Springer Monographs in Mathematics

Kazuaki Taira

Functional Analytic Techniques for Diffusion Processes

Springer

Springer Monographs in Mathematics

This series publishes advanced monographs giving well-written presentations of the "state-of-the-art" in fields of mathematical research that have acquired the maturity needed for such a treatment. They are sufficiently self-contained to be accessible to more than just the intimate specialists of the subject, and sufficiently comprehensive to remain valuable references for many years. Besides the current state of knowledge in its field, an SMM volume should ideally describe its relevance to and interaction with neighbouring fields of mathematics, and give pointers to future directions of research.

More information about this series at https://link.springer.com/bookseries/3733

Kazuaki Taira

Functional Analytic Techniques for Diffusion Processes

Springer

Kazuaki Taira
Tsuchiura, Japan

ISSN 1439-7382 ISSN 2196-9922 (electronic)
Springer Monographs in Mathematics
ISBN 978-981-19-1098-2 ISBN 978-981-19-1099-9 (eBook)
https://doi.org/10.1007/978-981-19-1099-9

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

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Dedicated to Prof. Kiyosi Itô (1915–2008) in appreciation of his constant encouragement

Foreword

There are dozens of books about Markov processes, some of them very good, but none match the depth and broad coverage of Kazuaki Taira's books. Let me try to put this into context.

Sometimes a massive study is done and leads to a major volume or volumes that redefine a field of study. For instance, the three-volume work of Nelson Dunford and Jack Schwartz did this for abstract mathematical analysis. The famed Charles Misner, Kip Thorne and John Wheeler book did this for general relativity.

Taira's work does this for Markov processes from a broad perspective. A simple view of Markov processes is that they deal with classes of dependent random variables that have both a nice theory and useful applications. But the general theory of Markov processes turns out to be extremely complicated. It is essential for applications to fields including mathematical biology, ecology, diffusion, statistical physics, etc. The mathematics needed for the hard parts of Markov processes require up-to-date versions of functional analysis, probability theory, partial and pseudo-differential equations, differential geometry, Fourier analysis, and more.

Taira's books bring these topics all together. They are not easy to explain in their general forms, but Taira does this carefully and quite nicely. These topics are usually hard to follow, but Taira explains things in a more easily readable way than one normally expects. The scope of his work is vast; it has been and continues to be a major influence in stochastic analysis and related fields.

This book is a revised and expanded edition of the previous book [191] published in 1988. But is a new edition needed? In June 2019, Taira and I were both at a meeting in Cesena, Italy. His lecture was wonderful; it was on new, deep results. The topics he covered are among the new results in his new edition. In particular, the new material on the theory of pseudo-differential operators widens the scope of the book (which has a huge scope to begin with). This is nicely explained in Chap. 1 (Introduction and Summary) and Chap. 13 (L^2 Approach to the Construction of Feller Semigroups) of this edition.

This wonderful book will be a major influence in a very broad field of study for a long time. I thank both Taira and Springer for their great contribution to the mathematical research community in publishing this book.

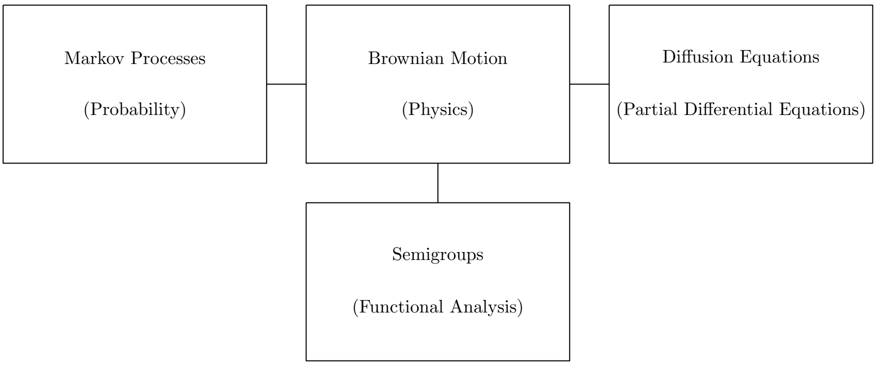
November 2021

Jerome Arthur Goldstein
University of Memphis
Memphis, Tennessee, USA

Preface

This book is devoted to the functional analytic approach to the problem of construction of diffusion processes in probability theory. It is well known that, by virtue of the Hille–Yosida theory of semigroups, the problem of construction of Markov processes can be reduced to the study of boundary value problems for degenerate elliptic integro-differential operators of second order. Several recent developments in the theory of partial differential equations have made possible further progress in the study of boundary value problems and hence of the problem of construction of Markov processes. The presentation of these new results is the main purpose of the present book. Unlike many other books on Markov processes, this book focuses on the relationship between Markov processes and elliptic boundary value problems with emphasis on the study of maximum principles. Our approach here is distinguished by the extensive use of the theory of partial differential equations.

Our functional analytic approach to diffusion processes is inspired by the following bird's-eye view of mathematical studies of Brownian motion (see Tables 1.1, 1.2 ans Figure 1.1 in Chap. 1):



This book grew out of lecture notes for graduate courses given by the author at Sophia University, Waseda University, Hokkaido University, Tôhoku University, Tokyo Metropolitan University, Tokyo Institute of Technology, Hiroshima University and University of Tsukuba. It is addressed to advanced undergraduates, graduate students and mathematicians with interest in probability, functional analysis and partial differential equations.

This book may be considered as the second edition of the book [191] published in 1988, which was found useful by a number of people, but it went out of print after several years. This augmented edition has been revised to streamline some of the analysis and to give better coverage of important examples and applications. I have endeavored to present it in such a way as to make it accessible to undergraduates as well. Moreover, in order to make the book more up-to-date, additional references have been included in the bibliography. This book is amply illustrated; 14 tables and 141 figures are provided.

The contents of the book are divided into five principal parts.

(1) The first part (Chaps. 2 through 6) provides the elements of the Lebesgue theory of measure and integration, probability theory, manifold theory, functional analysis and distribution theory which are used throughout the book. 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.

(2) In the second part (Chaps. 7–9), the basic definitions and results about Sobolev spaces are summarized and the calculus of pseudo-differential operators—a modern version of classical potentials—is developed. The theory of pseudo-differential operators forms a most convenient tool in the study of elliptic boundary value problems in Chap. 11. It should be emphasized that pseudo-differential operators provide a constructive tool to deal with existence and smoothness of solutions of partial differential equations. The full power of this very refined theory is yet to be exploited. Our approach is not far removed from the classical potential approach.

(3) Our subject proper starts with the third part (Chap. 10), where various maximum principles for degenerate elliptic differential operators of second order are studied. In particular, the underlying analytical mechanism of propagation of maxima is revealed here. This plays an important role in the interpretation and study of Markov processes in terms of partial differential equations in Chap. 12.

(4) The fourth part (Chap. 11) is devoted to general boundary value problems for second order elliptic differential operators. The basic questions of existence, uniqueness and regularity of solutions of general boundary value problems with a spectral parameter are studied in the framework of Sobolev spaces, using the calculus of pseudo-differential operators. A fundamental existence and uniqueness theorem is proved here. The importance of such a theorem is visible in constructing Markov processes in Chaps. 12 and 13.

(5) The fifth and final part (Chaps. 12 and 13) is devoted to the functional analytic approach to the problem of construction of Markov processes. This part is the heart of the subject. General existence theorems for Markov processes in terms of boundary value problems are proved in Chap. 12, and then the construction of Markov processes is carried out in Chap. 13, by solving general boundary value problems with a spectral parameter.

To make the material in Chaps. 10 through 13 accessible to a broad spectrum of readers, I have added an *Introduction and Summary* (Chap. 1). In this introductory chapter, I have included ten elementary (but important) examples of diffusion processes, and further I have attempted to state our problems and results in such a fashion that a broad spectrum of readers could understand, and also to describe how these problems can be solved, using the mathematics I present in Chaps. 2 through 9.

In the last Chap. 14, as concluding remarks, we give an overview on generation theorems for Feller semigroups proved by the author using the L^p theory of pseudo-differential operators and the Calderón–Zygmund theory of singular integral operators (Table 14.1).

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

In Appendix A, following Gilbarg–Trudinger [74], we present a brief introduction to the *potential theoretic approach* to the Dirichlet problem for Poisson's equation. The approach here can be traced back to the pioneering work of Schauder, [158] and [159], on the Dirichlet problem for second order elliptic differential operators. This appendix is included for the sake of completeness.

This book may be considered as an elementary introduction to the more advanced book *Boundary Value Problems and Markov Processes* (the third edition) which was published in the Lecture Notes in Mathematics series in 2020. In fact, we confined ourselves to the case when the differential operator A is elliptic on $\bar{D}$. The reason is that when A is not elliptic on $\bar{D}$ we do not know whether the operator $T(\alpha) = LP(\alpha)$, which plays a fundamental role in the proof, is a pseudo-differential operator or not. This book provides a powerful method for the analysis of elliptic boundary value problems in the framework of L^2 Sobolev spaces.

For advanced undergraduates working in functional analysis, partial differential equations and probability, this book may serve as an effective introduction to these three interrelated fields of analysis. For beginning graduate students about to major in the subject and mathematicians in the field looking for a coherent overview, I hope that the readers will find this book a useful entrée to the subject.

The presentation on some results of this book was given in "Mathematisch-Physikalisches Kolloquium" which was held on November 3rd, 2015 at Leibniz Universität Hannover (Germany) while I was on leave from Waseda University. I take this opportunity to express my sincere gratitude to these institutions.

In preparing this book, I am indebted to many friends, colleagues and students. It is my great pleasure to thank all of them. In particular, I would like to express my hearty thanks to Kenji Asada, Sunao Ōuchi, Bernard Helffer, Jacques Camus, Charles Rockland, Junjiro Noguchi, Yuji Kasahara, Masao Tanikawa, Yasushi Ishikawa, Elmar Schrohe, Seiichiro Wakabayashi, Silvia Romanelli and Angelo Favini. Kasahara, Tanikawa and Wakabayashi helped me to learn the material that was presented in the previous book [191]. Schrohe and Ishikawa have read and commented on portions of various preliminary drafts. I am deeply indebted to Professors Kôichi Uchiyama, Jean-Michel Bony, Minoru Motoo, Tadashi Ueno, Shinzo Watanabe, Francesco Altomare and Jerome Arthur Goldstein for their constant interest in my work. I am grateful to my students—especially Hideo Deguchi, Nobuyuki Sugino, Takayasu Ito and Yusuke Yoshida—for many comments and corrigenda concerning my original lecture notes.

Furthermore, I am very happy to acknowledge the influence of two of my teachers: Prof. Daisuke Fujiwara, from whose lectures I first learned this subject, and Prof. Hikosaburo Komatsu, who has done much to shape my viewpoint of analysis.

I would like to extend my warmest thanks to the late Prof. Richard Ernest Bellman (1920–1984) who originally suggested that my work be published in book form.

I am sincerely grateful to the four anonymous referees and a copyeditor for their many valuable suggestions and comments, which have substantially improved the presentation of this book. I would like to extend my hearty thanks to the staff of Springer-Verlag (Tokyo), who have generously complied with all my wishes.

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

Tsuchiura, Ibaraki, Japan Kazuaki Taira
November 2021

Contents

Notation and Conventions

The notation for set-theoretic concepts is standard. For example, the following notation is used for sets of numbers:

(1) $\mathbf{N}$: positive integers.
(2) $\mathbf{Z}$: integers.
(3) $\mathbf{Z}_0$: non-negative integers.
(4) $\mathbf{R}$: real numbers.
(5) $\mathbf{C}$: complex numbers.
(6) $[a, b]$: the closed interval $\{x \in \mathbf{R} : a \leq x \leq b\}$.
(7) $[a, b)$: the semiclosed interval $\{x \in \mathbf{R} : a \leq x < b\}$.
(8) (a, b): the open interval $\{x \in \mathbf{R} : a < x < b\}$.

The following notation and conventions are used for differentiation:

(1) $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_0^n$.
(2) $|\alpha| = \alpha_1 + \cdots + \alpha_n$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_0^n$.
(3) $\alpha! = \alpha_1! \ldots \alpha_n!$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_0^n$.
(4) $\alpha \geq \beta$ if and only if $\alpha_i \geq \beta_i$ for all $1 \leq i \leq n$.
(5) $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$.
(6) $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ for $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ and $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_0^n$.
(7) $\partial_x^\beta = \partial_{x_1}^{\beta_1} \ldots \partial_{x_n}^{\beta_n}$ for $\beta = (\beta_1, \ldots, \beta_n) \in \mathbf{Z}_0^n$.
(8) $D_{\xi_j} = -i \frac{\partial}{\partial_{\xi_j}}$ for $1 \leq j \leq n$, where $i = \sqrt{-1}$.
(9) $D_x^\beta = D_{x_1}^{\beta_1} \ldots D_{x_n}^{\beta_n}$ for $\beta = (\beta_1, \ldots, \beta_n) \in \mathbf{Z}_0^n$.
(10) $\langle y \rangle = \sqrt{1 + |y|^2}$ for $y = (y_1, \ldots, y_n) \in \mathbf{R}^n$.
(11) $\langle D_\xi \rangle^2 = \left(1 + \sum_{j=1}^n D_{\xi_j}^2\right) = 1 - \Delta_\xi$ (minus the usual Laplacian).

These conventions greatly simplify many expressions. For example, the Taylor series for a function $f(x)$ takes the form

$$f(x) = \sum_{\alpha \geq 0} \frac{1}{\alpha!} \partial^\alpha f(0) x^\alpha.$$

Chapter 1
Introduction and Summary

In this introductory chapter, ten elementary (but important) examples of diffusion processes are included. Furthermore, our problems and results are stated in such a fashion that a broad spectrum of readers could understand, and it is also described how these problems can be solved, using the mathematics presented in Chaps. 2–9.

(I) First, Table 1.1 below gives a bird's-eye view of strong Markov processes, Feller semigroups and degenerate elliptic Ventcel' (Wentzell) boundary value problems, and how these relate to each other.

(II) Secondly, our functional analytic approach to strong Markov processes through Feller semigroups may be visualized as in Fig. 1.1 below.

(III) Thirdly, Table 1.2 below gives a bird's-eye view of Markov transition functions, Feller semigroups and Green operators (resolvents), and how these relate to each other.

1.1 Markov Processes and Semigroups

This section is devoted to the functional analytic approach to the problem of construction of Markov processes in probability theory. General existence theorems for Markov processes are formulated in terms of elliptic boundary value problems with a spectral parameter.

1.1.1 Brownian Motion

In 1828, the English botanist Robert Brown observed that pollen grains suspended in water move chaotically, incessantly changing their direction of motion. The physical

K. Taira, *Functional Analytic Techniques for Diffusion Processes*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-981-19-1099-9_1

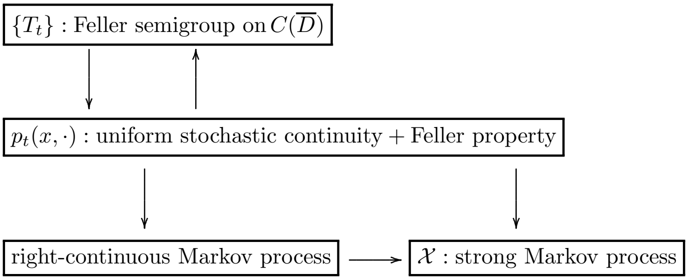


Fig. 1.1 A functional analytic approach to strong Markov processes

Table 1.1 A bird's-eye view of strong Markov processes, Feller semigroups and degenerate elliptic boundary value problems

Probability (Microscopic approach)	Functional analysis (Macroscopic approach)	Elliptic boundary value problems (Mesoscopic approach)
Strong Markov process $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$	Feller semigroup $\{T_t\}t \geq 0$	Infinitesimal generator $\mathfrak{A}$
Markov transition function $p_t(x, dy) = P_x\{x_t \in dy\}$	$T_t f(x) = \int_{\bar{D}} pt(x, dy) f(y)$	$T_t = e^t \mathfrak{A}$
Chapman–Kolmogorov equation $pt + s(x, dz) = \int_{\bar{D}} p_t(x, dy) p_s(y, dz)$	Semigroup property $T_t + s = Tt \cdot T_s$	Degenerate diffusion operator A
Absorption, reflection, viscosity phenomena, drift and diffusion along the boundary	Function space $C(\bar{D})$	Ventcel' (Wentzell) boundary condition L

Table 1.2 A bird's-eye view of Markov transition functions, Feller semigroups and Green operators

Markov transition function $pt(x, dy)$	Dynkin $\Longleftrightarrow$	Feller semigroup $T_t = e^t \mathfrak{A}$
Laplace transform $\Updownarrow$	$T_t f(x) = \int_{\bar{D}} pt(x, dy) f(y)$	$\Updownarrow$ Hille–Yosida
Green kernel $G_\alpha(x, y)$	$\Longleftrightarrow$ Riesz–Markov	Green operator $(\alpha I - \mathfrak{A})^{-1}$

explanation of this phenomenon is that a single grain suffers innumerable collisions with the randomly moving molecules of the surrounding water [27].

A mathematical theory for Brownian motion was put forward by Albert Einstein in 1905 [50]. 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:

$$p(t, x, y) = \frac{1}{\sqrt{2\pi Dt}} \exp^{-\frac{(y-x)^2}{2Dt}}.$$

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 number by observing particles undergoing Brownian motion. Einstein's theory was experimentally tested by Jean Baptiste Perrin between 1906 and 1909 [145].

Brownian motion was put on a firm mathematical foundation for the first time by Norbert Wiener in 1923 [237]. Let Ω be the space of continuous functions $\omega\colon [0, \infty) \mapsto \mathbf{R}$ 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 $\{\omega \in \Omega : a \le x_t(\omega) < b\}$, $t \ge 0$, $a < b$. Wiener constructed probability measures P_x, $x \in \mathbf{R}$, on $\mathcal{F}$ for which the following formula holds true:

$$\begin{aligned} &P_x \left\{\omega \in \Omega : a_1 \le x_{t_1}(\omega) < b_1, a_2 \le x_{t_2}(\omega) < b_2, \ldots, a_n \le x_{t_n}(\omega) < b_n\right\} \\ &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} p(t_1, x, y_1) p(t_2 - t_1, y_1, y_2) \ldots \\ &\qquad p(t_n - t_{n-1}, y_{n-1}, y_n)\, dy_1\, dy_2 \ldots dy_n, \\ &\quad 0 < t_1 < t_2 < \ldots < t_n < \infty. \end{aligned} \tag{1.1}$$

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.

Paul Lévy found another construction of Brownian motion in stochastic analysis, and gave a profound description of qualitative properties of the individual Brownian path in his book [114]: *Processus stochastiques et mouvement brownien.*

1.1.2 Markov Processes

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. (The reader may content himself with thinking of $\mathbf{R}$ while reading about 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

$$\{X \in E\} = X^{-1}(E) \in \mathcal{F} \quad \text{for all } E \in \mathcal{B}.$$

We express this by saying that X is $\mathcal{F}/\mathcal{B}$-measurable. A family $\{x_t\}_{t \ge 0}$ of random variables is called a *stochastic process*, and 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 *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, the (temporally homogeneous) *Markov property* is that the prediction of subsequent motion of a 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 particle starts afresh.

Now we introduce a class of Markov processes which we will deal with in this book (Definition 12.3).

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. We let

$$\begin{aligned} K_\partial &= K \cup \{\partial\}, \\ \mathcal{B}_\partial &= \text{ the } \sigma\text{-algebra in } K_\partial \text{ generated by } \mathcal{B}. \end{aligned}$$

(2) The space Ω of all mappings $\omega\colon [0, \infty] \to K_\partial$ such that $\omega(\infty) = \partial$ and that if $\omega(t) = \partial$ then $\omega(s) = \partial$ for all $s \geq t$. We 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 every $\omega \in \Omega$.
(4) For each $t \in [0, \infty]$, a mapping $\varphi_t\colon \Omega \to \Omega$ defined by $\varphi_t\omega(s) = \omega(t + s)$, $\omega \in \Omega$. Note that $\varphi_\infty\omega = \omega_\partial$ and $x_t \circ \varphi_s = x_{t+s}$ 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* $\mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x)$ if the following conditions are satisfied:

(i) For each $0 \leq t < \infty$, the function x_t is $\mathcal{F}_t/\mathcal{B}_\partial$-measurable, that is,

$$\{x_t \in E\} \in \mathcal{F}_t \quad \text{for all} \quad E \in \mathcal{B}_\partial.$$

(ii) For each $0 \leq t < \infty$ and $E \in \mathcal{B}$, the function

$$p_t(x, E) = P_x\{x_t \in E\} \tag{1.2}$$

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 the formula

$$P_x \{x_{t+h} \in E | \mathcal{F}_t\} = p_h (x_t, E), \tag{1.3}$$

or equivalently

$$P_x (A \cap \{x_{t+h} \in E\}) = \int_A p_h(x_t(\omega), E) \, dP_x(\omega) \text{ for every } A \in \mathcal{F}_t. \tag{1.3$'$}$$

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. The function p_t is called the *transition function* of the process $\mathcal{X}$. The transition function p_t 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. A particle moves in the space K until it "dies" at which time it reaches the point ∂; hence the point ∂ is called the *terminal point*.

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 $\mathcal{X}$.

Using the Markov property (1.3$'$) repeatedly, we easily obtain the following formula, analogous to formula (1.1):

$$\begin{aligned} &P_x \left\{\omega \in \Omega : x_{t_1}(\omega) \in A_1, x_{t_2}(\omega) \in A_2, \ldots, x_{t_n}(\omega) \in A_n\right\} \\ &= \int_{A_1} \int_{A_2} \cdots \int_{A_n} p_{t_1}(x, dy_1) p_{t_2 - t_1}(y_1, dy_2) \cdots p_{t_n - t_{n-1}}(y_{n-1}, dy_n), \\ &0 < t_1 < t_2 < \ldots < t_n < \infty, \quad A_1, A_2, \ldots, A_n \in \mathcal{B}. \end{aligned}$$

1.1.3 Transition Functions

From the viewpoint 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.

First, we give the precise definition of a transition function which is adapted to analysis (Definition 12.4):

Let K 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 if it satisfies the following four conditions:

(a) $p_t(x, \cdot)$ is a measure on $\mathcal{B}$ and $p_t(x, K) \leq 1$ for each $t \geq 0$ and $x \in K$.
(b) $p_t(\cdot, E)$ is a Borel measurable function for each $t \geq 0$ and $E \in \mathcal{B}$.
(c) $p_0(x, \{x\}) = 1$ for each $x \in K$.
(d) For any $t, s \geq 0$ and $E \in \mathcal{B}$, we have the formula

$$p_{t+s}(x, E) = \int_K p_t(x, dy)\, p_s(y, E). \tag{1.4}$$

Equation (1.4), called the *Chapman–Kolmogorov equation* [34, 103], 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. Thus it is just condition (d) which reflects the Markov property that a particle starts afresh.

The Chapman–Kolmogorov equation (1.4) asserts that the transition function $p_t(x, K)$ 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 transition function p_t is said to be *normal* if it satisfies the condition

$$p_{+0}(x, K) = 1 \quad \text{for all } x \in K.$$

The next theorem justifies our definition of a transition function, and hence it will be fundamental for our further study of Markov processes:

Theorem 1.1 (Dynkin) *For every Markov process, the function p_t, defined by* (1.2)*, is a transition function. Conversely, every normal transition function corresponds to some Markov process.*

Here are some important examples of normal transition functions on $\mathbf{R}$.

Example 1.2 (*uniform motion*) If $t \geq 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \chi_E(x + vt), \tag{1.5}$$

where v is a constant, and

$$\chi_E(y) = \begin{cases} 1 & \text{if } y \in E, \\ 0 & \text{if } y \notin E. \end{cases}$$

This process, starting at x, moves deterministically with constant velocity v.

Example 1.3 (*Poisson process*) If $t \geq 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \chi_E(x + n), \tag{1.6}$$

where λ is a positive constant.

This process, starting at x, advances one unit by jumps, and the probability of n jumps in time t is equal to $e^{-\lambda t}(\lambda t)^n/n!$.

Example 1.4 (*Brownian motion*) If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp\left[-\frac{(y - x)^2}{2t}\right] dy, \tag{1.7}$$

and

$$p_0(x, E) = \chi_E(x).$$

This is a mathematical model of one-dimensional Brownian motion.

Example 1.5 (*Brownian motion with constant drift*) If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp\left[-\frac{(y - mt - x)^2}{2t}\right] dy, \tag{1.8}$$

and

$$p_0(x, E) = \chi_E(x),$$

where m is a constant.

This represents Brownian motion with constant drift m: the process can be represented as $\{x_t + mt\}$, where $\{x_t\}$ is Brownian motion.

Example 1.6 (*Cauchy process*) If $t > 0$, $x \in \mathbf{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\pi} \int_E \frac{t}{t^2 + (y - x)^2} dy, \tag{1.9}$$

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.

1.1.4 Kolmogorov's Equations

Among 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 [103].

We now take a close look at Kolmogorov's work. Let p_t be a transition function on $\mathbf{R}$, and assume that the following two conditions are satisfied:

(i) For each $\varepsilon > 0$, we have the assertion

$$\lim_{t \downarrow 0} \frac{1}{t} \sup_{x \in \mathbf{R}} p_t(x, \mathbf{R} \setminus (x - \varepsilon, x + \varepsilon)) = 0.$$

(ii) The three 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 thought of as variance (over $\omega \in \Omega$) instantaneous (with respect to t) velocity when the process is at position x (see Sect. 1.1.2), and the limit $b(x)$ has a similar interpretation as a mean. The transition functions (1.5), (1.7) and (1.8) 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 (1.6) and (1.9) do not satisfy condition (i).

Furthermore, we assume that the transition function p_t 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 from 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:

$$\begin{cases} \dfrac{\partial p}{\partial t} = \dfrac{a(x)}{2} \dfrac{\partial^2 p}{\partial x^2} + b(x) \dfrac{\partial p}{\partial x} + c(x) p, \quad t > 0. \\ \lim_{t \downarrow 0} p(t, x, y) = \delta(x - y), \end{cases} \tag{1.10}$$

and is, for fixed x, the fundamental solution of the Cauchy problem: