beta x}\,e^{-sx^2}\,dx\right)ds \\ &= \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-s}}{\sqrt{s}}\,e^{-\frac{\beta^2}{4s}}\,ds.\end{aligned}$$

This proves formula (5.79).

The proof of Lemma 5.42 is complete.

Therefore, it follows from an application of Fubini's theorem and Lemma 5.42 with $\beta := x_n|\xi'|$ that

$$\begin{aligned}
&\frac{1}{(2\pi)^{n-1}} \int_{\mathbf{R}^{n-1}} e^{ix'\cdot\xi'} e^{-x_n|\xi'|} \hat{\varphi}(\xi')\, d\xi' \\
&= \int_{\mathbf{R}^{n-1}} \varphi(y') \left(\frac{1}{(2\pi)^{n-1}} \int_{\mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'} e^{-x_n|\xi'|}\, d\xi' \right) dy' \\
&= \int_{\mathbf{R}^{n-1}} \varphi(y') \left(\frac{1}{\pi^{n/2}} \frac{1}{x_n^{n-1}} \int_0^\infty e^{-s\left(1+|x'-y'|^2/x_n^2\right)} s^{n/2-1}\, ds \right) dy' \\
&= \int_{\mathbf{R}^{n-1}} \varphi(y') \left(\frac{1}{\pi^{n/2}} \frac{1}{x_n^{n-1}} \int_0^\infty e^{-\sigma} \sigma^{n/2-1}\, d\sigma\, \frac{x_n^n}{(|x'-y'|^2+x_n^2)^{n/2}} \right) dy' \\
&\qquad\qquad \left(\sigma = s\left(1+|x'-y'|^2/x_n^2\right)\right) \\
&= \frac{\Gamma(n/2)}{\pi^{n/2}} \int_{\mathbf{R}^{n-1}} \frac{x_n}{(|x'-y'|^2+x_n^2)^{n/2}} \varphi(y')\, dy' \\
&= \frac{2}{\omega_n} \int_{\mathbf{R}^{n-1}} \frac{x_n}{(|x'-y'|^2+x_n^2)^{n/2}} \varphi(y')\, dy'.
\end{aligned}$$

This proves the desired formula (5.77). □

5.7 Distribution Theory on a Manifold

This section gives a summary of the basic definitions and results from the theory of distributions on a manifold. The virtue of manifold theory is that it provides a geometric insight into the study of distributions, and intrinsic properties of distributions may be revealed.

5.7.1 Manifolds

In this subsection we summarize some basic facts about manifold theory. Manifolds are an abstraction of the idea of a surface in Euclidean space.

Let X be a set and $0 \leq r \leq \infty$. An *atlas* or *coordinate neighborhood system* of class C^r on X is a family of pairs $\mathcal{A} = \{(U_i, \varphi_i)\}_{i\in I}$ satisfying the following three conditions (MA1), (MA2) and (MA3):

(MA1) Each U_i is a subset of X and $X = \cup_{i \in I} U_i$.
(MA2) Each φ_i is a bijection of U_i onto an open subset of $\mathbf{R}^n$, and for every pair i, j of I with $U_i \cap U_j \neq \emptyset$ the set $\varphi_i(U_i \cap U_j)$ is open in $\mathbf{R}^n$.
(MA3) For each pair i, j of I with $U_i \cap U_j \neq \emptyset$, the mapping

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)$$

is a C^r diffeomorphism. Here a C^0 diffeomorphism means a homeomorphism.

In other words, X is a set which can be covered by subsets U_i, each of which is parametrized by an open subset of $\mathbf{R}^n$. Each pair (U_i, φ_i) is called a *chart* or *coordinate neighborhood* of $\mathcal{A}$. The mappings $\varphi_j \circ \varphi_i^{-1}$ in condition (MA3) are called *transition maps* or *coordinate transformations*.

Let (U, φ) be a chart on X. If p is a point of U, then $\varphi(p)$ is a point of $\mathbf{R}^n$ and hence an n-tuple of real numbers. We let

$$\varphi(p) = (x_1(p), x_2(p), \ldots, x_n(p)), \quad p \in U.$$

The n-tuple $(x_1(p), x_2(p), \ldots, x_n(p))$ of real numbers is called the *local coordinates* of p in the chart (U, φ), and the n-tuple $(x_1, x_2, \ldots, x_n)$ of real-valued functions on U is called the *local coordinate system* on (U, φ). We use the standard notation

$$\varphi(x) = (x_1, x_2, \ldots, x_n), \quad x \in U.$$

Two atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ on X are said to be *compatible* if the union $\mathcal{A}_1 \cup \mathcal{A}_2$ is an atlas on X. It is easy to see that the relation of compatibility between atlases is an equivalence relation. An equivalence class of atlases on X is said to define a C^r *structure* $\mathcal{D}$ on X. The union

$$\mathcal{A}_{\mathcal{D}} = \bigcup \{\mathcal{A} : \mathcal{A} \in \mathcal{D}\}$$

of the atlases in $\mathcal{D}$ is called the *maximal atlas* of $\mathcal{D}$, and a chart (U, φ) of $\mathcal{A}_{\mathcal{D}}$ is called an *admissible chart*.

An n-dimensional C^r *manifold* M is a pair consisting of a set X and a C^r structure $\mathcal{D}$ on X. We often identify M with the underlying set X for notational convenience.

Topology on Manifolds

Now we see how to define a topology on a manifold by means of atlases. Let M be an n-dimensional C^r manifold. A subset O of M is defined to be open if and only if, for each $x \in O$, there exists an admissible chart (U, φ) such that $x \in U$ and $U \subset O$. It is easy to verify that the open sets in M define a topology.

A C^r manifold is said to be *Hausdorff* if it is Hausdorff as a topological space. From now on, we assume that our manifolds are Hausdorff.

Let X be a topological space. A collection $\mathcal{C}$ of subsets of X is said to be *locally finite* if every point of X has a neighborhood which intersects only finitely many elements of $\mathcal{C}$. A covering $\{V_j\}$ of X is called a *refinement* of a covering $\{U_i\}$ of X if each V_j is contained in some U_i.

A topological space X is said to be *paracompact* if it is a Hausdorff space and every open covering of X has a locally finite refinement which is also an open covering of X. It is well known that a C^0 manifold M is paracompact and its number of connected components is at most countable if and only if M satisfies the second axiom of countability.

A subset N of a C^r manifold M, $0 \leq r \leq \infty$, is called a *submanifold* of M if, at each point x of N, there exists an admissible chart (U, φ) on M such that:

(SM) $\varphi : U \to V_1 \times V_2$, where V_1 is open in $\mathbf{R}^m$ and V_2 is open in $\mathbf{R}^{n-m}$, $1 \leq m \leq n$, and we have

$$\varphi(U \cap N) = V_1 \times \{0\}.$$

The number $n - m$ is called the *codimension* of N in M.

An open subset of M is a submanifold if we take $m = n$, and is called an *open submanifold*. A submanifold of M is called a *closed submanifold* if it is a closed subset of M.

Densities on a Manifold

Let E be an n-dimensional linear space over $\mathbf{R}$ and $E^* = L(E, \mathbf{R})$ its dual space. Let $\wedge^n E$ be the n-th exterior product of E and $\wedge^n E^*$ the n-th exterior product of E^*. Then the spaces $\wedge^n E$ and $\wedge^n E^*$ are both one-dimensional and are dual to each other. The non-zero elements of $\wedge^n E^*$ are called *volume elements* on E.

A complex-valued *density* on E is a mapping

$$\rho : \bigwedge^n E \longrightarrow \mathbf{C}$$

such that

$$\rho(\lambda\sigma) = |\lambda|\rho(\sigma), \quad \lambda \in \mathbf{R}.$$

The set of all densities on E is a complex linear space with the obvious operations of addition and scalar multiplication. This linear space is denoted by $\Omega(E^*)$, and is called the space of densities on E.

Densities can be constructed from volume elements in the following way: If $\omega \in \wedge^n E^*$, we define a mapping

$$|\omega| : \bigwedge^n E \longrightarrow \mathbf{C}$$

by the formula

$$|\omega|(\sigma) = |\langle \sigma, \omega \rangle|, \quad \sigma \in \bigwedge^n E,$$

where $\langle \cdot, \cdot \rangle$ is the pairing of $\wedge^n E$ and $\wedge^n E^*$. Then we have

$$|\omega| \in \Omega(E^*).$$

The space $\Omega(E^*)$ is one-dimensional. Indeed, if $(e_1, e_2, \ldots, e_n)$ is a basis of E and $(e^1, e^2, \ldots, e^n)$ is the corresponding dual basis of E^*, then every ρ of $\Omega(E^*)$ can be written in the form

$$\rho = \rho(e_1 \wedge \cdots \wedge e_n) \left| e^1 \wedge \cdots \wedge e^n \right|.$$

Now let M be an n-dimensional C^∞ manifold. We remark that if (U, φ) is a chart on M with $\varphi(x) = (x_1, x_2, \ldots, x_n)$, then the density

$$|dx_1 \wedge \cdots \wedge dx_n|$$

is a basis of the space $\Omega(T_x^*(M))$ of densities on the tangent space $T_x(M)$ of M at each point x of U.

We let

$$\Omega(T^*(M)) = \bigsqcup_{x \in M} \Omega(T_x^*(M))$$

be the disjoint union of the spaces $\Omega(T_x^*(M))$, and define a mapping

$$|\pi| : \Omega(T^*(M)) \longrightarrow M$$

by the formula

$$|\pi|(\rho) = x \quad \text{if } \rho \in \Omega(T_x^*(M)),$$

and define a mapping

$$|\varphi| : |\pi|^{-1}(U) \longrightarrow \varphi(U) \times \mathbf{R}^2$$

by the formula

$$|\varphi|(\rho) = \left(\varphi(x), \rho \left(\frac{\partial}{\partial x_1} \wedge \cdots \wedge \frac{\partial}{\partial x_n} \right) \right)$$

if $|\pi|(\rho) = x$. Here we identify $\mathbf{C}$ with $\mathbf{R}^2$.

We can make $\Omega(T^*(M))$ into an $(n+2)$-dimensional C^∞ manifold by giving natural charts for it. Indeed, if (U, φ) is a chart on M with $\varphi(x) = (x_1, x_2, \ldots, x_n)$, then the family of pairs $\{(|\pi|^{-1}(U), |\varphi|)\}$, where (U, φ) ranges over all admissible charts, is an atlas on $\Omega(T^*(M))$. We call $\Omega(T^*(M))$ the *fiber bundle of densities* on the tangent spaces of M.

A C^∞ *density* on M is a C^∞ mapping

$$\rho : M \longrightarrow \Omega(T^*(M))$$

such that $\rho(x) \in \Omega(T^*_x(M))$ for each $x \in M$. The set $C^\infty(|M|)$ of all C^∞ densities on M is a complex linear space with the obvious operations of addition and scalar multiplication.

5.7.2 Distributions on a Manifold

Let M be an n-dimensional C^∞ manifold (without boundary) which satisfies the second axiom of countability; hence M is paracompact. It is well known that every paracompact C^∞ manifold M has a partition of unity $\{\varphi_\lambda\}_{\lambda \in \Lambda}$ subordinate to any given open covering $\{U_\lambda\}_{\lambda \in \Lambda}$ of M. That is, the family $\{\varphi_\lambda\}_{\lambda \in \Lambda}$ in $C^\infty(M)$ satisfies the following three conditions (PU1), (PU2) and (PU3):

(PU1) $0 \leq \varphi_\lambda(x) \leq 1$ for all $x \in M$ and $\lambda \in \Lambda$.
(PU2) $\operatorname{supp} \varphi_\lambda \subset U_\lambda$ for each $\lambda \in \Lambda$.
(PU3) The collection $\{\operatorname{supp} \varphi_\lambda\}_{\lambda \in \Lambda}$ is locally finite and

$$\sum_{\lambda \in \Lambda} \varphi_\lambda(x) = 1 \quad \text{for every } x \in M.$$

We let

$$C^\infty(M) = \text{the space of } C^\infty \text{ functions on } M.$$

We equip the space $C^\infty(M)$ with the topology defined by the family of seminorms:

$$\varphi \longmapsto p(\varphi \circ \chi^{-1}), \quad \varphi \in C^\infty(M),$$

where (U, χ) ranges over all admissible charts on M and p ranges over all seminorms on $C^\infty(\chi(U))$ such as formula (5.1). By using a partition of unity, we can verify that the topology on $C^\infty(M)$ is defined by the family of seminorms associated with an atlas on M alone. However, since M satisfies the second axiom of countability, there exists an atlas on M consisting of countably many charts. This shows that $C^\infty(M)$ is metrizable. Furthermore, it is easy to see that $C^\infty(M)$ is complete; hence it is a Fréchet space.

If K is a compact subset of M, we let

$$C_K^\infty(M) = \text{the space of } C^\infty \text{ functions on } M \text{ with support in } K.$$

The space $C_K^\infty(M)$ is a closed subspace of $C^\infty(M)$. Furthermore, we let

$$C_0^\infty(M) = \bigcup_{K \subset M} C_K^\infty(M),$$

where K ranges over all compact subsets of M. We equip the space $C_0^\infty(M)$ with the inductive limit topology of the spaces $C_K^\infty(M)$.

We let

$$C^\infty(|M|) = \text{the space of } C^\infty \text{ densities on } M,$$
$$C_0^\infty(|M|) = \text{the space of } C^\infty \text{ densities on } M \text{ with compact support.}$$

Since M is paracompact, it is known that M admits a strictly positive C^∞ density μ. Hence we can identify $C^\infty(|M|)$ with $C^\infty(M)$ as linear topological spaces by the isomorphism

$$\begin{aligned} C^\infty(M) &\longrightarrow C^\infty(|M|) \\ \varphi &\longmapsto \varphi \cdot \mu. \end{aligned}$$

Similarly, the space $C_0^\infty(|M|)$ can be identified with the space $C_0^\infty(M)$.

A *distribution* on M is a continuous linear functional on $C_0^\infty(|M|)$. The space of distributions on M is denoted by $\mathcal{D}'(M)$. That is, $\mathcal{D}'(M)$ is the dual space $\left(C_0^\infty(|M|)\right)' = \mathcal{L}(C_0^\infty(|M|), \mathbf{C})$. If $\varphi \in C_0^\infty(M)$ and $u \in \mathcal{D}'(M)$, we denote the action of u on $\varphi \cdot \mu$ by $\langle u, \varphi \cdot \mu \rangle$ or sometimes by $\langle \varphi \cdot \mu, u \rangle$.

A function u defined on M is said to be in $L^1_{\text{loc}}(M)$ if, for any admissible chart (U, χ) on M, the local representative $u \circ \chi^{-1}$ of u is in $L^1_{\text{loc}}(\chi(U))$. The elements of $L^1_{\text{loc}}(M)$ are called *locally integrable functions* on M. Every element u of $L^1_{\text{loc}}(M)$ defines a distribution on M by the formula

$$\langle u, \varphi \cdot \mu \rangle = \int_M u\varphi \cdot \mu \quad \text{for every } \varphi \in C_0^\infty(M).$$

We list some basic properties of distributions on a manifold:

(1) If V is an open subset of M, then a distribution $u \in \mathcal{D}'(M)$ defines a distribution $u|_V \in \mathcal{D}'(V)$ by restriction to $C_0^\infty(|V|)$.

(2) The space $\mathcal{D}'(M)$ has the *sheaf property*; this means the following two properties (S1) and (S2) hold:

(S1) If $\{U_\lambda\}_{\lambda \in \Lambda}$ is an open covering of M and if a distribution $u \in \mathcal{D}'(M)$ is zero in each U_λ, then $u = 0$ in M.

(S2) Given an open covering $\{U_\lambda\}_{\lambda\in\Lambda}$ of M and a family of distributions $u_\lambda \in \mathcal{D}'(U_\lambda)$ such that $u_j = u_k$ in every $U_\lambda \cap U_\mu$, there exists a distribution $u \in \mathcal{D}'(M)$ such that $u = u_\lambda$ in each U_λ.

(3) The space of distributions with compact support can be identified with the dual space $\mathcal{E}'(M)$ of $C^\infty(|M|)$.

We have the same topological properties of $\mathcal{D}'(M)$ and $\mathcal{E}'(M)$ as those of $\mathcal{D}'(\Omega)$ and $\mathcal{E}'(\Omega)$ stated in Sect. 5.4.

5.7.3 *Differential Operators on a Manifold*

$$\begin{array}{ccc} C^\infty(U) & \xrightarrow{P|_U} & C^\infty(U) \\ \chi^* \uparrow & & \downarrow \chi_* \\ C^\infty(\chi(U)) & \xrightarrow[\chi_* P]{} & C^\infty(\chi(U)) \end{array}$$

Let M be an n-dimensional C^∞ manifold (without boundary). If P is a linear mapping of $C^\infty(M)$ into itself and if (U, χ) is a chart on M, we let

$$\chi_* P = \chi_* \circ (P|_U) \circ \chi^*,$$

where $P|_U$ is the restriction of P to U and $\chi^* v = v \circ \chi$ is the pull-back of v by χ and $\chi_* u = u \circ \chi^{-1}$ is the push-forward of u by χ, respectively. Then it follows that $\chi_* P$ is a linear mapping of $C^\infty(\chi(U))$ into itself. The situation can be visualized in the above commutative diagram.

A continuous linear mapping $P : C^\infty(M) \to C^\infty(M)$ is called a *differential operator* of order m on M if, for any chart (U, χ) on M, the mapping $\chi_* P$ is a differential operator of order m on $\chi(U) \subset \mathbf{R}^n$.

Example 5.43. Let (M, g) be an n-dimensional, Riemannian smooth manifold. The *Laplace–Beltrami operator* or simply the *Laplacian* Δ_M of M is a second-order differential operator defined (in local coordinates) by the formula

$$\begin{aligned} \Delta_M &= \operatorname{div}(\operatorname{grad} f) \\ &= \sum_{k,\ell=1}^{n} \frac{1}{\sqrt{\det(g_{ij})}} \frac{\partial}{\partial x_k} \left(\sqrt{\det(g_{ij})}\, g^{k\ell} \frac{\partial}{\partial x_\ell} \right) \quad \text{for every } f \in C^\infty(M), \end{aligned}$$

where

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right),$$
$$(g^{ij}) = \text{the inverse matrix of } (g_{ij}).$$

If $M = \mathbf{R}^n$ with standard Euclidean metric $(g_{ij}) = (\delta_{ij})$, then the Laplace–Beltrami operator Δ_M becomes the usual Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}.$$

5.7.4 Operators and Kernels on a Manifold

Let M and N be C^∞ manifolds equipped with strictly positive densities μ and ν, respectively.

If $K \in \mathcal{D}'(M \times N)$, we can define a continuous linear operator $A : C_0^\infty(N) \to \mathcal{D}'(M)$ by the formula

$$\langle A\psi, \varphi \cdot \mu \rangle = \langle K, \varphi \cdot \mu \otimes \psi \cdot \nu \rangle \quad \text{for all } \varphi \in C_0^\infty(M) \text{ and } \psi \in C_0^\infty(N).$$

If $A : C_0^\infty(N) \to \mathcal{D}'(M)$ is a continuous linear operator, we define its *transpose* A' by the formula

$$\left\langle A'\varphi, \psi \cdot \nu \right\rangle = \langle \varphi \cdot \mu, A\psi \rangle \quad \text{for all } \varphi \in C_0^\infty(M) \text{ and } \psi \in C_0^\infty(N).$$

Then the transpose A' is a continuous linear operator on $C_0^\infty(M)$ into $\mathcal{D}'(N)$. Also we have $(A')' = A$.

Similarly, we define the *adjoint* A^* of A by the formula

$$\left\langle A^*\varphi, \overline{\psi} \cdot \nu \right\rangle = \left\langle \varphi \cdot \mu, \overline{A\psi} \right\rangle \quad \text{for all } \varphi \in C_0^\infty(M) \text{ and } \psi \in C_0^\infty(N).$$

Then the adjoint A^* is a continuous linear operator on $C_0^\infty(M)$ into $\mathcal{D}'(N)$, and we have $(A^*)^* = A$.

It should be emphasized that the results in Sect. 5.5 extend to this case.

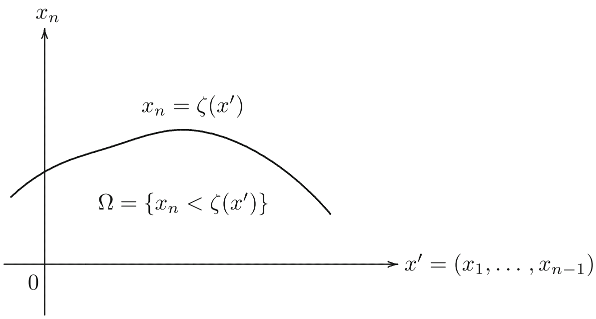


Fig. 5.8 The Lipschitz hypograph Ω

5.8 Domains of Class C^r

In this section we introduce the notion of domains of class C^r from the viewpoint of manifold theory.

An open set Ω in $\mathbf{R}^n$ is called a *Lipschitz hypograph* if its boundary $\partial\Omega$ can be represented as the graph of a Lipschitz continuous function. That is, there exists a Lipschitz continuous function $\zeta : \mathbf{R}^{n-1} \to \mathbf{R}$ such that (see Fig. 5.8)

$$\Omega = \{x = (x', x_n) \in \mathbf{R}^n : x_n < \zeta(x'), \quad x' \in \mathbf{R}^{n-1}\}. \tag{5.83}$$

An open subset of $\mathbf{R}^n$ is called a *domain* if it is also connected. Let $0 \leq r \leq \infty$. A domain Ω in $\mathbf{R}^n$ with boundary $\partial\Omega$ is said to be *of class* C^r or a C^r *domain* if, at each point x_0' of $\partial\Omega$, there exist a neighborhood U of x_0' in $\mathbf{R}^n$ and a bijection χ of U onto $B = \{x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n : |x| < 1\}$ such that (see Fig. 5.9)

$$\begin{aligned}
&\chi(U \cap \Omega) = B \cap \{x_n > 0\}, \\
&\chi(U \cap \partial\Omega) = B \cap \{x_n = 0\}, \\
&\chi \in C^r(U), \quad \chi^{-1} \in C^r(B).
\end{aligned}$$

More precisely, a C^r domain is an n-dimensional C^r manifold with boundary (see Sect. 7.1).

Sometimes, a different smoothness condition will be needed, so we broaden the above terminology as follows: For any non-negative integer k and any $0 < \theta \leq 1$, we say that the domain Ω defined by formula (5.83) is a $C^{k,\theta}$ *hypograph* if the function ζ is of class $C^{k,\theta}$, that is, if ζ is of class C^k and its k-th order partial derivatives are Hölder continuous with exponent θ.

The next definition requires that, roughly speaking, the boundary of Ω can be represented locally as the graph of a Lipschitz continuous function, by using different systems of Cartesian coordinates for different parts of the boundary:

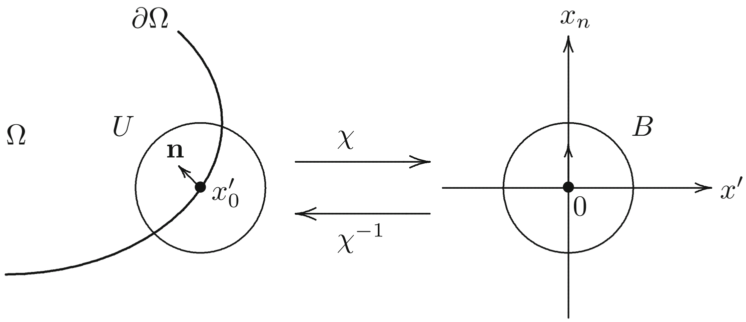


Fig. 5.9 The coordinate neighborhood (U, χ)

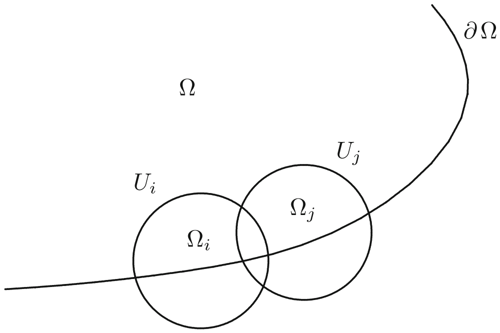


Fig. 5.10 The families $\{U_j\}$ and $\{\Omega_j\}$ in Definition 5.44

Definition 5.44. Let Ω be a bounded domain in Euclidean space $\mathbf{R}^n$ with boundary $\partial\Omega$. We say that Ω is a *Lipschitz domain* if there exist finite families $\{U_j\}_{j=1}^J$ and $\{\Omega_j\}_{j=1}^J$ having the following three properties (i), (ii) and (iii) (see Fig. 5.10):

(i) The family $\{U_j\}_{j=1}^J$ is a finite open covering of $\partial\Omega$.
(ii) Each Ω_j can be transformed to a Lipschitz hypograph by a rigid motion, that is, by a rotation plus a translation.
(iii) The set Ω satisfies the conditions

$$U_j \cap \Omega = U_j \cap \Omega_j, \quad 1 \leq j \leq J.$$

In the obvious way, we define a *domain of class* $C^{k,\theta}$ or $C^{k,\theta}$ *domain* by substituting "$C^{k,\theta}$" for "Lipschitz" throughout Definition 5.44. It should be emphasized that a Lipschitz domain is the same thing as a $C^{0,1}$ domain.

If Ω is a Lipschitz hypograph defined by (5.83), then we observe that its boundary

$$\partial\Omega = \left\{x = \left(x', \zeta(x')\right) : x' \in \mathbf{R}^{n-1}\right\}$$

is an $(n-1)$-dimensional, $C^{0,1}$ submanifold of $\mathbf{R}^n$ if we apply the following Rademacher theorem (see [MZ, Corollary 1.73], [Sn3, Theorem]):

Theorem 5.45 (Rademacher). *Any Lipschitz continuous function on $\mathbf{R}^n$ admits L^∞ first partial derivatives almost everywhere in $\mathbf{R}^n$.*

Indeed, it follows from an application of Rademacher's theorem that the function $\zeta(x')$ is Fréchet differentiable almost everywhere in $\mathbf{R}^{n-1}$ with

$$\|\nabla\zeta\|_{L^\infty(\mathbf{R}^{n-1})} \leq C, \tag{5.84}$$

where C is any Lipschitz constant for the function $\zeta(x')$. Hence the Riemannian metric (h_{ij}) of $\partial\Omega$ is given by

$$\begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1\,n-1} \\ h_{21} & h_{22} & \cdots & h_{2\,n-1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ h_{n-1\,1} & h_{n-1\,2} & \cdots & h_{n-1\,n-1} \end{pmatrix}$$
$$= \begin{pmatrix} 1+\zeta_{x_1}^2 & \zeta_{x_1}\zeta_{x_2} & \cdots & \zeta_{x_1}\zeta_{x_{n-1}} \\ \zeta_{x_2}\zeta_{x_1} & 1+\zeta_{x_2}^2 & \cdots & \zeta_{x_2}\zeta_{x_{n-1}} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \zeta_{x_{n-1}}\zeta_{x_1} & \zeta_{x_{n-1}}\zeta_{x_2} & \cdots & 1+\zeta_{x_{n-1}}^2 \end{pmatrix},$$

where

$$\zeta_{x_i} = \frac{\partial\zeta}{\partial x_i}, \quad 1 \leq i \leq n-1.$$

It is easy to see that

$$\det\left(h_{ij}\right) = 1 + \zeta_{x_1}^2 + \ldots + \zeta_{x_{n-1}}^2 = 1 + |\nabla\zeta(x')|^2.$$

Therefore, we obtain that the boundary $\partial\Omega$ has the surface measure $d\sigma$ and that the unit exterior normal $\boldsymbol{\nu}$ exists $d\sigma$-almost everywhere in $\mathbf{R}^{n-1}$ (see Fig. 5.11), where $d\sigma$ and $\boldsymbol{\nu}$ are given respectively by the formulas:

$$d\sigma = \sqrt{1 + |\nabla\zeta(x')|^2}\, dx',$$

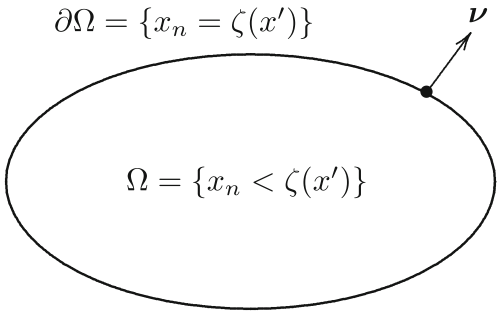


Fig. 5.11 The unit exterior normal $\boldsymbol{\nu}$ to $\partial\Omega$

$$\boldsymbol{\nu} = \frac{(-\nabla\zeta(x'), 1)}{\sqrt{1 + |\nabla\zeta(x')|^2}}.$$

Here it should be noted that we have, by inequality (5.84),

$$1 \leq \sqrt{1 + |\nabla\zeta(x')|^2} \leq \sqrt{1 + C^2},$$

so that the surface measure $d\sigma$ is equivalent locally to the Lebesgue measure dx'.

5.9 Notes and Comments

Schwartz [Sz] and Gelfand–Shilov [GS] are the classics for distribution theory. Our treatment here follows the expositions of Chazarain–Piriou [CP], Hörmander [Ho4] and Treves [Tv].

Sections 5.1 and 5.2: The material in these sections is taken from Gilbarg–Trudinger [GT] and Folland [Fo1].

Section 5.3: Peetre's theorem 5.7 is due to Peetre [Pe].

Section 5.4: For the Banach–Steinhaus theorem, see Treves [Tv, Chapter 33]. Example 5.29 is taken from Chazarain–Piriou [CP, Chapitre III, Lemme 9.4].

Remark 4.16 in Chap. 4 is based on the following approximation formula (5.85) for the Dirac measure $\delta(x-y)$:

$$\delta(x-y) = \frac{1}{(2\pi)^n} \lim_{t\downarrow 0} \int_{\mathbf{R}^n} e^{i(x-y)\cdot\xi}\, e^{-t|\xi|^2}\, d\xi. \tag{5.85}$$

Indeed, we obtain from Example 5.20 (2) that

$$\lim_{t\downarrow 0} K_t(x-y) = \lim_{t\downarrow 0} \left(\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i(x-y)\cdot\xi}\, e^{-t|\xi|^2}\, d\xi \right) = \delta(x-y).$$

Hence we have (formally) the assertion

$$\lim_{t\downarrow 0} w(x,t) = \lim_{t\downarrow 0}\int_{\mathbf{R}^n} K_t(x-y)\,u(y)\,dy = \int_{\mathbf{R}^n}\delta(x-y)\,u(y)\,dy = u(x).$$

We remark that formula (5.85) is an *approximation to the identity*, since the Dirac measure $\delta(x-y)$ is the distribution kernel of the identity operator.

Section 5.5: The proof of Schwartz's kernel theorem (Theorem 5.36) is taken from Chazarain–Piriou [CP, Chapitre I, Théorème 4.4].

Section 5.6: More detailed and concise accounts of layer potentials are given by the books of Folland [Fo1] and McLean [Mc].

The Poisson integral formula (5.76) is based on the following approximation formula (5.86) for the Dirac measure $\delta(x'-y')$:

$$\delta(x'-y') = \frac{1}{(2\pi)^{n-1}}\lim_{x_n\downarrow 0}\int_{\mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'}\,e^{-x_n|\xi'|}\,d\xi'. \tag{5.86}$$

Indeed, we find from the proof of formula (5.77) that

$$\begin{aligned}&\lim_{x_n\downarrow 0}\frac{2}{\omega_n}\int_{\mathbf{R}^{n-1}}\frac{x_n}{(|x'-y'|^2+x_n^2)^{n/2}}\,dy' \\ &= \lim_{x_n\downarrow 0}\left(\frac{1}{(2\pi)^{n-1}}\int_{\mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'}\,e^{-x_n|\xi'|}\,d\xi'\right) = \delta(x'-y').\end{aligned}$$

Hence we have (formally) the assertion

$$\begin{aligned}\lim_{x_n\downarrow 0} u(x',x_n) &= \lim_{x_n\downarrow 0}\left(\frac{2}{\omega_n}\int_{\mathbf{R}^{n-1}}\frac{x_n}{(|x'-y'|^2+x_n^2)^{n/2}}\varphi(y')\,dy'\right)\\ &= \int_{\mathbf{R}^{n-1}}\delta(x'-y')\,\varphi(y')\,dy'\\ &= \varphi(x').\end{aligned}$$

We remark that formula (5.86) is an *approximation to the identity*, since the Dirac measure $\delta(x'-y')$ is the distribution kernel of the identity operator.

Section 5.7: Distributions on a manifold were first studied by de Rham [De]. The material here is adapted from Abraham–Marsden–Ratiu [AMR], Chazarain–Piriou [CP] and Lang [Lg].

Section 5.8: The definition of a C^r domain is taken from McLean [Mc].

Chapter 6
Sobolev and Besov Spaces

Chapter 6 is devoted to the precise definitions and statements of Sobolev and Besov spaces of L^p type with some detailed proofs. One of the most useful ways of measuring differentiability properties of functions on $\mathbf{R}^n$ is in terms of L^p norms, and is provided by the Sobolev spaces on $\mathbf{R}^n$. The great advantage of this approach lies in the fact that the Fourier transform works very well in $L^p(\mathbf{R}^n)$. The function spaces we shall treat are the following:

(i) The generalized Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$ of L^p type on an open subset Ω of $\mathbf{R}^n$, which will be used in subsequent chapters. When Ω is a Lipschitz domain, these spaces coincide with each other.

(ii) The Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$ on $\mathbf{R}^{n-1}$ are function spaces defined in terms of the L^p modulus of continuity. The Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ of a Lipschitz domain Ω are defined to be locally the Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity.

In studying boundary value problems in the domain Ω, we need to make sense of the restriction $u|_{\partial\Omega}$ as an element of a function space on the boundary $\partial\Omega$ when u belongs to a Sobolev space of L^p type on Ω. In Sect. 6.1 we prove Hardy's inequality on the interval $(0, \infty)$ (Theorem 6.2) which is used systematically in the proof of a trace theorem (Theorem 6.6). In Sect. 6.2 we present some basic definitions and results of the Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$. In Sect. 6.3 we give the precise definition of the Besov space $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$. It should be emphasized that the Besov spaces $B^{s,p}(\partial\Omega)$ enter naturally in connection with boundary value problems in the framework of function spaces of L^p type. In Sect. 6.4 we prove a trace theorem (Theorem 6.6) which will play an important role in the study of boundary value problems in Chap. 7.

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_6,

6.1 Hardy's Inequality

First, we prove a general integral inequality on the interval $(0, \infty)$.

Theorem 6.1. *Let $K(x, y)$ be a Lebesgue measurable function defined on the space $(0, \infty) \times (0, \infty)$. Assume that $K(x, y)$ is positively homogeneous of degree -1, that is,*

$$K(\lambda x, \lambda y) = \lambda^{-1} K(x, y) \quad \text{for every } \lambda > 0,$$

and further that the integral

$$A_K := \int_0^\infty |K(1, y)| y^{-1/p} \, dy$$

is finite for some $1 \leq p \leq \infty$.

Then the operator Tf, defined by the formula

$$Tf(x) = \int_0^\infty K(x, y) f(y) \, dy, \quad y \in (0, \infty),$$

is bounded from $L^p(0, \infty)$ into itself. More precisely, we have the inequality

$$\|Tf\|_p \leq A_K \|f\|_p \quad \text{for all } f \in L^p(0, \infty).$$

Proof. Since we have, by the positive homogeneity of K, the formula

$$\begin{aligned} Tf(x) &= \int_0^\infty K(x, y) f(y) \, dy = \int_0^\infty K(x, zx) f(zx) \, x \, dz \quad (y = zx) \\ &= \int_0^\infty x^{-1} K(1, z) f(zx) \, x \, dz = \int_0^\infty K(1, z) f(zx) \, dz, \end{aligned}$$

by applying Minkowski's inequality for integrals [Fo2, Theorem 6.19] we obtain that

$$\|Tf\|_p \leq \int_0^\infty |K(1, z)| \|f(z \cdot)\|_p \, dz.$$

However, it is easy to see that

$$\begin{aligned} \|f(z \cdot)\|_p &= \left(\int_0^\infty |f(zx)|^p \, dx \right)^{1/p} = z^{-1/p} \left(\int_0^\infty |f(y)|^p \, dy \right)^{1/p} \\ &= z^{-1/p} \|f\|_p \quad \text{for each } z > 0. \end{aligned}$$

Therefore, we have the inequality

$$\|Tf\|_p \le \left(\int_0^\infty |K(1,z)| z^{-1/p}\, dz\right) \|f\|_p = A_K \|f\|_p.$$

The proof of Theorem 6.1 is complete.

Now we can prove Hardy's inequality, which will be used systematically.

Theorem 6.2 (Hardy's inequality). *Let* $1 \le p \le \infty$ *and* $\gamma \ne 0$*. If* $f(x)$ *is a non-negative, Lebesgue measurable function on the interval* $(0,\infty)$*, we define a function* $F(x)$ *by the formula*

$$F(x) = \begin{cases} \int_0^x f(y)\, dy & \text{if } \gamma < 0, \\ \int_x^\infty f(y)\, dy & \text{if } \gamma > 0. \end{cases}$$

Then we have the inequality

$$\left(\int_0^\infty (x^\gamma F(x))^p \frac{dx}{x}\right)^{1/p} \le \frac{1}{|\gamma|} \left(\int_0^\infty (x^{\gamma+1} f(x))^p \frac{dx}{x}\right)^{1/p}. \tag{6.1}$$

Proof. We only consider the case where $\gamma < 0$. The case where $\gamma > 0$ is proved similarly.

If we let

$$K(x,y) := \begin{cases} x^{\gamma-1/p} y^{-\gamma+1/p-1} & \text{if } 0 \le y \le x, \\ 0 & \text{if } x < y, \end{cases}$$

then it follows that $K(x,y)$ is positively homogeneous of degree -1 and satisfies the conditions

$$\int_0^\infty K(1,y) y^{-1/p}\, dy = \int_0^1 y^{-\gamma-1}\, dy = -\frac{1}{\gamma}.$$

If we introduce an integral operator

$$\begin{aligned} Tg(x) &:= \int_0^\infty K(x,y)\, g(y)\, dy \\ &= x^{\gamma-1/p} \int_0^x y^{-\gamma+1/p-1} g(y)\, dy \quad \text{for all } g \in L^p(0,\infty), \end{aligned}$$

then, by applying Theorem 6.1 to our situation we obtain that

$$\left(\int_0^\infty \left(x^{\gamma-1/p} \int_0^x y^{-\gamma+1/p-1} g(y)\, dy\right)^p dx\right)^{1/p} \le \frac{1}{|\gamma|} \left(\int_0^\infty g(y)^p\, dy\right)^{1/p}.$$

In particular, if we let

$$g(y) := y^{\gamma-1/p+1} f(y) \quad \text{for every } y \in (0, \infty),$$

then it follows that

$$\begin{aligned}\left(\int_0^\infty (x^\gamma F(x))^p \frac{dx}{x}\right)^{1/p} &= \left(\int_0^\infty \left(x^{\gamma-1/p} \int_0^x f(x)\, dy\right)^p dx\right)^{1/p} \\ &\leq \frac{1}{|\gamma|}\left(\int_0^\infty \left(y^{\gamma-1/p+1} f(y)\right)^p dy\right)^{1/p} \\ &= \frac{1}{|\gamma|}\left(\int_0^\infty (x^{\gamma+1} f(x))^p \frac{dx}{x}\right)^{1/p}.\end{aligned}$$

The proof of Theorem 6.2 is complete.

Example 6.3. If we let

$$\gamma =: \frac{1}{p} - 1, \quad 1 < p \leq \infty,$$

then we have, by inequality (6.1),

$$\left(\int_0^\infty \left(\frac{1}{x}\int_0^x f(y)\, dy\right)^p dx\right)^{1/p} \leq \frac{p}{p-1}\left(\int_0^\infty f(y)^p\, dy\right)^{1/p}.$$

6.2 Sobolev Spaces

In this section we present a brief description of the basic concepts and results from the theory of Sobolev spaces of L^p type, which will be used in subsequent chapters. Many problems in partial differential equations may be formulated in terms of abstract operators acting between suitable Sobolev and Besov spaces, and these operators are then analyzed by the methods of functional analysis.

6.2.1 First Definition of Sobolev Spaces

Let Ω be an open subset of $\mathbf{R}^n$. If $1 < p < \infty$ and if s is a non-negative integer, then the Sobolev space $W^{s,p}(\Omega)$ on Ω is defined to be the space of those functions $u \in L^p(\Omega)$ such that $D^\alpha u \in L^p(\Omega)$ for $|\alpha| \leq s$, and the norm $\|u\|_{W^{s,p}(\Omega)}$ is defined by the formula

$$\|u\|_{W^{s,p}(\Omega)} = \left(\sum_{|\alpha| \le s} \int_\Omega |D^\alpha u(x)|^p dx \right)^{1/p}. \tag{6.2}$$

If $1 < p < \infty$ and if $s = m + \theta$ with a non-negative integer m and $0 < \theta < 1$, then the Sobolev space $W^{s,p}(\Omega)$ on Ω is defined to be the space of those functions $u \in W^{m,p}(\Omega)$ such that, for $|\alpha| = m$, the integral (Slobodeckiĭ seminorm)

$$\iint_{\Omega \times \Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{n+p\theta}} \, dx \, dy$$

is finite. The norm $\|u\|_{W^{s,p}(\Omega)}$ of $W^{s,p}(\Omega)$ is defined by the formula

$$\|u\|_{W^{s,p}(\Omega)} = \left(\sum_{|\alpha| \le m} \int_\Omega |D^\alpha u(x)|^p \, dx + \sum_{|\alpha| = m} \iint_{\Omega \times \Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{n+p\theta}} \, dx \, dy \right)^{1/p}.$$

6.2.2 Second Definition of Sobolev Spaces

Next, we introduce a second family of Sobolev spaces, by using the Fourier transform.

Let $\mathcal{S}(\mathbf{R}^n)$ be the *Schwartz space* or space of smooth functions on $\mathbf{R}^n$ rapidly decreasing at infinity. We recall that the (direct) Fourier transform $\mathcal{F}$ and the inverse Fourier transform $\mathcal{F}^*$ are isomorphisms of $\mathcal{S}(\mathbf{R}^n)$ onto itself. The dual space $\mathcal{S}'(\mathbf{R}^n)$ of $\mathcal{S}(\mathbf{R}^n)$ is called the space of *tempered distributions* on $\mathbf{R}^n$. Roughly speaking, the tempered distributions are those distributions which grow at most polynomially at infinity, since the functions in $\mathcal{S}(\mathbf{R}^n)$ die out faster than any power of x at infinity. The importance of tempered distributions lies in the fact that they have Fourier transforms. More precisely, if $u \in \mathcal{S}'(\mathbf{R}^n)$, we define its (direct) Fourier transform $\mathcal{F}u$ by the formula

$$\langle \mathcal{F}u, \varphi \rangle = \langle u, \mathcal{F}\varphi \rangle \quad \text{for every } \varphi \in \mathcal{S}(\mathbf{R}^n),$$

where $\langle \cdot, \cdot \rangle$ is the pairing of $\mathcal{S}'(\mathbf{R}^n)$ and $\mathcal{S}(\mathbf{R}^n)$. Similarly, if $v \in \mathcal{S}'(\mathbf{R}^n)$, we define its inverse Fourier transform $\mathcal{F}^* v$ by the formula

$$\langle \mathcal{F}^* v, \psi \rangle = \langle v, \mathcal{F}^* \psi \rangle \quad \text{for every } \psi \in \mathcal{S}(\mathbf{R}^n).$$

It should be emphasized that the Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ are isomorphisms from $\mathcal{S}'(\mathbf{R}^n)$ onto itself.

If $s \in \mathbf{R}$, we define a linear map

$$\mathcal{J}^s = (I - \Delta)^{-s/2} : \mathcal{S}'(\mathbf{R}^n) \longrightarrow \mathcal{S}'(\mathbf{R}^n)$$

by the formula

$$\mathcal{J}^s u = \mathcal{F}^* \left((1 + |\xi|^2)^{-s/2} \mathcal{F} u\right) \quad \text{for every } u \in \mathcal{S}'(\mathbf{R}^n). \tag{6.3}$$

The operator $\mathcal{J}^s : \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ can be visualized in the following diagram:

$$\begin{array}{ccc}
u \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow{J^s=(1-\Delta)^{-s/2}} & \mathcal{S}'(\mathbf{R}^n) \ni J^s u \\
\mathcal{F}\Big\downarrow & & \Big\uparrow\mathcal{F}^* \\
\mathcal{F}u \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow[(1+|\xi|^2)^{-s/2}]{} & \mathcal{S}'(\mathbf{R}^n) \ni (1+|\xi|^2)^{-s/2}\mathcal{F}u
\end{array}$$

Then it is easy to see that the map $\mathcal{J}^s$ is an isomorphism from $\mathcal{S}'(\mathbf{R}^n)$ onto itself, and its inverse is the map $\mathcal{J}^{-s}$. The function $\mathcal{J}^s u$ is called the *Bessel potential* of order s of u.

We can calculate the convolution kernel $G_s(x)$ of the Bessel potential $\mathcal{J}^s u$ for all $s > 0$. More precisely, we have the following theorem:

Theorem 6.4. *Let* $s > 0$. *(i) The inverse Fourier transform*

$$\mathcal{F}^* \left((1 + |\xi|^2)^{-s/2}\right)$$

is equal to the function

$$G_s(x) = \frac{1}{(4\pi)^{s/2}} \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-\pi|x|^2/\delta}\, e^{-\delta/(4\pi)}\, \delta^{(s-n)/2}\, \frac{d\delta}{\delta}. \tag{6.4}$$

In other words, we have, by the Fourier inversion formula,

$$\mathcal{F}(G_s)(\xi) = \left(1 + |\xi|^2\right)^{-s/2}. \tag{6.5}$$

Moreover, we have

$$\int_{\mathbf{R}^n} G_s(x)\, dx = 1, \tag{6.6}$$

and so

$$G_s \in L^1(\mathbf{R}^n). \tag{6.7}$$

(ii) Let $1 \le p \le \infty$. The Bessel potential $\mathcal{J}^s$ can be expressed as follows:

$$\mathcal{J}^s u(x) = G_s * u(x) = \int_{\mathbf{R}^n} G_s(x-y)u(y)\,dy, \quad u \in L^p(\mathbf{R}^n). \tag{6.8}$$

Furthermore, the Bessel potential $\mathcal{J}^s$ is bounded from $L^p(\mathbf{R}^n)$ into itself. More precisely, we have the inequality

$$\|\mathcal{J}^s u\|_{L^p(\mathbf{R}^n)} \le \|u\|_{L^p(\mathbf{R}^n)}, \quad u \in L^p(\mathbf{R}^n). \tag{6.9}$$

Proof. (i) First, we have, for all $t > 0$,

$$\frac{1}{t^{s/2}} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t\delta}\, \delta^{s/2-1}\, d\delta.$$

In particular, by letting $t = 1 + |\xi|^2$ we obtain that

$$\frac{1}{(1+|\xi|^2)^{s/2}} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-(1+|\xi|^2)\delta}\, \delta^{s/2-1}\, d\delta.$$

Hence it follows that

$$\begin{aligned}
&\mathcal{F}^* \left((1+|\xi|^2)^{-s/2}\right) \\
&= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \frac{1}{(1+|\xi|^2)^{s/2}}\, d\xi \\
&= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \left(\frac{1}{\Gamma(s/2)} \int_0^\infty e^{-(1+|\xi|^2)\delta}\, \delta^{s/2-1}\, d\delta \right) d\xi \\
&= \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-\delta}\, \delta^{s/2-1} \left(\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, e^{-\delta|\xi|^2}\, d\xi \right) d\delta \\
&= \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-\delta}\, \delta^{s/2-1} \left\{ \frac{1}{(4\pi\delta)^{n/2}}\, e^{-|x|^2/(4\delta)} \right\} d\delta \\
&= \frac{1}{\Gamma(s/2)} \frac{1}{(4\pi)^{s/2}} \int_0^\infty e^{-\pi|x|^2/t}\, e^{-t/(4\pi)}\, t^{(s-n)/2}\, \frac{dt}{t} \\
&= G_s(x).
\end{aligned}$$

This proves formula (6.4), or equivalently, formula (6.5).

Moreover, we have, by Fubini's theorem

$$\begin{aligned}
&\int_{\mathbf{R}^n} G_s(x)\,dx \\
&= \frac{1}{(4\pi)^{s/2}}\,\frac{1}{\Gamma(s/2)}\int_0^\infty e^{-\delta/4\pi}\,\delta^{(s-n)/2-1}\left(\int_{\mathbf{R}^n} e^{-\pi|x|^2/\delta}\,dx\right)d\delta \\
&= \frac{1}{(4\pi)^{s/2}}\,\frac{1}{\Gamma(s/2)}\int_0^\infty e^{-\delta/4\pi}\,\delta^{(s-n)/2-1}\,\delta^{n/2}\,d\delta \quad (\delta = 4\pi t) \\
&= \frac{1}{(4\pi)^{s/2}}\,\frac{1}{\Gamma(s/2)}\int_0^\infty e^{-t}\,(4\pi)^{s/2-1}\,t^{s/2-1}(4\pi\,dt) \\
&= \frac{1}{\Gamma(s/2)}\int_0^\infty e^{-t}\,t^{s/2-1}\,dt = 1.
\end{aligned}$$

This proves formula (6.6) and assertion (6.7).

(ii) Therefore, by combining formulas (6.3) and (6.5) we obtain that

$$\begin{aligned}
\mathcal{J}^s u &= \mathcal{F}^*\left((1+|\xi|^2)^{-s/2}\mathcal{F}u(\xi)\right) \\
&= \mathcal{F}^*\left(\mathcal{F}(G_s)(\xi)\,\mathcal{F}u(\xi)\right) = \mathcal{F}^*\left(\mathcal{F}(G_s * u)(\xi)\right) \\
&= G_s * u, \quad u \in L^p(\mathbf{R}^n).
\end{aligned}$$

This proves formula (6.8).

Finally, in view of formula (6.8), we can apply Young's inequality (Corollary 5.3) and formula (6.6) to obtain that

$$\|\mathcal{J}^s u\|_{L^p(\mathbf{R}^n)} = \|G^s * u\|_{L^p(\mathbf{R}^n)} \le \|G^s\|_{L^1(\mathbf{R}^n)}\,\|u\|_{L^p(\mathbf{R}^n)} = \|u\|_{L^p(\mathbf{R}^n)}.$$

This proves inequality (6.9).

The proof of Theorem 6.4 is complete.

Now, if $s \in \mathbf{R}$ and $1 < p < \infty$, we let

$$\begin{aligned}
H^{s,p}(\mathbf{R}^n) &= \text{the image of } L^p(\mathbf{R}^n) \text{ under the mapping } \mathcal{J}^s \\
&= \{\mathcal{J}^s v : v \in L^p(\mathbf{R}^n)\}.
\end{aligned}$$

$$\boxed{u = J^s v \in H^{s,p}(\mathbf{R}^n) \xleftarrow{J^s} L^p(\mathbf{R}^n) \ni v = J^{-s}u}$$

We equip $H^{s,p}(\mathbf{R}^n)$ with the norm

$$\|u\|_{H^{s,p}(\mathbf{R}^n)} = \|\mathcal{J}^{-s}u\|_{L^p(\mathbf{R}^n)}, \quad u \in H^{s,p}(\mathbf{R}^n). \tag{6.10}$$

The space $H^{s,p}(\mathbf{R}^n)$ is called the *Bessel potential space* of order s or the *generalized Sobolev space* of order s.

We list some basic topological properties of $H^{s,p}(\mathbf{R}^n)$:

(1) The Schwartz space $\mathcal{S}(\mathbf{R}^n)$ is dense in each $H^{s,p}(\mathbf{R}^n)$.
(2) The space $H^{-s,p'}(\mathbf{R}^n)$ is the dual space of $H^{s,p}(\mathbf{R}^n)$, where $p' = p/(p-1)$ is the exponent conjugate to p.
(3) If $s > t$, then we have the inclusions

$$\mathcal{S}(\mathbf{R}^n) \subset H^{s,p}(\mathbf{R}^n) \subset H^{t,p}(\mathbf{R}^n) \subset \mathcal{S}'(\mathbf{R}^n),$$

with continuous injections.
(4) If s is a non-negative integer, then the space $H^{s,p}(\mathbf{R}^n)$ is isomorphic to the Sobolev space $W^{s,p}(\mathbf{R}^n)$, and the norm (6.10) is equivalent to the norm (6.2).

6.2.3 Definition of General Sobolev Spaces

Now we define the generalized Sobolev spaces $H^{s,p}(\Omega)$ for general domains Ω.

For each $s \in \mathbf{R}$ and $1 < p < \infty$, we let

$$H^{s,p}(\Omega) = \text{the space of restrictions to } \Omega \text{ of functions in } H^{s,p}(\mathbf{R}^n).$$

We equip the space $H^{s,p}(\Omega)$ with the norm

$$\|u\|_{H^{s,p}(\Omega)} = \inf \|U\|_{H^{s,p}(\mathbf{R}^n)},$$

where the infimum is taken over all $U \in H^{s,p}(\mathbf{R}^n)$ which equal u in Ω. The space $H^{s,p}(\Omega)$ is a Banach space with respect to the norm $\|\cdot\|_{s,p}$. It should be noted that

$$H^{0,p}(\Omega) = L^p(\Omega); \quad \|\cdot\|_{H^{0,p}(\Omega)} = \|\cdot\|_{L^p(\Omega)}.$$

Then we have the following important relationships between the Sobolev spaces $H^{s,p}(\Omega)$ and $W^{s,p}(\Omega)$:

Theorem 6.5. *If Ω is a bounded, Lipschitz domain, then we have, for all $s \geq 0$ and $1 < p < \infty$,*

$$H^{s,p}(\Omega) = W^{s,p}(\Omega).$$

6.3 Definition of Besov Spaces on the Boundary

In studying boundary value problems in a Lipschitz domain Ω of $\mathbf{R}^n$, we need to make sense of the restriction $u|_{\partial\Omega}$ as an element of a function space on the boundary $\partial\Omega$ when u belongs to a Sobolev space of L^p type on Ω. In this way, the Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ enter naturally in connection with boundary

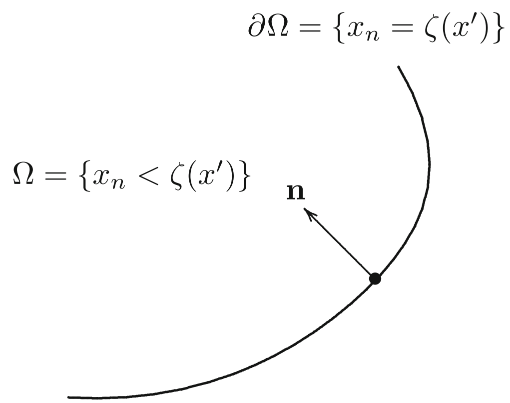


Fig. 6.1 The Lipschitz hypograph Ω

value problems. The Besov spaces $B^{s,p}(\partial\Omega)$ are defined to be locally the Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$ on $\mathbf{R}^{n-1}$, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity.

An open set Ω in $\mathbf{R}^n$ is called a *Lipschitz hypograph* if its boundary $\partial\Omega$ can be represented as the graph of a Lipschitz continuous function. In other words, there exists a Lipschitz continuous function $\zeta : \mathbf{R}^{n-1} \to \mathbf{R}$ such that (see Fig. 6.1)

$$\Omega = \{x = (x', x_n) \in \mathbf{R}^n : x_n < \zeta(x'), \quad x' \in \mathbf{R}^{n-1}\}. \tag{5.83}$$

We define *Besov spaces* $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ of a Lipschitz domain Ω, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity (see Sect. 5.7.2), in the following way.

Step 1: First, if $1 < p < \infty$, we let

$$\begin{aligned} &B^{1,p}(\mathbf{R}^{n-1}) \\ &= \text{the space of (equivalence classes of) functions } \varphi \in L^p(\mathbf{R}^{n-1}) \text{ for which} \\ &\iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi(x'+h') - 2\varphi(x') + \varphi(x'-h')|^p}{|h'|^{(n-1)+p}} dh'\, dx' < \infty. \end{aligned}$$

The space $B^{1,p}(\mathbf{R}^{n-1})$ is a Banach space with respect to the norm

$$\begin{aligned} &|\varphi|_{B^{1,p}(\mathbf{R}^{n-1})} \\ &= \left(\int_{\mathbf{R}^{n-1}} |\varphi(x')|^p \, dx' \right. \\ &\quad \left. + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi(x'+h') - 2\varphi(x') + \varphi(x'-h')|^p}{|h'|^{(n-1)+p}} \, dh'\, dx' \right)^{1/p}. \end{aligned}$$

If $s \in \mathbf{R}$ and $1 < p < \infty$, we let

$$\begin{aligned} B^{s,p}(\mathbf{R}^{n-1}) &= \text{the image of } B^{1,p}(\mathbf{R}^{n-1}) \text{ under the mapping } \mathcal{J}'^{s-1}, \\ &\quad \text{where } \mathcal{J}'^{s-1} \text{ is the Bessel potential of order } s-1 \\ &\quad \text{on } \mathbf{R}^{n-1} \\ &= \left\{ \mathcal{J}'^{s-1}\psi : \psi \in B^{1,p}(\mathbf{R}^{n-1}) \right\}. \end{aligned}$$

$$\boxed{\varphi = J'^{s-1}\psi \in B^{s,p}(\mathbf{R}^{n-1}) \overset{J'^{s-1}}{\longleftarrow} B^{1,p}(\mathbf{R}^{n-1}) \ni \psi = J'^{-s+1}\varphi}$$

We equip the space $B^{s,p}(\mathbf{R}^{n-1})$ with the norm

$$|\varphi|_{B^{s,p}(\mathbf{R}^{n-1})} = \left| \mathcal{J}'^{-s+1}\varphi \right|_{B^{1,p}(\mathbf{R}^{n-1})}, \quad \varphi \in B^{s,p}(\mathbf{R}^{n-1}).$$

The space $B^{s,p}(\mathbf{R}^{n-1})$ is called the *Besov space* of order s.

We list some basic topological properties of $B^{s,p}(\mathbf{R}^{n-1})$:

(1) The Schwartz space $\mathcal{S}(\mathbf{R}^{n-1})$ is dense in each $B^{s,p}(\mathbf{R}^{n-1})$.
(2) The space $B^{-s,p'}(\mathbf{R}^{n-1})$ is the dual space of $B^{s,p}(\mathbf{R}^{n-1})$, where $p' = p/(p-1)$.
(3) If $s > t$, then we have the inclusions

$$\mathcal{S}(\mathbf{R}^{n-1}) \subset B^{s,p}(\mathbf{R}^{n-1}) \subset B^{t,p}(\mathbf{R}^{n-1}) \subset \mathcal{S}'(\mathbf{R}^{n-1}),$$

with continuous injections.
(4) If $1 < p < \infty$ and if $s = m + \theta$ with a non-negative integer m and $0 < \theta < 1$, then the Besov space $B^{s,p}(\mathbf{R}^{n-1})$ coincides with the space of those functions $\varphi \in W^{m,p}(\mathbf{R}^{n-1})$ such that, for $|\alpha| = m$, the integral (Slobodeckiĭ seminorm)

$$\iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|D^\alpha\varphi(x') - D^\alpha\varphi(y')|^p}{|x'-y'|^{(n-1)+p\theta}}\, dx'\, dy'$$

is finite. Furthermore, the norm $|\varphi|_{s,p}$ is equivalent to the norm

$$\begin{aligned} &\left(\sum_{|\alpha|\le m} \int_{\mathbf{R}^{n-1}} |D^\alpha\varphi(x')|^p\, dx' \right. \qquad (6.11) \\ &\left. + \sum_{|\alpha|=m} \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|D^\alpha\varphi(x') - D^\alpha\varphi(y')|^p}{|x'-y'|^{(n-1)+p\theta}}\, dx'\, dy' \right)^{1/p}. \end{aligned}$$

Step 2: If Ω is a Lipschitz hypograph defined by (5.83), then we recall that its boundary

$$\partial\Omega = \{x = (x', \zeta(x')) : x' \in \mathbf{R}^{n-1}\}$$

is an $(n-1)$-dimensional $C^{0,1}$ submanifold of $\mathbf{R}^n$ and further that $\partial\Omega$ has a surface measure $d\sigma$ and a unit exterior normal $\boldsymbol{\nu}$ which exists $d\sigma$-almost everywhere in $\mathbf{R}^{n-1}$ (see Fig. 5.11):

$$d\sigma = \sqrt{1+|\nabla\zeta(x')|^2}\,dx',$$
$$\boldsymbol{\nu} = \frac{(-\nabla\zeta(x'),1)}{\sqrt{1+|\nabla\zeta(x')|^2}}.$$

Step 2-1: Now we can define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0<s\leq 1$ in the following way: For any function $\varphi \in L^p(\partial\Omega) = L^p(\partial\Omega, d\sigma)$, we define a function

$$\varphi_\zeta(x') := \varphi(x',\zeta(x')), \quad x' \in \mathbf{R}^{n-1},$$

and let, for $0<s<1$,

$$B^{s,p}(\partial\Omega) = \left\{\varphi \in L^p(\partial\Omega) : \varphi_\zeta \in B^{s,p}(\mathbf{R}^{n-1})\right\}.$$

We equip this space with the norm (the norm (6.11) with $m := 0$)

$$\begin{aligned} &|\varphi|_{B^{s,p}(\partial\Omega)} \\ &= |\varphi_\zeta|_{B^{s,p}(\mathbf{R}^{n-1})} \\ &= \left(\int_{\mathbf{R}^{n-1}} |\varphi_\zeta(x')|^p\,dx' + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x')-\varphi_\zeta(y')|^p}{|x'-y'|^{(n-1)+ps}}dx'\,dy'\right)^{1/p}. \end{aligned} \tag{6.12}$$

For $s=1$, we let

$$B^{1,p}(\partial\Omega)$$
$= $ the space of (equivalence classes of) functions $\varphi \in L^p(\partial\Omega, d\sigma)$ for which
$$\iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x'+h')-2\varphi_\zeta(x')+\varphi_\zeta(x'-h')|^p}{|h'|^{(n-1)+p}}dh'\,dx' < \infty.$$

The space $B^{1,p}(\partial\Omega)$ is a Banach space with respect to the norm

$$\begin{aligned} &|\varphi|_{B^{1,p}(\partial\Omega)} \\ &= |\varphi_\zeta|_{B^{1,p}(\mathbf{R}^{n-1})} \\ &= \left(\int_{\mathbf{R}^{n-1}} |\varphi_\zeta(x')|^p\,dx' \right. \\ &\qquad \left. + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x'+h')-2\varphi_\zeta(x')+\varphi_\zeta(x'-h')|^p}{|h'|^{(n-1)+p}}\,dh'\,dx'\right)^{1/p}. \end{aligned}$$

Step 2-2: If $\kappa(\Omega)$ is a Lipschitz hypograph for some rigid motion $\kappa : \mathbf{R}^n \to \mathbf{R}^n$, then we can define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ in the same way except that

$$\varphi_\zeta(x') := \varphi\left(\kappa^{-1}(x', \zeta(x'))\right), \quad x' \in \mathbf{R}^{n-1}.$$

Step 3: We consider the general case where Ω is a bounded Lipschitz domain. Using the notation of Definition 5.44 (see Fig. 5.10), we choose a partition of unity $\{\phi_j\}_{j=1}^J$ subordinate to the open covering $\{U_j\}_{j=1}^J$ of $\partial\Omega$, that is,

$$\phi_j \in C_0^\infty(U_j),$$
$$0 \leq \phi_j(x) \leq 1 \quad \text{in } U_j,$$
$$\sum_{j=1}^J \phi_j(x) = 1 \quad \text{on } \partial\Omega.$$

Then we define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ as follows:

$$B^{s,p}(\partial\Omega) = \left\{\varphi \in L^p(\partial\Omega) : \phi_j\varphi \in B^{s,p}(\partial\Omega_j),\ 1 \leq j \leq J\right\},$$

where the norm $|\varphi|_{B^{s,p}(\partial\Omega)}$ is defined by the formula

$$|\varphi|_{B^{s,p}(\partial\Omega)} = \sum_{j=1}^J \left|\phi_j\varphi\right|_{B^{s,p}(\partial\Omega_j)}.$$

It should be emphasized that the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ are independent of the open covering $\{U_j\}$ and the partition of unity $\{\phi_j\}$ used.

Step 4: Furthermore, we shall define Besov spaces $B^{s,p}(\partial\Omega)$ for $1 < s < 2$ on a bounded $C^{1,1}$ domain Ω.

Step 4-1: If Ω is a $C^{1,1}$ hypograph defined by formula (5.83) for some function $\zeta \in C^{1,1}(\mathbf{R}^{n-1})$, then we define the Besov spaces $B^{s,p}(\partial\Omega)$ for $s = 1 + \theta$ with $0 < \theta < 1$ in the same way by replacing the norm (6.12) by the norm (the norm (6.11) with $m := 1$)

$$\begin{aligned}
&|\varphi|_{B^{s,p}(\partial\Omega)} \\
&= \left|\varphi_\zeta\right|_{W^{s,p}(\mathbf{R}^{n-1})} \\
&= \left(\sum_{|\alpha|\leq 1} \int_{\mathbf{R}^{n-1}} |D^\alpha\varphi_\zeta(x')|^p\,dx' \right. \\
&\quad \left. + \sum_{|\alpha|=1} \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|D^\alpha\varphi_\zeta(x') - D^\alpha\varphi_\zeta(y')|^p}{|x'-y'|^{(n-1)+p\theta}}\,dx'\,dy'\right)^{1/p}.
\end{aligned}$$

Step 4-2: If Ω is a bounded $C^{1,1}$ domain, then the Besov spaces $B^{s,p}(\partial\Omega)$ for $1 < s < 2$ are defined to be locally the Besov spaces $B^{s,p}(\partial\Omega_j)$, $1 \leq j \leq J$, just as in Step 3. Here it should be emphasized that the boundary $\partial\Omega$ is an $(n-1)$-dimensional $C^{1,1}$ submanifold of $\mathbf{R}^n$.

The norm of $B^{s,p}(\partial\Omega)$ for $0 \leq s < 2$ will be denoted by $|\cdot|_{s,p}$.

6.4 Trace Theorems

In this section we prove an important trace theorem which will be used in the study of boundary value problems in the framework of function spaces of L^p type:

Theorem 6.6 (The trace theorem). *Let $1 < p < \infty$. For every function $f \in H^{s,p}(\mathbf{R}^n)$ with $s > 1/p$, the restriction*

$$g := Rf = f|_{\mathbf{R}^{n-1}} \tag{6.13}$$

is well defined almost everywhere in $\mathbf{R}^{n-1}$, and belongs to $B^{s-1/p,p}(\mathbf{R}^{n-1})$. Furthermore, the restriction mapping R so defined is continuous, that is, there exists a positive constant C such that

$$\|Rf\|_{B^{s-1/p,p}(\mathbf{R}^{n-1})} \leq C\|f\|_{H^{s,p}(\mathbf{R}^n)}, \quad f \in H^{s,p}(\mathbf{R}^n). \tag{6.14}$$

Proof. Without loss of generality, we may assume that

$$\frac{1}{p} < s < 1.$$

We let

$$\tilde{x} = (x, \xi) \in \mathbf{R}^n, \quad x = (x_1, x_2, \ldots, x_{n-1}) \in \mathbf{R}^{n-1}, \ \xi \in \mathbf{R}.$$

If $f(x, \xi)$ is a function in $H^{s,p}(\mathbf{R}^n)$, then it can be written in the form (see (6.8) and (6.9))

$$\begin{cases} f = G_s * \varphi, \\ \varphi \in L^p(\mathbf{R}^n), \end{cases} \tag{6.15}$$

with

$$\|f\|_{s,p} = \|\varphi\|_p.$$

Here and in the following, we simply write

$$\|f\|_{s,p} = \|f\|_{H^{s,p}(\mathbf{R}^n)},$$
$$\|\varphi\|_p = \|\varphi\|_{L^p(\mathbf{R}^n)}.$$

Hence we have, by (6.8),

$$g(x) = f(x,0) = \int_{\mathbf{R}^{n-1}} \left(\int_{\mathbf{R}} G_s(z,\xi)\varphi(x-z,\xi)\, d\xi \right) dz, \quad x \in \mathbf{R}^{n-1}.$$

Here we recall that

$$\|g\|_{B^{s-1/p,p}(\mathbf{R}^{n-1})} = \left(\int_{\mathbf{R}^{n-1}} |g(x)|^p\, dx + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|g(x-y)-g(x)|^p}{|y|^{(n-1)+(s-1/p)p}}\, dx\, dy \right)^{1/p}.$$

Therefore, to prove inequality (6.14) it suffices to show the following two inequalities (6.16) and (6.17):

$$\|g\|_p \leq A\|f\|_{s,p}. \tag{6.16}$$

$$\left(\int_{\mathbf{R}} \frac{\|g(\cdot - y) - g(\cdot)\|_p^p}{|y|^{n-2+sp}}\, dy \right)^{1/p} \leq B\|f\|_{s,p}. \tag{6.17}$$

Step 1: First, we prove inequality (6.16). By Hölder's inequality [Fo2, Theorem 6.2], it follows that

$$\left| \int_{\mathbf{R}} G_s(z,\xi)\varphi(x-z,\xi)\, d\xi \right| \leq \|G_s(z,\cdot)\|_q\, \|\varphi(x-z,\cdot)\|_p, \quad x,\, z \in \mathbf{R}^{n-1},$$

where

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

Hence we have the inequality

$$\begin{aligned} |g(x)| &= \left| \int_{\mathbf{R}^{n-1}} \left(\int_{\mathbf{R}} G_s(z,\xi)\varphi(x-z,\xi)\, d\xi \right) dz \right| \\ &\leq \int_{\mathbf{R}^{n-1}} \left| \int_{\mathbf{R}} G_s(z,\xi)\varphi(x-z,\xi)\, d\xi \right| dz \\ &\leq \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q\, \|\varphi(x-z,\cdot)\|_p\, dz. \end{aligned}$$

Moreover, by applying Minkowski's inequality for integrals [Fo2, Theorem 6.19] we obtain that

$$\begin{aligned} \|g\|_p &\leq \left\| \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, \|\varphi(x-z,\cdot)\|_p \, dz \right\|_p \\ &\leq \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, \|\varphi(-z,\cdot)\|_p \, dz \\ &= \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, \|\varphi(\cdot,\cdot)\|_p \, dz \\ &\leq \|\varphi\|_p \left(\int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, dz \right) \\ &= A \, \|f\|_{s,p}, \end{aligned}$$

with

$$A := \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, dz.$$

In order to prove inequality (6.16), we are reduced to the following lemma:

Lemma 6.7. *Let* $1/p < s < 1$. *Then we have*

$$A = \int_{\mathbf{R}^{n-1}} \|G_s(z,\cdot)\|_q \, dz < \infty. \tag{6.18}$$

Proof. It is known (see Aronszajn–Smith [AS]) that the Bessel kernel $G_s(x)$ can be expressed as follows:

$$\begin{aligned} &G_s(z,\xi) \qquad (6.19) \\ &= \frac{1}{2^{(n+s-2)/2}\, \pi^{n/2}\, \Gamma(n/2)} \, K_{(n-s)/2}\left(\sqrt{|z|^2+\xi^2}\right) \left(|z|^2+\xi^2\right)^{(s-n)/4}, \\ &\quad (z,\xi) \in \mathbf{R}^{n-1} \times \mathbf{R}, \end{aligned}$$

where K_ν is the modified Bessel function of the third kind (Watson [Wt]). This proves that $G_s(x)$ has the asymptotics

$$G_s(z,\xi) \sim \frac{\Gamma((n-s)/2)}{2^s \pi^{n/2} \Gamma(n/2)} \left(\left(|z|^2+\xi^2\right)^{(s-n)/2} \right) \quad \text{as } |z|+|\xi| \to 0, \tag{6.20a}$$

$$G_s(z,\xi) \sim \frac{1}{2^{(n+s-1)/2}\, \pi^{(n-1)/2} \Gamma(n/2)} \left(|z|^2+\xi^2\right)^{(s-n-1)/4} \times e^{-\sqrt{|z|^2+\xi^2}} \quad \text{as } |z|+|\xi| \to \infty. \tag{6.20b}$$

(a) First, we consider the case where $0 < |z| \leq 1$.

(a-1) If $|\xi| \geq 1$, then we have, by estimate (6.20b),

$$|G_s(z, \xi)| \leq C_1 e^{\sqrt{-|z|^2+|\xi|^2}} \leq C_1 e^{-|\xi|}, \tag{6.21}$$

since $|z|^2 + |\xi|^2 \geq 1$.

(a-2) If $|\xi| \leq 1$, then we have, by estimate (6.20a),

$$|G_s(z, \xi)| \leq C_2 \left(|z|^2 + |\xi|^2\right)^{(s-n)/2}. \tag{6.22}$$

However, it is easy to see that

$$\begin{aligned}
&\left(\int_{|\xi|\leq 1} (|z|^2 + |\xi|^2)^{(s-n)q/2}\, d\xi\right)^{1/q} \\
&= \left(\int_{\mathbf{R}} (|z|^2 + |\xi|^2\eta^2)^{(s-n)q/2} |z|\, d\eta\right)^{1/q} \\
&= \left(\int_{\mathbf{R}} (1 + \eta^2)^{(s-n)q/2} |z|^{(-n+s)q+1}\, d\eta\right)^{1/q} \\
&= |z|^{-n+s+1/q} \left(\int_{\mathbf{R}} (1 + \eta^2)^{-(n-s)q/2}\, d\eta\right)^{1/q},
\end{aligned}$$

and further that

$$\int_{\mathbf{R}} (1 + \eta^2)^{-(n-s)q/2}\, d\eta < \infty,$$

since we have the inequality

$$(n - s)q > (n - 1)\frac{p}{p - 1} > \frac{p}{p - 1} > 1.$$

Therefore, by using Hölder's inequality [Fo2, Theorem 6.2] we obtain from inequalities (6.21) and (6.22) that

$$\begin{aligned}
&\int_{0<|z|\leq 1} \|G_s(z, \cdot)\|_q\, dz \\
&\leq C_1 \int_{0<|z|\leq 1} \left(\int_{|\xi|\geq 1} e^{-q|\xi|}\, d\xi\right)^{1/q} dz \\
&\quad + C_2 \int_{0<|z|\leq 1} \left(\int_{|\xi|\leq 1} (|z|^2 + |\xi|^2)^{(s-n)q/2}\, d\xi\right)^{1/q} dz
\end{aligned} \tag{6.23}$$

$$
\begin{aligned}
&= C_3 + C_4 \int_{0<|z|\leq 1} |z|^{s-n+1/q}\, dz \\
&= C_3 + C_4 \int_0^1 \int_{\Sigma_{n-1}} r^{n-2} r^{-n+s+1-1/p}\, dr\, d\sigma \\
&= C_3 + C_4\omega_{n-1} \int_0^1 r^{s-1-1/p}\, dr = C_3 + C_4\omega_{n-1} \left[\frac{r^{s-1/p}}{s-1/p} \right]_0^1 \\
&= C_3 + C_4\omega_{n-1} \left(\frac{1}{s-1/p} \right) < \infty, \quad \frac{1}{p} < s < 1,
\end{aligned}
$$

where Σ_{n-1} is the unit sphere in $\mathbf{R}^{n-1}$ and ω_{n-1} is its surface area

$$
\omega_{n-1} := \frac{2\pi^{(n-1)/2}}{\Gamma((n-1)/2)}.
$$

(b) Next we consider the case where $|z| > 1$. Since we have, by estimate (6.20b),

$$
|G_s(z,\xi)| \leq C_5 e^{-\frac{1}{2}\sqrt{|z|^2+|\xi|^2}} e^{-\frac{1}{2}\sqrt{|z|^2+|\xi|^2}} \leq C_5 e^{-\frac{1}{2}|z|} e^{-\frac{1}{2}|\xi|},
$$

it follows from an application of Hölder's inequality [Fo2, Theorem 6.2] that

$$
\begin{aligned}
\int_{|z|>1} \|G_s(z,\xi)\|_q\, dz &\leq C_5 \int_{1<|z|} \left(\int_{\mathbf{R}} e^{-\frac{1}{2}q|z|} e^{-\frac{1}{2}q|\xi|}\, d\xi \right)^{1/q} dz \qquad (6.24) \\
&= C_5 \left(\int_{|z|>1} e^{-\frac{1}{2}|z|}\, dz \right) \left(\int_{\mathbf{R}} e^{-\frac{1}{2}q|\xi|}\, d\xi \right)^{1/q} \\
&< \infty.
\end{aligned}
$$

Therefore, the desired assertion (6.18) follows by combining the two inequalities (6.23) and (6.24).

The proof of Lemma 6.7 is complete.

Step 2: Secondly, we prove inequality (6.17). If we let

$$
g_\xi(x) := \int_{\mathbf{R}^{n-1}} G_s(z,\xi)\varphi(x-z,\xi)\, dz,
$$

then we can write the function $g(x)$ in the form

$$
g(x) = \int_{\mathbf{R}} \int_{\mathbf{R}^{n-1}} G_s(z,\xi)\varphi(x-z,\xi)\, dz d\xi = \int_{\mathbf{R}} g_\xi(x)\, d\xi.
$$

Hence we have

$$g(x-y)-g(x)=\int_{\mathbf{R}}\big(g_\xi(x-y)-g_\xi(x)\big)\,d\xi.$$

On the other hand, it is easy to verify that

$$\begin{aligned}
&g_\xi(x-y)-g_\xi(x)\\
&=\int_{\mathbf{R}^{n-1}}[\varphi(x-y-z,\xi)-\varphi(x-z,\xi)]\,G_s(z,\xi)\,dz\\
&=\int_{\mathbf{R}^{n-1}}\varphi(x-w,\xi)G_s(w-y,\xi)\,dw\\
&\quad-\int_{\mathbf{R}^{n-1}}\varphi(x-z,\xi)G_s(z,\xi)\,dz\\
&=\int_{\mathbf{R}^{n-1}}\varphi(x-z,\xi)\,[G_s(z-y,\xi)-G_s(z,\xi)]\,dz.
\end{aligned}$$

Thus, by applying Minkowski's inequality for integrals [Fo2, Theorem 6.19] we obtain that

$$\begin{aligned}
&\|g_\xi(\cdot-y)-g_\xi(\cdot)\|_p\\
&\leq\int_{\mathbf{R}^{n-1}}\|\varphi(\cdot-z,\xi)\|_p\;|G_s(z-y,\xi)-G_s(z,\xi)|\,dz\\
&=\int_{\mathbf{R}^{n-1}}\|\varphi(\cdot,\xi)\|_p\;|G_s(z-y,\xi)-G_s(z,\xi)|\,dz\\
&=\|\varphi(\cdot,\xi)\|_p\int_{\mathbf{R}^{n-1}}|G_s(z-y,\xi)-G_s(z,\xi)|\,dz.
\end{aligned}$$

In order to estimate the last integral

$$\int_{\mathbf{R}^{n-1}}|G_s(z-y,\xi)-G_s(z,\xi)|\,dz,$$

we decompose it into the two terms

$$\begin{aligned}
&\int_{\mathbf{R}^{n-1}}|G_s(z-y,\xi)-G_s(z,\xi)|\,dz\\
&\leq\int_{|z|\leq 2|y|}|G_s(z-y,\xi)-G_s(z,\xi)|\,dz\\
&\quad+\int_{|z|\geq 2|y|}|G_s(z-y,\xi)-G_s(z,\xi)|\,dz\\
&:=I(y,\xi)+J(y,\xi).
\end{aligned}$$

Step 2-1: The estimate of the term $I(y, \xi)$.

(a) If $|y| \geq |\xi|$, it follows that

$$\begin{aligned}
&\int_{|z| \leq 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \\
&\leq \int_{|z| \leq 2|y|} |G_s(z-y,\xi)|\, dz + \int_{|z| \leq 2|y|} |G_s(z,\xi)|\, dz \\
&\leq 2 \int_{|z| \leq 3|y|} |G_s(z,\xi)|\, dz.
\end{aligned}$$

Since we have, by estimate (6.20a),

$$|G_s(z,\xi)| \leq C \left(|z|^2 + |\xi|^2\right)^{(s-n)/2} \leq C' \left(|z| + |\xi|\right)^{s-n},$$

we obtain that

$$\begin{aligned}
\int_{|z| \leq 3|y|} |G_s(z,\xi)|\, dz &\leq C' \int_{|z| \leq 3|y|} |\xi|^{s-n} \left(1 + \frac{|z|}{|\xi|}\right)^{s-n} dz \\
&\leq C' |\xi|^{s-n} \int_{\mathbf{R}^{n-1}} (1 + |w|)^{s-n} |\xi|^{n-1}\, dw \\
&= C' |\xi|^{s-1} \left(\int_{\mathbf{R}^{n-1}} \frac{1}{(1+|w|)^{n-1+(1-s)}}\, dw \right) \\
&= C'' |\xi|^{s-1}.
\end{aligned}$$

(b) We consider the case where $|\xi| \geq |y|$.
(b-1) If $|y| \leq |\xi|/2$, then, by the mean value theorem it follows that, for $0 < \theta < 1$,

$$\begin{aligned}
&|G_s(z-y,\xi) - G_s(z,\xi)| \\
&= |G(z+(-y),\xi) - G_s(z,\xi)| \\
&\leq \left| \frac{\partial G_s}{\partial x_1}(z_1 - \theta y_1, z_2 - y_2, \ldots, z_{n-1} - y_{n-1}, \xi) \right| |y_1| \\
&\quad + \left| \frac{\partial G_s}{\partial x_2}(z_1, z_2 - \theta y_2, z_3 - y_3, \ldots, z_{n-1} - y_{n-1}, \xi) \right| |y_2| + \cdots \\
&\quad + \left| \frac{\partial G_s}{\partial x_{n-1}}(z_1, z_2, \ldots, z_{n-2}, z_{n-1} - \theta y_{n-1}, \xi) \right| |y_{n-1}|.
\end{aligned}$$

However, we have, by estimates (6.20),

$$\begin{aligned}&\left|\frac{\partial G_s}{\partial x_1}(z_1-\theta y_1, z_2-y_2,\ldots,z_{n-1}-y_{n-1},\xi)\right| |y_1| \\ &\leq C\left(|z_1-\theta y_1|+|z_2-y_2|+\cdots+|z_{n-1}-y_{n-1}|+|\xi|\right)^{s-n-1},\end{aligned}$$

and, for $0<\theta<1$ and $|y|\leq|\xi|/2$,

$$\begin{aligned}&|z_1-\theta y_1|+|z_2-y_2|+\ldots+|z_{n-1}-y_{n-1}|+|\xi| \\ &\geq |z|-|y|+|\xi|\geq|\xi|\left(\frac{|z|}{|\xi|}+1-\frac{|y|}{|\xi|}\right) \\ &\geq |\xi|\left(\frac{|z|}{|\xi|}+\frac{1}{2}\right).\end{aligned}$$

Similarly, we have, for the partial derivative $\frac{\partial G_s}{\partial x_2}$,

$$\begin{aligned}&\left|\frac{\partial G_s}{\partial x_2}(z_1, z_2-\theta y_2, z_3-y_3,\ldots,z_{n-1}-y_{n-1},\xi)\right| \\ &\leq C\left(|z_1|+|z_2-\theta y_2|+|z_3-y_3|+\cdots+|z_{n-1}-y_{n-1}|+|\xi|\right)^{s-n-1},\end{aligned}$$

and

$$\begin{aligned}&|z_1|+|z_2-\theta y_2|+|z_3-y_3|+\cdots+|z_{n-1}-y_{n-1}|+|\xi| \\ &\geq |\xi|\left(\frac{|z|}{|\xi|}+\frac{1}{2}\right),\end{aligned}$$

and, for the partial derivative $\frac{\partial G_s}{\partial x_{n-1}}$,

$$\begin{aligned}&\left|\frac{\partial G_s}{\partial x_{n-1}}(z_1, z_2,\ldots,z_{n-2}, z_{n-1}-\theta y_{n-1},\xi)\right| \\ &\leq C\left(|z_1|+|z_2|+\cdots+|z_{n-2}|+|z_{n-1}-\theta y_{n-1}|+|\xi|\right)^{s-n-1},\end{aligned}$$

and

$$|z_1|+|z_2|+\cdots+|z_{n-2}+|z_{n-1}-\theta y_{n-1}|+|\xi|\geq|\xi|\left(\frac{|z|}{|\xi|}+\frac{1}{2}\right).$$

Therefore, we obtain that

$$\begin{aligned}
&\int_{|z|\le 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\,dz \\
&\le C\,|y|\,|\xi|^{s-n-1} \int_{|z|\le 2|y|} \left(\frac{|z|}{|\xi|} + \frac{1}{2}\right)^{s-n-1} dz \\
&\le C'|\xi|^{s-n-1}|y| \int_{\mathbf{R}^{n-1}} \left(|w| + \frac{1}{2}\right)^{s-n-1} |\xi|^{n-1}\,dw \\
&= C'|\xi|^{s-2}|y| \int_{\mathbf{R}^{n-1}} \frac{1}{(|w|+1/2)^{(n-1)+(2-s)}} dw \\
&= C''|\xi|^{s-2}|y|.
\end{aligned}$$

(b-2) If $|\xi|/2 \le |y| \le |\xi|$, just as in the case (b-1) we obtain that

$$\begin{aligned}
\int_{|z|\le 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\,dz &\le 2\int_{|z|\le 3|y|} |G_s(z,\xi)|\,dz \\
&\le C\;|\xi|^{s-1}.
\end{aligned}$$

However, since we have the inequality

$$|\xi|^{s-1} = |y|\;|\xi|^{s-2}\left(\frac{|\xi|}{|y|}\right) \le 2\;|y|\;|\xi|^{s-2},$$

it follows that, for $|y| \le |\xi|$,

$$\int_{|z|\le 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\,dz \le C\;|y|\;|\xi|^{s-2}.$$

Summing up, we have proved that

$$\begin{aligned}
I(y,\xi) &:= \int_{|z|\le 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\,dz \\
&\le \begin{cases} C\;|\xi|^{s-1} & \text{if } |y| \ge |\xi|, \\ C\;|y|\,|\xi|^{s-2} & \text{if } |y| \le |\xi|. \end{cases}
\end{aligned} \tag{6.25}$$

Step 2-2: The estimate of the term $J(y,\xi)$. Just as in Step (b-1), we have, by the mean value theorem,

$$|G_s(z-y,\xi) - G_s(z,\xi)| \le C\left(\frac{1}{2}|z| + |\xi|\right)^{s-n-1} |y|, \quad \frac{1}{p} < s < 1. \tag{6.26}$$

(a) We consider the case where $|y| \geq |\xi|$: Since we have $|z| \geq 2|y|$, it follows that

$$\frac{1}{2}|z| + |\xi| \geq \frac{1}{2}|z| = \frac{1}{2}|y|\frac{|z|}{|y|} = \frac{1}{2}|y|\left(\frac{1}{2}\frac{|z|}{|y|} + \frac{1}{2}\frac{|z|}{|y|}\right)$$
$$\geq \frac{1}{2}|y|\left(1 + \frac{1}{2}\frac{|z|}{|y|}\right).$$

Hence, by using inequality (6.26) we obtain that

$$\int_{|z|\geq 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \tag{6.27}$$
$$\leq C\,|y| \int_{|z|\geq 2|y|} \left(\frac{1}{2}|z| + |\xi|\right)^{s-n-1} dz$$
$$\leq 2^{n+1-s} C\,|y|^{s-n} \int_{|z|\geq 2|y|} \left(\frac{1}{2}\frac{|z|}{|y|} + 1\right)^{s-n-1} dz$$
$$= 2^{n+1-s} C\,|y|^{s-n} \int_{|w|\geq 2} \frac{1}{(|w|/2+1)^{n-1+(2-s)}}\,|y|^{n-1}\, dw$$
$$= C'|y|^{s-1}.$$

However, we note that

$$|y|^{s-1} = \left(\frac{1}{|y|}\right)^{1-s} \leq \left(\frac{1}{|\xi|}\right)^{1-s} = |\xi|^{s-1}, \quad \frac{1}{p} < s < 1.$$

Therefore, we have, by inequality (6.27),

$$\int_{|z|\geq 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \leq C'|\xi|^{s-1}. \tag{6.28}$$

(b) We consider the case where $|\xi| \leq |y|$: By inequality (6.26), it follows that

$$\int_{|z|\geq 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \tag{6.29}$$
$$\leq C\,|y| \int_{|z|\geq 2|y|} \frac{1}{(|z|/2 + |\xi|)^{s-n+1}} dz$$
$$= C\,|y| \int_{|z|\geq 2|y|} \frac{1}{|\xi|^{n-s+1}\,(|z|/(2|\xi|)+1)^{n-s+1}} dz$$
$$\leq 2^{n-1} C\,|y| \int_{\mathbf{R}^{n-1}} |\xi|^{s-n-1}|\xi|^{n-1} \frac{1}{(1+|w|)^{n-1+(2-s)}} dw$$
$$= C'\,|y|\,|\xi|^{s-2}.$$

By combining inequalities (6.28) and (6.29), we have proved that

$$J(y,\xi) := \int_{|z|\geq 2|y|} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \le \begin{cases} C\ |\xi|^{s-1} & \text{if } |y| \geq |\xi|, \\ C\ |y|\,|\xi|^{s-2} & \text{if } |y| \leq |\xi|. \end{cases} \tag{6.30}$$

By combining inequalities (6.25) and (6.30), we obtain that, for any $y \in \mathbf{R}^{n-1}$,

$$\begin{aligned} &\|g(\cdot - y) - g(\cdot)\|_p \\ &\leq \int_{\mathbf{R}} \|g_\xi(\cdot - y) - g_\xi(\cdot)\|_p \, d\xi \\ &\leq \int_{\mathbf{R}} \|\varphi(\cdot,\xi)\|_p \left(\int_{\mathbf{R}^{n-1}} |G_s(z-y,\xi) - G_s(z,\xi)|\ dz \right) d\xi \\ &= \int_{|\xi|\leq|y|} \|\varphi(\cdot,\xi)\|_p \left(\int_{\mathbf{R}^{n-1}} |G_s(z-y,\xi) - G_s(z,\xi)|\ dz \right) d\xi \\ &\quad + \int_{|\xi|\geq|y|} \|\varphi(\cdot,\xi)\|_p \left(\int_{\mathbf{R}^{n-1}} |G_s(z-y,\xi) - G_s(z,\xi)|\, dz \right) d\xi \\ &\leq A \left[\int_{|\xi|\leq|y|} \|\varphi(\cdot,\xi)\|_p |\xi|^{s-1} d\xi + \int_{|\xi|\geq|y|} \|\varphi(\cdot,\xi)\|_p |y|\ |\xi|^{s-2}\, d\xi \right]. \end{aligned}$$

Hence, we have, for any $y \in \mathbf{R}^{n-1} \setminus \{0\}$ and any $\xi \in \mathbf{R}$,

$$\begin{aligned} \frac{\|g(\cdot - y) - g(\cdot)\|_p}{|y|^{s+(n-2)/p}} \leq A \Bigg[&\int_{|\xi|\leq|y|} \frac{|\xi|^{-1+s}}{|y|^{s+(n-2)/p}} \|\varphi(\cdot,\xi)\|_p \, d\xi \\ &+ \int_{|\xi|\leq|y|} \frac{|y|\ |\xi|^{-2+s}}{|y|^{s+(n-2)/p}} \|\varphi(\cdot,\xi)\|_p \, d\xi \Bigg] \end{aligned}$$

Furthermore, by using polar coordinates we find that

$$\begin{aligned} &\int_{\mathbf{R}^{n-1}} \left(\frac{\|g(\cdot - y) - g(\cdot)\|_p}{|y|^{s+(n-2)/p}} \right)^p dy \\ &\leq 2^{p-1} A^p \left[\int_{\mathbf{R}^{n-1}} \left(\int_{|\xi|\leq|y|} \frac{|\xi|^{-1+s}}{|y|^{s+(n-2)/p}} \|\varphi(\cdot,\xi)\|_p \, d\xi \right)^p dy \right] \\ &\quad + 2^{p-1} A^p \left[\int_{\mathbf{R}^{n-1}} \left(\int_{|\xi|\leq|y|} \frac{|y|\ |\xi|^{-2+s}}{|y|^{s+(n-2)/p}} \|\varphi(\cdot,\xi)\|_p \, d\xi \right)^p dy \right] \end{aligned} \tag{6.31}$$

$$= 2^{p-1}A^p\left[\int_0^\infty \int_{\Sigma_{n-1}} \left(\int_{|\xi|\le r} \frac{|\xi|^{-1+s}}{r^{s+(n-2)/p}}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p r^{n-2}\,dr\,d\sigma\right]$$
$$+2^{p-1}A^p\left[\int_0^\infty \int_{\Sigma_{n-1}} \left(\int_{|\xi|\ge r} \frac{r\,|\xi|^{-2+s}}{r^{s+(n-2)/p}}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p r^{n-2}\,dr\,d\sigma\right].$$

Step 2-3: Now we estimate the last two integrals in inequality (6.31).

(a) If $\xi > 0$, it follows that

$$\int_0^\infty \frac{1}{r^{n-2+sp}}\left(\int_0^r |\xi|^{-1+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p dr$$
$$+\int_0^\infty \frac{r^p}{r^{n-2+sp}}\left(\int_r^\infty |\xi|^{-2+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p dr$$
$$= \int_0^\infty \left(r^{-s+1/p}\right)^p \left(\int_0^r \xi^{-1+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p \frac{dr}{r}$$
$$+\int_0^\infty \left(r^{1-s+1/p}\right)^p \left(\int_r^\infty \xi^{-2+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p \frac{dr}{r}$$
$$:= I + II.$$

However, for the integral I, by applying Hardy's inequality (Theorem 6.2) with $\gamma := -s + 1/p < 0$ we obtain that

$$I \le A\int_0^\infty \left(\xi^{-s+1/p+1}\xi^{-1+s}\|\varphi(\cdot,\xi)\|_p\right)^p \frac{d\xi}{\xi} \tag{6.32}$$
$$= A\int_0^\infty \|\varphi(\cdot,\xi)\|_p^p\,d\xi.$$

Similarly, for the integral II, by applying Hardy's inequality (Theorem 6.2) with $\gamma := 1 - s + 1/p > 0$ we obtain that

$$II \le A'\int_0^\infty \left(\xi^{1-s+1/p+1}\xi^{-2+s}\|\varphi(\cdot,\xi)\|_p\right)^p \frac{d\xi}{\xi} \tag{6.33}$$
$$= A'\int_0^\infty \|\varphi(\cdot,\xi)\|_p^p\,d\xi.$$

(b) If $\xi < 0$, it follows that

$$\int_0^\infty \left(r^{-s+1/p}\right)^p \left(\int_{-r}^0 |\xi|^{-1+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p \frac{dr}{r}$$
$$+\int_0^\infty \left(r^{1-s+1/p}\right)^p \left(\int_{-\infty}^{-r} |\xi|^{-2+s}\|\varphi(\cdot,\xi)\|_p\,d\xi\right)^p \frac{dr}{r},$$

so that, with $\eta := -\xi$,

$$\int_0^\infty \left(r^{-s+1/p}\right)^p \left(\int_0^r \eta^{-1+s} \|\varphi(\cdot, -\eta)\|_p \, d\eta\right)^p \frac{dr}{r}$$
$$+ \int_0^\infty \left(r^{1-s+1/p}\right)^p \left(\int_r^\infty \eta^{-2+s} \|\varphi(\cdot, -\eta)\|_p \, d\eta\right)^p \frac{dr}{r}$$
$$:= III + IV.$$

However, for the integral III, by applying Hardy's inequality (Theorem 6.2) with $\gamma := -s + 1/p < 0$ we obtain that

$$\begin{aligned} III &\leq A \int_0^\infty \left(\eta^{1-s+1/p} \eta^{-1+s} \|\varphi(\cdot, -\eta)\|_p\right)^p \frac{d\eta}{\eta} \\ &= A'' \int_0^\infty \|\varphi(\cdot, -\eta)\|_p \, d\eta = A'' \int_{-\infty}^0 \|\varphi(\cdot, \xi)\|_p \, d\xi. \end{aligned} \tag{6.34}$$

Similarly, for the integral IV, by applying Hardy's inequality (Theorem 6.2) with $\gamma := 1 - s + 1/p > 0$ we obtain that

$$\begin{aligned} IV &\leq A''' \int_0^\infty \left(\eta^{1-s+1/p+1} \eta^{-2+s} \|\varphi(\cdot, -\eta)\|_p\right)^p \frac{d\eta}{\eta} \\ &= A''' \int_0^\infty \|\varphi(\cdot, -\eta)\|_p \, d\eta = A''' \int_{-\infty}^0 \|\varphi(\cdot, \xi)\|_p \, d\xi. \end{aligned} \tag{6.35}$$

Step 2-3: Therefore, by combining inequalities (6.32)–(6.35) we have proved that

$$\int_{\mathbf{R}^{n-1}} \frac{\|g(\cdot - y) - g(\cdot)\|_p^p}{|y|^{n-2+sp}} \, dy \leq C \int_{-\infty}^\infty \|\varphi(\cdot, \xi)\|_p^p \, d\xi = C \|\varphi\|_p^p = C \|f\|_{s,p}^p.$$

The proof of inequality (6.17) and hence that of inequality (6.14) is complete.

Under certain hypotheses on the domain Ω, functions in Sobolev spaces $H^{s,p}(\Omega)$ may be extended as functions in $H^{s,p}(\mathbf{R}^n)$. In this way, the trace theorem (Theorem 6.6) remains valid for $H^{s,p}(\Omega)$ and $B^{s-1/p,p}(\partial\Omega)$. More precisely, we have the following trace theorem (see Adams–Fournier [AF, Remarks 7.45]):

Theorem 6.8 (the trace theorem). *Let Ω be a bounded $C^{1,1}$ domain of $\mathbf{R}^n$. If $1 < p < \infty$, then the trace map*

$$\begin{gathered} \gamma = (\gamma_0, \gamma_1) : W^{2,p}(\Omega) \longrightarrow B^{2-1/p,p}(\partial\Omega) \oplus B^{1-1/p,p}(\partial\Omega) \\ u \longmapsto \left(u|_{\partial\Omega}, \left.\frac{\partial u}{\partial \mathbf{n}}\right|_{\partial\Omega}\right) \end{gathered}$$

is continuous. Here

$$\mathbf{n} = -\boldsymbol{\nu} = \frac{(\nabla\zeta(x'), -1)}{\sqrt{1 + |\nabla\zeta(x')|^2}}$$

is the unit interior normal to the boundary $\partial\Omega$ (see Fig. 6.1).

Indeed, it suffices to note that we have, by Theorem 6.5,

$$H^{2,p}(\Omega) = W^{2,p}(\Omega).$$

6.5 Notes and Comments

Our treatment of Sobolev and Besov spaces is adapted from Adams–Fournier [AF], Stein [Sn2] and Taibleson [Tb]. For more thorough treatments of function spaces, the reader might be referred to Adams–Fournier [AF], Aronszajn–Smith [AS], Bergh–Löfström [BL], Calderón [Ca], Stein [Sn2], Taibleson [Tb] and Triebel [Tr].

Section 6.1: Theorem 6.1 is taken from Folland [Fo2, Theorem 6.20] and Example 6.3 is taken from Folland [Fo2, Corollary 6.21].

Section 6.2: Theorem 6.4 is adapted from Stein [Sn2]. For the proof of Theorem 6.5, see Adams–Fournier [AF, Theorem 5.24].

Section 6.3: Theorem 6.6 is adapted from Stein [Sn1].

Section 6.4: Theorem 6.8 is adapted from Stein [Sn1] and [Sn2].

Chapter 7
Theory of Pseudo-differential Operators

In this chapter we present a brief description of the basic concepts and results from the theory of pseudo-differential operators – a modern theory of potentials –which will be used in the subsequent chapters. The development of the theory of pseudo-differential operators has greatly advanced our understanding of partial differential equations, and the pseudo-differential calculus has become an indispensable tool in contemporary analysis, in particular, in the study of elliptic boundary value problems. The calculus of pseudo-differential operators will be applied to elliptic boundary value problems in Chaps. 10 and 11.

The purpose of Sect. 7.1 is to summarize the basic facts about manifolds with boundary and the double of a manifold which are most frequently used in the theory of partial differential equations. Let Ω be a bounded domain of Euclidean space $\mathbf{R}^n$ with smooth boundary $\partial\Omega$. In Sect. 7.1 we formulate two fundamental theorems on smooth manifolds with boundary. The first theorem (Theorem 7.2) states that the boundary $\partial\Omega$ has an open neighborhood in Ω which is diffeomorphic to $\partial\Omega \times [0, 1)$. The second theorem (Theorem 7.3) states that Ω is a submanifold of some n-dimensional, smooth manifold M without boundary. This manifold M is called the double of Ω. In Sect. 7.2 we define the generalized Sobolev spaces $H^{s,p}(M)$ and the Besov spaces $B^{s,p}(\partial\Omega)$ where $M = \hat{\Omega}$ is the double of Ω. In Sect. 7.3 we introduce the Fourier integral distribution

$$K(x) = \int_{\mathbf{R}^N} e^{i\varphi(x,\theta)} a(x, \theta)\, d\theta$$

associated with the phase function $\varphi(x, \theta)$ and the amplitude $a(x, \theta)$. The operator A is called the Fourier integral operator associated with the phase function $\varphi(x, y, \theta)$ and the amplitude $a(x, y, \theta)$ if its distribution kernel $K_A(x, y)$ is given by the Fourier integral distribution

$$K_A(x, y) = \int_{\mathbf{R}^N} e^{i\varphi(x,y,\theta)} a(x, y, \theta)\, d\theta.$$

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_7,

In Sect. 7.4 we define pseudo-differential operators. A pseudo-differential operator of order m is a Fourier integral operator associated with the phase function $\varphi(x, y, \xi) = (x - y) \cdot \xi$ and some amplitude $a(x, y, \xi) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^n)$. In this section we study their basic properties such as the behavior of transposes, adjoints and compositions of such operators, and the effect of a change of coordinates on such operators. It should be emphasized that Theorem 7.18 contains all the machinery necessary for the theory of pseudo-differential operators, and its proof is based on Example 5.29 and the stationary phase theorem. By using the multiplier theorem of Marcinkiewicz just as in Coifman–Meyer [CM], Bourdaud [Bd] proved an L^p boundedness theorem for pseudo-differential operators (Theorem 7.24) which plays a fundamental role throughout the book. A *global version* of Theorem 7.24 will be proved in Appendix A, due to its length.

In Sect. 7.5 we describe the classical surface and volume potentials arising in boundary value problems for elliptic differential operators in terms of pseudo-differential operators (Theorems 7.28 and 7.29). One of the important questions in the theory of elliptic boundary value problems is that of the smoothness of a solution near the boundary. In Sect. 7.6 we introduce a condition about symbols in the normal direction at the boundary (the transmission property) in order to ensure the boundary regularity property. Moreover, it should be noticed that the notion of transmission property is invariant under a change of coordinates which preserves the boundary. Hence this notion can be transferred to manifolds with boundary. Section 7.7 is devoted to the Boutet de Monvel calculus. Elliptic boundary value problems cannot be treated directly by pseudo-differential operator methods. It was Boutet de Monvel who brought in the operator-algebraic aspect with his calculus in 1971. He constructed a relatively small "algebra" which contains the boundary value problems for elliptic differential operators as well as their parametrices. In Sect. 7.8 we prove (Theorem 7.36) that the distribution kernel $s(x, y)$ of a pseudo-differential operator $S \in L^m_{1,0}(\mathbf{R}^n)$ satisfies the estimate

$$|s(x, y)| \leq \frac{C}{|x - y|^{m+n}}, \quad x, y \in \mathbf{R}, \quad x \neq y.$$

7.1 Manifolds with Boundary and the Double of a Manifold

We denote by $\mathbf{R}^n_+$ the open half-space

$$\mathbf{R}^n_+ = \{x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n : x_n > 0\}.$$

We let

$$\overline{\mathbf{R}^n_+} = \{x \in \mathbf{R}^n : x_n \geq 0\},$$

and let U be an open set in $\overline{\mathbf{R}^n_+}$ in the topology induced on $\overline{\mathbf{R}^n_+}$ from $\mathbf{R}^n$. We define the *boundary* ∂U of U to be the intersection of U with $\mathbf{R}^{n-1} \times \{0\}$ and the *interior* $\operatorname{Int} U$ of U to be the complement of ∂U in U, that is,

$$\partial U = U \cap \{x \in \mathbf{R}^n : x_n = 0\},$$
$$\operatorname{Int} U = U \cap \{x \in \mathbf{R}^n : x_n > 0\}.$$

It is easy to see that $\operatorname{Int} U$ is open in U and that ∂U is closed in U, but not in $\mathbf{R}^n$. This inconsistent use of the notation ∂U is temporary.

Let U and V be two open sets in $\overline{\mathbf{R}^n_+}$. We say that a mapping $f : U \to V$ is *of class* C^r $(0 \le r \le \infty)$ if, for each point x of U, there exist a neighborhood U_1 of x in $\mathbf{R}^n$ and a neighborhood V_1 of $f(x)$ in $\mathbf{R}^n$, and a C^r mapping $f_1 : U_1 \to V_1$ such that $f_1|_{U \cap U_1} = f|_{U \cap U_1}$.

Then we have the following lemma:

Lemma 7.1. *Let U, V be open sets in $\overline{\mathbf{R}^n_+}$ and $f : U \to V$ a C^r diffeomorphism with $1 \le r \le \infty$. Then the mapping f induces two C^r diffeomorphisms*

$$\operatorname{Int} f : \operatorname{Int} U \longrightarrow \operatorname{Int} V$$
$$\partial f : \partial U \longrightarrow \partial V.$$

Now we can define a C^r manifold $(1 \le r \le \infty)$ with boundary in the following way: Let M be a set. An *atlas of charts with boundary* on M is a family of pairs $\{(U_i, \varphi_i)\}_{i \in I}$ satisfying the following three conditions (MB1)–(MB3):

(MB1) Each U_i is a subset of M and $M = \cup_{i \in I} U_i$.

(MB2) Each φ_i is a bijection of U_i onto an open subset of $\overline{\mathbf{R}^n_+}$, and for every pair i, j of I with $U_i \cap U_j \neq \emptyset$ the set $\varphi_i(U_i \cap U_j)$ is open in $\overline{\mathbf{R}^n_+}$.

(MB3) For each pair i, j of I with $U_i \cap U_j \neq \emptyset$, the mapping

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)$$

is a C^r diffeomorphism.

Each pair (U_i, φ_i) is called a *chart with boundary* of the atlas.

An n-dimensional C^r *manifold with boundary* is a set M together with an atlas of charts with boundary on M. By virtue of Lemma 7.1, we can define two sets $\operatorname{Int} M$ and ∂M as follows:

$$\operatorname{Int} M = \bigcup_{i \in I} (\operatorname{Int} \varphi_i)^{-1} \left(\operatorname{Int} (\varphi_i(U_i))\right),$$

$$\partial M = \bigcup_{i \in I} (\partial \varphi_i)^{-1} \left(\partial(\varphi_i(U_i))\right).$$

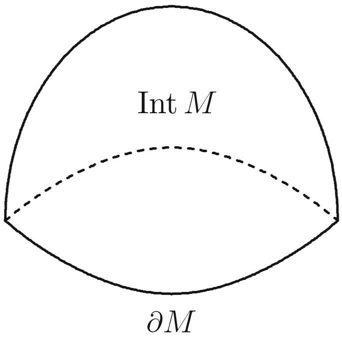


Fig. 7.1 The interior Int M and the boundary ∂M of M

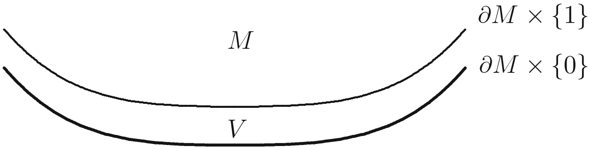


Fig. 7.2 The product neighborhood V of ∂M

We call Int M the *interior* of M and ∂M the *boundary* of M, respectively (see Fig. 7.1). The set Int M is an n-dimensional C^r manifold (without boundary) with atlas obtained from (U_i, φ_i) by replacing $\varphi_i(U_i)$ by Int $(\varphi_i(U_i))$, and the set ∂M is an $(n-1)$-dimensional C^r manifold (without boundary) with atlas obtained from (U_i, φ_i) by replacing $\varphi_i(U_i)$ by $\partial(\varphi_i(U_i))$.

We give two fundamental theorems on smooth manifolds with boundary. The first theorem states that ∂M has an open neighborhood in M which is diffeomorphic to $\partial M \times [0, 1)$ (see Fig. 7.2):

Theorem 7.2 (The product neighborhood theorem). *Let M be an n-dimensional paracompact smooth manifold with boundary ∂M. Then there exists a C^∞ diffeomorphism φ of $\partial M \times [0, 1)$ onto an open neighborhood V of ∂M in M which is the identity map on ∂M: $V \cong \partial M \times [0, 1)$.*

The diffeomorphism φ is called a *collar* for M and the neighborhood V is called a *product neighborhood* of ∂M, respectively.

The second theorem states that M is a submanifold of some n-dimensional, smooth manifold without boundary. Let $M_0 = M \times \{0\}$ and $M_1 = M \times \{1\}$ be two copies of M. The *double* $\hat{M}$ of M is the topological space obtained from the union $M_0 \cup M_1$ by identifying $(x, 0)$ with $(x, 1)$ for each x in ∂M (see Fig. 7.3). By using the product neighborhood theorem (Theorem 7.2), we have the following theorem:

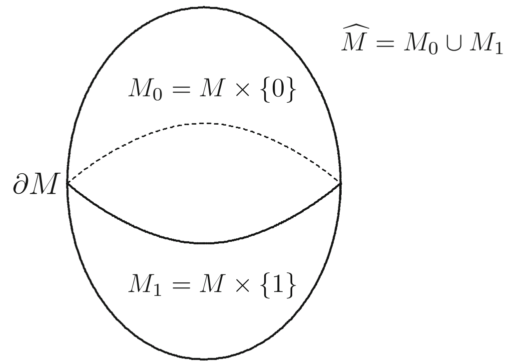


Fig. 7.3 The double $\hat{M}$ of M

Theorem 7.3. *If M is an n-dimensional, paracompact smooth manifold with boundary ∂M, then its double $\hat{M}$ is an n-dimensional, smooth manifold without boundary, and is uniquely determined up to C^∞ diffeomorphisms.*

7.2 Function Spaces

Let Ω be a bounded domain of Euclidean space $\mathbf{R}^n$ with smooth boundary $\partial\Omega$. Its closure $\overline{\Omega} = \Omega \cup \partial\Omega$ is an n-dimensional, compact smooth manifold with boundary. By virtue of Theorems 7.2 and 7.3, we may assume that the following three conditions (a)–(c) are satisfied (see Figs. 7.4 and 7.5):

(a) The domain Ω is a relatively compact, open subset of an n-dimensional compact smooth manifold M without boundary.
(b) In a tubular neighborhood W of $\partial\Omega$ in M a normal coordinate t is chosen so that the points of W are represented as (x', t), $x' \in \partial\Omega$, $-1 < t < 1$; $t > 0$ in Ω, $t < 0$ in $M \setminus \overline{\Omega}$ and $t = 0$ only on $\partial\Omega$.
(c) The manifold M is equipped with a strictly positive density μ which, on W, is the product of a strictly positive density ω on $\partial\Omega$ and the Lebesgue measure dt on $(-1, 1)$. The manifold $M = \hat{\Omega}$ is the double of Ω.

Now we define the generalized Sobolev spaces $H^{s,p}(M)$ and the Besov spaces $B^{s,p}(\partial\Omega)$ where $M = \hat{\Omega}$ is the double of Ω.

First, we introduce distributions which behave locally just like the distributions in $H^{s,p}(\mathbf{R}^n)$:

$$H^{s,p}_{\text{loc}}(\Omega) = \text{ the space of distributions } u \in \mathcal{D}'(\Omega) \text{ such that } \varphi u \in H^{s,p}(\mathbf{R}^n) \text{ for all } \varphi \in C_0^\infty(\Omega).$$

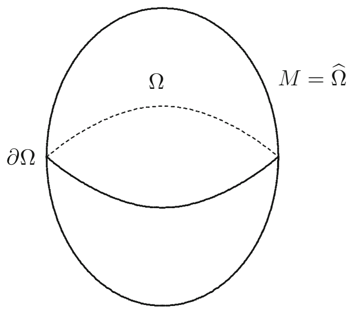


Fig. 7.4 The double $M = \hat{\Omega}$ of Ω

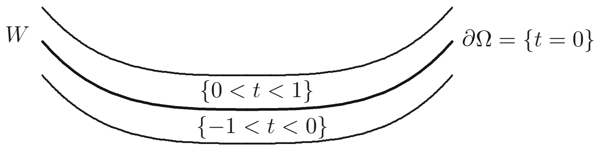


Fig. 7.5 The tubular neighborhood W of $\partial\Omega$

We equip the space $H^{s,p}_{\text{loc}}(\Omega)$ with the topology defined by the seminorms $u \mapsto \|\varphi u\|_{s,p}$ as φ ranges over $C_0^\infty(\Omega)$. It is easy to verify that the *localized Sobolev space* $H^{s,p}_{\text{loc}}(\Omega)$ is a Fréchet space.

The *localized Besov space* $B^{s,p}_{\text{loc}}(\Omega)$ is defined similarly, with $H^{s,p}(\mathbf{R}^n)$ replaced by $B^{s,p}(\mathbf{R}^n)$.

Secondly, we show the invariance of the space $H^{s,p}_{\text{loc}}(\Omega)$ under C^∞ diffeomorphisms. To do this, let Ω_1, Ω_2 be two open subsets of $\mathbf{R}^n$ and $\chi : \Omega_1 \to \Omega_2$ a C^∞ diffeomorphism. If $v \in \mathcal{D}'(\Omega_2)$, then we can define a distribution $\chi^* v \in \mathcal{D}'(\Omega_1)$ by the formula

$$\langle \chi^* v, \varphi \rangle = \left\langle v, \varphi \circ \chi^{-1} \cdot \left|\det\left(J(\chi^{-1})\right)\right| \right\rangle, \quad \varphi \in C_0^\infty(\Omega_1),$$

where $J(\chi^{-1})$ is the Jacobian matrix of χ^{-1}. The distribution $\chi^* v$ is called the inverse image of v under χ.

It is easy to see that the mapping $v \mapsto \chi^* v$ is an *isomorphism* of $H^{s,p}_{\text{loc}}(\Omega_2)$ onto $H^{s,p}_{\text{loc}}(\Omega_1)$ and that its inverse is the mapping $u \mapsto (\chi^{-1})^* u$. The situation can be visualized in the following diagram:

$$\begin{array}{ccc}
(\chi^{-1})^* u \in \mathcal{D}'(\Omega_2) & \xleftarrow{(\chi^{-1})^*} & \mathcal{D}'(\Omega_1) \ni u \\
\uparrow & & \uparrow \\
v \in H^{s,p}_{\text{loc}}(\Omega_2) & \xrightarrow[\chi^*]{} & H^{s,p}_{\text{loc}}(\Omega_1) \ni \chi^* v
\end{array}$$

Thirdly, the Sobolev spaces $H^{s,p}(M)$ on the double $M = \hat{\Omega}$ can be defined to be locally the Sobolev spaces $H^{s,p}(\mathbf{R}^n)$, upon using local coordinate systems flattening out M, together with a partition of unity, in the following way: We let

$$H^{s,p}(M) = \begin{array}{l} \text{the space of distributions } u \in \mathcal{D}'(M) \text{ such that, for any} \\ \text{admissible chart } (U, \chi) \text{ on } M, \text{ the inverse image } (\chi^{-1})^*(u|_U) \\ \text{of } u|_U \text{ under } \chi^{-1} \text{ belongs to } H^{s,p}_{\text{loc}}(\chi(U)). \end{array}$$

We equip the space $H^{s,p}(M)$ with the topology defined by the family of seminorms

$$u \longmapsto \|\tilde{\varphi} \cdot \left(\chi^{-1}\right)^* (u|_U)\|_{s,p}$$

where (U, χ) ranges over all admissible charts on M and $\tilde{\varphi}$ ranges over the space $C_0^\infty(\chi(U))$.

By the compactness of M, we can find an atlas $\{(U_j, \chi_j)\}_{j=1}^N$ consisting of finitely many charts on M. Let $\{\varphi_j\}_{j=1}^N$ be a partition of unity subordinate to the covering $\{U_j\}_{j=1}^N$. Then the norm of $H^{s,p}(M)$ can be defined by the formula

$$\|u\|_{s,p} = \sum_{j=1}^N \left\| \left(\chi_j^{-1}\right)^* (\varphi_j u) \right\|_{s,p},$$

where $\|\cdot\|_{s,p}$ on the right-hand side is the norm of $H^{s,p}(\mathbf{R}^n)$. Hence the Sobolev space $H^{s,p}(M)$ is a Banach space.

The Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ are defined similarly, with $H^{s,p}(\mathbf{R}^n)$ replaced by $B^{s,p}(\mathbf{R}^{n-1})$ (see Sect. 6.3). The norm of $B^{s,p}(\partial\Omega)$ will be denoted by $|\cdot|_{s,p}$.

Finally, we state two important theorems which will be used in the study of elliptic boundary value problems:

(i) (**The trace theorem**) Let $1 < p < \infty$. Then the trace map

$$\begin{aligned} \gamma_0 : H^{s,p}(\Omega) &\longrightarrow B^{s-1/p,p}(\partial\Omega) \\ u &\longmapsto u|_{\partial\Omega} \end{aligned}$$

is continuous for every $s > 1/p$ (Theorem 6.5), and is surjective.

(ii) (**The Rellich–Kondrachov theorem**) If $s > t$, then the injections

$$\begin{aligned} H^{s,p}(M) &\longrightarrow H^{t,p}(M), \\ B^{s,p}(\partial\Omega) &\longrightarrow B^{t,p}(\partial\Omega) \end{aligned}$$

are both compact (or completely continuous).

7.3 Fourier Integral Operators

7.3.1 Symbol Classes

Let Ω be an open subset of $\mathbf{R}^n$. If $m \in \mathbf{R}$ and $0 \le \delta < \rho \le 1$, we let

$$\begin{aligned} S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N) = {} & \text{the set of all functions } a(x,\theta) \in C^\infty(\Omega \times \mathbf{R}^N) \text{ with} \\ & \text{the property that, for any compact } K \subset \Omega \text{ and any} \\ & \text{multi-indices } \alpha, \beta, \text{ there exists a constant } C_{K,\alpha,\beta} > 0 \\ & \text{such that we have, for all } x \in K \text{ and } \theta \in \mathbf{R}^N, \\ & \quad \left|\partial_\theta^\alpha \partial_x^\beta a(x,\theta)\right| \le C_{K,\alpha,\beta}(1+|\theta|)^{m-\rho|\alpha|+\delta|\beta|}. \end{aligned}$$

The elements of $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ are called *symbols* of order m. We drop the $\Omega \times \mathbf{R}^N$ and use $S^m_{\rho,\delta}$ when the context is clear.

If K is a compact subset of Ω and j is a non-negative integer, we define a seminorm $p_{K,j,m}$ on $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ by the formula

$$S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N) \ni a \longmapsto p_{K,j,m}(a) = \sup_{\substack{x \in K \\ \theta \in \mathbf{R}^N \\ |\alpha|+|\beta| \le j}} \frac{\left|\partial_\theta^\alpha \partial_x^\beta a(x,\theta)\right|}{(1+|\theta|)^{m-\rho|\alpha|+\delta|\beta|}}.$$

We equip the space $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ with the topology defined by the family $\{p_{K,j,m}\}$ of seminorms where K ranges over all compact subsets of Ω and $j = 0, 1, \ldots$. It is easy to see that the space $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ is a Fréchet space.

We give some simple examples of symbols:

Example 7.4. (1) A polynomial $p(x,\xi) = \sum_{|\alpha| \le m} a_\alpha(x)\xi^\alpha$ of order m with coefficients $a_\alpha(x)$ in $C^\infty(\Omega)$ is in the class $S^m_{1,0}(\Omega \times \mathbf{R}^n)$.

(2) If $m \in \mathbf{R}$, the function

$$\Omega \times \mathbf{R}^n \ni (x,\xi) \longmapsto \langle\xi\rangle^m$$

is in the class $S^m_{1,0}(\Omega \times \mathbf{R}^n)$, where

$$\langle\xi\rangle = (1+|\xi|^2)^{1/2} = \left(1 + \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2\right)^{1/2}.$$

(3) A function $a(x,\theta) \in C^\infty(\Omega \times (\mathbf{R}^N \setminus \{0\}))$ is said to be *positively homogeneous* of degree m in θ if it satisfies the condition

$$a(x,t\theta) = t^m a(x,\theta) \quad \text{for all } t > 0.$$

If $a(x, \theta)$ is positively homogeneous of degree m in θ and if $\varphi(\theta)$ is a smooth function such that $\varphi(\theta) = 0$ for $|\theta| \leq 1/2$ and $\varphi(\theta) = 1$ for $|\theta| \geq 1$, then the function $\varphi(\theta)a(x, \theta)$ is in the class $S^m_{1,0}(\Omega \times \mathbf{R}^N)$.

We set

$$S^{-\infty}(\Omega \times \mathbf{R}^N) = \bigcap_{m \in \mathbf{R}} S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N).$$

For example, every function $\varphi(\xi) \in \mathcal{S}(\mathbf{R}^N)$ is in the class $S^{-\infty}(\Omega \times \mathbf{R}^N)$. More precisely, we have

$$S^{-\infty}(\Omega \times \mathbf{R}^N) = C^\infty(\Omega)\hat{\otimes}_\pi \mathcal{S}(\mathbf{R}^N),$$

where the space $C^\infty(\Omega)\hat{\otimes}_\pi \mathcal{S}(\mathbf{R}^N)$ is the completed π-topology (or projective topology) tensor product of the Fréchet spaces $C^\infty(\Omega)$ and $\mathcal{S}(\mathbf{R}^N)$ (see Schaefer [Sa, Chapter III, Section 6], Treves [Tv, Chapter 45]).

The next theorem gives a meaning to a formal sum of symbols of decreasing order:

Theorem 7.5. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. If $a_j(x, \theta) \in S^{m_j}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ with $m_j \downarrow -\infty$, $j = 0, 1, \ldots$, then there exists a symbol $a(x, \theta) \in S^{m_0}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$, unique modulo $S^{-\infty}(\Omega \times \mathbf{R}^N)$, such that we have, for all $k \geq 1$,*

$$a(x, \theta) - \sum_{j=0}^{k-1} a_j(x, \theta) \in S^{m_k}_{\rho,\delta}(\Omega \times \mathbf{R}^N). \tag{7.1}$$

If formula (7.1) holds true, we write

$$a(x, \theta) \sim \sum_{j=0}^{\infty} a_j(x, \theta).$$

The formal sum $\sum_j a_j(x, \theta)$ is called an *asymptotic expansion* of $a(x, \theta)$.

A symbol $a(x, \theta)$ in $S^m_{1,0}(\Omega \times \mathbf{R}^N)$ is said to be *classical* if there exist smooth functions $a_j(x, \theta)$, positively homogeneous of degree $m - j$ in θ for $|\theta| \geq 1$, such that

$$a(x, \theta) \sim \sum_{j=0}^{\infty} a_j(x, \theta).$$

The homogeneous function $a_0(x, \theta)$ of degree m is called the *principal part* of $a(x, \theta)$.

We let

$$S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^N) = \text{the set of all classical symbols of order } m.$$

It should be emphasized that the subspace $S^m_{\mathrm{cl}}(\mathbf{R}^N)$, defined as the set of all x-independent elements of $S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^N)$, is closed in the induced topology, and we have

$$S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^N) = C^\infty(\Omega)\hat{\otimes}_\pi S^m_{\mathrm{cl}}(\mathbf{R}^N).$$

We give some simple examples of classical symbols:

Example 7.6. The symbols in Examples 7.4 are all classical, and they have respectively as principal part the following functions:

(1) $p_m(x,\xi) = \sum_{|\alpha|=m} a_{\alpha(x)}\xi^\alpha$ for $p(x,\xi) = \sum_{|\alpha|\le m} a_\alpha(x)\xi^\alpha$.
(2) $|\xi|^m$ for $(1+|\xi|^2)^{m/2}$.
(3) $a(x,\theta)$ for $\varphi(\theta)a(x,\theta)$.

A symbol $a(x,\theta)$ in $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ is said to be *elliptic* of order m if, for any compact $K \subset \Omega$, there exists a constant $C_K > 0$ such that

$$|a(x,\theta)| \ge C_K(1+|\theta|)^m, \quad x \in K, \quad |\theta| \ge \frac{1}{C_K}.$$

There is a simple criterion for ellipticity in the case of classical symbols:

Theorem 7.7. *Let $a(x,\theta)$ be in $S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^N)$ with principal part $a_0(x,\theta)$. Then $a(x,\theta)$ is elliptic if and only if we have the condition*

$$a_0(x,\theta) \ne 0 \quad \textit{for all } x \in \Omega \textit{ and } |\theta| = 1.$$

7.3.2 Phase Functions

Let Ω be an open subset of $\mathbf{R}^n$. A function $\varphi(x,\theta)$ in $C^\infty(\Omega \times (\mathbf{R}^N \setminus \{0\}))$ is called a *phase function* on $\Omega \times (\mathbf{R}^N \setminus \{0\})$ if it satisfies the following three conditions (a)–(c):

(a) $\varphi(x,\theta)$ is real-valued.
(b) $\varphi(x,\theta)$ is positively homogeneous of degree one in the variable θ.
(c) The differential $d\varphi$ does not vanish on the space $\Omega \times (\mathbf{R}^N \setminus \{0\})$.

We give some important examples of phase functions:

Example 7.8. (a) Let U be an open subset of $\mathbf{R}^p$ and $\Omega = U \times U$. The function

$$\varphi(x,y,\xi) = (x-y)\cdot\xi = \sum_{k=1}^{p}(x_k - y_k)\xi_k$$

is a phase function on the space $\Omega \times (\mathbf{R}^p \setminus \{0\})$ $(n := 2p,\ N := p)$. The phase function $\varphi(x, y, \xi)$ will be used for pseudo-differential operators (see Sect. 7.4).

(b) Let Ω' be an open subset of $\mathbf{R}^{n-1}$ and $\Omega = (\Omega' \times \mathbf{R}^+) \times \Omega'$. The function

$$\varphi_K(x, y', \xi', \xi_n) = (x' - y') \cdot \xi' + x_n\, \xi_n = \sum_{k=1}^{n-1} (x_k - y_k)\xi_k + x_n\, \xi_n$$

is a phase function on the space $\Omega \times (\mathbf{R}^n \setminus \{0\})$ $(n := 2n-1,\ N := n)$. The phase function $\varphi_K(x, y', \xi', \xi_n)$ will be used for potential operators (see Sect. 7.7.1).

(c) Let Ω' be an open subset of $\mathbf{R}^{n-1}$ and $\Omega = \Omega' \times (\Omega' \times \mathbf{R}^+)$. The function

$$\varphi_T(x', y, \xi', \xi_n) = (x' - y') \cdot \xi' - y_n\, \xi_n$$

is a phase function on the space $\Omega \times (\mathbf{R}^n \setminus \{0\})$ $(n := 2n-1,\ N := n)$. The phase function $\varphi_T(x', y, \xi', \xi_n)$ will be used for trace operators (see Example 7.13(3)).

(d) Let Ω' be an open subset of $\mathbf{R}^{n-1}$ and $\Omega = (\Omega' \times \mathbf{R}^+) \times (\Omega' \times \mathbf{R}^+)$. The function

$$\varphi_B(x, y, \xi', \xi_n, \eta_n) = (x' - y') \cdot \xi' - y_n\, \eta_n + x_n\, \xi_n$$

is a phase function on the space $\Omega \times (\mathbf{R}^{n+1} \setminus \{0\})$ $(n := 2n,\ N := n + 1)$. The phase function $\varphi_B(x, y, \xi', \xi_n, \eta_n)$ will be used for singular Green operators (see Sect. 7.7.1).

The next lemma will play a fundamental role in defining oscillatory integrals:

Lemma 7.9. *If $\varphi(x, \theta)$ is a phase function on $\Omega \times (\mathbf{R}^N \setminus \{0\})$, then there exists a first-order differential operator*

$$L = \sum_{j=1}^{N} a_j(x, \theta) \frac{\partial}{i\, \partial \theta_j} + \sum_{k=1}^{n} b_k(x, \theta) \frac{\partial}{i\, \partial x_k} + c(x, \theta)$$

such that

$$L(e^{i\varphi}) = e^{i\varphi},$$

and its coefficients $a_j(x, \theta)$, $b_k(x, \theta)$, $c(x, \theta)$ enjoy the properties

$$a_j(x, \theta) \in S^0_{1,0}, \quad b_k(x, \theta) \in S^{-1}_{1,0}, \quad c(x, \theta) \in S^{-1}_{1,0}.$$

Furthermore, the transpose L' of L has coefficients $a'_j(x, \theta)$, $b'_k(x, \theta)$, $c'(x, \theta)$ in the same symbol classes as $a_j(x, \theta)$, $b_k(x, \theta)$, $c(x, \theta)$, respectively.

For example, if $\varphi(x, y, \xi)$ is a phase function as in Example 7.8(a)

$$\varphi(x, y, \xi) = (x - y) \cdot \xi, \quad (x, y) \in U \times U, \quad \xi \in (\mathbf{R}^p \setminus \{0\}),$$

then the operator L is given by the formula

$$L = \frac{1-\rho(\xi)}{2+|x-y|^2}\left\{\sum_{j=1}^{p}(x_j - y_j)\frac{\partial}{i\,\partial\xi_j} + \sum_{k=1}^{p}\frac{\xi_k}{|\xi|^2}\frac{\partial}{i\,\partial x_k} + \sum_{k=1}^{p}\frac{-\xi_k}{|\xi|^2}\frac{\partial}{i\,\partial y_k}\right\}$$
$$+\rho(\xi),$$

where $\rho(\xi)$ is a function in $C_0^\infty(\mathbf{R}^p)$ such that $\rho(\xi) = 1$ for $|\xi| \leq 1$.

7.3.3 Oscillatory Integrals

Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. We let

$$S^{\infty}_{\rho,\delta}\left(\Omega \times \mathbf{R}^N\right) = \bigcup_{m\in\mathbf{R}} S^m_{\rho,\delta}\left(\Omega \times \mathbf{R}^N\right).$$

If $\varphi(x,\theta)$ is a phase function on $\Omega \times (\mathbf{R}^N \setminus \{0\})$, we wish to give a meaning to the integral

$$I_\varphi(au) = \iint_{\Omega\times\mathbf{R}^N} e^{i\varphi(x,\theta)} a(x,\theta) u(x)\, dx\, d\theta, \quad u \in C_0^\infty(\Omega), \tag{7.2}$$

for each symbol $a(x,\theta) \in S^{\infty}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$.

By Lemma 7.9, we can replace $e^{i\varphi}$ in formula (7.2) by $L(e^{i\varphi})$. Then a *formal* integration by parts gives us that

$$I_\varphi(au) = \iint_{\Omega\times\mathbf{R}^N} e^{i\varphi(x,\theta)} L'(a(x,\theta)u(x))\, dx\, d\theta.$$

However, the properties of the coefficients of L' imply that L' maps $S^r_{\rho,\delta}$ continuously into $S^{r-\eta}_{\rho,\delta}$ for every $r \in \mathbf{R}$, where $\eta = \min(\rho, 1-\delta)$. By continuing this process, we can reduce the growth of the integrand at infinity until it becomes integrable, and give a meaning to the integral (7.2) for each symbol $a(x,\theta) \in S^{\infty}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$.

More precisely, we have the following theorem:

Theorem 7.10. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. Then we have the following two assertions (i) and (ii):*

(i) The linear functional

$$S^{-\infty}\left(\Omega \times \mathbf{R}^N\right) \ni a \longmapsto I_\varphi(au) \in \mathbf{C}$$

extends uniquely to a linear functional ℓ *on* $S^{\infty}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ *whose restriction to each* $S^{m}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ *is continuous. Furthermore, the restriction to* $S^{m}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ *of* ℓ *is expressed in the form*

$$\ell(a) = \iint_{\Omega \times \mathbf{R}^N} e^{i\varphi(x,\theta)} (L')^k (a(x,\theta)u(x))\, dx\, d\theta,$$

where $k > (m+N)/\eta$, $\eta = \min(\rho, 1-\delta)$.

(ii) For any fixed $a(x,\theta) \in S^{m}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$, *the mapping*

$$C_0^{\infty}(\Omega) \ni u \longmapsto I_{\varphi}(au) = \ell(a) \in \mathbf{C} \tag{7.3}$$

is a distribution of order $\leq k$ *for* $k > (m+N)/\eta$.

We call the linear functional ℓ on $S^{\infty}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ an *oscillatory integral*, but use the standard notation as in formula (7.2). The distribution (7.3) is called the *Fourier integral distribution* associated with the phase function $\varphi(x,\theta)$ and the amplitude $a(x,\theta)$, and is denoted by the formula

$$\int_{\mathbf{R}^N} e^{i\varphi(x,\theta)} a(x,\theta)\, d\theta.$$

A consequence of Theorem 7.10 is that the usual rules of integration theory such as Fubini's theorem and partial integration hold true for oscillatory integrals.

We give some simple examples of Fourier integral distributions:

Example 7.11. (a) The Dirac measure $\delta(x)$ may be expressed as an oscillatory integral in the form

$$\delta(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, d\xi.$$

This formula is called the *plane-wave expansion* of the delta function (cf. Gel'fand–Shilov [GS]).

(b) The distributions v.p. $(x_j/|x|^{n+1})$, $1 \leq j \leq n$, are expressed in the form

$$\text{v.p.}\, \frac{x_j}{|x|^{n+1}} = -\frac{i}{2^n \Gamma((n+1)/2)\pi^{(n-1)/2}} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \frac{\xi_j}{|\xi|}\, d\xi, \quad 1 \leq j \leq n.$$

Indeed, it suffices to note that $\widehat{R_j}(\xi) = \xi_j/|\xi|$, $1 \leq j \leq n$, as stated in Examples 5.30 in Chap. 5.

If u is a distribution on Ω, then the *singular support* of u is the smallest closed subset of Ω outside of which u is smooth. The singular support of u is denoted by $\operatorname{sing\,supp} u$.

The next theorem estimates the singular support of a Fourier integral distribution:

Theorem 7.12. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. If $\varphi(x, \theta)$ is a phase function on $\Omega \times (\mathbf{R}^N \setminus \{0\})$ and if $a(x, \theta)$ is in $S^{\infty}_{\rho,\delta}(\Omega \times \mathbf{R}^N)$, then the Fourier integral distribution*

$$K(x) = \int_{\mathbf{R}^N} e^{i\varphi(x,\theta)} a(x, \theta)\, d\theta \in \mathcal{D}'(\Omega)$$

satisfies the condition

$$\operatorname{sing\,supp} K \subset \left\{x \in \Omega : d_\theta \varphi(x, \theta) = 0 \ \textit{for some}\ \theta \in \mathbf{R}^N \setminus \{0\}\right\}.$$

7.3.4 Definitions and Basic Properties of Fourier Integral Operators

Let U and V be open subsets of $\mathbf{R}^p$ and $\mathbf{R}^q$, respectively, and $0 \leq \delta < \rho \leq 1$. If $\varphi(x, y, \theta)$ is a phase function on $U \times V \times (\mathbf{R}^N \setminus \{0\})$ and if $a(x, y, \theta)$ is a symbol in $S^{\infty}_{\rho,\delta}(U \times V \times \mathbf{R}^N)$, then there is associated a distribution $K \in \mathcal{D}'(U \times V)$ defined by the formula

$$K(x, y) = \int_{\mathbf{R}^N} e^{i\varphi(x,y,\theta)} a(x, y, \theta)\, d\theta.$$

By applying Theorem 7.12 with $\Omega := U \times V$ and $n := p + q$, we obtain that

$$\operatorname{sing\,supp} K \subset \left\{(x, y) \in U \times V : d_\theta \varphi(x, y, \theta) = 0 \ \text{ for some } \theta \in \mathbf{R}^N \setminus \{0\}\right\}.$$

The distribution $K(x, y) \in \mathcal{D}'(U \times V)$ defines a continuous linear operator

$$A : C_0^{\infty}(V) \longrightarrow \mathcal{D}'(U)$$

by the formula

$$\langle Av, u\rangle = \langle K, u \otimes v\rangle, \quad u \in C_0^{\infty}(U), \ \ v \in C_0^{\infty}(V).$$

The operator A is called the *Fourier integral operator* associated with the phase function $\varphi(x, y, \theta)$ and the amplitude $a(x, y, \theta)$, and is denoted as follows:

$$Av(x) = \iint_{V \times \mathbf{R}^N} e^{i\varphi(x,y,\theta)} a(x, y, \theta) v(y)\, dy\, d\theta, \quad v \in C_0^{\infty}(V).$$

We give some simple examples of Fourier integral operators:

Example 7.13. (1) Let U be an open subset of $\mathbf{R}^n$. The *identity operator* of the space $C_0^\infty(U)$ can be expressed as a Fourier integral operator in the form

$$u(x) = \frac{1}{(2\pi)^n} \iint_{U \times \mathbf{R}^n} e^{i(x-y)\cdot\xi} u(y)\, dy\, d\xi, \quad u \in C_0^\infty(U).$$

(2) A *differential operator*

$$P(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha, \quad a_\alpha(x) \in C^\infty(U),$$

with characteristic polynomial

$$p(x, \xi) = \sum_{|\alpha| \leq m} a_\alpha(x) \xi^\alpha,$$

can be expressed in the form

$$Pu(x) = \frac{1}{(2\pi)^n} \iint_{U \times \mathbf{R}^n} e^{i(x-y)\cdot\xi}\, p(x, \xi)\, u(y)\, dy\, d\xi, \quad u \in C_0^\infty(U).$$

(3) Let V be an open subset of $\mathbf{R}^n_y$ and let U be an open subset of $\mathbf{R}^{n-1}_{x'}$ such that

$$V \cap \left(\mathbf{R}^{n-1} \times \{0\}\right) = U \times \{0\}.$$

If γ is the *trace operator*, defined by the formula

$$\gamma : C_0^\infty(V) \longrightarrow C^\infty(U)$$
$$v \longmapsto v(x', 0),$$

then it can be expressed in the form

$$(\gamma v)(x') = \frac{1}{(2\pi)^n} \iint_{V \times \mathbf{R}^n} e^{i(x'-y')\cdot\xi' - i y_n \xi_n} v(y)\, dy\, d\xi, \quad v \in C_0^\infty(V).$$

Trace operators for the classes of Fourier integral operators have a formal meaning here. In Sect. 7.7.1 we define trace, potential and singular Green operators in the half-space $\mathbf{R}^n_+$ as an essential tool for studying elliptic boundary value problems.

It is natural to ask whether the Fourier integral operator $A : C_0^\infty(V) \to \mathcal{D}'(U)$ induces a continuous linear operator

$$A : C_0^\infty(V) \longrightarrow C^\infty(U)$$

and when the operator A has a continuous extension

$$A : \mathcal{E}'(V) \longrightarrow \mathcal{D}'(U).$$

The next theorem summarizes some basic properties of the Fourier integral operator A with phase function $\varphi(x, y, \theta)$:

Theorem 7.14. *(i) If $d_{y,\theta}\varphi(x, y, \theta) \neq 0$ on $U \times V \times (\mathbf{R}^N \setminus \{0\})$, then the Fourier integral operator A maps $C_0^\infty(V)$ continuously into $C^\infty(U)$.*
(ii) If $d_{x,\theta}\varphi(x, y, \theta) \neq 0$ on $U \times V \times (\mathbf{R}^N \setminus \{0\})$, then the Fourier integral operator A extends to a continuous linear operator on $\mathcal{E}'(V)$ into $\mathcal{D}'(U)$.
(iii) If $d_{y,\theta}\varphi(x, y, \theta) \neq 0$ and $d_{x,\theta}\varphi(x, y, \theta) \neq 0$ on $U \times V \times (\mathbf{R}^N \setminus \{0\})$, then we have, for all $v \in \mathcal{E}'(V)$,

$$\begin{aligned} &\operatorname{sing\,supp} Av \\ &\subset \left\{x \in U : d_\theta\varphi(x, y, \theta) = 0 \text{ for some } y \in \operatorname{sing\,supp} v \text{ and } \theta \in \mathbf{R}^N \setminus \{0\}\right\}. \end{aligned}$$

The situation can be visualized in the following diagram:

$$\begin{array}{ccc} \mathcal{E}'(V) & \xrightarrow{A} & \mathcal{D}'(U) \\ \uparrow & & \uparrow \\ C_0^\infty(V) & \xrightarrow[A]{} & C^\infty(U) \end{array}$$

7.4 Pseudo-differential Operators

In this section we define pseudo-differential operators and study their basic properties such as the behavior of transposes, adjoints and compositions of such operators, and the effect of a change of coordinates on such operators. Furthermore, we formulate classical surface and volume potentials in terms of pseudo-differential operators. This calculus of pseudo-differential operators will be applied to elliptic boundary value problems in Chaps. 11–13.

7.4.1 Definitions of Pseudo-differential Operators

Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. A *pseudo-differential operator* of order m on Ω is a Fourier integral operator of the form

$$Au(x) = \iint_{\Omega\times\mathbf{R}^n} e^{i(x-y)\cdot\xi} a(x, y, \xi)u(y)\, dy\, d\xi, \quad u \in C_0^\infty(\Omega), \tag{7.4}$$

with some $a(x, y, \xi) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^n)$. In other words, a pseudo-differential operator of order m is a Fourier integral operator associated with the phase function $\varphi(x, y, \xi) = (x - y) \cdot \xi$ and some amplitude $a(x, y, \xi) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^n)$.

For example, the Bessel potential $J^s = (I - \Delta)^{-s/2}$, $s > 0$, is a pseudo-differential operator of order $-s$ with amplitude

$$a(x, y, \xi) = \frac{1}{(2\pi)^n} \frac{1}{(1 + |\xi|^2)^{s/2}} \in S^{-s}_{1,0}(\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n).$$

Indeed, it suffices to note (see Examples 5.20 and 5.25) that

$$\begin{aligned} &(I - \Delta)^{-s/2} u(x) = G_s * u(x) \\ &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i x \cdot \xi} \widehat{G_s}(\xi) \hat{u}(\xi)\, d\xi \\ &= \iint_{\mathbf{R}^n \times \mathbf{R}^n} e^{i(x-y)\cdot\xi} \left(\frac{1}{(2\pi)^n} \frac{1}{(1 + |\xi|^2)^{s/2}} \right) u(y)\, dy\, d\xi, \quad u \in C_0^\infty(\mathbf{R}^n). \end{aligned}$$

We let

$$L^m_{\rho,\delta}(\Omega) = \text{the set of all pseudo-differential operators of order } m \text{ on } \Omega.$$

By applying Theorems 7.12 and 7.14 to our situation, we obtain the following three assertions (1)–(3):

(1) A pseudo-differential operator A maps the space $C_0^\infty(\Omega)$ continuously into the space $C^\infty(\Omega)$ and extends to a continuous linear operator $A : \mathcal{E}'(\Omega) \to \mathcal{D}'(\Omega)$. The situation may be represented by the following diagram:

$$\begin{array}{ccc} \mathcal{E}'(\Omega) & \xrightarrow{A} & \mathcal{D}'(\Omega) \\ \uparrow & & \uparrow \\ C_0^\infty(\Omega) & \xrightarrow[A]{} & C^\infty(\Omega) \end{array}$$

(2) The distribution kernel $K_A(x, y)$, defined by the formula

$$K_A(x, y) = \int_{\mathbf{R}^N} e^{i(x-y)\cdot\xi} a(x, y, \xi)\, d\xi,$$

of a pseudo-differential operator A satisfies the condition

$$\text{sing supp}\, K_A \subset \{(x, x) : x \in \Omega\},$$

that is, the kernel $K_A(x, y)$ is smooth off the diagonal $\Delta_\Omega = \{(x, x) : x \in \Omega\}$ in $\Omega \times \Omega$.

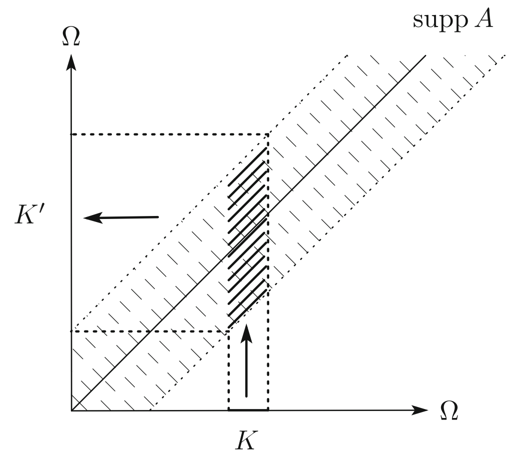


Fig. 7.6 Condition (a) on supp A

(3) sing supp $Au \subset$ sing supp u for every $u \in \mathcal{E}'(\Omega)$. In other words, Au is smooth whenever u is. This property is referred to as the *pseudo-local property*.

We set

$$L^{-\infty}(\Omega) = \bigcap_{m \in \mathbf{R}} L^m_{\rho,\delta}(\Omega).$$

The next theorem characterizes the class $L^{-\infty}(\Omega)$:

Theorem 7.15. *The following three conditions (i)–(iii) are equivalent:*

(i) $A \in L^{-\infty}(\Omega)$.
(ii) *A is written in the form* (7.4) *with some* $a(x, y, \xi) \in S^{-\infty}(\Omega \times \Omega \times \mathbf{R}^n)$.
(iii) *A is a* regularizer*, or equivalently, its distribution kernel* $K_A(x, y)$ *is in the space* $C^\infty(\Omega \times \Omega)$ *(see Theorem* 5.36*).*

A continuous linear operator $A : C_0^\infty(\Omega) \to \mathcal{D}'(\Omega)$ is said to be *properly supported* if the following two conditions (a) and (b) are satisfied (see Figs. 7.6 and 7.7):

(a) For any compact subset K of Ω, there exists a compact subset K' of Ω such that

$$\operatorname{supp} v \subset K \Longrightarrow \operatorname{supp} Av \subset K'.$$

(b) For any compact subset K' of Ω, there exists a compact subset $K \supset K'$ of Ω such that

$$\operatorname{supp} v \cap K = \emptyset \Longrightarrow \operatorname{supp} Av \cap K' = \emptyset.$$

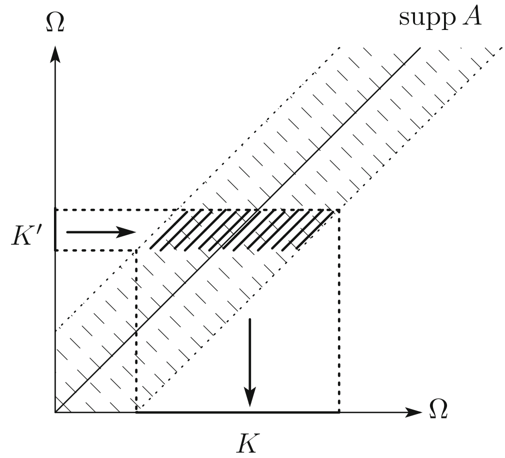


Fig. 7.7 Condition (b) on supp A

If A is properly supported, then it maps $C_0^\infty(\Omega)$ (resp. $C^\infty(\Omega)$) continuously into itself, and further it extends to a continuous linear operator on $\mathcal{E}'(\Omega)$ (resp. $\mathcal{D}'(\Omega)$) into itself. The situation may be represented by the following diagrams:

$$\begin{array}{ccc} \mathcal{E}'(\Omega) & \xrightarrow{A} & \mathcal{E}'(\Omega) \\ \uparrow & & \uparrow \\ C_0^\infty(\Omega) & \xrightarrow[A]{} & C_0^\infty(\Omega) \end{array}$$

$$\begin{array}{ccc} \mathcal{D}'(\Omega) & \xrightarrow{A} & \mathcal{D}'(\Omega) \\ \uparrow & & \uparrow \\ C^\infty(\Omega) & \xrightarrow[A]{} & C^\infty(\Omega) \end{array}$$

The next theorem states that every pseudo-differential operator can be written as the sum of a properly supported operator and a regularizer:

Theorem 7.16. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. If $A \in L^m_{\rho,\delta}(\Omega)$, then we have the decomposition*

$$A = A_0 + R,$$

where $A_0 \in L^m_{\rho,\delta}(\Omega)$ is properly supported and $R \in L^{-\infty}(\Omega)$.

Proof. Choose a function $\rho(x, y)$ in $C^\infty(\Omega \times \Omega)$ that satisfies the following two conditions (a) and (b) (cf. Figs. 7.6 and 7.7):

(a) $\rho(x, y) = 1$ in a neighborhood of the diagonal $\{(x, x) : x \in \Omega\}$ in $\Omega \times \Omega$.
(b) The restrictions to $\operatorname{supp} \rho$ of the projections $p_i : \Omega \times \Omega \ni (x_1, x_2) \mapsto x_i \in \Omega$, $i = 1, 2$, are proper mappings.

Then the operators A_0 and R, defined respectively by the kernels

$$K_{A_0}(x, y) = \rho(x, y) K_A(x, y),$$
$$K_R(x, y) = (1 - \rho(x, y)) K_A(x, y),$$

are the desired ones.

The proof of Theorem 7.16 is complete.

If $p(x, \xi) \in S^m_{\rho,\delta}(\Omega \times \mathbf{R}^n)$, then the operator $p(x, D)$, defined by the formula

$$p(x, D)u(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, p(x, \xi)\hat{u}(\xi)\, d\xi, \quad u \in C_0^\infty(\Omega), \tag{7.5}$$

is a pseudo-differential operator of order m on Ω, that is, $p(x, D) \in L^m_{\rho,\delta}(\Omega)$.

The next theorem asserts that every properly supported pseudo-differential operator can be reduced to the form (7.5):

Theorem 7.17. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \le \delta < \rho \le 1$. If $A \in L^m_{\rho,\delta}(\Omega)$ is properly supported, then we have*

$$p(x, \xi) = e^{-ix\cdot\xi} A(e^{ix\cdot\xi}) \in S^m_{\rho,\delta}(\Omega \times \mathbf{R}^n),$$

and

$$A = p(x, D).$$

Furthermore, if $a(x, y, \xi) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^n)$ is an amplitude for A, then we have the asymptotic expansion

$$p(x, \xi) \sim \sum_{\alpha \ge 0} \frac{1}{\alpha!} \partial_\xi^\alpha D_y^\alpha \left(a(x, y, \xi)\right)\Big|_{y=x}.$$

The function $p(x, \xi)$ is called the complete symbol *of A.*

We extend the notion of a complete symbol to the whole space $L^m_{\rho,\delta}(\Omega)$. If $A \in L^m_{\rho,\delta}(\Omega)$, we choose a properly supported operator $A_0 \in L^m_{\rho,\delta}(\Omega)$ such that $A - A_0 \in L^{-\infty}(\Omega)$, and define

$$\sigma(A) = \text{the equivalence class of the complete symbol of } A_0 \text{ in}$$
$$S^m_{\rho,\delta}(\Omega \times \mathbf{R}^n)/S^{-\infty}(\Omega \times \mathbf{R}^n).$$

In view of Theorems 7.15 and 7.16, it follows that $\sigma(A)$ does not depend on the properly supported operator A_0 chosen. The equivalence class $\sigma(A)$ is called the *complete symbol* of A. It is easy to see that the mapping

$$L^m_{\rho,\delta}(\Omega) \ni A \longmapsto \sigma(A) \in S^m_{\rho,\delta}(\Omega \times \mathbf{R}^n)/S^{-\infty}(\Omega \times \mathbf{R}^n)$$

induces an isomorphism

$$L^m_{\rho,\delta}/L^{-\infty} \longrightarrow S^m_{\rho,\delta}/S^{-\infty}.$$

We shall often identify the complete symbol $\sigma(A)$ with a representative in the class $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^n)$ for notational convenience, and call any member of $\sigma(A)$ a complete symbol of A.

A pseudo-differential operator $A \in L^m_{1,0}(\Omega)$ is said to be *classical* if its complete symbol $\sigma(A)$ has a representative in the class $S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^n)$.

We let

$$L^m_{\mathrm{cl}}(\Omega) = \text{the set of all classical pseudo-differential operators of order } m \text{ on } \Omega.$$

Then the mapping

$$L^m_{\mathrm{cl}}(\Omega) \ni A \longmapsto \sigma(A) \in S^m_{\mathrm{cl}}(\Omega \times \mathbf{R}^n)/S^{-\infty}(\Omega \times \mathbf{R}^n)$$

induces an isomorphism

$$L^m_{\mathrm{cl}}/L^{-\infty} \longrightarrow S^m_{\mathrm{cl}}/S^{-\infty}.$$

Also we have

$$L^{-\infty}(\Omega) = \bigcap_{m \in \mathbf{R}} L^m_{\mathrm{cl}}(\Omega).$$

If $A \in L^m_{\mathrm{cl}}(\Omega)$, then the principal part of $\sigma(A)$ has a canonical representative $\sigma_A(x,\xi) \in C^\infty(\Omega \times (\mathbf{R}^n \setminus \{0\}))$ which is positively homogeneous of degree m in the variable ξ. The function $\sigma_A(x,\xi)$ is called the *homogeneous principal symbol* of A. For example, the Bessel potential $J^s = (I - \Delta)^{-s/2}$, $s \in \mathbf{R}^n$, is a classical pseudo-differential operator on $\mathbf{R}^n$ with homogeneous principal symbol $|\xi|^{-s}$.

7.4.2 Basic Properties of Pseudo-differential Operators

Let Ω be an open subset of $\mathbf{R}^n$. A subset Γ of $\Omega \times \mathbf{R}^N$ is said to be *conic* if it satisfies the condition

$$(x,\theta) \in \Gamma \Longrightarrow (x, t\theta) \in \Gamma \quad \text{for all } t > 0.$$

Let $a(x, \theta)$ be a function in $C^\infty(\Omega \times \mathbf{R}^N)$. Then the *conic support* of a, denoted by conic supp a, is the smallest closed conic subset of $\Omega \times \mathbf{R}^N$ outside of which $a(x, \theta)$ vanishes.

The next theorem plays a crucial role in the theory of pseudo-differential operators:

Theorem 7.18. *Let Ω be an open subset of $\mathbf{R}^n$ and $1 - \rho \leq \delta < \rho \leq 1$. We are given the following:*

(a) A pseudo-differential operator $P \in L^m_{\rho,\delta}(\Omega)$.
(b) A symbol $a(y, \theta) \in S^q_{\rho,\delta}(\Omega \times \mathbf{R}^N)$.
(c) A phase function $\psi(x, \theta) \in C^\infty(\Omega \times (\mathbf{R}^N \setminus \{0\}))$ which satisfies the condition

$$d_x\psi(x, \theta) \neq 0 \quad \textit{for all } (x, \theta) \in \text{conic supp}\, a \textit{ and } \theta \neq 0.$$

Assume that either the operator P is properly supported or the symbol $a(y, \theta)$ has a compact support with respect to y in Ω.

If we let

$$b(x, \theta) = e^{-i\psi(x,\theta)} P(a(y, \theta)e^{i\psi(y,\theta)})(x, \theta), \quad (x, \theta) \in \Omega \times \mathbf{R}^N,$$

then we have the following three assertions (i)–(iii):

(i) $b(x, \theta) \in S^{m+q}_{\rho,\delta}(\Omega \times (\mathbf{R}^N \setminus \{0\}))$.
(ii) The function $b(x, \theta)$ has an asymptotic expansion

$$b(x, \theta) \sim \sum_{\alpha \geq 0,\, \beta \geq 0} \frac{1}{\alpha!\beta!} \partial_\xi^{\alpha+\beta} D_y^\beta (p(x, y, d_x\psi(x, \theta)))\Big|_{y=x} D_y^\alpha (a(y, \theta)e^{ir(x,y,\theta)})\Big|_{y=x}.$$

Here $p(x, y, \theta) \in S^m_{\rho,\delta}(\Omega \times \Omega \times \mathbf{R}^N)$ is an amplitude for $P \in L^m_{\rho,\delta}(\Omega)$ and

$$r(x, y, \theta) = \psi(y, \theta) - \psi(x, \theta) - \langle d_x\psi(x, \theta), y - x\rangle,$$
$$(x, y, \theta) \in \Omega \times \Omega \times (\mathbf{R}^N \setminus \{0\}).$$

(iii) In particular, we have

$$b(x, \theta) \equiv p(x, x, d_x\psi(x, \theta))a(x, \theta) \quad \textit{mod } S^{m+q-1}_{\rho,\delta}(\Omega \times \Omega \times (\mathbf{R}^N \setminus \{0\})).$$

Theorem 7.18 contains all the machinery necessary for the theory of pseudo-differential operators. In fact, Theorems 7.17, 7.19, 7.20 and 7.22 can be easily obtained from Theorem 7.18 respectively by taking $p(x, y, \theta)$, $\psi(x, \theta)$ and $a(y, \theta)$ as in the following diagram:

Theorem 7.17	Theorem 7.19	Theorem 7.19	Theorem 7.20	Theorem 7.22
$a(x,y,\xi)$	$\sigma(A)(y,-\xi)$	$\sigma(A)(y,\xi)$	$\sigma(A)(x,\xi)$	$\sigma(A)(x,\xi)$
$< x,\xi >$	$< x,\xi >$	$< x,\xi >$	$< x,\xi >$	$\langle \chi(x),\xi \rangle$
1	1	1	$\sigma(B)(y,\xi)$	1

The proof of Theorem 7.18 is based on Example 5.29 and the stationary phase theorem.

The next two theorems assert that the class of pseudo-differential operators forms an algebra closed under the operations of composition of operators and taking the transpose or adjoint of an operator:

Theorem 7.19. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. If $A \in L^m_{\rho,\delta}(\Omega)$, then its transpose A' and its adjoint A^* are both in $L^m_{\rho,\delta}(\Omega)$, and the complete symbols $\sigma(A')$ and $\sigma(A^*)$ have respectively the following asymptotic expansions:*

$$\sigma(A')(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!} \partial_\xi^\alpha D_x^\alpha (\sigma(A)(x,-\xi)),$$

$$\sigma(A^*)(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!} \partial_\xi^\alpha D_x^\alpha (\overline{\sigma(A)(x,\xi)}).$$

Theorem 7.20. *Let Ω be an open subset of $\mathbf{R}^n$. If $A \in L^{m'}_{\rho',\delta'}(\Omega)$ and $B \in L^{m''}_{\rho'',\delta''}(\Omega)$ where $0 \leq \delta' < \rho'' \leq 1$ and if one of them is properly supported, then the composition AB is in $L^{m'+m''}_{\rho,\delta}(\Omega)$ with $\rho = \min(\rho', \rho'')$, $\delta = \max(\delta', \delta'')$, and we have the asymptotic expansion*

$$\sigma(AB)(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!} \partial_\xi^\alpha (\sigma(A)(x,\xi)) \cdot D_x^\alpha (\sigma(B)(x,\xi)) .$$

A pseudo-differential operator $A \in L^m_{\rho,\delta}(\Omega)$ is said to be *elliptic* of order m if its complete symbol $\sigma(A)$ is elliptic of order m. In view of Theorem 7.7, it follows that a classical pseudo-differential operator $A \in L^m_{\mathrm{cl}}(\Omega)$ is elliptic if and only if its homogeneous principal symbol $\sigma_A(x,\xi)$ does not vanish on the space $\Omega \times (\mathbf{R}^n \setminus \{0\})$.

The next theorem states that elliptic operators are the "invertible" elements in the algebra of pseudo-differential operators:

Theorem 7.21. *Let Ω be an open subset of $\mathbf{R}^n$ and $0 \leq \delta < \rho \leq 1$. An operator $A \in L^m_{\rho,\delta}(\Omega)$ is elliptic if and only if there exists a properly supported operator $B \in L^{-m}_{\rho,\delta}(\Omega)$ such that*

$$\begin{cases} AB \equiv I \mod L^{-\infty}(\Omega), \\ BA \equiv I \mod L^{-\infty}(\Omega). \end{cases}$$

Such an operator B is called a *parametrix* for A. In other words, a parametrix for A is a two-sided inverse of A modulo $L^{-\infty}(\Omega)$. We observe that a parametrix is unique modulo $L^{-\infty}(\Omega)$.

The next theorem proves the invariance of pseudo-differential operators under a change of coordinates:

Theorem 7.22. *Let Ω_1, Ω_2 be two open subsets of $\mathbf{R}^n$ and let $\chi : \Omega_1 \to \Omega_2$ be a C^∞ diffeomorphism. If $A \in L^m_{\rho,\delta}(\Omega_1)$, where $1-\rho \leq \delta < \rho \leq 1$, then the mapping*

$$\begin{aligned} \chi_* A : C_0^\infty(\Omega_2) &\longrightarrow C^\infty(\Omega_2) \\ v &\longmapsto A(v \circ \chi) \circ \chi^{-1} \end{aligned}$$

is in the class $L^m_{\rho,\delta}(\Omega_2)$, and we have the asymptotic expansion

$$\sigma(\chi_* A)(y,\eta) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!} \left(\partial_\xi^\alpha \sigma(A)\right)(x, {}^t\chi'(x)\cdot\eta) \cdot D_z^\alpha \left(e^{ir(x,z,\eta)}\right)\Big|_{z=x}, \qquad (7.6)$$

with

$$r(x,z,\eta) = \left\langle \chi(z) - \chi(x) - \chi'(x)\cdot(z-x), \eta \right\rangle.$$

Here $x = \chi^{-1}(y)$, $\chi'(x)$ is the derivative of χ at x and ${}^t\chi'(x)$ is its transpose.

The situation may be represented by the following diagram:

$$\begin{array}{ccc} C_0^\infty(\Omega_1) & \xrightarrow{\ A\ } & C^\infty(\Omega_1) \\ {\scriptstyle \chi^*}\big\uparrow & & \big\downarrow{\scriptstyle \chi_*} \\ C_0^\infty(\Omega_2) & \xrightarrow[\ \chi_* A\ ]{} & C^\infty(\Omega_2) \end{array}$$

Here $\chi^* v = v \circ \chi$ is the pull-back of v by χ and $\chi_* u = u \circ \chi^{-1}$ is the push-forward of u by χ, respectively.

Remark 7.23. Formula (7.6) shows that

$$\sigma(A_\chi)(y,\eta) \equiv \sigma(A)\left(x, {}^t\chi'(x)\cdot\eta\right) \qquad \text{mod } S^{m-(\rho-\delta)}_{\rho,\delta}.$$

Note that the mapping

$$\Omega_2 \times \mathbf{R}^n \ni (y,\eta) \longmapsto \left(x, {}^t\chi'(x)\cdot\eta\right) \in \Omega_1 \times \mathbf{R}^n$$

is just a transition map of the cotangent bundle $T^*(\mathbf{R}^n)$. This implies that the principal symbol $\sigma_m(A)$ of $A \in L^m_{\rho,\delta}(\mathbf{R}^n)$ can be invariantly defined on $T^*(\mathbf{R}^n)$ when $1-\rho \leq \delta < \rho \leq 1$.

A differential operator of order m with smooth coefficients on Ω is continuous on $H^{s,p}_{\text{loc}}(\Omega)$ (resp. $B^{s,p}_{\text{loc}}(\Omega)$) into $H^{s-m,p}_{\text{loc}}(\Omega)$ (resp. $B^{s-m,p}_{\text{loc}}(\Omega)$) for all $s \in \mathbf{R}$. These results extend to pseudo-differential operators:

Theorem 7.24. *Every properly supported operator $A \in L^m_{1,\delta}(\Omega)$, $0 \leq \delta < 1$, extends to a continuous linear operator $A : H^{s,p}_{\text{loc}}(\Omega) \to H^{s-m,p}_{\text{loc}}(\Omega)$ for all $s \in \mathbf{R}$ and $1 < p < \infty$, and it also extends to a continuous linear operator $A : B^{s,p}_{\text{loc}}(\Omega) \to B^{s-m,p}_{\text{loc}}(\Omega)$ for all $s \in \mathbf{R}$ and $1 \leq p \leq \infty$.*

The situation may be represented by the following diagrams:

$$\begin{array}{ccc}
\mathcal{D}'(\Omega) & \xrightarrow{A} & \mathcal{D}'(\Omega) \\
\uparrow & & \uparrow \\
H^{s,p}_{\text{loc}}(\Omega) & \xrightarrow{A} & H^{s-m,p}_{\text{loc}}(\Omega) \\
\uparrow & & \uparrow \\
C^\infty(\Omega) & \xrightarrow[A]{} & C^\infty(\Omega)
\end{array}$$

$$\begin{array}{ccc}
\mathcal{D}'(\Omega) & \xrightarrow{A} & \mathcal{D}'(\Omega) \\
\uparrow & & \uparrow \\
B^{s,p}_{\text{loc}}(\Omega) & \xrightarrow{A} & B^{s-m,p}_{\text{loc}}(\Omega) \\
\uparrow & & \uparrow \\
C^\infty(\Omega) & \xrightarrow[A]{} & C^\infty(\Omega)
\end{array}$$

7.4.3 Pseudo-differential Operators on a Manifold

Now we define the concept of a pseudo-differential operator on a manifold, and transfer all the machinery of pseudo-differential operators to manifolds.Theorem 7.22 leads us to the following definition:

Definition 7.25. Let M be an n-dimensional, *compact* smooth manifold without boundary, and $1 - \rho \leq \delta < \rho \leq 1$. A continuous linear operator $A : C^\infty(M) \to C^\infty(M)$ is called a *pseudo-differential operator* of order $m \in \mathbf{R}$ if it satisfies the following two conditions (i) and (ii):

(i) The distribution kernel of A is smooth off the diagonal $\Delta_M = \{(x,x) : x \in M\}$ in $M \times M$.
(ii) For any chart (U, χ) on M, the mapping

$$\begin{aligned} \chi_* (A|_U) : C_0^\infty(\chi(U)) &\longrightarrow C^\infty(\chi(U)) \\ u &\longmapsto A(u \circ \chi) \circ \chi^{-1} \end{aligned}$$

belongs to the class $L^m_{\rho,\delta}(\chi(U))$.

The situation may be represented by the following commutative diagram:

$$\begin{array}{ccc} C_0^\infty(U) & \xrightarrow{A|_U} & C^\infty(U) \\ \chi^* \uparrow & & \downarrow \chi_* \\ C_0^\infty(\chi(U)) & \xrightarrow[\chi_*(A|_U)]{} & C^\infty(\chi(U)) \end{array}$$

Here the operator $A|_U$ is defined as follows:

$$A|_U : C_0^\infty(U) \longrightarrow C^\infty(M) \overset{A}{\longrightarrow} C^\infty(M) \longrightarrow C^\infty(U),$$

where the first arrow is the natural injection and the last one is the restriction to U.

We let

$$L^m_{\rho,\delta}(M) = \text{the set of all pseudo-differential operators of order } m \text{ on } M,$$

and set

$$L^{-\infty}(M) = \bigcap_{m \in \mathbf{R}} L^m_{\rho,\delta}(M).$$

Some results about pseudo-differential operators on $\mathbf{R}^n$ stated above are also true for pseudo-differential operators on M. In fact, pseudo-differential operators on M are defined to be locally pseudo-differential operators on $\mathbf{R}^n$.

For example, we have the following five results (1)–(5):

(1) A pseudo-differential operator A extends to a continuous linear operator $A : \mathcal{D}'(M) \to \mathcal{D}'(M)$.
(2) $\operatorname{sing\,supp} Au \subset \operatorname{sing\,supp} u$, $u \in \mathcal{D}'(M)$.
(3) A continuous linear operator $A : C^\infty(M) \to \mathcal{D}'(M)$ is a *regularizer* if and only if it is in $L^{-\infty}(M)$.
(4) Let $m \in \mathbf{R}$ and $0 \leq \delta < \rho \leq 1$. The class $L^m_{\rho,\delta}(M)$ is stable under the operations of composition of operators and taking the transpose or adjoint of an operator.
(5) A pseudo-differential operator $A \in L^m_{1,\delta}(M)$, where $0 \leq \delta < 1$, extends to a continuous linear operator $A : H^{s,p}(M) \to H^{s-m,p}(M)$ for all $s \in \mathbf{R}$ and $1 < p < \infty$ and also a continuous linear operator $A : B^{s,p}(M) \to B^{s-m,p}(M)$ for all $s \in \mathbf{R}$ and $1 \leq p \leq \infty$.

The situation may be represented by the following diagrams:

$$\begin{array}{ccc} \mathcal{D}'(M) & \xrightarrow{A} & \mathcal{D}'(M) \\ \uparrow & & \uparrow \\ H^{s,p}(M) & \xrightarrow{A} & H^{s-m,p}(M) \\ \uparrow & & \uparrow \\ C^\infty(M) & \xrightarrow[A]{} & C^\infty(M) \end{array}$$

$$\begin{array}{ccc} \mathcal{D}'(M) & \xrightarrow{A} & \mathcal{D}'(M) \\ \uparrow & & \uparrow \\ B^{s,p}(M) & \xrightarrow{A} & B^{s-m,p}(M) \\ \uparrow & & \uparrow \\ C^\infty(M) & \xrightarrow[A]{} & C^\infty(M) \end{array}$$

A pseudo-differential operator $A \in L^m_{1,0}(M)$ is said to be *classical* if, for any chart (U, χ) on M, the mapping $\chi_* (A|_U) : C_0^\infty(\chi(U)) \to C^\infty(\chi(U))$ belongs to the class $L^m_{\mathrm{cl}}(\chi(U))$.

We let

$$L^m_{\mathrm{cl}}(M) = \text{the set of all classical pseudo-differential operators of order } m \text{ on } M.$$

We observe that

$$L^{-\infty}(M) = \bigcap_{m \in \mathbf{R}} L^m_{\mathrm{cl}}(M).$$

Let $A \in L^m_{\mathrm{cl}}(M)$. If (U, χ) is a chart on M, there is associated a homogeneous principal symbol $\sigma_{A_\chi} \in C^\infty(\chi(U) \times (\mathbf{R}^n \setminus \{0\}))$. In view of Remark 7.23, by smoothly patching together the functions σ_{A_χ} we can obtain a smooth function $\sigma_A(x, \xi)$ on $T^*(M) \setminus \{0\} = \{(x, \xi) \in T^*(M) : \xi \neq 0\}$, which is positively homogeneous of degree m in the variable ξ. The function $\sigma_A(x, \xi)$ is called the *homogeneous principal symbol* of A.

A classical pseudo-differential operator $A \in L^m_{\mathrm{cl}}(M)$ is said to be *elliptic* of order m if its homogeneous principal symbol $\sigma_A(x, \xi)$ does not vanish on the bundle $T^*(M) \setminus \{0\}$ of non-zero cotangent vectors.

Then we have the following result (6):

(6) An operator $A \in L^m_{\mathrm{cl}}(M)$ is elliptic if and only if there exists a parametrix $B \in L^{-m}_{\mathrm{cl}}(M)$ for A:

$$\begin{cases} AB \equiv I \mod L^{-\infty}(M), \\ BA \equiv I \mod L^{-\infty}(M). \end{cases}$$

7.4.4 Hypoelliptic Pseudo-differential Operators

Let Ω be an open subset of $\mathbf{R}^n$. A properly supported pseudo-differential operator A on Ω is said to be *hypoelliptic* if it satisfies the condition

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp} Au, \quad u \in \mathcal{D}'(\Omega).$$

It is easy to see that this condition is equivalent to the following condition: For any open subset Ω_1 of Ω, we have

$$u \in \mathcal{D}'(\Omega), \; Au \in C^\infty(\Omega_1) \implies u \in C^\infty(\Omega_1).$$

For example, Theorem 7.21 tells us that elliptic operators are hypoelliptic.

We say that A is *globally hypoelliptic* if it satisfies the weaker condition

$$u \in \mathcal{D}'(\Omega), \; Au \in C^\infty(\Omega) \implies u \in C^\infty(\Omega).$$

It should be noted that these two notions may be transferred to manifolds. The following criterion for hypoellipticity is due to Hörmander:

Theorem 7.26. *Let Ω be an open subset of $\mathbf{R}^n$ and let $A = p(x, D)$ be a properly supported pseudo-differential operator in $L^m_{\rho,\delta}(\Omega)$ with $1 - \rho \leq \delta < \rho \leq 1$. Assume that, for any compact $K \subset \Omega$ and any multi-indices α, β there exist constants $C_{K,\alpha,\beta} > 0$, $C_K > 0$ and $\mu \in \mathbf{R}$ such that we have, for all $x \in K$ and $|\xi| \geq C_K$,*

$$\left| D_\xi^\alpha D_x^\beta p(x, \xi) \right| \leq C_{K,\alpha,\beta} |p(x, \xi)| (1 + |\xi|)^{-\rho|\alpha| + \delta|\beta|}, \tag{7.7a}$$

$$|p(x, \xi)|^{-1} \leq C_K (1 + |\xi|)^\mu. \tag{7.7b}$$

Then there exists a parametrix $B \in L^\mu_{\rho,\delta}(\Omega)$ *for A.*

We give a typical example of Theorem 7.26 which plays an essential role in the proof of [Ta9, Theorem 1.1]:

Example 7.27. Let (M, g) be an n-dimensional, compact Riemannian smooth manifold without boundary, and let Δ_M be the Laplace–Beltrami operator of M (see Example 5.43). If $\alpha(x)$ is a smooth function defined on M such that $0 \leq \alpha(x) \leq 1$ on M, then a pseudo-differential operator

$$A(x, D) = \alpha(x)\sqrt{-\Delta_M} + (1 - \alpha(x))$$

satisfies conditions (7.7a) and (7.7b) with $\mu := 0$, $\rho := 1$ and $\delta := 1/2$ (see [Ta9, Lemma 5.2]; [Ka]).

7.5 Potentials and Pseudo-differential Operators

The purpose of this section is to describe, in terms of pseudo-differential operators, the classical surface and volume potentials arising in boundary value problems for elliptic differential operators.

We give a formal description of a background. Let Ω be a bounded domain in Euclidean space $\mathbf{R}^n$ with smooth boundary $\partial\Omega$. Its closure $\overline{\Omega} = \Omega \cup \partial\Omega$ is an n-dimensional, compact smooth manifold with boundary. We may assume that $\overline{\Omega}$ is the closure of a relatively compact, open subset Ω of an n-dimensional, compact smooth manifold M without boundary in which Ω has a smooth boundary $\partial\Omega$ (see Fig. 7.8).

Let P be a differential operator of order m with smooth coefficients on M. Then we have the jump formula (see formula (5.13))

$$P(u^0) = (Pu)^0 + \tilde{P}\gamma u, \quad u \in C^\infty(\overline{\Omega}), \tag{7.8}$$

where u^0 is the extension of u to M by zero outside of $\overline{\Omega}$

$$u^0(x) = \begin{cases} u(x) & \text{for } x \in \overline{\Omega}, \\ 0 & \text{for } x \in M \setminus \overline{\Omega}, \end{cases}$$

and $\tilde{P}\gamma u$ is a distribution on M with support in $\partial\Omega$. If P admits an "inverse" Q, then the function u may be expressed as follows:

$$u = Q((Pu)^0)\big|_{\Omega} + Q(\tilde{P}\gamma u)\big|_{\Omega}. \tag{7.9}$$

The first term on the right-hand side is a volume potential and the second term is a surface potential with m "layers". For example, if P is the usual Laplacian Δ and if $\Omega = \mathbf{R}^n_+$, then formulas (7.8) and (7.9) coincide with formulas (5.12) and (5.49), respectively. The proof of formula (7.9), based on jump formula (7.8), may conceivably be new.

First, we state a theorem which covers surface potentials:

Theorem 7.28. *Let $A \in L^m_{\mathrm{cl}}(M)$ be properly supported. Assume that*

$$\textit{Every term in the complete symbol } \sum_{j=0}^{\infty} a_j(x,\xi) \textit{ of } A \tag{7.10}$$

is a rational *function of* ξ.

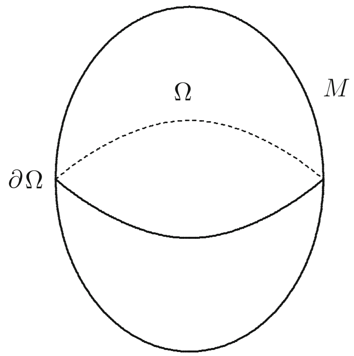


Fig. 7.8 The double M of Ω

Then we have the following two assertions (i) and (ii):

(i) The operator

$$H : v \longmapsto A(v \otimes \delta)|_{\Omega}$$

is continuous from $C^\infty(\partial\Omega)$ *into* $C^\infty(\overline{\Omega})$. *Moreover, if* $v \in \mathcal{D}'(\partial\Omega)$, *the distribution* Hv *has sectional traces on the boundary* $\partial\Omega$ *of any order.*

(ii) The operator

$$S : C^\infty(\partial\Omega) \longrightarrow C^\infty(\partial\Omega)$$
$$v \longmapsto Hv|_{\partial\Omega}$$

belongs to the class $L_{\mathrm{cl}}^{m+1}(\partial\Omega)$. *Furthermore, its homogeneous principal symbol is given by the formula*

$$(x', \xi') \longmapsto \frac{1}{2\pi} \int_\Gamma a_0(x', 0, \xi', \xi_n)\, d\xi_n$$

where $a_0(x', x_n, \xi', \xi_n) \in C^\infty(T^*(M) \setminus \{0\})$ *is the homogeneous principal symbol of* A, *and* Γ *is a circle in the half-plane* $\{\xi_n \in \mathbf{C} : \operatorname{Im} \xi_n > 0\}$ *which encloses the poles* ξ_n *of* $a_0(x', 0, \xi', \xi_n)$ *there (see Fig. 7.9).*

(iii) If $1 < p < \infty$, *then the operator* H *extends to a continuous linear operator*

$$H : B^{s,p}(\partial\Omega) \longrightarrow H^{s-m-1/p,p}(\Omega)$$

for every $s \in \mathbf{R}$.

Part (i) of Theorem 7.28 asserts that the surface potential H preserves smoothness up to the boundary $\partial\Omega$. The situation of Theorem 7.28 can be visualized in the following diagram:

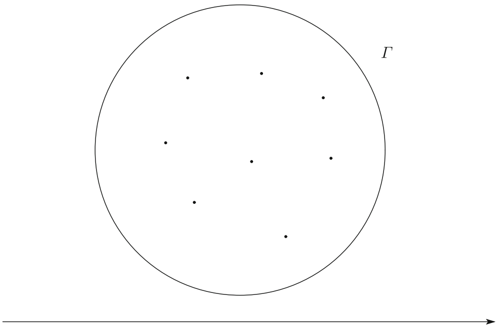


Fig. 7.9 The circle Γ in the half-plane $\{\operatorname{Im} \xi_n > 0\}$

$$\begin{array}{ccc} B^{s,p}(\partial\Omega) & \xrightarrow{H} & H^{s-m-1/p,p}(\Omega) \\ \uparrow & & \uparrow \\ C^\infty(\partial\Omega) & \xrightarrow[H]{} & C^\infty(\overline{\Omega}) \end{array}$$

It should be emphasized that condition (7.10) is invariant under a change of coordinates. Furthermore, it is easy to see that every parametrix for an elliptic differential operator satisfies condition (7.10).

The next theorem covers volume potentials:

Theorem 7.29. *Let $A \in L^m_{\mathrm{cl}}(M)$ be as in Theorem 7.28. Then we have the following two assertions (i) and (ii):*

(i) The operator

$$A_\Omega : f \longmapsto A(f^0)\big|_\Omega$$

is continuous from $C^\infty(\overline{\Omega})$ into itself.

(ii) If $1 < p < \infty$, then the operator A_Ω extends to a continuous linear operator

$$A_\Omega : H^{s,p}(\Omega) \longrightarrow H^{s-m,p}(\Omega)$$

for every $s > -1/p$.

The operator A_Ω can be visualized as follows:

$$A_\Omega : C^\infty(\overline{\Omega}) \longrightarrow \mathcal{D}'(M) \xrightarrow{A} \mathcal{D}'(M) \longrightarrow C^\infty(\Omega),$$

where the first arrow is the zero extension to M and the last one is the restriction to Ω. In view of the pseudo-local property of A, we note that A_Ω maps $C^\infty(\overline{\Omega})$ into $C^\infty(\Omega)$. Part (i) of Theorem 7.29 asserts that the volume potential A_Ω preserves smoothness up to the boundary $\partial\Omega$. Moreover, the situation of Theorem 7.29 can be visualized in the following diagram:

$$\begin{array}{ccc} H^{s,p}(\Omega) & \xrightarrow{A_\Omega} & H^{s-m,p}(\Omega) \\ \uparrow & & \uparrow \\ C^\infty(\overline{\Omega}) & \xrightarrow[A_\Omega]{} & C^\infty(\overline{\Omega}) \end{array}$$

7.6 The Transmission Property

One of the important questions in the theory of pseudo-differential operators in a domain Ω is that of the smoothness of a solution near the boundary $\partial\Omega$. Due to the non-local character of pseudo-differential operators, we find more difficulties in the bounded domain Ω than in the whole space $\mathbf{R}^n$. In fact, when considering the Dirichlet problem in Ω, it is natural to use the zero-extension u^0 of functions u defined in the interior Ω outside of the closure $\overline{\Omega} = \Omega \cup \partial\Omega$ as in Sect. 7.5. This extension has a probabilistic interpretation. Namely, it corresponds to stopping the diffusion process with jumps in the whole space $\mathbf{R}^n$ at the first exit time of the closure $\overline{\Omega}$.

Given an arbitrary pseudo-differential operator A, it is in general not true that the operator

$$A_\Omega : u \longmapsto A(u^0)\big|_\Omega$$

maps functions u which are smooth up to the boundary $\partial\Omega$ into functions $A_\Omega u$ with the same property. The crucial requirement here is that the symbol of A has the transmission property. On one hand, this restricts the class of boundary value problems in the calculus, on the other hand, however, it ensures that solutions to elliptic equations with smooth data are smooth; it therefore helps to avoid problems with singularities of solutions at the boundary.

Following Boutet de Monvel [Bo], we impose a condition about symbols in the normal direction at the boundary in order to ensure the stated regularity property (see Rempel–Schulze [RS]).

If $x = (x_1, x_2, \ldots, x_n)$ is the variable in $\mathbf{R}^n$, we write

$$x = (x', x_n), \quad x' = (x_1, x_2, \ldots, x_{n-1}),$$

so $x' \in \mathbf{R}^{n-1}$ is the tangential component of x with dual variables $\xi' = (\xi_1, \xi_2, \ldots, \xi_{n-1})$, and $x_n \in \mathbf{R}$ is its normal component with dual variable ξ_n.

If $m \in \mathbf{R}$, we let

$$S^m_{1,0}(\overline{\mathbf{R}^n_+} \times \mathbf{R}^n) = \text{the space of symbols } a(x,\xi) \text{ in } S^m_{1,0}(\mathbf{R}^n_+ \times \mathbf{R}^n)$$
$$\text{which have an extension in } S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^n).$$

A symbol $a(x,\xi) \in S^m_{1,0}(\overline{\mathbf{R}^n_+} \times \mathbf{R}^n)$ is said to have the *transmission property* with respect to the boundary $\mathbf{R}^{n-1}$ if all its derivatives

$$\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} a(x', 0, \xi', \nu), \quad \alpha_n \geq 0,$$

admit an expansion of the form

$$\begin{aligned}
&\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} a(x', 0, \xi', \nu) \qquad (7.11)\\
&= \sum_{j=0}^{m} b_j(x', \xi')\nu^j + \sum_{k=-\infty}^{\infty} a_k(x', \xi') \frac{\left(\langle \xi' \rangle - \sqrt{-1}\nu\right)^k}{\left(\langle \xi' \rangle + \sqrt{-1}\nu\right)^{k+1}}, \quad \nu \in \mathbf{R},
\end{aligned}$$

where (cf. [RS, 2.2.2.1, Proposition 3])

$$b_j(x', \xi') \in S^{m-j}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}),$$

and

$$a_k(x', \xi') \in S^{m+1}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$$

form a rapidly decreasing sequence with respect to k, and

$$\langle \xi' \rangle = \left(1 + |\xi'|^2\right)^{1/2}.$$

For classical pseudo-differential symbols of integer order, the transmission property can be expressed via homogeneity properties of the terms in the asymptotic expansion. However, this is no longer true for non-classical pseudo-differential symbols. For details, see the analysis by Grubb–Hörmander [GH].

For example, if $a(x,\xi) \in S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^n)$ is a classical symbol of order $m \in \mathbf{Z}$ with an asymptotic expansion

$$a(x,\xi) \sim \sum_{j=0}^{\infty} a_{m-j}(x,\xi),$$

where $a_{m-j}(x,\xi) \in S^{m-j}_{1,0}$ is positively homogeneous of degree $m-j$ for $|\xi| \geq 1$, then it is easy to verify (see [RS, 2.2.2.3, Proposition 1], [Ho4, Lemma 18.2.14]) that $a(x,\xi)$ has the transmission property if and only if we have, for all multi-indices $\alpha = (\alpha', \alpha_n)$,

$$\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}\left(\frac{\partial}{\partial \xi'}\right)^{\alpha'} a_{m-j}(x',0,0,+1)$$
$$= (-1)^{m-j-|\alpha'|}\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}\left(\frac{\partial}{\partial \xi'}\right)^{\alpha'} a_{m-j}(x',0,0,-1).$$

First, we have the following lemma:

Lemma 7.30. *Regularizing symbols always have the transmission property, and so do symbols $a(x,\xi)$ which vanish to infinite order at $\{x_n = 0\}$. Moreover, all symbols $a(x,\xi)$ which are polynomial in ξ have the transmission property.*

There are symbols with the transmission property of arbitrary order:

Example 7.31. Let $b(x,\xi') \in S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^{n-1})$ and $\varphi \in \mathcal{S}(\mathbf{R})$. Then it is easy to verify that the symbol

$$a(x,\xi) := b(x,\xi')\varphi(\xi_n/\langle\xi'\rangle)$$

has the transmission property with respect to the boundary $\mathbf{R}^{n-1}$.

The next proposition asserts the stability of the transmission property under the usual pseudo-differential constructions:

Proposition 7.32. *(a) If $a(x,\xi)$ and $b(x,\xi)$ has the transmission property, then all derivatives $\partial_\xi^\alpha \partial_x^\beta a(x,\xi)$ and the product $a(x,\xi)b(x,\xi)$ has the transmission property, respectively.*
(b) If $a_j(x,\xi)$ are symbols of order $m-j$ with the transmission property and if $a(x,\xi) \sim \sum_j a_j(x,\xi)$, then $a(x,\xi)$ has the transmission property.
(c) If $a(x,\xi)$ is elliptic with the transmission property, then every parametrix of $a(x,\xi)$ has the transmission property.

If $m \in \mathbf{R}$, we let

$$L^m_{1,0}(\overline{\mathbf{R}^n_+}) = \text{the space of pseudo-differential operators in } L^m_{1,0}(\mathbf{R}^n_+) \text{ which can be extended to a pseudo-differential operator in } L^m_{1,0}(\mathbf{R}^n).$$

A pseudo-differential operator $A \in L^m_{1,0}(\overline{\mathbf{R}^n_+})$ is said to have the *transmission property* with respect to the boundary $\mathbf{R}^{n-1}$ if any complete symbol of A has the transmission property with respect to the boundary $\mathbf{R}^{n-1}$.

As a first useful example, we formulate the pseudo-differential operators that have the transmission property in dimension one [Bo, Theorem 2.7]:

Theorem 7.33. *Let $A(x,D)$ be a pseudo-differential operator defined in a neighborhood of the half-line $x \geq 0$. In order that the* transmission property *with respect to the origin holds true for A, it is necessary and sufficient that A admits a decomposition*

$$A = A_0 + A_1 + A_2.$$

Here:

(1) The symbol of A_0 *vanishes to the infinite order at the origin* $x = 0$.
(2) A_1 *is a differential operator with smooth coefficients.*
(3) The distribution kernel of A_2 *is a function* $f(x, y)$ *which is smooth up to the diagonal for* $x > y$*, and also for* $x < y$.

Now we illustrate how the transmission property of the symbol ensures that the associated operator preserves smoothness up to the boundary. If A is a pseudo-differential operator in $L^m_{1,0}(\overline{\mathbf{R}^n_+})$, then we define a new operator

$$\begin{aligned} A_{\mathbf{R}^n_+} : C_0^\infty\left(\overline{\mathbf{R}^n_+}\right) &\longrightarrow C^\infty\left(\mathbf{R}^n_+\right) \\ u &\longmapsto A(u^0)\big|_{\mathbf{R}^n_+}, \end{aligned}$$

where u^0 is the extension of u to $\mathbf{R}^n$ by zero outside of $\overline{\mathbf{R}^n_+}$

$$u^0(x) = \begin{cases} u(x) & \text{for } x \in \overline{\mathbf{R}^n_+}, \\ 0 & \text{for } x \in \mathbf{R}^n \setminus \overline{\mathbf{R}^n_+}. \end{cases}$$

The operator $A_{\mathbf{R}^n_+}$ can be visualized as follows:

$$A_{\mathbf{R}^n_+} : C_0^\infty(\overline{\mathbf{R}^n_+}) \longrightarrow \mathcal{E}'(\mathbf{R}^n) \xrightarrow{A} \mathcal{D}'(\mathbf{R}^n) \longrightarrow C^\infty(\mathbf{R}^n_+),$$

where the first arrow is the zero extension to $\mathbf{R}^n$ and the last one is the restriction to $\mathbf{R}^n_+$. In view of the pseudo-local property of A, we note that $A_{\mathbf{R}^n_+}$ maps $C_0^\infty(\overline{\mathbf{R}^n_+})$ into $C^\infty(\mathbf{R}^n_+)$.

The transmission property implies that if u is smooth up to the boundary, then so is $A_{\mathbf{R}^n_+} u$. More precisely, we have the following theorem (see [RS, 2.3.1.2, Corollary 3; 2.3.2.6, Theorem 8]):

Theorem 7.34. *Let* $A \in L^m_{1,0}(\overline{\mathbf{R}^n_+})$*. Then we have the following two assertions (i) and (ii):*

(i) If A has the transmission property *with respect to the boundary* $\mathbf{R}^{n-1}$*, then* $A_{\mathbf{R}^n_+}$ *maps* $C_0^\infty\left(\overline{\mathbf{R}^n_+}\right)$ *continuously into* $C^\infty\left(\overline{\mathbf{R}^n_+}\right)$.

(ii) If A has the transmission property *with respect to the boundary* $\mathbf{R}^{n-1}$*, then* $A_{\mathbf{R}^n_+}$ *maps* $C^{k+\theta}_{\text{com}}\left(\overline{\mathbf{R}^n_+}\right)$ *continuously into* $C^{k-m+\theta}_{\text{loc}}\left(\overline{\mathbf{R}^n_+}\right)$ *for any integer* $k \geq m$*. Here* $C^{k+\theta}_{\text{com}}\left(\overline{\mathbf{R}^n_+}\right)$ *is the space of functions in* $C^k\left(\overline{\mathbf{R}^n_+}\right)$ *with compact support in* $\overline{\mathbf{R}^n_+}$ *and all of whose k-th order derivatives are Hölder continuous with*

exponent θ*, and* $C^{k-m+\theta}_{\text{loc}}\left(\overline{\mathbf{R}^n_+}\right)$ *is the space of functions in* $C^k\left(\overline{\mathbf{R}^n_+}\right)$ *all of whose* k*-th order derivatives are locally Hölder continuous with exponent* θ*, respectively.*

The situation of Theorem 7.34 can be visualized in the following diagram:

$$\begin{array}{ccc} C^{k+\theta}_{\text{comp}}(\overline{\mathbf{R}^n_+}) & \xrightarrow{A_{\mathbf{R}^n_+}} & C^{k-m+\theta}_{\text{loc}}(\overline{\mathbf{R}^n_+}) \\ \uparrow & & \uparrow \\ C^\infty_0(\overline{\mathbf{R}^n_+}) & \xrightarrow[A_{\mathbf{R}^n_+}]{} & C^\infty(\overline{\mathbf{R}^n_+}) \end{array}$$

Moreover, it should be noticed that the notion of transmission property is invariant under a change of coordinates which preserves the boundary. Hence this notion can be transferred to manifolds with boundary as follows. Indeed, if Ω is a relatively compact, open subset of an n-dimensional paracompact smooth manifold M without boundary (see Fig. 7.8), then the notion of transmission property can be extended to the class $L^m_{1,0}(M)$, upon using local coordinate systems flattening out the boundary $\partial\Omega$.

Then we have the following theorem (see [RS, 2.3.1.2, Theorem 4; 2.3.3.3, Theorem 1]):

Theorem 7.35. *Let* $A \in L^m_{1,0}(M)$*. Then we have the following two assertions (i) and (ii):*

(i) If A has the transmission property *with respect to the boundary* $\partial\Omega$*, then the operator*

$$\begin{aligned} A_\Omega : C^\infty(\overline{\Omega}) &\longrightarrow C^\infty(\Omega) \\ u &\longmapsto A(u^0)\big|_\Omega \end{aligned}$$

maps $C^\infty(\overline{\Omega})$ *continuously into itself, where* u^0 *is the extension of u to M by zero outside of* $\overline{\Omega}$*.*

(ii) If A has the transmission property *with respect to the boundary* $\partial\Omega$*, then the operator* A_Ω *maps* $C^{k+\theta}(\overline{\Omega})$ *continuously into* $C^{k-m+\theta}(\overline{\Omega})$ *for any integer* $k \geq m$ *and* $0 < \theta < 1$.

The situation of Theorem 7.35 can be visualized in the following diagram:

$$\begin{array}{ccc} C^{k+\theta}(\overline{\Omega}) & \xrightarrow{A_\Omega} & C^{k-m+\theta}(\overline{\Omega}) \\ \uparrow & & \uparrow \\ C^\infty(\overline{\Omega}) & \xrightarrow[A_\Omega]{} & C^\infty(\overline{\Omega}) \end{array}$$

7.7 The Boutet de Monvel Calculus

Elliptic boundary value problems cannot be treated directly by pseudo-differential operator methods. It was Boutet de Monvel [Bo] who brought in the operator-algebraic aspect with his calculus in 1971. He constructed a relatively small "algebra", called the Boutet de Monvel algebra, which contains the boundary value problems for elliptic differential operators as well as their parametrices.

7.7.1 Trace, Potential and Singular Green Operators on the Half-Space $\mathbf{R}^n_+$

In this subsection we give basic definitions and properties of classes of trace, potential and singular Green operators on the half-space $\mathbf{R}^n_+$. The presentation here is based on Rempel–Schulze [RS, Section 2.3.2].

Pseudo-differential trace operators T are a natural generalization of the usual differential trace operators from elliptic boundary value problems, while potential operators K can be described as the adjoints of trace operators T with respect to the L^2 inner products. Singular Green operators G are introduced in order to get an algebra of matrices of operators of the form

$$\mathcal{A} = \begin{pmatrix} P_\Omega + G & K \\ T & S \end{pmatrix},$$

called the *Boutet de Monvel algebra*. In fact, a typical example of a singular Green is the composition of a trace operator T and a potential operator K.

(1) A function $k(x', y', \xi', \nu) \in C^\infty(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R})$ is called a *potential symbol* of order m if it satisfies the condition

$$k(x', y', \xi', \nu) = \sum_{j=0}^{\infty} k_j(x', y', \xi') \frac{(1 - i\nu \langle \xi' \rangle^{-1})^j}{(1 + i\nu \langle \xi' \rangle^{-1})^{j+1}}.$$

Here the symbols $k_j(x', y', \xi')$ form a rapidly decreasing sequence in the class $S^m_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$.

Then the *potential operator*

$$K : C_0^\infty(\mathbf{R}^{n-1}) \longrightarrow C^\infty(\overline{\mathbf{R}^n_+})$$

is defined as an oscillatory integral by the formula

$$(Kv)(x', x_n) = \frac{1}{(2\pi)^n} \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'} \left(\int_{\mathbf{R}^{n-1}} k(x', y', \xi', \nu)\, d\nu \right) v(y')\, dy'\, d\xi', \quad v \in C_0^\infty(\mathbf{R}^{n-1}).$$

(2) A function $t(x', y', \xi', \nu) \in C^\infty(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R})$ is called a *trace symbol* of order m and type d if it satisfies the condition

$$t(x', y', \xi', \nu) = \sum_{k=0}^{d-1} b_k(x', y', \xi') \left(\nu \langle \xi' \rangle^{-1} \right)^k + \sum_{j=0}^{\infty} t_j(x', y', \xi') \frac{(1 + i\nu \langle \xi' \rangle^{-1})^j}{(1 - i\nu \langle \xi' \rangle^{-1})^{j+1}}.$$

Here $b_k(x', y', \xi') \in S^m_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$ and the symbols $t_j(x', y', \xi')$ form a rapidly decreasing sequence in the class $S^m_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$.

Then the *trace operator*

$$T : C^\infty_{(0)}(\overline{\mathbf{R}^n_+}) \longrightarrow C^\infty(\mathbf{R}^{n-1})$$

is defined as an oscillatory integral by the formula

$$(Tu)(x') = \frac{1}{(2\pi)^n} \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'} \times \left[\int_\gamma t(x', y', \xi', \nu) \left(\int_{\mathbf{R}} e^{-iy_n\nu} u^0(y', y_n) dy_n \right) d\nu \right] dy'\, d\xi', \quad u \in C^\infty_{(0)}(\overline{\mathbf{R}^n_+}).$$

Here γ is a large circle in the upper complex half-plane $\{\nu \in \mathbf{C} : \operatorname{Im} \nu > 0\}$, and u^0 is the extension of u to $\mathbf{R}^n$ by zero outside of $\overline{\mathbf{R}^n_+}$

$$u^0(y', y_n) = \begin{cases} u(y', y_n) & \text{for } y_n \geq 0, \\ 0 & \text{for } y_n < 0, \end{cases}$$

and

$$C^\infty_{(0)}(\overline{\mathbf{R}^n_+}) = \left\{ u|_{\overline{\mathbf{R}^n_+}} : u \in C_0^\infty(\mathbf{R}^n) \right\}.$$

(3) A function $b(x', y', \xi', \nu, \tau) \in C^\infty(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R} \times \mathbf{R})$ is called a *singular Green symbol* of order m and type d if it satisfies the condition

$$
\begin{aligned}
& b(x', y', \xi', \nu, \tau) \\
&= \sum_{\ell=0}^{d-1} \left(\sum_{j=0}^{\infty} c_{j\ell}(x', y', \xi') \frac{(1 - i\nu \langle \xi' \rangle^{-1})^j}{(1 + i\nu \langle \xi' \rangle^{-1})^{j+1}} \right) \left(\tau \langle \xi' \rangle^{-1} \right)^{\ell} \\
&\quad + \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} b_{j\ell}(x', y', \xi') \frac{(1 - i\nu \langle \xi' \rangle^{-1})^j}{(1 + i\nu \langle \xi' \rangle^{-1})^{j+1}} \frac{(1 + i\tau \langle \xi' \rangle^{-1})^{\ell}}{(1 - i\tau \langle \xi' \rangle^{-1})^{\ell+1}}.
\end{aligned}
$$

Here the symbols $c_{j\ell}(x', y', \xi')$ form a rapidly decreasing sequence in the class $S^m_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$ with respect to j, for each $\ell = 0, 1, 2, \ldots, d-1$, and the symbols $b_{j\ell}(x', y', \xi')$ form a rapidly decreasing double sequence in the class $S^m_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$.

Then the *singular Green operator*

$$
G : C^{\infty}_{(0)}(\overline{\mathbf{R}^n_+}) \longrightarrow C^{\infty}(\overline{\mathbf{R}^n_+})
$$

is defined as an oscillatory integral by the formula

$$
\begin{aligned}
& (Gu)(x', x_n) \\
&= \frac{1}{(2\pi)^{n+1}} \iint_{\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}} e^{i(x'-y')\cdot\xi'} \\
&\quad \times \left[\int_{\gamma} \left(\int_{\mathbf{R}} e^{ix_n\nu} b(x', y', \xi', \nu, \tau)\, d\nu \right) \left(\int_{\mathbf{R}} e^{-iy_n\tau} u^0(y', y_n) dy_n \right) d\tau \right] dy'\, d\xi', \\
&\qquad u \in C^{\infty}_{(0)}(\overline{\mathbf{R}^n_+}).
\end{aligned}
$$

Here γ is a large circle in the upper complex half-plane $\{\tau \in \mathbf{C} : \operatorname{Im} \tau > 0\}$.

7.7.2 The Boutet de Monvel Algebra

Let Ω be a relatively compact, open subset of an n-dimensional compact smooth manifold M without boundary (see Fig. 7.4). Boutet de Monvel [Bo] introduced matrices of operators

$$
\mathcal{A} = \begin{pmatrix} P_{\Omega} + G & K \\ T & S \end{pmatrix} : \begin{matrix} C^{\infty}(\overline{\Omega}) \\ \oplus \\ C^{\infty}(\partial\Omega) \end{matrix} \longrightarrow \begin{matrix} C^{\infty}(\overline{\Omega}) \\ \oplus \\ C^{\infty}(\partial\Omega) \end{matrix}
$$

Here:

(1) P is a pseudo-differential operator on the full manifold M and

$$P_\Omega u = P(u^0)|_\Omega \quad \text{for all } u \in C^\infty(\overline{\Omega}),$$

where u^0 is the extension of u by zero to M

$$u^0(x) = \begin{cases} u(x) & \text{for } x \in \overline{\Omega}, \\ 0 & \text{for } x \in M \setminus \overline{\Omega}. \end{cases}$$

The operator P_Ω can be visualized as follows:

$$P_\Omega : C^\infty(\overline{\Omega}) \longrightarrow \mathcal{D}'(M) \stackrel{P}{\longrightarrow} \mathcal{D}'(M) \longrightarrow C^\infty(\Omega),$$

where the first arrow is the zero extension to M and the last one is the restriction to Ω. In view of the pseudo-local property of P, we note that P_Ω maps $C^\infty(\overline{\Omega})$ into $C^\infty(\Omega)$.

The crucial requirement is that the symbol of P has the *transmission property* in order that P_Ω maps $C^\infty(\overline{\Omega})$ into itself.

(2) S is a pseudo-differential operator on $\partial\Omega$.

(3) The potential operator K and trace operator T are generalizations of the potentials and trace operators known from the classical theory of elliptic boundary value problems, respectively.

(4) The entry G, a *singular Green operator*, is an operator which is smoothing in the interior Ω while it acts like a pseudo-differential operator in directions tangential to the boundary $\partial\Omega$. As an example, we may take the difference of two solution operators to (invertible) classical boundary value problems with the same differential part in the interior but different boundary conditions.

Boutet de Monvel [Bo] proved that these operator matrices form an algebra in the following sense (see [RS, 2.3.3.2, Theorem 1; 3.1.1.1, Theorem 2]): Given another element of the calculus, say,

$$\mathcal{A}' = \begin{pmatrix} P'_\Omega + G' & K' \\ T' & S' \end{pmatrix} : \begin{matrix} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{matrix} \longrightarrow \begin{matrix} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{matrix}$$

the composition $\mathcal{A}'\mathcal{A}$ is again an operator matrix of the type described above. It is worth pointing out here that the product $P'_\Omega P_\Omega$ does not coincide with $(P'P)_\Omega$; in fact, the difference $P'_\Omega P_\Omega - (P'P)_\Omega$ turns out to be a singular Green operator (see [RS, 2.1.2.3, Lemma 5]).

For example, we consider the Dirichlet problem

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u = \varphi & \text{on } \partial\Omega. \end{cases} \tag{D}$$

Then problem (D) corresponds to the map

$$\mathcal{D} = \begin{pmatrix} \Delta_\Omega \\ \gamma_0 \end{pmatrix} : C^\infty(\overline{\Omega}) \longrightarrow \begin{array}{c} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{array}$$

where

$$\gamma_0 u = u|_{\partial\Omega}, \quad u \in C^\infty(\overline{\Omega}),$$

is the trace operator on $\partial\Omega$. The map $\mathcal{D}$ defines an isomorphism and its inverse $\mathcal{D}^{-1}$ has the form

$$\mathcal{D}^{-1} = \begin{pmatrix} P_\Omega + G & L \end{pmatrix} : \begin{array}{c} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{array} \longrightarrow C^\infty(\overline{\Omega}),$$

with a pseudo-differential operator P, a singular Green operator G and a potential operator L. More precisely, if $g(x,z)$ and $\ell(x,y)$ are the classical Green and Poisson kernels of problem (D) respectively, then it follows that

$$u(x) = \int_\Omega g(x,z) f(z)\, dz + \int_{\partial\Omega} \ell(x,y)\varphi(y)\, d\omega(y),$$

so that

$$\begin{aligned} (P_\Omega + G) f &= \int_\Omega g(x,z) f(z)\, dz, \\ L\varphi &= \int_{\partial\Omega} \ell(x,y)\varphi(y)\, d\omega(y). \end{aligned}$$

Furthermore, it should be noted that

$$\begin{pmatrix} \Delta_\Omega \\ \gamma_0 \end{pmatrix} \begin{pmatrix} P_\Omega + G & L \end{pmatrix} = \begin{pmatrix} \Delta_\Omega(P_\Omega + G) & \Delta_\Omega L \\ \gamma_0(P_\Omega + G) & \gamma_0 L \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \tag{7.12}$$

since $\mathcal{D}^{-1}$ is the two-sided inverse of $\mathcal{D}$.

If the map

$$\mathcal{A} = \begin{pmatrix} \Delta_\Omega \\ B \end{pmatrix} : C^\infty(\overline{\Omega}) \longrightarrow \begin{array}{c} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{array}$$

is another elliptic boundary value problem such as the Neumann problem, then we have, by formula (7.12),

$$\mathcal{A}\mathcal{D}^{-1} = \begin{pmatrix} \Delta_\Omega \\ B \end{pmatrix} \left(P_\Omega + G \ \ L \right) = \begin{pmatrix} I & 0 \\ B(P_\Omega + G) & BL \end{pmatrix},$$

and the right lower corner

$$Q = BL : C^\infty(\partial\Omega) \longrightarrow C^\infty(\partial\Omega)$$

is a pseudo-differential operator on $\partial\Omega$. The ellipticity of $\mathcal{A}$ and $\mathcal{D}$ implies that the operator Q is elliptic. Let R be a *parametrix* for Q:

$$\begin{cases} RQ \equiv I \mod L^{-\infty}(\partial\Omega), \\ QR \equiv I \mod L^{-\infty}(\Omega). \end{cases}$$

If we let

$$\mathcal{C} = \begin{pmatrix} I & 0 \\ -RB(P_\Omega + G) & R \end{pmatrix},$$

then we obtain that

$$\begin{aligned} \mathcal{A}\left(\mathcal{D}^{-1}\mathcal{C}\right) &= \begin{pmatrix} I & 0 \\ B(P_\Omega + G) & BL \end{pmatrix} \begin{pmatrix} I & 0 \\ -RB(P_\Omega + G) & R \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ B(P_\Omega + G) - (BL)RB(P_\Omega + G) & (BL)R \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ B(P_\Omega + G) - (QR)B(P_\Omega + G) & QR \end{pmatrix} \\ &\equiv \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \end{aligned}$$

This proves that the map

$$\mathcal{D}^{-1}\mathcal{C} = \left(P_\Omega + G - LRB(P_\Omega + G) \ \ LR \right) : \begin{matrix} C^\infty(\overline{\Omega}) \\ \oplus \\ C^\infty(\partial\Omega) \end{matrix} \longrightarrow C^\infty(\overline{\Omega})$$

is a parametrix for the map $\mathcal{A}$.

Therefore, starting at the map $\mathcal{D}$ we can find the parametrix $\mathcal{D}^{-1}\mathcal{C}$ of another elliptic map $\mathcal{A}$ by calculating the parametrix R of the elliptic pseudo-differential operator $Q = BL$ on the boundary $\partial\Omega$. In Sect. 10.4 we will return this reduction to the boundary in a more general setting (see Theorems 10.19 and 10.21).

7.8 Distribution Kernel of a Pseudo-differential Operator

In this section, following Coifman–Meyer [CM, Chapitre IV, Proposition 1] we prove that the distribution kernel $s(x, y)$ of a pseudo-differential operator $S \in L^m_{1,0}(\mathbf{R}^n)$ satisfies the estimate

$$|s(x, y)| \leq \frac{C}{|x - y|^{m+n}}, \quad x, y \in \mathbf{R}, \;\; x \neq y.$$

More precisely, we can obtain the following fundamental result:

Theorem 7.36. *Let $\sigma(x, \xi)$ be a symbol in the class $S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^n)$ such that*

$$\left|\sigma^{(\alpha)}_{(\beta)}(x, \xi)\right| \leq C_{\alpha\beta} \left(1 + |\xi|\right)^{m-|\alpha|}, \quad (x, \xi) \in \mathbf{R}^n \times \mathbf{R}^n. \tag{7.13}$$

We let

$$r(x, z) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{iz\cdot\xi} \sigma(x, \xi)\, d\xi,$$

where the integral is taken in the sense of oscillatory integrals (see Theorem 7.10). Then the distribution $r(x, z)$ satisfies the condition

$$r(x, z) \in C^\infty \left(\mathbf{R}^n \times \mathbf{R}^n \setminus \{0\}\right),$$

and the estimate

$$|r(x, z)| \leq \frac{C}{|z|^{n+m}}, \quad z \neq 0. \tag{7.14}$$

Proof. First, if we let the usual Laplacian

$$\Delta_\xi = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} + \ldots + \frac{\partial^2}{\partial \xi_n^2},$$

then it follows from condition (7.13) that

$$\left|\Delta_\xi^N (\sigma(x, \xi))\right| \leq C_N \left(1 + |\xi|^2\right)^{m/2-N}, \quad (x, \xi) \in \mathbf{R}^n \times \mathbf{R}^n. \tag{7.15}$$

The proof is divided into two steps.

Step 1: The case where $|z| \geq 1$: By taking a large integer N such that

$$N > \frac{n + m}{2},$$

we have, by integration by parts,

$$r(x,z) = \frac{(-1)^N}{|z|^{2N}} \int_{\mathbf{R}^n} e^{iz\cdot\xi} \Delta_\xi^N (\sigma(x,\xi))\, d\xi.$$

Hence, by using the polar coordinates we obtain from estimate (7.15) that

$$\begin{aligned} |r(x,z)| &\le \frac{1}{|z|^{2N}} \int_{\mathbf{R}^n} \left| \Delta_\xi^N (\sigma(x,\xi)) \right| d\xi \\ &\le \frac{C_N}{|z|^{2N}} \int_{\mathbf{R}^n} \left(1 + |\xi|^2\right)^{m/2-N} d\xi \\ &= \frac{C_N}{|z|^{2N}} \int_0^\infty \int_{\Sigma_n} \left(1 + r^2\right)^{m/2-N} r^{n-1}\, dr\, d\sigma \\ &= \frac{C_N}{|z|^{2N}} \omega_n \int_0^\infty \left(1 + r^2\right)^{m/2-N} r^{n-1}\, dr \\ &\le \frac{C'_N}{|z|^{n+m}} \quad \text{for all } |z| \ge 1, \end{aligned} \tag{7.16}$$

since we have $2N > n + m$. Here

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad n \ge 3,$$

is the surface area of the unit sphere Σ_n.

Step 2: The case where $0 < |z| \le 1$: Let

$$R = \frac{1}{|z|} \ge 1,$$

and take a smooth function $\varphi(t) \in C_0^\infty(\mathbf{R})$ such that

$$\varphi(t) = \begin{cases} 1 & \text{for } |t| \le \frac{1}{2}, \\ 0 & \text{for } |t| \ge 1. \end{cases}$$

Then we can decompose the kernel $r(x,z)$ into the following two terms $r_1(x,z)$ and $r_2(x,z)$:

$$\begin{aligned} r(x,z) &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{iz\cdot\xi} \varphi\left(\frac{\xi}{R}\right) \sigma(x,\xi)\, d\xi \\ &\quad + \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{iz\cdot\xi} \left(1 - \varphi\left(\frac{\xi}{R}\right)\right) \sigma(x,\xi)\, d\xi \\ &:= r_1(x,z) + r_2(x,z). \end{aligned}$$

In this way we are reduced to the estimate of the two terms $r_1(x,z)$ and $r_2(x,z)$.

(a) The estimate of the term $r_1(x,z)$: Since we have the inequality

$$\left| e^{iz\cdot\xi}\varphi\left(\frac{\xi}{R}\right)\sigma(x,\xi)\right| \leq C\,|\xi|^m \quad \text{for } |\xi| \leq R,$$

it follows that

$$\begin{aligned} |r_1(x,z)| &\leq \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\left| e^{iz\cdot\xi}\varphi\left(\frac{\xi}{R}\right)\sigma(x,\xi)\right| d\xi \\ &\leq \frac{1}{(2\pi)^n}\int_{|\xi|\leq R}\left|\varphi\left(\frac{\xi}{R}\right)\sigma(x,\xi)\right| d\xi \\ &\leq C'\int_{|\xi|\leq R}|\xi|^m\, d\xi = C'\int_0^R\int_{\Sigma_n} r^m r^{n-1}\,dr\,d\sigma \\ &= C'\omega_n\,\frac{R^{m+n}}{m+n} = \frac{C''}{|z|^{n+m}}. \end{aligned} \tag{7.17}$$

(b) The estimate of the term $r_2(x,z)$: Take

$$N > \frac{n+m}{2}$$

just as in Step 1. Then, by integration by parts it follows that

$$r_2(x,z) = \frac{(-1)^N}{|z|^{2N}}\int_{\mathbf{R}^n} e^{iz\cdot\xi}\Delta_\xi^N\left(\sigma(x,\xi)\left(1-\varphi\left(\frac{\xi}{R}\right)\right)\right)d\xi.$$

Hence we have the inequality

$$\begin{aligned} &|r_2(x,z)| \\ &\leq \frac{1}{|z|^{2N}}\int_{\mathbf{R}^n}\left|\Delta_\xi^N\left(\sigma(x,\xi)\left(1-\varphi\left(\frac{\xi}{R}\right)\right)\right)\right| d\xi \\ &\leq \frac{C_N}{|z|^{2N}}\left\{\int_{|\xi|\geq R}\left(1+|\xi|^2\right)^{m/2-N} d\xi + \frac{1}{R}\int_{|R/2\leq\xi|\leq R}|\xi|^{2m-2N-1}\,d\xi\right\} \\ &= \frac{C_N}{|z|^{2N}}\left\{\int_R^\infty\int_{\Sigma_n}\left(1+r^2\right)^{m/2-N} r^{n-1}dr\,d\sigma\right. \\ &\quad \left. +\frac{1}{R}\int_{R/2}^R\int_{\Sigma_n} r^{m-2N-1}r^{n-1}dr\,d\sigma\right\} \end{aligned} \tag{7.18}$$

$$\leq \frac{C'_N}{|z|^{2N}} \left(\frac{1}{R^{2N-m-n}} + \frac{1}{R^{2N-m-n+2}} \right)$$
$$\leq \frac{2C'_N}{|z|^{2N}} \left(\frac{1}{R} \right)^{2N-m-n} = \frac{2C'_N}{|z|^{2N}} |z|^{2N-m-n} = \frac{2C'_N}{|z|^{n+m}}.$$

Therefore, we obtain from estimates (7.17) and (7.18) that

$$|r(x,z)| \leq |r_1(x,z)| + |r_2(x,z)| \leq \frac{C'_N}{|z|^{n+m}} \quad \text{for all } 0 < |z| \leq 1. \tag{7.19}$$

The desired estimate (7.14) follows by combining the two estimates (7.16) and (7.19).

The proof of Theorem 7.36 is complete.

7.9 Notes and Comments

The development of the theory of pseudo-differential operators has greatly advanced our understanding of partial differential equations, and the pseudo-differential calculus has become an indispensable tool in contemporary analysis, in particular on compact manifolds without boundary. Our treatment of pseudo-differential operators follows the book of Chazarain–Piriou [CP]. For detailed studies of pseudo-differential operators, the reader might be referred to Chazarain–Piriou [CP], Èskin [Es], Hörmander [Ho4], Kumano-go [Ku], Rempel–Schulze [RS], Schulze [Su2] and Taylor [Ty].

Section 7.1: The material in this section is taken from Abraham–Marsden–Ratiu [AMR], Lang [Lg] and Munkres [Mu]. This section is given for completeness, to minimize the necessity of consulting too many outside references. Theorems 7.2 and 7.3 are taken from Munkres [Mu].

Section 7.2: For the double $M = \hat{\Omega}$ of Ω, see Munkres [Mu]. The trace theorem and the Rellich–Kondrachov theorem are taken from Bergh–Löfström [BL], Stein [Sn1], Taibleson [Tb] and Triebel [Tr]. The following diagram summarizes three pillars for *sequential compactness* in analysis:

Sequences of	Compactness theorems
real numbers	the Bolzano–Weierstrass theorem
continuous functions	the Ascoli–Arzelà theorem
distributions	the Rellich–Kondrachov theorem

Section 7.3: The symbol classes $S^m_{\rho,\delta}(\Omega \times \mathbf{R}^N)$ were first introduced by Hörmander [Ho2] (see also Seeley [Se1]). In Appendix B we shall introduce matrix-valued symbols in order to study systems of pseudo-differential operators.

For the theory of Fourier integral operators, see Hörmander [Ho3], Duistermaat–Hörmander [DH] and Duistermaat [Du].

Section 7.4: A proof of Theorem 7.18 is given in Chazarain–Piriou [CP, Chapitre IV, Théorème 2.1]. For the stationary phase theorem, see Chazarain–Piriou [CP, Chapitre III, Théorème 9.3]. Theorem 7.24 is due to Bourdaud [Bd, Theorem 1]. The notion of hypoellipticity was introduced by Schwartz [Sz]. Theorem 7.26 is due to Hörmander [Ho2, Theorem 4.2]. Hypoelliptic second-order differential operators have been studied in great detail by Oleĭnik–Radkevič [OR] and many others.

Section 7.5: Theorem 7.28 is taken from Chazarain–Piriou [CP, Chapitre V, Théorème 2.4] and Theorem 7.29 is taken from Chazarain–Piriou [CP, Chapitre V, Théorème 2.5], respectively. The reader might be referred to the original works Hörmander [Ho1] and Seeley [Se2]. See Polking [Po] for the parabolic case.

Section 7.6: Following Schrohe [Sr5, Definition 2.3], we can introduce an equivalent definition of the transmission property for general symbols $a(x, y, \xi)$ in the class $S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n)$ as in Sect. 7.4.1. To do this, let H be the linear space of all complex-valued functions $f(t)$ on the real line $\mathbf{R}$ which are of class C^∞ and have a regular pole at infinity. More precisely, a function $f(t) \in C^\infty(\mathbf{R})$ belongs to H if and only if it has a unique expansion

$$f(t) = \sum_{s=1}^{N} \alpha_s t^s + \sum_{k=-\infty}^{\infty} \alpha_k \left(\frac{1-it}{1+it} \right)^k, \quad i = \sqrt{-1},$$

where the coefficients α_k form a rapidly decreasing sequence (see [RS, Chapter 2, Section 2.1.1]). A symbol

$$a(x, y, \xi) \in S^m_{1,0}(\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n)$$

is said to have the *transmission property* at $\{x_n = y_n = 0\}$, provided that we have, for all non-negative integers k and ℓ,

$$\frac{\partial^{k+\ell} a}{\partial x_n^k \partial y_n^\ell}(x', 0, y', 0, \xi', \langle \xi' \rangle \xi_n) \in S^m_{1,0}(\mathbf{R}^{n-1}_{x'} \times \mathbf{R}^{n-1}_{y'} \times \mathbf{R}^{n-1}_{\xi'}) \hat{\otimes}_\pi H_{\xi_n},$$

where

$$\langle \xi' \rangle = \left(1 + |\xi'|^2\right)^{1/2}.$$

The subscripts x', y', ξ' and ξ_n are used in order to indicate the variable for which we have the corresponding property.

Hörmander [Ho4] introduces a similar notion: the bilateral transmission property. An operator on the half-space $\mathbf{R}^n_+$ has the *bilateral transmission property* if it maps functions which are smooth up to the boundary to functions with the same property (see [Ho4, Definition 18.2.13]). The difference between both notions is that the bilateral transmission property assures the transmission property also for the opposite half-space $\mathbf{R}^n_-$. This is needed in order to obtain a calculus which is closed under taking adjoints.

Section 7.7: Boundary value problems cannot be treated directly by pseudo-differential methods. Already in the 1960s, however, first essential steps were taken to provide a similar framework allowing the construction of parametrices to elliptic elements (see Višik–Èskin [VE]). More detailed and concise accounts of the transmission property and the Boutet de Monvel calculus are given in the books by Rempel–Schulze [RS], Grubb [Gb] and Schrohe [Sr5]. See also Èskin [Es, Chapter III] and Hörmander [Ho4, Section 18.2].

The operators in the Boutet de Monvel calculus may be regarded as operator-valued pseudo-differential operators. This point of view, going back to an idea of Schulze, was first sketched by Schrohe–Schulze [SS1]. In Appendix B, following Schrohe [Sr5], we show the pseudo-differential spirit of Boutet de Monvel's construction more closely than the older descriptions. This concept has been applied successfully to the analysis of boundary value problems on singular manifolds. In fact, in this operator-valued set-up, the Boutet de Monvel calculus can be combined very well with pseudo-differential calculi for cone and edge singularities (see [Su1, SS1, SS2, SS3, SS4, SS5]).

Section 7.8: This section is adapted from Coifman–Meyer [CM, Chapitre IV] and Nagase [Na].

Chapter 8
Waldenfels Operators and Maximum Principles

In this chapter, following Bony–Courrège–Priouret [BCP], we prove various maximum principles for second-order elliptic Waldenfels operators which play an essential role throughout the book. In Sect. 8.1 we give complete characterizations of linear operators which satisfy the positive maximum principle related to condition (β') in Theorem 9.50 in Chap. 9 (Theorems 8.2, 8.4 and 8.8). In Sect. 8.2 we prove the weak and strong maximum principles and Hopf's boundary point lemma for second-order elliptic Waldenfels operators (Theorems 8.11, 8.13 and 8.15).

8.1 Borel Kernels and Maximum Principles

In this section we give complete characterizations of linear operators which satisfy the positive maximum principle related to condition (β') in Theorem 9.50. It should be emphasized that these characterizations give, as a byproduct, the following characterizations of distributions of order 2 which are *non-negative* outside the origin (see Bony–Courrège–Priouret [BCP, Proposition I.1.2]):

Proposition 8.1. *(I) Let T be a linear operator (functional) from $C_0^2(\mathbf{R}^n)$ into* $\mathbf{R}$. *Then $T : C_0^2(\mathbf{R}^n) \to \mathbf{R}$ is continuous and satisfies the condition*

$$u \in C_0^2(\mathbf{R}^n),\ u \geq 0 \ \textit{in}\ \mathbf{R}^n \ \textit{and}\ 0 \notin \operatorname{supp} u \Longrightarrow T(u) \geq 0$$

if and only if there exist real numbers a^{ij}, b^i and c and a positive Radon measure *μ on $\mathbf{R}^n \setminus \{0\}$ such that the linear functional T can be written in the form*

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_8,

$$\begin{aligned} &T(u) \qquad (8.1)\\ &= \sum_{i,j=1}^{n} a^{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(0) + \sum_{i=1}^{n} b^i \frac{\partial u}{\partial x_i}(0) + c\, u(0) \\ &\quad + \int_{\mathbf{R}^n \setminus \{0\}} \mu(dy) \left[u(y) - \varphi(y) \left(u(0) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(0)\, y_i \right) \right], \quad u \in C_0^2(\Omega). \end{aligned}$$

Here:

(a) *$\varphi \in C_0^\infty(\mathbf{R}^n)$ and $\varphi = 1$ in a neighborhood of* 0.
(b) *The Radon measure μ satisfies the condition*

$$\int_{0<|y|^2\le 1} |y|^2\, \mu(dy) < \infty.$$

(II) *If a linear operator (functional) $T : C_0^\infty(\mathbf{R}^n) \to \mathbf{R}$ satisfies the condition*

$$u \in C_0^2(\mathbf{R}^n),\ u \ge 0 \ \text{ in } \mathbf{R}^n \text{ and } \ u(0) = 0 \Longrightarrow T(u) \ge 0,$$

then the linear functional T can be extended uniquely to a continuous *linear functional $T : C_0^2(\mathbf{R}^n) \to \mathbf{R}$ of the form* (8.1)*. In this case, the matrix (a^{ij}) is non-negative definite:*

$$\sum_{i,j=1}^{n} a^{ij}\xi_i\xi_j \ge 0 \quad \text{for all } \xi \in \mathbf{R}^n.$$

Let Ω be an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$. More precisely, if Ω is an open set in $\overline{\mathbf{R}^n_+}$ in the topology induced on $\overline{\mathbf{R}^n_+}$ from $\mathbf{R}^n$, then we define the *boundary* $\partial\Omega$ of Ω to be the intersection of Ω with $\mathbf{R}^{n-1} \times \{0\}$ and the *interior* $\overset{\circ}{\Omega}$ of Ω to be the complement of $\partial\Omega$ in Ω, that is,

$$\begin{aligned} \partial\Omega &= \Omega \cap \{x \in \mathbf{R}^n : x_n = 0\}, \\ \overset{\circ}{\Omega} &= \Omega \cap \{x \in \mathbf{R}^n : x_n > 0\}. \end{aligned}$$

We note (see Figs. 8.1 and 8.2) that

$$\Omega = \begin{cases} \overset{\circ}{\Omega} & \text{if } \Omega \text{ is open in } \mathbf{R}^n, \\ \overset{\circ}{\Omega} \cup \partial\Omega & \text{if } \Omega \text{ is open in } \overline{\mathbf{R}^n_+}. \end{cases}$$

Then we let

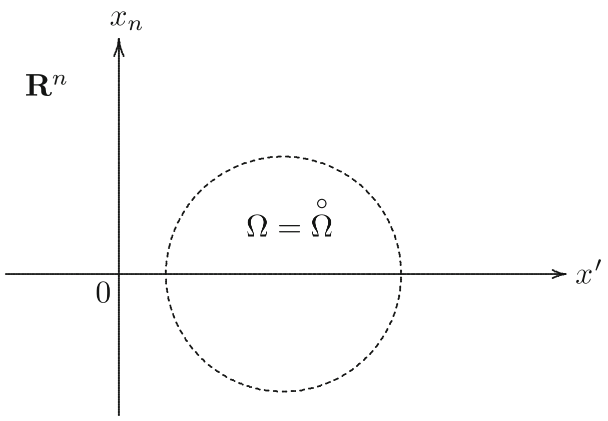


Fig. 8.1 Ω is open in $\mathbf{R}^n$

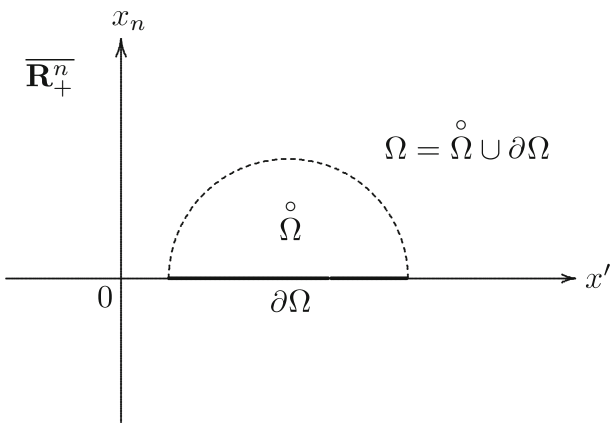


Fig. 8.2 Ω is open in $\overline{\mathbf{R}^n_+}$

$$B_{\text{loc}}(\overset{\circ}{\Omega}) = \text{the space of Borel measurable functions in } \overset{\circ}{\Omega}$$
$$\text{which are bounded on compact subsets of } \overset{\circ}{\Omega},$$

and

$$C_0(\Omega) = \text{the space of continuous functions in } \Omega$$
$$\text{with compact support in } \Omega.$$

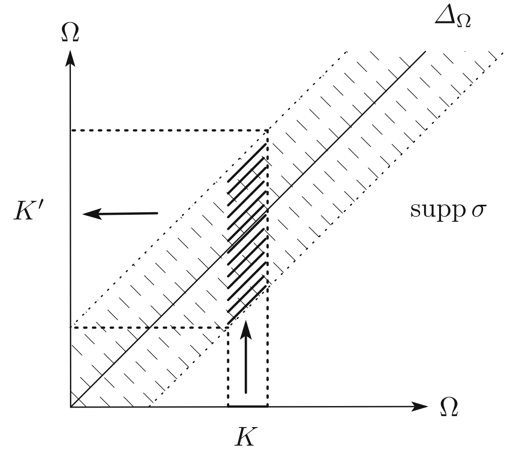


Fig. 8.3 Condition (LU3) on supp σ

Let $\mathcal{B}_\Omega$ and $\mathcal{B}_{\mathring{\Omega}}$ be the σ-algebra of all Borel sets in Ω and the σ-algebra of all Borel sets in $\mathring{\Omega}$, respectively. A *positive Borel kernel* of $\mathring{\Omega}$ into Ω is a mapping

$$\mathring{\Omega} \ni x \longmapsto s(x, dy)$$

of $\mathring{\Omega}$ into the space of non-negative measures on $\mathcal{B}_\Omega$ such that, for each $X \in \mathcal{B}_\Omega$, the function

$$\mathring{\Omega} \ni x \longmapsto s(x, X) = \int_X s(x, dy)$$

is Borel measurable in $\mathring{\Omega}$.

A *local unity function* on Ω is a smooth function $\sigma(x, y)$ in $\Omega \times \Omega$ which satisfies the following three conditions (LU1)–(LU3) (see Fig. 8.3):

(LU1) $0 \leq \sigma(x, y) \leq 1$ in $\Omega \times \Omega$.
(LU2) $\sigma(x, y) = 1$ in a neighborhood of the *diagonal* $\Delta_\Omega = \{(x, x) : x \in \Omega\}$ in $\Omega \times \Omega$.
(LU3) For any compact subset K of Ω, there exists a compact subset K' of Ω such that the functions $\sigma_x(\cdot) = \sigma(x, \cdot)$, $x \in K$, have their support in K'.

We can construct a local unity function $\sigma(x, y)$ in the following way: Let $\{U_i\}_{i \in I}$ be an open covering of Ω and let $\{\varphi_i\}_{i \in I}$ be a partition of unity subordinate to the covering $\{U_i\}$ (see Sect. 5.7.2). That is, the family $\{\varphi_i\}_{i \in I}$ satisfies the following three conditions (PU1)–(PU3) (see Fig. 8.4):

(PU1) $0 \leq \varphi_i(x) \leq 1$ for all $x \in \Omega$ and $i \in I$.
(PU2) $\operatorname{supp} \varphi_i \subset U_i$ for each $i \in I$.
(PU3) The collection $\{\operatorname{supp} \varphi_i\}_{i \in I}$ is locally finite and

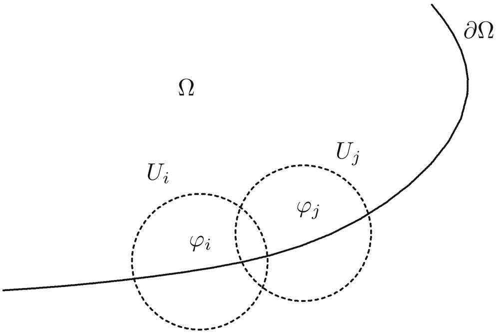


Fig. 8.4 The open covering $\{U_i\}$ and the partition of unity $\{\varphi_i\}$

$$\sum_{i \in I} \varphi_i(x) = 1 \quad \text{for every } x \in \Omega.$$

Here $\operatorname{supp} \varphi_i$ is the support of φ_i, i.e. the closure in Ω of the set $\{x \in \Omega : \varphi_i(x) \neq 0\}$.

If we take a smooth function $\psi_i(x)$ in Ω such that

$$\begin{cases} 0 \leq \psi_i(x) \leq 1 & \text{for all } x \in \Omega, \\ \psi_i(x) = 1 & \text{on } \operatorname{supp} \varphi_i, \end{cases}$$

then it is easy to verify that the function

$$\sigma(x, y) = \sum_{i \in I} \varphi_i(x)\, \psi_i(y), \quad (x, y) \in \Omega \times \Omega,$$

satisfies the desired conditions (LU1)–(LU3).

8.1.1 Linear Operators having Positive Borel Kernel

Now we assume that a positive Borel kernel $s(x, dy)$ satisfies the following two conditions (NS.1) and (NS.2):

(NS.1) $s(x, \{x\}) = 0$ for every $x \in \overset{\circ}{\Omega}$.
(NS.2) For every non-negative function $f(x)$ in $C_0(\Omega)$, the function

$$x \longmapsto \int_\Omega s(x, dy)|y - x|^2 f(y), \quad x \in \overset{\circ}{\Omega},$$

belongs to the space $B_{\mathrm{loc}}(\overset{\circ}{\Omega})$.

By using Taylor's formula and condition (NS.2), we can define a linear operator

$$S : C_0^2(\Omega) \longrightarrow B_{\text{loc}}(\overset{\circ}{\Omega})$$

by the formula (see Example 8.9)

$$\begin{aligned} &Su(x) \\ &= \int_\Omega s(x,dy)\left[u(y) - \sigma(x,y)\left(u(x) + \sum_{i=1}^n \frac{\partial u}{\partial x_i}(x)(y_i - x_i)\right)\right], \\ &\qquad x \in \overset{\circ}{\Omega},\ \ u \in C_0^2(\Omega). \end{aligned} \tag{8.2}$$

Here $C_0^2(\Omega)$ is the space of functions in $C^2(\Omega)$ with compact support in Ω (see Sect. 5.2.4).

First, we give a complete characterization of linear continuous operators $W : C_0^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ which have positive Borel kernels in the case where Ω is an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$:

Theorem 8.2. *Let Ω be an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$. If W is a linear operator from $C_0^2(\Omega)$ into $B_{\text{loc}}(\overset{\circ}{\Omega})$, then the following two assertions (p_0) and (w) are equivalent:*

(p_0) $W : C_0^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ is continuous and satisfies the condition

$$x \in \overset{\circ}{\Omega},\ \ u \in C_0^2(\Omega),\ \ u \geq 0 \ \text{ in } \Omega \text{ and } \ x \notin \operatorname{supp} u \Longrightarrow Wu(x) \geq 0. \tag{8.3}$$

(w) There exist a second-order differential operator $P : C^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ *and* positive Borel kernels $s(x,dy)$*, having properties (NS.1) and (NS.2), such that the operator W is written in the form*

$$\begin{aligned} Wu(x) &= Pu(x) + Su(x) \\ &= \left(\sum_{i,j=1}^n a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^n b^i(x)\frac{\partial u}{\partial x_i}(x) + c(x)u(x)\right) \\ &\quad + \int_\Omega s(x,dy)\left[u(y) - \sigma(x,y)\left(u(x) + \sum_{i=1}^n \frac{\partial u}{\partial x_i}(x)(y_i - x_i)\right)\right], \\ &\qquad x \in \overset{\circ}{\Omega},\ \ u \in C_0^2(\Omega). \end{aligned} \tag{8.4}$$

Here the coefficients $a^{ij}(x)$, $b^i(x)$ and $c(x)$ belong to the space $B_{\text{loc}}(\overset{\circ}{\Omega})$.

Proof. $(w) \implies (p_0)$: For $u \in C_0^2(\Omega)$, we make use of Taylor's formula

$$u(y) - u(x) - \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)\,(y_i - x_i) = \sum_{i,j=1}^{n} (y_i - x_i)\left(y_j - x_j\right) R_{ij}u(x,y),$$

where

$$R_{ij}u(x,y) = \int_0^1 \frac{\partial^2 u}{\partial x_i \partial x_j}(x + t(y-x))(1-t)\,dt, \quad x, y \in \Omega.$$

Here we observe that $R_{ij}u \in C(\Omega \times \Omega)$ and further that

$$\left\| R_{ij}u \right\|_{C(\Omega\times\Omega)} \leq \frac{1}{2} \|u\|_{C^2(\Omega)}.$$

Therefore, we find that assertion (w) implies assertion (p_0).

$(p_0) \implies (w)$: Conversely, we assume that $W : C_0^2(\Omega) \to B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ is a linear operator which satisfies condition (p_0). Then it follows from assertion (8.3) that there exists a positive Borel kernel $s(x, dy)$ from $\overset{\circ}{\Omega}$ into Ω which satisfies condition

$$s(x, \{x\}) = 0 \quad \text{for every } x \in \overset{\circ}{\Omega} \tag{NS.1}$$

and the condition

$$\int_{\Omega} s(x,dy)u(y) = Wu(x) \quad \text{if } u \in C_0^2(\Omega) \text{ and } x \in \overset{\circ}{\Omega} \setminus \operatorname{supp} u. \tag{8.5}$$

We shall show that this kernel $s(x, dy)$ satisfies condition (NS.2). To do this, it suffices to prove that the function

$$x \longmapsto \int_{\Omega} s(x,dy)|y-x|^2 f(y)$$

is locally bounded on $\overset{\circ}{\Omega}$ for every non-negative function $f \in C_0^\infty(\Omega)$. The proof is divided into three steps.

Step 1: Let $\Phi(x, y)$ be an arbitrary non-negative, continuous function on $\Omega \times \Omega$ which satisfies the following three conditions (i)–(iii):

(i) The functions $\Phi_x(y) = \Phi(x, y)$ is a function in $C^2(\Omega)$ for each $x \in \Omega$.
(ii) The functions

$$(x, y) \longmapsto \frac{\partial \Phi_x}{\partial y_i}(y) \quad (1 \leq i \leq n)$$

and

$$(x, y) \longmapsto \frac{\partial^2 \Phi_x}{\partial y_i \partial y_j}(y) \quad (1 \leq i, j \leq n)$$

are continuous on $\Omega \times \Omega$.

(iii) We have, for each $x \in \Omega$,

$$\Phi_x(x) = \frac{\partial \Phi_x}{\partial y_i}(x) = \frac{\partial^2 \Phi_x}{\partial y_i \partial y_j}(x) = 0 \quad (1 \leq i, j \leq n).$$

We prove that the function

$$x \longmapsto \int_\Omega s(x, dy)\Phi(x, y) f(y)$$

is locally bounded on $\overset{\circ}{\Omega}$ for every non-negative function $f \in C_0^\infty(\Omega)$. Since we have, for every non-negative function $f \in C_0^\infty(\Omega)$,

$$\Phi_x f \in C_0^2(\Omega)$$

and so

$$W(\Phi_x f) \in B_{\text{loc}}(\overset{\circ}{\Omega}),$$

we have only to show that

$$\int_\Omega s(x, dy)\Phi(x, y) f(y) = W(\Phi_x f)(x) \quad \text{for all } x \in \overset{\circ}{\Omega}. \tag{8.6}$$

In order to prove (8.6), for each $x \in \overset{\circ}{\Omega}$ we consider an increasing sequence $\{\theta_j\}$ of non-negative functions in $C^\infty(\Omega)$ which satisfies the following three conditions (a)–(c):

(a) $\theta_j(y) = 0$ in a neighborhood of x for all $j \in \mathbf{N}$.

(b) We have, for all $y \neq x$,

$$\sup_{j \geq 1} \theta(y) = 1.$$

(c) We have

$$\lim_{j \to \infty} \left\| \theta_j \Phi_x f - \Phi_x f \right\|_{C^2(\Omega)} = 0.$$

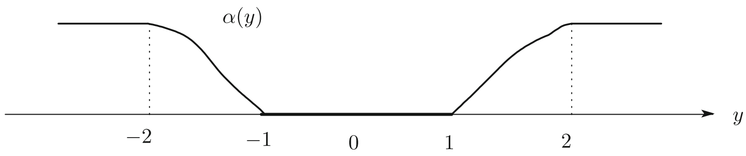


Fig. 8.5 The function $\alpha(y)$

For example, by choosing a smooth function $\alpha(y)$ on $\mathbf{R}^n$ such that (see Fig. 8.5)

$$0 \leq \alpha(y) \leq 1 \quad \text{on } \mathbf{R}^n,$$

and that

$$\alpha(y) = \begin{cases} 0 & \text{if } |y| \leq 1, \\ 1 & \text{if } |y| \geq 2, \end{cases}$$

we may take

$$\theta_j(y) = \alpha\left(j(y-x)\right) = \begin{cases} 0 & \text{if } |y-x| \leq \frac{1}{j}, \\ 1 & \text{if } |y-x| \geq \frac{2}{j}. \end{cases}$$

It should be emphasized that we have, for all $j \in \mathbf{N}$,

$$|\Phi_x(y)| \leq \frac{C_1}{j^3} \text{ and } |\nabla\alpha(y)| \leq C_2 j \quad \text{if } \nabla\alpha(y) \neq 0,$$
$$|\Phi_x(y)| \leq \frac{C_1}{j^3} \text{ and } |\nabla^2\alpha(y)| \leq C_3 j^2 \quad \text{if } \nabla^2\alpha(y) \neq 0,$$

where C_1, C_2 and C_3 are positive constants independent of $j \in \mathbf{N}$.

Then, by virtue of assertion (8.5), it follows that

$$\int_\Omega s(x, dy)\theta_j(y)\Phi_x(y)f(y) = W\left(\theta_j \Phi_x f\right)(x) \quad \text{for every } n \in \mathbf{N}. \tag{8.7}$$

We note that each term of the left-hand side of formula (8.7) is monotone increasing, while the operator W on the right-hand side of formula (8.7) is continuous. Therefore, the desired formula (8.6) follows by letting $j \to \infty$ in formula (8.7).

Step 2: We show that the Borel kernel $s(x, dy)$ satisfies condition (NS.2). The proof is divided into two steps.

Step 2-1: First, we show that if $g(x)$ is a non-negative, continuous function on $\mathbf{R}^n$ such that $g(0) = 0$, then the function

$$x \longmapsto \int_{\Omega} s(x, dy)|y - x|^2 g(y - x) f(y)$$

is locally bounded on $\overset{\circ}{\Omega}$.

Step 2-1a: In order to prove the boundedness of $s(x, dy)$, we need the following lemma:

Lemma 8.3. *Let $B = B_r(0) = \{z \in \mathbf{R}^n : |z| < r\}$ be the open ball about 0 radius $r > 0$ and let $g(x)$ be a bounded, non-negative function on B such that*

$$\lim_{x \to 0} g(x) = 0. \tag{8.8}$$

Then there exists a bounded, non-negative function $k \in C^2(B)$ such that

$$|z|^2 g(z) \leq k(z) \quad \textit{for all } z \in B. \tag{8.9}$$

$$k(0) = \frac{\partial k}{\partial x_i}(0) = \frac{\partial^2 k}{\partial x_i \partial x_j}(0) = 0 \quad \textit{for all } 1 \leq i, j \leq n. \tag{8.10}$$

Proof. The proof is divided into two steps,

(1) First, we consider the case where $n = 1$. To do this, it suffices to construct a function $k(x)$ on the interval $[0, r)$.

We consider the inferior envelope $\tilde{g}(x)$ of the family of linear functions, $ax + b$, with $a, b \in \mathbf{R}$, which majorize the function $g(x)$ on the interval $[0, r)$. Since the function $\tilde{g}(x)$ is concave, it follows that $\tilde{g}(x)$ is continuous on the open interval $(0, r)$. Moreover, it is easy to verify the following three assertions (a)–(c):

(a) $\tilde{g}(x)$ is non-negative and bounded.
(b) $\tilde{g}(0) = 0$ and $\lim_{x \downarrow 0} \tilde{g}(x) = 0$.
(c) $g(x) \leq \tilde{g}(x)$ for all $x \in [0, r)$.

Now, by letting

$$k(x) = 6 \int_0^x \left(\int_0^t \tilde{g}(s)\, ds \right) dt, \quad 0 \leq x < r, \tag{8.11}$$

we shall show the inequality

$$x^2 g(x) \leq k(x) \quad \text{for all } x \in [0, r); \tag{8.12}$$

which proves the desired inequality (8.9) in the case where $n = 1$.

If $\varphi(x)$ is a function in $C[0, r)$, we define its primitive $P\varphi(x)$ by the formula

$$P\varphi(x) = \int_0^x \varphi(t)\, dt, \quad 0 \leq x < r.$$

Since the function $\tilde{g}(x)$ is concave and $\tilde{g}(0) = 0$, we have the inequality

$$\tilde{g}(t) \geq \frac{t}{x}\,\tilde{g}(x) \quad \text{for all } 0 < x < r \text{ and } 0 \leq t \leq x. \tag{8.13}$$

Therefore, by applying the operation P twice we obtain from inequality (8.13) that

$$P\tilde{g}(x) = \int_0^x \tilde{g}(t)\,dt \geq \int_0^x \frac{t}{x}\,\tilde{g}(x)\,dt = \frac{1}{2}\,x\,\tilde{g}(x), \quad 0 \leq x < r,$$

and further from condition (c) that

$$\begin{aligned} P^2\tilde{g}(x) = P(P\tilde{g})(x) &\geq \frac{1}{2}\int_0^x t\,\tilde{g}(t)\,dt \geq \frac{1}{2}\int_0^x \frac{t^2}{x}\,\tilde{g}(x)\,dt = \frac{1}{6}\,x^2\,\tilde{g}(x) \\ &\geq \frac{1}{6}\,x^2\,g(x), \quad 0 < x < r. \end{aligned}$$

In view of (8.11), this proves (8.12).

(2) We can prove the general case where $n \geq 2$ as follows: For $0 \leq \rho < r$, we let

$$\gamma(\rho) = \sup_{|z|=\rho} g(z), \quad 0 < \rho < r.$$

We see that

$$g(z) \leq \gamma(|z|) \quad \text{for all } z \in B.$$

Moreover, it follows from condition (8.8) that $\gamma(\rho)$ is bounded on the interval $[0, \rho)$ and that

$$\lim_{\rho \downarrow 0} \gamma(\rho) = 0.$$

Therefore, just as in Step (1) we can construct a function (see formula (8.11))

$$\chi(\rho) = 6\,P^2\tilde{\gamma}(\rho), \quad 0 < \rho < r,$$

associated with the function $\gamma(\rho)$. Then it is easy to verify that the function $k(z)$, defined by the formula

$$k(z) = \chi(|z|), \quad z \in B,$$

satisfies the desired conditions (8.9) and (8.10).

The proof of Lemma 8.3 is complete.

Step 2-1b: If $g(x)$ is a non-negative, continuous function on $\mathbf{R}^n$ such that $g(0) = 0$, then it follows from an application of Lemma 8.3 that there exists a non-negative function $k \in C^2(\mathbf{R}^n)$ such that

$$|z|^2 g(z) \leq k(z) \quad \text{for all } z \in \mathbf{R}^n,$$
$$k(0) = \frac{\partial k}{\partial x_i}(0) = \frac{\partial^2 k}{\partial x_i \partial x_j}(0) = 0 \quad \text{for all } 1 \leq i, j \leq n.$$

However, by applying the result of Step 1 to the function $\Phi(x, y) := k(x - y)$ we obtain that the function

$$x \longmapsto \int_\Omega s(x, dy) k(y - x) f(y)$$

is locally bounded on $\overset{\circ}{\Omega}$.
Since we have the inequality

$$|y - x|^2 g(y - x) \leq k(y - x) \quad \text{for all } x, y \in \Omega,$$

it follows that the function

$$x \longmapsto \int_\Omega s(x, dy) |y - x|^2 g(y - x) f(y)$$

is also locally bounded on $\overset{\circ}{\Omega}$.

Step 2-2: We assume, to the contrary, that there exist a compact subset K of $\overset{\circ}{\Omega}$ and a non-negative function $f \in C_0(\Omega)$ such that

$$\sup_{x \in K} \int_\Omega s(x, dy) |y - x|^2 f(y) = +\infty.$$

Then we can find a sequence $\{x_j\}$ of points in K and a sequence $\{g_j\}$ of non-negative, continuous functions on $\mathbf{R}^n$ such that we have, for each $j \in \mathbf{N}$,

$$g_j(0) = 0,$$
$$\|g_j\|_\infty = 1,$$
$$\int_\Omega s(x_j, dy) |y - x_j|^2 g_j(y - x_j) f(y) \geq j!.$$

If we let

$$g(z) = \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} g_\ell(z), \quad z \in \mathbf{R}^n,$$

it follows that $g(z)$ is a non-negative, continuous function on $\mathbf{R}^n$ such that

$$g(0) = 0,$$

and that

$$\begin{aligned}
&\int_\Omega s(x_j, dy)|y - x_j|^2 g(y - x_j) f(y) \\
&= \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} \int_\Omega s(x_j, dy)|y - x_j|^2 g_\ell(y - x_j) f(y) \\
&\geq \frac{1}{2^j} \int_\Omega s(x_j, dy)|y - x_j|^2 g_j(y - x_j) f(y) \geq \frac{j!}{2^j} \quad \text{for all } j \in \mathbf{N},
\end{aligned}$$

so that

$$\sup_{j \geq 1} \int_\Omega s(x_j, dy)|y - x_j|^2 g(y - x_j) f(y) = +\infty.$$

This contradicts the assertion proved in Step 2-1.

Summing up, we have proved that the Borel kernel $s(x, dy)$ satisfies condition (NS.2).

Step 3: By starting at the kernel $s(x, dy)$ and a unity local function $\sigma(x, y)$ we can define a continuous linear operator

$$S : C_0^2(\Omega) \longrightarrow B_{\mathrm{loc}}(\overset{\circ}{\Omega})$$

by formula (8.2)

$$\begin{aligned}
&Su(x) \\
&= \int_\Omega s(x, dy) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)(y_i - x_i) \right) \right], \\
&\qquad\qquad x \in \overset{\circ}{\Omega}, \quad u \in C_0^2(\Omega).
\end{aligned}$$

We denote by $\tilde{\Omega}$ an open subset of $\mathbf{R}^n$ such that

$$\begin{cases} \tilde{\Omega} \cap \overline{\mathbf{R}^n_+} = \Omega & \text{if } \Omega \text{ is an open subset of } \mathbf{R}^n_+, \\ \tilde{\Omega} = \Omega & \text{if } \Omega \text{ is an open subset of } \mathbf{R}^n. \end{cases}$$

Then it follows from condition (8.5) that the mapping

$$C_0^\infty(\tilde{\Omega}) \ni u \longmapsto W\left(u|_{\tilde{\Omega}}\right)(x) - S\left(u|_{\tilde{\Omega}}\right)(x), \quad x \in \overset{\circ}{\Omega},$$

defines a distribution of order 2 on $\tilde{\Omega}$ whose support is reduced to the point x. Since $P = W - S$ maps $C_0^2(\Omega)$ into $B_{\mathrm{loc}}(\overset{\circ}{\Omega})$, we obtain that P is a second-order differential operator of the form

$$Pu(x) = \sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} b^i(x)\frac{\partial u}{\partial x_i}(x) + c(x)u(x), \quad x \in \overset{\circ}{\Omega},$$

where the coefficients $a^{ij}(x)$, $b^i(x)$ and $c(x)$ belong to $B_{\mathrm{loc}}(\overset{\circ}{\Omega})$.

Now the proof of Theorem 8.2 is complete.

Secondly, First, we give a useful characterization of linear continuous operators $W : C_0^2(\Omega) \to B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ which have positive Borel kernels in the case where Ω is an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$:

Theorem 8.4. *Let Ω be an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$. Let $\mathcal{V}$ be a linear subspace of $C_0^2(\Omega)$ which contains $C_0^\infty(\Omega)$, that is, $C_0^\infty(\Omega) \subset \mathcal{V} \subset C_0^2(\Omega)$. Assume that W is a linear operator from $\mathcal{V}$ into $B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ and satisfies the condition*

$$x \in \overset{\circ}{\Omega},\ u \in \mathcal{V},\ u \geq 0 \ \text{ in } \Omega \text{ and } \ u(x) = 0 \Longrightarrow Wu(x) \geq 0. \tag{8.14}$$

Then the operator W can be extended uniquely to a continuous *linear operator $W : C_0^2(\Omega) \to B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ which still satisfies condition (8.14) for all $u \in C_0^2(\Omega)$:*

$$x \in \overset{\circ}{\Omega},\ u \in C_0^2(\Omega),\ u \geq 0 \ \text{ in } \Omega \text{ and } \ u(x) = 0 \Longrightarrow Wu(x) \geq 0. \tag{p_1}$$

Proof. The proof is divided into three steps.

Step 1: First, we consider the case where $\mathcal{V} = C_0^2(\Omega)$. In order to prove the continuity of W, by the closed graph theorem (Theorem 3.51) it suffices to show that the operator

$$W : C_0^2(\Omega) \longrightarrow B_{\mathrm{loc}}(\overset{\circ}{\Omega})$$

is closed.

Step 1-1: The next lemma is an essential step in the proof:

Lemma 8.5. *For any point $x \in \overset{\circ}{\Omega}$ and any compact subset K of Ω, there exists a positive constant $C(x, K)$ such that*

$$|Wu(x)| \leq C(x, K)\, \|u\|_{C^2(\Omega)} \quad \textit{for all } u \in C_0^2(K). \tag{8.15}$$

Here

$$C_0^2(K) = \left\{u \in C_0^2(\Omega) : \operatorname{supp} u \subset K\right\}.$$

Proof. For a point $x \in \overset{\circ}{\Omega}$ and a compact set $K \subset \Omega$, we choose a local unity function $\sigma(x, y)$ such that

$$K \subset \operatorname{supp} \sigma(x, \cdot) \quad \text{for every } x \in \overset{\circ}{\Omega},$$

and let

$$\begin{aligned} \sigma_x^0(y) &= \sigma(x, y), \\ \sigma_x^i(y) &= \sigma_x^0(y)\,(y_i - x_i) = \sigma(x, y)\,(y_i - x_i)\,, \quad 1 \leq i \leq n. \end{aligned}$$

Moreover, if $u \in C_0^2(K)$, we let

$$\begin{aligned} u_x(y) &= u(y) - u(x)\sigma_x^0(y) - \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)\sigma_x^i(y) \\ &= u(y) - \sigma(x, y)\left(u(x) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)(y_i - x_i)\right), \quad x, y \in \Omega. \end{aligned} \tag{8.16}$$

We observe that

$$u(y) = u(x)\sigma_x^0(y) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)\sigma_x^i(y) + u_x(y), \quad u_x \in C_0^2(K), \tag{8.17}$$

and that

$$u_x(x) = \frac{\partial u_x}{\partial x_i}(x) = 0, \quad 1 \leq i \leq n. \tag{8.18}$$

Therefore, by applying the operator W to both sides of (8.17) we obtain that

$$\begin{aligned} &Wu(x) \\ &= u(x)W\left(\sigma_x^0\right)(x) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x)W\left(\sigma_x^i\right)(x) + W\left(u_x\right)(x), \quad x \in \overset{\circ}{\Omega}. \end{aligned} \tag{8.19}$$

On the other hand, by Taylor's formula it follows from condition (8.18) that

$$\begin{aligned} u_x(y) &= \sum_{i,j=1}^{n} (y_i - x_i)(y_j - x_j)R_{ij}(u_x)(x, y), \\ \left\| R_{ij}(u_x) \right\|_{C(\Omega \times \Omega)} &\leq \frac{1}{2} \left\| u_x \right\|_{C^2(\Omega)}. \end{aligned}$$

Therefore, we can find a non-negative function $\Psi_x \in C_0^\infty(\Omega)$ such that

$$\Psi_x(y) = |y - x|^2 \quad \text{for all } y \in K \subset \operatorname{supp} \sigma_x^0,$$

and further that

$$|u_x(y)| \leq n \, \|u_x\|_{C^2(\Omega)} \, \Psi_x(y) \quad \text{for all } y \in \Omega. \tag{8.20}$$

If we let

$$U_\pm(y) = n \, \|u_x\| \, \Psi_x(y) \pm u_x(y)$$

then we have, by inequality (8.20) and condition (8.18),

$$\begin{aligned} &U_\pm \in C_0^2(\Omega), \\ &U_\pm(y) \geq 0 \quad \text{in } \Omega, \\ &U_\pm(x) = 0. \end{aligned}$$

Hence, by applying condition (p_1) to the function $U_\pm$ we obtain that

$$WU_\pm(x) = n \, \|u_x\|_{C^2(\Omega)} \, W(\Psi_x)(x) \pm W(u_x)(x) \geq 0, \quad x \in \overset{\circ}{\Omega},$$

so that

$$|W(u_x)(x)| \leq n \, \|u_x\|_{C^2(\Omega)} \, W(\Psi_x)(x), \quad x \in \overset{\circ}{\Omega}. \tag{8.21}$$

However, by Taylor's formula it follows from formula (8.16) that there exists a positive constant $C_1(x, K)$ such that

$$\|u_x\|_{C^2(\Omega)} \leq C_1(x, K) \, \|u\|_{C^2(\Omega)} \quad \text{for all } u \in C_0^2(K). \tag{8.22}$$

Therefore, we have, by inequalities (8.21) and (8.22),

$$|W(u_x)(x)| \leq n \, C_1(x, K) \, \|u\|_{C^2(\Omega)} \, W(\Psi_x)(x), \quad x \in \overset{\circ}{\Omega}. \tag{8.23}$$

By combining formula (8.19) and inequality (8.23), we have proved that

$$\begin{aligned} &|W(u)(x)| \\ &\leq |u(x)| \left|W\left(\sigma_x^0\right)(x)\right| + \sum_{i=1}^n \left|\frac{\partial u}{\partial x_i}(x)\right| \left|W\left(\sigma_x^i\right)(x)\right| + |W(u_x)(x)| \end{aligned}$$

$$\leq \|u\|_{C^2(\Omega)} \left|W\left(\sigma_x^0\right)(x)\right| + \|u\|_{C^2(\Omega)} \sum_{i=1}^{n} \left|W\left(\sigma_x^i\right)(x)\right|$$
$$+n\, C_1(x,K)\, \|u\|_{C^2(\Omega)}\, W\left(\Psi_x\right)(x)$$
$$= \left(\left|W\left(\sigma_x^0\right)(x)\right| + \sum_{i=1}^{n} \left|W\left(\sigma_x^i\right)(x)\right| + n\, C_1(x,K) W\left(\Psi_x\right)(x) \right) \|u\|_{C^2(\Omega)}.$$

This proves the desired inequality (8.15).

The proof of Lemma 8.5 is complete.

Step 1-2: Now we show the following assertion:

$$u_j \longrightarrow u \text{ in } C_0^2(\Omega), \quad Wu_j \longrightarrow v \text{ in } B_{\text{loc}}(\overset{\circ}{\Omega})$$
$$\Longrightarrow v = Wu.$$

Since the topology of $C_0^2(\Omega)$ is the inductive limit topology of the spaces $C_0^2(K)$, there exists a compact subset $K \subset \Omega$ such that

$$u_j \longrightarrow u \quad \text{in } C_0^2(K).$$

Therefore, by applying inequality (8.15) to the functions $u_j - u$, we obtain that, as $j \to \infty$,

$$Wu_j(x) \longrightarrow Wu(x) \quad \text{for each } x \in \overset{\circ}{\Omega}.$$

This proves that $v = Wu$.

Summing up, we have proved Theorem 8.4 in the case where $\mathcal{V} = C_0^2(\Omega)$.

Step 2: We consider the general case where $\mathcal{V}$ is only a subspace of $C_0^2(\Omega)$ which contains $C_0^\infty(\Omega)$. In this case we cannot use the closed graph theorem, as in Step 1.

Step 2-1: Our proof is based on the following lemma:

Lemma 8.6. *For any point $x_0 \in \overset{\circ}{\Omega}$ and any compact set $K \subset \Omega$ such that $x_0 \in \overset{\circ}{K}$, there exist a compact set $L \subset \Omega$, a neighborhood V of x_0 ($V \subset \overset{\circ}{\Omega}$) and smooth functions $\eta_x^0(y)$, $\eta_x^1(y)$, ..., $\eta_x^n(y)$, $\varphi_x(y)$ in $C_0^\infty(\Omega)$ with support in L such that, for each $x \in V$,*

$$\eta_x^0(x) = 1,\ \eta_x^i(x) = 0;\ \frac{\partial \eta_x^0}{\partial x_k}(x) = 0,\ \frac{\partial \eta_x^i}{\partial x_k}(x) = \delta_k^i\ (1 \leq i,k \leq n). \tag{8.24}$$

$$\varphi_x(x) = 0 \text{ and } |y-x|^2 \leq \varphi_x(y) \text{ for all } y \in K. \tag{8.25}$$

Moreover, we have the estimates

$$\sup_{x \in V} \left\| \eta_x^i \right\|_{C^2(\Omega)} < \infty \quad (0 \leq i \leq n), \tag{8.26a}$$

$$\sup_{x \in V} \left| W\left(\eta_x^i\right)(x) \right| < \infty \quad (0 \leq i \leq n), \tag{8.26b}$$

$$\sup_{x \in V} \left| W\left(\varphi_x\right)(x) \right| < \infty. \tag{8.26c}$$

Proof. (1) First, we construct functions $\eta_x^0(y)$, $\eta_x^1(y)$, ..., $\eta_x^n(y)$ in the space $C_0^\infty(\Omega)$ which satisfy conditions (8.24).

To do so, we consider the following $(n+1) \times (n+1)$ matrix:

$$M(x; x_0) = \begin{pmatrix} \sigma_{x_0}^0(x) & \partial_1 \sigma_{x_0}^0(x) & \cdots & \partial_k \sigma_{x_0}^0(x) & \cdots & \partial_n \sigma_{x_0}^0(x) \\ \sigma_{x_0}^1(x) & \partial_1 \sigma_{x_0}^1(x) & \cdots & \partial_k \sigma_{x_0}^1(x) & \cdots & \partial_n \sigma_{x_0}^1(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \sigma_{x_0}^j(x) & \partial_1 \sigma_{x_0}^j(x) & \cdots & \partial_k \sigma_{x_0}^j(x) & \cdots & \partial_n \sigma_{x_0}^j(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \sigma_{x_0}^n(x) & \partial_1 \sigma_{x_0}^n(x) & \cdots & \partial_k \sigma_{x_0}^n(x) & \cdots & \partial_n \sigma_{x_0}^n(x) \end{pmatrix}, \qquad x \in \Omega.$$

Here

$$\partial_k \sigma_{x_0}^j(x) = \frac{\partial}{\partial y_k} \left(\sigma_{x_0}^j(y) \right) \bigg|_{y=x} \qquad (1 \leq i, k \leq n).$$

We note that the matrix $M(x; x_0)$ is the identity matrix at $x = x_0$. Hence we can find a compact neighborhood V of x_0 ($V \subset \overset{\circ}{\Omega}$) such that $M(x_0; x)$ is invertible on V. We express the inverse matrix $N(x; x_0)$ of $M(x; x_0)$ in the form

$$N(x; x_0) = \begin{pmatrix} \tau_0^0(x) & \tau_1^0(x) & \cdots & \tau_k^0(x) & \cdots & \tau_n^0(x) \\ \tau_0^1(x) & \tau_1^1(x) & \cdots & \tau_k^1(x) & \cdots & \tau_n^1(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \tau_0^j(x) & \tau_1^j(x) & \cdots & \tau_k^j(x) & \cdots & \tau_n^j(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \tau_0^n(x) & \tau_1^n(x) & \cdots & \tau_k^n(x) & \cdots & \tau_n^n(x) \end{pmatrix}, \qquad x \in V.$$

Then we can define functions $\eta_x^0(y)$, $\eta_x^1(y)$, ..., $\eta_x^n(y)$ by the formula

$$\begin{pmatrix} \eta_x^0(y) \\ \eta_x^1(y) \\ \cdot \\ \cdot \\ \eta_x^j(y) \\ \cdot \\ \cdot \\ \eta_x^n(y) \end{pmatrix} = \begin{pmatrix} \tau_0^0(x) & \tau_1^0(x) & \cdots & \tau_k^0(x) & \cdots & \tau_n^0(x) \\ \tau_0^1(x) & \tau_1^1(x) & \cdots & \tau_k^1(x) & \cdots & \tau_n^1(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \tau_0^j(x) & \tau_1^j(x) & \cdots & \tau_k^j(x) & \cdots & \tau_n^j(x) \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \tau_0^n(x) & \tau_1^n(x) & \cdots & \tau_k^n(x) & \cdots & \tau_n^n(x) \end{pmatrix} \begin{pmatrix} \sigma_x^0(y) \\ \sigma_x^1(y) \\ \cdot \\ \cdot \\ \sigma_x^k(y) \\ \cdot \\ \cdot \\ \sigma_x^n(y) \end{pmatrix}, \quad x \in V, \ y \in \Omega.$$

It is easy to see that the functions $\eta_x^0(y), \eta_x^1(y), \ldots, \eta_x^n(y)$ satisfy conditions (8.24), for each $x \in V$.

(2) Secondly, we construct functions $\varphi_x(y)$ in the space $C_0^\infty(\Omega)$ by using the functions $\eta_x^0(y), \eta_x^1(y), \ldots, \eta_x^n(y)$. We choose a non-negative smooth function $\theta(y)$ in $C_0^\infty(\Omega)$ such that

$$\theta(y) = 1 \quad \text{on } K \cup V.$$

For each $x \in V$, we let

$$\psi_x(y) = |y-x|^2\, \theta(y) \quad \text{for all } y \in \Omega,$$

and

$$\begin{aligned} &\tilde{\varphi}_x(y) \\ &= \psi_{x_0}(y) - \psi_{x_0}(x)\eta_x^0(y) - \sum_{i=1}^{n} \frac{\partial}{\partial y_i} (\psi_{x_0}(y))\bigg|_{y=x} \eta_x^i(y) \quad \text{for all } y \in \Omega. \end{aligned}$$

Then, by using Taylor's formula and conditions (8.24) we can find a positive constant C, independent of $x \in V$, such that

$$\tilde{\varphi}_x(y) \leq C\, |y-x|^2 \quad \text{for all } y \in \Omega.$$

Therefore, it is easy to see that the functions

$$\varphi_x(y) = \frac{1}{C}\, \tilde{\varphi}_x(y) \quad \text{for all } y \in \Omega,$$

satisfy conditions (8.25), for each $x \in V$.

(3) Finally, we can easily verify that the functions $\eta_x^0(y), \eta_x^1(y), \ldots, \eta_x^n(y)$ and $\varphi_x(y)$ satisfy the desired estimates (8.26a), (8.26b) and (8.26c).

The proof of Lemma 8.6 is complete.

Step 2-2: In the general case, we substitute the functions $\sigma_x^0(y), \sigma_x^1(y), \ldots, \sigma_x^n(y)$ and $\Psi_x(y)$ by the functions $\eta_x^0(y), \eta_x^1(y), \ldots, \eta_x^n(y)$ and $\varphi_x(y)$, respectively.

If $v \in \mathcal{V}$ with $\operatorname{supp} v \subset K$ for some compact subset $K \subset \Omega$, we let

$$v_x(y) = v(y) - v(x)\eta_x^0(y) - \sum_{i=1}^{n} \frac{\partial v}{\partial x_i}(x)\eta_x^i(y) \quad x, y \in \Omega. \tag{8.27}$$

We note that

$$v(y) = v(x)\eta_x^0(y) + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i}(x)\eta_x^i(y) + v_x(y), \quad v_x \in \mathcal{V}, \tag{8.28}$$

and that

$$v_x(x) = \frac{\partial v_x}{\partial x_i}(x) = 0, \quad 1 \le i \le n. \tag{8.29}$$

Therefore, by applying the operator W to both sides of (8.28) we obtain that

$$\begin{aligned} & Wv(x) \qquad (8.30) \\ &= v(x)W\left(\eta_x^0\right)(x) + \sum_{i=1}^{n} \frac{\partial v}{\partial x_i}(x)W\left(\eta_x^i\right)(x) + W\left(v_x\right)(x), \quad x \in \overset{\circ}{\Omega}. \end{aligned}$$

On the other hand, by Taylor's formula it follows from condition (8.29) that

$$v_x(y) = \sum_{i,j=1}^{n} (y_i - x_i)(y_j - x_j)R_{ij}(v_x)(x, y),$$

$$\left\| R_{ij}(v_x) \right\|_{C(\Omega \times \Omega)} \le \frac{1}{2} \left\| v_x \right\|_{C^2(\Omega)}.$$

However, by virtue of assertion (8.25) we obtain that

$$|v_x(y)| \le n \left\| v_x \right\|_{C^2(\Omega)} \varphi_x(y) \quad \text{for all } y \in \Omega. \tag{8.31}$$

If we let

$$V_\pm(y) = n \left\| v_x \right\| \varphi_x(y) \pm v_x(y)$$

then we have, by inequality (8.31) and condition (8.29),

$$\begin{aligned} & V_\pm \in \mathcal{V}, \\ & V_\pm(y) \ge 0 \quad \text{in } \Omega, \\ & V_\pm(x) = 0. \end{aligned}$$

Hence, by applying condition (p_1) to the function $V_{\pm}$ we obtain that

$$W V_{\pm}(x) = n \, \|v_x\|_{C^2(\Omega)} \, W(\varphi_x)(x) \pm W(v_x)(x) \geq 0, \quad x \in \overset{\circ}{\Omega},$$

so that

$$|W(v_x)(x)| \leq n \, \|v_x\|_{C^2(\Omega)} \, W(\varphi_x)(x), \quad x \in \overset{\circ}{\Omega}. \tag{8.32}$$

However, by Taylor's formula it follows from (8.27) that there exists a positive constant $C_1(x, K)$ such that

$$\|v_x\|_{C^2(\Omega)} \leq C_1(x, K) \, \|v\|_{C^2(\Omega)} \quad \text{for all } v \in \mathcal{V} \text{ with } \operatorname{supp} v \subset K. \tag{8.33}$$

Therefore, we have, by inequalities (8.32) and (8.33),

$$|W(v_x)(x)| \leq n \, C_1(x, K) \, \|v\|_{C^2(\Omega)} \, W(\varphi_x)(x), \quad x \in \overset{\circ}{\Omega}. \tag{8.34}$$

By combining formula (8.30) and inequality (8.34), we obtain that

$$\begin{aligned}
&|W(v)(x)| \\
&\leq |v(x)| \left| W\left(\eta_x^0\right)(x) \right| + \sum_{i=1}^{n} \left| \frac{\partial v}{\partial x_i}(x) \right| \left| W\left(\eta_x^i\right)(x) \right| + |W(v_x)(x)| \\
&\leq \|v\|_{C^2(\Omega)} \left| W\left(\eta_x^0\right)(x) \right| + \|v\|_{C^2(\Omega)} \sum_{i=1}^{n} \left| W\left(\eta_x^i\right)(x) \right| \\
&\quad + n \, C_1(x, K) \, \|v\|_{C^2(\Omega)} \, W(\varphi_x)(x) \\
&= \left(\sum_{i=0}^{n} \left| W\left(\eta_x^i\right)(x) \right| + n \, C_1(x, K) W(\varphi_x)(x) \right) \|v\|_{C^2(\Omega)} .
\end{aligned}$$

In view of estimates (8.26b) and (8.26c), we have proved that the operator $W : \mathcal{V} \to B_{\text{loc}}(\overset{\circ}{\Omega})$ can be extended uniquely to a continuous linear operator

$$W : C_0^2(\Omega) \longrightarrow B_{\text{loc}}(\overset{\circ}{\Omega}).$$

Step 3: Finally, it remains to show that the extended operator W still satisfies condition (8.14) for all $u \in C_0^2(\Omega)$:

$$x \in \overset{\circ}{\Omega}, \; u \in C_0^2(\Omega), \; u \geq 0 \quad \text{in } \Omega \text{ and } u(x) = 0 \Longrightarrow Wu(x) \geq 0. \tag{p_1}$$

Let $\{u_k\}$ be a sequence of smooth functions in the space $C_0^\infty(\Omega)$ such that

$$u_k \geq 0 \text{ in } \Omega,$$
$$u_k(x) = 0 \text{ for all } k \in \mathbf{N},$$
$$u_k \longrightarrow u \text{ in } C_0^2(\Omega).$$

For example, by choosing a smooth function $\alpha(y)$ on $\mathbf{R}^n$ such that

$$0 \leq \alpha(y) \leq 1 \quad \text{on } \mathbf{R}^n,$$

and that

$$\alpha(y) = \begin{cases} 0 & \text{if } |y| \leq 1, \\ 1 & \text{if } |y| \geq 2, \end{cases}$$

we may take the functions $\{u_k\}$ of the form

$$u_k(y) = \frac{1}{2}\left(1 - \alpha\left(k(y-x)\right)\right) \sum_{i,j=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_j}(x)\,(y_i - x_i)\left(y_j - x_j\right)$$
$$+\alpha\left(k(y-x)\right) J_{\varepsilon_k} u(y), \quad y \in \Omega.$$

Here the $J_{\varepsilon_k} u$ are Friedrichs' mollifiers with $\varepsilon_k \downarrow 0$ (see Sect. 5.2.6).

Then, by applying condition (8.14) to the functions $\{u_k\}$ we obtain from the continuity of W that

$$Wu(x) = \lim_{k\to\infty} Wu_k(x) \geq 0.$$

This proves the desired condition (p_1).

Now the proof of Theorem 8.4 is complete.

By combining Theorems 8.2 and 8.4, we have the following simple characterization of W:

Corollary 8.7. *Let Ω be an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$. If a linear operator $W : C_0^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ satisfies condition (p_1), then it follows that $W : C_0^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ is continuous and can be written in the form* (8.4)*, where a^{ij}, b^i, $c \in B_{\text{loc}}(\overset{\circ}{\Omega})$ and the positive Borel kernels $s(x, dy)$ enjoy properties (NS.1) and (NS.2).*

Proof. If $W : C_0^2(\Omega) \to B_{\text{loc}}(\overset{\circ}{\Omega})$ is a linear operator, then it follows from an application of Theorem 8.4 that condition (p_1) implies the continuity of W

and condition (p_0). Therefore, we obtain from Theorem 8.2 that condition (w) is satisfied.

We remark that if the integro-differential operator $W = P + S$ enjoys property (w), then the kernel $s(x, dy)$ and the principal part $\left(a^{ij}(x)\right)_{1 \le i,j \le n}$ of P are uniquely determined by W. Indeed, it suffices to note the following formulas:

$$\int_\Omega s(x, dy)u(y) = Wu(x) \text{ if } u \in C_0^2(\Omega) \text{ and } x \in \overset{\circ}{\Omega} \setminus \operatorname{supp} u, \tag{8.5}$$

$$2u(x) \sum_{i,j=1}^n a^{ij}(x)\xi_i\xi_j = W\left(\Phi_x^\xi u\right)(x) - \int_\Omega s(x, dy)\Phi_x^\xi(y)u(y) \tag{8.35}$$

$$\text{if } x \in \overset{\circ}{\Omega}, u \in C_0^2(\Omega) \text{ and } \xi = (\xi_i) \in \mathbf{R}^n,$$

where

$$\Phi_x^\xi(y) = \left(\sum_{i=1}^n \xi_i\,(y_i - x_i)\right)^2, \quad x, y \in \Omega \text{ and } \xi = (\xi_i) \in \mathbf{R}^n.$$

Finally, we characterize linear operators $W : C_0^2(\Omega) \to B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ which satisfy the positive maximum principle (PM):

Theorem 8.8. *Let Ω be an open subset of $\mathbf{R}^n$ or of $\overline{\mathbf{R}^n_+}$. If W is a linear operator from $C_0^2(\Omega)$ into $B_{\mathrm{loc}}(\overset{\circ}{\Omega})$ of the form (8.4), then we have the following two assertions (i) and (ii):*

(i) The operator W satisfies condition (8.14) if and only if the principal symbol $-\sum_{i,j=1}^n a^{ij}(x)\xi_i\xi_j$ of P is non-positive on $\overset{\circ}{\Omega} \times \mathbf{R}^n$.
(ii) The operator W satisfies the positive maximum principle *(PM)*

$$x_0 \in \overset{\circ}{\Omega},\ v \in C_0^2(\Omega) \text{ and } v(x_0) = \sup_\Omega v \ge 0 \Longrightarrow Wv(x_0) \le 0 \tag{PM}$$

if and only if the principal symbol of P is non-positive on $\overset{\circ}{\Omega} \times \mathbf{R}^n$ and the following conditions (8.36) hold true:

$$P1(x) = c(x) \le 0 \text{ for all } x \in \overset{\circ}{\Omega}, \tag{8.36a}$$

$$W1(x) = c(x) + \int_\Omega s(x, dy)[1 - \sigma(x, y)] \le 0 \text{ for all } x \in \overset{\circ}{\Omega}. \tag{8.36b}$$

In particular, the positive Borel kernels $s(x, dy)$ enjoy the following property (NS.3):

(NS.3) For any open subset Ω' of $\overset{\circ}{\Omega}$, the function

$$\Omega' \ni x \longmapsto s(x, \Omega \setminus \Omega') = \int_{\Omega \setminus \Omega'} s(x, dy)$$

belongs to the space $B_{\mathrm{loc}}(\Omega')$.

Proof. The proof is divided into two steps.

Step 1: It is easy to see that if the principal symbol of P is non-positive, then the integro-differential operator $W = P + S$ satisfies the condition (p_1). Indeed, we have

$$\begin{aligned} & x \in \overset{\circ}{\Omega}, \ u \in C_0^2(\Omega), \ u \geq 0 \text{ in } \Omega, \ u(x) = 0 \\ & \Longrightarrow \\ & Wu(x) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \int_\Omega s(x, dy) u(y) \geq 0. \end{aligned}$$

Conversely, it suffices to show that if condition (p_1) is satisfied, then we have

$$\begin{aligned} & x \in \overset{\circ}{\Omega}, \, h \in C_0^2(\Omega), \, h \geq 0 \text{ in } \Omega, \, h(x) = \frac{\partial h}{\partial x_i}(x) = 0 \ (1 \leq i \leq n) \\ & \Longrightarrow Wh(x) - \int_\Omega s(x, dy) h(y) \geq 0. \end{aligned} \tag{8.37}$$

Indeed, by virtue of formula (8.35) it follows from assertion (8.37) with $h := \Phi_x^\xi u$ that

$$\begin{aligned} & x \in \overset{\circ}{\Omega}, \ u \in C_0^2(\Omega), \ u \geq 0 \text{ in } \Omega \text{ and } \xi = (\xi_i) \in \mathbf{R}^n \\ & \Longrightarrow \\ & 2u(x) \left(\sum_{i,j=1}^{n} a^{ij}(x) \xi_i \xi_j \right) = W \left(\Phi_x^\xi u \right)(x) - \int_\Omega s(x, dy) \Phi_x^\xi(y) u(y) \geq 0. \end{aligned}$$

This proves that

$$\sum_{i,j=1}^{n} a^{ij}(x) \xi_i \xi_j \geq 0 \quad \text{for all } (x, \xi) \in \overset{\circ}{\Omega} \times \mathbf{R}^n.$$

Now we choose a real-valued, smooth function $\varphi \in C_0^\infty(\Omega)$ such that

$$\begin{cases} 0 \leq \varphi(y) \leq 1 & \text{in } \Omega, \\ \varphi(y) = 1 & \text{in a neighborhood of } x. \end{cases}$$

Then we have, by condition (p_1) with $u := \varphi\, h$ and (8.5),

$$\begin{aligned} 0 \leq W(\varphi\, h)(x) &= Wh(x) - W((1-\varphi)h)(x) \\ &= Wh(x) - \int_\Omega s(x, dy)(1 - \varphi(y))h(y) \\ &= Wh(x) - \int_\Omega s(x, dy)h(y) - \int_\Omega s(x, dy)\varphi(y)h(y), \end{aligned}$$

since

$$x \notin \operatorname{supp}(1-\varphi)h.$$

This proves that

$$Wh(x) - \int_\Omega s(x, dy)h(y) \geq -\int_\Omega s(x, dy)\varphi(y)h(y). \tag{8.38}$$

On the other hand, we have, by Taylor's formula,

$$\begin{aligned} &h(y) \\ &= h(x) - \sum_{i=1}^n \frac{\partial h}{\partial x_i}(x)\,(y_i - x_i) - \sum_{i,j=1}^n (y_i - x_i)\left(y_j - x_j\right) R_{ij}h(x, y) \\ &= \frac{1}{2} \sum_{i,j=1}^n (y_i - x_i)\left(y_j - x_j\right) \int_0^1 \frac{\partial^2 h}{\partial x_i\, \partial x_j}(x + t(y-x))(1-t)\, dt. \end{aligned}$$

If V is an open neighborhood of the support $\operatorname{supp}\varphi$ of φ, then it follows that

$$\begin{aligned} &\varphi(y)h(y) \\ &= \frac{1}{2}\varphi(y) \sum_{i,j=1}^n (y_i - x_i)\left(y_j - x_j\right) \int_0^1 \frac{\partial^2 h}{\partial x_i\, \partial x_j}(x + t(y-x))(1-t)\, dt \\ &\qquad \text{for all } y \in V. \end{aligned}$$

Hence we have the inequality

$$\varphi(y)h(y) \leq \frac{n}{4}\, \|h\|_{C^2(\Omega)}\, \chi_V(y) \quad \text{for all } y \in \Omega,$$

and so

$$\int_{\Omega} s(x, dy)\varphi(y)h(y) \leq n \, \|h\|_{C^2(\Omega)} \int_V s(x, dy) \, |y - x|^2 \, .$$

However, by condition (NS.2) it follows from an application of the Lebesgue dominated convergence theorem ([Fo2, Theorem 2.24]) that, for any given positive number ε, we can find a function φ with a sufficiently small neighborhood of x such that

$$\int_{\Omega} s(x, dy)\varphi(y)h(y) \leq \varepsilon. \tag{8.39}$$

Therefore the desired assertion (8.37) follows by combining inequalities (8.38) and (8.39), since ε is arbitrary.

Step 2: If the operator S is of the form

$$Su(x) = \int_{\Omega} s(x, dy) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x) \, (y_i - x_i) \right) \right],$$

then we have, by condition (8.36),

$$\begin{aligned}
& x \in \overset{\circ}{\Omega}, \; u \in C_0^2(\Omega), \; u(x) = \sup_{\Omega} u \geq 0 \\
& \Longrightarrow \\
& Wu(x) \\
= \; & Pu(x) + Su(x) \\
= \; & \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + c(x)u(x) + \int_{\Omega} s(x, dy) \, (u(y) - \sigma(x, y)u(x)) \\
\leq \; & c(x)u(x) + \int_{\Omega} s(x, dy)(1 - \sigma(x, y))u(x) \\
= \; & \left(P1(x) + \int_{\Omega} s(x, dy)(1 - \sigma(x, y) \right) u(x) \\
\leq \; & 0.
\end{aligned}$$

Conversely, we assume that the integro-differential operator $W = P + S$ satisfies the positive maximum principle (PM). Then it follows that W satisfies condition (p_1), so that the principal symbol of P is non-positive.

On the other hand, we can choose an increasing sequence $\{\theta_j\}$ of smooth functions in $C_0^{\infty}(\Omega)$ such that

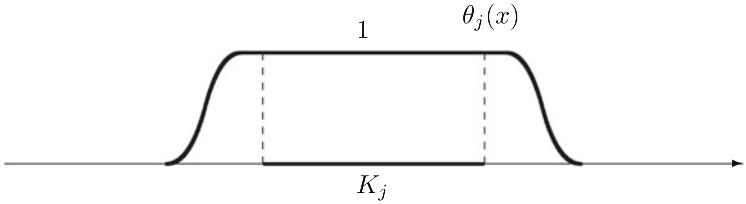


Fig. 8.6 The function $\theta_j(x)$

$$\begin{cases} 0 \le \theta_j(y) \le 1 & \text{in } \Omega, \\ \theta_j(y) = 1 & \text{in a neighborhood of } x, \\ \sup_{j \ge 1} \theta_j(y) = 1 & \text{for each point } y \in \Omega. \end{cases}$$

For example, if $\{K_j\}$ is an increasing sequence of compact subsets of Ω such that K_j is contained in the interior $\overset{\circ}{K}_{j+1}$ of K_{j+1} for each j and that

$$\Omega = \bigcup_{j=1}^{\infty} K_j,$$

then it follows from an application of Corollary 5.6 that there exists a function $\theta_j \in C_0^\infty(\Omega)$ such that (see Fig. 8.6)

$$\begin{cases} 0 \le \theta_j(y) \le 1 & \text{in } \Omega, \\ \theta_j(y) = 1 & \text{on } K_j. \end{cases}$$

Then we have, by condition (MP),

$$\begin{aligned} 0 \ge W\theta_j(x) &= c(x)\theta_j(x) + \int_\Omega s(x, dy)\left(\theta_j(y) - \sigma(x, y)\theta_j(x)\right) \qquad (8.40) \\ &= c(x) + \int_\Omega s(x, dy)(\theta_j(y) - \sigma(x, y)) \\ &= P1(x) + \int_\Omega s(x, dy)(\theta_j(y) - \sigma(x, y)). \end{aligned}$$

Therefore, the desired condition (8.36) follows by letting $j \to \infty$ in inequality (8.40).

The proof of Theorem 8.8 is complete.

In Sect. 7.4, we give a precise definition of the *principal symbol* of a differential operator.

8.1.2 Positive Borel Kernels and Pseudo-Differential Operators

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$ with smooth boundary ∂D. Now we give two important examples of positive Borel kernels in terms of pseudo-differential operators:

Example 8.9. Let $s(x, y)$ be the distribution kernel of a properly supported, pseudo-differential operator $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$, $\kappa > 0$, and $s(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$. Then the integro-differential operator S_r, defined by the formula (see formula (8.2))

$$\begin{aligned} & S_r u(x) \\ &= \int_D s(x, y) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy, \quad x \in D, \end{aligned}$$

is absolutely convergent.

Proof. Indeed, we can write the integral $S_r u(x)$ in the form

$$\begin{aligned} & S_r u(x) \\ &= \int_D s(x, y) \left[1 - \sigma(x, y) \right] u(y)\, dy \\ &\quad + \int_D s(x, y) \sigma(x, y) \left(u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) dy. \end{aligned}$$

By using Taylor's formula

$$\begin{aligned} & u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \\ &= \sum_{i,j=1}^N (y_i - x_i)(y_j - x_j) \left(\int_0^1 (1 - t) \frac{\partial^2 u}{\partial x_i \partial x_j}(x + t(y - x)) dt \right), \end{aligned}$$

we can find a constant $C_1 > 0$ such that

$$\left| u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right| \leq C_1 |x - y|^2, \quad x, y \in \overline{D}.$$

On the other hand, it follows from an application of Theorem 7.36 that, for any compact $K \subset \mathbf{R}^N$, there exists a constant $C_2 > 0$ such that the distribution kernel $s(x, y)$ of $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$, $\kappa > 0$, satisfies the estimate

$$0 \leq s(x, y) \leq \frac{C_2}{|x-y|^{N+2-\kappa}}, \quad x, y \in \overline{D}, \;\; x \neq y.$$

Therefore, we have, for some constant $C_3 > 0$,

$$\begin{aligned}
&\left| \int_D s(x, y)\sigma(x, y) \left(u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) dy \right| \\
&\leq C_3 \|u\|_{C^2(\overline{D})} \int_D \frac{1}{|x-y|^{N+2-\kappa}} \cdot |x-y|^2 \, dy \\
&= C_3 \|u\|_{C^2(\overline{D})} \int_D \frac{1}{|x-y|^{N-\kappa}} \, dy.
\end{aligned}$$

Similarly, we have, for some constant $C_4 > 0$,

$$\left| \int_D s(x, y) \left[1 - \sigma(x, y)\right] u(y) \, dy \right| \leq C_4 \|u\|_{C(\overline{D})} \int_D \frac{1}{|x-y|^{N-\kappa}} \, dy,$$

since we have

$$\begin{aligned}
&\sigma(x, y) - 1 \\
&= \sigma(x, y) - \sigma(x, x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial \sigma}{\partial x_j}(x, x) \\
&= \sum_{i,j=1}^N (y_i - x_i)(y_j - x_j) \left(\int_0^1 (1-t) \frac{\partial^2 \sigma}{\partial x_i \partial x_j}(x, x + t(y-x)) dt \right).
\end{aligned}$$

Therefore, we obtain that the integral $S_r u(x)$ is absolutely convergent.

The proof of Example 8.9 is complete.

Example 8.10. Let $r(x', y')$ be the distribution kernel of a pseudo-differential operator $R \in L^{2-\kappa_1}_{1,0}(\partial D)$, $\kappa_1 > 0$, and $r(x', y') \geq 0$ off the diagonal $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$ in $\partial D \times \partial D$. Let $t(x, y)$ be the distribution kernel of a properly supported, pseudo-differential operator $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, $\kappa_2 > 0$, and $t(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N} = \{(x, x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$. Then the integro-differential operator Γ_r, defined by the formula

$$
\begin{aligned}
&\Gamma_r u(x') \\
&= \int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\
&\quad + \int_D t(x', y) \left[u(y) - u(x') \right] dy, \quad x' \in \partial D,
\end{aligned}
$$

is absolutely convergent.

Proof. Since $R \in L^{2-\kappa_1}_{1,0}(\partial D)$ and $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, it follows from an application of Theorem 7.36 that the kernels $r(x', y')$ and $t(x', y)$ satisfy respectively the estimates

$$
\begin{aligned}
&0 \le r(x', y') \le \frac{C'}{|x' - y'|^{(N-1)+2-\kappa_1}}, \quad x', y' \in \partial D, \ x' \neq y', \\
&0 \le t(x', y) \le \frac{C''}{|x' - y|^{N+1-\kappa_2}}, \quad x' \in \partial D, \ y \in D,
\end{aligned}
$$

where $|x' - y'|$ denotes the geodesic distance between x' and y' with respect to the Riemannian metric of ∂D. Therefore, we obtain that the integrals

$$
\begin{aligned}
&R_r u(x') \\
&= \int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy', \\
&T_r u(x') = \int_D t(x', y) \left[u(y) - u(x') \right] dy
\end{aligned}
$$

are both absolutely convergent.

The proof of Example 8.10 is complete.

8.2 Maximum Principles for Waldenfels Operators

In this section we prove the strong maximum principle and Hopf's boundary point lemma for second-order elliptic Waldenfels operators which play an essential role in Chaps. 9 and 10.

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$ with smooth boundary ∂D. We consider a second-order elliptic *Waldenfels integro-differential operator W* with real coefficients such that

$$
\begin{aligned}
Wu(x) &= Pu(x) + Su(x) \\
&:= \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\
&\quad + \sum_{j=1}^{N} a_\sigma^j(x) \frac{\partial u}{\partial x_j}(x) + a_\sigma(x)u(x) \\
&\quad + \int_{\overline{D}} s(x,dy) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^{N} \frac{\partial u}{\partial x_j}(x) \left(y_j - x_j\right) \right) \right].
\end{aligned}
$$

Here:

(1) $a^{ij}(x) \in C(\overline{D})$, $a^{ij}(x) = a^{ji}(x)$ for all $x \in \overline{D}$ and $1 \leq i, j \leq N$, and there exists a constant $a_0 > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i \xi_j \geq a_0 |\xi|^2 \quad \text{for all } (x,\xi) \in \overline{D} \times \mathbf{R}^N.$$

(2) $b^i(x) \in C(\overline{D})$ for all $1 \leq i \leq N$.
(3) $c(x) \in C(\overline{D})$, and $c(x) \leq 0$ in D, but $c(x) \not\equiv 0$ in D.
(4) The positive Borel kernel $s(x,dy)$ satisfies the following three conditions (NS.1′), (NS.2′) and (NS.3′):

(NS.1′) $s(x,\{x\}) = 0$ for every $x \in \overline{D}$.
(NS.2′) For every non-negative function $f \in C(\overline{D})$, the function

$$\overline{D} \ni x \longmapsto \int_{\overline{D}} s(x,dy)|y-x|^2 f(y)$$

belongs to the space $C(\overline{D})$.
(NS.3′) For every open subset Ω of $\overline{D}$, the function

$$\Omega \ni x \longmapsto s(x, \overline{D} \setminus \Omega)$$

belongs to the space $B_{\text{loc}}(\Omega)$.

(5) $a_\sigma^j(x) = S(\sigma_x^j)(x)$ where $\sigma_x^j(y) = \sigma(x,y)(y_j - x_j)$ for all $1 \leq j \leq N$.
(6) $a_\sigma(x) = S(\sigma_x)(x)$ where $\sigma_x(y) = \sigma(x,y)$.
(7) Finally, we assume that

$$
\begin{aligned}
W1(x) &= P1(x) + S1(x) \qquad (8.41) \\
&= c(x) + a_\sigma(x) + \int_{\overline{D}} s(x,dy)\left(1 - \sigma(x,y)\right) \leq 0 \quad \text{in D.}
\end{aligned}
$$

8.2.1 The Weak Maximum Principle

First, we prove the weak maximum principle:

Theorem 8.11 (the weak maximum principle). *Let $W = P + S$ be a second-order elliptic Waldenfels operator. If S maps $C^2(\overline{D})$ continuously into $C(D)$, then we have the following two assertions (i) and (ii):*

(i) If a function $u(x) \in C(\overline{D}) \cap C^2(D)$ satisfies the condition

$$Wu(x) \geq 0 \quad \text{in } D$$

and if the function $W1$ satisfies the condition

$$W1(x) = P1(x)+S1(x) = c(x)+a_\sigma(x)+\int_{\overline{D}} s(x,dy)(1-\sigma(x,y)) < 0 \text{ in } D,$$

then the function $u(x)$ may take its positive maximum *only on the boundary ∂D.*

(ii) If a function $u(x) \in C(\overline{D}) \cap C^2(D)$ satisfies the condition

$$Wu(x) > 0 \quad \text{in } D$$

and if the function $W1$ satisfies the condition

$$W1(x) = P1(x)+S1(x) = c(x)+a_\sigma(x)+\int_{\overline{D}} s(x,dy)(1-\sigma(x,y)) \leq 0 \text{ in } D,$$

then the function $u(x)$ may take its non-negative maximum *only on the boundary ∂D.*

Proof. Assume, to the contrary, that there exists a point x_0 of D such that

$$u(x_0) = \max_{x \in \overline{D}} u(x).$$

Then it follows that

$$\frac{\partial u}{\partial x_i}(x_0) = 0, \quad 1 \leq i \leq N;$$

$$\sum_{i,j=1}^{N} a^{ij}(x_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) \leq 0.$$

Hence we have the inequalities

$$Pu(x_0) = \sum_{i,j=1}^{N} a^{ij}(x_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) + c(x_0)u(x_0) \leq c(x_0)u(x_0), \tag{8.42}$$

and

$$
\begin{aligned}
Su(x_0) &= \int_{\overline{D}} s(x_0, dy)(u(y) - \sigma(x_0, y)u(x_0)) + a_\sigma(x_0)u(x_0) \qquad (8.43)\\
&= \int_{\overline{D}} s(x_0, dy)(u(y) - u(x_0)) \\
&\quad + \left(\int_{\overline{D}} s(x_0, dy)(1 - \sigma(x_0, y)) + a_\sigma(x_0) \right) u(x_0) \\
&= \int_{\overline{D}} s(x_0, dy)(u(y) - u(x_0)) + S1(x_0)u(x_0) \\
&\leq S1(x_0)u(x_0).
\end{aligned}
$$

Assertion (i): If $Wu(x) \geq 0$ in D, $W1(x) < 0$ in D and if $u(x_0) = \max_{\overline{D}} u > 0$, then it follows from inequalities (8.41)–(8.43) that

$$
\begin{aligned}
0 &\leq Wu(x_0) = Pu(x_0) + Su(x_0) \leq (c(x_0) + S1(x_0))\, u(x_0) \\
&= \max_{x \in \overline{D}} u(x) \cdot W1(x_0) < 0.
\end{aligned}
$$

This is a contradiction.

Assertion (ii): Similarly, if $Wu(x) > 0$, $W1(x) \leq 0$ in D and if $u(x_0) = \max_{\overline{D}} u \geq 0$, then it follows from inequalities (8.42) and (8.43) that

$$
0 < Wu(x_0) = Pu(x_0) + Su(x_0) \leq \max_{x \in \overline{D}} u(x) \cdot W1(x_0) \leq 0.
$$

This is also a contradiction.

The proof of Theorem 8.11 is complete.

As an application of the weak maximum principle, we can obtain a pointwise estimate for solutions of the non-homogeneous equation $Wu = f$:

Theorem 8.12. *Let $W = P + S$ be a second-order elliptic Waldenfels operator such that S maps $C^2(\overline{D})$ continuously into $C(D)$, and assume that*

$$
W1(x) < 0 \quad \text{on } \overline{D} = D \cup \partial D.
$$

Then we have, for all $u \in C(\overline{D}) \cap C^2(D)$,

$$
\max_{\overline{D}} |u| \leq \max \left\{ \left(\frac{1}{\min_{\overline{D}}(-W1)} \right) \sup_D |Wu|, \ \max_{\partial D} |u| \right\}. \qquad (8.44)
$$

Proof. We let

$$
M = \max \left\{ \left(\frac{1}{\min_{\overline{D}}(-W1)} \right) \sup_D |Wu|, \ \max_{\partial D} |u| \right\},
$$

and consider two functions

$$v_\pm(x) = M \pm u(x).$$

Then it follows that

$$Wv_\pm(x) = M \cdot W1(x) \pm Wu(x) \leq 0 \quad \text{in } D.$$

Hence, by applying part (i) of Theorem 8.11 to the functions $-v_\pm(x)$ we obtain that the functions $v_\pm(x)$ may take their negative minimums only on the boundary ∂D. However, we have the inequality

$$v_\pm(x) = M \pm u(x) \geq 0 \quad \text{on } \partial D.$$

Therefore, we obtain that

$$v_\pm(x) \geq 0 \quad \text{on } \overline{D}.$$

This proves the desired estimate (8.44).

The proof of Theorem 8.12 is complete.

8.2.2 The Strong Maximum Principle

The next theorem is a generalization of the strong maximum principle for the Laplacian to the integro-differential operator case:

Theorem 8.13 (the strong maximum principle). *Let $W = P + S$ be a second-order elliptic Waldenfels operator such that S maps $C^2(\overline{D})$ continuously into $C(D)$. Assume that a function $u(x) \in C^2(\overline{D})$ satisfies the conditions*

$$Wu(x) \geq 0 \quad \textit{in } D$$

and that $M = \max_{\overline{D}} u \geq 0$. If the function $u(x)$ takes its non-negative maximum M at an interior point of D, then it is a constant.

Proof. The proof is divided into four steps.

Step 1: We let

$$M = \max_{x \in \overline{D}} u(x) \geq 0,$$
$$F = \{x \in D : u(x) = M\},$$

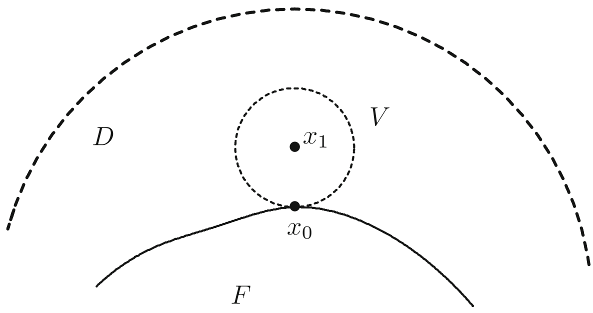


Fig. 8.7 The open ball V

and assume, to the contrary, that

$$F \subsetneqq D.$$

Since F is closed in D, we can find a point x_0 of F and an open ball V contained in the set $D \setminus F$, centered at x_1, such that (see Fig. 8.7)

(a) $V \subset D \setminus F$;
(b) x_0 is on the boundary ∂V of V.

Step 2: The next lemma on the existence of "barriers" is an essential step in the proof of Theorem 8.13:

Lemma 8.14. *Assume that*

$$s(x_0, \overline{V}) = 0. \tag{8.45}$$

Then there exists a function $v(x) \in C^\infty(\overline{D})$ which satisfies the following four properties (i) through (iv):

(i) $v(x) > 0$ *in* V.
(ii) $v(x) < 0$ *on* $\overline{D} \setminus \overline{V}$.
(iii) $v(x) = 0$ *on* ∂V.
(iv) $Wv(x_0) > 0$.

Proof. We take a function $\varphi(x) \in C_0^\infty(D)$ such that

$$\begin{cases} 0 \leq \varphi(x) \leq 1 & \text{in } D, \\ \varphi(x) = 1 & \text{on } \overline{V}, \end{cases}$$

and define a smooth function $v_q(x)$ by the formula

$$v_q(x) = \left(\exp\left[-q|x - x_1|^2\right] - \exp\left[-q\rho^2\right]\right)\varphi(x), \quad \rho = |x_0 - x_1|, \tag{8.46}$$

where q is a positive constant to be chosen later on. Then it is easy to see that the function $v_q(x)$ satisfies conditions (i) through (iii). Hence it suffices to show that the function $Wv_q(x)$ satisfies condition (iv) for q sufficiently large.

First, we estimate the function $Pv_q(x_0)$ from below. To do this, it should be noted that

$$v_q(x_0) = 0, \tag{8.47a}$$

$$\nabla v_q(x_0) = 2q(x_1 - x_0)\exp\left[-q\rho^2\right] \neq 0. \tag{8.47b}$$

Hence we have

$$Pv_q(x_0) = \left\{4q^2 \sum_{i,j=1}^{N} a^{ij}(x_0)(x_{1i} - x_{0i})(x_{1j} - x_{0j}) \right.$$
$$\left. -2q \sum_{i=1}^{N} \left(a^{ii}(x_0) + b^i(x_0)(x_{0i} - x_{1i})\right)\right\} \exp\left[-q\rho^2\right].$$

Since the matrix (a^{ij}) is positive definite, we can estimate the function $Pv(x_0)$ from below as follows:

$$Pv_q(x_0) \geq \left(C_1 q^2 + C_2 q\right)\exp\left[-q\rho^2\right], \tag{8.48}$$

where

$$C_1 = 4a_0\rho^2 > 0$$

and C_2 are constants independent of q.

Secondly, we estimate the function $Sv_q(x_0)$. By condition (8.45), we can write the function $Sv_q(x_0)$ in the form

$$Sv_q(x_0) \tag{8.49}$$
$$= \sum_{i=1}^{N} a^i(x_0)\frac{\partial v_q}{\partial x_i}(x_0) + a(x_0)v_q(x_0)$$
$$+ \int_{\overline{D}} s(x_0, dy)\left[v_q(y) - \sigma(x_0, y)\left(v_q(x_0) + \sum_{j=1}^{N}\frac{\partial v_q}{\partial x_j}(x_0)\left(y_j - x_{0j}\right)\right)\right]$$
$$= \sum_{i=1}^{N} a^i(x_0)\frac{\partial v_q}{\partial x_i}(x_0) + \int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\,(1 - \sigma(x_0, y))\,v_q(y)$$
$$+ \int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\sigma(x_0, y)\left[v_q(y) - v_q(x_0) - \sum_{j=1}^{N}\frac{\partial v_q}{\partial x_j}(x_0)\left(y_j - x_{0j}\right)\right].$$

By using formula (8.46) and Taylor's formula, we can find a positive constant C_3, independent of q, such that

$$\left|\frac{\partial v_q}{\partial x_i}(x_0)\right| \leq C_3 q \exp\left[-q\rho^2\right] \quad (1 \leq i \leq N), \tag{8.50}$$

$$\left|v_q(y)\right| \leq C_3 \exp\left[-q\rho^2\right] \quad \text{for all } y \in \overline{D} \setminus \overline{V}, \tag{8.51}$$

$$\left|v_q(y) - v_q(x_0) - \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0)\left(y_j - x_{0j}\right)\right| \tag{8.52}$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right] |y - x_0|^2 \quad \text{for all } y \in \overline{D} \setminus \overline{V}.$$

By virtue of inequalities (8.50) and (8.51), we can find positive constants C_4 and C_5, independent of q, such that

$$\left|\sum_{i=1}^{N} a^i(x_0) \frac{\partial v_q}{\partial x_i}(x_0)\right| \leq C_4 q \exp\left[-q\rho^2\right],$$

$$\left|\int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\,(1 - \sigma(x_0, y))\, v_q(y)\right| \leq C_5 \exp\left[-q\rho^2\right].$$

Moreover, by using condition (NS.2′) we can choose a local unity function $\sigma(x, y)$ such that

$$\int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\sigma(x_0, y)\, |y - x_0|^2 \leq \frac{C_1}{2C_3}, \tag{8.53}$$

if the support of $\sigma(x_0, \cdot)$ is sufficiently close to x_0. Hence it follows from inequalities (8.52) and (8.53) that

$$\int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\sigma(x_0, y) \left[v_q(y) - v_q(x_0) - \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0)\left(y_j - x_{0j}\right)\right]$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right] \int_{\overline{D}\setminus\overline{V}} s(x_0, dy)\sigma(x_0, y)\, |y - x_0|^2$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right] \times \frac{C_1}{2C_3} = \frac{C_1}{2} q^2 \exp\left[-q\rho^2\right].$$

Summing up, we have proved that

$$\left|S v_q(x_0)\right| \leq \left[\frac{C_1}{2} q^2 + C_4 q + C_5\right] \exp\left[-q\rho^2\right]. \tag{8.54}$$

Therefore, it follows from inequalities (8.48) and (8.54) that

$$\begin{aligned} Wv_q(x_0) &= Pv_q(x_0) + Sv_q(x_0) \geq P_q v(x_0) - \left|S_q v(x_0)\right| \\ &\geq \left[\frac{C_1}{2}q^2 + C_4 q + C_5\right] \exp\left[-q\rho^2\right] \\ &> 0, \end{aligned} \tag{8.55}$$

if we take a positive constant q so large that

$$q > \frac{-C_4 + \sqrt{C_4^2 - 2C_1C_5}}{C_1}.$$

The desired assertion (iv) follows from inequality (8.55) with $v(x) := v_q(x)$.

The proof of Lemma 8.14 is complete.

Step 3: We recall that

$$u(x_0) = M = \max_{x \in \overline{D}} u(x) \geq 0.$$

Hence we have the inequality

$$Pu(x_0) = \sum_{i,j=1}^{N} a^{ij}(x_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) + c(x_0)u(x_0) \leq c(x_0)u(x_0) \leq 0,$$

and, by condition (8.41),

$$\begin{aligned} Su(x_0) &= \int_{\overline{D}} s(x_0, dy)\,(u(y) - \sigma(x_0, y)u(x_0)) + a(x_0)u(x_0) \\ &= \int_{\overline{D}} s(x_0, dy)\,(u(y) - u(x_0)) \\ &\quad + \int_{\overline{D}} s(x_0, dy)\,(1 - \sigma(x_0, y)u(x_0)) + a(x_0)u(x_0) \\ &= \int_{\overline{D}} s(x_0, dy)\,(u(y) - u(x_0)) + a(x_0)u(x_0) \\ &\leq \int_{\overline{D}} s(x_0, dy)\,(u(y) - u(x_0)) \\ &\quad + \left(\int_{\overline{D}} s(x_0, dy)(1 - \sigma(x_0, y)) + a(x_0)\right) u(x_0) \\ &\leq S1(x_0)u(x_0) \leq 0. \end{aligned}$$

This implies that

$$Pu(x_0) = 0, \quad Su(x_0) = 0.$$

Indeed, it suffices to note that

$$0 \leq Wu(x_0) = Pu(x_0) + Su(x_0) \leq 0.$$

Thus we have

$$0 = Su(x_0) \leq \int_{\overline{D}} s(x_0, dy)\, (u(y) - u(x_0)) \leq 0,$$

and hence

$$\int_{\overline{D}} s(x_0, dy)\, (u(y) - u(x_0)) = 0. \tag{8.56}$$

However, we have

$$u(y) - u(x_0) < 0 \quad \text{for all } y \in \overline{V} \setminus \{x_0\}.$$

By assertion (8.56), this implies that the condition (8.45)

$$s(x_0, \overline{V}) = 0$$

holds true, since we have

$$s(x_0, \{x_0\}) = 0.$$

Therefore, we can apply Lemma 8.14 to our situation.

Step 4: If v is the function given in Lemma 8.14, we introduce a function

$$u_\lambda(x) = u(x) + \lambda v(x),$$

where λ is a positive constant to be chosen later on. Then, by Lemma 8.14 we have the following four assertions (a) through (d) (see Fig. 8.8):

(a) There exists a neighborhood V' of x_0 such that

$$Wv > 0 \quad \text{in } V',$$

since $Wv = Pv + Sv$ is a continuous function in D.

(b) $Wu_\lambda = Wu + \lambda Wv \geq \lambda Wv > 0$ in V'.

(c) $u_\lambda = u + \lambda v \leq u \leq M$ on $\overline{D} \setminus V$, since $v \leq 0$ on $\overline{D} \setminus V$.

(d) $u_\lambda = u + \lambda v \leq M$ on $\overline{V} \setminus V'$ for λ sufficiently small, since $u < M$ on $\overline{V} \setminus V'$.

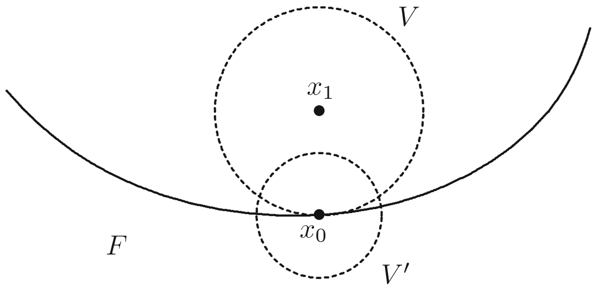


Fig. 8.8 The ball V and the neighborhood V' of x_0

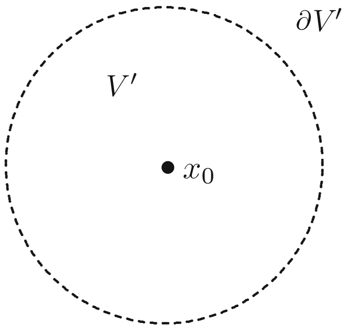


Fig. 8.9 The neighborhood V' of x_0

Therefore, we obtain the following assertions (8.57) (see Fig. 8.9):

$$Wu_\lambda > 0 \quad \text{in } V', \tag{8.57a}$$

$$u_\lambda \leq M \quad \text{on } \partial V', \tag{8.57b}$$

$$u_\lambda(x_0) = M. \tag{8.57c}$$

However, these assertions (8.57) contradict part (ii) of Theorem 8.11 (the weak maximum principle).

Summing up, we have proved that

$$F = D,$$

that is,

$$u(x) \equiv M \quad \text{in } D.$$

Now the proof of Theorem 8.13 is complete.

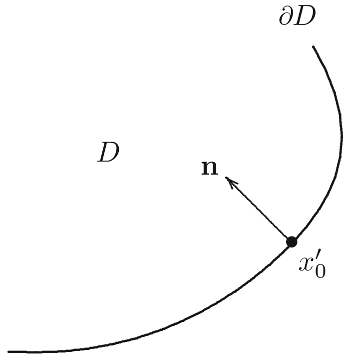


Fig. 8.10 The unit interior normal $\mathbf{n}$ at x'_0

8.2.3 The Hopf Boundary Point Lemma

Finally, we consider the unit interior normal derivative $(\partial u)/(\partial \mathbf{n})$ at a boundary point where the function $u(x)$ takes its non-negative maximum.

The Hopf boundary point lemma reads as follows:

Theorem 8.15 (Hopf's boundary point lemma). *Let $W = P + S$ be a second-order elliptic Waldenfels operator such that S maps $C^2(\overline{D})$ continuously into $C(\overline{D})$. Assume that a function $u(x) \in C^2(\overline{D})$ satisfies the conditions*

$$Wu(x) \geq 0 \quad \text{in } D,$$

$$\max_{\overline{D}} u \geq 0,$$

and further that there exists a point $x'_0 \in \partial D$ such that $u(x'_0) = \max_{\overline{D}} u$. Then we have the inequality (see Fig. 8.10)

$$\frac{\partial u}{\partial \mathbf{n}}(x'_0) < 0,$$

unless the function $u(x)$ is a constant in D.

Proof. By Theorem 8.13, we have only to consider the following case:

$$\begin{cases} u(x'_0) = M = \max_{x \in \overline{D}} u(x) \geq 0, \\ u(y) < u(x'_0) \quad \text{for all } y \in D. \end{cases} \tag{8.58}$$

The proof is divided into three steps.

Step 1: Now we assume, to the contrary, that

$$\frac{\partial u}{\partial \mathbf{n}}(x'_0) = 0. \tag{8.59}$$

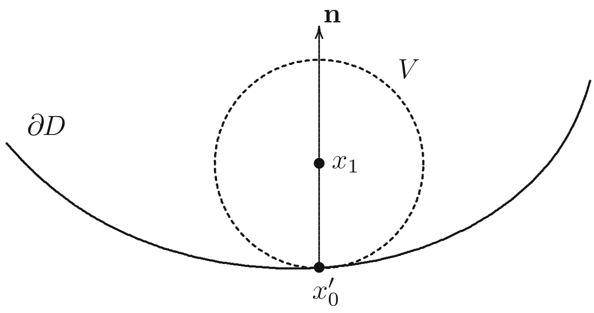


Fig. 8.11 The open ball V

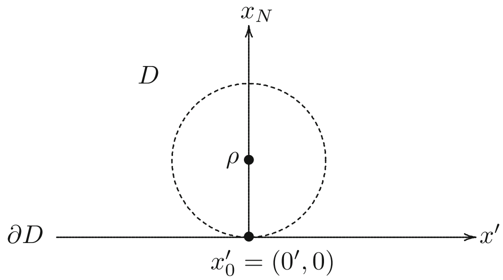


Fig. 8.12 The local coordinate system (x', x_N)

We can find an open ball V contained in the domain D, centered at x_1, such that (see Fig. 8.11)

(a) The point x'_0 is on the boundary ∂V of V.
(b) $\mathbf{n} = s(x_1 - x'_0)$ for some $s > 0$.

Step 2: The next lemma on the existence of "barriers" is an essential step in the proof of Theorem 8.15, just as in the proof of Theorem 8.13:

Lemma 8.16. *There exists a function* $v(x) \in C^\infty(\overline{D})$ *which satisfies the following five properties (i) through (v):*

(i) $v(x) > 0$ *in* V.
(ii) $v(x) < 0$ *on* $\overline{D} \setminus \overline{V}$.
(iii) $v(x) = 0$ *on* ∂V.
(iv) $Wv(x'_0) > 0$.
(v) $(\partial v/\partial \mathbf{n})(x'_0) > 0$.

Proof. Near the boundary point x'_0, we introduce local coordinate systems (x', x_N) such that $x' = (x_1, x_2, \ldots, x_{N-1})$ give local coordinates for the boundary ∂D and that (see Fig. 8.12)

$$D = \{(x', x_N) : x_N > 0\},$$
$$\partial D = \{(x', x_N) : x_N = 0\},$$

$$x_0' = (0', 0) = (0, \dots, 0, 0),$$
$$x_1 = (0', \rho) = (0, \dots, 0, \rho).$$

We take a function $\psi(x) \in C^\infty(\overline{D})$ such that

$$\begin{cases} 0 \le \psi(x) \le 1 & \text{on } \overline{D}, \\ \psi(x) = 1 & \text{in a tubular neighborhood of } \partial D, \end{cases}$$

and define a function $v_q(x) = v_q(x', x_N)$ by the formula

$$v_q(x) = \left(\exp\left[-q(|x'|^2 + (x_N - \rho)^2)\right] - \exp\left[-q\rho^2\right]\right)\psi(x), \tag{8.60}$$
$$\rho = |x_0' - x_1|,$$

where q is a positive constant to be chosen later on. Then it is easy to see that the function $v_q(x)$ satisfies conditions (i) through (iii) and (v). Hence it suffices to show that the function $Wv_q(x)$ satisfies condition (iv) for q sufficiently large.

First, we estimate the function $Pv_q(x_0')$ from below. To do this, it should be noticed that

$$v_q(x_0') = 0, \tag{8.61a}$$
$$\frac{\partial v_q}{\partial x_i}(x_0') = 0 \quad (1 \le i \le N-1), \tag{8.61b}$$
$$\frac{\partial v_q}{\partial x_N}(x_0') = 2\rho q \exp\left[-q\rho^2\right], \tag{8.61c}$$

and that

$$\frac{\partial^2 v_q}{\partial x_i \partial x_j}(x_0') = -2q\delta_{ij} \exp\left[-q\rho^2\right] \quad (1 \le i, j \le N-1),$$
$$\frac{\partial^2 v_q}{\partial x_N^2}(x_0') = (4q^2\rho^2 - 2q)\exp\left[-q\rho^2\right].$$

Hence we have

$$Pv_q(x_0') = \left\{ 4a^{NN}(x_0')\rho^2 q^2 + 2\left(b^{NN}(x_0')\rho - \sum_{i=1}^N a^{ii}(x_0')\right) q \right\} \exp\left[-q\rho^2\right].$$

Since the matrix (a^{ij}) is positive definite, we can estimate the function $Pv_q(x_0')$ from below as follows:

$$Pv_q(x_0') \ge \left(C_1 q^2 + C_2 q\right)\exp\left[-q\rho^2\right], \tag{8.62}$$

where

$$C_1 = 4a_0\rho^2 > 0$$

and C_2 are constants independent of q.

Secondly, in order to estimate the function $Sv_q(x_0')$ we study the Borel kernel $s(x_0', dy)$: By conditions (8.58) and (8.59), it follows that

$$\frac{\partial u}{\partial x_i}(x_0') = 0, \quad 1 \leq i \leq N,$$
$$\frac{\partial^2 u}{\partial x_N^2}(x_0') \leq 0.$$

Hence we have the inequality

$$\begin{aligned} Pu(x_0') &= \sum_{i,j=1}^{N} a^{ij}(x_0') \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0') + c(x_0')u(x_0') \\ &= a^{NN}(x_0') \frac{\partial^2 u}{\partial x_N^2}(x_0') + \sum_{i,j=1}^{N-1} a^{ij}(x_0') \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0') + c(x_0')u(x_0') \leq 0. \end{aligned}$$

Moreover, by condition (8.41) it follows that

$$\begin{aligned} Su(x_0') &= \int_{\overline{D}} s(x_0', dy) \left(u(y) - \sigma(x_0', y)u(x_0') \right) + a_\sigma(x_0')u(x_0') \\ &= \int_{\overline{D}} s(x_0', dy) \left(u(y) - u(x_0') \right) \\ &\quad + \left(a_\sigma(x_0') + \int_{\overline{D}} s(x_0', dy) \left(1 - \sigma(x_0', y) \right) \right) u(x_0') \\ &\leq \int_{\overline{D}} s(x_0', dy) \left(u(y) - u(x_0') \right) \leq 0. \end{aligned}$$

This implies that

$$Pu(x_0') = 0, \quad Su(x_0') = 0.$$

Indeed, it suffices to note that

$$0 \leq Wu(x_0') = Pu(x_0') + Su(x_0') \leq 0.$$

Thus we obtain that

$$0 = Su(x_0') \leq \int_{\overline{D}} s(x_0', dy) \left(u(y) - u(x_0')\right) \leq 0,$$

so that

$$\int_{\overline{D}} s(x_0', dy) \left(u(y) - u(x_0')\right) = 0. \tag{8.63}$$

However, if we let

$$G = \{x' \in \partial D : u(x') = \max_{x \in \overline{D}} u(x)\},$$

we obtain from condition (8.58) that

$$u(y) - u(x_0') < 0 \quad \text{for all } y \in D \cup (\partial D \setminus G).$$

Hence it follows from condition (8.63) that

$$s(x_0', \overline{V}) = 0,$$

since we have

$$u(y) - u(x_0') < 0 \quad \text{for all } y \in \overline{V} \setminus \{x_0'\},$$
$$s(x_0', \{x_0'\}) = 0.$$

Therefore, we can write the function $Sv_q(x_0')$ in the form

$$\begin{aligned}
&Sv_q(x_0') \qquad (8.64)\\
&= \sum_{i=1}^{N} a^i(x_0') \frac{\partial v_q}{\partial x_i}(x_0') + a(x_0') v_q(x_0') \\
&\quad + \int_{\overline{D}} s(x_0', dy) \left[v_q(y) - \sigma(x_0', y) \left(v_q(x_0') + \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0') \left(y_j - x_{0j}' \right) \right) \right] \\
&= \sum_{i=1}^{N} a^i(x_0') \frac{\partial v_q}{\partial x_i}(x_0') + \int_{\overline{D} \setminus \overline{V}} s(x_0', dy) \left(1 - \sigma(x_0', y)\right) v_q(y) \\
&\quad + \int_{\overline{D} \setminus \overline{V}} s(x_0', dy) \sigma(x_0', y) \left[v_q(y) - v_q(x_0') - \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0') \left(y_j - x_{0j}' \right) \right].
\end{aligned}$$

By using formula (8.60) and Taylor's formula, we can find a positive constant C_3, independent of q, such that

$$\left|\frac{\partial v_q}{\partial x_i}(x_0')\right| \leq C_3 q \exp\left[-q\rho^2\right] \quad (1 \leq i \leq N-1), \tag{8.65}$$

$$\left|v_q(y)\right| \leq C_3 \exp\left[-q\rho^2\right] \quad \text{for all } y \in \overline{D} \setminus \overline{V}, \tag{8.66}$$

$$\left|v_q(y) - v_q(x_0') - \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0')\left(y_j - x_{0j}'\right)\right| \tag{8.67}$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right]\left|y - x_0'\right|^2 \quad \text{for all } y \in \overline{D} \setminus \overline{V}.$$

By virtue of inequalities (8.65) and (8.66), we can find positive constants C_4 and C_5, independent of q, such that

$$\left|\sum_{i=1}^{N} a^i(x_0')\frac{\partial v_q}{\partial x_i}(x_0')\right| \leq C_4 q \exp\left[-q\rho^2\right],$$

$$\left|\int_{\overline{D}\setminus\overline{V}} s(x_0', dy)\left(1 - \sigma(x_0', y)\right) v_q(y)\right| \leq C_5 \exp\left[-q\rho^2\right].$$

Moreover, by using condition (NS.2$'$) we can choose a local unity function $\sigma(x, y)$ such that

$$\int_{\overline{D}\setminus\overline{V}} s(x_0', dy)\sigma(x_0', y)\left|y - x_0'\right|^2 \leq \frac{C_1}{2C_3}, \tag{8.68}$$

if the support of $\sigma(x_0', \cdot)$ is sufficiently close to x_0'. Hence it follows from inequalities (8.67) and (8.68) that

$$\int_{\overline{D}\setminus\overline{V}} s(x_0', dy)\sigma(x_0', y)\left[v_q(y) - v_q(x_0') - \sum_{j=1}^{N} \frac{\partial v_q}{\partial x_j}(x_0')\left(y_j - x_{0j}'\right)\right]$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right] \int_{\overline{D}\setminus\overline{V}} s(x_0', dy)\sigma(x_0', y)\left|y - x_0'\right|^2$$

$$\leq C_3 q^2 \exp\left[-q\rho^2\right] \times \frac{C_1}{2C_3} = \frac{C_1}{2} q^2 \exp\left[-q\rho^2\right].$$

Summing up, we have proved that

$$\left|S v_q(x_0')\right| \leq \left[\frac{C_1}{2} q^2 + C_4 q + C_5\right] \exp\left[-q\rho^2\right]. \tag{8.69}$$

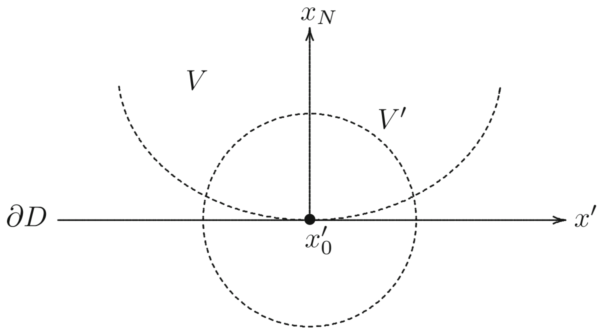


Fig. 8.13 The ball V and the neighborhood V' of x'_0

Therefore, it follows from inequalities (8.62) and (8.69) that

$$\begin{aligned} Wv_q(x'_0) &= Pv_q(x'_0) + Sv_q(x'_0) \geq Pv_q(x'_0) - \left|Sv_q(x'_0)\right| \\ &\geq \left[\frac{C_1}{2}q^2 + C_4 q + C_5\right] \exp\left[-q\rho^2\right] \\ &> 0, \end{aligned} \tag{8.70}$$

if we take a positive constant q so large that

$$q > \frac{-C_4 + \sqrt{C_4^2 - 2C_1C_5}}{C_1}.$$

The desired assertion (iv) follows from inequality (8.70) with $v(x) := v_q(x)$.
The proof of Lemma 8.16 is complete.

Step 3: If we introduce a function

$$u_\lambda(x) = u(x) + \lambda v(x)$$

for a positive constant λ, then we obtain from Lemma 8.16 the following four assertions (a) through (d) (see Fig. 8.13):

(a) There exists a neighborhood V' of x'_0 in $\overline{D}$ such that

$$Wv > 0 \quad \text{in } V' \cap \overline{D},$$

since $Wv = Pv + Sv$ is a continuous function on $\overline{D}$.
(b) $Wu_\lambda = Wu + \lambda Wv \geq \lambda Wv > 0$ in $V' \cap \overline{D}$.
(c) $u_\lambda = u + \lambda v \leq u \leq M$ on $\overline{D} \setminus V$, since $v \leq 0$ on $\overline{D} \setminus V$.
(d) $u_\lambda = u + \lambda v \leq M$ on $\overline{V} \setminus V'$ for λ sufficiently small, since $u < M$ on $\overline{V} \setminus V'$.

Hence it follows from an application of part (ii) of Theorem 8.11 with $u := u_\lambda$ that

$$u_\lambda \leq M \quad \text{in } V' \cap D,$$

so that

$$u(y) + \lambda v(y) = u_\lambda(y) \leq M = u_\lambda(x_0') = u(x_0') + \lambda v(x_0') \\ \text{for all } y \in V' \cap D.$$

This proves that

$$u(y) - u(x_0') \leq -\lambda(v(y) - v(x_0')) \quad \text{for all } y \in V' \cap D.$$

Therefore, we obtain that

$$\frac{\partial u}{\partial \mathbf{n}}(x_0') = \lim_{y \to x_0'} \frac{u(y) - u(x_0')}{y - x_0'} \leq -\lambda \lim_{y \to x_0'} \frac{v(y) - v(x_0')}{y - x_0'} = -\lambda \frac{\partial v}{\partial \mathbf{n}}(x_0') \\ < 0.$$

This contradicts hypothesis (8.59).

Now the proof of Theorem 8.15 is complete.

8.3 Notes and Comments

Chapter 8 is an expanded and revised version of Sect. 3.4 and Appendix C of the first edition of the present monograph. For a general study of maximum principles, the reader might refer to Protter–Weinberger [PW] and Bony–Courrège–Priouret [BCP].

Section 8.1: Proposition 8.1 is essentially due to Bony–Courrège–Priouret [BCP, Proposition I.1.5] and Theorem 8.2 is taken from Bony–Courrège–Priouret [BCP, Théorème I].

Section 8.2: Theorem 8.4 is taken from Bony–Courrège–Priouret [BCP, Théorème II] and Theorem 8.8 is taken from Bony–Courrège–Priouret [BCP, Théorème III], respectively. The boundary point lemma (Theorem 8.15) was proved independently by Hopf [Hp] and Oleĭnik [Ol].

In Taira [Ta5, Chapter 7], we prove various maximum principles for second-order *degenerate* elliptic differential operators, and reveal the underlying analytical mechanism of propagation of maximums in terms of subunit vectors introduced by Fefferman–Phong [FP].

Part III
Markov Processes, Semigroups and Boundary Value Problems

Chapter 9
Markov Processes, Transition Functions and Feller Semigroups

The content of this chapter may be summarized in the following diagram:

Probability	**Functional Analysis**	**Boundary Value Problems**
Strong Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$	Feller semigroup $\{T_t\}$	Infinitesimal generator $\mathfrak{A}$
Markov transition function $p_t(\cdot, dy)$	$T_t f(\cdot) = \int p_t(\cdot, dy) f(y)$	$T_t = \exp[t\mathfrak{A}]$
Chapman–Kolmogorov equation	Semigroup property $T_{t+s} = T_t \cdot T_s$	Waldenfels integro-differential operator $W = P + S$
Various diffusion phenomena	Function spaces $C(K)$, $C_0(K)$	Ventcel' boundary condition L

In this chapter we introduce a class of (temporally homogeneous) Markov processes which we will deal with in this book (Definition 9.3). Intuitively, the Markov property is that the prediction of subsequent motion of a physical particle, knowing its position at time t, depends neither on the value of t nor on what has been observed during the time interval $[0, t)$; that is, a physical particle "starts afresh". From the point of view of analysis, however, the transition function of a Markov process is something more convenient than the Markov process itself. In fact, it can be shown that the transition functions of Markov processes generate solutions of certain parabolic partial differential equations such as the classical diffusion equation; and, conversely, these differential equations can be used to construct and study the transition functions and the Markov processes themselves. In Sect. 9.1 we give the precise definition of a Markov transition function adapted to the theory of semigroups (Definition 9.4). A Markov process is called a strong Markov process if the "starting afresh" property holds not only for every fixed moment but also for suitable random times. In Sect. 9.1.7 we formulate precisely this "strong" Markov property (Definition 9.25), and give a useful criterion for the strong Markov property (Theorem 9.26). In Sect. 9.1.8 we introduce the basic notion of uniform stochastic

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_9,

continuity of transition functions (Definition 9.27), and give simple criteria for the strong Markov property in terms of transition functions (Theorems 9.28 and 9.29). In Sect. 9.2 we introduce a class of semigroups associated with Markov processes (Definition 9.30), called Feller semigroups, and we give a characterization of Feller semigroups in terms of Markov transition functions (Theorems 9.33 and 9.34). Section 9.3 is devoted to a version of the Hille–Yosida theorem (Theorem 3.10) adapted to the present context. We prove generation theorems for Feller semigroups (Theorems 9.35 and 9.50) which form a functional analytic background for the proof of Theorem 1.2 in Chap. 10. In particular, Theorem 9.50 and Corollary 9.51 give useful criteria in terms of *maximum principles*. In Sects. 9.4 and 9.5, following Ventcel' (Wentzell) [We] we study the problem of determining all possible boundary conditions for multi-dimensional diffusion processes. More precisely, we describe analytically the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}$ in the case where the state space is the closure $\overline{D}$ of a bounded domain D in Euclidean space $\mathbf{R}^N$ (Theorems 9.52 and 9.53). Theorems 9.52 and 9.53 are essentially due to Ventcel' [We]. Our proof of these theorems follows Bony–Courrège–Priouret [BCP], where the infinitesimal generators of Feller semigroups are studied in great detail in terms of the maximum principle (see Chap. 8). Analytically, a Markovian particle in $\overline{D}$ is governed by an integro-differential operator W, called a Waldenfels operator, in the interior D of the domain, and it obeys a boundary condition L, called a Ventcel' boundary condition, on the boundary ∂D of the domain. Probabilistically, a Markovian particle moves both by jumps and continuously in the state space and it obeys the Ventcel' boundary condition which consists of six terms corresponding to the diffusion along the boundary, the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary. For the probabilistic meanings of Ventcel' boundary conditions, the reader might refer to Dynkin–Yushkevich [DY].

In this way, we can reduce the problem of existence of Feller semigroups to the unique solvability of the boundary value problem for Waldenfels integro-differential operators W with Ventcel' boundary conditions L in the theory of partial differential equations.

9.1 Markov Processes

In 1828 the English botanist R. Brown observed that pollen grains suspended in water move chaotically, incessantly changing their direction of motion (see Fig. 9.1). The physical explanation of this phenomenon is that a single grain suffers innumerable collisions with the randomly moving molecules of the surrounding water. A mathematical theory for Brownian motion was put forward by A. Einstein [Ei] in 1905. Let $p(t, x, y)$ be the probability density function that a one-dimensional Brownian particle starting at position x will be found at position y at time t. Einstein derived the following formula from statistical mechanical considerations:

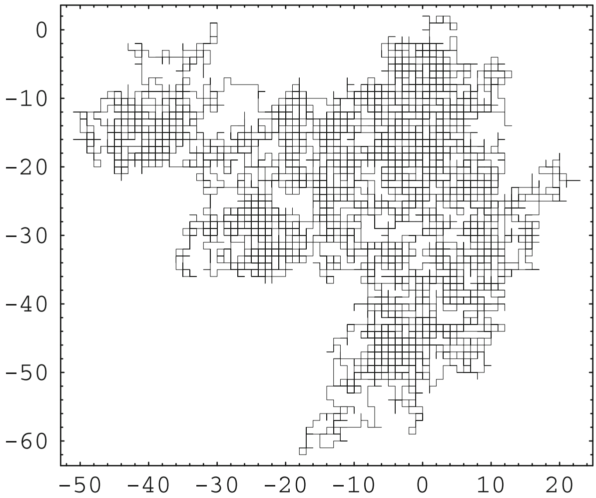


Fig. 9.1 Brownian motion

$$p(t, x, y) = \frac{1}{\sqrt{2\pi Dt}} \exp\left[-\frac{(y-x)^2}{2Dt}\right].$$

Here D is a positive constant determined by the radius of the particle, the interaction of the particle with surrounding molecules, temperature and the Boltzmann constant. This gives an accurate method of measuring Avogadro's number by observing particles. Einstein's theory was experimentally tested by J. Perrin [Pr] between 1906 and 1909.

In Sect. 9.1 we give the precise definition of a Markov transition function adapted to the theory of semigroups (Definition 9.4).

9.1.1 Definitions of Markov Processes

Brownian motion was put on a firm mathematical foundation for the first time by N. Wiener [Wi] in 1923. Let Ω be the space of continuous functions $\omega : [0, \infty) \mapsto \mathbf{R}$ with coordinates $x_t(\omega) = \omega(t)$ and let $\mathcal{F}$ be the smallest σ-algebra in Ω which contains all sets of the form

$$\{\omega \in \Omega : a \leq x_t(\omega) < b\}, \quad t \geq 0, a < b.$$

Wiener constructed probability measures P_x, $x \in \mathbf{R}$, on $\mathcal{F}$ for which the following holds true:

$$\begin{aligned} & P_x\{\omega \in \Omega : a_1 \leq x_{t_1}(\omega) < b_1, a_2 \leq x_{t_2}(\omega) < b_2, \ldots, \\ & \qquad a_n \leq x_{t_n}(\omega) < b_n\} \\ &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \ldots \int_{a_n}^{b_n} p(t_1, x, y_1) p(t_2 - t_1, y_1, y_2) \ldots \\ & \qquad p(t_n - t_{n-1}, y_{n-1}, y_n)\, dy_1\, dy_2 \ldots dy_n, \\ & \quad 0 < t_1 < t_2 < \ldots < t_n < \infty. \end{aligned} \tag{9.1}$$

This formula expresses the "starting afresh" property of Brownian motion that if a Brownian particle reaches a position, then it behaves subsequently as though that position had been its initial position. The measure P_x is called the *Wiener measure* starting at x.

Let $(\Omega, \mathcal{F})$ be a measurable space. A non-negative measure P on $\mathcal{F}$ is called a *probability measure* if $P(\Omega) = 1$. The triple $(\Omega, \mathcal{F}, P)$ is called a *probability space*. The elements of Ω are known as sample points, those of $\mathcal{F}$ as events and the values $P(A)$, $A \in \mathcal{F}$, are their probabilities. An extended real-valued, $\mathcal{F}$-measurable function X on Ω is called a *random variable*. The integral

$$\int_\Omega X(\omega)\, dP$$

(if it exists) is called the *expectation* of X, and is denoted by $E(X)$.

We begin with a review of conditional probabilities and conditional expectations (see Sects. 3.5 and 3.6). Let $\mathcal{G}$ be a σ-algebra contained in $\mathcal{F}$. If X is an integrable random variable, then the *conditional expectation* of X for given $\mathcal{G}$ is any random variable Y which satisfies the following two conditions (CE1) and (CE2):

(CE1) The function Y is $\mathcal{G}$-measurable.
(CE2) $\int_A Y(\omega)\, dP = \int_A X(\omega)\, dP$ for all $A \in \mathcal{G}$.

We recall that conditions (CE1) and (CE2) determine Y up to a set in $\mathcal{G}$ of measure zero. We shall write

$$Y = E(X \mid \mathcal{G}).$$

When X is the characteristic function χ_B of a set $B \in \mathcal{F}$, we shall write

$$P(B \mid \mathcal{G}) = E\left(\chi_B \mid \mathcal{G}\right).$$

The function $P(B \mid \mathcal{G})$ is called the *conditional probability* of B for given $\mathcal{G}$. This function can also be characterized as follows:

(CP1) The function $P(B \mid \mathcal{G})$ is $\mathcal{G}$-measurable.
(CP2) $P(A \cap B) = E(P(B \mid \mathcal{G}); A)$ for every $A \in \mathcal{G}$. That is, we have, for every $A \in \mathcal{G}$,

$$P(A \cap B) = \int_A P(B \mid \mathcal{G})(\omega)\, dP.$$

It should be emphasized that the function $P(B \mid \mathcal{G})$ is determined up to a set in $\mathcal{G}$ of P-measure zero, that is, it is an equivalence class of $\mathcal{G}$-measurable functions on Ω with respect to the measure P.

Markov processes are an abstraction of the idea of Brownian motion. Let K be a locally compact, separable metric space and $\mathcal{B}$ the σ-algebra of all Borel sets in K, that is, the smallest σ-algebra containing all open sets in K. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function X defined on Ω taking values in K is called a *random variable* if it satisfies the condition

$$X^{-1}(E) = \{X \in E\} \in \mathcal{F} \quad \text{for all } E \in \mathcal{B}.$$

We express this by saying that X is $\mathcal{F}/\mathcal{B}$-measurable. A family $\{x_t\}_{t \geq 0}$ of random variables is called a *stochastic process*, and it may be thought of as the motion in time of a physical particle. The space K is called the *state space* and Ω the *sample space*. For a fixed $\omega \in \Omega$, the function $x_t(\omega)$, $t \geq 0$, defines in the state space K a *trajectory* or a *path* of the process corresponding to the sample point ω.

In this generality the notion of a stochastic process is of course not so interesting. The most important class of stochastic processes is the class of Markov processes which is characterized by the Markov property. Intuitively, this is the principle of the lack of any "memory" in the system. Markov processes are an abstraction of the idea of Brownian motion. More precisely, the *temporally homogeneous Markov property* or simply *Markov property* is that the prediction of subsequent motion of a physical particle, knowing its position at time t, depends neither on the value of t nor on what has been observed during the time interval $[0, t)$; that is, a physical particle "starts afresh".

Now we introduce a class of Markov processes which we will deal with in this book. This vague idea can be made precise and effective in several ways.

If $\{Z_\lambda\}_{\lambda \in \Lambda}$ is a family of random variables, we let (see Proposition 2.12)

$$\sigma(Z_\lambda; \lambda \in \Lambda) = \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F}$$
$$\text{with respect to which all } Z_\lambda \text{ are measurable.}$$

If $\{x_t\}_{t \geq 0}$ is a stochastic process, we introduce three sub-σ-algebras $\mathcal{F}_{\leq t}$, $\mathcal{F}_{=t}$ and $\mathcal{F}_{\geq t}$ of $\mathcal{F}$ as follows:

$$\begin{cases} \mathcal{F}_{\le t} &= \sigma(x_s : 0 \le s \le t) \\ &= \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ &\quad \text{with respect to which all } x_s,\ 0 \le s \le t, \text{ are measurable,} \\ \mathcal{F}_{=t} &= \sigma(x_t) \\ &= \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ &\quad \text{with respect to which } x_t \text{ is measurable,} \\ \mathcal{F}_{\ge t} &= \sigma(x_s : t \le s < \infty) \\ &= \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ &\quad \text{with respect to which all } x_s,\ t \le s < \infty, \text{ are measurable.} \end{cases}$$

Intuitively, an event in $\mathcal{F}_{\le t}$ is determined by the behavior of the process $\{x_s\}$ up to time t and an event in $\mathcal{F}_{\ge t}$ by its behavior after time t. Thus they represent respectively the "past" and "future" relative to the "present" moment.

Definition 9.1. A stochastic process $\mathcal{X} = \{x_t\}$ is called a *Markov process* if it satisfies the condition

$$P(B \mid \mathcal{F}_{\le t}) = P(B \mid \mathcal{F}_{=t}) \quad \text{for any "future" set } B \in \mathcal{F}_{\ge t}.$$

More precisely, we have, for any "future" set $B \in \mathcal{F}_{\ge t}$,

$$P(A \cap B) = \int_A P(B \mid \mathcal{F}_{=t})(\omega)\, dP \quad \text{for every "past" set } A \in \mathcal{F}_{\le t}.$$

Intuitively, this means that the conditional probability of a "future" event B given the "present" is the same as the conditional probability of B given the "present" and the "past".

An observer may record not only the trajectories of the process, but also some other occurrences, only indirectly related or entirely unrelated to the process. Thus we obtain a broader and more flexible formulation of the Markov property if we enlarge the "past" as follows.

Let $\{\mathcal{F}_t\}_{t\ge 0}$ be a family of sub-σ-algebras of $\mathcal{F}$ which satisfies the following two conditions (a) and (b):

(a) If $s < t$, then $\mathcal{F}_s \subset \mathcal{F}_t$.
(b) For each $t \ge 0$, the function x_t is $\mathcal{F}_t/\mathcal{B}$-measurable, that is,

$$\{x_t \in E\} \in \mathcal{F}_t \quad \text{for all } E \in \mathcal{B}.$$

We express property (a) by saying that the family $\{\mathcal{F}_t\}$ is increasing, and property (b) by saying that the stochastic process $\{x_t\}$ is *adapted* to $\{\mathcal{F}_t\}$. We remark that the family $\{\mathcal{F}_{\le t}\}_{t\ge 0}$ satisfies both conditions and is the minimal possible one.

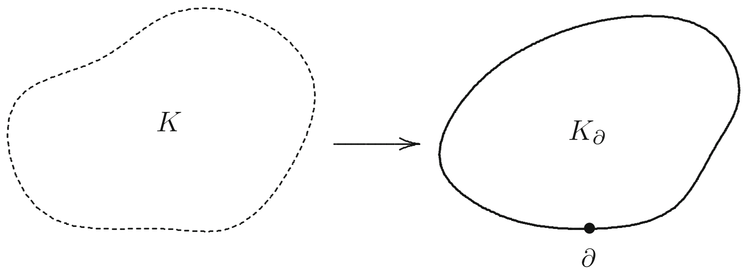


Fig. 9.2 The compactification K_∂ of K

Definition 9.2. Let $\{x_t\}_{t\geq 0}$ be a stochastic process and let $\{\mathcal{F}_t\}_{t\geq 0}$ be an increasing family of sub-σ-algebras of $\mathcal{F}$. We say that $\{x_t\}$ is a *Markov process* with respect to $\{\mathcal{F}_t\}$ if it satisfies the following two conditions (i) and (ii):

(i) $\{x_t\}$ is adapted to $\{\mathcal{F}_t\}$.
(ii) $P(B \mid \mathcal{F}_t) = P(B \mid \mathcal{F}_{=t})$ for all $B \in \mathcal{F}_{\geq t}$.

It should be noted that Definition 9.2 reduces to Definition 9.1 if we take $\mathcal{F}_t := \mathcal{F}_{\leq t}$. Moreover, by choosing the family $\{\mathcal{F}_t\}$ as the "past" has the effect of making it harder for the Markov property to hold true, while the property becomes more powerful.

Now we define a class of (temporally homogeneous) Markov processes which we will deal with in this book:

Definition 9.3. Assume that we are given the following:

(1) A locally compact, separable metric space K and the σ-algebra $\mathcal{B}$ of all Borel sets in K. A point ∂ is adjoined to K as the *point at infinity* if K is not compact, and as an isolated point if K is compact (see Fig. 9.2). We let

$$K_\partial = K \cup \{\partial\},$$

$$\mathcal{B}_\partial = \text{the } \sigma\text{-algebra in } K_\partial \text{ generated by } \mathcal{B}.$$

The space $K_\partial = K \cup \{\partial\}$ is called the *one-point compactification* of K.
(2) The space Ω of all mappings $\omega : [0,\infty] \to K_\partial$ such that $\omega(\infty) = \partial$ and that if $\omega(t) = \partial$ then $\omega(s) = \partial$ for all $s \geq t$. Let ω_∂ be the constant map $\omega_\partial(t) = \partial$ for all $t \in [0,\infty]$.
(3) For each $t \in [0,\infty]$, the coordinate map x_t defined by $x_t(\omega) = \omega(t)$ for all $\omega \in \Omega$.
(4) For each $t \in [0,\infty]$, a pathwise shift mapping $\theta_t : \Omega \to \Omega$ defined by the formula $\theta_t\omega(s) = \omega(t+s)$ for all $\omega \in \Omega$. Note that $\theta_\infty\omega = \omega_\partial$ and that $x_t \circ \theta_s = x_{t+s}$ for all $t, s \in [0,\infty]$.
(5) A σ-algebra $\mathcal{F}$ in Ω and an increasing family $\{\mathcal{F}_t\}_{0\leq t\leq\infty}$ of sub-σ-algebras of $\mathcal{F}$.
(6) For each $x \in K_\partial$, a probability measure P_x on $(\Omega, \mathcal{F})$.

We say that these elements define a *temporally homogeneous Markov process* or simply *Markov process* $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ if the following four conditions (i)–(iv) are satisfied:

(i) For each $0 \leq t < \infty$, the function x_t is $\mathcal{F}_t/\mathcal{B}_\partial$- measurable, that is,

$$\{x_t \in E\} \in \mathcal{F}_t \quad \text{for all } E \in \mathcal{B}_\partial.$$

(ii) For each $0 \leq t < \infty$ and $E \in \mathcal{B}$, the function

$$p_t(x, E) = P_x\{x_t \in E\}$$

is a Borel measurable function of $x \in K$.

(iii) $P_x\{\omega \in \Omega : x_0(\omega) = x\} = 1$ for each $x \in K_\partial$.

(iv) For all $t, h \in [0, \infty]$, $x \in K_\partial$ and $E \in \mathcal{B}_\partial$, we have

$$P_x\{x_{t+h} \in E \mid \mathcal{F}_t\} = p_h(x_t, E) \quad \text{a.e.},$$

or equivalently

$$P_x(A \cap \{x_{t+h} \in E\}) = \int_A p_h(x_t(\omega), E)\, dP_x(\omega) \quad \text{for all } A \in \mathcal{F}_t. \tag{9.2}$$

In this definition, the term ‘Markov process’ means a family of Markov processes over the measure space $(\Omega, \mathcal{F}, P_x)$ with respect to $\{\mathcal{F}\}_t$, one Markov process for each of the measures P_x corresponding to all possible initial positions $x \in K_\partial$.
Here is an intuitive way of thinking about the above definition of a Markov process. The sub-σ-algebra $\mathcal{F}_t$ may be interpreted as the collection of events which are observed during the time interval $[0, t]$. The value $P_x(A)$, $A \in \mathcal{F}$, may be interpreted as the probability of the event A under the condition that a particle starts at position x; hence the value $p_t(x, E)$ expresses the transition probability that a particle starting at position x will be found in the set E at time t (see Fig. 9.3). The function $p_t(x, \cdot)$ is called the *transition function* of the process $\mathcal{X}$. The transition function $p_t(x, \cdot)$ specifies the probability structure of the process. The intuitive meaning of the crucial condition (iv) is that the future behavior of a particle, knowing its history up to time t, is the same as the behavior of a particle starting at $x_t(\omega)$, that is, a particle starts afresh.

By using the Markov property (9.2) repeatedly, we easily obtain the following formula (9.3), analogous to formula (9.1):

$$\begin{aligned} &P_x\{\omega \in \Omega : x_{t_1}(\omega) \in A_1, x_{t_2}(\omega) \in A_2, \ldots, x_{t_n}(\omega) \in A_n\} \\ &= \int_{A_1} \int_{A_2} \cdots \int_{A_n} p_{t_1}(x, dy_1) p_{t_2 - t_1}(y_1, dy_2) \cdots p_{t_n - t_{n-1}}(y_{n-1}, dy_n), \\ &\qquad 0 < t_1 < t_2 < \ldots < t_n < \infty, \quad A_1, A_2, \ldots, A_n \in \mathcal{B}. \end{aligned} \tag{9.3}$$

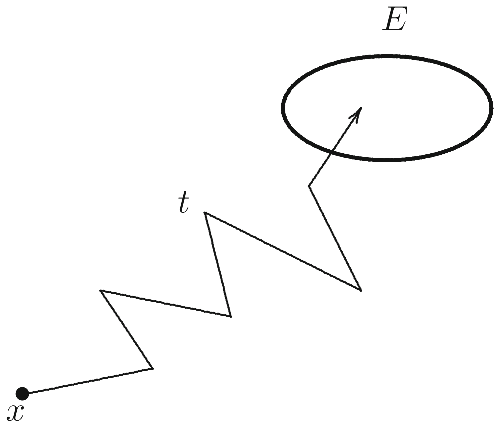


Fig. 9.3 The transition probability $p_t(x, E)$

A Markovian particle moves in the space K until it "dies" or "disappears" at the time when it reaches the point ∂; hence the point ∂ is called the *terminal point* or *cemetery*. With this interpretation in mind, we let

$$\zeta(\omega) = \inf\{t \in [0, \infty] : x_t(\omega) = \partial\}.$$

The random variable ζ is called the *lifetime* of the process $\mathcal{X}$. The process $\mathcal{X}$ is said to be *conservative* if it satisfies the condition

$$P_x\{\zeta = \infty\} = 1 \quad \text{for each } x \in K.$$

9.1.2 Transition Functions

From the point of view of analysis, the transition function is something more convenient than the Markov process itself. In fact, it can be shown that the transition functions of Markov processes generate solutions of certain parabolic partial differential equations such as the classical diffusion equation; and, conversely, these differential equations can be used to construct and study the transition functions and the Markov processes themselves.

Our first job is thus to give the precise definition of a transition function adapted to the theory of semigroups:

Definition 9.4. Let (K, ρ) be a locally compact, separable metric space and $\mathcal{B}$ the σ-algebra of all Borel sets in K.

A function $p_t(x, E)$, defined for all $t \geq 0$, $x \in K$ and $E \in \mathcal{B}$, is called a *temporally homogeneous Markov transition function* on K or simply *Markov transition function* on K if it satisfies the following four conditions (a)–(d):

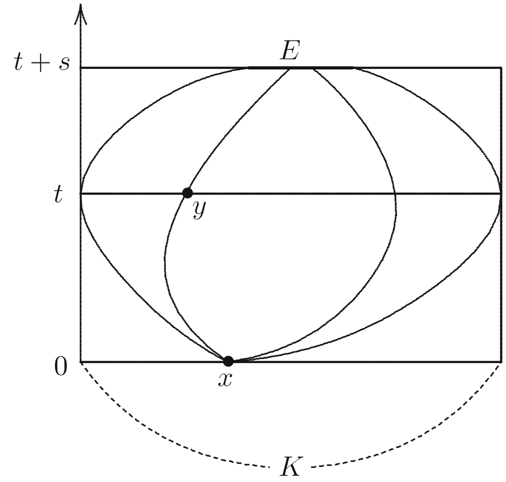


Fig. 9.4 The intuitive meaning of formula (9.4)

(a) $p_t(x, \cdot)$ is a non-negative measure on $\mathcal{B}$ and $p_t(x, K) \leq 1$ for each $t \geq 0$ and each $x \in K$.
(b) $p_t(\cdot, E)$ is a Borel measurable function for each $t \geq 0$ and each $E \in \mathcal{B}$.
(c) $p_0(x, \{x\}) = 1$ for each $x \in K$.
(d) (The Chapman–Kolmogorov equation) For any $t, s \geq 0$, any $x \in K$ and any $E \in \mathcal{B}$, we have

$$p_{t+s}(x, E) = \int_K p_t(x, dy) p_s(y, E). \tag{9.4}$$

It is just condition (d) which reflects the Markov property that a particle starts afresh. Here is an intuitive way of thinking about the above definition of a Markov transition function. The value $p_t(x, E)$ expresses the transition probability that a physical particle starting at position x will be found in the set E at time t. Equation (9.4) expresses the idea that a transition from the position x to the set E in time $t + s$ is composed of a transition from x to some position y in time t, followed by a transition from y to the set E in the remaining time s; the latter transition has probability $p_s(y, E)$ which depends only on y (see Fig. 9.4). Thus a physical particle "starts afresh"; this property is called the *Markov property*.

The Chapman–Kolmogorov equation (9.4) tells us that $p_t(x, K)$ is monotonically increasing as $t \downarrow 0$, so that the limit

$$p_{+0}(x, K) = \lim_{t \downarrow 0} p_t(x, K)$$

exists.

A Markov transition function $p_t(x, \cdot)$ is said to be *normal* if it satisfies the condition

$$p_{+0}(x, K) = \lim_{t \downarrow 0} p_t(x, K) = 1 \quad \text{for every } x \in K.$$

The next theorem justifies the definition of a transition function, and hence it will be fundamental for our further study of Markov processes:

Theorem 9.5. *For every Markov process, the function p_t, defined by the formula*

$$p_t(x, E) = P_x\{x_t \in E\}, \quad x \in K, \ E \in \mathcal{B}, \ t \geq 0,$$

is a Markov transition function. Conversely, every normal Markov transition function corresponds to some Markov process.

Here are some important examples of normal transition functions on the line $\mathbf{R}$ (see Lamperti [La, Chapter 7, Section 8]):

Example 9.6 (uniform motion). If $t \geq 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \chi_E(x + vt),$$

where v is a constant, and $\chi_E(y) = 1$ if $y \in E$ and $= 0$ if $y \notin E$.

This process, starting at x, moves deterministically with constant velocity v.

Example 9.7 (Poisson process). If $t \geq 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \chi_E(x + n),$$

where λ is a positive constant.

This process, starting at x, advances one unit by jumps, and the probability of n jumps during the time 0 and t is equal to $e^{-\lambda t}(\lambda t)^n / n!$.

Example 9.8 (Brownian motion). If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp\left[-\frac{(y - x)^2}{2t}\right] dy,$$

and

$$p_0(x, E) = \chi_E(x).$$

This is a mathematical model of one-dimensional Brownian motion. Its character is quite different from that of the Poisson process; the transition function $p_t(x, E)$ satisfies the condition

$$p_t(x, (x - \varepsilon, x + \varepsilon)) = 1 - o(t)$$

or equivalently,

$$p_t(x, \mathbf{R} \setminus (x-\varepsilon, x+\varepsilon)) = o(t)$$

for every $\varepsilon > 0$ and every $x \in \mathbf{R}$. This means that, unlike the Poisson process, this process never stands still. In fact, this process changes state not by jumps but by *continuous* motion. A Markov process with this property is called a *diffusion process*.

Example 9.9 (Brownian motion with constant drift). If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp\left[-\frac{(y-mt-x)^2}{2t}\right] dy,$$

and

$$p_0(x, E) = \chi_E(x),$$

where m is a constant.

This represents Brownian motion with a constant drift of magnitude m superimposed; the process can be represented as $\{x_t + mt\}$, where $\{x_t\}$ is Brownian motion on $\mathbf{R}$.

Example 9.10 (Cauchy process). If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\pi} \int_E \frac{t}{t^2 + (y-x)^2}\, dy,$$

and

$$p_0(x, E) = \chi_E(x).$$

This process can be thought of as the "trace" on the real line of trajectories of two-dimensional Brownian motion, and it moves by jumps (see Knight [Kn, Lemma 2.12]). More precisely, if $B_1(t)$ and $B_2(t)$ are two independent Brownian motions and if T is the first passage time of $B_1(t)$ to x, then $B_2(T)$ has the Cauchy density

$$\frac{1}{\pi}\frac{|x|}{x^2 + y^2}, \quad -\infty < y < \infty.$$

Here are two examples of diffusion processes on the closed half-line $K = \overline{\mathbf{R}^+} = [0, \infty)$ in which we must take account of the effect of the boundary point 0 of K (see [La, Chapter 7, Section 8]):

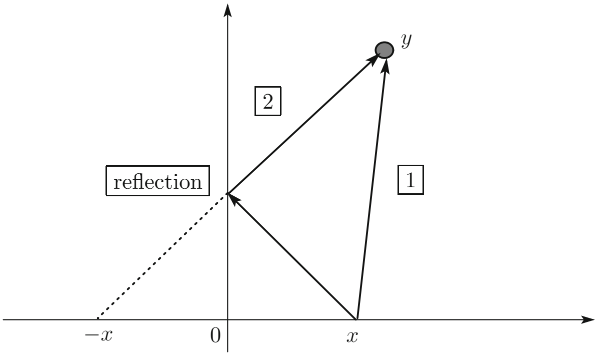


Fig. 9.5 The reflecting barrier

Example 9.11 (reflecting barrier Brownian motion). If $t > 0$, $x \in K = [0, \infty)$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \left(\int_E \exp\left[-\frac{(y-x)^2}{2t} \right] dy + \int_E \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right), \tag{9.5}$$

and

$$p_0(x, E) = \chi_E(x).$$

This represents Brownian motion with a reflecting barrier at $x = 0$; the process may be represented as $\{|x_t|\}$, where $\{x_t\}$ is Brownian motion on $\mathbf{R}$. Indeed, since $\{|x_t|\}$ goes from x to y if $\{x_t\}$ goes from x to $\pm y$ due to the symmetry of the transition function in Example 9.8 about $x = 0$, we find that (see Fig. 9.5)

$$\begin{aligned} p_t(x, E) &= P_x\{|x_t| \in E\} \\ &= \frac{1}{\sqrt{2\pi t}} \left(\int_E \exp\left[-\frac{(y-x)^2}{2t} \right] dy + \int_E \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right). \end{aligned}$$

Example 9.12 (sticking barrier Brownian motion). If $t > 0$, $x \in K = [0, \infty)$ and $E \in \mathcal{B}$, we let

$$\begin{aligned} p_t(x, E) = {} & \frac{1}{\sqrt{2\pi t}} \left(\int_E \exp\left[-\frac{(y-x)^2}{2t} \right] dy - \int_E \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right) \\ & + \left(1 - \frac{1}{\sqrt{2\pi t}} \int_{-x}^{x} \exp\left[-\frac{z^2}{2t} \right] dz \right) \chi_E(0), \end{aligned}$$

and

$$p_0(x, E) = \chi_E(x).$$

This represents Brownian motion with a sticking barrier at $x = 0$. When a Brownian particle reaches the boundary point 0 for the first time, instead of reflecting it sticks there forever; in this case the state 0 is called a *trap*.

Here is a typical example of diffusion processes on the closed interval $K = [0, 1]$ in which we must take account of the effect of the two boundary points 0 and 1 of K:

Example 9.13 (reflecting barrier Brownian motion). If $t > 0$, $x \in K = [0, 1]$ and $E \in \mathcal{B}$, we let

$$\begin{aligned} &p_t(x, E) \\ &= \frac{1}{\sqrt{2\pi t}} \int_E \left(\sum_{n=-\infty}^{\infty} \left[\exp\left[-\frac{(y-x+2n)^2}{2t} \right] + \exp\left[-\frac{(y+x+2n)^2}{2t} \right] \right] \right) dy, \end{aligned}$$

and

$$p_0(x, E) = \chi_E(x).$$

This represents Brownian motion with two reflecting barriers at $x = 0$ and $x = 1$.

It has been assumed so far that $p_t(x, K) \leq 1$ for each $t \geq 0$ and each $x \in K$. This implies that a Markovian particle may die or disappear in a finite time.

Here are three typical examples of absorbing barrier Brownian motion.

Example 9.14 (absorbing barrier Brownian motion). If $t > 0$, $x \in K = [0, \infty)$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \left(\int_E \exp\left[-\frac{(y-x)^2}{2t} \right] dy - \int_E \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right), \tag{9.6}$$

and

$$p_0(x, E) = \chi_E(x).$$

This represents Brownian motion with an absorbing barrier at $x = 0$; a Brownian particle dies at the first moment when it hits the boundary point $x = 0$ (see Fig. 9.6). That is, the boundary point 0 of K is the terminal point.

Example 9.15 (absorbing barrier Brownian motion). If $t > 0$, $x \in K = [0, 1]$ and $E \in \mathcal{B}$, we let

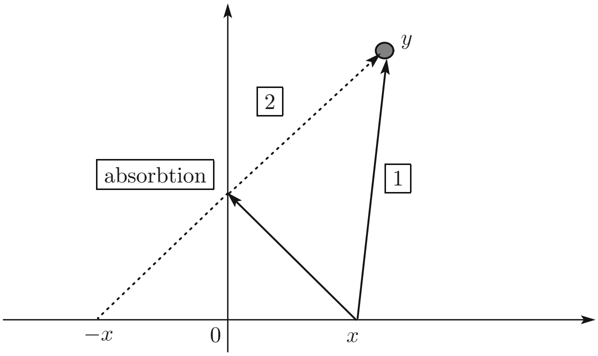


Fig. 9.6 The absorbing barrier

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \left(\sum_{n=-\infty}^{\infty} \left[\exp\left[-\frac{(y-x+2n)^2}{2t} \right] - \exp\left[-\frac{(y+x+2n)^2}{2t} \right] \right] \right) dy,$$

and

$$p_0(x, E) = \chi_E(x).$$

This represents Brownian motion with two absorbing barriers at $x = 0$ and $x = 1$.

Example 9.16 (absorbing–reflecting barrier Brownian motion). Let λ be a constant such that $0 < \lambda < 1$. If $t > 0$, $x \in K = [0, \infty)$ and $E \in \mathcal{B}$, we let

$$\begin{aligned} &p_t(x, E) \\ &= \frac{1}{\sqrt{2\pi t}} \left(\int_E \exp\left[-\frac{(y-x)^2}{2t} \right] dy + \int_E \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right) \\ &\quad - \frac{1}{\sqrt{2\pi t}} \left(\frac{2(1-\lambda)}{\lambda} \right) \\ &\quad \times \int_E \exp\left[\frac{(1-\lambda)}{\lambda} y \right] \left(\int_{-\infty}^{-y} \exp\left[\frac{(1-\lambda)}{\lambda} z \right] \exp\left[-\frac{(z-x)^2}{2t} \right] dz \right) dy, \end{aligned} \tag{9.7}$$

and

$$p_0(x, E) = \chi_E(x).$$

This process $\{x_t\}$ may be thought of as a "combination" of the absorbing and reflecting Brownian motions; the absorbing and reflecting cases are formally obtained by letting $\lambda \to 0$ and $\lambda \to 1$, respectively. In fact, it is easy to verify that formulas (9.6) and (9.5) may be obtained from formula (9.7) by letting $\lambda \to 0$ and $\lambda \to 1$, respectively.

A Markov transition function $p_t(x, \cdot)$ is said to be *conservative* if it satisfies the condition that we have, for all $t > 0$,

$$p_t(x, K) = 1 \quad \text{for each } x \in K.$$

For example, the reflecting barrier Brownian motion of Example 9.11 is conservative. Indeed, it suffices to note that we have, for all $t > 0$,

$$\begin{aligned} p_t(x, [0, \infty)) &= \frac{1}{\sqrt{2\pi t}} \left(\int_0^\infty \exp\left[-\frac{(y-x)^2}{2t} \right] dy + \int_0^\infty \exp\left[-\frac{(y+x)^2}{2t} \right] dy \right) \\ &= \frac{1}{\sqrt{2\pi t}} \left(\int_{-\infty}^\infty \exp\left[-\frac{(y-x)^2}{2t} \right] dy \right) \\ &= 1 \quad \text{for every } x \in K = [0, \infty). \end{aligned}$$

There is a simple trick which allows us to turn the general case into the conservative case. We add a new point ∂ to the locally compact space K as the point at infinity if K is not compact, and as an isolated point if K is compact; so the space $K_\partial = K \cup \{\partial\}$ is compact. Then we can extend a Markov transition function $p_t(x, \cdot)$ on K to a Markov transition function $p_t'(x, \cdot)$ on K_∂ by the formulas

$$\begin{cases} p_t'(x, E) = p_t(x, E), & x \in K, \quad E \in \mathcal{B}; \\ p_t'(x, \{\partial\}) = 1 - p_t(x, K), & x \in K; \\ p_t'(\partial, K) = 0, \quad p_t'(\partial, \{\partial\}) = 1. \end{cases}$$

Intuitively, this means that a Markovian particle moves in the space K until it dies at which time it reaches the point ∂; hence the point ∂ is the terminal point or cemetery.

In the sequel, we will not distinguish in our notation between $p_t(x, \cdot)$ and $p_t'(x, \cdot)$; in the cases of interest for us the point ∂ will be absorbing.

9.1.3 Kolmogorov's Equations

In the first works devoted to Markov processes, the most fundamental was A.N. Kolmogorov's work (1931) where the general concept of a Markov transition function was introduced for the first time and an analytic method of describing Markov transition functions was proposed.

We now take a close look at Kolmogorov's work (see Lamperti [La, Chapter 6, Section 5]). Let $p_t(x, \cdot)$ be a transition function on $\mathbf{R}$, and assume that the following two conditions (i) and (ii) are satisfied:

(i) For each $\varepsilon > 0$, we have

$$\lim_{t \downarrow 0} \frac{1}{t} \sup_{x \in \mathbf{R}} p_t(x, \mathbf{R} \setminus (x - \varepsilon, x + \varepsilon)) = 0.$$

(ii) The limits

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{x-\varepsilon}^{x+\varepsilon} p_t(x, dy)(y - x)^2 = a(x),$$

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{x-\varepsilon}^{x+\varepsilon} p_t(x, dy)(y - x) = b(x),$$

$$\lim_{t \downarrow 0} \frac{1}{t} (p_t(x, \mathbf{R}) - 1) = c(x)$$

exist for each $x \in \mathbf{R}$.

Physically, the limit $a(x)$ may be interpreted as a variance instantaneous (with respect to t) velocity at position x, and the limit $b(x)$ has a similar interpretation as a mean. The transition functions in Examples 9.6, 9.8 and 9.9 satisfy conditions (i) and (ii) with $a(x) = 0$, $b(x) = v$, $c(x) = 0$; $a(x) = 1$, $b(x) = c(x) = 0$; $a(x) = 1$, $b(x) = m$, $c(x) = 0$, respectively, whereas the transition functions in Examples 9.7 and 9.10 do not satisfy condition (i).

Furthermore, we assume that the transition function $p_t(x, \cdot)$ has a density $p(t, x, y)$ with respect to the Lebesgue measure dy. Intuitively, the density $p(t, x, y)$ represents the state of the process at position y at time t, starting at the initial state that a unit mass is at position x. Under certain regularity conditions, Kolmogorov showed that the density $p(t, x, y)$ is, for fixed y, the fundamental solution of the Cauchy problem

$$\begin{cases} \dfrac{\partial p}{\partial t} = \dfrac{a(x)}{2} \dfrac{\partial^2 p}{\partial x^2} + b(x) \dfrac{\partial p}{\partial x} + c(x) p, \quad t > 0. \\ \lim_{t \downarrow 0} p(t, x, y) = \delta_y(x), \end{cases} \tag{9.8}$$

and is, for fixed x, the fundamental solution of the Cauchy problem

$$\begin{cases} \frac{\partial p}{\partial t} = \frac{\partial^2}{\partial y^2} \left(\frac{a(y)}{2} \right) - \frac{\partial}{\partial y} (b(y) p) + c(y) p, \quad t > 0. \\ \lim_{t \downarrow 0} p(t, x, y) = \delta_x(y). \end{cases} \tag{9.9}$$

Here δ is the Dirac measure (see Example 5.11), and δ_y and δ_x represent unit masses at position y and x, respectively. Equation (9.8) is called *Kolmogorov's backward equation*, since we consider the terminal state (the variable y) to be fixed and vary the initial state (the variable x). In this context, Eq. (9.9) is called *Kolmogorov's forward equation*. These equations are also called the *Fokker–Planck* *partial differential equations*. In the case of Brownian motion (Example 9.8), Eqs. (9.8) and (9.9) become the classical diffusion (or heat) equations

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial x^2}, \quad t > 0.$$
$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2}, \quad t > 0.$$

Conversely, Kolmogorov raised the problem of construction of Markov transition functions by solving the given Fokker–Planck partial differential equations (9.8) and (9.9).

It is worth pointing out here that the forward equation (9.9) is given in a more intuitive form than the backward equation (9.8), but regularity conditions on the functions $a(y)$ and $b(y)$ are more stringent than those needed in the backward case. This suggests that the backward approach is more convenient than the forward approach from the viewpoint of analysis.

In 1936, W. Feller treated this problem by classical analytic methods, and proved that Eq. (9.8) (or (9.9)) has a unique solution $p(t, x, y)$ under certain regularity conditions on the functions $a(x)$, $b(x)$ and $c(x)$, and that this solution $p(t, x, y)$ determines a Markov process. In 1943, R. Fortet proved that these solutions correspond to Markov processes with continuous paths.

On the other hand, S.N. Bernstein [Be] (1938) and P. Lévy [Le] (1948) made probabilistic approaches to this problem, by using stochastic differential equations.

9.1.4 Feller and C_0 Transition Functions

Let (K, ρ) be a locally compact, separable metric space and $\mathcal{B}$ the σ-algebra of all Borel sets in K. Let $B(K)$ be the space of real-valued, bounded Borel measurable functions on K; $B(K)$ is a Banach space with the supremum norm

$$\|f\|_\infty = \sup_{x\in K} |f(x)|.$$

If p_t is a transition function on K, we let

$$T_t f(x) = \int_K p_t(x, dy) f(y) \quad \text{for every } f \in B(K).$$

Then, by applying Theorem 2.7 with

$$\mathcal{F} := \mathcal{B},$$
$$\mathcal{H} := \{f \in B(K) : T_t f \text{ is Borel measurable}\},$$

we obtain that $\mathcal{H} = B(K)$, that is, the function $T_t f$ is Borel measurable whenever $f \in B(K)$. Indeed, it suffices to note the following two facts (i) and (ii):

(i) Condition (b) of Definition 9.4 implies condition (i) of Theorem 2.7.
(ii) An application of the monotone convergence theorem [Fo2, Theorem 2.14] gives that condition (ii) of Theorem 2.7 is satisfied.

In view of condition (a) of Definition 9.4, it follows that, for each $t \geq 0$, the operator T_t is non-negative and contractive on $B(K)$ into itself:

$$f \in B(K),\ 0 \leq f(x) \leq 1 \quad \text{on } K \Longrightarrow 0 \leq T_t f(x) \leq 1 \quad \text{on } K.$$

Furthermore, we have, by condition (d) of Definition 9.4 and Fubini's theorem,

$$\begin{aligned} T_{t+s} f(x) &= \int_K p_{t+s}(x, dy) f(y) = \int_K \int_K p_t(x, dz) p_s(z, dy) f(y) \\ &= \int_K p_t(x, dz) \left(\int_K p_s(z, dy) f(y) \right) = \int_K p_t(x, dz) T_s f(z) \\ &= T_t(T_s f)(x), \end{aligned}$$

so that the operators T_t form a *semigroup*

$$T_{t+s} = T_t \cdot T_s, \quad t, s \geq 0.$$

We also have, by condition (c) of Definition 9.4,

$$T_0 = I = \text{the identity operator.}$$

The Hille–Yosida theory of semigroups requires the strong continuity of $\{T_t\}_{t \geq 0}$:

$$\lim_{t \downarrow 0} \|T_t f - f\|_\infty = 0 \quad \text{for every } f \in B(K), \tag{9.10}$$

that is,

$$\lim_{t \downarrow 0} \sup_{x \in K} \left| \int_K p_t(x, dy) f(y) - f(x) \right| = 0 \quad \text{for every } f \in B(K). \tag{9.10$'$}$$

Now, by taking $f := \chi_{\{x\}} \in B(K)$ in (9.10′), we obtain that

$$\lim_{t \downarrow 0} p_t(x, \{x\}) = 1 \quad \text{for every } x \in K. \tag{9.11}$$

However, the Brownian motion transition function in Example 9.8, the most important and interesting example, does not satisfy condition (9.11). Thus we shift our attention to continuous functions on K, instead of Borel measurable functions on K.

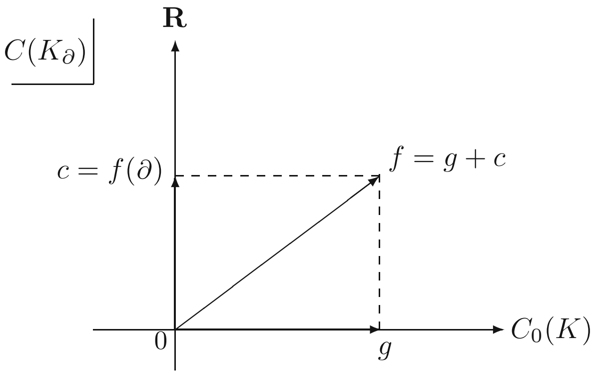


Fig. 9.7 The spaces $C(K_\partial)$ and $C_0(K)$

Now let $C(K)$ be the space of real-valued, bounded continuous functions on K; $C(K)$ is a normed linear space with the supremum norm

$$\|f\|_\infty = \sup_{x \in K} |f(x)|.$$

We add a new point ∂ to the locally compact space K as the point at infinity if K is not compact, and as an isolated point if K is compact; so the space $K_\partial = K \cup \{\partial\}$ is compact. Then we say that a function $f \in C(K)$ converges to $a \in \mathbf{R}$ as $x \to \partial$ (see Sect. 4.4.1) if, for each $\varepsilon > 0$, there exists a compact subset E of K such that

$$|f(x) - a| < \varepsilon \quad \text{for all } x \in K \setminus E.$$

We shall write

$$\lim_{x \to \partial} f(x) = a.$$

Let $C_0(K)$ be the subspace of $C(K)$ which consists of all functions satisfying the condition $\lim_{x \to \partial} f(x) = 0$; $C_0(K)$ is a closed subspace of $C(K)$. We remark that $C_0(K)$ may be identified with $C(K)$ if K is compact.

Moreover, we introduce a useful convention

Any real-valued function f on K is extended to the compact space $K_\partial = K \cup \{\partial\}$ by setting $f(\partial) = 0$.

From this point of view, the space $C_0(K)$ is identified with the subspace of $C(K_\partial)$ which consists of all functions f satisfying the condition $f(\partial) = 0$. More precisely, it should be emphasized that (see Fig. 9.7)

$$C(K_\partial) = \{\text{constant functions}\} + C_0(K).$$

Now we can introduce two important conditions on the measures $p_t(x, \cdot)$ related to continuity in $x \in K$, for fixed $t \geq 0$:

Definition 9.17. A transition function p_t is called a *Feller transition function* or simply *Feller function* if the function

$$T_t f(x) = \int_K p_t(x, dy) f(y)$$

is a continuous function of $x \in K$ whenever $f(x)$ is bounded and continuous on K. In other words, the Feller property is equivalent to saying that the space $C(K)$ is an invariant subspace of $B(K)$ for the operators T_t. Moreover, we say that p_t is a C_0 *transition function* if the space $C_0(K)$ is an invariant subspace of $C(K)$ for the operators T_t:

$$f \in C_0(K) \Longrightarrow T_t f \in C_0(K).$$

Remark 9.18. The Feller property is equivalent to saying that the measures $p_t(x, \cdot)$ depend continuously on $x \in K$ in the usual weak topology, for every fixed $t \geq 0$.

9.1.5 Path Functions of Markov Processes

It is naturally interesting and important to ask the following question:

Question 9.19. Given a Markov transition function p_t, under which conditions on p_t does there exist a Markov process with transition function p_t whose paths are almost surely continuous?

A Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ is said to be *right-continuous* provided that, for each $x \in K$,

$$P_x\{\omega \in \Omega : \text{the mapping } t \mapsto x_t(\omega) \text{ is a right-continuous function from } [0, \infty) \text{ into } K_\partial\} = 1.$$

Furthermore, we say that $\mathcal{X}$ is *continuous* provided that, for each $x \in K$,

$$P_x\{\omega \in \Omega : \text{the mapping } t \mapsto x_t(\omega) \text{ is a continuous function from } [0, \zeta) \text{ into } K_\partial\} = 1.$$

Here ζ is the lifetime of the process $\mathcal{X}$.

Now we give some useful criteria for path-continuity in terms of transition functions:

Theorem 9.20. *Let K be a locally compact, separable metric space and let p_t be a normal transition function on K.*

(i) Assume that the following two conditions (L) and (M) are satisfied:

(L) For each $s > 0$ and each compact $E \subset K$, we have

$$\lim_{x \to \partial} \sup_{0 \le t \le s} p_t(x, E) = 0.$$

(M) For each $\varepsilon > 0$ and each compact $E \subset K$, we have

$$\lim_{t \downarrow 0} \sup_{x \in E} p_t(x, K \setminus U_\varepsilon(x)) = 0,$$

where $U_\varepsilon(x) = \{y \in K : \rho(y, x) < \varepsilon\}$ is an ε-neighborhood of x.

Then there exists a Markov process $\mathcal{X}$ with transition function p_t whose paths are right-continuous on $[0, \infty)$ and have left-hand limits on $[0, \zeta)$ almost surely.

(ii) Assume that condition (L) and the following condition (N) (replacing condition (M)) are satisfied:

(N) For each $\varepsilon > 0$ and each compact $E \subset K$, we have

$$\lim_{t \downarrow 0} \frac{1}{t} \sup_{x \in E} p_t(x, K \setminus U_\varepsilon(x)) = 0.$$

Then there exists a Markov process $\mathcal{X}$ with transition function p_t whose paths are almost surely continuous on $[0, \zeta)$.

Remark 9.21. 1. Condition (L) is trivially satisfied, if the state space K is compact.
2. It is known (see Dynkin [Dy1, Lemma 6.2]) that if the paths of a Markov process are right-continuous, then the transition function p_t satisfies the condition

$$\lim_{t \downarrow 0} p_t(x, U_\varepsilon(x)) = 1 \quad \text{for all } x \in K.$$

9.1.6 Stopping Times

In this subsection we formulate the starting afresh property for suitable random times τ, that is, the events $\{\omega \in \Omega : \tau(\omega) < a\}$ should depend on the process $\{x_t\}$ only "up to time a", but not on the "future" after time a. This idea leads us to the following definition:

Definition 9.22. Let $\{\mathcal{F}_t : t \ge 0\}$ be an increasing family of σ-algebras in a probability space $(\Omega, \mathcal{F}, P)$. A mapping $\tau : \Omega \to [0, \infty]$ is called a *stopping time* or *Markov time* with respect to $\{\mathcal{F}_t\}$ if it satisfies the condition

$$\{\tau < a\} = \{\omega \in \Omega : \tau(\omega) < a\} \in \mathcal{F}_a \quad \text{for all } a > 0. \tag{9.12}$$

If we introduce another condition

$$\{\tau \leq a\} = \{\omega \in \Omega : \tau(\omega) \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0, \tag{9.13}$$

then condition (9.13) implies condition (9.12); hence we obtain a smaller family of stopping times. Indeed, we have, for all $a > 0$,

$$\begin{aligned}\{\tau < a\} = \{\omega \in \Omega : \tau(\omega) < a\} &= \bigcup_{n=1}^{\infty} \left\{\omega \in \Omega : \tau(\omega) \leq a - \frac{1}{n}\right\} \\ &= \bigcup_{n=1}^{\infty} \left\{\tau \leq a - \frac{1}{n}\right\},\end{aligned}$$

and it follows from condition (9.13) that each set in the union belongs to $\mathcal{F}_{a-1/n}$. Hence we obtain from the monotonicity of $\{\mathcal{F}_t\}$ that

$$\{\tau < a\} \in \bigcup_{n=1}^{\infty} \mathcal{F}_{a-1/n} \subset \mathcal{F}_a \quad \text{for all } a > 0.$$

This proves that condition (9.12) is satisfied.

Conversely, we can prove the following lemma:

Lemma 9.23. *Assume that the family* $\{\mathcal{F}_t\}$ *is* right-continuous*, that is,*

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s \quad \textit{for each } t \geq 0.$$

Then condition (9.12) *implies condition* (9.13)*.*

Proof. First, we have, for all $a > 0$,

$$\begin{aligned}\{\tau \leq a\} = \{\omega \in \Omega : \tau(\omega) \leq a\} &= \bigcap_{n=1}^{\infty} \left\{\omega \in \Omega : \tau(\omega) < a + \frac{1}{n}\right\} \\ &= \bigcap_{n=1}^{\infty} \left\{\tau < a + \frac{1}{n}\right\}.\end{aligned}$$

However, it follows from condition (9.12) that each set in the intersection belongs to $\mathcal{F}_{a+1/n}$. Hence we obtain from the right-continuity of $\{\mathcal{F}_t\}$ that

$$\{\tau \leq a\} \in \bigcap_{n=1}^{\infty} \mathcal{F}_{a+1/n} = \mathcal{F}_a \quad \text{for all } a > 0.$$

This proves that condition (9.13) is satisfied.

The proof of Lemma 9.23 is complete.

Summing up, we have proved that conditions (9.12) and (9.13) are equivalent provided that the family $\{\mathcal{F}_t\}$ is right-continuous.

If τ is a stopping time with respect to the right-continuous family $\{\mathcal{F}_t\}$ of σ-algebras, we let

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0\}\,.$$

Intuitively, we may think of $\mathcal{F}_\tau$ as the "past" up to the random time τ. Then we have the following lemma:

Lemma 9.24. *$\mathcal{F}_\tau$ is a σ-algebra.*

Proof. (1) It is clear that $\emptyset \in \mathcal{F}_\tau$.

(2) If $A \in \mathcal{F}_\tau$, then we have, by condition (9.13),

$$A^c \cap \{\tau \leq a\} = \{\tau \leq a\} \setminus (A \cap \{\tau \leq a\}) \in \mathcal{F}_a \quad \text{for all } a > 0.$$

This proves that $A^c \in \mathcal{F}_\tau$.

(3) If $A_k \in \mathcal{F}_\tau$ for $k = 1, 2, \ldots$, then we have, by condition (9.13),

$$\left(\bigcup_{k=1}^{\infty} A_k\right) \cap \{\tau \leq a\} = \bigcup_{k=1}^{\infty} (A_k \cap \{\tau \leq a\}) \in \mathcal{F}_a \quad \text{for all } a > 0.$$

This proves that $\cup_{k=1}^{\infty} A_k \in \mathcal{F}_\tau$.

The proof of Lemma 9.24 is complete.

Now we list some elementary properties of stopping times and their associated σ-algebras:

(i) Any non-negative constant mapping is a stopping time. More precisely, if $\tau \equiv t_0$ for some constant $t_0 \geq 0$, then it follows that τ is a stopping time and that $\mathcal{F}_\tau$ reduces to $\mathcal{F}_{t_0}$.

Proof. Since we have

$$\{\tau \equiv t_0 \leq a\} = \begin{cases} \emptyset \in \mathcal{F}_a & \text{if } 0 < a < t_0, \\ \Omega \in \mathcal{F}_a & \text{if } a \geq t_0, \end{cases}$$

it follows that τ is a stopping time and further from the right-continuity of $\{\mathcal{F}_t\}$ that

$$\begin{aligned} \mathcal{F}_\tau &= \{A \in \mathcal{F} : A \cap \{\tau \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0\} \\ &= \{A \in \mathcal{F} : A \in \mathcal{F}_a \quad \text{for all } a \geq t_0\} \\ &= \bigcap_{a \geq t_0} \mathcal{F}_a = \mathcal{F}_{t_0}. \end{aligned}$$

The proof of Assertion (i) is complete.

(ii) If $\{\tau_n\}$ is a finite or denumerable collection of stopping times for the family $\{\mathcal{F}_t\}$, then it follows that

$$\tau = \inf_n \tau_n$$

is also a stopping time.

Proof. Since each τ_n is a stopping time, we have, for all $a > 0$,

$$\{\tau = \inf_n \tau_n < a\} = \bigcup_n \{\tau_n < a\} \in \mathcal{F}_a.$$

Indeed, it suffices to note that each set in the union belongs to $\mathcal{F}_a$.

(iii) If $\{\tau_n\}$ is a finite or denumerable collection of stopping times for the family $\{\mathcal{F}_t\}$, then it follows that

$$\tau = \sup_n \tau_n$$

is also a stopping time.

Proof. Since each τ_n is a stopping time and since $\{\mathcal{F}_t\}$ is increasing, we have, for all $a > 0$,

$$\{\tau = \sup_n \tau_n < a\} = \bigcup_{k=1}^{\infty} \bigcap_n \left\{\tau_n < a - \frac{1}{k}\right\} \in \bigcup_{k=1}^{\infty} \mathcal{F}_{a-1/k} \subset \mathcal{F}_a.$$

Indeed, it suffices to note that each set in the intersection belongs to $\mathcal{F}_{a-1/k}$.

(iv) If τ is a stopping time and t_0 is a positive constant, then it follows that $\tau + t_0$ is also a stopping time.

Proof. Since the stopping time τ is non-negative, we have, by the monotonicity of $\{\mathcal{F}_t\}$,

$$\begin{aligned}\{\tau + t_0 < a\} &= \{\tau < a - t_0\} \\ &= \begin{cases} \emptyset \in \mathcal{F}_a & \text{if } 0 < a \leq t_0, \\ \{\tau < a - t_0\} \in \mathcal{F}_{a-t_0} \subset \mathcal{F}_a & \text{if } a > t_0. \end{cases}\end{aligned}$$

This proves that $\tau + t_0$ is a stopping time.

(v) Let τ_1 and τ_2 be stopping times for the family $\{\mathcal{F}_t\}$ such that $\tau_1 \leq \tau_2$ on Ω. Then it follows that

$$\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}.$$

This is a generalization of the monotonicity of the family $\{\mathcal{F}_t\}$.

Proof. If A is an arbitrary element of $\mathcal{F}_{\tau_1}$, then it satisfies the condition

$$A \cap \{\tau_1 \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0.$$

Since we have $\{\tau_2 \leq a\} \subset \{\tau_1 \leq a\}$ for all $a > 0$, it follows that

$$A \cap \{\tau_2 \leq a\} = (A \cap \{\tau_1 \leq a\}) \cap \{\tau_2 \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0. \tag{9.14}$$

This proves that $A \in \mathcal{F}_{\tau_2}$.

(vi) Let $\{\tau_n\}_{n=1}^{\infty}$ be a sequence of stopping times for the family $\{\mathcal{F}_t\}$ such that $\tau_{n+1} \leq \tau_n$ on Ω. Then it follows that the limit

$$\tau = \lim_{n\to\infty} \tau_n = \inf_{n\geq 1} \tau_n$$

is a stopping time and further that

$$\mathcal{F}_\tau = \bigcap_{n\geq 1} \mathcal{F}_{\tau_n}.$$

This property generalizes the right-continuity of the family $\{\mathcal{F}_t\}$.

Proof. First, by assertion (ii) it follows that τ is a stopping time. Moreover, we have, for each $n = 1, 2, \ldots$,

$$\tau = \inf_{k\geq 1} \tau_k \leq \tau_n.$$

Hence it follows from assertion (v) that

$$\mathcal{F}_\tau \subset \mathcal{F}_{\tau_n} \quad \text{for each } n = 1, 2, \ldots,$$

so that

$$\mathcal{F}_\tau \subset \bigcap_{n\geq 1} \mathcal{F}_{\tau_n}.$$

Conversely, let A be an arbitrary element of $\cap_{n\geq 1}\mathcal{F}_{\tau_n}$. Then it follows that, for each $n = 1, 2, \ldots$,

$$A \cap \{\tau_n \leq a\} \in \mathcal{F}_a \quad \text{for all } a > 0.$$

However, since $\tau_n \downarrow \tau$ as $n \to \infty$, we have

$$\{\tau \leq a\} = \bigcap_{m\in\mathbf{N}} \bigcup_{n\in\mathbf{N}} \left\{\tau_n \leq a + \frac{1}{m}\right\}.$$

Hence we obtain from assertion (9.14) that

$$A \cap \{\tau \leq a\} = A \bigcap \left(\bigcap_{m \in \mathbf{N}} \bigcup_{n \in \mathbf{N}} \left\{ \tau_n \leq a + \frac{1}{m} \right\} \right)$$
$$= \bigcap_{m \in \mathbf{N}} \bigcup_{n \in \mathbf{N}} \left[A \cap \left\{ \tau_n \leq a + \frac{1}{m} \right\} \right],$$

where each member in the union belongs to $\mathcal{F}_{a+1/m}$.

Therefore, it follows from the right-continuity of $\{\mathcal{F}_t\}$ that

$$A \cap \{\tau \leq a\} \in \bigcap_{m \in \mathbf{N}} \mathcal{F}_{a+\frac{1}{m}} = \mathcal{F}_a \quad \text{for all } a > 0.$$

This proves that $A \in \mathcal{F}_\tau$.

The proof of Assertion (vi) is complete.

9.1.7 Definition of Strong Markov Processes

A Markov process is called a *strong Markov process* if the "starting afresh" property holds not only for every fixed moment but also for suitable random times. In this subsection we formulate this "strong" Markov property precisely (Definition 9.25), and give a useful criterion for the strong Markov property (Theorem 9.26).

Let (K, ρ) be a locally compact, separable metric space. We add a new point ∂ to the locally compact space K as the point at infinity if K is not compact, and as an isolated point if K is compact; so the space $K_\partial = K \cup \{\partial\}$ is compact.

Let $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ be a Markov process. For each $t \in [0, \infty]$, we define a mapping

$$\Phi_t : [0, t] \times \Omega \longrightarrow K_\partial$$

by the formula

$$\Phi_t(s, \omega) = x_s(\omega).$$

A Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ is said to be *progressively measurable* with respect to $\{\mathcal{F}_t\}$ if the mapping Φ_t is $\mathcal{B}_{[0,t]} \times \mathcal{F}_t / \mathcal{B}_\partial$-measurable for each $t \in [0, \infty]$, that is, if we have the condition

$$\Phi_t^{-1}(E) = \{\Phi_t \in E\} \in \mathcal{B}_{[0,t]} \times \mathcal{F}_t \quad \text{for all } E \in \mathcal{B}_\partial.$$

Here $\mathcal{B}_{[0,t]}$ is the σ-algebra of all Borel sets in the interval $[0, t]$ and $\mathcal{B}_\partial$ is the σ-algebra in K_∂ generated by $\mathcal{B}$. It should be noted that if $\mathcal{X}$ is progressively measurable and if τ is a stopping time, then the mapping $x_\tau : \omega \mapsto x_{\tau(\omega)}(\omega)$ is $\mathcal{F}_\tau/\mathcal{B}_\partial$- measurable.

Definition 9.25. We say that a progressively measurable Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ has the *strong Markov property* with respect to $\{\mathcal{F}_t\}$ if the following condition is satisfied:

For all $h \geq 0$, $x \in K_\partial$, $E \in \mathcal{B}_\partial$ and all stopping times τ, we have

$$P_x\{x_{\tau+h} \in E \mid \mathcal{F}_\tau\} = p_h(x_\tau, E),$$

or equivalently,

$$P_x(A \cap \{x_{\tau+h} \in E\}) = \int_A p_h(x_{\tau(\omega)}(\omega), E)\, dP_x(\omega) \quad \text{for all } A \in \mathcal{F}_\tau.$$

This expresses the idea of "starting afresh" at random times.

The next result gives a useful criterion for the strong Markov property:

Theorem 9.26. *If the transition function of a right-continuous Markov process has the C_0-property, then it is a strong Markov process.*

9.1.8 The Strong Markov Property and Uniform Stochastic Continuity

In this subsection we introduce the basic notion of uniform stochastic continuity of transition functions (Definition 9.27), and give simple criteria for the strong Markov property in terms of transition functions (Theorems 9.28 and 9.29).

Let (K, ρ) be a locally compact, separable metric space. We begin with the following definition:

Definition 9.27. A transition function p_t on K is said to be *uniformly stochastically continuous* on K if it satisfies the following condition:

For each $\varepsilon > 0$ and each compact $E \subset K$, we have

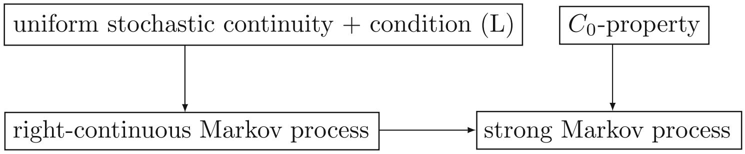
$$\lim_{t \downarrow 0} \sup_{x \in E} [1 - p_t(x, U_\varepsilon(x))] = 0, \tag{9.15}$$

where $U_\varepsilon(x) = \{y \in K : \rho(y, x) < \varepsilon\}$ is an ε-neighborhood of x.

It should be noted that every uniformly stochastically continuous transition function p_t is normal and satisfies condition (M) in Theorem 9.20. Therefore, by combining part (i) of Theorems 9.20 and 9.26 we obtain the following theorem:

Theorem 9.28. *If a uniformly stochastically continuous, C_0 transition function satisfies condition (L), then it is the transition function of some strong Markov process whose paths are right-continuous and have no discontinuities other than jumps.*

Theorems 9.20 and 9.26 can be visualized as follows:



A continuous strong Markov process is called a *diffusion process*. The next result states a sufficient condition for the existence of a diffusion process with a prescribed transition function:

Theorem 9.29. *If a uniformly stochastically continuous, C_0 transition function satisfies conditions (L) and (N), then it is the transition function of some diffusion process.*

This theorem is an immediate consequence of part (ii) of Theorems 9.20 and 9.28.

9.2 Feller Semigroups and Transition Functions

In Sect. 9.2 we introduce a class of semigroups associated with Markov processes (Definition 9.30), called Feller semigroups, and we give a characterization of Feller semigroups in terms of Markov transition functions (Theorems 9.33 and 9.34).

9.2.1 Definition of Feller Semigroups

Let (K, ρ) be a locally compact, separable metric space and let $C(K)$ be the Banach space of real-valued, bounded continuous functions on K with the supremum norm

$$\|f\|_\infty = \sup_{x \in K} |f(x)|.$$

Recall (see Sect. 9.1.4) that $C_0(K)$ is the closed subspace of $C(K)$ which consists of all functions satisfying the condition $\lim_{x \to \partial} f(x) = 0$, and further that $C_0(K)$ may be identified with $C(K)$ if K is compact.

Now we introduce a class of semigroups associated with Markov processes:

Definition 9.30. A family $\{T_t\}_{t \geq 0}$ of bounded linear operators acting on the space $C_0(K)$ is called a *Feller semigroup* on K if it satisfies the following three conditions (i), (ii) and (iii):

(i) $T_{t+s} = T_t \cdot T_s$ for all $t, s \geq 0$ (the semigroup property); $T_0 = I$.
(ii) The family $\{T_t\}$ is strongly continuous in t for $t \geq 0$:

$$\lim_{s \downarrow 0} \|T_{t+s} f - T_t f\|_\infty = 0 \quad \text{for every } f \in C_0(K).$$

(iii) The family $\{T_t\}$ is non-negative and contractive on $C_0(K)$:

$$f \in C_0(K), \ 0 \leq f(x) \leq 1 \quad \text{on } K \implies 0 \leq T_t f(x) \leq 1 \quad \text{on } K.$$

9.2.2 Characterization of Feller Semigroups in Terms of Transition Functions

In Sect. 9.1.3, we proved the following theorem:

Theorem 9.31. *If p_t is a Feller transition function on K, then the associated operators $\{T_t\}_{t \geq 0}$, defined by the formula*

$$T_t f(x) = \int_K p_t(x, dy) f(y) \quad \textit{for every } f \in C(K),$$

form a non-negative and contraction semigroup on $C(K)$:

(i) $T_{t+s} = T_t \cdot T_s$, t, $s \geq 0$ (the semigroup property); $T_0 = I$.
(ii) $f \in C(K)$, $0 \leq f(x) \leq 1$ on $K \implies 0 \leq T_t f(x) \leq 1$ on K.

The purpose of this subsection is to prove a converse:

Theorem 9.32. *If $\{T_t\}_{t \geq 0}$ is a non-negative and contraction semigroup on the space $C_0(K)$, then there exists a unique C_0 transition function p_t on K such that the formula*

$$T_t f(x) = \int_K p_t(x, dy) f(y) \tag{9.16}$$

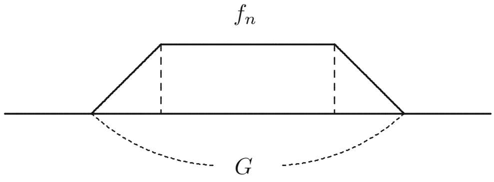
holds true for all $f \in C_0(K)$.

Proof. We fix $t \geq 0$ and $x \in K$, and define a linear functional F on $C_0(K)$ as follows:

$$F(f) = T_t f(x) \quad \text{for all } f \in C_0(K).$$

Then it follows that F is non-negative and bounded with norm $\|F\| \leq 1$, since T_t is a non-negative and contractive operator on $C_0(K)$. Therefore, by applying the Riesz–Markov representation theorem (Theorem 3.41) to the functional F we obtain that there exists a unique Radon measure $p_t(x, \cdot)$ on K such that

Fig. 9.8 The function f_n



$$T_t(x) = F(f) = \int_K p_t(x, dy) f(y) \quad \text{for all } f \in C_0(K). \tag{9.17}$$

We show that the measures p_t satisfy conditions (a)–(d) of Definition 9.4.

(a) First, we have the inequality

$$p_t(x, K) = \sup\{F(f) : f \in C_0(K),\ 0 \leq f \leq 1 \text{ on } K\} = \|F\| \leq 1, x \in K,$$

since F is contractive.

(b) Since $T_0 = I$, it follows that

$$f(x) = T_0 f(x) = \int_K p_0(x, dy) f(y) \quad \text{for all } f \in C_0(K).$$

This proves that $p_0(x, \{x\}) = 1$ for each $x \in K$.

(c) We prove that the function $p_t(\cdot, E)$ is *Borel measurable* for each $E \in \mathcal{B}$. To do this, it suffices to show that the collection

$$\mathcal{A} = \{E \in \mathcal{B} : p_t(\cdot, E) \text{ is } \mathcal{B}\text{-measurable}\}$$

coincides with the σ-algebra $\mathcal{B}$. The proof is divided into five steps.

Step 1: The collection $\mathcal{A}$ contains the collection $\mathcal{O}$ of all open subsets of K:

$$\mathcal{A} \supset \mathcal{O}. \tag{9.18}$$

Indeed, if $G \in \mathcal{O}$, we let (see Fig. 9.8)

$$f_n(x) = \min\{n\rho(x, K \setminus G), 1\}, \quad n = 1, 2, \ldots.$$

Then f_n is a function in $C_0(K)$, and satisfies the condition

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 1 & \text{if } x \in G, \\ 0 & \text{if } x \in K \setminus G. \end{cases}$$

Thus, by virtue of the dominated convergence theorem we obtain from formula (9.17) with $f := f_n$ that

$$\lim_{n\to\infty} T_t f_n = \lim_{n\to\infty} \int_K p_t(x, dy) f_n(y) = p_t(x, G).$$

Since the functions $T_t f_n$ are continuous, this proves that the limit function $p_t(\cdot, G)$ is $\mathcal{B}$-measurable and so $G \in \mathcal{A}$.

Step 2: We have, by assertion (9.18),

$$d(\mathcal{O}) \subset d(\mathcal{A}). \tag{9.19}$$

Step 3: The collection $\mathcal{A}$ is a d-system

$$d(\mathcal{A}) = \mathcal{A}. \tag{9.20}$$

Indeed, it is easy to verify the following three assertions (i), (ii) and (iii):

(i) By assertion (9.15), it follows that

$$K \in \mathcal{O} \subset \mathcal{A}.$$

(ii) If $A, B \in \mathcal{A}$ and $A \subset B$, then it follows that the function

$$p_t(\cdot, B \setminus A) = p_t(\cdot, B) - p_t(\cdot, A)$$

is $\mathcal{B}$-measurable. This proves that $B \setminus A \in \mathcal{A}$.

(iii) If $\{A_n\}_{n=1}^{\infty}$ is an increasing sequence of elements of $\mathcal{A}$, then it follows that the function

$$p_t\left(\cdot, \cup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} p_t(\cdot, A_n)$$

is $\mathcal{B}$-measurable. This proves that

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Step 4: Since $\mathcal{O}$ is a π-system, it follows from an application of the monotone class theorem (Theorem 2.4) that

$$d(\mathcal{O}) = \sigma(\mathcal{O}) = \mathcal{B}. \tag{9.21}$$

Step 5: By combining assertions (9.21), (9.19) and (9.20), we obtain that

$$\mathcal{B} = d(\mathcal{O}) \subset d(\mathcal{A}) = \mathcal{A} \subset \mathcal{B},$$

so that

$$\mathcal{A} = \mathcal{B}.$$

(d) In view of the semigroup property and Fubini's theorem, it follows from formula (9.17) that we have, for all $f \in C_0(K)$,

$$\begin{aligned}\int_K p_{t+s}(x,dz)f(z) &= T_{t+s}f(x) = T_t(T_s f)(x) = \int_K p_t(x,dy)\int_K p_s(y,dz)f(z)\\ &= \int_K \left(\int_K p_t(x,dy)p_s(y,dz)\right) f(z).\end{aligned}$$

Hence the uniqueness part of the Riesz–Markov representation theorem (Theorem 3.41) gives that

$$p_{t+s}(x,E) = \int_K p_t(x,dy)p_s(y,E) = \int_K p_t(x,dy)p_s(y,E) \quad \text{for all } E \in \mathcal{B}.$$

Finally, the C_0-property of p_t comes automatically, since $T_t : C_0(K) \to C_0(K)$. The proof of Theorem 9.32 is now complete.

It should be emphasized that the Feller or C_0-property concerns only the continuity of a Markov transition function $p_t(x,E)$ in x, and not, by itself, continuity in t.

Now we give a necessary and sufficient condition on $p_t(x,E)$ in order that its associated operators $\{T_t\}_{t\geq 0}$ are strongly continuous in t on the space $C_0(K)$:

$$\lim_{s\downarrow 0} \|T_{t+s}f - T_t f\|_\infty = 0 \quad \text{for all } f \in C_0(K). \tag{9.22}$$

Theorem 9.33. *Let $p_t(x,\cdot)$ be a C_0 transition function on K. Then the associated operators $\{T_t\}_{t\geq 0}$, defined by formula (9.16), are strongly continuous in t on $C_0(K)$ if and only if $p_t(x,\cdot)$ is uniformly stochastically continuous on K and satisfies the following condition (L):*

(L) For each $s > 0$ and each compact $E \subset K$, we have

$$\lim_{x\to\partial} \sup_{0\leq t\leq s} p_t(x,E) = 0. \tag{9.23}$$

Proof. The proof is divided into two steps.

Step 1: First, we prove the "if" part of the theorem. Since continuous functions with compact support are dense in $C_0(K)$, it suffices to prove the strong continuity of $\{T_t\}$ at $t = 0$

$$\lim_{t\downarrow 0} \|T_t f - f\|_\infty = 0 \tag{9.22$'$}$$

for all such functions f.

For any compact subset E of K containing $\operatorname{supp} f$, we have

$$\begin{aligned}\|T_t f - f\|_\infty &\le \sup_{x\in E} |T_t f(x) - f(x)| + \sup_{x\in K\setminus E} |T_t f(x)| \\ &\le \sup_{x\in E} |T_t f(x) - f(x)| + \|f\|_\infty \cdot \sup_{x\in K\setminus E} p_t(x, \operatorname{supp} f).\end{aligned} \tag{9.24}$$

However, condition (L) implies that, for each $\varepsilon > 0$, we can find a compact subset E of K such that, for all sufficiently small $t > 0$,

$$\sup_{x\in K\setminus E} p_t(x, \operatorname{supp} f) < \varepsilon. \tag{9.25}$$

On the other hand, we have, for each $\delta > 0$,

$$\begin{aligned}T_t f(x) - f(x) &= \int_{U_\delta(x)} p_t(x, dy)(f(y) - f(x)) \\ &\quad + \int_{K\setminus U_\delta(x)} p_t(x, dy)(f(y) - f(x)) - f(x)(1 - p_t(x, K)),\end{aligned}$$

and hence

$$\begin{aligned}&\sup_{x\in E} |T_t f(x) - f(x)| \\ &\le \sup_{\rho(x,y)<\delta} |f(y) - f(x)| + 3\|f\|_\infty \cdot \sup_{x\in E} [1 - p_t(x, U_\delta(x))] .\end{aligned}$$

Since $f(x)$ is uniformly continuous, we can choose a constant $\delta > 0$ such that

$$\sup_{\rho(x,y)<\delta} |f(y) - f(x)| < \varepsilon.$$

Furthermore, it follows from condition (9.15) with $\varepsilon := \delta$ that, for all sufficiently small $t > 0$,

$$\sup_{x\in E} [1 - p_t(x, U_\delta(x))] < \varepsilon.$$

Hence we have, for all sufficiently small $t > 0$,

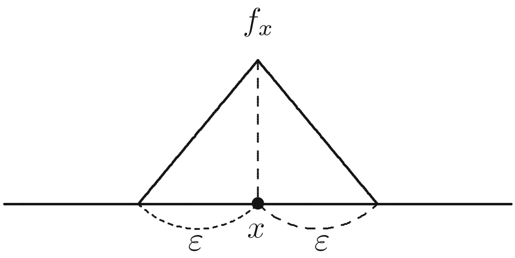
$$\sup_{x\in E} |T_t f(x) - f(x)| < \varepsilon\,(1 + 3\|f\|_\infty)\,. \tag{9.26}$$

Therefore, by carrying inequalities (9.25) and (9.26) into inequality (9.24) we obtain that, for all sufficiently small $t > 0$,

$$\|T_t f - f\|_\infty < \varepsilon\,(1 + 4\|f\|_\infty)\,.$$

This proves (9.22′), that is, the strong continuity of $\{T_t\}$.

Fig. 9.9 The function f_x



Step 2: Next, we prove the "only if" part of the theorem.

1. Let x be an arbitrary point of K. For any $\varepsilon > 0$, we define (see Fig. 9.9)

$$f_x(y) = \begin{cases} 1 - \dfrac{1}{\varepsilon}\rho(x, y) & \text{if } \rho(x, y) \leq \varepsilon, \\ 0 & \text{if } \rho(x, y) > \varepsilon. \end{cases} \tag{9.27}$$

If E is a compact subset of K, then the functions f_x, $x \in E$, are in $C_0(K)$, for all sufficiently small $\varepsilon > 0$, and satisfy the condition

$$\| f_x - f_z \|_\infty \leq \frac{1}{\varepsilon}\rho(x, z), \quad x, z \in E. \tag{9.28}$$

However, for any $\delta > 0$, by the compactness of E we can find a finite number of points $x_1, x_2, \ldots, x_n$ of E such that

$$E \subset \bigcup_{k=1}^{n} U_{\delta\varepsilon/4}(x_k),$$

and hence

$$\min_{1 \leq k \leq n} \rho(x, x_k) \leq \frac{\delta\varepsilon}{4} \quad \text{for all } x \in E.$$

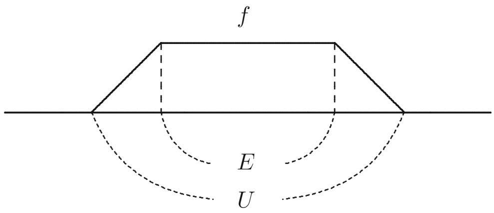
Thus, by combining this inequality with inequality (9.28) we obtain that

$$\min_{1 \leq k \leq n} \| f_x - f_{x_k} \|_\infty \leq \frac{\delta}{4} \quad \text{for all } x \in E. \tag{9.29}$$

Now we have, by (9.27),

$$\begin{aligned} 0 &\leq 1 - p_t(x, U_\varepsilon(x)) \\ &\leq 1 - \int_{K_\partial} p_t(x, dy) f_x(y) = f_x(x) - T_t f_x(x) \end{aligned}$$

Fig. 9.10 The function f



$$\leq \|f_x - T_t f_x\|_\infty$$
$$\leq \|f_x - f_{x_k}\|_\infty + \|f_{x_k} - T_t f_{x_k}\|_\infty + \|T_t f_{x_k} - T_t f_x\|_\infty$$
$$\leq 2\|f_x - f_{x_k}\|_\infty + \|f_{x_k} - T_t f_{x_k}\|_\infty.$$

In view of inequality (9.29), the first term in the last inequality is bounded by $\delta/2$ for the right choice of k. Furthermore, it follows from the strong continuity (9.22′) of $\{T_t\}$ that the second term tends to zero as $t \downarrow 0$, for each $k = 1, \cdots, n$.
Consequently, we have, for all sufficiently small $t > 0$,

$$\sup_{x \in E} [1 - p_t(x, U_\varepsilon(x))] \leq \delta.$$

This proves condition (9.15), that is, the uniform stochastic continuity of $p_t(x, \cdot)$.

2. It remains to verify condition (L). We assume, to the contrary, that:

For some $s > 0$ and some compact $E \subset K$, there exist a constant $\varepsilon_0 > 0$, a sequence $\{t_k\}$, $t_k \downarrow t$ $(0 \leq t \leq s)$ and a sequence $\{x_k\}$, $x_k \to \partial$, such that

$$p_{t_k}(x_k, E) \geq \varepsilon_0. \tag{9.30}$$

Now we take a relatively compact subset U of K containing E, and let (see Fig. 9.10)

$$f(x) = \frac{\rho(x, K \setminus U)}{\rho(x, E) + \rho(x, K \setminus U)}.$$

Then it follows that the function $f(x)$ is in $C_0(K)$ and satisfies the condition

$$T_t f(x) = \int_K p_t(x, dy) f(y) \geq p_t(x, E) \geq 0.$$

Therefore, by combining this inequality with inequality (9.30) we obtain that

$$T_{t_k} f(x_k) = \int_K p_{t_k}(x_k, dy) f(y) \geq p_{t_k}(x_k, E) \geq \varepsilon_0. \tag{9.31}$$

However, we have the inequality

$$T_{t_k} f(x_k) \leq \|T_{t_k} f - T_t f\|_\infty + T_t f(x_k). \tag{9.32}$$

Since the semigroup $\{T_t\}$ is strongly continuous and $T_t f \in C_0(K)$, we can let $k \to \infty$ in inequality (9.32) to obtain that

$$\limsup_{k \to \infty} T_{t_k} f(x_k) = 0.$$

This contradicts condition (9.31).

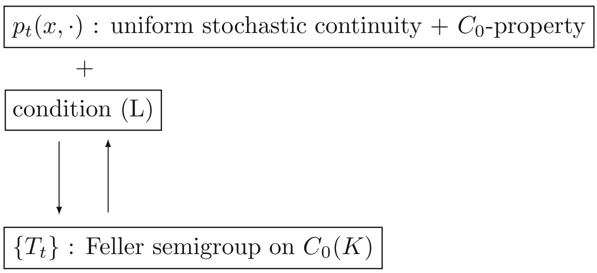
The proof of Theorem 9.33 is complete.

Rephrased, Theorem 9.33 gives a characterization of Feller semigroups in terms of transition functions:

Theorem 9.34. *If $p_t(x, \cdot)$ is a uniformly stochastically continuous, C_0 transition function on K and satisfies condition (L), then its associated operators $\{T_t\}_{t \geq 0}$ form a Feller semigroup on K.*

Conversely, if $\{T_t\}_{t \geq 0}$ is a Feller semigroup on K, then there exists a uniformly stochastically continuous, C_0 transition function $p_t(x, \cdot)$ on K, satisfying condition (L), such that (9.16) holds true for all $f \in C_0(K)$.

Theorem 9.34 can be visualized as follows:



9.3 The Hille–Yosida Theory of Feller Semigroups

Section 9.3 is devoted to a version of the Hille–Yosida theorem (Theorem 3.10) adapted to the present context. In particular, we prove generation theorems for Feller semigroups (Theorems 9.35 and 9.50) which form a functional analytic background for the proof of Theorem 1.2 in Chap. 10.

9.3.1 Generation Theorems for Feller Semigroups

Let (K, ρ) be a locally compact, separable metric space and let $C(K)$ be the space of real-valued, bounded continuous functions on K; $C(K)$ is a normed linear space with the supremum norm

$$\|f\|_\infty = \sup_{x \in K} |f(x)|.$$

We add a new point ∂ to the locally compact space K as the point at infinity if K is not compact, and as an isolated point if K is compact; so the space $K_\partial = K \cup \{\partial\}$ is compact. Recall that $C_0(K)$ is the closed subspace of $C(K)$ which consists of all functions satisfying the condition $\lim_{x \to \partial} f(x) = 0$, and that $C_0(K)$ may be identified with $C(K)$ if K is compact.

If $\{T_t\}_{t \geq 0}$ is a Feller semigroup on K, we define its *infinitesimal generator* A by the formula

$$Au = \lim_{t \downarrow 0} \frac{T_t u - u}{t}, \tag{9.33}$$

provided that the limit (9.30) exists in the space $C_0(K)$. More precisely, the generator A is a linear operator from $C_0(K)$ into itself defined as follows:

(1) The domain $D(A)$ of A is the set

$$D(A) = \{u \in C_0(K) : \text{the limit (9.33) exists in } C_0(K)\}.$$

(2) $Au = \lim_{t \downarrow 0} \dfrac{T_t u - u}{t}$ for every $u \in D(A)$.

The next theorem is a version of the Hille–Yosida theorem (Theorem 3.10) adapted to the present context:

Theorem 9.35 (Hille–Yosida).

(i) Let $\{T_t\}_{t \geq 0}$ be a Feller semigroup on K and A its infinitesimal generator. Then we have the following four assertions (a)–(d):

(a) The domain $D(A)$ is dense in the space $C_0(K)$.

(b) For each $\alpha > 0$, the equation $(\alpha I - A)u = f$ has a unique solution u in $D(A)$ for any $f \in C_0(K)$. Hence, for each $\alpha > 0$ the Green operator $(\alpha I - A)^{-1} : C_0(K) \to C_0(K)$ can be defined by the formula

$$u = (\alpha I - A)^{-1} f \quad \text{for every } f \in C_0(K).$$

(c) For each $\alpha > 0$, the operator $(\alpha I - A)^{-1}$ is non-negative on $C_0(K)$:

$$f \in C_0(K), \quad f \geq 0 \quad \text{on } K \Longrightarrow (\alpha I - A)^{-1} f \geq 0 \quad \text{on } K.$$

(d) *For each* $\alpha > 0$*, the operator* $(\alpha I - A)^{-1}$ *is bounded on* $C_0(K)$ *with norm*

$$\|(\alpha I - A)^{-1}\| \leq \frac{1}{\alpha}.$$

(ii) *Conversely, if* A *is a linear operator from* $C_0(K)$ *into itself satisfying condition (a) and if there is a constant* $\alpha_0 \geq 0$ *such that, for all* $\alpha > \alpha_0$*, conditions (b)–(d) are satisfied, then* A *is the infinitesimal generator of some Feller semigroup* $\{T_t\}_{t\geq 0}$ *on* K.

Proof. In view of Theorem 3.10, it suffices to show that the semigroup $\{T_t\}_{t\geq 0}$ is non-negative if and only if its resolvents $\{(\alpha I - A)^{-1}\}_{\alpha>\alpha_0}$ are non-negative.

The "only if" part is an immediate consequence of the integral expression of $(\alpha I - A)^{-1}$ in terms of the semigroup $\{T_t\}$ (see formula (4.11)):

$$(\alpha I - A)^{-1} = \int_0^\infty e^{-\alpha t} T_t \, dt, \quad \alpha > 0.$$

On the other hand, the "if" part follows from expression (4.18) of the semigroup $T_t(\alpha)$ in terms of the Yosida approximation $J_\alpha = \alpha(\alpha I - A)^{-1}$:

$$T_t(\alpha) = e^{-\alpha t} \exp[\alpha t J_\alpha] = e^{-\alpha t} \sum_{n=0}^{\infty} \frac{(\alpha t)^n}{n!} J_\alpha^n,$$

and definition (4.19) of the semigroup T_t:

$$T_t = \lim_{\alpha\to\infty} T_t(\alpha).$$

The proof of Theorem 9.35 is complete.

Corollary 9.36. *Let* K *be a* compact *metric space and let* A *be the infinitesimal generator of a Feller semigroup on* K*. Assume that the constant function* 1 *belongs to the domain* $D(A)$ *of* A *and that we have, for some constant* c,

$$A1(x) \leq -c \quad \text{on } K. \tag{9.34}$$

Then the operator $A' = A + cI$ *is the infinitesimal generator of some Feller semigroup on* K.

Proof. It follows from an application of part (i) of Theorem 9.35 that, for all $\alpha > c$, the operators

$$(\alpha I - A')^{-1} = ((\alpha - c)I - A)^{-1}$$

are defined and non-negative on the whole space $C(K)$. However, in view of inequality (9.34) we obtain that

$$\alpha \leq \alpha - (A1 + c) = (\alpha I - A')1 \quad \text{on } K,$$

so that

$$\alpha(\alpha I - A')^{-1}1 \leq (\alpha I - A')^{-1}(\alpha I - A')1 = 1 \quad \text{on } K.$$

Hence we have, for all $\alpha > c$,

$$\|(\alpha I - A')^{-1}\| = \|(\alpha I - A')^{-1}1\|_\infty \leq \frac{1}{\alpha}.$$

Therefore, by applying part (ii) of Theorem 9.35 to the operator $A' = A + cI$ we find that A' is the infinitesimal generator of some Feller semigroup on K.

The proof of Corollary 9.36 is complete.

Now we write down explicitly the infinitesimal generators of Feller semigroups associated with the transition functions in Examples 9.6–9.12.

Example 9.37 (uniform motion). Let $K = \mathbf{R}$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^1(K) : f' \in C_0(K)\}, \\ Af = v f' \quad \text{for every } f \in D(A), \end{cases}$$

where v is a positive constant. Then the resolvents $\{(\alpha I - A)^{-1}\}_{\alpha>0}$ are given by the formula

$$(\alpha I - A)^{-1} g = \frac{1}{v} \int_x^\infty e^{-\frac{\alpha}{v}(y-x)} g(y)\, dy \quad \text{for every } g \in C_0(K).$$

Example 9.38 (Poisson process). Let $K = \mathbf{R}$ and

$$\begin{cases} D(A) = C_0(K), \\ Af(x) = \lambda(f(x+1) - f(x)) \quad \text{for every } f \in D(A). \end{cases}$$

The operator A is not "local"; the value $Af(x)$ depends on the values $f(x)$ and $f(x+1)$. This reflects the fact that the Poisson process changes state by jumps.

Example 9.39 (Brownian motion). Let $K = \mathbf{R}$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ f'' \in C_0(K)\}, \\ Af = \frac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

The operator A is "local", that is, the value $Af(x)$ is determined by the values of f in an arbitrary small neighborhood of x. This reflects the fact that Brownian motion changes state by continuous motion. The resolvents $\{(\alpha I - A)^{-1}\}_{\alpha>0}$ are given by the formula

$$(\alpha I - A)^{-1} g = \frac{1}{\sqrt{2\alpha}} \int_{-\infty}^{\infty} e^{-\sqrt{2\alpha}|x-y|} g(y)\, dy \quad \text{for every } g \in C_0(K).$$

Example 9.40 (Brownian motion with constant drift). Let $K = \mathbf{R}$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ f'' \in C_0(K)\}, \\ Af = \dfrac{1}{2} f'' + m f' \quad \text{for every } f \in D(A). \end{cases}$$

Example 9.41 (Cauchy process). Let $K = \mathbf{R}$ and the domain $D(A)$ contain C^2 functions on K with compact support, and let the infinitesimal generator A be of the form

$$Af(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} (f(x+y) - f(x))\, \frac{dy}{y^2}.$$

The operator A is not "local", which reflects the fact that the Cauchy process changes state by jumps.

Example 9.42 (reflecting barrier Brownian motion). Let $K = [0, \infty)$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ f'' \in C_0(K),\ f'(0) = 0\}, \\ Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

The resolvents $\{(\alpha I - A)^{-1}\}_{\alpha>0}$ are given by

$$\begin{aligned} &(\alpha I - A)^{-1} g \qquad (9.35) \\ &= \frac{1}{\sqrt{2\alpha}} \int_0^{\infty} e^{\sqrt{2\alpha}(x-y)} g(y)\, dy + \frac{1}{\sqrt{2\alpha}} \int_0^{\infty} e^{-\sqrt{2\alpha}(x+y)} g(y)\, dy \\ &\quad - \frac{1}{\sqrt{2\alpha}} \int_0^{x} \left[e^{\sqrt{2\alpha}(x-y)} - e^{-\sqrt{2\alpha}(x-y)} \right] g(y)\, dy \quad \text{for every } g \in C_0(K). \end{aligned}$$

Example 9.43 (sticking barrier Brownian motion). Let $K = [0, \infty)$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ f'' \in C_0(K),\ f''(0) = 0\}, \\ Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

The resolvents $\{(\alpha I - A)^{-1}\}_{\alpha>0}$ are given by

$$
\begin{aligned}
&(\alpha I - A)^{-1} g \qquad (9.36)\\
&= \frac{1}{\sqrt{2\alpha}} \int_x^\infty e^{\sqrt{2\alpha}(x-y)} g(y)\, dy + \frac{1}{\alpha}\, g(0)\, e^{-\sqrt{2\alpha}x} \\
&\quad - \frac{1}{\alpha} \int_0^\infty e^{-\sqrt{2\alpha}(x+y)} g(y)\, dy + \frac{1}{\sqrt{2\alpha}} \int_0^x e^{-\sqrt{2\alpha}(x-y)} g(y)\, dy \\
&\qquad \text{for every } g \in C_0(K).
\end{aligned}
$$

Moreover, we can obtain the following:

Example 9.44 (reflecting barrier Brownian motion). Let $K = [0, 1]$ and

$$
\begin{cases}
D(A) = \{f \in C^2(K) : \ f'(0) = f'(1) = 0\}, \\
Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A).
\end{cases}
$$

Example 9.45 (absorbing barrier Brownian motion). Let $K = [0, \infty)$ and

$$
\begin{cases}
D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K), \ f'' \in C_0(K), \ f(0) = 0\}, \\
Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A).
\end{cases}
$$

This represents Brownian motion with an absorbing barrier at $x = 0$; a Brownian particle dies at the first moment when it hits the boundary $x = 0$. That is, the point 0 is the terminal point.

Example 9.46 (absorbing barrier Brownian motion). Let $K = [0, 1]$ where the boundary points 0 and 1 are identified with the point at infinity ∂. More precisely, we introduce a subspace $C_0(K)$ of $C(K)$ as follows:

$$
C_0(K) = \{f \in C(K) : f(0) = f(1) = 0\}.
$$

Then we define a linear operator $A : C_0(K) \to C_0(K)$ by the formula

$$
\begin{cases}
D(A) = \{f \in C^2(K) : f' \in C_0(K), \ f'' \in C_0(K), \ f(0) = f(1) = 0\}, \\
Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A).
\end{cases}
$$

This represents Brownian motion with two absorbing barriers at $x = 0$ and $x = 1$; a Brownian particle dies at the first moment when it hits the boundary points $x = 0$ and $x = 1$. That is, the two points 0 and 1 are the terminal points.

Example 9.47 (absorbing–reflecting barrier Brownian motion). Let $K = [0, \infty)$, and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ f'' \in C_0(K), \\ \qquad\qquad \lambda f'(0) - (1-\lambda) f(0) = 0\}, \\ Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

Here λ is a constant such that $0 < \lambda < 1$. This process $\{x_t\}$ may be thought of as a "combination" of the absorbing and reflecting Brownian motions; the absorbing and reflecting cases are formally obtained by letting $\lambda \to 0$ and $\lambda \to 1$, respectively.

Here is an example where it is difficult to begin with a transition function and the infinitesimal generator is the basic tool used to describe the process.

Example 9.48 (sticky barrier Brownian motion). Let $K = [0, \infty)$. We define a linear operator $A : C_0(K) \to C_0(K)$ by the formula

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K), \\ \qquad f'' \in C_0(K),\ f'(0) - \lambda f''(0) = 0\}, \\ Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

Here λ is a positive constant. This process $\{x_t\}$ may be thought of as a "combination" of the reflecting and sticking Brownian motions; the reflecting and sticking cases are formally obtained by letting $\lambda \to 0$ and $\lambda \to \infty$, respectively. Upon hitting $x = 0$, a Brownian particle leaves immediately, but it spends a positive duration of time there. We remark that the set $\{t > 0 : x_t = 0\}$ is somewhat analogous to Cantor-like sets of positive Lebesgue measure. The resolvents $\{(\alpha I - A)^{-1}\}_{\alpha > 0}$ are given by the formula

$$\begin{aligned} &(\alpha I - A)^{-1} g \qquad\qquad (9.37) \\ &= \frac{1}{\sqrt{2\alpha}} \int_0^\infty e^{\sqrt{2\alpha}(x-y)} g(y)\, dy + C\, e^{-\sqrt{2\alpha} x} \\ &\quad - \frac{1}{\sqrt{2\alpha}} \int_0^x \left[e^{\sqrt{2\alpha}(x-y)} - e^{-\sqrt{2\alpha}(x-y)} \right] g(y)\, dy \quad \text{for every } g \in C_0(K), \end{aligned}$$

where C is a constant given by the formula

$$C = \left(\frac{\frac{1}{\lambda} - \sqrt{2\alpha}}{\frac{1}{\lambda} + \sqrt{2\alpha}} \right) \int_0^\infty e^{-\sqrt{2\alpha} y} g(y)\, dy + \frac{\sqrt{\frac{2}{\alpha}}\, g(0)}{\frac{1}{\lambda} + \sqrt{2\alpha}}.$$

It should be noted that (9.35) and (9.36) may be obtained from (9.37) by letting $\lambda \to 0$ and $\lambda \to \infty$, respectively.

Finally, it is worth pointing out here that a strong Markov process cannot stay at a single position for a positive length of time and then leave that position by continuous motion; it must either jump away or leave instantaneously. We give a simple example of a strong Markov process which changes state not by continuous motion but by jumps when the motion reaches the boundary:

Example 9.49. Let $K = [0, \infty)$ and

$$\begin{cases} D(A) = \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K),\ \ f'' \in C_0(K), \\ \qquad f''(0) = 2c \int_0^\infty (f(y) - f(0))\, dF(y)\}, \\ Af = \dfrac{1}{2} f'' \quad \text{for every } f \in D(A). \end{cases}$$

Here c is a positive constant and F is a distribution function on $(0, \infty)$.

This process $\{x_t\}$ may be interpreted as follows. When a Brownian particle reaches the boundary $x = 0$, it stays there for a positive length of time and then jumps back to a random point, chosen with the function F, in the interior $(0, \infty)$. The constant c is the parameter in the "waiting time" distribution at the boundary $x = 0$. Note that the boundary condition

$$f''(0) = 2c \int_0^\infty (f(y) - f(0))\, dF(y)$$

depends on the values of $f(y)$ far away from the boundary $x = 0$, unlike the boundary conditions in Examples 9.11–9.14.

9.3.2 Generation Theorems for Feller Semigroups in Terms of Maximum Principles

Although Theorem 9.35 tells us precisely when a linear operator A is the infinitesimal generator of some Feller semigroup, it is usually difficult to verify conditions (b)–(d). So we give useful criteria in terms of *maximum principles* [Ta9, Theorem 2.18 and Corollary 2.19]:

Theorem 9.50 (Hille–Yosida–Ray). *Let K be a* compact *metric space. Then we have the following two assertions (i) and (ii):*

(i) Let B be a linear operator from $C(K) = C_0(K)$ into itself, and assume that

- (α) *The domain $D(B)$ of B is dense in the space $C(K)$.*
- (β) *There exists an open and dense subset K_0 of K such that if $u \in D(B)$ takes a* positive *maximum at a point x_0 of K_0, then we have the inequality*

$$Bu(x_0) \leq 0.$$

Then the operator B is closable in the space $C(K)$.

(ii) Let B be as in part (i), and further assume that

(β') *If $u \in D(B)$ takes a* positive *maximum at a point x' of K, then we have the inequality*

$$Bu(x') \leq 0.$$

(γ) *For some $\alpha_0 \geq 0$, the range $\mathcal{R}(\alpha_0 I - B)$ of $\alpha_0 I - B$ is dense in the space $C(K)$.*

Then the minimal closed extension $\overline{B}$ *of B is the infinitesimal generator of some Feller semigroup on K.*

Corollary 9.51. *Let B be the infinitesimal generator of a Feller semigroup on a compact metric space K and C a bounded linear operator from $C(K)$ into itself. Assume that either C or $B + C$ satisfies condition (β'). Then the operator $A = B + C$ is the infinitesimal generator of some Feller semigroup on K.*

9.4 Infinitesimal Generators of Feller Semigroups on a Bounded Domain (i)

In the early 1950s, W. Feller completely characterized the analytic structure of one-dimensional diffusion processes; he gave an intrinsic representation of the infinitesimal generator $\mathfrak{A}$ of a one-dimensional diffusion process and determined all possible boundary conditions which describe the domain $D(\mathfrak{A})$ of $\mathfrak{A}$. The probabilistic meaning of Feller's work was clarified by E.B. Dynkin [Dy1, Dy2], K. Itô and H.P. McKean, Jr. [IM], D. Ray [Ra] and others. One-dimensional diffusion processes are fully understood, both from the analytic and the probabilistic viewpoints.

Now we take a close look at Feller's work. Let $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ be a one-dimensional diffusion process with state space K. A point x of K is called a right (resp. left) *singular point* if $x_t(\omega) \geq x$ (resp. $x_t(\omega) \leq x$) for all $t \in [0, \zeta(\omega))$ with P_x-measure one. A right and left singular point is called a *trap*. For example, the point at infinity ∂ is a trap. A point which is neither right nor left singular is called a *regular point*.

For simplicity, we assume that the state space K is the half-line

$$K = [0, \infty),$$

and all its interior points are regular. Feller proved that there exist a strictly increasing, continuous function s on $(0, \infty)$ and Borel measures m and k on $(0, \infty)$ such that the infinitesimal generator $\mathfrak{A}$ of the process $\mathcal{X}$ can be expressed as follows:

$$\mathfrak{A} f(x) = \lim_{y \downarrow x} \frac{f^+(y) - f^+(x) - \int_{(x,y]} f(z)\, dk(z)}{m((x, y])}. \tag{9.38}$$

Here:

(1) $f^{+}(x) = \lim_{\varepsilon \downarrow 0} \frac{f(x+\varepsilon)-f(x)}{s(x+\varepsilon)-s(x)}$, the right-derivative of f at x with respect to s.
(2) The measure m is positive for non-empty open subsets, and is finite for compact sets.
(3) The measure k is finite for compact subsets.

The function s is called a *canonical scale*, and the measures m and k are called a *canonical measure* (or speed measure) and a *killing measure* for the process $\mathcal{X}$, respectively. They determine the behavior of a Markovian particle in the interior of the state space K.

We remark that the right-hand side of (9.38) is a generalization of the second-order differential operator

$$a(x) f'' + b(x) f' + c(x) f,$$

where $a(x) > 0$ and $c(x) \leq 0$ on K. For example, the formula

$$\mathfrak{A} f = a(x) f'' + b(x) f'$$

can be written in the form (9.13), if we take

$$\begin{aligned} s(x) &= \int_0^x \exp\left[-\int_0^y \frac{b(z)}{a(z)}\, dz\right] dy, \\ dm(x) &= \frac{1}{a(x)} \exp\left[\int_0^x \frac{b(y)}{a(y)}\, dy\right] dx, \\ dk(x) &= 0. \end{aligned}$$

The boundary point 0 is called a *regular boundary* if we have, for a point $r \in (0, \infty)$,

$$\int_{(0,r)} [s(r) - s(x)][dm(x) + dk(x)] < \infty,$$

$$\int_{(0,r)} [m((x,r)) + k((x,r))]\, ds(x) < \infty.$$

It can be shown that this notion is independent of the point r used. Intuitively, the regularity of the boundary point means that a Markovian particle approaches the boundary in finite time with positive probability, and also enters the interior from the boundary.

The behavior of a Markovian particle at the boundary point is characterized by boundary conditions. In the case of regular boundary points, Feller determined all possible boundary conditions which are satisfied by the functions $f(x)$ in the domain $D(\mathfrak{A})$ of $\mathfrak{A}$. A general boundary condition is of the form

$$\gamma f(0) - \delta A f(0) + \mu f^{+}(0) = 0,$$

where γ, δ and μ are constants such that $\gamma \leq 0$, $\delta \geq 0$, $\mu \geq 0$, $\mu + \delta > 0$. If we admit jumps from the boundary into the interior, then a general boundary condition takes the form

$$\gamma f(0) - \delta \mathfrak{A} f(0) + \mu f^{+}(0) + \int_{(0,\infty)} [f(x) - f(0)]\, d\nu(x) = 0 \tag{9.39}$$

where ν is a Borel measure with respect to which the function $\min(1, s(x) - s(+0))$ is integrable. It should be noted that boundary condition (9.39) is a "combination" of the boundary conditions in Examples 9.11, 9.14 and 9.49 if we take

$$s(x) = x, \quad dm(x) = 2dx, \quad dk(x) = 0.$$

A Markov process is said to be one-dimensional or multi-dimensional according as the state space is a subset of $\mathbf{R}$ or $\mathbf{R}^N$ ($N \geq 2$).

The main purpose of this book is to generalize Feller's work to the multi-dimensional case. In 1959 A.D. Ventcel' [We] studied the problem of determining all possible boundary conditions for multi-dimensional diffusion processes. In this section and the next section, we shall describe analytically the infinitesimal generator of a Feller semigroup in the case when the state space K is the closure $\overline{D}$ of a *bounded domain* D in $\mathbf{R}^N$ (Theorems 9.52 and 9.53).

Let K be a *compact* metric space and let $C(K) = C_0(K)$ be the Banach space of real-valued continuous functions $f(x)$ on K with the maximum norm

$$\|f\|_\infty = \max_{x \in K} |f(x)|.$$

A sequence $\{\mu_n\}_{n=1}^\infty$ of real Borel measures on K is said to *converge weakly* to a real Borel measure μ on K if it satisfies the condition

$$\lim_{n \to \infty} \int_K f(x)\, d\mu_n(x) = \int_K f(x)\, d\mu(x) \quad \text{for every } f \in C(K). \tag{9.40}$$

We remark (see Theorem 3.46) that the space of all real Borel measures μ on K is a normed linear space with the norm

$$\|\mu\| = \text{the total variation } |\mu|(K) \text{ of } \mu,$$

and further (see formula (3.17)) that the weak convergence (9.40) of real Borel measures is just the weak* convergence of the dual space $C(K)'$ of $C(K)$.

Now we recall that a Feller semigroup $\{T_t\}_{t \geq 0}$ on K is a strongly continuous semigroup of bounded linear operators T_t acting on $C(K)$ such that

$$f \in C(K),\ 0 \leq f(x) \leq 1 \quad \text{on } K \Longrightarrow 0 \leq T_t f(x) \leq 1 \quad \text{on } K.$$

The infinitesimal generator $\mathfrak{A}$ of $\{T_t\}$ is defined by the formula

$$\mathfrak{A}u = \lim_{t \downarrow 0} \frac{T_t u - u}{t}, \tag{9.41}$$

provided that the limit (9.41) exists in $C(K)$. That is, the generator $\mathfrak{A}$ is a linear operator from $C(K)$ into itself whose domain $D(\mathfrak{A})$ consists of all $u \in C(K)$ for which the limit (9.41) exists.

Theorem 9.35, a version of the Hille–Yosida theorem, asserts that a Feller semigroup is completely characterized by its infinitesimal generator. Therefore, we are reduced to the study of the infinitesimal generators of Feller semigroups.

Our first job is to derive an explicit formula in the *interior* D of $\overline{D}$ for the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}_{t \geq 0}$ on the closure $\overline{D}$.

The next result is essentially due to Ventcel' [We] and [Ta5, Theorem 9.4.1]:

Theorem 9.52. *Let D be a bounded domain in $\mathbf{R}^N$ and let $\{T_t\}_{t \geq 0}$ be a Feller semigroup on $\overline{D}$ with infinitesimal generator $\mathfrak{A}$. Assume that, for every point x^0 of D, there exist a local coordinate system $(x_1, x_2, \ldots, x_N)$ on a neighborhood of x^0 and continuous functions $\chi_1, \chi_2, \ldots, \chi_N$ on $\overline{D}$ such that $\chi_i = x_i$ in a neighborhood of x^0 and that the functions $1, \chi_1, \chi_2, \ldots, \chi_N$ and $\sum_{i=1}^N \chi_i$ belong to the domain $D(\mathfrak{A})$ of $\mathfrak{A}$. Then we have, for all $u \in D(\mathfrak{A}) \cap C^2(\overline{D})$,*

$$\begin{aligned} \mathfrak{A}u(x^0) = & \sum_{i,j=1}^N a^{ij}(x^0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x^0) + \sum_{i=1}^N b^i(x^0) \frac{\partial u}{\partial x_i}(x^0) + c(x^0) u(x^0) \\ & + \int_{\overline{D}} e(x^0, dy) \left[u(y) - u(x^0) - \sum_{i=1}^N \frac{\partial u}{\partial x_i}(x^0) \left(\chi_i(y) - \chi_i(x^0) \right) \right]. \end{aligned} \tag{9.42}$$

Here:

(1) The matrix $(a^{ij}(x^0))$ is symmetric and positive semi-definite.
(2) $b^i(x^0) = \mathfrak{A}(\chi_i - \chi_i(x^0))(x^0)$ for all $1 \leq i \leq N$.
(3) $c(x^0) = \mathfrak{A}1(x^0)$.
(4) $e(x^0, \cdot)$ is a non-negative Borel measure on $\overline{D}$ such that, for any neighborhood U of x^0,

$$e(x^0, \overline{D} \setminus U) < \infty, \tag{9.43a}$$

$$\int_U e(x^0, dy) \left[\sum_{i=1}^N \left(\chi_i(y) - \chi_i(x^0) \right)^2 \right] < \infty. \tag{9.43b}$$

Proof. The proof is divided into three steps.

Step 1: By applying Theorem 9.33 with $C_0(K) := C(\overline{D})$, we obtain that there corresponds to a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D}$ a unique uniformly stochastically continuous Feller transition function p_t on $\overline{D}$ in the following manner:

$$T_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y) \quad \text{for every } f \in C(\overline{D}).$$

Since the functions 1, χ_1, χ_2, ..., χ_N and $\sum_{i=1}^N \chi_i^2$ belong to the domain $D(\mathfrak{A})$, it follows that

$$\chi_1 - \chi_1(x^0), \dots, \chi_N - \chi_N(x^0), \ \sum_{i=1}^N \left(\chi_i - \chi_i(x^0)\right)^2 \in D(\mathfrak{A}).$$

Thus we have

$$\begin{aligned}
&\mathfrak{A}u(x^0) \qquad\qquad (9.44)\\
&= \lim_{t \downarrow 0} \frac{1}{t} \left(T_t u(x^0) - u(x^0)\right) \\
&= \lim_{t \downarrow 0} \frac{1}{t} \left(\int_{\overline{D}} p_t(x^0, dy) u(y) - u(x^0)\right) \\
&= \lim_{t \downarrow 0} \left\{ \frac{1}{t} \left(p_t(x^0, \overline{D}) - 1\right) u(x^0) \right. \\
&\quad + \frac{1}{t} \sum_{i=1}^N \int_D p_t(x^0, dy) \left(\chi_i(y) - \chi_i(x^0)\right) \frac{\partial u}{\partial x_i}(x^0) \\
&\quad \left. + \frac{1}{t} \int_{\overline{D}} p_t(x^0, dy) \left[u(y) - u(x^0) - \sum_{i=1}^N \frac{\partial u}{\partial x_i}(x^0) \left(\chi_i(y) - \chi_i(x^0)\right) \right] \right\} \\
&= c(x^0) u(x^0) + \sum_{i=1}^N b^i(x^0) \frac{\partial u}{\partial x_i}(x^0) \\
&\quad + \lim_{t \downarrow 0} \frac{1}{t} \int_{\overline{D} \setminus \{x^0\}} p_t(x^0, dy) \tilde{u}(x^0, y) d(x^0, y),
\end{aligned}$$

where

$$c(x^0) = \lim_{t \downarrow 0} \frac{T_t 1(x^0) - 1}{t} = \mathfrak{A}1(x^0),$$

$$b^i(x^0) = \lim_{t \downarrow 0} \frac{T_t(\chi_i - \chi_i(x^0))(x^0)}{t} = \mathfrak{A}(\chi_i - \chi_i(x^0))(x^0),$$

and

$$d(x^0, y) = \sum_{i=1}^{N} \left(\chi_i(y) - \chi_i(x^0)\right)^2,$$

$$\tilde{u}(x^0, y) = \frac{u(y) - u(x^0) - \sum_{i=1}^{N} \frac{\partial u}{\partial x_i}(x^0)(\chi_i(y) - \chi_i(x^0))}{d(x^0, y)}, \quad y \in \overline{D} \setminus \{x^0\}.$$

To rewrite the last term of 9.44, we define a non-negative measure $\tilde{p}_t(x^0, \cdot)$ on $\overline{D}$ by the formula

$$\tilde{p}_t(x^0, E) = \frac{1}{t} \int_E p_t(x^0, dy) d(x^0, y) \quad \text{for all } E \in \mathcal{B}_D.$$

Here and in the following $\mathcal{B}_K$ denotes the σ-algebra of all Borel sets in a metric space K. Then we can rewrite (9.44) as follows:

$$\mathfrak{A}u(x^0) = c(x^0)u(x^0) + \sum_{i=1}^{N} b^i(x^0) \frac{\partial u}{\partial x_i}(x^0) + \lim_{t \downarrow 0} \int_D \tilde{p}_t(x^0, dy)\tilde{u}(x^0, y). \tag{9.44$'$}$$

Note that, for all sufficiently small $t > 0$,

$$\begin{aligned} \tilde{p}_t(x^0, \overline{D}) &\leq \lim_{t \downarrow 0} \tilde{p}_t(x^0, \overline{D}) + 1 = \lim_{t \downarrow 0} \frac{1}{t} \int_{\overline{D}} p_t(x^0, dy) d(x^0, y) + 1 \qquad (9.45) \\ &= \mathfrak{A} \left(\sum_{i=1}^{N} \left(\chi_i - \chi_i(x^0)\right)^2 \right)(x^0) + 1. \end{aligned}$$

Step 2: Now we introduce a compactification of the space $\overline{D} \setminus \{x^0\}$ to which the function $\tilde{u}(x^0, \cdot)$ may be continuously extended. We let

$$z^{ij}(x^0, y) = \frac{(\chi_i(y) - \chi_i(x^0))(\chi_j(y) - \chi_j(x^0))}{d(x^0, y)}, \quad y \in \overline{D} \setminus \{x^0\}.$$

Then it is easy to see that the functions $z^{ij}(x^0, \cdot)$ satisfy the condition

$$|z^{ij}(x^0, y)| \leq 1,$$

and that the matrix $(z^{ij}(x^0, \cdot))$ is symmetric and positive semi-definite. We define a compact subspace M of symmetric, positive semi-definite matrices by the formula

$$M = \{(z^{ij})_{1 \leq i, j \leq N} : z^{ij} = z^{ji}, \ (z^{ij}) \geq 0, \ |z^{ij}| \leq 1\},$$

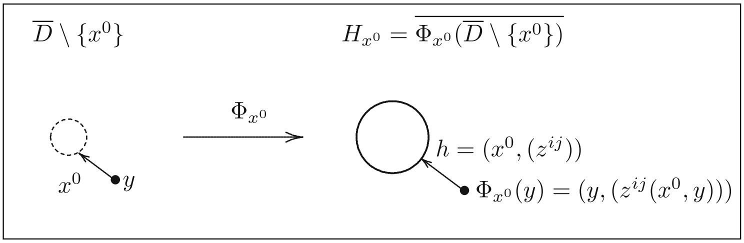


Fig. 9.11 The compactification H_{x^0} of $\overline{D} \setminus \{x^0\}$

and consider an injection

$$\Phi_{x^0} : \overline{D} \setminus \{x^0\} \ni y \longmapsto \big(y, (z^{ij}(x^0, y))\big) \in \overline{D} \times M.$$

Then the function $\tilde{u}(x^0, \Phi_{x^0}^{-1}(\cdot))$, defined on $\Phi_{x^0}(\overline{D} \setminus \{x^0\})$, can be extended to a continuous function $\hat{u}(x, \cdot)$ on the closure

$$H_{x^0} = \overline{\Phi_{x^0}(\overline{D} \setminus \{x^0\})}$$

of $\Phi_{x^0}(\overline{D} \setminus \{x^0\})$ in $\overline{D} \times M$. Indeed, by using Taylor's formula we have, in a neighborhood of x^0,

$$\begin{aligned} u(y) &= u(x^0) + \sum_{i=1}^{N} \frac{\partial u}{\partial x_i}(x^0) \left(\chi_i(y) - \chi_i(x^0)\right) \\ &\quad + \sum_{i,j=1}^{N} \int_0^1 \frac{\partial^2 u}{\partial x_i \partial x_j}(x^0 + \theta(y - x^0))(1 - \theta)\, d\theta \\ &\qquad \times \left(\chi_i(y) - \chi_i(x^0)\right) \left(\chi_j(y) - \chi_j(x^0)\right), \end{aligned}$$

and hence

$$\begin{aligned} \tilde{u}(x^0, y) &= \sum_{i,j=1}^{N} \int_0^1 \frac{\partial^2 u}{\partial x_i \partial x_j}(x^0 + \theta(y - x^0))(1 - \theta) d\theta z^{ij}(x^0, y) \\ &\longrightarrow \hat{u}(x, h) = \frac{1}{2} \sum_{i,j=1}^{N} \frac{\partial^2 u}{\partial x_i \partial x_j}(x^0) z^{ij}, \end{aligned} \tag{9.46}$$

as $\Phi_{x^0}(y) = (y, (z^{ij}(x^0, y))) \to h = (x^0, (z^{ij}))$ (see Fig. 9.11).
We define a non-negative measure $\hat{p}_t(x^0, \cdot)$ on H_{x^0} by the formula

$$\hat{p}_t(x^0, \hat{E}) = \tilde{p}_t(x^0, \Phi_{x^0}^{-1}(\hat{E})) \quad \text{for all } \hat{E} \in \mathcal{B}_{H_{x^0}}. \tag{9.47}$$

Then it follows from inequality (9.36) that we have, for all sufficiently small $t > 0$,

$$\hat{p}_t(x^0, H_{x^0}) \leq \tilde{p}_t(x^0, \overline{D}) \leq \mathfrak{A}\left(\sum_{i=1}^{N}(\chi_i - \chi_i(x^0))^2\right)(x^0) + 1.$$

Hence, by applying Theorem 3.48 to our situation we obtain that there exists a sequence $\{t_n\}$, $t_n \downarrow 0$, such that the measures $\hat{p}_{t_n}(x^0, \cdot)$ converge weakly to a finite non-negative Borel measure $\hat{p}(x, \cdot)$ on H_{x^0}.
Therefore, in view of (9.46) and (9.47), we can pass to the limit in (9.44″) to obtain the following:

$$\begin{aligned} &\mathfrak{A}u(x^0) \qquad (9.44'') \\ &= c(x^0)u(x^0) + \sum_{i=1}^{N} b^i(x^0)\frac{\partial u}{\partial x_i}(x^0) + \lim_{t\downarrow 0}\int_{\overline{D}\setminus\{x^0\}} \tilde{p}_t(x^0, dy)\tilde{u}(x^0, y) \\ &= c(x^0)u(x^0) + \sum_{i=1}^{N} b^i(x^0)\frac{\partial u}{\partial x_i}(x^0) + \lim_{n\to\infty}\int_{H_{x^0}} \hat{p}_{t_n}(x^0, dh)\hat{u}(x, h) \\ &= c(x^0)u(x^0) + \sum_{i=1}^{N} b^i(x^0)\frac{\partial u}{\partial x_i}(x^0) + \int_{H_{x^0}} \hat{p}(x, dh)\hat{u}(x, h). \end{aligned}$$

To rewrite the last term of (9.44″), we define a non-negative Borel measure $\tilde{p}(x^0, \cdot)$ on $\overline{D} \setminus \{x^0\}$ by the formula

$$\tilde{p}(x^0, E) = \hat{p}(x, \Phi_{x^0}(E)) \quad \text{for all } E \in \mathcal{B}_{\overline{D}\setminus\{x^0\}},$$

and let

$$Z : \overline{D} \times M \ni h = (y, (z^{ij})) \longmapsto (z^{ij}) \in M.$$

Then we have

$$\begin{aligned} &\int_{H_{x^0}} \hat{p}(x, dh)\hat{u}(x, h) \qquad (9.48) \\ &= \int_{H_{x^0}\setminus\Phi_{x^0}(\overline{D}\setminus\{x^0\})} \hat{p}(x, dh)\hat{u}(x, h) + \int_{\Phi_{x^0}(\overline{D}\setminus\{x^0\})} \hat{p}(x, dh)\hat{u}(x, h) \\ &= \frac{1}{2}\sum_{i,j=1}^{N}\int_{H_{x^0}\setminus\Phi_{x^0}(\overline{D}\setminus\{x^0\})} \hat{p}(x, dh)Z^{ij}(h)\frac{\partial^2 u}{\partial x_i \partial x_j}(x^0) \\ &\quad + \int_{\overline{D}\setminus\{x^0\}} \tilde{p}(x^0, dy)\tilde{u}(x^0, y) \end{aligned}$$

$$= \sum_{i,j=1}^{N} a^{ij}(x^0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x^0)$$
$$+ \int_D e(x^0, dy) \left[u(y) - u(x^0) - \sum_{i=1}^{N} \frac{\partial u}{\partial x_i}(x^0) \left(\chi_i(y) - \chi_i(x^0) \right) \right],$$

where

$$a^{ij}(x^0) = \frac{1}{2} \int_{H_{x^0} \setminus \Phi_{x^0}(\overline{D} \setminus \{x^0\})} \hat{p}(x, dh) Z^{ij}(h),$$

and

$$e(x^0, \{x^0\}) = 0,$$
$$e(x^0, E) = \int_{E \setminus \{x^0\}} \hat{p}(x^0, dy) \left(\frac{1}{d(x^0, y)} \right), \quad E \in \mathcal{B}_{\overline{D}}.$$

Therefore, by combining (9.44″) and (9.48) we obtain the desired expression (9.42) for $\mathfrak{A}u$ in the interior D of $\overline{D}$.

Step 3: Properties (9.43) follow from our construction of $a^{ij}(x^0)$, $b^i(x^0)$, $c(x^0)$ and $e(x^0, \cdot)$.

The proof of Theorem 9.52 is now complete.

Theorem 9.52 asserts that the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D}$ is written in the interior D of $\overline{D}$ as the sum W of a degenerate elliptic differential operator P of second order and an integro-differential operator S:

$$\begin{aligned} Wu(x) &= Pu(x) + Su(x) \qquad (9.49) \\ &:= \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\ &\quad + \int_{\overline{D}} e(x, dy) \left[u(y) - u(x) - \sum_{i=1}^{N} \frac{\partial u}{\partial x_i}(x) \left(\chi_i(y) - \chi_i(x) \right) \right]. \end{aligned}$$

The differential operator P is called a *diffusion operator* which describes analytically a strong Markov process with continuous paths in the interior D such as Brownian motion, and the functions $a^{ij}(x)$, $b^i(x)$ and $c(x)$ are called the diffusion coefficients, the drift coefficients and the termination coefficient, respectively. The operator S is called a second-order *Lévy operator* which corresponds to the jump phenomenon in the interior D; a Markovian particle moves by jumps to a random point, chosen with kernel $e(x, dy)$, in the interior D. Therefore, the operator $W = P + S$, called a *Waldenfels integro-differential operator* or simply a

Waldenfels operator, corresponds to the physical phenomenon where a Markovian particle moves both by jumps and continuously in the state space D (see Fig. 1.2).

Intuitively, the above result may be interpreted as follows: By Theorems 9.33 and 9.5, there correspond to a Feller semigroup $\{T_t\}_{t\geq 0}$ a unique transition function p_t and a Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ in the following manner:

$$T_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y), \quad f \in C(\overline{D});$$

$$p_t(x, E) = P_x\{x_t \in E\}, \quad E \in \mathcal{B}_{\overline{D}}.$$

In view of Theorem 9.20 and Remark 9.21, it will be true that if the paths of $\mathcal{X}$ are continuous, then the transition function p_t has local character such as condition (N) of Theorem 9.20; hence the infinitesimal generator $\mathfrak{A}$ is local, that is, the value $\mathfrak{A}u(x^0)$ at an interior point x^0 is determined by the values of u in an arbitrary small neighborhood of x^0. However, it is well known (see Theorem 5.8) that a linear operator is local if and only if it is a differential operator. Therefore, we have the following assertion: *The infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}_{t\geq 0}$ on $\overline{D}$ is a differential operator in the interior D of $\overline{D}$ if the paths of its corresponding Markov process $\mathcal{X}$ are continuous.*

In the general case when the paths of $\mathcal{X}$ may have discontinuities such as jumps, the infinitesimal generator $\mathfrak{A}$ takes the form of the sum W of a differential operator P and an integro-differential (non-local) operator S, as proved in Theorem 9.52.

9.5 Infinitesimal Generators of Feller Semigroups on a Bounded Domain (ii)

In this section, we shall derive an explicit formula on the boundary ∂D of $\overline{D}$ for the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}_{t\geq 0}$ on the closure $\overline{D}$ (Theorem 9.53).

Let D be a bounded domain in Euclidean space $\mathbf{R}^N$, with smooth boundary ∂D, and choose, for each point x' of ∂D, a neighborhood U of x' in $\mathbf{R}^N$ and a local coordinate system $(x_1, \ldots, x_{N-1}, x_N)$ on U such that (see Fig. 9.12)

$$x \in U \cap D \Longleftrightarrow x \in U, \; x_N(x) > 0;$$

$$x \in U \cap \partial D \Longleftrightarrow x \in U, \; x_N(x) = 0,$$

and such that the functions $(x_1, x_2, \ldots, x_{N-1})$, restricted to $U \cap \partial D$, form a local coordinate system of ∂D on $U \cap \partial D$ (see Sect. 7.1). Furthermore, we may assume that the functions $x_1, x_2, \ldots, x_{N-1}, x_N$ can be extended to smooth functions $\chi_1, \ldots, \chi_{N-1}, \chi_N$ on $\mathbf{R}^N$, respectively, so that

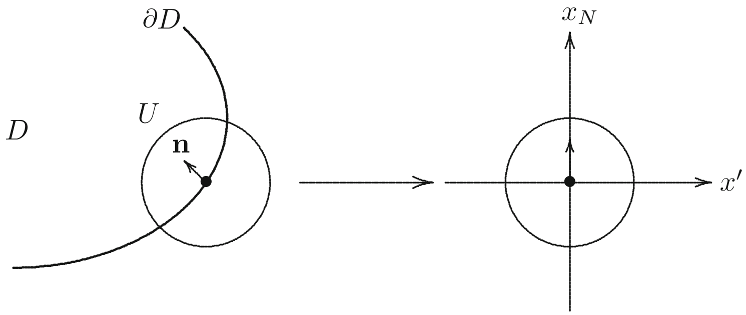


Fig. 9.12 The local coordinate system (x', x_N)

$$d(x', y) = \chi_N(y) + \sum_{i=1}^{N-1} \left(\chi_i(y) - \chi_i(x')\right)^2 > 0 \tag{9.50}$$

if $x' \in U \cap \partial D$ and $y \in \overline{D} \setminus \{x'\}$.

The next theorem, due to Ventcel' [We], asserts that every C^2 function in the domain $D(\mathfrak{A})$ of $\mathfrak{A}$ must obey a boundary condition at each point of ∂D [Ta5, Theorem 9.5.1]:

Theorem 9.53. *Let D be a bounded domain in $\mathbf{R}^N$, with smooth boundary ∂D, and let $\{T_t\}_{t\ge 0}$ be a Feller semigroup on $\overline{D}$ and $\mathfrak{A}$ its infinitesimal generator. Then every function u in $D(\mathfrak{A}) \cap C^2(\overline{D})$ satisfies, at each point x' of ∂D, the boundary condition of the form*

$$\begin{aligned} &\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x') \\ &+ \gamma(x') u(x') + \mu(x') \frac{\partial u}{\partial x_N}(x') - \delta(x') \mathfrak{A} u(x') \\ &+ \int_{\overline{D}} \nu(x', dy) \left[u(y) - u(x') - \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x') \left(\chi_i(y) - \chi_i(x')\right) \right] = 0. \end{aligned} \tag{9.51}$$

Here:

(1) The matrix $(\alpha^{ij}(x'))$ is symmetric and positive semi-definite.
(2) $\gamma(x') \le 0$.
(3) $\mu(x') \ge 0$.
(4) $\delta(x') \ge 0$.
(5) $\nu(x', \cdot)$ is a non-negative Borel measure on $\overline{D}$ such that, for any neighborhood V of x' in $\mathbf{R}^N$,

$$\nu(x', \overline{D} \setminus V) < \infty, \tag{9.52a}$$

$$\int_{V \cap \overline{D}} \nu(x', dy) \left[\chi_N(y) + \sum_{i=1}^{N-1} \left(\chi_i(y) - \chi_i(x') \right)^2 \right] < \infty. \tag{9.52b}$$

Proof. The proof is essentially the same as that of Theorem 9.52, and is divided into five steps.

Step 1: By Theorem 9.33, there corresponds to a Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D}$ a unique uniformly stochastically continuous Feller transition function p_t on $\overline{D}$ in the following manner:

$$T_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y) \quad \text{for all } f \in C(\overline{D}).$$

Thus we have

$$\begin{aligned} &\frac{1}{t}\left(T_t u(x') - u(x')\right) \qquad (9.53)\\ &= \frac{1}{t}\left(p_t(x', \overline{D}) - 1\right) u(x') + \frac{1}{t} \sum_{i=1}^{N-1} \int_{\overline{D}} p_t(x', dy) \left(\chi_i(y) - \chi_i(x')\right) \frac{\partial u}{\partial x_i}(x') \\ &\quad + \frac{1}{t} \int_{\overline{D}} p_t(x', dy) \left[u(y) - u(x') - \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x') \left(\chi_i(y) - \chi_i(x')\right) \right] \\ &= \gamma_t(x') u(x') + \sum_{i=1}^{N-1} \beta_t^j(x') \frac{\partial u}{\partial x_i}(x') + \frac{1}{t} \int_{\overline{D}} p_t(x', dy) \tilde{u}(x', y) d(x', y), \end{aligned}$$

where

$$\gamma_t(x') = \frac{1}{t}\left(p_t(x', \overline{D}) - 1\right),$$

$$\beta_t^j(x') = \frac{1}{t} \int_{\overline{D}} p_t(x', dy) \left(\chi_j(y) - \chi_j(x')\right),$$

and

$$d(x', y) = \chi_N(y) + \sum_{i=1}^{N-1} \left(\chi_i(y) - \chi_i(x')\right)^2, \quad y \in \overline{D},$$

$$\tilde{u}(x', y) = \frac{u(y) - u(x') - \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x')(\chi_i(y) - \chi_i(x'))}{d(x', y)}, \quad y \in \overline{D} \setminus \{x'\}.$$

We rewrite the last term of (9.53). To do this, we introduce a non-negative function

$$\ell_t(x') = \frac{1}{t}\int_{\overline{D}} p_t(x', dy) d(x', y),$$

and consider two cases.

Case A: $\ell_t(x') > 0$. In this case we can write

$$\frac{1}{t}\int_{\overline{D}\setminus\{x'\}} p_t(x', dy)\tilde{u}(x', y) d(x', y) = \ell_t(x')\int_{\overline{D}\setminus\{x'\}} \tilde{q}_t(x', dy)\tilde{u}(x', y),$$

where

$$\tilde{q}_t(x', E) = \frac{1}{t\ell_t(x')}\int_E p_t(x', dy) d(x', y), \quad E \in \mathcal{B}_{\overline{D}}.$$

Here and in the following $\mathcal{B}_K$ denotes the σ-algebra of all Borel sets in K. We note that

$$\begin{aligned}\tilde{q}_t(x', \overline{D}\setminus\{x'\}) &= \frac{1}{t\ell_t(x')}\int_{\overline{D}\setminus\{x'\}} p_t(x', dy) d(x', y)\\ &= \frac{\int_{\overline{D}\setminus\{x'\}} p_t(x', dy) d(x', y)}{\int_{\overline{D}} p_t(x', dy) d(x', y)} = 1,\end{aligned}$$

since it follows from condition (9.50) that $d(x', x') = 0$.

Case B: $\ell_t(x') = 0$. In this case, we have

$$0 = \ell_t(x') = \frac{1}{t}\int_{\overline{D}} p_t(x', dy) d(x', y),$$

and so, by condition (9.50),

$$p_t(x', \overline{D}\setminus\{x'\}) = 0.$$

Hence we can write

$$\frac{1}{t}\int_{\overline{D}\setminus\{x'\}} p_t(x', dy)\tilde{u}(x', y) d(x', y) = \ell_t(x')\int_{\overline{D}\setminus\{x'\}} \tilde{q}_t(x', dy)\tilde{u}(x', y) = 0,$$

where (for example)

$$\tilde{q}_t(x', \cdot) = \text{the unit mass at a point of } D,$$

so that

$$\tilde{q}_t(x', \overline{D} \setminus \{x'\}) = 1.$$

Summing up, we obtain from Case A and Case B that

$$\frac{1}{t}\left(T_t u(x') - u(x')\right) \tag{9.53$'$}$$
$$= \gamma_t(x')u(x') + \sum_{j=1}^{N-1} \beta_t^j(x') \frac{\partial u}{\partial x_j}(x') + \ell_t(x') \int_{\overline{D}\setminus\{x'\}} \tilde{q}_t(x', dy)\tilde{u}(x', y).$$

Step 2: Now we introduce a compactification of $\overline{D} \setminus \{x'\}$ to which the function $\tilde{u}(x', \cdot)$ may be continuously extended. We let

$$w(x', y) = \frac{\chi_N(y)}{d(x', y)}, \quad y \in \overline{D} \setminus \{x'\},$$
$$z^{ij}(x', y) = \frac{(\chi_i(y) - \chi_i(x'))(\chi_j(y) - \chi_j(x'))}{d(x', y)}, \quad y \in \overline{D} \setminus \{x'\}.$$

Then it is easy to see that the functions $w(x', \cdot)$ and $z^{ij}(x', \cdot)$ satisfy the conditions

$$0 \leq w(x', y) \leq 1,$$
$$|z^{ij}(x', y)| \leq 1,$$
$$w(x', y) + \sum_{i=1}^{N-1} z^{ii}(x', y) = 1,$$

and the matrix $(z^{ij}(x', \cdot))$ is symmetric and positive semi-definite. We define a compact subspace M of symmetric, positive semi-definite matrices by

$$M = \{(z^{ij})_{1 \leq i,j \leq N-1} : z^{ij} = z^{ji},\ (z^{ij}) \geq 0,\ |z^{ij}| \leq 1\},$$

and a compact subspace H of $\overline{D} \times [0, 1] \times M$ by the formula

$$H = \left\{ (y, w, (z^{ij})) \in \overline{D} \times [0, 1] \times M : w + \sum_{i=1}^{N-1} z^{ii} = 1 \right\}, \tag{9.54}$$

and consider an injection

$$\Phi_{x'} : \overline{D} \setminus \{x'\} \ni y \longmapsto \left(y, w(x', y), (z^{ij}(x', y))\right) \in H.$$

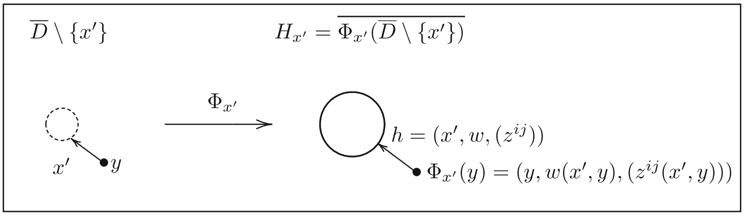


Fig. 9.13 The compactification $H_{x'}$ of $\overline{D} \setminus \{x'\}$

Then the function $\tilde{u}(x', \Phi_{x'}(\cdot))$, defined on $\Phi_{x'}(\overline{D} \setminus \{x'\})$, can be extended to a continuous function $\hat{u}(x', \cdot)$ on the closure

$$H_{x'} = \overline{\Phi_{x'}(\overline{D} \setminus \{x'\})},$$

of $\Phi_{x'}(\overline{D} \setminus \{x'\})$ in H. Indeed, by using Taylor's formula we have, in a neighborhood of x',

$$\begin{aligned} u(y) = u(x') &+ \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x') \left(\chi_i(y) - \chi_i(x')\right) + \frac{\partial u}{\partial x_N}(x') \chi_N(y) \\ &+ \sum_{i,j=1}^{N} \int_0^1 \frac{\partial^2 u}{\partial x_i \partial x_j}(x' + \theta(y - x'))(1 - \theta)\, d\theta \\ &\times \left(\chi_i(y) - \chi_i(x')\right) \left(\chi_j(y) - \chi_j(x')\right), \end{aligned}$$

and hence (see Fig. 9.13)

$$\begin{aligned} \tilde{u}(x', y) &= \frac{\partial u}{\partial x_N}(x') w(x', y) \\ &\quad + \sum_{i,j=1}^{N} \int_0^1 \frac{\partial^2 u}{\partial x_i \partial x_j}(x' + \theta(y - x'))(1 - \theta) d\theta \times z^{ij}(x', y) \\ &\quad \longrightarrow \hat{u}(x', h) = \frac{\partial u}{\partial x_N}(x') w + \frac{1}{2} \sum_{i,j=1}^{N} \frac{\partial^2 u}{\partial x_i \partial x_j}(x') z^{ij}, \end{aligned} \tag{9.55}$$

as

$$\Phi_{x'}(y) = \left(y, w(x', y), (z^{ij}(x', y))\right) \longrightarrow h = (x', w, (z^{ij})).$$

We define a non-negative measure $\hat{q}_t(x', \cdot)$ on $H_{x'}$ by the formula

$$\hat{q}_t(x', \hat{E}) = \tilde{q}_t(x', \Phi_{x'}^{-1}(\hat{E})) \quad \text{for all } \hat{E} \in \mathcal{B}_{H_{x'}}.$$

Then we can write (9.53′) as follows:

$$\frac{1}{t}\left(T_t u(x') - u(x')\right) \tag{9.53''}$$
$$= \gamma_t(x')u(x') + \sum_{j=1}^{N-1} \beta_t^j(x')\frac{\partial u}{\partial x_j}(x') + \ell_t(x')\int_{H_{x'}} \hat{q}_t(x', dh)\hat{u}(x', h).$$

We remark that the measure $\hat{q}_t(x', \cdot)$ is a probability measure on $H_{x'}$.

Step 3: We pass to the limit in (9.53″). To do this, we introduce non-negative functions

$$\theta_m(x') = -\gamma_{1/m}(x') + \sum_{j=1}^{N-1} |\beta_{1/m}^j(x')| + \ell_{1/m}(x'), \tag{9.56}$$
$$m = 1, 2, \ldots,$$

and consider two cases.

Case I: $\liminf_{m\to\infty} \theta_m(x') = 0$. In this case, there exists a subsequence $\{\theta_{m_k}(x')\}$ of $\{\theta_m(x')\}$ such that

$$\lim_{k\to\infty} \theta_{m_k}(x') = 0.$$

Thus, by passing to the limit in (9.53″) with $t := 1/m_k$ we obtain that

$$\mathfrak{A}u(x') = 0.$$

Hence we have condition (9.51), if we take

$$\alpha^{ij}(x') = \beta^i(x') = \gamma(x') = \mu(x') = 0,$$
$$\delta(x') = 1,$$
$$\nu(x', dx) = 0.$$

Case II: $\liminf_{m\to\infty} \theta_m(x') > 0$. In this case, there exist a subsequence $\{\theta_{m_k}(x')\}$ of $\{\theta_m(x')\}$ and a function $\theta(x')$ such that

$$\lim_{k\to\infty} \theta_{m_k}(x') = \theta(x') > 0. \tag{9.57}$$

Then, by dividing both sides of (9.53″) with $t := 1/m_k$ by the function $\theta_{m_k}(x')$ we obtain that

$$\begin{aligned}
&\overline{\delta}_k(x')\left(\frac{T_{t_k}u(x')-u(x')}{t_k}\right) \qquad (9.58)\\
&= \overline{\gamma}_k(x')u(x') + \sum_{j=1}^{N-1}\overline{\beta}_k^j(x')\frac{\partial u}{\partial x_j}(x') + \overline{\ell}_k(x')\int_{H_{x'}}\overline{q}_k(x',dh)\hat{u}(x',h),
\end{aligned}$$

where

$$\begin{aligned}
t_k &= \frac{1}{m_k},\\
\overline{\delta}_k(x') &= \frac{1}{\theta_{m_k}(x')}, \quad \overline{\gamma}_k(x') = \frac{\gamma_{t_k}(x')}{\theta_{m_k}(x')},\\
\overline{\beta}_k^j(x') &= \frac{\beta_{t_k}^j(x')}{\theta_m(x')}, \quad \overline{\ell}_k(x') = \frac{\ell_{t_k}(x')}{\theta_{m_k}(x')},\\
\overline{q}_k(x',\cdot) &= \hat{q}_{t_k}(x',\cdot).
\end{aligned}$$

However, we have, by (9.57),

$$0 \le \overline{\delta}_k(x') < \infty,$$

and further, by (9.55),

$$\begin{aligned}
&0 \le -\overline{\gamma}_k(x') \le 1, \quad -1 \le \overline{\beta}_k^j(x') \le 1, \quad 0 \le \overline{\ell}_k(x') \le 1,\\
&-\overline{\gamma}_k(x') + \sum_{j=1}^{N-1}|\overline{\beta}_k^j(x')| + \overline{\ell}_k(x') = 1.
\end{aligned}$$

We remark that the measures $\overline{q}_k(x',\cdot)$ are probability measures on $H_{x'}$.

Since the metric spaces $[0,+\infty]$, $[0,1]$ and $[-1,1]$ are compact and since the space of probability measures on $H_{x'}$ is also compact (see Theorem 3.48), we can pass to the limit in (9.58) to obtain the following:

$$\begin{aligned}
&\delta(x')\mathfrak{A}u(x') \qquad (9.59)\\
&= \gamma(x')u(x') + \sum_{j=1}^{N-1}\beta^j(x')\frac{\partial u}{\partial x_j}(x') + \ell(x')\int_{H_{x'}}\hat{q}(x',dh)\hat{u}(x',h).
\end{aligned}$$

Here the functions $\delta(x')$, $\gamma(x')$, $\beta^j(x')$ and $\ell(x')$ satisfy the conditions

$$\begin{aligned}
&0 \le \delta(x') < \infty,\\
&0 \le -\gamma(x') \le 1,
\end{aligned}$$

$$
\begin{aligned}
&-1 \leq \beta^j(x') \leq 1, \\
&0 \leq \ell(x') \leq 1,
\end{aligned}
$$

and

$$
-\gamma(x') + \sum_{j=1}^{N-1} |\beta^j(x')| + \ell(x') = 1, \tag{9.60}
$$

and the measure $\hat{q}(x', \cdot)$ is a probability measure on $H_{x'}$.

To rewrite the last term of formula (9.59), we define a non-negative Borel measure $\tilde{q}(x', \cdot)$ on $\overline{D} \setminus \{x'\}$ by the formula

$$
\tilde{q}(x', E) = \hat{q}(x', \Phi_{x'}(E)) \quad \text{for all } E \in \mathcal{B}_{\overline{D}\setminus\{x'\}},
$$

and let

$$
\begin{aligned}
W &: \overline{D} \times [0,1] \times M \ni h = (y, w, (z^{ij})) \longmapsto w \in [0,1], \\
Z &: \overline{D} \times [0,1] \times M \ni h = (y, w, (z^{ij})) \longmapsto (z^{ij}) \in M.
\end{aligned}
$$

Then, in view of assertion (9.55) it follows that

$$
\begin{aligned}
&\ell(x') \int_{H_{x'}} \hat{q}(x', dh) \hat{u}(x', h) \qquad (9.61) \\
&= \ell(x') \int_{H_{x'}\setminus \Phi_{x'}(\overline{D}\setminus\{x'\})} \hat{q}(x', dh) \hat{u}(x', h) + \ell(x') \int_{\Phi_{x'}(\overline{D}\setminus\{x'\})} \hat{q}(x', dh) \hat{u}(x', h) \\
&= \ell(x') \Bigg\{ \int_{H_{x'}\setminus \Phi_{x'}(\overline{D}\setminus\{x'\})} \hat{q}(x', dh) W(h) \frac{\partial u}{\partial x_N}(x') \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{N} \int_{H_{x'}\setminus \Phi_{x'}(\overline{D}\setminus\{x'\})} \hat{q}(x', dh) Z^{ij}(h) \frac{\partial^2 u}{\partial x_i \partial x_j}(x') \Bigg\} \\
&\quad + \ell(x') \int_{\overline{D}\setminus\{x'\}} \tilde{q}(x', dy) \tilde{u}(x', y) \\
&= \mu(x') \frac{\partial u}{\partial x_N} + \sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') \\
&\quad + \int_{\overline{D}} \nu(x', dy) \left[u(y) - u(x') - \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x') \left(\chi_i(y) - \chi_i(x')\right) \right],
\end{aligned}
$$

where

$$\mu(x') = \ell(x') \int_{H_{x'} \setminus \Phi_{x'}(\overline{D} \setminus \{x'\})} \hat{q}(x', dh) W(h), \tag{9.62}$$

$$\alpha^{ij}(x') = \frac{\ell(x')}{2} \int_{H_{x'} \setminus \Phi_{x'}(\overline{D} \setminus \{x'\})} \hat{q}(x', dh) Z^{ij}(h), \tag{9.63}$$

and

$$\nu(x', \{x'\}) = 0, \tag{9.64a}$$

$$\nu(x', E) = \ell(x') \int_{E \setminus \{x'\}} \hat{q}(x', dh) Z^{ij}(h), \quad E \in \mathcal{B}_{\overline{D}}. \tag{9.64b}$$

Therefore, by combining (9.59) and (9.64) we obtain the desired boundary condition (9.51) in Case II.

Step 4: Properties (9.52a) and (9.52b) follow from our construction of $\alpha^{ij}(x')$, $\beta^i(x')$, $\gamma(x')$, $\mu(x')$, $\delta(x')$ and $\nu(x', \cdot)$.

Step 5: Finally, we show that the boundary condition (9.51) is consistent, that is, condition (9.51) does not take the form $0 = 0$.

In Case I, we have taken

$$\delta(x') = 1.$$

In Case II, we assume that

$$\gamma(x') = \beta^i(x') = 0, \quad \nu(x', \cdot) = 0.$$

Then we have, by Eq. (9.60),

$$\ell(x') = 1,$$

and hence, by (9.64),

$$\hat{q}(x', \Phi_{x'}(\overline{D} \setminus \{x'\})) = \tilde{q}(x', \overline{D} \setminus \{x'\}) = 0.$$

This implies that

$$\hat{q}(x', H_{x'} \setminus \Phi_{x'}(\overline{D} \setminus \{x'\})) = 1,$$

since the measure $\hat{q}(x', \cdot)$ is a probability measure on $H_{x'}$. Therefore, in view of definition (9.54) it follows from (9.62) and (9.63) that

$$\mu(x') + 2\sum_{i=1}^{N-1} \alpha^{ii}(x') = \ell(x') \int_{H_{x'}} \hat{q}(x', dh) \left(W(h) + \sum_{i=1}^{N-1} Z^{ii}(h) \right)$$
$$= \ell(x')\hat{q}(x', H_{x'} \setminus \Phi_{x'}(\overline{D} \setminus \{x'\}))$$
$$= 1.$$

The proof of Theorem 9.53 is now complete.

Probabilistically, Theorems 9.52 and 9.53 may be interpreted as follows: A Markovian particle in a *Markov process* $\mathcal{X}$ on the state space $\overline{D}$ is governed by an integro-differential operator W of the form (9.49) in the interior D of $\overline{D}$, and it obeys a boundary condition L of the form (9.51) on the boundary ∂D of $\overline{D}$:

$$\begin{aligned} Lu(x') &:= \sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x') \\ &\quad + \gamma(x')u(x') + \mu(x') \frac{\partial u}{\partial x_N}(x') - \delta(x') W u(x') \\ &\quad + \int_{\overline{D}} \nu(x', dy) \left[u(y) - u(x') - \sum_{i=1}^{N-1} \frac{\partial u}{\partial x_i}(x') \left(\chi_i(y) - \chi_i(x')\right) \right] \\ &= 0. \end{aligned} \tag{9.65}$$

The pseudo-differential boundary condition L is called a second-order *Ventcel' boundary condition* (cf. [We]). It should be emphasized that the six terms of L

$$\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x'),$$

$$\gamma(x')u(x'), \quad \mu(x') \frac{\partial u}{\partial x_N}(x'), \quad \delta(x') W u(x'),$$

$$\int_{\partial D} \nu(x', dy') \left[u(y') - u(x') - \sum_{j=1}^{N-1} (\chi_j(y) - \chi_j(x')) \right],$$

$$\int_{D} \nu(x', dy) \left[u(y) - u(x') - \sum_{j=1}^{N-1} (\chi_j(y) - \chi_j(x')) \right]$$

correspond to the diffusion phenomenon along the boundary (like Brownian motion on ∂D), the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary, respectively (see Figs. 1.4–1.6).

Analytically, via a version of the Hille–Yosida theorem (Theorem 9.35), Theorems 9.52 and 9.53 may be interpreted as follows: A *Feller semigroup* $\{T_t\}_{t \geq 0}$ on $\overline{D}$ is described by an integro-differential operator W of the form (9.49) and a boundary condition L of the form (9.65). Therefore, we are reduced to the study of *boundary value problems* for Waldenfels integro-differential operators W with Ventcel' boundary conditions L in the theory of partial differential equations.

9.6 Notes and Comments

Chapter 9 is a refinement of Chapter 9 of Taira [Ta5]. The results discussed here are adapted from Blumenthal–Getoor [BG], Dynkin [Dy1, Dy2], Dynkin–Yushkevich [DY], Ethier–Kurtz [EK], Feller [Fe1, Fe2], Ikeda–Watanabe [IW], Itô–McKean, Jr. [IM], Lamperti [La], Revuz–Yor [RY] and Stroock–Varadhan [SV]. In particular, our treatment of temporally homogeneous Markov processes follows the expositions of Dynkin [Dy1, Dy2] and Blumenthal–Getoor [BG]. However, unlike many other books on Markov processes, this chapter focuses on the interrelationship between three subjects: Feller semigroups, transition functions and Markov processes. Our approach to the problem of constructing Markov processes with Ventcel' boundary conditions is distinguished by the extensive use of ideas and techniques characteristic of recent developments in functional analysis.

Section 9.1: Theorem 9.5 is taken from Dynkin [Dy1, Chapter 4, Section 2], while Theorem 9.20 is taken from Dynkin [Dy1, Chapter 6] and [Dy2, Chapter 3, Section 2]. Theorem 9.26 is due to Dynkin [Dy1, Theorem 5.10] and Theorem 9.28 is due to Dynkin [Dy1, Theorem 6.3], respectively. Theorem 9.28 is a non-compact version of Lamperti [La, Chapter 8, Section 3, Theorem 1]. Section 9.1.6 is adapted from Lamperti [La, Chapter 9, Section 2].

On the other hand, Bernstein [Be] and Lévy [Le] made probabilistic approaches to Markov processes with discontinuous paths, in terms of characteristic functions of additive processes (processes with stationary, independent increments having right continuous trajectories with left limits). Indeed, the exponents of these characteristic functions are closely related to Waldenfels integro-differential operators which can be the infinitesimal generators of such jump type processes. By using such an infinitesimal generator, we can obtain an extended form of forward and backward Kolmogorov's equations for the transition density function, as parabolic integro-differential equations.

Section 9.2: The semigroup approach to Markov processes can be traced back to the work of Kolmogorov [Kl]. It was substantially developed in the early 1950s, with Feller [Fe1] and [Fe2] doing the pioneering work. Our presentation here follows the book of Dynkin [Dy2] and also part of Lamperti's [La]. Theorem 9.34 is a non-compact version of Lamperti [La, Chapter 7, Section 7, Theorem 1].

Section 9.3: Theorem 9.50 is due to Sato–Ueno [SU, Theorem 1.2] and Bony–Courrège–Priouret [BCP, Théorème de Hille–Yosida–Ray] (cf. [It, Ra, Ta5]).

Section 9.4: Theorem 9.52 is adapted from Sato–Ueno [SU], while the main idea of its proof is due to Ventcel' [We]. Moreover, Bony, Courrège and Priouret [BCP] give a more precise characterization of the infinitesimal generators of Feller semigroups in terms of the maximum principle (see Chap. 8).

Section 9.5: Theorem 9.53 is due to Ventcel' [We]. We can reconstruct the functions $\alpha^{ij}(x')$, $\beta^{i}(x')$, $\gamma(x')$, $\mu(x')$ and $\delta(x')$ so that they are bounded and Borel measurable on the boundary ∂D (see Bony–Courrège–Priouret [BCP, Théorème XIII]). For the probabilistic meanings of Ventcel' boundary conditions, the reader might refer to Dynkin–Yushkevich [DY]. G. Goldstein [Gg] gives a physical interpretation of general Ventcel' boundary conditions for both the heat and wave equations.

Chapter 10
Semigroups and Boundary Value Problems for Waldenfels Operators

In the early 1950s, W. Feller [Fe1, Fe2] completely characterized the analytic structure of one-dimensional diffusion processes; he gave an intrinsic representation of the infinitesimal generator $\mathfrak{A}$ of a one-dimensional diffusion process and determined all possible boundary conditions which describe the domain $D(\mathfrak{A})$ of the infinitesimal generator $\mathfrak{A}$. The probabilistic meaning of Feller's work was clarified by E.B. Dynkin [Dy1, Dy2], K. Itô and H.P. McKean, Jr. [IM], D. Ray [Ra] and others. One-dimensional diffusion processes are fully understood both from the analytic and probabilistic points of view.

Chapter 10 is the heart of the subject, and is devoted to the functional analytic approach to the problem of constructing (temporally homogeneous) Markov processes with Ventcel' (Wentzell) boundary conditions in probability theory. In Sects. 10.1 and 10.2 we formulate our problem and Theorem 1.2 (Theorem 10.2), generalizing Feller's work to the multi-dimensional case. The approach here is adapted from Bony–Courrège–Priouret [BCP], Cancelier [Cn], Sato–Ueno [SU] and Taira [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9, Ta10].

Our functional analytic approach is as follows: (1) We reduce the problem of existence of Feller semigroups to the unique solvability of boundary value problems for Waldenfels integro-differential operators W with Ventcel' boundary conditions L and then prove existence theorems for Feller semigroups (Theorems 10.2 and 10.21). In Sect. 10.3 we consider Dirichlet problem (D) for Waldenfels integro-differential operators $W = P + S$, and prove that if $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$ has the *transmission property* with respect to ∂D due to Boutet de Monvel [Bo] (see Sect. 7.6), then Dirichlet problem (D) is uniquely solvable in the framework of Hölder spaces (Theorem 10.4). Then, by using the Green operator G^0_α and harmonic operator H_α for Dirichlet problem (D) we can reduce the study of the boundary value problems to that of the classical pseudo-differential operator LH_α on the boundary (Theorem 10.19) in Sect. 10.4. This is a generalization of the classical Fredholm integral equation on the boundary. (2) Sect. 10.5 is devoted to the proof of Theorem 10.21 and hence that of Theorem 1.2.

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_10,

$$\boxed{\text{Theorem 10.21}} \Longrightarrow \boxed{\text{Theorem 10.2}} \Longrightarrow \boxed{\text{Theorem 1.2}}$$

The first essential step in the proof is to show that if $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$ has the *transmission property* with respect to ∂D, then the operator LH_α is the sum of a second-order degenerate elliptic differential operator P_α and a classical pseudo-differential operator S_α with *non-negative distribution kernel* on ∂D. The second essential step in the proof is to calculate the *complete symbol* of the classical pseudo-differential operator LH_α on ∂D. This is carried out in Sect. 10.7 due to its length. In the third essential step in the proof we formulate an existence and uniqueness theorem for a class of pseudo-differential operators in the framework of Hölder spaces (Theorem 10.23) which enters naturally in the study of the pseudo-differential operator LH_α. The proof of Theorem 10.23 is based on a method of *elliptic regularizations* essentially due to Oleĭnik–Radkevič [OR] developed for second-order differential operators with non-negative characteristic form. In order to prove estimate (10.28), we need an interpolation argument. Moreover, we remark that Corollary 3.26 to Mazur's theorem (Theorem 3.25) in Chap. 3 plays an important role in the proof of estimate (10.28). (3) In this way, by using the Hölder space theory of pseudo-differential operators we can prove that if the Ventcel' boundary conditions L are *transversal* on the boundary, then we can verify all the conditions of the generation theorems of Feller semigroups (Theorems 9.35 and 9.50) discussed in Sect. 9.3.

10.1 Formulation of the Problem

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$, with smooth boundary ∂D; its closure $\overline{D} = D \cup \partial D$ is an N-dimensional, compact smooth manifold with boundary. We may assume that $\overline{D}$ is the closure of a relatively compact open subset D of an N-dimensional, compact smooth manifold $\hat{D}$ without boundary in which D has a smooth boundary ∂D. This manifold $\hat{D}$ is called the *double* of D (see Fig. 10.1).

Let $C(\overline{D})$ be the space of real-valued, continuous functions f on $\overline{D}$. We equip the space $C(\overline{D})$ with the topology of uniform convergence on the whole $\overline{D}$; hence it is a Banach space with the maximum norm

$$\|f\|_\infty = \max_{x \in \overline{D}} |f(x)|.$$

A strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on the space $C(\overline{D})$ is called a *Feller semigroup* on $\overline{D}$ if it is non-negative and contractive on $C(\overline{D})$:

$$f \in C(\overline{D}), \ 0 \leq f(x) \leq 1 \quad \text{on } \overline{D} \Longrightarrow 0 \leq T_t f(x) \leq 1 \quad \text{on } \overline{D}.$$

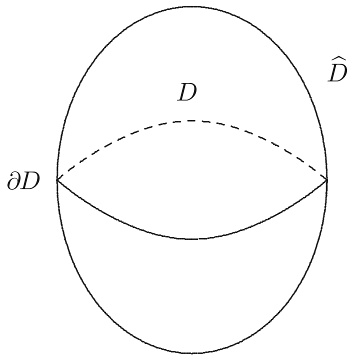


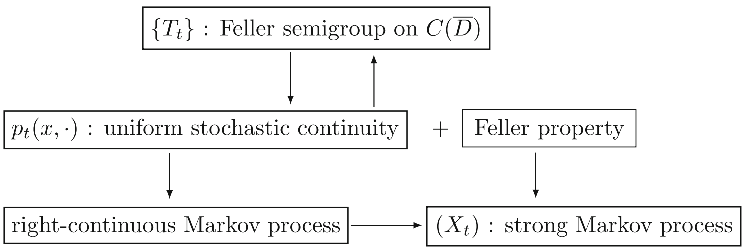
Fig. 10.1 The double $\hat{D}$ of D

It follows from an application of Theorem 9.32 that if T_t is a Feller semigroup on $\overline{D}$, then there exists a unique (temporally homogeneous) Markov transition function $p_t(x, \cdot)$ on $\overline{D}$ such that

$$T_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y) \quad \text{for every } f \in C(\overline{D}).$$

It can be shown (see [Dy2, Chapter III, Section 3]) that the function $p_t(x, \cdot)$ is the transition function of some *strong Markov process* (X_t); hence the value $p_t(x, E)$ expresses the transition probability that a Markovian particle starting at position x will be found in the set E at time t (see Fig. 9.3).

Our approach can be visualized as follows (see Chap. 9):



Furthermore, it is shown in Chap. 9 (see Theorems 9.52 and 9.53) that the infinitesimal generator $\mathfrak{A}$ of a Feller semigroup $\{T_t\}_{t \geq 0}$ is described analytically by a Waldenfels integro-differential operator W and a Ventcel' boundary condition L, which we shall formulate precisely.

Let W be a second-order *elliptic* integro-differential operator with real smooth coefficients such that

$$\begin{aligned} W_D u(x) &= Pu(x) + S_D u(x) \qquad (10.1)\\ &:= \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\ &\quad + \sum_{j=1}^{N} a_\sigma^j(x) \frac{\partial u}{\partial x_j}(x) + a_\sigma(x)u(x) \\ &\quad + \int_D s(x,y) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^{N} (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy, \\ &\qquad x \in D. \end{aligned}$$

Here:

(1) $a^{ij}(x) \in C^\infty(\mathbf{R}^N)$, $a^{ij}(x) = a^{ji}(x)$ for all $x \in \mathbf{R}^N$ and $1 \le i, j \le N$, and there exists a constant $a_0 > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i \xi_j \ge a_0 |\xi|^2 \quad \text{for all } (x, \xi) \in \mathbf{R}^N \times \mathbf{R}^N.$$

(2) $b^i(x) \in C^\infty(\mathbf{R}^N)$ for all $1 \le i \le N$.

(3) $P1(x) = c(x) \in C^\infty(\mathbf{R}^N)$ and $c(x) \le 0$ in D.

(4) The integral kernel $s(x,y)$ is the distribution kernel of a properly supported, classical pseudo-differential operator $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$, $\kappa > 0$, which has the *transmission property* with respect to ∂D, and $s(x,y) \ge 0$ off the *diagonal* $\Delta_{\mathbf{R}^N} = \{(x,x) : x \in \mathbf{R}^N\}$ in $\mathbf{R}^N \times \mathbf{R}^N$. The measure dy is the Lebesgue measure on $\mathbf{R}^N$. Here it should be noticed (see Sect. 7.6) that the operator S_D can be *formally* written in the form

$$S_D v(x) := S(v^0)\big|_D = \int_D s(x,y)v(y)\,dy, \quad x \in D,$$

where v^0 is the extension of v to $\mathbf{R}^N$ by zero outside $\overline{D}$

$$v^0(x) = \begin{cases} v(x) & \text{for } x \in \overline{D}, \\ 0 & \text{for } x \in \mathbf{R}^N \setminus \overline{D}. \end{cases}$$

(5) The function $\sigma(x,y)$ is a local unity function on $\overline{D}$, that is, $\sigma(x,y)$ is a smooth function on $\overline{D} \times \overline{D}$ such that $\sigma(x,y) = 1$ in a neighborhood of the *diagonal* $\Delta_{\overline{D}} = \{(x,x) : x \in \overline{D}\}$ in $\overline{D} \times \overline{D}$ (see Sect. 9.1). The function $\sigma(x,y)$ depends on the shape of the domain D. More precisely, it depends on a family of local

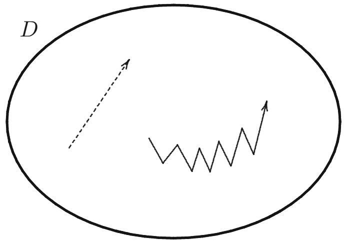


Fig. 10.2 A Markovian particle moves by jumps and continuously

charts on D in each of which the Taylor expansion is valid for functions u. For example, if D is *convex*, we may take $\sigma(x, y) \equiv 1$ on $\overline{D} \times \overline{D}$.

(6) $a_\sigma^j(x) = (S_D \sigma_x^j)(x)$ where $\sigma_x^j(y) = \sigma(x, y)(y_j - x_j)$ for all $1 \leq j \leq N$.

(7) $a_\sigma(x) = (S_D \sigma_x)(x)$ where $\sigma_x(y) = \sigma(x, y)$.

(8) The operator W_D satisfies the condition (1.6)

$$\begin{aligned} W_D 1(x) &= P1(x) + S_D 1(x) \\ &= c(x) + a_\sigma(x) + \int_D s(x, y)\,[1 - \sigma(x, y)]\,dy \leq 0 \quad \text{in } D. \end{aligned}$$

The operator W is called a second-order *Waldenfels integro-differential operator* or simply a *Waldenfels operator* (cf. [Wa]). The differential operator P is called a diffusion operator which describes analytically a strong Markov process with continuous paths in the interior D such as Brownian motion, and the functions $a^{ij}(x)$, $b^i(x)$ and $c(x)$ are called the diffusion coefficients, the drift coefficients and the termination coefficient, respectively. The integro-differential operator

$$S_r u(x) = \int_D s(x, y) \left[u(y) - \sigma(x, y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy$$

is called a second-order *Lévy operator* which corresponds to the jump phenomenon in the interior D; a Markovian particle moves by jumps to a random point, chosen with kernel $s(x, y)$, in the interior D. More precisely, we find from inequalities (10.2) and (10.3) below that the Lévy measure $s(\cdot, y)\,dy$ has a singularity of order $N - \kappa$ at the diagonal $\Delta_{\overline{D}}$, and this singularity at the diagonal is produced by the accumulation of *small jumps* of Markovian particles.

Therefore, the Waldenfels integro-differential operator W corresponds to the physical phenomenon where a Markovian particle moves both by jumps and continuously in the state space D (see Fig. 10.2).

The intuitive meaning of condition (6) is that the jump phenomenon from a point $x \in D$ to the outside of a neighborhood of x in the interior D is "dominated" by the absorption phenomenon at x. In particular, if $a(x) \equiv 0$ in D, then condition (6) implies that any Markovian particle does not move by jumps from $x \in D$ to the outside of a neighborhood $V(x)$ of x in the interior D, since we have

$$\int_D s(x,y)\,[1-\sigma(x,y)]\,dy = 0,$$

and so, by conditions (4) and (5),

$$s(x,y) = 0 \quad \text{for all } y \in D \setminus V(x).$$

It should be noted that the integro-differential operator

$$S_r u(x) = \int_D s(x,y) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy$$

is a “regularization” of S, since the integrand is absolutely convergent (see Example 8.9). Indeed, we can write $S_r u(x)$ in the form

$$\begin{aligned} &S_r u(x) \\ &= \int_D s(x,y) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy \\ &= \int_D s(x,y)\,[1-\sigma(x,y)]\,u(y)\,dy \\ &\quad + \int_D s(x,y)\sigma(x,y) \left(u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) dy. \end{aligned}$$

By using Taylor’s formula

$$\begin{aligned} &u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \\ &= \sum_{i,j=1}^N (y_i - x_i)(y_j - x_j) \left(\int_0^1 (1-t) \frac{\partial^2 u}{\partial x_i \partial x_j}(x + t(y-x))dt \right), \end{aligned}$$

we can find a constant $C_1 > 0$ such that

$$\left| u(y) - u(x) - \sum_{j=1}^N (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right| \leq C_1 |x-y|^2, \quad x, y \in \overline{D}.$$

On the other hand, by arguing as in the proof of Theorem 7.36 in Chap. 7 we can find a constant $C_2 > 0$ such that the kernel $s(x, y)$ of $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$ satisfies the estimate

$$0 \leq s(x, y) \leq \frac{C_K}{|x-y|^{N+2-\kappa}}, \quad x, y \in \overline{D}, \ \ x \neq y.$$

Therefore, we have, with some constant $C_3 > 0$,

$$\left| \int_D s(x, y)\sigma(x, y) \left(u(y) - u(x) - \sum_{j=1}^{N} (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) dy \right| \tag{10.2}$$
$$\leq C_3 \|u\|_{C^2(\overline{D})} \int_D \frac{1}{|x-y|^{N+2-\kappa}} \cdot |x-y|^2 \, dy$$
$$= C_3 \|u\|_{C^2(\overline{D})} \int_D \frac{1}{|x-y|^{N-\kappa}} \, dy.$$

Similarly, we have, with some constant $C_4 > 0$,

$$\left| \int_D s(x, y) \left[1 - \sigma(x, y)\right] u(y) \, dy \right| \leq C_4 \|u\|_{C(\overline{D})} \int_D \frac{1}{|x-y|^{N-\kappa}} \, dy, \tag{10.3}$$

since we have

$$\sigma(x, y) - 1$$
$$= \sigma(x, y) - \sigma(x, x) - \sum_{j=1}^{N} (y_j - x_j) \frac{\partial \sigma}{\partial x_j}(x, x)$$
$$= \sum_{i,j=1}^{N} (y_i - x_i)(y_j - x_j) \left(\int_0^1 (1-t) \frac{\partial^2 \sigma}{\partial x_i \partial x_j}(x, x + t(y-x)) dt \right).$$

In a tubular neighborhood B of the boundary ∂D, we can introduce local coordinate systems (x', t) (see Sect. 7.1) such that $x' = (x_1, x_2, \ldots, x_{N-1})$ give local coordinates for the boundary ∂D and

$$D = \{(x', t) : t > 0\}.$$

We further normalize the coordinates by assuming the curves $x(t) = (x_0', t)$, $x_0' \in \partial D$, are unit speed geodesics perpendicular to ∂D for $0 \leq t < \varepsilon$ (see Fig. 10.3).

Let L be a second-order boundary condition such that, in terms of local coordinates $(x_1, x_2, \ldots, x_{N-1})$,

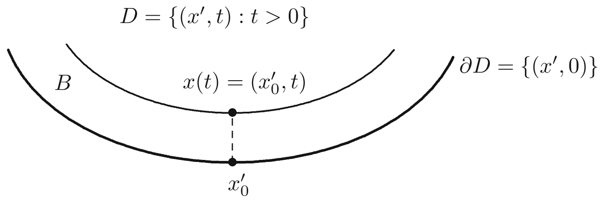


Fig. 10.3 The tubular neighborhood B of ∂D

$$\begin{aligned}
&Lu(x') \qquad (10.4)\\
&= Qu(x') + \mu(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x') W_D u(x') + \Gamma u(x')\\
&:= \left(\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x') + \gamma(x') u(x')\right)\\
&\quad + \mu(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \delta(x') W_D u(x')\\
&\quad + \sum_{j=1}^{N-1} \eta_\tau^j(x') \frac{\partial u}{\partial x_j}(x') + \eta_\tau(x') u(x')\\
&\quad + \int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy'\\
&\quad + \int_D t(x', y) \left[u(y) - u(x') \right] dy, \quad x' \in \partial D,
\end{aligned}$$

where:

(1) The operator Q is a second-order *degenerate* elliptic differential operator on ∂D with non-positive principal symbol. In other words, the $\alpha^{ij}(x')$ are the components of a smooth symmetric contravariant tensor of type $\binom{2}{0}$ on ∂D satisfying the condition

$$\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \xi_i \xi_j \geq 0, \quad x' \in \partial D, \;\; \xi' = \sum_{j=1}^{N-1} \xi_j \, dx_j \in T^*_{x'}(\partial D).$$

Here $T^*_{x'}(\partial D)$ is the cotangent space of ∂D at x'.

(2) $\beta(x') = \sum_{i=1}^{N-1} \beta^i(x') \partial u / \partial x_i$ is a smooth vector field on ∂D (see Fig. 10.4).
(3) $Q1(x') = \gamma(x') \in C^\infty(\partial D)$ and $\gamma(x') \leq 0$ on ∂D.
(4) $\mu(x') \in C^\infty(\partial D)$ and $\mu(x') \geq 0$ on ∂D.
(5) $\delta(x') \in C^\infty(\partial D)$ and $\delta(x') \geq 0$ on ∂D.
(6) $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ is the unit interior normal to the boundary ∂D.

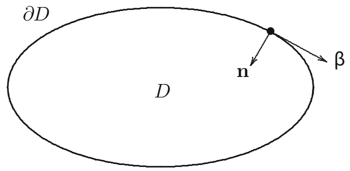


Fig. 10.4 The vector field β and the unit interior normal $\mathbf{n}$ to ∂D

(7) The integral kernel $r(x', y')$ is the distribution kernel of a classical pseudo-differential operator $R \in L^{2-\kappa_1}_{1,0}(\partial D)$, $\kappa_1 > 0$, and $r(x', y') \geq 0$ off the *diagonal* $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$ in $\partial D \times \partial D$. The density dy' is a strictly positive density on ∂D.

(8) The integral kernel $t(x, y)$ is the distribution kernel of a properly supported, classical pseudo-differential operator $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, $\kappa_2 > 0$, which has the *transmission property* with respect to the boundary ∂D, and $t(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N}$.

(9) The function $\tau(x, y)$ is a local unity function on $\overline{D}$; more precisely, $\tau(x, y)$ is a smooth function on $\overline{D} \times \overline{D}$, with compact support in a neighborhood of the diagonal $\Delta_{\partial D}$, such that, at each point x' of ∂D, $\tau(x', y) = 1$ for y in a neighborhood of x' in $\overline{D}$. The function $\tau(x, y)$ depends on the shape of the boundary ∂D.

(10) $\eta^j_\tau(x') = R(\tau^j_{x'})(x')$ where $\tau^j_{x'}(y') = \tau(x', y')(y_j - x_j)$ for all $1 \leq j \leq N - 1$.

(11) $\eta_\tau(x') = R(\tau_{x'})(x')$ where $\tau_{x'}(y') = \tau(x', y')$.

(12) The operator Γ is a boundary condition of order $2 - \kappa_1$, and satisfies the condition (1.6)

$$\begin{aligned} &Q1(x') + \Gamma 1(x') \\ &= \gamma(x') + \eta_\tau(x') + \int_{\partial D} r(x', y') \left[1 - \tau(x', y')\right] dy' \leq 0 \quad \text{on } \partial D. \end{aligned}$$

The boundary condition L is called a second-order *Ventcel' boundary condition* (cf. [We]). The six terms of L

$$\sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial u}{\partial x_i}(x'),$$

$$\gamma(x')u(x'), \quad \mu(x') \frac{\partial u}{\partial \mathbf{n}}(x'), \quad \delta(x') W_D u(x'),$$

$$\int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy',$$

$$\int_D t(x', y) \left[u(y) - u(x')\right] dy$$

correspond to the diffusion phenomenon along the boundary (like Brownian motion on ∂D), the absorption phenomenon, the reflection phenomenon, the sticking (or viscosity) phenomenon and the jump phenomenon on the boundary and the inward jump phenomenon from the boundary, respectively (see Figs. 1.4–1.6).

The intuitive meaning of condition (10) is that the jump phenomenon from a point $x' \in \partial D$ to the outside of a neighborhood of x' on the boundary ∂D is "dominated" by the absorption phenomenon at x'. In particular, if $\gamma(x') \equiv 0$ on ∂D, then condition (10) implies that any Markovian particle does not move by jumps from $x' \in \partial D$ to the outside of a neighborhood $V(x')$ of x' on the boundary ∂D, since we have

$$\int_{\partial D} r(x', y') \left[1 - \tau(x', y')\right] dy' = 0,$$

and so, by conditions (7) and (9),

$$r(x', y') = 0 \quad \text{for all } y' \in \partial D \setminus V(x').$$

It should be noted that the integro-differential operator

$$\begin{aligned}
&\Gamma_r u(x') \\
&= \int_{\partial D} r(x', y') \left[u(y') - \tau(x', y') \left(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x') \right) \right] dy' \\
&\quad \int_D t(x', y) \left[u(y) - u(x')\right] dy, \quad x' \in \partial D,
\end{aligned}$$

is a "regularization" of $R \in L^{2-\kappa_1}_{1,0}(\partial D)$ and $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, since the integrals

$$R_r u(x') = \int_{\partial D} r(x', y') \big[u(y') - \tau(x', y')\big(u(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial u}{\partial x_j}(x')\big)\big] dy',$$

$$T_r u(x') = \int_D t(x', y) \left[u(y) - u(x')\right] dy$$

are both absolutely convergent (see Example 8.10). Indeed, it suffices to note (see Theorem 7.36) that the kernels $r(x', y')$ and $t(x', y)$ satisfy respectively the estimates

$$0 \le r(x', y') \le \frac{C'}{|x' - y'|^{(N-1)+2-\kappa_1}}, \quad x', y' \in \partial D, \ \ x' \neq y', \tag{10.5}$$

$$0 \le t(x', y) \le \frac{C''}{|x' - y|^{N+1-\kappa_2}}, \quad x' \in \partial D, \ \ y \in D, \tag{10.6}$$

where $|x' - y'|$ denotes the geodesic distance between x' and y' with respect to the Riemannian metric of the manifold ∂D. We find from inequalities (10.5) and (10.6) that the Lévy measures $r(\cdot, y')\, dy'$ and $t(\cdot, y)\, dy$ have singularities of order $(N-1)-\kappa_1$ and $N-\kappa_2$, respectively, and these singularities are produced by the accumulation of *small jumps* of Markovian particles.

This chapter is devoted to the functional analytic approach to the problem of constructing (temporally homogeneous) Markov processes with Ventcel' boundary conditions in probability. More precisely, we consider the following problem:

Problem 10.1. Conversely, given analytic data (W, L), can we construct a Feller semigroup $\{T_t\}_{t\geq 0}$ whose infinitesimal generator is characterized by (W, L)?

Our approach is distinguished by the extensive use of ideas and techniques characteristic of recent developments in the theory of partial differential equations. It focuses on the relationship between three interrelated subjects in analysis; Feller semigroups, pseudo-differential operators and elliptic boundary value problems, providing powerful methods for future research (see [Ta3, Ta4, Ta5, Ta6, Ta7, Ta8, Ta9, Ta10]).

10.2 A Generation Theorem for Feller Semigroups on a Bounded Domain

In this section we consider the transversal case and prove Theorem 1.2. The boundary condition L is said to be *transversal* on the boundary ∂D if it satisfies the condition

$$\int_D t(x', y)\, dy = +\infty \quad \text{if } \mu(x') = \delta(x') = 0. \tag{10.7}$$

The intuitive meaning of condition (10.7) is that a Markovian particle jumps away "instantaneously" from the points $x' \in \partial D$ where neither the reflection phenomenon nor the sticking phenomenon (which is similar to the reflection phenomenon) occurs. Probabilistically, this means that every Markov process on the boundary ∂D is the "trace" on ∂D of trajectories of some Markov process on the closure $\overline{D} = D \cup \partial D$ (see Remark 10.22 below).

Intuitively, Theorem 1.2 asserts that there exists a Feller semigroup on $\overline{D}$ corresponding to the physical phenomenon where one of reflection, sticking or inward jump from the boundary occurs at each point of the boundary ∂D (see Fig. 10.5):

The next generation theorem for Feller semigroups proves Theorem 1.2:

Theorem 10.2. *We define a linear operator* $\mathfrak{A}$ *from the space* $C(\overline{D})$ *into itself as follows:*

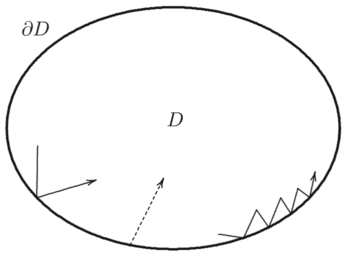


Fig. 10.5 The intuitive meaning of Theorem 1.2

(a) The domain of definition $D(\mathfrak{A})$ *of* $\mathfrak{A}$ *is the set*

$$D(\mathfrak{A}) = \left\{u \in C(\overline{D}) : W_D u \in C(\overline{D}),\ \ Lu = 0\right\}. \tag{10.8}$$

(b) $\mathfrak{A}u = W_D u$ *for every* $u \in D(\mathfrak{A})$.

Here $W_D u$ *and* Lu *are taken in the sense of* distributions.

Assume that the boundary condition L *is transversal on the boundary* ∂D*. Then the operator* $\mathfrak{A}$ *generates a Feller semigroup* $\{T_t\}_{t\geq 0}$ *on* $\overline{D}$*.*

Remark 10.3. Bony, Courrège and Priouret [BCP] proved Theorem 10.2 in the case where the differential operator Q in formula (10.4) is *elliptic* (see [BCP, Théorème XIX]). Theorem 10.2 is proved by Cancelier [Cn, Théorème 3.2] and also by Taira [Ta6, Theorem 1]. It should be emphasized that Takanobu and Watanabe give a probabilistic version of Theorem 10.2 in the case where the domain D is the half-space $\mathbf{R}^N_+$ (see [TW, Corollary]).

The idea of our proof of Theorem 10.2 is as follows (see Bony–Courrège–Priouret [BCP], Sato–Ueno [SU], Taira [Ta5]).

First, we consider the Dirichlet problem for the Waldenfels integro-differential operator $W_D = P + S_D$: For given functions $f(x)$ and $\varphi(x')$ defined in D and on ∂D, respectively, find a function $u(x)$ in D such that

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ \gamma_0 u = \varphi & \text{on } \partial D, \end{cases} \tag{D}$$

where α is a positive parameter. We show that if $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$ has the transmission property with respect to ∂D, then Dirichlet problem (D) is uniquely solvable in the framework of Hölder spaces (Theorem 10.4). Secondly, we reduce the problem of construction of Feller semigroups to that of *unique solvability* of the following boundary value problem for the Ventcel' boundary condition L:

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ (\lambda - L)u = \varphi & \text{on } \partial D, \end{cases} \tag{*}$$

where λ is a non-negative constant (Theorem 10.19). In this way we prove existence theorems for Feller semigroups (Theorems 10.2 and 10.21).

10.3 The Dirichlet Problem for Waldenfels Operators

In this section we study Dirichlet problem (D) for the Waldenfels integro-differential operator $W_D = P + S_D$:

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ \gamma_0 u = \varphi & \text{on } \partial D. \end{cases} \tag{D}$$

We prove an existence and uniqueness theorem for Dirichlet problem (D) in the framework of *Hölder spaces* (Theorem 10.4) which plays an important role in the proof of existence theorems for Feller semigroups (Theorems 10.2 and 10.21).

10.3.1 Existence and Uniqueness Theorem for the Dirichlet Problem

We prove the following existence and uniqueness theorem for Dirichlet problem (D) (cf. [BCP, Théorème XV]):

Theorem 10.4. *Let W be a second-order elliptic Waldenfels operator and $\alpha > 0$. Assume that*

$$W_D 1(x) \leq 0 \quad \text{in } D.$$

Let k be an arbitrary non-negative integer and $0 < \theta < 1$. Then Dirichlet problem (D) has a unique solution u in the space $C^{k+2+\theta}(\overline{D})$ for any $f \in C^{k+\theta}(\overline{D})$ and any $\varphi \in C^{k+2+\theta}(\partial D)$.

The proof of Theorem 10.4 will be proved in the next subsection due to its length.

10.3.2 Proof of Theorem 10.4

The next theorem proves Theorem 10.4:

$$\boxed{\text{Theorem 10.5}} \Longrightarrow \boxed{\text{Theorem 10.4}}$$

Theorem 10.5. *Let W be a second-order elliptic Waldenfels operator. Assume that*

$$W_D 1(x) \leq 0 \quad \text{in } D.$$

Let k be an arbitrary non-negative integer and $0 < \theta < 1$. Then we have the following three assertions (i), (ii) and (iii):

(i) (Uniqueness) If $f \in C(\overline{D})$ and $\varphi \in C(\partial D)$, then a solution $u \in C(\overline{D}) \cap C^{2+\theta}(D)$ of problem (D) is unique.
(ii) (Existence) For any $f \in C^{k+\theta}(\overline{D})$ and any $\varphi \in C^{k+2+\theta}(\partial D)$, problem (D) has a unique solution u in the space $C^{k+2+\theta}(\overline{D})$.
(iii) (Global Regularity) If $f \in C^{k+\theta}(\overline{D})$ and $\varphi \in C^{k+2+\theta}(\partial D)$ for some non-negative integer k, then a solution $u \in C(\overline{D}) \cap C^2(D)$ of problem (D) belongs to the space $C^{k+2+\theta}(\overline{D})$.

Proof. The proof of Theorem 10.5 is divided into three steps.

Step 1: The uniqueness theorem for problem (D) (Assertion (i)) follows from an application of Theorem 8.12 in Chap. 8 with $W := W_D - \alpha$ for $\alpha > 0$, since we have the inequality

$$\max_{\overline{D}} |u| \leq \max \left\{ \frac{1}{\alpha} \sup_D |(W_D - \alpha)u|, \ \max_{\partial D} |u| \right\}. \tag{10.9}$$

Indeed, it suffices to note that

$$W_D 1(x) - \alpha = c(x) + a_\sigma(x) + \int_D s(x,y)\,[1 - \sigma(x,y)]\,dy - \alpha \leq -\alpha \quad \text{on } \overline{D}.$$

Step 2: The essential point in the proof of the existence theorem for problem (D) (Assertion (ii)) is how to show that the operator

$$(W_D - \alpha, \gamma_0) = (P - \alpha + S_D, \gamma_0) : C^{2+\theta}(\overline{D}) \longrightarrow C^\theta(\overline{D}) \oplus C^{2+\theta}(\partial D)$$

is an algebraic and topological *isomorphism* for $\alpha > 0$. To do this, we estimate the integral operator S in terms of Hölder norms, and show that the pseudo-differential operator case (W_D, γ_0) may be considered as a perturbation of a *compact operator* to the differential operator case (P, γ_0) in the framework of Hölder spaces. Namely, our proof can be reduced to the following differential operator case:

$$\begin{cases} (\alpha - P)v = f & \text{in } D, \\ \gamma_0 v = \varphi & \text{on } \partial D. \end{cases} \tag{D$'$}$$

The existence and uniqueness theorem in this case is well established in the framework of Hölder spaces. In fact, the next theorem summarizes the basic facts about the Dirichlet problem in the framework of *Hölder spaces* (see Gilbarg–Trudinger [GT]):

Theorem 10.6. *Let P be a second-order elliptic differential operator. Assume that*

$$P1(x) \leq 0 \quad \text{in } D.$$

Let k be an arbitrary non-negative integer and $0 < \theta < 1$. Then we have the following three assertions (i), (ii) and (iii):

(i) (Existence and Uniqueness) If $f \in C^{\theta}(D)$ with $0 < \theta < 1$ and $\varphi \in C(\partial D)$, then problem (D') has a unique solution u in $C(\overline{D}) \cap C^{2+\theta}(D)$.
(ii) (Interior Regularity) If $u \in C^2(D)$ and $f \in C^{k+\theta}(D)$ for some non-negative integer k, then it follows that $u \in C^{k+2+\theta}(D)$.
(iii) (Global Regularity) If $f \in C^{k+\theta}(\overline{D})$ and $\varphi \in C^{k+2+\theta}(\partial D)$ for some non-negative integer k, then a solution $u \in C(\overline{D}) \cap C^2(D)$ of problem (D') belongs to the space $C^{k+2+\theta}(\overline{D})$.

First, it follows from an application of Theorem 10.6 with $P := P - \alpha$ that the operator

$$(P - \alpha, \gamma_0) : C^{2+\theta}(\overline{D}) \longrightarrow C^{\theta}(\overline{D}) \oplus C^{2+\theta}(\partial D)$$

is an algebraic and topological *isomorphism* for $\alpha > 0$. In particular, we have

$$\text{ind}\,(P - \alpha, \gamma_0) = 0.$$

On the other hand, since $S \in L^{2-\kappa}_{1,0}(\mathbf{R}^N)$ for $\kappa > 0$ and has the *transmission property* with respect to ∂D, we obtain that the integro-differential operator S_D, defined by the formula

$$\begin{aligned} S_D u(x) = &\sum_{i=1}^{N} a_{\sigma}^{i}(x) \frac{\partial u}{\partial x_i}(x) + a_{\sigma}(x) u(x) \\ &+ \int_D s(x,y) \left[u(y) - \sigma(x,y) \left(u(x) + \sum_{j=1}^{N} (y_j - x_j) \frac{\partial u}{\partial x_j}(x) \right) \right] dy, \end{aligned}$$

maps $C^{2+\theta}(\overline{D})$ continuously into $C^{\kappa+\theta}(\overline{D})$. Hence it follows from an application of the Ascoli–Arzelà theorem that the operator

$$(S_D, 0) : C^{2+\theta}(\overline{D}) \longrightarrow C^{\theta}(\overline{D}) \oplus C^{2+\theta}(\partial D)$$

is *compact*. That is, we find that the operator

$$(W_D - \alpha, \gamma_0) = (P - \alpha, \gamma_0) + (S_D, 0)$$

is a perturbation of the compact operator $(S, 0)$ to the operator $(P - \alpha, \gamma_0)$. By applying the stability theorem for indices (Theorem 3.65) to our situation, we obtain that

$$\operatorname{ind}(W_D - \alpha, \gamma_0) = \operatorname{ind}(P - \alpha, \gamma_0) = 0.$$

Hence, in order to show the bijectivity of $(W_D - \alpha, \gamma_0)$ it suffices to prove its *injectivity*:

$$\begin{cases} u \in C^{2+\theta}(\overline{D}),\ (W_D - \alpha)u = 0 \text{ in } D,\ \gamma_0 u = 0 \text{ on } \partial D \\ \Longrightarrow\ u = 0 \text{ in } D. \end{cases}$$

However, this assertion is an immediate consequence of inequality (10.9).

Therefore, we have proved the existence and uniqueness theorem for problem (D) (Assertion (ii)).

Step 3: Finally, we prove the global regularity theorem for problem (D) (Assertion (iii)).

Assume that

$$\begin{cases} u \in C(\overline{D}) \cap C^2(D), \\ (W_D - \alpha)u = f \in C^{k+\theta}(\overline{D}), \\ u = \varphi \in C^{k+2+\theta}(\partial D). \end{cases}$$

By virtue of Assertion (ii), we can find a unique function $v \in C^{k+2+\theta}(\overline{D})$ such that

$$\begin{cases} (W_D - \alpha)v = f \in C^{k+\theta}(\overline{D}), \\ v = \varphi \in C^{k+2+\theta}(\partial D). \end{cases}$$

If we let

$$w = u - v,$$

then we have

$$\begin{cases} w \in C(\overline{D}) \cap C^2(D), \\ (W_D - \alpha)w = 0 \quad \text{in } D, \\ w = 0 \quad \text{on } \partial D. \end{cases}$$

Therefore, by applying Assertion (i) to the function w we obtain that

$$w = 0,$$

so that

$$u = v \in C^{k+2+\theta}(\overline{D}).$$

The proof of Theorem 10.5 is complete.

Now the proof of Theorem 10.4 is complete.

10.4 Construction of Feller Semigroups and Boundary Value Problems

In this section we reduce the problem of constructing Feller semigroups to that of the unique solvability of boundary value problems for the Waldenfels integro-differential operator $W_D = P + S_D$ with Ventcel' boundary conditions L (Theorem 10.19). Our approach here is divided into three Sects. 10.4.1–10.4.3.

10.4.1 Green Operators G_α^0 and Harmonic Operators H_α

By using Theorem 10.4 with $k := 0$, we obtain that Dirichlet problem (D) has a unique solution u in the space $C^{2+\theta}(\overline{D})$ for any $f \in C^\theta(\overline{D})$ and any $\varphi \in C^{2+\theta}(\partial D)$, $0 < \theta < 1$. Therefore, we can introduce two linear operators

$$G_\alpha^0 : C^\theta(\overline{D}) \longrightarrow C^{2+\theta}(\overline{D}),$$

and

$$H_\alpha : C^{2+\theta}(\partial D) \longrightarrow C^{2+\theta}(\overline{D})$$

as follows.

(a) For any $f \in C^\theta(\overline{D})$, the function $G_\alpha^0 f \in C^{2+\theta}(\overline{D})$ is the unique solution of the problem

$$\begin{cases} (\alpha - W_D)G_\alpha^0 f = f & \text{in } D, \\ G_\alpha^0 f = 0 & \text{on } \partial D. \end{cases} \tag{10.10}$$

(b) For any $\varphi \in C^{2+\theta}(\partial D)$, the function $H_\alpha\varphi \in C^{2+\theta}(\overline{D})$ is the unique solution of the problem

$$\begin{cases} (\alpha - W_D)H_\alpha\varphi = 0 & \text{in } D, \\ H_\alpha\varphi = \varphi & \text{on } \partial D. \end{cases} \tag{10.11}$$

The operator G_α^0 is called the *Green operator* for Dirichlet problem (D) and the operator H_α is called the *harmonic operator* for Dirichlet problem (D), respectively. We remark that the operators G_α^0 and H_α are generalizations of the classical Green representation formula (5.75) and Poisson integral formula (5.76), respectively.

Then we have the following results (cf. [BCP, Proposition III.1.6], [Ta5, Lemmas 9.6.2 and 9.6.3]):

Lemma 10.7. *The operator G_α^0, $\alpha > 0$, considered from $C(\overline{D})$ into itself, is non-negative and continuous with norm*

$$\left\|G_\alpha^0\right\| = \left\|G_\alpha^0 1\right\|_\infty = \max_{x \in \overline{D}} G_\alpha^0 1(x).$$

Proof. Let $f(x)$ be an arbitrary function in $C^\theta(\overline{D})$ such that $f(x) \geq 0$ on $\overline{D}$. Then, by applying the weak maximum principle (Theorem 8.11 in Chap. 8) with $W := W - \alpha$ to the function $-G_\alpha^0 f$ we obtain from (10.10) that

$$G_\alpha^0 f \geq 0 \quad \text{on } \overline{D}.$$

This proves the non-negativity of G_α^0.

Since G_α^0 is non-negative, we have, for all $f \in C^\theta(\overline{D})$,

$$-G_\alpha^0 \|f\|_\infty \leq G_\alpha^0 f \leq G_\alpha^0 \|f\|_\infty \quad \text{on } \overline{D}.$$

This implies the continuity of G_α^0 with norm

$$\left\|G_\alpha^0\right\| = \left\|G_\alpha^0 1\right\|_\infty.$$

The proof of Lemma 10.7 is complete.

Lemma 10.8. *The operator H_α, $\alpha > 0$, considered from $C(\partial D)$ into $C(\overline{D})$, is non-negative and continuous with norm*

$$\|H_\alpha\| = \|H_\alpha 1\|_\infty = \max_{x \in \overline{D}} H_\alpha 1(x).$$

Proof. Let $\varphi(x')$ be an arbitrary function in $C^{2+\theta}(\partial D)$ such that $\varphi(x') \geq 0$ on ∂D. Then, by applying the weak maximum principle (Theorem 8.11) with $W := W - \alpha$ to the function $-H_\alpha \varphi$ we obtain from formula (10.11) that

$$H_\alpha \varphi \geq 0 \quad \text{on } \overline{D}.$$

This proves the non-negativity of H_α.

Since H_α is non-negative, we have, for all $\varphi \in C^{2+\theta}(\partial D)$,

$$-H_\alpha \|\varphi\|_\infty \leq H_\alpha \varphi \leq H_\alpha \|\varphi\|_\infty \quad \text{on } \overline{D}.$$

This implies the continuity of H_α with norm

$$\|H_\alpha\| = \|H_\alpha 1\|_\infty .$$

The proof of Lemma 10.8 is complete.

More precisely, we have the following theorem (see [BCP, Proposition III.1.6]):

Theorem 10.9. *(i) (a) The operator G_α^0, $\alpha > 0$, can be uniquely extended to a non-negative, bounded linear operator on $C(\overline{D})$ into itself, denoted again by G_α^0, with norm*

$$\left\|G_\alpha^0\right\| = \left\|G_\alpha^0 1\right\|_\infty \leq \frac{1}{\alpha}. \tag{10.12}$$

(b) For any $f \in C(\overline{D})$, we have

$$G_\alpha^0 f\big|_{\partial D} = 0.$$

(c) For all α, $\beta > 0$, the resolvent equation holds true:

$$G_\alpha^0 f - G_\beta^0 f + (\alpha - \beta) G_\alpha^0 G_\beta^0 f = 0 \quad \text{for every } f \in C(\overline{D}). \tag{10.13}$$

(d) For any $f \in C(\overline{D})$, we have

$$\lim_{\alpha \to +\infty} \alpha G_\alpha^0 f(x) = f(x), \quad x \in D. \tag{10.14}$$

Furthermore, if $f = 0$ on ∂D, then this convergence is uniform in $x \in \overline{D}$, that is,

$$\lim_{\alpha \to +\infty} \alpha G_\alpha^0 f = f \quad \text{in } C(\overline{D}). \tag{10.14$'$}$$

(e) The operator G_α^0 maps $C^{k+\theta}(\overline{D})$ into $C^{k+2+\theta}(\overline{D})$ for any non-negative integer k.

(ii)

(a′) The operator H_α, $\alpha > 0$, can be uniquely extended to a non-negative, bounded linear operator from $C(\partial D)$ into $C(\overline{D})$, denoted again by H_α, with norm

$$\|H_\alpha\| \leq 1.$$

(b′) For any $\varphi \in C(\partial D)$, we have

$$H_\alpha \varphi|_{\partial D} = \varphi.$$

(c′) *For all* $\alpha, \beta > 0$, *we have*

$$H_\alpha\varphi - H_\beta\varphi + (\alpha - \beta)G_\alpha^0 H_\beta\varphi = 0 \quad \textit{for every } \varphi \in C(\partial D). \tag{10.15}$$

(d′) *For any* $\varphi \in C(\partial D)$, *we have*

$$\lim_{\alpha\to+\infty} H_\alpha\varphi(x) = 0, \quad x \in D.$$

(e′) *The operator* H_α *maps* $C^{k+2+\theta}(\partial D)$ *into* $C^{k+2+\theta}(\overline{D})$ *for any non-negative integer* k.

Proof. (i) Assertion (a): By making use of Friedrichs' mollifiers (see Theorem 6.5), we find that the space $C^\theta(\overline{D})$ is dense in $C(\overline{D})$ and further that non-negative functions can be approximated by non-negative smooth functions. Hence, by Lemma 10.7 it follows that the operator $G_\alpha^0 : C^\theta(\overline{D}) \to C^{2+\theta}(\overline{D})$ can be uniquely extended to a non-negative, bounded linear operator $G_\alpha^0 : C(\overline{D}) \to C(\overline{D})$ with norm

$$\left\| G_\alpha^0 \right\| = \left\| G_\alpha^0 1 \right\|_\infty .$$

Furthermore, since the function $G_\alpha^0 1$ satisfies the conditions

$$\begin{cases} (W_D - \alpha)G_\alpha^0 1 = -1 & \text{in } D, \\ G_\alpha^0 1 = 0 & \text{on } \partial D, \end{cases}$$

by applying inequality (10.9) we obtain that

$$\left\| G_\alpha^0 \right\| = \left\| G_\alpha^0 1 \right\|_\infty \leq \frac{1}{\alpha}.$$

Assertion (b): This follows from (10.10), since the space $C^\theta(\overline{D})$ is dense in $C(\overline{D})$ and since the operator $G_\alpha^0 : C(\overline{D}) \to C(\overline{D})$ is bounded.

Assertion (c): We find from the uniqueness theorem for problem (D) (Theorem 10.4) that Eq. (10.13) holds true for all $f \in C^\theta(\overline{D})$. Hence it holds true for all $f \in C(\overline{D})$, since the space $C^\theta(\overline{D})$ is dense in $C(\overline{D})$ and since the operators G_α^0 are bounded.

Assertion (d): First, let f be an arbitrary function in $C^\theta(\overline{D})$ satisfying $f|_{\partial D} = 0$. Then it follows from the uniqueness theorem for problem (D) (Theorem 10.4) that we have, for all α, β,

$$f - \alpha G_\alpha^0 f = G_\alpha^0 ((\beta - W_D) f) - \beta G_\alpha^0 f.$$

Thus we have, by estimate (10.12),

$$\|f - \alpha G_\alpha^0 f\|_\infty \leq \frac{1}{\alpha} \|(\beta - W_D) f\|_\infty + \frac{\beta}{\alpha} \|f\|_\infty,$$

so that

$$\lim_{\alpha \to +\infty} \|f - \alpha G_\alpha^0 f\|_\infty = 0.$$

Now let $f(x)$ be an arbitrary function in $C(\overline{D})$ satisfying $f|_{\partial D} = 0$. By means of Friedrichs' mollifiers (see Theorem 6.5), we can find a sequence $\{f_j\}$ in $C^\theta(\overline{D})$ such that

$$\begin{cases} f_j \longrightarrow f & \text{in } C(\overline{D}) \text{ as } j \to \infty, \\ f_j = 0 & \text{on } \partial D. \end{cases}$$

Then we have, by estimate (10.12),

$$\begin{aligned} \|f - \alpha G_\alpha^0 f\|_\infty &\leq \|f - f_j\|_\infty + \|f_j - \alpha G_\alpha^0 f_j\|_\infty + \|\alpha G_\alpha^0 f_j - \alpha G_\alpha^0 f\|_\infty \\ &\leq 2\|f - f_j\|_\infty + \|f_j - \alpha G_\alpha^0 f_j\|_\infty, \end{aligned}$$

and hence

$$\limsup_{\alpha \to +\infty} \|f - \alpha G_\alpha^0 f\|_\infty \leq 2\|f - f_j\|_\infty.$$

This proves assertion (10.14′), since $\|f - f_j\|_\infty \to 0$ as $j \to \infty$.

To prove assertion (10.14), let $f(x)$ be an arbitrary function in $C(\overline{D})$ and x an arbitrary point of D. Take a function $\psi(y) \in C(\overline{D})$ such that

$$\begin{cases} 0 \leq \psi(y) \leq 1 & \text{on } \overline{D}, \\ \psi(y) = 0 & \text{in a neighborhood of } x, \\ \psi(y) = 1 & \text{near the boundary } \partial D. \end{cases}$$

Then it follows from the non-negativity of G_α^0 and estimate (10.12) that

$$0 \leq \alpha G_\alpha^0 \psi(x) + \alpha G_\alpha^0 (1 - \psi)(x) = \alpha G_\alpha^0 1(x) \leq 1. \tag{10.16}$$

However, by applying assertion (10.13′) to the function $1 - \psi$ we have

$$\lim_{\alpha \to +\infty} \alpha G_\alpha^0 (1 - \psi)(x) = (1 - \psi)(x) = 1.$$

In view of inequalities (10.16), this implies that

$$\lim_{\alpha \to +\infty} \alpha G_\alpha^0 \psi(x) = 0.$$

Thus, since we have the inequalities

$$-\|f\|_\infty \psi \leq f\psi \leq \|f\|_\infty \psi \quad \text{on } \overline{D},$$

it follows that

$$|\alpha G_\alpha^0(f\psi)(x)| \leq \|f\|_\infty \cdot \alpha G_\alpha^0 \psi(x) \longrightarrow 0 \quad \text{as } \alpha \to +\infty.$$

Therefore, by applying assertion (10.14′) to the function $(1-\psi)f$ we obtain that

$$f(x) = ((1-\psi)f)(x) = \lim_{\alpha \to +\infty} \alpha G_\alpha^0 ((1-\psi)f)(x) = \lim_{\alpha \to +\infty} \alpha G_\alpha^0 f(x).$$

Assertion (e): This is an immediate consequence of part (iii) of Theorem 10.4.

(ii) Assertion (a′): Since the space $C^{2+\theta}(\partial D)$ is dense in $C(\partial D)$, by Lemma 10.8 it follows that the operator $H_\alpha : C^{2+\theta}(\partial D) \to C^{2+\theta}(\overline{D})$ can be uniquely extended to a non-negative, bounded linear operator $H_\alpha : C(\partial D) \to C(\overline{D})$. Moreover, by applying inequality (10.9) we have

$$\|H_\alpha\| = \|H_\alpha 1\|_\infty = 1.$$

Assertion (b′): This follows from (10.11), since the space $C^{2+\theta}(\partial D)$ is dense in $C(\partial D)$ and since the operator $H_\alpha : C(\partial D) \to C(\overline{D})$ is bounded.

Assertion (c′): We find from the uniqueness theorem for problem (D) that (10.15) holds true for all $\varphi \in C^{2+\theta}(\partial D)$. Hence it holds true for all $\varphi \in C(\partial D)$, since the space $C^{2+\theta}(\partial D)$ is dense in $C(\partial D)$ and since the operators G_α^0 and H_α are bounded.

Assertion (d′): This is an immediate consequence of part (iii) of Theorem 10.4.

The proof of Theorem 10.9 is now complete.

10.4.2 Boundary Value Problems and Reduction to the Boundary

Now we consider the following general boundary value problem in the framework of the spaces of *continuous functions*:

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ Lu = 0 & \text{on } \partial D, \end{cases} \tag{$**$}$$

where $\alpha > 0$ is a parameter. In this subsection, by using the Green operator G_α^0 and harmonic operator H_α for Dirichlet problem (D) we reduce the study of the boundary value problems to that of the classical pseudo-differential operator LH_α on the boundary (Theorem 10.19). This is a generalization of the classical Fredholm integral equation on the boundary.

To do this, we introduce three operators $\overline{W}$, $\overline{LG_\alpha^0}$ and $\overline{LH_\alpha}$ associated with problem $(**)$.

Step 1: First, we can introduce a linear operator

$$W : C(\overline{D}) \longrightarrow C(\overline{D})$$

as follows.

(a) The domain $D(W)$ of W is the space $C^{2+\theta}(\overline{D})$.
(b) $Wu = W_D u = Pu + S_D u$ for every $u \in D(W)$.

Then we have the following result (cf. [Ta5, Lemma 9.6.5]):

Lemma 10.10. *The operator W has its minimal closed extension $\overline{W}$ in the space $C(\overline{D})$.*

Proof. We apply part (i) of Theorem 9.50 to the operator W.

Assume that a function $u \in C^{2+\theta}(\overline{D})$ takes a positive maximum at a point x_0 of D:

$$u(x_0) = \max_{x \in \overline{D}} u(x) > 0.$$

Then it follows that

$$\frac{\partial u}{\partial x_i}(x_0) = 0, \quad 1 \leq i \leq N,$$
$$\sum_{i,j=1}^N a^{ij}(x_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) \leq 0,$$

since the matrix $(a^{ij}(x))$ is positive definite. Hence we have the inequality

$$\begin{aligned} & W_D u(x_0) \\ &= Pu(x_0) + (S_D u)(x_0) \\ &= \sum_{i,j=1}^N a^{ij}(x_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) + c(x_0) u(x_0) \end{aligned}$$

$$
\begin{aligned}
&+a_\sigma(x_0)u(x_0) + \int_D s(x_0, y)\,[(u(y) - u(x_0))]\,dy \\
&+ \int_D s(x_0, y)\,[1 - \sigma(x_0, y)]\,dy \cdot u(x_0) \\
\leq\ & c(x_0)u(x_0) + a_\sigma(x_0)u(x_0) + \int_D s(x_0, y)\,[1 - \sigma(x_0, y)]\,dy \cdot u(x_0) \\
=\ & W_D 1(x_0) \cdot u(x_0) \\
\leq\ & 0.
\end{aligned}
$$

This implies that the operator W satisfies condition (β) of Theorem 9.50 with $K_0 := D$ and $K := \overline{D}$, so that W has its minimal closed extension $\overline{W}$ in $C(\overline{D})$.

The proof of Lemma 10.10 is complete.

Remark 10.11. The situation may be visualized by the diagram below.

$$
\begin{array}{ccc}
C(\overline{D}) & \xrightarrow{\ \overline{W}\ } & C(\overline{D}) \\
\uparrow & & \uparrow \\
C^{2+\theta}(\overline{D}) & \xrightarrow[\ W\ ]{} & C^{\theta}(\overline{D})
\end{array}
$$

The extended operators $G_\alpha^0 : C(\overline{D}) \to C(\overline{D})$ and $H_\alpha : C(\partial D) \to C(\overline{D})$, $\alpha > 0$, still satisfy (10.10) and (10.11) respectively in the following sense (cf. [Ta5, Lemma 9.6.7 and Corollary 9.6.8]):

Lemma 10.12. *(i) For any $f \in C(\overline{D})$, we have*

$$
\begin{cases}
G_\alpha^0 f \in D(\overline{W}), \\
(\alpha I - \overline{W})G_\alpha^0 f = f \quad \text{in } D.
\end{cases}
$$

(ii) For any $\varphi \in C(\partial D)$, we have

$$
\begin{cases}
H_\alpha \varphi \in D(\overline{W}), \\
(\alpha I - \overline{W})H_\alpha \varphi = 0 \quad \text{in } D.
\end{cases}
$$

Here $D(\overline{W})$ is the domain of the closed extension $\overline{W}$.

Proof. Assertion (i): Choose a sequence $\{f_j\}$ in $C^\theta(\overline{D})$ such that $f_j \to f$ in $C(\overline{D})$ as $j \to \infty$. Then it follows from the boundedness of G_α^0 that

$$
G_\alpha^0 f_j \longrightarrow G_\alpha^0 f \quad \text{in } C(\overline{D}),
$$

and also

$$(\alpha - W_D) G_\alpha^0 f_j = f_j \longrightarrow f \quad \text{in } C(\overline{D}).$$

Hence we have

$$\begin{cases} G_\alpha^0 f \in D(\overline{W}), \\ (\alpha I - \overline{W}) G_\alpha^0 f = f \quad \text{in } D. \end{cases}$$

since the operator $\overline{W} : C(\overline{D}) \to C(\overline{D})$ is closed.

Assertion (ii): Similarly, part (ii) is proved, since the space $C^{2+\theta}(\partial D)$ is dense in $C(\partial D)$ and since the operator $H_\alpha : C(\partial D) \to C(\overline{D})$ is bounded.

The proof of Lemma 10.12 is complete.

Corollary 10.13. *Assume that*

$$W_D 1(x) < 0 \quad \text{on } \overline{D} = D \cup \partial D.$$

Then we have the following two assertions (i) and (ii):

(i) If a function $u \in D(\overline{W})$ satisfies the conditions

$$\begin{cases} \overline{W} u = 0 & \text{in } D, \\ u = 0 & \text{on } \partial D, \end{cases}$$

then it follows that $u = 0$ on $\overline{D}$.

(ii) Every function u in the domain $D(\overline{W})$ can be written in the following form:

$$u = G_\alpha^0 \big((\alpha I - \overline{W}) u \big) + H_\alpha (u|_{\partial D}), \quad \alpha > 0. \tag{10.17}$$

Proof. (i) Choose a sequence $\{u_j\}$ in $D(W) = C^{2+\theta}(\overline{D})$ such that

$$\begin{cases} W u_j \longrightarrow \overline{W} u & \text{in } C(\overline{D}), \\ u_j \longrightarrow u & \text{in } C(\overline{D}). \end{cases}$$

Then it follows from an application of inequality (9.44) with $u := u_j$ that

$$\max_{\overline{D}} |u_j| \leq \max \left\{ \left(\frac{1}{\min_{\overline{D}} (-W_D 1)} \right) \max_{\overline{D}} |W u_j|, \ \max_{\partial D} |u_j| \right\}. \tag{10.18}$$

Since we have

$$\overline{W} u = 0 \text{ in } D, \quad u = 0 \text{ on } \partial D,$$

by passing to the limit in inequality (10.18) we obtain that

$$\max_{\overline{D}} |u| = \lim_{j\to\infty} \max_{\overline{D}} |u_j| = 0,$$

so that

$$u = 0 \quad \text{on } \overline{D}.$$

(ii) We let

$$w = u - G_\alpha^0\left((\alpha I - \overline{W})u\right) - H_\alpha(u|_{\partial D}).$$

Then it follows from Lemma 10.12 that the function $w \in D(\overline{W})$ satisfies the conditions

$$\begin{cases} (\alpha I - \overline{W})w = 0 & \text{in } D, \\ w = 0 & \text{on } \partial D. \end{cases}$$

Thus we can apply part (i) with $W := W - \alpha$ to obtain that

$$w = 0.$$

This proves the desired formula (10.17).

The proof of Corollary 10.13 is complete.

Step 2: Secondly, we introduce a linear operator

$$LG_\alpha^0 : C(\overline{D}) \longrightarrow C(\partial D)$$

as follows.

(a) The domain $D\left(LG_\alpha^0\right)$ of LG_α^0 is the space $C^\theta(\overline{D})$.
(b) $LG_\alpha^0 f = L\left(G_\alpha^0 f\right)$ for every $f \in D\left(LG_\alpha^0\right)$.

Then we have the following result (cf. [BCP, Lemme III.2.4], [Ta5, Lemma 9.6.9]):

Lemma 10.14. *The operator LG_α^0, $\alpha > 0$, can be uniquely extended to a non-negative, bounded linear operator $\overline{LG_\alpha^0} : C(\overline{D}) \to C(\partial D)$.*

Proof. Let $f(x)$ be an arbitrary function in $D(LG_\alpha^0)$ such that $f(x) \geq 0$ on $\overline{D}$. Then we have

$$\begin{cases} G_\alpha^0 f \in C^2(\overline{D}), \\ G_\alpha^0 f \geq 0 & \text{on } \overline{D}, \\ G_\alpha^0 f = 0 & \text{on } \partial D, \end{cases}$$

and so

$$\begin{aligned} LG_\alpha^0 f(x') &= \mu(x') \frac{\partial}{\partial \mathbf{n}} (G_\alpha^0 f)(x') + \delta(x') f(x') + \int_D t(x', y)\, G_\alpha^0 f(y)\, dy \\ &\geq 0 \quad \text{on } \partial D. \end{aligned}$$

This proves that the operator LG_α^0 is non-negative.

By the non-negativity of LG_α^0, we have, for all $f \in D(LG_\alpha^0)$,

$$-LG_\alpha^0 \|f\|_\infty \leq LG_\alpha^0 f \leq LG_\alpha^0 \|f\|_\infty \quad \text{on } \partial D.$$

This implies the boundedness of LG_α^0 with norm

$$\left\| LG_\alpha^0 \right\| = \left\| LG_\alpha^0 1 \right\|_\infty .$$

Recall that the space $C^\theta(\overline{D})$ is dense in $C(\overline{D})$ and that non-negative functions can be approximated by non-negative smooth functions. Hence we find that the operator LG_α^0 can be uniquely extended to a non-negative, bounded linear operator $\overline{LG_\alpha^0} : C(\overline{D}) \to C(\partial D)$.

The proof of Lemma 10.14 is complete.

The situation may be visualized by the following diagram:

$$\begin{array}{ccc} C(\overline{D}) & \xrightarrow{\overline{LG_\alpha^0}} & C(\partial D) \\ \uparrow & & \uparrow \\ C^\theta(\overline{D}) & \xrightarrow[LG_\alpha^0]{} & C^\theta(\partial D) \end{array}$$

The next lemma states a fundamental relationship between the operators $\overline{LG_\alpha^0}$ and $\overline{LG_\beta^0}$ for $\alpha, \beta > 0$ (cf. [Ta5, Lemma 9.6.10]):

Lemma 10.15. *For any $f \in C(\overline{D})$, we have*

$$\overline{LG_\alpha^0} f - \overline{LG_\beta^0} f + (\alpha - \beta) \overline{LG_\alpha^0}\, G_\beta^0 f = 0, \quad \alpha, \beta > 0. \tag{10.19}$$

Proof. Choose a sequence $\{f_j\}$ in $C^\theta(\overline{D})$ such that $f_j \to f$ in $C(\overline{D})$ as $j \to \infty$. Then, by using the resolvent Eq. (10.13) with $f := f_j$ we have

$$LG_\alpha^0 f_j - LG_\beta^0 f_j + (\alpha - \beta) LG_\alpha^0 G_\beta^0 f_j = 0.$$

Hence (10.19) follows by letting $j \to \infty$, since the operators $\overline{LG_\alpha^0}$, $\overline{LG_\beta^0}$ and G_β^0 are all bounded.

The proof of Lemma 10.15 is complete.

Step 3: Finally, we introduce a linear operator

$$LH_\alpha : C(\partial D) \longrightarrow C(\partial D)$$

as follows.

(a) The domain $D(LH_\alpha)$ of LH_α is the space $C^{2+\theta}(\partial D)$.
(b) $LH_\alpha \psi = L(H_\alpha \psi)$ for every $\psi \in D(LH_\alpha)$.

The operators LH_α are a pseudo-differential operator version of the classical Fredholm integral equation on the boundary ∂D.

Then we have the following result (cf. [Ta5, Lemma 9.6.11]):

Lemma 10.16. *The operator LH_α, $\alpha > 0$, has its minimal closed extension $\overline{LH_\alpha}$ in the space $C(\partial D)$.*

Proof. We apply part (i) of Theorem 9.50 to the operator LH_α. To do this, it suffices to show that the operator LH_α satisfies condition (β') with $K := \partial D$ (or condition (β) with $K := K_0 = \partial D$) of the same theorem.

Assume that a function ψ in $D(LH_\alpha) = C^{2+\theta}(\partial D)$ takes its positive maximum at some point x' of ∂D. Since the function $H_\alpha\psi$ is in $C^{2+\theta}(\overline{D})$ and satisfies the conditions

$$\begin{cases} (W_D - \alpha)H_\alpha\psi = 0 & \text{in } D, \\ H_\alpha\psi = \psi & \text{on } \partial D, \end{cases}$$

by applying the weak maximum principle (Theorem 8.11) with $W := W - \alpha$ to the function $H_\alpha\psi$, we find that the function $H_\alpha\psi$ takes its positive maximum at a boundary point $x_0' \in \partial D$:

$$\psi(x_0') = \max_{y \in \overline{D}} H_\alpha\psi(y) > 0,$$

$$\frac{\partial\psi}{\partial x_i}(x_0') = 0, \quad 1 \le i \le N-1.$$

Thus we can apply the Hopf boundary point lemma (Theorem 8.15 in Chap. 8) to obtain that

$$\frac{\partial}{\partial \mathbf{n}}(H_\alpha\psi)(x_0') < 0.$$

Hence we have, by condition (1.6),

$$
\begin{aligned}
LH_\alpha \psi(x_0') &= \sum_{i,j=1}^{N-1} \alpha^{ij}(x_0') \frac{\partial^2 \psi}{\partial x_i \partial x_j}(x_0') + \gamma(x_0')\psi(x_0') + \mu(x_0') \frac{\partial}{\partial \mathbf{n}} (H_\alpha \psi)(x_0') \\
&\quad -\alpha\, \delta(x_0')\psi(x_0') + \int_D t(x_0', y) \left[H_\alpha \psi(y) - \psi(x_0') \right] dy \\
&\quad + \eta_\tau(x_0')\psi(x_0') + \int_{\partial D} r(x_0', y') \left[\psi(y') - \psi(x_0')\tau(x_0', y') \right] dy' \\
&= \sum_{i,j=1}^{N-1} \alpha^{ij}(x_0') \frac{\partial^2 \psi}{\partial x_i \partial x_j}(x_0') + \mu(x_0') \frac{\partial}{\partial \mathbf{n}} (H_\alpha \psi)(x_0') \\
&\quad -\alpha \delta(x_0')\psi(x_0') + \int_D t(x_0', y) \left[H_\alpha \psi(y) - \psi(x_0') \right] dy \\
&\quad + \int_{\partial D} r(x_0', y') \left[\psi(y') - \psi(x_0') \right] dy' \\
&\quad + \left(\gamma(x_0') + \eta_\tau(x_0') + \int_{\partial D} r(x_0', y') \left[1 - \tau(x_0', y') \right] dy' \right) \psi(x_0') \\
&\leq 0.
\end{aligned}
$$

This verifies condition (β') of Theorem 9.50, so that LH_α has its minimal closed extension $\overline{LH_\alpha}$ in $C(\partial D)$.

The proof of Lemma 10.16 is complete.

Remark 10.17. The situation may be visualized by the following diagram:

$$
\begin{array}{ccc}
C(\partial D) & \xrightarrow{\overline{LH_\alpha}} & C(\partial D) \\
\uparrow & & \uparrow \\
C^{2+\theta}(\partial D) & \xrightarrow[LH_\alpha]{} & C^{\theta}(\partial D)
\end{array}
$$

Moreover, the operator $\overline{LH_\alpha}$ enjoys the following property (10.20):

$$
\text{If a function } \psi \text{ in the domain } D\left(\overline{LH_\alpha}\right) \text{ takes its } \textit{positive} \text{ maximum at some point } x' \text{ of } \partial D, \text{ then we have the inequality } \overline{LH_\alpha}\psi(x') \leq 0. \tag{10.20}
$$

The next lemma states a fundamental relationship between the operators $\overline{LH_\alpha}$ and $\overline{LH_\beta}$ for $\alpha, \beta > 0$ (cf. [Ta5, Lemma 9.6.13]):

Lemma 10.18. *The domain $D\left(\overline{LH_\alpha}\right)$ of $\overline{LH_\alpha}$ does not depend on $\alpha > 0$; so we denote by $\mathcal{D}$ the common domain. Then we have*

$$
\overline{LH_\alpha}\psi - \overline{LH_\beta}\psi + (\alpha - \beta)\overline{LG_\alpha^0}\, H_\beta \psi = 0, \quad \alpha, \beta > 0, \quad \psi \in \mathcal{D}. \tag{10.21}
$$

Proof. Let $\psi(x')$ be an arbitrary function in $D(\overline{LH_\beta})$, and choose a sequence $\{\psi_j\}$ in $D(LH_\beta) = C^{2+\theta}(\partial D)$ such that

$$\begin{cases} \psi_j \longrightarrow \psi & \text{in } C(\partial D), \\ LH_\beta \psi_j \longrightarrow \overline{LH_\beta}\psi & \text{in } C(\partial D). \end{cases}$$

Then it follows from the boundedness of H_β and $\overline{LG_\alpha^0}$ that

$$LG_\alpha^0(H_\beta\psi_j) = \overline{LG_\alpha^0}(H_\beta\psi_j) \longrightarrow \overline{LG_\alpha^0}(H_\beta\psi) \quad \text{in } C(\partial D).$$

Therefore, by using (10.15) with $\varphi := \psi_j$ we obtain that

$$\begin{aligned} LH_\alpha\psi_j &= LH_\beta\psi_j - (\alpha-\beta)LG_\alpha^0(H_\beta\psi_j) \\ &\longrightarrow \overline{LH_\beta}\psi - (\alpha-\beta)\overline{LG_\alpha^0}(H_\beta\psi) \quad \text{in } C(\partial D). \end{aligned}$$

This implies that

$$\begin{cases} \psi \in D(\overline{LH_\alpha}), \\ \overline{LH_\alpha}\psi = \overline{LH_\beta}\psi - (\alpha-\beta)\overline{LG_\alpha^0}(H_\beta\psi). \end{cases}$$

Conversely, by interchanging α and β we have the inclusion

$$D(\overline{LH_\alpha}) \subset D(\overline{LH_\beta}),$$

and so

$$D(\overline{LH_\alpha}) = D(\overline{LH_\beta}).$$

The proof of Lemma 10.18 is complete.

Step 4: Now we can state a general existence theorem for Feller semigroups on ∂D in terms of the boundary value problem

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ (\lambda - L)u = \varphi & \text{on } \partial D. \end{cases} \tag{$*$}$$

The next theorem asserts that the operator $\overline{LH_\alpha}$ is the infinitesimal generator of some Feller semigroup on ∂D if and only if problem ($*$) is solvable for *sufficiently many* functions φ in the space $C(\partial D)$ (cf. [BCP, Théorème XX], [Ta5, Theorem 9.6.15]):

Theorem 10.19. *(i) If the operator $\overline{LH_\alpha}$, $\alpha > 0$, is the infinitesimal generator of a Feller semigroup on ∂D, then, for each constant $\lambda > 0$, the boundary value problem*

$$\begin{cases} (\alpha - W_D)u = 0 & \text{in } D, \\ (\lambda - L)u = \varphi & \text{on } \partial D \end{cases} \tag{$*'$}$$

has a solution $u \in C^{2+\theta}(\overline{D})$ for any φ in some dense subset of $C(\partial D)$.

(ii) Conversely, if, for some constant $\lambda \geq 0$, problem ($*'$) has a solution $u \in C^{2+\theta}(\overline{D})$ for any φ in some dense subset of $C(\partial D)$, then the operator $\overline{LH_\alpha}$ is the infinitesimal generator of some Feller semigroup on ∂D.

Proof. Assertion (i): If the operator $\overline{LH_\alpha}$ generates a Feller semigroup on ∂D, by applying part (i) of Theorem 9.35 with $K := \partial D$ to the operator $\overline{LH_\alpha}$ we obtain that

$$R\left(\lambda I - \overline{LH_\alpha}\right) = C(\partial D) \quad \text{for each } \lambda > 0.$$

This implies that the range $R(\lambda I - LH_\alpha)$ is a dense subset of $C(\partial D)$ for each $\lambda > 0$. However, if $\varphi \in C(\partial D)$ is in the range $R(\lambda I - LH_\alpha)$, and if $\varphi = (\lambda I - LH_\alpha)\psi$ with $\psi \in C^{2+\theta}(\partial D)$, then the function $u = H_\alpha\psi \in C^{2+\theta}(\overline{D})$ is a solution of problem ($*'$). This proves assertion (i).

Assertion (ii): We apply part (ii) of Theorem 9.35 with $K := \partial D$ to the operator LH_α. To do this, it suffices to show that the operator LH_α satisfies condition (γ) of the same theorem, since it satisfies condition (β'), as is shown in the proof of Lemma 10.16.

By the uniqueness theorem for problem (D), it follows that any function $u \in C^{2+\theta}(\overline{D})$ which satisfies the equation

$$(\alpha - W_D)u = 0 \quad \text{in } D$$

can be written in the form

$$u = H_\alpha\left(u|_{\partial D}\right), \quad u|_{\partial D} \in C^{2+\theta}(\partial D) = D\left(LH_\alpha\right).$$

Thus we find that if there exists a solution $u \in C^{2+\theta}(\overline{D})$ of problem ($*'$) for a function $\varphi \in C(\partial D)$, then we have

$$(\lambda I - LH_\alpha)\left(u|_{\partial D}\right) = \varphi,$$

and so

$$\varphi \in R\left(\lambda I - LH_\alpha\right).$$

Therefore, if there exists a constant $\lambda \geq 0$ such that problem $(*')$ has a solution u in $C^{2+\theta}(\overline{D})$ for any φ in some dense subset of $C(\partial D)$, then the range $R(\lambda I - LH_\alpha)$ is dense in $C(\partial D)$. This verifies condition (γ) (with $\alpha_0 := \lambda$) of Theorem 9.50. Hence assertion (ii) follows from an application of the same theorem.

The proof of Theorem 10.19 is complete.

We can give a precise meaning to the boundary conditions Lu for functions u in the domain $D(\overline{W})$.

We let

$$D(L) = \left\{u \in D(\overline{W}) : u|_{\partial D} \in \mathcal{D}\right\},$$

where $\mathcal{D}$ is the common domain of the operators $\overline{LH_\alpha}$, $\alpha > 0$. It should be noted that the domain $D(L)$ contains the space $C^{2+\theta}(\overline{D})$, since we have

$$C^{2+\theta}(\partial D) = D\left(LH_\alpha\right) \subset \mathcal{D}.$$

Corollary 10.13 tells us that every function u in $D(L) \subset D(\overline{W})$ can be written in the form

$$u = G_\alpha^0\left((\alpha I - \overline{W})u\right) + H_\alpha\left(u|_{\partial D}\right), \quad \alpha > 0. \tag{10.13}$$

Then we define Lu by the formula

$$Lu = \overline{LG_\alpha^0}\left((\alpha I - \overline{W})u\right) + \overline{LH_\alpha}\left(u|_{\partial D}\right). \tag{10.22}$$

The next lemma justifies definition (10.22) of Lu for all $u \in D(L)$ (cf. [Ta5, Lemma 9.6.16]):

Lemma 10.20. *The right-hand side of* (10.22) *depends only on u, not on the choice of expression* (10.17)*.*

Proof. Assume that

$$\begin{aligned} u &= G_\alpha^0\left(\left(\alpha I - \overline{W}\right)u\right) + H_\alpha\left(u|_{\partial D}\right) \\ &= G_\beta^0\left(\left(\beta I - \overline{W}\right)u\right) + H_\beta\left(u|_{\partial D}\right), \end{aligned}$$

where $\alpha > 0$, $\beta > 0$. Then it follows from (10.19) with $f := \left(\alpha I - \overline{W}\right)u$ and (10.22) with $\psi := u|_{\partial D}$ that

$$\begin{aligned} &\overline{LG_\alpha^0}\left(\left(\alpha I - \overline{W}\right)u\right) + \overline{LH_\alpha}\left(u|_{\partial D}\right) \\ &= \overline{LG_\beta^0}\left(\left(\alpha I - \overline{W}\right)u\right) - (\alpha - \beta)\overline{LG_\alpha^0}G_\beta^0\left(\left(\alpha I - \overline{W}\right)u\right) \\ &\quad + \overline{LH_\beta}\left(u|_{\partial D}\right) - (\alpha - \beta)\overline{LG_\alpha^0}H_\beta\left(u|_{\partial D}\right) \end{aligned} \tag{10.23}$$

$$= \overline{LG^0_\beta}\left((\beta I - \overline{W})u\right) + \overline{LH_\beta}\left(u|_{\partial D}\right)$$
$$+(\alpha-\beta)\left\{\overline{LG^0_\beta}u - \overline{LG^0_\alpha}G^0_\beta\left(\alpha I - \overline{W}\right)u - \overline{LG^0_\alpha}H_\beta\left(u|_{\partial D}\right)\right\}.$$

However, the last term of (10.23) vanishes. Indeed, it follows from (10.19) with $f := u$ that

$$\begin{aligned}
&\overline{LG^0_\beta}u - \overline{LG^0_\alpha}\left(G^0_\beta\left(\alpha I - \overline{W}\right)u\right) - \overline{LG^0_\alpha}H_\beta\left(u|_{\partial D}\right)\\
&= \overline{LG^0_\beta}u - \overline{LG^0_\alpha}\left(G^0_\beta\left(\beta I - \overline{W}\right)u + H_\beta\left(u|_{\partial D}\right) + (\alpha-\beta)G^0_\beta u\right)\\
&= \overline{LG^0_\beta}u - \overline{LG^0_\alpha}u - (\alpha-\beta)\overline{LG^0_\alpha}G^0_\beta u\\
&= 0.
\end{aligned}$$

Therefore, we obtain from (10.23) that

$$\overline{LG^0_\alpha}\left(\left(\alpha I - \overline{W}\right)u\right) + \overline{LH_\alpha}\left(u|_{\partial D}\right) = \overline{LG^0_\beta}\left(\left(\beta I - \overline{W}\right)u\right) + \overline{LH_\beta}\left(u|_{\partial D}\right).$$

The proof of Lemma 10.20 is complete.

10.4.3 A Generation Theorem for Feller Semigroups in Terms of Green Operators

Now we formulate a generation theorem for Feller semigroups in terms of Green operators G_α which implies Theorem 10.2:

$$\boxed{\text{Theorem } 10.21} \Longrightarrow \boxed{\text{Theorem } 10.2}$$

Theorem 10.21. *We define a linear operator*

$$\mathfrak{A} : C(\overline{D}) \longrightarrow C(\overline{D})$$

as follows (see formula (10.8)).

(a) The domain $D(\mathfrak{A})$ of $\mathfrak{A}$ is the set

$$D(\mathfrak{A}) = \left\{u \in D(\overline{W}) : u|_{\partial D} \in \mathcal{D},\ \ Lu = 0\right\}, \tag{10.24}$$

where $\mathcal{D}$ is the common domain of the operators $\overline{LH_\alpha}$, $\alpha > 0$.
(b) $\mathfrak{A}u = \overline{W}u$ for every $u \in D(\mathfrak{A})$.

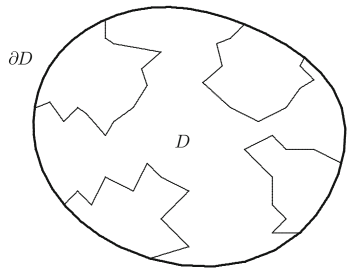


Fig. 10.6 The intuitive meaning of formula (10.25)

If the boundary condition L is transversal on the boundary ∂D, then the operator $\mathfrak{A}$ is the infinitesimal generator of some Feller semigroup on $\overline{D}$, and the Green operator $G_\alpha = (\alpha I - \mathfrak{A})^{-1}$, $\alpha > 0$, is given by the formula

$$G_\alpha f = G_\alpha^0 f - H_\alpha \left(\overline{LH_\alpha}^{-1} \left(\overline{LG_\alpha^0 f} \right) \right) \quad \textit{for every } f \in C(\overline{D}). \tag{10.25}$$

Remark 10.22. Intuitively, formula (10.25) asserts that if the boundary condition L is transversal on the boundary ∂D, then we can "piece together" a Markov process (Feller semigroup) on the boundary ∂D with W-process in the interior D to construct a Markov process (Feller semigroup) on the whole $\overline{D} = D \cup \partial D$. The situation may be represented schematically by Fig. 10.6.

10.5 Proof of Theorem 1.2

This section is devoted to the proof of Theorem 10.21 and hence that of Theorem 1.2. We shall apply a version of the Hille–Yosida theorem (Theorem 9.35) to the operator $\mathfrak{A}$ defined by formula (10.17). To do this, we verify that $\mathfrak{A}$ satisfies conditions (a)–(d) in the same theorem.

10.5.1 Proof of Theorem 10.21

The proof of Theorem 10.21 is divided into several steps.

Theorem 10.23	$\Longrightarrow$	Theorem 10.19	$\Longrightarrow$	Theorem 10.21

Step 1: First, by the Boutet de Monvel calculus [Bo] (see Sect. 7.7) we know that if $T \in L_{1,0}^{1-\kappa_2}(\mathbf{R}^N)$ has the *transmission property* with respect to the boundary ∂D, then the operator

$$\int_D t(x', y) H_\alpha \varphi(y)\, dy = T_D \left(H_\alpha \varphi \right)|_{\partial D}$$

is a classical pseudo-differential operator of order $1 - \kappa_2$ on the boundary ∂D. Therefore, we find that the operator LH_α is the sum of a second-order degenerate elliptic differential operator and a classical pseudo-differential operator of order $\max(2 - \kappa_1, 1 - \kappa_2)$:

$$\begin{aligned}
&LH_\alpha\varphi(x') \\
&= \sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2\varphi}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial\varphi}{\partial x_i}(x') + \gamma(x')\varphi(x') \\
&\quad + \mu(x') \frac{\partial}{\partial \mathbf{n}} (H_\alpha\varphi)(x') - \alpha\, \delta(x')\varphi(x') \\
&\quad + \int_D t(x', y) \left[H_\alpha\varphi(y) - \varphi(x') \right] dy \\
&\quad + \sum_{j=1}^{N-1} \eta_\tau^j(x') \frac{\partial\varphi}{\partial x_j}(x') + \eta_\tau(x')\varphi(x') \\
&\quad + \int_{\partial D} r(x', y') \left[\varphi(y') - \tau(x', y') \left(\varphi(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial\varphi}{\partial x_j}(x') \right) \right] dy'.
\end{aligned}$$

Now we prove that

For *all* $\alpha > 0$, the operator $\overline{LH_\alpha}$ generates a Feller semigroup on the boundary ∂D.

First, we have the following six assertions (i)–(vi):

(i) The operator

$$Q\varphi = \sum_{i,j=1}^{N-1} \alpha^{ij}(x') \frac{\partial^2\varphi}{\partial x_i \partial x_j} + \sum_{i=1}^{N-1} \beta^i(x') \frac{\partial\varphi}{\partial x_i} + \gamma(x')\varphi$$

is a second-order degenerate elliptic differential operator on ∂D with non-positive principal symbol, and $Q1(x') = \gamma(x') \leq 0$ on ∂D.

(ii) The operator

$$\Pi_\alpha\varphi = \left. \frac{\partial}{\partial \mathbf{n}} (H_\alpha\varphi) \right|_{\partial D}$$

is a first-order classical pseudo-differential operator on ∂D (see [Ho1, RS, Se2]). For example, if A is the usual Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_N^2},$$

then we can write down the symbol of Π_α as follows (see formula (10.60)):

$$-|\xi'| - \frac{1}{2}\left(\frac{\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1)M(x')\right) + \sqrt{-1}\,\frac{1}{2}\operatorname{div}\delta_{(\xi')}(x')$$
$$+ \text{ terms of order} \leq -1/2 \text{ depending on } \alpha.$$

Here $\operatorname{div}\delta_{(\xi')}$ is the *divergence* of a real smooth vector field $\delta_{(\xi')}$ on ∂D defined (in local coordinates) by the formula

$$\delta_{(\xi')} = \sum_{j=1}^{N-1} \frac{\partial|\xi'|}{\partial\xi_j}\frac{\partial}{\partial x_j} \quad \text{for } \xi' \neq 0.$$

Moreover, it should be noted (cf. condition (8.3)) that

$$\varphi \in C^{2+\theta}(\partial D),\ \varphi \geq 0 \text{ on } \partial D \text{ and } x_0'' \in \partial D,\ x_0' \notin \operatorname{supp}\varphi \tag{10.26}$$
$$\Longrightarrow \quad \Pi_\alpha\varphi(x_0') = \frac{\partial}{\partial\mathbf{n}}(H_\alpha\varphi)(x_0') \geq 0.$$

Indeed, if we let

$$u = H_\alpha\varphi \in C^{2+\theta}(\overline{D}),$$

then we have

$$\begin{cases} (\alpha - W_D)u = 0 & \text{in } D, \\ u = \varphi & \text{on } \partial D. \end{cases}$$

Since $\varphi \geq 0$ on ∂D, it follows from an application of the weak maximum principle (see Theorem 8.11) that

$$H_\alpha\varphi = u \geq 0 \quad \text{in } D.$$

This implies that

$$\Pi_\alpha\varphi(x_0') = \frac{\partial}{\partial\mathbf{n}}(H_\alpha\varphi)(x_0') \geq 0,$$

since $H_\alpha\varphi(x_0') = \varphi(x_0') = 0$.

(iii) The integro-differential operator

$$\int_{\partial D} r(x',y')\left[\varphi(y') - \tau(x',y')\left(\varphi(x') + \sum_{j=1}^{N-1}(y_j - x_j)\frac{\partial\varphi}{\partial x_j}(x')\right)\right]dy'$$

is a classical pseudo-differential operator of order $2 - \kappa_1$ on the boundary ∂D. Indeed, it can be *formally* written in the form

$$\int_{\partial D} r(x', y') \left[\varphi(y') - \tau(x', y') \left(\varphi(x') + \sum_{j=1}^{N-1} (y_j - x_j) \frac{\partial \varphi}{\partial x_j}(x') \right) \right] dy'$$

$$= R\varphi(x') - R(\tau_{x'}) \cdot \varphi(x') - \sum_{j=1}^{N-1} R(\tau_{x'}^j) \cdot \frac{\partial \varphi}{\partial x_j}(x'),$$

where

$$\tau_{x'}(y') = \tau(x', y'), \quad y' \in \partial D,$$
$$\tau_{x'}^j(y') = \tau(x', y') \left(y_j - x_j \right), \quad y' \in \partial D.$$

(iv) The integral operator

$$\int_D t(x', y) \left[H_\alpha \varphi(y) - \varphi(x') \right] dy = T_D \left(H_\alpha \varphi \right)|_{\partial D} - (T_D 1)|_{\partial D} \cdot \varphi(x')$$

is a classical pseudo-differential operator of order $1 - \kappa_2$ on the boundary ∂D, since $T \in L_{1,0}^{1-\kappa_2}(\mathbf{R}^N)$ has the *transmission property* with respect to the boundary ∂D (see [Bo, RS]). For example, if P is the usual Laplacian Δ and if the symbol of T is given by the formula (see formula (7.11))

$$\frac{a_0(x', \xi')}{|\xi'| - \sqrt{-1}\,\xi_n} + a_1(x', \xi') \frac{|\xi'| + \sqrt{-1}\,\xi_n}{(|\xi'| - \sqrt{-1}\,\xi_n)^2} + \ldots,$$
$$a_0(x', \xi'),\ a_1(x', \xi') \in S_{1,0}^{2-\kappa_2}(\mathbf{R}^{N-1} \times \mathbf{R}^{N-1}),$$

then we can write down the symbol of the operator

$$T_D \left(H_\alpha \varphi \right)|_{\partial D}$$

concretely as follows:

$$\frac{a_0(x', \xi')}{2|\xi'|} + \frac{1}{4} \left(\frac{\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1) M(x') \right) \frac{a_0(x', \xi')}{2|\xi'|^2}$$
$$+ \sqrt{-1}\, \frac{a_0(x', \xi')}{4|\xi'|} \operatorname{div} \delta_{(\xi')}(x')$$
$$+ \text{terms of order} \le -\kappa_2 - 1 \text{ depending on } \alpha.$$

(v) We remark (cf. condition (8.3)) that

$$\varphi \in C^{2+\theta}(\partial D), \ \ \varphi \geq 0 \ \text{ on } \partial D \text{ and } \ x_0'' \in \partial D, \ x_0' \notin \operatorname{supp}\varphi$$
$$\Longrightarrow \ \ LH_\alpha\varphi(x_0') \geq 0.$$

Indeed, we have, by assertion (10.26),

$$LH_\alpha\varphi(x_0') = \mu(x_0')\frac{\partial}{\partial \mathbf{n}}(H_\alpha\varphi)(x_0') + \int_{\partial D} r(x_0', y')\varphi(y')\,dy'$$
$$+ \int_D t(x_0', y)H_\alpha\varphi(y)\,dy \geq 0.$$

Therefore, by applying Theorem 9.5 to our situation we obtain that the pseudo-differential operator LH_α may be written in the form

$$LH_\alpha\varphi = P_\alpha\varphi + S_\alpha\varphi, \quad \varphi \in C^2(\partial D).$$

Here:

(a) The operator P_α is a second-order *degenerate* elliptic differential operator on ∂D with non-positive principal symbol and $P_\alpha 1(x') = \gamma_\alpha(x') \leq 0$ on ∂D.
(b) The operator S_α is an integro-differential operator given by the formula

$$S_\alpha\varphi(x')$$
$$= \int_{\partial D} s(x', y')\Big[\varphi(y') - \sigma(x', y')\Big(\varphi(x') + \sum_{i=1}^{N-1}\frac{\partial\varphi}{\partial x_i}(x')(y_i - x_i)\Big)\Big]dy'.$$

It should be emphasized that $s(x', y') \geq 0$ off the diagonal $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$, and that

$$LH_\alpha 1(x') = \gamma_\alpha(x') + \int_{\partial D} s(x', y')\left[1 - \sigma(x', y')\right]dy' \leq 0 \quad \text{on } \partial D.$$

(vi) Finally, since the function $H_\alpha 1$ takes its positive maximum 1 only on the boundary ∂D, it follows from an application of the Hopf boundary point lemma (see Theorem 8.15) that

$$H_\alpha 1(y) - 1 < 0 \quad \text{in } D,$$
$$\Pi_\alpha 1(x') = \frac{\partial}{\partial \mathbf{n}}(H_\alpha 1)(x') < 0 \quad \text{on } \partial D.$$

Hence, it follows from transversality condition (10.7) and condition (1.6) that

$$\begin{aligned}
LH_\alpha 1(x') &= \gamma(x') + \mu(x')\,\Pi_\alpha 1(x') - \alpha\delta(x') \\
&\quad + \eta_\tau(x') + \int_{\partial D} r(x', y')\left[1 - \tau(x', y')\right] dy' \\
&\quad + \int_D t(x', y)\,[H_\alpha 1(y) - 1]\,dy \\
&= \left(\gamma(x') + \eta_\tau(x') + \int_{\partial D} r(x', y')\left[1 - \tau(x', y')\right] dy'\right) - \alpha\delta(x') \\
&\quad + \mu(x')\Pi_\alpha 1(x') + \int_D t(x', y)\,[H_\alpha 1(y) - 1]\,dy \\
&\leq -\alpha\delta(x') + \mu(x')\Pi_\alpha 1(x') + \int_D t(x', y)\,[H_\alpha 1(y) - 1]\,dy \\
&< 0 \quad \text{on } \partial D.
\end{aligned} \tag{10.27}$$

Step 2: The next unique solvability theorem for pseudo-differential operators in the framework of Hölder spaces will play an essential role in the construction of Feller semigroups (see [Cn, Théorème 4.5], [Ta6, Theorem 2.1]):

Theorem 10.23. *Let T be a second-order classical pseudo-differential operator on an n-dimensional, compact smooth manifold M without boundary such that*

$$T = P + S.$$

Here:

(a) The operator P is a second-order degenerate elliptic differential operator on M with non-positive principal symbol $p(x, \xi)$, and $P1(x) \leq 0$ on M.
(b) The operator S is a classical pseudo-differential operator of order $2-\kappa$, $\kappa > 0$, on M and its distribution kernel $s(x, y)$ is non-negative off the diagonal $\Delta_M = \{(x, x) : x \in M\}$ in $M \times M$.
(c) $T1(x) = P1(x) + S1(x) \leq 0$ on M.

Then, for each integer $k \geq 1$ there exists a constant $\lambda = \lambda(k) > 0$ such that, for any $f \in C^{k+\theta}(M)$, we can find a function $\varphi \in C^{k+\theta}(M)$ satisfying the equation

$$(T - \lambda I)\varphi = f \quad \text{on } M,$$

and the estimate

$$\|\varphi\|_{C^{k+\theta}(M)} \leq C(\lambda)\|f\|_{C^{k+\theta}(M)}. \tag{10.28}$$

Here $C(\lambda) > 0$ is a constant independent of f.

Theorem 10.23 will be proved in the next Sect. 10.6 due to its length.

Step 3: By applying Theorem 10.23 to the operator LH_α, we find that

If $\lambda > 0$ is sufficiently large, then the range $R(LH_\alpha - \lambda I)$ contains the space $C^{2+\theta}(\partial D)$. (10.29)

This implies that the range $R(LH_\alpha - \lambda I)$ is a *dense* subset of $C(\partial D)$. Therefore, by applying part (ii) of Theorem 10.19 to the operator L we obtain that the operator $\overline{LH_\alpha}$ is the infinitesimal generator of some Feller semigroup on ∂D, for any $\alpha > 0$.

Step 4: Now we prove that

The equation (10.30)

$$\overline{LH_\alpha}\psi = \varphi$$

has a unique solution ψ in $D(\overline{LH_\alpha})$ for any $\varphi \in C(\partial D)$; hence the inverse $\overline{LH_\alpha}^{-1}$ of $\overline{LH_\alpha}$ can be defined on the whole space $C(\partial D)$. Further the operator $-\overline{LH_\alpha}^{-1}$ is non-negative and bounded on $C(\partial D)$.

We have, by inequality (10.27) and transversality condition (10.7),

$$\ell_\alpha = -\sup_{x' \in \partial D} LH_\alpha 1(x') > 0.$$

Furthermore, by using Corollary 10.13 with $K := \partial D$, $A := \overline{LH_\alpha}$ and $c := \ell_\alpha$ we obtain that the operator $\overline{LH_\alpha} + \ell_\alpha I$ is the infinitesimal generator of some Feller semigroup on ∂D. Therefore, since $\ell_\alpha > 0$, it follows from an application of part (i) of Theorem 9.35 with $A := \overline{LH_\alpha} + \ell_\alpha I$ that the equation

$$-\overline{LH_\alpha}\,\psi = \left(\ell_\alpha I - (\overline{LH_\alpha} + \ell_\alpha I)\right)\psi = \varphi$$

has a unique solution $\psi \in D(\overline{LH_\alpha})$ for any $\varphi \in C(\partial D)$, and further that the operator $-\overline{LH_\alpha}^{-1} = \left(\ell_\alpha I - (\overline{LH_\alpha} + \ell_\alpha I)\right)^{-1}$ is non-negative and bounded on the space $C(\partial D)$ with norm

$$\left\| -\overline{LH_\alpha}^{-1} \right\| = \left\| \left(\ell_\alpha I - (\overline{LH_\alpha} + \ell_\alpha I)\right)^{-1} \right\|_\infty \leq \frac{1}{\ell_\alpha}.$$

Step 5: By assertion (10.30), we can define

$$G_\alpha f = G_\alpha^0 f - H_\alpha\left(\overline{LH_\alpha}^{-1}\left(\overline{LG_\alpha^0 f}\right)\right) \quad \text{for every } f \in C(\overline{D}). \tag{10.21}$$

Step 5-1: First, we prove that

$$G_\alpha = (\alpha I - \mathfrak{A})^{-1}, \quad \alpha > 0. \tag{10.31}$$

In view of Lemmas 10.10 and 10.20, it follows that we have, for all $f \in C(\overline{D})$,

$$\begin{cases} G_\alpha f = G_\alpha^0 f - H_\alpha \left(\overline{LH_\alpha}^{-1} \left(\overline{LG_\alpha^0} f \right) \right) \in D(\overline{W}), \\ G_\alpha f|_{\partial D} = -\overline{LH_\alpha}^{-1} \left(\overline{LG_\alpha^0} f \right) \in D\left(\overline{LH_\alpha}\right) = \mathcal{D}, \\ LG_\alpha f = \overline{LG_\alpha^0} f - \overline{LH_\alpha} \left(\overline{LH_\alpha}^{-1} \left(\overline{LG_\alpha^0} f \right) \right) = 0, \end{cases}$$

and that

$$(\alpha I - \overline{W}) G_\alpha f = f.$$

This proves that

$$\begin{cases} G_\alpha f \in D(\mathfrak{A}), \\ (\alpha I - \mathfrak{A}) G_\alpha f = f, \end{cases}$$

that is,

$$(\alpha I - \mathfrak{A}) G_\alpha = I \quad \text{on } C(\overline{D}).$$

Therefore, in order to prove (10.31) it suffices to show the injectivity of the operator $\alpha I - \mathfrak{A}$ for $\alpha > 0$.
Assume that

$$u \in D(\mathfrak{A}) \text{ and } (\alpha I - \mathfrak{A}) u = 0.$$

Then, by Corollary 10.13, the function u can be written as

$$u = H_\alpha (u|_{\partial D}), \quad u|_{\partial D} \in \mathcal{D} = D\left(\overline{LH_\alpha}\right).$$

Thus we have

$$\overline{LH_\alpha} (u|_{\partial D}) = Lu = 0.$$

In view of assertion (10.30), this implies that

$$u|_{\partial D} = 0,$$

so that

$$u = H_\alpha (u|_{\partial D}) = 0 \quad \text{in } D.$$

Step 5-2: The non-negativity of G_α, $\alpha > 0$, follows immediately from formula (10.25), since the operators G_α^0, H_α, $-\overline{LH_\alpha}^{-1}$ and $\overline{LG_\alpha^0}$ are all non-negative.

Step 5-3: We prove that the operator G_α is bounded on the space $C(\overline{D})$ with norm

$$\|G_\alpha\| \leq \frac{1}{\alpha}, \quad \alpha > 0. \tag{10.32}$$

To do this, it suffices to show that

$$G_\alpha 1 \leq \frac{1}{\alpha} \quad \text{on } \overline{D}, \tag{10.28$''$}$$

since G_α is non-negative on $C(\overline{D})$.

First, it follows from the uniqueness property of solutions of problem (D$''$) that

$$\alpha G_\alpha^0 1 + H_\alpha 1 = 1 + G_\alpha^0 (W_D 1) \quad \text{on } \overline{D}. \tag{10.33}$$

In fact, both sides have the same boundary value 1 and satisfy the same equation: $(\alpha - W_D)u = \alpha$ in D.
By applying the operator L to both sides of equality (10.33), we obtain from condition (1.6) that

$$\begin{aligned}
&-LH_\alpha 1(x') \\
&= -L1(x') - LG_\alpha^0(W_D 1)(x') + \alpha L G_\alpha^0 1(x') \\
&= -\left(\gamma(x') + \eta_\tau(x') + \int_{\partial D} r(x', y') \left[1 - \tau(x', y')\right] dy'\right) \\
&\quad -\mu(x') \frac{\partial}{\partial \mathbf{n}} (G_\alpha^0 (W_D 1))(x') - \int_D t(x', y) G_\alpha^0 (W_D 1)(y) dy + \alpha L G_\alpha^0 1(x') \\
&\geq \alpha L G_\alpha^0 1(x') \quad \text{on } \partial D,
\end{aligned}$$

since we have

$$\begin{aligned}
G_\alpha^0(W_D 1) &= 0 \quad \text{on } \partial D, \\
G_\alpha^0(W_D 1) &\leq 0 \quad \text{on } \overline{D}.
\end{aligned}$$

Hence we have, by the non-negativity of $-\overline{LH_\alpha}^{-1}$,

$$-\overline{LH_\alpha}^{-1}\left(L G_\alpha^0 1\right) \leq \frac{1}{\alpha} \quad \text{on } \partial D. \tag{10.34}$$

By using (10.18) with $f := 1$, inequality (10.34) and equality (10.33), we obtain that

$$\begin{aligned} G_\alpha 1 &= G_\alpha^0 1 + H_\alpha \left(-\overline{LH_\alpha}^{-1} \left(LG_\alpha^0 1 \right) \right) \\ &\le G_\alpha^0 1 + \frac{1}{\alpha} H_\alpha 1 = \frac{1}{\alpha} + \frac{1}{\alpha} G_\alpha^0 (W_D 1) \\ &\le \frac{1}{\alpha} \quad \text{on } \overline{D}, \end{aligned}$$

since the operators H_α and G_α^0 are non-negative and since $W_D 1 \le 0$ in D.

Step 5-4: Finally, we prove that

$$\text{The domain } D(\mathfrak{A}) \text{ is } \textit{dense} \text{ in the space } C(\overline{D}). \tag{10.35}$$

Before the proof, we need some lemmas on the behavior of G_α^0, H_α and $-\overline{LH_\alpha}^{-1}$ as $\alpha \to +\infty$ (see [BCP, Proposition III.1.6]; [Ta5, Lemmas 9.6.19 and 9.6.20]):

Lemma 10.24. *For every $f \in C(\overline{D})$, we have*

$$\lim_{\alpha \to +\infty} \left[\alpha G_\alpha^0 f + H_\alpha \left(f|_{\partial D} \right) \right] = f \quad \textit{in } C(\overline{D}). \tag{10.36}$$

Proof. Choose a constant $\beta > 0$ and let

$$g = f - H_\beta (f|_{\partial D}).$$

Then, by using (10.11) with $\varphi := f|_{\partial D}$ we obtain that

$$\alpha G_\alpha^0 g - g = \left[\alpha G_\alpha^0 f + H_\alpha (f|_{\partial D}) - f \right] - \beta G_\alpha^0 H_\beta (f|_{\partial D}). \tag{10.37}$$

However, we have, by estimate (10.32),

$$\lim_{\alpha \to +\infty} G_\alpha^0 H_\beta (f|_{\partial D}) = 0 \quad \text{in } C(\overline{D}),$$

and by assertion (10.14′)

$$\lim_{\alpha \to +\infty} \alpha G_\alpha^0 g = g \quad \text{in } C(\overline{D}),$$

since $g|_{\partial D} = 0$. Therefore, the desired formula (10.36) follows by letting $\alpha \to +\infty$ in (10.37).

The proof of Lemma 10.24 is complete.

Lemma 10.25. *The function*

$$\Pi_\alpha 1(x') = \frac{\partial}{\partial \mathbf{n}} (H_\alpha 1)(x'), \quad x' \in \partial D,$$

diverges to $-\infty$ uniformly and monotonically as $\alpha \to +\infty$.

Proof. First, formula (10.15) with $\varphi := 1$ gives that

$$H_\alpha 1 = H_\beta 1 - (\alpha - \beta) G_\alpha^0 H_\beta 1.$$

Thus, in view of the non-negativity of G_α^0 and H_α it follows that

$$\alpha \geq \beta \quad \Longrightarrow \quad H_\alpha 1 \leq H_\beta 1 \quad \text{on } \overline{D}.$$

Since $H_\alpha 1|_{\partial D} = H_\beta 1|_{\partial D} = 1$, this implies that the functions

$$\frac{\partial}{\partial \mathbf{n}} (H_\alpha 1)(x'), \quad x' \in \partial D,$$

are monotonically non-increasing in α. Furthermore, by using formula (10.14) with $f := H_\beta 1$ we find that the function

$$H_\alpha 1(x) = H_\beta 1(x) - \left(1 - \frac{\beta}{\alpha}\right) \alpha G_\alpha^0 H_\beta 1(x)$$

converges to zero monotonically as $\alpha \to +\infty$, for each interior point x of D.

Now, for any given constant $K > 0$ we can construct a function $u \in C^2(\overline{D})$ such that

$$u = 1 \quad \text{on } \partial D, \tag{10.38a}$$

$$\frac{\partial u}{\partial \mathbf{n}} \leq -K \quad \text{on } \partial D. \tag{10.38b}$$

Indeed, it follows from an application of Theorem 10.4 that, for any integer $m > 0$, the function

$$u = (H_{\alpha_0} 1)^m, \quad \alpha_0 > 0,$$

belongs to $C^{2+\theta}(\overline{D})$ and satisfies condition (10.38a). Furthermore, we have the inequality

$$\frac{\partial u}{\partial \mathbf{n}} = m \frac{\partial}{\partial \mathbf{n}} (H_{\alpha_0} 1) \leq m \sup_{x' \in \partial D} \frac{\partial}{\partial \mathbf{n}} (H_{\alpha_0} 1)(x').$$

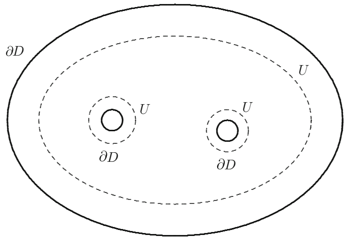


Fig. 10.7 The neighborhood U of ∂D

However, since the function $H_{\alpha_0}1$ takes its positive maximum 1 only on the boundary ∂D, we can apply the Hopf boundary point lemma (see Theorem 8.15) to obtain that

$$\frac{\partial}{\partial \mathbf{n}}(H_{\alpha_0}1) < 0 \quad \text{on } \partial D.$$

In view of this inequality, we obtain that the function $u = (H_{\alpha_0}1)^m$ satisfies condition (10.38b) for m sufficiently large.

Take a function $u(x)$ in $C^2(\overline{D})$ satisfying conditions (10.38a) and (10.38b), and choose a neighborhood U of ∂D, relative to $\overline{D}$, with smooth boundary ∂U such that (see Fig. 10.7)

$$u \geq \frac{1}{2} \quad \text{on } U. \tag{10.39}$$

Recall that the function $H_\alpha 1$ converges to zero in D monotonically as $\alpha \to +\infty$. Since $u = H_\alpha 1 = 1$ on ∂D, by using Dini's theorem we can find a constant $\alpha > 0$ (depending on u and hence on K) such that

$$H_\alpha 1 \leq u \quad \text{on } \partial U \setminus \partial D, \tag{10.40a}$$

$$\alpha > 2\|W_D u\|_\infty. \tag{10.40b}$$

It follows from inequalities (10.39) and (10.40b) that

$$(W_D - \alpha)(H_\alpha 1 - u) = \alpha u - W_D u \geq \frac{\alpha}{2} - \|W_D u\|_\infty > 0 \quad \text{in } U.$$

Thus, by applying the weak maximum principle (Theorem 8.11) with $W := W - \alpha$ to the function $H_\alpha 1 - u$ we obtain that the function $H_\alpha 1 - u$ may take its positive maximum only on the boundary ∂U. However, conditions (10.38a) and (10.40a) imply that

$$H_\alpha 1 - u \leq 0 \quad \text{on } \partial U = (\partial U \setminus \partial D) \cup \partial D.$$

Therefore, we have the inequality

$$H_\alpha 1 \leq u \quad \text{on } \overline{U} = U \cup \partial U,$$

and hence

$$\frac{\partial}{\partial \mathbf{n}}(H_\alpha 1) \leq \frac{\partial u}{\partial \mathbf{n}} \leq -K \quad \text{on } \partial D,$$

since $u|_{\partial D} = H_\alpha 1|_{\partial D} = 1$.

The proof of Lemma 10.25 is complete.

Corollary 10.26. *If the boundary condition L is transversal on the boundary ∂D, then we have*

$$\lim_{\alpha \to +\infty} \left\| -\overline{LH_\alpha}^{-1} \right\| = 0.$$

Proof. First, we have, by condition (1.6),

$$\begin{aligned}
&LH_\alpha 1(x') \\
&= \gamma(x') + \mu(x')\frac{\partial}{\partial \mathbf{n}}(H_\alpha 1)(x') - \alpha\delta(x') \\
&\quad + \eta_\tau(x') + \int_{\partial D} r(x', y')\left[1 - \tau(x', y')\right] dy' + \int_D t(x', y)\left[H_\alpha 1(y) - 1\right] dy \\
&= \mu(x')\frac{\partial}{\partial \mathbf{n}}(H_\alpha 1)(x') - \alpha\delta(x') \\
&\quad + \left(\gamma(x') + \eta_\tau(x') + \int_{\partial D} r(x', y')\left[1 - \tau(x', y')\right] dy'\right) \\
&\quad + \int_D t(x', y)\left[H_\alpha 1(y) - 1\right] dy \\
&\leq \mu(x')\frac{\partial}{\partial \mathbf{n}}(H_\alpha 1)(x') - \alpha\delta(x') + \int_D t(x', y)\left[H_\alpha 1(y) - 1\right] dy.
\end{aligned}$$

However, it follows from an application of Beppo Levi's theorem that

$$\lim_{\alpha \to +\infty} \int_D t(x', y)\left[H_\alpha 1(y) - 1\right] dy = -\int_D t(x', y) dy,$$

since the function $H_\alpha 1$ converges to zero in D monotonically as $\alpha \to +\infty$.

Hence we obtain from Lemma 10.25 that if the boundary condition L is *transversal* on the boundary ∂D, that is, if we have the condition

$$\int_D t(x', y) dy = +\infty \quad \text{if } \mu(x') = \delta(x') = 0,$$

then the function $LH_\alpha 1$ diverges to $-\infty$ monotonically as $\alpha \to +\infty$. By Dini's theorem, this convergence is uniform in $x' \in \partial D$. Thus it follows that the function

$$\frac{1}{LH_\alpha 1(x')}$$

converges to zero uniformly in $x' \in \partial D$ as $\alpha \to +\infty$. This proves that

$$\left\| -\overline{LH_\alpha}^{-1} \right\| = \left\| -\overline{LH_\alpha}^{-1} 1 \right\| \leq \left\| \frac{1}{LH_\alpha 1} \right\| \longrightarrow 0 \quad \text{as } \alpha \to +\infty,$$

since we have the inequality

$$1 = \frac{-LH_\alpha 1(x')}{|LH_\alpha 1(x')|} \leq \left\| \frac{1}{LH_\alpha 1} \right\| \left(-LH_\alpha 1(x') \right) \quad \text{for all } x' \in \partial D.$$

The proof of Corollary 10.26 is complete.

10.5.2 End of Proof of Theorem 1.2

In view of formula (10.31) and inequality (10.32), it suffices to prove that

$$\lim_{\alpha \to +\infty} \|\alpha G_\alpha f - f\|_\infty = 0 \quad \text{for all } f \in C^{2+\theta}(\overline{D}), \tag{10.41}$$

since the space $C^{2+\theta}(\overline{D})$ is dense in $C(\overline{D})$.

First, we observe that

$$\begin{aligned}
\|\alpha G_\alpha f - f\|_\infty &= \left\| \alpha G_\alpha^0 f - \alpha H_\alpha \left(\overline{LH_\alpha}^{-1} \left(LG_\alpha^0 f \right) \right) - f \right\|_\infty \\
&\leq \left\| \alpha G_\alpha^0 f + H_\alpha \left(f|_{\partial D} \right) - f \right\|_\infty \\
&\quad + \left\| -\alpha H_\alpha \left(\overline{LH_\alpha}^{-1} \left(LG_\alpha^0 f \right) \right) - H_\alpha \left(f|_{\partial D} \right) \right\|_\infty \\
&\leq \left\| \alpha G_\alpha^0 f + H_\alpha \left(f|_{\partial D} \right) - f \right\|_\infty \\
&\quad + \left\| -\alpha \overline{LH_\alpha}^{-1} \left(LG_\alpha^0 f \right) - f|_{\partial D} \right\|_\infty .
\end{aligned}$$

Thus, in view of (10.36) it suffices to show that

$$\lim_{\alpha \to +\infty} \left[-\alpha \overline{LH_\alpha}^{-1} \left(LG_\alpha^0 f \right) - f|_{\partial D} \right] = 0 \quad \text{in } C(\partial D). \tag{10.42}$$

Take a constant β such that $0 < \beta < \alpha$, and write

$$f = G_\beta^0 g + H_\beta \varphi,$$

where (cf. (10.17)):

$$\begin{cases} g = (\beta - W_D) f \in C^\theta(\overline{D}), \\ \varphi = f|_{\partial D} \in C^{2+\theta}(\partial D). \end{cases}$$

Then, by using Eqs. (10.13) (with $f := g$) and (10.15) we obtain that

$$G_\alpha^0 f = G_\alpha^0 G_\beta^0 g + G_\alpha^0 H_\beta \varphi = \frac{1}{\alpha - \beta} \left(G_\beta^0 g - G_\alpha^0 g + H_\beta \varphi - H_\alpha \varphi \right).$$

Hence we have the inequality

$$\begin{aligned}
&\left\| -\alpha \overline{LH_\alpha}^{-1} \left(LG_\alpha^0 f \right) - f|_{\partial D} \right\|_\infty \\
&= \left\| \frac{\alpha}{\alpha - \beta} \left(-\overline{LH_\alpha}^{-1} \right) \left(LG_\beta^0 g - LG_\alpha^0 g + LH_\beta \varphi \right) + \frac{\alpha}{\alpha - \beta} \varphi - \varphi \right\|_\infty \\
&\leq \frac{\alpha}{\alpha - \beta} \left\| -\overline{LH_\alpha}^{-1} \right\| \cdot \left\| LG_\beta^0 g + LH_\beta \varphi \right\|_\infty \\
&\quad + \frac{\alpha}{\alpha - \beta} \left\| -\overline{LH_\alpha}^{-1} \right\| \cdot \left\| LG_\alpha^0 \right\|_\infty \cdot \|g\|_\infty + \frac{\beta}{\alpha - \beta} \|\varphi\|_\infty .
\end{aligned}$$

By Corollary 10.26, it follows that the first term on the last inequality converges to zero as $\alpha \to +\infty$. For the second term, by using (10.13) with $f := 1$ and the non-negativity of G_β^0 and LG_α^0 we find that

$$\left\| LG_\alpha^0 \right\| = \left\| LG_\alpha^0 1 \right\|_\infty = \left\| LG_\beta^0 1 - (\alpha - \beta) LG_\alpha^0 G_\beta^0 1 \right\|_\infty \leq \left\| LG_\beta^0 1 \right\|_\infty .$$

Hence the second term also converges to zero as $\alpha \to +\infty$. It is clear that the third term converges to zero as $\alpha \to +\infty$. This completes the proof of assertion (10.42) and hence that of assertion (10.41).

The proof of assertion (10.35) is complete.

Step 6: Summing up, we have proved that the operator $\mathfrak{A}$, defined by formula (10.24), satisfies conditions (a)–(d) in Theorem 9.35. Hence it follows from an application of the same theorem that the operator $\mathfrak{A}$ is the infinitesimal generator of some Feller semigroup on $\overline{D}$.

Now the proof of Theorem 10.21 and hence that of Theorem 1.2 is complete. □

10.6 Unique Solvability for Second-Order Pseudo-differential Operators

In this section we give a sketch of the proof of a unique solvability theorem for a class of second-order classical pseudo-differential operators in the framework of Hölder spaces (Theorem 10.23). The proof of Theorem 10.23 is based on a method of *elliptic regularizations* essentially due to Oleĭnik–Radkevič [OR, Chapter I] developed for second-order differential operators with non-negative characteristic form, just as in the proof of Cancelier [Cn, Théorème 4.5]. In order to prove estimate (10.28), we need an interpolation argument. Moreover, we remark that Corollary 3.26 to Mazur's theorem (Theorem 3.25) in Chap. 3 plays an important role in the proof of estimate (10.28).

10.6.1 Fundamental Results for Second-Order Pseudo-differential Operators

Let M be an n-dimensional compact smooth manifold M without boundary. We begin with the following two fundamental results for second-order classical pseudo-differential operators T on M:

Theorem 10.27. *Let $T = P + S$ be a second-order classical pseudo-differential operator on M, as in Theorem 10.23. Assume that*

$$T1(x) = P1(x) + S1(x) < 0 \quad \text{on } M.$$

Then we have, for all $\varphi \in C^2(M)$,

$$\|\varphi\|_{C(M)} \leq \left(\frac{1}{\min_M(-T1)} \right) \|T\varphi\|_{C(M)}.$$

Theorem 10.27 is a compact manifold version of Theorem 8.12 formulated in Chap. 8.

Theorem 10.28. *Let $T = P + S$ be a second-order classical pseudo-differential operator on M, as in Theorem 10.23. Assume that the operator T is elliptic on M and satisfies the condition*

$$T1 = P1 + S1 < 0 \quad \text{on } M.$$

Then, for every integer $k \geq 0$ the operator

$$T : C^{k+2+\theta}(M) \longrightarrow C^{k+\theta}(M)$$

is bijective.

Proof. Since T is elliptic and its principal symbol is *real*, it follows from an application of [Ta5, Corollary 6.7.12] that T is a Fredholm operator with index zero

$$\operatorname{ind} T = \dim N(T) - \operatorname{codim} R(T) = 0.$$

Here we recall the following definitions (see Sect. 3.8.1):

(i) The null space $N(T) = \{\varphi \in C^{k+2+\theta}(M) : T\varphi = 0\}$ of T has finite dimension, that is, $\dim N(T) < \infty$.
(ii) The range $R(T) = \{T\varphi : \varphi \in C^{k+2+\theta}(M)\}$ of T has finite codimension, that is,

$$\operatorname{codim} R(T) = \dim C^{k+\theta}(M)/R(T) < \infty.$$

However, Theorem 10.27 asserts that T is injective, that is, $\dim N(T) = 0$. Therefore, we obtain that $\operatorname{codim} R(T) = 0$. This proves that T is surjective.

10.6.2 *Proof of Theorem 10.23*

If k is a positive integer, we have the following characterization of the Sobolev space $W^{k,\infty}(M)$:

$$W^{k,\infty}(M) = \left\{\varphi \in C^{k-1}(M) : \max_{|\alpha|\leq k-1\, \sup_{x,y\in M x\neq y}} \frac{|\partial^\alpha\varphi(x) - \partial^\alpha\varphi(y)|}{|x-y|} < \infty\right\},$$

where $|x - y|$ is the geodesic distance between x and y with respect to the Riemannian metric of the manifold M.

The proof of Theorem 10.23 is divided into three steps.

Step I: First, we prove Theorem 10.23 for the Sobolev space $W^{1,\infty}(M)$ ($k = 1$):

Lemma 10.29. *There exists a constant* $\lambda = \lambda(1) > 0$ *such that, for any* $f \in W^{1,\infty}(M)$*, we can find a function* $\varphi \in W^{1,\infty}(M)$ *satisfying the equation*

$$(T - \lambda I)\varphi = f \quad \text{on } M,$$

and the estimate

$$\|\varphi\|_{1,\infty} \leq C_1(\lambda)\|f\|_{1,\infty}.$$

Here $C_1(\lambda) > 0$ *is a constant independent of* f.

Proof. The proof is divided into three steps.

Step 1: First, we construct a smooth function $\sigma(x, y)$ on $M \times M$ which satisfies the following two conditions (a) and (b):

(a) $0 \leq \sigma(x, y) \leq 1$ on $M \times M$.
(b) $\sigma(x, y) = 1$ in a neighborhood of the diagonal Δ_M in $M \times M$.

Let $\{U_\alpha\}_{\alpha=1}^{\ell}$ be a finite open covering of M and let $\{\varphi_\alpha\}_{\alpha=1}^{\ell}$ be a partition of unity subordinate to the covering $\{U_\alpha\}$ (see Sect. 5.7.2). That is, the family $\{\varphi_\alpha\}$ in $C^\infty(M)$ satisfies the following three conditions (1), (2) and (3):

(1) $0 \leq \varphi_\alpha(x) \leq 1$ for all $x \in M$.
(2) $\operatorname{supp} \varphi_\alpha \subset U_\alpha$ for each α.
(3) $\sum_{\alpha=1}^{\ell} \varphi_\alpha(x) = 1$ for each $x \in M$.

If we take a smooth function $\psi_\alpha(x)$ in M such that

$$\begin{cases} 0 \leq \psi_\alpha(x) \leq 1 & \text{for all } x \in M, \\ \psi_\alpha(x) = 1 & \text{on } \operatorname{supp} \varphi_\alpha, \end{cases}$$

then it is easy to verify that the function

$$\sigma(x, y) = \sum_{\alpha=1}^{\ell} \varphi_\alpha(x)\, \psi_\alpha(y), \quad (x, y) \in M \times M,$$

satisfies the desired conditions (a) and (b).

Now we find that the operator $T = P + S$ can be written, in terms of local coordinates $(x_1, x_2, \ldots, x_n)$, in the form

$$\begin{aligned} T\varphi(x) &= \sum_{i,j=1}^{n} \alpha^{ij}(x) \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} \beta^i(x) \frac{\partial \varphi}{\partial x_i}(x) + \gamma(x)\varphi(x) \\ &\quad + \int_M s(x, y) \left[\varphi(y) - \sigma(x, y) \left(\varphi(x) + \sum_{i=1}^{n} (y_i - x_i) \frac{\partial \varphi}{\partial x_i}(x) \right) \right] dy \\ &:= Q\varphi(x) + S\varphi(x). \end{aligned}$$

Here:

(a) The differential operator

$$\sum_{i,j=1}^{n} \alpha^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}$$

is the principal part of P; more precisely, the $\alpha^{ij}(x)$ are the components of a smooth symmetric contravariant tensor of type $\binom{2}{0}$ on M which satisfies the condition

$$\sum_{i,j=1}^{n} \alpha^{ij}(x)\xi_i\xi_j \geq 0, \quad x \in M, \ \xi = \sum_{j=1}^{n} \xi_j dx_j \in T_x^*(M),$$

where $T_x^*(M)$ is the cotangent space of M at x.

(b) The function $\sigma(x,y) = \sum_{\alpha=1}^{N} \sigma_\alpha(x,y)$ is a local unity function on M.

(c) The density dy is a strictly positive density on M.

(d) $T1(x) = Q1(x) + S1(x) = \gamma(x) + \int_M s(x,y)[1 - \sigma(x,y)]\,dy \leq 0$ on M.

Furthermore, it should be emphasized that there exists a constant $C > 0$ such that the distribution kernel $s(x,y)$ of $S \in L_{\mathrm{cl}}^{2-\kappa}(M)$, $\kappa > 0$, satisfies the estimate (see Theorem 7.36)

$$|s(x,y)| \leq \frac{C}{|x-y|^{n+2-\kappa}} \quad \text{for all } (x,y) \in (M \times M) \setminus \Delta_M.$$

Hence we find that the integral

$$S_r\varphi(x) = \int_M s(x,y) \left[\varphi(y) - \sigma(x,y) \left(\varphi(x) + \sum_{i=1}^{n} (y_i - x_i) \frac{\partial \varphi}{\partial x_i}(x) \right) \right] dy$$

is absolutely convergent, since $\kappa > 0$ and $\sigma(x,y) = 1$ in a neighborhood of the diagonal Δ_M.

Now, if $\varphi \in C^1(M)$, we define a continuous function $B_T(\varphi,\varphi)$ on M by the formula

$$\begin{aligned} B_T(\varphi,\varphi)(x) &= 2 \sum_{i,j=1}^{n} \alpha^{ij}(x) \frac{\partial \varphi}{\partial x_i}(x) \frac{\partial \varphi}{\partial x_j}(x) \\ &\quad + \int_M s(x,y)\,(\varphi(y) - \varphi(x))^2\,dy - T1(x) \cdot \varphi(x)^2, \quad x \in M. \end{aligned}$$

It should be noted that the function $B_T(\varphi,\varphi)$ is *non-negative* on M for all $\varphi \in C^1(M)$.

The next result may be proved just as in the proof of Cancelier [Cn, Théorème 4.1].

Lemma 10.30. *Let $\{X_j\}_{j=1}^{r}$ be a family of real smooth vector fields on M such that the X_j span the tangent space $T_x(M)$ at each point x of M. If $\varphi \in C^\infty(M)$, we let*

$$p_1(x) = \sum_{j=1}^{r} |X_j \varphi(x)|^2, \quad x \in M,$$

and

$$R_1(x) = T p_1(x) - \sum_{j=1}^{r} B_T(X_j \varphi, X_j \varphi)(x), \quad x \in M.$$

Then, for each $\eta > 0$ *there exist constants* $\beta_0 > 0$ *and* $\beta_1 > 0$ *such that we have, for all* $\varphi \in C^\infty(M)$*,*

$$|R_1(x)| \leq \eta \sum_{j=1}^{r} B_T(X_j \varphi, X_j \varphi)(x) + \beta_0 \|\varphi\|^2_{C(M)} \tag{10.43}$$

$$+ \beta_1 \|\varphi\|^2_{C^1(M)} + \frac{1}{2} \|T\varphi\|^2_{C^1(M)}, \quad x \in M.$$

Remark 10.31. The constants β_0 and β_1 are *uniform* for the operators $T + \varepsilon\Lambda - \lambda I$, $0 \leq \varepsilon \leq 1$, $\lambda \geq 0$, where Λ is a second-order *elliptic* differential operator on M defined by the formula

$$\Lambda = -\sum_{j=1}^{r} X_j^* X_j = \sum_{j=1}^{r} X_j^2 + \sum_{j=1}^{r} \div X_j \cdot X_j.$$

Step 2: First, let $f(x)$ be an arbitrary element of $C^\infty(M)$. Since the operator $T + \varepsilon\Lambda - \lambda I$ is elliptic for all $\varepsilon > 0$ and $(T + \varepsilon\Lambda - \lambda I)1 = T1 - \lambda \leq -\lambda < 0$ on M for $\lambda > 0$, it follows from an application of Theorem 10.28 that we can find a unique function $\varphi_\varepsilon \in C^\infty(M)$ such that

$$(T + \varepsilon\Lambda - \lambda I)\varphi_\varepsilon = f \quad \text{on } M.$$

Furthermore, by applying Theorem 10.27 to the operator $T + \varepsilon\Lambda - \lambda I$ we obtain that

$$\|\varphi_\varepsilon\|_{C(M)} \leq \frac{1}{\lambda} \|f\|_{C(M)}, \tag{10.44}$$

since $\min_M (-(T + \varepsilon\Lambda - \lambda I)1) \geq \lambda$.

Let x_0 be a point of M at which the function

$$p_1^\varepsilon(x) = \sum_{j=1}^{r} |X_j \varphi_\varepsilon(x)|^2$$

attains its positive maximum. Then we have the inequality

$$\Lambda p_1^\varepsilon(x_0) = \left(\sum_{j=1}^r X_j^2\right) p_1^\varepsilon(x_0) \leq 0,$$

and also

$$\begin{aligned} T p_1^\varepsilon(x_0) &= \sum_{i,j=1}^n \alpha^{ij}(x_0) \frac{\partial^2 p_1^\varepsilon}{\partial x_i \partial x_j}(x_0) + \gamma(x_0) p_1^\varepsilon(x_0) \\ &\quad + \int_M s(x_0, y)[p_1^\varepsilon(y) - \sigma(x_0, y) p_1^\varepsilon(x_0)]\, dy \\ &\leq \left(\gamma(x_0) + \int_M s(x_0, y)[1 - \sigma(x_0, y)]\, dy\right) p_1^\varepsilon(x_0) \\ &\quad + \int_M s(x_0, y)[p_1^\varepsilon(y) - p_1^\varepsilon(x_0)]\, dy \\ &\leq T1(x_0) \cdot p_1^\varepsilon(x_0). \end{aligned}$$

Hence, by using inequality (10.43) with $\eta := 1/2$ and inequality (10.44) we obtain that

$$\begin{aligned} \lambda p_1^\varepsilon(x_0) &\leq (\lambda - T1(x_0)) p_1^\varepsilon(x_0) - \varepsilon \Lambda p_1^\varepsilon(x_0) \\ &\leq (\lambda - T - \varepsilon\Lambda) p_1^\varepsilon(x_0) \\ &= -\left((T + \varepsilon\Lambda - \lambda) p_1^\varepsilon(x_0) - \sum_{j=1}^r B_{T+\varepsilon\Lambda-\lambda I}(X_j\varphi_\varepsilon, X_j\varphi_\varepsilon)(x_0)\right) \\ &\quad - \sum_{j=1}^r B_{T+\varepsilon\Lambda-\lambda I}(X_j\varphi_\varepsilon, X_j\varphi_\varepsilon)(x_0) \\ &\leq -\frac{1}{2} \sum_{j=1}^r B_{T+\varepsilon\Lambda-\lambda I}(X_j\varphi_\varepsilon, X_j\varphi_\varepsilon)(x_0) + \beta_0 \|\varphi_\varepsilon\|_{C(M)}^2 \\ &\quad + \beta_1 \|\varphi_\varepsilon\|_{C^1(M)}^2 + \frac{1}{2} \|f\|_{C^1(M)}^2 \\ &\leq \frac{\beta_0}{\lambda^2} \|f\|_{C(M)}^2 + \beta_1 \|\varphi_\varepsilon\|_{C^1(M)}^2 + \frac{1}{2} \|f\|_{C^1(M)}^2 . \end{aligned}$$

This proves that

$$(\lambda - \beta_1)\|\varphi_\varepsilon\|^2_{C^1(M)} \leq \lambda \left(\|\varphi_\varepsilon\|^2_{C(M)} + p_1^\varepsilon(x_0)\right) - \beta_1\|\varphi_\varepsilon\|^2_{C^1(M)} \quad (10.45)$$
$$\leq \frac{1}{\lambda}\|f\|^2_{C(M)} + \frac{\beta_0}{\lambda^2}\|f\|^2_{C(M)} + \frac{1}{2}\|f\|^2_{C^1(M)}.$$

Here it should be emphasized (see Remark 10.31) that the constants β_0 and β_1 are independent of $\varepsilon > 0$ and $\lambda > 0$. Therefore, if $\lambda > 0$ is so large that

$$\lambda > \beta_1,$$

then it follows from inequality (10.45) that

$$\|\varphi_\varepsilon\|^2_{C^1(M)} \leq C(\lambda)\|f\|^2_{C^1(M)}, \quad (10.46)$$

where $C(\lambda) > 0$ is a constant *independent* of $\varepsilon > 0$.

Step 3: Now let $f(x)$ be an arbitrary element of $W^{1,\infty}(M)$. Then we can find a sequence $\{f_p\}_{p=1}^\infty$ in $C^\infty(M)$ such that

$$\begin{cases} f_p \longrightarrow f & \text{in } C(M), \\ \|f_p\|_{C^1(M)} \leq \|f\|_{1,\infty}. \end{cases}$$

If $\varphi_{\varepsilon p}$ is a unique solution in $C^\infty(M)$ of the equation

$$(T + \varepsilon\Lambda - \lambda I)\varphi_{\varepsilon p} = f_p \quad \text{on M}, \quad (10.47)$$

it follows from an application of inequality (10.46) that

$$\|\varphi_{\varepsilon p}\|^2_{C^1(M)} \leq C(\lambda)\|f_p\|^2_{C^1(M)} \leq C(\lambda)\|f\|^2_{1,\infty}.$$

This proves that the sequence $\{\varphi_{\varepsilon p}\}$ is uniformly bounded and equicontinuous on M. Hence, by virtue of the Ascoli–Arzelà theorem we can choose a subsequence $\{\varphi_{\varepsilon' p'}\}$ which converges uniformly to a function $\varphi \in C(M)$, as $\varepsilon' \downarrow 0$ and $p' \to \infty$. Furthermore, since the unit ball in the Hilbert space $L^2(M)$ is *sequentially weakly compact* (see Yosida [Yo, Chapter V, Section 2, Theorem 1]), we may assume that the sequence $\{\partial_j \varphi_{\varepsilon' p'}\}$ converges weakly to a function ψ_j in $L^2(M)$, for each $1 \leq j \leq n$. Then it follows that

$$\partial_j \varphi = \psi_j \in L^2(M), \quad 1 \leq j \leq n.$$

On the other hand, it is easy to verify that the set

$$K = \left\{g \in L^2(M) : \|g\|_\infty \leq \sqrt{C(\lambda)}\,\|f\|_{1,\infty}\right\}$$

is convex, strongly closed and *balanced* in $L^2(M)$. Therefore, by applying Corollary 3.26 with

$$X := L^2(M), \quad M := K,$$

we obtain that the set K is *weakly closed* in $L^2(M)$.

However, we have

$$\begin{cases} \partial_j \varphi_{\varepsilon' p'} \in K, \\ \partial_j \varphi_{\varepsilon' p'} \longrightarrow \psi_j \quad \text{weakly in } L^2(M) \text{ for each } 1 \leq j \leq n. \end{cases}$$

Therefore, since the set K is weakly closed in $L^2(M)$, it follows that

$$\partial_j \varphi = \psi_j \in K, \quad 1 \leq j \leq n,$$

that is,

$$\|\partial_j \varphi\|_\infty \leq \sqrt{C(\lambda)}\, \|f\|_{1,\infty}, \quad 1 \leq j \leq n.$$

Summing up, we have proved that

$$\begin{cases} \varphi \in W^{1,\infty}(M), \\ \|\varphi\|_{1,\infty} \leq C_1(\lambda) \|f\|_{1,\infty}, \end{cases}$$

where $C_1(\lambda) > 0$ is a constant *independent* of f.

Finally, by letting $\varepsilon' \downarrow 0$ and $p' \to \infty$ in the equation

$$(T + \varepsilon' \Lambda - \lambda I)\varphi_{\varepsilon' p'} = f_{p'} \quad \text{on M},$$

we obtain that

$$(T - \lambda I)\varphi = f \quad \text{on } M.$$

The proof of Lemma 10.29 is now complete.

Step II: Similarly, we can prove Theorem 10.23 for the general Sobolev space $W^{m,\infty}(M)$ where $m \geq 2$:

Lemma 10.32. *For each integer $m \geq 2$, there exists a constant $\lambda = \lambda(m) > 0$ such that, for any $f \in W^{m,\infty}(M)$, we can find a function $\varphi \in W^{m,\infty}(M)$ satisfying the equation*

$$(T - \lambda I)\varphi = f \quad \text{on } M,$$

and the estimate

$$\|\varphi\|_{m,\infty} \leq C_m(\lambda)\|f\|_{m,\infty}.$$

Here $C_m(\lambda) > 0$ *is a constant independent of* f.

Step III: Theorem 10.23 follows from Lemmas 10.29 and 10.32 by a well-known interpolation argument.

First, by combining Lemmas 10.29 and 10.32 we obtain the following two inequalities (10.48) and (10.49):

$$\|(T - \lambda I)^{-1} f\|_{k,\infty} \leq C_k(\lambda)\|f\|_{k,\infty} \text{ for all } f \in W^{k,\infty}(M). \tag{10.48}$$

$$\|(T - \lambda I)^{-1} f\|_{k+1,\infty} \leq C_{k+1}(\lambda)\|f\|_{k+1,\infty} \text{ for all } f \in W^{k,\infty}(M). \tag{10.49}$$

If we define a real interpolation space

$$\left(W^{k,\infty}(M), W^{k+1,\infty}(M)\right)_{\theta,\infty}, \quad k \in \mathbf{N},\ 0 < \theta < 1,$$

between the Sobolev spaces $W^{k,\infty}(M)$ and $W^{k+1,\infty}(M)$ as follows:

$$\begin{aligned}&\left(W^{k,\infty}(M), W^{k+1,\infty}(M)\right)_{\theta,\infty}\\ &= \left\{u \in W^{k,\infty}(M) : \|u\|_{\theta,\infty} := \sup_{t>0} \frac{K(t,u)}{t^\theta} < \infty\right\},\end{aligned}$$

where

$$\begin{aligned}K(t,u) = \inf\Big\{ &\|u_0\|_{k,\infty} + t\|u_1\|_{k+1,\infty} : u = u_0 + u_1,\\ &u_0 \in W^{k,\infty}(M),\ u_1 \in W^{k+1,\infty}(M)\Big\}.\end{aligned}$$

Then it is known (Bergh–Löfström [BL, Theorem 6.4.5], Triebel [Tr, Theorem 2.4.2]) that the Hölder space $C^{k+\theta}(M)$ coincides with the interpolation space

$$C^{k+\theta}(M) = \left(W^{k,\infty}(M), W^{k+1,\infty}(M)\right)_{\theta,\infty}$$

with the norm $\|\cdot\|_{\theta,\infty}$.

Indeed, it suffices to note the following three assertions (i), (ii) and (iii):

(i) For any integer $m \in \mathbf{N}$, we have the inclusions

$$C^m(M) \subset W^{m,\infty}(M) \subset B^m_{\infty,\infty}(M).$$

(ii) For any integer $k \in \mathbf{N}$ and any $0 < \theta < 1$, we have

$$\left(B^k_{\infty,\infty}(M), B^{k+1}_{\infty,\infty}(M)\right)_{\theta,\infty} = B^{k+\theta}_{\infty,\infty}(M) = C^{k+\theta}(M),$$
$$\left(C^k(M), C^{k+1}(M)\right)_{\theta,\infty} = C^{k+\theta}(M).$$

(iii) For any integer $k \in \mathbf{N}$ and $0 < \theta < 1$, we have the inclusions

$$\begin{aligned} C^{k+\theta}(M) &= \left(C^k(M), C^{k+1}(M)\right)_{\theta,\infty} \\ &\subset \left(W^{k,\infty}(M), W^{k+1,\infty}(M)\right)_{\theta,\infty} \\ &\subset \left(B^k_{\infty,\infty}(M), B^{k+1}_{\infty,\infty}(M)\right)_{\theta,\infty} \\ &= C^{k+\theta}(M). \end{aligned}$$

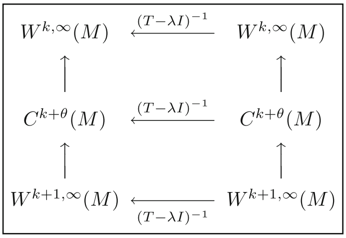
Therefore, by an interpolation argument we obtain from inequalities (10.48) and (10.49) that

$$\|(T - \lambda I)^{-1} f\|_{C^{k+\theta}(M)} \leq C_{k+1}(\lambda)^{\theta}\, C_k(\lambda)^{1-\theta} \|f\|_{C^{k+\theta}(M)} \quad \text{for all } f \in C^{k+\theta}(M).$$

This proves the desired estimate (10.28) with

$$C(\lambda) := C_{k+1}(\lambda)^{\theta}\, C_k(\lambda)^{1-\theta}.$$

The situation may be visualized by the following diagram:



Now the proof of Theorem 10.23 is complete. □

10.7 The Symbol of the First-Order Pseudo-differential Operator Π_α

The purpose of this last section is to explicitly calculate the symbol of the harmonic operator H_α and that of the pseudo-differential operator Π_α in the special case where $P = \Delta$.

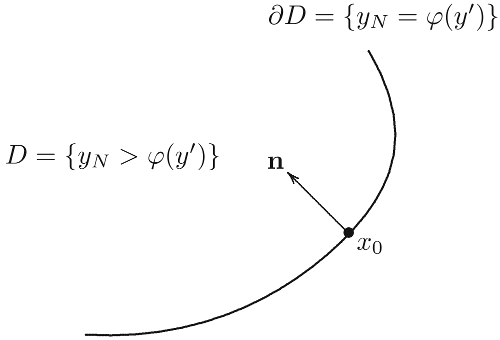


Fig. 10.8 The local coordinate system (y', y_N)

Let x_0 be an arbitrary point of the boundary ∂D. We take the following local coordinate system: Make the y_N-axis coincide with the interior normal **n** at x_0 and the hyperplane $y_N = 0$ coincide with the tangent hyperplane of ∂D at x_0 (see Fig. 10.8). Then the domain D is given by the formula

$$y_N - \varphi(y') > 0,$$

where $\varphi(y')$ is a smooth function of the variables $y' = (y_1, y_2, \ldots, y_{N-1})$.

We may assume that the Taylor expansion of $\varphi(y')$ is given as follows:

$$\varphi(y') = \sum_{j=1}^{N-1} \omega_j y_j^2 + \sum_{i,j,k=1}^{N-1} \omega_{ijk} y_i y_j y_k + O\left(\left|y'\right|^4\right),$$

where the ω_i and ω_{ijk} are constants satisfying the conditions

$$\omega_{kij} = \omega_{ijk} = \omega_{jik}.$$

For any two vectors $\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$ and $\beta' = (\beta_1, \beta_2, \ldots, \beta_{N-1})$ tangent to ∂D at x_0, the bilinear form

$$\omega_{x_0}(\alpha', \beta') = 2 \sum_{j=1}^{N-1} \omega_j \alpha_j \beta_j$$

is the *second fundamental form*, while

$$M(x_0) = \frac{2}{N-1} \sum_{j=1}^{N-1} \omega_j$$

is the *mean curvature* of the boundary ∂D at x_0.

Now we can introduce a new local coordinate system $x = (x', x_N)$ with $x' = (x_1, x_2, \ldots, x_{N-1})$ by the formulas

$$y_i = x_i - \frac{x_N \dfrac{\partial \varphi}{\partial x_i}}{\sqrt{1 + \sum_{i=1}^{N-1} \left(\frac{\partial \varphi}{\partial x_i}\right)^2}}, \quad 1 \leq i \leq N-1,$$

$$y_N = \varphi(x') + \frac{x_N}{\sqrt{1 + \sum_{i=1}^{N-1} \left(\frac{\partial \varphi}{\partial x_i}\right)^2}}.$$

Indeed, it suffices to note that the Jacobian satisfies the condition

$$\frac{D(y_1, y_2, \ldots, y_N)}{D(x_1, x_2, \ldots, x_N)} = 1$$

at the origin, so that $(x_1, x_2, \ldots, x_N)$ can be chosen as a local coordinate system in some neighborhood of the origin (see Fig. 10.9).

Then we have the following Taylor expansions for the local coordinate system $(x_1, x_2, \ldots, x_N)$:

$$\begin{aligned} x_i &= y_i + 2\omega_i y_i y_N + 4\omega_i^2 y_i y_N^2 - 2\omega_i y_i \left(\sum_{j=1}^{N-1} \omega_j y_j^2 \right) \\ &\quad + 3 y_N \sum_{j,k=1}^{N-1} \omega_{ijk} y_j y_k + O\left(|y|^4\right), \\ x_N &= y_N - \sum_{i=1}^{N-1} \omega_i y_i^2 - 2 \left(\sum_{j=1}^{N-1} \omega_j^2 y_j^2 \right) y_N - \sum_{i,j,k=1}^{N-1} \omega_{ijk} y_i y_j y_k + O\left(|y|^4\right). \end{aligned}$$

The Riemannian metric is given by the following formula (10.50):

$$\begin{aligned} ds^2 &= dy_1^2 + dy_2^2 + \ldots + dy_N^2 \\ &= \sum_{i=1}^{N-1} (1 - 2\omega_i x_N)^2 dx_i^2 + 4 \left(\sum_{i=1}^{N-1} \omega_i x_i dx_i \right)^2 \\ &\quad - 12 x_N \sum_{i,j,k=1}^{N-1} \omega_{ijk} x_k dx_i dx_j + dx_N^2 + O\left(|x|^3 |dx|^2\right). \end{aligned} \tag{10.50}$$

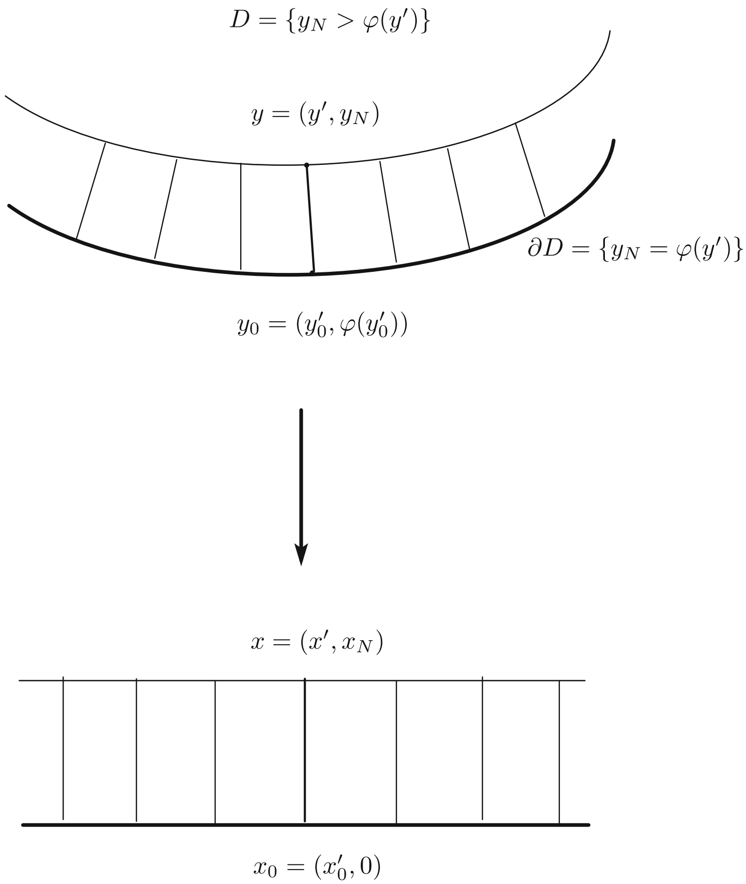


Fig. 10.9 The local coordinate system (x', x_N)

Therefore, we obtain from (10.50) that the symbol of the minus Laplacian

$$-\Delta = -\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_N^2}\right)$$

on D is given by the formula

$$\sum_{j=1}^{N-1} \left(1 + 4\omega_j x_N + 12\omega_j^2 x_N^2 + O\left(|x|^3\right)\right) \xi_j^2 + \xi_N^2 \tag{10.51}$$
$$+ \sum_{i,j=1}^{N-1} \left(12 x_N \left(\sum_{k=1}^{N-1} \omega_{ijk} x_k\right) - 4\omega_i \omega_j x_i x_j + O\left(|x|^3\right)\right) \xi_i \xi_j$$

$$-2\sqrt{-1}\sum_{j=1}^{N-1}\left(-2\left(\sum_{i=1}^{N-1}\omega_i\omega_j x_j\right)+3x_N\left(\sum_{i=1}^{N-1}\omega_{iij}\right)+O\left(|x|^2\right)\right)\xi_j$$
$$+2\sqrt{-1}\left(\sum_{j=1}^{N-1}\omega_j+2x_N\sum_{j=1}^{N-1}\omega_j^2+3\sum_{i,j=1}^{N-1}\omega_{iij}x_j+O\left(|x|^2\right)\right)\xi_N.$$

We remark that the principal symbol $A_2(x,\xi',\xi_N)$ of $-\Delta$ is a polynomial of ξ_N of degree 2. We denote the roots of $A_2(x',0,\xi',\xi_N)$ with positive imaginary part and negative imaginary part by $\tau^+(x',0,\xi')$ and $\tau^-(x',0,\xi')$, respectively. Then, in view of (10.47) it is easy to verify that the roots $\tau^\pm(x',0,\xi')$ have the following Taylor expansions (10.52):

$$\tau^\pm(x',0,\xi')=\pm\sqrt{-1}\left(|\xi'|-\frac{2\sum_{i,j=1}^{N-1}\omega_i\omega_j x_i x_j\xi_i\xi_j}{|\xi'|}+O\left(|x'|^3\right)\right), \tag{10.52}$$

where

$$\xi'=(\xi_1,\xi_2,\ldots,\xi_{N-1}),\quad |\xi'|=\sqrt{\xi_1^2+\xi_2^2+\ldots+\xi_{N-1}^2}.$$

Moreover, we can write down the symbol

$$f(x,\xi)=f_{-2}(x,\xi)+f_{-3}(x,\xi)+\ldots$$

of the fundamental solution $\mathcal{F}$ of $-\Delta$ as follows:

$$\begin{aligned} f_{-2}(x,\xi) &= \frac{1}{|\xi|^2}-\frac{4x_N\left(\sum_{j=1}^{N-1}\omega_j\xi_j^2\right)}{|\xi|^4} \qquad (10.53a)\\ &+\frac{16x_N^2\left(\sum_{j=1}^{N-1}\omega_j\xi_j^2\right)^2}{|\xi|^6}-\frac{12x_N^2\left(\sum_{j=1}^{N-1}\omega_j^2\xi_j^2\right)}{|\xi|^4}\\ &-\frac{\sum_{i,j=1}^{N-1}\left(12x_N\left(\sum_{k=1}^{N-1}\omega_{ijk}x_k\right)-4\omega_i\omega_j x_i x_j\right)\xi_i\xi_j}{|\xi|^4}\\ &+O\left(|x|^3\right)|\xi|^{-2},\\ f_{-3}(x,\xi) &= -2\sqrt{-1}\,\xi_N\left(\frac{4\sum_{j=1}^{N-1}\omega_j\xi_j^2}{|\xi|^6}+\frac{\sum_{j=1}^{N-1}\omega_j}{|\xi|^4}\right) \qquad (10.53b)\\ &+O(|x|)\,|\xi|^{-3}, \end{aligned}$$

where

$$\xi = (\xi', \xi_N) = (\xi_1, \xi_2, \ldots, \xi_{N-1}, \xi_N), \quad |\xi| = \sqrt{|\xi'|^2 + \xi_N^2}.$$

We calculate the symbol of the Poisson operator $\mathcal{P}$. To do this, we denote by T^+ a first-order pseudo-differential operator on ∂D having the symbol $\tau^+(x', 0, \xi')$, and consider a mapping $Q : C_0^\infty(\partial D) \to C^\infty(\overline{D})$ defined by the formula

$$Q : \varphi \longmapsto \mathcal{F}\left(-\sqrt{-1}\, \frac{\partial \delta_{\partial D}}{\partial \mathbf{n}} \otimes \varphi - \delta_{\partial D} \otimes T^+\varphi\right).$$

Here $\delta_{\partial D} \otimes \psi$ is a distribution on $\mathbf{R}^N$ defined by the formula

$$\langle \delta_{\partial D} \otimes \psi, \phi \rangle = \int_{\partial D} \phi|_{\partial D} \cdot \psi \, d\gamma, \quad \phi \in C_0^\infty\left(\mathbf{R}^N\right),$$

where $d\gamma$ is the hypersurface element of ∂D. It should be emphasized that the mapping Q is a pseudo-Poisson operator in the sense of Boutet de Monvel [Bo], and further that its symbol has an asymptotic expansion with respect to homogeneous degree of $\left(x_N^{-1}, \xi'\right)$. Moreover, by using (10.52) and (10.53) we can calculate the symbol of Q as follows:

$$\begin{aligned}
&\sqrt{-1}\, e^{\sqrt{-1} x_N \tau^+(x',0,\xi')} \\
&+ \frac{1}{2\pi} \int_{-\infty}^{\infty} x_N \frac{\partial f_{-2}}{\partial x_N}(x', 0, \xi', \xi_N) \left(\xi_N - \tau^-(x', 0, \xi')\right) e^{\sqrt{-1} x_N \xi_N} \, d\xi_N \\
&- \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{j=1}^{N-1} \frac{\partial f_{-2}}{\partial \xi_j}(x', 0, \xi', \xi_N) \left(-\sqrt{-1}\, \frac{\partial \tau^-}{\partial x_j}(x', 0, \xi')\right) e^{\sqrt{-1} x_N \xi_N} \, d\xi_N \\
&+ \frac{1}{2\pi} \int_{-\infty}^{\infty} f_{-3}(x', 0, \xi', \xi_N) \left(\xi_N - \tau^-(x', 0, \xi')\right) e^{\sqrt{-1} x_N \xi_N} \, d\xi_N \\
&+ O\left(\left(x_N \left|\xi'\right|^{-1}\right)^2\right) \\
= {} & \sqrt{-1}\, e^{-x_N \left(|\xi'| - 2|\xi'|^{-1} \left(\sum_{i,j=1}^{N-1} \omega_i \omega_j x_i x_j \xi_i \xi_j\right) + O\left(|x'|^3\right)\right)} \\
&- \sqrt{-1}\, x_N \left(\sum_{j=1}^{N-1} \omega_j \xi_j^2 + 3 \sum_{i,j,k=1}^{N-1} \omega_{ijk} x_k \xi_i \xi_j + O\left(\left|x'\right|^2\right)\right) \\
&\times \left(2 x_N \left|\xi'\right|^{-1} + \left|\xi'\right|^{-2}\right) e^{-x_N |\xi'|} \\
&+ \left(\sum_{j,k=1}^{N-1} \omega_j \omega_k \xi_k x_k \xi_j^2 \left|\xi'\right|^{-1}\right) \left(2 x_N \left|\xi'\right|^{-2} + 2 \left|\xi'\right|^{-3} + O\left(\left|x'\right|\right)\right) e^{-x_N |\xi'|}
\end{aligned}$$

$$+\sqrt{-1}\left(\left(\sum_{j=1}^{N-1}\omega_j\xi_j^2\right)\left(x_N^2|\xi'|^{-1}-\frac{1}{2|\xi'|^3}\right)\right.$$
$$\left.+\sqrt{-1}\left(\sum_{i=1}^{N-1}\omega_i\right)\left(x_N-\frac{1}{2|\xi'|}\right)\right)e^{-x_N|\xi'|}+O\left(\left(x_N\left|\xi'\right|^{-1}\right)^2\right).$$

This proves that the mapping

$$K:\varphi\longmapsto Q\varphi|_{\partial D}$$

is an *elliptic* pseudo-differential operator of order zero, and its complete symbol $k(x',\xi')$ is given by

$$k(x',\xi')=\sqrt{-1}\left(1-\frac{\sum_{j=1}^{N-1}\omega_j\xi_j^2}{2\,|\xi'|^3}-\frac{\sum_{j=1}^{N-1}\omega_j}{2\,|\xi'|}+\dots\right).$$

If L is a parametrix for K, then it is known that the operator

$$QL\varphi=\mathcal{F}\left(-\sqrt{-1}\,\frac{\partial\delta_{\partial D}}{\partial\mathbf{n}}\otimes L\varphi-\delta_{\partial D}\otimes T^+L\varphi\right)$$

coincides with the harmonic operator H modulo smooth operators. The symbol of the harmonic operator H is given by

$$\begin{aligned}&e^{-x_N\left(|\xi'|-2|\xi'|^{-1}\left(\sum_{j=1}^{N-1}\omega_j x_j\xi_j\right)^2+O\left(|x'|^3\right)\right)}\\&-x_N\left(\left(\sum_{j=1}^{N-1}\omega_j\xi_j^2\right)\left|\xi'\right|^{-2}+O\left(\left|x'\right|\right)\right)e^{-x_N|\xi'|}\\&+x_N\left(\sum_{j=1}^{N-1}\omega_j+O\left(\left|x'\right|\right)\right)e^{-x_N|\xi'|}+O\left(\left(x_N\left|\xi'\right|^{-1}\right)^2\right).\end{aligned}\tag{10.54}$$

Therefore, we find that the mapping

$$\Pi:\varphi\longmapsto\left.\frac{\partial(\mathcal{P}\varphi)}{\partial\mathbf{n}}\right|_{\partial D}$$

is a first-order classical pseudo-differential operator on the boundary ∂D, and further that its principal symbol $p_1(x',\xi')$ is given by

$$p_1(x', \xi') = -\left|\xi'\right| + \frac{2\left(\sum_{j=1}^{N-1} \omega_j x_j \xi_j\right)^2}{|\xi'|} + O\left(\left|x'\right|^3\right),$$

and the real part of the second symbol $p_0(x', \xi')$ is given by

$$\text{Re}\, p_0(x', \xi') = -\frac{\sum_{j=1}^{N-1} \omega_j \xi_j^2}{|\xi'|^2} + \sum_{j=1}^{N-1} \omega_j + O\left(\left|x'\right|\right).$$

Moreover, we can write down the complete symbol $p(x', \xi')$ of Π in the following coordinate-free form (10.55) (see [Ta2]):

Lemma 10.33. *The symbol $p(x', \xi')$ of the pseudo-differential operator Π is given by*

$$p(x', \xi') = -|\xi'| - \frac{1}{2}\left(\frac{\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1)M(x')\right) + \sqrt{-1}\,\frac{1}{2}\,\text{div}\,\delta_{(\xi')}(x') + \textit{terms of order} \leq -1. \tag{10.55}$$

Here:

(a) $|\xi'|$ is the length of ξ' with respect to the Riemannian metric $(g_{ij}(x'))$ of ∂D induced by the natural metric of $\mathbf{R}^N$.
(b) $M(x')$ is the mean curvature of ∂D at x'.
(c) $\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})$ is the second fundamental form of ∂D at x', while $\widehat{\xi'} \in T_{x'}(\partial D)$ is the tangent vector corresponding to the cotangent vector $\xi' \in T^*_{x'}(\partial D)$ by the duality between $T_{x'}(\partial D)$ and $T^*_{x'}(\partial D)$ with respect to the Riemannian metric $(g_{ij}(x'))$ of ∂D.
(d) $\text{div}\,\delta_{(\xi')}$ is the divergence of a real smooth vector field $\delta_{(\xi')}$ on ∂D defined (in local coordinates) by the formula

$$\delta_{(\xi')} = \sum_{j=1}^{N-1} \frac{\partial|\xi'|}{\partial \xi_j} \frac{\partial}{\partial x_j} \quad \textit{for } \xi' \neq 0.$$

Proof. First, we remark that the symbol $p(x', \xi')$ of the operator Π is given by

$$p(x', \xi') = -\left|\xi'\right| - \frac{1}{2}\left(\frac{\omega_x(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1)M(x')\right) + \sqrt{-1}\, r_0(x', \xi') + \text{terms of order} \leq -1, \tag{10.56}$$

where $r_0(x', \xi')$ is a real term of order zero.

In order to calculate the term $r_0(x', \xi')$, let (U, χ) be a local chart of ∂D. Then we have, for all $\varphi, \psi \in C_0^\infty(U)$,

$$\begin{aligned}
&(\varphi, \Pi\psi)_{L^2(\partial D)} \\
&= (\Pi\varphi, \psi)_{L^2(\partial D)} \\
&= \int_{\mathbf{R}^{N-1}} \Pi\varphi\left(\chi^{-1}(x')\right) \cdot \overline{\psi\left(\chi^{-1}(x')\right)} \rho(x')\, dx \\
&= \int_{\mathbf{R}^{N-1}} \varphi\left(\chi^{-1}(x')\right) \cdot \overline{(\Pi_0)^* \left(\psi\left(\chi^{-1}(x')\right) \rho(x')\right)}\, dx \\
&= \int_{\mathbf{R}^{N-1}} \varphi\left(\chi^{-1}(x')\right) \cdot \overline{\left(\frac{1}{\rho(x')} (\Pi_0)^* \rho(x')\right) \psi\left(\chi^{-1}(x')\right)} \rho(x')\, dx,
\end{aligned}$$

where

$$\rho(x') = \sqrt{\det\left(g_{ij}(x')\right)},$$

and $(\Pi_0)^*$ is the adjoint of Π in the space $L^2(\mathbf{R}^{N-1})$. By using Theorems 7.19 and 7.20, we obtain from (10.56) that the symbol of the pseudo-differential operator

$$\Pi^* = \frac{1}{\rho(x')} (\Pi_0)^* \rho(x')$$

is given by

$$\begin{aligned}
&-|\xi'| - \frac{1}{2}\left(\frac{\omega_x(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2} - (N-1)M(x')\right) \qquad (10.57)\\
&-\sqrt{-1}\left(r_0(x', \xi') - \sum_{j=1}^{N-1} \frac{\partial^2 |\xi'|}{\partial x_j \partial \xi_j} - \frac{1}{\rho(x')} \sum_{j=1}^{N-1} \frac{\partial |\xi'|}{\partial \xi_j} \frac{\partial \rho(x')}{\partial x_j}\right) \\
&+ \text{terms of order} \leq -1.
\end{aligned}$$

Since $\Pi = \Pi^*$, it follows from (10.56) and (10.57) that

$$\begin{aligned}
&\sqrt{-1}\, r_0(x', \xi') \\
&= -\sqrt{-1}\left(r_0(x', \xi') - \sum_{j=1}^{N-1} \frac{\partial^2 |\xi'|}{\partial x_j \partial \xi_j} - \frac{1}{\rho(x')} \sum_{j=1}^{N-1} \frac{\partial |\xi'|}{\partial \xi_j} \frac{\partial \rho(x')}{\partial x_j}\right),
\end{aligned}$$

so that

$$r_0(x', \xi') = \frac{1}{2}\left(\frac{1}{\rho(x')}\sum_{j=1}^{N-1} \frac{\partial\,|\xi'|}{\partial \xi_j}\frac{\partial \rho(x')}{\partial x_j} + \sum_{j=1}^{N-1} \frac{\partial^2\,|\xi'|}{\partial x_j\,\partial \xi_j}\right). \tag{10.58}$$

However, we recall that the divergence div v of a real smooth vector field

$$v = \sum_{j=1}^{N-1} v^j \frac{\partial}{\partial x_j}$$

is given (in local coordinates) by

$$\begin{aligned}\operatorname{div} v &= \frac{1}{\rho(x')}\sum_{j=1}^{N-1} \frac{\partial}{\partial x_j}\left(\rho(x')\, v^j\right) \\ &= \sum_{j=1}^{N-1} \frac{\partial v^j}{\partial x_j} + \frac{1}{\rho(x')}\sum_{j=1}^{N-1} \frac{\partial \rho(x')}{\partial x_j}\, v^j.\end{aligned}$$

Hence we can rewrite (10.58) as follows:

$$r_0(x', \xi') = \frac{1}{2}\operatorname{div} \delta_{(\xi')}(x'). \tag{10.59}$$

Therefore, the desired formula (10.55) follows by combining (10.56) and (10.59).
The proof of Lemma 10.33 is complete.

It should be emphasized that the distribution kernel $\ell(x', y')$ of Π is given by

$$\ell(x', y') = \frac{\Gamma(N/2)}{\pi^{N/2}}\frac{1}{|x' - y'|^N} + \dots,$$

where $|x' - y'|$ denotes the geodesic distance between x' and y' with respect to the Riemannian metric $(g_{ij}(x'))$ of the boundary ∂D.

Now we are in a position to calculate the symbol of the pseudo-differential operator Π_α in the special case where $P = \Delta$. To do this, we remark that the operator Π_α coincides with the operator Π when $\alpha = 0$. Therefore, by replacing the operator Δ by $\Delta - \alpha$ we can write down the complete symbol $p(x', \xi'; \alpha)$ of Π_α for $\alpha \geq 0$ as follows (cf. formula (10.55)):

$$\begin{aligned}p(x', \xi'; \alpha) &= -\sqrt{|\xi'|^2 + \alpha} - \frac{1}{2}\left(\frac{\omega_{x'}(\widehat{\xi'}, \widehat{\xi'})}{|\xi'|^2 + \alpha} - (N-1)M(x')\right) \\ &\quad + \sqrt{-1}\,\frac{1}{2}\operatorname{div} \delta_{(\xi', \alpha)}(x') + \text{ terms of order} \leq -1.\end{aligned}$$

Here div $\delta_{(\xi',\alpha)}$ is the *divergence* of a real smooth vector field $\delta_{(\xi',\alpha)}$ on ∂D defined (in local coordinates) by the formula

$$\delta_{(\xi',\alpha)} = \sum_{j=1}^{N-1} \frac{\partial \sqrt{|\xi'|^2+\alpha}}{\partial \xi_j} \frac{\partial}{\partial x_j} \quad \text{for } \xi' \neq 0.$$

Hence we find that the complete symbol $p(x',\xi';\alpha)$ is given by

$$\begin{aligned} p(x',\xi';\alpha) &= -|\xi'| - \frac{1}{2}\left(\frac{\omega_{x'}(\widehat{\xi'},\widehat{(\xi')})}{|\xi'|^2} - (N-1)M(x')\right) \\ &\quad + \sqrt{-1}\,\frac{1}{2}\,\mathrm{div}\,\delta_{(\xi')}(x') \\ &\quad + \text{terms of order} \leq -1/2 \text{ depending on } \alpha. \end{aligned} \tag{10.60}$$

Indeed, it suffices to note the following asymptotic expansions for any parameter $\alpha \geq 0$:

$$\begin{aligned} \sqrt{|\xi'|^2+\alpha} &= |\xi'| - \frac{\alpha}{2}|\xi'|^{-1} - \frac{\alpha^2}{8}|\xi'|^{-3} + \dots \\ &= |\xi'| + \text{ terms of order} \leq -1 \text{ depending on } \alpha, \\ \frac{1}{|\xi'|^2+\alpha} &= \frac{1}{|\xi'|^2} + \alpha\,|\xi'|^{-4} + \dots \\ &= \frac{1}{|\xi'|^2} + \text{ terms of order} \leq -4 \text{ depending on } \alpha, \\ \frac{\partial\sqrt{|\xi'|^2+\alpha}}{\partial \xi_j} &= \frac{\xi_j}{\sqrt{|\xi'|}} + \alpha\,\frac{\xi_j}{|\xi'|^{3/2}} + \dots \\ &= \frac{\partial|\xi'|}{\partial \xi_j} + \text{ terms of order} \leq -1/2 \text{ depending on } \alpha. \end{aligned}$$

10.8 Notes and Comments

The results discussed in this chapter are adapted from Taira [Ta4, Ta5, Ta6, Ta7, Ta8, Ta9, Ta10]. Agmon [Ag] and Lions–Magenes [LM] are the classics for elliptic boundary value problems by variational methods.

Section 10.2: Theorem 10.2 is a generalization of Taira [Ta6, Theorem 1]. It should be noted that Theorem 1.2 was proved before by Taira [Ta10, Theorem 10.1.3] under some additional conditions, and also by Cancelier [Cn, Théorème 3.2]. On the other hand, Takanobu and Watanabe [TW] proved a probabilistic version of Theorem 1.2 in the case where the domain D is the half-space $\mathbf{R}^N_+$ (see [TW, Corollary]). It seems that our method of construction of Feller semigroups is, in spirit, not far removed from the probabilistic method used by Watanabe [Wb] (see Remark 10.3).

Section 10.3: Theorem 10.4 is a pseudo-differential operator version of Bony–Courrège–Priouret [BCP, Théorème XV]. Since the pseudo-differential operator $T \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$ has the transmission property with respect to ∂D, we can make use of the zero-extension of functions in the interior D outside of the closure $\overline{D} = D \cup \partial D$. This extension has a probabilistic interpretation that any Markovian particle does not move by jumps from D into the outside $\mathbf{R}^N \setminus \overline{D}$ *across* the boundary ∂D. Therefore, we can prove that every solution w of the Dirichlet problem

$$\begin{cases} (\alpha - W_D)w = 0 & \text{in } D, \\ w = \psi & \text{on } \partial D \end{cases} \tag{10.11}$$

is expressed in the form

$$w = H_\alpha \psi.$$

This formula is a pseudo-differential operator version of the de la Vallée-Poussin formula obtained by the balayage method. If a Markovian particle may move by jumps from the interior D into the outside $\mathbf{R}^N \setminus \overline{D}$ across the boundary ∂D, then we can apply *balayage potential theory* to solve Dirichlet problem (10.11) in an extended context (see [BH, Ha]).

Section 10.4: Theorem 10.19 is a pseudo-differential operator version of Bony–Courrège–Priouret [BCP, Théorème XX].

Section 10.5: Theorem 10.21 is a revised version of Taira [Ta6, Theorem 1].

Section 10.6: Theorem 10.23 is inspired by Oleĭnik–Radkevič [OR, Chapter I] and Cancelier [Cn, Théorème 4.5]. This Sect. 10.6 is a refinement of Appendix B of the first edition of the present monograph.

Section 10.7: The results discussed here are adapted from Fujiwara–Uchiyama [FU] and Taira [Ta2].

Chapter 11
Proof of Theorem 1.3

In this chapter we consider the non-transversal case, and prove Theorem 1.3 (Theorem 11.1). The idea of the proof can be traced back to the work of Taira [Ta9] and [Ta6].

We assume the following condition (A):

(A) There exists a second-order Ventcel' boundary condition L_ν such that

$$Lu = m(x')\, L_\nu u + \gamma(x')u \quad \text{on } \partial D,$$

where

(3′) $m(x') \in C^\infty(\partial D)$ and $m(x') \geq 0$ on ∂D,

and L_ν is given, in terms of local coordinates $(x_1, x_2, \ldots, x_{N-1})$, by the formula

$$\begin{aligned}
& L_\nu u(x') \\
&= \overline{Q}u(x') + \overline{\mu}(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \overline{\delta}(x')W_D u(x') + \overline{\Gamma}u(x') \\
&:= \sum_{i,j=1}^{N-1} \overline{\alpha}^{ij}(x')\frac{\partial^2 u}{\partial x_i \partial x_j}(x') + \sum_{i=1}^{N-1} \overline{\beta}^i(x')\frac{\partial u}{\partial x_i}(x') + \overline{\gamma}(x') \\
&\quad + \overline{\mu}(x')\frac{\partial u}{\partial \mathbf{n}}(x') - \overline{\delta}(x')\, W_D u(x') \\
&\quad + \sum_{j=1}^{N-1} \overline{\eta}_\tau^j(x')\frac{\partial u}{\partial x_j}(x') + \overline{\eta}_\tau(x')u(x')
\end{aligned}$$

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*, Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_11,

$$+\int_{\partial D} \overline{r}(x', y') \left[u(y') - \overline{\tau}(x', y') \left(u(x') + \sum_{j=1}^{N-1} \left(y_j - x_j\right) \frac{\partial u}{\partial x_j}(x') \right) \right] dy'$$
$$+\int_{D} \overline{t}(x', y) \left[u(y) - u(x') \right] dy, \quad x' \in \partial D,$$

and satisfies the *transversality* condition

$$\int_D \overline{t}(x', y)\, dy = +\infty \quad \text{if } \overline{\mu}(x') = \overline{\delta}(x') = 0. \tag{1.10}$$

Moreover, we recall the following four conditions (7)–(10):

(7) The integral kernel $\overline{r}(x', y')$ is the distribution kernel of a classical pseudo-differential operator $\overline{R} \in L^{2-\kappa_1}_{1,0}(\partial D)$, $\kappa_1 > 0$, and $\overline{r}(x', y') \geq 0$ off the diagonal $\Delta_{\partial D} = \{(x', x') : x' \in \partial D\}$ in $\partial D \times \partial D$. The density dy' is a strictly positive density on ∂D.
(8) The integral kernel $\overline{t}(x, y)$ is the distribution kernel of a properly supported, classical pseudo-differential operator $\overline{T} \in L^{1-\kappa_2}_{1,0}(\mathbf{R}^N)$, $\kappa_2 > 0$, which has the *transmission property* with respect to the boundary ∂D, and $\overline{t}(x, y) \geq 0$ off the diagonal $\Delta_{\mathbf{R}^N}$.
(9) The function $\overline{\tau}(x, y)$ is a local unity function on $\overline{D}$; more precisely, $\overline{\tau}(x, y)$ is a smooth function on $\overline{D} \times \overline{D}$, with compact support in a neighborhood of the diagonal $\Delta_{\partial D}$, such that, at each point x' of ∂D, $\overline{\tau}(x', y) = 1$ for y in a neighborhood of x' in $\overline{D}$.
(10) The operator $\overline{\Gamma}_r$ is a boundary condition of order $2 - \kappa_1$, and satisfies the condition

$$\overline{Q}1(x') + \overline{\Gamma}_r 1(x') = \overline{\gamma}(x') + \int_{\partial D} \overline{r}(x', y') \left[1 - \overline{\tau}(x', y')\right] dy' \leq 0 \quad \text{on } \partial D. \tag{1.9}$$

We let

$$M = \{x' \in \partial D : \mu(x') = \delta(x') = 0, \int_D t(x', y)\, dy < \infty\}.$$

Then, by condition (1.10) it follows that

$$M = \{x' \in \partial D : m(x') = 0\},$$

since we have

$$\mu(x') = m(x')\, \overline{\mu}(x'), \quad \delta(x') = m(x')\, \overline{\delta}(x'), \quad t(x', y) = m(x')\, \overline{t}(x', y).$$

Hence we find that the boundary condition L is *not* transversal on ∂D.

Furthermore, we assume the following condition (H):

(H) $m(x') + |\gamma(x')| > 0$ on ∂D.

The intuitive meaning of conditions (A) and (H) is that a Markovian particle does not stay on ∂D for any period of time until it "dies" when it reaches the set M where the particle is definitely absorbed.

In Sect. 11.1 we consider a one-point compactification $K_\partial = K \cup \{\partial\}$ of the space $K = \overline{D} \setminus M$, where

$$M = \{x' \in \partial D : m(x') = 0\},$$

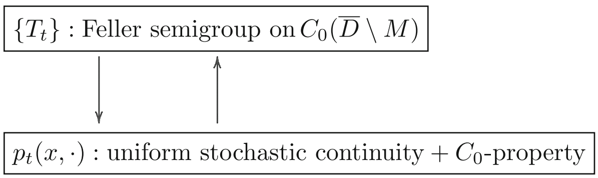
and introduce a closed subspace of $C(K_\partial)$ by

$$C_0(K) = \{u \in C(K_\partial) : u(\partial) = 0\}.$$

Then we have the isomorphism

$$C_0(K) \cong C_0(\overline{D} \setminus M) = \left\{u \in C(\overline{D}) : u = 0 \text{ on } M\right\}.$$

In Sect. 11.2 we apply part (ii) of Theorem 9.35 to the operator $\mathfrak{W}$ defined by formula (1.10). Our functional analytic approach may be visualized as follows (see Sect. 9.1):



11.1 The Space $C_0(\overline{D} \setminus M)$

First, we consider the one-point compactification $K_\partial = K \cup \{\partial\}$ of the space $K = \overline{D} \setminus M$, where

$$M = \{x' \in \partial D : m(x') = 0\}.$$

We say that two points x and y of $\overline{D}$ are *equivalent* modulo M if $x = y$ or $x, y \in M$; we then write $x \sim y$. We denote by $\overline{D}/M$ the totality of equivalence classes modulo M. On the set $\overline{D}/M$, we define the quotient topology induced by the projection $q : \overline{D} \to \overline{D}/M$. It is easy to see that the topological space $\overline{D}/M$

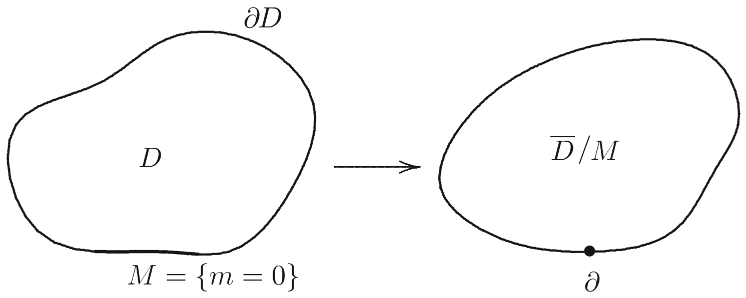


Fig. 11.1 The compactification $\overline{D}/M$ of $\overline{D} \setminus M$

is a *one-point compactification* of the space $\overline{D} \setminus M$ and that the *point at infinity* ∂ corresponds to the set M (see Fig. 11.1):

$$K_\partial = \overline{D}/M, \quad \partial = M.$$

Furthermore, we have the following isomorphism (11.1):

$$C(K_\partial) \cong \left\{u \in C(\overline{D}) : u \text{ is constant on } M\right\}. \tag{11.1}$$

Now we introduce a closed subspace of $C(K_\partial)$ as in Sect. 9.1:

$$C_0(K) = \{u \in C(K_\partial) : u(\partial) = 0\}.$$

Then we have, by assertion (11.1),

$$C_0(K) \cong C_0(\overline{D} \setminus M) = \left\{u \in C(\overline{D}) : \ u = 0 \text{ on } M\right\}. \tag{11.2}$$

11.2 End of Proof of Theorem 1.3

We shall apply part (ii) of Theorem 9.35 to the operator $\mathfrak{W}$ defined by formula (1.10).

First, we show that if condition (A) is satisfied, then the operator LH_α is *bijective* in the framework of Hölder spaces. This is proved by applying Theorem 10.23 just as in the proof of Theorem 1.2. Therefore, we find that a unique solution u of the boundary value problem

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ Lu = m(x')\, L_\nu u + \gamma(x')u = 0 & \text{on } \partial D \end{cases} \tag{$**$}$$

can be expressed as follows:

$$u = G_\alpha f = G_\alpha^\nu f - H_\alpha \left(\overline{LH_\alpha}^{-1} \left(LG_\alpha^\nu f \right) \right).$$

This formula allows us to verify all the conditions of the generation theorems of Feller semigroups discussed in Sect. 9.3.

To do this, we simplify the boundary condition

$$Lu = 0 \quad \text{on } \partial D.$$

If conditions (A) and (H) are satisfied, then we may assume that the boundary condition L is of the form

$$Lu = m(x')\, L_\nu u + (m(x') - 1)u, \tag{11.3}$$

with

$$0 \leq m(x') \leq 1 \quad \text{on } \partial D.$$

Indeed, it suffices to note that the boundary condition

$$Lu = m(x')\, L_\nu u + \gamma(x')\, (u|_{\partial D}) = 0 \quad \text{on } \partial D$$

is equivalent to the condition

$$\left(\frac{m(x')}{m(x') - \gamma(x')} \right) L_\nu u + \left(\frac{\gamma(x')}{m(x') - \gamma(x')} \right) (u|_{\partial D}) = 0 \quad \text{on } \partial D.$$

Furthermore, we note that

$$\overline{LG_\alpha^0} f = m(x')\, \overline{L_\nu G_\alpha^0} f,$$

and that

$$\overline{LH_\alpha} \varphi = m(x')\, \overline{L_\nu H_\alpha} \varphi + (m(x') - 1)\varphi.$$

Hence, in view of definition (10.21) it follows that

$$Lu = m(x')\, L_\nu u + (m(x') - 1)\, (u|_{\partial D}), \quad u \in D(L). \tag{11.3'}$$

Therefore, the next generation theorem for Feller semigroups implies Theorem 1.3:

$$\boxed{\text{Theorem 11.1}} \Longrightarrow \boxed{\text{Theorem 1.3}}$$

Theorem 11.1. *We define a linear operator*

$$\mathfrak{W} : C_0(\overline{D} \setminus M) \longrightarrow C_0(\overline{D} \setminus M)$$

as follows (cf. formula (10.23)*).*

(a) The domain $D(\mathfrak{W})$ *of* $\mathfrak{W}$ *is the set*

$$D(\mathfrak{W}) = \{u \in C_0(\overline{D} \setminus M) : \overline{W}u \in C_0(\overline{D} \setminus M), \quad Lu = m(x')\, L_\nu u + (m(x') - 1)\, (u|_{\partial D}) = 0\}. \tag{11.4}$$

(b) $\mathfrak{W}u = \overline{W}u$ *for every* $u \in D(\mathfrak{W})$.

Assume that the following condition (A′) is satisfied:

(A′) $0 \leq m(x') \leq 1$ *on* ∂D.

Then the operator $\mathfrak{W}$ *is the infinitesimal generator of some Feller semigroup* $\{T_t\}_{t \geq 0}$ *on* $\overline{D} \setminus M$*, and the Green operator* $G_\alpha = (\alpha I - \mathfrak{W})^{-1}$*,* $\alpha > 0$*, is given by*

$$G_\alpha f = G_\alpha^\nu f - H_\alpha \left(\overline{LH_\alpha}^{-1} \left(LG_\alpha^\nu f \right) \right), \quad f \in C_0(\overline{D} \setminus M). \tag{11.5}$$

Here G_α^ν *is the Green operator for the boundary condition* L_ν *given by formula* (10.22)*:*

$$G_\alpha^\nu f = G_\alpha^0 f - H_\alpha \left(\overline{L_\nu H_\alpha}^{-1} \left(\overline{L_\nu G_\alpha^0 f} \right) \right), \quad f \in C(\overline{D}).$$

Proof. We apply part (ii) of Theorem 9.35 to the operator $\mathfrak{W}$ defined by formula (11.4), just as in the proof of Theorem 1.2. The proof is divided into several steps.

Step 1: First, we prove that

> For *all* $\alpha > 0$, the operator $\overline{LH_\alpha}$ generates a Feller semigroup
>
> on the boundary ∂D.

By virtue of the *transmission property* of $\overline{T} \in L_{1,0}^{2-\kappa_2}(\mathbf{R}^N)$, it follows (see Boutet de Monvel [Bo], Rempel–Schulze [RS, Chapter 3]) that the operator LH_α is the sum of a degenerate elliptic differential operator of second order and a *classical pseudo-differential operator* of order $2 - \min(\kappa_1, \kappa_2)$:

$$\begin{aligned} &LH_\alpha\varphi(x') \\ &= m(x')\, L_\nu H_\alpha\varphi(x') + (m(x') - 1)\varphi(x') \end{aligned}$$

$$= m(x')\Bigg(\sum_{i,j=1}^{N-1} \overline{\alpha}^{ij}(x')\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}(x') + \sum_{i=1}^{N-1}\overline{\beta}^i(x')\frac{\partial\varphi}{\partial x_i}(x') + \overline{\gamma}(x')\varphi(x')\Bigg)$$

$$+m(x')\,\overline{\mu}(x')\frac{\partial}{\partial\mathbf{n}}\,(H_\alpha\varphi)\,(x') - \alpha\, m(x')\,\overline{\delta}(x')\varphi(x') + (m(x')-1)\varphi(x')$$

$$+m(x')\Bigg(\sum_{j=1}^{N-1}\overline{\eta}_\tau^j(x')\frac{\partial u}{\partial x_j}(x') + \overline{\eta}_\tau(x')u(x')\Bigg)$$

$$+m(x')\int_{\partial D}\overline{r}(x',y')\big[\varphi(y') - \overline{\tau}(x',y')\Big(\varphi(x') + \sum_{j=1}^{N-1}(y_j - x_j)\Big)\frac{\partial\varphi}{\partial x_j}(x'))\big]dy'$$

$$+\int_D \overline{t}(x',y)\left[H_\alpha\varphi(y) - \varphi(x')\right]dy\Bigg).$$

Furthermore, it follows from an application of Hopf's boundary point lemma (see Theorem 8.15) that

$$H_\alpha 1(y) - 1 < 0 \quad \text{in } D,$$

$$\Pi_\alpha 1(x') = \frac{\partial}{\partial\mathbf{n}}(H_\alpha 1)(x') < 0 \quad \text{on } \partial D.$$

This implies that

$$LH_\alpha 1(x') = m(x')\,L_\nu H_\alpha 1(x') + (m(x')-1) < 0 \quad \text{on } \partial D.$$

Indeed, it suffices to note that

$$\begin{aligned} L_\nu H_\alpha 1(x') &= -\alpha\,\overline{\delta}(x') + \overline{\mu}(x')\,\Pi_\alpha 1(x') \\ &\quad + \left(\overline{\gamma}(x') + \overline{\eta}_\tau(x') + \int_{\partial D}\overline{r}(x',y')\left[1 - \overline{\tau}(x',y')\right]dy'\right) \\ &\quad + \int_D \overline{t}(x',y)\,[H_\alpha 1(y) - 1]\,dy \\ &\le -\alpha\,\overline{\delta}(x') + \overline{\mu}(x')\,\Pi_\alpha 1(x') + \int_D \overline{t}(x',y)\,[H_\alpha 1(y) - 1]\,dy \\ &< 0 \quad \text{on } \partial D, \end{aligned} \tag{11.6}$$

since we have, by condition (1.9),

$$\overline{\gamma}(x') + \overline{\eta}_\tau(x') + \int_{\partial D}\overline{r}(x',y')\left[1 - \overline{\tau}(x',y')\right]dy' \le 0 \quad \text{on } \partial D.$$

Therefore, by applying Theorem 10.23 to the operator LH_α we obtain that

$$\text{If } \lambda > 0 \text{ is sufficiently large, then the range } R(LH_\alpha - \lambda I) \text{ contains the space } C^{2+\theta}(\partial D). \tag{11.7}$$

This implies that the range $R(LH_\alpha - \lambda I)$ is a *dense* subset of $C(\partial D)$. Therefore, by applying part (ii) of Theorem 10.19 to the operator L we obtain that the operator $\overline{LH_\alpha}$ is the infinitesimal generator of some Feller semigroup on ∂D, for all $\alpha > 0$.

Step 2: Now we prove that

If condition (A′) is satisfied, then the equation (11.8)

$$\overline{LH_\alpha}\,\psi = \varphi$$

has a unique solution ψ in $D\left(\overline{LH_\alpha}\right)$ for any $\varphi \in C(\partial D)$; hence

the inverse $\overline{LH_\alpha}^{-1}$ of $\overline{LH_\alpha}$ can be defined on the whole space $C(\partial D)$.

Further, the operator $-\overline{LH_\alpha}^{-1}$ is non-negative and bounded on $C(\partial D)$.

Since we have, by inequality (11.6) and condition (A′),

$$LH_\alpha 1(x') = m(x') L_\nu H_\alpha 1(x') + (m(x') - 1) < 0 \quad \text{on } \partial D,$$

it follows that

$$k_\alpha = - \sup_{x' \in \partial D} LH_\alpha 1(x') > 0,$$

and further that the constants k_α are increasing in $\alpha > 0$:

$$\alpha \geq \beta > 0 \implies k_\alpha \geq k_\beta.$$

Indeed, it suffices to note that $H_\alpha 1$ converges to zero and $\Pi_\alpha 1$ diverges to $-\infty$ *monotonically* as $\alpha \to +\infty$, respectively. Moreover, by using Corollary 9.36 with $K := \partial D$, $A := \overline{LH_\alpha}$ and $c := k_\alpha$ we obtain that the operator $\overline{LH_\alpha} + k_\alpha I$ is the infinitesimal generator of some Feller semigroup on ∂D. Therefore, since $k_\alpha > 0$, it follows from an application of part (i) of Theorem 9.35 with $A := \overline{LH_\alpha} + k_\alpha I$ that the equation

$$-\overline{LH_\alpha}\,\psi = \left(k_\alpha I - (\overline{LH_\alpha} + k_\alpha I)\right)\psi = \varphi$$

has a unique solution $\psi \in D\left(\overline{LH_\alpha}\right)$ for any $\varphi \in C(\partial D)$, and further that the operator

$$-\overline{LH_\alpha}^{-1} = \left(k_\alpha I - (\overline{LH_\alpha} + k_\alpha I)\right)^{-1}$$

is non-negative and bounded on the space $C(\partial D)$ with norm

$$\left\| -\overline{LH_\alpha}^{-1} \right\| = \left\| \left(k_\alpha I - (\overline{LH_\alpha} + k_\alpha I)\right)^{-1} \right\| \leq \frac{1}{k_\alpha}. \tag{11.9}$$

Step 3: By assertion (11.8), we can define the operator G_α by formula (11.5) for all $\alpha > 0$. We prove that

$$G_\alpha = (\alpha I - \mathfrak{W})^{-1}, \quad \alpha > 0. \tag{11.10}$$

By virtue of Lemma 10.24, it follows that we have, for all $f \in C_0(\overline{D} \setminus M)$,

$$G_\alpha f \in D(\overline{W}),$$

and

$$\overline{W} G_\alpha f = \alpha G_\alpha f - f.$$

Furthermore, we have

$$LG_\alpha f = LG_\alpha^\nu f - \overline{LH_\alpha} \left(\overline{LH_\alpha}^{-1} \left(LG_\alpha^\nu f \right) \right) = 0 \quad \text{on } \partial D. \tag{11.11}$$

However, we recall the formula (11.3′)

$$Lu = m(x')\, L_\nu u + (m(x') - 1)\, (u|_{\partial D})\,, \quad u \in D(L).$$

Hence we find that (11.11) is equivalent to the following:

$$m(x')\, L_\nu(G_\alpha f) + (m(x') - 1)\, (G_\alpha f|_{\partial D}) = 0 \quad \text{on } \partial D. \tag{11.11′}$$

This implies that

$$G_\alpha f = 0 \quad \text{on } M = \{x' \in \partial D : m(x') = 0\},$$

so that

$$\overline{W} G_\alpha f = \alpha G_\alpha f - f = 0 \quad \text{on } M.$$

Summing up, we have proved that

$$G_\alpha f \in D(\mathfrak{W}) = \left\{ u \in C_0(\overline{D} \setminus M) : \overline{W} u \in C_0(\overline{D} \setminus M),\ Lu = 0 \right\},$$

and

$$(\alpha I - \mathfrak{W})G_\alpha f = f, \quad f \in C_0(\overline{D} \setminus M),$$

that is,

$$(\alpha I - \mathfrak{W})G_\alpha = I \quad \text{on } C_0(\overline{D} \setminus M).$$

Therefore, in order to prove (11.10) it suffices to show the injectivity of the operator $\alpha I - \mathfrak{W}$ for $\alpha > 0$.
Assume that

$$u \in D(\mathfrak{W}) \text{ and } (\alpha I - \mathfrak{W})u = 0.$$

Then, by Corollary 10.13 it follows that the function u can be written in the form

$$u = H_\alpha \left(u|_{\partial D}\right), \quad u|_{\partial D} \in \mathcal{D} = D\left(\overline{LH_\alpha}\right).$$

Thus we have

$$\overline{LH_\alpha}\left(u|_{\partial D}\right) = Lu = 0.$$

In view of assertion (11.8), this implies that

$$u|_{\partial D} = 0,$$

so that

$$u = H_\alpha \left(u|_{\partial D}\right) = 0 \quad \text{in } D.$$

Step 4: Now we prove the following three assertions (i)–(iii):

(i) The operator G_α is non-negative on the space $C_0(\overline{D} \setminus M)$:

$$f \in C_0(\overline{D} \setminus M),\ f \geq 0 \quad \text{on } \overline{D} \setminus M \implies G_\alpha f \geq 0 \quad \text{on } \overline{D} \setminus M.$$

(ii) The operator G_α is bounded on the space $C_0(\overline{D} \setminus M)$ with norm

$$\|G_\alpha\| \leq \frac{1}{\alpha}, \quad \alpha > 0.$$

(iii) The domain $D(\mathfrak{W})$ is dense in the space $C_0(\overline{D} \setminus M)$.

Step 4-1: First, we show the non-negativity of G_α on the space $C(\overline{D})$:

$$f \in C(\overline{D}),\ f \geq 0 \quad \text{on } \overline{D} \implies G_\alpha f \geq 0 \quad \text{on } \overline{D}.$$

Recall that the Dirichlet problem

$$\begin{cases} (\alpha - W_D)u = f & \text{in } D, \\ u = \varphi & \text{on } \partial D \end{cases} \tag{D$'$}$$

is uniquely solvable. Hence it follows that

$$G_\alpha^\nu f = H_\alpha \left(G_\alpha^\nu f|_{\partial D} \right) + G_\alpha^0 f \quad \text{on } \overline{D}. \tag{11.12}$$

Indeed, both sides have the same boundary values $G_\alpha^\nu f|_{\partial D}$ and satisfy the same equation: $(\alpha - W_D)u = f$ in D.
Thus, by applying the operator L to both sides of (11.12) we obtain that

$$LG_\alpha^\nu f = \overline{LH_\alpha} \left(G_\alpha^\nu f|_{\partial D} \right) + \overline{LG_\alpha^0} f.$$

Since the operators $-\overline{LH_\alpha}^{-1}$ and $\overline{LG_\alpha^0}$ are non-negative, it follows that

$$\begin{aligned} \left(-\overline{LH_\alpha}^{-1} \right) \left(LG_\alpha^\nu f \right) &= -G_\alpha^\nu f|_{\partial D} + \left(-\overline{LH_\alpha}^{-1} \right) \left(\overline{LG_\alpha^0} f \right) \\ &\geq -G_\alpha^\nu f|_{\partial D} \quad \text{on } \partial D. \end{aligned}$$

Therefore, by the non-negativity of H_α and G_α^0 we find that

$$\begin{aligned} G_\alpha f &= G_\alpha^\nu f + H_\alpha \left(-\overline{LH_\alpha}^{-1} \left(LG_\alpha^\nu f \right) \right) \geq G_\alpha^\nu f - H_\alpha \left(G_\alpha^\nu f|_{\partial D} \right) \\ &= G_\alpha^0 f \geq 0 \quad \text{on } \overline{D}. \end{aligned}$$

Step 4-2: Next we prove the boundedness of G_α on the space $C_0(\overline{D} \setminus M)$ with norm

$$\|G_\alpha\| \leq \frac{1}{\alpha}, \quad \alpha > 0. \tag{11.13}$$

To do this, it suffices to show that

$$f \in C_0(\overline{D} \backslash M),\ f \geq 0 \quad \text{on } \overline{D} \implies \alpha G_\alpha f \leq \max_{\overline{D}} f \quad \text{on } \overline{D}, \tag{11.13$'$}$$

since G_α is non-negative on the space $C(\overline{D})$.
We observe (cf. (11.3′)) that

$$LG_\alpha^\nu f = m(x')\, L_\nu G_\alpha^\nu f + (m(x') - 1) \left(G_\alpha^\nu f|_{\partial D} \right) = (m(x') - 1) \left(G_\alpha^\nu f|_{\partial D} \right),$$

so that

$$G_\alpha f = G_\alpha^\nu f - H_\alpha\left(\overline{LH_\alpha}^{-1}\left(LG_\alpha^\nu f\right)\right) \tag{11.5'}$$
$$= G_\alpha^\nu f + H_\alpha\left(-\overline{LH_\alpha}^{-1}\left((m(x')-1)G_\alpha^\nu f|_{\partial D}\right)\right).$$

Therefore, by the non-negativity of H_α and $-\overline{LH_\alpha}^{-1}$ it follows that

$$G_\alpha f = G_\alpha^\nu f + H_\alpha\left(-\overline{LH_\alpha}^{-1}\left((m(x')-1)G_\alpha^\nu f|_{\partial D}\right)\right) \leq G_\alpha^\nu f$$
$$\leq \frac{1}{\alpha}\max_{\overline{D}} f \quad \text{on } \overline{D},$$

since we have the inequalities

$$(m(x')-1)G_\alpha^\nu f|_{\partial D} \leq 0 \quad \text{on } \partial D,$$
$$\|G_\alpha^\nu\| \leq 1/\alpha.$$

This proves assertion (11.13′) and hence assertion (11.13).

Step 4-3: Finally, we prove the density of $D(\mathfrak{W})$ in the space $C_0(\overline{D}\setminus M)$. In view of (11.10), it suffices to show that

$$\lim_{\alpha\to+\infty}\|\alpha G_\alpha f - f\|_\infty = 0, \quad f \in C_0(\overline{D}\setminus M)\cap C^\infty(\overline{D}). \tag{11.14}$$

We recall (cf. (11.5′)) that

$$\alpha G_\alpha f - f = \alpha G_\alpha^\nu f - f - \alpha H_\alpha\left(\overline{LH_\alpha}^{-1}\left(LG_\alpha^\nu f\right)\right) \tag{11.15}$$
$$= \left(\alpha G_\alpha^\nu f - f\right) + H_\alpha\left(\overline{LH_\alpha}^{-1}\left(\alpha(1-m(x'))G_\alpha^\nu f|_{\partial D}\right)\right).$$

We estimate each term on the right-hand side of (11.15).

Step 4-3-1: First, by applying Theorem 1.2 to the boundary condition L_ν we find from assertion (10.28) that the first term on the right-hand side of (11.15) tends to zero:

$$\lim_{\alpha\to+\infty}\|\alpha G_\alpha^\nu f - f\|_\infty = 0. \tag{11.16}$$

Step 4-3-2: To estimate the second term on the right-hand side of (11.15), we note that

$$H_\alpha\left(\overline{LH_\alpha}^{-1}\left(\alpha(1-m(x'))G_\alpha^\nu f|_{\partial D}\right)\right)$$
$$= H_\alpha\left(\overline{LH_\alpha}^{-1}\left((1-m(x'))f|_{\partial D}\right)\right)$$
$$+H_\alpha\left(\overline{LH_\alpha}^{-1}\left((1-m(x'))(\alpha G_\alpha^\nu f - f)|_{\partial D}\right)\right).$$

However, we have, by inequality (11.9) and assertion (11.16),

$$\begin{aligned}
&\left\| H_\alpha \left(\overline{LH_\alpha}^{-1} \left((1 - m(x')) \left(\alpha G_\alpha^\nu f - f \right) |_{\partial D} \right) \right) \right\|_\infty \qquad (11.17)\\
&\leq \left\| -\overline{LH_\alpha}^{-1} \right\| \cdot \left\| (1 - m(x')) \left(\alpha G_\alpha^\nu f - f \right) |_{\partial D} \right\|_\infty \\
&\leq \frac{1}{k_\alpha} \left\| (1 - m(x')) \left(\alpha G_\alpha^\nu f - f \right) |_{\partial D} \right\|_\infty \\
&\leq \frac{1}{k_1} \| \alpha G_\alpha^\nu f - f \|_\infty \longrightarrow 0 \quad \text{as } \alpha \to +\infty.
\end{aligned}$$

Here we have used the fact that

$$k_1 = - \sup_{x' \in \partial D} LH_1 1(x') \leq k_\alpha = - \sup_{x' \in \partial D} LH_\alpha 1(x') \quad \text{for all } \alpha \geq 1.$$

Thus we are reduced to the study of the term

$$H_\alpha \left(\overline{LH_\alpha}^{-1} \left((1 - m(x')) f |_{\partial D} \right) \right).$$

Now, for any given $\varepsilon > 0$, we can find a function $h \in C^\infty(\partial D)$ such that

$$\begin{cases} h = 0 \quad \text{near } M = \{x' \in \partial D : m(x') = 0\}, \\ \|(1 - m(x')) f |_{\partial D} - h\|_\infty < \varepsilon. \end{cases}$$

Then we have the inequality

$$\begin{aligned}
&\left\| H_\alpha \left(\overline{LH_\alpha}^{-1} \left((1 - m(x')) f |_{\partial D} \right) \right) - H_\alpha \left(\overline{LH_\alpha}^{-1} h \right) \right\|_\infty \qquad (11.18)\\
&\leq \left\| -\overline{LH_\alpha}^{-1} \right\| \cdot \left\| (1 - m(x')) f |_{\partial D} - h \right\|_\infty \leq \frac{\varepsilon}{k_\alpha} \\
&\leq \frac{\varepsilon}{k_1} \quad \text{for all } \alpha \geq 1.
\end{aligned}$$

Furthermore, we can find a function $\theta \in C_0^\infty(\partial D)$ such that

$$\begin{cases} \theta = 1 & \text{near } M, \\ (1 - \theta) h = h & \text{on } \partial D. \end{cases}$$

Then we have the inequality

$$\begin{aligned}h(x') &= \left(1-\theta(x')\right)h(x')\\ &= \left(-LH_\alpha 1(x')\right)\left(\frac{1-\theta(x')}{-LH_\alpha 1(x')}\right)h(x')\\ &\le \left[\sup_{x'\in\partial D}\left(\frac{1-\theta(x')}{-LH_\alpha 1(x')}\right)\right]\|h\|_\infty\left(-LH_\alpha 1(x')\right).\end{aligned}$$

Since the operator $-\overline{LH_\alpha}^{-1}$ is non-negative on the space $C(\partial D)$, it follows that

$$-\overline{LH_\alpha}^{-1}h \le \sup_{x'\in\partial D}\left(\frac{1-\theta(x')}{-LH_\alpha 1(x')}\right)\cdot\|h\|_\infty \quad \text{on } \partial D,$$

so that

$$\begin{aligned}\left\|H_\alpha\left(\overline{LH_\alpha}^{-1}h\right)\right\|_\infty &\le \left\|-\overline{LH_\alpha}^{-1}h\right\|_\infty \\ &\le \sup_{x'\in\partial D}\left(\frac{1-\theta(x')}{-LH_\alpha 1(x')}\right)\cdot\|h\|_\infty.\end{aligned} \tag{11.19}$$

However, there exists a constant $c_0 > 0$ such that

$$0 \le \frac{1-\theta(x')}{m(x')} \le c_0, \quad x' \in \partial D.$$

Hence we have the inequality

$$\begin{aligned}\frac{1-\theta(x')}{-LH_\alpha 1(x')} &\le \left(\frac{1-\theta(x')}{m(x')\left(-L_\nu H_\alpha 1(x')\right)+(1-m(x'))}\right)\\ &\le c_0\left\|\frac{1}{-L_\nu H_\alpha 1}\right\|_\infty \quad \text{for all } \alpha \ge 1.\end{aligned}$$

In view of Lemma 10.25, this implies that

$$\lim_{\alpha\to+\infty}\left[\sup_{x'\in\partial D}\left(\frac{1-\theta(x')}{-LH_\alpha 1(x')}\right)\right] = 0.$$

Summing up, we obtain from inequalities (11.18) and (11.19) that

$$\begin{aligned}&\limsup_{\alpha\to+\infty}\left\|H_\alpha\left(\overline{LH_\alpha}^{-1}\left((1-m(x'))f|_{\partial D}\right)\right)\right\|_\infty\\ &\le \limsup_{\alpha\to+\infty}\left[\left\|H_\alpha\left(\overline{LH_\alpha}^{-1}h\right)\right\|_\infty\right.\end{aligned}$$

$$+ \left\| H_\alpha \left(\overline{LH_\alpha}^{-1} \left((1 - m(x')) f|_{\partial D} \right) \right) - H_\alpha \left(\overline{LH_\alpha}^{-1} h \right) \right\|_\infty \Bigg]$$

$$\leq \lim_{\alpha \to +\infty} \left[\sup_{x' \in \partial D} \left(\frac{1 - \theta(x')}{-LH_\alpha 1(x')} \right) \right] \|h\|_\infty + \frac{\varepsilon}{k_1}$$

$$\leq \frac{\varepsilon}{k_1}.$$

Since ε is arbitrary, this proves that

$$\lim_{\alpha \to +\infty} \left\| H_\alpha \left(\overline{LH_\alpha}^{-1} \left((1 - m(x')) f|_{\partial D} \right) \right) \right\|_\infty = 0. \tag{11.20}$$

Therefore, by combining assertions (11.17) and (11.20) we find that the second term on the right-hand side of (11.15) also tends to zero:

$$\lim_{\alpha \to +\infty} \left\| H_\alpha \left(\overline{LH_\alpha}^{-1} \left(\alpha (1 - m(x')) G_\alpha^\nu f|_{\partial D} \right) \right) \right\|_\infty = 0.$$

This completes the proof of assertion (11.14) and hence that of assertion (iii).

Step 5: Summing up, we have proved that the operator $\mathfrak{W}$, defined by formula (11.4), satisfies conditions (a)–(d) in Theorem 9.35. Hence, in view of assertion (11.2) it follows from an application of part (ii) of the same theorem that the operator $\mathfrak{W}$ is the infinitesimal generator of some Feller semigroup $\{T_t\}_{t \geq 0}$ on $\overline{D} \setminus M$.

The proof of Theorem 11.1 and hence that of Theorem 1.3 is now complete. □

Remark 11.2. It is worth pointing out that if instead of G_α^ν we had used the Green operator G_α^0 for the Dirichlet problem, as in the proof of Theorem 10.2, then our proof would have broken down.

11.3 Notes and Comments

The results discussed in this chapter are adapted from Taira [Ta6] and [Ta9].

Section 11.1: The space $C_0(\overline{D} \setminus M)$ is introduced by Taira [Ta6].

Section 11.2: Theorem 11.1 is a revised version of Taira [Ta6, Theorem 2], and is a generalization of Taira [Ta9, Theorem 1.4]. It should be noted that Taira [Ta8] has proved Theorem 11.1 under the condition that $L_\nu = \partial/\partial \mathbf{n}$ and $\delta(x') \equiv 0$ on ∂D, by using the L^p theory of pseudo-differential operators (see [Ta8, Theorem 2]).

Chapter 12
Markov Processes Revisited

In this book we have mainly studied (temporally homogeneous) Markov transition functions with only informal references to the random variables which actually form the Markov processes themselves (see Sect. 9.1). In this chapter we study this neglected side of our subject. The discussion will have a more measure-theoretical flavor than hitherto.

Section 12.1 is devoted to a review of the basic definitions and properties of (temporally homogeneous) Markov processes. In Sect. 12.2 we consider when the paths of a Markov process are actually continuous, and prove Theorem 9.20 (Corollary 12.8). In Sect. 12.3 we give a useful criterion for path-continuity of a Markov process $\{x_t\}$ in terms of the infinitesimal generator $\mathfrak{A}$ of the associated Feller semigroup $\{T_t\}$ (Theorem 12.10). Section 12.4 is devoted to examples of multi-dimensional diffusion processes. More precisely, we prove that (1) reflecting barrier Brownian motion (Theorem 12.12), (2) reflecting and absorbing barrier Brownian motion (Theorem 12.15), (3) reflecting, absorbing and drifting barrier Brownian motion (Theorem 12.16) are typical examples of multi-dimensional diffusion processes, that is, examples of continuous strong Markov processes.

12.1 Basic Definitions and Properties of Markov Processes

First, we recall the basic definitions of stochastic processes (see Sect. 9.1.1). Let K be a locally compact, separable metric space and $\mathcal{B}$ the σ-algebra of all Borel sets in K. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function X defined on Ω taking values in K is called a random variable if it satisfies the condition

$$X^{-1}(E) = \{X \in E\} \in \mathcal{F} \quad \text{for all } E \in \mathcal{B},$$

that is, X is $\mathcal{F}/\mathcal{B}$-measurable.

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_12,

A family $\mathcal{X} = \{x(t, \omega)\}$, $t \in [0, \infty)$, $\omega \in \Omega$, of random variables is called a stochastic process. We regard the process $\mathcal{X}$ primarily as a function of t whose values $x(t, \cdot)$ for each t are random variables defined on Ω taking values in K. More precisely, we are dealing with one function of two variables, that is, for each fixed t the function $x_t(\cdot)$ is $\mathcal{F}/\mathcal{B}$-measurable. If, instead of t, we fix an $\omega \in \Omega$, then we obtain a function $x(\cdot, \omega) : [0, \infty) \to K$ which may be thought of as the motion in time of a physical particle. In this context, the space K is called the state space and Ω the sample space. The function $x_t(\omega) = x(t, \omega)$, $t \in [0, \infty)$, defines in the state space K a trajectory or a path of the process corresponding to the sample point ω.

Sometimes it is useful to think of a stochastic process $\mathcal{X}$ specifically as a function of two variables $x(t, \omega)$ where $t \in [0, \infty)$ and $\omega \in \Omega$. One powerful tool in this connection is Fubini's theorem. To do this, we introduce a class of stochastic processes which we will deal with in this chapter.

Definition 12.1. A stochastic process $\mathcal{X} = \{x_t\}_{t \geq 0}$ is said to be *measurable* provided that the function $x(\cdot, \cdot) : [0, \infty) \times \Omega \to K$ is measurable with respect to the product σ-algebra $\mathcal{A} \times \mathcal{F}$, where $\mathcal{A}$ is the σ-algebra of all Borel sets in the interval $[0, \infty)$.

We remark that the condition that the function $x_t(\cdot) = x(t, \cdot)$ is $\mathcal{F}/\mathcal{B}$-measurable for each t does not guarantee the measurability of the process $x(\cdot, \cdot)$.

Now let p_t be a (temporally homogeneous) Markov transition function on the metric space K (see Definition 9.4). The idea behind Definition 9.4 of a (temporally homogeneous) Markov transition function suggests the following definition (cf. formula (9.3)):

Definition 12.2. A stochastic process $\mathcal{X} = \{x_t\}_{t \geq 0}$ is said to be *governed by the (temporally homogeneous) transition function* p_t provided that we have, for all $0 \leq t_1 < t_2 < \ldots < t_n < \infty$ and all Borel sets $B_1, B_2, \ldots, B_n \in \mathcal{B}$,

$$P(\omega \in \Omega : x_{t_1}(\omega) \in B_1, x_{t_2}(\omega) \in B_2, \ldots, x_{t_n}(\omega) \in B_n) \tag{12.1}$$
$$= \int_{y_n \in B_n} \cdots \int_{y_1 \in B_1} \int_{y_1 \in B_1} \int_{x \in K} \mu(dx) p_{t_1}(x, dy_1)$$
$$p_{t_2 - t_1}(y_1, dy_2) \cdots p_{t_n - t_{n-1}}(y_{n-1}, dy_n),$$

where μ is some probability measure on the measurable space $(K, \mathcal{B})$, and is called the *initial distribution* of the process $\{x_t\}$.

Remark 12.3. Formula (12.1) expresses the "starting afresh" property of a stochastic process that if a Markovian particle reaches a position, then it behaves subsequently as though that position had been its initial position.

The notion of the (temporally homogeneous) Markov property is introduced and discussed in Sect. 9.1: If $\mathcal{X} = \{x_t\}_{t \geq 0}$ is a stochastic process, we introduce three sub-σ-algebras of $\mathcal{F}$ as follows:

$$\begin{cases} \mathcal{F}_{\leq t} & = \sigma(x_s : 0 \leq s \leq t) \\ & = \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ & \quad \text{with respect to which all } x_s, 0 \leq s \leq t, \text{ are measurable,} \\ \mathcal{F}_{=t} & = \sigma(x_t) \\ & = \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ & \quad \text{with respect to which } x_t \text{ is measurable,} \\ \mathcal{F}_{\geq t} & = \sigma(x_s : t \leq s < \infty) \\ & = \text{the smallest } \sigma\text{-algebra contained in } \mathcal{F} \\ & \quad \text{with respect to which all } x_s, t \leq s < \infty, \text{ are measurable.} \end{cases}$$

We recall that an event in $\mathcal{F}_{\leq t}$ is determined by the behavior of the process $\{x_s\}$ up to time t and an event in $\mathcal{F}_{\geq t}$ by its behavior after time t. Thus they represent respectively the "past" and "future" relative to the "present" moment.

A stochastic process $\mathcal{X} = \{x_t\}$ is called a (temporally homogeneous) Markov process if it satisfies the condition

$$P(B \mid \mathcal{F}_{\leq t}) = P(B \mid \mathcal{F}_{=t}) \quad \text{for any "future" set } B \in \mathcal{F}_{\geq t}.$$

More precisely, we have, for any "future" set $B \in \mathcal{F}_{\geq t}$,

$$P(A \cap B) = \int_A P(B \mid \mathcal{F}_{=t})(\omega)\, dP(\omega) \quad \text{for every "past" set } A \in \mathcal{F}_{\leq t}.$$

Intuitively, this means that the conditional probability of a "future" event B given the "present" is the same as the conditional probability of B given the "present" and "past".

The next theorem justifies Definition 12.2, and hence it will be fundamental for our further study of (temporally homogeneous) Markov processes:

Theorem 12.4. *Let $\mathcal{X} = \{x_t\}_{t \geq 0}$ be any stochastic process with values in the metric space K which is governed by a (temporally homogeneous) Markov transition function p_t. Then it follows that $\{x_t\}$ is a* (temporally homogeneous) Markov process.

Our study of Markov processes is based on formula (12.1) which shows how the finite-dimensional distributions of the process $\mathcal{X} = \{x_t\}$ are calculated from the Markov transition function p_t. However, knowledge of all the finite-dimensional distributions may not be sufficient to precisely determine the path functions of a Markov process. Therefore, it is important to ask the following question:

Question 12.5. Given a (temporally homogeneous) Markov transition function p_t and an initial distribution μ, does there exist a (temporally homogeneous) Markov process $\mathcal{X} = \{x_t\}$ having the corresponding finite-dimensional distributions whose paths are almost surely "nice" in some sense?

We say that two (temporally homogeneous) Markov processes $\mathcal{X} = \{x_t\}_{t\geq 0}$ and $\mathcal{Y} = \{y_t\}_{t\geq 0}$ defined on the same probability space $(\Omega, \mathcal{F}, P)$ is *equivalent* provided that we have, for all $t \in [0, \infty)$,

$$P(\{\omega \in \Omega : x_t(\omega) = y_t(\omega)\}) = 1.$$

The next theorem asserts that, under quite general conditions there does exist a Markov process with "nice" paths equivalent to any given process:

Theorem 12.6. *Let* (K, ρ) *be a* compact *metric space and let* $\mathcal{X} = \{x_t\}_{t\geq 0}$ *be a stochastic process with values in* K *which is governed by a (temporally homogeneous)* normal *Markov transition function* p_t. *Then there exists a (temporally homogeneous) Markov process* $\mathcal{Y} = \{y_t\}$, *equivalent to the process* $\mathcal{X} = \{x_t\}$, *such that*

$$\begin{aligned} P(\{\omega \in \Omega :\ & \textit{the function } y_t(\omega) \textit{ are right-continuous and} \\ & \textit{have left-hand limits for all } t \geq 0\}) = 1. \end{aligned}$$

12.2 Path-Continuity of Markov Processes

It is naturally interesting and important to consider when the paths of a (temporally homogeneous) Markov process $\{x_t\}$ are actually continuous for all $t \geq 0$. The purpose of this section is to establish some useful sufficient conditions for path-continuity of the Markov process $\{x_t\}$.

First, we have the following theorem:

Theorem 12.7. *Let* (K, ρ) *be a locally compact metric space and let* $\mathcal{X} = \{x_t\}_{t\geq 0}$ *be a measurable stochastic process with values in* K. *Assume that, for each* $\varepsilon > 0$ *and each* $M > 0$, *the condition*

$$P(\{\omega \in \Omega : \rho(x_t(\omega), \rho_{t+h}(\omega)) \geq \varepsilon\}) = o(h) \quad \textit{as } h \downarrow 0 \tag{12.2}$$

holds uniformly for all $t \in [0, M]$, *i.e., we have, for all* $t \in [0, M]$,

$$\lim_{h\downarrow 0} \frac{P(\{\omega \in \Omega : \rho(x_t(\omega), \rho_{t+h}(\omega)) \geq \varepsilon\})}{h} = 0.$$

Then it follows that

$$P(\{\omega \in \Omega :\ x_t(\omega) \textit{ has a jump discontinuity somewhere}\}) = 0. \tag{12.3}$$

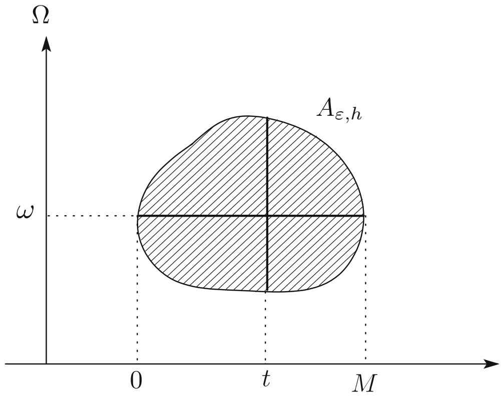


Fig. 12.1 The set $A_{\varepsilon,h}$

Proof. The proof is divided into two steps.

Step 1: First, we show that if we let

$$J_{\varepsilon,h}(\omega) = \{t \in [0, M] : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon\}, \quad \omega \in \Omega, \tag{12.4}$$

then it follows from condition (12.2) that

$$E\,[m(J_{\varepsilon,h})] = o(h) \quad \text{as } h \downarrow 0, \tag{12.5}$$

where $m = dt$ is the Lebesgue measure on $\mathbf{R}$.
To do this, we define the set (see Fig. 12.1)

$$A_{\varepsilon,h} = \{(t, \omega) \in [0, M] \times \Omega : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon\}.$$

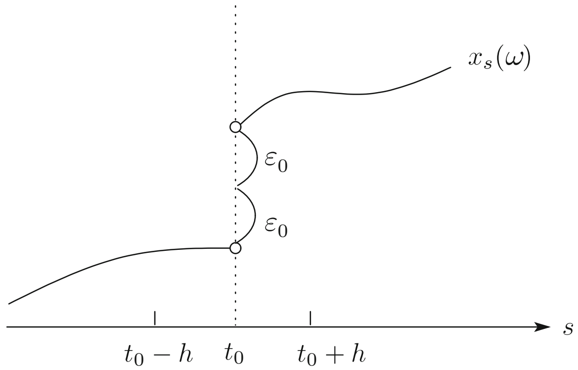
We remark that the set $A_{\varepsilon,h}$ is measurable with respect to the product σ-algebra $\mathcal{A} \times \mathcal{F}$. Indeed, it suffices to note the following two facts (a) and (b):

(a) The mapping $(t, \omega) \mapsto (x_t(\omega), x_{t+h}(\omega))$ is measurable from the product space $[0, M] \times \Omega$ into the product space $K \times K$, for each $h \geq 0$.
(b) The metric ρ is a continuous function on the product space $K \times K$.

By virtue of Fubini's theorem, we can compute the product measure $m \times P$ of $A_{\varepsilon,h}$ by integrating the measure of a cross section as follows:

$$(m \times P)(A_{\varepsilon,h}) = \int_0^M P(\{\omega \in \Omega : (t, \omega) \in A_{\varepsilon,h}\})\, dt.$$

Fig. 12.2 The trajectory $x_t(\omega)$



By condition (12.2), it follows that

$$(m \times P)(A_{\varepsilon,h}) = o(h) \quad \text{as } h \downarrow 0. \tag{12.6}$$

Moreover, by integrating in the other order we obtain from definition (12.4) that

$$\begin{aligned}(m \times P)(A_{\varepsilon,h}) &= \int_\Omega m(\{t \in \mathbf{R} : (t, \omega) \in A_{\varepsilon,h}\})\, dP(\omega) \\ &= E\left[m(\{t \in [0, M] : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon\})\right] \\ &= E\left[m(J_{\varepsilon,h})\right].\end{aligned} \tag{12.7}$$

Therefore, the desired assertion (12.5) follows by combining assertions (12.6) and (12.7).

Step 2: Now we show that the existence of jumps in the trajectories of $\{x_t\}$ contradicts assertion (12.5).

Step 2-1: We assume that, for some $\varepsilon_0 > 0$,

$$P(\{\omega \in \Omega : x_t(\omega) \text{ has a jump with gap greater than } 2\varepsilon_0\}) > 0.$$

If $x_t(\omega)$ is a trajectory having such a jump at $t = t_0$ (see Fig. 12.2), we obtain that the two limits

$$x_{t_0^+}(\omega) = \lim_{h \downarrow 0} x_{t_0+h}(\omega), \quad x_{t_0^-}(\omega) = \lim_{h \downarrow 0} x_{t_0-h}(\omega)$$

exist and satisfy the condition

$$\rho(x_{t_0^-}(\omega), x_{t_0^+}(\omega)) \geq 2\varepsilon_0.$$

Then we have, for sufficiently small $h > 0$,

$$\begin{cases} \rho(x_{t_0^-}(\omega), x_s(\omega)) \leq \dfrac{\varepsilon_0}{2} & \text{for all } s \in (t_0 - h, t_0), \\ \rho(x_{t_0^+}(\omega), x_s(\omega)) \leq \dfrac{\varepsilon_0}{2} & \text{for all } s \in (t_0, t_0 + h). \end{cases}$$

Moreover, by the triangle inequality it follows that

$$\begin{aligned} 2\varepsilon_0 &\leq \rho(x_{t_0^-}(\omega), x_{t_0^+}(\omega)) \\ &\leq \rho(x_{t_0^-}(\omega), x_t(\omega)) + \rho(x_t(\omega), x_{t+h}(\omega)) + \rho(x_{t+h}(\omega), x_{t_0^+}(\omega)) \\ &\leq \frac{\varepsilon_0}{2} + \rho(x_t(\omega), x_{t+h}(\omega)) + \frac{\varepsilon_0}{2} \quad \text{for all } t \in (t_0 - h, t_0), \end{aligned}$$

so that

$$\rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon_0 \quad \text{for all } t \in (t_0 - h, t_0).$$

This implies that, for all sufficiently small $h > 0$,

$$m\left(\{t \in [0, M] : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon_0\}\right) \geq m\left((t_0 - h, t_0)\right) = h,$$

or equivalently,

$$\frac{m(J_{\varepsilon_0,h})(\omega)}{h} \geq 1 \quad \text{for all sufficiently small } h > 0. \tag{12.8}$$

Step 2-2: Now we assume, to the contrary, that

$$P(\{\omega \in \Omega : x_t(\omega) \text{ has a jump discontinuity somewhere}\}) > 0.$$

Then it follows from Step 2-1 that there exists a positive number ε_0 such that assertion (12.8) holds true.

Therefore, we can find a positive constant $\delta = \delta(\varepsilon_0)$ such that

$$\begin{aligned} E[m(J_{\varepsilon_0,h})] = \int_\Omega m(J_{\varepsilon_0,h})(\omega)\, dP(\omega) &\geq \int_{L_{\varepsilon_0}} m(J_{\varepsilon_0,h})(\omega)\, dP(\omega) \\ &\geq h\,\delta \quad \text{for all sufficiently small } h > 0, \end{aligned} \tag{12.9}$$

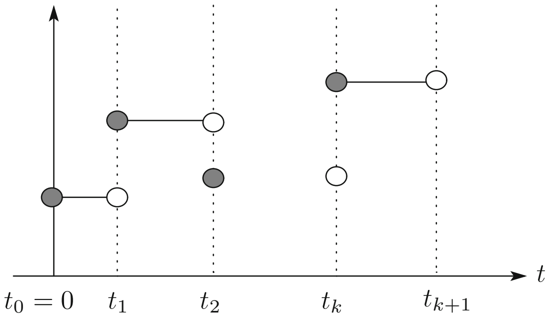
where

$$L_{\varepsilon_0} = \{\omega \in \Omega : x_t(\omega) \text{ has a jump with gap greater than } 2\varepsilon_0\}.$$

Assertion (12.9) contradicts condition (12.5).

The proof of Theorem 12.7 is complete.

Fig. 12.3 The stochastic process $\{y_t\}$



The next corollary proves part (ii) of Theorem 9.20 under condition (N) with $E = K$:

Corollary 12.8. *Let (K, ρ) be a locally compact metric space and let $\mathcal{X} = \{x_t\}_{t \geq 0}$ be a right-continuous Markov process governed by a Markov transition function p_t. Assume that, for each $\varepsilon > 0$, the condition*

$$p_h(x, K \setminus U_\varepsilon(x)) = o(h) \tag{12.10}$$

holds uniformly in $x \in K$ as $h \downarrow 0$. In other words, for each $\varepsilon > 0$ we have the condition

$$\lim_{h \downarrow 0} \frac{1}{h} \sup_{x \in K} p_t(x, K \setminus U_\varepsilon(x)) = 0.$$

Here $U_\varepsilon(x) = \{y \in K : \rho(y, x) < \varepsilon\}$ is an ε-neighborhood of x.

Then it follows that

$$P(\{\omega \in \Omega : x_t(\omega) \text{ is continuous for all } t \geq 0\}) = 1.$$

Proof. We$\{x_t\}$ is right-continuous and has limits from the left as well. The proof of Corollary 12.8 is divided into two steps.

Step 1: First, we prove the following lemma (see Dynkin [Dy1, Lemma 5.9]):

Lemma 12.9. *Every right-continuous stochastic process $\{x_t\}$ is measurable.*

Proof. Let $\{t_n\}$ be an arbitrary increasing sequence such that

$$0 = t_0 < t_1 < t_2 < \ldots \to \infty.$$

Then we can define a stochastic process $\{y_t\}$ by the formula (see Fig. 12.3)

$$y_t(\omega) = \begin{cases} x_{t_0}(\omega) = x(t_0, \omega), \ \omega \in \Omega & \text{for } t_0 \leq t < t_1, \\ x_{t_1}(\omega) = x(t_1, \omega), \ \omega \in \Omega & \text{for } t_1 \leq t < t_2, \\ \quad \cdot \\ \quad \cdot \\ x_{t_k}(\omega) = x(t_k, \omega), \ \omega \in \Omega & \text{for } t_k \leq t < t_{k+1}, \\ \quad \cdot \\ \quad \cdot \end{cases}$$

Since we have, for any $a \in \mathbf{R}$,

$$\{(t, \omega) : y_t(\omega) < a\} = \bigcup_{i=0}^{\infty} [t_i, t_{i+1}) \times \{\omega : x_{t_i}(\omega) < a\},$$

it follows that the process $\{y_t\}$ is measurable.

For each integer $n \in \mathbf{N}$, we choose a non-negative integer k such that

$$\frac{k}{2^n} \leq t < \frac{k+1}{2^n},$$

and let

$$\phi_n(t) = \frac{k+1}{2^n}.$$

It is clear that $\phi_n(t) \downarrow t$ as $n \to \infty$. Now, if we define a stochastic process $\{x_n\}$ by the formula

$$x_n(t, \omega) = x(\phi_n(t), \omega) \quad \text{for every } \omega \in \Omega,$$

then we have the following two assertions (a) and (b):

(a) The process $x_n(t)$ is measurable.
(b) $x_n(t, \omega) \to x(t, \omega)$ as $n \to \infty$.

Indeed, assertion (a) is proved just as in the case of the process $\{y_t\}$, while assertion (b) follows from the right-continuity of the process $\{x_t\}$.

Therefore, we obtain from assertions (a) and (b) that the process $\{x_t\}$ is measurable.

The proof of Lemma 12.9 is complete.

Step 2: Let μ be the initial distribution of the process $\{x_t\}$ in Definition 12.2. Namely, we have, for all $0 \leq t_1 < t_2 < \ldots < t_n < \infty$ and all Borel sets B_1, B_2, ..., $B_n \in \mathcal{B}$,

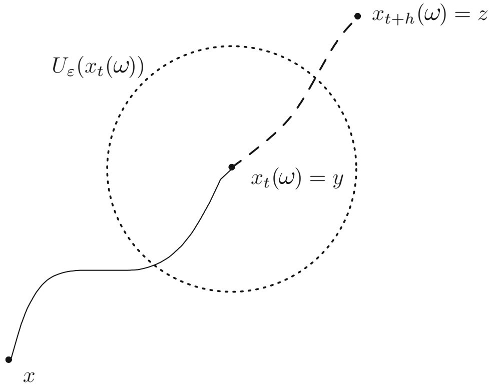


Fig. 12.4 $\rho(y, z) \geq \varepsilon \Leftrightarrow y \in K,\, z \in K \setminus U_\varepsilon(y)$

$$\begin{aligned}
&P(\omega \in \Omega : x_{t_1}(\omega) \in B_1, x_{t_2}(\omega) \in B_2, \ldots, x_{t_n}(\omega) \in B_n) \\
&= \int_{y_n \in B_n} \cdots \int_{y_1 \in B_1} \int_{y_1 \in B_1} \int_{x \in K} \mu(dx)\, p_{t_1}(x, dy_1) \\
&\qquad p_{t_2 - t_1}(y_1, dy_2) \cdots p_{t_n - t_{n-1}}(y_{n-1}, dy_n).
\end{aligned}$$

Then we have, by Fubini's theorem (see Fig. 12.4),

$$\begin{aligned}
&P(\{\omega \in \Omega : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon\}) \qquad (12.11) \\
&= P(\{\omega \in \Omega : x_t(\omega) \in K, x_{t+h}(\omega) \in K \setminus U_\varepsilon(x_t(\omega))\}) \\
&= \int_K \int_{K \setminus U_\varepsilon(y)} \left(\int_K p_t(x, dy) \right) p_h(y, dz)\, \mu(dx) \\
&= \iint_K \left(\int_{K \setminus U_\varepsilon(y)} p_h(y, dz) \right) p_t(x, dy)\, \mu(dx) \\
&= \iint_K p_t(x, dy)\, p_h(y, K \setminus U_\varepsilon(y))\, \mu(dx).
\end{aligned}$$

In view of condition (12.10), we obtain from formula (12.11) that the assertion

$$\begin{aligned}
&\lim_{h \downarrow 0} \frac{P(\{\omega \in \Omega : \rho(x_t(\omega), x_{t+h}(\omega)) \geq \varepsilon\})}{h} \\
&= \lim_{h \downarrow 0} \frac{\iint_K p_t(x, dy)\, p_h(y, K \setminus U_\varepsilon(y))\, \mu(dx)}{h} \\
&= \iint_K p_t(x, dy) \left(\lim_{h \downarrow 0} \frac{p_h(y, K \setminus U_\varepsilon(y))}{h} \right) \mu(dx) \\
&= 0
\end{aligned}$$

holds uniformly for all $t \geq 0$. This assertion implies that condition (12.2) holds uniformly for all $t \geq 0$, as $h \downarrow 0$. Hence it follows from an application of Theorem 12.7 that

$$P(\{\omega : x_t(\omega) \text{ has a jump discontinuity somewhere}\}) = 0. \tag{12.12}$$

However, we recall that the stochastic process $\{x_t\}$ is right-continuous and has limits from the left as well.

Therefore, we obtain from assertion (12.12) that

$$P(\{\omega \in \Omega : x_t(\omega) \text{ is continuous for all } t \geq 0\}) = 1.$$

The proof of Corollary 12.8 is complete.

12.3 Path-Continuity of Markov Processes Associated with Semigroups

It is usually difficult to verify condition (12.10) directly, since it is rather exceptional when any simple formula for the transition probability function p_t is available. The purpose of this section is to give a useful criterion for path-continuity of the Markov process $\{x_t\}$ in terms of the infinitesimal generator $\mathfrak{A}$ of the associated Feller semigroup $\{T_t\}$.

Let (K, ρ) be a compact metric space and let $C(K)$ be the space of real-valued, bounded continuous functions on K; $C(K)$ is a Banach space with the maximum norm

$$\|f\|_\infty = \max_{x \in K} |f(x)|.$$

A strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on the space $C(K)$ is called a Feller semigroup if it is non-negative and contractive on $C(K)$, that is,

$$f \in C(K),\ 0 \leq f(x) \leq 1 \quad \text{on } K \implies 0 \leq T_t f(x) \leq 1 \quad \text{on } K.$$

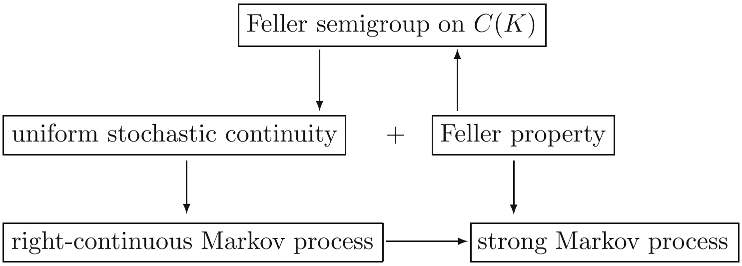
We recall (see Theorem 9.34) that if p_t is a uniformly stochastically continuous Feller function on K, then its associated operators $\{T_t\}_{t \geq 0}$, defined by formula

$$T_t f(x) = \int_K p_t(x, dy) f(y) \quad \text{for every } f \in C(K), \tag{12.13}$$

form a Feller semigroup on K. Conversely, if $\{T_t\}_{t \geq 0}$ is a Feller semigroup on K, then there exists a uniformly stochastically continuous Feller function p_t on K such that (12.13) holds true.

Furthermore, we know that the function p_t is the transition function of some *strong Markov process* $\mathcal{X} = \{x_t\}_{t \geq 0}$ whose paths are right-continuous and have no discontinuities other than jumps.

Our approach can be visualized as follows:



The next theorem gives some useful sufficient conditions for path-continuity of the Markov process $\{x_t\}$ in terms of the infinitesimal generator $\mathfrak{A}$ of the associated Feller semigroup $\{T_t\}$:

Theorem 12.10. *Let* (K, ρ) *be a* compact *metric space and let* $\mathcal{X} = \{x_t\}_{t \geq 0}$ *be a right-continuous Markov process governed by a uniformly stochastically continuous Feller transition function* p_t *on* K. *Assume that the infinitesimal generator* $\mathfrak{A}$ *of the associated Feller semigroup* $\{T_t\}_{t \geq 0}$*, defined by formula* (12.13)*, satisfies the condition that, for each* $\varepsilon > 0$ *and each point* $x \in K$*, there exists a function* $f \in \mathcal{D}(\mathfrak{A})$ *which satisfies the following three conditions (i)–(iii):*

(i) $f(x) \geq 0$ *on* K.
(ii) $f(y) > 0$ *for all* $y \in K \setminus U_\varepsilon(x)$.
(iii) $f(y) = \mathfrak{A} f(y) = 0$ *in some neighborhood of* x.

Here $U_\varepsilon(x) = \{z \in K : \rho(z, x) < \varepsilon\}$ *is an* ε*-neighborhood of* x.

Then we have

$$P(\{\omega \in \Omega : x_t(\omega) \text{ is continuous for all } t \geq 0\}) = 1.$$

Proof. We shall apply Corollary 12.8 to our situation.

To do this, we assume, to the contrary, that condition (12.10) does not hold. Then we can find a positive number ε_0 such that

$$p_h(x, K \setminus U_{2\varepsilon_0}(x)) \quad \text{is not of order } o(h) \text{ uniformly in } x \text{ as } h \downarrow 0.$$

More precisely, for this ε_0 there exist a positive constant δ, a decreasing sequence $\{h_n\}$ of positive numbers, $h_n \downarrow 0$, and a sequence $\{x_n\}$ of points in K such that

$$p_{h_n}(x_n, K \setminus U_{2\varepsilon_0}(x_n)) \geq \delta\, h_n \quad \text{for all sufficiently large } n. \tag{12.14}$$

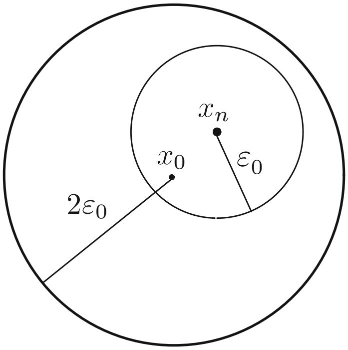


Fig. 12.5 The open balls $U_{2\varepsilon_0}(x_0)$ and $U_{\varepsilon_0}(x_n)$

However, since K is compact, we may assume that the sequence $\{x_n\}$ itself converges to some point x_0 of K. If $\rho(x_n, x_0) < \varepsilon_0$, then it follows that (see Fig. 12.5)

$$U_{\varepsilon_0}(x_n) \subset U_{2\varepsilon_0}(x_0).$$

Hence we have, by assertion (12.14),

$$\begin{aligned} p_{h_n}(x_n, K \setminus U_{\varepsilon_0}(x_0)) &\geq p_{h_n}(x_n, K \setminus U_{2\varepsilon_0}(x_n)) \\ &\geq \delta\, h_n \quad \text{for all sufficiently large } n. \end{aligned} \tag{12.15}$$

Now, for these ε_0 and x_0 we can construct a function $f \in \mathcal{D}(\mathfrak{A})$ which satisfies conditions (i)–(iii). It follows from condition (ii) that

$$c = \min_{x \in K \setminus U_{\varepsilon_0}(x_0)} f(x) > 0.$$

Then we have, by assertion (12.15),

$$\begin{aligned} T_{h_n} f(x_n) &= \int_K p_{h_n}(x_n, dy) f(y) \geq \int_{K \setminus U_{\varepsilon_0}(x_0)} p_{h_n}(x_n, dy) f(y) \\ &\geq \left(\min_{K \setminus U_{\varepsilon_0}(x_0)} f \right) \cdot p_{h_n}(x_n, K \setminus U_{\varepsilon_0}(x_0)) \\ &\geq c\delta\, h_n \quad \text{for all sufficiently large } n. \end{aligned} \tag{12.16}$$

However, since $f(x_n) = 0$ for x_n close to the limit point x_0, we obtain from assertion (12.16) that

$$\frac{T_{h_n} f(x_n) - f(x_n)}{h_n} = \frac{T_{h_n} f(x_n)}{h_n} \geq c\delta \quad \text{for all } x_n \in U(x_0). \tag{12.17}$$

On the other hand, since we have, for $f \in \mathcal{D}(\mathfrak{A})$,

$$\mathfrak{A} f(y) = \lim_{n \to \infty} \frac{T_{h_n} f(y) - f(y)}{h_n} \quad \text{uniformly in } y \in K,$$

it follows from condition (iii) that

$$\frac{T_{h_n} f(y) - f(y)}{h_n} \longrightarrow \mathfrak{A} f(y) = 0 \quad \text{uniformly in } y \in U(x_0).$$

Hence we have, for all $y \in U(x_0)$,

$$\left| \frac{T_{h_n} f(y) - f(y)}{h_n} \right| < \frac{c\delta}{2} \quad \text{for all sufficiently large } n. \tag{12.18}$$

However, since $f(x_n) = 0$ and $T_{h_n} f \geq 0$ on K, by letting $y := x_n$ in inequality (12.18) we obtain that

$$\frac{T_{h_n} f(x_n) - f(x_n)}{h_n} = \left| \frac{T_{h_n} f(x_n) - f(x_n)}{h_n} \right| < \frac{c\delta}{2} \quad \text{for all } x_n \in U(x_0).$$

This assertion contradicts assertion (12.17).

Therefore, we have proved that condition (12.10) holds true. Theorem 12.10 follows from an application of Corollary 12.8.

The proof of Theorem 12.10 is complete.

12.4 Examples of Diffusion Processes on a Bounded Domain

In this section we prove that (1) reflecting barrier Brownian motion, (2) reflecting and absorbing barrier Brownian motion, (3) reflecting, absorbing and drifting barrier Brownian motion are typical examples of multi-dimensional diffusion processes, that is, examples of continuous strong Markov processes on a bounded domain. It should be emphasized that these three Brownian motions correspond to the Neumann boundary value problem, the Robin boundary value problem and the oblique derivative boundary value problem for the Laplacian Δ in terms of elliptic boundary value problems, respectively.

The following diagram gives a bird's eye view of diffusion processes and elliptic boundary value problems and how these relate to each other:

Diffusion Processes	Elliptic Boundary Value Problems
Brownian motion	Laplacian Δ
Reflecting barrier	Neumann boundary condition $\mathcal{B}_N$
Reflecting and absorbing barrier	Robin boundary condition $\mathcal{B}_R$
Reflecting, absorbing and drifting barrier	Oblique derivative boundary condition $\mathcal{B}_O$

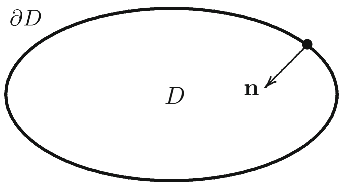


Fig. 12.6 The bounded domain D with smooth boundary ∂D

12.4.1 The Neumann Case

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$ with smooth boundary ∂D; its closure $\overline{D} = D \cup \partial D$ is an N-dimensional, compact smooth manifold with boundary (see Fig. 12.6).

First, we consider the *Neumann boundary condition*

$$\mathcal{B}_N u = \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial D,$$

where $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ is the unit interior normal to the boundary ∂D.

We introduce a linear operator $\mathfrak{A}_N$ as follows:

(a) $D(\mathfrak{A}_N) = \left\{ u \in C(\overline{D}) : \Delta u \in C(\overline{D}),\ \dfrac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial D \right\}$, where Δu is taken in the sense of *distributions*.

(b) $\mathfrak{A}_N u = \Delta u$ for every $u \in D(\mathfrak{A}_N)$.

Then it follows from an application of Theorem 1.2 with $L := \mathcal{B}_N$ that the operator $\mathfrak{A}_N$ is the infinitesimal generator of a Feller semigroup $\{S_t\}_{t \geq 0}$. Let $\{x_t(\omega)\}$ be the strong Markov process corresponding to the Feller semigroup $\{S_t\}_{t \geq 0}$ with Neumann boundary condition. In this subsection we study the path-continuity of the Markov process $\{x_t(\omega)\}$. In order to make use of Theorem 12.10, we shall construct a function $f \in D(\mathfrak{A}_N)$ which satisfies conditions (i)–(iii) of the same theorem. Our construction of the function $f(x)$ may be visualized by Fig. 12.7.

(1) The case where x_0 is an arbitrary (interior) point of D: By applying Theorem [Ta10, Theorem 1.5] with $\mu(x') := 1$ and $\gamma(x') := 0$, we can find a function

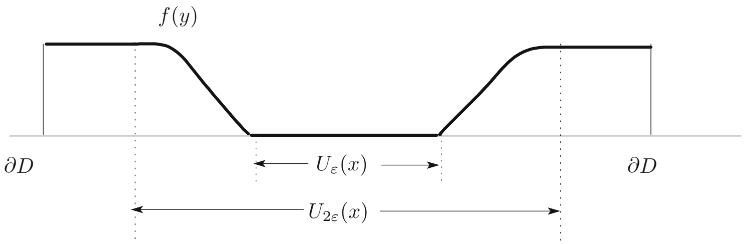


Fig. 12.7 The function $f \in D(\mathfrak{A}_N)$

$\phi \in C^\infty(\overline{D})$ such that

$$\begin{cases} \Delta\phi = -1 & \text{in } D, \\ \dfrac{\partial \phi}{\partial \mathbf{n}} = 0 & \text{on } \partial D. \end{cases} \tag{12.19}$$

Then we have the following lemma:

Lemma 12.11. *The function ϕ satisfies the condition*

$$\phi(x) > 0 \quad \text{on } \overline{D}.$$

Proof. We assume, to the contrary, that

$$\min_{\overline{D}} \phi \leq 0.$$

(1) First, we consider the case: There exists a point $x_0 \in D$ such that

$$\phi(x_0) = \min_{\overline{D}} \phi \leq 0.$$

Since we have, by condition (12.19),

$$\Delta(-\phi) = 1 > 0 \quad \text{in } D, \tag{12.20}$$

it follows from an application of the strong maximum principle (Theorem 8.13 in Chap. 8) with $W := \Delta$ and $u := -\phi$ that

$$\phi(x) \equiv \phi(x_0) \quad \text{in } D.$$

Hence we have

$$\Delta(-\phi) = 0 \quad \text{in } D.$$

This contradicts condition (12.20).

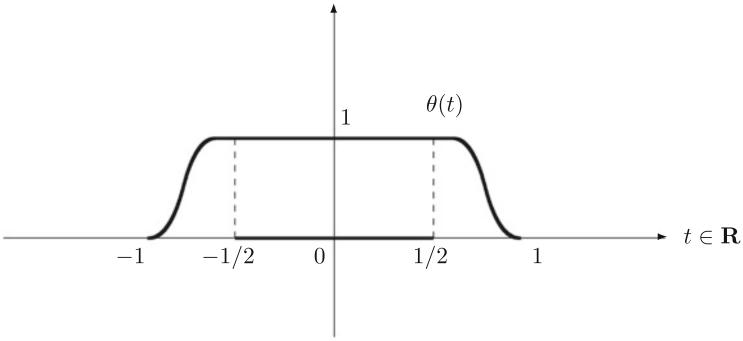


Fig. 12.8 The function $\theta(t)$

(2) Next we consider the case: There exists a point $x_0' \in \partial D$ such that

$$\phi(x_0') = \min_{\overline{D}} \phi \leq 0.$$

Then, by applying the boundary point lemma (Theorem 8.15) with $W := \Delta$ to the function $u := -\phi$ we obtain from inequality (12.20) that

$$\frac{\partial \phi}{\partial \mathbf{n}}(x_0') > 0.$$

This contradicts the boundary condition

$$\frac{\partial \phi}{\partial \mathbf{n}} = 0 \quad \text{on } \partial D.$$

The proof of Lemma 12.11 is complete.

Now we choose a real-valued, smooth function $\theta \in C^\infty(\mathbf{R})$ such that (see Fig. 12.8)

$$\begin{cases} 0 \leq \theta(t) \leq 1 & \text{on } \mathbf{R}, \\ \operatorname{supp} \theta \subset (-1, 1), & \\ \theta(t) = 1 & \text{for all } t \in \left[-\frac{1}{2}, \frac{1}{2}\right], \end{cases} \tag{12.21}$$

and define a function $f \in C^\infty(\overline{D})$ by the formula

$$f(x) = \left(1 - \theta\left(\frac{|x - x_0|}{\varepsilon}\right)\right)\phi(x),$$

where

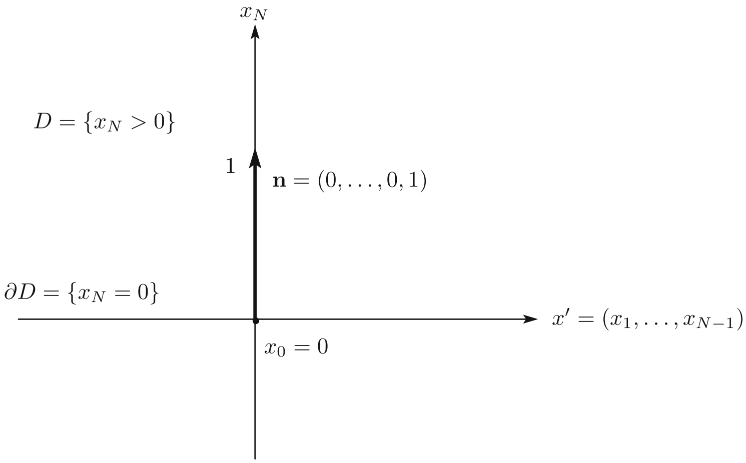


Fig. 12.9 The local coordinate system (x', x_N)

$$0 < \varepsilon < \operatorname{dist}(x, \partial D) = \inf_{z \in \partial D} |x - z|.$$

Then, in view of Lemma 12.11 it is easy to verify that $f \in D(\mathfrak{A}_N)$ and satisfies conditions (i)–(iii) of Theorem 12.10.

(2) The case where x_0 is an arbitrary (boundary) point of ∂D: By change of coordinates, we may assume that

$$x_0 = (0', 0) = (0, \ldots, 0, 0),$$
$$\mathbf{n} = (0', 1) = (0, \ldots, 0, 1) \in \mathbf{R}^N.$$

The situation may be represented schematically by Fig. 12.9.

If we define a function $f \in C^\infty(\overline{D})$ by the formula

$$f(x', x_N) = \left(1 - \theta\left(\frac{x_N}{\varepsilon}\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\phi(x', x_N),$$
$$x' = (x_1, x_2, \ldots, , x_{N-1}) \in \mathbf{R}^{N-1}, \; x_N \in \mathbf{R},$$

then it follows from conditions (12.21) and (12.19) that

$$\frac{\partial f}{\partial \mathbf{n}} = \left.\frac{\partial f}{\partial x_N}\right|_{x_N=0}$$
$$= -\left(\frac{\partial}{\partial x_N}\left(\theta\left(\frac{x_N}{\varepsilon}\right)\right)\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\Bigg|_{x_N=0} \cdot \phi(x', 0)$$

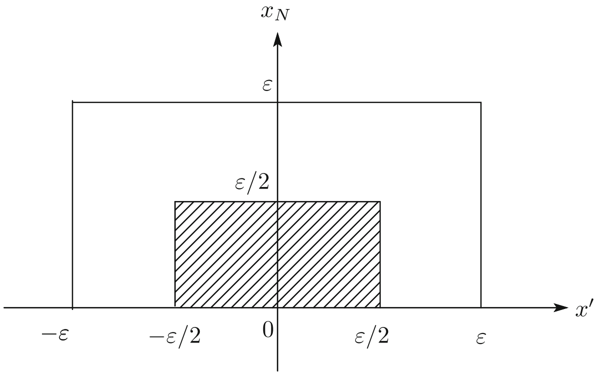


Fig. 12.10 The neighborhoods $U_{\varepsilon/2}(x_0)$ and $U_\varepsilon(x_0)$

$$\begin{aligned}&+\left(1-\theta\left(\frac{x_N}{\varepsilon}\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\cdot\frac{\partial\phi}{\partial x_N}(x',x_N)\bigg|_{x_N=0}\\&=-\frac{1}{\varepsilon}\theta'(0)\theta\left(\frac{|x'|}{\varepsilon}\right)\cdot\phi(x',0)+\left(1-\phi\left(\frac{|x'|}{\varepsilon}\right)\right)\frac{\partial\phi}{\partial\mathbf{n}}\\&=0\quad\text{on }\partial D.\end{aligned}$$

This proves that $f \in D(\mathfrak{A}_N)$. Moreover, we have the following three assertions (a)–(c) (see Fig. 12.10):

(a) Since $f(x) = 0$ in the $\varepsilon/2$ neighborhood $U_{\varepsilon/2}(x_0)$ of x_0, it follows that

$$f(y) = \mathfrak{A}_N f(y) = 0 \quad \text{for all } y \in U_{\varepsilon/2}(x_0).$$

This proves that condition (iii) of Theorem 12.10 is satisfied.

(b) It is clear that $f(x) \geq 0$ on $\overline{D}$, so that condition (i) in Theorem 12.10 is satisfied.

(c) Since $f(y) = \phi(y)$ outside the $\sqrt{N}\,\varepsilon$ neighborhood $U_{\sqrt{N}\,\varepsilon}(x_0)$ of x_0, it follows from Lemma 12.11 that

$$f(y) > 0 \quad \text{for all } y \in \overline{D} \setminus U_{\sqrt{N}\,\varepsilon}(x_0).$$

This proves that condition (ii) of Theorem 12.10 is satisfied.

Therefore, by applying Theorem 12.10 to the operator $\mathfrak{A}_N$ we obtain the following generalization of Examples 9.11 and 9.13 to the multi-dimensional case:

Lemma 12.11 $\Longrightarrow$ Theorem 12.10 $\Longrightarrow$ Theorem 12.12

Theorem 12.12. *The strong Markov process $\{x_t(\omega)\}$ associated with the Feller semigroup $\{S_t\}_{t\geq 0}$ enjoys the property*

$$P\left(\{\omega : x_t(\omega) \text{ is continuous for all } t \geq 0\}\right) = 1.$$

Since $\{S_t\}_{t\geq 0}$ is a Feller semigroup on the compact set $\overline{D}$, it follows from an application of Theorem 9.34 that there exists a uniformly stochastically continuous, Feller transition function $p_t(x, \cdot)$ on $\overline{D}$ such that

$$S_t f(x) = \int_{\overline{D}} p_t(x, dy) f(y) \tag{12.22}$$

holds true for all $f \in C(\overline{D})$.

Furthermore, we can prove the following important result:

Proposition 12.13. *The Feller transition function $p_t(x, \cdot)$ is* conservative*, that is, we have, for all $t > 0$,*

$$p_t(x, \overline{D}) = 1 \quad \text{for each } x \in \overline{D}. \tag{12.23}$$

Proof. First, we note that

$$\begin{cases} 1 \in D(\mathfrak{A}_N), \\ \mathfrak{A}_N 1 = 0 \quad \text{in } D. \end{cases}$$

Hence, by applying Theorem 4.30 with $T_t := S_t$ we obtain that the function $u(t) = S_t 1$ is a unique solution of the initial-value problem

$$\begin{cases} \dfrac{du}{dt} = \mathfrak{A}_N u \quad \text{for all } t > 0, \\ u(0) = 1 \end{cases} \tag{$*$}$$

which satisfies the following three conditions (a)–(c):

(a) The function $u(t)$ is continuously differentiable for all $t > 0$.
(b) $\|u(t)\| \leq 1$ for all $t \geq 0$.
(c) $u(t) \to 1$ as $t \downarrow 0$.

However, it is easy to see that the function $u(t) \equiv 1$ is also a solution of problem ($*$) which satisfies three conditions (a)–(c).

Therefore, it follows from the uniqueness theorem for problem ($*$) that

$$S_t 1 = 1 \quad \text{for all } t > 0.$$

In view of formula (12.22), we obtain the desired assertion (12.23) as follows:

$$1 = S_t 1(x) = \int_{\overline{D}} p_t(x, dy) = p_t(x, \overline{D}) \quad \text{for each } x \in \overline{D}.$$

The proof of Proposition 12.13 is complete.

12.4.2 The Robin Case

Secondly, we consider the *Robin boundary condition*

$$\mathcal{B}_R u = a(x')\frac{\partial u}{\partial \mathbf{n}} + b(x')u = 0 \quad \text{on } \partial D, \tag{12.24}$$

where

$$\begin{cases} a \in C^\infty(\partial D), & a(x') > 0 \quad \text{on } \partial D, \\ b \in C^\infty(\partial D), & b(x') \leq 0 \quad \text{on } \partial D. \end{cases}$$

We introduce a linear operator $\mathfrak{A}_R$ as follows:

(a) $D(\mathfrak{A}_R) = \left\{ u \in C(\overline{D}) : \Delta u \in C(\overline{D}),\ a(x')\frac{\partial u}{\partial \mathbf{n}} + b(x')u = 0 \text{ on } \partial D \right\}$, where Δu is taken in the sense of *distributions*.

(b) $\mathfrak{A}_R f = \Delta f$ for every $f \in D(\mathfrak{A}_R)$.

However, since $a(x') > 0$ on ∂D, by letting

$$\tilde{b}(x') = \frac{b(x')}{a(x')}$$

we find that the boundary condition (12.24) is equivalent to the following boundary condition:

$$\frac{\partial f}{\partial \mathbf{n}} + \tilde{b}(x') f = 0 \quad \text{on } \partial D.$$

In other words, without loss of generality we may assume that

$$\begin{aligned} a(x') &= 1 \quad \text{on } \partial D, \\ \tilde{b}(x') &\leq 0 \quad \text{on } \partial D. \end{aligned}$$

Then it follows from an application of Theorem 1.2 with $L := \mathcal{B}_R$ that the operator $\mathfrak{A}_R$ is the infinitesimal generator of a Feller semigroup $\{T_t\}_{t \geq 0}$ with Robin boundary condition. Let $\{y_t(\omega)\}$ be the strong Markov process corresponding to the Feller semigroup $\{T_t\}_{t \geq 0}$. In this subsection we study the path-continuity of the Markov process $\{y_t(\omega)\}$. To do this, we shall make use of Theorem 12.10.

(1) The case where x_0 is an arbitrary (interior) point of D: By applying Theorem [Ta10, Theorem 1.5] with $\mu(x') := 1$ and $\gamma(x') := \tilde{b}(x')$, we can find a function $\phi \in C^\infty(\overline{D})$ such that

$$\begin{cases} \Delta\psi = -1 & \text{in } D, \\ \dfrac{\partial\psi}{\partial\mathbf{n}} + \tilde{b}(x')\psi = 0 & \text{on } \partial D. \end{cases} \tag{12.25}$$

Then we have the following lemma:

Lemma 12.14. *The function ψ satisfies the condition*

$$\psi(x) > 0 \quad \textit{on } \overline{D}.$$

Proof. We assume, to the contrary, that

$$\min_{\overline{D}} \psi \leq 0.$$

(1) First, we consider the case: There exists a point $x_0 \in D$ such that

$$\psi(x_0) = \min_{\overline{D}} \psi \leq 0.$$

Since we have, condition (12.25),

$$\Delta(-\psi) = 1 > 0 \quad \text{in } D, \tag{12.26}$$

it follows from an application of the strong maximum principle (Theorem 8.13) with $W := \Delta$ and $u := -\psi$ that

$$\psi(x) \equiv \psi(x_0) \quad \text{in } D.$$

Hence we have

$$\Delta(-\psi) = 0 \quad \text{in } D.$$

This contradicts condition (12.26).

(2) Next we consider the case: There exists a point $x_0' \in \partial D$ such that

$$\psi(x_0') = \min_{\overline{D}} \psi \leq 0.$$

Then, by applying the boundary point lemma (Theorem 8.15) with $W := \Delta$ to the function $u := -\psi$ we obtain from inequality (12.26) that

$$\frac{\partial\psi}{\partial\mathbf{n}}(x_0') > 0.$$

This contradicts the boundary condition

$$0 = \frac{\partial \psi}{\partial \mathbf{n}}(x_0') + \tilde{b}(x_0')\psi(x_0') \geq \frac{\partial \psi}{\partial \mathbf{n}}(x_0') > 0 \quad \text{on } \partial D,$$

since $\tilde{b}(x_0') \leq 0$ on ∂D.

The proof of Lemma 12.14 is complete.

Now we choose a real-valued, smooth function $\theta \in C^\infty(\mathbf{R})$ that satisfies condition (12.21) (see Fig. 12.8), and define a function $f \in C^\infty(\overline{D})$ by the formula

$$f(x) = \left(1 - \theta\left(\frac{|x - x_0|}{\varepsilon}\right)\right)\psi(x),$$

where

$$0 < \varepsilon < \text{dist}\,(x, \partial D).$$

Then, in view of Lemma 12.14 it is easy to verify that $f \in D(\mathfrak{A}_R)$ and satisfies conditions (i)–(iii) of Theorem 12.10.

(2) The case where x_0 is an arbitrary (boundary) point of ∂D: By change of coordinates, we may assume that

$$x_0 = (0', 0) = (0, \ldots, 0, 0),$$
$$\mathbf{n} = (0', 1) = (0, \ldots, 0, 1) \in \mathbf{R}^N.$$

If we define a function $f \in C^\infty(\overline{D})$ by the formula

$$f(x', x_N) = \left(1 - \theta\left(\frac{x_N}{\varepsilon}\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\psi(x', x_N),$$

it follows from conditions (12.21) and (12.25) that

$$\begin{aligned}
&\frac{\partial f}{\partial \mathbf{n}} + \tilde{b}(x')f \\
&= \frac{\partial}{\partial x_N}\left(1 - \theta\left(\frac{x_N}{\varepsilon}\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\psi(x', x_N)\Big|_{x_N=0} \\
&\quad + \tilde{b}(x')\left(1 - \theta\left(\frac{x_N}{\varepsilon}\right)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\psi\Big|_{x_N=0} \\
&= -\frac{1}{\varepsilon}\theta'(0)\theta\left(\frac{|x'|}{\varepsilon}\right)\cdot\psi(x', 0) + \left(1 - \theta(0)\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\frac{\partial \psi}{\partial x_N}(x', x_N)\Big|_{x_N=0} \\
&\quad + \tilde{b}(x')\left(1 - \theta\,(0)\,\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\psi(x', 0)
\end{aligned}$$

$$
\begin{aligned}
&= \left(1-\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\frac{\partial\psi}{\partial x_N}(x',x_N)\bigg|_{x_N=0} + \left(1-\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\tilde{b}(x')\psi(x',0) \\
&= \left(1-\theta\left(\frac{|x'|}{\varepsilon}\right)\right)\left(\frac{\partial\psi}{\partial\mathbf{n}} + \tilde{b}(x')\psi\right) \\
&= 0 \quad \text{on } \partial D.
\end{aligned}
$$

This proves that $f \in D(\mathfrak{A}_R)$. Moreover, we have the following three assertions (a)–(c) (see Fig. 12.9):

(a) Since $f(x) = 0$ in the $\varepsilon/2$ neighborhood $U_{\varepsilon/2}(x_0)$ of x_0, it follows from Lemma 12.14 that

$$f(y) = \mathfrak{A}_R f(y) = 0 \quad \text{for all } y \in U_{\varepsilon/2}(x_0).$$

This proves that condition (iii) of Theorem 12.10 is satisfied.

(b) It is clear that $f(x) \geq 0$ on $\overline{D}$, so that condition (i) in Theorem 12.10 is satisfied.

(c) Since $f(y) = \psi(y)$ outside the $\sqrt{N}\,\varepsilon$ neighborhood $U_{\sqrt{N}\,\varepsilon}(x_0)$ of x_0, it follows from Lemma 12.11 that

$$f(y) > 0 \quad \text{for all } y \in \overline{D} \setminus U_{\sqrt{N}\,\varepsilon}(x_0).$$

This proves that condition (ii) of Theorem 12.10 is satisfied.

Therefore, by applying Theorem 12.10 to the operator $\mathfrak{A}_R$ we obtain the following generalization of Example 9.16 to the multi-dimensional case:

$$\boxed{\text{Lemma 12.14}} \Longrightarrow \boxed{\text{Theorem 12.10}} \Longrightarrow \boxed{\text{Theorem 12.15}}$$

Theorem 12.15. *The strong Markov process $\{y_t(\omega)\}$ associated with the Feller semigroup $\{T_t\}_{t\geq 0}$ enjoys the property*

$$P\left(\{\omega :\ y_t(\omega) \text{ is continuous for all } t \geq 0\}\right) = 1.$$

12.4.3 The Oblique Derivative Case

Finally, we consider the *oblique derivative boundary condition*

$$\mathcal{B}_O u = \frac{\partial u}{\partial \mathbf{n}} + \beta(x')u + b(x')u = 0 \quad \text{on } \partial D,$$

where $\mathbf{n}$ is the unit inner normal vector to the boundary ∂D,

$$b \in C^\infty(\partial D), \quad b(x') \leq 0 \ \text{ on } \partial D,$$

and

$$\beta(x')u := \sum_{i=1}^{N-1} \beta_i(x') \frac{\partial u}{\partial x_i}$$

is a tangent vector field of class C^1 on the boundary ∂D.

We introduce a linear operator $\mathfrak{A}_O$ as follows:

(a) The domain $D(\mathfrak{A}_O)$ is the space

$$D(\mathfrak{A}_O) = \left\{ u \in C(\overline{D}) : \Delta u \in C(\overline{D}),\ \frac{\partial u}{\partial \mathbf{n}} + \beta(x')u + b(x')u = 0 \text{ on } \partial D \right\}.$$

Here Δu is taken in the sense of *distributions*.

(b) $\mathfrak{A}_O u = \Delta u$ for every $u \in D(\mathfrak{A}_O)$.

Then it follows from an application of Theorem 1.2 with $L := \mathcal{B}_O$ that the operator $\mathfrak{A}_O$ is the infinitesimal generator of a Feller semigroup $\{U_t\}_{t\geq 0}$ with oblique derivative boundary condition. Let $\{z_t(\omega)\}$ be the strong Markov process corresponding to the Feller semigroup $\{U_t\}_{t\geq 0}$. In this subsection we study the path-continuity of the Markov process $\{z_t(\omega)\}$. To do this, we shall make use of Theorem 12.10.

If we introduce a vector field ℓ by the formula

$$\ell = \mathbf{n} + \beta,$$

then the oblique derivative boundary condition

$$\mathcal{B}_O u = \frac{\partial u}{\partial \mathbf{n}} + \beta(x')u + b(x')u = 0 \quad \text{on } \partial D$$

can be written in the form (see Fig. 12.11)

$$\frac{\partial u}{\partial \ell} + b(x')u = 0 \quad \text{on } \partial D.$$

Furthermore, for each initial point $y' = (y_1, y_2, \ldots, y_{N-1}) \in \mathbf{R}^{N-1}$ we consider the following initial-value problem for ordinary differential equations:

$$\begin{cases} \dfrac{d\gamma_1}{dt} = \beta_1(\gamma(y',t)), & \gamma_1(y',0) = y_1, \\ \vdots & \vdots \\ \dfrac{d\gamma_{N-1}}{dt} = \beta_{N-1}(\gamma(y',t)), & \gamma_{N-1}(y',0) = y_{N-1}, \\ \dfrac{d\gamma_N}{dt} = 1, & \gamma_N(y',0) = 0. \end{cases} \tag{12.27}$$

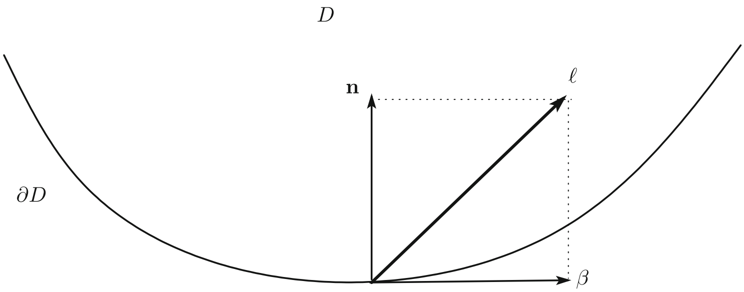


Fig. 12.11 The vector field $\ell = \mathbf{n} + \beta$

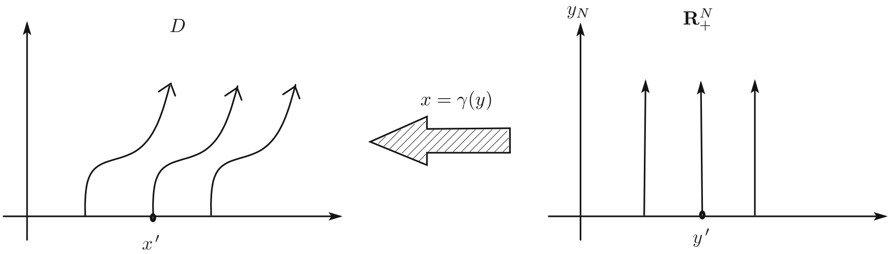


Fig. 12.12 The change of variables $x = \gamma(y)$

We remark that the initial-value problem (12.27) has a unique local solution

$$\gamma(y', t) = (\gamma_1(y', t), \gamma_2(y', t), \dots, \gamma_N(y', t)),$$

since the vector field $\beta(x')$ is Lipschitz continuous. Hence we can introduce a new change of variables by the formula

$$(x_1, x_2, \dots, x_N) = (\gamma_1(y', y_N), \gamma_2(y', y_N), \dots, \gamma_N(y', y_N)),$$
$$y = (y', y_N) \in \mathbf{R}^N.$$

The situation may be represented schematically by Fig. 12.12.

Then it follows that the vector field ℓ can be simplified as follows:

$$\frac{\partial}{\partial y_N} = \sum_{j=1}^{N-1} \frac{\partial x_j}{\partial y_N} \frac{\partial}{\partial x_j} + \frac{\partial x_N}{\partial y_N} \frac{\partial}{\partial x_N} = \sum_{j=1}^{N-1} \beta_j(x) \frac{\partial}{\partial x_j} + \frac{\partial}{\partial x_N} = \frac{\partial}{\partial \ell}.$$

Hence we obtain that

$$\frac{\partial u}{\partial \mathbf{n}} + \beta(x')u + b(x')u = \frac{\partial u}{\partial \ell} + b(x')u = \left(\frac{\partial \tilde{u}}{\partial y_N} + \tilde{b}(y')\tilde{u} \right)\Bigg|_{y_N=0},$$

where

$$\tilde{u}(y) = u(\gamma(y)) = u(x),$$
$$\tilde{b}(y') = b\left(\gamma(y')\right) = b(x').$$

In this way, we are reduced to the Robin boundary condition case in the half-space $\mathbf{R}^N_+$:

$$\mathcal{B}_R \tilde{u} = \frac{\partial \tilde{u}}{\partial y_N} + \tilde{b}(y')\tilde{u} = 0 \quad \text{on } \partial\mathbf{R}^N_+.$$

Therefore, by applying Theorem 12.10 to the operator $\mathfrak{A}_O$ we obtain the following theorem:

$$\boxed{\text{Theorem 12.15}} \Longrightarrow \boxed{\text{Theorem 12.16}}$$

Theorem 12.16. *The strong Markov process $\{z_t(\omega)\}$ associated with the Feller semigroup $\{U_t\}_{t\geq 0}$ enjoys the property*

$$P\left(\{\omega : z_t(\omega) \text{ is continuous for all } t \geq 0\}\right) = 1.$$

12.5 Notes and Comments

The results discussed in this chapter are adapted from Dynkin [Dy2], Kinney [Ki], Lamperti [La] and Ray [Ra].

Section 12.1: Theorem 12.4 is due to Lamperti [La, Chapter 8, Section 2] and Theorem 12.6 is due to Lamperti [La, Section 8.3, Theorem 1], respectively.

Section 12.2: Theorem 12.7 is adapted from Lamperti [La, Section 8.3, Theorem 2] and Corollary 12.8 is adapted from Dynkin [Dy2, Theorem 3.5] (see also [La, Section 8.3, Corollary]).

Section 12.3: Theorem 12.10 is taken from Dynkin [Dy2, Theorem 3.9] and Lamperti [La, Section 8.3, Theorem 3].

Section 12.4: Theorem 12.12 (the Neumann case), Theorem 12.15 (the Robin case) and Theorem 12.16 (the oblique derivative case) may be new.

Chapter 13
Concluding Remarks

In this final chapter we summarize the contents of the first edition of the present monograph "Semigroups, boundary value problems and Markov processes" which was published in 2004. In Sect. 13.1 we study a class of *degenerate* boundary value problems for second-order elliptic differential operators which includes as particular cases the Dirichlet and Robin problems. We state existence and uniqueness theorems for this class of degenerate elliptic boundary value problems (Theorems 13.1 and 13.2). The crucial point is how to define modified boundary spaces $B_{L_0}^{1-1/p,p}(\partial D)$ and $C_{L_0}^{1+\theta}(\partial D)$ in which our boundary value problems are uniquely solvable. The purpose of Sect. 13.2 is to study our degenerate elliptic boundary value problems from the viewpoint of the theory of analytic semigroups, and is to generalize generation theorems for *analytic semigroups* both in the L^p topology and in the topology of uniform convergence (Theorems 13.3 and 13.4). As an application, we state generation theorems for Feller semigroups corresponding to the diffusion phenomenon where a Markovian particle moves continuously until it "dies" at the time when it reaches the set where the particle is definitely absorbed (Theorem 13.5). In Sect. 13.3 we assume that the domain D is *convex*, and extend the existence and uniqueness theorems for degenerate elliptic boundary value problems in Sect. 13.1 (Theorems 13.6) and the generation theorems for analytic and Feller semigroups in Sect. 13.2 to the *integro-differential operator* case (Theorems 13.7, 13.9 and 13.10). Due to the non-local character of integro-differential operators, we are forced to impose various conditions on the structure of jumps of Markovian particles such as the *moment condition*. Moreover, in order to remove a singularity of solutions at the boundary ∂D, we impose the condition that no jumps outside the closure $\overline{D}$ are allowed.

The following diagram gives a bird's eye view of existence and uniqueness theorems for degenerate elliptic boundary value problems and generation theorems for analytic and Feller semigroups in Chap. 13 proved by the author ([Ta7], [Ta8] and [Ta9]) using the L^p theory of pseudo-differential operators:

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7_13,

Waldenfels operator W	Theorems of Chapter 13	proved by
differential operator $W = A$	Theorem 13.1 Theorem 13.2	[Ta7, Theorem 1] [Ta8, Theorem 1.1]
differential operator $W = A$	Theorem 13.3 Theorem 13.4 Theorem 13.5	[Ta9, Theorem 1.2] [Ta9, Theorem 1.3] [Ta9, Theorem 1.4]
integro-differential operator $W = A + S_r$	Theorem 13.6 Theorem 13.7	[Ta8, Theorem 1] [Ta8, Theorem 2]
integro-differential operator $W = A + S_r$	Theorem 13.9 Theorem 13.10	[Ta8, Theorem 3] [Ta8, Theorem 4]

The second edition [Ta9] of "Boundary value problems and Markov processes" may be considered as a short introduction to this augmented second edition.

13.1 Existence and Uniqueness Theorems for Boundary Value Problems

Let D be a bounded domain of Euclidean space $\mathbf{R}^N$ with smooth boundary ∂D; its closure $\overline{D} = D \cup \partial D$ is an N-dimensional compact smooth manifold with boundary (see Fig. 1.1). We let

$$Au(x) = \sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x)$$

be a second-order *elliptic* differential operator with real coefficients such that

(1) $a^{ij}(x) \in C^\infty(\overline{D})$, $a^{ij}(x) = a^{ji}(x)$ for all $x \in \overline{D}$ and $1 \leq i, j \leq N$, and there exists a constant $a_0 > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i \xi_j \geq a_0 |\xi|^2 \quad \text{for all } (x, \xi) \in \overline{D} \times \mathbf{R}^N.$$

(2) $b^i(x) \in C^\infty(\overline{D})$ for all $1 \leq i \leq N$.
(3) $c(x) \in C^\infty(\overline{D})$ and $c(x) \leq 0$ on $\overline{D}$, but $c(x) \not\equiv 0$ in D.

The differential operator A is called a diffusion operator which describes analytically a strong Markov process with continuous paths (diffusion process) in the interior D (see Fig. 13.1).

We let

$$L_0 u(x') = \mu(x') \frac{\partial u}{\partial \mathbf{n}}(x') + \gamma(x')u(x')$$

be a first-order boundary condition with real coefficients such that

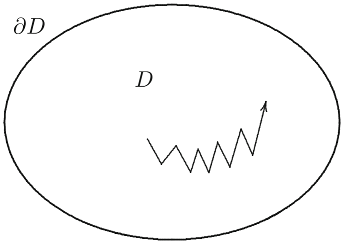


Fig. 13.1 A Markovian particle moves continuously

(1) $\mu(x') \in C^\infty(\partial D)$ and $\mu(x') \geq 0$ on ∂D.
(2) $\gamma(x') \in C^\infty(\partial D)$ and $\gamma(x') \leq 0$ on ∂D.
(3) $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ is the unit interior normal to the boundary ∂D.

Now we consider the following boundary value problem: Given functions f and φ defined in D and on ∂D, respectively, find a function u in D such that

$$\begin{cases} Au = f & \text{in } D, \\ L_0 u = \mu(x') \dfrac{\partial u}{\partial \mathbf{n}} + \gamma(x') u = \varphi & \text{on } \partial D. \end{cases} \tag{13.1}$$

It should be noted that if $\mu(x') \equiv 0$ and $\gamma(x') \equiv -1$ on ∂D (resp. $\mu(x') > 0$ on ∂D), then the boundary condition L_0 is the so-called Dirichlet (resp. Robin) condition. It is easy to see that problem (13.1) is non-degenerate (or coercive) if and only if either $\mu(x') > 0$ on ∂D or $\mu(x') \equiv 0$ and $\gamma(x') < 0$ on ∂D. The generation theorem for analytic semigroups is well established in the non-degenerate case both in the L^p topology (see Friedman [Fr1], Tanabe [Tn]) and in the topology of uniform convergence (see Masuda [Ma] for the Dirichlet case).

The fundamental hypothesis is the following condition (FH):

(FH) $\mu(x') - \gamma(x') > 0$ on ∂D.

The intuitive meaning of condition (FH) is that either reflection or absorption occurs at each point of the boundary ∂D. More precisely, condition (FH) implies that absorption occurs at each point of the set

$$M = \{x' \in \partial D : \mu(x') = 0\},$$

while reflection occurs at each point of the set $\partial D \setminus M = \{x' \in \partial D : \mu(x') > 0\}$. In other words, a Markovian particle moves continuously in the space $\overline{D} \setminus M$ until it "dies" when it reaches the set M (see Fig. 13.2).

The first purpose of the first edition is to prove existence and uniqueness theorems for problem (13.1) in the framework of L^p spaces and Hölder spaces. In the case where $N = 3$, $\mu(x')$ may be a function such that, in terms of local coordinates (x_1, x_2) of ∂D,

$$\mu(x') = e^{-1/x_1^2} \sin^2 \frac{1}{x_1}\, e^{-1/x_2^2} \sin^2 \frac{1}{x_2}.$$

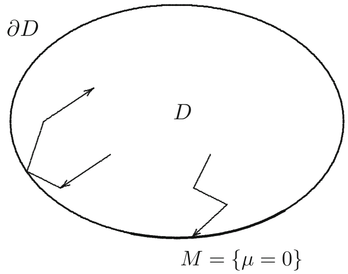


Fig. 13.2 A Markovian particle dies when it reaches the set M

Therefore, the crucial point in our approach is how to define function subspaces in which the boundary value problem (13.1) is uniquely solvable.

If k is a positive integer and $1 < p < \infty$, we define the Sobolev space of L^p type

$$\begin{aligned} H^{k,p}(D) = &\text{ the space of (equivalence classes of) functions} \\ &u \in L^p(D) \text{ whose derivatives } D^\alpha u, |\alpha| \leq k, \\ &\text{in the sense of distributions are all in } L^p(D), \end{aligned}$$

and the space

$$\begin{aligned} &B^{k-1/p,p}(\partial D) \\ &= \text{the space of the boundary values } \varphi \text{ of functions } u \in H^{k,p}(D). \end{aligned}$$

In the space $B^{k-1/p,p}(\partial D)$, we introduce a norm

$$|\varphi|_{B^{k-1/p,p}(\partial D)} = \inf \left\{ \|u\|_{H^{k,p}(D)} : u \in H^{k,p}(D), \ u|_{\partial D} = \varphi \right\}.$$

The space $B^{k-1/p,p}(\partial D)$ is a Banach space with respect to the norm $|\cdot|_{B^{k-1/p,p}(\partial D)}$. More precisely, it is a Besov space (see Bergh–Löfström [BL], Taibleson [Tb], Triebel [Tr]).

We introduce a subspace of $B^{k-1/p,p}(\partial D)$ which is associated with the boundary condition L_0 in the following way: We let

$$\begin{aligned} &B_{L_0}^{1-1/p,p}(\partial D) \\ &= \left\{ \varphi = \mu(x')\varphi_1 - \gamma(x')\varphi_2 : \varphi_1 \in B^{1-1/p,p}(\partial D), \ \varphi_2 \in B^{2-1/p,p}(\partial D) \right\}, \end{aligned}$$

and define a norm

$$|\varphi|_{B^{1-1/p,p}_{L_0}(\partial D)} = \inf\left\{|\varphi_1|_{B^{1-1/p,p}(\partial D)} + |\varphi_2|_{B^{2-1/p,p}(\partial D)} : \varphi = \mu(x')\varphi_1 - \gamma(x')\varphi_2\right\}.$$

Then it is easy to verify (see [Ta7, Lemma 4.7]) that the space $B^{1-1/p,p}_{L_0}(\partial D)$ is a Banach space with respect to this norm $|\cdot|_{B^{1-1/p,p}_{L_0}(\partial D)}$. It should be noted that the space $B^{1-1/p,p}_{L_0}(\partial D)$ is an "interpolation space" between the Besov spaces $B^{2-1/p,p}(\partial D)$ and $B^{1-1/p,p}(\partial D)$. More precisely, we have

$$\begin{cases} B^{1-1/p,p}_{L_0}(\partial D) = B^{2-1/p,p}(\partial D) & \text{if } \mu(x') \equiv 0 \text{ on } \partial D, \\ B^{1-1/p,p}_{L_0}(\partial D) = B^{1-1/p,p}(\partial D) & \text{if } \mu(x') > 0 \text{ on } \partial D. \end{cases}$$

Now we can state an existence and uniqueness theorem for the boundary value problem (13.1) in the framework of L^p spaces ([Ta7, Theorem 1], [RF, Theorem 1.1]):

Theorem 13.1. *Let* $1 < p < \infty$. *Assume that condition (FH) is satisfied. Then the mapping*

$$(A, L_0) : H^{2,p}(D) \longrightarrow L^p(D) \bigoplus B^{1-1/p,p}_{L_0}(\partial D)$$

is an algebraic and topological isomorphism. *In particular, for any* $f \in L^p(D)$ *and any* $\varphi \in B^{1-1/p,p}_{L_0}(\partial D)$, *there exists a unique solution* $u \in H^{2,p}(D)$ *of problem* (13.1).

Furthermore, in order to study problem (13.1) in the framework of Hölder spaces, we introduce a subspace of $C^{1+\theta}(\partial D)$, $0 < \theta < 1$, which is a Hölder space version of $B^{1-1/p,p}_{L_0}(\partial D)$. We let

$$C^{1+\theta}_{L_0}(\partial D) = \left\{\varphi = \mu(x')\varphi_1 - \gamma(x')\varphi_2 : \varphi_1 \in C^{1+\theta}(\partial D),\ \varphi_2 \in C^{2+\theta}(\partial D)\right\},$$

and define a norm

$$|\varphi|_{C^{1+\theta}_{L_0}(\partial D)} = \inf\left\{|\varphi_1|_{C^{1+\theta}(\partial D)} + |\varphi_2|_{C^{2+\theta}(\partial D)} : \varphi = \mu(x')\varphi_1 - \gamma(x')\varphi_2\right\}.$$

Then it is easy to verify (see [Ta7, Lemma 4.7]) that the space $C^{1+\theta}_{L_0}(\partial D)$ is a Banach space with respect to the norm $|\cdot|_{C^{1+\theta}_{L_0}(\partial D)}$. We remark that

$$\begin{cases} C^{1+\theta}_{L_0}(\partial D) = C^{2+\theta}(\partial D) & \text{if } \mu(x') \equiv 0 \text{ on } \partial D, \\ C^{1+\theta}_{L_0}(\partial D) = C^{1+\theta}(\partial D) & \text{if } \mu(x') > 0 \text{ on } \partial D. \end{cases}$$

The next theorem is a Hölder space version of Theorem 13.1 [Ta8, Theorem 1.1]:

Theorem 13.2. *Let* $0 < \theta < 1$. *If condition (FH) is satisfied, then the mapping*

$$(A, L_0) : C^{2+\theta}(\overline{D}) \longrightarrow C^{\theta}(\overline{D}) \bigoplus C^{1+\theta}_{L_0}(\partial D)$$

is an algebraic and topological isomorphism. *In particular, for any* $f \in C^{\theta}(\overline{D})$ *and any* $\varphi \in C^{1+\theta}_{L_0}(\partial D)$, *there exists a unique solution* $u \in C^{2+\theta}(\overline{D})$ *of problem* (13.1).

13.2 Generation Theorems for Analytic Semigroups on a Bounded Domain

The second purpose of the first edition is to study the boundary value problem (13.1) from the point of view of analytic semigroup theory in functional analysis, and to generalize the generation theorem for analytic semigroups to the *degenerate* case.

We associate with problem (13.1) an unbounded linear operator A_p from $L^p(D)$ into itself as follows:

(a) The domain of definition $D(A_p)$ of A_p is the set

$$D(A_p) = \left\{ u \in H^{2,p}(D) : L_0 u = \mu(x') \frac{\partial u}{\partial \mathbf{n}} + \gamma(x') u = 0 \text{ on } \partial D \right\}.$$

(b) $A_p u = Au$ for every $u \in D(A_p)$.

Then we can prove a generation theorem for analytic semigroups in the framework of L^p spaces [Ta9, Theorem 1.2]:

Theorem 13.3. *Let* $1 < p < \infty$. *If condition (FH) is satisfied, then we have the following two assertions (i) and (ii):*

(i) For every $\varepsilon > 0$, *there exists a constant* $r_p(\varepsilon) > 0$ *such that the resolvent set of* A_p *contains the set*

$$\Sigma_p(\varepsilon) = \left\{ \lambda = r^2 e^{i\vartheta} : r \geq r_p(\varepsilon),\ -\pi + \varepsilon \leq \vartheta \leq \pi - \varepsilon \right\},$$

and that the resolvent $(A_p - \lambda I)^{-1}$ *satisfies the estimate*

$$\left\| (A_p - \lambda I)^{-1} \right\| \leq \frac{c_p(\varepsilon)}{|\lambda|} \quad \text{for all } \lambda \in \Sigma_p(\varepsilon), \tag{13.2}$$

where $c_p(\varepsilon) > 0$ *is a constant depending on* ε.

(ii) The operator A_p *generates a semigroup* e^{zA_p} *on the space* $L^p(D)$ *which is* analytic *in the sector*

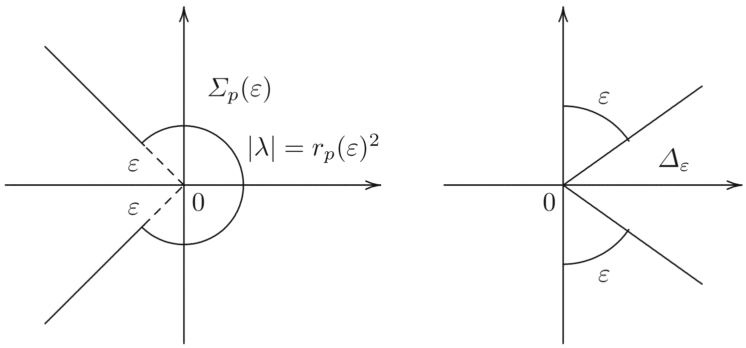


Fig. 13.3 The set $\Sigma_p(\varepsilon)$ and the sector Δ_ε

$$\Delta_\varepsilon = \{z = t + is : z \neq 0, |\arg z| < \pi/2 - \varepsilon\}$$

for any $0 < \varepsilon < \pi/2$ *(see Fig. 13.3).*

It should be noted that Theorem 13.3 for $p = 2$ is proved by Taira [Ta1, Theorem 1].

Secondly, we state a generation theorem for analytic semigroups in the topology of uniform convergence.

Let $C(\overline{D})$ be the space of real-valued, continuous functions f on $\overline{D}$. We equip the space $C(\overline{D})$ with the topology of uniform convergence on the whole $\overline{D}$; hence it is a Banach space with the maximum norm

$$\|f\|_\infty = \max_{x \in \overline{D}} |f(x)|.$$

We introduce a subspace of $C(\overline{D})$ which is associated with the boundary condition L_0. We remark that the boundary condition

$$L_0 u = \mu(x') \frac{\partial u}{\partial \mathbf{n}} + \gamma(x') u = 0 \quad \text{on } \partial D$$

includes the condition

$$u = 0 \quad \text{on } M = \{x' \in \partial D : \mu(x') = 0\},$$

if $\gamma(x') \neq 0$ on M. With this fact in mind, we let

$$C_0(\overline{D} \setminus M) = \left\{u \in C(\overline{D}) : u = 0 \text{ on } M\right\}.$$

The space $C_0(\overline{D} \setminus M)$ is a closed subspace of $C(\overline{D})$; hence it is a Banach space. Furthermore, we introduce a linear operator $\mathfrak{A}$ from $C_0(\overline{D} \setminus M)$ into itself as follows:

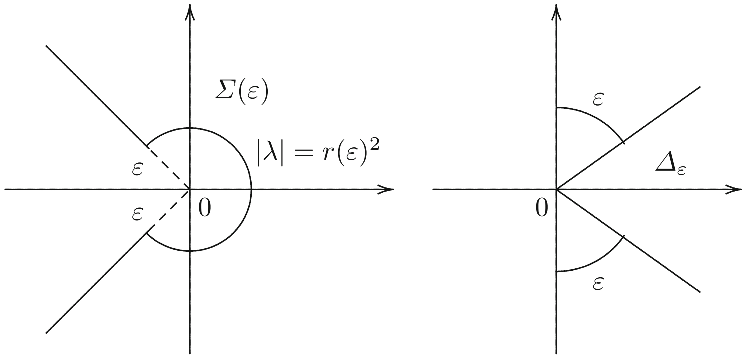


Fig. 13.4 The set $\Sigma(\varepsilon)$ and the sector Δ_ε

(a) The domain of definition $D(\mathfrak{A})$ of $\mathfrak{A}$ is the set

$$D(\mathfrak{A}) = \left\{u \in C_0(\overline{D} \setminus M) : Au \in C_0(\overline{D} \setminus M),\ L_0 u = 0 \text{ on } \partial D\right\}. \quad (13.3)$$

(b) $\mathfrak{A}u = Au$ for every $u \in D(\mathfrak{A})$.

Here Au and $L_0 u$ are taken in the sense of *distributions*.

Then Theorem 13.3 remains valid with $L^p(D)$ and A_p replaced by $C_0(\overline{D} \setminus M)$ and $\mathfrak{A}$, respectively. More precisely, we can prove the following theorem [Ta9, Theorem 1.3]:

Theorem 13.4. *If condition (FH) is satisfied, then we have the following two assertions (i) and (ii):*

(i) For every $\varepsilon > 0$*, there exists a constant* $r(\varepsilon) > 0$ *such that the resolvent set of* $\mathfrak{A}$ *contains the set*

$$\Sigma(\varepsilon) = \left\{\lambda = r^2 e^{i\vartheta} : r \geq r(\varepsilon),\ -\pi + \varepsilon \leq \vartheta \leq \pi - \varepsilon\right\},$$

and such that the resolvent $(\mathfrak{A} - \lambda I)^{-1}$ *satisfies the estimate*

$$\|(\mathfrak{A} - \lambda I)^{-1}\| \leq \frac{c(\varepsilon)}{|\lambda|} \quad \text{for all } \lambda \in \Sigma(\varepsilon), \quad (13.4)$$

where $c(\varepsilon) > 0$ *is a constant depending on* ε.

(ii) The operator $\mathfrak{A}$ *generates a semigroup* $e^{z\mathfrak{A}}$ *on the space* $C_0(\overline{D} \setminus M)$ *which is* analytic *in the sector*

$$\Delta_\varepsilon = \{z = t + is : z \neq 0, |\arg z| < \pi/2 - \varepsilon\}$$

for any $0 < \varepsilon < \pi/2$ *(see Fig. 13.4).*

Moreover, the operators $\{e^{t\mathfrak{A}}\}_{t\geq 0}$ *are non-negative and contractive on the space* $C_0(\overline{D} \setminus M)$*:*

$$f \in C_0(\overline{D}\setminus M),\ 0 \leq f(x) \leq 1 \ \textit{on}\ \overline{D} \setminus M \Longrightarrow 0 \leq e^{t\mathfrak{A}} f(x) \leq 1 \ \textit{on}\ \overline{D} \setminus M.$$

Theorems 13.3 and 13.4 express a *regularizing effect* for the parabolic differential operator $\partial/\partial t - A$ with homogeneous boundary condition L_0.

Furthermore, we can reformulate part (ii) of Theorem 13.4 as follows [Ta9, Theorem 1.4]:

Theorem 13.5. *If condition (FH) is satisfied, then the operator* $\mathfrak{A}$ *generates a Feller semigroup* $\{T_t\}_{t\geq 0}$ *on the space* $\overline{D} \setminus M$.

Theorem 13.5 generalizes Bony–Courrège–Priouret [BCP, Théorème XIX] to the case where $\mu(x') \geq 0$ on the boundary ∂D (cf. [Ta5, Theorem 10.1.3]).

13.3 The Integro-Differential Operator Case

More generally, it should be emphasized that Theorems 13.2–13.5 may be extended to the *integro-differential operator* case. For simplicity, we assume that the domain D is *convex*. Let W be a second-order *elliptic* integro-differential operator with real coefficients such that

$$\begin{aligned} Wu(x) &= Au(x) + S_r u(x) \\ &:= \left(\sum_{i,j=1}^{N} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{N} b^i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x) \right) \\ &\quad + \int_{\mathbf{R}^N\setminus\{0\}} \left(u(x+z) - u(x) - \sum_{j=1}^{N} z_j \frac{\partial u}{\partial x_j}(x) \right) s(x,z)\, m(dz). \end{aligned}$$

Here:

(1) $a^{ij}(x) \in C^\infty(\overline{D})$, $a^{ij}(x) = a^{ji}(x)$ for all $x \in \overline{D}$ and $1 \leq i, j \leq N$, and there exists a constant $a_0 > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i\xi_j \geq a_0|\xi|^2 \quad \text{for all } (x,\xi) \in \overline{D} \times \mathbf{R}^N.$$

(2) $b^i(x) \in C^\infty(\overline{D})$ for all $1 \leq i \leq N$.

(3) $c(x) \in C^\infty(\overline{D})$ and $c(x) \leq 0$ in D, but $c(x) \not\equiv 0$ in D.

(4) $s(x,z) \in C(\overline{D} \times \mathbf{R}^N)$ and $0 \leq s(x,z) \leq 1$ on $\overline{D} \times \mathbf{R}^N$, and there exist constants $C_0 > 0$ and $0 < \theta < 1$ such that

$$|s(x,z) - s(y,z)| \leq C_0 |x-y|^\theta, \quad x,\ y \in \overline{D},\ z \in \mathbf{R}^N, \tag{13.5}$$

and

$$s(x,z) = 0 \quad \text{if } x + z \notin \overline{D}. \tag{13.6}$$

Condition (13.6) implies that the integral operator S_r may be considered as an operator acting on functions u defined on the closure $\overline{D}$ (see Garroni–Menaldi [GM2, Chapter 2, Section 2.3]).

(5) The measure $m(dz)$ is a Radon measure on the space $\mathbf{R}^N \setminus \{0\}$ which satisfies the *moment condition*

$$\int_{\{|z| \leq 1\}} |z|^2 \, m(dz) + \int_{\{|z|>1\}} |z| \, m(dz) < \infty. \tag{13.7}$$

Condition (13.7) is a standard condition for the measure $m(dz)$ (see Garroni–Menaldi [GM2, Chapter 2, Section 2.1]).

The operator $W = A + S_r$ is called a second-order *Waldenfels integro-differential operator* or simply a *Waldenfels operator* (cf. [BCP, Wa]). The integro-differential operator S_r is called a second-order *Lévy operator* which corresponds to the jump phenomenon in the closure $\overline{D}$ (see [St]). In this context, condition (13.6) implies that any Markovian particle does not move by jumps from the interior D into the outside of $\overline{D}$. On the other hand, condition (13.7) imposes various conditions on the structure of jumps for the Lévy operator S_r. More precisely, the condition

$$\int_{\{|z| \leq 1\}} |z|^2 \, m(dz) < \infty$$

implies that the measure $m(\cdot)$ admits a singularity of order 2 at the origin, and this singularity at the origin is produced by the accumulation of *small jumps* of Markovian particles. The condition

$$\int_{\{|z|>1\}} |z| \, m(dz) < \infty$$

implies that the measure $m(\cdot)$ admits a singularity of order 1 at infinity, and this singularity at infinity is produced by the accumulation of *large jumps* of Markovian particles.

First, we consider the following boundary value problem: Given functions f and φ defined in D and on ∂D, respectively, find a function u in D such that

$$\begin{cases} Wu = f & \text{in } D, \\ L_0 u = \varphi & \text{on } \partial D. \end{cases} \tag{13.8}$$

Due to the non-local character of the Waldenfels integro-differential operator W, we find more difficulties in the bounded domain D than in the whole space $\mathbf{R}^N$. In fact, when considering the Dirichlet problem in D, it is natural to use the zero-extension of functions in the interior D outside of the closure $\overline{D} = D \cup \partial D$. This extension has a probabilistic interpretation. Namely, it corresponds to stopping the diffusion process with jumps in the whole space $\mathbf{R}^N$ at the first exit time of the closure $\overline{D}$.

However, the zero-extension produces a singularity of solutions at the boundary ∂D. In order to remove this singularity, we introduce various conditions on the structure of jumps for the Waldenfels integro-differential operator W. More precisely, if only jumps of order 1 at most are allowed outside of the closure $\overline{D}$, then we obtain regular solutions of the Dirichlet problem for the Waldenfels integro-differential operator W [GM2, Chapter 3].

For the boundary value problem (13.8), we are forced to impose a stronger condition (13.6) in order to remove a singularity of solutions at the boundary ∂D. Namely, no jumps outside of $\overline{D}$ are allowed. Probabilistically, this is a condition on the support of the Lévy measure $s(x, \cdot)\, m(\cdot)$ associated with the pseudo-differential operator S.

The next theorem is a generalization of Theorem 13.2 to the integro-differential operator case [Ta8, Theorem 1]:

Theorem 13.6. *Assume that condition (FH) is satisfied. Then the mapping*

$$(W, L_0) : C^{2+\theta}(\overline{D}) \longrightarrow C^{\theta}(\overline{D}) \bigoplus C^{1+\theta}_{L_0}(\partial D)$$

is an algebraic and topological isomorphism. *In particular, for any $f \in C^{\theta}(\overline{D})$ and any $\varphi \in C^{1+\theta}_{L_0}(\partial D)$, there exists a unique solution $u \in C^{2+\theta}(\overline{D})$ of problem* (13.8).

As an application of Theorem 13.6, we can construct a Feller semigroup corresponding to the physical phenomenon where a Markovian particle moves both by jumps and continuously in the state space until it "dies" at the time when it reaches the set where the particle is definitely absorbed, generalizing Theorem 13.5.

To do this, we define a linear operator $\mathcal{W}$ from the Banach space $C_0(\overline{D} \setminus M)$ into itself as follows:

(a) The domain of definition $D(\mathcal{W})$ is the set

$$\begin{aligned} D(\mathcal{W}) = \{u \in C^2(\overline{D}) \cap C_0(\overline{D} \setminus M) : Wu \in C_0(\overline{D} \setminus M), \\ L_0 u = 0 \quad \text{on } \partial D\}. \end{aligned}$$

(b) $\mathcal{W}u = Wu$ for every $u \in D(W)$.

The next theorem asserts that there exists a Feller semigroup on $\overline{D} \setminus M$ corresponding to the physical phenomenon where a Markovian particle moves both by jumps and continuously in the state space $\overline{D} \setminus M$ until it "dies" at the time when it reaches the set M as in Theorem 1.3 [Ta8, Theorem 2]:

Theorem 13.7. *If condition (FH) is satisfied, then the operator $\mathcal{W}$ is closable in the space $C_0(\overline{D} \setminus M)$, and its minimal closed extension $\overline{\mathcal{W}}$ is the infinitesimal generator of some Feller semigroup $\{e^{t\overline{\mathcal{W}}}\}_{t\geq 0}$ on $\overline{D} \setminus M$.*

Remark 13.8. For the non-degenerate case, the reader is referred to Komatsu [Ko, Theorem 5.2], Stroock [St, Theorem 2.2], Garroni–Menaldi [GM1, Chapter VIII, Theorem 3.3] and also Galakhov–Skubachevskiĭ [GB, Theorem 5.1].

Secondly, we study the boundary value problem (13.8) from the point of view of analytic semigroup theory, generalizing Theorems 13.3 and 13.4.

To do this, we associate with problem (13.8) an unbounded linear operator W_p from $L^p(D)$ into itself as follows:

(a) The domain of definition $D(W_p)$ is the set

$$D(W_p) = \{u \in H^{2,p}(D) : L_0 u = 0 \quad \text{on } \partial D\}.$$

(b) $W_p u = Wu$ for every $u \in D(W_p)$.

The next theorem is a generalization of Theorem 13.3 to the integro-differential operator case [Ta8, Theorem 3]:

Theorem 13.9. *Let $1 < p < \infty$. Assume that condition (FH) is satisfied. Then we have the following two assertions (i) and (ii):*

(i) For every $\varepsilon > 0$, there exists a constant $r_p(\varepsilon) > 0$ such that the resolvent set of W_p contains the set

$$\Sigma_p(\varepsilon) = \left\{\lambda = r^2 e^{i\vartheta} : r \geq r_p(\varepsilon),\ -\pi + \varepsilon \leq \vartheta \leq \pi - \varepsilon\right\},$$

and that the resolvent $(W_p - \lambda I)^{-1}$ satisfies the estimate

$$\left\|(W_p - \lambda I)^{-1}\right\| \leq \frac{c_p(\varepsilon)}{|\lambda|}, \quad \lambda \in \Sigma_p(\varepsilon), \tag{13.9}$$

where $c_p(\varepsilon) > 0$ is a constant depending on ε.

(ii) The operator W_p generates a semigroup e^{zW_p} on the space $L^p(D)$ which is analytic *in the sector*

$$\Delta_\varepsilon = \{z = t + is : z \neq 0,\ |\arg z| < \pi/2 - \varepsilon\}$$

for any $0 < \varepsilon < \pi/2$ (see Fig. 13.3).

Moreover, we introduce a linear operator $\mathfrak{W}$ from $C_0(\overline{D} \setminus M)$ into itself as follows:

(a) The domain of definition $D(\mathfrak{W})$ is the set

$$D(\mathfrak{W}) = \{u \in C_0(\overline{D} \setminus M) \cap H^{2,p}(D) : Wu \in C_0(\overline{D} \setminus M),$$
$$L_0 u = 0 \quad \text{on } \partial D\}, \qquad N < p < \infty.$$

(b) $\mathfrak{W}u = Wu$ for every $u \in \mathcal{D}(\mathfrak{W})$.

Here it should be noted that the domain $D(\mathfrak{W})$ is *independent* of $N < p < \infty$.

The next theorem is a generalization of Theorem 13.4 to the integro-differential operator case [Ta8, Theorem 4]:

Theorem 13.10. *Let $N < p < \infty$. If condition (FH) is satisfied, then we have the following two assertions (i) and (ii):*

(i) For every $\varepsilon > 0$, there exists a constant $r(\varepsilon) > 0$ such that the resolvent set of $\mathfrak{W}$ contains the set

$$\Sigma(\varepsilon) = \left\{\lambda = r^2 e^{i\vartheta} : r \geq r(\varepsilon),\ -\pi + \varepsilon \leq \vartheta \leq \pi - \varepsilon\right\},$$

and that the resolvent $(\mathfrak{W} - \lambda I)^{-1}$ satisfies the estimate

$$\|(\mathfrak{W} - \lambda I)^{-1}\| \leq \frac{c(\varepsilon)}{|\lambda|}, \quad \lambda \in \Sigma(\varepsilon), \tag{13.10}$$

where $c(\varepsilon) > 0$ is a constant depending on ε.

(ii) The operator $\mathfrak{W}$ generates a semigroup $e^{z\mathfrak{W}}$ on the space $C_0(\overline{D} \setminus M)$ which is analytic *in the sector $\Delta_\varepsilon = \{z = t + is : z \neq 0,\ |\arg z| < \pi/2 - \varepsilon\}$ for any $0 < \varepsilon < \pi/2$ (see Fig. 13.4).*

Remark 13.11. By combining Theorems 13.7 and 13.10, we can prove that the operator $\mathfrak{W}$ coincides with the minimal closed extension $\overline{\mathcal{W}}$: $\mathfrak{W} = \overline{\mathcal{W}}$.

Theorems 13.9 and 13.10 express a *regularizing effect* for the parabolic integro-differential operator $\partial/\partial t - W$ with homogeneous boundary condition L_0 (cf. [GM1, Chapter VIII, Theorem 3.1]).

13.4 Notes and Comments

It should be emphasized that the Calderón–Zygmund theory of singular integral operators with non-smooth kernels provides a powerful tool to deal with smoothness of solutions of elliptic boundary value problems, with minimal assumptions of regularity on the coefficients. The theory of singular integrals continues to be one of the most influential topics in the modern history of analysis, and is a very refined mathematical tool whose full power is yet to be exploited [Sn4].

Finally, we give an overview of general results on generation theorems for Feller semigroups proved by the author [Ta5,Ta6,Ta7,Ta8,Ta9,Ta10,Ta11] using the theory

Diffusion operator A	Lévy operator S_r	Ventcel' condition L	using the theory of	proved by
degenerate elliptic smooth case	null	second order case	pseudo-differential operators	[Ta5]
elliptic smooth case	null	degenerate Robin case	pseudo-differential operators	[Ta7] [Ta9]
elliptic smooth case	Hölder continuous case	degenerate Robin case	pseudo-differential operators	[Ta8]
elliptic discontinuous case	general case	Dirichlet case	singular integral operators	[Ta10]
elliptic discontinuous case	null	first order case	singular integral operators	[Ta11]
elliptic smooth case	general case	general case	pseudo-differential operators	[Ta6] This book

of pseudo-differential operators [Ho1, Se1, Se2] and the theory of singular integral operators [CZ].

Appendix A
Boundedness of Pseudo-differential Operators

In this appendix we prove an L^p boundedness theorem for pseudo-differential operators – a global version of Theorem 7.24 – which plays a fundamental role throughout the book. Bourdaud [Bd] proved the L^p boundedness by using the multiplier theorem of Marcinkiewicz, just as in Coifman–Meyer [CM]: The method of proof consists of the following two steps (1) and (2):

(1) A characterization of L^p functions in terms of the Littlewood-Paley series.
(2) A reduction of the problem to elementary symbols of the form

$$\sigma(x, \xi) = \sum_{j=0}^{\infty} M_j(2^{j\delta}x)\,\psi(2^{-j}\xi),$$

where $\psi(\xi)$ is a smooth function with compact support which does not contain the origin and $\{M_j\}$ is a bounded sequence in an appropriate Hölder space.

This appendix is a refinement of Appendix A of the first edition of the present monograph. For a general study of the L^p theory of pseudo-differential operators, the reader might refer to Coifman–Meyer [CM], Kumano-go [Ku] and Taylor [Ty].

A.1 The Littlewood–Paley Series of a Tempered Distribution

In this section we characterize L^p functions by their Littlewood–Paley series. We begin with the following elementary lemma (cf. [BL, Lemma 6.1.7]):

Lemma A.1. *For a given constant $a > 1$, there exists a function $\varphi(\xi) \in C_0^\infty(\mathbf{R}^n)$ such that*

$$\operatorname{supp} \varphi = \left\{ \xi \in \mathbf{R}^n : \frac{1}{a} \leq |\xi| \leq a \right\}, \tag{A.1a}$$

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*,
Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7,

$$\sum_{j\in\mathbf{Z}} \varphi(a^{-j}\,\xi) = 1 \quad \textit{for all } \xi \neq 0. \tag{A.1b}$$

Proof. If $h(\xi)$ is a non-negative function in $C_0^\infty(\mathbf{R}^n)$ such that

$$\text{supp } h = \left\{\xi \in \mathbf{R}^n : \frac{1}{a} \leq |\xi| \leq a\right\},$$

then we have, for all $k \in \mathbf{Z}$,

$$\text{supp } h(a^{-k}\,\cdot) = \{\xi \in \mathbf{R}^n : a^{k-1} \leq |\xi| \leq a^{k+1}\}.$$

Thus we can define a function $H(\xi) \in C_0^\infty(\mathbf{R}^n)$ by the formula

$$H(\xi) = \sum_{k\in\mathbf{Z}} h(a^{-k}\,\xi), \quad \xi \in \mathbf{R}^n.$$

It should be noted that the sum is locally finite, and that

$$\begin{cases} H(0) = 0, \\ H(\xi) > 0 \quad \text{for all } \xi \neq 0. \end{cases}$$

Therefore, it is easy to verify that the function

$$\varphi(\xi) = \frac{h(\xi)}{H(\xi)}, \quad \xi \in \mathbf{R}^n,$$

enjoys the desired properties (A.1) and (A.1a).

Furthermore, if we define a function $\psi_0(\xi)$ by the formula

$$\psi_0(\xi) = \sum_{j=-\infty}^{0} \varphi(a^{-j}\,\xi), \quad \xi \in \mathbf{R}^n,$$

then it follows from properties (A.1) and (A.1a) that

$$\psi_0 \in C_0^\infty(\mathbf{R}^n), \tag{A.2a}$$

$$\text{supp } \psi_0 = \{\xi \in \mathbf{R}^n : |\xi| \leq a\}, \tag{A.2b}$$

$$\psi_0(\xi) = 1 - \sum_{j=1}^{\infty} \varphi(a^{-j}\,\xi) \quad \text{for all } \xi \neq 0. \tag{A.2c}$$

Now, by using the Fourier transform we can introduce a family of linear operators

$$\Delta_0 : \mathcal{S}'(\mathbf{R}^n) \longrightarrow \mathcal{S}'(\mathbf{R}^n),$$
$$\Delta_j ; \mathcal{S}'(\mathbf{R}^n) \longrightarrow \mathcal{S}'(\mathbf{R}^n), \quad j = 1, 2, \ldots,$$

by the formulas

$$\widehat{\Delta_0 f}(\xi) = \psi_0(\xi) \hat{f}(\xi),$$
$$\widehat{\Delta_j f}(\xi) = \varphi(a^{-j}\xi) \hat{f}(\xi), \quad j = 1, 2, \ldots.$$

The operators Δ_j (and Δ_0) can be visualized in the following commutative diagram:

$$\begin{array}{ccc} f \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow{\Delta_j} & \mathcal{S}'(\mathbf{R}^n) \ni \Delta_j f \\ {\scriptstyle \mathcal{F}}\big\downarrow & & {\scriptstyle \mathcal{F}}\big\downarrow \\ \widehat{f}(\xi) \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow[\varphi(a^{-j}\xi)]{} & \mathcal{S}'(\mathbf{R}^n) \ni \varphi(a^{-j}\xi)\widehat{f}(\xi) \end{array}$$

Then, by properties (A.1) and (A.2) it is easy to see that

$$f = \sum_{j=0}^{\infty} \Delta_j f, \quad f \in \mathcal{S}'(\mathbf{R}^n). \tag{A.3}$$

Indeed, it suffices to note that, for all $\xi \neq 0$,

$$\begin{aligned} \sum_{j=0}^{\infty} \widehat{\Delta_j f}(\xi) &= \widehat{\Delta_0 f}(\xi) + \sum_{j=1}^{\infty} \widehat{\Delta_j f}(\xi) \\ &= \left(\psi_0(\xi) + \sum_{j=1}^{\infty} \varphi(a^{-j}\xi) \right) \hat{f}(\xi) = \left(\sum_{j \in \mathbf{Z}} \varphi(a^{-j}\xi) \right) \hat{f}(\xi) \\ &= \hat{f}(\xi). \end{aligned}$$

The series (A.3) is called the *Littlewood–Paley series* of f.

A.2 Peetre's Definition of Besov and Generalized Sobolev Spaces

In this section, following Bergh–Löfström [BL], we give Peetre's equivalent definition of Besov spaces and generalized Sobolev spaces.

To do this, we choose $a = 2$ in Lemma A.1. Namely, $\varphi(\xi)$ is a function in $C_0^\infty(\mathbf{R}^n)$ which satisfies the conditions

$$\text{supp}\ \varphi = \left\{\xi \in \mathbf{R}^n : \frac{1}{2} \leq |\xi| \leq 2\right\}, \tag{A.4a}$$

$$\sum_{k \in \mathbf{Z}} \varphi(2^{-k}\xi) = 1 \quad \text{for all } \xi \neq 0. \tag{A.4b}$$

Then we can define functions $\varphi_k(x)$, $\psi(x) \in \mathcal{S}(\mathbf{R}^n)$ by the formulas

$$\widehat{\varphi_k}(\xi) = \varphi(2^{-k}\xi), \quad k \in \mathbf{Z},$$

$$\hat{\psi}(\xi) = \psi_0(\xi) = 1 - \sum_{j=1}^{\infty} \varphi(2^{-j}\xi).$$

It should be noted that

$$\text{supp}\ \widehat{\varphi_k} = \{\xi \in \mathbf{R}^n : 2^{k-1} \leq |\xi| \leq 2^{k+1}\}, \quad k \in \mathbf{Z},$$

$$\text{supp}\ \hat{\psi} = \{\xi \in \mathbf{R}^n : |\xi| \leq 2\}.$$

In this appendix, if $s \in \mathbf{R}$ we define the Bessel potential J^s by the formula

$$J^s = (I - \Delta)^{s/2} : \mathcal{S}'(\mathbf{R}^n) \longrightarrow \mathcal{S}'(\mathbf{R}^n).$$

More precisely, we let

$$J^s f = \mathcal{F}^*((1 + |\xi|^2)^{s/2} \hat{f}(\xi)), \quad f \in \mathcal{S}'(\mathbf{R}^n).$$

The operators J^s can be visualized in the following diagram:

$$\begin{array}{ccc} f \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow{J^s = (1-\Delta)^{s/2}} & \mathcal{S}'(\mathbf{R}^n) \ni J^s f \\ \mathcal{F} \big\downarrow & & \big\uparrow \mathcal{F}^* \\ \widehat{f} \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow[(1+|\xi|^2)^{s/2}]{} & \mathcal{S}'(\mathbf{R}^n) \ni (1 + |\xi|^2)^{s/2} \widehat{f} \end{array}$$

Then we have the following basic properties of the objects just defined (see [BL, Lemma 6.2.1]):

Lemma A.2. *(i) Assume that a distribution $f \in \mathcal{S}'(\mathbf{R}^n)$ satisfies the condition*

$$\varphi_k * f \in L^p(\mathbf{R}^n), \quad k \in \mathbf{Z},$$

with $1 \le p \le \infty$*. Then we have, for all* $s \in \mathbf{R}$*,*

$$\|J^s(\varphi_k * f)\|_{L^p} \le C\, 2^{sk}\, \|\varphi_k * f\|_{L^p}, \quad k = 1, 2, \ldots,$$

with a constant $C > 0$ *independent of* p *and* k*.*

(ii) If a distribution $f \in \mathcal{S}'(\mathbf{R}^n)$ *satisfies the condition*

$$\psi * f \in L^p(\mathbf{R}^n),$$

with $1 \le p \le \infty$*, then we have, for all* $s \in \mathbf{R}$*,*

$$\|J^s(\psi * f)\|_{L^p} \le C'\, \|\psi * f\|_{L^p},$$

with a constant $C' > 0$ *independent of* p *and* k*.*

Now, by virtue of Lemma A.2 we can make the following Definition A.3 of the Besov and generalized Sobolev spaces (see [BL, Definition 6.2.2]):

Definition A.3. Let $s \in \mathbf{R}$, and $1 \le p, q \le \infty$. If $f \in \mathcal{S}'(\mathbf{R}^n)$, we let

$$\|f\|_{B^s_{p,q}} = \begin{cases} \|\psi * f\|_{L^p} + \left(\sum_{k=1}^{\infty} \left(2^{sk}\, \|\varphi_k * f\|_{L^p}\right)^q\right)^{1/q} & \text{if } 1 \le q < \infty, \\ \|\psi * f\|_{L^p} + \sup_{k\ge 1} \left(2^{sk}\, \|\varphi_k * f\|_{L^p}\right) & \text{if } q = \infty, \end{cases}$$

$$\|f\|_{H^s_p} = \|J^s f\|_{L^p}.$$

Then the *Besov space* $B^s_{p,q}(\mathbf{R}^n)$ and the *generalized Sobolev space* $H^s_p(\mathbf{R}^n)$ are defined respectively as follows:

$$B^s_{p,q}(\mathbf{R}^n) = \{f \in \mathcal{S}'(\mathbf{R}^n) : \|f\|_{B^s_{p,q}} < \infty\},$$
$$H^s_p(\mathbf{R}^n) = \{f \in \mathcal{S}'(\mathbf{R}^n) : \|f\|_{H^s_p} < \infty\}.$$

Remark A.4. It is known (see Bergh–Löfström [BL], Triebel [Tr]) that the Sobolev space $H^{s,p}(\mathbf{R}^n)$ and the Besov space $B^{s,p}(\mathbf{R}^{n-1})$ introduced in Sect. 6.2 are respectively equivalent to the following:

$$H^{s,p}(\mathbf{R}^n) = H^s_p(\mathbf{R}^n), \quad s \in \mathbf{R},\ 1 < p < \infty,$$
$$B^{s,p}(\mathbf{R}^{n-1}) = B^s_{p,p}(\mathbf{R}^{n-1}), \quad s \in \mathbf{R},\ 1 \le p \le \infty.$$

It should be noted that

$$H^0_p(\mathbf{R}^n) = L^p(\mathbf{R}^n), \quad 1 \le p \le \infty.$$

Furthermore, it is easy to verify the following two assertions (I) and (II):

(I) The spaces $B^s_{p,q}(\mathbf{R}^n)$ and $H^s_p(\mathbf{R}^n)$ are Banach spaces with norms $\|\cdot\|_{B^s_{p,q}}$ and $\|\cdot\|_{H^s_p}$, respectively.

(II) The Bessel potential J^σ is a topological isomorphism of $B^s_{p,q}(\mathbf{R}^n)$ onto $B^{s-\sigma}_{p,q}(\mathbf{R}^n)$ for each $\sigma \in \mathbf{R}$, and is also a topological isomorphism of $H^s_p(\mathbf{R}^n)$ onto $H^{s-\sigma}_p(\mathbf{R}^n)$ for each $\sigma \in \mathbf{R}$, respectively.

The next theorem characterizes the spaces $B^s_{p,q}(\mathbf{R}^n)$ and $H^s_p(\mathbf{R}^n)$ in terms of the Littlewood–Paley series (see [BL, Theorem 6.4.3]):

Theorem A.5. *(i) Let $s \in \mathbf{R}$ and $1 < p < \infty$. Then we have, for any $f \in \mathcal{S}'(\mathbf{R}^n)$,*

$$f \in H^s_p(\mathbf{R}^n) \Longleftrightarrow \left(\sum_{j=0}^{\infty} 2^{2sj} \, |\Delta_j f|^2 \right)^{1/2} \in L^p(\mathbf{R}^n).$$

(ii) Let $s \in \mathbf{R}$ and $1 \leq p, q \leq \infty$. Then we have, for any $f \in \mathcal{S}'(\mathbf{R}^n)$,

$$f \in B^s_{p,q}(\mathbf{R}^n) \Longleftrightarrow \sum_{j=0}^{\infty} 2^{sqj} \, \|\Delta_j f\|_{B^s_{p,q}} < \infty.$$

Remark A.6. Theorem A.5 remains valid with the constant 2 replaced by a general constant $a > 1$.

A.3 Non-regular Symbols

In this section we introduce a class of non-regular symbols $\sigma(x, \xi)$ which are Hölder continuous with respect to the variable x and belong to the Hörmander class $S^0_{1,\delta}$ with respect to the variable ξ.

Let $0 \leq \delta \leq 1$, $N \in \mathbf{N}$ and r a *non-integral* positive number. A function $\sigma(x, \xi)$ defined on $\mathbf{R}^n \times \mathbf{R}^n$ belongs to the class $S^0_{1,\delta}(N, r)$ if it satisfies the following two conditions (a) and (b):

(a) For each $|\alpha| \leq N$ and $|\beta| < r$, there exists a constant $C_{\alpha,\beta} > 0$ such that

$$\left| \partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi) \right| \leq C_{\alpha,\beta} (1 + |\xi|)^{-|\alpha| + \delta|\beta|}, \quad x, \; \xi \in \mathbf{R}^n.$$

(b) For each $|\alpha| \leq N$ and $|\beta| = [r]$, there exists a constant $C'_{\alpha,\beta} > 0$ such that

$$\left| \partial_\xi^\alpha \partial_x^\beta \sigma(x + h, \xi) - \partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi) \right| \leq C'_{\alpha,\beta} \, |h|^{r-[r]} (1 + |\xi|)^{-|\alpha| + r\delta}, \qquad \text{(A.5)}$$

$$x, \; \xi, \; h \in \mathbf{R}^n.$$

Here $[r]$ is the integral part of r.

Remark A.7. If r is integral, then condition (A.5) should be replaced by the *Zygmund condition* for $|\beta| = r - 1$

$$\left|\partial_\xi^\alpha \partial_x^\beta \sigma(x+h,\xi) + \partial_\xi^\alpha \partial_x^\beta \sigma(x-h,\xi) - 2\partial_\xi^\alpha \partial_x^\beta \sigma(x,\xi)\right|$$
$$\leq C'_{\alpha,\beta}\, |h|(1+|\xi|)^{-|\alpha|+r\delta}, \quad x,\ \xi,\ h \in \mathbf{R}^n.$$

It is easy to verify that the class $S^0_{1,\delta}(N,r)$ is a Banach space with respect to the norm

$$\|\sigma\| = \sum_{\substack{|\alpha|\leq N \\ |\beta|<r}} \sup_{x,\xi} \left\{ \frac{|\partial_\xi^\alpha \partial_x^\beta \sigma(x,\xi)|}{(1+|\xi|)^{-|\alpha|+\delta|\beta|}} \right\}$$
$$+ \sum_{\substack{|\alpha|\leq N \\ |\beta|=[r]}} \sup_{x,\xi,h} \left\{ \frac{|\partial_\xi^\alpha \partial_x^\beta \sigma(x+h,\xi) - \sigma(x,\xi)|}{|h|^{r-[r]}(1+|\xi|)^{-|\alpha|+r\delta}} \right\}.$$

The next lemma asserts that every symbol $\sigma(x,\xi)$ in $S^0_{1,\delta}(N,r)$ can be decomposed into elementary symbols (cf. [CM, Chapter II, Proposition 5]):

Lemma A.8. *Let* $0 \leq \delta \leq 1$, $r > 0$ *and let* N *be an* even *integer greater than* n. *Then every symbol* $\sigma(x,\xi) \in S^0_{1,\delta}(N,r)$ *can be decomposed into the form*

$$\sigma(x,\xi) = \tilde{\sigma}(x,\xi) + \sum_{k\in\mathbf{Z}^n} (1+|k|^2)^{-(n+1)/2} \sigma_k(x,\xi). \tag{A.6}$$

Here:

(i) The symbol $\tilde{\sigma}(x,\xi) \in S^0_{1,\delta}(N,r)$ *satisfies the condition*

$$\tilde{\sigma}(x,\xi) = 0 \quad \textit{for all } |\xi| \geq 2, \tag{A.7}$$

and there exists a constant $C > 0$, *depending on n and* N, *such that*

$$\|\partial_\xi^\alpha \tilde{\sigma}(\cdot,\xi)\|_{\Lambda_r} \leq C\|\sigma\|, \quad |\alpha| \leq N.$$

(ii) Every elementary symbol $\sigma_k(x,\xi)$ *is written in the form*

$$\sigma_k(x,\xi) = \sum_{j=0}^{\infty} M_{k,j}(2^{j\delta}x)\, \psi_k(2^{-j}\xi), \tag{A.8}$$

where

$$\psi_k(\xi) = (1+|k|^2)^{(n+1-N)/2} e^{ik\xi}\,\theta(\xi)$$

$$\theta(\xi) \in C_0^\infty(\mathbf{R}^n), \quad \operatorname{supp}\theta = \left\{\xi \in \mathbf{R}^n : \frac{1}{3} \le |\xi| \le 3\right\},$$

$$M_{k,j}(x) = (1+|k|^2)^{N/2} C_{k,j}(x) \in \Lambda_r(\mathbf{R}^n).$$

Moreover, we have, with some constant $C' > 0$ *independent of* k *and* j,

$$\|\psi_k\|_{L^\infty_{N-n-1}} \le C', \tag{A.9a}$$

$$\|M_{k,j}\|_{\Lambda_r} \le C'\|\sigma\|. \tag{A.9b}$$

Here $\Lambda_r(\mathbf{R}^n) = B^r_{\infty,\infty}(\mathbf{R}^n)$ *is the classical Hölder space and* $L^\infty_m(\mathbf{R}^n)$ *is the space of functions on* $\mathbf{R}^n$ *whose distribution derivatives of order* $\le m$ *belong to* $L^\infty(\mathbf{R}^n)$, *respectively.*

Proof. The proof is divided into three steps.

Step 1: First, we take a non-negative function $\mu(\xi) \in C_0^\infty(\mathbf{R}^n)$ which satisfies the condition

$$\mu(\xi) = \begin{cases} 1 & \text{if } |\xi| \le 1, \\ 0 & \text{if } |\xi| \ge 2, \end{cases} \tag{A.10}$$

and let

$$\begin{aligned} \sigma(x,\xi) &= \mu(\xi)\sigma(x,\xi) + (1-\mu(\xi))\sigma(x,\xi) \\ &:= \tilde{\sigma}(x,\xi) + \tau(x,\xi). \end{aligned} \tag{A.11}$$

By condition (A.10) and the Leibniz rule, it is easy to verify that

$$\tilde{\sigma}(x,\xi) = \mu(\xi)\sigma(x,\xi) = 0 \quad \text{for all } |\xi| \ge 2,$$

and that we have, for all $|\alpha| \le N$,

$$\begin{aligned} \|\partial_\xi^\alpha \tilde{\sigma}(\cdot,\xi)\|_{\Lambda_r} &\le C_1\|\sigma\|(1+|\xi|)^{-|\alpha|+r\delta} \le C_1\|\sigma\|(1+|\xi|)^{r\delta} \\ &\le 3^{r\delta}\, C_1\|\sigma\|, \end{aligned}$$

with a constant $C_1 > 0$.
Similarly, we have the estimate

$$\|\tau\| \le C_2\|\sigma\|,$$

with a constant $C_2 > 0$.

Step 2: Secondly, we take a non-negative function $\lambda(\xi) \in C_0^\infty(\mathbf{R}^n)$ which satisfies the conditions

$$\text{supp } \lambda = \left\{ \xi \in \mathbf{R}^n : \frac{1}{2} \leq |\xi| \leq 2 \right\},$$

$$\sum_{j=0}^{\infty} \lambda(2^{-j}\xi) = 1 \quad \text{for all } |\xi| \geq 1,$$

and let

$$\begin{aligned} \tau_j(x,\xi) &= \lambda(\xi)\, \tau(2^{-j\delta}x, 2^j\xi) \\ &= \lambda(\xi)\,(1-\mu(2^j\xi))\sigma(2^{-j\delta}x, 2^j\xi), \quad j = 0,\ 1,\ 2, \ldots. \end{aligned}$$

Then the symbols $\tau_j(x,\xi)$ are estimated as follows:

Lemma A.9. *For each $|\alpha| \leq n$, there exists a constant $C > 0$, independent of j, such that*

$$\|D_\xi^\alpha \tau_j(\cdot,\xi)\|_{\Lambda_r} \leq C\, \|\sigma\|, \quad j = 0,\ 1,\ 2, \ldots. \tag{A.12}$$

Proof. For example, we consider the case where $\alpha = 0$: It should be noted that we have, for $|\beta| = [r]$,

$$\begin{aligned} &\left(\frac{1}{2^{j\delta}}\right)^{|\beta|} \left| \partial_x^\beta \sigma\left(\frac{x+h}{2^{j\delta}}, 2^j\xi\right) - \partial_x^\beta \sigma\left(\frac{x}{2^{j\delta}}, 2^j\xi\right)\right| \\ &\leq \|\sigma\| \left(\frac{1}{2^{j\delta}}\right)^{|\beta|} \left(\frac{|h|}{2^{j\delta}}\right)^{r-[r]} (1+|2^j\xi|)^{r\delta} \\ &= \|\sigma\| |h|^{r-[r]} \left(\frac{1+|2^j\xi|}{2^j}\right)^{r\delta}. \end{aligned}$$

However, it follows that

$$1 + 2^{j-1} \leq 1 + |2^j\xi| \leq 1 + 2^{j+1}, \quad \frac{1}{2} \leq |\xi| \leq 2,$$

so that

$$\frac{1}{2} \leq \frac{1}{2^j} + \frac{1}{2} \leq \frac{1+|2^j\xi|}{2^j} \leq \frac{1}{2^j} + 2 \leq 3, \quad \xi \in \text{supp } \lambda.$$

Hence we have estimate (A.12) for $\alpha = 0$.

Step 3: Thirdly, we take a non-negative function $\theta(\xi) \in C_0^\infty(\mathbf{R}^n)$ which satisfies the conditions

$$\text{supp } \theta = \left\{ \xi \in \mathbf{R}^n : \frac{1}{3} \leq |\xi| \leq 3 \right\},$$
$$\theta(\xi) = 1 \quad \text{on supp } \lambda.$$

Then we can expand the symbol $\tau_j(x,\xi) = \tau_j(x,\xi)\,\theta(\xi)$ in the Fourier series form

$$\tau_j(x,\xi) = \left(\sum_{k \in \mathbf{Z}^n} C_{k,j}(x)\, e^{ik\cdot\xi} \right) \theta(\xi),$$

where

$$C_{k,j}(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \tau_j(x,\eta)\, e^{-ik\cdot\eta}\, d\eta.$$

Moreover, we can rewrite the symbol $\tau_j(x,\xi)$ in the form (see decomposition (A.8)):

$$\begin{aligned} \tau_j(x,\xi) &= \sum_{k \in \mathbf{Z}^n} (1+|k|^2)^{-(n+1)/2} \left\{ (1+|k|^2)^{N/2} C_{k,j}(x) \right\} \\ &\quad \times \left\{ (1+|k|^2)^{(n+1-N)/2} e^{ik\xi}\, \theta(\xi) \right\} \\ &:= \sum_{k \in \mathbf{Z}^n} u_k\, M_{k,j}(x)\, \psi_k(\xi). \end{aligned} \tag{A.13}$$

Here it should be noted that

$$N \text{ is an } \textit{even} \text{ integer greater than } n,$$
$$\sum_{k \in \mathbf{Z}^n} u_k = \sum_{k \in \mathbf{Z}^n} (1+|k|^2)^{-(n+1)/2} < \infty.$$

Hence we have, by decomposition (A.13),

$$\begin{aligned} \tau(x,\xi) &= \left(\sum_{j=0}^{\infty} \lambda(2^{-j}\xi) \right) \tau(x,\xi) \\ &= \sum_{j=0}^{\infty} \tau_j(2^{j\delta}x, 2^{-j}\xi) = \sum_{j=0}^{\infty} \sum_{k \in \mathbf{Z}^n} u_k\, M_{k,j}(2^{j\delta}x)\, \psi_k(2^{-j}\xi) \end{aligned} \tag{A.14}$$

$$= \sum_{k\in\mathbf{Z}^n} u_k \left\{ \sum_{j=0}^{\infty} M_{k,j}(2^{j\delta} x)\, \psi_k(2^{-j}\xi) \right\}$$
$$:= \sum_{k\in\mathbf{Z}^n} (1+|k|^2)^{-(n+1)/2}\, \sigma_k(x,\xi).$$

The desired decomposition (A.5) follows by combining decompositions (A.10) and (A.14).

It remains to prove estimates (A.9) and (A.9a) for the functions $\psi_k(\xi)$ and $M_{k,j}(x)$.

(1) The estimate (A.9) of $\psi_k(\xi)$: Since we have, for $|\alpha| \le N-n-1$,

$$\left|\partial_\xi^\alpha \psi_k(\xi)\right| \le C_3 \|\theta\|_{L^\infty_{N-n-1}},$$

it follows that

$$\|\psi_k\|_{L^\infty_{N-n-1}} \le C_4.$$

(2) The estimate (A.9a) of $M_{k,j}(x)$: By integration by parts, it follows that

$$\begin{aligned} M_{k,j}(x) &= \frac{(1+|k|^2)^{N/2}}{(2\pi)^n} \int_{\mathbf{R}^n} e^{-ik\xi}\, \tau_j(x,\xi)\, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \tau_j(x,\xi) \cdot (I-\Delta_\xi)^{N/2} e^{-ik\xi}\, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{-ik\xi} (I-\Delta_\xi)^{N/2} \tau_j(x,\xi)\, d\xi. \end{aligned}$$

Hence, by Lemma A.9 it is easy to see that

$$\|M_{k,j}\|_{\Lambda_r} \le C_5 \|\sigma\|.$$

Now the proof of Lemma A.8 is complete.

A.4 The L^p Boundedness Theorem

In this section we formulate an L^p boundedness theorem for pseudo-differential operators with non-regular symbols due to Bourdaud [Bd, Theorem 1]:

Theorem A.10. *Let* $0 \le \delta \le 1$, $r > 0$ *and let* N *be an* even *integer greater than* $(3n/2)+1$. *If* $\sigma(x,\xi)$ *is a symbol in the class* $S^0_{1,\delta}(N,r)$, *then the pseudo-differential operator* $\sigma(x,D)$, *defined by the formula*

$$\sigma(x, D) f(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \sigma(x, \xi) \hat{f}(\xi)\, d\xi,$$

is bounded on the generalized Sobolev space $H^s_p(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 < p < \infty$, *and is bounded on the Besov space* $B^s_{p,q}(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 \leq p, q \leq \infty$, *respectively.*

By virtue of Lemma A.8, the proof of Theorem A.10 is reduced to the proof of the following two Propositions A.11 and A.12:

Proposition A.11. *Let* $0 \leq \delta \leq 1$, $r > 0$ *and let* N *be an* even *integer greater than* n. *Assume that a symbol* $\sigma(x, \xi) \in S^0_{1,\delta}(N, r)$ *satisfies the condition*

$$\sigma(x, \xi) = 0 \quad \textit{for all } |\xi| \geq 2. \tag{A.15}$$

Then the pseudo-differential operator

$$\sigma(x, D) f(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \sigma(x, \xi) \hat{f}(\xi)\, d\xi$$

is bounded on the generalized Sobolev space $H^s_p(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 < p < \infty$, *and is bounded on the Besov space* $B^s_{p,q}(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 \leq p, q \leq \infty$. *More precisely, we have the estimate for the operator norm*

$$\|\sigma(x, D)\| \leq C \|\sigma\|,$$

with a constant $C > 0$ *depending on* n, p, q, r, s *and* δ.

Proposition A.12. *Let* $0 \leq \delta \leq 1$, $r > 0$ *and let* N *be an* even *integer greater than* n. *Assume that a symbol* $\sigma(x, \xi) \in S^0_{1,\delta}(N, r)$ *is an elementary symbol of the form*

$$\sigma(x, \xi) = \sum_{j=0}^{\infty} M_j(2^{j\delta} x)\, \psi(2^{-j}\xi), \tag{A.16}$$

where

$$\psi(\xi) \in C_0^\infty(\mathbf{R}^n), \quad \operatorname{supp} \psi = \left\{ \xi \in \mathbf{R}^n : \frac{1}{3} \leq |\xi| \leq 3 \right\}.$$

The sequence $\{M_j\}$ *is bounded in the Hölder space* $\Lambda_r(\mathbf{R}^n)$.

Then the pseudo-differential operator

$$\sigma(x, D) f(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \sigma(x, \xi) \hat{f}(\xi)\, d\xi$$

is bounded on the generalized Sobolev space $H_p^s(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 < p < \infty$*, and is bounded on the Besov space* $B_{p,q}^s(\mathbf{R}^n)$ *for all* $(\delta - 1)r < s < r$ *and* $1 \leq p, q \leq \infty$*. More precisely, we have the estimate for the operator norm*

$$\|\sigma(x, D)\| \leq C \left(\sup_{j \geq 0} \|M_j\|_{\Lambda_r} \right) \|\psi\|_{L^\infty_{N-n-1}},$$

with a constant $C > 0$ *depending on* n*,* p*,* q*,* r*,* s *and* δ*.*

In this appendix we shall only prove the generalized Sobolev space case. The proof of the Besov space case is left to the reader.

A.5 Proof of Proposition A.11

First, it should be noted that

$$\begin{aligned} &\|\sigma(x, D)u\|_{H_p^s} = \|J^s(\sigma(x, D)u)\|_{L^p} = \|\sigma(x, D)J^{-s}v\|_{L^p}, \\ &u = J^{-s}v \in H_p^s(\mathbf{R}^n),\ v \in L^p(\mathbf{R}^n). \end{aligned}$$

In other words, we have the following commutative diagram:

$$\begin{array}{ccc} H_p^s & \xrightarrow{\sigma(x,D)} & H_p^s \\ \uparrow J^{-s} & & J^s \downarrow \\ L^p & \xrightarrow[J^s\sigma(x,D)J^{-s}]{} & L^p \end{array}$$

Furthermore, we find that the symbol $\tau(x, \xi)$ of $J^s\sigma(x, D)J^{-s}$ is written in the form

$$\tau(x, \xi) = \sigma(x, \xi) + r(x, \xi),$$

where $r(x, \xi)$ belongs to the Hörmander class $S^{-1}_{1,\delta}$ with respect to the variable ξ.

Therefore, the proof of Proposition A.11 is reduced to the proof of the L^p boundedness of the two operators $\sigma(x, D)$ and $r(x, D)$ which will be proved in Lemmas A.13 and A.17, respectively.

Lemma A.13. *Let* $0 \leq \delta \leq 1$*,* $r > 0$ *and let* N *be an* even *integer greater than* n*. Assume that a symbol* $\sigma(x, \xi) \in S^0_{1,\delta}(N, r)$ *satisfies the condition*

$$\sigma(x, \xi) = 0 \quad \text{for all } |\xi| \geq 2. \tag{A.17}$$

Then there exists a constant $C > 0$ *such that*

$$\|\sigma(x, D)f\|_{L^p} \leq C\|f\|_{L^p}, \quad f \in L^p(\mathbf{R}^n). \tag{A.18}$$

Proof. First, we have, by condition (A.17),

$$\begin{aligned}\sigma(x, D)f(x) &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{ix\cdot\xi}\sigma(x,\xi)\hat{f}(\xi)\,d\xi \\ &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\left(\int_{\mathbf{R}^n} e^{i(x-y)\cdot\xi}\sigma(x,\xi)\,d\xi\right)f(y)\,dy \\ &= \int_{\mathbf{R}^n} K(x-y, y)f(y)\,dy,\end{aligned}$$

where

$$K(x,z) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{iz\cdot\xi}\sigma(x,\xi)\,d\xi.$$

However, by integration by parts it follows that, for any integer $\ell \in \mathbf{N}$,

$$\begin{aligned}(1+|z|^2)^\ell\,K(x,z) &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{iz\cdot\xi}(1+|z|^2)^\ell\,\sigma(x,\xi)\,d\xi \\ &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}(I-\Delta_\xi)^\ell e^{iz\cdot\xi}\cdot\sigma(x,\xi)\,d\xi \\ &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{iz\cdot\xi}\,(I-\Delta_\xi)^\ell\sigma(x,\xi)\,d\xi.\end{aligned}$$

Hence we can find a constant $C_1 > 0$ such that

$$(1+|z|^2)^\ell\,|K(x,z)| \leq C_1\|\sigma\|.$$

Now, if we take the integer ℓ satisfying the condition

$$n < 2\ell \leq N,$$

then we have the inequality

$$|K(x,z)| \leq \frac{C_1\|\sigma\|}{(1+|z|^2)^\ell}.$$

Thus it follows that

$$
\begin{aligned}
|\sigma(x, D) f(x)| &\leq \int_{\mathbf{R}^n} |K(x-y, y)|\, |f(y)|\, dy \\
&\leq C_1 \|\sigma\| \int_{\mathbf{R}^n} \frac{|f(y)|}{(1+|x-y|^2)^\ell}\, dy \\
&:= K_0 * |f|(x),
\end{aligned}
$$

where

$$
K_0(x) = \frac{C_1 \|\sigma\|}{(1+|x|^2)^\ell} \in L^1(\mathbf{R}^n).
$$

Therefore, by applying Young's inequality (Corollary 5.3) we obtain that

$$
\|\sigma(x, D) f\|_{L^p} \leq \|K_0 * |f|\|_{L^p} \leq \|K_0\|_{L^1} \, \|f\|_{L^p}.
$$

This proves inequality (A.18) with $C := \|K_0\|_{L^1}$.

The next lemma, due to Nagase [Na, Theorem 2], plays an essential role in the proof of an L^p boundedness theorem for pseudo-differential operators $r(x, D)$ of *negative order*:

Lemma A.14. *Assume that a symbol $r(x, \xi) \in S^0_{1,\delta}(N, r)$ satisfies the following condition (A.19):*

There exists a constant $\rho > 0$ such that, for each $|\alpha| \leq n+1$, we have, with some constant $C > 0$,

$$
|\partial_\xi^\alpha r(x, \xi)| \leq C(1+|\xi|)^{-|\alpha|-\rho}. \tag{A.19}
$$

Then there exists a function $K(x, z) \in C(\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\}))$ such that

$$
r(x, D) f = \int_{\mathbf{R}^n} K(x, x-y) f(y)\, dy, \quad f \in L^p(\mathbf{R}^n). \tag{A.20}
$$

Moreover, for each $0 < \rho' < \min(1, \rho)$, there exists a constant $C' > 0$ such that

$$
|K(x, z)| \leq C' \frac{1}{1+|z|} \frac{1}{|z|^{n-\rho'}}. \tag{A.21}
$$

Proof. The proof is divided into five steps.

Step 1: First, we take a function $\chi(\xi) \in C_0^\infty(\mathbf{R}^n)$ such that

$$
\chi(\xi) = \begin{cases} 1 & \text{if } |\xi| \leq \dfrac{1}{2}, \\ 0 & \text{if } |\xi| \geq 1, \end{cases}
$$

and let

$$r_\varepsilon(x,\xi) = \chi(\varepsilon\xi) r(x,\xi), \quad 0 < \varepsilon < 1.$$

Then it is clear that the support of $r_\varepsilon(x,\xi)$ is compact with respect to the variable ξ.

Step 2: If we let

$$K_\varepsilon(x,z) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{iz\cdot\xi} r_\varepsilon(x,\xi)\, d\xi, \quad 0 < \varepsilon < 1,$$

then it follows from an application of the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] that

$$\begin{aligned} r(x,D)f(x) &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} r(x,\xi) \hat{f}(\xi)\, d\xi \\ &= \frac{1}{(2\pi)^n} \lim_{\varepsilon\downarrow 0} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \chi(\varepsilon\xi) r(x,\xi) \hat{f}(\xi)\, d\xi \\ &= \frac{1}{(2\pi)^n} \lim_{\varepsilon\downarrow 0} \iint_{\mathbf{R}^n\times\mathbf{R}^n} e^{i(x-y)\cdot\xi} r_\varepsilon(x,\xi) f(y)\, dy\, d\xi \\ &= \lim_{\varepsilon\downarrow 0} \left(\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i(x-y)\cdot\xi} r_\varepsilon(x,\xi)\, d\xi \right) f(y)\, dy, \end{aligned}$$

that is,

$$r(x,D)f(x) = \lim_{\varepsilon\downarrow 0} \int_{\mathbf{R}^n} K_\varepsilon(x,x-y) f(y)\, dy, \quad f \in L^p(\mathbf{R}^n). \tag{A.22}$$

However, by integration by parts it follows that

$$\begin{aligned} z^\alpha K_\varepsilon(x,z) &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} z^\alpha e^{iz\cdot\xi} r_\varepsilon(x,\xi)\, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} D_\xi^\alpha \left(e^{iz\cdot\xi}\right) \cdot r_\varepsilon(x,\xi)\, d\xi \\ &= \frac{(-1)^{|\alpha|}}{(2\pi)^n} \int_{\mathbf{R}^n} e^{iz\cdot\xi} D_\xi^\alpha r_\varepsilon(x,\xi)\, d\xi \\ &= \frac{(-1)^{|\alpha|}}{(2\pi)^n} \int_{\mathbf{R}^n} \left(e^{iz\cdot\xi} - 1\right) D_\xi^\alpha r_\varepsilon(x,\xi)\, d\xi, \end{aligned} \tag{A.23}$$

since we have

$$\int_{\mathbf{R}^n} D_\xi^\alpha r_\varepsilon(x,\xi)\,d\xi = \int_{\mathbf{R}^n} D_\xi^\alpha\left(\chi(\varepsilon\xi)\,r(x,\xi)\right)d\xi = 0.$$

Here we make the following elementary claim:

Lemma A.15. *(a)* $D_\xi^\alpha r_\varepsilon(x,\xi) \to D_\xi^\alpha r(x,\xi)$ *as* $\varepsilon \downarrow 0$.
(b) For each $|\alpha| = n$*, we have the inequality*

$$|D_\xi^\alpha r_\varepsilon(x,\xi)| \le \|r\|(1+|\xi|)^{-n-\rho}, \tag{A.24}$$

where

$$\|r\| = \sum_{|\alpha|\le n} \sup_{x,\xi} \left\{ \frac{|\partial_\xi^\alpha r(x,\xi)|}{(1+|\xi|)^{-|\alpha|-\rho}} \right\}.$$

Therefore, by letting $\varepsilon \downarrow 0$ in formula (A.23) it follows from an application of the Lebesgue dominated convergence theorem [Fo2, Theorem 2.24] that

$$\lim_{\varepsilon\downarrow 0} z^\alpha\, K_\varepsilon(x,z) = \frac{(-1)^{|\alpha|}}{(2\pi)^n} \int_{\mathbf{R}^n} \left(e^{iz\cdot\xi} - 1\right) D_\xi^\alpha r(x,\xi)\,d\xi. \tag{A.25}$$

If we let

$$K(x,z) = \frac{(-1)^{|\alpha|}}{(2\pi)^n\, z^\alpha} \int_{\mathbf{R}^n} \left(e^{iz\cdot\xi} - 1\right) D_\xi^\alpha r(x,\xi)\,d\xi, \quad z \ne 0, \tag{A.26}$$

we obtain from formula (A.25) that

$$\lim_{\varepsilon\downarrow 0} z^\alpha\, K_\varepsilon(x,z) = z^\alpha\, K(x,z).$$

Step 3: To prove inequality (A.21), we estimate the function $z^\alpha\, K(x,z)$ for $|\alpha| = n$ and $|\alpha| = n+1$. To do this, we need the following elementary inequality (A.27):

Lemma A.16. *We have, for* $0 < \rho' < 1$*,*

$$|e^{iz\cdot\xi} - 1| \le 2|z|^{\rho'}|\xi|^{\rho'}. \tag{A.27}$$

Proof. By the mean value theorem, it follows that

$$|e^{iz\cdot\xi} - 1| \le \min\{2, |z|\,|\xi|\} \le 2\min\{1, |z|\,|\xi|\}. \tag{A.28}$$

However, we have the inequalities

$$\begin{cases} |z|\,|\xi| \le 1 \Longrightarrow |z|\,|\xi| \le (|z|\,|\xi|)^{\rho'}, \\ |z|\,|\xi| \ge 1 \Longrightarrow 1 \le (|z|\,|\xi|)^{\rho'}, \end{cases} \tag{A.29}$$

since $0 < \rho' < 1$.

The desired inequality (A.27) follows by combining inequalities (A.28) and (A.29).

Case (a): $|\alpha| = n$. By using formula (A.25) and inequalities (A.24) and (A.27), we obtain that

$$\begin{aligned} |z^\alpha K_\varepsilon(x,z)| &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} \left|e^{iz\cdot\xi}-1\right| |D_\xi^\alpha r_\varepsilon(x,\xi)|\,d\xi \\ &\le \frac{2\|r\||z|^{\rho'}}{(2\pi)^n}\int_{\mathbf{R}^n} |\xi|^{\rho'}(1+|\xi|)^{-n-\rho}\,d\xi \\ &\le \frac{2\|r\||z|^{\rho'}}{(2\pi)^n}\int_{\mathbf{R}^n} \frac{1}{(1+|\xi|)^{n+(\rho-\rho')}}\,d\xi \\ &\le C_1\|r\||z|^{\rho'}, \end{aligned} \tag{A.30}$$

with some constant $C_1 > 0$.

Case (b): $|\alpha| = n+1$. Just as in the case (a), we have, with some constant $C_2 > 0$,

$$\begin{aligned} |z^\alpha K_\varepsilon(x,z)| &\le \frac{2\|r\||z|^{\rho'}}{(2\pi)^n}\int_{\mathbf{R}^n} \frac{1}{(1+|\xi|)^{n+\rho+(1-\rho')}}\,d\xi \\ &\le C_2\|r\||z|^{\rho'}. \end{aligned} \tag{A.31}$$

Therefore, by letting $\varepsilon \downarrow 0$ in inequalities (A.30) and (A.31) it follows that, for $|\alpha| = n$ and $|\alpha| = n+1$, there exists a constant $C_3 > 0$ such that

$$|z^\alpha K(x,z)| \le C_3|z|^{\rho'}.$$

This proves inequality (A.21).

Step 4: On the other hand, by combining Lemmas A.15 and A.16 we obtain that, for $|\alpha| = n$ and $|\alpha| = n+1$,

$$\left|\left(e^{ix\cdot\xi}-1\right) D_\xi^\alpha r(x,\xi)\right| \le \frac{C_3|z|^{\rho'}}{(1+|\xi|)^{n+(\rho-\rho')}}.$$

Thus, by (A.26) it follows from an application of the Lebesgue dominated convergence theorem that the function

$$z^\alpha K(x,z) = \frac{(-1)^{|\alpha|}}{(2\pi)^n} \int_{\mathbf{R}^n} \left(e^{iz\cdot\xi} - 1\right) D_\xi^\alpha r(x,\xi)\, d\xi$$

is continuous on $\mathbf{R}^n \times \mathbf{R}^n$. In particular, we have

$$K(x,z) \in C(\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})).$$

Step 5: Finally, by inequality (A.21) we can let $\varepsilon \downarrow 0$ in (A.22) to obtain the integral formula (A.20).

As a corollary of Lemma A.14, we can prove an L^p boundedness theorem for pseudo-differential operators $r(x,D)$ of negative order (see Nagase [Na, Theorem 3]):

Lemma A.17. *If a bounded continuous symbol* $r(x,\xi)$ *satisfies condition* (A.19), *then there exists a constant* $C > 0$ *such that*

$$\|r(x,D)f\|_{L^p} \le C\|f\|_{L^p}, \quad f \in L^p(\mathbf{R}^n). \tag{A.32}$$

Proof. By inequality (A.21), it follows that

$$\begin{aligned}
\int_{\mathbf{R}^n} |K(x,x-y)|\, dx &\le C' \int_{\mathbf{R}^n} \frac{1}{1+|x-y|} \frac{1}{|x-y|^{n-\rho'}}\, dx \\
&= C' \int_{\mathbf{R}^n} \frac{1}{1+|x|} \frac{1}{|x|^{n-\rho'}}\, dx \\
&\le C' \left(\int_{|x|\le 1} \frac{1}{|x|^{n-\rho'}}\, dx + \int_{|x|>1} \frac{1}{|x|^{n+(1-\rho')}}\, dx \right) \\
&< \infty.
\end{aligned}$$

Similarly, we have the estimate

$$\begin{aligned}
\int_{\mathbf{R}^n} |K(x,x-y)|\, dy &\le C' \left(\int_{|y|\le 1} \frac{1}{|y|^{n-\rho'}}\, dy + \int_{|y|>1} \frac{1}{|y|^{n+(1-\rho')}}\, dy \right) \\
&< \infty.
\end{aligned}$$

Therefore, the desired inequality (A.32) follows by applying Schur's lemma (Theorem 5.2).

A.6 Proof of Proposition A.12

The next lemma plays an essential role in the proof of Proposition A.12:

Lemma A.18. *Let $1 < p < \infty$, $s \in \mathbf{R}$, $a, \gamma > 1$ and $m > n/2$.*

(i) If $\psi(\xi)$ is a function in $C_0^\infty(\mathbf{R}^n)$ which satisfies the condition

$$\operatorname{supp} \psi \subset \left\{ \xi \in \mathbf{R}^n : \frac{1}{\gamma} \leq |\xi| \leq \gamma \right\},$$

we let

$$\begin{aligned} f_j(x) &= \psi(a^{-j} D)(x) \\ &:= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \psi\left(a^{-j}\xi\right) \hat{f}(\xi)\, d\xi, \quad f \in H_p^s(\mathbf{R}^n). \end{aligned}$$

Then there exists a constant $C > 0$, independent of f, such that

$$\left\| \left(\sum_{j=0}^\infty a^{2sj} |f_j|^2 \right)^{1/2} \right\|_{L^p} \leq C \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}. \tag{A.33}$$

(ii) If $\{f_j\}$ is a sequence in $\mathcal{S}'(\mathbf{R}^n)$ which satisfies the condition

$$\operatorname{supp} \widehat{f_j} \subset \left\{ \xi \in \mathbf{R}^n : \frac{a^j}{\gamma} \leq |\xi| \leq \gamma a^j \right\},$$

then there exists a constant $C > 0$, independent of $\{f_j\}$, such that

$$\left\| \sum_{j=0}^\infty f_j \right\|_{H_p^s} \leq C \left\| \left(\sum_{j=0}^\infty a^{2sj} |f_j|^2 \right)^{1/2} \right\|_{L^p}. \tag{A.34}$$

(iii) If $s > 0$, then inequality (A.34) remains valid for every sequence $\{f_j\}$ in $\mathcal{S}'(\mathbf{R}^n)$ which satisfies the condition

$$\operatorname{supp} \widehat{f_j} \subset \left\{ \xi \in \mathbf{R}^n : |\xi| \leq \gamma a^j \right\}.$$

Parts (i) and (ii) are essentially proved in Bergh–Löfström [BL, Theorem 6.4.3], while part (iii) is proved in Meyer [Me, Lemme 5].

A.6.1 Proof of the Case $\delta = 1$

In this subsection, we prove the case where $\delta = 1$ and $0 < s < r$. The proof is divided into five steps.

Step 1: Let $\{M_j\}$ be a bounded sequence in the Hölder space $\Lambda_r(\mathbf{R}^n) = B^r_{\infty,\infty}(\mathbf{R}^n)$ (see [Tr, Theorem 2.5.7]). We take $a = 2$ and let

$$M_j(x) = \sum_{\ell=0}^{\infty} M_{j,\ell}(x) \tag{A.35}$$

be the Littlewood–Paley series of the function $M_j(x)$, that is,

$$\widehat{M_{j,0}}(\xi) = \psi_0(\xi)\widehat{M_j}(\xi),$$
$$\widehat{M_{j,k}}(\xi) = \varphi(2^{-k}\xi)\widehat{M_j}(\xi), \quad k = 1, 2, \ldots,$$

where

$$\text{supp } \varphi = \left\{\xi \in \mathbf{R}^n : \frac{1}{2} \leq |\xi| \leq 2\right\}, \tag{A.36a}$$

$$\sum_{j\in\mathbf{Z}} \varphi(2^{-j}\xi) = 1 \quad \text{for all } \xi \neq 0 \tag{A.36b}$$

and

$$\psi_0 \in C_0^\infty(\mathbf{R}^n), \tag{A.37a}$$

$$\text{supp } \psi_0 = \{\xi \in \mathbf{R}^n : |\xi| \leq 2\}, \tag{A.37b}$$

$$\psi_0(\xi) = 1 - \sum_{j=1}^{\infty} \varphi(2^{-j}\xi) \quad \text{for all } \xi \neq 0. \tag{A.37c}$$

If we introduce functions $\varphi_\ell(x) \in \mathcal{S}(\mathbf{R}^n)$ by the formulas

$$\widehat{\varphi_\ell}(\xi) = \begin{cases} \psi_0(\xi) & \text{if } \ell = 0, \\ \varphi(2^{-k}\xi), & \text{if } \ell = 1, 2, \ldots, \end{cases}$$

then the functions $M_{j,\ell}(x)$ can be expressed in the convolution form

$$M_{j,\ell}(x) = \varphi_\ell * M_j(x), \quad \ell = 0, 1, 2, \ldots.$$

Moreover, the functions $M_{j,\ell}(x)$ are estimated as follows:

Lemma A.19. *Let $r' > 0$ be an arbitrary number such that*

$$[r] < r' < r, \quad [r'] = [r]. \tag{A.38}$$

If we let

$$A = \sup_{j \geq 0} \|M_j\|_{\Lambda^r},$$

then we have, for all $\ell = 0, 1, 2, \ldots,$

$$\|M_{j,\ell}\|_{L^\infty} \leq C\, A\, 2^{-r'\ell}, \tag{A.39}$$

with some constant $C > 0$ *independent of* ℓ.

Proof. By Part (i) of Lemma A.2, it follows that

$$\begin{aligned} \|M_{j,\ell}\|_{L^\infty} &= \|\varphi_\ell * M_j\|_{L^\infty} = \|J^{-r'} \cdot J^{r'}(\varphi_\ell * M_j)\|_{L^\infty} \\ &\leq C_1\, 2^{-r'\ell} \|\varphi_\ell * (J^{r'} M_j)\|_{L^\infty}. \end{aligned} \tag{A.40}$$

However, we have, by Young's inequality (Corollary 5.3),

$$\|\varphi_\ell * (J^{r'} M_j)\|_{L^\infty} \leq \|\varphi_\ell\|_{L^1}\, \|J^{r'} M_j\|_{L^\infty} = \|\hat{\varphi}\|_{L^1}\, \|M_j\|_{H^{r'}_\infty}. \tag{A.41}$$

Since $[r'] = [r] < r' < r$, we have the inclusions

$$\Lambda_r = B^r_{\infty,\infty} \subset B^{r'}_{\infty,1} \subset H^{r'}_\infty,$$

and hence, with some constants $C_2, C_3 > 0$,

$$\|M_j\|_{H^{r'}_\infty} \leq C_2\, \|M_j\|_{B^{r'}_{\infty,1}} \leq C_3\, \|M_j\|_{\Lambda_r} \leq C_3\, A. \tag{A.42}$$

Therefore, by combining inequalities (A.40), (A.41) and (A.42) we obtain that

$$\begin{aligned} \|M_{j,\ell}\|_{L^\infty} &\leq C_1\, 2^{-r'\ell} \|J^{r'} M_j\|_{L^\infty} \leq (C_1\, C_3\, \|\hat{\varphi}\|_{L^1})\, 2^{-r'\ell} \|M_j\|_{\Lambda_r} \\ &\leq (C_1\, C_3\, \|\hat{\varphi}\|_{L^1})\, A\, 2^{-r'\ell}. \end{aligned}$$

This proves the desired inequality (A.39).

Step 2: If we let

$$N_{j,h}(x) = M_{j,h-j}(2^j x), \quad h \geq j, \tag{A.43}$$

then we have the following two Lemmas A.20 and A.21:

Lemma A.20. *The functions* $N_{j,h}(x)$ *are estimated as follows:*

$$\|N_{j,h}\|_{L^\infty} \leq C\, A\, 2^{-(h-j)r'}, \quad h \geq j.$$

Proof. Indeed, it follows from an application of Lemma A.19 that

$$\|N_{j,h}\|_{L^\infty} = \|M_{j,h-j}\|_{L^\infty} \leq C\, A\, 2^{-(h-j)r'}.$$

Lemma A.21. *The spectra of the functions $N_{j,h}(x)$ are estimated as follows:*

$$\operatorname{supp}\, \widehat{N_{j,h}} \subset \left\{\xi \in \mathbf{R}^n : 2^{h-1} \leq |\xi| \leq 2^{h+1}\right\}, \quad h \geq j. \tag{A.44}$$

Proof. By formula (A.43), we can rewrite the function $\widehat{N_{j,h}}(\xi)$ as follows:

$$\begin{aligned}\widehat{N_{j,h}}(\xi) &= \int_{\mathbf{R}^n} e^{-ix\cdot\xi} M_{j,h-j}(2^j x)\, dx \\ &= \int_{\mathbf{R}^n} e^{-iy\cdot(\xi/2^j)} M_{j,h-j}(y)\, \frac{dy}{2^{jn}} = \frac{1}{2^{jn}} \widehat{M_{j,h-j}}(2^{-j}\xi) \\ &= \frac{1}{2^{jn}} \varphi(2^{-h}\xi)\, \widehat{M_j}(2^{-j}\xi).\end{aligned}$$

However, since we have

$$\operatorname{supp}\, \varphi = \left\{\xi \in \mathbf{R}^n : \frac{1}{2} \leq |\xi| \leq 2\right\},$$

it follows that

$$\operatorname{supp}\, \widehat{N_{j,h}} \subset \{\xi \in \mathbf{R}^n : 2^{h-1} \leq |\xi| \leq 2^{h+1}\}.$$

Step 3: Now, by using decomposition (A.8) with $\delta := 1$ we express the function $\sigma(x, D) f$ in terms of the functions $N_{j,h}$ and f_j.

By (A.35) and (A.43), it follows that

$$\begin{aligned}\sigma(x, D) f(x) &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, \sigma(x, \xi)\, \hat{f}(\xi)\, d\xi \\ &= \sum_{j=0}^{\infty} \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, M_j(2^j x)\, \psi(2^{-j}\xi)\, \hat{f}(\xi)\, d\xi \\ &= \sum_{j=0}^{\infty} \frac{M_j(2^j x)}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi}\, \psi(2^{-j}\xi)\, \hat{f}(\xi)\, d\xi \\ &= \sum_{j=0}^{\infty} M_j(2^j x)\, \psi(2^{-j} D) f(x) = \sum_{j=0}^{\infty} M_j(2^j x)\, f_j(x)\end{aligned} \tag{A.45}$$

$$= \sum_{j=0}^{\infty} \left(\sum_{\ell=0}^{\infty} M_{j,\ell}(2^j x) \right) f_j(x) = \sum_{j=0}^{\infty} \left(\sum_{h=j}^{\infty} M_{j,h-j}(2^j x) \right) f_j(x)$$
$$= \sum_{j=0}^{\infty} \left(\sum_{h=j}^{\infty} N_{j,h}(x) \right) f_j(x),$$

where

$$f_j(x) = \psi(2^{-j} D)(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \psi\left(2^{-j}\xi\right) \hat{f}(\xi)\, d\xi.$$

However, the spectra of the functions $N_{j,h}(x) f_j(x)$ are estimated as follows:

Lemma A.22. *We have, for all* $h \geq j$,

$$\operatorname{supp} \widehat{N_{j,h} f_j} \subset \begin{cases} \left\{\xi \in \mathbf{R}^n : \frac{1}{3} \cdot 2^h \leq |\xi| \leq 3 \cdot 2^h \right\} & \text{if } h \geq j+5, \\ \left\{\xi \in \mathbf{R}^n : |\xi| \leq 35 \cdot 2^j \right\}. & \text{if } j \leq h \leq j+4. \end{cases}$$

Proof. First, it should be noted that

$$\widehat{N_{j,h} f_j}(\xi) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \widehat{N_{j,h}}(\xi - \eta)\, \psi(2^{-j}\eta)\, \hat{f}(\eta)\, d\eta,$$

and that

$$\operatorname{supp} \widehat{N_{j,h}}(\cdot - \eta) \subset \{\xi \in \mathbf{R}^n : 2^{h-1} \leq |\xi - \eta| \leq 2^{h+1}\},$$
$$\operatorname{supp} \psi(2^{-j}\cdot) = \left\{\eta \in \mathbf{R}^n : \frac{1}{3} \cdot 2^j \leq |\eta| \leq 3 \cdot 2^j \right\}.$$

Therefore, we can estimate the support of $\widehat{N_{j,h} f_j}$ as follows.

Case (A): $h \geq j+5$. In this case, we have the estimate

$$\begin{aligned} |\xi| &\leq |\xi - \eta| + |\eta| \leq 2^{h+1} + 3 \cdot 2^j = 2^h(2 + 3 \cdot 2^{j-h}) \\ &\leq 2^h \left(2 + 3 \cdot 2^{-5}\right) \\ &\leq 3 \cdot 2^h. \end{aligned}$$

Similarly, we have the estimate

$$\begin{aligned}|\xi| &\geq |\xi-\eta| - |\eta| \\ &\geq 2^{h-1} - 3\cdot 2^j = 2^h\left(\frac{1}{2} - 3\cdot 2^{j-h}\right) \geq 2^h\left(\frac{1}{2} - 3\cdot\frac{1}{32}\right) \\ &\geq \frac{1}{3}\cdot 2^h.\end{aligned}$$

This proves that

$$\operatorname{supp}\widehat{N_{j,h}f_j} \subset \left\{\xi \in \mathbf{R}^n : \frac{1}{3}\cdot 2^h \leq |\xi| \leq 3\cdot 2^h\right\}$$

if $h \geq j+5$.

Case (B): $j \leq h \leq j+4$. In this case, we have the estimates

$$\begin{aligned}|\xi-\eta| &\leq 2^{h+1} \leq 2^{j+5}, \\ |\eta| &\leq 3\cdot 2^j,\end{aligned}$$

and hence

$$|\xi| \leq |\xi-\eta| + |\eta| \leq 2^{j+5} + 3\cdot 2^j = 35\cdot 2^j.$$

This proves that

$$\operatorname{supp}\widehat{N_{j,h}f_j} \subset \left\{\xi \in \mathbf{R}^n : |\xi| \leq 35\cdot 2^j\right\}$$

if $j \leq h \leq j+4$.

The proof of Lemma A.22 is complete.

Step 4: By virtue of Lemma A.22, we can rewrite (A.45) in the form

$$\begin{aligned}\sigma(x,D)f &= \sum_{j=0}^{\infty}\left(\sum_{h=j}^{\infty} N_{j,h}\right) f_j \\ &= \sum_{\nu=0}^{4}\left(\sum_{j=0}^{\infty} N_{j,j+\nu} f_j\right) + \sum_{h=5}^{\infty}\left(\sum_{j=0}^{h-5} N_{j,h} f_j\right) \\ &:= g_1 + g_2.\end{aligned}$$

(i) The *estimate* of $g_1(x)$: Since $0 < s < r$, by applying part (iii) of Lemma A.18 with $\gamma := 35$ and $a := 2$ and then Lemma A.20 we obtain that

$$\|g_1\|_{H_p^s} \leq \sum_{\nu=0}^{4} \left\| \sum_{j=0}^{\infty} N_{j,j+\nu} f_j \right\|_{H_p^s} \tag{A.46}$$
$$\leq \sum_{\nu=0}^{4} \left\| \left(\sum_{j=0}^{\infty} 4^{sj} |N_{j,j+\nu} f_j|^2 \right)^{1/2} \right\|_{L^p}$$
$$\leq C_1 A \left\| \left(\sum_{j=0}^{\infty} 4^{sj} |f_j|^2 \right)^{1/2} \right\|_{L^p}.$$

However, we have, by inequality (A.29) with $\gamma := 3$, $a := 2$ and $m := N - n - 1$,

$$\left\| \left(\sum_{j=0}^{\infty} 4^{sj} |f_j|^2 \right)^{1/2} \right\|_{L^p} \leq C_2 \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}.$$

Therefore, by combining this inequality with inequality (A.46) we obtain that

$$\|g_1\|_{H_p^s} \leq C_1 C_2 A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}. \tag{A.47}$$

(ii) The *estimate* of $g_2(x)$: By applying inequality (A.34) with $\gamma := 3$ and $a := 2$ and then Lemma A.20, we obtain that

$$\|g_2\|_{H_p^s} \leq C_3 \left\| \left(\sum_{h=5}^{\infty} 4^{hs} \left| \sum_{j=0}^{h-5} N_{j,h} f_j \right|^2 \right)^{1/2} \right\|_{L^p} \tag{A.48}$$
$$\leq C_4 A \left\| \left(\sum_{h=5}^{\infty} 4^{hs} \left(\sum_{j=0}^{h-5} 2^{(j-h)r'} |f_j| \right)^2 \right)^{1/2} \right\|_{L^p}$$
$$= C_4 A \left\| \left(\sum_{h=5}^{\infty} 4^{h(s-r')} \left(\sum_{j=0}^{h-5} 2^{j(r'-s)} \left(2^{sj} |f_j|\right) \right)^2 \right)^{1/2} \right\|_{L^p}.$$

To estimate the function $g_2(x)$, we need the following elementary result:

Lemma A.23. *If $b > 1$ and $\xi = (\xi_j)$ is an element of the Hilbert space ℓ^2, we let*

$$\eta = (\eta_k),$$
$$\eta_k = \frac{\sum_{j=0}^{k} b^j \xi_j}{b^k}, \ k = 0,\ 1,\ 2, \ldots.$$

Then it follows that $\eta \in \ell^2$ and there exists a constant $C > 1$, independent of ξ, such that

$$\|\eta\|_{\ell^2} \leq C \|\xi\|_{\ell^2}. \tag{A.49}$$

For example, we may take

$$C = \frac{\sqrt{2}\,\sqrt{b^2+1}}{b^2-1}\, b.$$

Proof. (1) First, we have, for any given $\varepsilon > 0$, the elementary inequality

$$\begin{aligned}
&(a_1 + a_2 + \ldots + a_n)^2 \\
&\leq \left(1+\varepsilon^2\right)^{n-1} {a_1}^2 + \left(1+\varepsilon^2\right)^{n-2} \left(1 + \frac{1}{\varepsilon^2}\right) {a_2}^2 \\
&\quad + \ldots + \left(1+\varepsilon^2\right) \left(1 + \frac{1}{\varepsilon^2}\right) {a_{n-1}}^2 \\
&\quad + \left(1 + \frac{1}{\varepsilon^2}\right) {a_n}^2, \qquad a_1,\ a_2, \ldots, a_n \in \mathbf{R}.
\end{aligned} \tag{A.50}$$

(2) Now, by applying inequality (A.50) we obtain that

$$\begin{aligned}
\eta_k^2 &= \left(\sum_{j=0}^{k} b^{j-k} \xi_j\right)^2 \\
&\leq (1+\varepsilon^2)^k \left(\frac{\xi_0}{b^k}\right)^2 + (1+\varepsilon^2)^k \left(\frac{1}{\varepsilon^2}\right) \left(\frac{\xi_1}{b^{k-1}}\right)^2 \\
&\quad + \ldots + (1+\varepsilon^2)^3 \left(\frac{1}{\varepsilon^2}\right) \left(\frac{\xi_{k-2}}{b^2}\right)^2 + (1+\varepsilon^2)^2 \left(\frac{1}{\varepsilon^2}\right) \left(\frac{\xi_{k-1}}{b}\right)^2 \\
&\quad + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) {\xi_k}^2 \\
&= \left(\frac{1+\varepsilon^2}{b^2}\right)^k {\xi_0}^2 + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \Bigg\{ \left(\frac{1+\varepsilon^2}{b^2}\right)^{k-1} {\xi_1}^2
\end{aligned}$$

$$+ \dots + \left(\frac{1+\varepsilon^2}{b^2}\right)^2 {\xi_{k-2}}^2 + \left(\frac{1+\varepsilon^2}{b^2}\right)^2 {\xi_{k-1}}^2 + {\xi_k}^2 \Bigg\}$$

$$= \left(\frac{1+\varepsilon^2}{b^2}\right)^k {\xi_0}^2 + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \sum_{j=1}^{k} \left(\frac{1+\varepsilon^2}{b^2}\right)^{k-j} {\xi_j}^2.$$

Hence it follows that

$$\sum_{k=0}^{\infty} \eta_k^2 \tag{A.51}$$

$$\le \sum_{k=0}^{\infty} \left(\frac{1+\varepsilon^2}{b^2}\right)^k {\xi_0}^2 + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \sum_{k=1}^{\infty} \sum_{j=1}^{k} \left(\frac{1+\varepsilon^2}{b^2}\right)^{k-j} {\xi_j}^2$$

$$= \sum_{k=0}^{\infty} \left(\frac{1+\varepsilon^2}{b^2}\right)^k {\xi_0}^2 + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \sum_{j=1}^{\infty} \left(\sum_{k=j}^{\infty} \left(\frac{1+\varepsilon^2}{b^2}\right)^{k-j} \right) {\xi_j}^2.$$

If we take

$$\varepsilon = \left(\frac{b^2-1}{2}\right)^{1/2},$$

then we have, by inequality (A.51),

$$\sum_{k=0}^{\infty} \eta_k^2 \le \left(\frac{1}{1-((1+\varepsilon^2)/b^2)}\right) {\xi_0}^2 + \left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \left(\frac{1}{1-((1+\varepsilon^2)/b^2)}\right) \sum_{j=1}^{\infty} {\xi_j}^2$$

$$= \left(\frac{2b^2}{b^2-1}\right) {\xi_0}^2 + \left(\frac{b^2+1}{b^2-1}\right) \left(\frac{2b^2}{b^2-1}\right) \sum_{j=1}^{\infty} {\xi_j}^2$$

$$\le \left(\frac{b^2+1}{b^2-1}\right) \left(\frac{2b^2}{b^2-1}\right) \sum_{j=0}^{\infty} {\xi_j}^2.$$

This proves the desired inequality (A.49).

Now, by applying Lemma A.23 with

$$b := 2^{r'-s} > 1, \quad 0 < s < r',$$

$$\xi_j := 2^{sj} \, |f_j|,$$

we obtain that

$$\begin{aligned}
&\sum_{h=5}^{\infty} 4^{h(s-r')} \left(\sum_{j=0}^{h-5} 2^{j(r'-s)} \left(2^{sj} |f_j|\right) \right)^2 \\
&= \sum_{h=5}^{\infty} \left(b^{-(h-5)} \left(\sum_{j=0}^{h-5} b^j \xi_j \right) \right)^2 \times b^{-10} = \sum_{\ell=0}^{\infty} \left(b^{-\ell} \left(\sum_{j=0}^{\ell} b^j \xi_j \right) \right)^{1/2} \times b^{-10} \\
&\le {C_5}^2 \sum_{j=0}^{\infty} |\xi_j|^2 \\
&= {C_5}^2 \sum_{j=0}^{\infty} 4^{sj} |f_j|^2.
\end{aligned}$$

Thus we have, by inequality (A.48),

$$\begin{aligned}
\|g_2\|_{H_p^s} &\le C_4\, C_5\, A \left\| \left(\sum_{j=0}^{\infty} 4^{sj} |f_j|^2 \right)^{1/2} \right\|_{L^p} \qquad \text{(A.52)} \\
&\le C_6\, A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}.
\end{aligned}$$

Step 5: By combining inequalities (A.47) and (A.52), we have proved that, for $0 < s < r'$,

$$\begin{aligned}
\|\sigma(x, D) f\|_{H_p^s} &= \left\| \sum_{j=0}^{\infty} \sum_{h=j}^{\infty} N_{j,h}\, f_j \right\|_{H_p^s} = \|g_1 + g_2\|_{H_p^s} \\
&\le C_7\, A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}.
\end{aligned}$$

Now the proof of Proposition A.12 for $\delta = 1$ and $0 < s < r$ is complete, since $r' > 0$ may be chosen arbitrarily close to r. □

A.6.2 *Proof of the Case* $0 \le \delta < 1$

In this subsection, we prove the case where $0 \le \delta < 1$ and $(\delta - 1)r < s < r$: To do this, it should be noted that

$$S^0_{1,\delta}(N, r) \subset S^0_{1,1}(N, r) \quad \text{for } 0 \le \delta < 1.$$

Hence it follows from the proof of the case $\delta = 1$ that if a symbol $\sigma(x, \xi)$ is in the class $S^0_{1,\delta}(N, r)$, then the operator $\sigma(x, D)$ is bounded on $H^s_p(\mathbf{R}^n)$ for all $0 < s < r$ and $1 < p < \infty$.

Therefore, we have only to consider the case where $(\delta - 1)r < s < 0$, since Proposition A.12 for $s = 0$ follows from an interpolation argument. Indeed, it suffices to note that the space $H^0_p(\mathbf{R}^n) = L^p(\mathbf{R}^n)$ is a complex interpolation space between the spaces $H^\sigma_p(\mathbf{R}^n)$ and $H^{-\sigma}_p(\mathbf{R}^n)$ with $0 < \sigma < (1 - \delta)r$ (see [BL, Theorem 6.4.5] and [Tr, Theorem 2.4.7]). The proof is divided into four steps.

Step 1: Now we assume that $\sigma(x, \xi) \in S^0_{1,\delta}(N, r)$ is an elementary symbol of the form

$$\sigma(x, \xi) = \sum_{j=0}^{\infty} M_j(2^{j\delta} x)\, \psi(2^{-j} \xi), \tag{A.53}$$

where

$$\psi(\xi) \in C_0^\infty(\mathbf{R}^n), \quad \operatorname{supp} \psi = \left\{ \xi \in \mathbf{R}^n : \frac{1}{3} \leq |\xi| \leq 3 \right\}.$$

The sequence $\{M_j\}$ is bounded in the Hölder space $\Lambda_r(\mathbf{R}^n)$.

We take

$$a = 2^{1-\delta} > 1, \quad 0 \leq \delta < 1,$$

and let

$$M_j(x) = \sum_{\ell=0}^{\infty} M_{j,\ell}(x) \tag{A.54}$$

be the Littlewood–Paley series of $M_j(x)$, that is,

$$\widehat{M_{j,0}}(\xi) = \psi_0(\xi) \widehat{M_j}(\xi),$$
$$\widehat{M_{j,k}}(\xi) = \varphi(a^{-k} \xi) \widehat{M_j}(\xi), \quad k = 1, 2, \ldots,$$

where

$$\operatorname{supp} \varphi = \left\{ \xi \in \mathbf{R}^n : \frac{1}{a} \leq |\xi| \leq a \right\}, \tag{A.55a}$$

$$\sum_{j \in \mathbf{Z}} \varphi(a^{-j} \xi) = 1 \quad \text{for all } \xi \neq 0 \tag{A.55b}$$

and

$$\psi_0 \in C_0^\infty(\mathbf{R}^n), \tag{A.56a}$$

$$\text{supp } \psi_0 = \{\xi \in \mathbf{R}^n : |\xi| \leq a\}, \tag{A.56b}$$

$$\psi_0(\xi) = 1 - \sum_{j=1}^{\infty} \varphi(2^{-j}\,\xi) \quad \text{for all } \xi \neq 0. \tag{A.56c}$$

If we introduce functions $\varphi_\ell(x) \in \mathcal{S}(\mathbf{R}^n)$ by the formulas

$$\widehat{\varphi_\ell}(\xi) = \begin{cases} \psi_0(\xi) & \text{if } \ell = 0, \\ \varphi(a^{-\ell}\,\xi), & \text{if } \ell = 1,\ 2, \ldots, \end{cases}$$

then the functions $M_{j,\ell}$ can be expressed in the convolution form

$$M_{j,\ell}(x) = \varphi_\ell * M_j(x), \quad \ell = 0,\ 1,\ 2, \ldots.$$

Moreover, the functions $M_{j,\ell}$ are estimated as follows:

Lemma A.24. *If we let*

$$A = \sup_{j \geq 0} \|M_j\|_{\Lambda^r},$$

then we have, for all $\ell = 0, 1, 2, \ldots,$

$$\|M_{j,\ell}\|_{L^\infty} \leq C\,A\,a^{-r\ell}, \tag{A.57}$$

with some constant $C > 0$ *independent of* ℓ.

Proof. By Part (i) of Lemma A.2, it follows that

$$\begin{aligned} \|M_{j,\ell}\|_{L^\infty} = \|\varphi_\ell * M_j\|_{L^\infty} &= \left\| J^{(\delta-1)r} \cdot J^{(1-\delta)r}(\varphi_\ell * M_j)\right\|_{L^\infty} \\ &\leq C_1\,a^{-r\ell}\|\varphi_\ell * (J^{(1-\delta)r} M_j)\|_{L^\infty}. \end{aligned} \tag{A.58}$$

However, we have, by Young's inequality (Corollary 5.3),

$$\begin{aligned} \|\varphi_\ell * (J^{(1-\delta)r} M_j)\|_{L^\infty} &\leq \|\varphi_\ell\|_{L^1}\,\|J^{(1-\delta)r} M_j\|_{L^\infty} \\ &= \|\hat{\varphi}\|_{L^1}\,\|M_j\|_{H_\infty^{(1-\delta)r}}. \end{aligned} \tag{A.59}$$

Since $0 \leq \delta < 1$, we have the inclusions

$$\Lambda_r = B^r_{\infty,\infty} \subset B^{(1-\delta)r}_{\infty,1} \subset H^{(1-\delta)r}_\infty,$$

and hence, with some constants $C_2, C_3 > 0$,

$$\|M_j\|_{H_\infty^{(1-\delta)r}} \leq C_2 \|M_j\|_{B_{\infty,1}^{(1-\delta)r}} \leq C_3 \|M_j\|_{\Lambda_r} \leq C_3 A. \tag{A.60}$$

Therefore, by combining inequalities (A.58), (A.59) and (A.60) we obtain that

$$\begin{aligned}\|M_{j,\ell}\|_{L^\infty} &\leq C_1 a^{-r\ell} \|J^{(1-\delta)r} M_j\|_{L^\infty} \leq (C_1 C_3 \|\hat{\varphi}\|_{L^1}) a^{-r\ell} \|M_j\|_{\Lambda_r} \\ &\leq (C_1 C_3 \|\hat{\varphi}\|_{L^1}) A a^{-r\ell}.\end{aligned}$$

This proves inequality (A.57).

Step 2: By (A.8) and (A.54), it follows that

$$\begin{aligned}\sigma(x, D) f &= \sum_{j=0}^{\infty} M_j(2^{j\delta} x) \psi(2^{-j} D) f \\ &= \sum_{j=0}^{\infty} \left(\sum_{\ell=0}^{\infty} M_{j,\ell}(2^{j\delta} x) \right) f_j .\end{aligned} \tag{A.61}$$

However, since $0 \leq \delta < 1$, we can find a number $\nu > 2$ such that

$$2^{(1-\nu)(1-\delta)} \leq \frac{1}{12}.$$

Then the spectra of the functions $M_{j,\ell}(2^{j\delta} x) f_j(x)$ are estimated as follows:

Lemma A.25. *We have, for all $\ell \geq 0$,*

$$\operatorname{supp} \widehat{M_{j,\ell} f_j} \subset \begin{cases} \left\{ \xi \in \mathbf{R}^n : \frac{1}{4} \cdot 2^j \leq |\xi| \leq 4 \cdot 2^j \right\} & \text{if } j \geq \ell + \nu, \\ \left\{ \xi \in \mathbf{R}^n : |\xi| \leq \gamma \cdot 2^\ell \right\}. & \text{if } 0 \leq j \leq \ell + \nu - 1, \end{cases}$$

where

$$\gamma = 3 \cdot 2^{(\nu-1)} + 2^{(\nu-2)\delta+1}.$$

Proof. First, it should be noted that

$$\begin{aligned}\widehat{M_{j,\ell} f_j}(\xi) &= \int_{\mathbf{R}^n} e^{-ix\cdot\xi} M_{j,\ell}(2^{j\delta} x) \, f_j(x) \, dx = \widehat{M_{j,\ell} f_j} * \widehat{f_j}(\xi) \\ &= \frac{1}{(2^{j\delta})^n} \int_{\mathbf{R}^n} \widehat{M_{j,\ell}}(2^{-j\delta}(\xi - \eta)) \, \psi(2^{-j}\eta) \, \hat{f}(\eta) \, d\eta,\end{aligned}$$

and that

$$2^{(1-\delta)(\ell-1)} \le \frac{|\xi-\eta|}{2^{j\delta}} \le 2^{(1-\delta)(\ell+1)},$$

$$\frac{1}{3} \le \frac{|\eta|}{2^j} \le 3.$$

Therefore, we can estimate the support of $\widehat{M_{j,h} f_j}$ as follows.

Case (A): $j \ge \ell+\nu$. In this case, we have the estimate

$$\begin{aligned}
|\xi| &\le |\xi-\eta| + |\eta| \\
&\le 2^{(1-\delta)(\ell+1)+j\delta} + 3\cdot 2^j = 2^j\left(3+2^{(1-\delta)(\ell+1)+j(\delta-1)}\right) \\
&\le 2^j\left(3+2^{(1-\delta)(1-\nu)}\right) \le 2^j\left(3+\frac{1}{12}\right) \\
&\le 4\cdot 2^j.
\end{aligned}$$

Similarly, we have the estimate

$$\begin{aligned}
|\xi| &\ge |\xi-\eta| - |\eta| \\
&\ge \frac{1}{3}\cdot 2^j - 2^{(1-\delta)(\ell+1)+j\delta} = 2^j\left(\frac{1}{3}-2^{(1-\delta)(\ell-j)+1-\delta}\right) \\
&\ge 2^j\left(\frac{1}{3}-2^{(1-\delta)(1-\nu)}\right) \ge 2^j\left(\frac{1}{3}-\frac{1}{12}\right) \\
&= \frac{1}{4}\cdot 2^j.
\end{aligned}$$

This proves that

$$\operatorname{supp}\widehat{M_{j,\ell} f_j} \subset \left\{\xi\in\mathbf{R}^n : \frac{1}{4}\cdot 2^j \le |\xi| \le 4\cdot 2^j\right\} \quad \text{if } j \ge \ell+\nu,$$

with $\gamma = 3\cdot 2^{(\nu-1)} + 2^{(\nu-2)\delta+1}$.

Case (B): $0 \le j \le \ell+\nu-1$. In this case, we have the estimate

$$\begin{aligned}
|\xi| &\le |\xi-\eta| + |\eta| \le 2^{(1-\delta)(\ell+1)+j\delta} + 3\cdot 2^j \\
&\le 2^\ell\left(3\cdot 2^{(\nu-1)} + 2^{(\nu-2)\delta+1}\right).
\end{aligned}$$

This proves that

$$\operatorname{supp}\widehat{M_{j,\ell} f_j} \subset \left\{\xi\in\mathbf{R}^n : |\xi| \le \gamma\cdot 2^\ell\right\} \quad \text{if } 0 \le j \le \ell+\nu-1.$$

The proof of Lemma A.25 is complete.

Step 3: By virtue of Lemma A.25, we can rewrite (A.61) in the form

$$\begin{aligned}\sigma(x, D) f &= \sum_{j=\nu}^{\infty} \left(\sum_{\ell=0}^{j-\nu} M_{j,\ell}(2^{j\delta} x) \right) f_j \\ &\quad + \sum_{\ell=0}^{\infty} \left(\sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta} x) f_j \right) \\ &:= g_1 + g_2.\end{aligned}$$

(i) The *estimate* of $g_1(x)$: By applying inequality (A.34) with $\gamma := 4$ and $a := 2$, we obtain that

$$\|g_1\|_{H_p^s} \le C_1 \left\| \left(\sum_{j=\nu}^{\infty} 4^{sj} \left| \sum_{\ell=0}^{j-\nu} M_{j,\ell}(2^{j\delta} x) \right|^2 |f_j|^2 \right)^{1/2} \right\|_{L^p}.$$

However, by Lemma A.24 it follows that

$$\begin{aligned}\left| \sum_{\ell=0}^{j-\nu} M_{j,\ell}(2^{j\delta} x) \right| &\le \sum_{\ell=0}^{j-\nu} \|M_{j,\ell}\|_{L^\infty} \le C_2 A \sum_{\ell=0}^{j-\nu} \left(\frac{1}{2^{(1-\delta)r}} \right)^{\ell} \\ &\le C_3 A.\end{aligned}$$

Hence we have, by inequality (A.33) with $\gamma := 3$, $a := 2$ and $m := N - n - 1$,

$$\begin{aligned}\|g_1\|_{H_p^s} &\le C_1 C_3 A \left\| \left(\sum_{j=\nu}^{\infty} 4^{sj} |f_j|^2 \right)^{1/2} \right\|_{L^p} \\ &\le C_4 A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}.\end{aligned} \tag{A.62}$$

(ii) The *estimate* of $g_2(x)$: First, it follows that

$$\begin{aligned}&\sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left| \sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta} x) f_j(x) \right|^2 \\ &\le \sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left(\sum_{j=0}^{\ell+\nu-1} |M_{j,\ell}(2^{j\delta} x)| \, |f_j(x)|^2 \right)^2.\end{aligned}$$

However, we have, by Lemma A.24,

$$\begin{aligned}\sum_{j=0}^{\ell+\nu-1} |M_{j,\ell}(2^{j\delta}x)|\,|f_j(x)|^2 &\leq \sum_{j=0}^{\ell+\nu-1} \|M_{j,\ell}\|_{L^\infty}\,|f_j(x)| \\ &\leq C_5\,A\,2^{(\delta-1)r\ell} \sum_{j=0}^{\ell+\nu-1} |f_j(x)|.\end{aligned}$$

Hence we obtain that

$$\begin{aligned}&\sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left| \sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta}x)\,f_j(x) \right|^2 \qquad \text{(A.63)} \\ &\leq C_5{}^2\,A^2 \sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell}\,4^{(\delta-1)r\ell} \left(\sum_{j=0}^{\ell+\nu-1} |f_j(x)| \right)^2 \\ &= C_5{}^2\,A^2 \sum_{\ell=0}^{\infty} 4^{s\ell} \left(\sum_{j=0}^{\ell+\nu-1} |f_j(x)| \right)^2.\end{aligned}$$

If we let

$$\begin{aligned}&b := 2^{-s} > 1, \quad (\delta-1)r < s < 0, \\ &\xi_j := 2^{sj}\,|f_j|,\end{aligned}$$

then we can write the last term in inequality (A.63) in the form

$$\begin{aligned}&\sum_{\ell=0}^{\infty} 4^{s\ell} \left(\sum_{j=0}^{\ell+\nu-1} |f_j(x)| \right)^2 \\ &= \sum_{\ell=0}^{\infty} b^{-2(\ell+\nu-1)} \left(\sum_{j=0}^{\ell+\nu-1} b^j\,\xi_j| \right)^2 \times 4^{s(1-\nu)}.\end{aligned}$$

Therefore, by applying Lemma A.23 we obtain from inequality (A.63) that

$$\sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left| \sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta}x)\,f_j(x) \right|^2 \leq C_5{}^2\,A^2 \sum_{j=0}^{\infty} 4^{sj}\,|f_j|^2.$$

By raising this inequality to the power $p/2$ and then integrating, it follows that

$$\left\| \left(\sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left| \sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta}x) f_j(x) \right|^2 \right)^{1/2} \right\|_{L^p}^p \tag{A.64}$$

$$\leq C_5{}^p A^p \left\| \sum_{j=0}^{\infty} 4^{sj} |f_j|^2 \right\|_{L^p}^p .$$

On the other hand, since we have, by Lemma A.25,

$$g_2(x) = \sum_{\ell=0}^{\infty} \left(\sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta}x) f_j(x) \right),$$

$$\operatorname{supp} \widehat{M_{j,\ell} f_j} \subset \left\{ \xi \in \mathbf{R}^n : |\xi| \leq \gamma \cdot 2^{\ell} \right\} \quad \text{if } 0 \leq j \leq \ell + \nu - 1,$$

it follows from an application of part (iii) of Lemma A.18 with $s := s+(1-\delta)r > 0$ that

$$\|g_2\|_{H_p^{s+(1-\delta)r}} \leq C_6 \left\| \left(\sum_{\ell=0}^{\infty} 4^{(s+(1-\delta)r)\ell} \left| \sum_{j=0}^{\ell+\nu-1} M_{j,\ell}(2^{j\delta}x) f_j \right|^2 \right)^{1/2} \right\|_{L^p} . \tag{A.65}$$

By combining inequalities (A.65) and (A.64) and using inequality (A.33) with $a := 2$ and $\gamma := 3$, we obtain that

$$\|g_2\|_{H_p^{s+(1-\delta)r}} \leq C_5\, C_6\, A \left\| \left(\sum_{j=0}^{\infty} 4^{sj} |f_j|^2 \right)^{1/2} \right\|_{L^p}$$

$$\leq C_7\, A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}.$$

This implies the estimate

$$\|g_2\|_{H_p^s} \leq C_8 \|g_2\|_{H_p^{s+(1-\delta)r}} \leq C_7\, C_8\, A \|\psi\|_{L^\infty_{N-n-1}} \|f\|_{H_p^s}, \tag{A.66}$$

since for $H_p^{s+(1-\delta)r} \subset H_p^s$ for $0 \leq \delta < 1$ and $r > 0$.

Step 4: Summing up, we obtain from inequalities (A.62) and (A.66) that

$$\|\sigma(x, D) f\|_{H_p^s} = \left\| \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} M_{j,\ell}(2^{j\delta})\, f_j \right\|_{H_p^s} = \|g_1 + g_2\|_{H_p^s}$$
$$\leq C_9\, A \|\psi\|_{L^{\infty}_{N-n-1}} \|f\|_{H_p^s}.$$

The proof of Proposition A.12 for $0 \leq \delta < 1$ and $(\delta - 1)r < s < 0$ is now complete. □

Appendix B
The Boutet de Monvel Calculus via Operator-Valued Pseudo-differential Operators

In this appendix, following faithfully Schrohe [Sr5], we present a short introduction to the Boutet de Monvel calculus on the half-space $\mathbf{R}^n_+$ in the framework of a pseudo-differential calculus with operator-valued symbols, which shows the pseudo-differential spirit of Boutet de Monvel's construction more clearly than the older descriptions. We modify Schrohe's paper [Sr5] in such a fashion that a broad spectrum of readers will be able to understand the Boutet de Monvel calculus.

B.1 Introduction

Let X be an n-dimensional compact smooth manifold with boundary ∂X. As in Sect. 7.1, we may assume that $\overline{X} = X \cup \partial X$ is the closure of a relatively compact subset X of an n-dimensional compact smooth manifold $\hat{X}$ without boundary in which X has a smooth boundary ∂X (see Fig. B.1). This manifold $\hat{X}$ is the *double* of X.

Boutet de Monvel considers matrices of operators

$$A = \begin{pmatrix} P_+ + G & K \\ T & S \end{pmatrix} : \begin{matrix} C^\infty(X, E_1) \\ \oplus \\ C^\infty(\partial X, F_1) \end{matrix} \longrightarrow \begin{matrix} C^\infty(X, E_2) \\ \oplus \\ C^\infty(\partial X, F_2) \end{matrix} \tag{B.1}$$

Here E_1, E_2 are vector bundles over X, and F_1, F_2 are vector bundles over ∂X; each of them might be zero.

(1) P is a pseudo-differential operator on the double $\hat{X}$ of X; the subscript $+$ indicates that the action of P_+ is defined by extending the function by zero to the full manifold X, applying P, and then restricting the result to X. That is, $P_+ u$ is defined by the formula

$$P_+ u = P(u^0)|_X ,$$

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*, Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7,

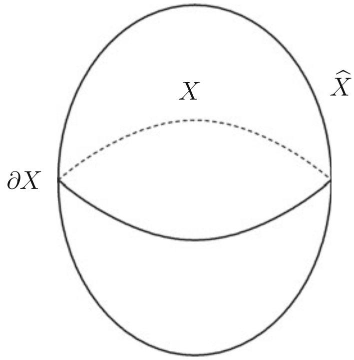


Fig. B.1 The double $\hat{X}$ of X

where u^0 is the extension of u to $\hat{X}$ by zero outside of X

$$u^0(x) = \begin{cases} u(x) & \text{if } x \in X, \\ 0 & \text{if } x \in \hat{X} \setminus X. \end{cases}$$

(2) S is a usual pseudo-differential operator on the boundary ∂X.
(3) K and T are generalizations of the potential and trace operators known from the theory of boundary value problems.
(4) The entry G, a so-called singular Green operator, is an operator which is smoothing in the interior while it acts like a pseudo-differential operator in directions tangential to the boundary ∂X. As an example, we may take the difference of two solution operators to (invertible) classical boundary value problems with the same differential part in the interior but different boundary conditions.

Given an arbitrary pseudo-differential operator P, it is in general not true that P_+ maps functions which are smooth up to the boundary ∂X into functions with the same property. Therefore, the above mapping property will not hold true if we admit all pseudo-differential operators. The crucial requirement here is that the symbol of P has the transmission property, which will be discussed below in detail. On one hand, this restricts the class of boundary value problems in the calculus, on the other hand, however, it ensures that solutions of elliptic equations with smooth data are smooth. Therefore, the transmission property helps to avoid problems with singularities of solutions at the boundary.

It is also a central point that these operator matrices form an algebra in the following sense: Given another element of the calculus, say,

$$A' = \begin{pmatrix} P'_+ + G' & K' \\ T' & S' \end{pmatrix} : \begin{array}{c} C^\infty(X, E_2) \\ \oplus \\ C^\infty(\partial X, F_2) \end{array} \longrightarrow \begin{array}{c} C^\infty(X, E_3) \\ \oplus \\ C^\infty(\partial X, F_3) \end{array}$$

acting from vector bundles E_2, F_2 to vector bundles E_3, F_3, the composition $A'A$ is again an operator matrix of the type described above. This is far from being obvious. For example, consider one of the terms arising in the matrix composition, that is, the product $P'_+ P_+$. Except for special cases, it will not coincide with $(P'P)_+$; in fact, the difference

$$L(P', P) = P'_+ P_+ - (P'P)_+$$

turns out to be a singular Green operator.

The presentation in Boutet de Monvel's original paper is rather concise. This appendix provides a self-contained introduction to the calculus on the half-space $\mathbf{R}^n_+$ in terms of operator-valued symbol classes satisfying uniform estimates. This text addresses primarily those readers who are familiar with the standard pseudo-differential calculus as it is presented, for example, in the book of Kumano-go [Ku]. We have not included a section on coordinate invariance and the construction of the manifold. For one thing, this allowed us to keep the exposition short. Moreover, there is no new aspect to be developed in this direction.

B.2 Symbol Classes

B.2.1 General Notation

In the sequel, $H^s(\mathbf{R}^q)$, $s \in \mathbf{R}$, will denote the usual Sobolev space of L^2 type on Euclidean space $\mathbf{R}^q$. For any $s = (s_1, s_2) \in \mathbf{R}^2$, we introduce a weighted Sobolev space by the formula

$$H^{s}\left(\mathbf{R}^q\right) = H^{(s_1,s_2)}\left(\mathbf{R}^q\right) = \left\{\langle x\rangle^{-s_2} u : u \in H^{s_1}\left(\mathbf{R}^q\right)\right\},$$

where

$$x = (x_1, x_2, \dots, x_q) \in \mathbf{R}^q, \quad \langle x\rangle = (1 + |x|^2)^{1/2}.$$

In particular, we have, for any $s_1 \in \mathbf{R}$,

$$H^{(s_1,0)}\left(\mathbf{R}^q\right) = H^{s_1}\left(\mathbf{R}^q\right).$$

We let

$$\mathcal{S}\left(\mathbf{R}^q\right) = \text{ the Schwartz space of all rapidly decreasing functions on } \mathbf{R}^q.$$

For a Fréchet space E, $\mathcal{S}(\mathbf{R}^q, E)$ is the vector-valued analog. The dual spaces are $\mathcal{S}'(\mathbf{R}^q)$ and $\mathcal{S}'(\mathbf{R}^q, E) = \mathcal{L}(\mathcal{S}(\mathbf{R}^q), E)$, respectively. The Fourier transform

$$\mathcal{F} : \mathcal{S}\left(\mathbf{R}^q, E\right) \longrightarrow \mathcal{S}\left(\mathbf{R}^q, E\right)$$

will in general be indicated by a hat: For every $u \in \mathcal{S}(\mathbf{R}^q, E)$, we let

$$\hat{u}(\xi) = (\mathcal{F}u)(\xi) = \int_{\mathbf{R}^q} e^{-ix\cdot\xi} u(x)\, dx, \quad \xi \in \mathbf{R}^q.$$

Given a distribution u on $\mathbf{R}^q$, we write r^+u for its restriction to the upper half-space $\mathbf{R}^q_+$:

$$r^+u = u|_{\mathbf{R}^q_+},$$

where

$$\mathbf{R}^q_+ = \left\{x = (x_1, x_2, \ldots, x_{q-1}, x_q) \in \mathbf{R}^q : x_q > 0\right\}.$$

For any $s = (s_1, s_2) \in \mathbf{R}^2$, we define a weighted Sobolev space on $\mathbf{R}^q_+$ by the formula

$$H^s\left(\mathbf{R}^q_+\right) = \left\{r^+u : u \in H^s\left(\mathbf{R}^q\right)\right\},$$

and also

$$\mathcal{S}\left(\mathbf{R}^q_+\right) = \left\{r^+u : u \in \mathcal{S}\left(\mathbf{R}^q\right)\right\}.$$

Moreover, $H^s_0(\overline{\mathbf{R}^q_+})$ denotes the weighted Sobolev space of all $u \in H^s(\mathbf{R}^q)$ which are supported in $\overline{\mathbf{R}^q_+}$:

$$H^s_0\left(\overline{\mathbf{R}^q_+}\right) = \left\{u \in H^s\left(\mathbf{R}^q\right) : \operatorname{supp} u \subset \overline{\mathbf{R}^q_+}\right\}.$$

We note that

$$\mathcal{S}\left(\mathbf{R}^q_+\right) = \text{proj} - \lim_{s_1, s_2 \to \infty} H^{(s_1, s_2)}\left(\mathbf{R}^q_+\right),$$

$$\mathcal{S}'\left(\mathbf{R}^q_+\right) = \text{ind} - \lim_{s_1, s_2 \to -\infty} H^{(s_1, s_2)}_0\left(\overline{\mathbf{R}^q_+}\right).$$

B.2.2 Group Actions

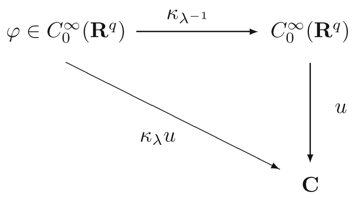
Let E, F be Banach spaces and let $\mathcal{L}(E, F)$ be the space of all continuous linear operators on E into F. A strongly continuous group action on E is a family $\kappa = \{\kappa_\lambda : \lambda \in \mathbf{R}_+\}$ of isomorphisms in $\mathcal{L}(E, E)$ such that $\kappa_\lambda \kappa_\mu = \kappa_{\lambda\mu}$ and the mapping: $\lambda \to \kappa_\lambda e$ is continuous for every $e \in E$. For all Sobolev spaces $H^s(\mathbf{R}^q)$ and $H^s(\mathbf{R}^q_+)$, we shall use the group action defined on functions by the formula

$$(\kappa_\lambda u)(x) = \lambda^{q/2} u(\lambda x), \quad \lambda \in \mathbf{R}_+. \tag{B.2}$$

This group action extends to distributions $u \in \mathcal{D}'(\mathbf{R}^q)$ by the formula

$$(\kappa_\lambda u)(\varphi) = u(\kappa_{\lambda^{-1}} \varphi) \quad \text{for all } \varphi \in C_0^\infty(\mathbf{R}^q).$$

The situation can be visualized in the following diagram:



It should be noted that if u is a function on $\mathbf{R}^q$, we have

$$\begin{aligned}(\kappa_\lambda u)(\varphi) &= \int_{\mathbf{R}^q} \lambda^{q/2} u(\lambda x)\, \varphi(x)\, dx = \int_{\mathbf{R}^q} \lambda^{-q/2} u(z)\, \varphi(\lambda^{-1} z)\, dz \\ &= u(\kappa_{\lambda^{-1}} \varphi) \quad \text{for all } \varphi \in C_0^\infty(\mathbf{R}^q).\end{aligned}$$

On $E = \mathbf{C}^\ell$ with $\ell \in \mathbf{N}$, we use the trivial group action $\kappa_\lambda \equiv \mathrm{id}$. Sums of spaces of the above kind will be endowed with the sum of the group actions.

Lemma B.1. *There are constants c and M such that*

$$\|\kappa_\lambda\|_{\mathcal{L}(E,E)} \le c \max\left\{\lambda, \lambda^{-1}\right\}^M \quad \text{for all } \lambda \in \mathbf{R}_+.$$

Indeed, Lemma B.1 follows easily from the well-known statement for additive semigroups (see Hirschmann [Hi, Remark 2.2]).

B.2.3 Operator-Valued Symbols

Let E, F be Banach spaces with strongly continuous group actions κ and $\tilde{\kappa}$, respectively. If $a(y, \tilde{y}, \eta)$ is a function in $C^\infty(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \mathcal{L}(E, F))$ and $\mu \in \mathbf{R}$, then we shall write

$$a \in S^\mu \left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right),$$

provided that, for all multi-indices α, β, γ there exists a positive constant $C = C(\alpha, \beta, \gamma)$ such that

$$\left\| \tilde{\kappa}_{\langle\eta\rangle^{-1}} D_\eta^\alpha D_y^\beta D_{\tilde{y}}^\gamma a(y, \tilde{y}, \eta) \kappa_{\langle\eta\rangle} \right\|_{\mathcal{L}(E,F)} \leq C \left\langle\eta\right\rangle^{\mu-|\alpha|}, \quad \langle\eta\rangle = (1 + |\eta|^2)^{1/2}.$$

The situation can be visualized in the following diagram:

$$\begin{array}{ccc} E & \xrightarrow{a(y,\tilde{y},\eta)} & F \\ \uparrow \kappa_{\langle\eta\rangle} & & \tilde{\kappa}_{\langle\eta\rangle^{-1}} \downarrow \\ E & \longrightarrow & F \end{array}$$

If the operator-valued symbol $a(y, \tilde{y}, \eta)$ is independent of y or $\tilde{y}$, we shall write $a \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; E, F)$.

For $E = F = \mathbf{C}$, we recover the definition of the usual symbol class $S^\mu_{1,0}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q)$ as follows:

$$S^\mu \left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \mathbf{C}, \mathbf{C}\right) = S^\mu_{1,0} \left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q\right).$$

Example B.2 (Trace operators). Let $\gamma_j : \mathcal{S}(\mathbf{R}_+) \to \mathbf{C}$ be a mapping defined by the formula

$$\gamma_j f = \lim_{t \downarrow 0} \frac{\partial^j f}{\partial x_n^j}(t) = f^{(j)}(0), \quad j = 0, 1, \ldots.$$

The trace theorem for Sobolev spaces (Theorem 6.6 with $p := 2$) guarantees that γ_j extends to an element of $\mathcal{L}(H^\sigma(\mathbf{R}_+), \mathbf{C})$ provided that $\sigma_1 > j + 1/2$. We can also view γ_j as an operator-valued symbol independent of the variables y and η; then we have

$$\gamma_j \in S^{j+1/2} \left(\mathbf{R}^q \times \mathbf{R}^q; H^\sigma \left(\mathbf{R}_+\right), \mathbf{C}\right).$$

Indeed, recalling that the group action on $H^\sigma(\mathbf{R}_+)$ is given by formula (B.2) with $q := 1$ while on $\mathbf{C}$ we choose the identity, we only have to verify that

$$\left\| \gamma_j \kappa_{\langle \eta \rangle} \right\|_{\mathcal{L}(H^\sigma(\mathbf{R}_+),\mathbf{C})} = O(\langle \eta \rangle^{j+1/2}).$$

This is immediate, since we have

$$\frac{\partial^j}{\partial t^j}\left(\langle \eta \rangle^{1/2} f(\langle \eta \rangle\, t) \right) = \langle \eta \rangle^{j+1/2} \frac{\partial^j f}{\partial t^j}(\langle \eta \rangle\, t) \quad \text{for all } f \in \mathcal{S}(\mathbf{R}_+).$$

Definition B.3. For an operator-valued symbol $a(y, \tilde{y}, \eta) \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$, the pseudo-differential operator

$$\mathrm{Op}(a) : \mathcal{S}\left(\mathbf{R}^q, E\right) \longrightarrow \mathcal{S}\left(\mathbf{R}^q, F\right)$$

is defined by the formula

$$[\mathrm{Op}(a)]\, u(y) = \frac{1}{(2\pi)^q} \iint_{\mathbf{R}^q \times \mathbf{R}^q} e^{i(y-\tilde{y})\cdot\eta} a(y, \tilde{y}, \eta) u(\tilde{y})\, d\tilde{y}\, d\eta, \quad u \in \mathcal{S}\left(\mathbf{R}^q, E\right).$$

If the operator-valued symbol $a(y, \tilde{y}, \eta)$ is independent of $\tilde{y}$, this formula reduces to the following:

$$[\mathrm{Op}(a)]\, u(y) = \frac{1}{(2\pi)^q} \int_{\mathbf{R}^q} e^{iy\cdot\eta} a(y, \eta) \hat{u}(\eta)\, d\eta.$$

In this case, we call $a(y, \eta)$ a *left symbol* for $\mathrm{Op}(a)$. If $a(y, \tilde{y}, \eta)$ is independent of y, then we have

$$[\mathrm{Op}(a)]\, u(y) = \frac{1}{(2\pi)^q} \iint_{\mathbf{R}^q \times \mathbf{R}^q} e^{i(y-\tilde{y})\cdot\eta} a(\tilde{y}, \eta) u(\tilde{y})\, d\tilde{y}\, d\eta,$$

and $a(\tilde{y}, \eta)$ is called a *right symbol* for $\mathrm{Op}(a)$.

Example B.4 (Action in the normal direction). We let

$$a(x', x_n, \xi', \xi_n) \in S^\mu_{1,0}\left(\mathbf{R}^n \times \mathbf{R}^n\right)$$

with $\mu \in \mathbf{R}$. For fixed $(x', \xi') \in \mathbf{R}^{n-1} \times \mathbf{R}^{n-1}$, the function $a(x', \cdot, \xi', \cdot)$ is an element of the symbol class

$$S^\mu_{1,0}\left(\mathbf{R} \times \mathbf{R}\right).$$

For every $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$, the function $a(x', \cdot, \xi', \cdot)$ induces a bounded linear operator

$$\operatorname{Op}_{x_n}(a) = \operatorname{Op}_{x_n} a(x', x_n, \xi', \xi_n) : H^{(\sigma_1,\sigma_2)}(\mathbf{R}) \longrightarrow H^{(\sigma_1-\mu,\sigma_2)}(\mathbf{R}),$$

that is,

$$\operatorname{Op}_{x_n}(a) = \operatorname{Op}_{x_n} a(x', x_n, \xi', \xi_n) : H^{\sigma}(\mathbf{R}) \longrightarrow H^{\sigma-(\mu,0)}(\mathbf{R})$$

(cf. Schrohe [Sr1, Theorem 1.7]). Here the subscript x_n indicates that the action is with respect to the variable x_n and the covariable ξ_n only. Then we have

$$\kappa_{\langle\xi'\rangle^{-1}} \left[\operatorname{Op}_{x_n}(a)\right] \kappa_{\langle\xi'\rangle} = \operatorname{Op}_{x_n} a(x', x_n/\langle\xi'\rangle, \xi', \langle\xi'\rangle \xi_n). \tag{B.3}$$

The situation can be visualized in the following diagram:

$$\begin{array}{ccc}
H^{\sigma}(\mathbf{R}) & \xrightarrow{\quad \operatorname{Op}_{x_n}(a) \quad} & H^{\sigma-(\mu,0)}(\mathbf{R}) \\
\uparrow \kappa_{\langle\xi'\rangle} & & \kappa_{\langle\xi'\rangle^{-1}} \downarrow \\
H^{\sigma}(\mathbf{R}) & \xrightarrow[\operatorname{Op}_{x_n} a(x', x_n/\langle\xi'\rangle, \xi', \langle\xi'\rangle\xi_n)]{} & H^{\sigma-(\mu,0)}(\mathbf{R})
\end{array}$$

Indeed, we have, for all $u \in \mathcal{S}(\mathbf{R})$,

$$\begin{aligned}
&\kappa_{\langle\xi'\rangle^{-1}} \left[\operatorname{Op}_{x_n}(a)\right] (\kappa_{\langle\xi'\rangle} u)(x_n) \\
&= \frac{1}{2\pi} \int_{\mathbf{R}} \langle\xi'\rangle^{-1/2} e^{i x_n/\langle\xi'\rangle \cdot \eta_n} a(x', x_n/\langle\xi'\rangle, \xi', \eta_n) \langle\xi'\rangle^{-1/2} \hat{u}(\eta_n/\langle\xi'\rangle)\, d\eta_n \\
&= \frac{1}{2\pi} \int_{\mathbf{R}} e^{i x_n \xi_n} a(x', x_n/\langle\xi'\rangle, \xi', \langle\xi'\rangle \xi_n) \hat{u}(\xi_n)\, d\xi_n \\
&= \left[\operatorname{Op}_{x_n} a(x', x_n/\langle\xi'\rangle, \xi', \langle\xi'\rangle \xi_n)\right] u(x_n).
\end{aligned}$$

The next theorem shows that $\operatorname{Op}_{x_n}(a)$ is an operator-valued symbol in the sense of Sect. B.2.3:

Theorem B.5. *If $a(x,\xi) \in S^{\mu}_{1,0}(\mathbf{R}^n \times \mathbf{R}^n)$ and $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$, then we have*

$$\operatorname{Op}_{x_n}(a) \in S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}), H^{\sigma-(\mu,0)}(\mathbf{R})\right).$$

Proof. Given multi-indices α and β, it follows from Example B.4 that we have to estimate the following:

$$\begin{aligned}
&\sup_{x',\xi'} \left\| \langle\xi'\rangle^{|\alpha|} \kappa_{\langle\xi'\rangle^{-1}} [\operatorname{Op}_{x_n}(D^{\alpha}_{\xi'} D^{\beta}_{x'} a)] \kappa_{\langle\xi'\rangle} \right\|_{\mathcal{L}(H^{\sigma}(\mathbf{R}), H^{\sigma-(\mu,0)}(\mathbf{R}))} \\
&= \sup_{x',\xi'} \left\| \langle\xi'\rangle^{|\alpha|} [\operatorname{Op}_{x_n}(D^{\alpha}_{\xi'} D^{\beta}_{x'} a)](x', x_n/\langle\xi'\rangle, \xi', \xi_n \langle\xi'\rangle) \right\|_{\mathcal{L}(H^{\sigma}(\mathbf{R}), H^{\sigma-(\mu,0)}(\mathbf{R}))}.
\end{aligned}$$

Since $D_{\xi'}^{\alpha} D_{x'}^{\beta} a$ is of order $\mu - |\alpha|$, we may assume that $\alpha = \beta = 0$. Now we find that the operator

$$\mathrm{Op}_{x_n} a(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle) : H^{\sigma}(\mathbf{R}) \longrightarrow H^{\sigma-(\mu,0)}(\mathbf{R})$$

is continuous, and a bound for its norm is given by the suprema

$$\sup\left\{\left|D_{\xi_n}^{\gamma} D_{x_n}^{\delta}\{a(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle)\}\right| \langle\xi_n\rangle^{-\mu} : x_n, \xi_n \in \mathbf{R}\right\}$$

for a finite number of derivatives. Since each of them is finite, the desired estimate follows.

The proof of Theorem B.5 is complete.

As a consequence, we easily obtain order-reducing operators for the full space situation;

Corollary B.6. *For every $\mu \in \mathbf{R}$, the symbol $r^{\mu}(\xi)$, defined by the formula*

$$r^{\mu}(\xi) = \langle\xi\rangle^{\mu} = (1 + |\xi|^2)^{\mu/2}, \quad \xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbf{R}^n,$$

belongs to the symbol class $S_{1,0}^{\mu}(\mathbf{R}^n \times \mathbf{R}^n)$, and induces the operator-valued symbol

$$\mathrm{Op}_{x_n}(r^{\mu}) \in S_{1,0}^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}), H^{\sigma-(\mu,0)}(\mathbf{R})\right)$$

for $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$.

Remark B.7. The definitions in Sect. B.2.3 extend to projective and inductive limits. Let $\tilde{E}$ and $\tilde{F}$ be Banach spaces with group actions. If $F_1 \hookleftarrow F_2 \hookleftarrow \dots F_k \hookleftarrow \dots$ and $E_1 \hookrightarrow E_2 \hookrightarrow \dots E_k \hookrightarrow \dots$ are sequences of Banach spaces with the same group action and

$$F = \bigcap_{k\in\mathbf{N}} F_k = \mathrm{proj} - \lim_{k\to\infty} F_k,$$

$$E = \bigcup_{k\in\mathbf{N}} E_k = \mathrm{ind} - \lim_{k\to\infty} E_k.$$

Then we let

$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \tilde{E}, F\right) = \mathrm{proj} - \lim_{k\to\infty} S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \tilde{E}, F_k\right),$$

$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, \tilde{F}\right) = \mathrm{proj} - \lim_{k\to\infty} S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E_k, \tilde{F}\right),$$

$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right) = \mathrm{proj} - \lim_{k,\ell\to\infty} S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E_k, F_{\ell}\right).$$

In particular, it makes sense to speak of the following symbol classes:

$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \mathcal{S}'(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right),$$
$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \mathcal{S}'(\mathbf{R}_+), \mathbf{C}\right),$$
$$S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right).$$

We shall write

$$S^{-\infty}(\cdots) = \bigcap_{\mu \in \mathbf{R}} S^{\mu}(\cdots).$$

Theorem B.8. *Given symbols* $a_j(y, \tilde{y}, \eta) \in S^{\mu-j}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$, $j = 0, 1, 2, \ldots$, *there exists a symbol* $a(y, \tilde{y}, \eta) \in S^{\mu}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$ *such that*

$$a(y, \tilde{y}, \eta) \sim \sum_{j=0}^{\infty} a_j(y, \tilde{y}, \eta).$$

As usual, the equivalence relation $\sim$ *is defined by the fact that we have, for every positive integer* J,

$$a(y, \tilde{y}, \eta) - \sum_{j \leq J} a_j(y, \tilde{y}, \eta) \in S^{\mu-J-1}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right).$$

Moreover, the symbol $a(y, \tilde{y}, \eta)$ *is unique modulo* $S^{-\infty}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$.

The proof follows from the standard argument, just as in Hörmander [Ho4, Proposition 18.1.3].

Definition B.9. A symbol $a(y, \tilde{y}, \eta) \in S^{\mu}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$ is said to be *classical* if it has an asymptotic expansion

$$a(y, \tilde{y}, \eta) \sim \sum_{j=0}^{\infty} a_j(y, \tilde{y}, \eta)$$

with

$$a_j(y, \tilde{y}, \eta) \in S^{\mu-j}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right)$$

satisfying the homogeneity relation

$$a_j(y, \tilde{y}, \lambda\eta) = \lambda^{\mu-j}\tilde{\kappa}_{\lambda} a_j(y, \tilde{y}, \eta)\kappa_{\lambda^{-1}} \tag{B.4}$$

for all $\lambda \geq 1$, $|\eta| \geq R$ with a suitable constant R. Then we write

$$a(y, \tilde{y}, \eta) \in S^{\mu}_{\mathrm{cl}}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right).$$

For $E = \mathbf{C}^k$ and $F = \mathbf{C}^{\ell}$, we recover the standard notion.

The symbols γ_j in Example B.2 are homogeneous of degree $j + 1/2$ in the sense of formula (B.4).

The next lemma is straightforward to prove:

Lemma B.10. *We let*

$$a(y, \tilde{y}, \eta) \in S^{\mu}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right),$$
$$b(y, \tilde{y}, \eta) \in S^{\tilde{\mu}}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; F, G\right).$$

Then we have the following two assertions (a) and (b):

(a) $D^{\alpha}_{\eta} D^{\beta}_{y} D^{\gamma}_{\tilde{y}} a \in S^{\mu - |\alpha|}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F\right)$ *for all multi-indices* α, β *and* γ.
(b) The pointwise composition $(ba)(y, \tilde{y}, \eta) = b(y, \tilde{y}, \eta) a(y, \tilde{y}, \eta)$ *yields an element*

$$ba \in S^{\mu + \tilde{\mu}}\left(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, G\right).$$

Theorem B.11. *(i) Let* $a(y, \tilde{y}, \eta) \in S^{\mu}(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E, F)$*. Then there exist a (unique) left symbol* $a_L(y, \eta)$ *and a (unique) right symbol* $a_R(\tilde{y}, \eta)$ *for* $\mathrm{Op}(a)$.
(ii) Let $a(y, \tilde{y}, \eta) \in S^{\mu}(\mathbf{R}^q \times \mathbf{R}^q; E, F)$ *and* $b(y, \tilde{y}, \eta) \in S^{\tilde{\mu}}(\mathbf{R}^q \times \mathbf{R}^q; F, G)$*. Then there is a left symbol* $c_L(y, \eta) \in S^{\mu + \tilde{\mu}}(\mathbf{R}^q \times \mathbf{R}^q; E, G)$ *such that*

$$\mathrm{Op}(b) \circ \mathrm{Op}(a) = \mathrm{Op}\left(c_L\right).$$

Moreover, we have the asymptotic expansion

$$c_L(y, \eta) \sim \sum_{\alpha} \frac{1}{\alpha!} \partial^{\alpha}_{\eta} b_L(y, \eta) D^{\alpha}_{y} a_L(y, \eta).$$

Here $b_L(y, \eta)$ *is a left symbol for* $\mathrm{Op}(b)$.

Proof. The proof of part (i) is analogous to that of Kumano-go [Ku, Chapter 2, Theorem 2.5]. For part (ii), we choose a right symbol $a_R(\tilde{y}, \eta)$ for $\mathrm{Op}(a)$ and a left symbol $b_L(y, \eta)$ for $\mathrm{Op}(b)$, respectively. Then we have

$$\mathrm{Op}(b) \circ \mathrm{Op}(a) = \mathrm{Op}(b) \circ \mathrm{Op}(a_R) = \mathrm{Op}(\tilde{c}),$$

with $\tilde{c}(y, \tilde{y}, \eta) = b_L(y, \eta) a_R(\tilde{y}, \eta)$.

By switching to the left symbol $c_L(y,\eta)$ of $\mathrm{Op}(\tilde{c})$, we obtain

$$\begin{aligned}
[\mathrm{Op}(b)\circ \mathrm{Op}(a)]\,u(y) &= [\mathrm{Op}(\tilde{c})]\,u(y) \\
&= \frac{1}{(2\pi)^q}\iint_{\mathbf{R}^q\times\mathbf{R}^q} e^{i(y-\tilde{y})\cdot\eta}\tilde{c}(y,\tilde{y},\eta)u(\tilde{y})\,d\tilde{y}\,d\eta \\
&= \frac{1}{(2\pi)^q}\int_{\mathbf{R}^q} e^{iy\cdot\eta}c_L(y,\eta)\hat{u}(\eta)\,d\eta \\
&= [\mathrm{Op}\,(c_L)]\,u(y),
\end{aligned}$$

and the asymptotic expansion (see Theorems 7.17 and 7.20)

$$\begin{aligned}
c_L(y,\eta) &\sim \sum_\alpha \frac{1}{\alpha!}\partial_\eta^\alpha D_{\tilde{y}}^\alpha\left(\tilde{c}(y,\tilde{y},\eta)\right)\Big|_{\tilde{y}=y} = \sum_\alpha \frac{1}{\alpha!}\partial_\eta^\alpha \left(b_L(y,\eta)D_{\tilde{y}}^\alpha a_R(\tilde{y},\eta)\right)\Big|_{\tilde{y}=y} \\
&\sim \sum_\alpha \frac{1}{\alpha!}\partial_\eta^\alpha b_L(y,\eta)D_y^\alpha a_L(y,\eta).
\end{aligned}$$

The proof of Theorem B.11 is complete.

B.2.4 Duality

Let (E_-, E_0, E_+) be a triple of Hilbert spaces. We assume that all are embedded in a common vector space V and that $E_0 \cap E_+ \cap E_-$ is dense in $E_\pm$ as well as in E_0. Moreover, we assume that there is a continuous, non-degenerate sesquilinear form $(\cdot,\cdot)_E : E_+ \times E_- \to \mathbf{C}$ which coincides with the inner product $(\cdot,\cdot)_E$ of E_0 on $(E_+ \cap E_0) \times (E_- \cap E_0)$. We ask that, via $(\cdot,\cdot)_E$, we may identify E_+ with the dual of E_- and vice versa, and that the quantities

$$\|e\|'_{E_-} = \sup_{\|f\|_{E_+}} |(f,e)_E|\,, \quad \|f\|'_{E_+} = \sup_{\|e\|_{E_-}} |(f,e)_E|$$

furnish equivalent norms on the spaces E_- and E_+, respectively. Assume that there is a group action κ on V which has strongly continuous restrictions to each of the spaces, unitary on E_0, that is, $(\kappa_\lambda e, f)_E = (e, \kappa_{\lambda^{-1}} f)_E$ for all e and $f \in E_0$. Then we have

$$(\kappa_\lambda e, f)_E = (e, \kappa_{\lambda^{-1}} f)_E \quad \text{for all } (e,f) \in E_+ \times E_-,$$

since the identity holds true on the dense set $(E_+ \cap E_0) \times (E_- \cap E_0)$. In other words, the action κ on E_+ is dual to the action κ on E_- and vice versa.

Typical examples for the above situation are given by the following triples:

$$(E_-, E_0, E_+) = \big(H^{-\sigma}(\mathbf{R}), L^2(\mathbf{R}), H^\sigma(\mathbf{R})\big),$$
$$(E_-, E_0, E_+) = \big(H_0^{-\sigma}(\overline{\mathbf{R}_+}), L^2(\mathbf{R}_+), H^\sigma(\mathbf{R}_+)\big),$$

with $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$.

Let (F_-, F_0, F_+) be an analogous triple of Hilbert spaces with group action $\tilde{\kappa}$, and let $a(\tilde{y}, y, \eta) \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; E_-, F_-)$. We define $a^*(y, \tilde{y}, \eta)$ by the formula

$$a^*(y, \tilde{y}, \eta) = a(\tilde{y}, y, \eta)^* \in \mathcal{L}(F_+, E_+),$$

where the last asterisk $*$ denotes the adjoint operator with respect to the sesquilinear forms $(\cdot,\cdot)_E$ and $(\cdot,\cdot)_F$:

$$\big(a(\tilde{y}, y, \eta)^* f, e\big)_E = (f, a(\tilde{y}, y, \eta)e)_F \quad \text{for all } (e, f) \in E_- \times F_+.$$

It is not difficult to verify that $a^* \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q \times \mathbf{R}^q; F_+, E_+)$.

Moreover, we may introduce a continuous, non-degenerate sesquilinear form

$$(\cdot,\cdot)_{\mathcal{S}_E} : \mathcal{S}(\mathbf{R}^q, E_+) \times \mathcal{S}(\mathbf{R}^q, E_-) \longrightarrow \mathbf{C}$$

by the formula

$$(u, v)_{\mathcal{S}_E} = \int_{\mathbf{R}^q} (u(y), v(y))_E \, dy.$$

Analogously, we may introduce a continuous, non-degenerate sesquilinear form

$$(\cdot,\cdot)_{\mathcal{S}_F} : \mathcal{S}(\mathbf{R}^q, F_+) \times \mathcal{S}(\mathbf{R}^q, F_-) \longrightarrow \mathbf{C}$$

by the formula

$$(u, v)_{\mathcal{S}_F} = \int_{\mathbf{R}^q} (u(y), v(y))_F \, dy.$$

The symbol $a^*(y, \tilde{y}, \eta)$ induces a continuous mapping

$$\mathrm{Op}(a^*) : \mathcal{S}(\mathbf{R}^q, F_+) \longrightarrow \mathcal{S}(\mathbf{R}^q, E_+).$$

This is the unique operator satisfying the condition

$$\big([\mathrm{Op}(a^*)]u, v\big)_{\mathcal{S}_E} = (u, [\mathrm{Op}(a)]v)_{\mathcal{S}_F}$$
$$\text{for all } u \in \mathcal{S}(\mathbf{R}^q, F_+) \text{ and } v \in \mathcal{S}(\mathbf{R}^q, E_+).$$

B.2.5 Wedge Sobolev Spaces

Let E be a Banach space with a group action κ. The wedge Sobolev space $\mathcal{W}^s(\mathbf{R}^q, E)$, $s \in \mathbf{R}$, is the completion of the Schwartz space $\mathcal{S}(\mathbf{R}^q, E)$ with respect to the norm

$$\|u\|^2_{\mathcal{W}^s(\mathbf{R}^q,E)} = \int_{\mathbf{R}^q} \langle\eta\rangle^{2s} \left\| \kappa_{\langle\eta\rangle^{-1}} \hat{u}(\eta) \right\|_E^2 \, d\eta.$$

The wedge Sobolev space $\mathcal{W}^s(\mathbf{R}^q, E)$ is a subset of $\mathcal{S}'(\mathbf{R}^q, E)$ and is a Hilbert space with the natural inner product. We remark that if F is a Banach space such that the embedding $E \hookrightarrow F$ is compact, then the embedding $\mathcal{W}^s(\mathbf{R}^q, E) \hookrightarrow \mathcal{W}^{s'}(\mathbf{R}^q, F)$ is compact whenever $s > s'$.

For any $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ and $s \in \mathbf{R}$, we have the following two assertions (1) and (2):

(1) The dual space of $\mathcal{W}^s(\mathbf{R}^q, H^\sigma(\mathbf{R}_+))$ is $\mathcal{W}^{-s}(\mathbf{R}^q, H_0^{-\sigma}(\overline{\mathbf{R}_+}))$ and vice versa.
(2) The dual of $\mathcal{W}^s(\mathbf{R}^q, H^\sigma(\mathbf{R}))$ is $\mathcal{W}^{-s}(\mathbf{R}^q, H^{-\sigma}(\mathbf{R}))$.

Proofs of these statements are given in Hirschmann [Hi, Corollary 6.5].

For any $s = (s_1, s_2) \in \mathbf{R}^2$, we can define the wedge Sobolev space

$$\mathcal{W}^s(\mathbf{R}^q, E) = \left\{ \langle y\rangle^{-s_2} u : u \in \mathcal{W}^{s_1}(\mathbf{R}^q, E) \right\}.$$

Then we obtain, in particular, the following two assertions (cf. Schrohe [Sr2, Lemma 1.8, Corollary 1.10]):

$$\text{proj} - \lim_{s_1,s_2,\sigma_1,\sigma_2\to\infty} \mathcal{W}^{(s_1,s_2)}\left(\mathbf{R}^q, H^{(\sigma_1,\sigma_2)}(\mathbf{R}_+)\right) = \mathcal{S}(\mathbf{R}_+^{q+1}),$$

$$\text{ind} - \lim_{s_1,s_2,\sigma_1,\sigma_2\to-\infty} \mathcal{W}^{(s_1,s_2)}\left(\mathbf{R}^q, H_0^{(\sigma_1,\sigma_2)}(\overline{\mathbf{R}_+})\right) = \mathcal{S}'(\mathbf{R}_+^{q+1}).$$

Lemma B.12. *For any $s \in \mathbf{R}$, we have the following four assertions (a)–(d):*

(a) $\mathcal{W}^s(\mathbf{R}^q, \mathbf{C}) = H^s(\mathbf{R}^q)$.
(b) $\mathcal{W}^s(\mathbf{R}^q, H^s(\mathbf{R}^k)) = H^s(\mathbf{R}^{q+k})$.
(c) $\mathcal{W}^s(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+})) = H_0^s(\overline{\mathbf{R}^{q+1}})$.
(d) $\mathcal{W}^s(\mathbf{R}^q, H^s(\mathbf{R}_+)) = H^s(\mathbf{R}_+^{q+1})$.

Proof. Assertion (a) is obvious.

For Assertion (b), we remark that the behavior of the Fourier transform under dilations yields that

$$\|u\|^2_{\mathcal{W}^s(\mathbf{R}^q,H^s(\mathbf{R}^k))}$$
$$= \int_{\mathbf{R}^q} \langle\eta\rangle^{2s} \, \| \langle\eta\rangle^{k/2} \langle\tau\rangle^s \, (F_{t\to\tau} F_{y\to\eta} u)(\eta, \langle\eta\rangle \, \tau) \|^2_{L^2(\mathbf{R}^k_\tau)} \, d\eta$$

$$= \int_{\mathbf{R}^q} \langle\eta\rangle^{2s+k} \int_{\mathbf{R}^k} \langle\tau\rangle^{2s} \left|(F_{t\to\tau} F_{y\to\eta} u)(\eta, \langle\eta\rangle\, \tau)\right|^2 d\tau\, d\eta$$
$$= \int_{\mathbf{R}^q} \langle\eta\rangle^{2s} \int_{\mathbf{R}^k} \left\langle \langle\eta\rangle^{-1}\, \tau' \right\rangle^{2s} \left|(F_{t\to\tau'} F_{y\to\eta} u)(\eta, \tau')\right|^2 d\tau'\, d\eta.$$

The desired statement follows, since we have

$$\langle\eta\rangle^{2s} \left\langle \langle\eta\rangle^{-1}\, \tau' \right\rangle^{2s} = \left\langle \eta, \tau' \right\rangle^{2s} = \left(1 + |\eta|^2 + |\tau'|^2\right)^s.$$

Assertion (c) follows from Assertion (b): The space $H_0^s(\overline{\mathbf{R}_+^{q+1}})$ is the closure of $C_0^\infty(\mathbf{R}_+^{q+1})$ in the topology of $H^s(\mathbf{R}^{q+1})$. Since the space $C_0^\infty(\mathbf{R}_+^{q+1})$ is dense in $\mathcal{S}(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+}))$ and since the norms of the spaces

$$\mathcal{W}^s\left(\mathbf{R}^q, H_0^s\left(\overline{\mathbf{R}_+}\right)\right), \quad \mathcal{W}^s(\mathbf{R}^q, H^s(\mathbf{R})), \quad H^s(\mathbf{R}^{q+1})$$

coincide on the space $C_0^\infty(\mathbf{R}_+^{q+1})$, we obtain the desired assertion:

$$H_0^s\left(\overline{\mathbf{R}_+^{q+1}}\right) = \left[C_0^\infty(\mathbf{R}_+^{q+1})\right]_{H^s(\mathbf{R}^{q+1})} = \left[C_0^\infty(\mathbf{R}_+^{q+1})\right]_{\mathcal{W}^s(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+}))}$$
$$= \left[\mathcal{S}(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+}))\right]_{\mathcal{W}^s(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+}))}$$
$$= \mathcal{W}^s(\mathbf{R}^q, H_0^s(\overline{\mathbf{R}_+})).$$

Assertion (d) follows from Assertion (c) by duality:

$$\mathcal{W}^s(\mathbf{R}^q, H^s(\mathbf{R}_+)) = \left(\mathcal{W}^{-s}\left(\mathbf{R}^q, H_0^{-s}\left(\overline{\mathbf{R}_+}\right)\right)\right)' = \left(H_0^{-s}\left(\overline{\mathbf{R}_+^{q+1}}\right)\right)'$$
$$= H^s(\mathbf{R}_+^{q+1}).$$

The proof of Lemma B.12 is complete.

The following theorem is proved by Seiler [Si, Theorem 3.14]:

Theorem B.13. *Let E, F be Hilbert spaces and $a(y, \eta) \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; E, F)$. Then the operator*

$$\operatorname{Op}(a) : \mathcal{W}^s(\mathbf{R}^q, E) \longrightarrow \mathcal{W}^{s-\mu}(\mathbf{R}^q, F)$$

is bounded for every $s \in \mathbf{R}$.

Corollary B.14. *Under the assumptions of Theorem B.13, the operator*

$$\mathrm{Op}(a) : \mathcal{W}^s(\mathbf{R}^q, E) \longrightarrow \mathcal{W}^{s-(\mu,0)}(\mathbf{R}^q, F)$$

is even bounded for every $s = (s_1, s_2) \in \mathbf{R}^2$.

Proof. By an interpolation argument, it suffices to treat the case: $s_2 \in 2\mathbf{Z}$.
The boundedness of the operator

$$\mathrm{Op}(a) : \mathcal{W}^s(\mathbf{R}^q, E) \longrightarrow \mathcal{W}^{s-(\mu,0)}(\mathbf{R}^q, F)$$

is equivalent to the boundedness of the operator

$$\langle y \rangle^{s_2} \, \mathrm{Op}(a) \, \langle y \rangle^{-s_2} : \mathcal{W}^{s_1}(\mathbf{R}^q, E) \longrightarrow \mathcal{W}^{s_1-\mu}(\mathbf{R}^q, F).$$

Either $\langle y \rangle^{s_2}$ or $\langle y \rangle^{-s_2}$ is a polynomial. Without loss of generality, we may take $s_2 \geq 0$. By using the identity

$$\left[\mathrm{Op}(a), y_j\right] = \mathrm{Op}(D_{y_j} a),$$

we may shift the powers of y from the left to the right-hand side. According to Theorem B.13, both $\mathrm{Op}(D_y^\alpha a)$ and $y^\beta \langle y \rangle^{-s_2}$, $|\beta| \leq s_2$, are bounded operators on the respective spaces, and the desired assertion follows.
The proof of Corollary B.14 is complete.

B.3 The Transmission Property

It has been pointed out in the Introduction that we will have to impose a condition on the pseudo-differential entry P in formula (B.1) in order to ensure the stated mapping property. Following Boutet de Monvel [Bo], we ask that the symbol of P has the transmission property.

Definition B.15. Given a function f on $\mathbf{R}^n_+$, we denote by $e^+ f$ its extension by zero to a function on $\mathbf{R}^n$:

$$e^+ f(x) = \begin{cases} f(x) & \text{if } x \in \mathbf{R}^n_+, \\ 0 & \text{if } x \in \mathbf{R}^n \setminus \mathbf{R}^n_+. \end{cases}$$

Extension by zero also makes sense for distributions u in $H^{(s_1,s_2)}(\mathbf{R}^n)$ if $s_1 > -1/2$. Hence it follows that $e^+ u$ is an element of $H^{(s_1,s_2)}(\mathbf{R}^n)$ for $-1/2 < s_1 < 1/2$. If u in $H^{(s_1,s_2)}(\mathbf{R}^n)$ for $s_1 \geq 1/2$, then we have

$$e^+ u \in H^\sigma(\mathbf{R}^n)$$

for all $\sigma = (\sigma_1, \sigma_2)$ with $\sigma_1 < 1/2$ and $\sigma_2 = s_2$.

For classical pseudo-differential symbols of integer order, the transmission property can be expressed via homogeneity properties of the terms in the asymptotic expansion (see [Ho4, Theorem 18.2.18]). In fact, this property also ensures condition (B.6) below. For non-classical symbols, this is no longer true. For details, see the analysis by Grubb–Hörmander [GH]. In Schrohe [Sr3], it is shown that the transmission property can be characterized via commutator estimates on wedge Sobolev spaces.

We let

$$\begin{aligned} H^+ &= \left\{(e^+u)^\wedge : u \in \mathcal{S}(\mathbf{R}_+)\right\} = \mathcal{F}\left(\mathcal{S}(\mathbf{R}_+)\right), \\ H_0^- &= \left\{(e^-v)^\wedge : v \in \mathcal{S}(\mathbf{R}_-)\right\} = \mathcal{F}\left(\mathcal{S}(\mathbf{R}_-)\right). \end{aligned}$$

Here H^+ and H_0^- are spaces of smooth functions on $\mathbf{R}$ decaying to the first order near infinity, and

$$e^-g(x) = \begin{cases} g(x) & \text{if } x \in \mathbf{R}^n_-, \\ 0 & \text{if } x \in \mathbf{R}^n \setminus \mathbf{R}^n_-, \end{cases}$$

where $\mathbf{R}^n_-$ is the lower half-space

$$\mathbf{R}^n_- = \{(x_1, x_2, \ldots, x_{n-1}, x_n) \in \mathbf{R}^n : x_n < 0\}.$$

We denote by H' the space of all polynomials. Then we let

$$\begin{aligned} H &= H^+ \oplus H_0^- \oplus H', \\ H^- &= H_0^- \oplus H', \\ H_0 &= H^+ \oplus H_0^-. \end{aligned}$$

The next proposition provide fundamental characterizations of the spaces H^+, H_0^- and H_0 in terms of complex analysis:

Proposition B.16. *(a) The spaces H^+, H_0^-, H^-, H_0 and H are algebras.*

(b) We can give Paley–Wiener type characterizations of the spaces H^+, H_0^- and H_0 as follows.

(b.1) A function $h \in C^\infty(\mathbf{R})$ belongs to H^+ if and only if it has an analytic extension to the lower open half-plane $\{\operatorname{Im}\zeta < 0\}$, continuous in the lower closed half-plane $\{\operatorname{Im}\zeta \leq 0\}$, together with an asymptotic expansion

$$h(\zeta) \sim \sum_{k=1}^{\infty} a_k\, \zeta^{-k}, \tag{B.5}$$

for $|\zeta| \to \infty$ *in the lower closed half-plane* $\{\mathrm{Im}\,\zeta \le 0\}$*, which can be differentiated formally.*

A function $h \in C^\infty(\mathbf{R})$ *belongs to* H^+ *if and only if it has a unique expansion*

$$h(t) = \sum_{k=0}^{\infty} \alpha_k \frac{(1-it)^k}{(1+it)^{k+1}},$$

where the coefficients α_k *form a rapidly decreasing sequence.*

(b.2) A function $h \in C^\infty(\mathbf{R})$ *belongs to* H_0^- *if and only if it has an analytic extension to the upper open half-plane* $\{\mathrm{Im}\,\zeta > 0\}$*, continuous in the upper closed half-plane* $\{\mathrm{Im}\,\zeta \ge 0\}$*, together with an asymptotic expansion*

$$h(\zeta) \sim \sum_{k=1}^{\infty} a_k\, \zeta^{-k},$$

for $|\zeta| \to \infty$ *in the upper closed half-plane* $\{\mathrm{Im}\,\zeta \ge 0\}$*, which can be differentiated formally.*

A function $h \in C^\infty(\mathbf{R})$ *belongs to* H_0^- *if and only if it has a unique expansion*

$$h(t) = \sum_{k=-1}^{-\infty} \alpha_k \frac{(1+it)^{|k|-1}}{(1-it)^{|k|}},$$

where the coefficients α_k *form a rapidly decreasing sequence.*

(b.3) A function $h \in C^\infty(\mathbf{R})$ *belongs to* $H_0 = H^+ \oplus H_0^-$ *if and only if it has an asymptotic expansion*

$$h(\zeta) \sim \sum_{k=1}^{\infty} a_k\, \zeta^{-k},$$

for $|\zeta| \to \infty$ *in* $\mathbf{R}$*, which can be differentiated formally.*

A function $h \in C^\infty(\mathbf{R})$ *belongs to* H_0 *if and only if it has a unique expansion*

$$h(t) = \sum_{k=-\infty}^{\infty} \alpha_k \frac{(1-it)^k}{(1+it)^{k+1}},$$

where the coefficients α_k *form a rapidly decreasing sequence.*

Part (a) of Proposition B.16 is easily verified. Proofs of parts (b.1), (b.2) and (b.3) can be found in Rempel–Schulze [RS, 2.1.1.1].

Definition B.17. A pseudo-differential symbol

$$p = p(x, y, \xi) \in S^{\mu}_{1,0}(\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n)$$

has the *transmission property* at $\{x_n = y_n = 0\}$, provided that we have, for all k and ℓ,

$$\partial^{\ell}_{y_n} \partial^{k}_{x_n} p(x', 0, y', 0, \xi', \langle \xi' \rangle \xi_n) \in S^{\mu}_{1,0}(\mathbf{R}^{n-1}_{x'} \times \mathbf{R}^{n-1}_{y'} \times \mathbf{R}^{n-1}_{\xi'}) \hat{\otimes}_{\pi} H_{\xi_n}.$$

The subscripts x', y', ξ', and ξ_n are used only for the moment in order to indicate the variable for which we have the corresponding property.

We write

$$p \in S^{\mu}_{\mathrm{tr}} (\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n) .$$

We assume that $p = p(x, \xi)$ is a classical symbol of order $\mu \in \mathbf{Z}$ with an asymptotic expansion

$$p(x, \xi) \sim \sum_{j=0}^{\infty} p_{\mu-j}(x, \xi),$$

where the symbol

$$p_{\mu-j}(x, \xi) \in S^{\mu-j} (\mathbf{R}^n \times \mathbf{R}^n)$$

is positively homogeneous of degree $\mu - j$ in ξ for $|\xi| \geq 1$, that is, we have, for all $t \geq 1$ and $|\xi| \geq 1$,

$$p_{\mu-j}(x, t\xi) = t^{\mu-j} p_{\mu-j}(x, \xi).$$

We sketch the argument why the classical symbol $p(x, \xi)$ of integer order has the transmission property if and only if we have, for all multi-indices $\alpha = (\alpha', \alpha_n)$,

$$\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} \left(\frac{\partial}{\partial \xi'}\right)^{\alpha'} p_{\mu-j}(x', 0, 0, +1) \tag{B.6}$$
$$= (-1)^{\mu-j-|\alpha'|} \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} \left(\frac{\partial}{\partial \xi'}\right)^{\alpha'} p_{\mu-j}(x', 0, 0, -1).$$

Indeed, a Taylor expansion gives that

$$\partial^{\alpha_n}_{x_n} p_{\mu-j}(x', 0, \xi', \langle \xi' \rangle \nu)$$
$$= \langle \xi' \rangle^{\mu-j} \partial^{k}_{x_n} p_{\mu-j}(x', 0, \xi'/\langle \xi' \rangle, \nu)$$

$$= \langle \xi' \rangle^{\mu-j} \left[\sum_{|\alpha'| \le N} \frac{1}{\alpha'!} \partial_{x_n}^{\alpha_n} \partial_{\xi'}^{\alpha'} p_{\mu-j}(x', 0, 0, \nu)(\xi'/\langle \xi' \rangle)^{\alpha'} + r_N(x', \xi', \nu) \right].$$

Here $r_N(x', \xi', \nu)$ can be estimated in terms of the quantity

$$\sup \left\{ |\partial_{\xi'}^{\beta'} p_{\mu-j}(x', 0, \sigma\xi'/\langle \xi' \rangle, \nu)| : |\beta'| = N + 1,\ 0 \le \sigma \le 1 \right\},$$

which is of order $O(\langle \nu \rangle^{\mu-j-N-1})$, uniformly in x' and ξ'. Moreover, we have

$$\partial_{x_n}^{\alpha_n} \partial_{\xi'}^{\alpha'} p_{\mu-j}(x', 0, 0, \nu) = \partial_{x_n}^{\alpha_n} \partial_{\xi'}^{\alpha'} p_{\mu-j}(x', 0, 0, \pm 1)|\nu|^{\mu-j-|\alpha'|},$$

so that, eventually, Assertion (b.3) of Proposition B.16 gives the desired assertion.

The next lemma is obvious:

Lemma B.18. *Regularizing symbols always have the transmission property, and so do symbols which vanish to infinite order at* $\{x_n = 0\}$. *Moreover, all symbols which are polynomial in* ξ *have the transmission property according to condition* (B.6).

Example B.19. The symbol $\langle \xi \rangle$ does not have the transmission property. Indeed, since we have the identity

$$\begin{aligned} &\langle (\xi', \xi_n \langle \xi' \rangle) \rangle \\ &= \left(1 + |\xi'|^2 + \xi_n^2 \langle \xi' \rangle^2\right)^{1/2} = \left(\langle \xi' \rangle^2 + \xi_n^2 \langle \xi' \rangle^2\right)^{1/2} = \langle \xi' \rangle \sqrt{1 + \xi_n^2} \\ &= \langle \xi' \rangle \langle \xi_n \rangle, \end{aligned}$$

this would require that $\nu \mapsto \langle \nu \rangle \in H$. However, by writing

$$\langle \nu \rangle = |\nu| \left(1 + \frac{1}{\nu^2}\right) = |\nu| \sum_{j=0}^{\infty} \binom{1/2}{j} \nu^{-2j} = \sum_{j=0}^{\infty} \binom{1/2}{j} (\operatorname{sgn} \nu)\, \nu^{-2j+1},$$

we see that condition (B.6) is violated.

There are symbols with the transmission property of arbitrary order:

Example B.20. Let $q(x, \xi') \in S^{\mu}_{1,0}(\mathbf{R}^n \times \mathbf{R}^{n-1})$ and $\varphi \in \mathcal{S}(\mathbf{R})$. Then we have

$$p(x, \xi) := q(x, \xi')\varphi(\xi_n/\langle \xi' \rangle) \in S^{\mu}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n).$$

Indeed, since we have

$$p(x', 0, \xi', \langle \xi' \rangle \xi_n) = q(x', 0, \xi')\varphi(\xi_n) \in S^{\mu}_{1,0}(\mathbf{R}^{n-1}_{x'} \times \mathbf{R}^{n-1}_{\xi'}) \otimes H_{\xi_n},$$

it is straightforward to verify the symbol estimates for $p(x, \xi)$.

The next proposition asserts the stability of the transmission property under the usual pseudo-differential constructions:

Proposition B.21. *(a) If $p(x, y, \xi)$ and $q(x, y, \xi)$ have the transmission property, then all derivatives $D_\xi^\alpha D_x^\beta D_y^\gamma p(x, y, \xi)$, the product $p(x, y, \xi)q(x, y, \xi)$, and the left symbols $p_L(x, \xi)$ and right symbols $p_R(y, \xi)$, have the transmission property.*

(b) If $p_j(x, y, \xi)$ are symbols of order $\mu - j$ with the transmission property and if $p(x, y, \xi) \sim \sum_j p_j(x, y, \xi)$, then $p(x, y, \xi)$ has the transmission property.

(c) If $p(x, y, \xi)$ is elliptic with the transmission property, then every parametrix of $p(x, y, \xi)$ has the transmission property.

The proof of Proposition B.21 is straightforward.

The next lemma, together with Theorem B.24 below, shows that there are also order-reducing symbols for the half-space situation. The construction goes back to Grubb [Gb].

Lemma B.22. *Choose $\chi \in \mathcal{S}(\mathbf{R})$ with $\operatorname{supp}(\mathcal{F}^{-1}\chi) \subset \mathbf{R}_-$ and $\chi(0) = 1$. For every $\mu \in \mathbf{Z}$ and a positive constant a with $a >> \sup \|\chi'\|$ (the supremum norm of the first derivative of χ), we let*

$$r_-^\mu(\xi) = \left(\chi\left(\frac{\xi_n}{a\,\langle\xi'\rangle}\right)\langle\xi'\rangle - i\,\xi_n\right)^\mu, \quad \xi = (\xi', \xi_n) \in \mathbf{R}^n. \tag{B.7}$$

Then the symbol $r_-^\mu(\xi)$ is an elliptic element in the symbol class $S_{\mathrm{tr}}^\mu(\mathbf{R}^n \times \mathbf{R}^n)$. The same assertion holds true for the symbol

$$r_+^\mu(\xi) = \overline{r_-^\mu(\xi)} = \left(\chi\left(\frac{\xi_n}{a\,\langle\xi'\rangle}\right)\langle\xi'\rangle + i\,\xi_n\right)^\mu, \quad \xi = (\xi', \xi_n) \in \mathbf{R}^n.$$

Proof. The above definition makes sense, since we have

$$\begin{aligned}\frac{\chi\left(\frac{\xi_n}{a\langle\xi'\rangle}\right)\langle\xi'\rangle - i\,\xi_n}{\langle\xi'\rangle - i\,\xi_n} &= 1 + \langle\xi'\rangle\,\frac{\chi\left(\frac{\xi_n}{a\langle\xi'\rangle}\right) - \chi(0)}{\langle\xi'\rangle - i\,\xi_n} \\ &:= 1 + r(\xi),\end{aligned}$$

where the symbol $r(\xi)$ is small:

$$|r(\xi)| = \left|\langle\xi'\rangle\,\frac{\chi\left(\frac{\xi_n}{a\langle\xi'\rangle}\right) - \chi(0)}{\langle\xi'\rangle - i\,\xi_n}\right| \le \frac{\sup\|\chi'\|}{a}\cdot\frac{|\xi_n|}{\langle\xi\rangle}.$$

An application of Example B.20 shows that $r_-^\mu(\xi)$ is an elliptic symbol of order μ. Moreover, we have

$$r_-^\mu(\xi', \langle\xi'\rangle\,\xi_n) = \langle\xi'\rangle^\mu\,(\chi(\xi_n/a) - i\,\xi_n)^\mu\,.$$

By using Proposition B.16, it is not difficult to verify that $\chi(\xi_n/a) - i\xi_n$ as well as its integer powers belong to the space $H^- = H_0^- \oplus H'$. Indeed, it suffices to note that $\operatorname{supp} \mathcal{F}^{-1}\chi \subset \mathbf{R}_-$. Hence the symbol $r_-^{\mu}(\xi)$ has the transmission property.

The proof of Lemma B.22 is complete.

Definition B.23. Given $s = (s_1, s_2) \in \mathbf{R}^2$ with $s_1 > -1/2$ and a symbol p, we define the operator

$$\mathrm{Op}^+ p : H^s(\mathbf{R}_+^n) \longrightarrow \mathcal{D}'(\mathbf{R}_+^n) \tag{B.8}$$

by the formula

$$\left[\mathrm{Op}^+ p\right] u = r^+ \left(\mathrm{Op}(p)(e^+ u)\right).$$

For $P = \mathrm{Op}(p)$, we write $P_+ = \mathrm{Op}^+ p$.

It is obvious that we may replace $\mathcal{D}'(\mathbf{R}_+^n)$ on the right-hand side of (B.8) by $H^\sigma(\mathbf{R}_+^n)$, where $\sigma_1 = \min\{s_1, 0\} - \mu$, $\sigma_2 = s_2$ and μ is the order of p.

Theorem B.24. *Let $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$. For the symbols $r_\pm^\mu(\xi)$, we have the following six assertions (a)–(f):*

(a) $\mathrm{Op}_{x_n}(r_\pm^\mu) \in S^\mu(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}), H^{\sigma-(\mu,0)}(\mathbf{R}))$.

(b) $\mathrm{Op}_{x_n}^+ r_-^\mu \in S^\mu(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+))$ *with* $\sigma_1 > -1/2$.

(c) Let $e_\sigma : H^\sigma(\mathbf{R}_+) \to H^\sigma(\mathbf{R})$ be an arbitrary extension operator. Then it follows that

$$r^+ \left[\mathrm{Op}_{x_n}\left(r_-^\mu\right)\right] e_\sigma \in S^\mu\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+)\right).$$

The operator is independent of the particular choice of e_σ.

(d) $\mathrm{Op}_{x_n}(r_+^\mu) \in S^\mu(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^{\sigma-(\mu,0)}(\overline{\mathbf{R}_+}))$. *Note that we use neither restriction nor extension, for $H_0^\sigma(\overline{\mathbf{R}_+}) \hookrightarrow H^\sigma(\mathbf{R})$.*

(e) For every $\mu' \in \mathbf{Z}$, we have

$$\left[\mathrm{Op}_{x_n}\left(r_+^\mu\right)\right]\left[\mathrm{Op}_{x_n}\left(r_+^{\mu'}\right)\right] = \mathrm{Op}_{x_n}\left(r_+^{\mu+\mu'}\right)$$

on each space $H_0^\sigma(\overline{\mathbf{R}_+})$. In particular, as operators

$$H_0^\sigma(\overline{\mathbf{R}_+}) \longrightarrow H_0^{\sigma-(\mu,0)}(\overline{\mathbf{R}_+}),$$

we have

$$\left[\mathrm{Op}_{x_n}\left(r_+^\mu\right)\right]^{-1} = \mathrm{Op}_{x_n}\left(r_+^{-\mu}\right).$$

(f) For every $\mu' \in \mathbf{Z}$*, we have*

$$\left[\mathrm{Op}^+_{x_n} r^\mu_-\right]\left[\mathrm{Op}^+_{x_n} r^{\mu'}_-\right] = \mathrm{Op}^+_{x_n} r^{\mu+\mu'}_-$$

on each space $H^\sigma(\mathbf{R}_+)$*. Here we tacitly assume that the zero-extension* e^+ *is replaced by an arbitrary extension operator if* $\sigma_1 \leq -1/2$ *or* $\sigma_1 - \mu' \leq -1/2$*. In particular, we have*

$$\left[\mathrm{Op}^+_{x_n} r^\mu_-\right]^{-1} = \mathrm{Op}^+_{x_n} r^{-\mu}_-.$$

Proof. Assertion (a) follows from Theorem B.5. For Assertions (b) and (c), we make the following observation: If $e : H^\sigma(\mathbf{R}_+) \to H^\sigma(\mathbf{R})$ is an arbitrary extension operator, then it follows that the operator

$$r^+\left[\mathrm{Op}_{x_n}(r^\mu_-)(eu)\right]$$

is independent of the choice of e. Indeed, it suffices to note that $\mathrm{supp}\, \mathcal{F}^{-1}\chi \subset \mathbf{R}_-$ in Lemma B.22. Moreover, we can find a continuous operator

$$e^s : H^s(\mathbf{R}_+) \longrightarrow H^s(\mathbf{R})$$

satisfying the condition

$$\tilde{\kappa}_{\lambda^{-1}} e^s \kappa_\lambda = e^s$$

for the standard group actions κ and $\tilde{\kappa}$ on $H^s(\mathbf{R}_+)$ and $H^s(\mathbf{R})$, respectively. The situation can be visualized in the following diagram:

$$\begin{array}{ccc} H^s(\mathbf{R}_+) & \xrightarrow{e^s} & H^s(\mathbf{R}) \\ \uparrow \kappa_\lambda & & \tilde{\kappa}_{\lambda^{-1}} \downarrow \\ H^s(\mathbf{R}_+) & \xrightarrow[e^s]{} & H^s(\mathbf{R}) \end{array}$$

Therefore, Assertion (a) gives the desired assertions, since we have

$$\mathrm{Op}^+_{x_n} r^\mu_-(u) = r^+\left[\mathrm{Op}_{x_n}(r^\mu_-)(e^+u)\right] = r^+\left[\mathrm{Op}_{x_n}(r^\mu_-)(e^s u)\right], \quad u \in H^\sigma(\mathbf{R}_+).$$

For every $u \in H^\sigma_0(\overline{\mathbf{R}_+})$, it follows that $[\mathrm{Op}_{x_n}(r^\mu_+)](u)$ has support in the space $\overline{\mathbf{R}_+}$. Hence Assertion (d) follows from Assertion (a) as follows:

$$\mathrm{Op}^+_{x_n} r^\mu_+ : H^\sigma_0(\overline{\mathbf{R}_+}) \longrightarrow H^\sigma(\mathbf{R}) \overset{\mathrm{Op}_{x_n}(r^\mu_+)}{\longrightarrow} H^{\sigma-(\mu,0)}(\mathbf{R}) \longrightarrow H^{\sigma-(\mu,0)}_0(\overline{\mathbf{R}_+}).$$

Assertion (e) is trivial. Assertion (f) is first shown for large σ_1; then it extends to the general case.

The proof of Theorem B.24 is complete.

The next lemma illustrates the effect of the transmission property:

Lemma B.25. *Let $p \in S^{\mu}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$. For every $\ell \in \mathbf{N}$, the formula*

$$k_\ell = k_\ell(x', \xi') = r^+ \left[\mathrm{Op}_{x_n}(p)\right] (\delta_0^{(\ell)}) \tag{B.9}$$

yields an element

$$k_\ell \in S^{\mu+l+1/2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right).$$

Here $\delta_0^{(\ell)}(x_n)$ is the ℓ-th derivative of the Dirac measure $\delta_0(x_n)$ at the origin 0, and we consider k_ℓ as the operator which associates to a complex number c the function $c\,[\mathrm{Op}_{x_n}(p)]\delta_0^{(\ell)}$ on $\mathbf{R}_+$.

Note that $k_\ell = 0$ if p is a polynomial in ξ.

Proof. Take a right symbol $p_R = p_R(x', y_n, \xi)$ for $\mathrm{Op}_{x_n}(p)$. Fix a cut-off function $\omega \in C_0^\infty(\mathbf{R})$ such that $\omega(t) = 1$ near $t = 0$, and write

$$\begin{aligned}
&p_R(x', y_n, \xi) \\
&= \omega(y_n) p_R(x', y_n, \xi) + (1 - \omega(y_n))\, p_R(x', y_n, \xi) \\
&= \sum_{j=0}^{\ell} \frac{y_n^j}{j!} \omega(y_n) \partial_{y_n}^j p_R(x', 0, \xi) + y_n^{\ell+1} p_{R\ell}(x', y_n, \xi) + (1 - \omega(y_n))\, p_R(x', y_n, \xi)
\end{aligned}$$

with a suitable term $p_{R\ell}$. The operators associated with the second and the third summand vanish, so we can focus on the first one. By part (a) of Proposition B.21, it follows that the right symbol $p_R(x', y_n, \xi)$ has the transmission property. Hence we have

$$\partial_{y_n}^j p_R(x', 0, \xi', \langle \xi' \rangle \xi_n) \in S^{\mu}_{1,0}\left(\mathbf{R}^{n-1}_{x'} \times \mathbf{R}^{n-1}_{\xi'}\right) \hat{\otimes}_\pi H_{\xi_n},$$

and we can write

$$\begin{aligned}
&\partial_{y_n}^j p_R(x', 0, \xi', \xi_n) \\
&= \sum_{k=0}^{\mu} s_{jk}(x', \xi') \xi_n^k + \sum_{k=0}^{\infty} \lambda_{jk} b_{jk}(x', \xi') h_{jk}(\xi_n / \langle \xi' \rangle)
\end{aligned}$$

with (see Schaefer [Sa, Chapter III, Section 6.4, Theorem], Treves [Tv, Theorem 45.2])

$$s_{jk} \in S^{\mu-k}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}),$$

and

$$\sum_{k=1}^{\infty} |\lambda_{jk}| < \infty,$$

$$h_{jk} \in H_0 = H^+ \oplus H_0^- = \mathcal{F}(\mathcal{S}(\mathbf{R}_+)) \oplus \mathcal{F}(\mathcal{S}(\mathbf{R}_-)).$$

Here the sequences $\{b_{jk}\}_{k=1}^{\infty}$ converge to zero in the symbol class $S^{\mu}_{1,0}(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1})$ and the sequences $\{h_{jk}\}_{k=1}^{\infty}$ converge to zero in the space H_0, respectively. The polynomial part gives no contribution to formula (B.9). Hence it suffices to consider a single term

$$b(x', \xi')h(\xi_n/\langle\xi'\rangle)$$

under the summation to show that its contribution to formula (B.9) is an element of the symbol class

$$S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right),$$

and to verify that the seminorms for this element depend continuously on those for $b(x', \xi')$ and $h(\nu)$. Since $b(x', \xi')$ is of order μ and since we have

$$y_n^j \delta_0^{(\ell)} = (-1)^{\ell} \binom{\ell}{j} j!\, \delta_0^{(\ell-j)},$$

we have only to show that, for all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$, the operator

$$r^+\kappa_{\langle\xi'\rangle^{-1}} \left[\mathrm{Op}_{x_n} D^{\alpha}_{\xi'} \left[h(\xi_n/\langle\xi'\rangle)\right]\right] \delta_0^{(\ell)} : \mathbf{C} \longrightarrow H^{\sigma}(\mathbf{R}_+)$$

has a norm of order $O(\langle\xi'\rangle^{-|\alpha|+\ell+1/2})$.

Now the operator

$$D^{\alpha}_{\xi'} \left[h(\xi_n/\langle\xi'\rangle)\right]$$

is a linear combination of terms of the form

$$(\xi_n/\langle\xi'\rangle)^k h^{(k')}(\xi_n/\langle\xi'\rangle)s(\xi'),$$

where

$$s \in S^{-|\alpha|}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}), \quad 0 \le k \le k' \le |\alpha|.$$

The function $\nu^{k'} h^{(k')}(\nu)$ is an element of $H_0 = H^+ \oplus H_0^-$, so we may focus on the case $\alpha = 0$. We observe that

$$\begin{aligned}\kappa_{\langle\xi'\rangle^{-1}} \mathrm{Op}_{x_n} h(\xi_n/\langle\xi'\rangle)\delta_0^{(\ell)} &= i^\ell \langle\xi'\rangle^{-1/2} \mathcal{F}^{-1}_{\xi_n \to x_n} \left[h(\xi_n/\langle\xi'\rangle)\xi_n^\ell\right](x_n/\langle\xi'\rangle) \\ &= i^\ell \langle\xi'\rangle^{1/2+\ell} \mathcal{F}^{-1}\left[h(\nu)\nu^\ell\right](x_n).\end{aligned}$$

This gives the desired result, since $r^+\mathcal{F}^{-1}(h(\nu)\nu^\ell)$ is a function in $\mathcal{S}(\mathbf{R}_+)$.

The proof of Lemma B.25 is complete.

Example B.26. The polynomial $1 + |\xi|^2$ is the symbol of the differential operator $1 - \Delta$. Hence, by Lemma B.18 and part (c) of Proposition B.21 it follows that its inverse $a(\xi) = (1 + |\xi|^2)^{-1}$ has the transmission property. For $x_n > 0$, we have

$$\begin{aligned}\left[\mathrm{Op}_{x_n}(a)\right]\delta_0(x_n) &= \frac{1}{2\pi}\int_{\mathbf{R}} e^{ix_n\xi_n} \frac{1}{1 + |\xi'|^2 + \xi_n^2}\, d\xi_n \\ &= \frac{1}{2(2\pi)\langle\xi'\rangle}\int_{\mathbf{R}} e^{ix_n\xi_n}\left(\frac{1}{\langle\xi'\rangle + i\xi_n} + \frac{1}{\langle\xi'\rangle - i\xi_n}\right) d\xi_n.\end{aligned}$$

However, complex integration around a large half circle in the upper half-plane shows that the integral is equal to the following:

$$2\pi e^{ix_n\xi_n}|_{\xi_n = i\langle\xi'\rangle} = 2\pi e^{-\langle\xi'\rangle x_n}.$$

Hence, in this case, $k_0(x', \xi') = [\mathrm{Op}_{x_n}(a)]$ is the operator which assigns to $c \in \mathbf{C}$ the function

$$c\varphi(x_n) \in \mathcal{S}(\mathbf{R}_+),$$

where

$$\varphi(x_n) = \frac{1}{2\langle\xi'\rangle} e^{-x_n\langle\xi'\rangle}.$$

Then we have

$$\left(\kappa_{\langle\xi'\rangle^{-1}}\varphi\right)(x_n) = \frac{1}{2}\langle\xi'\rangle^{-3/2} e^{-x_n}.$$

All its semi-norms of the space $\mathcal{S}(\mathbf{R}_+)$ are of order $O(\langle\xi'\rangle^{-3/2})$, which shows that $k_0(x', \xi')$ is an element of the symbol class

$$S^{-3/2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}, \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right).$$

The next theorem illustrates how the transmission property of the symbol ensures that the associated operator preserves smoothness up to the boundary:

Theorem B.27. *Let* $a \in S^{\mu}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$ *for* $\mu \in \mathbf{R}$ *and* $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ *with* $\sigma_1 > -1/2$. *Then we have*

$$\mathrm{Op}^{+}_{x_n} a \in S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+)\right).$$

Proof. For $-1/2 < \sigma_1 < 1/2$, the assertion is immediate from Theorem B.5. Indeed, it suffices to note that the extension operator

$$e^+ : H^{\sigma}(\mathbf{R}_+) \longrightarrow H^{\sigma}(\mathbf{R})$$

is continuous.

By using an interpolation argument, we may assume that $\sigma_1 \in \mathbf{N}$. We proceed by induction on σ_1. Recall that the norm of a function $v \in H^{\sigma+(1-\mu,0)}(\mathbf{R}_+)$ can be estimated by the norm of v in $H^{\sigma-(\mu,0)}(\mathbf{R}_+)$ and the norm of $\partial_{x_n} v$ in $H^{\sigma-(\mu,0)}(\mathbf{R}_+)$. We note that

$$\partial_{x_n}\left(e^+ u\right) = e^+(\partial_{x_n} u) + u(0)\delta_0 \quad \text{for all } u \in H^{\sigma+(1,0)}(\mathbf{R}).$$

Hence we have

$$\begin{aligned}
&\partial_{x_n}\left(\left[\mathrm{Op}^{+}_{x_n} a\right] u\right) \\
&= \partial_{x_n}\left(r^+\left[\mathrm{Op}_{x_n}(a)\right] e^+ u\right) \\
&= \left[\mathrm{Op}^{+}_{x_n}(\partial_{x_n} a)\right] u + \left[\mathrm{Op}^{+}_{x_n} a\right]\left(\partial_{x_n}(e^+ u)\right) \\
&= \left[\mathrm{Op}^{+}_{x_n}(\partial_{x_n} a)\right] u + \left[\mathrm{Op}^{+}_{x_n} a\right] e^+(\partial_{x_n} u) + \gamma_0 u\, r^+\left[\mathrm{Op}_{x_n}(a)\right]\delta_0.
\end{aligned}$$

Since we have

$$\partial_{x_n} \kappa_{\langle\xi'\rangle^{-1}} = \langle\xi'\rangle^{-1} \kappa_{\langle\xi'\rangle^{-1}} \partial_{x_n},$$

we obtain that

$$\begin{aligned}
&\left\|\partial_{x_n} \kappa_{\langle\xi'\rangle^{-1}}\left[\mathrm{Op}^{+}_{x_n} a\right]\kappa_{\langle\xi'\rangle}\right\|_{\mathcal{L}(H^{\sigma+(1,0)}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+))} \\
&= \left\|\langle\xi'\rangle^{-1}\kappa_{\langle\xi'\rangle^{-1}}\partial_{x_n}\left[\mathrm{Op}^{+}_{x_n} a\right]\kappa_{\langle\xi'\rangle}\right\|_{\mathcal{L}(H^{\sigma+(1,0)}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+))} \\
&\le \langle\xi'\rangle^{-1}\left\|\kappa_{\langle\xi'\rangle^{-1}}\left[\mathrm{Op}^{+}_{x_n}(\partial_{x_n} a)\right]\kappa_{\langle\xi'\rangle}\right\|_{\mathcal{L}(H^{\sigma+(1,0)}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+))} \\
&\quad + \left\|\kappa_{\langle\xi'\rangle^{-1}}\left[\mathrm{Op}^{+}_{x_n} a\right]\kappa_{\langle\xi'\rangle}\right\|_{\mathcal{L}(H^{\sigma}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+))}
\end{aligned}$$

$$+\langle\xi'\rangle^{-1}\left\|\kappa_{\langle\xi'\rangle^{-1}}r^{+}\left[\mathrm{Op}_{x_n}(a)\right]\delta_0\right\|_{\mathcal{L}(\mathbf{C},H^{\sigma-(\mu,0)}(\mathbf{R}_+))}$$
$$\times\left\|\gamma_0\kappa_{\langle\xi'\rangle}\right\|_{\mathcal{L}(H^{\sigma+(1,0)}(\mathbf{R}_+),\mathbf{C})}.$$

By induction on σ_1, all terms are of order $O(\langle\xi'\rangle^{\mu})$; for the last term, we have only to apply Lemma B.25 with $\ell := 0$ and Example B.2 with $j := 0$.

The proof of Theorem B.27 is complete.

B.4 Symbol Classes for the Boutet de Monvel Calculus

B.4.1 The Operator ∂_+

In the subsequent text we shall denote by ∂_+ the usual derivative considered as a differential operator on distributions over $\mathbf{R}_+$. We choose this notation in order to distinguish ∂_+ from ∂_t which also acts on distributions on the full line. For any $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ with $\sigma_1 > -1/2$, we have

$$\partial_+ = r^+\partial_t e^+ : H^{\sigma}(\mathbf{R}_+) \longrightarrow H^{\sigma-(1,0)}(\mathbf{R}_+).$$

On the other hand, the operator ∂_+ acts on all spaces $H^{\sigma}(\mathbf{R}_+)$ and defines an (x', ξ')-independent element of $S^1(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), H^{\sigma-(1,0)}(\mathbf{R}_+))$ for every $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$:

$$\partial_+ \in S^1\left(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), H^{\sigma-(1,0)}(\mathbf{R}_+)\right).$$

B.4.2 Boundary Symbols on the Half-Space $\mathbf{R}^n_+$

(a) A *potential symbol* k of order μ is an element of the symbol class

$$S^{\mu}\left(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}; \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right) \tag{B.10}$$
$$= \bigcap_{\sigma=(\sigma_1,\sigma_2)\in\mathbf{R}^2} S^{\mu}\left(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}; \mathbf{C}, H^{\sigma}(\mathbf{R}_+)\right).$$

This is a Fréchet space with the topology of the projective limit. The group action is the identity on $\mathbf{C}$ and given by formula (B.2) on $H^{\sigma}(\mathbf{R}_+)$.

(b) A *trace symbol* t of order μ and type zero is an element of the symbol class

$$
\begin{aligned}
&S^{\mu}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathbf{C}) \\
&= \bigcap_{\sigma=(\sigma_1,\sigma_2)\in\mathbf{R}^2} S^{\mu}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^{\sigma}(\overline{\mathbf{R}_+}), \mathbf{C}).
\end{aligned}
\tag{B.11}
$$

Again this is a Fréchet space with the projective limit topology.

For every $\sigma_1 > -1/2$, the weighted Sobolev space $H^{(\sigma_1,\sigma_2)}(\mathbf{R}_+)$ is embedded in the space $H_0^{\tau}(\overline{\mathbf{R}_+})$ with $\tau = (\min\{\sigma_1, 0\}, \sigma_2)$, by using the zero-extension. Therefore, a trace symbol t of order μ and type 0 defines an element of $S^{\mu}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{(\sigma_1,\sigma_2)}(\mathbf{R}_+), \mathbf{C})$ whenever $\sigma_1 > -1/2$.

A trace symbol t of order μ and type $d \in \mathbf{N}_0$ is a sum of operator-valued symbols

$$
t = \sum_{j=0}^{d} t_j \partial_+^j,
$$

where $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$, and each t_j is a trace symbol of order $\mu - j$ and type zero and the summation is in the symbol class

$$
S^{\mu}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{(\sigma_1,\sigma_2)}(\mathbf{R}_+), \mathbf{C})
$$

for $\sigma_1 > d - 1/2$. We endow the space of trace symbols of order μ and type d with the topology of the *non-direct sum* of Fréchet spaces (see Definition B.28 below). Grubb [Gb] uses 'class' instead of 'type'.

(c) A *singular Green symbol* g of order μ and type zero is an element of the symbol class

$$
\begin{aligned}
&S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right) \\
&= \bigcap_{\sigma,\tau\in\mathbf{R}^2} S^{\mu}\big(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^{\sigma}(\overline{\mathbf{R}_+}), H^{\tau}(\mathbf{R}_+)\big),
\end{aligned}
\tag{B.12}
$$

endowed with the Fréchet topology of the projective limit.

A singular Green symbol of order μ and type zero furnishes an element of the symbol class

$$
S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right), \quad \sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2,
$$

provided that $\sigma_1 > -1/2$. We define the singular Green symbols of order μ and type d as the sums

$$
g = \sum_{j=0}^{d} g_j \partial_+^j,
$$

where each g_j is a singular Green symbol of order $\mu - j$ and type zero, and the summation is in the symbol class

$$S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right), \quad \sigma_1 > d - 1/2.$$

The resulting space carries the Fréchet topology of the non-direct sum.

(d) A *boundary symbol* a in the Boutet de Monvel calculus of order μ and type d is an operator-valued symbol of the form

$$a = \begin{pmatrix} \mathrm{Op}_{x_n}^{+} p + g & k \\ t & s \end{pmatrix}, \tag{B.13}$$

where $p \in S^{\mu}(\mathbf{R}^n \times \mathbf{R}^n)$, g is a singular Green symbol of order μ and type d, k is a potential symbol of order μ, t is a trace symbol of order μ and type d, and $s \in S^{\mu}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$. We know from Theorem B.27 that

$$\mathrm{Op}_{x_n}^{+} p \in S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), H^{\sigma-(\mu,0)}(\mathbf{R}_+)\right),$$
$$\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2,$$

provided that $\sigma_1 > -1/2$. Also it follows that the symbol

$$s \in S^{\mu}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}) = S^{\mu}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathbf{C}, \mathbf{C})$$

is an operator-valued symbol. Therefore, for each $\sigma_1 > d - 1/2$ a boundary symbol of order μ and type d can be considered an element of the symbol class

$$S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+) \oplus \mathbf{C}, H^{\sigma-(\mu,0)}(\mathbf{R}_+) \oplus \mathbf{C}\right). \tag{B.14}$$

That is, we have

$$a = \begin{pmatrix} \mathrm{Op}_{x_n}^{+} p + g & k \\ t & s \end{pmatrix} : \begin{matrix} H^{\sigma}(\mathbf{R}_+) \\ \oplus \\ \mathbf{C} \end{matrix} \longrightarrow \begin{matrix} H^{\sigma-(\mu,0)}(\mathbf{R}_+) \\ \oplus \\ \mathbf{C} \end{matrix}$$

We endow the space of boundary symbols of order μ and type d with the Fréchet topology of the non-direct sum of the Fréchet spaces involved.

(e) We obtain the notions of regularizing potential, trace, singular Green, and boundary symbols by taking the intersection of the corresponding spaces over all $\mu \in \mathbf{R}$.

(f) The definitions in (a), (b), and (c) extend easily to double symbols.
We obtain classical symbol classes by taking $S^{\mu}_{\mathrm{cl}}(\ldots)$.

(g) Since we eventually want to treat operators acting on sections of vector bundles E_1, E_2 over a compact manifold X and F_1, F_2 over its boundary ∂X, we will have to replace the spaces

$$\mathbf{C}, \ \mathcal{S}(\mathbf{R}_+), \ H^{\sigma}(\mathbf{R}_+), \ H^{\sigma}_0(\overline{\mathbf{R}_+}), \ \mathcal{S}'(\mathbf{R}_+),$$

in general by N-fold Cartesian products for a suitable positive integer N:

$$\mathbf{C}^{N_1},\ \mathcal{S}(\mathbf{R}_+)^{N_2},\ H^\sigma(\mathbf{R}_+)^{N_3},\ H_0^\sigma(\overline{\mathbf{R}_+})^{N_4},\ \mathcal{S}'(\mathbf{R}_+)^{N_5}.$$

In order to avoid superfluous notation, we shall not write the N_j unless clarity demands it.

The following notion of the non-direct sum has already been used in the above definitions (b) and (d):

Definition B.28. Let E, F be Fréchet spaces. Assume that E and F are continuously embedded in the same Hausdorff vector space. The exterior direct sum

$$E \oplus F$$

is a Fréchet space and has the closed subspace

$$\mathcal{N} = \{(a, -a) : a \in E \cap F\}.$$

Then the *non-direct sum* of E and F is the Fréchet space

$$E + F := (E \oplus F)\,/\mathcal{N}.$$

Definition B.29. We call the boundary symbol a in (B.13) a *generalized singular Green symbol* of order μ and type d, if $p = 0$. For $d = 0$, we obtain an element of the symbol class

$$S^\mu(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+)^{N_1} \oplus \mathbf{C}^{N_2}, \mathcal{S}(\mathbf{R}_+)^{N_3} \oplus \mathbf{C}^{N_4})$$

with suitable non-negative integers N_1, N_2, N_3, $N_4 \in \mathbf{N}_0$:

$$a = \begin{pmatrix} g & k \\ t & s \end{pmatrix} : \begin{matrix} \mathcal{S}'(\mathbf{R}_+)^{N_1} \\ \oplus \\ \mathbf{C}^{N_2} \end{matrix} \longrightarrow \begin{matrix} \mathcal{S}(\mathbf{R}_+)^{N_3} \\ \oplus \\ \mathbf{C}^{N_4} \end{matrix}$$

The next proposition lays the foundation for Theorem B.32 below:

Proposition B.30. *Let $u \in L^2(\mathbf{R} \times \mathbf{R})$, and assume that $x_j^k D_{x_j}^\ell u(x_1, x_2)$ is an element of $L^2(\mathbf{R} \times \mathbf{R})$ for $j = 1, 2$ and all $k, \ell \in \mathbf{N}_0$. Then it follows that $u \in \mathcal{S}(\mathbf{R} \times \mathbf{R})$.*

Proof. For each $\ell \in 2\mathbf{N}$, we have

$$\left[\mathrm{Op}\left(1 + \xi_1^\ell + \xi_2^\ell\right)\right] u \in L^2(\mathbf{R}^2).$$

Since $1 + \xi_1^\ell + \xi_2^\ell$ is an elliptic symbol of order ℓ, it follows that

$$u \in H^{(\ell,0)}(\mathbf{R}^2) \quad \text{for all } \ell \in \mathbf{N}.$$

On the other hand, the assumption with $\ell := 0$ implies that

$$u \in H^{(0,k)}(\mathbf{R}^2) \quad \text{for all } k \in \mathbf{N}.$$

Therefore, we have only to show that

$$\bigcap_{s \geq 0} \left[H^{(s,0)}(\mathbf{R}^2) \cap H^{(0,s)}(\mathbf{R}^2) \right] = \mathcal{S}(\mathbf{R}^2).$$

Denote by F the Fréchet space on the left-hand side:

$$F := \bigcap_{s \geq 0} \left[H^{(s,0)}(\mathbf{R}^2) \cap H^{(0,s)}(\mathbf{R}^2) \right].$$

A system $\{p_k : k \in \mathbf{N}\}$ of semi-norms for F is given by

$$p_k(u) = \left\| \langle x \rangle^{2k} u \right\|_{L^2} + \left\| \langle D \rangle^{2k} u \right\|_{L^2},$$

where

$$\langle x \rangle^{2k} = \left(1 + x_1^2 + x_2^2\right)^k, \quad \langle D \rangle^{2k} = \left(1 + \partial_1^2 + \partial_2^2\right)^k.$$

It is clear that the Schwartz space $\mathcal{S}(\mathbf{R}^2)$ is a subset of F. Moreover, $\mathcal{S}(\mathbf{R}^2)$ is dense in F. Indeed, fix a function $\varphi \in C_0^\infty(\mathbf{R}^2)$ such that $\varphi(0) = 1$, and let

$$\varphi_\varepsilon(x) = \varphi(\varepsilon x), \quad \varepsilon > 0.$$

Then it follows that $\varphi_\varepsilon u \in \mathcal{S}(\mathbf{R}^2)$ for each $u \in F$. By using the Lebesgue dominated convergence theorem, we obtain that

$$\varphi_\varepsilon u \longrightarrow u \quad \text{in } F \text{ as } \varepsilon \to 0.$$

This proves the density of $\mathcal{S}(\mathbf{R}^2)$ in F.

For every $u \in \mathcal{S}(\mathbf{R}^2)$, integration by parts and the Schwarz inequality imply that

$$\left\| x^\alpha D_x^\beta u \right\|_{L^2}^2 = \int_{\mathbf{R}^2} x^\alpha D_x^\beta u \cdot \overline{x^\alpha D_x^\beta u}\, dx \leq C p_k(u)^2,$$

provided that $k \geq \max\{|\alpha|, |\beta|\}$. Here C is a generic positive constant depending only on α and β. We obtain that $x^\alpha D_x^\beta$ extends to a continuous operator on F, and that $F \subset \mathcal{S}(\mathbf{R}^2)$.

The proof of Proposition B.30 is complete.

Definition B.31. We shall write $\mathbf{R}^2_{++} = \mathbf{R}_+ \times \mathbf{R}_+$.

The next theorem asserts the equivalence of the operator-valued approach and the standard definition: The estimates in part (i) of Theorem B.32 are those required in the usual presentation of singular Green operators (cf. Grubb [Gb, (2.3.28)]).

Theorem B.32 (Singular Green symbol kernels). *Let $\mu \in \mathbf{R}$ and let $\{g(y,\eta) : y,\ \eta \in \mathbf{R}^q\}$ be a family of operators $L^2(\mathbf{R}_+) \to L^2(\mathbf{R}_+)$. Then the following three conditions (i), (ii) and (iii) are equivalent:*

(i) Each $g(y,\eta)$ is an integral operator with a symbol kernel $\tilde{g}(y,\eta;u,v)$ satisfying the following estimates: For all $k,\ k',\ \ell,\ \ell' \in \mathbf{N}_0$ and $\alpha,\ \beta \in \mathbf{N}_0^q$, there exists a positive constant c depending on $k,\ k',\ \ell,\ \ell',\ \alpha,\ \beta$, with the estimate

$$\left\| u^k D_u^{k'} v^\ell D_v^{\ell'} D_\eta^\alpha D_y^\beta \tilde{g}(y,\eta;u,v) \right\|_{L^2(\mathbf{R}_{++})} \leq c\, \langle \eta \rangle^{\mu - |\alpha| - k + k' - \ell + \ell'} .$$

(ii) $g(y,\eta) \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; \mathbf{S}'(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+))$.
(iii) $g \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; L^2(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+))$ and $g^* \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; L^2(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+))$ where $g^*(y,\eta) = \{g(y,\eta)^* : y,\ \eta \in \mathbf{R}^q\}$ is the family of pointwise adjoints.

Proof. (i) $\Longrightarrow$ (ii): It is easy to verify that $\kappa_{\langle\eta\rangle^{-1}} D_\eta^\alpha D_y^\beta g(y,\eta) \kappa_{\langle\eta\rangle}$ is the integral operator with the symbol kernel

$$h_{\alpha,\beta}(y,\eta;u,v) = \left(D_\eta^\alpha D_y^\beta \tilde{g} \right) \left(y,\eta; \langle\eta\rangle^{-1} u, \langle\eta\rangle^{-1} v \right) \langle\eta\rangle^{-1} .$$

The estimates for $\tilde{g}$ imply that $h_{\alpha,\beta}(y,\eta;\cdot,\cdot)$ is a function in the space $\mathcal{S}(\mathbf{R}^2_{++})$, and all its semi-norms are of order $O(\langle\eta\rangle^{\mu-|\alpha|})$. In particular, the symbol kernel $h_{\alpha,\beta}(y,\eta;u,v)$ induces an operator from $\mathcal{S}'(\mathbf{R}_+)$ into $\mathcal{S}(\mathbf{R}_+)$, and we have condition (ii).

(ii) $\Longrightarrow$ (iii): It is trivial that

$$g \in S^\mu \left(\mathbf{R}^q \times \mathbf{R}^q; L^2(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+) \right) .$$

Moreover, we have, for all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$,

$$g \in S^\mu \left(\mathbf{R}^q \times \mathbf{R}^q; H_0^\sigma(\overline{\mathbf{R}_+}), L^2(\mathbf{R}_+) \right) .$$

Hence Sect. B.2.4 shows the asserted property of $g^*(y,\eta)$.

(iii) $\Longrightarrow$ (i): The operator

$$\kappa_{\langle\eta\rangle^{-1}} g(y,\eta)\kappa_{\langle\eta\rangle} : L^2(\mathbf{R}_+) \longrightarrow \mathcal{S}(\mathbf{R}_+)$$

is continuous. In particular, it is a Hilbert–Schmidt operator on $L^2(\mathbf{R}_+)$ and thus has an integral kernel $h_1(y,\eta;\cdot,\cdot) \in L^2(\mathbf{R}^2_{++})$ with the norm

$$\|h_1(y,\eta;\cdot,\cdot)\|_{L^2(\mathbf{R}^2_{++})} = \left\|\kappa_{\langle\eta\rangle^{-1}} g(y,\eta)\kappa_{\langle\eta\rangle}\right\|_{HS(L^2(\mathbf{R}_+))}.$$

The last norm is bounded by the norm in the space $\mathcal{L}(L^2(\mathbf{R}_+), H^{(1,1)}(\mathbf{R}_+))$. By a direct calculation, the operator $g(y,\eta)$ then has the integral kernel

$$\tilde{g}_1(y,\eta;u,v) = h_1(y,\eta;\langle\eta\rangle\, u, \langle\eta\rangle\, v)\,\langle\eta\rangle\,. \tag{B.15}$$

Correspondingly, the operator $\kappa_{\langle\eta\rangle^{-1}} g^*(y,\eta)\kappa_{\langle\eta\rangle}$ has the kernel $h_2(y,\eta;u,v)$, and we have

$$h_1(y,\eta;u,v) = \overline{h_2(y,\eta;v,u)}. \tag{B.16}$$

The mapping

$$u^k D_u^{k'} \kappa_{\langle\eta\rangle^{-1}} D_\eta^\alpha D_y^\beta g(y,\eta)\kappa_{\langle\eta\rangle} : L^2(\mathbf{R}_+) \longrightarrow \mathcal{S}(\mathbf{R}_+)$$

is also continuous. Therefore, we have the estimate

$$\left\|u^k D_u^{k'} D_\eta^\beta D_y^\beta h_1(y,\eta;u,v)\right\|_{L^2(\mathbf{R}^2_{++})} = O(\langle\eta\rangle^{\mu-|\alpha|}).$$

By using (B.16), we also have the estimate

$$\left\|v^\ell D_v^{\ell'} D_\eta^\alpha D_y^\beta h_1(y,\eta;u,v)\right\|_{L^2(\mathbf{R}^2_{++})} = O(\langle\eta\rangle^{\mu-|\alpha|}).$$

These estimates, together with Proposition B.30, show that

$$\left\|u^k D_u^{k'} v^\ell D_v^{\ell'} D_\eta^\beta D_y^\beta h_1(y,\eta;u,v)\right\|_{L^2(\mathbf{R}^2_{++})} = O(\langle\eta\rangle^{\mu-|\alpha|}).$$

By combining this fact with (B.15), we obtain condition (i).
The proof of Theorem B.32 is now complete.

Lemma B.33. *We may replace the estimates in part (i) of Theorem B.32 by the following:*

$$\sup_{u,v}\left|u^k D_u^{k'} v^\ell D_v^{\ell'} D_\eta^\alpha D_y^\beta \tilde{g}(y,\eta;u,v)\right| \leq c\,\langle\eta\rangle^{1+\mu-|\alpha|-k+k'-\ell+\ell'}.$$

Proof. Lemma B.33 is a consequence of the following two estimates (B.17) and (B.18) for all functions $f \in \mathcal{S}(\mathbf{R}_+)$:

$$\sup_{t>0} |f(t)|^2 \leq 2\|f\|_{L^2(\mathbf{R}_+)} \|D_t f\|_{L^2(\mathbf{R}_+)}. \tag{B.17}$$

$$\|f\|_{L^2(\mathbf{R}_+)} \leq \langle\eta\rangle^{-1/2} \sup_{t>0}(1 + \langle\eta\rangle\, t)|f(t)|. \tag{B.18}$$

Estimate (B.17) follows by applying the Schwarz inequality to

$$\begin{aligned} |f(t)|^2 &= -\int_t^\infty \frac{\partial}{\partial s}\left\{f(s)\cdot\overline{f(s)}\right\} ds \\ &= -\int_t^\infty \frac{\partial f}{\partial s}\cdot\overline{f(s)}\, ds - \int_t^\infty f(s)\cdot\overline{\frac{\partial f}{\partial s}}\, ds. \end{aligned}$$

In order to prove estimate (B.18), we have only to note that

$$\begin{aligned} \|f\|^2_{L^2(\mathbf{R}_+)} &\leq \sup_{t>0}\left((1 + \langle\eta\rangle\, t)^2\, |f(t)|^2\right) \int_0^\infty \frac{1}{(1 + \langle\eta\rangle\, s)^2}\, ds \\ &= \sup_{t>0}\left((1 + \langle\eta\rangle\, t)^2\, |f(t)|^2\right) \langle\eta\rangle^{-1}. \end{aligned}$$

The proof of Lemma B.33 is complete.

Theorem B.34 (Potential and trace symbol kernels).

(a) Let $\mu \in \mathbf{R}$ and let $k = \{k(y,\eta) : y, \eta \in \mathbf{R}^q\}$ be a family of operators in $\mathcal{L}(\mathbf{C}, L^2(\mathbf{R}_+))$. Then the following four assertions (i)–(iv) are equivalent:

(i) The operators $\{k(y,\eta) : y, \eta \in \mathbf{R}^q\}$ act on $\mathbf{C}$ by multiplication by functions $\tilde{k}(y,\eta;\cdot)$ satisfying the following estimates: For all ℓ, $\ell' \in \mathbf{N}_0$ and all multi-indices α, $\beta \in \mathbf{N}_0^{n-1}$, there is a constant c, depending on ℓ, ℓ', α and β, such that

$$\left\| u^\ell D_u^{\ell'} D_\eta^\alpha D_y^\beta \tilde{k}(y,\eta;u)\right\|_{L^2(\mathbf{R}_+)} \leq c\, \langle\eta\rangle^{\mu-|\alpha|-\ell+\ell'}.$$

(ii) $k(y,\eta) \in S^\mu(\mathbf{R}^q \times \mathbf{R}^q; \mathbf{C}, \mathcal{S}(\mathbf{R}_+))$, that is, k is a potential symbol.

(iii) The family k^* of pointwise adjoints $k^* = \{k(y,\eta)^* : y, \eta \in \mathbf{R}^q\}$ is an element of $S^\mu(\mathbf{R}^q \times \mathbf{R}^q; \mathcal{S}'(\mathbf{R}_+), \mathbf{C})$.

(iv) We may replace the estimates in Assertion (i) by the following:

$$\sup_u \left| u^\ell D_u^{\ell'} D_\eta^\alpha D_y^\beta \tilde{k}(y,\eta;u)\right| \leq c\, \langle\eta\rangle^{1/2+\mu-|\alpha|-\ell+\ell'}.$$

(b) A trace symbol $t(y,\eta) \in S^{\mu}(\mathbf{R}^q \times \mathbf{R}^q; S'(\mathbf{R}_+), \mathbf{C}))$ *of order* μ *and type zero is given by*

$$t(y,\eta)f = \int_0^\infty \tilde{t}(y,\eta;u)\, f(u)\, du, \quad f \in \mathcal{S}(\mathbf{R}_+),$$

where $\tilde{t}(y,\eta;u)$ *satisfies the estimates in Assertion (i) or Assertion (iv).*

In particular, potential and trace symbols are dual to each other.

Proof. The equivalence of Assertions (i) and (ii) can be shown just as in Proposition B.30. Furthermore, we find that Assertions (ii) and (iii) are equivalent, since we have (cf. Sect. B.2.4)

$$H_0^{-\sigma}(\overline{\mathbf{R}_+}) = \text{ the dual space } H^{\sigma}(\mathbf{R}_+)' \text{ of } H^{\sigma}(\mathbf{R}_+).$$

Finally, we may use sup-norm estimates by the same arguments as in Lemma B.33. Assertion (b) is an immediate consequence of Assertion (a).

The proof of Theorem B.34 is complete.

Example B.35. In Example B.2 we proved that

$$\gamma_j \in S^{j+1/2}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}_+), \mathbf{C})$$

whenever $\sigma_1 > j + 1/2$. Now we show that $\gamma_j(x', \xi')$ is a trace symbol of order $j + 1/2$ and type $j + 1$. Indeed, we can write

$$\gamma_0(x',\xi')f = \int_0^\infty \langle \xi' \rangle\, e^{-y_n \langle \xi' \rangle} f(y_n)\, dy_n - \int_0^\infty e^{-y_n \langle \xi' \rangle}\, \partial_{y_n} f(y_n)\, dy_n \tag{B.19}$$

for all $f \in \mathcal{S}(\mathbf{R}_+)$. Hence it follows that

$$\gamma_0 = t_0 + t_1\, \partial_+ = t_0 + t_1\, r^+ \partial_t e^+,$$

where

$$t_0(x',\xi')f = \langle \xi' \rangle \int_0^\infty e^{-y_n \langle \xi' \rangle} f(y_n)\, dy_n,$$

$$t_1(x',\xi')f = -\int_0^\infty e^{-y_n \langle \xi' \rangle} f(y_n)\, dy_n.$$

In particular, we have

$$t_0 \kappa_{\langle \xi' \rangle} f = \langle \xi' \rangle^{1/2} \int_0^\infty e^{-t} f(t)\, dt = \langle \xi' \rangle^{1/2} \left\langle e^{-\cdot}, f \right\rangle_{\mathcal{S}(\mathbf{R}_+), \mathcal{S}'(\mathbf{R}_+)},$$

and so

$$t_0 \in S^{1/2}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathbf{C}).$$

In the same way, it follows that

$$t_1 \in S^{-1/2}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathbf{C}).$$

By applying (B.19) to the function $\partial_{x_n}^j f$, we obtain the desired result by iteration.

Remark B.36. It is obvious that there are many different ways to write

$$\gamma_j = \sum_{\ell=0}^{j+1} t_\ell \partial_+^\ell$$

with trace symbols t_ℓ of order $j + 1/2 - \ell$ and type zero.

B.5 The Analysis of Compositions

We may compose two boundary symbols of orders $\mu_1, \mu_2 \in \mathbf{Z}$ and types $d_1, d_2 \in \mathbf{N}_0$, say

$$a_1 = \begin{pmatrix} \mathrm{Op}_{x_n}^+ p_1 + g_1 & k_1 \\ t_1 & s_1 \end{pmatrix},$$

$$a_2 = \begin{pmatrix} \mathrm{Op}_{x_n}^+ p_2 + g_2 & k_2 \\ t_2 & s_2 \end{pmatrix},$$

provided that the dimensions of the matrices are compatible. According to part (b) of Lemma B.10, we obtain that

$$a_1 a_2 \in S^{\mu_1+\mu_2}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}_+)^{N_1}, \mathbf{C}^{N_2}, H^{\sigma-(\mu_1+\mu_2,0)}(\mathbf{R}_+)^{N_3}, \mathbf{C}^{N_4})$$

for suitable non-negative integers N_1, N_2, N_3 and N_4, assuming that $\sigma_1 > d_2 - 1/2$ and that $\sigma_1 - \mu_2 > d_1 - 1/2$. We can compute the composition

$$a_1 a_2 = \begin{pmatrix} p_1^+ + g_1 & k_1 \\ t_1 & s_1 \end{pmatrix} \begin{pmatrix} p_2^+ + g_2 & k_2 \\ t_2 & s_2 \end{pmatrix}$$

as follows:

$$\begin{pmatrix} (p_1 \sharp_n p_2)^+ + \ell(p_1, p_2) + p_1^+ g_2 + g_1 p_2^+ + g_1 g_2 + k_1 t_2 & p_1^+ k_2 + g_1 k_2 + k_1 s_2 \\ t_1 p_2^+ + t_1 g_2 + s_1 t_2 & t_1 k_2 + s_1 s_2 \end{pmatrix}.$$

Here we have written p_j^+ instead of $\mathrm{Op}_{x_n}^+ p_j$, $j = 1, 2$, in order to save space, and $(p_1 \sharp_n p_2)^+$ instead of $\mathrm{Op}_{x_n}^+(p_1 \sharp_n p_2)$; the notation $\sharp_n$ indicates composition with respect to the variable x_n.

We will verify the following 13 assertions (i)–(xiii):

(i) $\ell(p_1, p_2) = p_1^+ p_2^+ - (p_1 \sharp_n p_2)^+$ is a singular Green symbol of type $\mu_{2+} := \max\{\mu_2, 0\}$.
(ii) $p_1^+ g_2$ is a singular Green symbol of type d_2.
(iii) $g_1 p_2^+$ is a singular Green symbol of type $(\mu_2 + d_1)_+ = \max\{\mu_2 + d_1, 0\}$.
(iv) $g_1 g_2$ is a singular Green symbol of type d_2.
(v) $k_1 t_2$ is a singular Green symbol of type d_2.
(vi) $p_1^+ k_2$ is a potential symbol.
(vii) $g_1 k_2$ is a potential symbol.
(viii) $k_1 s_2$ is a potential symbol.
(ix) $t_1 p_2^+$ is a trace symbol of type $(\mu_2 + d_1)_+$.
(x) $t_1 g_2$ is a trace symbol of type d_2.
(xi) $s_1 t_2$ is a trace symbol of type d_2.
(xii) $t_1 k_2$ is a pseudo-differential symbol.
(xiii) $s_1 s_2$ is a pseudo-differential symbol.

In all cases the order of the respective symbols is $\mu_1 + \mu_2$. When referring to the symbols in Assertions (xii) and (xiii) as 'pseudo-differential', we stress that the Banach spaces E, $\tilde{E}$ in the sense of Sect. B.2.3 are simply $\mathbf{C}^{N_2}$ and $\mathbf{C}^{N_4}$, respectively. Therefore, we obtain the following result:

Theorem B.37. *The pointwise composition $a_1 a_2$ of two boundary symbols a_1 and a_2 of orders μ_1 and μ_2 and types d_1 and d_2, respectively, is a boundary symbol of order $\mu_1 + \mu_2$ and type* $\max\{\mu_2 + d_1, d_2\}$. *Its pseudo-differential part is $p_1 \sharp_n p_2$.*

The proof is rather long, and we shall break it up into a sequence of partial results.

Step 1: First, we deal with the easy cases, namely, Assertions (ii), (iv), (v), (vi), (vii), (viii), (x), (xi), (xii) and (xiii). We may assume that $N_1 = N_2 = N_3 = N_4 = 1$. We write

$$g_1 = \sum_{j=0}^{d_1} g_{1j}\, \partial_+^j, \quad g_2 = \sum_{j=0}^{d_2} g_{2j}\, \partial_+^j.$$

For Assertion (ii), we note that

$$p_1^+ g_2 = \sum_{j=0}^{d_2} \left(p_1^+ g_{2j}\right) \partial_+^j,$$

where, according to Theorem B.27 and Lemma B.10, we obtain that

$$p_1^+ g_{2j} \in S^{\mu_1+\mu_2-j}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right).$$

For Assertion (iv), we observe that

$$\partial_+^j \in S^j \left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}, H^\sigma(\mathbf{R}_+), H^{\sigma-(j,0)}(\mathbf{R}_+)\right) \tag{B.20}$$

for all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$. Then Lemma B.10 yields the desired result.

The proof of Assertions (v), (vi), (vii), (x) and (xii) is analogous.

For Assertions (viii), (xi) and (xiii), we recall additionally that

$$s_2 \in S^{\mu_2}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}) = S^{\mu_2}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathbf{C}, \mathbf{C}).$$

Step 2: The compositions in Assertions (iii) and (ix) are slightly more delicate. For example, we consider Assertion (iii): In order to show that $g_1 p_2^+$ is a singular Green operator, we first note that

$$\begin{aligned}\partial_+ \left(p_2^+\right) &= \left(r^+ \partial_{x_n} e^+\right) \left(r^+ \left[\mathrm{Op}_{x_n}(p_2)\right] e^+\right) = r^+ \left(\partial_{x_n} \left[\mathrm{Op}_{x_n}(p_2)\right] e^+\right) \\ &= \mathrm{Op}^+_{x_n} \left[i p_2(x,\xi)\xi_n + \partial_{x_n} p_2(x,\xi)\right].\end{aligned}$$

By iteration, we have

$$\partial_+^j \left(p_2^+\right) = \mathrm{Op}^+_{x_n} q_j$$

for a suitable symbol $q_j \in S^{\mu_2+j}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$. So it is no restriction to assume that $d_1 = 0$.

Next we shall establish a central point in Sect. B.5.1 below: For a fixed positive integer N, we may split p_2 into a 'differential' part and one that acts on the H_0^σ-spaces:

$$\mathrm{Op}_{x_n}(p_2) = \sum_{j=0}^{N} \sum_{k=0}^{\mu_2+} x_n^j s_{jk} \partial_{x_n}^k + a$$

with

$$\begin{aligned}&s_{jk} \in S^{\mu_2-k}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}),\\ &a \in S^{\mu_2} \left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}), H^{\sigma-(\mu_2+,0)}(\mathbf{R})\right) \text{ for all } \sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2.\end{aligned}$$

Moreover, we have

$$\chi_+ a \in S^{\mu_2} \left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^{\sigma-(\mu_2+,0)}(\overline{\mathbf{R}_+})\right), \quad -N \leq \sigma_1 \leq 0.$$

where χ_+ denotes the restriction to $\mathbf{R}_+$ followed by embedding into the space $H_0^0(\overline{\mathbf{R}_+})$. This makes sense: For every $u \in C_0^\infty(\mathbf{R}_+)$ and x', ξ' fixed, it follows that $a(x', \xi')u$ is smooth on $\mathbf{R}$. As we shall show, $\chi_+ a(x', \xi')$ extends to an element of the space $\mathcal{L}(H_0^\sigma(\overline{\mathbf{R}_+}), H_0^{\sigma-(\mu_{2+},0)}(\overline{\mathbf{R}_+}))$ for every $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ and satisfies the required symbol estimates.

After the above reduction, it follows that g_1 is a singular Green symbol of order μ_1 and type 0, and we have

$$g_1 p_2^+ = \sum_{j=0}^{N} \sum_{k=0}^{\mu_{2+}} \left(g_1 x_n^j s_{jk}\right) \partial_+^k + g_1 \chi_+ a.$$

It is clear that s_{jk} induces an element of the symbol class

$$S^{\mu_2-k}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^\sigma(\overline{\mathbf{R}_+})\right)$$

and further that x_n^j induces an element of the symbol class

$$x_n^j \in S^{-j}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^{\sigma-(0,j)}(\overline{\mathbf{R}_+})\right).$$

Therefore, the summation yields a singular Green symbol of order $\mu_1 + \mu_2$

$$\sum_{k=0}^{\mu_{2+}} \left(\sum_{j=0}^{N} \left(g_1 x_n^j s_{jk}\right) \right) \partial_+^k \in S^{\mu_1+\mu_2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathcal{S}(\mathbf{R}_+)\right),$$

while

$$g_1 \chi_+ a \in S^{\mu_1+\mu_2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), \mathcal{S}(\mathbf{R}_+)\right), \quad \sigma_1 \geq -N.$$

By Lemma B.38 below with $\mu := \mu_1 + \mu_2$ and $d := \mu_{2+}$, it follows that $g_1 p_2^+$ is a singular Green symbol of order $\mu_1 + \mu_2$. The type is μ_{2+}, since we assumed that $d_1 = 0$; in the general case it is $(\mu_2 + d_1)_+$ by the above consideration.

Also the composition in Assertion (ix) follows from the above representation for $\mathrm{Op}_{x_n}(p_2)$.

Step 3: Assertion (i) presents additional complications. We shall deal with them below.

First, however, we state a lemma which was employed in the above proof:

Lemma B.38. *Let $d \in \mathbf{N}_0$ and $g \in S^\mu(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}_+), L^2(\mathbf{R}_+))$ with $\sigma_1 > d - 1/2$. Assume that, for each $N \in \mathbf{N}$, there exist symbols g_{kN} such that*

$$g_{kN} \in S^{\mu-k}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^{-(N,N)}(\overline{\mathbf{R}_+}), H^{(N.N)}(\mathbf{R}_+)\right)$$

with

$$g = \sum_{k=0}^{d} g_{kN} \partial_+^k .$$

Then g is a singular Green symbol of order μ and type d.

Lemma B.38 is easily established.

Now we turn to the composition in Assertion (i).

Lemma B.39. *For each $h \in H_0 = H^+ \oplus H_0^-$ and all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ with $\sigma_1 \leq 0$, we have*

$$\chi_+ \mathrm{Op}_{x_n} h(\xi_n / \langle \xi' \rangle) \in S^0\big(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^\sigma(\overline{\mathbf{R}_+})\big).$$

Proof. The crucial point here is the boundedness of the operator

$$\chi_+ \mathrm{Op}_{x_n}(h)(\xi_n) : H_0^{-N}(\overline{\mathbf{R}_+}) \longrightarrow H_0^{-N}(\overline{\mathbf{R}_+}) \quad \text{for every } N \in \mathbf{N}.$$

We let

$$\lambda = \mathrm{Op}_{x_n}(1 + i\xi_n)^N .$$

It is clear that the operator

$$\lambda : H_0^0(\overline{\mathbf{R}_+}) \longrightarrow H_0^{-N}(\overline{\mathbf{R}_+})$$

is an isomorphism with the inverse

$$\lambda^{-1} = \mathrm{Op}_{x_n}(1 + i\xi_n)^{-N} : H_0^{-N}(\overline{\mathbf{R}_+}) \longrightarrow H_0^0(\overline{\mathbf{R}_+}).$$

Thus it suffices to show that the operator

$$\lambda^{-1}\left[\chi_+ \mathrm{Op}_{x_n}(h)\right] \lambda : C_0^\infty(\mathbf{R}_+) \longrightarrow H_0^0(\overline{\mathbf{R}_+})$$

extends to a bounded operator on the space $H_0^0(\overline{\mathbf{R}_+})$. The situation can be visualized in the following diagram:

$$\begin{array}{ccc}
H_0^{-N}(\overline{\mathbf{R}_+}) & \xrightarrow{\chi_+ \mathrm{Op}_{x_n}(h)} & H_0^{-N}(\overline{\mathbf{R}_+}) \\
\lambda \Big\uparrow & & \Big\downarrow \lambda^{-1} \\
H_0^0(\overline{\mathbf{R}_+}) & \xrightarrow[\lambda^{-1}[\chi_+ \mathrm{Op}_{x_n}(h)]\lambda]{} & H_0^0(\overline{\mathbf{R}_+})
\end{array}$$

Now we have

$$\left[\mathrm{Op}_{x_n}(h)\right]\lambda = \mathrm{Op}_{x_n}(h(1+i\xi_n)^N) = \mathrm{Op}_{x_n}(h_0 + p_0)$$

for some function $h_0 \in H_0$ and a polynomial p_0 of degree $\leq N$. Therefore, we easily find that the operator

$$\lambda^{-1}\left[\chi_+\mathrm{Op}_{x_n}(h)\right]\lambda = \lambda^{-1}\chi_+\mathrm{Op}_{x_n}(h_0 + p_0)$$

has a bounded extension to the space $H_0^0(\overline{\mathbf{R}_+})$.

The proof of Lemma B.39 is complete.

B.5.1 Decomposing $\mathrm{Op}_{x_n}^+ p$

In this subsection we prove the decomposition result used for Assertions (iii) and (ix). Let $p \in S^\mu(\mathbf{R}^n \times \mathbf{R}^n)$ with $\mu \in \mathbf{Z}$. Fix $N \in \mathbf{N}$ and a cut-off function $\omega \in C_0^\infty(\mathbf{R})$ with $\omega(t) = 1$ near $t = 0$. We write

$$\begin{aligned} &p(x,\xi) \\ &= \sum_{j=0}^{N-1} \frac{x_n^j}{j!}\omega(x_n)\partial_{x_n}^j p(x',0,\xi) + x_n^N\omega(x_n)p_N(x,\xi) + (1-\omega(x_n))p(x,\xi), \end{aligned}$$

with the Taylor remainder $p_N(x,\xi)$. As a consequence of the transmission property, we obtain that $\partial_{x_n}^j p(x',0,\langle\xi'\rangle\,\xi_n)$ is an element of the symbol class

$$S^\mu\left(\mathbf{R}_{x'}^{n-1} \times \mathbf{R}_{\xi'}^{n-1}\right)\hat{\otimes}_\pi H_{\xi_n}.$$

Since we have

$$H = H' \oplus H_0,$$

it follows that

$$\partial_{x_n}^j p(x',0,\xi) = \sum_{k=0}^{\mu} s_{jk}(x',\xi')\xi_n^k + q_j(x',\xi),$$

where

$$s_{jk} \in S_{1,0}^{\mu-k}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}),$$

and the symbol $q_j \in S^{\mu}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^n)$ has a representation

$$q_j(x', \xi', \xi_n) = \sum_{k=0}^{\infty} \lambda_{jk} c_{jk}(x,' \xi') h_{jk}(\xi_n / \langle \xi' \rangle) \tag{B.21}$$

with (see Schaefer [Sa, Chapter III, Section 6.4, Theorem], Treves [Tv, Theorem 45.2])

$$\sum_{k=0}^{\infty} |\lambda_{jk}| < \infty,$$

$$h_{jk} \in H_0 = H^+ \oplus H_0^- = \mathcal{F}(\mathcal{S}(\mathbf{R}_+)) \oplus \mathcal{F}(\mathcal{S}(\mathbf{R}_-)).$$

Here the sequences $\{c_{jk}\}_{k=1}^{\infty}$ converge to zero in the symbol class $S^{\mu}_{1,0}(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1})$ and the sequences $\{h_{jk}\}_{k=1}^{\infty}$ converge to zero in the space H_0, respectively.

Therefore, we obtain that

$$\mathrm{Op}_{x_n}(p) = \sum_{k=0}^{\mu} \left(\sum_{j=0}^{N-1} \frac{x_n^j}{j!} \omega(x_n) s_{jk}(x', \xi') \right) \partial_{x_n}^k + a$$

where

$$\begin{aligned} & a(x, \xi) \\ &= \mathrm{Op}_{x_n} \left(\sum_{j=0}^{N-1} \frac{x_n^j}{j!} \omega(x_n) q_j + x_n^N \omega(x_n) p_N(x, \xi) + (1 - \omega(x_n)) p(x, \xi) \right). \end{aligned} \tag{B.22}$$

It is clear that $a(x', x_n, \xi', \xi_n)$ is an element of the symbol class

$$S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^{\sigma}(\mathbf{R}), H^{\sigma-(\mu_+,0)}(\mathbf{R})\right)$$

for all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ and $\mu_+ = \max\{\mu, 0\}$. Moreover, we obtain that

$$\chi_+ a \in S^{\mu}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^{\sigma}(\overline{\mathbf{R}_+}), H_0^{\sigma-(\mu_+,0)}(\overline{\mathbf{R}_+})\right)$$

provided that $-N \leq \sigma_1 \leq 0$. Indeed, this is rather straightforward for the last two terms on the right-hand side of formula (B.22); for the terms under the summation we employ the representation formula (B.21) which allows us to focus on a single term. Then we apply Lemma B.39.

As explained above, this decomposition of $\mathrm{Op}_{x_n}(p)$ furnishes the statement regarding compositions (iii) and (ix).

B.5.2 The Analysis of the Leftover Term

The leftover term $\ell(p, q)$ arises from the composition of the boundary symbols associated with two pseudo-differential operators $p(x, \xi)$ and $q(x, \xi)$ of orders μ_1 and μ_2:

$$\begin{aligned}\ell(p, q) &= \left[\mathrm{Op}^+_{x_n} p\right]\left[\mathrm{Op}^+_{x_N} q\right] - r^+\left[\mathrm{Op}_{x_n}(p)\right]\left[\mathrm{Op}_{x_n}(q)\right] e^+ \\ &= r^+\left[\mathrm{Op}_{x_n}(p)\right]\left(e^+ r^+ - 1\right)\left[\mathrm{Op}_{x_n}(q)\right] e^+.\end{aligned}$$

In this subsection we show that the leftover term $\ell(p, q)$ is a singular Green symbol of order $\mu_1 + \mu_2$ and type $\mu_{2+} = \max\{\mu_2, 0\}$.

As in Sect. B.5.1, we can find symbols

$$s_j(x, \xi') \in S^{\mu_1 - j}_{1,0}\left(\mathbf{R}^n \times \mathbf{R}^{n-1}\right), \quad \tilde{s}_k(x, \xi') \in S^{\mu_2 - k}_{1,0}\left(\mathbf{R}^n \times \mathbf{R}^{n-1}\right)$$

such that

$$\mathrm{Op}_{x_n}(p) = \sum_{j=0}^{\mu_1} s_j\, \partial^j_{x_n} + a,$$

$$\mathrm{Op}_{x_n}(q) = \sum_{k=0}^{\mu_2} \tilde{s}_k\, \partial^k_{x_n} + b$$

(with the corresponding notation). In view of the identity

$$\partial_{x_n}\left(e^+ f\right) = e^+\left(\partial_{x_n} f\right) + f(0)\,\delta_0(x_n) \quad \text{for all } f \in \mathcal{S}(\mathbf{R}_+), \tag{B.23}$$

we conclude that

$$\ell\left(\sum_{j=0}^{\mu_1} s_j\, \xi_n^j, q\right) = 0. \tag{B.24}$$

Therefore, we have

$$\begin{aligned}\ell(p, q) &= r^+\left[\mathrm{Op}_{x_n}(p)\right]\left(e^+ r^+ - 1\right)\left[\mathrm{Op}_{x_n}(q)\right] e^+ \\ &= r^+ a\left(e^+ r^+ - 1\right)\left(\sum_{k=0}^{\mu_2} \tilde{s}_k\, \partial^k_{x_n} e^+ + b e^+\right).\end{aligned}$$

However, by iterating identity (B.23) we see that

$$\partial^k_{x_n}\left(e^+ f\right) = e^+\left(\partial^k_{x_n} f\right) + \sum_{\ell=0}^{k-1} f^{(\ell)}(0)\delta_0^{k-\ell-1}(x_n).$$

Hence it follows that

$$\left(e^+r^+ - 1\right)\sum_{k=0}^{\mu_2} \tilde{s}_k\, \partial_{x_n}^k (e^+ f) = e^+r^+ \left(\sum_{k=0}^{\mu_2}\sum_{\ell=0}^{k-1} \tilde{s}_k\, f^{(\ell)}(0)\delta_0^{k-\ell-1}\right) - \sum_{k=0}^{\mu_2}\sum_{\ell=0}^{k-1} \tilde{s}_k\, f^{(\ell)}(0)\delta_0^{k-\ell-1} = -\sum_{k=0}^{\mu_2}\sum_{\ell=0}^{k-1} \tilde{s}_k(x', 0, \xi')\gamma_\ell(f)\delta_0^{k-\ell-1}.$$

We know from Example B.35 that γ_ℓ is a trace symbol of order $1/2 + \ell$ and type $\ell + 1$:

$$\gamma_\ell \in S^{\ell+1/2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathcal{S}'(\mathbf{R}_+), \mathbf{C}\right).$$

So we can write

$$\begin{aligned} &\ell(p,q) \\ &= -r^+a\left(\sum_{k=0}^{\mu_2}\sum_{\ell=0}^{k-1} \tilde{s}_k(x', 0, \xi')\gamma_\ell(f)\delta_0^{k-\ell-1}\right) + r^+a\left(e^+r^+ - 1\right) be^+ \\ &= \sum_{\ell=0}^{\mu_2-1} k_\ell\gamma_\ell + r^+a\left(e^+r^+ - 1\right) be^+, \end{aligned} \tag{B.25}$$

where

$$k_\ell = -\sum_{k=\ell+1}^{\mu_2} r^+a\tilde{s}_k(x', 0, \xi')\delta_0^{k-\ell-1}.$$

By using (B.24), we may replace a by p. Since we have

$$\tilde{s}_k(x', 0, \xi') \in S_{1,0}^{\mu_2-k}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}\right),$$

it follows from Lemma B.25 with $\ell := k - \ell - 1$ and $\mu := \mu_1 + \mu_2 - k$ that k_ℓ is a potential symbol of order $\mu_1 + \mu_2 - \ell - 1/2$

$$k_\ell \in S^{\mu_1+\mu_2-\ell-1/2}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; \mathbf{C}, \mathcal{S}(\mathbf{R}_+)\right),$$

so that $k_\ell\gamma_\ell$ is a singular Green symbol of order $\mu_1 + \mu_2$ and type μ_{2+} as asserted.

Now we consider the second summand in (B.25). On the space $C_0^\infty(\mathbf{R}_+)$, we have

$$r^+a\left(e^+r^+ - 1\right)be^+ = -r^+ae^-r^-be^+ = -\left(r^+ae^-J\right)\left(Jr^-be^+\right),$$

where r^- is the restriction to $\mathbf{R}_-$, e^- is the extension from $\mathbf{R}_-$ to $\mathbf{R}$ by zero, and J is the reflection operator, respectively:

$$r^-u = u|_{\mathbf{R}_-},$$

$$e^-u(x_n) = \begin{cases} u(x_n) & \text{if } x_n \in \mathbf{R}_-, \\ 0 & \text{if } x_n \in \mathbf{R} \setminus \mathbf{R}_-, \end{cases}$$

$$(Ju)(x_n) = u(-x_n).$$

Note that, for every function $u \in C_0^\infty(\mathbf{R}_+) \subset \mathcal{S}(\mathbf{R})$, be^+u is a function in $\mathcal{S}(\mathbf{R})$, so that there are no problems with the compositions.

We shall show that r^+ae^-J and Jr^-be^+ are type zero singular Green operators of orders μ_1 and μ_2, respectively.

First, we analyze the operator r^+ae^-J. From (B.22), we know that

$$a = \mathrm{Op}_{x_n}\left(\sum_{j=0}^{N-1} \frac{x_n^j}{j!}\omega(x_n)q_j + x_n^N\omega(x_n)p_N + (1-\omega(x_n))p\right). \tag{B.26}$$

Here q_j is the projection onto the space $H_0 = H^+ \oplus H_0^-$ of $\partial_{x_n}^j p(x', 0, \xi', \cdot)$, and $x_n^N p_N$ is the Taylor remainder; the positive integer N as well as the cut-off function ω are fixed.

In the argument below, we shall only need the fact that $p_N \in S_{1,0}^{\mu_1}(\mathbf{R}^n \times \mathbf{R}^n)$, which is obvious from Taylor's formula. Since the function $(1-\omega(x_n))p$ vanishes to arbitrary order at $x_n = 0$, we can find a symbol $r \in S_{1,0}^{\mu_1}(\mathbf{R}^n \times \mathbf{R}^n)$ such that

$$x_n^N p_N(x,\xi) + (1-\omega(x_n))p(x,\xi) = x_n^N r(x,\xi).$$

Step 1: The operator of multiplication by $x_n^j\omega(x_n)$ is an element of the symbol class

$$S^{-j}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H^\sigma(\mathbf{R}_+), H^{\sigma-(0,j)}(\mathbf{R}_+)\right)$$

for all $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$: This follows from the identity

$$\kappa_{\langle\xi'\rangle^{-1}} x_n^j \omega(x_n) \kappa_{\langle\xi'\rangle} = \langle\xi'\rangle^{-j} x_n^j \omega\left(x_n/\langle\xi'\rangle\right),$$

together with the observation that the family

$$\left\{\omega(\cdot/\langle\xi'\rangle) : \xi' \in \mathbf{R}^{n-1}\right\}$$

is uniformly bounded on the weighted Sobolev space $H^\sigma(\mathbf{R}_+)$.

Step 2: $r^+[\mathrm{Op}_{x_n}(q_j)]e^-J \in S^{\mu_1}(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1};\mathcal{S}'(\mathbf{R}_+),\mathcal{S}(\mathbf{R}_+))$: We write the symbol $q_j(x',\xi')$ in the form (B.21). Then it suffices to show that

$$r^+\left[\mathrm{Op}_{x_n}h\left(\xi_n/\langle\xi'\rangle\right)\right]e^-J \in S^0\left(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1};\mathcal{S}'(\mathbf{R}_+),\mathcal{S}(\mathbf{R}_+)\right).$$

However, we note that

$$\kappa_{\langle\xi'\rangle^{-1}}\mathrm{Op}_{x_n}h\left(\xi_n/\langle\xi'\rangle\right)\kappa_{\langle\xi'\rangle} = \mathrm{Op}_{x_n}h(\xi_n).$$

Thus it suffices to prove the continuity of the operator

$$r^+\left[\mathrm{Op}_{x_n}(h)\right]e^-J : \mathcal{S}'(\mathbf{R}_+) \longrightarrow \mathcal{S}(\mathbf{R}_+),$$

since derivatives can be treated in the same way.

The operator $\mathrm{Op}_{x_n}(h)$ has the integral kernel

$$k(x_n,y_n) = \int_{\mathbf{R}} e^{i(x_n-y_n)\xi_n}h(\xi_n)\,d\xi_n = (\mathcal{F}^{-1}h)(x_n-y_n).$$

Hence the operator $r^+[\mathrm{Op}_{x_n}(h)]e^-J$ is given via the integral kernel

$$k(x_n,-y_n) = (\mathcal{F}^{-1}h)(x_n+y_n) \quad \text{on } \mathbf{R}^2_{++} = \mathbf{R}_+\times\mathbf{R}_+.$$

Since we have

$$h \in H_0 = H^+ \oplus H_0^- = \mathcal{F}\left(\mathcal{S}(\mathbf{R}_+)\right)\oplus\mathcal{F}\left(\mathcal{S}(\mathbf{R}_-)\right),$$
$$(\mathcal{F}^{-1}h)|_{\mathbf{R}_+} \in \mathcal{S}(\mathbf{R}_+),$$

we obtain the desired assertion.

Steps 1 and 2 imply that the terms under the summation in (B.26) are singular Green operators of order $\mu_1 - j$ and type zero.

Step 3: Fix $K \in \mathbf{N}$ and $r \in S^{\mu_1}_{1,0}(\mathbf{R}^n\times\mathbf{R}^n)$. Then it follows that the operator $r^+[\mathrm{Op}_{x_n}x_n^N r]e^-J$ defines an element of the symbol class

$$S^{-K}\left(\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}; H_0^{-(K,K)}(\overline{\mathbf{R}_+}), H^{(K,K)}(\mathbf{R}_+)\right)$$

provided that N is sufficiently large: We show that the norm of the operator

$$\begin{aligned}&\kappa_{\langle\xi'\rangle^{-1}} r^+ \left[\mathrm{Op}_{x_n}\left(x_n^N D_{\xi'}^\alpha D_{x'}^\beta r\right)\right] e^- J \kappa_{\langle\xi'\rangle} \\ &= \langle\xi'\rangle^{-N} r^+ \left[\mathrm{Op}_{x_n} x_n^N \left(D_{\xi'}^\alpha D_{x'}^\beta r\right)(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle)\right] e^- J\end{aligned} \tag{B.27}$$

in the space $\mathcal{L}(H_0^{-(K,K)}(\overline{\mathbf{R}_+}), H^{(K,K)}(\mathbf{R}_+))$ is of order $O(\langle\xi'\rangle^{-K-|\alpha|})$.

It is clearly no restriction to assume that $\alpha = \beta = 0$. Then the operator in (B.27) has the integral kernel

$$k(x_n, y_n) = \langle\xi'\rangle^{-N} \int_{\mathbf{R}} e^{i(x_n+y_n)\xi_n} x_n^N r\left(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle\right) d\xi_n$$

on $\mathbf{R}^2_{++} = \mathbf{R}_+ \times \mathbf{R}_+$. This makes sense as an oscillatory integral: We may choose $\ell > (\mu_1+1)/2$ and regularize it as follows:

$$\int_{\mathbf{R}} e^{i(x_n+y_n)\xi_n} x_n^N (x_n+y_n)^{-2\ell} \langle\xi'\rangle^{2\ell} \left(\Delta_{\xi_n}^\ell r\right)\left(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle\right) d\xi_n. \tag{B.28}$$

Then the integrand is of order

$$O\left(x_n^N (x_n+y_n)^{-2\ell} \langle\xi'\rangle^{2\ell} \langle(\xi', \xi_n\langle\xi'\rangle)\rangle^{\mu_1-2\ell}\right).$$

In view of the identity

$$\langle(\xi', \xi_n\langle\xi'\rangle)\rangle = \left(1 + |\xi'|^2 + \xi_n^2\langle\xi'\rangle^2\right)^{1/2} = \langle\xi'\rangle\langle\xi_n\rangle,$$

we obtain that

$$\begin{aligned}k(x_n, y_n) &= \langle\xi'\rangle^{-N} \int_{\mathbf{R}} e^{i(x_n+y_n)\xi_n} x_n^N r\left(x', x_n/\langle\xi'\rangle, \xi', \xi_n\langle\xi'\rangle\right) d\xi_n \\ &= O\left(x_n^N (x_n+y_n)^{-2\ell} \langle\xi'\rangle^{\mu_1-N}\right) \quad \text{for } \ell > \frac{\mu_1+1}{2}.\end{aligned}$$

For $x_n, y_n \leq 1$, we choose $\ell = N/2$ (N is large and may be assumed to be even) and conclude that

$$k(x_n, y_n) = O\left(\langle\xi'\rangle^{\mu_1-N}\right).$$

Otherwise, we let $2\ell > N+2$; then we have

$$k(x_n, y_n) = O\left(\langle x_n+y_n\rangle^{-2} \langle\xi'\rangle^{\mu_1-N}\right).$$

Consequently, we have the following two assertions:

$$\sup_{y_n \geq 0} \int_0^\infty |k(x_n, y_n)|\, dx_n = O\left(\langle \xi' \rangle^{\mu_1 - N}\right),$$
$$\sup_{x_n \geq 0} \int_0^\infty |k(x_n, y_n)|\, dy_n = O\left(\langle \xi' \rangle^{\mu_1 - N}\right).$$

By using Schur's lemma (Theorem 5.2) with $p := 2$, we obtain that the norm in the space $\mathcal{L}(L^2(\mathbf{R}_+), L^2(\mathbf{R}_+))$ of the operator family in formula (B.27) is of order $O(\langle \xi' \rangle^{\mu_1 - N})$.

Next, we recall a general fact: We can estimate the norm of an operator

$$T \in \mathcal{L}\left(H_0^{-(K,K)}(\overline{\mathbf{R}_+}), H^{(K,K)}(\mathbf{R}_+)\right)$$

in terms of the quantity

$$\left\{ \left\| x_n^m \partial_{x_n}^{m'} T x_n^l \partial_{x_n}^{l'} \right\|_{\mathcal{L}(L^2(\mathbf{R}_+), L^2(\mathbf{R}_+))} : m, m', l, l' \leq K \right\},$$

and these operators have integral kernels

$$(-1)^{l'} x_n^m \partial_{x_n}^{m'} y_n^l \partial_{x_n}^{l'} k(x_n, y_n)$$

if $k(x_n, y_n)$ is the integral kernel for T. We plug this into our regularized expression (B.28) for the integral kernel k. Then we have to ask $l > (\mu_1 + 2K + 1)/2$ to make the integral converge. For $N > \mu_1 + 2K + 1$, we can apply Schur's Lemma as before and we obtain the assertion.

Steps 1, 2, and 3 show that $r^+ a e^- J$ is a singular Green operator of order μ_1 and type zero.

Step 4: In virtually the same way we can treat $J r^- b e^+$, and prove that it is a singular Green symbol of order μ_2 and type zero. Altogether we then know that the leftover term

$$\begin{aligned} \ell(p, q) &= \sum_{l=0}^{\mu_2 - 1} k_l \gamma_l + r^+ a \left(e^+ r^+ - 1\right) b e^+ \\ &= \sum_{l=0}^{\mu_2 - 1} k_l \gamma_l - \left(r^+ a e^- J\right)\left(J r^- b e^+\right) \end{aligned}$$

is a singular Green symbol of order $\mu_1 + \mu_2$ and type $\mu_{2+} = \max\{\mu_2, 0\}$.

Remark B.40. If a_1 and a_2 are classical, then so is the composition $a_1 a_2$.

B.6 Operators on the Half-Space $\mathbf{R}^n_+$

Now we introduce the Boutet de Monvel algebra on the half-space $\mathbf{R}^n_+$.

B.6.1 Operators in the Boutet de Monvel Calculus

Given a boundary symbol a of order $\mu \in \mathbf{Z}$ and type $d \in \mathbf{N}_0$ as in (B.13), we call $A = \mathrm{Op}(a)$ an operator of order μ and type d in the Boutet de Monvel calculus and write

$$A \in \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+}).$$

Therefore, the operator A is a 2×2 matrix of operators

$$A = \begin{pmatrix} P_+ + G & K \\ T & S \end{pmatrix}. \tag{B.29}$$

Here:

(1) We call $P_+ = \mathrm{Op}^+(p) = \mathrm{Op}_{x'}\mathrm{Op}^+_{x_n} p$ the pseudo-differential part of A.
(2) The operator $G = \mathrm{Op}(g)$ is a so-called singular Green operator.
(3) The operator $T = \mathrm{Op}(t)$ a trace operator.
(4) The operator $K = \mathrm{Op}(k)$ a potential (or Poisson) operator.
(5) The operator $S = \mathrm{Op}(s)$ is the pseudo-differential part on the boundary.

The classical elements in the Boutet de Monvel calculus are the operators $\mathrm{Op}(a)$, where a is classical in the sense of Definition B.9. The notation is $\mathcal{B}^{\mu,d}_{\mathrm{cl}}(\overline{\mathbf{R}^n_+})$.

We call A a generalized singular Green operator if a is a generalized singular Green symbol as in Definition B.29.

The intersection

$$\bigcap_{\mu \in \mathbf{Z}} \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+})$$

is the space of regularizing operators of type d.

We endow these spaces with the topology inherited from the topology on the associated boundary symbols (see Sect. B.4.2).

Remark B.41. The sum in the upper left corner of formula (B.29) is not direct: For example, let P be a regularizing pseudo-differential operator which has an integral kernel in the Schwartz space $\mathcal{S}(\mathbf{R}^n \times \mathbf{R}^n)$. Then it is easy to see from Theorem B.32 that P_+ coincides with a regularizing singular Green operator.

On the other hand, Theorem B.48 shows that singular Green operators are regularizing when localized to the interior. If, in addition, we assume P to be classical, we conclude that P_+ can only coincide with a singular Green operator if it is regularizing.

By combining Assertion (B.14) with Corollary B.14, we obtain the following result for wedge Sobolev spaces:

Theorem B.42. *An operator $A \in \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+})$ induces a continuous mapping*

$$A : \begin{array}{c} \mathcal{W}^s(\mathbf{R}^{n-1}, H^\sigma(\mathbf{R}_+))^{M_1} \\ \oplus \\ \mathcal{W}^s(\mathbf{R}^{n-1}, \mathbf{C})^{M_2} \end{array} \longrightarrow \begin{array}{c} \mathcal{W}^{s-(m,0)}(\mathbf{R}^{n-1}, H^{\sigma-(m,0)}(\mathbf{R}_+))^{N_1} \\ \oplus \\ \mathcal{W}^{s-(m,0)}(\mathbf{R}^{n-1}, \mathbf{C})^{N_2} \end{array}$$

for every $s \in \mathbf{R}^2$ and for every $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$ with $\sigma_1 > d - 1/2$.

For unweighted Sobolev spaces, the following statement is an immediate consequence of Theorem B.42 and part (b) of Lemma B.12. The statement for weighted Sobolev spaces can be obtained by the commutator technique just as in the proof of Corollary B.14.

Theorem B.43. *Let $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$, and $A \in \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+})$. Then the operator*

$$A : \begin{array}{c} H^\sigma(\mathbf{R}^n_+)^{M_1} \\ \oplus \\ H^\sigma(\mathbf{R}^{n-1})^{M_2} \end{array} \longrightarrow \begin{array}{c} H^{\sigma-(m,0)}(\mathbf{R}^n_+)^{N_1} \\ \oplus \\ H^{\sigma-(m,0)}(\mathbf{R}^{n-1})^{N_2} \end{array}$$

is bounded for every $\sigma_1 > d - 1/2$.

Corollary B.44. *The operators in Theorem B.43 have continuous restrictions*

$$A : \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{M_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{M_2} \end{array} \longrightarrow \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{N_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{N_2}. \end{array}$$

Theorem B.45 (Compositions). *Consider two operators*

$$A_1 : \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{M_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{M_2} \end{array} \longrightarrow \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{N_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{N_2}. \end{array}$$

and

$$A_2 : \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{L_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{L_2} \end{array} \longrightarrow \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{M_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{M_2}. \end{array}$$

with $A_\ell \in \mathcal{B}^{\mu_\ell, d_\ell}(\overline{\mathbf{R}^n_+})$ *for* $\ell = 1, 2$. *Then we have the following three assertions (a), (b) and (c):*

(a) *The composition* $A_1 A_2$ *is an element of* $\mathcal{B}^{\mu_1+\mu_2, d}(\overline{\mathbf{R}^n_+})$ *with*

$$d = \max\{\mu_2 + d_1, d_2\}.$$

(b) *The composition is a regularizing operator whenever one of the factors is, and it is a generalized singular Green operator whenever this is the case for* A_1 *or* A_2. *In particular,* $\mathcal{B}^{0,0}(\overline{\mathbf{R}^n_+})$ *is an algebra, and* $\mathcal{B}^{-\infty,0}(\overline{\mathbf{R}^n_+})$ *as well as the generalized singular Green operators are ideals.*

(c) $A_1 A_2$ *is a classical operator if both* A_1 *and* A_2 *are classical.*

Proof. First, we write $A_\ell = \mathrm{Op}(a_\ell)$ with

$$a_\ell = \begin{pmatrix} \mathrm{Op}^+_{x_n} p_\ell + \sum_{j=0}^{d_\ell} g_{\ell j}\, \partial_+^j & k_\ell \\ \sum_{j=0}^{d_\ell} t_{\ell j}\, \partial_+^j & s_\ell \end{pmatrix}, \quad \ell = 1, 2.$$

Next, we choose left symbols for p_1, g_{1j}, t_{1j}, k_1, s_1, and choose right symbols for p_2, g_{2j}, t_{2j}, k_2, s_2 with respect to the x'-action. Then we have

$$\mathrm{Op}(b_L(y,\eta))\mathrm{Op}(b_R(\tilde{y},\eta)) = \mathrm{Op}(b_L(y,\eta) b_R(\tilde{y},\eta)).$$

Therefore, the desired assertions follow from Theorems B.11 and B.37.

The proof of Theorem B.45 is complete.

Theorem B.46 (Adjoints). *Let* $A \in \mathcal{B}^{\mu,0}(\overline{\mathbf{R}^n_+})$ *for* $\mu \in \mathbf{Z}$ *with* $\mu \leq 0$. *Then the formal adjoint* A^* *of the operator*

$$A : \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{M_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{M_2} \end{array} \longrightarrow \begin{array}{c} \mathcal{S}(\mathbf{R}^n_+)^{N_1} \\ \oplus \\ \mathcal{S}(\mathbf{R}^{n-1})^{N_2} \end{array}$$

with respect to the inner products in the Hilbert spaces

$$L^2(\mathbf{R}^n_+)^{M_1} \oplus L^2(\mathbf{R}^{n-1})^{M_2}$$

and

$$L^2(\mathbf{R}^n_+)^{N_1} \oplus L^2(\mathbf{R}^{n-1})^{N_2},$$

respectively, is an element of $\mathcal{B}^{\mu,0}(\overline{\mathbf{R}^n_+})$. *If* $A = \mathrm{Op}(a)$ *with a boundary symbol a as in* (B.13)*, then we have*

$$A^* = \begin{pmatrix} \mathrm{Op}^+_{x_n} p^* + g^* & k^* \\ t^* & s^* \end{pmatrix}.$$

Here

$$g^*(x', y', \xi') = g(y', x', \xi')^*, \quad t^*(x', y', \xi') = k(y', x', \xi')^*,$$
$$k^*(x', y', \xi') = t(y', x', \xi')^*, \quad s^*(x', y', \xi') = s(y', x', \xi')^*.$$

Note that the assertion is no longer true if d or μ are positive.

Proof. Let $(\mathrm{Op}(p))^* = \mathrm{Op}(p^*)$ be the formal adjoint of $\mathrm{Op}(p)$ with respect to the inner product of $L^2(\mathbf{R}^n)$. Then the formal adjoint $(\mathrm{Op}^+ p)^*$ of $\mathrm{Op}^+ p$ with respect to the inner product of $L^2(\mathbf{R}^n_+)$ is given by

$$(\mathrm{Op}^+ p)^* = \mathrm{Op}^+ p^*.$$

Indeed, since $\mu \leq 0$, it follows that $\mathrm{Op}^+ p^*$ is bounded on the space $L^2(\mathbf{R}_+)$, and we have

$$\begin{aligned}\left(r^+[\mathrm{Op}(p^*)]e^+u, v\right) &= \left([\mathrm{Op}(p^*)]e^+u, e^+v\right) = \left(e^+u, [\mathrm{Op}(p)]e^+v\right)\\ &= \left(u, r^+[\mathrm{Op}(p)]e^+v\right).\end{aligned}$$

For the symbol g, we first apply Theorem B.32 to see that the pointwise adjoint $g(x', y', \xi')^*$ defines a singular Green symbol. Next, we deduce from Sect. B.2.4 that the formal adjoint of $\mathrm{Op}(g)$ is $\mathrm{Op}(g^*)$ with $g^*(x', y', \xi') = g(y', x', \xi')^*$.

Similarly, we see from Theorem B.34 that the pointwise adjoint of a trace symbol of type zero is a potential symbol and vice versa; then we employ Sect. B.2.4.

For a pseudo-differential symbol s, the assertion follows from the standard pseudo-differential calculus or, alternatively, also from Sect. B.2.4.

The proof of Theorem B.46 is complete.

Theorem B.47 (Asymptotic expansions for compositions and adjoints). *Let $N \in \mathbf{N}$. By adopting the notation of Theorems B.45 and B.46, we can find two boundary symbols c_N and d_N such that*

$$A_1 A_2 - \sum_{|\alpha|=0}^{N-1} \frac{1}{\alpha!} \mathrm{Op}(\partial^\alpha_{\xi'} a_1(x', \xi') D^\alpha_{x'} a_2(x', \xi')) = \mathrm{Op}(c_N)$$

and that

$$A^* - \sum_{|\alpha|=0}^{N-1} \frac{1}{\alpha!} \mathrm{Op}(\partial^\alpha_{\xi'} D^\alpha_{x'} a(x', \xi')^*) = \mathrm{Op}(d_N).$$

The boundary symbol c_N is of order $\mu_1 + \mu_2 - N$ and type $\max\{\mu_2 + d_1, d_2\}$, while d_N is of order $\mu - N$ and type zero.

Indeed, Theorem B.47 follows from Theorem B.11.

Theorem B.48. *Let G be a singular Green operator of order μ and type d, T a trace operator of order μ and type d, and K a potential operator of order μ. Furthermore, we take a constant $\varepsilon > 0$ and a function $\varphi \in C_b^\infty(\overline{\mathbf{R}^n_+})$ which vanishes for all $x_n < \varepsilon$. Denote, just for the moment, by M_φ the operator of multiplication by φ. Then we have the following four assertions (a)–(d):*

(a) GM_φ is a regularizing singular Green operator of type zero.
(b) $M_\varphi G$ is a regularizing singular Green operator of type d.
(c) TM_φ is a regularizing trace operator of type zero.
(d) $M_\varphi K$ is a regularizing potential operator.

Proof. (a) Without loss of generality, we may assume that the symbols are scalar. For any $N \in \mathbf{N}$, we let

$$\varphi_N(x', x_n) = \frac{\varphi(x', x_n)}{x_n^N}, \quad (x', x_n) \in \mathbf{R}^n_+.$$

Then it follows that $\varphi_N \in C_b^\infty(\mathbf{R}^n_+)$ and that

$$GM_\varphi = GM_{\varphi_N} M_{x_n^N}.$$

Since $M_{x_n^N}$ is an element of the symbol class

$$S^{-N}\left(\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}; H_0^\sigma(\overline{\mathbf{R}_+}), H_0^{\sigma-(0,N)}(\overline{\mathbf{R}_+})\right)$$

for each $\sigma = (\sigma_1, \sigma_2) \in \mathbf{R}^2$, Assertion (a) follows easily.

The arguments for Assertions (b), (c) and (d) are analogous.

The proof of Theorem B.48 is complete.

Now we introduce ellipticity and parametrices. As usual, ellipticity is the property of the symbol (and the associated operator) that enables us to find an inverse up to an error of order -1, while a parametrix is an inverse modulo regularizing elements. Both notions are easily seen to be equivalent as a consequence of the symbolic calculus.

Finally, we shall state a simpler condition for ellipticity. Namely, the ellipticity of the interior pseudo-differential symbol together with the uniform invertibility of the boundary symbol with the pseudo-differential symbol p replaced by the x_n-independent symbol p_0, where $p_0(x', \xi) = p(x', 0, \xi)$.

Definition B.49. Let $A \in \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+})$ with $\mu \in \mathbf{Z}$, $d \leq \mu_+ = \max\{\mu, 0\}$, and $A = \mathrm{Op}(a)$.

(a) We say that A is *elliptic* if there exist two boundary symbols b_ℓ and b_r of order $-\mu$ and type $(-\mu)_+ = \max\{-\mu, 0\}$ such that the boundary symbols $b_\ell a - I$ and $ab_r - I$ have order -1. Also the boundary symbol a is said to be *elliptic*.

Since we have

$$b_\ell = b_\ell(I - ab_r) + (b_\ell a - I)b_r + b_r,$$

it follows that the boundary symbols b_ℓ and b_r differ by a symbol of order $-\mu - 1$ only, and we may choose $b_\ell = b_r$. We remark that b_ℓ and b_r are also elliptic. The composition rules imply that the type of $b_\ell a - I$ is μ_+, while that of $ab_r - I$ is $(-\mu)_+$.

(b) We say that the operator $B \in \mathcal{B}^{-\mu,(-\mu)_+}(\overline{\mathbf{R}^n_+})$ is a *parametrix* for A if both $BA - I$ and $AB - I$ are regularizing. The types of $BA - I$ and $AB - I$ are μ_+ and $(-\mu)_+$, respectively.

Theorem B.50. *Let $A \in \mathcal{B}^{\mu,d}(\overline{\mathbf{R}^n_+})$ with $d \leq \mu_+$. Then A has a parametrix if and only if it is elliptic.*

Proof. If $A = \mathrm{Op}(a)$ has a parametrix $B = \mathrm{Op}(b)$, then Theorem B.47 implies that both $ba - I$ and $ab - I$ are boundary symbols of order -1.

Conversely, assume that we find two boundary symbols b_ℓ and b_r such that $b_\ell a - I$ and $ab_r - I$ have order -1. If we let $B_\ell = \mathrm{Op}(b_\ell)$ and $B_r = \mathrm{Op}(b_r)$, then we can find a boundary symbol r_ℓ of order -1 and type μ_+ and a boundary symbol r_r of order -1 and type $(-\mu)_+$ such that

$$B_\ell A - I = \mathrm{Op}(r_\ell), \quad AB_r - I = \mathrm{Op}(r_r).$$

Indeed, this follows immediately from an application of Theorem B.47. Next we choose two boundary symbols $\tilde{b}_\ell$ and $\tilde{b}_r$ such that

$$\tilde{b}_\ell \sim \sum_{j=0}^{\infty} (-r_\ell)^{\sharp j} \sharp b_\ell,$$

$$\tilde{b}_r \sim b_r \sharp \sum_{j=0}^{\infty} (-r_r)^{\sharp j}.$$

In this notation, the '$\sharp$' indicates that we pick a boundary symbol for the corresponding composition; for example,

$$\mathrm{Op}\left((-r_\ell)^{\sharp j}\right) = [\mathrm{Op}(-r_\ell)]^j .$$

The type of $r_\ell^{\sharp j}$ is μ_+ for all j, while that of $r_r^{\sharp j}$ is $(-\mu)_+$.

Finally, we carry out the asymptotic summation modulo regularizing boundary symbols. If we let

$$\tilde{B}_\ell = \mathrm{Op}(\tilde{b}_\ell), \quad \tilde{B}_r = \mathrm{Op}(\tilde{b}_r),$$

then we find that

$$\tilde{B}_\ell A - I = R_\ell \in \mathcal{B}^{-\infty,\mu_+}\left(\overline{\mathbf{R}^n_+}\right),$$

$$A\tilde{B}_r - I = R_r \in \mathcal{B}^{-\infty,(-\mu)_+}\left(\overline{\mathbf{R}^n_+}\right).$$

Therefore, we obtain that

$$\begin{aligned}\tilde{B}_\ell &\equiv \tilde{B}_\ell \left(A\tilde{B}_r - R_r\right) = \left(\tilde{B}_\ell A\right)\tilde{B}_r - \tilde{B}_\ell R_r \\ &= (I + R_\ell)\,\tilde{B}_r - \tilde{B}_\ell R_r = \tilde{B}_r + R_\ell \tilde{B}_r - \tilde{B}_\ell R_r \\ &\equiv \tilde{B}_r \quad \text{modulo } \mathcal{B}^{-\infty,(-\mu)_+}(\overline{\mathbf{R}^n_+}).\end{aligned}$$

This implies that both $\tilde{B}_\ell$ and $\tilde{B}_r$ furnish a parametrix for A.

The proof of Theorem B.50 is complete.

With a little more work, we find the following simpler ellipticity criterion which we shall not prove here:

Theorem B.51. *Let*

$$a = \begin{pmatrix} \mathrm{Op}^+_{x_n} p + & gk \\ t & s \end{pmatrix}$$

be a boundary symbol of order μ and type $d \le \mu_+ = \max\{\mu, 0\}$. Let $p_0(x', \xi) = p(x', 0, \xi)$ and

$$a^0 = \begin{pmatrix} \mathrm{Op}^+_{x_n} p_0 + & gk \\ t & s \end{pmatrix}.$$

Then the boundary symbol a is elliptic if and only if the following two conditions (1) and (2) are satisfied:

(1) $p \in S^\mu_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$ is an elliptic, $N \times N$ matrix-valued symbol on $\overline{\mathbf{R}^n_+}$.
(2) $a^0(x', \xi') : H^{(\mu,0)}(\mathbf{R}_+)^N \oplus \mathbf{C}^{M'} \to L^2(\mathbf{R}_+)^N \oplus \mathbf{C}^M$ is an isomorphism for all x' and ξ' with $|\xi'| \ge 1$, satisfying the condition

$$\begin{pmatrix} \kappa_{\langle \xi' \rangle} & 0 \\ 0 & 1 \end{pmatrix}^{-1} a^0(x', \xi') \begin{pmatrix} \kappa_{\langle \xi' \rangle} & 0 \\ 0 & 1 \end{pmatrix} = O(\langle \xi' \rangle^{-\mu}).$$

By referring to p as "elliptic of order μ on $\overline{\mathbf{R}^n_+}$", we mean that there is an element $q \in S^{-\mu}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$ such that $pq - I$ and $qp - I$ coincide with a symbol in the symbol class $S^{-1}_{\mathrm{tr}}(\mathbf{R}^n \times \mathbf{R}^n)$ on $\overline{\mathbf{R}^n_+} \times \mathbf{R}^n$.

The key ingredient in the proof is the fact that, according to Theorem B.48, the non-pseudo-differential entries of a are regularizing outside a neighborhood of the boundary.

B.6.2 Outlook

It remains to establish the invariance of the symbol classes under suitable changes of coordinates in order to introduce the Boutet de Monvel calculus on manifolds with boundary. We shall omit this part, since there are no new aspects to be developed from the present point of view.

The results on compositions, adjoints, and mapping properties carry over to the case of operators acting on sections of vector bundles E_1, E_2 over a compact manifold X and F_1, F_2 over its boundary ∂X. In particular, if A is an operator of order μ and type d, then it follows that the operator

$$A : \begin{array}{c} H^s(X, E_1) \\ \oplus \\ H^s(\partial X, F_1) \end{array} \longrightarrow \begin{array}{c} H^{s-\mu}(X, E_2) \\ \oplus \\ H^{s-\mu}(\partial X, F_2) \end{array} \tag{B.30}$$

is bounded for every $s > -1/2$, so that we have the mapping property (B.1). In view of the well-known imbedding properties of Sobolev spaces on compact manifolds, we see immediately that ellipticity implies the Fredholm property of A in formula (B.30).

A few more features which might be of interest can be found in Schrohe [Sr4] which also deals with weighted symbols on certain non-compact manifolds. If $E_1 = E_2 = E$ and $F_1 = F_2 = F$, for example, the algebra $\mathcal{B}^{0,0}(X)$ of all operators of order and type zero is a Fréchet sub-algebra of the Banach algebra of all bounded operators on the Hilbert space $L^2(X, E) \oplus L^2(\partial X, F)$. It is closed under holomorphic functional calculus in several complex variables, hence a pre-C^*-algebra and moreover a Ψ^*-algebra in the sense of Gramsch [Gh].

By using the fact that there exist order-reducing operators in the calculus (Theorem B.24), we can show that the calculus is closed under inversion: If the operator A in (B.30) is bijective, then its inverse is again an element of the calculus.

Moreover, ellipticity is not only sufficient but also necessary for the Fredholm property of the operator

$$A : \begin{array}{c} H^\mu(X, E_1) \\ \oplus \\ H^\mu(\partial X, F_1) \end{array} \longrightarrow \begin{array}{c} H^0(X, E_2) \\ \oplus \\ H^0(\partial X, F_2). \end{array}$$

References

[AMR] Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, Tensor Analysis, and Applications. Global Analysis Pure and Applied: Series B, vol. 2. Addison-Wesley, Reading (1983)

[AF] Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics, vol. 140, 2nd edn. Elsevier/Academic, Amsterdam (2003)

[Ag] Agmon, S.: Lectures on Elliptic Boundary Value Problems. Van Nostrand Mathematical Studies, vol. 2. Van Nostrand, Princeton (1965)

[ADN] Agmon, S., Douglis, A., Nirenberg, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I. Commun. Pure Appl. Math. **12**, 623–727 (1959)

[AV] Agranovich, M.S., Vishik, M.I.: Elliptic problems with a parameter and parabolic problems of general type. Uspehi Mat. Nauk **19**(3)(117), 53–161 (1964, in Russian); English translation: Russ. Math. Surv. **19**(3), 53–157 (1964)

[Ap] Applebaum, D.: Lévy Processes and Stochastic Calculus. Cambridge Studies in Advanced Mathematics, vol. 116, 2nd edn. Cambridge University Press, Cambridge (2009)

[AS] Aronszajn, N., Smith, K.T.: Theory of Bessel potentials I. Ann. Inst. Fourier (Grenoble), **11**, 385–475 (1961)

[Ba1] Bass, R.F.: Probabilistic Techniques in Analysis. Probability and its Applications. Springer, New York (1995)

[Ba2] Bass, R.F.: Diffusions and Elliptic Operators. Probability and its Applications. Springer, New York (1998)

[BL] Bergh, J., Löfström, J.: Interpolation Spaces: An Introduction. Grundlehren der Mathematischen Wissenschaften. Springer, Berlin/New York (1976)

[Be] Bernstein, S.N.: Equations différentielles stochastiques. In: Actualités Sci. et Ind. vol. 738, pp. 5–31. Conf. intern. Sci. Math. Univ. Genève. Hermann, Paris (1938)

[Bi] Bichteler, K.: Stochastic Integration with Jumps. Encyclopedia of Mathematics and its Applications, vol. 89. Cambridge University Press, Cambridge (2002)

[BH] Bliedtner, J., Hansen, W.: Potential Theory: An Analytic and Probabilistic Approach to Balayage. Universitext. Springer, Berlin/Heidelberg/New York/Tokyo (1986)

[BG] Blumenthal, R.M., Getoor, R.K.: Markov Processes and Potential Theory. Pure and Applied Mathematics, vol. 29. Academic, New York/London (1968)

[BCP] Bony, J.-M., Courrège, P., Priouret, P.: Semi-groupes de Feller sur une variété à bord compacte et problèmes aux limites intégro-différentiels du second ordre donnant lieu au principe du maximum. Ann. Inst. Fourier (Grenoble) **18**, 369–521 (1968)

[Bd] Bourdaud, G.: L^p-estimates for certain non-regular pseudo-differential operators. Commun. Partial Differ. Equ. **7**, 1023–1033 (1982)

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*, Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7,

[Bo] Boutet de Monvel, L.: Boundary problems for pseudo-differential operators. Acta Math. **126**, 11–51 (1971)

[Ca] Calderón, A.P.: Lebesgue spaces of differentiable functions and distributions. In: Proceedings of Symposia in Pure Mathematics, IV, pp. 33–49. American Mathematical Society, Providence, R.I. (1961)

[CZ] Calderón, A.P., Zygmund, A.: On the existence of certain singular integrals. Acta Math. **88**, 85–139 (1952)

[Cn] Cancelier, C.: Problèmes aux limites pseudo-différentiels donnant lieu au principe du maximum. Commun. Partial Differ. Equ. **11**, 1677–1726 (1986)

[CP] Chazarain, J., Piriou, A.: Introduction à la théorie des équations aux dérivées partielles linéaires. Gauthier-Villars, Paris (1981)

[CM] Coifman, R.R., Meyer, Y.: Au-delà des opérateurs pseudo-différentiels. Astérisque, vol. 57. Société Mathématique de France, Paris (1978)

[De] de Rham, G.: Variétés différentiables. Hermann, Paris (1955)

[Dg] Dugundji, J.: An extension of Tietze's theorem. Pac. J. Math. **1**, 353–367 (1951)

[Du] Duistermaat, J.J.: Fourier Integral Operators. Courant Institute of Mathematical Sciences, New York University, New York (1973)

[DH] Duistermaat, J.J., Hörmander, L.: Fourier integral operators II. Acta Math. **128**, 183–269 (1972)

[Dy1] Dynkin, E.B.: Foundations of the theory of Markov processes. Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow (1959) (in Russian); English translation: Pergamon Press, Oxford/London/New York/Paris (1960); German translation: Springer, Berlin/Göttingen/Heidelberg (1961); French translation: Dunod, Paris (1963)

[Dy2] Dynkin, E.B.: Markov Processes I, II. Grundlehren der Mathematischen Wissenschaften. Springer, Berlin/Göttingen/Heidelberg (1965)

[DY] Dynkin, E.B., Yushkevich, A.A.: Markov Processes, Theorems and Problems. Plenum Press, New York (1969)

[Ei] Einstein, A.: Investigations on the Theory of the Brownian Movement. Dover, New York (1956)

[EN] Engel, K.-J., Nagel, R.: One-Parameter Semigroups for Linear Evolution Equations. Graduate Texts in Mathematics, vol. 194. Springer, New York/Berlin/Heidelberg (2000)

[Es] Èskin, G.I.: Boundary value problems for elliptic pseudodifferential equations. Nauka, Moscow (1973, in Russian); English translation: American Mathematical Society, Providence (1981)

[EK] Ethier, S.N., Kurtz, T.G.: Markov Processes, Characterization and Convergence. Wiley, New York/Chichester/Brisbane/Toronto/Singapore (1986)

[FP] Fefferman, C., Phong, D.H.: Subelliptic eigenvalue problems. In: Conference on Harmonic Analysis (1981: Chicago), pp. 590–606. Wadsworth, Belmont (1983)

[Fe1] Feller, W.: The parabolic differential equations and the associated semigroups of transformations. Ann. Math. **55**, 468–519 (1952)

[Fe2] Feller, W.: On second order differential equations. Ann. Math. **61**, 90–105 (1955)

[Fo1] Folland, G.B.: Introduction to Partial Differential Equations, 2nd edn. Princeton University Press, Princeton (1995)

[Fo2] Folland, G.B.: Real Analysis: Modern Techniques and Their Applications, 2nd edn. Wiley, New York (1999)

[Fr1] Friedman, A.: Partial Differential Equations. Holt, Rinehart and Winston Inc., New York/Montreal/London (1969)

[Fr2] Friedman, A.: Foundations of Modern Analysis. Holt, Rinehart and Winston Inc., New York/Montreal/London (1970)

[FU] Fujiwara, D., Uchiyama, K.: On some dissipative boundary value problems for the Laplacian. J. Math. Soc. Jpn. **23**, 625–635 (1971)

[GB] Galakhov, E.I., Skubachevskiĭ, A.L.: On Feller semigroups generated by elliptic operators with integro-differential boundary conditions. J. Differ. Equ. **176**, 315–355 (2001)

[GM1] Garroni, M.G., Menaldi, J.-L.: Green Functions for Second Order Parabolic Integro-Differential Problems. Pitman Research Notes in Mathematics Series, vol. 275. Longman Scientific & Technical, Harlow (1992)
[GM2] Garroni, M.G., Menaldi, J.-L.: Second Order Elliptic Integro-Differential Problems. Chapman & Hall/CRC Research Notes in Mathematics, vol. 430. Chapman & Hall/CRC, Boca Raton (2002)
[GS] Gel'fand, I.M., Shilov, G.E.: Generalized Functions I. Properties and Operations. Academic, New York/London (1964)
[GT] Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Classics in Mathematics. Springer, Berlin (2001). Reprint of the 1998 edition
[GK] Gohberg, I.C., Kreĭn, M.G.: The basic propositions on defect numbers, root numbers and indices of linear operators. Uspehi Mat. Nauk **12**, 43–118 (1957, in Russian); English translation: Am. Math. Soc. Transl. **13**, 185–264 (1960)
[Gg] Goldstein, G.R.: Derivation and physical interpretation of general boundary conditions. Adv. Differ. Equ. **11**, 457–480 (2006)
[Gj] Goldstein, J.A.: Semigroups of Linear Operators and Applications. Oxford Mathematical Monographs. Clarendon Press/Oxford University Press, New York (1985)
[Gh] Gramsch, B.: Relative Inversion in der Störungstheorie von Operatoren und Ψ-Algebren. Math. Ann. **269**, 27–71 (1984)
[Gb] Grubb, G.: Functional Calculus for Pseudodifferential Boundary Value Problems. Progress in Mathematics, vol. 65, 2nd edn. Birkhäuser, Boston (1996)
[GH] Grubb, G., Hörmander, L.: The transmission property. Math. Scand. **67**, 273–289 (1990)
[Ha] Hansen, W.: Restricted mean value property for Balayage spaces with jumps. Potential Anal. **36**, 263–273 (2012)
[HP] Hille, E., Phillips, R.S.: Functional Analysis and Semi-groups. American Mathematical Society/Colloquium, Providence (1974). Third printing of the revised edition of 1957
[Hi] Hirschmann, T.: Functional analysis in cone and edge Sobolev spaces. Ann. Global Anal. Geom. **8**, 167–192 (1990)
[Hp] Hopf, E.: A remark on linear elliptic differential equations of second order. Proc. Am. Math. Soc. **3**, 791–793 (1952)
[Ho1] Hörmander, L.: Pseudo-differential operators and non-elliptic boundary problems. Ann. Math. **83**, 129–209 (1966)
[Ho2] Hörmander, L.: Pseudo-differential operators and hypoelliptic equations. In: Calderón, A.P. (ed.) Proceedings of Symposia in Pure Mathematics: Singular Integrals, vol. X, Chicago, pp. 138–183. American Mathematical Society, Providence (1967)
[Ho3] Hörmander, L.: Fourier integral operators I. Acta Math. **127**, 79–183 (1971)
[Ho4] Hörmander, L.: The Analysis of Linear Partial Differential Operators III. Pseudo-Differential Operators. Reprint of the 1994 edition, Grundlehren der Mathematischen Wissenschaften. Springer, Berlin/Heidelberg/New York/Tokyo (2007)
[IW] Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, 2nd edn. North-Holland/Kodansha, Amsterdam/Tokyo (1989)
[It] Itô, K.: Stochastic processes. Iwanami Shoten, Tokyo (1957, in Japanese)
[IM] Itô, K., McKean, H.P. Jr.: Diffusion Processes and Their Sample Paths. Grundlehren der Mathematischen Wissenschaften, Second printing. Springer, Berlin/New York (1974)
[Ja] Jacob, N.: Pseudo Differential Operators and Markov Processes. Fourier Analysis and Semigroups, vol. I. Imperial College Press, London (2001); Generators and Their Potential Theory, vol. II. Imperial College Press, London (2002); Markov Processes and Applications, vol. III. Imperial College Press, London (2005)
[Ka] Kannai, Y.: Hypoellipticity of certain degenerate elliptic boundary value problems. Trans. Am. Math. Soc. **217**, 311–328 (1976)
[Ki] Kinney, J.R.: Continuity properties of sample functions of Markov processes. Trans. Am. Math. Soc. **74**, 280–302 (1953)
[Kn] Knight, F.B.: Essentials of Brownian Motion and Diffusion. Mathematical Surveys, vol. 18. American Mathematical Society, Providence (1981)

[Kl] Kolmogorov, A.N.: Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. **104**, 415–458 (1931)

[Ko] Komatsu, T.: Markov processes associated with certain integro-differential operators. Osaka J. Math. **10**, 271–303 (1973)

[Kr] Kreĭn, S.G.: Linear Differential Equations in Banach Space. Nauka, Moscow (1967, in Russian); English translation: Translations of Mathematical Monographs, vol. 29. American Mathematical Society, Providence (1971); Japanese translation: Yoshioka Shoten, Kyoto (1972)

[Ku] Kumano-go, H.: Pseudodifferential Operators. MIT, Cambridge (1981)

[La] Lamperti, J.: Stochastic Processes, A Survey of the Mathematical Theory. Applied Mathematical Sciences, vol. 23. Springer, New York/Heidelberg (1977)

[Lg] Lang, S.: Differential Manifolds. Addison-Wesley, Reading (1972)

[Le] Lévy, P.: Processus Stochastiques et Mouvement Brownien. Suivi d'une note de M. Loéve. (French) Gauthier-Villars, Paris, 365 (1948)

[LM] Lions, J.-L., Magenes, E.: Problèmes aux Limites Non-homogènes et Applications 1, 2. Dunod, Paris (1968); English translation: Non-homogeneous Boundary Value Problems and Applications 1, 2. Springer, Berlin/Heidelberg/New York (1972)

[MZ] Malý, J., Ziemer, W.P.: Fine Regularity of Solutions of Elliptic Partial Differential Equations. Mathematical Surveys and Monographs, vol. 51. American Mathematical Society, Providence (1997)

[MH] Marsden, J.E., Hughes, T.J.R.: Mathematical Foundations of Elasticity. Prentice-Hall, Englewood Cliffs (1983)

[Ma] Masuda, K.: Evolution Equations (in Japanese). Kinokuniya Shoten, Tokyo (1975)

[Mc] McLean, W.: Strongly Elliptic Systems and Boundary Integral Equations. Cambridge University Press, Cambridge (2000)

[Me] Meyer, Y.: Remarques sur un théorème de J.-M. Bony. Suppl. Rend. Circ. Mat. Palermo **II**-1, 1–20 (1981)

[Mu] Munkres, J.R.: Elementary Differential Topology. Annals of Mathematics Studies, vol. 54. Princeton University Press, Princeton (1966)

[Na] Nagase, M.: The L^p-boundedness of pseudo-differential operators with non-regular symbols. Commun. Partial Differ. Equ. **2**, 1045–1061 (1977)

[Ni] Nishio, M.: Probability Theory (in Japanese). Jikkyo Shuppan, Tokyo (1978)

[Ol] Oleĭnik, O.A.: On properties of solutions of certain boundary problems for equations of elliptic type. Mat. Sbornik **30**, 595–702 (1952, in Russian)

[OR] Oleĭnik, O.A., Radkevič, E.V.: Second Order Equations with Nonnegative Characteristic Form. Itogi Nauki, Moscow (1971, in Russian); English translation: American Mathematical Society/Plenum Press, Providence, New York/London (1973)

[Pl] Palais, R.S.: Seminar on the Atiyah–Singer Index Theorem. Annals of Mathematics Studies, vol. 57. Princeton University Press, Princeton (1965)

[Pa] Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences, vol. 44. Springer, New York (1983)

[Pe] Peetre, J.: Réctification à l'article "Une caractérisation abstraite des opérateurs différentiels". Math. Scand. **8**, 116–120 (1960)

[Pr] Perrin, J.: Les Atomes. Gallimard, Paris (1970)

[Po] Polking, J.C.: Boundary value problems for parabolic systems of partial differential equations. In: Calderón, A.P. (ed.) Proceedings of Symposia in Pure Mathematics: Singular Integrals, Chicago, vol. X, pp. 243–274. American Mathematical Society, Providence (1967)

[PW] Protter, M.H., Weinberger, H.F.: Maximum Principles in Differential Equations. Prentice-Hall, Englewood Cliffs (1967)

[Ra] Ray, D.: Stationary Markov processes with continuous paths. Trans. Am. Math. Soc. **82**, 452–493 (1956)

[RS] Rempel, S., Schulze, B.-W.: Index Theory of Elliptic Boundary Problems. Akademie-Verlag, Berlin (1982)

[RY] Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Grundlehren der Mathematischen Wissenschaften, 3rd edn. Springer, Berlin (1999)

[Ru] Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1987)

[RF] Runst, T., Youssfi, A.: Boundary value problems for Waldenfels operators. Indiana Univ. Math. J. **54**, 237–255 (2005)

[SU] Sato, K., Ueno, T.: Multi-dimensional diffusion and the Markov process on the boundary. J. Math. Kyoto Univ. **14**, 529–605 (1964, 1965)

[Sa] Schaefer, H.H.: Topological Vector Spaces. Graduate Texts in Mathematics, Third printing, vol. 3. Springer, New York/Berlin (1971)

[Sh] Schechter, M.: Principles of Functional Analysis. Graduate Studies in Mathematics, vol. 36, 2nd edn. American Mathematical Society, Providence (2002)

[Sr1] Schrohe, E.: Boundedness and spectral invariance for standard pseudodifferential operators on anisotropically weighted L^p Sobolev spaces. Integral Equ. Oper. Theory **13**, 271–284 (1990)

[Sr2] Schrohe, E.: A characterization of the singular Green operators in Boutet de Monvel's calculus via wedge Sobolev spaces. Commun. Partial Differ. Equ. **19**, 677–699 (1994)

[Sr3] Schrohe, E.: A characterization of the uniform transmission property for pseudodifferential operators. In: Demuth, M., et al. (eds.) Pseudodifferential Calculus and Mathematical Physics. Advances in Partial Differential Equations, vol. 1, pp. 210–234. Akademie Verlag, Berlin (1994)

[Sr4] Schrohe, E.: Fréchet algebra techniques for boundary value problems on noncompact manifolds: Fredholm criteria and functional calculus via spectral invariance. Math. Nachr. **199**, 145–185 (1999)

[Sr5] Schrohe, E.: A short introduction to Boutet de Monvel's calculus. In: Gil, J., Grieser, D., Lesch, M. (eds.) Approaches to Singular Analysis. Operator Theory, Advances and Applications, vol. 125, pp. 85–116. Birkhäuser, Basel (2001)

[SS1] Schrohe, E., Schulze, B.-W.: Boundary value problems in Boutet de Monvel's calculus for manifolds with conical singularities I. In: Michael Demuth, Elmar Schrohe and Bert-Wolfgang Schulze (eds.) Pseudo-Differential Calculus and Mathematical Physics. Mathematical Topics, vol. 5, pp. 97–209. Akademie Verlag, Berlin (1994)

[SS2] Schrohe, E., Schulze, B.-W.: Boundary value problems in Boutet de Monvel's calculus for manifolds with conical singularities II. In: Michael Demuth, Elmar Schrohe and Bert-Wolfgang Schulze (eds.) Boundary Value Problems, Schrödinger Operators, Deformation Quantization. Mathematical Topics, vol. 8, pp. 70–205. Akademie Verlag, Berlin (1995)

[SS3] Schrohe, E., Schulze, B.-W.: Mellin operators in a pseudodifferential calculus for boundary value problems on manifolds with edges. In: I. Gohberg, R. Mennicken and C. Tretter (eds.) Differential and Integral Operators (Regensburg, 1995). Operator Theory: Advances and Applications, vol. 102, pp. 255–285. Birkhäuser, Basel (1998)

[SS4] Schrohe, E., Schulze, B.-W.: A symbol algebra for pseudodifferential boundary value problems for manifolds with edges. In: Michael Demuth and Bert-Wolfgang Schulze (eds.) Differential Equations, Asymptotic Analysis, and Mathematical Physics (Potsdam, 1996). Mathematical Research, vol. 100, pp. 292–324. Akademie Verlag, Berlin (1997)

[SS5] Schrohe, E., Schulze, B.-W.: Mellin and Green symbols for boundary value problems on manifolds with edges. Integral Equ. Oper. Theory **34**, 339–363 (1999)

[Su1] Schulze, B.-W.: Pseudo-Differential Operators on Manifolds with Singularities. Studies in Mathematics and its Applications, vol. 24. North-Holland, Amsterdam (1991)

[Su2] Schulze, B.-W.: Boundary Value Problems and Singular Pseudo-Differential Operators. Pure and Applied Mathematics (New York). Wiley, Chichester (1998)

[Sz] Schwartz, L.: Théorie des distributions. Nouvelle édition, Hermann, Paris (1966)

[Se1] Seeley, R.T.: Refinement of the functional calculus of Calderón and Zygmund. Nederl. Akad. Wetensch. Proc. Ser. A **68**, 521–531 (1965)

[Se2] Seeley, R.T.: Singular integrals and boundary value problems. Am. J. Math. **88**, 781–809 (1966)

[Si] Seiler, J.: Continuity of edge and corner pseudo-differential operators. Math. Nachr. **205**, 163–182 (1999)
[Sk] Skubachevskii, A.L.: Elliptic Functional-Differential Equations and Applications, Operator Theory: Advances and Applications, vol. 91. Birkhäuser Verlag, Basel (1997)
[Sn1] Stein, E.M.: The characterization of functions arising as potentials II. Bull. Am. Math. Soc. **68**, 577–582 (1962)
[Sn2] Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton Mathematical Series, vol. 30. Princeton University Press, Princeton (1970)
[Sn3] Stein, E.M.: The differentiability of functions in $\mathbf{R}^n$. Ann. Math. **113**, 383–385 (1981)
[Sn4] Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
[St] Stroock, D.W.: Diffusion processes associated with Lévy generators. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 209–244 (1975)
[SV] Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Grundlehren der Mathematischen Wissenschaften. Springer, Berlin/New York (1979)
[Tb] Taibleson, M.H.: On the theory of Lipschitz spaces of distributions on Euclidean n-space I. Principal properties. J. Math. Mech. **13**, 407–479 (1964)
[Ta1] Taira, K.: On some degenerate oblique derivative problems. J. Fac. Sci. Univ. Tokyo Sect. IA **23**, 259–287 (1976)
[Ta2] Taira, K.: Sur le problème de la dérivée oblique II. Ark. Mat. **17**, 177–191 (1979)
[Ta3] Taira, K.: Sur l'existence de processus de diffusion. Ann. Inst. Fourier (Grenoble) **29**, 99–126 (1979)
[Ta4] Taira, K.: Un théorème d'existence et d'unicité des solutions pour des problèmes aux limites non-elliptiques. J. Funct. Anal. **43**, 166–192 (1981)
[Ta5] Taira, K.: Diffusion Processes and Partial Differential Equations. Academic, Boston (1988)
[Ta6] Taira, K.: On the existence of Feller semigroups with boundary conditions. Mem. Am. Math. Soc. **99**(475) (1992). American Mathematical Society, Providence, vii+65
[Ta7] Taira, K.: Analytic Semigroups and Semilinear Initial-Boundary Value Problems. London Mathematical Society Lecture Note Series, vol. 223. Cambridge University Press, Cambridge (1995)
[Ta8] Taira, K.: Boundary value problems for elliptic integro-differential operators. Math. Z. **222**, 305–327 (1996)
[Ta9] Taira, K.: Boundary Value Problems and Markov Processes. Lecture Notes in Mathematics, vol. 1499, 2nd edn. Springer, Berlin (2009)
[Ta10] Taira, K.: On the existence of Feller semigroups with discontinuous coefficients. Acta Math. Sin. (English Series), **22**, 595–606 (2006)
[Ta11] Taira, K.: On the existence of Feller semigroups with discontinuous coefficients II. Acta Math. Sin. (English Series), **25**, 715–740 (2009)
[Ta12] Taira, K.: A mixed problem of linear elastodynamics. J. Evol. Equ. **13**, 481–507 (2013)
[TW] Takanobu, S., Watanabe, S.: On the existence and uniqueness of diffusion processes with Wentzell's boundary conditions. J. Math. Kyoto Univ. **28**, 71–80 (1988)
[Tn] Tanabe, H.: Equations of Evolution. Iwanami Shoten, Tokyo (1975, in Japanese); English translation: Monographs and Studies in Mathematics, vol. 6. Pitman, Boston/London (1979)
[Ty] Taylor, M.: Pseudodifferential Operators. Princeton Mathematical Series, vol. 34. Princeton University Press, Princeton (1981)
[Tv] Treves, F.: Topological Vector Spaces, Distributions and Kernels. Academic, New York/London (1967)
[Tr] Triebel, H.: Theory of Function Spaces. Monographs in Mathematics, vol. 78. Birkhäuser Verlag, Basel (1983)
[VE] Višik, M.I., Èskin, G.I.: Normally solvable problems for elliptic systems of convolution equations. Mat. Sb. (N.S.) **74**(116), 326–356 (1967, in Russian)

[Wa] von Waldenfels, W.: Positive Halbgruppen auf einem n-dimensionalen Torus. Archiv der Math. **15**, 191–203 (1964)

[Wb] Watanabe, S.: Construction of diffusion processes with Wentzell's boundary conditions by means of Poisson point processes of Brownian excursions. In: Probability Theory, Banach Center Publications, vol. 5, pp. 255–271. PWN, Warsaw (1979)

[Wt] Watson, G.N.: A Treatise on the Theory of Bessel Functions. Cambridge Mathematical Library. Cambridge University Press, Cambridge (1995). Reprint of the second edition (1944)

[We] Wentzell (Ventcel'), A.D.: On boundary conditions for multidimensional diffusion processes. Teoriya Veroyat. i ee Primen. **4**, 172–185 (1959, in Russian); English translation: Theory Prob. Appl. **4**, 164–177 (1959)

[Wi] Wiener, N.: Differential space. J. Math. Phys. **2**, 131–174 (1923)

[Yo] Yosida, Y.: Functional Analysis. Classics in Mathematics. Springer, Berlin (1995). Reprint of the sixth edition (1980)

Index

K. Taira, *Semigroups, Boundary Value Problems and Markov Processes*, Springer Monographs in Mathematics, DOI 10.1007/978-3-662-43696-7,